EXPRESSION OF THE *PASTEURELLA HAEMOLYTICA*

O-SIALOGLYCOPROTEIN ENDOPEPTIDASE AS A GST FUSION PROTEIN

A Thesis

Presented to

The Faculty of Graduate Studies

of

The University of Guelph

by

STEVEN ARI BEASLEY

In partial fulfillment of requirements

for the degree of

Master of Science

May, 1998

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

Canada

ABSTRACT

EXPRESSION OF THE *PASTEURELLA HAEMOLYTICA* *O*-SIALOGLYCOPROTEIN ENDOPEPTIDASE AS A GST FUSION PROTEIN

Steven Ari Beasley   Advisor:
University of Guelph, 1998   Dr. Alan Mellors

Abstract.

The *Pasteurella haemolytica* *O*-sialoglycoprotein endopeptidase specifically cleaves cell-membrane glycoproteins that bear negatively charged clusters such as sialo- and sulfoglycans. This thesis describes the heterologous expression in *E. coli* of the *gcp* gene, in fusion with the gene for glutathione-*S*-transferase (GST) of *Schistosoma japonicum*. The GST fusion protein product, rGgcp, was purified by affinity chromatography on a glutathione-Sepharose column, and thrombin cleavage of rGgcp liberated the rGcp moiety. The purified recombinant protein was used as a substrate for *in vitro* refolding by the molecular chaperones PDI and DnaK/DnaJ/GrpE under a variety of experimental conditions. Glycoprotease enzyme activity was not reproducibly generated in any of the rGgcp samples tested.

A novel shuttle vector, pNF2176, was used for the expression of rGcp and rGgcp in *P. haemolytica* serotypes A1 and A11. The glycoprotease antigen was expressed in *P. haemolytica* serotypes A1 and A11, but there was no increase in glycoprotease activity in either serotype associated with the gene expression. In contrast with previous reports, serotype A11 was found to constitutively express an enzymatically active glycoprotease.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AMP	Ampicillin
ATP	Adenosine 5'-triphosphate
BAP	Blood agar plate
BCIP	5-bromo-4-chloro-3-indoyl phosphate
BODIPY-FL	4,4-difluor-5,7-dimethyl-4-bora-3α,4α-diaza-s-indacene-3-propionic acid
BHIB	Brain heart infusion broth
C/S	Culture supernatant
CCS	Concentrated culture supernatant
CHAPS	3-[(3-cholamidopropyl)dimethylammonio]-1-propane sulphonate
CPS	Capsular polysaccharide
CYT	Cytoplasmic
DNA	Deoxyribonucleic acid
DTT	Dithiothreotol
EDTA	Ethylenediaminetetraacetic acid
GCP	*P. haemolytica O*-sialoglycoprotein endopeptidase protein
GGCP	Fusion protein of GST and GCP
GPA	Glycophorin A
GST	Glutathione-*S*-transferase
GuHCl	Guanidinium chloride
HEPES	4-(2-hydroxyethyl)-1-piperazine-ethane-sulphonate
IPTG	Isopropyl *β-D*-thiogalactoside
kb	kilobase
kDa	kilodalton
LB	Luria-Bertani
LKT	Leukotoxin
LPS	Lipopolysaccharide
LT	Luria-thymine
MCS	Multiple cloning site
NBT	*p*-Nitro blue tetrazolium
OMP	Outer membrane protein
PBS	Phosphate buffer saline
PER	Periplasm
P.I.C.	Protease Inhibitor Cocktail
PSGL-1	P-selectin glycoprotein ligand-1
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SRL	Sucrose-RNAse-lysozyme solution
TA(E)	Tris-acetate-(EDTA) buffer
TB(E)	Tris-borate-(EDTA) buffer
TE	Tris-EDTA buffer
TEMED	N,N,N′,N′-tetramethylene-ethylenediamine
TTBS	Tween 20-Tris-HCl buffer saline

Chapter 1. Literature Review.

A. *Pasteurella haemolytica* and *pneumonic pasteurellosis.*

L. *Pasteurella haemolytica.*

The members of the eubacterial taxonomic family *Pasteurellaceae* comprise the species *Haemophilus*, *Actinobacillus*, and *Pasteurella* (the HAP family). Most members of this family are commensals of mucous membranes in vertebrates, and have been identified as the etiological agents of a number of respiratory diseases (Biberstein, 1990). These pathogenic bacteria have a devastating economic impact upon the livestock industry which has prompted much effort to determine means of control (Barnum, 1990).

One member of this family is *P. haemolytica*, a bipolar staining, Gram-negative, facultatively anaerobic, non-motile, ovoid, cocci bacterium found in the nasal flora of ruminants. Biochemical characteristics of this species include growth on MacConkey's agar, haemolysis on blood agar, fermentation of mannitol, xylose, and arabinose, and positive responses to catalase and cytochrome oxidase enzymes (Bergey's Manual of Determinative Bacteriology, 1994). The species was first described in 1932 after isolation from a pneumonic calf (Newsome and Cross, 1932). The species was originally divided into two biotypes, A and T, based on their ability to ferment arabinose and trehalose respectively (Smith, 1961). However, the T biotype has been recently re-classified as a separate species, *P. trehalosi*, based on DNA sequence homology (Sneath and Stevens, 1990). Among the two biotypes, there are seventeen serologically distinct strains, plus several untypeable strains (Davies *et al.*, 1996). The strains are routinely distinguished by a haemagglutination assay using bovine red blood cells (Biberstein, 1978), a slide

agglutination assay using typing serum (Frank and Wessman, 1978), or more recently, a coagglutination assay (Fodor *et al.*, 1996).

II. Etiology and pathology of pneumonic pasteurellosis.

Numerous bacterial strains are found in the nasopharynx of healthy cattle. These include *P. haemolytica*, *P. multocida*, *Actinomyces* spp., *Diplococcus pneumoniae*, *Staphylococcus aureus*, and *Pseudomonas aeruginosa* (Barbour *et al.*, 1997). In cattle infected with bovine respiratory tract disease, the predominant isolate recovered from pneumonic lungs is *P. haemolytica* serotype A1 (Yates, 1982; Purdy *et al.*, 1997). This bacterium has been implicated as the primary causative agent of morbidity and mortality in feedlot cattle. Pneumonic pasteurellosis can be induced experimentally in healthy calves after the introduction of *P. haemolytica* A1 into the lungs by aerosol spray, intra-tracheal inoculation, or trans-thoracic injection (Collier, 1968; Friend *et al.*, 1977; Panciera and Corstvet, 1984). *P. haemolytica* is readily cleared from the lungs in healthy cattle, but when an animal becomes stressed and immuno-compromised, due to environmental factors (shipping; dietary changes; viral infections such as bovine rhinotracheitis), *P. haemolytica* proliferates throughout the respiratory tract and colonizes the lungs to cause pneumonic pasteurellosis (Frank and Smith, 1983). This disease is characterized by the clinical observations of severe depression, reduced appetite, nasal discharge, and fever. Histopathological changes consist of severe fibrinocellular alveolitis, necrosis of lobules, the accumulation of neutrophils and macrophages, and varying degrees of thrombosis and inflammation (Friend *et al.*, 1977, Rehmtulla and Thomson, 1981). The damage to bovine lung is due not only to *P. haemolytica* virulence factors, but also to the recruitment of neutrophils to the lung, which is stimulated by the *P. haemolytica* culture supernatant

2

(Brunner *et al.*, 1989). Neutrophils have been implicated in the initiation of pulmonary injury, partly due to the generation of toxic reactive oxygen species that could account for the microscopic lesions typically found in pneumonic lungs (Watson *et al.*, 1995).

III. Treatment and prevention of pneumonic pasteurellosis.

The prevalence of pneumonic pasteurellosis throughout the beef industry has led to extensive research on the prevention and treatment of the disease. The development of a more effective vaccine against the disease is an area of active research. There are at least sixteen vaccines commercially available today (Srinand *et al.*, 1996). Vaccines based on killed bacteria have proven to be ineffective. However, live bacteria isolated from pneumonic calves, which were introduced transthoracically, by intrapulmonary challenge, or subcutaneousically, have been shown to be more efficacious in protecting cattle against an experimental challenge. This protection correlates with an increase in anti-*P. haemolytica* antibodies in the serum (Panciera *et al.*, 1984; McBride *et al.*, 1996). However, the use of live bacteria in a vaccine can produce severe side effects. The observation that serum antibodies correlate with decreased lung lesions has led to the development of antigen-based vaccines (Sreevatsan *et al.*, 1996). One commercial vaccine, "Presponse™" uses the secreted soluble antigens of *P. haemolytica* (Shewen *et al.*, 1988). It can be supplemented with recombinant leukotoxin to maximize its effect in reducing the clinical manifestations of pasteurellosis (Conlon *et al.*, 1991). The efficacy of the vaccine can be further enhanced by the addition of another recombinant virulence factor of *P. haemolytica*, a fusion protein incorporating an enzyme, *O*-sialoglycoprotein endopeptidase (E.C. 3.4.24.57) (Shewen *et al.*, 1994). The addition of a third virulence factor, capsular polysaccharide (CPS), to the "Presponse™" vaccine did not result in any

3

improvement; conversely, an undesirable consequence was anaphylaxis, which developed in 36% of cattle receiving CPS (Conlon and Shewen, 1993). In another study, a combination of specific adjuvants and CPS was found to provide enhanced immunity when compared with the commercial vaccines "PH-K™", and "1-Shot™" (Brogden *et al.*, 1995). Vaccination of pregnant cows with "Presponse™" produced an increase in antibody titers against CPS in their calves. This preparturient vaccination scheme is a promising approach in the protection of cattle herds (Hodgins and Shewen, 1996). The efficacy of some vaccines has been enhanced by chromium salt supplementation which gives rise to an increase in serum antibody levels (Chang *et al.*, 1996). No vaccine has provided complete protection to pasteurellosis and, since the best available vaccines give about 70% protection, further research in vaccine improvement is needed.

IV. Plasmids and antibiotic resistance in *P. haemolytica*.

Treatment of cattle infected with bovine respiratory tract disease has been based largely on the use of antibiotics. However, this approach has led to an increase in the incident of drug-resistant bacteria. In most cases, the resistance is encoded by plasmids and is exacerbated by the extensive use of antibiotics in veterinary practice. Many strains have been shown to display multi-drug resistance, due to the presence of multiple genes on either a single plasmid, or compatible plasmids. Plasmid-borne resistance to streptomycin and tetracycline was reported in a *P. haemolytica* A1 strain isolated from an infected calf (Zimmerman and Hirsh, 1980). Resistance to other antibiotics such as β-lactams (e.g. ampicillin) or sulfonamides is widespread in isolates recovered during outbreaks of pasteurellosis (Livrelli *et al.*, 1991; Chang *et al.*, 1992; Murphy *et al.*, 1993). Chloramphenicol acetyltransferase activity was recently implicated after a 1992 study

4

found that 30% of *P. haemolytica* and *P. multocida* isolates were resistant to chloramphenicol. This finding, in addition to the high toxicity of chloramphenicol, led to the prohibition in 1994 of chloramphenicol as a veterinary therapeutic in Europe (Vassort-Bruneau *et al.*, 1996). New antibiotics are currently being tested, such as the β-lactamase-resistant cephalosporin, ceftiofur, which has proven to be effective against *in vitro* cultures of *Pasteurella* spp. (Blackall *et al.*, 1996). However, unless novel antibiotics are used judiciously, these recent developments will be rendered useless by the spread of plasmid-borne resistance.

A beneficial aspect of the existence of antibiotic-resistant plasmids is that they have allowed the creation of cloning vectors that can be used in the expression and characterization of native genes, as well as the creation of isogenic mutants through allelic exchange. However, the development of useful shuttle vector plasmids for the genetic manipulation of this bacterial family has not been easy. This could be due to the restriction-modification barrier that cloned DNA must overcome in the host bacterium, or perhaps due to host-incompatibility for the plasmid origin of replication.

A number of plasmids have been constructed in the past decade that can "shuttle" between *Escherichia coli* and members of the HAP family of organisms, including *P. haemolytica*. One broad host range cloning vector, pJFF224-NX was derived from *Actinobacillus pleuropneumoniae* (Frey, 1992). In our laboratory several attempts were made to use this vector to express the *P. haemolytica* Gcp, however these efforts were unsuccessful (Watt, personal communication). Another *Actinobacillus pleuropneumoniae* plasmid, pGZRS-1, which encodes for sulfonamide and streptomycin resistance, was used to construct a family of shuttle vectors, pGZRS-18/19 and pGZRS-38/39 based on the

cloning site of pUC-18/19 and either ampicillin or kanamycin selection markers, respectively (West *et al.*, 1995). These plasmids are readily replicated in *P. haemolytica*, and were used to express an ampicillin resistance gene product in *P. haemolytica*. Recently, a native *Pasteurella multocida* plasmid was used to create the cloning vector pIG112. Although this plasmid replicates stably in *P. haemolytica*, the potential of this vector for protein expression in *P. haemolytica* was not investigated (Wright *et al.*, 1997). The HAP plasmids are apparently derived from a common ancestral origin, in that a nucleotide sequence analysis of pIG112 found a high degree of homology with two *P. haemolytica* plasmids, pYFC1 and pAB2, and the *Haemophilus ducreyi* plasmid pLS88.

Several attempts have been made to create a suitable cloning vector from a *P. haemolytica* derived plasmid. The first potential vector was the ampicillin-resistant plasmid, pAB2, which was found to replicate both in *E. coli* and *P. haemolytica* (Craig *et al.*, 1989). The use of this plasmid as a cloning vector is limited, as a subclone containing a fragment of the leukotoxin A gene was unable to transform *P. haemolytica* A1 (Wood *et al.*, 1995). Another ampicillin resistance-bearing plasmid, pPH843, was used to develop three new constructs, pAKA-16, -19, and -22 (Azad *et al.*, 1994). The latter two failed to replicate in *P. haemolytica* and were suitable only as suicide vectors, and successful protein expression has not been reported for pAKA-16. Recently, a *P. haemolytica* sulfonamide (Su^r) and streptomycin (Sm^r) resistance-bearing plasmid, pYFC-1 (Chang *et al.*, 1992), was used to construct a new *E. coli-P. haemolytica* shuttle vector. The new vector, pNF2176, was shown to express a promoterless gene *in vivo* under the control of the sulfonamide promoter (Fedorova and Highlander, 1997a).

6

V. *P. haemolytica* serotype A1 virulence factors.

The most promising vaccines against bovine pasteurellosis are based on the virulence factors of *P. haemolytica*. A number of cellular components involved in the pathogenesis of pasteurellosis have been identified, including capsular polysaccharide (CPS), lipopolysaccharide (LPS), outer membrane proteins and iron-regulated outer membrane proteins (OMPs & IROMPs), a neuraminidase, a leukotoxin (Lkt), and an extracellular glycoprotease (Gcp).

A significant factor in bacterial pathogenicity is the presence and composition of a capsular polysaccharide layer (glycocalyx). The glycocalyx is important in protecting the bacteria from host phagocytosis and complement-mediated killing (Confer *et al.*, 1990). When *P. haemolytica* is isolated from pneumonic lungs, a distinct CPS layer can be seen by electron microscopy (Brogden and Clarke, 1997). The composition of the CPS is serotype-specific and may play a significant role in the strain-specificity of the virulence. The glycocalyx may protect the bacterium against attack by neutrophils. The colonization of the lungs by the bacteria is also aided by CPS-mediated adhesion to the alveolar epithelium. Furthermore, the adherence of the bacteria to the epithelial lining may be facilitated by fimbriae expressed by *P. haemolytica* (Morck *et al.*, 1989).

P. haemolytica LPS (endotoxin) is composed of biologically active lipid A, core oligosaccharide, and an antigenic polysaccharide chain specific for the different serotypes. The LPS from the different serotypes is recognized by monoclonal antibody raised against serotype-specific endotoxin (Durham *et al.*, 1988). The LPS alters the surface tension of sheep surfactant, perhaps leading to alveolar hemorrhaging. Pathologically, LPS may be a significant factor in pasteurellosis by inducing lung inflammation, damaging the

7

endothelium, and modifying leukocyte function. An increased concentration of endotoxin correlates with pulmonary hemorrhage, edema, and acute inflammation (Slocombe *et al.*, 1990). LPS induces pathological changes in a two-phase reaction; first there is a release of arachidonic acid metabolites causing pulmonary hypertension, followed by neutrophil infiltration and resultant damage caused by leukocyte mediators (proteases, oxygen radicals, etc.) (Confer *et al.*, 1990).

Outer membrane proteins are recognized by the antisera from pneumonic cattle (Davies *et al.*, 1994). Three major proteins of 70 kDa, 77 kDa, and 100 kDa are observed by SDS-PAGE analysis when *P. haemolytica* is grown *in vivo*, or under iron-depleted conditions *in vitro*. A 100 kDa IROMP has been identified as the transferrin receptor. Vaccination of cattle, with *P. haemolytica* fractions enriched with these three proteins, produces an increase in resistance to challenge (Morck *et al.*, 1991; Confer *et al.*, 1995). Other significant OMPs that react with bovine antisera have been identified and characterized, including proteins of 94 kDa, 84 kDa, 53.5 kDa, 49 kDa, 43 kDa, 41 kDa, and 16 kDa and two 30 kDa proteins (PomA and PomB) (Morton *et al.*, 1996; Parameshwar *et al.*, 1997). Although these proteins are immunogenic, any correlation between serum antibody levels against OMPs and protection from pasteurellosis has not been conclusively demonstrated.

The leukotoxin of *P. haemolytica* has been shown to be a major factor in the pathogenesis of pneumonic pasteurellosis. The genes for the 102 kDa Lkt protein (*lktA*), a 20 kDa activating protein (*lktC*), and the leukotoxin secretory apparatus (*lktB* & *lktD*) have been cloned and characterized (Lo *et al.*, 1987; Strathdee and Lo, 1989). The *P. haemolytica* Lkt determinant is highly homologous with the *E. coli* hemolysin

determinant. Biochemically, Lkt and hemolysin belong to the RTX group of toxins, which act as transmembrane pores and destroy the integrity of eukaryotic cell membranes. Lkt has a deleterious effect on the bovine immune system by targeting both leukocytes and platelet functions in a Ca^{2+}-dependent manner. At low concentrations, Lkt causes leukocytes to release inflammatory mediators and initiates apoptosis. At higher concentrations, Lkt causes target cell membrane pore formation, resulting in the rapid depletion of small biomolecules such as ATP, loss of K^+ ions and the entry of Ca^{2+}. This causes the cells to swell and lyse (Clinkenbeard et al., 1989; Maheswaran et al., 1993).

Flow cytometry studies have shown that Lkt binding is specific for bovine leukocytes. The binding appears to be to an unidentified leukocyte membrane protein, since pretreatment of leukocytes with proteinase K abolished Lkt binding (Brown et al., 1997). Bovine platelets are also rapidly lysed by Lkt, releasing fibrinogen and other factors that could be responsible for thrombus formation and the fibrinous exudates seen in pneumonic lungs (Clinkenbeard and Upton, 1991). Significant reductions in mortality and in lung lesions are found in cattle infected with P. haemolytica mutants deficient in Lkt (Petras et al., 1995). The protective effect of rLkt and Lkt in vaccines has been well documented (Conlon et al., 1991; Confer et al., 1997).

Another extracellular enzyme excreted by P. haemolytica is neuraminidase (sialidase, EC. 3.2.1.18) (Frank and Tabatabai, 1981). This enzyme desialylates glycoproteins, including fetuin and bovine submaxillary mucin (Straus et al., 1993a), and potentially it could interfere with cell-cell interactions during the host immune response. Neuraminidase activity has been detected in every P. haemolytica A1 field strain examined (372 in one study) as well as in other Pasteurella species (Straus and Purdy, 1995).

P. haemolytica culture supernatants have also been found to contain a proteolytic enzyme specific for sialylated glycoprotein substrates, named the *P. haemolytica* O-sialoglycoprotein endopeptidase (EC. 3.4.24.57) (Otulakowski *et al.*, 1983). This enzyme is immunogenic in that neutralizing antibodies have been detected in the sera from infected calves (Abdullah *et al.*, 1990; Lee *et al.*, 1994b). The glycoprotease enzyme and its genetic manipulation are the subject of this thesis.

B. *P. haemolytica* O-sialoglycoprotein endopeptidase.

I. Characterization of Gcp.

Glycoprotease activity has been found in the supernatant of logarithmic phase cultures of all *P. haemolytica* A biotypes except the non-pathogenic strain A11, but was not found in the T biotypes (*P. trehalosi*) (Abdullah *et al.*, 1991; Lee *et al.*, 1994a). There are four broad classes of proteases: serine proteases, thiol proteases, acid proteases, and neutral metalloproteases (Neurath, 1989). The glycoprotease enzyme has been classified as a neutral metalloprotease (EC. 3.4.24.57) since it contains a putative zinc ion binding site and it can be inhibited by EDTA, although no metal ion activator has yet been identified. The observation that Gcp inhibition by EDTA is partially reversible by dialysis, suggests that inhibition is due to the masking of the active site of the enzyme by the inhibitor, and not due to the removal of any metal ion. Inhibition of the glycoprotease by citrate and ascorbate ions supports this hypothesis. The enzyme is not inhibited by either serine protease inhibitors (phenylmethanesulfonyl fluoride, di-isopropylphosphofluoridate, aprotinin, or soybean trypsin inhibitor), thiol protease inhibitors (antipain, leupeptin, *N*-ethylmaleimide, *p*-chloromercuribenzoate, or *p*-hydroxymercuribenzoate), or the acid protease inhibitor pepstatin (Abdullah *et al.*, 1992). However, unlike the well-known

Zn^{2+}-dependent bacterial metalloprotease thermolysin, Gcp is neither inhibited by phosphoramidon, nor by the thermolysin substrate furoylacryloylglycylleucinamide (Abdullah, 1991).

The glycoprotease appears to be released into culture supernatant when outer membrane vesicles are blebbed off the cell surface (P. Jiang, personal communication). Attempts to purify the protein to homogeneity, free from contaminating virulence factors such as the leukotoxin and LPS, have proven challenging. Enzyme-rich fractions are now prepared by subculturing a 4.5 h brain-heart infusion broth culture of *P. haemolytica* A1 into RPMI 1640 broth supplemented with 0.2% CHAPS for 3-4 h at 37°C with aeration. The culture supernatant is filtered to remove cellular debris before concentration by diafiltration. The resulting concentrated culture supernatant (CCS) is used as an enzyme extract, and can be partially purified by chromatographic separation on CM-cellulose chromatography (Jiang and Mellors, unpublished).

II. Substrate specificity of Gcp.

A unique feature of the glycoprotease is its high specificity for glycoproteins carrying extensive clusters of negatively charged sugars, i.e. sialoglycan and sulfoglycan residues. Such substrates are usually eukaryotic transmembrane proteins with approximately 30% serine or threonine residues in their extracellular domains. Typically these substrates are *O*-sialoglycoproteins and are known as cell-surface mucins, although not all *O*-sialoglycoproteins are substrates. The enzyme does not cleave *N*-linked sialoglycoproteins or desialylated substrates (Abdullah *et al.*, 1992). The best characterized substrate for Gcp is human glycophorin A (GPA); it is currently used for Gcp activity assays as either a radioiodinated or fluorescently labeled substrate. GPA is

the major sialoglycoprotein on the human red cell surface and bears one N-glycan and 15

O-glycans, the latter being conjugated to serine and threonine residues of the N-terminal

extracellular region of the protein. The serine or threonine linked O-glycans were

identified as a major disialylated tetrasaccharide and a minor monosialo linear trisaccharide

(Blanchard, 1990). The major site of cleavage of GPA by Gcp is found at Arg31-Asp32,

however other peptide bonds are also cleaved after treatment with Gcp extracts (Abdullah

et al., 1992). Radioiodinated or fluorescently labeled GPA substrates have been used to

determine the K_m values for Gcp, which are in the range of 2.5 μM to 5 μM.

A useful feature of Gcp is that it can cleave a narrow range of cell surface O-

sialoglycoprotein substrates on living cells, without affecting cell viability in culture. A

number of human cell substrates have been identified in the past six years including the

leukocyte cell-surface proteins CD34 (a pan-hematopoietic stem cell antigen), CD43

(leukosialin), CD44 (hyaluronic acid receptor), CD45 (leukocyte activation receptor

common antigen, and a transmembrane tyrosine-phosphatase) (Sutherland *et al.*, 1992).

The ability of Gcp to cleave CD34 on the surface of live human bone marrow stem cells

has been exploited for the immuno-magnetic purification of functional hematopoietic

progenitor cells used in bone marrow transplants (Marsh *et al.*, 1992).

Other human cell surface O-sialoglycoprotein substrates for Gcp include the

ligands for binding to selectins. Selectins are a group of cell adhesion molecules that can

bind to glycoprotein ligands on the surface of neighbouring cells. The three major classes

of selectins are the P-selectin (platelet selectin), L-selectin (lymphocyte selectin), and E-

selectin (endothelial cell selectin). The binding of HL-60 tumor cells and neutrophils to P-

selectin and L-selectin, but not E-selectin, was abolished after the tumor cells were

pretreated with Gcp. Three P-selectin ligands of molecular weight ~100 kDa, 115 kDa, and 125 kDa were identified by autoradiography of glycoprotease treated [^{35}S]-cysteine labeled HL-60 cells (Steininger *et al.*, 1992). Colon cancer cell lines also possess ligands susceptible to cleavage by Gcp, and which normally bind P- and L-selectin in cell-binding assays. However in these cells, a ligand for E-selectin was also found to be cleaved by Gcp (Mannori *et al.*, 1995). An L-selectin ligand sensitive to cleavage by Gcp was detected when the homotypic aggregation of neutrophils was inhibited by Gcp treatment (Bennett *et al.*, 1995). The probable identity of the ligand recognized by both the P- and L-selectin is the 220 kDa dimeric P-selectin glycoprotein ligand-1 (PSGL-1). This protein is widely distributed on a number of different hemopoietic and myeloid cells, and treatment with either Gcp or anti-PSGL-1 antibodies abolishes both P- and L-selectin binding (Spertini *et al.*, 1996). In order for PSGL-1 to bind P-selectin, it was shown that the protein required two *O*-linked sialylated and fucosylated glycans and three sulfated tyrosine residues (Li *et al.*, 1996). Another mucin-type glycoprotein, CD24, was also found to be a ligand for P-selectin on PSGL-1 negative cells. This glycoprotein is modified with sulfate-containing carbohydrate moieties and is resistant to sialidase treatment (Aigner *et al.*, 1997). These findings indicate that Gcp recognizes negatively charged *O*-linked glycosylated substrates, i.e. proteins bearing extensive clusters of sialoglycans or sulfoglycans.

Other *O*-linked sialoglycoprotein Gcp substrates include: VAP-1, a 170 kDa protein that mediates lymphocyte binding to endothelial cells (Salmi and Jalkanen, 1996); the mucins, epiglycanin and epitectin from carcinoma cells (Kemperman *et al.*, 1994; Hu *et al.*, 1994); the platelet glycoprotein Ibα (Yeo and Sutherland, 1995); the related laminin-

binding proteins dystroglycan and cranin (Smalheiser and Kim, 1995); the human interleukin-7 receptor (Titley and Healey, personal communication), and two Epstein-Barr viral glycoproteins gp350 and gp150 (Hutt-Fletcher, personal communication). Soluble sialoglycoproteins such as fetuin or tenascin are not substrates for Gcp (Cladman *et al.*, 1996; Clark *et al.*, 1997).

Studies are currently under way to determine the bovine substrates important for the pathogenic role of the glycoprotease. A host humoral factor, bovine immunoglobulin IgG$_1$, was found to be proteolysed by *P. haemolytica* culture supernatants, although the proteolytic agent has not been identified (Lee and Shewen, 1996). Human IgG$_1$ is not a substrate for Gcp. Flow cytometry studies using fluorescently labeled cells indicate that bovine IgG$_1$ could be a Gcp substrate (Beeston and Shewen, personal communication).

III. Genetic characterization of Gcp.

Gcp exhibits a pI of 5.2 and is biologically active within a pH range of 4.5-8.0, with optimum activity at neutral pH. A recent re-appraisal of the nucleotide sequence analysis has shown that the protein is 343 amino acids in length with a predicted molecular weight of 37 kDa (Hills, personal communication). This confirms the evidence from chemiluminescent Western blots of Gcp protein which shows that Gcp has a higher apparent molecular weight than the 36 kDa *E. coli orfX* gene product (Watt *et al.*, 1997b).

Many open reading frames, which share significant homology with that of Gcp, have been sequenced from a wide variety of organisms, including archaebacteria, eubacteria and eukaryotes. The analysis shown in Figure 1.1 was performed by a Gapped BLAST search of the NCBI databases (Altschul *et al.*, 1990; Altschul *et al.*, 1997). The

organism whose gene has the highest homology is the closely related *Haemophilus influenzae* RD *gcp* with a 92% sequence homology (Table 1.1). Chemiluminescent Western blotting of an overnight *H. influenzae* culture has revealed an antigen of the same molecular weight as the *P. haemolytica* Gcp (Watt *et al.*, 1997b). In the same study, antigen of a slightly different size was also found in *E. coli* and *Neisseria meningitidis* cultures. The *E. coli* antigen is probably the putative 36 kDa *orfX* gene product, a gene found upstream from the *rpsU-dnaG-rpoD* macromolecular synthesis operon (Nesin *et al.*, 1987). The *orfX* product shares a 76% amino acid homology with *P. haemolytica* Gcp.

The wide distribution of the *gcp* gene throughout the biosphere, from *Archaebacter* to *Mammalia*, strongly suggests an essential role for the gene product, apart from its only observed enzymatic activity, which is restricted to a few strains of *Pasteurellaceae*. Since extracellular glycoprotease activity has yet to be found in species other than *P. haemolytica*, the exact function of the protein in other species is a matter of speculation. The N-terminal region (not the kinase domain) of a protein kinase from the archaebacterium *Methanococcus vannielii* has been found to possess homology with the *P. haemolytica* Gcp (Smith and King, 1995). However, the authors did not attribute any function to the Gcp-like domain.

Potential functions for Gcp homologues can be predicted from several highly conserved amino acid motifs. The most striking homology can be found in the N-terminal region, which suggests a conserved signal sequence. However, there is no known consensus secretion signal sequence or signal peptide cleavage site which corresponds to the N-terminal region of Gcp and its many homologues (Nielsen *et al.*, 1997). Another

Table 1.1. A comparison of nucleotide homology and amino acid homology (identity and similarity) between open reading frames in various species and the gene for *P. haemolytica* O-sialoglycoprotein endopeptidase, as determined by a Gapped Blast search of Genbank.

