Molecular Study of the PE-binding adhesin candidates from *Haemophilus influenzae*
Desheng Lu
Degree of Doctor of Philosophy, 1998
Department of Molecular and Medical Genetics, University of Toronto

ABSTRACT

In order to characterize a putative 37 kDa phosphatidylethanolamine (PE)-binding adhesin, its coding sequence was amplified from a nontypeable *H. influenzae* (NTHI6564) by PCR and sequenced. Recombinant adhesin was expressed in *E. coli* as a GST and as a 6XHis tagged fusion protein. Polyclonal antisera were raised against both purified fusion proteins in rabbits, and both antisera recognized a single protein band of 37 kDa from *H. influenzae* whole cell extracts by Western blotting. However, neither fusion protein bound to PE by TLC overlay assay, nor did antisera against this protein inhibit *H. influenzae* binding to PE in vitro. Subsequent studies demonstrated that the putative PE-binding adhesin can not bind to PE, but has potent celite binding ability.

The celite binding protein has a C-terminal disulfide bond between Cys254 and Cys308. A Cys308 to Ser mutant was expressed as a GST fusion protein, and a C-terminal truncated mutant was also constructed and expressed as a 6XHis tagged fusion protein. Both mutants lost the disulfide bond but retained celite binding capability, indicating that the C-terminus and disulfide bond of this protein are not involved in celite binding. Twenty three *H. influenzae* clinical strains were analyzed by Western blotting and PCR. The results indicated that the celite binding protein gene is highly conserved and its protein product is generally expressed in *H. influenzae* clinical isolates.
Subcellular fractionation was used to localize the celite binding protein to the periplasmic region both in NTHi6564 and in a transfected recombinant E. coli strain. Recombinant protein was purified from periplasmic extracts by chromatofocusing and gel filtration chromatography. The observed isoelectric point (pI) of this protein, as estimated by chromatofocusing, was 6.4, which is close to the theoretical pI (6.7). The CD spectrum of the purified protein suggested that the major secondary structure is α-helix. The purified protein contained about two zinc atoms per protein molecule as determined by neutron activation analysis and atomic absorption spectroscopy. The zinc atoms could be removed by incubation with EDTA, and, by further addition of zinc, a total of five zinc atoms per protein molecule could be bound. Direct binding of 65Zn to the recombinant protein was demonstrated by zinc blotting. Thus, this celite binding protein is in fact a periplasmic zinc binding protein (Pzp1).

A pzp1 deficient mutant was constructed. This mutant was defective for growth under aerobic conditions and grew poorly under anaerobic conditions. The growth defect was specifically rescued by supplementing the growth medium with high concentrations of zinc (100 μM). This result indicates that the pzp1 gene is essential for growth under aerobic conditions and also is important for anaerobic growth of H. influenzae. Taken together, our results provide direct evidence that Pzp1 is a key protein for zinc uptake in H. influenzae. This is the first description of a protein involved in zinc uptake in prokaryotes.
To identify the real PE-binding adhesin(s) from *H. influenzae*, whole cell extracts of NTHI6564 were prepared and applied to a PE affinity column. A predominant protein band with similar migration on SDS-PAGE to Pzp1 was eluted, but Western blotting showed that it was not Pzp1. The N-terminal amino acid sequence of this protein matched the outer membrane protein F2 of *H. influenzae*. Subsequently, OmpP2 was expressed as a GST-fusion protein (GST-P2). However, only a small portion of GST-P2 could be purified in soluble form. Purified GST-P2 specifically bound to PE by TLC overlay. In addition, OmpP5 was also eluted from the PE affinity matrix. We conclude that OmpP2, and possibly OmpP5 of *H. influenzae* are PE-binding proteins.
ACKNOWLEDGMENTS

I would like to express my sincere gratitude to my supervisor Dr. Lingwood for his constant support and scientific guidance in my graduate studies. I deeply appreciate his understanding and consideration at difficult times during the course of my Ph.D training.

My special thanks go to Ms. Beth Boyd for her scientific advise, excellent technical help and friendship. I will remember her sympathy, encouragement and inspiration in most difficult times of my training forever.

I am very grateful to Dr. Bognar for his understanding and support. I would like to thank my Ph.D committee members, Dr. Ellen and Dr. Kain, for their guidance and assistance.

I would like to express my appreciate to my colleagues: Anita Nutikka, Murugesappillai Mylvaganam, Daniel Mamak, Eva Hartmann, Jason Busse, Mario Huesca and Deborah Foster for all their assistance and support. Many thanks to Shirley Thompson for her superb secretarial assistance.

I would like to thank Dr. R. Hancock for neutron activation analysis, M. DiDonato in Dr. Sarkar's lab for $^{65}$Zn blotting.

Finally, I share this thesis with my wife, Bing Yang, for her unconditional love, understanding and support.
This thesis is dedicated to my family
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>TITLE PAGE</td>
<td>i</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>ii</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>v</td>
</tr>
<tr>
<td>TABLE OF CONTENTS</td>
<td>vii</td>
</tr>
<tr>
<td>LIST OF ABBREVIATIONS</td>
<td>xi</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>xiii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>xvi</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1. Bacterial adherence</td>
<td>1</td>
</tr>
<tr>
<td>1.1. Adhesins</td>
<td>1</td>
</tr>
<tr>
<td>1.2. Carbohydrate receptors</td>
<td>7</td>
</tr>
<tr>
<td>1.3. Phospholipid receptor - Phosphatidylethanolamine (PE)</td>
<td>10</td>
</tr>
<tr>
<td>2. <em>Haemophilus influenzae</em></td>
<td>13</td>
</tr>
<tr>
<td>2.1. General characteristics</td>
<td>13</td>
</tr>
<tr>
<td>2.2. Pathogenesis</td>
<td>15</td>
</tr>
<tr>
<td>2.2.1. Capsule</td>
<td>16</td>
</tr>
<tr>
<td>2.2.2. Outer membrane proteins (OMP)</td>
<td>17</td>
</tr>
<tr>
<td>2.2.3. Lipooligosaccharide (LOS)</td>
<td>24</td>
</tr>
<tr>
<td>2.2.4. Immunoglobulin A1 protease</td>
<td>25</td>
</tr>
<tr>
<td>2.2.5. Hemagglutinating Pili</td>
<td>27</td>
</tr>
<tr>
<td>2.2.6. High-molecular-weight adhesins (HMW1 and HMW2)</td>
<td>29</td>
</tr>
<tr>
<td>2.2.7. Other <em>H. influenzae</em> adhesins</td>
<td>31</td>
</tr>
<tr>
<td>2.2.8. Putative PE-binding adhesin</td>
<td>32</td>
</tr>
<tr>
<td>3. Lipoprotein receptor antigen group (Lra1)</td>
<td>33</td>
</tr>
<tr>
<td>3.1. FimA</td>
<td>33</td>
</tr>
</tbody>
</table>
2.9. Celite or PE affinity chromatography
2.10. FITC labeling of GST-0119 and His-0119
2.11. Celite binding of FITC-conjugated GST-0119
2.12. Immunoprecipitation
2.13. Electrophoresis and Western blotting
2.14. Construction of mutants
2.15. Subcellular fractions of H. influenzae
2.16. Expression and purification of the putative PE-binding adhesin in E. coli
2.17. Metal composition of purified celite binding protein
2.18. Circular dichroism (CD) measurements
2.19. Construction of the pppl isogenic mutant
2.20. Growth curves
2.21. Expression and purification of OmpP2 fusion protein GST-P2

RESULTS
1. Expression and characterization of the putative PE-binding adhesin of H. influenzae
1.1 Cloning and sequence of the putative PE-binding adhesin
1.2. Expression of recombinant fusion proteins
1.3. The putative PE-binding adhesin has a disulfide bond in the C-terminus
1.4. Immunoprecipitation of the putative PE-binding adhesin in H. influenzae
1.5. The putative PE-binding adhesin cannot bind to Hep2 cells
1.6. The putative PE-binding adhesin cannot bind to PE by TLC overlay

ix
1.7. Expression of the putative PE-binding adhesin in E. coli

1.8. The putative PE-binding adhesin is not a PE-binding protein

1.9. Antiserum against His-0119 cannot inhibit H. influenzae

binding to PE by TLC overlay

2. Characterization of the celite binding protein

2.1. GST-0119 and His-019 bind to celite

2.2. Purification of native Haemophilus HI0119 protein by celite chromatography

2.3. Mutational analysis

2.4. Conservation of the celite binding protein in H. influenzae

3. The celite binding protein is a periplasmic zinc-binding protein (Pzp1)

3.1. Localization of the celite binding protein

3.2. Purification of recombinant celite binding protein in E. coli

3.3. Circular dichroism (CD) analysis of recombinant celite binding protein

3.4. The celite binding protein can bind zinc specifically

3.5. Regulation of the pzpl gene of H. influenzae

4. Pzp1 plays a key role for zinc uptake in H. influenzae

4.1. Construction of the pzpl negative mutant

4.2. Growth defect of the pzpl mutant can be specifically rescued by zinc

5. Outer membrane protein P2 is a PE-binding protein

DISCUSSION

SUMMARY AND FUTURE DIRECTIONS

REFERENCES
# LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABC</td>
<td>ATP-binding cassette</td>
</tr>
<tr>
<td>ATP</td>
<td>adenosine triphosphate</td>
</tr>
<tr>
<td>bp</td>
<td>basepair</td>
</tr>
<tr>
<td>BSA</td>
<td>bovine serum albumin</td>
</tr>
<tr>
<td>CD</td>
<td>circular dichroism</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxyribonucleic acid</td>
</tr>
<tr>
<td>dNTP</td>
<td>2′-deoxynucleotide triphosphates</td>
</tr>
<tr>
<td>E. coli</td>
<td><em>Escherichia coli</em></td>
</tr>
<tr>
<td>EDTA</td>
<td>ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>EDDHA</td>
<td>ethylenediamine-N,N′-diacetic acid</td>
</tr>
<tr>
<td>FITC</td>
<td>fluorescein isothiocyanate</td>
</tr>
<tr>
<td>G+C</td>
<td>guanidine plus cytidine</td>
</tr>
<tr>
<td>Gal</td>
<td>galactose</td>
</tr>
<tr>
<td>G3</td>
<td>globotriaosylceramide</td>
</tr>
<tr>
<td>G4</td>
<td>globotetraosylceramide</td>
</tr>
<tr>
<td>Gg3</td>
<td>gangliotriaosylceramide</td>
</tr>
<tr>
<td>Gg4</td>
<td>gangliotetraosylceramide</td>
</tr>
<tr>
<td>Glc</td>
<td>glucose</td>
</tr>
<tr>
<td>GM1</td>
<td>GM1-Ganglioside</td>
</tr>
<tr>
<td>GM2</td>
<td>GM2-Ganglioside</td>
</tr>
<tr>
<td>GST</td>
<td>glutathione S-transferase</td>
</tr>
<tr>
<td>Hib</td>
<td><em>Haemophilus influenzae</em> type b strain</td>
</tr>
<tr>
<td>His</td>
<td>histidine</td>
</tr>
<tr>
<td>HMW</td>
<td>high molecular weight protein</td>
</tr>
<tr>
<td>IPTG</td>
<td>isopropyl-β-D-thiogalactoside</td>
</tr>
</tbody>
</table>
kb  kilobasepair
kDa  kilodalton
LOS  lipooligosaccharide
LPS  lipopolysaccharide
Lra1 lipoprotein receptor antigen group
NAc  N-acetyl
NAD  nicotinamide adenine dinucleotide
OD  optical density
NTHI  *Haemophilus influenzae* nontypeable strain
Omp  outer membrane protein
ORF  open reading frame
PBS  phosphate buffered saline
PC  phosphatidylcholine
PE  phosphatidylethanolamine
PG  phosphatidylglycerol
pI  isoelectric point
Pzp1 periplasmic zinc binding protein
RNAase ribonuclease
SCHA saliva-coated hydroxyapatite
SDS-PAGE sodium dodecyl sulfate polyacrylamide gel electrophoresis
SGC sulfatogalactosylceramide
SGG sulfatogalactoglycerolipid
Tbp transferrin binding protein
TBS Tris-buffered saline
TLC thin layer chromatography
**LIST OF FIGURES**

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>PapG recognition of three Galα(1-4)Gal-containing isoreceptors.</td>
<td>3</td>
</tr>
<tr>
<td>2.</td>
<td>Schematic representation of a generic periplasmic permease.</td>
<td>39</td>
</tr>
<tr>
<td>3.</td>
<td>Cloning of the DNA sequence corresponding the mature putative PE-binding adhesin (HI0119).</td>
<td>73</td>
</tr>
<tr>
<td>4.</td>
<td>Nucleotide and deduced amino acid sequences corresponding to the mature putative PE-binding adhesin.</td>
<td>75</td>
</tr>
<tr>
<td>5.</td>
<td>Alignment of HI0119 and other bacterial proteins.</td>
<td>77</td>
</tr>
<tr>
<td>6.</td>
<td>Predicted structural properties of the putative PE-binding adhesin of <em>H. influenzae</em>.</td>
<td>79</td>
</tr>
<tr>
<td>7.</td>
<td>Construction of pGHA1 and pTHA1.</td>
<td>84</td>
</tr>
<tr>
<td>8.</td>
<td>Analysis of the fusion proteins by SDS-PAGE.</td>
<td>86</td>
</tr>
<tr>
<td>9.</td>
<td>Determination of the oligometric and intramolecular disulfide-bond status of purified GST-0119 and His-0119 using SDS-PAGE.</td>
<td>89</td>
</tr>
<tr>
<td>10.</td>
<td>Immunoprecipitation of the putative PE-binding adhesin from <em>H. influenzae</em>.</td>
<td>92</td>
</tr>
<tr>
<td>11.</td>
<td>TLC overlay of the purified GST-0119 and His-0119 fusion proteins.</td>
<td>95</td>
</tr>
<tr>
<td>12.</td>
<td>Construction of plasmid pTAP119.</td>
<td>97</td>
</tr>
<tr>
<td>13.</td>
<td>Expression of recombinant putative adhesin in <em>E. coli</em>.</td>
<td>99</td>
</tr>
<tr>
<td>14.</td>
<td>Purification of recombinant adhesin from PE-celite or</td>
<td></td>
</tr>
</tbody>
</table>
celite chromatography.

15. Inhibition of NTH16564 PE binding by antiserum against His-0119.


17. Purification of His-0119 by celite chromatography.

18. Purification of the putative adhesin of H. influenzae by celite chromatography.

19. Construction and expression of the cysteine mutant.

20. Construction of the C-terminal truncated mutant.

21. SDS-PAGE analysis of purified His-0119cm.

22. Purification of His-0119cm by celite chromatography.

23. Determination of the conservation and expression of the celite binding protein in various clinical isolates H. influenzae.

24. SDS-PAGE and Western blotting of the culture supernatant of NTHI6564.

25. Localization of the celite binding protein in NTHI6564.


27. Gel filtration of recombinant celite binding protein.

28. CD spectrum of recombinant celite binding protein.

29. $^{65}$Zinc-blotting with the celite binding protein.