Species	Nucleotide homology	Amino acid homology	Reference
Eubacteria			
Pasteurella haemolytica	100%	343/343 (100%)	Abdullah *et al.*, 1991.
Haemophilus influenzae	92%	292/319 (92%)	Fleischmann *et al.*, 1995.
Escherichia coli	68%	282/319 (89%)	Nesin *et al.*, 1987.
Bacillus subtilis	-	209/315 (66%)	Sadaie *et al.*, 1997.
Aquifex aeolicus	-	214/338 (62%)	Deckert *et al.*, 1998.
Mycobacterium tuberculosis	-	193/317 (61%)	Philipp *et al.*, 1996.
Borrelia burgdorferi	-	196/317 (61%)	Fraser *et al.*, 1997.
Mycobacterium leprae	-	191/318 (60%)	Genbank, S72817
Synechocystis sp.	-	183/314 (58%)	Kaneko *et al.*, 1996.
Helicobacter pylori	-	172/313 (55%)	Tomb *et al.*, 1997.
Mycoplasma pneumoniae	-	174/324 (54%)	Himmerreich *et al.*, 1996.
Mycoplasma genitalium	-	172/312 (55%)	Fraser *et al.*, 1995.
Eukaryota			
Arabidopsis thaliana	-	182/319 (57%)	Genbank, 2583127.
Saccharomyces cerevisiae	-	172/341 (50%)	Simon *et al.*, 1994.
Caenorhabditis elegans	-	168/331 (51%)	Wilson *et al.*, In press.
Archaebacteria			
Methanobacterium thermoautotrophicum	-	174/319 (55%)	Smith *et al.*, 1997.
Methanococcus jannaschii	-	166/324 (51%)	Bult *et al.*, 1996.
Archaeoglobus fulgidus	-	158/319 (50%)	Klenk *et al.*, 1997.

Figure 1.1.   Results of an amino acid homology search in the SwissProt and Genbank

databases utilizing the GAPPED BLAST method (dashes are introduced to maximize

alignments).   Organism abbreviations are defined in Table 1.1.   Identical amino acids are

darkly shaded, whereas similar amino acids are lightly shaded.   The numbers represent the

amino acid position.

P.haem, H.infl, E.coli, B.subt, A.aeol, M.tube, Syne.sp., B.burg, M.lepr, A.thal, H.pylo, M.pneu, M.geni, S.cere, M.ther, C.eleg, M.jann, A.fulg — multiple sequence alignment (three blocks).

Block 1

Name	Start		End
P.haem	1	MR...D-EDKG-----...HADYGGVV...DHIRK...-...QEA	62
H.infl	1	MK...U-BEKR-----...HADYGGVV...RDHIRK...-...KAA	62
E.coli	1	MRV...I...-DERK-----...AD...GVKL--...ADYGGVV...RDHVRK...-...QAA	62
B.subt	7	...AA...VK-NGKE--IIS.VVAS...ES--...KRFGG...I...SAREHTRNI...-IFDR.	68
A.aeol	1	...AV...AL...-DQK...--VLG.VIL...AVV--...SPFGG.VV...SAREHTRNI...-IFDR.	62
M.tube	4	V...AR-LDPDGTVT.LADEVA.SV.E--.VRFGG.VV...RAHLEALG.-AMRRA	65
Syne.sp.	4	...A...AV...VN--NRN-----VCS.VVSS...QT--.QIFGG.VV...RQH.LINT-CLDQA	62
B.burg	1	MKV...DCC...VVE-NGIH----ILS.IKLN.TE---.KK.YGIV...I...LHTEAIMS-VCIKA	61
M.lepr	11	...GAACLDDY.-TVT.LADEVA.SV.E--QARFGG...I...RAHLEALG.-TERCA	74
A.thal	86	V...DAA.VVS-PFNH----.SSS---CRAEL--LVQ.GGVA.KO.EEAHSRVIDK-VVQDA	142
H.pylo	2	...DSSL.LTRIS.AQ----.I.HFKIS.EKH--.SSYGG...L.-AEN.S-.LLER	61
M.pneu	7	...T...DSIGVIT-.SK-----VQ.HIVL.SAKL--.BQTGG...VVA.SHEQNL.-----KA	61
M.geni	7	V...T.D...LSI.VI-DQ.-----IKS.IVI.SANL--.VKTGG...I.ARCHEQNLF-----KA	61
S.cere	34	KV.A...DAC.SVL.RFS.S----AAP.V.ANLK.TLDSID.GG.II.TK.HIH.QARIG.-.TERA	97
M.ther	1	.LC...GTAEK.G.V.-.AGN----VLSLRGKPL.P-----EK.GIH.RE.AEH.AKWIPR-.LABA	59
C.eleg	23	VKV...DA...VN-.KRE----ILSSER.TERAI--QRQQ.GIN.SVCALO.-REN.R.GEKC	84
M.jann	5	.IC...GTAEK.G.VT-S.GE----VLF.KT---.MY--KPPKQGIN.RE.A-D.HAE.F.K.KEA	63
A.fulg	1	.IA...GTAWSLSIGVV.-.E-.----VI.LENDPY.PK--E---GGIH.RE.QHHSERLPS-.LSRV	58

Block 2

Name	Start		End
P.haem	63	...	125
H.infl	63	.E-...TAK-...SG...-...I...DD--.S--	125
E.coli	63	.A-.SG.TAK-...AV...VG...F.D...IP...-...D--.P--	125
B.subt	69	FR-K.GMTY.-...A.VV.G...G.VNA.KA.SF.Y.IELV...IA.IY.NR.V.--DI--	131
A.aeol	63	.A-.SRINLE-E...SF.LT...ILS.V.VAF.KA...EYRK.LVP...L.IY.VF.K--KV--	125
M.tube	68	.A-A.G.KQP-.I--V.A.I.G...A...VAA.KAYSA...G.FYA.NLG.A.ADV.---HG--	125
Syne.sp.	63	.Q-ASG.GWP-E.EA.V.VA.A...M...VAA.KT.MVHQK.F...L.IY.SY.SQ--PD--	125
B.burg	62	.A-K.TKI.-E...L.V.SR...IAS.I.LNF.KG.ISLKK.IICID.IL.Y.LMHS--KI--	124
M.lepr	75	.A-A.G.TGBAKP.VV.A.I.G...A...VAA.KAYSA...G.FYA.NLG.A.ADV.---HG--	137
A.thal	143	.D-K...TEK-.LSAV.V.I.G.SLC.R.VRK.RV.GNESL.IV...A.A.VAR.V.--QE--	205
H.pylo	62	I.ISL.KDF.-K.KA.I.NQ...SVT.IEG.MM.KA.SLSL.L.LILED.LR.VYSLFIN.--KQ--	125
M.pneu	62	.Q-QSGVVLE-Q.TH...A.N...P.C.H.A.F...SFLLDK.LEPIN.LYA.IFSALIDQDI.Q--	126
M.geni	62	IR-DL.FEIR-.LSH...ACN...A.C.H.A.F...SFLLDK.LEPIN.LYA.IFSCLIDQDL.K--	126
S.cere	98	.I-.SNAREG--.L.CV.RG.AMPGS.SG.LD.KG.V...K.LIG...L.LI.RMGT-.GK.	161
M.ther	61	CR-D.GVELG-E.GL.SF.SRG...GP.RTVA.AA.T.LSLD.IV.N.CI.IE.GR.TT--G.--	122
C.eleg	85	.N-D.GTS.K-.L.AV.V.VT...I...KE.ISA.IG.KKHRL.LIP...RA.A.SIL.VD--DS--	147
M.jann	64	FE-VV.D--KN-E.DL...FSQ.G...GP.ER.TA.VA.T.SLTLKK.IIG.N.CIA.IE.GK.TT--EA--	124
A.fulg	59	FE-K.D--KN-SI.VV.FSQGPGMGPCLRVVA.A.L...IKLEK.L.VG.N.CLA.VE.GRWQT--G.--	119

Block 3

Name	Start		End
P.haem	126	P.E.EV.A...BGG...DV...G...EL...T.GK...-------DY--...NSK	185
H.infl	126	.H.REV...VSGG...R.D...K.V-I...AR...------DY--...G.A.LSR	185
E.coli	126	P...VS.T.I...AR...------.-...G.PLLS.K	185
B.subt	131	-V..AL.VVSGG...YMKEH.SF.V-I.TL.Y.VARTM...------PY--.G.PQIDK	190
A.aeol	126	-.YP.L.I.G.D.YL.RDF.R.DF-.GTL.V.Y.VAG.M...------G.--.G.PIIDR	184
M.tube	126	.LPEC...VSGG...LH.RSL.EPIIE.STV.Y.VARS...------G.--.G.K.LD.	188
Syne.sp.	126	LQP.ELC...VSGG...IH.K.C.D.RQ-.TTR.VARS.D.------G.--.G.P.IDR	185
B.burg	125	-.YP.IS.LSG...LIA.QKNFDDV.I-.RTL.C...VA.HYDM------GF--.G.PNIEQ	183
M.lepr	139	.LPEC...VSGG...LQ.RSL.APIVE.STV.Y.VARS...------G.--.G.KVLD.	198
A.thal	206	LS.PEM...GGG...NL.LAHKL.G.TQ-.TTV.KTAR.W...------.MHRSG.P.VEE	267
H.pylo	126	TCM.LSV...VSGG...SLILEARDYENIKI-VAT.L.SF...S.VSK.M.D.------G.--.G.PIVEK	185
M.pneu	127	LKL.ALGEVVSGG...AIYLIKSLFDL.S-IA.TS...I.VY.VGRAM.F------PY--.G.PQLD.	186
M.geni	127	LQL.ALGEV.SGG...AIYL.KSFYEL.S-IA.TS...I.VY.IGRAM.F------DY--.G.SKID.	186
S.cere	162	.O.P.FVS.VB.GG...F.LSRAIDDH.I-.CDT.I.V.DSL.CGRE...F----------K.TMIAR	219
M.ther	123	SD.---VS.YVSGGN.V.IAFNE-.R.RV-E.TL.I.V.NML.QFARESG.------G.H--.GG.PVIE.	179
C.eleg	148	.REP.SAV.LSGG.ALISVAED.EKFK.-Y.Q.VSGSP.CI.VARQ.DLGSEF.G.--HV.A.VE.	213
M.jann	125	-.DP-LT.YVSGGN.V.IAYVSK-K.RV-E.TL.I.V.NCL.QFARYVN.------PH--.GG.PYIEE	181
A.fulg	120	RKP--VS.YVSGGNS.VIARR.G-NR.RV-E.TL.IGI.N.L.KLARHM.------KH--.GG.PKIEE	176

```
P.haem   186 L---------...-...------------...----...------------      230
H.infl   186 L---------A-...-...-A...------------VNQAIKNE          230
E.coli   186 M---------...-AQ...AGR-EVE...-...------------R-----...  225
B.subt   191 L---------...-...KGNDN---IPL...AW----LEEG-----SYN---...S-AVI LH- ASQK  233
A.aeol   185 L---------...-KEG-KKL-YPL K L----MEEG-----N N---...-AIL LL KEK VRK  228
M.tube   189 L---------...-RT DRDA-IV...RG...----SG-...---AD RYA...-AV RYVESHAADP  235
Syne.sp. 186 A---------...-KQ D GT-...KL EGK----ISL QGGYHPY...---S...-AMLRLTQELKQSSA  235
B.burg   184 I---------S-KNGDE...T-EQ...VTT----FKKK---ENWY...---Y...-ACIHQLEKFKSKD-  229
M.lepr   199 L---------...-RT DRDA-IV...RG...----...G-...-----AD LNA...-AV RYVESH-----  240
A.thal   268 D---------...-LE GDAKS-VK...V...----KYHK-----DCN---YA...-QVRLA EA K-----  303
H.pylo   213 L---------...LDYRH...EPLM...I L----KNS...-N A---...------AVRLEVEK A  229
M.pneu   187 L---------F-QPELVKS-HY...F...S-----...-K----WTK---Y...S-QCFTK -Q R RK  229
M.geni   187 L---------F-NKELVKP-HY...K...S-----...-K----WTK---Y...S-QCL K QISANKT  229
S.cere   220 EMEKFINQDIN-DQDFALK-LEM...S LKNSASK N-----M S---...AFI ALRT LT LGKT IQ  278
M.ther   180 L---------...-LKAS-EY-IEL YSV-----K-----...-...I...L -A LRKME GASLE-  219
C.eleg   214 ...---------SRASADGH-LRY...IFL----PNV...-----KAN---MN DQI G-SYL LLE-R RK S  258
M.jann   182 L---------...-RK...-----------KKL----V L...-YTVK...---IA...L -A MRAYD -----...  216
A.fulg   177 L---------...-KK QKYH-...L---...Y----VVK-----...---...MV -A QRLFDSGVRME-  216

P.haem   231 -...------------...----...------------...-----...------------  286
H.infl   231 -...I...---A...YI...DA...LA...K-----E-----...I...K...ET H  286
E.coli   226 --T D...R--A...RA...EDAV...LM...D-----...-----F...RT...K...  281
B.subt   234 -Q--...IAP--E LSAS...NSV...VLVT TA...AK-----E------YDV QVLL...A...RG...A EK  287
A.aeol   229 -...---------...YS...ET...EILZE SLW MK-----K-----...I...VV...SR...EVFKK  277
M.tube   236 -----FRT--A...AG...E AVA VLTMKAV AT-----A-----L VST LI...A...SR...ELATQ  287
Syne.sp. 236 -P PVD-----...LAS...DT ARSLTK TIQCVL-----D-----H LTTITVG...A...SR...YH QT  288
B.burg   230 ---NPT...--NN...AS...K AFENLITPL...IK-----D-----...QINK...I...AS...LY...----...  279
M.lepr   240 ---PDALP--A...AG...E AVA VLTMNAV AT-----G-----L VST LIV...A...SR...ELA Q  294
A.thal   304 ---EIRNR--A...AS...RVA LHLEEK E IDWALELE-----PSI HM IS...AS...YV LR NN  366
H.pylo   230 PN N A I...--QK GYH...S A IEHLIQQT...YFK-----I-----KRP IFGIV...A Q LA...KAFEN  285
M.pneu   230 -GFNP...HDWNEF SN...ATI IDHYINHV D IQ-----...-----HQPQM LLG...YE...----...  278
M.geni   230 -RI...S-----EL ASN...ATI IDHYIDHV N IK-----K-----FAP M LVG...SY----...SN  278
S.cere   278 -...P REI--RS...YQV ESY...H...INK L HV K-----SQPEKFKNVREF CS...S QRE TK ET  338
M.ther   220 ----------...L YSI...ETAFSMLVEVTE...A-----Y-----...EKNQVLLC...AV RR R DM R...  267
C.eleg   259 -...TSIDIP---...FCASL NT ARH...SS LHIFF...-----SLSEQEKLP...Q...IG...A...QYIFGAISK  316
M.jann   217 -...RL...------...CYSL EYAFSMLTEITE...A-----H-----...NKGEVMLV...A...NR...EM KA  269
A.fulg   217 ----------...V...FS...ETAFAMLTEVTE...A-----Y-----LDLNEVLLV...A...KR QEM RI  264

P.haem   287 ...------------...------------...------------  343
H.infl   287 L...QN...G...---...-...------------QGQHS...A D-...-...  342
E.coli   282 ...RR...---...------------...-AV

...------AGATA GV...-...R...PLA...  335
B.subt   288 EFAQHE...ITLV----I...PLALC...-...AA TIAFE------KGIRGAYDMN-GQ GLEL...  339
A.aeol   278 ASQEYGF-...LY----I...H SLS...AL...A...ME F...------RGVVAP DVN-PQ NIPLE FGR...  327
M.tube   288 RCGEAGR-TLR----I...S RL...------------  312
Syne.sp. 289 AAQEHQL-Q...----F...PLK...G...A...CAAADHFQ------...GDRSP TLG-...QS LSVEQV  341
B.burg   280 KID...I-QTY----...PLDL...------GL...  308
M.lepr   295 RCAAAGL-MLR----I...G R...------...  319
A.thal   367 IVENKNL-KLV----C...SLC...V MVAW...LEHFRVGRYDPAT...PE YVYD-LR...PLG...  427
H.pylo   286 LCDAFDC-KLV----LAPLE...CS...A...  309
M.pneu   279 QVTQ...QL-PYL----IAPLKYTS...A...GF  308
M.geni   279 RIST...NL-PFL----IADSKYTS...A...GFYAS LIN------G DK  314
S.cere   339 ELGT...NS-TS...FNFY...PMDLC S...SI...GWA...I...------------VS...DIC-PIRQ PLND...  396
M.ther   268 MCQEHHV-...FH----M...PEYCG...A...DAWLGO VY...  300
C.eleg   319 LSAAHNV-TTI----KVLLSLC...AEM...YSG LEM V------...  352
M.jann   270 MCEGQNV-DFY----V...PKE CG...A...PAWLGL MH...------...  303
A.fulg   265 MCEDRGA-KFY----V...PKELAG...A...MIA...GL MY...------HGH...PVEK...Y R DFRIEDV  319
```

region of interest is the conserved histidine pair at positions 111 and 115 in Gcp. These histidine residues were thought to be part of a potential zinc ion binding site (Abdullah *et al.*, 1991), since the sequence HXXXH is common to all twelve classes of Zn^{2+} metalloproteases where at least one of the X residues is a glutamic acid (Hooper, 1994). Through a Ψ-BLAST analysis, which is a powerful database search function for the detection of super-families of proteins, it was found that the GHXXA motif at positions 114-118 is common to the polyketide synthase family. The polyketide synthase complex is a multi-enzyme system found in fungi, that synthesizes a wide variety of antibiotics. One of its many catalytic domains is homologous with the glycoprotease, especially at the Gcp N-terminal region. The super-family also shows a number of conserved proline and glycine residues. These similarities constitute a cluster of orthologous groups (COG) and suggest a common tertiary structure for the glycoprotease and its homologues.

IV. Gcp expressed as a recombinant protein.

The gene encoding Gcp was first identified in *E. coli* HB101/pPH1 from a clone bank of *P. haemolytica* that was screened for glycoprotease activity (Abdullah *et al.*, 1991). A series of subclones were constructed, each containing selected fragments of the pPH1 insert, and analyzed by maxicell analysis for protein expression. Subclones in the high expression vector pTTQ19, pPH1.1 and pPH1.1ΔE, were found to produce plasmid-encoded protein identical in size to the pPH1 clone. The DNA fragment from pPH1.1 was then sequenced and the coding region for the *gcp* gene was determined. The *gcp* gene was subcloned into the high expression vector pTTQ18 to create the pTTQ18 derivative pGP1. Glycoprotease activity was detected in *E. coli* culture lysates carrying pPH1 and pGP1 (Abdullah *et al.*, 1991). However, the production of active rGcp enzyme from

these plasmid constructs was not consistent. The recombinant glycoprotease expressed from *E. coli* HB101/pGP1 was localized to the periplasm. Lysates from *E. coli* cultures carrying pGP1 were used to raise polyclonal antibodies in rabbits that neutralized the wild type *P. haemolytica* Gcp (Abdullah *et al.*, 1991). The recombinant protein was also used to produce three monoclonal antibodies, mAbs 48, 91, and 110, that recognize and neutralize the authentic glycoprotease (Lo *et al.*, 1994).

A system was developed for the large scale production of rGcp, secreted to the culture supernatant of *E. coli*, to investigate its antigenic properties in animals. The plasmid, pGcp-F, was constructed by creating an in-frame fusion of the *gcp* gene with the *E. coli* hemolysin secretion signal *hlyA*. The plasmid was transformed into *E. coli* DH1, with the secretion functions HlyB and HlyD supplied *in trans* on a separate compatible plasmid, pWAM716. The recombinant fusion protein, rGcp-F, was successfully secreted into the culture supernatant of *E. coli* DH1/pGcp-F/pWAM716 and was recognized by mAb110. The relatively homogenous preparation did not display any biological activity. Recently, it has been determined that the fusion protein resulted in a C-terminal truncation of Gcp lacking 36 amino acids. The efficacy of the Presponse™ vaccine was increased when supplemented with both rGcp-F and recombinant Lkt compared with the vaccine alone, or supplemented with only one of the two components. Calves challenged with rGcp-F alone were protected from experimental challenge at a vaccine efficacy comparable to Presponse™ (Watt, 1997). rGcp-F was also used as a source of antigen in the production of polyclonal anti-Gcp antibodies in rabbits.

The heterologous rGcp expressed in *E. coli* HB101/pPH1.1ΔE was localized to the periplasm, where it accumulates as a high molecular mass, disulphide-linked aggregate;

the chaperone protein thioredoxin was used as an N-terminal carrier for the glycoprotease in order to circumvent this problem. Thioredoxin is a protein that can catalyze the reduction and isomerization of disulphide bonds. The fusion protein, rTRXGCP, was a monomeric, cytoplasmic protein that was found to be biologically active after enterokinase cleavage of the carrier, or after partial purification by DEAE ion exchange chromatography (Watt *et al.*, 1997a). It is not known whether enzymatic activity was dependent on the removal of the N-terminal fusion partner, since both the fusion protein and a degradation product of rTRXGCP of similar size to rGcp were found in the enzymatically-active chromatographically-purified samples.

The molecular chaperones, protein disulphide isomerase (PDI) and the DnaK/DnaJ/GrpE system were also used successfully for *in vitro* refolding of rGcp. The recombinant protein, rGcp, was isolated from the periplasm fraction of *E. coli* HB101/pPH1.1ΔE cultures by osmotic shock (Watt *et al.*, 1997b). Mammalian PDI is a homo-dimeric protein that, like bacterial thioredoxin, catalyzes thiol/disulphide interchange reactions *in vivo* and *in vitro*. Each PDI monomer contains duplicate domains with strong homology with thioredoxin (Freedman *et al.*, 1989). PDI was able to refold denatured and DTT-reduced rGcp under appropriate redox conditions (Watt *et al.*, 1997b). The use of chaperones for *in vitro* refolding has been well documented (Mendoza *et al.*, 1991; Schröder *et al.*, 1993; Ziemienowicz *et al.*, 1993; Weissman and Kim, 1993). Incubation of denatured reduced rGcp in the presence of the DnaK, DnaJ, and GrpE chaperones resulted in a partial recovery of biological activity (Watt *et al.*, 1997b). The chaperone DnaJ has been proposed to act like thioredoxin, PDI, and DsbA, in the

isomerization of substrate protein dithiol/disulphide groups (de Crouy-Chanel *et al.*, 1995).

The recovery of glycoprotease activity from the aforementioned recombinants was not readily duplicated. Furthermore, the total and specific activities observed were significantly lower than wild type activity. However, the results of these studies confirmed that the *gcp* gene product, under appropriate conditions, could give rise to glycoprotease activity. The lack of substantial enzymatic activity was apparently due to a combination of factors; the major impediment appearing to be the misfolding of the heterologous protein when expressed in *E. coli*. Another important factor was the formation of disulphide-linked rGcp aggregates during partial secretion into the more aerobic environment of the host cell periplasm. The reactivation of rGcp by agents with disulphide bond reducing activity indicated that the oxidation of cysteine thiols in *E. coli* abrogates biological activity. A minor influence on the lack of observable activity was the inhibition of the enzyme by components of the host cell lysates. The production of an active recombinant glycoprotease will depend on strategies to overcome these obstacles.

C. Recombinant Proteins.

I. The expression of recombinant proteins in *E. coli*.

The use of *E. coli* as a host for the production of recombinant proteins is widespread because of the development of genetically amenable host strains. These systems are well-characterized and the bacteria can be grown quickly and inexpensively. Almost $5 billion of recombinant protein products (including insulin, human growth hormone, interferons, and many others) were produced in *E. coli* in 1993 (Swartz, 1996). Many heterologous proteins can be produced in a biologically active form. Variables such

as the growth temperature, aeration, growth media, and choice of host strain, all influence the production of recombinant proteins, however, a number of heterologous proteins do not fold correctly in high-level expression systems, and some form aggregates known as inclusion bodies, or are degraded by cellular proteases. Another problem with heterologous protein expression in *E. coli* is the absence of post-translational modifications, such as glycosylation, which are required for many eukaryotic proteins. An increasingly popular method of producing soluble, biologically active products involves the *in vivo* co-expression of folding catalysts, or the *in vitro* refolding of the recombinant protein (Thomas *et al.*, 1997).

The native conformation of many proteins will form spontaneously; that is, it is a thermodynamically favoured state, and is specified by the protein's amino acid sequence (Anfinsen, 1973). First, the nascent, unfolded polypeptide partially folds into native secondary structure motifs, in a kinetically favourable intermediate known as the molten globule. The molten globule is an incompletely folded structure with some secondary structure, loosely packed tertiary structure, and possibly some exposed hydrophobic regions. The unproductive aggregation of the highly populated molten globule intermediate state may result in inclusion body formation, to give an insoluble aggregate of unfolded proteins. The refolding of the molten globule into the native conformation is thought to be the rate-limiting step in the process of folding. However, it is now known that many proteins require the presence of molecular chaperones to increase the rate of formation of a correctly folded tertiary structure (Ellis and Hemmingsen, 1989). Chaperone proteins play an important role in catalyzing the folding of the native conformation from the nascent peptide or the molten globule intermediate, by stabilizing

24

the unfolded protein, preventing unproductive interactions and facilitating productive folding (Figure 1.2).

The heat shock chaperones (Hsp70 and cofactors), DnaK, DnaJ, and GrpE operate on linear, nascent polypeptides in a cyclical ATP-dependent manner. A putative mechanism of their action involves the binding of a nascent protein to DnaK or DnaJ, chaperone proteins that interact with one another to maintain the peptide in an unfolded or partially folded state. The DnaK/DnaJ complex is thought to shield hydrophobic stretches, preventing unproductive aggregation of the polypeptide, or the formation of misfolded intermediates. The polypeptide is either released in a folded or partially folded state after GrpE-mediated ATP hydrolysis, or it is passed on to the GroEL/S chaperone system (Szabo *et al.*, 1994). The GroEL/S complex (Hsp60) facilitates the isomerization of partially folded polypeptides in an ATP-dependent manner. The compact folding intermediates of up to approximately 90 kDa are made competent for folding by binding within the cavity of two stacked, heptameric GroEL rings to which GroES binds in a 1:1 stoichiometry. GroES regulates the ATPase activity of the complex and permits the dissociation of the folding substrate from the GroEL/S complex. The process of GroEL/S folding requires multiple cycles of protein release and binding (Fedorov and Baldwin, 1997). There are other chaperones that influence protein folding in *E. coli* who may work in conjunction with, or independently of the Hsp70, and Hsp60 proteins. These include the highly conserved Hsp90 family of chaperones. The mechanisms of action of these other chaperones in the *in vivo* folding of proteins are poorly defined.

Two of the rate-limiting steps in protein folding are the *cis-trans* isomerization of the peptide bond adjacent to proline residues, and the proper configuration of disulfide

bonds. The first step is catalyzed by a ubiquitous family of enzymes known as the peptidyl prolyl *cis-trans* isomerases (PPI). It has been shown recently that the *E. coli* ribosome associated protein, Trigger Factor, has PPI activity (Hesterkamp *et al.*, 1996). The attainment of a properly disulphide-linked native conformation is contingent on the subcellular localization of the protein. The *E. coli* cytoplasm is a reducing redox environment compared with the periplasmic space. The enzyme thioredoxin is found in the cytoplasm of *E. coli* and is responsible for the reduction and isomerization of oxidized cysteine groups. Proteins that are secreted to the more oxidizing environment of the periplasm are subject to the action of the Dsb proteins (Figure 1.3). The Dsb proteins are responsible for disulphide bond formation in the periplasm of *E. coli*. The first protein, DsbA thiol:disulphide oxidase, oxidizes reduced proteins entering the periplasmic space. The oxidizing potential of DsbA is regenerated by the membrane bound DsbB protein. Another protein in the periplasm that can function in a manner similar to DsbA is DsbC. This protein can function independently of the DsbA/DsbB system in the oxidation or isomerization of periplasmic proteins. Similar to the relationship between DsbB and DsbA, DsbC is regenerated by a membrane bound thiol:disulphide reductase, DsbD. Two other periplasmic Dsb proteins, DsbE and DsbG, have been recently identified. The former exhibits reductase activity whereas the function of the latter has not been well elucidated.

Figure 1.2   A model of the DnaK/DnaJ/GrpE and GroEL/GroES chaperone-mediated

folding of proteins in *E. coli* (Thomas *et al.*, 1997).   DnaK and DnaJ bind to the nascent

polypeptide, maintaining the protein in a folding-competent state.   GrpE causes the release

of the protein in an ATP-dependent manner, resulting in the folding of the protein into the

native conformation, or for further cycles of chaperone-assisted refolding.   The GroES/L

system facilitates the ATP-dependent isomerization of a partially folded intermediate.

Trigger factor (TF) is a ribosome associated protein that possesses peptidyl-prolyl-

cis/trans isomerase activity that may act upon the nascent polypeptide.   Proteins that do

not fold productively eventually accumulate as an aggregate, or are targeted for

proteolysis.

3'

5'

DnaJ

DnaK

TF

Proteolytic
Products

DnaK

DnaJ

GrpE

+ ATP

ADP

Native
conformation

Aggregation

GroES

GroEL

Proteolytic
Products
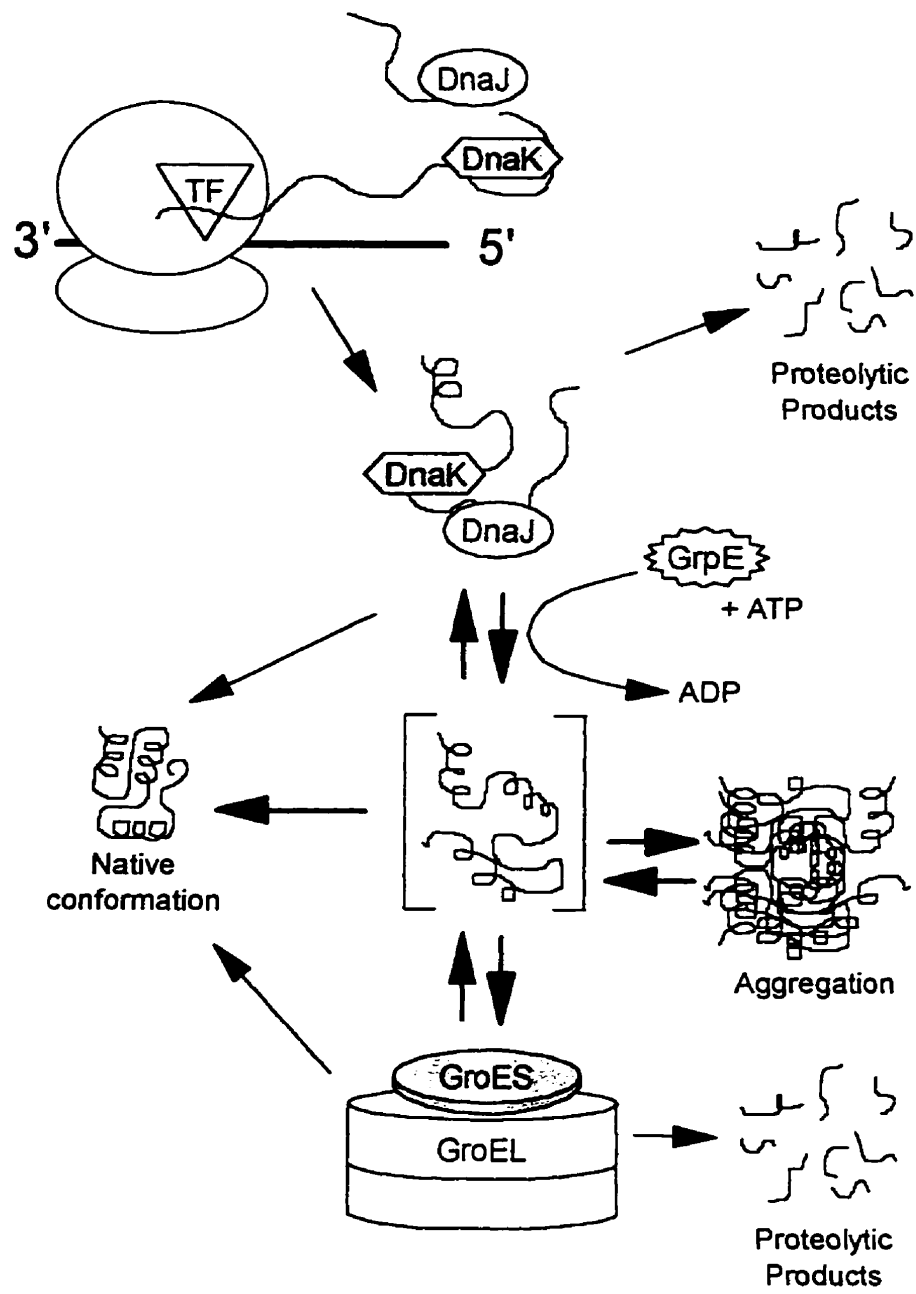
Figure 1.3.   A model for the formation and isomerization of disulphide bonds in a newly

synthesized polypeptide, in the periplasm of *E. coli* (Missiakas and Raina, 1997).   The

redox state of all enzymes is given by the suffix –OX for oxidized, or –RED for reduced.

The ovals represent select members of the *dsb* operon.   The enzymes thioredoxin (TrxA)

and thioredoxin reductase (TrxB), are represented by the rectangles.

Periplasm

Cytoplasm

30
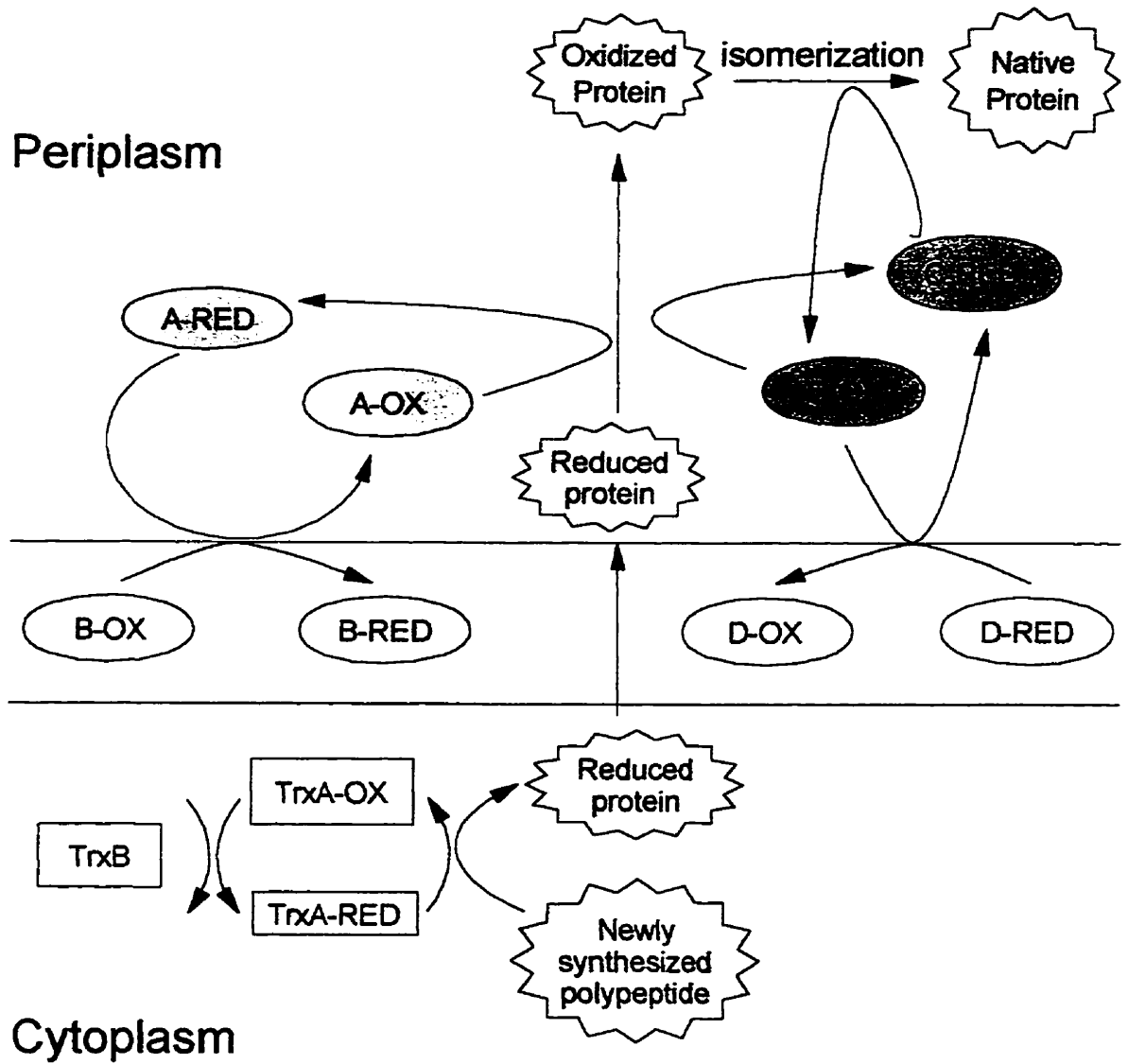
II. Fusion proteins.

Another means of manipulating the environment and conditions of heterologous protein expression involves the creation of fusion proteins. The benefits of producing fusion proteins in *E. coli* include: increasing the solubility of the product; engineering the cellular localization of the product; protection of the heterologous proteins from proteolysis; the addition of a carrier as a reporter for protein expression; the production of bifunctional enzymes; and the addition of tags to facilitate purification (Uhlén *et al*, 1992).

The heterologous protein of interest can be fused to either the N- or C-terminal of the carrier protein. C-terminal fusions are most common since the expression level is more predictable and the transcriptional and translational start sequences do not have to be engineered into the 5' end. Dual fusion systems are also used to combine the advantages of various carrier proteins, which could include a specific secretion signal peptide in addition to an affinity tag, or the use of a two affinity tags. Dual fusion proteins can also be used for the protection, at both the N- and C-termini, of heterologous proteins that are susceptible to proteolysis (Uhlén and Moks, 1990).

The major application of fusion proteins is for the one-step purification of recombinant proteins. Several systems have been developed that involve the binding of the carrier to a ligand on an affinity matrix in order to specifically isolate the recombinant protein (Table 1.2). Contaminating proteins are washed away from the bound fusion protein, thereby allowing for the recovery of a homogenous product. The carrier protein can often be removed from the heterologous recombinant protein by the use of a site-specific protease recognition site engineered in between the two. A recent advance in affinity purification includes the incorporation of a self-splicing protease as a part of the

31

carrier protein itself as the means of eluting the recombinant protein (Chong *et al.*, 1997). This circumvents the contamination of the purified sample with the carrier protein and cleavage factor. The pGEX family of vectors has been widely used for the production of proteins fused to the enzyme glutathione-*S*-transferase (GST) (Smith and Johnson, 1988). GST fusion proteins can be selectively bound on glutathione-Sepharose affinity columns, allowing for the one-step purification of recombinant proteins and the removal of the carrier by a site-specific protease. The construction of a GST fusion protein comprising *P. haemolytica O*-sialoglycoprotein endopeptidase was a main objective of this thesis.

Table 1.2. Common fusion partners that are used in the affinity purification of fusion proteins (Uhlén *et al.*, 1992).

Fusion Partner	Size (kDa)	Ligand	Elution condition
Glutathione-S-transferase	28	Glutathione	Reduced glutathione
Maltose binding peptide	40	Starch	maltose
Histidine tail	1-7	IMAC	imidazole
Histidine-patch thioredoxin	12	IMAC	imidazole
Strep-tag	13	Streptavidin	iminobiotin
PinPoint™	13	Streptavidin	biotin
Flag™ peptide	2-5	Specific antibody	low calcium
ZZ	14	IgG	low pH
β-galactosidase	116	APTG	

Abbreviations: IMAC: immobilized metal ion affinity column; ZZ: IgG-binding fragment based on staphylococcal protein A; IgG: Immunoglobulin G; APTG: p-aminophenyl-β-D-thiogalactoside.

D. Research Objectives.

The objectives of this thesis are as follows:

1) To produce a purified, biologically active recombinant glycoprotease as a fusion with glutathione-S-transferase.

2) To express the glycoprotease as an active recombinant in *P. haemolytica* by the use of a shuttle vector.

Chapter 2. General materials and methods.

All chemicals were purchased from Sigma Chemical Co. (St. Louis, MO) or Fischer Scientific Canada Ltd. (Nepean, ON) unless otherwise specified. The compositions of various buffers and reagents are described in Appendix A.

I. Bacterial strains and culture conditions.

The bacteria used were laboratory strains obtained from glycerol stocks stored at $-70°C$, or maintained on agar plates and stored at $4°C$ unless otherwise mentioned. The bacterial strains and plasmids are listed in Tables 2.1 and 2.2 respectively. The *E. coli* strains were maintained on Luria-Bertani (LB) plates and re-streaked monthly; whereas *P. haemolytica* strains were grown on brain-heart infusion (BHI) plates or blood agar plates and re-streaked weekly. All of the strains containing plasmids were selected by ampicillin (Amp) supplementation at either 100 µg/ml or 5 µg/ml for *E. coli* and *P. haemolytica* respectively. Glycerol stocks of all strains were prepared and maintained at $-70°C$ (Sambrook *et al.*, 1989).

II. Recombinant DNA techniques.

Enzymes used for DNA modification were obtained from Pharmacia Biotech (Quebec City, PQ), Gibco-BRL (Burlington, ON), or Boehringer Mannheim Canada (Laval, PQ). Plasmid purification was performed using the QIAprep™ spin miniprep kit (Qiagen Inc., Chatsworth, CA) which is a modification of the alkaline lysis protocol of Birnboim and Doly (1979). For the low-copy number plasmids, pNF2176 and its derivatives, chloramphenicol was added at 180 µg/ml to overnight cultures and incubated for another 3 h to increase the total yield. Following purification, the plasmids were stored at -20°C in TE buffer.

Table 2.1. Bacterial strains used in this study.

Strain	Genotype[1]	Special comments	Source/Reference
E. coli			
191	D10 (*dsbA*::Kan)	A strain deficient in periplasmic disulphide bond formation.	N. Martin.
BL21	F- *ompT* r$_B$-, m$_B$- *gal dcm*	A protease deficient strain recommended for GST fusion protein expression.	Pharmacia Biotech.
CA274	*trp*-49(am) *lacZ125*(am) λ- *relA1 spoT1*		D. Mangroo; Seong *et al.*, 1989.
CSR603	F- *uvrA6 recA1 phr*-1 *thi*-1 *thr*-1 *leuB6 lacY1 galK2 ara14 xyl15 mtl1 proA2 argE3 rpsL31 tsk-33 supE44 gyrA98* λ'	A strain used for maxicell analysis.	R.Y.C. Lo; Sancar *et al.*, 1979.
D10	*hfr met* λ' RNase$_{19}$'		N. Martin; Gesteland, 1966.
DH1	*supE44 hsdR17 recA1 endA1 gyrA46 thi relA1*		R.Y.C. Lo; Hanahan, 1983.
DH5α	DH1 (Δ*lacU169* (φ80 *lacZ*ΔM15))		R.Y.C. Lo; Hanahan, 1983.
HB101	*supE44 hsdS20*(r$_B$-, m$_B$-)	An *E. coli* B and K12 cross.	R.Y.C. Lo; Boyer and Roulland-Dussoix. 1969.
JM83	F- Δ(*pro-lac*) *phi80 lacZ*ΔM15 *ara rpsL thi* λ-		R.Y.C. Lo; Yanisch-Perron *et al.*, 1985.
JM101	*supE thi* Δ(*lac-proAB*) F' [*traD36 proAB*⁻ *lacI*q *lacZ*ΔM15]		D. Mangroo: Sambrook *et al.*, 1987.
JM109	*recA1 supE44 endA1 hsdR17 gyrA46 relA1 thi* Δ(*lac-proAB*)		R.Y.C. Lo; Yanisch-Perron *et al.*, 1985.
MV1190	JM101 [Δ(*srl-recA*)306::Tn10(*tet*r)]		D. Mangroo; Biorad.
Novablue	*endA1 hsdR17* (r$_{K12}$-m$_{K12}$+) *supE44 thi*-1 *recA1 gyrA46 relA1* F' [*proAB*⁻ *lacI*q *lacZ*ΔM15::Tn10]		D. Mangroo: Novagen.
Sure2	*mcrA endA1 supE44* F' [*proAB*⁺ *lacI*q *lacZ*ΔM15 Tn10(*tet*r)]		Stratagene, Inc.
TG1	JM101 (*hsd*Δ5)		D. Mangroo: Sambrook *et al.*, 1987.
XL1-blue	DH1 [*lac*⁻ F' [*proAB*⁺ *lacI*q *lacZ*ΔM15 Tn10(*tet*r)]]		D. Mangroo: Sambrook *et al.*, 1987.
P. haemolytica			
A1	Wild type		ATCC# 43270.
A11	Wild type		R.Y.C. Lo.

1 -- (Sambrook *et al.*, 1989)

Table 2.2. Plasmids used in this study.

Plasmids	Size (bp)	Comments	Reference/Source
pBR322	4361	Low copy number cloning vector.	Lab stocks.
pTTQ19	4562	High level expression vector.	Stark, 1987; Amersham
pPH1	~8300	A pBR322 derivative that contains the *gcp* gene; derived from a *P. haemolytica* clone bank.	Abdullah *et al.*, 1991
pPH1.1	~7800	A 3.3 kb insert from pPH1 subcloned into pTTQ19.	Abdullah *et al.*, 1991
pPH1.1ΔE	~6000	A subclone derived from pPH1.1 that produces rGcp; also a pTTQ19 derivative.	Abdullah *et al.*, 1991
pGEX-4T-3	4968	GST fusion vector.	Smith and Johnson, 1988.; Pharmacia Biotech
pGGCP	6100	A pGEX-4T-3 construct that produces the rGgcp fusion protein.	This study.
pNF2176	5162	An *E. coli* - *P. haemolytica* shuttle vector.	Fedorova and Highlander, 1997a.
pNFG1	6364	Subclone of pNF2176 containing the *gcp* insert.	This study.
pNFG2	6364	Similar to above, but the *gcp* gene is in the opposite orientation.	This study.
pNFGG1	7060	Subclone of pNF2176 containing the *ggcp* insert.	This study.
pNFGG2	7060	Similar to above, but the *ggcp* gene is in the opposite orientation.	This study.

Purified plasmid DNA was digested for one hour at 37°C using the appropriate restriction enzyme and buffer conditions, according to the manufacturer's instructions. The digestions were terminated by one of two means: In the case of DNA to be ligated, the restriction enzyme was heat-inactivated at either 65°C or 85°C; alternatively, the reaction was halted by the addition of a half volume of tracking dye (Appendix A) to the reaction mixture, before analysis by agarose gel electrophoresis.

The gels comprised from 0.8-1.0% agarose in TAE buffer depending on the expected size of the DNA fragments. The gels were subjected to electrophoresis in a horizontal gel apparatus (Tyler Research, Edmonton, AB) in TAE running buffer containing 1 µg/ml ethidium bromide. The samples were run for 3 hours at 40V-55V or overnight at 10V-15V at room temperature. The DNA bands were then visualized and photographed under a UV lamp using Molecular Analyst™ software (BioRad Laboratories).

Digested DNA samples to be used in ligation reactions were heat-inactivated and, in the case of single-cut vector, treated with calf intestinal alkaline phosphatase for 30 min at 37°C to prevent re-ligation of the cohesive ends. This was followed by the purification of the DNA with GeneClean™ (Bio/Can Scientific, Mississauga, ON), and isolated DNA was resuspended in a minimal volume of sterile distilled water. The purified vector and insert DNA were ligated with T4 ligase (Gibco-BRL) combined at relative molar concentrations of approximately 1:1 and 1:4 in a reaction volume of 10 µl. The reaction was for 4 h at room temperature or overnight at 14°C. The ligation mixture was then used to transform $CaCl_2$–treated competent E. coli (usually DH5α), plated onto LB + Amp plates and incubated overnight at 37°C.

The cells were made competent by treating subcultured *E. coli* with successive washings of cold sterile 50 mM $CaCl_2$ by a modified version of the protocol of Cohen *et al.* (1972). These cells were either stored at 4°C and used within three days, or stored for longer periods at -70°C after the addition of an equal volume of 50% glycerol. The competent cells were transformed by incubating a 100 μl aliquot with ~0.1 μg of plasmid DNA for 30-60 minutes on ice. The cells were warmed to 42°C for one minute before the addition of 50 μl of prewarmed LB broth. The cells were incubated for at least one hour at 37°C to allow for plasmid-borne β-lactamase production before they were plated onto selective media.

Single colonies of prospective transformants were re-patched to exclude satellite colonies and screened according to a modified procedure first described by Barnes (1977). A toothpick stab of the patched colony was resuspended in 10 μl TE, followed by the addition of 20 μl SRL buffer. A 20 μl aliquot of this solution was loaded onto a 0.8% agarose gel containing 0.2% SDS in TB buffer. The sample was equilibrated for 5-30 min in the gel wells to allow for cell lysis before electrophoresis at 55V for 3 h. The samples were stained in 1 μg/ml ethidium bromide and visualized under UV light. Putative recombinant plasmids were selected relative to plasmids of a known size. The identity of the plasmids was confirmed later by restriction enzyme analysis.

III. Protein Techniques.

Protein concentrations were assayed by a dye-binding method using bovine serum albumin as a standard according to the procedure outlined in the BioRad kit (Bradford, 1976). The absorbance of the samples was measured at 595 nm in a Beckman DU-20 spectrophotometer, or in a Molecular Devices Thermomax™ microplate reader. Proteins

were analyzed by SDS-PAGE on 12% (w/v) polyacrylamide separating gels with a 5% stacking gel (Laemmli, 1970). The samples were diluted with sample buffer, boiled for 5 min, and subjected to electrophoresis for ~1 h at 160-170V in a BioRad Mini Protean™ apparatus. For non-reducing conditions, the sample buffer did not contain any β-mercaptoethanol. The gels were then stained with Coomassie Blue, or by silver-staining (Bollag and Edelstein, 1991; Wray *et al.*, 1981). The relative amount of protein on SDS-PAGE gels was determined with a BioRad Model 620 video densitometer.

For immunoblot detection, the proteins were transferred onto nitrocellulose or PVDF membranes in a BioRad transblot apparatus for either 3 hours at 60V or overnight at 30V (Towbin *et al.*, 1979). After transfer, the membrane was blocked for 30 min in 3% gelatin-TTBS, followed by incubation with the primary antibody for at least three hours. For the detection of Gcp, specific monoclonal antibodies, mAb48, mAb91, and mAb110, were used singly or in combination as the primary antibodies at a dilution of 1:100 in 1% gelatin in TTBS (Lo *et al.*, 1994). For some immunoblots, a goat anti-mouse alkaline phosphatase-conjugated secondary antibody (BioRad) was used at 1:2500 dilution in 1% gelatin in TTBS. The blot was then washed three times for 10 min with TTBS before a 2 h incubation with the secondary antibody. After a final wash, the chromophore was generated by equilibrating the membrane in NBT buffer for 5 min before the addition of the two substrates 5-bromo-4-chloro-3-indoyl phosphate (BCIP) and *p*-nitro blue tetrazolium (NBT) at a concentration of 0.16 mg/ml and 0.08 mg/ml respectively. This method was also used for dot immunoblotting on nitrocellulose membranes.

For greater sensitivity of antigen detection in immunoblots, a chemiluminescent detection kit (Kirkegaard & Perry Laboratories, Gaithersburg, MD) was employed

according to the protocol of Kricka (1991). This protocol uses milk powder in lieu of gelatin as a blocking reagent, with a reduced time for the washing and incubation steps. Briefly, the protein is immobilized on PVDF membrane and blocked for 1 h in the milk solution. The membrane is then incubated for 1 h with the primary antibody, followed by three 5 min washes with the blocking reagent. The membrane was then incubated with the secondary antibody for 1 h, washed three times for 5 min, and rinsed with water before the addition of the chemiluminescent substrate. The conversion of luminol to a luminescent excited intermediate state in the presence of H_2O_2, by horseradish peroxidase-conjugated secondary antibody (Isacsson and Wettermark, 1974) was detected on Cronex 4 X-ray film (DuPont, Wilmington, DE) with exposure for 5 to 20 min.

IV. Radio-iodination of glycophorin A

GPA was prepared from erythrocyte ghosts according to established protocols (Dodge et al., 1963; Segrest et al. 1979) (work of either W. Cladman or P. Jiang) and was stored at -70°C. The tyrosine residues of human glycophorin A (GPA) were radio-iodinated (Markwell et al. 1982). First, two washed Iodo-beads™ (Pierce Chemical, Rockford, IL) were incubated with 200 μCi of Na[^{125}I] (sp. activity 20 mCi/ml) (ICN Biochemicals, Irvine, CA) in 250 μl 50 mM HEPES (pH 7.4) (Calbiochem, San Diego, CA) buffer at room temperature. After five minutes, 2 mg GPA in 250 μl 50 mM HEPES (pH 7.4) buffer was added to the Iodo-beads and radio-iodination was carried out for thirty minutes. The [^{125}I]-GPA was then separated from the low molecular mass reactants, including [^{125}I]-iodide, on a 10 ml bed volume Sephadex G25 desalting column. Approximately twenty 0.5 ml fractions were collected and 5 μl aliquots of each were counted in a gamma counter to identify the excluded [^{125}I]-GPA peaks. The first four

41

radioactive fractions (which contain the [^{125}I]-GPA) were pooled and divided into twenty 100 μl aliquots, which were stored at −20°C until needed. The extent of radiolabeling of GPA was assayed by SDS-PAGE and autoradiography.

V. Activity assays for the _O_-sialoglycoprotein endopeptidase.

Two methods have been used to measure glycoprotease activity by the hydrolysis of human glycophorin A as the substrate. The first method involves the degradation of [^{125}I]-GPA monomer and dimer bands and the subsequent appearance of specific product bands as visualized by autoradiography (Mellors and Lo, 1995). In this method, 3.5 μg of the substrate was incubated with enzyme at 37°C for 15 min in 25 μl of 50 mM HEPES (pH 7.4). For enzyme samples containing low levels of glycoprotease activity, overnight incubations (16-18 h) were performed, in the presence of a protease inhibitor cocktail (1 μg/ml each of antipain, aprotinin, leupeptin, and pepstatin) to inhibit bacterial proteases, and the antibiotics (400 IU/ml penicillin, 400 μg/ml streptomycin) to prevent bacterial growth. These agents do not inhibit the glycoprotease enzyme. The assays were diluted with 12.5 μl sample buffer, boiled for five minutes, and 10 μl aliquots were analyzed by SDS-PAGE. Another 10 μl aliquot was counted for total [^{125}I]-radioactivity [B dpm]. After electrophoresis, the gel was dried and exposed overnight to Cronex 4 X-ray film at −80°C (DuPont, Wilmington, DE). The located monomer and dimer [^{125}I]-GPA bands from each sample were excised from the gel with a razor blade and the radioactivity was measured in a gamma counter [A dpm]. Variations in the radioactivity between the different samples were corrected by dividing this value [A dpm] by the total number of counts for the 10 μl aliquot [B dpm]. The percent hydrolysis of the sample was determined by comparing the sample against an unhydrolyzed substrate control:

$$\% \text{ Hydrolysis} = [1-\{(A/B)_{sample}/(A/B)_{unhydrolyzed\ control}\}] \times 100\%$$

The radioactive method also reveals a characteristic pattern of hydrolysis products generated by the specific action by the glycoprotease, and this pattern can be used to rule out non-specific protease contamination.

The second method measures the increase in fluorescence resulting from the proteolysis of a fluorescence-quenched substrate, GPA conjugated with 4,4-difluor-5,7-dimethyl-4-bora-3α,4α-diaza-s-indacene-3-propionic acid (BODIPY-FL-GPA) (Jiang and Mellors, in press). The sample is incubated for 15 min at 37°C with 3.5 µg of the substrate in 50 mM HEPES (pH 7.4) in a total reaction volume of 25 µl. The reaction is terminated by a 20x dilution with cold 50 mM HEPES (pH 7.4), followed immediately by the measurement of fluorescence in a Hitachi F-2000 spectrophotofluorimeter, slit widths 2 nm, at an excitation wavelength of 485 nm and emission at 515 nm. Fluorescence quenching within the substrate is reduced by proteolysis. This method provides a higher reproducibility than the more labour-intensive radiochemical method, but it is not sensitive enough for samples with very low glycoprotease activity, which were assayed in a 16-hour [^{125}I]-GPA hydrolysis assay.

The effects of various samples on the CHAPS-supplemented CCS wild type glycoprotease enzyme activity was determined by incubation of the sample and the enzyme on ice for 1 h, followed by the addition of either [^{125}I]-GPA or BODIPY-FL-GPA. The enzyme assay was then carried out as previously described and the decrease (or increase) in glycoprotease activity was compared with controls.

CHAPTER 3. Cloning and purification of the *O*-sialoglycoprotease as a GST fusion protein.

I. Introduction

- The plasmid, pGEX-4T-3, allows the expression of heterologous genes in *E. coli* as a fusion to an N-terminal carrier protein, glutathione-*S*-transferase (GST) from the parasitic helminth *Schistosoma japonicum* (Smith and Johnson, 1988). The family of glutathione-*S*-transferase enzymes (EC. 2.5.1.18) catalyzes the addition of aliphatic, aromatic, or heterocyclic radicals, as well as epoxides and arene oxides, to the sulphydryl group of glutathione (Wilce and Parker, 1994). This function facilitates cellular detoxification by rendering xenobiotics and other physiological substances more water-soluble, allowing for the elimination of these glutathione conjugates through an ATP-dependent efflux pump, GS-X (Ishikawa, 1992). The enzyme GST possesses a high specificity for the tripeptide substrate glutathione (L-γ-glutamylcysteinylglycine), allowing for the rapid purification of GST fusion proteins by affinity chromatography with glutathione-Sepharose beads under non-denaturing, reducing conditions. Furthermore, the addition of a GST carrier may result in the expression of a more soluble gene product, localized in the *E. coli* cytosol.

The widespread use of available pGEX plasmids is an indication of the popularity and success of this expression system. Many of the commercial variants have been constructed via frameshifts of the multiple cloning site and may incorporate a kinase region for the *in vitro* labeling of the fusion protein with ^{32}P, or a proteolytic cleavage site. Currently, vectors are available with recognition sites for three site-specific proteases: thrombin, factor Xa, or a GST-human rhinovirus 3C protease fusion that allows for the

auto-catalytic proteolysis of the recombinant fusion protein on the affinity column. Other vectors have incorporated the addition of a glycine-rich sequence before or after the cleavage site to facilitate proteolysis (Guan and Dixon, 1991; Hakes and Dixon, 1992). A novel triple fusion protein system that includes a C-terminal 10 amino acid streptavidin affinity tag, in addition to the GST moiety, offers advantages in removing degradation products and the site-specific protease (Sun and Budde, 1995). Several new vectors have been derived from the original pGEX vector to allow for the expression of eukaryotic proteins with appropriate post-translational modifications. These expressions include: the production of the human nuclear phosphoprotein p53 in reticulocyte lysates (Chumakov and Koeffler, 1993); the phosphorylated intracellular domain of human CD95 in murine cells (Rudert *et al.*, 1996); a correctly phosphorylated human papilloma virus type 6 E7 protein in yeast (Romanos *et al.*, 1995); and a commercially available baculovirus system for GST fusion expression in insect cells (Pharmingen, San Diego, CA).

A GST fusion protein vector was employed in the present study, to generate fusion proteins containing the recombinant *P. haemolytica* O-sialoglycoprotease, in an attempt to circumvent several problems encountered in the production of the heterologous recombinant enzyme protein. These problems included the periplasmic aggregation of the recombinant protein, through inappropriate disulphide bond formation in the more aerobic environment of the *E. coli* periplasm. The GST fusion protein was also made to facilitate isolation of the recombinant product via affinity chromatography, thus eliminating contamination by *E. coli* lysate components which inhibit the enzyme. The objective of this expression was to yield a high level of a purified, biologically active recombinant glycoprotease that could be used as an alternative to the wild type protein.

II. Experimental Procedures.

Construction of the vector pGGCP.

A fusion protein between GST and *P. haemolytica* Gcp was constructed by cloning the *gcp* gene, amplified by PCR, into the plasmid pGEX-4T-3 (Figure 3.1) (Pharmacia Biotech, Montreal, PQ). Two oligonucleotide primers were synthesized by B. Cooney (OMAFRA, Guelph, ON) and purified using OPC™ columns according to the manufacturer's protocol (PE Applied Biosystems, Norwalk, CT). The primers incorporate the restriction enzyme sites for *Eco*RI and *Sal*I respectively in order to facilitate the cloning into the MCS of the vector pGEX-4T-3 (Table 3.1).

Three hundred nanograms each of these two primers were used in the polymerase chain reaction (Saiki *et al.*, 1988) to amplify the *gcp* fragment using 20 ng of the plasmid pPH1.1 as the template. The reaction mixture contained 2.5 mM of each dNTP, 2.5 mM $MgCl_2$, and 1x PCR buffer (Boehringer Mannheim, Laval, PQ) in a total volume of 100 µl under a layer of mineral oil. PCR was carried out in a Perkin-Elmer Cetus 480 thermal cycler™. The reaction was initiated by the addition of 5 units of AmpliTaq DNA polymerase (Gibco BRL), after a 5 minute hot-start at 95°C. This was followed by 30 cycles of 1 min at 95°C, 1 min at 43°C, and 1 min at 72°C.