30. Construction of the pzp1 deficient mutant.
31. SDS-PAGE and Western blotting of wild type and the \textit{pzpI} mutants.  
32. Zinc suppression of the growth defect of the \textit{pzpI} mutant.  
33. Recovery of growth defect of the \textit{pzpI} mutant is zinc specific.  
34. Purification of \textit{H. influenzae} adhesin(s) from a PE affinity matrix.  
35. Separation of amplified mature OmpP2 gene and restriction enzyme digestions of \textit{pGHP2} by 0.7% agarose gel electrophoresis.  
36. SDS-PAGE and Western blotting of whole cell lysates of induced and noninduced \textit{E. coli} TOP10 harboring \textit{pGHP2}.  
37. Purification of the fusion protein GST-P2.  
38. TLC overlay of the fusion protein GST-P2.  
39. Comparison of OmpP2 and OmpP5 amino acid sequences.
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Microorganisms tested for binding to ganglio-series glycolipids and PE.</td>
<td>12</td>
</tr>
<tr>
<td>2.</td>
<td>Amino acid analysis of recombinant celite binding protein.</td>
<td>132</td>
</tr>
<tr>
<td>3.</td>
<td>Typical chemical analysis of celite 545.</td>
<td>139</td>
</tr>
<tr>
<td>4.</td>
<td>Analysis of zinc content by atomic absorption spectroscopy.</td>
<td>140</td>
</tr>
</tbody>
</table>
INTRODUCTION

1. Bacterial adherence

Human diseases caused by bacterial pathogens result from an interplay between the invading microorganism and the host. The pathogenesis of bacterial infections typically involves a multitude of virulence factors that interfere in the sophisticated physiological processes of the host (Finlay & Falkow, 1989). The first major interaction between microorganism and its host entails attachment to the eukaryotic cell surface. Two potential fates await the microorganism after it binds to a host cell. Some microorganisms multiply at and remain on the surface of the host cell. Other organisms are to be internalized by the targeted cell after adhesion (Isberg, 1991). Bacterial adherence to host cell is a critical phase in the development of many infections or diseases (Gibbons, 1977; Beachey, 1981). Specificity of bacterial adherence to host cell depends primarily on two factors, an adhesin and a receptor.

1.1. Adhesins

For many bacterial species, adherence is mediated by fimbriae (pili), which are protein filaments protruding from the bacterial cell surface (Beachey, 1981). Most gram-negative bacteria assemble adhesive fimbriae that generally adopt one of two basic morphologies: rod-like fibers with a diameter of approximately 7 nm or flexible and thin fibrillae with a diameter of 2-5 nm (de Graaf & Mooi, 1986). The adhesive interaction of pili can be mediated by major subunits or minor subunits, with the adhesin being located either at the tip of the pilus or along the length of the pilus shaft (Smyth et al., 1996). Bacterial pill have been divided into different classes due to their morphological and structural features, assembly process and specificity of
binding (Smyth et al., 1996). The best characterized of these are P pili, E. coli pyelonephritis-associated pili (Pap). These pili are found in about 5-10% of human fecal E. coli isolates and in up to 90% of E. coli strains associated with urinary tract infections (Plos et al., 1990; Marklund et al., 1992). Genes involved in the biosynthesis and expression of functional P pili are clustered in an operon structure. The DNA sequence of the entire gene cluster has been determined and was shown to encode 11 genes (Hultgren et al., 1993). P pili are composite heteropolymeric fibers consisting of flexible adhesive fibrillae joined end to end to pilus rods (Kuehn et al., 1992). The pilus rod consists of repeating PapA protein subunits arranged in a right-handed helical cylinder (Fig. 1). Tip fibrillae are composed mostly of repeating subunits of PapE arranged in an open helical conformation, which may facilitate interactions with extracellular matrix protein such as fibronectin (Westerlund et al., 1991). The PapG protein, which is localized to the distal ends of tip fibrillae, is responsible for binding to the \(\alpha\)-D-galactopyranosyl-(1\-4)-\(\beta\)-D-galactopyranoside (Gal\(\alpha\)(1\-4)Gal) moiety present in the globo series of glycolipids on cells lining the upper urinary tract (Fig. 1) (Bock et al., 1985; Lindberg et al., 1987; Lund et al., 1988; Hultgren et al., 1993).

Type 1 pili are the most common adhesive appendages found on members of the family *Enterobacteriaceae* (Duguid & Old, 1980). Type 1 pili bind the simple sugar mannose and a number of its structure analogs (Hanson & Brinton, 1988). This binding can be studied at a very fine level, and interactions of this type are crucial to the ability of many members of the family *Enterobacteriaceae* to colonize plant and animal tissues. The adhesive component of type 1 pili is the product of the *fimH* gene. The pilus structure likely plays a role in the sugar specificity and affinity of the *fimH* product, probably by affecting its conformation (Hultgren et al., 1991).
Figure 1. PapG recognition of three Galα(1-4)Gal-containing isoreceptors. The architecture of the P pilus reveals a strategy used by pyelonephritic E. coli to present an adhesin in an accessible configuration that allows it to recognize Galα(1-4)Gal-containing isoreceptors. The β-galabiose portion of the globoseries of glycolipids was found to bind to the tip fibrillum-located PapG adhesin. The G-I, G-II, and GIII adhesins direct the high affinity recognition of three different Galα(1-4)Gal-containing isoreceptors; Gb3, Gb4 and globopentaosylceramide, respectively. Adapted from Hultgren et al., 1993.
Type 4 pili are produced by a variety of important human bacterial pathogens, including *Pseudomonas aeruginosa* (Pasloske et al., 1985), *Neisseria gonorrhoeae* (Meyer et al., 1984), enteropathogenic and enterotoxigenic *E. coli* (Giron et al., 1991; Donnenberg et al., 1992) and *Vibrio cholerae* (Taylor et al., 1987). They are flexible, with a diameter of 6 to 7 nm and a length of up to 20 μm. They are produced at the polar position of the bacterial cell (Strom & Lory, 1993). The main structural subunits of these pili, termed pilins, have N-methylphenylalanine at the start of the mature pilin and share a highly conserved amino-terminal domain of 25 to 30 amino acid residues, while the C-terminal part contains variable regions. The pili of *P. aeruginosa* bind specifically to the carbohydrate sequence βGalNAc(1-4)βGal found in glycolipids asialo-GM1 and asialo-GM2 (Sheth et al., 1994). The adherence of the *P. aeruginosa* pili to glycolipid receptors is a tip-associated phenomenon involving a tip-exposed C-terminal region of the pilin structural subunit (Lee et al., 1994).

Bacterial adhesins are not invariably assembled into polymeric pilus rods. Many bacteria do not have pili, but still cause infectious diseases. Thus, nonpilus adhesins must be used by these organisms. For example, an EPEC adhesin, intimin, is an outer membrane protein encoded by *eaeA* that mediates close attachment of enteropathogenic bacteria to apical surfaces of epithelial cells (Jerse et al., 1990; Yu & Kapper, 1992). Intimin is required for formation of the attaching-effacing lesions and for full pathogenesis of the bacteria. The cell-binding domain of intimin may bind to β1 integrins (Frankel et al., 1994). The filamentous hemagglutinin (FHA) of *Bordetella pertussis* is derived from a very large precursor and forms large complexes which are loosely associated with the cell surface (Locht et al., 1993). The FHA molecule has at least three distinct binding activities which have been localized to different regions of the molecule. The region involved in binding to
sulfated saccharides has been mapped at the N terminus (Menozzi et al., 1994). FHA binds to the integrin CR3 which is present on macrophages, and it has been shown that the RGD sequence present in FHA is important for this binding (Relman et al., 1989). The carbohydrate recognition domain of FHA, which has been localized between amino acid residues 1141 and 1279, binds to lactosylceramides and confers the ability to bind to ciliary cells and macrophages (Prasad et al., 1993).

An important feature of bacterial adhesins is their multiplicity. One bacterial strain frequently carries several adhesins of different specificity (Pinlay & Falkow, 1989; Hacker, 1992). Often the adhesins that are expressed vary depending on growth conditions (Mekalanos, 1992). In particular, many bacteria are able to switch from the production of one type of adhesin to another. This periodic alternation of a surface structure is called phase variation. For example, the different pili on an E. coli strain are under phase variation: they are mostly expressed by separate cells that can rapidly switch their pilus synthesis (Nowicki et al., 1986). Due to phase variation, the infecting E. coli population in the urinary tract is heterogeneous, consisting of cells with different pili and of nonpiliated cells (Pere et al., 1987).

Microbial attachment to host cells or tissue surfaces is characterized by a high degree of specificity (Sharon & Lis, 1993). For example, uropathogenic E. coli is abundant in tissues surrounding the ducts that connect the kidneys and the bladder, yet it is seldom found in the upper respiratory tract. In contrast, group A streptococci, which colonize the upper respiratory tract and skin, rarely cause urinary tract infections. Since bacterial adhesins bind exclusively to certain surface carbohydrates, these interactions determine which tissues are susceptible to bacterial invasion (Reid & Sobel, 1987; Karlsson, 1989; Sharon & Lis, 1993).
1.2. Carbohydrate receptors

Eukaryotic cell membranes are rich in carbohydrates in the form of oligosaccharides covalently linked to lipids or proteins, which constitute glycolipids or glycoproteins respectively (Krivan et al., 1992). The carbohydrates are characterized by great structural diversity and variability related to individual, species and cell type. Furthermore, the glycoconjugates within any one cell have many different types of carbohydrates (Hakomori, 1983b). The high diversity of the carbohydrates accord with the host tissue or cell specificities of bacterial infection. For example, uroepithelial cells from those rare individuals who lack the P blood-group substance do not bind to P piliated *E. coli* (Ofek & Doyle, 1994). Such individuals are much less susceptible to infections from those bacteria than the rest of the population is. However, the bacteria will bind if the epithelial cells are first coated with a synthetic glycolipid containing galabiose (Sharon & Lis, 1993).

Glycolipids are found present at the external surface of the lipid bilayer and form large clusters, rather than being distributed homogeneously. The hydrophilic carbohydrate moieties may lie along the surface of the membrane in an aqueous medium while the hydrophobic ceramide region anchors it in the bilayer (Hakomori, 1983b). The chains of oligosaccharides are synthesized in the Golgi apparatus by the sequential action of glycosyltransferases, each one specifically catalyzing the addition of one carbohydrate (Fenderson et al., 1990). The sugar sequence is determined by the glycosyltransferase that sequentially recognizes the sugar chain. Sugars are later modified by sulfation (Guerrant & Lingwood, 1991). Based on the chemical structure of the core carbohydrate, three major series of glycolipids have been identified: the globoseries (Galα1-4Galβ1-4Glcβ1-cer), the lacto-series (GlcNAcβ1-3Galβ1-4Glcβ1-cer) and the ganglio-series (GlcNAcβ1-4Galβ1-
Glycoproteins are synthesized by co- and post-translational modifications, in which, by way of a complex series of reactions, carbohydrate chains are attached to special sites on a newly synthesized protein (Fessenden et al., 1986). There are two types of glycosylation, called N-type or O-type depending on the atom of the amino acid to which the carbohydrate is attached. N-type glycosylation occurs exclusively on the nitrogen atom of Asn side chains, whereas O-glycosylation occurs on the oxygen atoms of hydroxyls of Ser and Thr residues. N-glycosylation is found in both plants and animals, while O-glycosylation is confined chiefly to animals (Einspahr, 1991).

As a receptor for microorganisms, glycolipids are less complicated since they have only one carbohydrate chain per molecule whereas glycoproteins can have several N- and/or O-linked oligosaccharides (Krivan et al., 1992). In many cases thus far described, the carbohydrates recognized by bacterial adhesins are conjugated to lipid moieties and not to protein (Karlsson, 1989). There are several possible reasons. Since the carbohydrate moieties of glycolipids are closer to the plasma membrane surface than those of glycoproteins, microorganisms bound to glycolipids may be more intimately associated with the host membrane which provides them with a better nutritional environment (Krivan et al., 1992). Furthermore, a characteristic of the glycolipid is that it is always membrane bound and hence absent in most cases from secretion. Thus, the selection by a microbe of a specifically lipid-linked oligosaccharide may assure membrane attachment (Karlsson, 1995).

The first well-characterized carbohydrate moiety to be established as a microbial receptor is the ganglioside GM1, used by cholera toxin for attachment to
enterocytes (Spangler, 1992). Since then, an overwhelming number of reports have established the ability of microbial ligands to attach to carbohydrates. Bacterial binding to specific carbohydrate moieties includes P piliated E. coli binding to Galα(1-4)βGal sequences (Lindberg et al., 1987); S piliated E. coli binding to NeuNAca(2-3)Galβ (Schmoll et al., 1989); Propionibacterium granulosum binding to Galβ(1-4)Glc sequences (Hansson et al., 1985); Actinomyces naeslundii binding to GalNAcβ-containing glycoconjugate (Brennan et al., 1984). By thin-layer chromatography (TLC) overlay, our group have shown that many bacterial pathogens share a similar lipid-binding specificity in that phosphatidylethanolamine (PE), gangliotriaosylceramide (GalNAcβ(1-4)Galβ(1-4)Glc Ceramide, Gg3), and gangliotetraosylceramide (Galβ(1-3)GalNAcβ(1-4)Galβ(1-4)Glc ceramide, Gg4) are recognized (Table 1) (Lingwood et al., 1991, 1992; Jagannatha et al., 1991; Krivan et al., 1991; Yu et al., 1994). In addition, several respiratory pathogens including Bordetella pertussis, Mycoplasma pneumoniae and Haemophilus influenzae, share a sulfatogalactosylceramide (SGC)-binding specificity (Brennan et al., 1991; Zhang et al., 1994; Busse et al., 1997).

In protein-oligosaccharide interactions the binding specificity and strength are mainly based on the hydrogen bonds between the oligosaccharide and the protein (Quirocho, 1986). Other factors that influence carbohydrate recognition and binding are the lipid moiety of glycolipid, the architecture of the receptor binding site, hydrophobic interactions, and the orientation and conformation of the saccharide isomers and their synthetic derivatives (Quirocho, 1986; Lingwood, 1992). The lipid environment can modify the ability of the carbohydrate to be recognized. Phospholipids containing long-chain fatty acids were found to obscure the carbohydrate moiety of glycolipids, whereas short-chain fatty phospholipids promote exposure. Conversely, short-chain fatty acids within the glycolipid

-9-
ceramide moiety were found to reduce exposure whereas longer chain species enhanced anti-glycolipid antibody binding. Consequently, the glycolipid may be indented or protrude from the plane of the plasma membrane according to its fatty acid chain lengths and the surrounding phospholipid hydrocarbon chain lengths (Lingwood, 1992). The influence of the lipid moiety may allow the differential presentation of epitopes on the same carbohydrate sequence to permit multiple ligand binding, thereby increasing the receptor repertoire of a given sugar sequence (Lingwood, 1992).

1.3. Phospholipid receptor - Phosphatidylethanolamine (PE)

In various eukaryotic cells, membrane phospholipid asymmetry is ubiquitous. Typically, the inner leaflet of plasma membranes contains essentially all phosphatidylserine (PS), most phosphatidylethanolamine (PE) and phosphatidylinositol (PI), while sphingomyelin (SM) is largely confined to the outer leaflet, and phosphatidylcholine (PC) is distributed equally between both leaflets (Devaux, 1991; van Meer, 1993). This asymmetric distribution of membrane phospholipids is maintained partly by an enzyme, ATP-dependent aminophospholipid translocase, which specifically catalyzes the transportation of PS and PE from the outer leaflet to the inner membrane (Seigneuret & Devaux, 1984).