The PCR product was extracted by washing the aqueous layer with an equal volume of chloroform, followed by an equal volume of chloroform/phenol equilibrated in TE buffer. The DNA was precipitated by the addition of twice the volume of ethanol with 0.1 M sodium acetate and an incubation for 3 h at −20°C. The pellet was washed in ethanol and air dried before being resuspended in a tenth volume of TE buffer. The DNA fragment was then digested with the restriction enzymes, *Eco*RI and *Sal*I, to allow for an

46

Figure 3.1. The plasmid map of pGEX (reproduced from Pharmacia Biotech). The figure shows the relative positions of the open reading frames for ampicillin resistance, GST, and the repressor lacI^q; the position of the tac promoter, the origin of replication, and the multiple cloning site for pGEX-4T-3 are also shown.
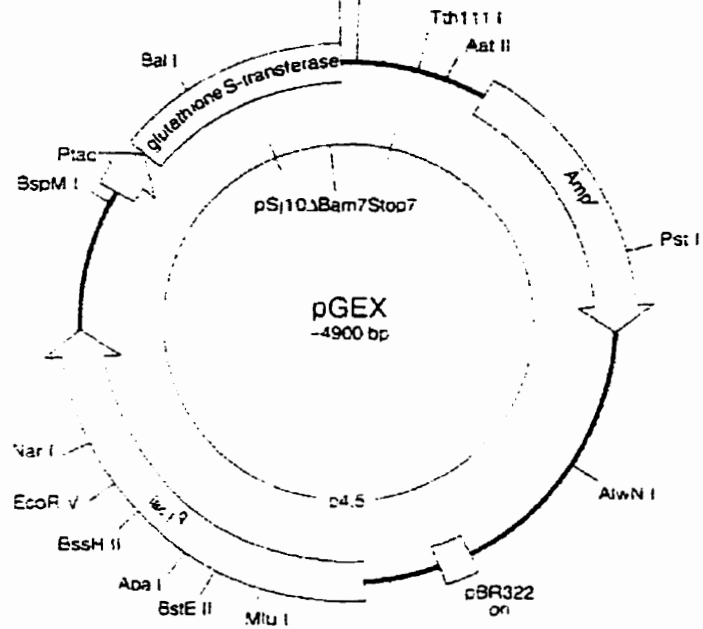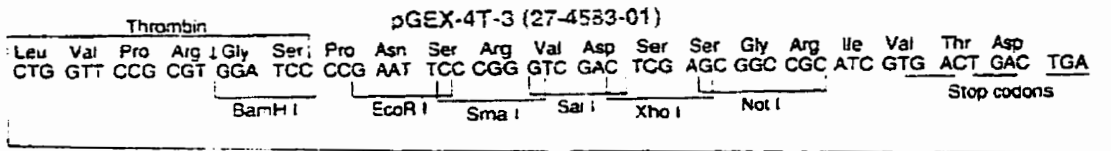
pGEX-4T-3 (27-4583-01)

Thrombin																							
Leu	Val	Pro	Arg ↓	Gly	Ser	Pro	Asn	Ser	Arg	Val	Asp	Ser	Ser	Gly	Arg	Ile	Val	Thr	Asp				
CTG	GTT	CCG	CGT	GGA	TCC	CCG	AAT	TCC	CGG	GTC	GAC	TCG	AGC	GGC	CGC	ATC	GTG	ACT	GAC	TGA			

BamH I   EcoR I   Sma I   Sal I   Xho I   Not I

Stop codons

48

Table 3.1. DNA primers used for the amplification of *gcp* during the construction of the vector pGGCP.

Primer	Sequence[1]	Comments[2]
L (#938)	5' CTTATTTTG↓AATTCTCCAACTATGCG 3'	This primer corresponds to nucleotides –21 to +5 relative to the *gcp* gene.
R (#939)	5' TGCGGACG↓TCGACTAATTACATAATT 3'	This primer is found downstream of the stop codon, corresponding to the complimentary strand of nucleotides 1145 to 1120.

1   The underlined regions represent the recognition site for the restriction enzymes *EcoR*I and *Sal*I respectively. The arrows represent the cleavage site. The bold letters indicate where a mutated base has been introduced into the PCR product.

2   The nucleotide sequences are found in Appendix C.

in-frame fusion into the vector pGEX-4T-3 digested with *Eco*RI and *Sal*I, by the techniques outlined in Chapter 2. The new plasmid, pGGCP, encodes the 64 kDa fusion protein rGgcp (as elucidated using Generunner™), and was transformed into *E. coli* DH5α and maintained on LB+Amp agar plates.

Conditions for rGgcp production.

The recombinant protein, rGgcp, was produced by inoculating *E. coli* DH5α/pGGCP in LB broth supplemented with 100 μg/ml ampicillin (LB+Amp). Parallel cultures of *E. coli* DH5α/pGEX-4T-3 and *E. coli* DH5α were grown concurrently, as vector and negative controls for the rGgcp. The cultures were grown overnight to stationary phase, at either 37°C or 30°C, with moderate aeration (150 rpm). They were then subcultured 1:10 into LB+Amp broth and grown at 30°C to an OD_{600} of about 0.5 (approximately 1 h) before protein expression was induced by the addition of 1 mM IPTG. The culture was induced for 1.5 h at 30°C and the cells were pelleted by centrifugation for 15 min at 7000 ×g at 4°C. The pellet was washed in a half culture volume of PBS buffer, and resuspended in a tenth culture volume of PBS, supplemented with 5 mM DTT, 0.2% CHAPS, and a protease inhibitor cocktail of 25 μg/ml each of antipain, aprotinin, leupeptin, and pepstatin (the column-loading buffer). The cells were then lysed, either by three passages through a French pressure cell at 1000 lbs/in^2, or with a Heat Systems Microson™ sonicator using four 15 sec bursts at power level 10, interspersed with an equal time of cooling in an ice/ethanol bath. Cellular debris was cleared by centrifugation at 13000 ×g, 4°C. Aliquots of the culture supernatant, whole cell samples before and after IPTG induction, and the lysates, were stored at -20°C prior to analysis.

50

Purification of the rGgcp from the lysate was performed by affinity column

chromatography using glutathione-Sepharose beads (Pharmacia Biotech, Laval, PQ)

equilibrated with the loading buffer. The protein elution was monitored by recording the

absorbance at 280 nm, at a flow rate of 0.5 ml/min, to yield 1 ml fractions. The cellular

lysate was added to the column and washed immediately with the column-loading buffer

until all unbound protein was removed. The rGgcp fusion protein was eluted from the

column by the addition of an elution buffer (50 mM HEPES, 10 mM reduced glutathione,

1 mM DTT, 0.2% CHAPS, and in some experiments by the addition of 0.1 mM each of

$CaCl_2$, $ZnCl_2$, and/or $MgCl_2$). All fractions were analyzed for rGgcp content by dot

immunoblotting on nitrocellulose membranes. GST activity was measured with the

colourimetric substrate 1-chloro-2,4-dinitrobenzene which yields a product detected by its

absorbance at 340 nm (Pharmacia Biotech protocol). The purified fractions were assayed

for potential glycoprotease activity, and stored at -20°C.

Thrombin cleavage of rGgcp.

The fusion protein contains a thrombin-sensitive site in the linker region between

the GST moiety and the glycoprotease. The glycoprotease was cleaved from the carrier

by the addition of 1 µl of a thrombin stock solution (1 NIH unit/µl, stored at -70°C) to 24

µl of the purified rGgcp fraction (approximately 1 µg protein) with incubation for at least

two hours at room temperature. The samples were assayed for Gcp activity by the

immediate addition of 3.5 µg of [125I]-GPA and incubated as described in Chapter 2. The

thrombin cleavage was also performed in the presence of 3.5 µg of substrate in an attempt

to stabilize and detect glycoprotease biological activity.

Subcellular localization of the fusion protein.

A modified osmotic shock protocol was used to determine the subcellular localization of the recombinant protein (Neu and Heppel, 1965; Cornelis *et al.*, 1982). Briefly, plasmid protein expression was induced as described previously in a 20 ml culture volume, followed by centrifugation for 10 min at 6000 x*g* to yield a cell pellet and a culture supernatant fraction. The cell pellet was then washed three times in 10 mM HEPES before being equilibrated for 10 min in 10 ml of a sterile 20% sucrose (w/v), 30 mM Tris-HCl (pH 8.0), 10 mM EDTA solution at room temperature. The cells were centrifuged again and suspended in 250 µl 25% sterile sucrose (w/v) before being quickly squirtted into 10 ml of ice-cold, sterile water and stirred for 10 min. The suspension was centrifuged for 10 min at 9000 x*g* and the supernatant was carefully decanted (periplasmic fraction). The remaining cells were suspended in 10 ml of 10 mM Tris-HCl (pH 8.0) and sonicated to release the cytoplasmic fraction. The various aliquots were subjected to Western immunoblot analysis with appropriate monoclonal antibodies to determine the localization and size of the rGcp and rGgcp antigens. Subcellular fractionation was monitored by assaying fractions for the cytosolic marker enzyme malate dehydrogenase (Charnock *et al.*, 1992) and the periplasmic marker enzyme β-lactamase (Baumann *et al.*, 1989).

Renaturation and refolding procedures.

The binding of rGgcp to the glutathione-Sepharose column allowed for the removal of contaminating proteins with the purified protein eluted from the column by a reduced glutathione buffer or denaturants. Two different denaturants were tested as a means of removing the bound protein. This was accomplished through the addition of

either guanidinium hydrochloride (GuHCl) in a linear gradient from 0 to 6 M, or with a low pH glycine buffer (50 mM glycine (pH 2.5), 1 mM DTT, 0.2% CHAPS). The denatured protein was immediately dialyzed against 50mM HEPES (pH 7.5) supplemented with 1 mM DTT and 0.2% CHAPS, in the presence or absence of substrate (to stabilize the proper conformation). Alternatively, the denatured protein was desalted on a Sephadex G25 column, equilibrated and eluted with the buffer used in the dialysis experiment (Werner *et al.*, 1994).

The use of molecular chaperones has been successful in the *in vitro* refolding of some proteins into a biologically active conformations (Mendoza *et al.*, 1991; Schröder *et al.*, 1993; Ziemienowicz *et al.*, 1993; Weissman and Kim, 1993). The use of the DnaK/DnaJ/GrpE chaperone system or PDI was shown to partially refold an inactive periplasmic rGcp osmotic shock lyophilate preparation (Watt *et al.*, 1997a). In the present study, rGgcp was denatured and eluted from an affinity column by a linear gradient of 0 to 6 M GuHCl. Three fractions containing the purified protein were flushed with N_2, and kept at room temperature overnight in darkness. The denaturant was removed by a G25 desalting column equilibrated with 10 mM HCl. The protein fractions were then pooled, flushed with N_2, and either used immediately or stored at -20°C as a source of denatured, reduced substrate for chaperone-mediated refolding. Refolding of the rGgcp product was attempted with the DnaK/DnaJ/GrpE or PDI chaperones (StressGen Biotechnologies Corp., Victoria, B.C.) according to the methods used with some limited success by Watt *et al.* (1997a). Variations tested here included: the presence or absence of [^{125}I]-GPA; the addition of thrombin, 0.2% CHAPS and/or 1 mM $ZnCl_2$; a sample size of either 2 or 10 µl; and the inclusion of a ATP regeneration system (10 mM

creatine phosphate and 3.5 units/100μl creatine phosphate kinase). The DnaK/DnaJ/GrpE chaperones were present at concentrations of 1 μM, 0.1 μM, and 1 μM respectively, whereas PDI was present at 0.5 μg/100 μl. In one experiment, the two chaperone systems were included, separately and together, in a final reaction volume of 100 μl. A 12 μl aliquot was removed from each sample, at t = 0 and t = 12 hours, and these were assayed for biological activity. In another experiment, a 50 μl reaction volume was used and a 7 μl aliquot was removed for the activity assay at 0, 1, 2, 3, 4, and 5 h, a time span which allowed for some enzyme re-activation in the studies of Watt *et al.* (1997a).

The DnaK homologue, Hsp70, of a rabbit reticulocyte lysate system, commonly used for the *in vitro* expression of mRNA, was shown to refold denatured firefly luciferase (Schumacher *et al.*, 1996). This system was employed in the attempted refolding of purified rGgcp, and rGgcp denatured by guanidinium chloride or by low pH as described above. Other variations included the removal of the N-terminal GST domain by thrombin cleavage as described previously. Sample sizes of either 2 or 10 μl were incubated at 37°C in reticulocyte lysate (Promega Corp., Madison, WI) supplemented with 2 mM ATP in a total volume of 50 μl. Every hour, a 10 μl aliquot was removed and assayed for glycoprotease activity by the [^{125}I]-GPA assay. CCS Gcp from *P. haemolytica* A1 was heat-denatured to 50% activity by incubation at 40°C for a half-hour (Udoh, 1986) as a positive control, while sonicates from DH5α/pGEX-4T-3 were used as negative controls in the refolding experiment.

III. Results

Production of rGgcp.

The cloning of *P. haemolytica* *gcp* fragment into the fusion vector pGEX-4T-3 produced the 6.1 kb plasmid pGGCP. The 1.2 kb DNA fragment was cloned into the *Eco*RI and *Sal*I restriction enzyme sites of the vector to give an in-frame fusion of the gene to the C-terminus of GST, separated by a thrombin cleavage linker region (Figure 3.2). The construct should express a 64 kDa product, that after thrombin cleavage, should yield a recombinant Gcp protein, with a seven amino acid peptide preceding the N-terminal methionine of Gcp. After several putative clones were isolated, the identity of the construct was verified by restriction enzyme analysis. Purified plasmid was digested with *Eco*RI and *Sal*I to remove the insert, or the plasmid was digested with a restriction enzyme that had a single site in either the vector or insert (Figure 3.3).

E. coli DH5α was transformed with the various plasmids (Table 2.2) and used as the host for the expression of plasmid-borne proteins. Positive chromogenic reactions with alkaline phosphatase-conjugated second antibody were seen for dot immunoblots of IPTG-induced whole cell samples of strains bearing pPH1.1, pPH1.1ΔE, or pGGCP (data not shown). The appearance of a major 64 kDa band in silver-stained SDS-PAGE gels, which was recognized by the monoclonal antibody mAb110, confirmed the production of the fusion protein in *E. coli*/pGGCP, hereafter known as rGgcp (Figure 3.5).

The fusion protein, rGgcp, was expressed and purified by affinity chromatography using a 1 ml bed volume of glutathione-Sepharose beads. The fusion protein bound to the

Figure 3.2.   A section of the fusion plasmid pGGCP. The construction of the plasmid

involved the cloning of a 1.2 kb PCR fragment into the *Eco*RI and *Sal*I sites (underlined)

of the plasmid pGEX-4T-3 (Smith and Johnson, 1988).   The rectangles represent the

coding region of the *gst* and *gcp* genes respectively.   The solid arrow indicates the

thrombin cleavage site.   The numbers represent nucleotide positions for pGEX-4T-3 and
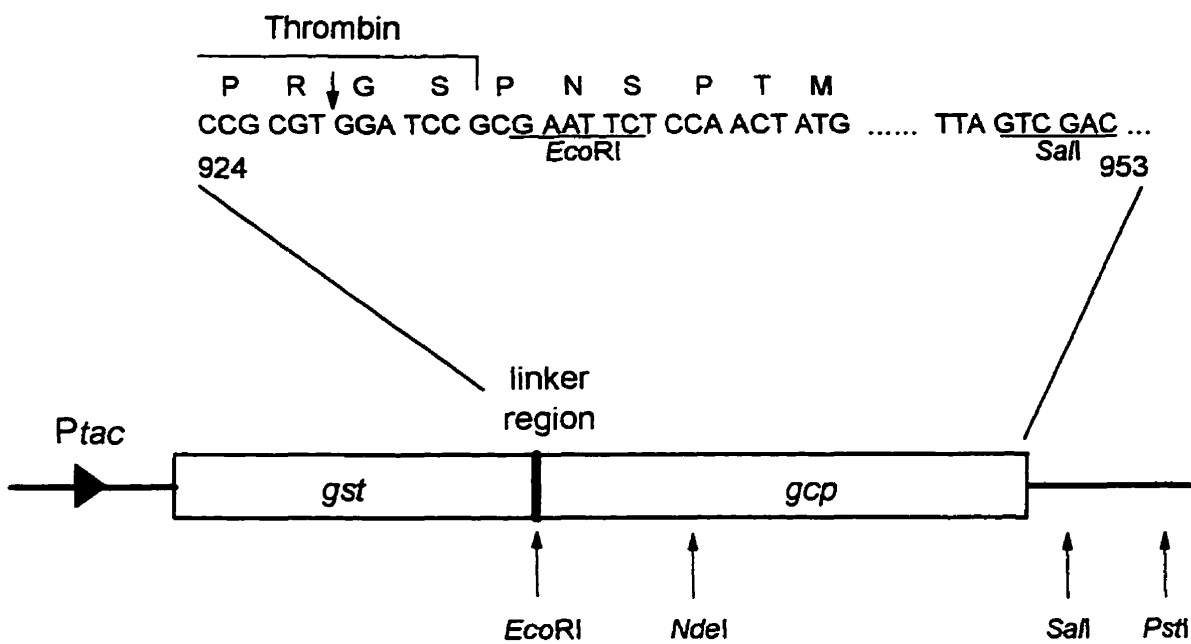
restriction enzyme sites are indicated by open arrows.

Thrombin

P   R↓G   S P   N   S   P   T   M

CCG CGT GGA TCC GC<u>G AAT TCT</u> CCA ACT ATG ...... TTA <u>GTC GAC</u> ...

924   *EcoRI*   *SaII* 953

Ptac

gst

linker
region

gcp

EcoRI   Ndel   SaII   Pstl

Figure 3.3.   The cloning of pGGCP. (A) The PCR amplification of *P. haemolytica gcp*

fragment. Lane M, 1 kbp molecular weight ladder; lane 1, *P. haemolytica* genomic DNA

template; lane 2, parental plasmid pTTQ19 template; lane 3-9, recombinant plasmid

pPH1.1 template; lane 10, blank. (B) Restriction enzyme digest of pGGCP constructs

isolated from four different colonies. Lane M, 1 kbp molecular weight ladder; lane 1,

*Eco*RI and *Sal*I digested pGEX-4T-3; lane 2, *gcp* PCR product; lane 3-6, *Eco*RI and *Sal*I

digested pGGCP from colonies 1-4 respectively; lane 7-10, *Nde*I digested pGGCP from

colonies 1-4 respectively; lane 11, *Pst*I digested pGGCP from colony 1.
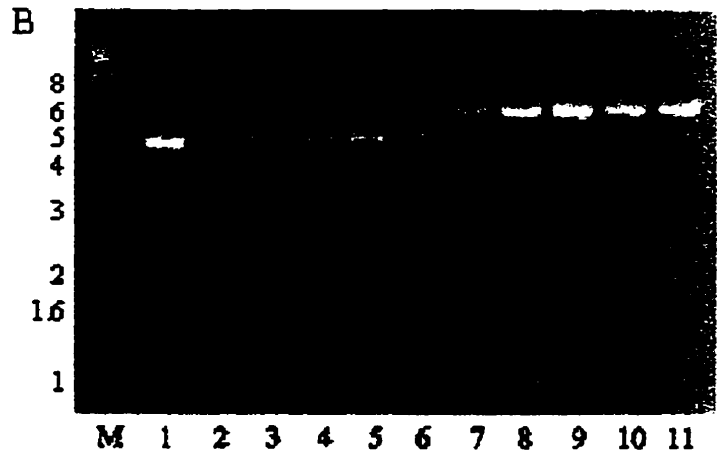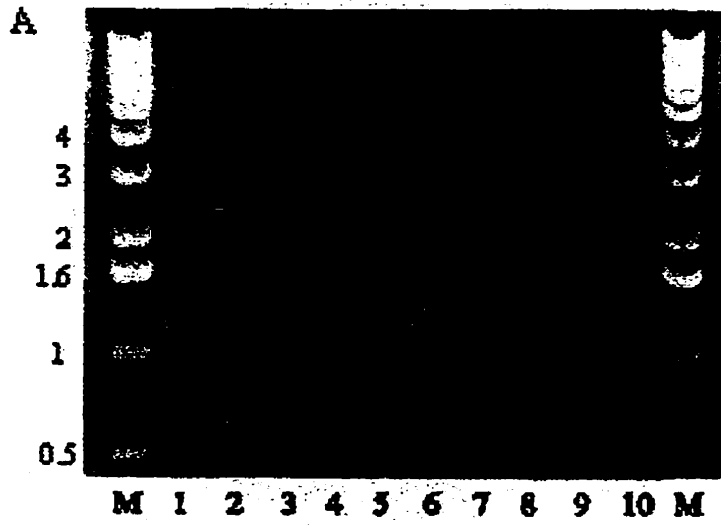
column, albeit not completely, only when 0.2% CHAPS was present in the loading buffer, as determined by dot immunoblotting (data not shown). The chromatographic separation of rGgcp lysates was monitored by continuously measuring absorbance at 280 nm (Figure 3.4). Unbound protein eluted prior to fraction 10, as confirmed by silver-stained 12% polyacrylamide gels (data not shown). The eluted proteins were separated by 12% SDS-PAGE, one gel was silver-stained to show protein, and a similar gel was immunoblotted with monoclonal antibody against Gcp (Figure 3.5). The fusion protein comprised 16% of total cellular protein in the lysate after 1.5 hours of induction as determined by densitometry of the silver-stained gel, and rGgcp accounted for the only detectable protein in the latter chromatographic fractions.

The purified rGgcp from fraction 20 was cleaved by the site-specific protease thrombin to remove the N-terminal GST domain. At a thrombin concentration of 0.1 NIH units/μl, the treatment yielded a 28 kDa band that corresponds to the expected size of GST and a 37 kDa band corresponding to Gcp (Figure 3.6). Glycoprotease enzyme assays for the rGgcp lysates, purified eluates, or thrombin cleaved fractions did not reveal any significant biological activity (data not shown).

Several methods of cell lysis were tested, of which the two most efficient means (as measured by release of protein) were sonication and French press lysis. Since sonication has been implicated in the denaturation and inactivation of some GST fusion proteins, the effect of sonication on the wild type CCS Gcp was tested. It was found that excessive sonication had no effect on the enzymatic activity of CCS Gcp (data not shown), nor did it affect the binding of rGgcp to the affinity column.

Figure 3.4.   Isolation of rGgcp by glutathione-Sepharose affinity chromatography of *E.*

coli DH5α/pGGCP lysates. The column was washed with PBS buffer supplemented with

0.2% CHAPS, and 5 mM DTT. The bound protein was eluted from the column by a

buffer consisting of 50 mM HEPES (pH 8.0), 10 mM reduced glutathione, 1 mM DTT,

and 0.2% CHAPS. The asterisks depict fractions containing rGgcp as determined by dot

immunoblotting with mAb110 against Gcp. Each fraction contained a 1 ml eluate.
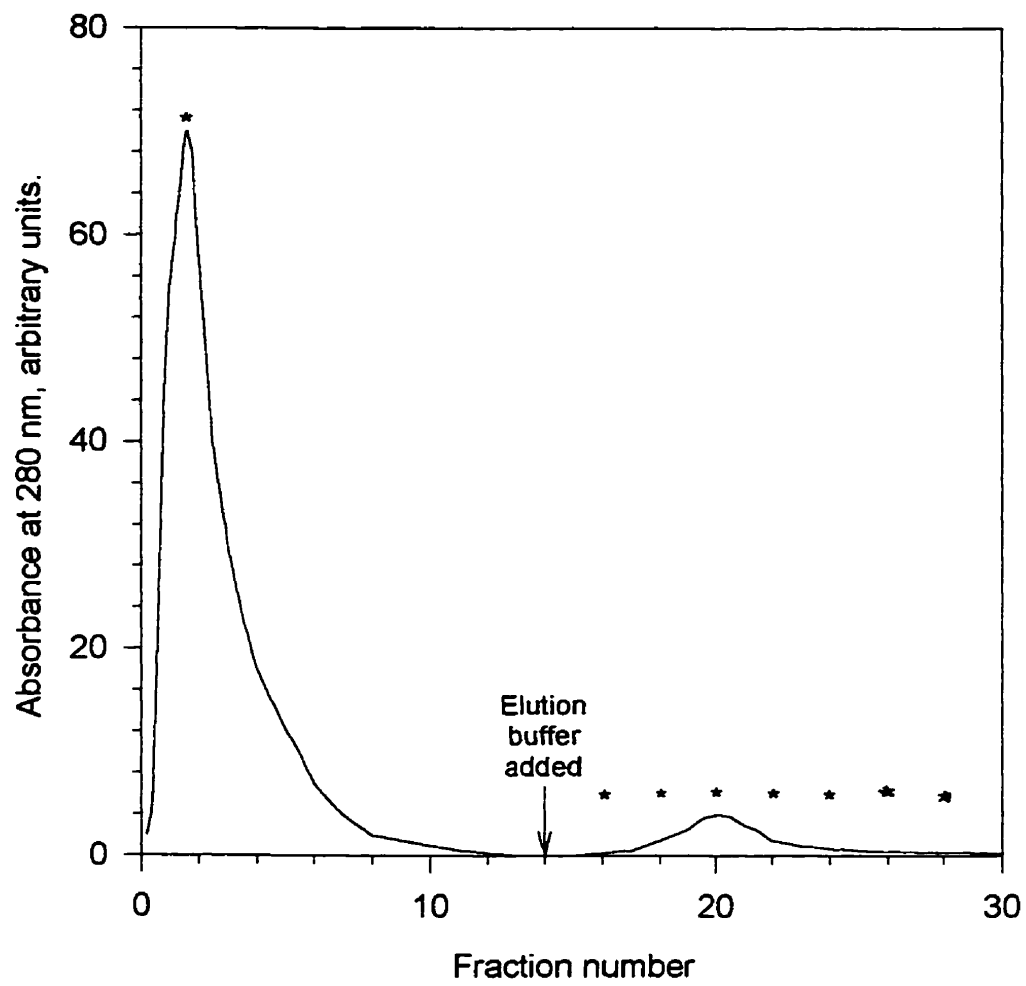
Figure 3.5.   Duplicate 12% SDS-polyacrylamide gels were either silver-stained (A) or

Western immunoblotted for Gcp (B) to show the isolation of rGgcp by GST-affinity

chromatography.   Protein expression from the vector control *E. coli* DH5α/pGEX-4T-3

(lane 1) and DH5α/pGGCP (lane 2) was induced by a 1.5 h incubation with 1 mM IPTG

followed by sonication.   The lysate from DH5α/pGGCP was separated on 1 ml

glutathione-Sepharose beads and 1 ml fractions were collected.   Fraction 2 (lane 3),

fraction 20 (lane 4), fraction 23 (lane 5), and fraction 28 (lane 6), are shown (refer to

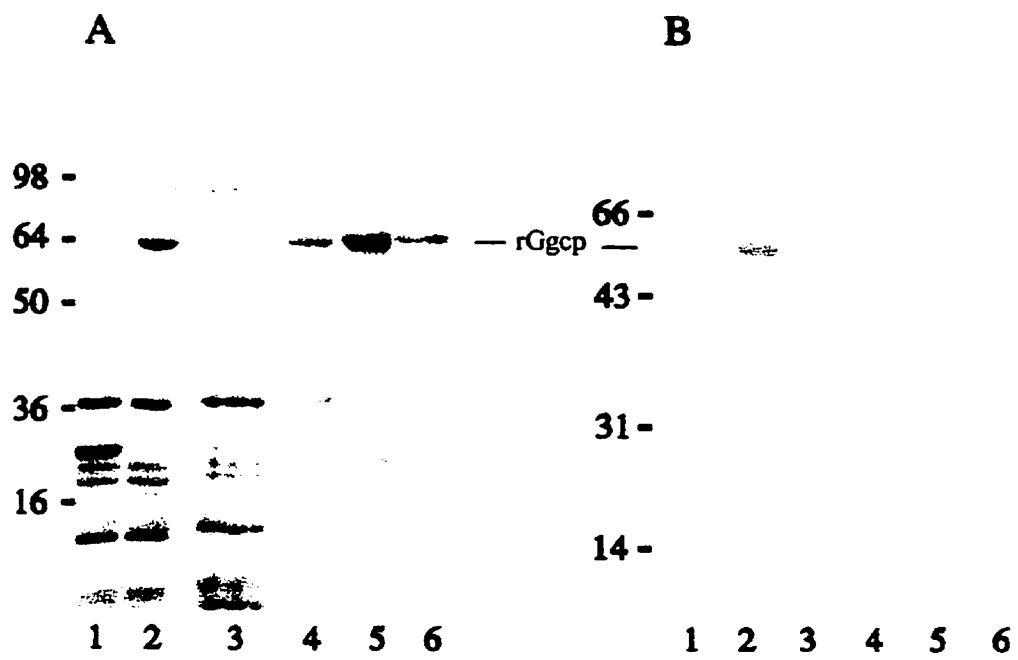Figure 3.4).   Molecular sizes (kDa) were determined from molecular weight standards.

A   B

98 -

                                                    66 -
64 -   —   — —— — — rGgcp —

50 -   43 -

36 - —— ——   ——   31 -

16 -   14 -

      1   2   3   4   5   6   1 2 3 4 5 6

64

Figure 3.6.   Thrombin cleavage of purified rGgcp.   A silver-stained 12% SDS-polyacrylamide gel shows a 2 h cleavage, by various concentrations of thrombin, of purified rGgcp from fraction 20, Figure 3.5.   Molecular weight markers in kDa (lane M); buffer control (lane 1); untreated rGgcp (lane 2); rGgcp treatment with 1 NIH unit/μl thrombin (lane 3); 0.1 NIH unit/μl thrombin (lane 4); 0.05 NIH unit/μl thrombin (lane 5); 0.01 NIH unit/μl thrombin (lane 6); 0.001 NIH unit/μl thrombin (lane 7); and 0.0001 NIH unit/μl thrombin (lane 8).

66-

Ggcp
43-

Gcp
31-

Gst

14-

M   1   2   3   4   5   6   7   8
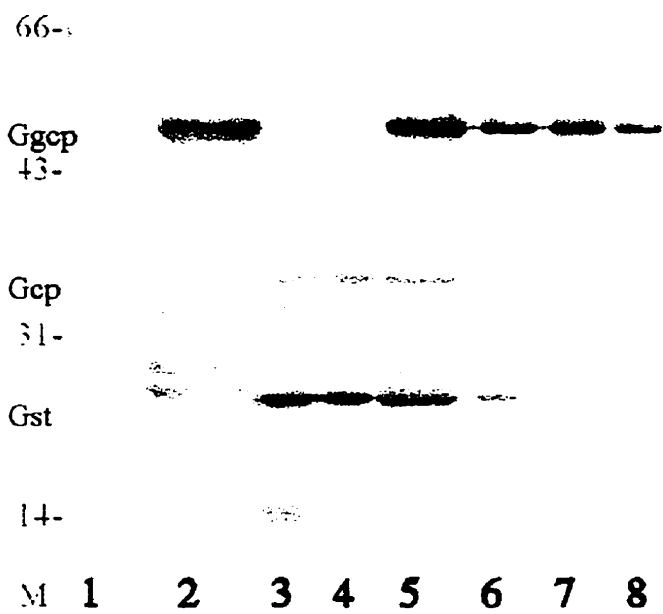
The cellular localization of the rGgcp enzyme was examined to determine the environment of the recombinant protein in the *E. coli* host. In contrast to the periplasmic localization of rGcp, the fusion protein rGgcp was found in the cytoplasm, by Western immunoblot analysis of subcellular fractions (Figure 3.7). The sub-cellular separation was monitored by marker enzyme assays, malate dehydrogenase (cytosolic) and β-lactamase for the periplasmic fraction (Table 3.2).

A major problem in the production of a recombinant glycoprotease is the expression of *gcp* to yield rGcp as a periplasmic disulphide-bridged aggregate. The wild type protein appears as a monomer by Western immunoblots under both reducing and non-reducing conditions, whereas rGcp is a disulphide-linked protein aggregate of large size, greater than a dimer or trimer (Watt *et al.*, 1997a). GST fusion proteins are often found as disulphide-linked dimers in solution due to the presence of three surface exposed cysteine residues in the GST domain. The cytoplasmic rGgcp also appears in a dimeric form under non-reducing conditions (Figure 3.8) suggesting that multimeric aggregation through disulphide bonds has been abolished for the Gcp domain, but that some dimerization can still occur through the GST region.

Renaturation and refolding of rGgcp.

After the removal of contaminating, unbound protein, rGgcp was eluted from the column by the addition of either glycine buffer or GuHCl. The recombinant protein eluted at 2.4 M GuHCl. Several attempts to renature the eluted rGgcp either by dialysis or by G25 desalting column chromatography did not yield any biological activity. An inhibition assay showed that the dialyzed eluates were not inhibitory towards the wild type CCS Gcp, which indicates that the denaturing agents had been removed and that any

Figure 3.7. Subcellular localization of the rGcp and rGgcp antigens. Western immunoblot of an osmotic shock preparation of periplasmic and cytoplasmic fractions from *E. coli* DH5α clones expressing rGcp (pPH1.1ΔE) (lanes 1-4) and rGgcp (pGGCP) (lanes 5-8). The fractions, in order, are the culture supernatant (lanes 1 and 5), cell pellet (lanes 2 and 6), the periplasm (lanes 3 and 7), and the cytoplasm (lanes 4 and 8). The vector control for pGGCP, pGEX-4T-3, showed no antigens recognized by the monoclonal antibody 110 specific for Gcp, and is not shown.
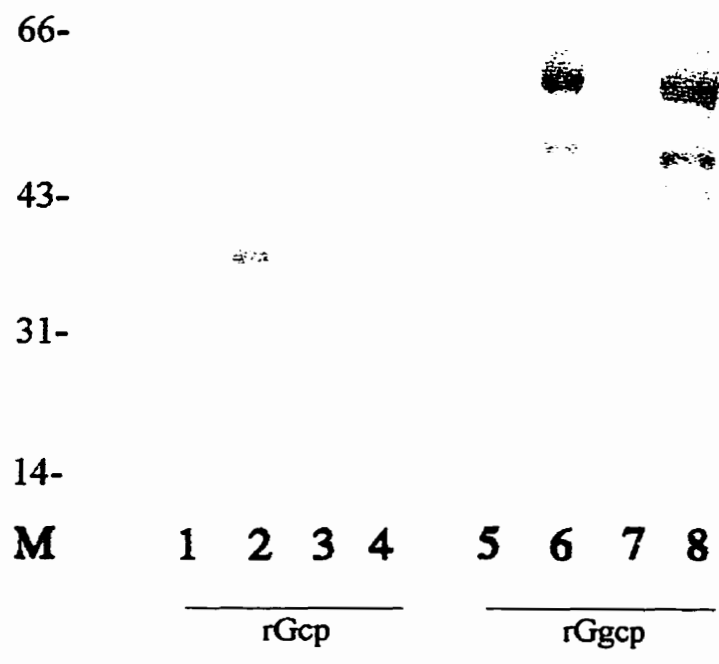
66-

43-

31-

14-

M   1 2 3 4   5 6 7 8

rGcp   rGgcp

Table 3.2.   Subcellular localization of rGcp and rGgcp in *E. coli* DH5α clones as determined by Western blotting with the mAb110 specific for Gcp.

DH5α Clones	Fraction	MDH[1]		β-lactamase[2]		Antigen
pPH1.1ΔE	SUP	2.9	(6.1%)	6.6	(4.2%)	-
"	PER	10.4	(21.8%)	130.2	(82.2%)	+
"	CYT	34.4	(72.1%)	21.6	(13.6%)	-
pGEX-4T-3	SUP	2.4	(5.5%)	12.6	(7.2%)	-
"	PER	10.4	(23.8%)	138.0	(78.5%)	-
"	CYT	30.9	(70.7%)	25.2	(14.3%)	-
pGGCP	SUP	1.5	(5.6%)	8.4	(5.5%)	-
"	PER	4.2	(15.8%)	124.2	(81.5%)	-
"	CYT	20.9	(78.6%)	19.8	(13.0%)	+

[1] MDH activity, nmol NADH consumed/min/ml

[2] β-lactamase activity, 3 pmol penicillin G hydrolyzed/min/ml

The values in parenthesis indicate the percentage of the total enzyme activity recovered in that fraction. Abbreviations: MDH, malate dehydrogenase; SUP, culture supernatant; PER, periplasm; CYT, cytoplasm.

70

Figure 3.8.   Western immunoblots of IPTG-induced whole cell samples separated by SDS-PAGE on 12% polyacrylamide gels under reducing (A) and non-reducing (B) conditions.   Molecular weight markers, kDa (lane M), cell pellet from *E. coli* DH5α/pPH1.1ΔE (lane 1), pGEX-4T-3 (lane 2), and pGGCP (lane 3).

	Reducing						Non-reducing			
A					**B**					
97-					97-					
66-					66-					
43-					43-					
31-					31-					
14-					14-					
M	1	2	3		M	1	2	3	

contamination with inhibitory components of the host cell lysates had also been eliminated. Controls showed that at least 20 mM GuHCl was required to inactivate the wild type Gcp enzyme by 50% (data not shown), so that contamination with traces of this denaturant would not cause problems with the assay of a reactivated enzyme.

The rGgcp protein was also tested as a substrate for chaperone-mediated refolding *in vitro*. Of approximately 200 samples tested under a variety of conditions (see Appendix B), only two samples indicated a very low level of glycoprotease activity. The first weakly positive sample was chromatographically purified rGgcp that had been treated with PDI for a 4 h incubation. However, no activity was seen in the equivalent 3 or 5 h treatments, so that this positive result was considered to be anomalous. A second sample which showed a weakly positive glycoprotease activity reaction was another PDI refolded preparation, this time for thrombin cleaved rGgcp in the presence of 0.2% CHAPS, after a 12 h incubation with PDI (Figure 3.9). False positives for glycoprotease activity can sometimes arise by bacterial contamination of enzyme assays despite the presence of protease inhibitors and antibiotics, so that the examples of apparent glycoprotease activity seen in Figure 3.9, may be artefacts. Several attempts were made to reproduce the positive results under similar conditions, but without success.

Rabbit reticulocyte lysate contains a complex mixture of proteins required for mammalian protein folding. The use of this system did not give rise to any refolding of denatured rGgcp, or heat inactivated wild type CCS Gcp, as indicated by glycoprotease enzyme assays. Components of the reticulocyte lysate also inhibited glycoprotease activity for the wild type enzyme and so the reticulocyte system was not used further.

Figure 3.9.   Chaperone-mediated refolding of rGgcp. Autoradiographs of the overnight

[^{125}I]-GPA activity assay for glycoprotease activity show a degradation product in A7 and

B6. The negative (substrate only) and positive (+ CCS Gcp) controls are shown in gel A,

lanes 1 and 2 respectively. Gel A shows PDI-treated rGgcp at t = 0, 1, 2, 3, 4, and 5 h for

lanes 3-8 (thrombin cleaved rGgcp) and lanes 9-14 (rGgcp). Gel B shows PDI-treated

rGgcp under a variety of experimental conditions. rGgcp was used as the substrate

directly (lanes 1, 2, and 7-10) or treated with thrombin (lanes 3-6 and 11-14). The

samples were supplemented with 0.2% CHAPS in lanes 1-2, 5-6, 9-10, and 13-14. The

folding reaction was carried out in the absence (lanes 1-6) or presence of [^{125}I]-GPA

(lanes 7-14). Unincubated control assays are shown (odd number lanes) with 18 h

incubations (even number lanes). Reaction conditions are described under Experimental

Procedures.

A

GPA₂

GPA

1 2 3 4 5 6 7 8 9 10 11 12 13 14

B

GPA₂

GPA

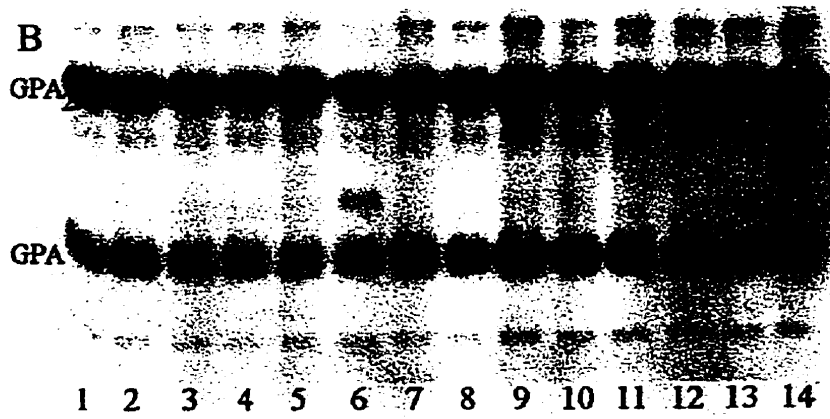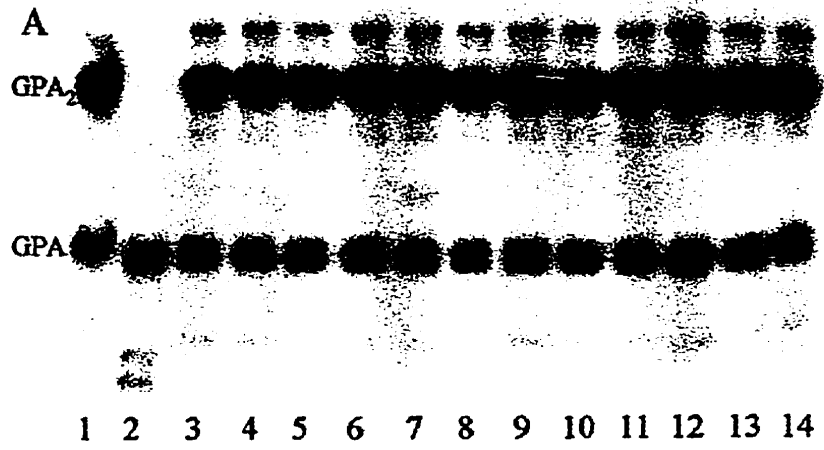1 2 3 4 5 6 7 8 9 10 11 12 13 14

Table 3.3.   The expression of rGcp and rGgcp in various *E. coli* strains transformed with pTTQ19, pPH1.1ΔE, pGEX-4T-3, or pGGCP. Sonicates from each sample were tested for glycoprotease activity by the [^{125}I]-GPA assay. A 10 µl aliquot of the sonicate was also tested for inhibition of wild type CCS Gcp by the BODIPY-FL-GPA assay (n=3).

- *E. coli* Strain	IPTG-induced Antigen[1]	Gcp Activity
191	+	-
BL21	+	-
CA274[2]	+	-
CSR603[2]	+	-
D10[2]	+	-
DH1	+	-
DH5α	+	-
HB101	+	-
JM83	+	-
JM101	+	-
JM109[2]	+	-
MV1190	+	-
Novablue	+	-
Sure2	+	-
TG1[2]	+	-
XL1-blue	+	-

1 – Antigen expression was detected only in strains that contained either pPH1.1ΔE or pGGCP, and not in any strains bearing the negative vector controls, nor in uninduced strains.

2 – Slight inhibition of wild type CCS Gcp was seen in the lysates of these strains.

The expression of rGcp and rGgcp in various *E. coli* hosts.

Sixteen *E. coli* strains were transformed with the plasmid pPH1.1ΔE and its parental vector pTTQ19, and with pGGCP and its parental vector pGEX-4T-3. The purpose of the study was to determine whether other *E. coli* strains, selected for protease deficiency and for other characteristics, would prove more amenable as hosts for the production of a biologically active recombinant protein. As shown in Table 3.3, while all strains tested showed IPTG-induced expression of rGcp and rGgcp, no strains showed enzyme activity in the whole cell lysate. Some strains showed inhibition of wild type Gcp when the sonicated lysate was included in an enzyme assay for the wild-type glycoprotease.

IV. Discussion

A GST-Gcp fusion protein, rGgcp, was constructed by cloning the *P. haemolytica gcp* gene into pGEX-4T-3. The expression of the recombinant protein in *E. coli* was dependent on IPTG induction, and yielded a protein of 64 kDa with few proteolytic products detectable by antibodies against Gcp. The product was readily purified by affinity chromatography to over 90% homogeneity as determined by silver-stained SDS-PAGE gels (Figure 3.5).

The binding of rGgcp to the affinity column had an absolute requirement for the presence of 0.2% CHAPS. This suggests that the recombinant fusion protein in the absence of detergent exists in an aggregated form, with its GST domain masked in a manner that does not permit affinity-binding to glutathione. The detergent disrupts the aggregate and permits specific binding to the ligand. It is not clear whether such aggregation occurs *in vivo* within the bacterium or if it is an artefact of the lysis process.

The aggregation may indicate aberrant misfolding of the protein during expression, or the dimerization of the protein through interaction of the GST domains. Either process could result in the absence of glycoprotease activity. Other GST fusion proteins require solubilization with detergent before binding to the affinity matrix (Frangioni and Neel, 1993). GST fusion proteins have been shown to dimerize in solution (Yan et al., 1995), through self-association of the GST domain. GST alone exists as a dimer in aqueous solutions (Wilce and Parker, 1994). Some proportion of rGgcp was observed as a dimer in the host cell cytoplasm when analyzed under non-reducing conditions (Figure 3.8). Under non-denaturing conditions, the rGgcp was in the form of aggregate that was too large to enter polyacrylamide gels (data not shown). Thus, even in the relatively reducing environment of the cytoplasm, disulphide bond formation occurs between the fusion protein monomers. Other studies on the GST fusion carrier have indicated that reversible oxidative aggregation does occur between the GST dimers through the four highly exposed cysteine residues per subunit (Kaplan et al., 1997). The local oxidative environment that nascent rGgcp encounters may have a major deleterious effect on the folding and activation of the glycoprotease domain.

A variety of refolding strategies was employed to produce active enzyme using the unpurified and the purified preparations of rGcp. The protocols were based on the expectation that a highly purified fusion protein would be more amenable to refolding than the relatively unpurified material extracted in previous studies (Watt et al., 1997b) (see Appendix B). No variations of the general scheme outlined in the methods section produced any significant or reproducible biological activity (data not shown). However, some general principles can be formulated, based on these and previous attempts to

produce an active recombinant enzyme. The presence of an N-terminal fusion partner may have a deleterious effect upon the binding and recognition of substrate by the enzyme. The rTRXGCP fusion protein displayed slight activity after enterokinase treatment to remove the thioredoxin moiety. Furthermore, active chromatographically-purified rTRXGCP contained several smaller degradation products that may have accounted for the observed weak activity (Watt *et al.*, 1997a). In the present study, the traces of substrate degradation that could be attributed to an enzymatically active glycoprotease were found only in the PDI refolded samples that had been treated by thrombin to remove the N-terminal GST moiety. If this is indeed the case, this suggests that the N-terminal region sterically hinders the binding of the substrate to the C-terminal catalytic domain. However, this is highly speculative due to the lack of reproducibility in any of the samples.

Numerous attempts at renaturation of denatured rGgcp did not generate biological activity. The GST molecule is thought to unfold and refold in a two-state process, without any intermediates or any requirement for chaperone proteins (Kaplan *et al.*, 1997). The Gcp portion may require chaperones to correctly fold, but the *in vitro* use of purified chaperones was not successful. If the trace of biological activity seen for two PDI-treated samples is due to productive folding, then it is a rare event under these experimental conditions, and suggests a transient, unstable, folding intermediate with a short half-life.

The expression of both rGcp and rGgcp in a variety of *E. coli* strains indicate that while this organism is generally suitable for heterologous gene expression, it is a poor host for the productive folding of the glycoprotease. Several of the strains tested were common laboratory host strains which were nuclease and recombination deficient. The

protease negative strain BL21 is suitable for minimizing intracellular proteolysis, and the

dsb mutant strain, *E. coli* 191, has a low capacity for forming periplasmic disulphide

bonds. However, neither strain allowed expression of enzymatically active rGcp or rGgcp.

Therefore another strategy was attempted, that is the use of a shuttle vector, pNF2176, to

express the *E. coli*-derived plasmid constructs in *P. haemolytica* serotypes.

CHAPTER 4. The cloning and characterization of rGcp and rGgcp in *Pasteurella haemolytica* A1 and A11 utilizing the vector pNF2176.

L. Introduction.

A shuttle vector, pNF2176, has recently been described in the literature (Fedorova and Highlander, 1997a) for the transformation of *E. coli* and *P. haemolytica* strains, and was shown to be competent for gene expression. The vector, based on a sulfonamide (Su^r) and streptomycin (Sm^r) resistance-bearing plasmid, pYFC-1 (Chang *et al.*, 1992), was constructed by cloning the ROB-1 β-lactamase gene, *blaP* (Liverlli *et al.*, 1991) into pYFC-1, followed by the replacement of the dihydropteroate synthase gene, *sulII* (type II Su^r), and the streptomycin kinase gene, *styR* (Sm^r), with a multiple cloning site (MCS) downstream from the sulfonamide promoter, P$_{sulII}$. This vector has been used in the construction of a *P. haemolytica* strain that secretes an inactive leukotoxin by introducing a mutated *lktC* gene (leukotoxin activating protein) into the genome by allelic exchange. A pNF2176 derivative containing the *lktC* gene was able to complement the mutation, indicating that the plasmid can functionally express native proteins (Fedorova and Highlander, 1997b). This vector was tested as a vehicle for the expression of *P. haemolytica* Gcp and the GST fusion protein, Ggcp (described in Chapter 3), as correctly folded, biologically active, recombinant proteins.

The expression of recombinant native proteins in *P. haemolytica* A1 poses a problem in distinguishing between the action of the wild type and recombinant forms of a given gene product. *P. haemolytica* biotype A11 has been reported to lack *O*-sialoglycoprotease activity and to have a different genetic organization of the *gcp* gene as revealed by Southern blot analysis (Abdullah *et al.*, 1990; Lee *et al.*, 1994a). The use of

the A11 strain as a negative control for the A1 strain has a precedent in the literature: the culture supernatant of an A11 strain was used as a non-pathogenic control in an *in vitro* study of neutrophil activation by serotype A1 (Mdurvwa and Brunner, 1994). The ovine-derived A11 strain is phylogenetically distinct from the bovine pathogen A1, and A2, based on its 16S rRNA, OMP and LPS profile (Davies *et al.*, 1996). An analysis of allelic variation, by a technique known as multilocus enzyme electrophoresis, for a large number of *P. haemolytica* isolates confirmed that A11 is taxonomically distinct from either A1 or A2, and may represent a separate species (Davies *et al.*, 1997). The A11 strain is non-pathogenic when compared with either the A1 or A2 serotypes. A11 does not produce a neuraminidase, in addition to the lack of glycoprotease activity (Straus *et al.*, 1993b).

The pNF2176 vector was used in this study to express rGcp and rGgcp in both *P. haemolytica* serotypes A1 and A11. Although serotype A11 may be a distinct species, the protein expression and protein folding mechanisms (e.g. chaperones) would be more similar to serotype A1 than that of *E. coli*. The A11 serotype would be useful for detecting any enzymatic activity of the recombinant gene product, since this serotype is the most closely related strain to serotype A1 that apparently lacks glycoprotease activity. The expression of the glycoprotease in *P. haemolytica* may circumvent the folding problems encountered when producing the recombinant glycoprotease in *E. coli*.

II. Experimental Procedures.

Construction of the vectors pNFG1/2 and pNFGG1/2.

The *gcp* gene from the plasmid pPH1.1ΔE and the *ggcp* gene from the plasmid pGGCP (described in Table 2.1) were used in the construction and expression of a recombinant glycoprotease in *P. haemolytica* using the expression vector pNF2176

(Figure 4.1). The vectors containing the *gcp* gene are pNFG1 and pNFG2, whereas the vectors containing the *ggcp* gene are pNFGG1 and pNFGG2, based on the orientation of the cloned gene relative to the sulfonamide promoter (Table 4.1). The genes were amplified by PCR using oligonucleotide primers synthesized from OMAFRA (OMAFRA, Guelph, ON). The primers incorporate a restriction enzyme site for *Sal*I to allow for the cloning of the genes in two orientations into the MCS of pNF2176 (Table 4.2).

The PCR conditions were similar to those described in chapter 3, except that 200 ng of each of the primer pairs was used in the reactions. The plasmids, pPH1.1ΔE and pGGCP, were used as the templates for the gene *gcp* and *ggcp* respectively. The PCR products were cleaned with "QIAquick™" spin columns according to the manufacturer's protocol (Qiagen Inc., Chatsworth, CA). The restriction enzyme *Sal*I was used to digest the purified PCR products and pNF2176, followed by heat inactivation for 20 min at 65°C. The terminal phosphate groups of the digested vector were removed by calf intestine alkaline phosphatase treatment to prevent religation. The treated DNA samples were purified and ligated using the techniques outlined in Chapter 2. The ligation mixtures were used to transform *E. coli* DH5α and prospective clones were isolated from LB + Amp agar plates. Plasmid DNA was purified from each clone, subjected to restriction enzyme analysis, and stored at -20°C.

Transformation of *P. haemolytica* with pNFG1/2 and pNFGG1/2.

Plasmid isolated from *E. coli* DH5α was electroporated into *P. haemolytica* serotypes A1 and A11. The *P. haemolytica* cells were made competent for electroporation by subculturing 3 ml of an overnight culture into 250 ml of sterile BHIB, and grown at 37°C with aeration until an OD$_{600}$ of about 0.5 was reached. The cells were

Figure 4.1. The plasmid map for pNF2176 (reproduced from Fedorova and Highlander, 1997a). The plasmid carries the origin from pYFC (*ori*), and the selection marker for ampicillin resistance gene (*blaP*). The promoter for the sulfonamide gene (P_{suIII}) is followed by the multiple cloning site (MCS). The transcriptional terminator from the steptomycin resistance gene of the parental pYFC plasmid is marked *t*. Several restriction enzyme sites are listed and non-unique sites are indicated with an asterisk.
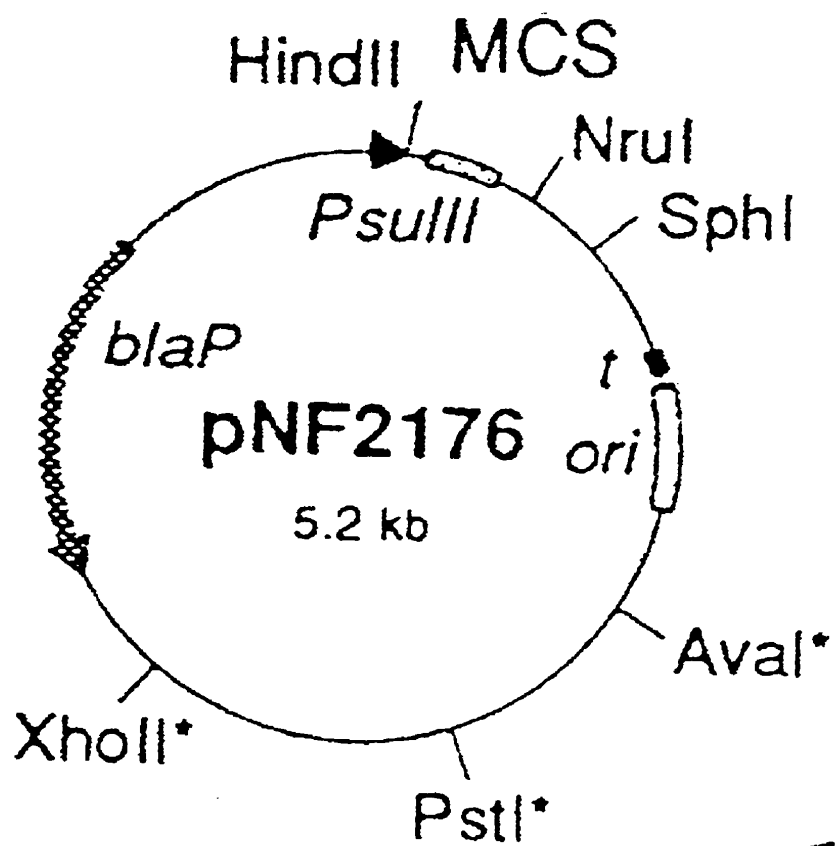
Table 4.1. Construction of the rGcp and rGgcp pNF2176 plasmid derivatives. The triangles represent the promoter sites, the vertical arrows indicate the restriction enzyme recognition sites, and the direction of transcription is indicated by the horizontal arrows.

Plasmid	Size (bp)	Map of Constructs	Size of R.E. Fragments (bp)
pNF2176	5162		SalI – 5162
pNFG1	6364		XhoI – 6364 SalI – 5162 & 1202 XhoI & ScaI - 4026, 2280, & 58
pNFG2	6364		XhoI – 6364 SalI – 5162 & 1202 XhoI & ScaI - 3254, 3052, & 58
pNFGG1	7060		XhoI – 7060 SalI – 5162 & 1898 ScaI - 4288, 2714, & 58
pNFGG2	7060		XhoI – 7060 SalI – 5162 & 1898 ScaI - 3684, 3318, & 58

86

Table 4.2. DNA primers used for the amplification of *gcp* (primers GCP797L and R) and *ggcp* (primers GGCP797L and R) during the construction of the vectors pNFG1/2 and pNFGG1/2.

Primer	Sequence[1]	Comments[2]
GCP797L	5' AACAATTCTAG↓TCGACCAGAGAATAA 3'	This primer corresponds to nucleotides −83 to −58 relative to the *gcp* gene.
GGCP797L	5' GAAATGAGCTG↓TCGACAATTAATCAT 3'	This primer corresponds to nucleotides −86 to −61 relative to the *ggcp* gene..
R (#939)	5' TGCGGACG↓TCGACTAATTACATAATT 3'	This primer is found downstream of the stop codon, corresponding to the complimentary sequence of nucleotides 1145 to 1120 for *gcp* & 1830 to 1812 for *ggcp*.

1   The underlined regions indicate the restriction site for *Sal*I. The arrows represent the cleavage site. Bold letters indicate where a mutated base has been introduced into the PCR product.

2   The nucleotide sequences are found in Appendix C.

cooled on ice for 15 min before being pelleted for 15 min at 4000 xg in a GSA rotor. The soft pellets were resuspended in 250 ml sterile 272 mM sucrose and centrifuged at the same speed. The cells were then washed in 125 ml 272 mM sucrose before being pelleted at 6000 xg. The extremely soft pellet was carefully resuspended in 5 ml of sterile 15% glycerol, and centrifuged for 15 min at 6000 xg in an SS34 rotor. The remaining cells were resuspended in 500 µl 15% glycerol and 100 µl aliquots were stored at -70°C. Purified plasmid DNA (5 or 10 µg) was incubated with 100 µl aliquots of the competent cells thawed on ice. The cells were electroporated in a BioRad Gene pulser™ set at 25 µF, 600 Ω, and 1.5 kV, in cooled cuvettes with a 0.1 cm electrode gap. Immediately after a one second pulse, 1 ml of sterile BHIB was added to the cuvettes, and the cells were allowed to recover for 6 h at 37°C with vigorous shaking. The transformants were plated onto BHI + Amp plates and grown overnight at 37°C. Putative clones were re-patched onto both BAP and BHI + Amp plates. Four criteria were used for the determination of *P. haemolytica* colonies: (a) haemolysis on BAP, (b) a characteristic odour compared with a positive control, (c) visual identification of an off-white opaque colony, and (d) a positive oxidase test with the substrate *p*-aminodimethylaniline oxalate. *P. haemolytica* strains were serotyped using a modification of the slide agglutination assay of Frank and Wessman (1978). Briefly, a 20 µl aliquot of type specific sera, raised in rabbits against either the A1 or A11 serotypes, was mixed on a clear glass slide with a 20 µl aliquot of a saturated *P. haemolytica* culture. The solution remained turbid in a negative reaction, whereas a positive reaction resulted in clumping and clearing of the mixture. Additional

serotyping information was obtained from an independent source (M°Bey, Department of Pathobiology, OVC, University of Guelph).

Growth and expression of pNFG1/2 and pNFGG1/2.

E. coli strains containing the vectors were grown in LB + Amp (100 µg/ml) whereas plasmid-bearing *P. haemolytica* strains were grown in BHI + Amp (5 µg/ml). Strains containing pNF2176 or no plasmid were used as negative controls for recombinant glycoproteolytic activity. Overnight cultures were grown for 2 h at 37°C, with or without 1 mM IPTG, and centrifuged to obtain the culture supernatant. The cell pellet was resuspended in a volume of 50 mM HEPES (pH 7.4) equal to that of the culture supernatant, and sonicated to yield a cell lysate. Gene product expression was tested by chemiluminescent Western immunoblotting with monoclonal antibody specific for Gcp, and the samples were assayed for Gcp activity. The $[^{125}I]$-GPA assay was performed in the presence or absence of the neuraminidase inhibitor, (2,3)-dehydro-2-deoxy-N-acetyl-neuraminic acid (10 mM).