In the early stages of apoptosis, PS is translocated from the inner side of the plasma membrane to the outer layer, which allows phagocytes to recognize and engulf the apoptotic cells (Fadoc et al., 1992). A recent study has shown that PE is also exposed on the cell surface in the early phase of apoptosis (Emoto, et al., 1997). It suggests that a complete loss of asymmetric distribution of plasma membrane phospholipids occurs during the apoptotic process resulting in aminophospholipid
The mitotic phase of the cell cycle ends as the cytoplasmic components are divided by the process of cytokinesis. During cytokinesis, cell plasma membranes are drawn inward to form cleavage furrows, which gradually deepen and fuse with each other, resulting in cell division. PE is found exposed on the cell surface specifically at the cleavage furrow during the late telophase or GI phase of cytokinesis (Emoto et al., 1996). The redistribution of the plasma membrane phospholipids is a crucial step for cytokinesis and the cell surface PE may play a pivotal role in mediating a coordinated movement between the contractile ring and plasma membrane to achieve successful cell division (Emoto et al., 1996).

Some studies suggested that PE is kept in the bilayer configuration by interaction with other phospholipids in biological membranes (Cullis et al., 1986). However, reorganization of the membrane phospholipids could lead to expression of the nonbilayer nature of PE and induce bilayer instability. Its preference for nonbilayer configuration is thought to play an essential role in membrane fusion (Verkleij, 1984). In addition, some studies demonstrated that PE is a major target of lupus anticoagulant activity (Rauch & Janoff, 1992) and plays a role in enhancing the inactivation of factor Va by activated protein C (Smirnov & Esmon, 1994), suggesting that PE may be important in thrombosis.
Table 1. Microorganisms tested for binding to ganglio-series glycolipids and PE

<table>
<thead>
<tr>
<th>Organism</th>
<th>Binding to Ganglio-series Glycolipids</th>
<th>Binding to PE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bordetella pertussis</td>
<td>+</td>
<td>not tested</td>
</tr>
<tr>
<td>Borrelia burgdorferi</td>
<td>+</td>
<td>not tested</td>
</tr>
<tr>
<td>Campylobacter upsaliensis</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Chlamydia pneumonia</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Chlamydia trachomatis</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Clostridium difficile</td>
<td>+</td>
<td>not tested</td>
</tr>
<tr>
<td>Clostridium perfringens</td>
<td>+</td>
<td>not tested</td>
</tr>
<tr>
<td>Coxiella burnetti</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Cryptococcus neoformans</td>
<td>+</td>
<td>not tested</td>
</tr>
<tr>
<td>EPEC</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Haemophilus influenzae</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Helicobacter mustelae</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Helicobacter pylori</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Klebsiella pneumoniae</td>
<td>+</td>
<td>not tested</td>
</tr>
<tr>
<td>Mycobacterium tuberculosis</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Mycobacterium avium-intracellulare</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Neisseria gonorrhoeae</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Neisseria meningitidis</td>
<td>+</td>
<td>not tested</td>
</tr>
<tr>
<td>Pasteurella multocida</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Pseudomonas aeruginosa</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Salmonella</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Shigella dysenteriae</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Staphylococcus aureus</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Streptococcus agalactiae (type b)</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Streptococcus pneumoniae</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>VTEC</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>
Recently, PE has been identified as a receptor for *H. pylori* and *H. influenzae* (Lingwood *et al.*, 1992; Busse *et al.*, 1997). Adhesion of Helicobacters, including *H. mustelae* and *H. pylori*, correlates with the quantity of PE present in the epithelial cell and with differences in the fatty acid composition of the phospholipid receptor (Gold *et al.*, 1995). Our group has found that many bacterial pathogens, including *Pseudomonas aeruginosa*, *Chlamydia trachomatis* and EPEC can specifically bind PE (Table 1). Furthermore, a PE-binding adhesin has been purified from *H. pylori*. The purified adhesin was an effective inhibitor of *H. pylori* binding to PE in vitro. TLC overlay confirmed the binding specificity of this purified adhesin for Gg3, Gg4 and PE (Lingwood *et al.*, 1993). Therefore, this adhesin is an appropriate focus for future studies on the role of adhesion in pathogenesis.

2. *Haemophilus influenzae*

2.1. General characteristics

*H. influenzae* is a small, nonmotile, Gram-negative, aerobic or facultatively anaerobic, rod-shaped or coccobacillary bacterium. This organism was first isolated by Pfeiffer during the 1892 influenzae pandemic and assumed to be the causative agent of this disease (Pfeiffer, 1893). The specific name of *Haemophilus influenzae* which was given to the organism is a permanent reminder of this erroneous association. The generic name *Haemophilus* evolved from the fact that this organism has an absolute nutritional requirement for hemin. Nevertheless, *H. influenzae* has been clearly implicated as the cause of a number of human diseases, including both local and serious systemic infections (Turk, 1984).
The best-known phenotypic trait of *H. influenzae* is its absolute requirement for both hemin and nicotinamide-adenine dinucleotide (NAD) under aerobic conditions, i.e. the classical X and V factors, respectively (Evans *et al.*, 1974). Hemin can serve as a source of both iron and porphyrin for this organism (White & Granick, 1963). *H. influenzae* is divided into six serotypes (designated a-f) on the basis of immunologically distinct capsular polysaccharide antigens. Additionally, strains may be nonencapsulated. These strains are defined on the basis of their failure to agglutinate with typing antisera against capsular serotypes and are considered nontypeable. In the past, nontypeable isolates of *H. influenzae* were considered to be phenotypic variants of type b strains (Huber *et al.*, 1985). However, studies by multilocus enzyme electrophoresis, indicated that most nontypeable isolates were distinguished from common type b strains (Porras *et al.*, 1986).

The encapsulated *H. influenzae*, predominantly type b, cause serious systemic infections such as meningitis, epiglottitis, septic arthritis. The unencapsulated, nontypeable strains mainly cause localized respiratory tract diseases, such as otitis media, sinusitis, bronchitis, conjunctivitis and pneumonia. Occasionally, nontypeable strains are a cause of serious systemic disease, including endocarditis, meningitis, and a recently described fulminant sepsis syndrome called Brazilian purpuric fever (Brazilian purpuric fever study group, 1987). In recent years the vaccines based on the type b polyribosyl-ribitol capsular polysaccharides are available and have dramatically reduced the incidence of systemic type b disease in developed and developing countries (Adams *et al.*, 1993; Steinhoff, 1997). On the other hand, existing *H. influenzae* vaccines fail to protect against nontypeable strains.

In 1995, the complete DNA sequence of the *H. influenzae* genome was reported (Fleischmann *et al.*, 1995). This was the first completed genome sequence of
a cellular life form. The *H. influenzae* Rd genome is a circular chromosome of 1830 kb. The overall G+C nucleotide content is approximately 38 percent. A total of 1743 predicted coding regions was identified. Of these, 736 have no role assignment (Fleischmann *et al.*, 1995).

2.2. Pathogenesis

*H. influenzae* is a strictly human pathogen that is passed from person to person by way of the respiratory tract. Up to 50% to 80% of individuals carry nontypeable *H. influenzae* (NTHI) asymptotically in the nasopharynx, whereas only fewer than 5% carry encapsulated type b organisms (Howard, 1992). Children are more likely to be colonized with *H. influenzae* than are adults (Turk, 1984). Asymptomatic carriage of encapsulated *H. influenzae* may induce protective immunity.

Disease caused by *H. influenzae* is believed to begin with colonization of upper respiratory mucosa. In most cases colonization persists in the absence of symptoms. In certain circumstances, colonization is followed by contiguous spread within the respiratory tract. Localization in the middle ear, the conjunctiva, or the lower respiratory tract results in disease at these sites. Occasionally bacteria penetrate the nasopharyngeal epithelial barrier and enter the bloodstream and induce systematic infections (Turk, 1984).

Successful pharyngeal colonization requires that an organism overcome the mucociliary clearance mechanism of the mucosal surfaces. *H. influenzae* elaborates a series of virulence factors, which are able to impair ciliary function, cause damage to the respiratory epithelium and facilitate the process of attachment to respiratory
epithelium (Moxon & Wilson, 1991).

2.2.1. Capsule

There are six immunologic capsular types (a-f). Since over 90% of all invasive *H. influenzae* disease in humans is due to type b strains, type b capsule has been studied extensively. The type b capsule, a linear polymer of ribosyl-ribitol phosphate (PRP), is a major determinant of virulence (Turk, 1984; Moxon & Vaughn, 1981). It confers virulence properties that are different from other serotypes (Moxon, 1992). The type b capsule is thought to contribute to the pathogenesis of *H. influenzae* infections by inhibiting neutrophil phagocytosis and resisting complement-mediated bactericidal activity, thus enhancing bloodstream survival of type b strains (Tunkel & Scheld, 1993).

The genes for type b capsule are chromosomal. The *cap* b locus contains an unstable direct repeat of 17 kb of DNA joined by a 1-1.3 kb bridge region containing the gene bexA (Kroll *et al.*, 1988). Each repeat contains genes necessary for polysaccharide synthesis, export, and surface expression. BexA is a critical component of the polysaccharide exporter, which is homologous to members of a family of ATP-dependent transport systems (Kroll, 1992). In addition, the *cap* locus has the structure of a compound transposon: copies of the insertion element IS1016 flank the gene cluster. This structure gives strains the capacity to amplify genes at *cap* by unequal homologous recombination (Kroll *et al.*, 1991).

As a consequence of the genetic configuration of the *cap* locus, type b strains become capsule deficient at a high frequency. The capsule-deficient mutants demonstrated significantly greater adherence and invasion than the encapsulated
parent (St. Geme & Falkow, 1992), suggesting that capsule loss by type b strains may be relevant to colonization of the nasopharynx and persistence within the respiratory tract.

### 2.2.2. Outer membrane proteins (OMP)

The outer membrane proteins (OMP) of *H. influenzae* show marked antigenic variation. Several SDS-PAGE subtyping schemes have been devised. At least 13 antigenic groups of type b strains have been defined on the basis of OMPs (Erwin & Kenny, 1984). Some studies have demonstrated similar antigenic diversity of the major OMPs among nontypeable *H. influenzae* (NTHI)(Murphy et al., 1983). In addition, antibodies directed against several OMPs have been shown to have protective or bactericidal activity (Munson & Granoff, 1985; Granoff & Munson, 1986). Some OMPs may be involved in adhesion of *H. influenzae* (Reddy et al., 1996).

#### i. Outer membrane protein P1

OmpP1 is a heat-modifiable protein. It has a larger apparent molecular size when solubilized in SDS at 100°C than when solubilized at room temperature. Antibodies directed against P1 have been shown to have protective activity in an infant rat bacteremic model (Granoff & Munson, 1986). Moreover, a murine P1-specific monoclonal antibody and rabbit antiserum raised against purified P1 from either typeable or nontypeable strains were also found to be protective in animal models (Loeb, 1987; Pelton et al., 1990). Recently, a surface-exposed epitope was mapped by using the protective monoclonal antibody. This epitope is located within residues 184 and 193 (Panezutti et al., 1993). The P1 gene from a type b strain has
been cloned and sequenced. The molecular weight of the P1 protein is approximately 47 kDa. This protein is 42% identical and 61.5% similar to the FadL protein of E. coli, a heat-modifiable OMP involved in long-chain fatty acid transport. Although P1 was unable to restore long-chain fatty acid transport efficiently in a fadL-deficient E. coli strain, it seems that P1 is the Haemophilus analog of the E. coli FadL protein (Muwon et al., 1989).

The P1 proteins from Haemophilus influenzae type b (Hib) are heterogeneous antigenically and with respect to apparent molecular weight in SDS-PAGE. The differences in the mobilities of the P1 proteins have been used to subtype Hib strains. In the subtyping scheme, these proteins were designated H, L, or U. The P1 gene has been analyzed from three OMP subtype strains (Munson et al., 1989). Each strain produces a single P1 protein. The P1 genes are highly conserved. Three variable regions account for the majority of the sequence heterogeneity.

ii. Outer membrane protein P2

OmpP2 is the predominant outer membrane protein, ranging in molecular mass from 36-42 kDa among H. influenzae strains. P2 has been shown to be a target of human bactericidal antibody. Antibody directed against P2 has protective activity in the infant rat bacteremic model (Munson et al., 1983). However, P2 demonstrates substantial strain heterogeneity and some antigenic drift (Groeneveld et al., 1989; Haase et al., 1991). The P2 gene has been cloned and sequenced from Hib and NTHI strains. Based on the hydrophobicity, amphiphilicity and turn propensity of the amino acid sequence, P2 is predicted to contain 16β strands which traverse the membrane and eight surface-exposed loops (Srikumar et al., 1992). Comparison of the sequences of the P2 genes of various H. influenzae strains revealed highly
conserved sequences coding for the membrane-spanning regions and highly variable sequences coding surface-exposed loops. Variation is particularly notable in loops L2, L4, L5, and L8 (Sikkema & Murphy, 1992; Forbes et al., 1992; Bell et al., 1994). It is clear that the molecular mass and antigenic heterogeneity of P2 is due to variations in gene sequence. The antigenic differences in the P2 proteins among H. influenzae strains may induce a strain-specific host immune response and allow other strains to cause recurrent infection. During persistent infections in patients with chronic bronchitis, changes in the molecular weight of P2 were observed (Groeneveld et al., 1989). This variation resulted in antigenic drift since the variants were no longer recognized by strain-specific antibodies recognizing P2 (Groeneveld et al., 1989; van Alphen et al., 1991). Duim et al. reported the sequence changes which were responsible for antigenic variation in OmpP2 were localized to L6, suggesting that a portion of L6 is also at the cell surface (Duim et al., 1994). Recently, bactericidal epitopes of P2 have been mapped to two different surface-exposed loops of the P2 molecule, i.e. loop L5 and loop L8 (Haase et al., 1994; Yi & Murphy, 1994). Furthermore, a strain-specific immune response to NTHI was found to be directed at an immunodominant epitope on the loop L5 region of the P2 molecule (Yi & Murphy, 1997).

P2 has been shown to function as a porin in H. influenzae (Vachon et al., 1986). Porins are water-filled channels formed by proteins that span the bacterial outer membrane and that confer on the outer membrane the properties of a molecular sieve (Nikaido, 1992). As the sole porin for H. influenzae, P2 exists as a trimer and is closely associated with lipooligosaccharide (Burns & Smith, 1987). The pores of P2 porin are somewhat larger than those produced by the major porins of E. coli: oligosaccharides with a molecular mass above 1.4 kDa are permeable through the P2 porin (Vachon et al., 1986).
An isogenic mutant of Hib lacking the ability to express the P2 protein has been constructed (Cope et al., 1990). The P2 mutant is viable, although it grows more slowly in vitro than the parent strain. However, the loss of P2 had a drastic effect on the virulence of Hib. The P2-deficient mutant is not virulent in the infant rat bacteremic model. In addition, the introduction of the functional NTHI OmpP2 gene into the mutant restored the growth phenotype and virulence to wild-type levels (Sanders et al., 1993), suggesting that a NTHI OmpP2 can be expressed and function properly in the Hib outer membrane.

Mucins are high-molecular-weight glycoproteins and major constituents of the mucus layer which covers the airway surface. Since H. influenzae normally resides in the nasopharynx, mucins of the nasopharyngeal mucus may function as receptor molecules for this organism and thus play an important role in the colonization of the nasopharynx. A recent study has shown that P2, P5 and unidentified lower molecular weight outer membrane proteins of NTHI appear to function as adhesins for human nasopharyngeal mucin and sialic acid-containing oligosaccharides which appear to be the receptors for NTHI (Reddy et al., 1996).

iii. Outer membrane protein P4 (e)

OmpP4 is an outer membrane lipoprotein of molecular weight of 29 kDa. The P4 gene has been cloned and sequenced (Green et al., 1991). The P4 protein is antigenically well conserved among Hib and NTHI strains. Antibodies directed against P4 have bactericidal activity to clinical isolates of H. influenzae (Green et al., 1991). Recently, P4 has been identified as one of the key components involved in the utilization by H. influenzae of hemin, protoporphyrin IX, and hemoglobin as
sources of porphyrin (Reidl & Mekalanos, 1996). The P4 gene has been found to complement hemA mutant of E. coli for growth on hemin or protoporphyrin IX as sole porphyrin sources. Construction of the P4 deficient mutant demonstrated that P4 is essential for growth under aerobic conditions but not under anaerobic conditions. The aerobic growth defect of P4 mutants could be reversed by providing exogenous hemin in the presence of outer membrane perturbers such as EDTA, suggesting that the mutants are defective in the transport of hemin through the outer membrane. The N-terminal region of the P4 gene may be involved in hemin binding and/or transport (Reidl & Mekalanos, 1996).

IV. Outer membrane protein P5

OmpP5 is the Haemophilus analog of the E. coli OmpA. The P5 protein has 50% identity and 65% similarity to the OmpA of E. coli, and also has heat modifiability properties similar to those of the OmpA (Munson et al., 1985). Additionally, antibody directed against P5, which was cross-reactive with the E. coli OmpA, did not protect rats against bacteremia when challenged with H. influenzae (Munson & Granoff, 1985). OmpA is a very abundant constituent of the outer membrane of many gram-negative bacteria including other Haemophilus species such as H. somnus and H. ducreyi (Tagawa et al., 1993; Spinola et al., 1993). This protein is thought to be important in the maintenance of the integrity of the E. coli outer membrane and has recently been shown to have porin activity (Sugawara & Nikaido, 1992). An OmpA deficient mutant of E. coli was shown to have reduced virulence in an infant rat model of bacteremia (Weiser & Gotschlich, 1991).

Recently, a thin, nonhemagglutinating pilus of NTHI strains has been characterized (Sirakova et al., 1994). This pilus refers to the thin (diameter,
approximately 2.4 nm), peritrichously arranged, flexible, non-hollow-core, and nonhemagglutinating filaments, which are expressed by all the NTHI isolates recovered from the middle ears and nasopharynges of children with chronic otitis media (Bakaletz et al., 1989). Transmission electron microscopic observations indicated that this pilus may be involved in the initial docking or adherence of NTHI to mucosal epithelial cells (Bakaletz et al., 1988). The gene which codes for pilin in NTHI strain has been cloned and sequenced (Sirakova et al., 1994). The pilin sequence showed 92% identity with P5 from Hib. This degree of homology strongly suggests that the pilin gene is the P5 homolog of NTHI. A pilin gene deficient mutant has been constructed. This mutant adhered significantly less to human oropharyngeal cells and middle ear mucus than the parent strain. The mutant also showed reduced virulence in two chinchilla models. In addition, the purified pilin adhered to middle ear mucus and blocked adherence of whole organisms (Miyamoto & Bakaletz, 1996). Furthermore, antibodies directed against pilin protein afforded protection against NTHI induced otitis media in a chinchilla model (Sirakova et al., 1994).

V. Outer membrane protein P6

OmpP6 is an approximately 16 kDa lipoprotein that is partially exposed on the bacterial surface and makes up about 1 to 5% of the outer membrane proteins in H. influenzae (Munson & Granoff, 1985; Murphy et al., 1986). The P6 protein contains regions which have sequence homology to the peptidoglycan-associated lipoprotein (Pal) of E. coli. E. coli Pal is important in maintaining the structural integrity of the outer membrane. P6 may perform a similar function in H. influenzae (Deich et al., 1988). It has been demonstrated that P6 is antigenically stable and highly conserved at both the amino acid and nucleotide levels in both Hib and NTHI strains of
diverse geographic and clinical origin (Nelson et al., 1991). Since there seems to be an inability to construct a H. influenzae P6 mutant, a functional P6 protein may be essential for the growth of the organism (Bogdan & Apicella, 1995).

Several lines of evidence suggest that P6 is an important protective antigen and a potential vaccine component. First, antibodies to P6 are protective in the infant rat model of invasive Hib infection (Munson & Granoff, 1985). Second, antibodies to P6 have been detected in sera, middle ear effusions, nasopharyngeal secretions and breast milk. Immunopurified antibody to P6 from human serum was bactericidal for H. influenzae (Murphy et al., 1986). It has been observed that prevention of colonization was most evident during breast-feeding (Harabuchi et al., 1994). Third, antiserum raised to purified recombinant P6 was bactericidal for H. influenzae, at least for all strains tested (Green et al., 1990). Fourth, following mucosal immunization of rats with P6 enhanced respiratory clearance of NTHI has been observed (Kyd et al., 1995). Finally, in the chinchilla model of otitis media, immunization with P6 from NTHI induced bactericidal antibody and afforded protection by a reduced incidence and duration of culture-positive middle ear fluids (DeMaria et al., 1996).

Recent studies have shown that P6 binds to its own gene and may regulate its own expression (Sikkema et al., 1992). This has potentially important implications as a mechanism for regulation of expression of this outer membrane protein. The region of the P6 molecule from amino acids 68 to 88 contains a helix-turn-helix sequence that is in good agreement with the consensus sequence shared by the family of helix-turn-helix DNA-binding proteins (Sikkema et al., 1992). In addition, a surface-exposed, conformational epitope of the P6 protein was mapped using monoclonal antibody 3B9. This epitope is composed of two discontinuous regions of
22.3. Lipooligosaccharide (LOS)

Lipopoly saccharides (LPS) are one of the major outer membrane constituents of all gram-negative bacteria. These glycolipids play key roles in the biology of these organisms and have been found to be important virulence factors in pathogenic species. The LOS of *H. influenzae* is analogous to the LPS of other gram-negative bacteria; it consists of lipid A linked by 2-keto-3-deoxyoctulosonic acid (KDO) to a conserved heptose-containing inner core trisaccharide moiety. Each heptose within this triad can provide a point for further oligosaccharide chain elongation, leading to a vast array of possible structures (Schweda *et al.*, 1995). Unlike LPS, LOS lacks the repeating polysaccharide O antigen. Various studies have demonstrated that the LOS of *H. influenzae* is an important factor in pathogenesis and virulence (Zwahlen *et al.*, 1986). LOS facilitated the survival of *H. influenzae* within the nasopharynx, dissemination from the nasopharynx to the blood, and induction of central nervous system injury (Moxon, 1992). The lipid A portion of LOS is responsible for the toxicity associated with this organism, but the role of the oligosaccharide portion of the LOS in pathogenesis of *H. influenzae* infection is less clear. Differences in carbohydrate residues have been found to confer differences in bacterial virulence (Moxon, 1992). In addition, the LOS phenotype has been demonstrated to be a critical determinant of virulence of the Brazilian purpuric fever clone of *H. influenzae* biogroup aegyptius for infant rats (Rubin & St. Geme, 1993).

The LOS of *H. influenzae* can undergo rapid switching or phase variation between defined structures which leads to an extensive repertoire of oligosaccharide epitopes within a single strain (Weiser, 1993). It is believed that phase variation
provides a mechanism whereby the pathogen can adapt to variation in environmental conditions within and between individual hosts. Phase variation has been found to enhance the invasive capacity of Hib (Weiser et al., 1990). Moreover, some of the phase-varying epitopes have been found to mimic human blood group antigens, such as the Pk antigen (Galα1-4Galβ1-4Glc) and paragloboside (Galβ1-4GInNAcβ1-3Galβ1-4Glc). One molecular mechanism of phase variation involves a translational on-off switch created by slipped-strand mispairing of highly repetitive DNA sequences found in multiple chromosomal loci for LPS biosynthesis (Weiser et al., 1989).

2.2.4. Immunoglobulin A1 protease

IgA is the predominant immunoglobulin in the external secretions. Secretory IgA (S-IgA) constitutes a complex of IgA dimers or tetramers with incorporated J (joining) chain linked to the secretory component (SC), an 80 kDa glycoprotein of epithelial origin. There are two IgA subclasses, IgA1 and IgA2. IgA1 is vastly predominating in blood, bone marrow, spleen, lymph nodes, and in the upper part of the respiratory tract, i.e. nasopharynx. In contrast, IgA2 may predominate in the distal parts of the gastrointestinal tract and in urogenital secretion (Brandtzaeg, 1992). *H. influenzae* can produce an extracellular protease rendering the microorganism capable of inactivating human IgA1. This property is also shared by other bacterial pathogens such as *Neisseria meningitidis* and *Streptococcus pneumoniae* (Plaut, 1983). One common characteristic of all bacterial IgA1 proteases is that they can specifically cleave the heavy chain of the IgA1 in the hinge region. The enzymatic products are monomeric Fab fragments with retained antigen binding capacity and the Fc portion in a monomeric or dimeric form associated with
the secretory component (Plaut, 1983). IgA2 is not cleaved by these proteases, as it lacks the portion of the hinge region (Plaut et al., 1974). The gene encoding IgA1 protease of H. influenzae has been cloned and sequenced (Poulsen et al., 1989). The mature IgA1 protease is approximately 100 kDa. Two types of IgA1 protease have been defined on the basis of the exact cleavage site in the hinge region in the human IgA1. H. influenzae type 1 and type 2 IgA1 proteases cleave between proline and serine (position 231 and 232) and between proline and threonine (position 235 and 236), respectively (Kilian et al., 1983). Several lines of indirect evidence indicate that IgA1 proteases enable bacteria to colonize mucosal surfaces in the presence of adherence-inhibiting S-IgA1 antibodies (Kilian et al., 1988). Cleavage of IgA1 is often extensive and may result in a local IgA deficiency that facilitates the colonization by microorganisms. Additionally, binding of Fab fragments to bacterial surface epitopes may protect bacteria by blocking access to intact antibody molecules (Kilian et al., 1988). Various studies have shown that IgA1 protease activity is important for successful colonization of H. influenzae on mucosal membranes (Lomholt et al., 1993).

H. influenzae IgA1 proteases show considerable antigenic polymorphism, in particular among nontypeable strains (Poulsen et al., 1992). At least 30 antigenic types of H. influenzae IgA1 proteases have been identified by using enzyme-neutralizing antisera (Lomholt et al., 1993). This antigenic polymorphism appears to be the result of horizontal gene transfer between clones expressing antigenically different IgA1 proteases. This seems to be the principal mechanism by which H. influenzae evades the host immune response against IgA1 protease.
2.2.5. Hemagglutinating Pili

The hemagglutinating pili of *H. influenzae* mediate its adherence to human cells and may be important in the establishment of colonization of the human upper respiratory tract. It has been clear that piliated *H. influenzae* strains can adhere to human epithelial cells *in vitro* in significantly greater numbers than their nonpiliated variants (Pichichero, 1984). Moreover, a piliated *H. influenzae* strain was found to be more effective than its nonpiliated variant in colonizing rats following intranasal inoculation (Anderson et al., 1985). Weber et al. showed that a piliated Hib strain colonized 1-year-old monkeys more effectively than a pilus-negative mutant (Weber et al. 1991). In addition, the hemagglutinating pili confer binding to mucus (Read et al., 1992) and cause agglutination of human erythrocytes via the AnWj blood group antigen (van Alphen et al., 1991). This antigen is present on the surfaces of most human erythrocytes. The adherence to epithelial cells involves a sialic acid-containing lactosylceramide eukaryotic receptor (van Alphen et al., 1991).

The hemagglutinating pili are helical structures approximately 5 nm in diameter and up to 450 nm in length and possess a hollow core. The pilus rods are composed of polymerized pilins, which are subunit proteins ranging in size between 22 to 27 kDa, depending on the strain (Gilsdorf et al., 1997). The gene *(hisA)* encoding pilin has been cloned and sequenced from Hib and NTHI strains (Coleman et al., 1991; Forney et al., 1991; Gilsdorf et al., 1992; St Geme & Falkow, 1993). These pilins are significantly homologous, with 63% to 81% amino acid identity. The first 15 amino acids of the leader sequence are completely conserved in all of these strains. The mature pilins demonstrate significant sequence similarity with several *E. coli* pilins, including PapA. Like pilin proteins from other gram-negative bacteria, *H. influenzae* pilins.
*H. influenzae* pilins also contain two cysteines at similar positions, and tyrosine and glycine residues at the C-terminus. The *hifA* locus contains five genes, from *hifA* to *hifE*, which comprise the *H. influenzae* pilus gene cluster (Gilsdorf et al., 1997). The *hifB* gene shows significant homology to the *papD* gene of the pilus operon, a periplasmic chaperone that functions to stabilize pilin subunits in assembly-competent complexes prior to their ordered incorporation into the growing pilus. HifB appears to have a similar function to PapD; it forms stable complexes with the structural pilin, holding this subunit in a native conformation (Watson et al., 1994; St Geme et al., 1996a). The *hifC* gene encodes a protein with significant homology to the pilus assembly platform ( usher), including PapC and FimD of *E. coli*. HifC presumably plays a role in receiving structural pilin subunits from the HifB chaperon and directing their incorporation into hemagglutinating pili (van Ham et al., 1994; Watson et al., 1994). The *hifD* and *hifE* gene products have homology to each other and to pilin. Mutation studies indicate that HifD and HifE are required for the expression of functional pili. Recent studies show that HifD and HifE are two minor tip proteins, and HifE is located at the tips of pili. HifE antiserum completely blocked pilus-mediated hemagglutination, suggesting that HifE mediates pilus adherence (McCrea et al., 1994; 1997).

At least 14 serotypes of hemagglutinating pili have been recognized, indicating the presence of type-specific epitopes (Gilsdorf, 1992). Various studies have indicated that linear epitopes on pilins are highly conserved among both type b and nontypeable piliated strains, but epitopes on polymerized pili are conformational and vary significantly from strain to strain (Gilsdorf et al., 1990; 1992). Antibodies to *H. influenzae* pili block buccal cell adherence of the homologous piliated strain and show either reduced or no interference of stains expressing immunologically heterologous pili. Furthermore, antipilus antibodies
are bactericidal for the homologous strain but are not bactericidal for strains expressing heterologous pili (LiPuma & Gilford, 1988).

An apparently important functional characteristic of *H. influenzae* pili is phase variation. During natural infection, nasopharyngeal isolates are often piliated, while their isogenic counterparts from systemic sites are virtually always nonpiliated, though they can be enriched for piliation in vitro (Pichechero *et al.*, 1984; Mason *et al.*, 1985). The rates of transition are believed to be roughly $10^{-3}$ to $10^{-4}$ in both directions (Farley *et al.*, 1990). The phase variation is transcriptionally regulated by variation of the length of the reiterated sequence that forms the promoter region of the subunit gene (van Ham *et al.*, 1993). The promoter regions of *hifA* and *hifB* overlay and the area of overlap contains a variable number of TA repeats. Nonpiliated variants contained 9 TA units. In contrast, piliated variants had 10 TA repeats and expressed stable transcripts. The changes in the number of TA repeats in the overlapping promoter region are caused by the process of slip-strand mispairing (van Ham *et al.*, 1993).