III. Results

The construction of the pNF2176 derivatives was initiated by the PCR amplification of the genes, *gcp* and *ggcp*. The incorporation of a single restriction enzyme site allowed for the bi-directional cloning of the genes with respect to the sulfonamide promoter, into the pNF2176 vector. The resulting pNF2176 derivatives included the *gcp*-coding vectors, pNFG1 and pNFG2, and the *ggcp*-coding vectors, pNFGG1 and pNFGG2 (Figure 4.2). The vectors were successfully transformed into *P. haemolytica* serotypes A1 and A11 as confirmed by a restriction enzyme digestion of plasmid DNA isolated from each strain (Figure 4.3). The identities of the *P. haemolytica* strains were verified by the

Figure 4.2.   Cloning of the pNF2176 derivatives. (A) PCR amplification of the genes

gcp and *ggcp*. Lane M, 1 kbp molecular weight ladder; lane 1 & 2, negative control; lane

3, pTTQ19 parental plasmid template; lane 4-7, pPH1.1ΔE template; lane 8, pGEX-4T-3

parental plasmid template; lane 9-12, pGGCP template. (B) Restriction enzyme digests

of pNF2176 derivatives isolated from *E. coli* DH5α. Lane M, 1 kbp molecular weight

ladder; plasmid pNF2176 (lane 1); pNFG1 (lanes 2-4); pNFG2 (lanes 5-7); pNFGG1

(lanes 8-10); and pNFGG2 (lanes 11-13). The plasmids were digested with *Sal*I (lanes 1,

3, 6, 9, & 12); *Xho*I (lanes 2, 5, 8, & 11); *Xho*I & *Sca*I (lanes 4 & 7); and *Sca*I (lanes 10 &

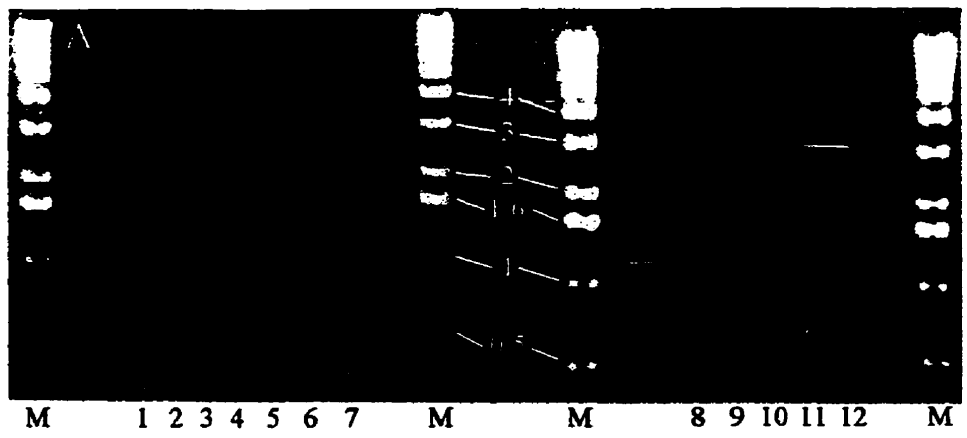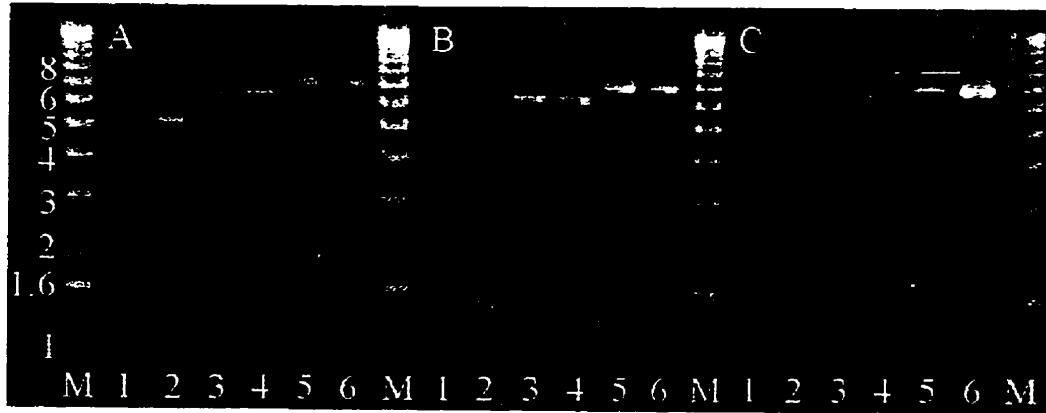13). The expected sizes of the fragments are listed in Table 4.1.

Figure 4.3.   Restriction enzyme digests of pNF2176 derivatives isolated from *E. coli* DH5α (A), and *P. haemolytica* serotypes A1 (B), and A11 (C). Lane M, 1 kbp molecular weight ladder; lanes 1-6: strain alone, pNF2176, pNFG1, pNFG2, pNFGG1, and pNFGG2.

determinative tests described in the methods section, and confirmed by serotyping. It was necessary to re-isolate the *P. haemolytica* strains from glycerol stocks after a single passage on BHI plates, since extraneous DNA bands began to appear on E-lyse gels in the passaged strains. These anomalous bands were never observed in the parental strain, but only in the strains that carried the pNF2176 vector or its derivatives. The origin or identity of the anomalous DNA was not determined. All of the following results are acquired from strains that did not show anomalous DNA.

All of the *P. haemolytica* strains were grown as described in the methods section. The culture supernatant fraction was collected from overnight cultures and tested for glycoprotease activity by both the $[^{125}I]$-GPA and BODIPY-FL-GPA assays (Figures 4.4 and 4.5). The transformed *E. coli* DH5α clones never exhibited glycoprotease activity as measured by either assay. There was no significant alteration of Gcp activity relative to the wild type strain, in any of the transformed *P. haemolytica* A1 clones. The original A11 transformants were negative for any glycoprotease activity, in agreement with previous observations (Abdullah *et al.*, 1990; Lee *et al.*, 1994a). However, a constitutive level of glycoprotease activity appeared in the wild type A11 strain that was used thereafter for the cloning of pNF2176 and its derivatives. Serotyping confirmed that the organism was the A11 biotype, and that the glycoprotease activity was not due to contamination with other serotypes.

A sensitive chemiluminescence-based Western blot analysis on whole cell fractions was carried out on the clones, for the detection of rGcp antigen. The amounts of expressed protein were too low to be detected colourimetrically by an alkaline phosphatase-conjugated secondary antibody. The results show that in both *P.*

Figure 4.4.   Glycoprotease activity in the culture supernatants of *P. haemolytica*

serotypes A11 and A1 transformed with pNF2176 and its derivatives. (A) Serotype A11:

lane 1, GPA substrate control; lane 2, glycoprotease enzyme positive control; lanes 3 and

4, two stocks of the parental *P. haemolytica* serotype A11 strain; lanes 5-7, un-

transformed descendants of the parental A11 strain; lanes 8-12 are the transformed A11

descendants containing pNF2176, pNFG1, pNFG2, pNFGG1, and pNFGG2 respectively.

(B) Serotype A1: control lanes 1 & 2 as in (A); lane 3-5, untransformed A1 strain

controls; lanes 6-10 are serotype A1 transformed with pNF2176, pNFG1, pNFG2,

pNFGG1, and pNFGG2 respectively.   The culture supernatants were assayed by [^{125}I]-

GPA hydrolysis for 30 min at 37°C.

A

1 2 3 4 5 6 7 8 9 10 11 12

B

1 2 3 4 5 6 7 8 9 10

Figure 4.5.   Glycoprotease activity, as determined by the BODIPY-FL-GPA activity

assay, for a 10 µl aliquot of the culture supernatants of *P. haemolytica* serotypes A1 and

A11 transformed with the pNF2176 vector and its derivatives. The pNF2176 derivatives

were transformed into Isolate C of both the A1 and A11 serotypes. The samples were

incubated for 15 minutes at 37°C as described in Chapter 2. All samples were tested in

triplicate; the standard deviation is shown as the error bar.
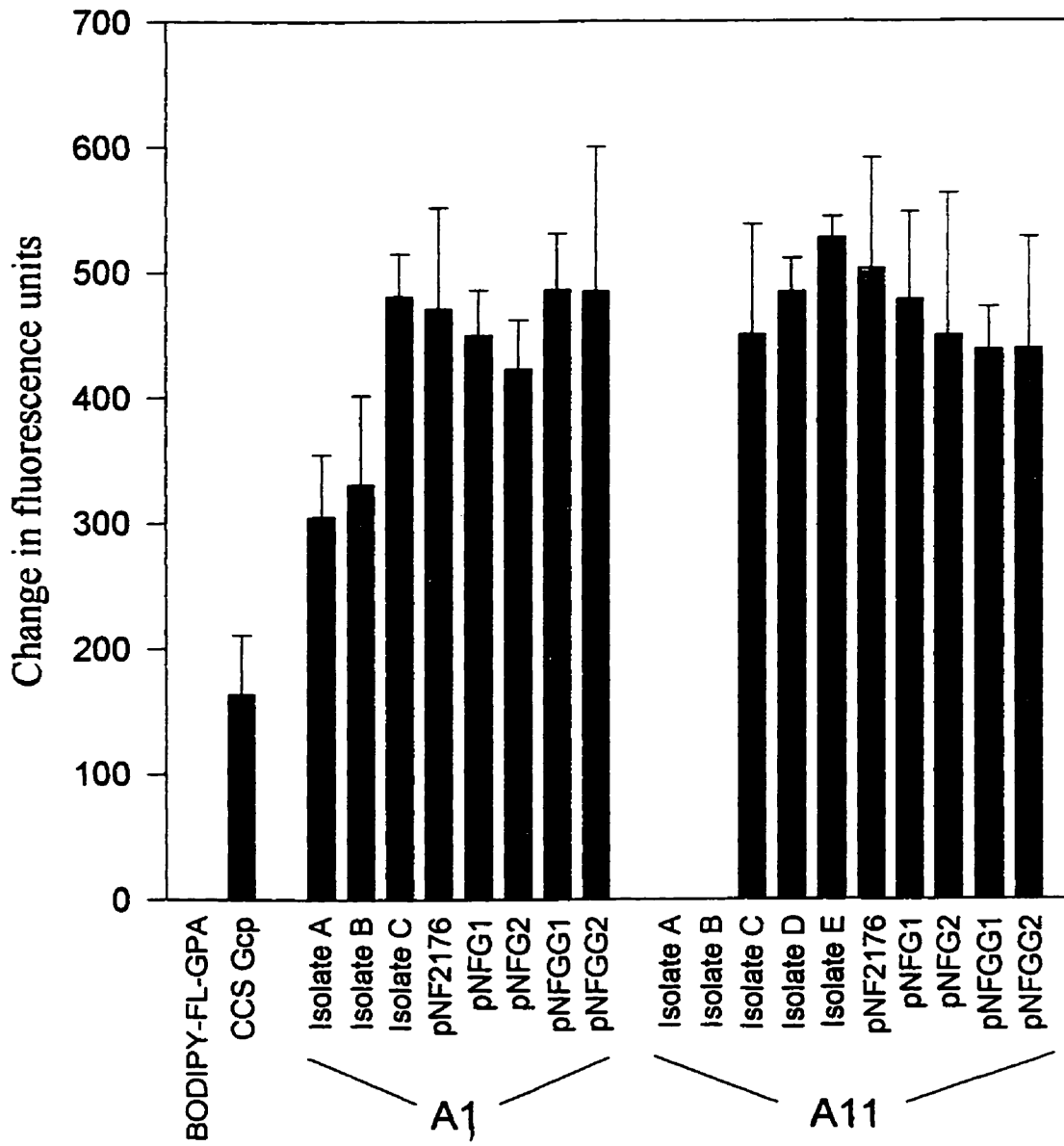
Figure 4.6. The glycoprotease antigen as detected by chemiluminescent Western blots of whole cell lysates from *P. haemolytica* serotype A1 (A), and serotype A11 (B), transformed with the pNF2176 vector and its derivatives. Lane 1, rGcp isolated from *E. coli* DH5α/pPH1.1ΔE; lane 2, rGgcp purified by affinity column chromatography; lane 3, host *P. haemolytica* strain; lanes 4-8, host strain transformed with pNF2176, pNFG1, pNFG2, pNFGG1, or pNFGG2.

A

1 2 3 4 5 6 7 8

B

1 2 3 4 5 6 7 8
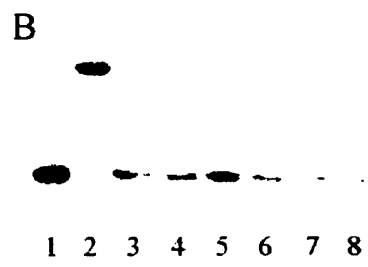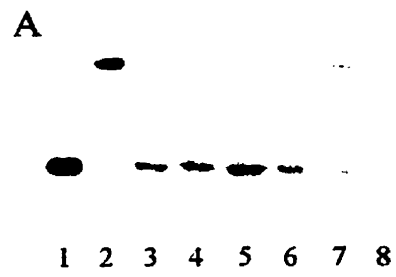
haemolytica A1 and A11, pNFG1 codes for an antigenic protein of 37 kDa, whereas both pNFGG1 and pNFGG2 code for an antigenic protein of 64 kDa (Figure 4.6). Comparable amounts of protein were loaded onto each gel for the Western immunoblot analysis, as determined by SDS-PAGE.

The PCR primer GGCP797L introduced a base modification (T→C) into an essential sequence of the *tac* promoter 74 nucleotides up-stream from the *ggcp* gene, which appears to prevent induction by IPTG. An attempt was made to induce gene expression in pNFGG1 and pNFGG2 in *E. coli* DH5α by the addition of 1 mM IPTG, but colourimetric dot immunoblots of induced strains did not reveal any protein expression (data not shown). The lack of a functional *tac* promoter precluded the attempted inducible expression of rGgcp in the *P. haemolytica* serotypes.

IV. Discussion

A major problem encountered in the production of recombinant *P. haemolytica* glycoprotease in *E. coli* is the aberrant folding of the recombinant protein into a biologically inactive, disulphide-bonded periplasmic aggregate. The expression of the GST-fusion protein, rGgcp, allowed for the ready isolation of the recombinant protein from other proteins, including enzyme inhibitors, found in *E. coli* lysates. The Gcp protein can be cleaved from the GST carrier, to provide a substrate for the *in vitro* refolding of recombinant protein. Despite the attainment of purities of rGcp much higher than those obtained in the previous study (Watt, 1995), this strategy did not yield a biologically active form of the glycoprotease.

The bi-directional cloning of the genes for rGcp and rGgcp in the shuttle vector pNF2176 resulted in the construction of vectors for the expression of *gcp* (pNFG1 and

pNFG2), and *ggcp* (pNFGG1 and pNFGG2) in *P. haemolytica* serotypes A1 and A11.

The expression of the cloned genes in pNFG1 and pNFGG1 was under the control of the plasmid sulfonamide promoter. The promoter was able to initiate transcription, as evidenced by the low levels of a distinct 64 kDa band corresponding to rGgcp expressed in strains bearing pNFGG1. A low level of plasmid protein expression was also observed in pNFGG2. This could be attributed to the promoter for the β-lactamase gene, which is oriented in the opposite direction to that of the sulfonamide promoter. High level expression from the inducible *E. coli tac* promoter was abolished during the engineering of a *Sal*I site into the gene. An approximately two-fold increase in the level of antigenic protein was also seen in strains bearing pNFG1, compared with the strain and vector control.

The expression of the recombinant proteins in their native host might be expected to correlate with increased biological activity. However, no enzymatic activity above that of the wild type activity was observed. This finding was supported by the lack of glycoprotease activity seen in the original transformed parental strain of serotype A11. The unexpected appearance of glycoprotease activity in culture supernatants of untransformed *P. haemolytica* A11, which were descendants of the parental strain, suggests that expression of the *gcp* gene in serotype A11 is under a different type of control than in those serotypes like A1, which show constitutive expression of *gcp*. Occasionally it has been observed in our laboratory that subculture of serotype A1 can lead to the loss of glycoprotease enzyme activity. The genetic or environmental factors that lead to loss or acquisition of glycoprotease secretion need further investigation.

CHAPTER 5. Summary and conclusions.

The *P. haemolytica* O-sialoglycoprotein endopeptidase as a GST fusion protein.

The primary objective of this thesis was the production of a purified biologically active recombinant glycoprotease comprising the glutathione-*S*-transferase of *Schistosoma japonicum* and the *P. haemolytica* glycoprotease (rGgcp). A number of difficulties encountered previously in the production of the recombinant glycoprotease in *E. coli* had to be overcome. The recombinant glycoprotease is not secreted as in its native host, but accumulates as a disulphide-linked aggregate in the periplasmic space. Two strategies were previously employed to prevent periplasmic localization of rGcp. The first involved the fusing of the *E. coli* hemolysin secretion signal to the C-terminus of the glycoprotease, to facilitate the export of the protein to the culture supernatant. A second method was the N-terminal fusion of thioredoxin to the glycoprotease (rTRXGCP) which resulted in the retention of the product within the reducing environment of the cytoplasm. However, though both methods greatly reduced aggregation via oxidation, neither resulted in productive folding to yield significant enzymatic activity. A small amount of glycoprotease activity was recovered after enterokinase cleavage of rTRXGCP, and by the partial purification of rTRXGCP (and its cleavage product, rGcp) by DEAE-chromatography. These results, together with *in vitro* refolding studies with the thiol reductase/isomerase, PDI, and the chaperones. DnaK/DnaJ/GrpE, indicate that rGcp is misfolded when expressed in *E. coli*.

The expression of a GST fusion protein, the 64 kDa rGgcp, resulted in the production of a soluble, cytosolic product (Figure 3.8). Unlike the thioredoxin fusion protein, the GST moiety forms a disulphide-linked dimer in solution, probably by

interactions between surface exposed cysteine residues on the GST domain (Kaplan *et al.*, 1997). Like many other GST fusions, rGgcp exists as a dimeric species under non-reducing conditions (Figure 3.8). The protein could be purified to homogeneity by affinity chromatography and the carrier GST could be selectively removed by thrombin, so that the objective of obtaining a more highly purified form of the recombinant glycoprotease was met. However, unlike the thioredoxin fusion protein, biological activity was not consistently observed in any of the purified or thrombin cleaved samples. The expression of rGcp and rGgcp in a number of *E. coli* host strains, including strains chosen for low protease or defective periplasmic disulphide bond formation, did not result in biological activity.

Numerous methods were tested for the *in vitro* refolding of the purified rGgcp, including the use of molecular chaperones. Despite extensive replication, and a wide range of experimental conditions employed, putative glycoprotease activity in the transformed *E. coli* clones was only observed in two samples. It is of interest to note that these two samples were both cleavage products from which the carrier protein, GST, was removed. A second common feature was that both active samples were obtained in the presence of the zwitterionic detergent CHAPS, which can affect folding by masking exposed hydrophobic regions during refolding. The third common aspect for this anomalous productive refolding is that it was in the presence of the enzyme PDI which possesses thiol isomerase and reductase activities. However, all efforts to reproduce the conditions of productive folding were fruitless, and no protocol for achieving the formation of the correct native conformation of rGcp was devised.

Expression of rGcp and rGgcp in *P. haemolytica*.

The *E. coli* - *P. haemolytica* shuttle vector pNF2176 was used for the expression of rGcp and rGgcp in *P. haemolytica* serotypes A1 and A11. Both recombinant gene products were expressed in *Pasteurella*, albeit at extremely low levels (< 1 µg protein/ml of culture). There was no increased glycoprotease activity in lysates or purified fractions from any of the transformed hosts. It was determined that a *P. haemolytica* A11 strain that had previously been shown to be negative for glycoprotease activity, gave rise to enzymatically active glycoprotease in the course of subsequent culture. This is the first report of glycoprotease activity from *P. haemolytica* serotype A11, and indicates that there are aspects of gene expression in *P. haemolytica* serotypes that are not fully understood. The expression of some *P. haemolytica* genes, such as the 100 kDa bovine transferrin receptor, depends on a low extracellular concentration of iron (Davies *et al.*, 1994). In our laboratories, previous studies have tried to modulate Gcp expression by the addition of various metal ions or metal ion chelators to the culture media, but no effects on Gcp expression were detected (Mellors, personal communication).

The low levels of recombinant protein expression induced from the sulfonamide promoter of pNF2176 may have hindered the detection of biological activity. However, in wild type *P. haemolytica* A1, the low level of constitutive expression of the Gcp protein, as determined by immunoblots, does not prevent the ready detection of enzymatic activity in the culture supernatant or subcellular fractions. Nevertheless, it may prove useful to identify an inducible high-expression promoter, recognized in *P. haemolytica*. This could be achieved by the construction of a novel vector, with an altered promoter region, but derived by replacement or alteration of the sulfonamide promoter of pNF2176. The

development of shuttle vectors sensitive to the induction of high levels of recombinant proteins in *P. haemolytica* would facilitate the further study of the glycoprotease and other *P. haemolytica* virulence factors.

This thesis has shown that the recombinant *P. haemolytica* glycoprotease can be purified as a GST-Gcp fusion protein by affinity chromatography and the rGcp moiety can be isolated by thrombin cleavage. Although the rGcp was not biologically active, it is expressed to high levels and can serve as another source of antigen for the production of monoclonal antibodies, or as a vaccine supplement. The purified rGcp could also serve as a suitable substrate for other studies on protein folding and molecular chaperones.

This thesis also demonstrates the first use of the shuttle vector, pNF2176, for the cloning and expression of the glycoprotease as a recombinant protein in *P. haemolytica*. The pNF2176 plasmid and its derivatives replicate and express heterologous genes in serotype A11, indicating the potential of this construct as a broad-range host vector. Here I report for the first time the presence of glycoprotease activity in cultures of *P. haemolytica* serotype A11, an organism which is known to possess a different genetic organisation of the gene, when compared to other A serotypes. Delineation of the environmental and genetic factors which affect the expression of the enzyme in serotype A11 could extend our knowledge of the control of virulence factors in *Pasteurellaceae*.

Literature Cited.

Abdullah, K.M., Lo, R.Y.C., and Mellors, A. 1990. Distribution of glycoprotease activity and the glycoprotease gene among serotypes of *Pasteurella haemolytica*. Biochem. Soc. Trans. 18: 901-3.

Abdullah, K.M. 1991. Biochemical and genetic characterization of a *Pasteurella haemolytica* glycoprotease. Ph. D. Thesis. University of Guelph, Guelph, Ontario, Canada.

Abdullah, K.M., Lo, R.Y.C., and Mellors, A. 1991. Cloning, nucleotide sequence, and expression of the *Pasteurella haemolytica* A1 glycoprotease gene. J. Bacteriol. 173: 5597-603.

Abdullah, K.M., Udoh, E.A., Shewen, P.E., and Mellors, A. 1992. A neutral glycoprotease of *Pasteurella haemolytica* A1 specifically cleaves *O*-sialoglycoproteins. Infect. Immun. 60: 56-62.

Aigner, S., Sthoeger, Z.M., Fogel, M., Weber, E., Zarn, J., Ruppert, M., Zeller, Y., Vestweber, D., Stahel, R., Sammar, M., and Altevogt, P. 1997. CD24, a mucin-type glycoprotein, is a ligand for P-selectin on human tumor cells. Blood. 89(9): 3385-95.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. 1990. Basic local alignment search tool. J. Mol. Biol. 215: 403-10.

Altschul, S.F., Madden, T.L., Schaumlaffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucl. Acids. Res. 25: 3389-3402.

Anfinsen, C.B. 1973. Principles that govern the folding of protein chains. Science 181: 223-30.

Azad, A.K., Coote, J.G., and Parton, R. 1994. Construction of conjugative shuttle and suicide vectors for *Pasteurella haemolytica* and *P. multocida*. Gene 145: 81-5.

Barbour, E.K., Nabbut, N.H., Hamadeh, S.K., and Al-Nakhli, H.M. 1997. Bacterial identity and characteristics in healthy and unhealthy respiratory tracts of sheep and calves. Vet. Res. Commun. 21: 421-30.

Bardwell, J.C., Lee, J.O., Jander, G., Martin, N., Belin, D., and Beckwith, J. 1993. A pathway for disulfide bond formation *in vivo*. Proc. Natl. Acad. Sci. 90: 1038-42.

Barnes, W.M. 1977. Plasmid detection and sizing in single colony lysates. Science. 195: 393-4.

Barnum, D.A. 1990. Socio-economic significance of the HAP group. Can. J. Vet. Res. 54: S1-S5.

Baumann, M., Simon, H., Schnieder, K.M., Danneel, H.J., Küster, U., and Giffhorn, F. 1989. Susceptibility of *Rhodobacter sphaeroides* to β-lactam antibiotics: Isolation and characterization of a periplasmic β-lactamase (Cephalosporinase). J. Bacteriol. 171: 308-313.

Bennett, T.A., Schammel, C.M., Lynam, E.B., Guyer, D.A., Mellors, A., Edwards, B., Rogelj, S., and Sklar, L.A. 1995. Evidence for a third component in neutrophil aggregation: potential roles of *O*-linked glycoproteins as L-selectin counter-structures. J. Leuk. Biol. 58: 510-8.

Bergey's Manual of Determinative Bacteriology. Holt, J.G., Krieg, N.R., Sneath, P.H.A., Staley, J.T., and Williams, S.T. (Eds.). Ninth Edition. 1994. Williams & Wilkins, Baltimore, MD.

Biberstein, E.L. 1978. Biotyping and serotyping of *P. haemolytica*. In: Bergan, T., Norris, J.R. (Eds.). Methods in Microbiology. Vol. 10. New York, Academic Press Inc. pp. 253-69.

Biberstein, E.L. 1990. Our understanding of the *Pasteurellaceae*. Can. J. Vet. Res. 54: S78-S82.

Birnboim, H.C., and Doly, J. 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucl. Acids Res. 7: 1513-1523.

Blackall, P.J., Pahoff, J.L., Stephens, C.P., and Darvill, F.M. 1996. *In vitro* activity of ceftiofur against Australian isolates of the family *Pasteurellaceae* associated with respiratory disease in cattle and pigs. Aust. Vet. J. 74: 71.

Blanchard, D. 1990. Human red cell glycophorins: Biochemical and antigenic properties. Transfusion Med. Rev. 4: 170-86.

Bollag, D.M., and Edelstein, S.J. 1991. Protein methods. Wiley-Liss. New York.

Boyer, H.W., and Roulland-Dussoix, D. 1969. A complementation analysis of the restriction and modification of DNA in *Escherichia coli*. J. Mol. Biol. 41: 459-72.

Bradford, M.M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye complex. Anal. Biochem. 72: 248-254.

Brogden, K.A., DeBey, B., Audibert, F., Lehmkuhl, H., and Chedid, L. 1995. Protection of ruminants by *Pasteurella haemolytica* A1 capsular polysaccharide vaccines containing muramyl dipeptide analogs. Vaccine. 13: 1677-84.

Brogden, K., and Clarke, C. 1997. Increase of glycocalyx and altered lectin agglutination profiles of *Pasteurella haemolytica* A1 after incubation in bovine subcutaneous tissue chambers *in vivo* or in ruminant serum *in vitro*. Infect. Immun. 65: 957-63.

Brown, J.F., Leite, F., and Czuprynski, C.J. 1997. Binding of *Pasteurella haemolytica* leukotoxin to bovine leukocytes. Infect. Immun. 65: 3719-24.

Brunner, C.J., George III, A.L., and Hsu, C.L. 1989. Chemotactic response of bovine neutrophils to *Pasteurella haemolytica* culture fluid. Vet. Immun. Immunopath. 21: 279-92.

Bult, C.J., White, O., Olsen, G.J., Zhou, L., Fleischmann, R.D., Sutton, G.G., Blake, J.A., Fitzgerald, L.M., Clayton, R.A., Gocayne, J.D., Kerlavage, A.R., Dougherty, B.A., Tomb, J.F., Adams, M.D., Reich, C.I., Overbeek, R., Kirkness, E.F., Weinstock, K.G., Merrick, J.M., Glodek, A., Scott, J.L., Geoghagen, N.S.M., Weidman, J.F., Fuhrmann, J.L., Nguyen, D., Utterback, T.R., Kelley, J.M., Peterson, J.D., Sadow, P.W., Hanna, M.C., Cotton, M.D., Roberts, K.M., Hurst, M.A., Kaine, B.P., Borodovsky, M., Klenk, H.P., Fraser, C.M., Smith, H.O., Woese, C.R. and Venter, J.C. 1996. Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*. Science. 273: 1058-1073.

Chang, G.X., Mallard, B.A., Mowat, D.N., and Gallo, G.F. 1996. Effect of supplemental chromium on antibody responses of newly arrived feeder calves to vaccines and ovalbumin. Can. J. Vet. Res. 60: 140-4.

Chang, Y.F., Ma, D.P., Bai, H.Q., Young, R., Struck, D.K., Shin, S.J., and Lein, D.H. 1992. Characterization of plasmids with antimicrobial resistance genes in *Pasteurella haemolytica* A1. DNA Seq. 3: 89-97.

Charnock, C., Refseth, U.H., and Sirevåg, R. 1992. Malate dehydrogenase from *Chlorobium vibrioforme*, *Chlorobium tepidum*, and *Heliobacterium gestii*: Purification, characterization, and investigation of dinucleotide binding by dehydrogenases by use of empirical methods of protein sequence analysis. J. Bacteriol. 174: 1307-1313.

Chong, S., Mersha, F.B., Comb, D.G., Scott, M.E., Landry, D., Vence, L.M., Perler, F.B., Benner, J., Kucera, R.B., Hirvonen, C.A., Pelletier, J.J., Paulus, H., and Xu, M.Q. 1997. Single-column purification of free recombinant proteins using a self-cleavable affinity tag derived from a protein splicing element. Gene 192: 271-81.

Chumakov, A., and Koeffler, H.P. 1993. Specific interaction between adenoviral 55-kDa E1B protein and *in vivo* produced p53 fusion proteins. Gene 131: 231-6.

Cladman, W.M., Watt, M.A.V., Dini, J.P., and Mellors, A. 1996. The *Pasteurella haemolytica* O-sialoglycoprotein endopeptidase is inhibited by zinc ions and does not cleave fetuin. Biochem. Biophys. Res. Comm. 220: 141-6.

. Clark, R.A., Erickson, H.P., and Springer, T.A. 1997. Tenascin supports lymphocyte rolling. J. Cell Biol. 137: 755-765.

Clinkenbeard, K.D., Mosier, D.A., Timko, A.L., and Confer, A.W. 1989. Effects of *Pasteurella haemolytica* leukotoxin on cultured bovine lymphoma cells. Am. J. Vet. Res. 50: 271-5.

Clinkenbeard, K.D., and Upton, M.L. 1991. Lysis of bovine platelets by *Pasteurella haemolytica* leukotoxin. Am. J. Vet. Res. 52: 453-7.

Cohen, S.N., Chang, A.C.Y., and Hsu, L. 1972. Nonchromosomal antibiotic resistance in bacteria: genetic transformations of *Escherichia coli* by R-factor DNA. Proc. Nat. Acad. Sci. 69: 2110-4.

Collier, J.R. 1968. Significance of bacteria in bovine respiratory disease. J. Am. Vet. Med. Assoc. 153: 1645-51.

Confer, A.W., Panciera, K.D., Clinkenbeard, K.D., and Mosier, D.A. 1990. Molecular aspects of virulence of *Pasteurella haemolytica*. Can. J. Vet. Res. 54: S48-52.

Confer, A.W., McCraw, R.D., Durham, J.A., Morton, R.J., and Panciera, R.J. 1995. Serum antibody responses of cattle to iron-regulated outer membrane proteins of *Pasteurella haemolytica* A1. Vet. Immun. Immunopath. 47: 101-10.

Confer, A.W., Clinkenbeard, K.D., Gatewood, D.M., Driskel, B.A., and Montelongo, M. 1997. Serum antibody responses of cattle vaccinated with partially purified native *Pasteurella haemolytica* leukotoxin. Vaccine. 15: 1423-9.

Conlon, J.A., Shewen, P.E., and Lo, R.Y.C. 1991. Efficacy of recombinant leukotoxin in protection against pneumonic challenge with live *Pasteurella haemolytica* A1. Infect. Immun. 59: 587-91.