2.2.6. High-molecular-weight adhesins (HMW1 and HMW2)

A group of high-molecular-weight (HMW) surface-exposed proteins of NTHI are major targets of human serum antibody (Barenkamp & Bodor, 1990). The genes encoding two such HMW proteins (HMW1 & HMW2) have been cloned and sequenced from the NTHI strain, and their derived amino acid sequences were found to be closely related to the filamentous hemagglutinin (FHA) protein of *Bordetella pertussis*, a critical adherence factor of that organism (Barenkamp & Leininger, 1992). Immunoblot studies show that the HMW1 and HMW2 proteins are also antigenically related to filamentous hemagglutinin, suggesting that they
may serve similar biological functions (Barenkamp & Leininger, 1992). Subsequently, St. Geme et al. demonstrated that loss of expression of HMW1 by the NTHI prototypic strain significantly decreased the capacity to adherence. The absence of expression of both HMW1 and HMW2 in the prototypic strain was associated with a further decrease in adherence. Expression of either HMW1 or HMW2 in nonadherent laboratory strains of E. coli resulted in acquisition of the capacity for adherence to human epithelial cells (St. Geme et al., 1993). In addition, it has been demonstrated that the HMW proteins of NTHI are composed of a filamentous material localized to discrete regions of the cell surface and are evidently released into the medium (Bakaletz & Barenkamp, 1994). The surface localization and presence of released protein would permit the filamentous material to be readily accessible for interaction with both target epithelial cells and phagocytes, as well as for immune recognition.

HMW1 protein is 125 kDa in size and is encoded by a 4.6 kb open reading frame, while HMW2 protein is 120 kDa and is encoded by a 4.4 kb DNA fragment. Both proteins have 70% identity. Approximately 75% of unrelated NTHI isolates have proteins antigenically related to the HMW1 gene product (Barenkamp & Leininger, 1992). The HMW1 protein appears to interact with a glycoprotein receptor containing N-linked oligosaccharide chains with sialic acid in an alpha 2-3 configuration on cultured human epithelial cells (St. Geme, 1994). Interestingly, Hib strains were found to lack cross-reactive high-molecular-weight proteins, suggesting that both proteins are unique to nontypeable strains (Barenkamp & Leininger, 1992). Using the chinchilla experimental otitis media model, Barenkamp has demonstrated that immunization with the combination of HMW1 and HMW2 provides moderate protection against the homologous strain (Barenkamp, 1996).
Both HMW1 and HMW2 proteins show cellular specificity in their binding, suggesting that they interact with specific receptor molecules whose distribution varies from one cell type to another (Hultgren et al., 1993). HMW1 mediates very efficient attachment to cells derived from human conjunctiva (Chang), larynx (HEp2), and oral cavity (KB) but only low to moderate levels of adherence to cells derived from human cervix (ME-180) and virtually no adherence to cells from human endometrium (Hec-1B). In contrast, expression of HMW2 was associated with moderate levels of attachment to conjunctival, laryngeal, and endometrial cells but low-level attachment to cervical cells and minimal binding to cells from the oral cavity (Hultgren et al., 1993).

2.2.7. Other H. influenzae adhesins

Since NTHI strains that lack HMW1/HMW2-like proteins remain capable of efficient attachment to cultured human epithelial cells, this organism must have additional adhesion molecules. A gene encoding such an adhesion protein was recently isolated designated his for H. influenzae adhesin. Transformation of a non-adherent E. coli strain with plasmids containing the genetic locus encoding this protein give rise to E. coli transformants that adhered avidly to Chang conjunctival cells. The his gene encodes a protein of 114 kDa. Subsequent study revealed that an his homolog existed in 13 of 15 HMW1/HMW2-deficient NTHI strains. In contrast, the his gene was not present in any of 23 NTHI strains which expressed HMW1/HMW2-like proteins (Barenkamp & St Geme, 1996).

Recently, St. Geme et al. isolated a locus involved in expression of short, thin surface fibrils by Hib and showed that surface fibrils promote attachment to human epithelial cells (St Geme et al., 1996b). The H. influenzae surface fibril locus is
designated *hsf*. The deduced amino acid sequence of the *hsf* product demonstrated 81% similarity and 72% identity to Hia. Consistent with the finding, the Hsf and Hia proteins demonstrated the same binding specificities, suggesting that *hsf* and *hia* are alleles of the same locus. Additionally, a *hsf* homolog was found to be ubiquitous among encapsulated *H. influenzae* strains (St Geme et al., 1996b).

### 2.2.8. Putative PE-binding adhesin

Recently, a putative PE-binding adhesin of approximately 46 kDa was purified from *H. influenzae* using a PE-celite affinity matrix (Busse et al., 1997). The purified adhesin was a potent inhibitor of *H. influenzae* Gg3 and PE binding *in vitro*. Polyclonal antibodies specific for this protein were able to reduce the attachment of *H. influenzae* to cultured HEp2 epithelial cells *in vitro* (Busse et al., 1997). Additionally, preincubation of *H. influenzae* with this antibody specifically inhibited the binding of the organism to PE and Gg3 as monitored by TLC overlay (Busse et al., 1997). The N-terminal sequence of the purified adhesin from a NTHI strain was determined chemically and the sequence was searched in the TIGR genomic sequence of *H. influenzae* Rd strain (Fleischmann et al., 1995). The gene HI0119, encoded a putative adhesin B precursor, was identified with 100% match corresponding to the N-terminal sequence. The gene product was identified as a putative adhesin since the derived amino acid sequence has 24.5% identity and 48.3% similarity to the FimA of *Streptococcus parasanguis* (Fleischmann et al., 1995, Busse, 1997). FimA has been demonstrated to be a fimbrial adhesin in *S. parasanguis* (Oligino & Fives-Taylor, 1993). This adhesin is a member of the lipoprotein receptor antigen group 1 (Jenkinson, 1994). This group consists of six homologous lipoproteins, including FimA, ScaA, PsaA, SnaB, EfaA and SchA (Jenkinson, 1994).
3. Lipoprotein receptor antigen group (Lra1)

3.1. FimA

FimA is a 36 kDa surface protein associated with the fimbriae of *S. parasanguis* FW213 (Fenno et al., 1989). But FimA is not the fimbrial structural subunit because strains mutated in *fimA* continued to produce fimbriae (Fenno et al., 1995). Immunoelectron microscopy revealed FimA was localized at the tips of the fimbriae of *S. parasanguis* FW213 (Fenno et al., 1995). The gene encoding FimA is highly conserved in all four genetic groups of *Streptococcus sanguis* (Fenno et al., 1995). The *fimA* product has been overexpressed in *E. coli* and did not appear as a solubilized monomer, but instead as a heterogeneous aggregate of >2×10^6 kDa. The protein could be purified under denatured conditions. The purified, renatured FimA had a high affinity for its substrate in the salivary pellicles and specifically inhibited the attachment of *S. parasanguis* FW213 to saliva-coated hydroxyapatite (SCHA). The ability of FimA to bind to a receptor on SCHA and block the adherence of FW213 appears to be dependent on a specific three-dimensional conformation of its peptide chain (Oligino & Fives-Taylor, 1993). However, it is not clear what is the active binding portion of FimA in SCHA.

FimA has been found to be a major virulence determinant associated with *S. parasanguis* endocarditis (Burnette-Curley et al., 1995). Mutants deficient in the production of FimA are significantly decreased in their ability to bring about endocarditis in the rat model. Although the exact mechanism by which FimA functions as a virulence factor of endocarditis has yet to be determined, FimA may play a role in adherence to fibrin deposits associated with damaged cardiac valves (Fenno et al., 1993; Burnette-Curley et al., 1995). Furthermore, immunization with FimA conferred protective immunity to *S. parasanguis* in the rat model of
endocarditis, probably as a result of anti-FimA-mediated inhibition of bacterial adherence to platelet-fibrin deposits (Burnette-Curley et al., 1995).

Recent studies showed that the fimA locus of S. parasanguis encodes an ATP-binding membrane transport system (Fenno et al., 1995). The amino acid sequence of FimA contains the consensus lipoprotein cleavage site (LXXC) common to the 'periplasmic' binding proteins of gram-positive transport systems. The nucleotide sequence directly upstream of fimA contains two open reading frames (ORF), ORF5 and ORF1, whose deduced protein products are homologous to members of a superfamily of ATP-binding cassette membrane transport systems. The derived amino acid sequence of ORF5 is a 28.6 kDa membrane-associated protein that has the consensus binding site for ATP (GXXGXGKS). The deduced product of ORF1 is an extremely hydrophobic integral membrane protein of 30.8 kDa with a pattern of six potential membrane-spanning regions, typical of a component of these types of transport system. The nucleotide sequence downstream of fimA, ORF3, encodes a 20 kDa protein having significant homology with bacteriaferritin co-migratory protein (Fcp) of E. coli K-12. Northern blot analysis suggest that the fimA locus was transcribed as a polycistronic message (Fenno et al., 1995).

3.2. ScaA

Streptococcal coaggregation adhesin ScaA is a 34.7 kDa lipoprotein expressed by S. gordonii PK488, which coaggregates with Actinomyces naeslundii PK606. The ScaA sequence shows 91, 80, and 80% identity to SsaB, FimA and PsaA, respectively (Kolenbrander et al., 1994). Southern and Western blot analysis demonstrated that ScaA homolog is present in 12 oral streptococci (Anderson et al., 1993; Kolenbrander et al., 1994). ScaA is not essential for growth in complex media containing glucose,
since ScaA-deficient, coaggregation-defective mutants of *S. gordonii* PK488 have no noticeable differences in growth rate or final growth yield (Kolenbrander et al., 1994). The *scaA* locus has been sequenced. The orientation of four ORFs, two upstream ORF1 and ORF2) and one downstream (ORF4) of *scaA*, indicate transcription in one direction. The putative 28,054 Da protein encoded by ORF1 contains a consensus binding site for ATP (GXXGKXGGK). ORF2 potentially encodes a hydrophobic protein of 29,705 Da with six potential membrane-spanning regions. ORF4 potentially encodes a 163-amino-acid protein of 17,912 Da, which was nearly identical to the downstream adjacent gene products of *ssnB*, *fimA* and *psaA*. The genetic organization of the *scaA* locus is similar to those of the bacterial periplasmic-binding protein-dependent transport systems of gram-negative bacteria (Kolenbrander et al., 1994).

3.3. *PsaA*

Pneumococcal surface adhesin A (*PsaA*) is a 37 kDa surface protein present on *Streptococcus pneumoniae*. The *psaA* gene product is 92.3% and 80% homologous to the FimA from *S. parasanguis* and SsaB from *S. sanguis* respectively (Sampson et al., 1994). Northern blot analysis of pneumococcal RNA suggested that *psaA* is transcribed as part of a polycistronic message, like *fimA*. Immunoblot analysis studies using anti-PsaA monoclonal antibody showed that PsaA is common to all 23 pneumococci vaccine serotypes (Russell et al., 1990). Based on restriction analysis of PCR products, Sampson et al. showed that *psaA* is highly conserved among pneumococci belonging to different capsular serotypes (Sampson et al., 1997). Recently, Talkington et al. demonstrated that PsaA is a protective immunogen in mice (Talkington et al. 1996). An isogenic mutant of *psaA* has been constructed. Mutagenesis of the *psaA* gene did not affect in vitro growth of pneumococci.
However, this mutant was found to be significantly less virulent than the wild-type strain, as judged by intranasal or intraperitoneal challenge of mice (Berry & Paton 1996). It suggested that a functional psaA gene is essential for full virulence of S. pneumoniae.

3.4. Other members of the Lral family

Saliva-binding protein (SsaB) is a 36 kDa lipoprotein from a salivary aggregating strain of S. sanguis 12 (Ganeshkumar et al., 1991, 1993). SsaB is 91% and 73% homologous to the ScaA from S. gordonii and FimA from S. parasanguis respectively. Immunogold bead labeling with antibody raised against SsaB showed that the protein is present on the cell surface of S. sanguis 12. The purified SsaB was demonstrated to inhibit the adhesion of S. sanguis 12 to saliva-coated hydroxyapatite (Ganeshkumar et al., 1988).

The efaA gene was identified using serum from an Enterococcus faecalis endocarditis (Lowe et al., 1995). Nucleotide sequence analysis of efaA revealed a 924 bp open reading frame encoding a protein with a predicted molecular weight of 34,768. The amino acid sequence of EfaA shows 55 to 60% homology to a group of streptococcal proteins, FimA from S. parasanguis, SsaB from S. sanguis, ScaA from S. gordonii, and PsaA from S. pneumoniae (Lowe et al., 1995).

A new member of the Lral family, scbA, was recently cloned from S. cristia CCSA (Correia et al., 1996). The scbA gene appears to be part of an ABC transport operon and encodes a putative peptide of 34.7 kDa. Compared with the five members of the Lral protein family, ScbA has an amino acid identity ranging from 56.5% with EfaA to 92.9% with ScaA. Surprisingly, ScbA does not exhibit adhesion
properties characteristic of other Lra1 proteins. Strain CCSA binds poorly to saliva-coated hydroxyapatite and does not coaggregate with *Actinomyces naeslundii* PK606. An *schA* insertion-duplication mutation that abolished expression of ScbA has been created. There is no difference in fibrin binding between this mutant and wild-type CCSA (Correia et al., 1996).

4. ATP-binding cassette (ABC) transport system

The ABC superfamily is one of the largest and most diverse families of proteins that mediate the selective movement of solutes across biological membranes. Over 100 different ABC transporters have been identified in both prokaryotes and eukaryotes (Higgins, 1992). Each is specific for a single substrate or group of related substrates that can range from inorganic ions to sugars and amino acids, complex polysaccharides and proteins. Members of this family include biomedically important proteins such as the mammalian multidrug resistance proteins (MDR), producing a drug resistance phenotype in cancer cells (Chen et al., 1986), and the human cystic fibrosis transmembrane conductance regulator (CFTR), the mutations of which cause the lethal disease cystic fibrosis (Riordan et al., 1989).

ABC transporters of all types generally consist of four domains that may be expressed as separate polypeptides or can be fused together into larger, multidomain proteins (Higgins, 1995). Two transmembrane domains span the membrane multiple times to form the pathway through which solute crosses the membrane and determine the substrate specificity of the transporter. Two ATP-binding domains, located at the cytoplasmic face of the membrane, couple ATP hydrolysis to solute movement. In some transporters, the four domains are expressed as separate polypeptides (e.g., the oligopeptide permease of *S. typhimurium*). In others, the
domains can be fused in a variety of configurations. For example, the two ATP-binding domains are fused into a single polypeptide in the ribose transporter of E. coli, and the two transmembrane domains are fused in the Fe-hydroxamate transporter of E. coli. Interestingly, MDR and CFTR have all four domains fused into a single multidomain polypeptide (Higgins, 1995).

The ATP-binding domains usually share 30-40% sequence identity with the equivalent domains from other ABC transporters, irrespective of their substrate specificity or species of origin (Mimura et al., 1989). These domains harbor the highly conserved Walker A and Walker B motifs responsible for the interaction with ATP. The Walker motifs are in fact signature sequences of the ABC transporter (Ames, 1986; Higgins, 1992).

Bacterial binding-protein-dependent transport systems belong to the superfamily of ABC transporters. They consist of at least three parts: one or two integral membrane proteins, one or two peripheral membrane ATP-binding proteins exposed to the cytoplasm, and a periplasmic substrate binding protein. In gram-positive bacteria, the substrate binding proteins are membrane-associated lipoproteins exposed to the cell surface. Usually these three subunits are encoded together in one operon (Ames, 1986). The integral membrane protein components are believed to form solute-specific channels; the peripheral membrane ATP-binding proteins energize the systems, and the ligand-bound binding proteins confer specificity and high affinity for the substrates to the transport system (Doige & Ames, 1993).
Figure 2. Schematic representation of a generic periplasmic permease. The outer membrane contains proteinaceous pores that allow the substrate (solid bars) to enter the periplasm, where it is bound by periplasmic binding protein (receptor), represented by a sphere with an empty binding site or a rectangle when liganded to the substrate. The four membrane components (hydrophobic, diagonal shading; hydrophilic ATP-binding, horizontal shading) form a complex within the cytoplasmic membrane. When the receptor changes conformation upon binding substrate it interacts with the membrane complex. The squiggle indicates the involvement of ATP in energy coupling. Adapted from Doige & Ames, 1993.
Periplasmic substrate binding proteins are generally soluble monomeric proteins with molecular weights varying from about 25 kDa to about 56 kDa. They have high affinity for their substrates (Doige & Ames, 1993). The ligand-bound binding protein undergoes a conformational change which allows its interaction with one of the membrane-bound components (Fig. 2) (Doige & Ames, 1993). Conformational changes in the membrane-bound apparatus triggered by the interaction with the liganded binding protein elicit both the release of substrate from the binding protein and the appearance of binding site(s) on the membrane-bound component(s), which allow passage of the substrate from one binding site to the other, into the inside of the cell (Ames, 1986; Doige & Ames, 1993).

Three-dimensional analyses of numerous periplasmic binding proteins of gram-negative bacteria have revealed that they are composed of two lobes, connected via two or three peptide stretches, and separated by a deep cleft in the unliganded protein; the binding site is located in the cleft. When the ligand is bound, the lobes are placed close to each other and the ligand is completely buried (Sack et al., 1989). The conformational change from the open to the closed form involves a large scale, rigid body movement of one lobe relative to the other. The bilobed, two-domain structures of the binding protein probably arose by a duplication event from a primordial gene encoding a simple one-domain protein (Oh et al., 1993). Interestingly, despite the close structural resemblance of the different periplasmic binding proteins, their homology based on their deduced amino acid sequence is surprisingly limited (Tam & Saier, 1993). In addition, although the periplasmic binding proteins are essential constituents of bacterial ABC-type uptake systems, extracellular substrate binding proteins have never been detected for bacterial efflux systems of the ABC type, and eukaryotic substrate binding proteins homologous to the bacterial proteins have not yet been found. The
periplasmic binding proteins appear to form a prokaryotic-specific family (Tam & Saier, 1993).

In bacterial binding-protein-dependent transport systems, the number of binding proteins in the periplasm nearly always exceeds the number of the cognate membrane compounds. The estimation in the maltose transport system is a 30 to 50-fold excess of binding protein over the membrane components (di Guan et al., 1988). This phenomenon may be related to their genetic organization. In the majority of the systems, the gene encoding the binding protein is located at the first position of the operon (Ames, 1986). The location of a binding protein gene at the 5' end of an operon may help to increase its expression relative to that of other genes further downstream. Limited processing of RNA polymerase together with exonucleolytic 3'-5' degradation of the mRNA may result in a gradient of expression, favoring the 5' genes in the respective operons (Newbury et al., 1987).

For some binding-protein-dependent transport systems, two distinct binding proteins with different substrate specificities interact with same inner membrane transport complex. The genes encoding the two alternative binding proteins are usually related. For example, the E. coli thiosulfate- and sulfate-binding proteins are highly similar, but encoded in different transcriptional units. They share the common membrane components (Hryniewicz et al., 1990).

Some periplasmic binding proteins have a dual function, as recognition sites of transport systems and as chemoreceptors. In their function as chemoreceptors they interact with signal transducers i.e., the integral membrane chemoreception proteins that transfer the chemotactic signal through the cytoplasmic membrane. For example, in addition to its role in transport, the periplasmic maltose-binding
protein (MBP) functions as the primary maltose chemoreceptor (Richarme, 1982). Generation of the chemotactic response to maltose requires a second protein, the chemotactic signal transducer Tar (Springer et al., 1977).

5. Metal ion uptake in bacteria

Studies in earlier days of bioenergetics suggested that the proton motive force and membrane potential were the driving forces of transport for metal ions. Selectivity of transport was a puzzling function of the walls of ion channels or carriers (Silver & Walderhaug, 1992). These simple models expanded with the discovery of ATP-dependent bacterial transporters. More and more studies have demonstrated the complexity of multiple transport systems (Silver & Walderhaug, 1992).

5.1. Iron

Bacteria have developed various strategies for acquiring iron while at the same time protecting themselves from iron's potential toxic effects. The major strategies used by bacteria to acquire iron include production and utilization of siderophores, and utilization of host iron compounds such as transferrin and lactoferrin (Mietzner & Morse, 1994).

Siderophores (Greek for iron carriers) are small molecules, generally less than 1,000 Da in mass, virtually ferric specific ligands that facilitate the solubilization and transport of Fe(III). Most siderophores fall into two chemical classes, phenolates (catechols) and hydroxamates (Crichton & Ward, 1992). The prototypical phenolate siderophore is enterochelin, a cyclic trimer of 2,3-dihydroxybenzoylserine that is
produced by many gram-negative pathogens. Aerobactin is prototypical of the hydroxamate class of siderophores. It is a conjugate of 6-(N-acetyl-N-hydroxyamino)-2-aminohexanoic acid and citric acid (Crosa, 1989). Despite the considerable structural variation found among siderophores, they all form six-coordinate octahedral complexes with Fe(III). Ferric siderophores are specifically recognized by their outer membrane receptors. After binding ferric siderophores, the outer membrane receptors are believed to undergo conformational change to internalize their ligands. This process depends absolutely on the TonB protein. TonB is anchored in the cytoplasmic membrane and likely spans the periplasm to contact outer membrane proteins involved in transport processes (Crosa, 1989). It is thought that TonB acts to transduce energy from the cytoplasmic membrane to specific energy-requiring processes in the outer membrane, resulting in the translocation of ligands bound by TonB-dependent outer membrane proteins into the periplasm. TonB-dependent outer membrane proteins share several conserved regions, including one near the amino terminus which likely interacts directly with the TonB proteins. The TonB-dependent systems include all of the ferric siderophore specific outer membrane proteins known for E. coli: FebA (enterobactin), IutA (aerobactin), FhuA (ferrichrome), FhuE (rhodoturulic acid), and FecA (ferric citrate) (Mietzner & Morse, 1994).
1990), *Neisseria gonorrhoeae* and *Neisseria meningitidis* (Mietzner *et al.*, 1987) possess a siderophore-independent mechanism for iron acquisition that involves the direct binding of iron-carrying serum glycoproteins (e.g., transferrin) to the bacterial cell. This binding of transferrin to the bacterial surface involves two transferrin binding proteins, Tbp1 and Tbp2. Tbp1 has been shown to possess similarity to TonB-dependent outer membrane proteins (Mietzner & Morse, 1994). Both Tbps of *H. influenzae* have been isolated and characterized (Gray-Owen *et al.*, 1995); Tbp1 of *H. influenzae* has a molecular weight of approximately 100 kDa, while Tbp2 is more variable in size but is generally between 75 and 85 kDa. The *tbp1* and *tbp2* genes appear to be organized into a single transcriptional unit. The loss of either protein correlated with a significantly decreased ability to bind transferrin and an inability to grow on media containing human transferrin as the sole iron source (Gray-Owen *et al.*, 1995).

Recently, a periplasmic binding protein-dependent iron transport system in *H. influenzae* has been identified (Sanders *et al.*, 1994; Adhikari *et al.*, 1995). This system is encoded by a *hitABC* operon. This operon can complement the siderophore-deficient *E. coli* strain to grow on medium containing d'pyridyl, a chemical iron chelator that sequesters free iron in the medium. HitA is an iron-regulated periplasmic iron binding protein, which is 37% and 69% identical to the periplasmic iron binding protein SfuA of *Serratia marcescens* and the ferric iron-binding protein Fbp of *Neisseria*, respectively. Purified HitA has the ability to compete for iron bound to dipyridyl both *in vitro* and within the periplasmic space of a siderophore-deficient strain of *E. coli* (Adhikari *et al.*, 1995). An isogenic mutant lacking a functional *hitC* gene could not utilize either free or protein-bound iron (Sanders *et al.*, 1994).
5.2. Copper

Copper is an essential metal trace element, which plays a vital role in the growth and physiology of aerobic organisms; however, excess of this metal results in cell death. Bacteria have evolved to possess effective means of achieving the fine balance between copper requirement and copper toxicity (Silver & Walderhaug, 1992).

Copper transport in *E. coli* is pictured as a combination of plasmid-encoded and chromosomally encoded transport activities. The chromosomally determined copper transport system includes the products of at least six genes, *cutA, cutB, cutC, cutD, cutE*, and *cutF* (Brown *et al.*, 1994). A mutation in one or more of these genes results in an increased copper sensitivity. Under normal physiological conditions, copper transport and equilibrium are mediated by CutA and CutB for uptake, followed by intracellular copper-binding proteins (CutE and CutF) and by the CutC and CutD protein for efflux. Under conditions of excess copper, plasmids which increase the resistance of gram-negative cells to copper are selected. The *E. coli* copper resistance system consists of an operon of four genes, *pcoA, pcoR, pcoB*, and *pcoC* (Silver & Walderhaug, 1992).

P-type ATPases are a large family of enzymes characterized by a phospho-aspartate intermediate in the ATP-driven cation transport cycle. They are found in prokaryotes as well as lower eukaryotes, plants and animals. Odermatt *et al.* cloned a copAB operon from the gram-positive bacterium *Enterococcus hirae* encoding two P-type ATPases of 727 and 745 amino acids, respectively (Odermatt *et al.*, 1993). Both enzymes display heavy metal ion binding motifs in their polar N-terminal region. Expression of the operon is regulated by the ambient copper concentration.
Disruption of the copA gene renders the cells dependent, whereas disruption of copB results in copper-sensitive phenotype. These data suggest that copB most likely serves in the extrusion of copper, while copA appears to be responsible for its uptake. Recently, the copper transport ATPase (ctaA) gene was cloned from the *Cyanobacterium synechococcus*. ctaA encodes a 790 amino acid polypeptide related to the CopA of *Enterococcus hirae*, to other known P-type ATPases, and to the candidate gene products for human diseases of copper metabolism, Menkes syndrome and Wilson disease (Phung et al., 1994). Menkes syndrome is primarily related to blockage of copper uptake across the intestinal mucosal lining. Wilson disease is related to inability to discharge copper from liver into blood circulation and bile. Disruption of the ctaA gene in *C. synechococcus* results in mutant cell with increased tolerance to copper compared with wild type. CtaA appears to be involved in copper transport into the cells (Phung et al., 1994).

5.3. Nickel

A nickel-specific high-affinity transporter has been analyzed in detail in *E. coli* (Navarro et al., 1993). This system is encoded by the nik locus, which contains five overlapping open reading frames that form a single transcription unit. The deduced amino acid sequence of the nik operon shows that its five gene products, NikA to NikE, are highly homologous to components of oligopeptide- and dipeptide-binding protein-dependent transport systems from several gram-negative and gram-positive species. NikA is a periplasmic binding protein, NikB and NikC are integral membrane components, and NikD and NikE contain ATP-binding sites. The nik operon is expressed under anaerobic growth conditions and provides nickel for the three hydrogenases involved in anaerobic metabolism. Insertion mutations in nik genes totally abolished the nickel-containing hydrogenase activity under nickel
limitation and markedly altered the rate of nickel transport (Navarro et al., 1993).

In *Alcaligenes eutrophus*, a gram-negative aquatic and soil bacterium, the transport systems for nickel appear to be different from that in *E. coli* (Eitinger & Friedrich, 1991). Nickel uptake occurs by a nonspecific magnesium transport system and a high-affinity nickel transporter, the product of gene hoxN. HoxN is a 33.1 kDa integral membrane protein with seven transmembrane helices that mediates energy-dependent uptake of nickel. Recently, a high-affinity nickel transporter (NixA) was identified from *Helicobacter pylori* (Mobley et al., 1995). NixA is a polypeptide of 34,317 Da that displays characteristics of an integral membrane protein. This protein has 40% identity to HoxN. NixA confers upon *E. coli* a high affinity nickel-transport system and is necessary for expression of catalytically active urease, which requires nickel ions for successful hydrolysis of urea (Bauerfeind et al., 1996).

5.4 Manganese

More than two decades ago, high-affinity and high-specificity manganese uptake were described for a number of bacteria (Silver & Jasper, 1977). Recently, Bartsevich and Pakraso isolated a cyanobacterial mutant strain impaired in the photosynthetic oxygen evolution process (Bartsevich & Pakraso, 1995). The growth rates of these mutant strains were significantly lower in a manganese-deficient medium and restored to near normal levels upon addition of micromolar concentrations of manganese, indicating the presence of a second transport system for manganese in this organism. The mutant strain could be complemented by a fragment of chromosomal DNA that contains three genes, *mntA*, *mntB* and *mntC*. In the *mntCAB* gene cluster, the *mntA* and the *mntB* genes are linked downstream
of the \textit{mntC} gene. Analysis of the deduced amino acid sequences of proteins encoded by these genes has indicated that these polypeptides are components of a binding protein-dependent ABC transporter complex. Interestingly, MntC is 30\% identical to the ScaA protein in \textit{S. gordonii} (Bartsevich \& Pakraso, 1995). The MntABC transporter was induced under manganese starvation conditions, whereas the second transporter systems was induced in the presence of micromolar levels of manganese (Bartsevich \& Pakraso, 1996).

6. Metal - Zinc

6.1. Biological functions

Zinc is the most widely used trace element in biology, although not the most available, since it is only the 27th most abundant element in the earth's crust (Vallee \& Falchuk, 1993). The earliest demonstration of a role for zinc in living systems occurred in 1869 with Raulin's discovery of essentiality of zinc for \textit{Aspergillus niger} (Raulin, 1869). Subsequently, the nutritional requirement of all forms of life for this metal was elucidated. Interestingly, the zinc requirement for bacterial growth was not demonstrated until 1947, well after its requirement for fungal (1869), algal (1900), plant (1914), and mammalian (1934) growth had been shown (Failla, 1978).

In prokaryotic and eukaryotic cells, zinc has been found to be a catalytic component of over 300 enzymes, including carbonic anhydrase, alkaline phosphatase, alcohol dehydrogenase, and several carboxypeptidases (Vallee \& Falchuk, 1993). Zinc also plays a critical structural role in enzymes and noncatalytic proteins. For example, several important motifs found in transcriptional regulatory
proteins are stabilized by zinc, including the zinc finger, zinc cluster, RING finger, and LIM domains (Schmiedeskamp & Klevit, 1994). It has been estimated that proteins containing these domains may constitute up to 1% of all human gene products. Furthermore, based on analysis of *Saccharomyces* genome, more than 2% of all yeast proteins, more than 150 of 6,000 total gene products, contain zinc-binding domains (Eide & Guerinot, 1997).

Through the interactions with biological macromolecules zinc plays a very important role in many cellular processes, such as gene expression, cell division and differentiation, programmed cell death, antioxidant defenses, and biomembrane functioning (Vallee & Falchuk, 1993). Due to its broad biological roles, zinc is considered a major element in assuring the correct function of an organism, from the very first embryonic stages to the last periods of life (Fabris & Mocchegiani, 1995). In addition, the significance of zinc in growth, development, and host defense have led to the universal inclusion of zinc in total parenteral nutrient solutions and infant formulas (Aggett & Comerford, 1995).

6.2. Efflux of zinc

Since excess zinc is toxic, both eukaryotes and prokaryotes have developed mechanisms to prevent overaccumulation of zinc. Proteins involved in zinc efflux have been reported in yeast and mammalian cells (Conklin *et al.*, 1992; Palmiter & Findley, 1995; Palmiter *et al.*, 1996a; 1996b). The plasmid-encoded CadA is a cadmium efflux ATPase found in Gram positive bacteria (Silver, 1996). This protein may transport zinc ions (Silver & Phung, 1996). Very recent studies have identified a P-type ATPase in *E. coli*, ZntA, which appears to play a role in zinc export (Board *et al.*, 1997; Rensing *et al.*, 1997b). Strains in which the zntA gene have been disrupted
exhibit increased sensitivity to zinc toxicity and increased cytosolic zinc concentration (Beard et al., 1997; Rensing et al., 1997b). P-type ATPase is a transmembrane protein that functions as efflux pump, which is different from bacterial binding-protein-dependent transport systems.