Conlon, J.A., and Shewen, P.E. 1993. Clinical and serological evaluation of a *Pasteurella haemolytica* A1 capsular polysaccharide vaccine. Vaccine. 11: 767-72.

Cornelis, P., Digneffe, C., and Willemot, K. 1982. Cloning and expression of a *Bacillus coagulans* amylase gene in *Escherichia coli*. Mol. Gen. Genet. 186: 507-511.

Craig, F.F., Coote, J.G., Parton, R., Freer, J.H., and Gilmour, N.J. 1989. A plasmid which can be transferred between *Escherichia coli* and *Pasteurella haemolytica* by electroporation and conjugation. J. Gen. Microbiol. 135: 2885-90.

Davies, R.L., McCluskey, J., Gibbs, H.A., Coote, J.G., Freer, J.H., and Parton, R. 1994. Comparison of outer-membrane proteins of *Pasteurella haemolytica* expressed *in vitro* and *in vivo* in cattle. Microbiol. 140: 3293-300.

Davies, R.L., Paster, B.J., and Dewhirst, F.E. 1996. Phylogenetic relationships and diversity within the *Pasteurella haemolytica* complex based on 16S rRNA sequence comparison and outer membrane protein and lipopolysaccharide analysis. Int. J. Syst. Bacteriol. 46: 736-44.

Davies, R.L., Arkinsaw, S., and Selander, R.K. 1997. Evolutionary genetics of *Pasteurella haemolytica* isolates recovered from cattle and sheep. Infect. Immun. 65: 3585-93.

de Crouy-Chanel, A., Kohiyama, M., and Richarme, G. 1995. A novel function of *Escherichia coli* chaperone DnaJ. Protein disulfide isomerase. J. Biol. Chem. 270: 22669-72.

Deckert, G., Warren, P.V., Gaasterland, T., Young, W.G., Lenox, A.L., Graham, D.E., Overbeek, R., Snead, M.A., Keller, M., Aujay, M., Huber, R., Feldman, R.A., Short, J.M., Olson, G.J., and Swanson, R.V. 1998. The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*. Nature 392: 353-358.

Dodge, J.T., Mitchell, C., and Hanahan, D.J. 1963. The preparation and chemical characteristics of hemoglobin-free ghosts of human erythrocytes. Arch. Biochem. Biophys. 100: 119-130.

Durham, J.A., Antone, S.M., Cunningham, M.W., and Confer, A.W. 1988. Monoclonal antibodies to *Pasteurella haemolytica* serotype 1. Lipopolysaccharide: demonstration of antigenic similarities among several serotypes. J. Clin. Microbiol. 26: 885-9.

Ellis, R.J., and Hemmingsen, S.E. 1989. Proteins as molecular chaperones. Nature 328: 378-9.

Fedorov, A.N., and Baldwin, T.O. 1997. GroE modulates kinetic partitioning of folding intermediates between alternative states to maximize the yield of biologically active protein. J. Mol. Biol. 268: 712-723.

Fedorova, N.D., and Highlander, S.K. 1997a. Plasmids for heterologous expression in *Pasteurella haemolytica*. Gene 186: 207-211.

Fedorova, N.D., and Highlander, S.K. 1997b. Generation of targeted nonpolar gene insertions and operon fusions in *Pasteurella haemolytica* and creation of a strain that produces and secretes inactive leukotoxin. Infect. Immun. 65: 2593-8.

Fleischmann, R.D., Adams, M.D., White, O., Clayton, R.A., Kirkness, E.F., Kerlavage, A.R., Bult, C.J., Tomb, J.F., Dougherty, B.A., Merrick, J.M., McKenney, K., Sutton, G., Fitzhugh, W., Fields, C.A., Gocayne, J.D., Scott, J.D., Shirley, R., Liu, L.I., Glodek, A., Kelley, J.M., Weidman, J.F., Phillips, C.A., Spriggs, T., Heldblom, E., Cotton, M.D., Utterback, T.R., Hanna, M.C., Nguyen, D.T., Saudek, D.M., Brandon, R.C., Fine, L.D., Fritchman, J.L., Fuhrmann, J.L., Geoghagen, N.S.M., Gnehm, C.L., McDonald, L.A., Small, K.V., Fraser, C.M., Smith, H.O., and Venter, J.C. 1995. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. Science. 269(5223): 496-512.

Fodor, L., Pénzes, Z., and Varga, J. 1996. Coagglutination test for serotyping *Pasteurella haemolytica*. J. Clin. Microbiol. 34: 393-7.

Frangioni, J. V., and Neel, B.G. 1993. Solubilization and purification of enzymatically active glutathione-*S*-transferase. Anal. Biochem. 210: 179-187.

Frank, G.H., and Wessman, G.E. 1978. Rapid plate agglutination procedure for serotyping *Pasteurella haemolytica*. J. Clin. Microbiol. 7: 142-5.

Frank, G.H., and Tabatabai, L.B. 1981. Neuraminidase activity of *Pasteurella haemolytica* isolates. Infect. Immun. 32: 1119-22.

Frank, G.H., and Smith, P.C. 1983. Prevalence of *Pasteurella haemolytica* in transported cattle. Am. J. Vet. Res. 44: 981-5.

Fraser, C.M., Casjens, S., Huang, W.M., Sutton, G.G., Clayton, R.A., Lathigra, R., White, O., Ketchum, K.A., Dodson, R., Hickey, E.K., Gwinn, M., Dougherty, B., Tomb, J.F., Fleischmann, R.D., Richardson, D., Peterson, J., Kerlavage, A.R., Quackenbush, J., Salzberg, S., Hanson, M., van Vugt, R., Palmer, N., Adams, M.D., Gocayne, J.D., Weidman, J., Utterback, T., Watthey, L., McDonald, L., Artiach, P., Bowman, C., Garland, S., Fujii, C., Cotton, M.D., Horst, K., Roberts, K., Hatch, B., Smith, H.O., and Venter, J.C. 1997. Genomic sequence of a Lyme disease spirochete, *Borrelia burgdorferi*. Nature 390: 580-586.

Fraser, C.M., Gocayne, J.D., White, O., Adams, M.D., Clayton, R.A., Fleischmann, R.D., Bult, C.J., Kerlavage, A.R., Sutton, G., Kelley, J.M., Fritchman, J.L., Weidman, J.F., Small, K.V., Sandusky, M., Fuhrmann, J.L., Nguyen, D.T., Utterback, T.R., Saudek, D.M., Phillips, C.A., Merrick, J.M., Tomb, J.F., Dougherty, B.A., Bott, K.F., Hu, P.C., Lucier, T.S., Peterson, S.N., Smith, H.O., Hutchinson III, C.A., and Venter, J.C. 1995. The minimal gene complement of *Mycoplasma genitalium*. Science. 270: 397-403.

Freedman, R.B. 1989. Protein disulphide isomerase: multiple roles in the modification of nascent secretory proteins. Cell. 57: 1069-72.

Frey, J. 1992. Construction of a broad host range shuttle vector for gene cloning and expression in *Actinobacillus pleuropneumoniae* and other *Pasteurellaceae*. Res. Microbiol. 143: 263-9.

Friend, S.C., Thomson, R.G., and Wilkie, B.N. 1977. Pulmonary lesions induced by *Pasteurella haemolytica* in cattle. Can. J. Comp. Med. 41: 219-223.

Gesteland, R.F. 1966. Isolation and characterization of ribonuclease I mutants of *Escherichia coli*. J. Mol. Biol. 16: 67-84.

Guan, K., and Dixon, J.E. 1991. Eukaryotic proteins expressed in *Escherichia coli*: an improved thrombin cleavage and purification procedure of fusion proteins with glutathione-*S*-transferase. Anal. Biochem. 192: 262-7.

Hakes, D.J., and Dixon, J.E. 1992. New vectors for high level expression of recombinant proteins in bacteria. Anal. Biochem. 202: 293-8.

Hanahan, D. 1983. Studies on transformation of *Escherichia coli* with plasmids. J. Mol. Biol. 166: 557-80.

Hesterkamp, T., Hauser, S., Lütcke, H., and Bukau, B. 1996. *Escherichia coli* trigger factor is a prolyl isomerase that associates with nascent polypeptide chains. Proc. Nat. Acad. Sci. 93: 4437-41.

Himmelreich, R., Hilbert, H., Plagens, H., Pirkl, E., Li, B.C., and Herrmann, R. 1996. Complete sequence analysis of the genome of the bacterium *Mycoplasma pneumoniae*. Nucleic Acids Res. 24: 4420-4449.

Hodgins, D.C., and Shewen, P.E. 1996. Preparturient vaccination to enhance passive immunity to the capsular polysaccharide of *Pasteurella haemolytica* A1. Vet. Immun. Immunopath. 50: 67-77.

Hooper, N.M. 1994. Families of zinc metalloproteases. FEBS Lett. 354: 1-6.

Hu, R.H., Mellors, A., and Bhavanandan, V.P. 1994. Cleavage of epitectin, a mucin-type sialoglycoprotein from the surface of human laryngeal carcinoma cells by a glycoprotease from *Pasteurella haemolytica*. Arch. Biochem. Biophys. 310: 300-309.

Isacsson, V., and Wettermark, G. 1974. Chemiluminescence in analytical chemistry. Anal. Chim. acta. 68: 339-62.

Ishikawa, T. 1992. The ATP-dependent glutathione S-conjugate export pump. Trends Biochem. Sci. 17: 463-8.

Jiang, P., and Mellors, A. 1998. Membrane protein proteolysis assayed by fluorescence quenching: Assay of O-sialoglycoprotein endopeptidase. Anal. Biochem. In Press.

Kaneko, T., Sato, S., Kotani, H., Tanaka, A., Asamizu, E., Nakamura, Y., Miyajima, N., Hirosawa, M., Sugiura, M., Sasamoto, S., Kimura, T., Hosouchi, T., Matsuno, A., Muraki, A., Nakazaki, N., Naruo, K., Okumura, S., Shimpo, S., Takeuchi, C., Wada, T., Watanabe, A., Yamada, M., Yasuda, M., and Tabata, S. 1996. Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis sp.* strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions. DNA Res. 3(3): 109-136.

Kaplan, W., Hüsler, P., Klump, H., Erhardt, J., Sluis-Cremer, N., and Dirr, H. 1997. Conformational stability of pGEX-expressed *Schistosoma japonicum* glutathione S-transferase: a detoxification enzyme and fusion-protein affinity tag. Protein Sci. 6: 399-406.

Kemperman, H., Wijnands, Y., Wesseling, J., Niessen, C.M., Sonnenberg, A., and Roos, E. 1994. The mucin epiglycanin on TA3/Ha carcinoma cells prevents $\alpha_6\beta_4$-mediated adhesion to laminin and kalinin and E-cadherin-mediated cell-cell interaction. J. Cell Biol. 127: 2071-80.

Klenk, H.P., Clayton, R.A., Tomb, J., White, O., Nelson, K.E., Ketchum, K.A., Dodson, R.J., Gwinn, M., Hickey, E.K., Peterson, J.D., Richardson, D.L., Kerlavage, A.R., Graham, D.E., Kyrpides, N.C., Fleischmann, R.D., Quackenbush, J., Lee, N.H., Sutton, G.G., Gill, S., Kirkness, E.F., Dougherty, B.A., McKenney, K., Adams, M.D., Loftus, B., Peterson, S., Reich, C.I., McNeil, L.K., Badger, J.H., Glodek, A., Zhou, L., Overbeek, R., Gocayne, J.D., Weidman, J.F., McDonald, L., Utterback, T., Cotton, M.D., Spriggs, T., Artiach, P., Kaine, B.P., Sykes, S.M., Sadow, P.W., D'Andrea, K.P., Bowman, C., Fujii, C., Garland, S.A., Mason, T.M., Olsen, G.J., Fraser, C.M., Smith, H.O., Woese, C.R., and Venter, J.C. 1997. The complete genome sequence of the hyperthermophilic, sulphate-reducing archaeon *Archaeoglobus fulgidus*. Nature. 390: 364-370.

Kricka, L.J. 1991. Chemiluminescent and bioluminescent techniques. Clin. Chem. 37: 1472-81.

Laemmli, U.K. 1970. Cleavage of structural proteins during the assembly of the head bacteriophage T4. Nature 227: 680-685.

Lee, C.W., Lo, R.Y.C., Shewen, P.E., and Mellors, A. 1994a. The detection of the sialoglycoprotease gene and assay for sialoglycoprotease activity among isolates of *Pasteurella haemolytica* A1 strains, serotypes A13, A14, T15 and A16. FEMS Microbiol. Lett. 121: 199-206.

Lee, C.W., Shewen, P.E., Cladman, W.M., Conlon, J.A.R., Mellors, A., and Lo, R.Y.C. 1994b. Sialoglycoprotease of *Pasteurella haemolytica* A1: Detection of anti-sialoglycoprotease antibodies in sera of calves. Can. J. Vet. Res. 58: 93-8.

Lee, C.W., and Shewen, P.E. 1996. Evidence of bovine immunoglobulin G1 (IgG1) protease activity in partially purified culture supernate of *Pasteurella haemolytica* A1. Can. J. Vet. Res. 60: 127-32.

Li, F., Wilkins, P.P., Crawley, S., Weinstein, J., Cummings, R.D., and McEver, R.P. 1996. Post-translational modifications of recombinant P-selectin glycoprotein ligand-1 required for binding to P- and E-selectin. J. Biol. Chem. 271: 3255-64.

Livrelli, V., Peduzzi, J., and Joly, B. 1991. Sequence and molecular characterization of the ROB-1 β-lactamase gene from *Pasteurella haemolytica*. Antimicrob. Agents Chemother. 35: 242-51.

Lo, R.Y.C., Strathdee, C.A., and Shewen, P.E. 1987. Nucleotide sequence of the leukotoxin genes of *Pasteurella haemolytica* A1. Infect. Immun. 55: 1987-96.

Lo, R.Y.C., Watt, M.A., Gyroffy, S., and Mellors, A. 1994. Preparation of recombinant glycoprotease of *Pasteurella haemolytica* A1 utilizing the *Escherichia coli* α-hemolysin secretion system. FEMS Microbiol. Lett. 116: 225-30.

Maheswaran, S.K., Kannan, M.S., Weiss, D.J., Reddy, K.R., Townsend, E.L., Yoo, H.S., Lee, B.W., and Whiteley, L.O. 1993. Enhancement of neutrophil-mediated injury to bovine pulmonary endothelial cells by *Pasteurella haemolytica* leukotoxin. Infect. Immun. 61: 2618-2625.

Mannori, G., Crottet, P., Cecconi, O., Hanasaki, K., Aruffo, A., Nelson, R.M., Varki, A., and Bevilacqua, M.P. 1995. Differential colon cancer cell adhesion to E-, P-, and L-selectin: role of mucin-type glycoproteins. Cancer Res. 55: 4425-31.

Markwell, M.A.K. 1982. A new solid-state reagent to iodinate proteins: conditions for the efficient labeling of antiserum. Anal. Biochem. 125: 427-32.

Marsh, J.C.W., Sutherland, D.R., Davidson, J., Mellors, A., and Keating, A. 1992. Retention of progenitor cell function in CD34⁻ cells purified using a novel O-sialoglycoprotease. Leukemia. 6(9): 926-934.

McBride, J.W., Corstvet, R.E., Paulsen, D.B., McClure, J.R., and Enright, F.M. 1996. Characterization of bovine pulmonary and serum antibody responses after parenteral or intrapulmonary vaccination with live *Pasteurella haemolytica*. Comp. Immun. Microbiol. Infect. Dis. 19: 99-115.

Mdurvwa, E.G., and Brunner, C.J. 1996. Bovine neutrophil activation by culture fluid from *Pasteurella haemolytica* serotypes A1 and A11. Vet. Microbiol. 41: 311-9.

Mellors, A., and Lo, R.Y.C. 1995. The *O*-sialoglycoprotease of *Pasteurella haemolytica*. Meth. Enzymol. 248: 728-739.

Mendoza, J.A., Rogers, E., Lorimer, G.H., and Horowitz, P.M. 1991. Unassisted refolding of urea-unfolded rhodanese. J. Biol. Chem. 266: 13587-91.

Missiakas, D., and Raina, S. 1997. Protein folding in the bacterial periplasm. J. Bacteriol. 2465-71.

Morck, D.W., Olson, M.E., Acres, S.D., Daoust, P.Y., and Costerton, J.W. 1989. Presence of bacterial glycocalyx and fimbriae on *P. haemolytica* in feedlot cattle with pneumonic pasteurellosis. Can. J. Vet. Res. 53: 167-71.

Morck, D.W., Ellis, B.D., Domingue, P.A.G., Olson, M.E., and Costerton, J.W. 1991. *In vivo* expression of iron regulated outer-membrane proteins in *Pasteurella haemolytica* A1. Microbial. Pathogen. 11: 1-6.

Morton, R.J., Simons, K.R., and Confer, A.W. 1996. Major outer membrane proteins of *P. haemolytica* serovars 1-15: comparison of separation techniques and surface-exposed proteins on selected serovars. Vet. Microbiol. 51: 319-30.

Murphy, G.L., Robinson, L.C., and Burrows, G.E. 1993. Restriction endonuclease analysis and ribotyping differentiate *Pasteurella haemolytica* serotype A1 isolates from cattle within a feedlot. J. Clin. Microbiol. 31: 2303-8.

Nesin, M., Lupski, J.R., Svec, P. and Godson, G.N. 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.

Neu, H.C., and Heppel, L.A. 1965. The release of enzymes from *Escherichia coli* by osmotic shock and during the formation of spheroplasts. J. Biol. Chem. 240(9): 3685-92.

Neurath, H. 1989. The diversity of proteolytic enzymes. In: Proteolytic enzymes: a practical approach. R.J. Beynon and J.S. Bond (Eds.). Oxford: IRL Press. pp. 1-13.

Newsome, I.E., and Cross, F. 1932. Some bipolar organisms found in pneumonia in sheep. J. Am. Vet. Med. Assoc. 80: 711-9.

Nielsen, H., Engelbrecht, J., Brunak, S., and von Heijne, G. 1997. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Engineering. 10: 1-6.

116

Otulakowski, G.L., Shewen, P.E., Udoh, A.E., Mellors, A., and Wilkie, B.N. 1983. Proteolysis of sialoglycoprotein by *Pasteurella haemolytica* cytotoxic culture supernatant. Infect. Immun. 42: 64-70.

Parameshwar, J.M., Murphy, G.L., Wyckoff III, J.H., Farmer, S., Hancock, R.E.W., and Confer, A.W. 1997. Purification and partial characterization of the OmpA family of proteins of *Pasteurella haemolytica*. Infect. Immun. 65: 211-8.

Panciera, R.J., and Corstvet, R.E. 1984. Bovine pneumonic pasteurellosis: model for *Pasteurella haemolytica*- and *Pasteurella multocida*-induced pneumonia in cattle. Am. J. Vet. Res. 45: 2532-7.

Panciera. R.J., Corstvet, R.E., Confer, A.W., and Gresham, C.N. 1984. Bovine pneumonic pasteurellosis. Effect of vaccination with live *Pasteurella* species. Am. J. Vet. Res. 45: 2538-42.

Petras, S.F., Chidambaram, M., Illyes, E.F., Froshauer, S., Weinstock, G.M., and Reese, C.P. 1995. Antigenic and virulence properties of *Pasteurella haemolytica* leukotoxin mutants. Infect. Immun. 63: 1033-9.

Philipp, W.J., Poulet, S., Eiglmeier, K., Pascopella, L., Balasubramanian, V., Heym, B., Bergh, S., Bloom, B.R., Jacobs Jr., W.R., and Cole, S.T. 1996. An integrated map of the genome of the tubercle bacillus, *Mycobacterium tuberculosis* H37Rv, and comparison with *Mycobacterium leprae*. Proc. Natl. Acad. Sci. 93: 3132-3137.

Purdy, C.W., Raleigh, R.H., Collins, J.K., Watts, J.L., and Straus, D.C. 1997. Serotyping and enzyme characterization of *Pasteurella haemolytica* and *Pasteurella multocida* isolates recovered from pneumonic lungs of stressed feeder calves. Curr. Microbiol. 34: 244-9.

Rehmtulla, A.J., and Thomson, R.G. 1981. A review of the lesions in shipping fever of cattle. Can. Vet. J. 22: 1-8.

Romanos, M.A., Hughes, F.J., Comerford, S.A., and Scorer, C.A. 1995. Production of a phosphorylated GST::HPV-6 E7 fusion protein using a yeast expression vector and glutathione-*S*-transferase fusions. Gene 152: 137-8.

Rounsley, S.D., Lin, X., Ketchum, K.A., Crosby, M.L., Brandon, R.C., Sykes, S.M., Mason, T.M., Kerlavage, A.R., Adams, M.D., Somerville, C.R., and Venter, J.C. 1997. *Arabidopsis thaliana* chromosome II BAC F4L23 genomic sequence. Unpublished. Genbank Accession 2583127.

Rudert, F., Visser, E., Gradl, G., Grandison, P., Shemshedini, L., Wang, Y., Grierson, A., and Watson, J. 1996. pLEF, a novel vector for expression for expression of glutathione-*S*-transferase fusion proteins in mammalian cells. Gene 169: 281-2.

Sadaie, Y., Yata, K., Fujita, M., Sagai, H., Itaya, M., Kasahara, Y., and Ogasawara, N. 1997. Nucleotide sequence and analysis of the *phoB-rrnE-groESL* region of the *Bacillus subtilis* chromosome. Microbiol. 143: 1861-1866.

Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B., and Erlich, H.A. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science. 239: 487-91.

Salmi, M., and Jalkanen, S. 1996. Human vascular adhesion protein 1 (VAP-1) is a unique sialoglycoprotein that mediates carbohydrate-dependent binding of lymphocytes to endothelial cells. J. Exp. Med. 183: 569-79.

Sambrook, J., E.F. Fritsch, and Maniatis, T. 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Sancar, A., Hack, A.M., and Rupp, W.D. 1979. Simple method for identification of plasmid-coded proteins. J. Bacteriol. 137: 692-3.

Schröder, H., Langer, T., Hartl, F.U., and Bukau, B. 1993. DnaK, DnaJ and GrpE form a cellular chaperone machinery capable of repairing heat-induced protein damage. EMBO J. 12: 4137-44.

Schumacher, R.J., Hansen, W.J., Freeman, B.C., Alnemri, E., Litwack, G., and Toft, D.O. 1996. Cooperative action of Hsp70, Hsp90, and DnaJ proteins in protein renaturation. Biochem. 35: 14889-98.

Segrest, J.P., Wilkinson T.M., and Sheng L. 1979. Isolation of glycophorin with deoxycholate. Biochim. Biophys. Acta 554: 533-7.

Seong, B.L., Lee, C.P., and RajBhandary, U.L. 1989. Suppression of amber codons *in vivo* as evidence that mutants derived from *Escherichia coli* initiator tRNA can act at the step of elongation in protein synthesis. J. Biol. Chem. 264: 6504-8.

Shewen, P., Sharp, A., and Wilkie, B.N. 1988. Efficacy testing a *P. haemolytica* extract vaccine. Vet. Med. 83: 1078-83.

Shewen, P., Perets, A., Lee, C.W., and Lo, R.Y.C. 1994. Evaluation of a *Pasteurella haemolytica* recombinant sialoglycoprotease fusion protein (rGcp-F) vaccine. HAP (*Haemophilus, Actinobacillus* and *Pasteurella*) International Conference. Edinburgh, Scotland, Abstract # P44.

Simon, M., Benit, P., Vassal, A., Dubois, C., and Faye, G. 1994. Sequence of the PHO2-POL3 (CDC2) region of chromosome IV of *Saccharomyces cerevisiae*. Yeast. 10: 1653-1656.

Slocombe, R.F., Mulks, M., Killingsworth, C.R., Derksen, F.J., and Robinson, N.E. 1990. Effect of *Pasteurella haemolytica*-derived endotoxin on pulmonary structure and function in calves. Am. J. Vet. Res. 51: 433-8.

Smalheiser, N.R., and Kim, E. 1995. Purification of cranin, a laminin binding membrane protein. J. Biol. Chem. 270: 15425-15433.

Smith, D.B., and Johnson, K.S. 1988. Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione-*S*-transferase. Gene 67: 31-40.

Smith, D.R., Doucette-Stamm, L.A., Deloughery, C., Lee, H.M., Dubois, J., Aldredge, T., Bashirzadeh, R., Blakely, D., Cook, R., Gilbert, K., Harrison, D., Hoang, L., Keagle, P., Lumm, W., Pothier, B., Qiu, D., Spadafora, R., Vicare, R., Wang, Y., Wierzbowski, J., Gibson, R., Jiwani, N., Caruso, A., Bush, D., Safer, H., Patwell, D., Prabhakar, S., McDougall, S., Shimer, G., Goyal, A., Pietrovski, S., Church, G.M., Daniels, C.J., Mao, J.I., Rice, P., Nolling, J., and Reeve, J.N. 1997. Complete genome sequence of *Methanobacterium thermoautotrophicum* ΔH: functional analysis and comparative genomics. J. Bacteriol. 179: 7135-7155.

Smith, G.R. 1961. The characteristics of two types of *Pasteurella haemolytica* associated with different pathological conditions of sheep. J. Path. Bacteriol. 81: 431-40.

Smith, R.F., and King, K.Y. 1995. Identification of a eukaryotic-like protein kinase gene in archaebacteria. Protein Sci. 4: 126-9.

Sneath, P.H.A. and Stevens, M. 1990. *Actinobacillus rossii* sp. nov., *Actinobacillus seminis* sp. nov., nom. rev. *Pasteurella bettii* sp. nov., *Pasteurella lymphangitidis* sp. nov., *Paseurellamairi* sp. nov., and *Pasteurella trehalosi* sp. nov. Int. J. Syst. Bacteriol. 40: 148-53.

Spertini, O., Cordey, A.S., Monai, N., Giuffre, L., and Schapira, M. 1996. P-selectin glycoprotein ligand 1 is a ligand for L-selectin on neutrophils, monocytes, and CD34+ hematopoietic progenitor cells. J. Cell Biol. 135: 523-31.

Sreevatsan, S., Ames, T.R., Werdin, R.E., Yoo, H.S., and Maheswaran, S.K. 1996. Evaluation of three experimental subunit vaccines against pneumonic pasteurellosis in cattle. Vaccine. 14: 147-54.

Srinand, S., Hsuan, S.L., Yoo, H.S., Maheswaran, S.K., Ames, T.R., and Werdin, R.E. 1996. Comparative evaluation of antibodies induced by commercial *Pasteurella haemolytica* vaccines using solid phase immunoassays. Vet. Microbiol. 49: 181-95.

Stark, M.J.R. 1987. Multicopy expression vectors carrying the *lac* repressor gene for regulated high-level expression of genes in *Escherichia coli*. Gene. 51: 255-267.

Steininger, C.N., Eddy, C.A., Leimgruber, R.M., Mellors, A., and Welply, J.K. 1992. The glycoprotease of *Pasteurella haemolytica* A1 eliminates binding of myeloid cells to P-selectin but not to E-selectin. Biochem. Biophys. Res. Comm. 188: 760-6.

Strathdee, C.A., and Lo, R.Y.C. 1989. Cloning, nucleotide sequence, and characterization of genes encoding the secretion function of the *Pasteurella haemolytica* leukotoxin determinant. J. Bacteriol. 171: 916-28.

Straus, D.C., Unbehagen, P.J., and Purdy, C.W. 1993a. Neuraminidase production by a *Pasteurella haemolytica* A1 strain associated with bovine pneumonia. Infect. Immun. 61: 253-9.

Straus, D.C., Jolley, W.L., and Purdy, C.W. 1993b. Characterization of neuraminidases produced by various serotypes of *Pasteurella haemolytica*. Infect. Immun. 61: 4669-74.

Straus, D.C., and Purdy, C.W. 1995. Extracellular neuraminidase production by *Pasteurella* species isolated from infected animals. Curr. Microbiol. 31: 312-5.

Sun, G., and Budde, R.J.A. 1995. A modified pGEX expression system that eliminates degradation products and thrombin from the recombinant protein. Anal. Biochem. 231: 458-60.

Sutherland, D.R., Abdullah, K.M., Cyopick, P., and Mellors, A. 1992. Cleavage of the cell-surface *O*-sialoglycoproteins CD34, CD43, CD44, and CD45 by a novel glycoprotease from *Pasteurella haemolytica*. J. Immunol. 148: 1458-64.

Swartz, J.R. 1996. Chapter 108: *Escherichia coli* recombinant DNA technology. In: *Escherichia coli* and *Salmonella* Cellular and Molecular Biology. Neidhardt, F.C., Ingraham, J.L., Low, K.B., Magasanik, B., Schaecter, M., Umbarger, H.E. Eds. American Society of Microbiology, Washington, D.C. pp. 1693-1711.

Szabo, A., Langer, T., Schröder, H., Flanagan, J., Budua, B., and Hartl, F.U. 1994. The ATP hydrolysis-dependent reaction cycle of the *Escherichia coli* Hsp70 system – DnaK, DnaJ, and GrpE. Proc. Nat. Acad. Sci. 91: 10345-9.

Thomas, J.G., Ayling, A., and Baneyx, F. 1997. Molecular chaperones, folding catalysts, and the recovery of active recombinant proteins from *Escherichia coli*. Appl. Biochem. Biotech. 66: 197-238.

Tomb, J.F., White, O., Kerlavage, A.R., Clayton, R.A., Sutton, G.G., Fleischmann, R.D., Ketchum, K.A., Klenk, H.P., Gill, S., Dougherty, B.A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E.F., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H.G., Glodek, A., McKenney, K., Fitzgerald, L.M., Lee, N., Adams, M.D., Hickey, E.K., Berg, D.E., Gocayne, J.D., Utterback, T.R., Peterson, J.D., Kelley, J.M., Cotton, M.D.,

Weidman, J.M., Fujil, C., Bowman, C., Watthey, L., Wallin, E., Hayes, W.S., Borodovsky, M., Karp, P.D., Smith, H.O., Fraser, C.M., and Venter, J.C. 1997. The complete genome sequence of the gastric pathogen *Helicobacter pylori*. Nature. 388: 539-547.

Towbin, H., Staehelin, T., and Gordon, J. 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc. Nat. Acad. Sci. 76: 4350-4354.

Udoh, A.E. 1986. The extracellular protease of *Pasteurella haemolytica*. Ph.D. thesis. University of Guelph, Guelph, Ontario, Canada.

Uhlén, M., and Moks, T. 1990. Gene fusions for purpose of expression: an introduction. Meth. Enzymol. 180: 129-43.

Uhlén, M., Forsberg, G., Moks, T., Hartmanis, M., and Nilsson, B. 1992. Fusion proteins in biotechnology. Curr. Opin. Biotechnol. 3: 363-9.

Vassort-Bruneau, C., Lesage-Descauses, M.C., Martel, J.L., Lafont, J.P., and Chaslus-Dancla, E. 1996. CAT III chloramphenicol resistance in *Pasteurella haemolytica* and *Pasteurella multocida* isolated from calves. J. Antimicrob. Chemother. 38: 205-13.