6.3. Uptake of zinc

In mammals a primary site of zinc regulation is the intestine. Absorption increases under conditions of zinc deficiency and decreases when zinc is in excess. After being absorbed, zinc is bound to albumin in the circulation, where zinc is maintained within a narrow range, typically about 1 μg/ml. Most of zinc is taken up by the liver before being redistributed to other organs. The zinc content in various organs ranges from 10 to 100 μg/g, with brain being lowest and bone highest; these amounts are resistant to large variations in dietary zinc (Cousins, 1985). The primary routes of zinc excretion are via pancreatic, biliary and intestinal secretions (Cunnane, 1988).

At the cellular level, zinc interacts with extracellular binding sites, which are likely to include binding sites involved in the subsequent translocation of this ion to the cell interior. Inside the cell, zinc binds to cytosolic and organelle binding sites or is taken up by intracellular organelles. However, very little is known about the molecular mechanisms cells use to obtain zinc (Vallee & Falchuk, 1993).

The yeast Saccharomyces cerevisiae is a useful model system for understanding metal ion uptake in a eukaryotic cell. Biochemical assays of zinc uptake in yeast indicated that this process was transporter-mediated, i.e., uptake was dependent on time, temperature, and concentration and required metabolic energy.
S. cerevisiae has at least two separate systems for zinc uptake (Zhou & Eide, 1996b). One system has high affinity for substrate and is induced in zinc-deficient cells. The second system has lower affinity and is not highly regulated by zinc status. A gene, zrt1 (zinc-regulated transporter), has been identified because of its significant similarity to irt1, an Fe(II) transporter gene from the plant Arabidopsis thaliana (Eide et al., 1996). The zrt1 gene encodes the zinc transporter protein of the high-affinity uptake system called Zrtlp. The levels of zrt1 expression correlated with activity of the high affinity system; overexpression of zrt1 increased high affinity uptake, whereas a zrt1 mutation eliminated high affinity activity and resulted in poor growth of the mutant on zinc-limited media (Zhao & Eide, 1996a). zrt1 is a member of a closely related family of transporter genes found in organisms as diverse as fungi, plants, nematodes, and humans. Like other members of this family, Zrtlp is predicted to be an integral membrane protein containing eight potential transmembrane domains. The zrt2 gene encodes the low affinity zinc transporter in S. cerevisiae. The amino acid sequence of Zrt2p is remarkably similar to those of Zrt1p and Irt1p. Overexpression of zrt2 increased low affinity uptake, whereas disrupting this gene eliminated that activity, but had little effect on the high affinity system. Because the zrt1zrt2 double mutant is viable, an additional zinc uptake pathway may exist in S. cerevisiae (Zhao & Eide, 1996b).

In contrast to yeast and other eukaryotic cells, there are very few studies concerning zinc uptake in bacteria. There are several reasons for limited study in this field (Failla, 1977; Hughes & Pode, 1989). First, extremely low zinc concentrations are required for optimal bacterial growth. Chemically defined culture media usually contains 0.5-1 μM zinc as an impurity. Generally bacteria achieve optimal cell yields in such media without zinc supplementation. For unknown reasons, the zinc requirement for bacteria seems to be tenfold lower than that for
algae or fungi. Second, elimination of zinc from media using solvents or alumina has generally been unsuccessful. The extreme difficulty in preparing a zinc-deficient medium has frustrated studies in this field. Finally, the high electrostatic affinity of zinc for anionic sites on the microbial surface is not readily distinguishable from zinc transport. The removal of nonspecifically bound zinc from the cell surface without damaging the membrane is one of the principal problems encountered in determining the quantity of zinc transported into a cell (Failla, 1977).

The few studies reported suggest that bacteria appear to possess a specific energy-dependent zinc transport system, in spite of the existence of problems already mentioned. Bucheder and Broda reported energy-dependent uptake of $^{65}$Zn by E. coli. This process was energy dependent as evidenced by a requirement for exogenous glucose, air, and by temperature dependence (Bucheder & Broda, 1974). Kung et al. measured the zinc content of synchronously dividing E. coli cells and found that cellular zinc increased in a step-like fashion, whereas cellular potassium, calcium, and magnesium content increased smoothly during the cell cycle (Kung et al., 1976). In addition, zinc is accumulated by Bacillus megaterium during exponential growth (Lee & Weinberg, 1971). Since studies concerned with zinc uptake have been confounded by both the nonspecific binding of zinc to the bacterial surface and the rapid exchange of cellular zinc with zinc in the medium, so far a specific zinc transport mechanism has not been properly demonstrated.

7. Rational and objectives

Recently, our group purified a putative PE-binding adhesin from H. influenzae. This protein has 24.5% identity and 48.3% similarity to FimA of S. parasanguis, a member of streptococcal adhesin family. The purified adhesin
preparation was found to be a potent inhibitor of \textit{H. influenzae} PE binding as monitored by TLC overlay. Furthermore, antisera raised against this preparation prevented the attachment of \textit{H. influenzae} to PE and Hep2 cells \textit{in vitro} (Busse, 1997). These preliminary results strongly encouraged us to further characterize this putative PE-binding adhesin. Thus, the original objectives in this project were:

1. To clone and express the putative PE-binding adhesin from the NTHI strain.
2. To characterize the PE-binding site of this putative adhesin.

Subsequent studies clearly demonstrated that the putative PE-binding adhesin was not an adhesin and is unable to bind PE \textit{in vitro}; it is a celite binding protein with unknown function. Although the function of this protein is probably not involved in attachment of \textit{H. influenzae} to PE, the study progress made at that time gave us a chance to identify this protein's function. A pilot project was therefore initiated with the following objectives:

1. To characterize the celite binding ability of this protein;
2. To localize the celite binding protein in \textit{H. influenzae};
3. To study the conservation of this protein in \textit{H. influenzae} clinical strains;
4. To determine the nature and amount of metals bound to the celite binding protein, if it is a metal binding protein;
5. To make an isogenic mutant of the HI0119 gene encoding the celite binding protein to clarify its function;
6. To identify the real PE-binding adhesin in \textit{H. influenzae}. 

-54-
MATERIALS AND METHODS

1. Materials

Clinical strains of *H. influenzae* were kindly provided by the Dept. of Microbiology, Hospital for Sick Children (HSC). Taq polymerase, dNTP, restriction enzymes and buffers except *DraIII*, DNA ligase, T4 DNA polymerase, RNase, proteinase K, vector pGEX-2T, and glutathione Sepharose 4B were from Pharmacia. PFU polymerase, restriction enzyme *DraIII*, and *E. coli* strain JM109 were from Stratagene. TA cloning kit, pTrc plasmids, Ni-NTA agarose, *E. coli* strain Top10 and INVα were from Invitrogen. DNA standards were from GibcoBRL. Goat anti-rabbit horseradish peroxidase conjugate, nitrocellulose membrane, PVDF membrane, β-mercaptoethanol, and protein standard were from BioRad. The celite 545 and SpectroPor dialysis tubing (MWCO 12,000-14,000 kDa) were from Fisher. The dialysis tubing was washed extensively with ddH2O before use. Fluorescein isothiocyanate (FITC), 4-chloro-1-naphthol, orcinol-H2SO4 reagent, CaCl2, MgCl2, CuCl3, NiCl2, Cd(NO3)2, MnCl2 FeCl3, ZnCl2, Zn(NO3)2, EDTA, kanamycin, ampicillin, X-gal, IPTG, Triton X-114, phosphatidylethanolamine (PE) from *E. coli*, PE from soybean, phosphatidylcholine (PC), and bovine serum albumin (BSA) were from Sigma. Gangliotriaosylceramide (Gg3) was purified from guinea pig erythrocytes. Gg4 was prepared by acetic acid hydrolysis of bovine brain ganglioside GM1 (Lingwood & Nutikka, 1994). Globotetraosylceramide (Gb4) was prepared from human kidney tissue as described previously (Boyd & Lingwood, 1989). BCA protein assay kit was from Pierce. 65ZnCl was from Amersham. New Zealand white rabbits were ordered from the animal facility, HSC. BBL GasPak Plus generator was from Baxter Healthcare. Plastic-backed silica gel (SILG) TLC plates were from Macherey-Nagel. Polyclonal antibody against the PE-binding preparation of *H. influenzae* was...
prepared as described previously (Busse et al., 1997). Polyclonal rabbit anti-OmpP2 was a generous gift from Dr. R.S. Munson (Washington University, St. Louis).

2. Methods

2.1. Bacterial culture

H. influenzae clinical strains were obtained immediately following isolation and stocks stored at -70°C in glycerol-citrate. The strains were plated from frozen stocks onto chocolate agar plates, and grown overnight at 37°C under aerobic conditions or anaerobic conditions using a BBL GasPak Plus generator with catalyst. For liquid cultures, H. influenzae strains were grown in 20% Levinthal broth with aeration by shaking at 180 rpm in a 37°C incubator. Each strain was subcultured 2 times before the assay. E. coli was grown in Luria-Bertani (LB) or BHI broth or agar, aerobically at 37°C and supplemented with 100 µg/ml of ampicillin or with 20 µg/ml kanamycin where appropriate.

2.2. Extraction of chromosomal DNA from H. influenzae

Large molecular weight chromosomal DNA was prepared from liquid cultures. This procedure was modified from that of Moxon et al. (Moxon et al., 1984). The bacterial pellet was lysed in 0.1 M NaCl, 10 mM Tris-HCl, and 10 mM EDTA (pH 8.0) containing 1% SDS and heated to 60°C for 10 min. The lysate was digested at 55°C for 60 min with addition of Proteinase K to 1 mg/ml. The mixture was extracted twice with an equal volume of phenol followed by chloroform before precipitation with cold ethanol. The DNA was resuspended in 10 mM Tris, 1 mM EDTA, pH 8.0, containing RNase at 40 µg/ml.
2.3. Amplification of the putative PE-binding adhesin gene

PCRs were performed in 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 2 mM MgCl₂, 200 μM all four dNTPs, 200 nM each primer, and 1 U of PFU polymerase using a Minicycler™ (Fisher). After a preamplification denaturation step of 94°C, 5 min, reactions were incubated for 30 cycles of 1 min at 94°C, 1 min at 55°C, 3 min at 72°C and a final extension step of 72°C for 10 min. Taq polymerase was added to the amplified reaction at the final extension step. The primers used to amplify the putative PE-binding adhesin (HI0119) gene were HIMA1, 5'CGG ATC CGT ATA GCA TAG TTA AAA CC 3' and HIMA2, 5'GAA TTC TTA TTT AGC TAA ACA TTC CAT GTA GC 3'. Synthetic BamHI and EcoRI restriction enzyme sites (underlined in primers) were engineered into the upstream and downstream primers, respectively.

The amplified 950 bp coding region of the HI0119 gene was ligated into a TA cloning vector pCRITM. Competent cells of INVα were transformed with DNA from the ligation reaction. Transformed colonies were selected on BHI agar plates which contained 100 μg/ml ampicillin and 40 μg/ml of X-gal. The resulting plasmid was designated as pTA119. The insert of pTA119 was sequenced at the University of Toronto Automated Sequencing Facility, Toronto, on a DNA sequencer using a SequiTherm™Long-Read™ Cycle Sequencing kit.

2.4. Expression and purification of the putative PE-binding adhesin as a GST fusion protein

The plasmid pTA119 was digested with the restriction enzymes BamHI and
EcoRI and ligated into the complementary restriction enzyme sites of the expression vector pGEX-2T. The resulting construct was designated pGHA1. Recombinant plasmid pGHA1 was transformed into competent cells of E. coli strain TOP10. The construct pGHA1 encodes a fusion protein containing glutathione S-transferase (GST) at the N terminus and the putative PE-binding adhesin at its C terminus and separated by a thrombin cleavage site.

E. coli strain TOP10 transformed with the recombinant plasmid pGHA1 was grown overnight in 2 ml of BHI or LB broth containing 100 μg/ml ampicillin. The aliquots of the overnight cultures were diluted 1:100 into 100 ml BHI or LB broth. When growth of the cultures reached an OD600 of about 0.5, expression of GST-0119 was induced by adding IPTG to a final concentration of 0.1 mM. Cells were grown in the presence of IPTG for 5 hours at 30°C. Cells were then harvested by centrifugation at 7,700 rpm for 10 min and taken up in 10 ml of PBS. Bacteria were disrupted by sonication on ice, and cellular debris pelleted by centrifugation at 12,000 rpm for 20 min at 4°C. The fusion protein in the supernatant was affinity purified with 100 μl of a 50% slurry of glutathione-Sepharose 4B in PBS. The mixture was incubated with gentle agitation at room temperature for 30 min. Then the glutathione beads were washed four times with PBS and collected by centrifugation at 700 rpm for 5 min after each washing step. The fusion proteins were eluted by incubating the beads for 10 mM reduced glutathione in 50 mM Tris-HCl (pH 8.0). The elution was repeated twice.

2.5. Expression and purification of the putative PE-binding adhesin as a 6XHis tagged fusion protein

The plasmid pTA119 was digested with BamHI and EcoRI. The resulting
fragment was subcloned into the expression vector pTrcHisA. The resulting plasmid was designated pTHA1. The plasmid pTHA1 encodes the putative PE-binding protein preceded at the N-terminus by six consecutive histidine residues and a 35 amino acid spacer.

E. coli strain TOP10 transformed with the plasmid pTHA1 was grown overnight in 2 ml of BHI or LB broth containing 100 µg/ml ampicillin. The aliquots of the overnight cultures were diluted 1:100 into 100 ml BHI or LB broth. Cells were grown at 37°C to an OD600 of about 0.5 and induced for 3-5 hours with 1 mM IPTG to express high levels of fusion protein. Cells were harvested by centrifugation at 7,700 rpm for 10 min. When the fusion protein was purified under native conditions, the bacterial pellets were suspended in 5 ml sonication buffer (50 mM Na-phosphate pH 8.0, 300 mM NaCl). The cell suspensions were quickly frozen in dry ice/methanol, and thawed in a 37°C water bath. This procedure was repeated three times. Bacteria were further disrupted by sonication. Cellular debris were removed by centrifugation at 12,000 rpm for 20 min. The fusion protein in the supernatant was affinity purified with 1 ml of a 50% slurry of Ni-NTA agarose in sonication buffer. The mixture was rocked for 1 hour at 4°C. After incubation, the matrix with adsorbed fusion protein was packed into a column, and washed with about 30 ml sonication buffer followed by 30 ml wash buffer (50 mM Na-phosphate, 300 mM NaCl, 10% glycerol, pH 6.0). The proteins were eluted with a 10 ml gradient of 0-0.5 M imidazole in wash buffer. 1 ml fractions were collected and analyzed on SDS-PAGE and Western blotting. When the fusion protein was purified under denaturing conditions, the pellets were suspended in 5 ml 6 M guanidine-HCl solution. The mixture was stirred for 1 hour at room temperature. The lysate was centrifuged at 12,000 rpm for 20 min. The supernatant was incubated with 1 ml of a 50% slurry of Ni-NTA agarose in 6 M guanidine HCl for 1 hour at room
temperature with gentle shaking. The Ni-NTA beads were washed by 8 M urea solutions with pH 8.0, 6.3, and 5.3, respectively. After final washing, the fusion proteins were eluted with 6 M guanidine-HCl solution (pH 4.0). The eluted proteins were dialyzed against PBS and analyzed by SDS-PAGE and Western blotting.