Watson, M.A. 1995. Definition of chemiluminescence and superoxide production responses of bovine neutrophils to selected soluble and particulate stimulants, and comparisons with the responses to *Pasteurella haemolytica*. Am. J. Vet. Res. 56: 1045-54.

Watt, M.A.V. 1997. Activation of the *Pasteurella haemolytica* O-sialoglycoprotease. Ph.D. thesis. University of Guelph, Guelph, Ontario, Canada.

Watt, M.AV., Lo, R.Y.C., and Mellors, A. 1997a. Refolding of recombinant *Pasteurella haemolytica* A1 glycoprotease expressed in an *Escherichia coli* thioredoxin gene fusion system. Cell stress and chaperone. 2(3): 80-90.

Watt, M.A.V., Mellors, A., and Lo, R.Y.C. 1997b. Comparison of the recombinant and authentic forms of the *Pasteurella haemolytica* A1 glycoprotease. FEMS Microbiol. Lett. 147: 37-43.

Weissman, J.S., and Kim, P.S. 1993. Efficient catalysis of disulphide bond rearrangements by protein disulphide isomerase. Nature. 365: 185-8.

Werner, M.H., Clore, G.M., Gronenborn, A.M., Kondoh, A., and Fisher, R.J. 1994. Refolding proteins by gel filtration chromatography. FEBS Lett. 345: 125-30.

West, S.E., Romero, M.J., Regassa, L.B., Zielinski, N.A., and Welch, R.A. 1995. Construction of *Actinobacillus pleuropneumoniae-Escherichia coli* shuttle vectors: expression of antibiotic-resistance genes. Gene 160: 81-6.

Wilce, M.C., and Parker, M.W. 1994. Structure and function of the glutathione-*S*-transferases. Biochim. Biophys. Acta 1205: 1-18.

Wilson, R., Ainscough, R., Anderson, K., Baynes, C., Berks, M., Bonfield, J., Burton, J., Connell, M., Copsey, T., Cooper, J., Coulson, A., Craxton, M., Dear, S., Du, Z., Durbin, R., Favello, A., Fulton, L., Gardner, A., Green, P., Hawkins, T., Hillier, L., Jier, M., Johnston, L., Jones, M., Kershaw, J., Kirsten, J., Laister, N., Latreille, P., Lightning, J., Lloyd, C., McMurray, A., Mortimore, B., O'Callaghan, M., Parsons, J., Percy, C., Rifken, L., Roopra, A., Saunders, D., Shownkeen, R., Smaldon, N., Smith, A., Sonnhammer, E., Staden, R., Sulston, J., Thierry-Mieg, J., Thomas, K., Vaudin, M., Vaughan, K., Waterston, R., Watson, A., Weinstock, L., Wilkinson-Sproat, J., and Wohldman, P. 1998. 2.2 Mb of contiguous nucleotide sequence from chromosome III of *Caenorhabditis elegans*. Nature. In press.

Wood, A.R., Lainson, F.A., Wright, F., Baird, G.D., and Donachie, W. 1995. A native plasmid of *Pasteurella haemolytica* serotype A1: DNA sequence analysis and investigation of its potential as a vector. Res. Vet. Sci. 58: 163-8.

Wray, W., Boulidas, T., Wray, V.P., and Hancock, R. 1981. Silver staining of proteins in polyacrylamide gels. Anal. Biochem. 118: 197-203.

Wright, C.L., Strugnell, R.A., and Hodgson, A.L. 1997. Characterization of a *Pasteurella multocida* plasmid and its use to express recombinant proteins in *P. multocida*. Plasmid 37: 65-79.

Yan, H., Lim, J.T.E., Contillo, L.G., and Krolewski, J.J. 1995. Glutathione-*S*-transferase fusion proteins mimic receptor dimerization in permeabilized cells. Anal. Biochem. 231: 455-8.

Yanisch-Perron, C., Vieira, J., and Messing, J. 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene 33: 103-19.

Yates, W.D.G. 1982. A review of infectious rhinotracheitis, shipping fever pneumonia and viral-bacterial synergism in respiratory disease of cattle. Can. J. Comp. Med. 46: 225-263.

Yeo, E., and Sutherland, D.R. 1995. Selective cleavage of platelet gbIbα (CD42b) but not gpIIb/IIIa (CD41/CD61), P-selectin (CD62), or CDw109 with a novel *O*-sialoglycoprotease. In: Leukocyte typing (V): white cell differentiation antigens. S.F. Schlossman, *et al.* (Eds.). Vol. 2. Oxford University Press. pp. 1313-15.

Ziemienowicz, A., Skowyra, D., Zeilstra-Ryalls, J., Fayet, O., Georgopoulos, C., and Zylicz, M. 1993. Both the *Escherichia coli* chaperone systems, GroEL/GroES and DnaK/DnaJ/GrpE, can reactivate heat-treated RNA polymerase. Different mechanisms for the same activity. J. Biol. Chem. 268: 25425-31.

Zimmerman, M.L., and Hirsh, D.C. 1980. Demonstration of an R plasmid in a strain of *Pasteurella haemolytica* isolated from feedlot cattle. Am. J. Vet. Res. 41: 166-9.

APPENDIX A: Recipes for Media and Solutions.

Media

LB (Luria-Bertani) broth
      10.0 g/L Difco Bacto-tryptone
       5.0 g/L Difco Bacto-yeast extract
      10.0 g/L NaCl
      (12.0 g/L agar for plates)

LT (Luria-thymine) broth
      12.0 g/L Difco Bacto-tryptone
       5.0 g/L Difco Bacto-yeast extract
       5.0 g/L NaCl
       1.0 g/L glucose
       50 mg/L thymine
      (12.0 g/L agar for plates)

BHI broth
      37.0 g/L Difco Bacto brain heart infusion
      (12.0 g/L agar for plates)

Solutions for DNA methods

TE buffer
      1.21 g/L Tris-HCl (10 mM)
      2 ml/L 0.5M EDTA (1 mM)
      → titrated to (pH 7.5) with HCl

Tracking dye
      50% glycerol
      0.1% bromophenol blue
      0.1% xylene cyanol FF
      → store at 4°C

1 kB ladder
      40 μl stock DNA ladder (Gibco BRL)
      70 μl TE buffer
      70 μl tracking dye
      → store at -20°C

TAE buffer (10x stock)
      48.5 g/L Tris (0.4M)
      3.8 g/L Na$_4$EDTA (0.01M)
      → titrated to pH 7.9 with acetic acid

<u>TB buffer (4x stock)</u>
  48.5 g/L Tris (0.4M)
  23.23 g/L Boric acid (0.376M)
  25 ml/L of 0.5M EDTA (pH 7.5) (12.5 mM)

<u>SRL solution (lysis buffer for E-lyse)</u>
  25% sucrose in TB buffer (autoclaved 5 minutes)
  2 units/ml RNase (boiled 10 min to remove contaminating DNAses)
  → incubate for 10 min at 90°C, stored at 4°C in 1 ml aliquots
  → add 1mg/ml of lysozyme before use

Solutions for SDS-PAGE

<u>5x Sample buffer</u>
  0.6 ml 1M Tris-HCl (pH 6.8)
  5.0 ml 50% glycerol
  2.0 ml 10% SDS
  0.5 ml β-mercaptoethanol
  1.0 ml 1% bromophenol blue
  0.9 ml dH$_2$0
  → store at -20°C

<u>12% Separating gel (for six gels)</u>
  8.5 ml dH$_2$0
  6.25 ml 1.5M Tris-HCl (pH 8.8)
  0.25 ml 10% SDS
  10.0 ml 30% Acrylamide/ 0.8% Bis-acrylamide
  125 µl 10% ammonium persulfate
  12.5 µl TEMED

<u>5% Stacking gel (for six gels)</u>
  4.3 ml dH$_2$0
  1.9 ml 1.5M Tris-HCl (pH 6.8)
  75 µl 10% SDS
  1.2 ml 30% Acrylamide/ 0.8% Bis-acrylamide
  37.5 µl 10% ammonium persulfate
  7.5 µl TEMED

Solutions for Immunoblotting

<u>1° Antibody solution</u>
8.3 ml 3% gelatin in TTBS
16.7 ml 1xTTBS
250 µl 100x mAB110 stock
125 µl 10% NaN₃ (0.02%)

<u>2° Antibody solution</u>
16.7 ml 3% gelatin in TTBS
33.3 ml 1xTTBS
20 µl Biorad AP conjugated rabbit α mouse 2° Ab (1/2500 dilution)
125 µl 10% NaN₃ (0.02%)

<u>5x TTBS</u>
12.11 g/L TRIS (0.1 M)
146.1 g/L NaCl (2.5 M)
→ pH to 7.5 with HCl
→ add 2.5 ml Tween 20

<u>5x NBT</u>
60.55 g/L TRIS (0.5 M)
29.22 g/L NaCl (0.5 M)
→ titrated to pH 9.5 with HCl
→ add 50 ml/L 1M MgCl₂ when making 1x NBT

Solutions used in the purification of Ggcp

<u>PBS buffer</u>
1.42 g/L Na₂HPO₄ (10mM)
216 mg/L NaH₂PO₄ (1.8 mM)
8.19 g/L NaCl (140 mM)
201.3 mg/L KCl (2.7 mM)
→ titrated to pH 7.4 with either HCl or NaOH if necessary

<u>Loading buffer</u>
PBS buffer
771.3 mg/L DTT (5 mM)
0.2 % CHAPS

<u>Elution buffer</u>
11.92 g/L HEPES (50 mM)
3.07 g/L reduced glutathione (10 mM)
154 mg/L DTT (1 mM)
0.2% CHAPS
→ titrated to pH 8.0 with either HCl or NaOH

Other Solutions

Protease Inhibitor Cocktail (P.I.C.)

      25X stocks (25 μg/ml) of antipain, aprotinin, leupeptin, and pepstatin A stored at –20°C in 50 mM HEPES (pH 7.4) were used in the glycoprotease activity assay.   -

      50X stocks (1.25 mg/ml) of antipain, aprotinin, leupeptin, and pepstatin A stored at -20 °C in 50 mM HEPES (pH 7.4) were used in cell lysates.

antipain -- inhibits papain and trypsin
aprotinin -- inhibits serine proteases
leupeptin -- inhibits serine and cysteine proteases
pepstatin -- inhibits acid proteases

APPENDIX B. Outline of experimental strategies employed in the attempted recovery of biological activity from rGgcp.

Variations on the production of rGgcp
        a) Growth temperature; 30°C or 37°C
        b) Lysis method; sonication or French press lysis
        c) Composition of the loading buffer
                → PBS vs. HEPES buffer
                → ± 100 mM NaCl
                → ± 1 mM $ZnCl_2$, $CaCl_2$, or $MgCl_2$

Removal of washed rGgcp from the affinity column

        a) Elution with a reduced glutathione buffer
                → TrisHCl or HEPES buffer
                → ± 100 mM NaCl
                → ± 0.2% CHAPS
                → ± 1 mM DTT
                → ± 1 mM $ZnCl_2$, $CaCl_2$, or $MgCl_2$
                → pH 8.0 or 7.5

        b) Denaturation of rGgcp off of the column
            i) Glycine buffer (pH 2.5) (50 mM)
            ii) Guanidium chloride (0 to 6 M)

Renaturation of rGgcp by the removal of the denaturant
        a) Dialysis against 50 mM HEPES (pH 7.5) buffer (± 3.5 µg of substrate)
        b) +50 mM HEPES (pH 7.5) buffer to the low pH denatured rGgcp
        c) G25 column (used for refolding studies)

In vitro refolding of purified, denatured rGgcp

        a) Rabbit reticulocyte lysate
        b) PDI chaperone
        c) DnaK/J/GrpE
        d) Both PDI & DnaK/J/GrpE together
                → ± thrombin cleavage
                → ± 1 mM $ZnCl_2$
                → ± 2 or 10 µl of sample in a 25 µl reaction volume
                → ± 0.2% CHAPS
                → ± substrate during the refolding process
                → ± creatine phosphate kinase ATP regeneration system

APPENDIX C. Nucleotide and amino acid sequence of the (A) *gcp* and (B) *ggcp* genes. The numbers above each line refer to the nucleotide position relative to the start codon. The nucleotide sequence is underlined where represented by the PCR primers.

1. The *gcp* gene. The *gcp* gene sequence is the re-appraised *BamHI-EcoRI* fragment of pPH1.1ΔE that was originally published by Abdullah *et al.* (1991). The restriction sites are underlined.

```
-140         -130         -120         -110         -100         -90
 *            ●            *            *            *            *
GGATCCAAGA   ATATGAAAGC   AAAGAGCTAC   CGAATCCTGA   AAAACTGAAG   TATGGCGAAC
  BamHI

-80          -70          -60          -50          -40          -30
 *            *            ●            *            *            ●
AATTCTAGTC   GTACAGAGAA   TAATGTGAGG   GGCGTTCTTC   GCCCCTTTTG   GTTTTCTAAC
       Primer GCP737L

-20          -10          1            11           21           31
                          M   R   I    L   G   I  E    T  S  C    D  E  T
TTATTTTGAC   TTCTCCAACT   ATGCGAATTT   TAGGTATTGA   AACCTCTTGT   GATGAAACCG
         Primer L #938?

41           51           61           71           81           91
G   V   A  I   Y  D   E   D   K  G    L  V  A  N   Q   L   Y    S  Q  I
GTGTTGCCAT   TTATGATGAA   GACAAAGGCT   TAGTGGCAAA   CCAGCTTTAT   AGCCAAATTG

101          111          121          131          141          151
D   M   H  A   D  Y   G   G   V  V    P  E  L  A   S   R   D    H  I  R
ATATGCACGC   CGATTACGGT   GGCGTAGTCC   CTGAACTGGC   TTCTCGAGAC   CATATCCGTA

161          171          181          191          201          211
K   T   L  P   L  I   Q   E   A  L    K  E  A  N   L   Q   P    S  D  I
AAACGTTGCC   ACTAATTCAA   GAAGCCTTAA   AAGAGGCCAA   TCTGCAACCC   TCGGATATTG

221          231          241          251          261          271
D   G   I  A   Y  T   A   G   P  G    L  V  G  A   L   L   V    G  S  T
ACGGCATTCG   GTATACTGCC   GGCCCAGGCT   TGGTCGGGGC   TTTATTGGTC   GGCTCAACCA

281          291          301          311          321          331
I   A   R  S   L  A   Y   A   W  N    V  P  A  L   G   V   H    H  M  E
TTGCCCGTTC   GCTGGCTTAT   GCTTGGAATG   TTCCGGCATT   GGGCGTTCAC   CATATGGAAG

341          351          361          371          381          391
G   H   L  L   A  P   M   L   E  E    N  A  P  E   F   P   F    V  A  L
GGCATTTACT   TGCCCCAATG   TTGGAAGAAA   ATGCCCCTGA   ATTTCCGTTT   GTGGCATTAT

401          411          421          431          441          451
L   I   S  G   G  H   T   Q   L  V    K  V  D  G   V   G   Q    Y  E  L
TGATTTCAGG   TGGACACACC   CAACTGGTAA   AAGTTGACGG   CGTTGGGCAA   TACGAACTAC

461          471          481          491          501          511
L   G   E  S   I  D   D   A   A  G    E  A  F  D   K   T   G    K  L  L
TCGGGGAATC   AATTGATGAT   GCTGCCGGTG   AAGCCTTTGA   CAAAACAGGC   AAACTACTCG
```

129

```
521         531         541         551         561         571
G   L   D   Y   P   A   G   V   A   M   S   K   L   A   E   S   G   T   P   N
GTTTGGATTA  CCCTGCCGGT  GTAGCGATGT  CAAAATTAGC  CGAATCCGGC  ACGCCAAATC        .

581         591         601         611         621         631
R   F   K   F   P   R   P   M   T   D   R   P   G   L   D   F   S   F   S   G
GTTTTAAATT  CCCTCGTCCA  ATGACCGACA  GACCGGGACT  GGATTTCAGT  TTCTCCGGTT

641         651         661         671         681         691
L   K   T   F   A   A   N   T   I   K   A   N   L   N   E   N   G   E   L   D
TAAAAACCTT  TGCTGCGAAT  ACGATTAAAG  CCAATCTTAA  TGAAAATGGT  GAACTCGATG

701         711         721         731         741         751
E   Q   T   K   C   D   I   A   H   A   F   Q   Q   A   V   V   D   T   I   L
AGCAAACCAA  ATGCGATATT  GCCCACGCAT  TCCAACAAGC  CGTGGTTGAT  ACTATTTTAA

761         771         781         791         801         811
I   K   C   K   R   A   L   E   Q   T   G   Y   K   R   L   V   M   A   G   G
TTAAATGCAA  GCGAGCGTTA  GAGCAAACCG  GCTATAAACG  CTTAGTAATG  GCAGGCGGCG

821         831         841         851         861         871
V   S   A   N   K   Q   L   R   A   D   L   A   E   M   M   K   K   L   K   G
TAAGTGCCAA  TAAACAATTA  CGAGCAGACC  TTGCGGAAAT  GATGAAAAAA  TTAAAGGCG

881         891         901         911         921         931
E   V   F   Y   P   R   P   Q   F   C   T   D   N   G   A   M   I   A   Y   T
AAGTATTCTA  CCCTCGCCCA  CAATTTTGCA  CTGACAACGG  CGCAATGATT  GCCTACACTG

941         951         961         971         981         991
G   F   L   R   L   K   N   D   E   Q   T   D   L   S   I   S   V   K   P   R
GCTTTCTTCG  CTTAAAAAAC  GATGAACAAA  CCGACTTAAG  CATTAGCGTA  AAACCCCGCT

1001        1011        1021        1031        1041        1051
W   A   M   T   E   L   P   P   I   N   •           *           *
GGGCTATGAC  CGAATTACCA  CCGATTAATT  AACCTTTCAA  GCGGTGAAAT  TTCTTGTTAA

1061        1071        1081        1091        1101        1111
•           *           *           *           *           *
TTTTGCAAAA  ATTTAATCAA  AAATAACCGC  TTGCTATATG  ATAGATTAAA  TTTATGAATA

1121        1131        1141        1151        1161        1171
*           *           *           *           *           *
ATTATGTAAT  TAGCCTACCT  CCGCACAGGA  GCGTAGAAAA  CATATTCAAG  CTGAATTC
                                                            EcoRI
          Primer R (#93?
```

130

2. The *ggcp* gene. The amino acid sequence for the GST carrier is a lighter shade.

```
-137        -127        -117        -107        -97         -87
*           *           *           *           *           *
TCTGGATAAT  GTTTTTTGCG  CCGACATCAT  AACGGTTCTG  GCAAATATTC  TGAAATGAGC

-77         -67         -57         -47         -37         -27
*           *           *           *           *           *
TGTTGACAAT  TAATCATCGG  CTCGTATAAT  GTGTGGAATT  GTGAGCGGAT  AACAATTTCA
   Primer GGCP????

-17         -7          +1  3       13          23          33
*           *           •
CACAGGAAAC  AGTATTCATG  TCCCCTATAC  TAGGTTATTG  GAAAATTAAG  GGCCTTGTGC

43          53          63          73          83          93

AACCCACTCG  ACTTCTTTTG  GAATATCTTG  AAGAAAAATA  TGAAGAGCAT  TTGTATGAGC

103         113         123         133         143         153

GCGATGAAGG  TGATAAATGG  CGAAACAAAA  AGTTTGAATT  GGGTTTGGAG  TTTCCCAATC

163         173         183         193         203         213

TTCCTTATTA  TATTGATGGT  GATGTTAAAT  TAACACAGTC  TATGGCCATC  ATACGTTATA

223         233         243         253         263         273

TAGCTGACAA  GCACAACATG  TTGGGTGGTT  GTCCAAAAGA  GCGTGCAGAG  ATTTCAATGC

283         293         303         313         323         333

TTGAAGGAGC  GGTTTTGGAT  ATTAGATACG  GTGTTTCGAG  AATTGCATAT  AGTAAAGACT

343         353         363         373         383         393

TTGAAACTCT  CAAAGTTGAT  TTTCTTAGCA  AGCTACCTGA  AATGCTGAAA  ATGTTCGAAG

403         413         423         433         443         453

ATCGTTTATG  TCATAAAACA  TATTTAAATG  GTGATCATGT  AACCCATCCT  GACTTCATGT

463         473         483         493         503         513

TGTATGACGC  TCTTGATGTT  GTTTTATACA  TGGACCCAAT  GTGCCTGGAT  GCGTTCCCAA

523         533         543         553         563         573

AATTAGTTTG  TTTTAAAAAA  CGTATTGAAG  CTATCCCACA  AATTGATAAG  TACTTGAAAT

583         593         603         613         623         633

CCAGCAAGTA  TATAGCATGG  CCTTTGCAGG  GCTGGCAAGC  CACGTTTGGT  GGTGGCGACC

643         653         663         673         683         693
                                                            M   R   I
ATCCTCCAAA  ATCGGATCTG  GTTCCGCGTG  GATCCCCGAA  TTCTCCAACT  ATGCGAATTT
```
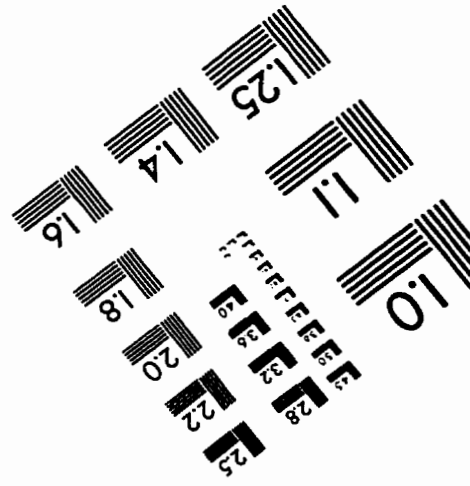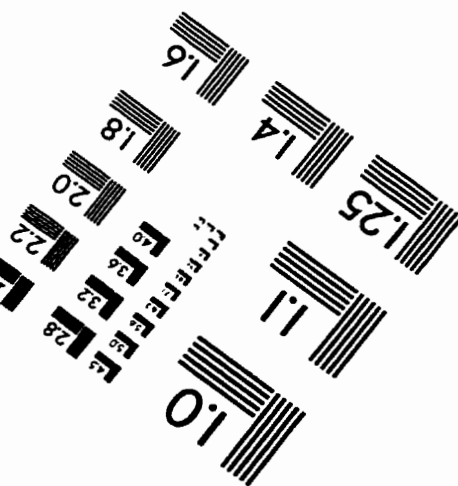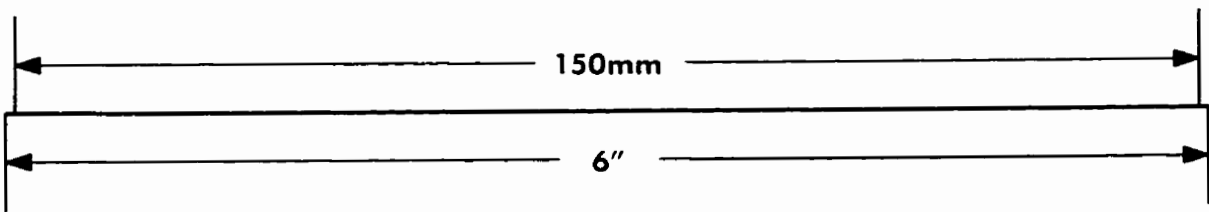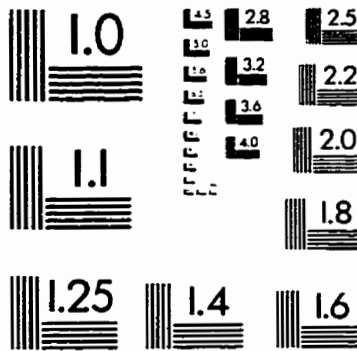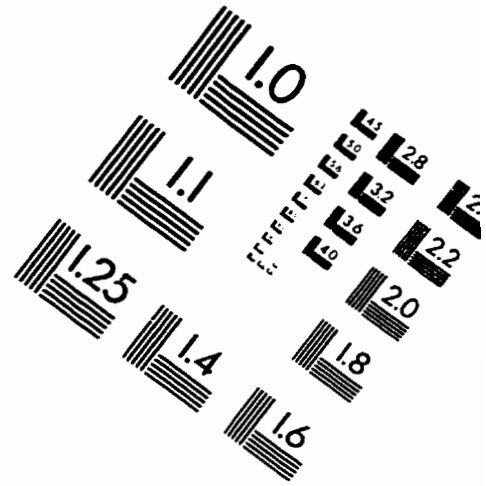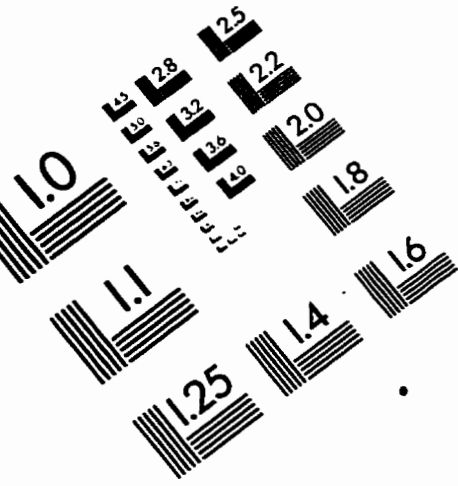
131

```
703          713          723          733          743          753
L  G  I  E    T  S  C    D  E  T    G  V  A  I    Y  D  E    D  K  G
TAGGTATTGA  AACCTCTTGT  GATGAAACCG  GTGTTGCCAT  TTATGATGAA  GACAAAGGCT

763          773          783          793          803          813
L  V  A  N    Q  L  Y    S  Q  I    D  M  H  A    D  Y  G    G  V  V
TAGTGGCAAA  CCAGCTTTAT  AGCCAAATTG  ATATGCACGC  CGATTACGGT  GGCGTAGTCC

823          833          843          853          863          873
P  E  L  A    S  R  D    H  I  R    K  T  L  P    L  I  Q    E  A  L
CTGAACTGGC  TTCTCGAGAC  CATATCCGTA  AAACGTTGCC  ACTAATTCAA  GAAGCCTTAA

883          893          903          913          923          933
K  E  A  N    L  Q  P    S  D  I    D  G  I  A    Y  T  A    G  P  G
AAGAGGCCAA  TCTGCAACCC  TCGGATATTG  ACGGCATTGC  CTATACTGCC  GGCCCAGGCT

943          953          963          973          983          993
L  V  G  A    L  L  V    G  S  T    I  A  R  S    L  A  Y    A  W  N
TGGTCGGGGC  TTTATTGGTC  GGCTCAACCA  TTGCCCGTTC  GCTGGCTTAT  GCTTGGAATG

1003         1013         1023         1033         1043         1053
V  P  A  L    G  V  H    H  M  E    G  H  L  L    A  P  M    L  E  E
TTCCGGCATT  GGGCGTTCAC  CATATGGAAG  GGCATTTACT  TGCCCCAATG  TTGGAAGAAA

1063         1073         1083         1093         1103         1113
N  A  P  E    F  P  F    V  A  L    L  I  S  G    G  H  T    Q  L  V
ATGCCCCTGA  ATTTCCGTTT  GTGGCATTAT  TGATTTCAGG  TGGACACACC  CAACTGGTAA

1123         1133         1143         1153         1163         1173
K  V  D  G    V  G  Q    Y  E  L    L  G  E  S    I  D  D    A  A  G
AAGTTGACGG  CGTTGGGCAA  TACGAACTAC  TCGGGGAATC  AATTGATGAT  GCTGCCGGTG

1183         1193         1203         1213         1223         1233
E  A  F  D    K  T  G    K  L  L    G  L  D  Y    P  A  G    V  A  M
AAGCCTTTGA  CAAAACAGGC  AAACTACTCG  GTTTGGATTA  CCCTGCCGGT  GTAGCGATGT

1243         1253         1263         1273         1283         1293
S  K  L  A    E  S  G    T  P  N    R  F  K  F    P  R  P    M  T  D
CAAAATTAGC  CGAATCCGGC  ACGCCAAATC  GTTTTAAATT  CCCTCGTCCA  ATGACCGACA

1303         1313         1323         1333         1343         1353
R  P  G  L    D  F  S    F  S  G    L  K  T  F    A  A  N    T  I  K
GACCGGGACT  GGATTTCAGT  TTCTCCGGTT  TAAAAACCTT  TGCTGCGAAT  ACGATTAAAG

1363         1373         1383         1393         1403         1413
A  N  L  N    E  N  G    E  L  D    E  Q  T  K    C  D  I    A  H  A
CCAATCTTAA  TGAAAATGGT  GAACTCGATG  AGCAAACCAA  ATGCGATATT  GCCCACGCAT

1423         1433         1443         1453         1463         1473
F  Q  Q  A    V  V  D    T  I  L    I  K  C  K    R  A  L    E  Q  T
TCCAACAAGC  CGTGGTTGAT  ACTATTTTAA  TTAAATGCAA  GCGAGCGTTA  GAGCAAACCG

1483         1493         1503         1513         1523         1533
G  Y  K  R    L  V  M    A  G  G    V  S  A  N    K  Q  L    R  A  D
GCTATAAACG  CTTAGTAATG  GCAGGCGGCG  TAAGTGCCAA  TAAACAATTA  CGAGCAGACC

1543         1553         1563         1573         1583         1593
L  A  E  M    M  K  K    L  K  G    E  V  F  Y    P  R  P    Q  F  C
TTGCGGAAAT  GATGAAAAAA  TTAAAAGGCG  AAGTATTCTA  CCCTCGCCCA  CAATTTTGCA
```

132

```
1603        1613        1623        1633        1643        1653
T   D   N   G   A   M   I   A   Y   T   G   F   L   R   L   K   N   D   E   Q
CTGACAACGG  CGCAATGATT  GCCTACACTG  GCTTTCTTCG  CTTAAAAAAC  GATGAACAAA

1663        1673        1683        1693        1703        1713
T   D   L   S   I   S   V   K   P   R   W   A   M   T   E   L   P   P   I   N
CCGACTTAAG  CATTAGCGTA  AAACCCCGCT  GGGCTATGAC  CGAATTACCA  CCGATTAATT

1723        1733        1743        1753        1763        1773
●           ●           ●           *           *           ●

AACCTTTCAA  GCGGTGAAAT  TTCTTGTTAA  TTTTGCAAAA  ATTTAATCAA  AAATAACCGC

1783        1793        1803        1813        1823        1833
*           *           *           *           *           *

TTGCTATATG  ATAGATTAAA  TTTATGAATA  ATTATGTAAT  TAGTCGACTC  GAGCGGCCGC
                                    Primer R  #939)

1843        1853        1863        1873        1883        1893
*           *           *           ●           ●           *

ATCGTGACTG  ACTGACGATC  TGCCTCGCGC  GTTTCGGTGA  TGACGGTGAA  AACCTCTGAC
```

133

IMAGE EVALUATION
TEST TARGET (QA-3)

150mm

6"