2.6. Production of anti-GST-0119 and anti-His-0119 antisera

The purified GST-0119 or His-0119 were run on separate SDS-PAGE gels and visualized by briefly staining with 0.05% Coomassie blue in water. The protein bands, containing approximately 200 μg of protein, were excised from the gel, crushed, and emulsified with Freund's complete adjuvant and injected subcutaneously into New Zealand white rabbits. Four weeks after the initial immunization, the rabbits were boosted with the appropriate fusion protein, prepared as described above, emulsified with Freund's incomplete adjuvant. The presence of specific antibodies in both antisera was confirmed two weeks post-boost using Western blotting.

2.7. TLC overlay

All lipids (about 1 μg) were separated by TLC with chloroform-methanol-0.88% aqueous KCl 60:40:9 (v/v/v). The plates were air dried and blocked by incubation with 3% gelatin in water for 2 hours at 37°C. The gelatin was poured off and the plates were washed with 50 mM Tris-buffered saline (TBS). Then recombinant fusion protein (about 10 μg/ml) in 50 mM TBS, pH 7.4, was incubated with the plate overnight at 4°C. The plates were washed and incubated with polyclonal anti-His-119 antisera diluted 1/2000 in 50 mM TBS for 3 hours at room temperature. Following washing, a 1/2000 dilution of goat anti-rabbit horse radish
peroxidase conjugate in 50 mM TBS was added and incubated for 2 hours at room temperature. The substrate was prepared by combining 2 ml of a 3 mg/ml solution of 4-chloro-1-naphthol methanol, 10 ml of 50 mM TBS and 5 µl of 30% hydrogen peroxide. After the color (indicative of binding) developed, usually in about 5 to 15 minutes, the reaction was stopped by extensive washing with ddH₂O (Lingwood et al., 1992b).

2.8. Preparation of the PE affinity matrix

A PE affinity matrix was prepared as previously described (Boulanger et al., 1994). Briefly, 50 mg of PE from E. coli in 50 ml of methanol was vortexed with 20 g of activated Celite 545. The methanol was evaporated using rotary evaporation. The PE matrix was air dried in the fume hood overnight, suspended in PBS, and added to a 50 ml glass column. The column was subsequently blocked by running through 100 ml of 1% glycine at room temperature. The gel was washed with PBS at 4°C prior to use.

2.9. Celite or PE affinity chromatography

Five g celite 545 was suspended in PBS buffer to form a column, and washed with 100 ml 1% glycine at room temperature. The column was washed with 200 ml of PBS before use. A 20 ml sonic lysate of bacteria was applied to the celite column or PE affinity column at a flow rate of 0.5 ml/min, and the column was washed with 1 L of PBS, followed by 1 L 2 M NaCl to remove weakly bound proteins. Bound protein was eluted with 10 mM EDTA and/or 1 M Tris (pH 11.2). Since EDTA elution was slow, the column was incubated in EDTA elution buffer for about 30 min before collecting. The eluate was dialyzed against 10 mM Tris-HCl (pH 7.5) at
2.10. FITC labeling of GST-0119 and His-0119

FITC was added directly to GST-0119 (in a 1:1 (w/w) ratio) in 0.5 M Na₂CO₃/NaHCO₃ conjugating buffer pH 9.5, and the mixture was gently rotated for 1-2 hours at room temperature in the dark. Free FITC was removed by extensive dialysis against PBS (Khine & Lingwood, 1994).

2.11. Celite binding of FITC conjugated GST-0119

10 mg Celite 545 was washed with 1 ml PBS two times and resuspended in 500 μl PBS containing 0.5% BSA. FITC conjugated GST-0119 (0.1 μg) was added and incubated at room temperature for 20-30 min. For inhibition experiments, 2 μg of unlabeled GST-0119, GST or GST-0119SII were mixed with FITC-GST-0119 before addition to the celite suspension. Unbound GST-0119 was removed by washing the celite at least 5 times with 1 ml of PBS. The preparation were examined and photographed under incident uv illumination using a Polyvar fluorescence microscope. FITC-BSA was used as a control for non-specific binding.

2.12. Immunoprecipitation

*H. influenzae* NTHI5564 was grown aerobically on chocolate agar plates overnight in a 37°C incubator. 5 plates were harvested into 20 ml PBS, centrifuged at 7,700 rpm for 10 min at 4°C to pellet the cells. The bacterial pellets were suspended in 5 ml PBS containing 1 mM PMSF to inhibit proteolytic enzymes. The cell suspension was disrupted by sonication on ice, and followed by the addition of
Triton X-100 to a final concentration of 1%. After shaking gently at room temperature for 1 hour, the lysate was centrifuged at 12,000 rpm for 20 min at 4°C. The antisera against His-0119 were added into the supernatant at 1:50 dilution, and the tube was incubated for 3 hours at 4°C. Then 200 μl of 10% suspension of Protein A-Sepharose were added to the tube and incubated with shaking for 2 hours at 4°C. The pellets were spun at 1,000 rpm for 5 min and washed five times with cold PBS with 1 mM PMSF, 0.5% Triton X-100. After final washing, sample was subjected to SDS-PAGE and immunoblot analysis.

2.13. Electrophoresis and Western blotting

SDS-PAGE and immunoblotting were carried out as described (Laemmli, 1970; Towbin, 1979). For Western blots, antiserum raised against His-0119 was used at a 1:2000 dilution. Nonreducing and “native” gel electrophoresis were carried out as described (Loh, 1994). Briefly, in nonreducing gels, the protein sample buffer did not contain β-mercaptoethanol, thereby preserving the disulfide linkages. In “native” gels, in addition to the omission of β-mercaptoethanol, the SDS content of the protein sample buffer was reduced to 0.1%, and the protein samples were not boiled prior to loading.

2.14. Construction of mutants

i. The cysteine mutant

The primer used to mutate cysteine 308 was H1cm, 5’ GAA TTC TTA TTT AGC TAA AGA TTC CAT GTA GC 3’, in which the codon for Cys308 was altered to encode serine (nucleotide change indicated in bold, EcoRI site is underlined). An
amplification was performed with primers HIMA1 and H1cm. The reaction product was digested with EcoRI and BamHI and ligated into the plasmid pGEX-2T for expression as GST fusion protein in *E. coli*. The mutation was confirmed by DNA sequencing.

ii. The C-terminal truncation mutant

Since an unique DraIII restriction site was located at 701 bp in the C-terminus of the NTHI0119 gene, restriction enzymes DraIII and EcoRI were used to delete the C-terminal 235 bp region from plasmid pTHA1, followed by blunt-ending with T4 DNA polymerase and ligation (Sambrook et al., 1989). The resulting plasmid was denoted as pTHAcm. The truncated mutant (His-0119cm) was expressed and purified using the same method as for His-0119.

2.15. Subcellular fractions of *H. influenzae*

i. Immunofluorescence

*NTHI6564* plate or broth cultures were suspended in PBS and incubated in antisera against His-0119 diluted in 1% BSA for one to four hours at 4°C. Bacteria were washed, and bound antibody was visualized with fluorescein-conjugated antirabbit antibodies under incident uv illumination.

ii. Preparation of extracellular proteins

A 100 ml overnight growth of the broth culture of *NTHI6564* was centrifuged at 9,000 rpm for 20 min. Then the supernatant was concentrated 100
fold in an Amicon concentrator, analyzed by SDS-PAGE and Western blotting.

iii. Triton X-114 extraction of integral membrane proteins

Integral membrane proteins of NTHI6564 were extracted with Triton X-114 essentially as described by Swancutt et al. (Swancutt et al., 1989). Briefly, bacterial cultures were grown to an OD600 of 0.5 (approximately 3 x 10^8 CFU/ml), harvested by centrifugation at 4°C and washed using 1 volume of 200 mM Tris-HCl (pH 8.0). The pellets were suspended in 20 mM Tris-HCl (pH 8.0), 10 mM EDTA, 2% Triton X-114. After incubation for 4 hours at 4°C, cellular debris was removed by centrifugation. The supernatants were warmed to 37°C to allow phase separation. After centrifugation at 12,000 rpm for 10 min, the aqueous phase was separated from the detergent phase. The detergent phases were washed three times using 20 mM Tris-HCl (pH 8.0), 10 mM EDTA.

IV. Osmotic shock extraction of periplasmic proteins

Proteins in the periplasmic space were obtained by the osmotic shock procedure described by Ames (Ames, 1994). Briefly, bacterial cultures were grown at 37°C overnight and harvested by centrifugation. The bacterial pellet was gently resuspended in 30 mM Tris-HCl (pH 8.0), 20% sucrose. EDTA was added to 1 mM and the bacteria were incubated at room temperature for 5-10 min with shaking. The cells were recovered by centrifugation and then resuspended in ice-cold 5 mM MgSO4 with shaking for 10 min at 4°C. The supernatant, containing periplasmic proteins, was collected as osmotic shock fluid.
2.16. Expression and purification of the putative PE-binding adhesin in *E. coli*

PCR amplification of the HI0119 gene and its upstream 590 bp region was performed as described above. The primers used in amplification were PZU500, 5' GAC TAC GTC ATT GAT GC 3' and HIM&., Fj'GAA TTT AGC TAA ACA TTC CAT GTA GC 3'. The 1.6 kb amplified product was ligated into a TA cloning vector, pCRII™. The resulting plasmid was designated as pTAp119. The orientation of the insert was determined by restriction enzyme *BamHI* digestion; 2 clones, containing either orientation, were checked to assess whether the native HI0119 promoter has function in *E. coli*.

To purify the native protein, the high resolution technique of chromatofocusing, which fractionates proteins based on pi, was used. The *E. coli* transformant containing the plasmid pTAp119 was grown overnight in 1 L of BHI medium with ampicillin. The cells were harvested by centrifugation and the periplasmic proteins were dialysed overnight against 25 mM imidazole-HCl buffer, pH 7.4, and was applied to a column of Polybuffer exchanger 94 (Pharmacia)(1.5 cm X 30 cm) equilibrated with the same buffer. Elution was carried out with 10 column volumes of a degassed solution of Polybuffer 74 (Pharmacia) diluted 1:8 with distilled water and adjusted to pH 4.0 with HCl. Fractions were monitored for absorbance at 280 nm, and for pH, and the recombinant protein positive fractions (identified by Western blotting) were pooled and lyophilized. To remove ampholytes, the lyophilized material was dissolved in 1 ml of water and applied to a G-75 fine column (1.5 cm x 80 cm) equilibrated with 50 mM Tris-HCl, pH 7.2. The column was run at 0.5 ml/min and fractions were monitored at 254 nm to detect the protein and ampholyte peaks.
A standardized stock solution of recombinant protein of known concentration was prepared as follows: purified protein was dialysed extensively against water, and the protein concentration was determined by amino acid analysis in triplicate. The absorbance at 280 nm was measured for dilutions of stock solution prepared in 20 mM Tris-HCl, pH 7.1, 20 mM NaCl, and from these values the extinction coefficient of recombinant protein was calculated to be 36,800 L mol⁻¹ cm⁻¹.

2.17. Metal composition of purified celite binding protein

i. Neutron activation analysis

Purified celite binding protein with a concentration of 0.56 mg/ml was heat sealed in a clean polyethylene vial. This was irradiated for 5 min at a neutron flux of 1.0x10¹² n cm⁻² s⁻¹ in the SLOWPOKE reactor of the University of Toronto. The irradiated solution was transferred to a clean vial, and after a delay time of 2.5 hours, to allow C¹³ to decay sufficiently, its radioactivity was assayed for 15 min with a hyperpure germanium detector-based gamma-ray spectrometer. The Mn content was established using the comparator method (Hancock, 1978). To determine the Cr, Ni, Sc, Fe, Zn, and Co contents of the sample, it was irradiated for 16 hours at a neutron flux of 2.5x10¹¹ n cm⁻² s⁻¹. After a delay time of 6 days to allow ²⁴Na to decay, the radioactivity in the sample was counted for 24 hours and the elemental concentration determined (Hancock, 1978).

ii. Atomic absorption analysis.

Aliquots of purified celite binding protein were incubated in the presence of 1
mM Zn(NO3)2 or 5 mM EDTA at 4°C overnight. The samples were then dialyzed extensively against 20 mM Tris-HCl (pH 7.1), 20 mM NaCl at 4°C for four days. Zinc content was determined by atomic absorption spectroscopy using a VARIAN SpectrAA 10 spectrometer in the Department of Clinical Biochemistry, HSC.

iii. Zinc blot

Zinc blot was carried out by the method of Schiff et al. (Schiff et al., 1988). Briefly, after transfer of proteins to nitrocellulose, the filters were washed in metal-binding buffer (100 mM Tris-HCl, pH 8.0, 50 mM NaCl) for 2 hours. The nitrocellulose was probed for 1 hr with 30 μCi of 65ZnCl2 in 20 ml of metal-binding buffer. The filter was washed with metal-binding buffer for 30 min with three changes of buffer. Nitrocellulose filters were wrapped in Saran Wrap and exposed -80°C to Kodak X-OMAT film.

2.18. Circular dichroism (CD) measurements

CD spectra were recorded between 190 and 250 nm on a Jasco-720A spectropolarimeter using a 1 mm quartz cell at 25°C. The concentration of Pzp1 was 0.56 mg/ml in 20 mM Tris-HCl (pH 7.1), 20 mM NaCl.

2.19. Construction of the pzp1 isogenic mutant

A 3.5 kb DNA fragment from the NTHi5564 chromosome, including the pzp1 gene and its 1.2 kb upstream and 1.3 kb downstream region, was amplified by PCR as described above. The primers used for PCR were HZU5, 5' CGG ATC CCT CTT GTA GCA ATG GCT TCA GTG 3' and HZD3, 5' GAA TTC CAT TGG GAT GTT GGT CTC -68-
The amplified product was cloned into vector pCRIP, generating plasmid pTA3.5. The 2.5 kb BamHI-EcoRI fragment from pTA3.5 was subcloned into vector pTrcA. The resulting plasmid, pTrc2.5, was digested with PstI to remove a 822 bp internal region of the pzp1 gene, and ligated to a 1.2 kb kanamycin resistant cassette from pUC4K. The resulting plasmid was designated as Ypzp::kan.

The plasmid Ypzp::kan was digested to completion with BamHI and EcoRI, and this digestion mixture was used to transform NTH16564 cells made competent for transformation with M-IV medium as described previously (Barcak, 1991). The transformants were selected on chocolate agar plates containing 20 μg of kanamycin per ml under either aerobic or anaerobic conditions.

Transformant colonies were screened by direct amplification of chromosomal DNA from single colonies by PCR. Oligonucleotide primers used in the screening PCR were primers P1, 5' GTA TAG CAT CAG TAA AAC C 3' and P2, 5' TTA TTT AGC TAA ACA TTC CAT GTA GC 3'.

2.20. Growth curves

Single colonies from NTH16564 wild-type strain and the pzp1 mutant were grown on chocolate agar plates under anaerobic conditions overnight. The bacteria from chocolate agar plates were suspended in 1 ml 20% Levinthal broth at an optical density at 600 nm of 0.7. 200 μl of this bacteria suspension was added into 50 ml 20% Levinthal broth with or without zinc supplement and aerobically grown at 37°C with shaking at 180 rpm. The OD600nm was determined at various points along the growth curve.
2.21. Expression and purification of OmpP2 fusion protein GST-P2

The primers used to amplify \textit{ompP2} gene were HIP2A, 5'C\textsc{CGG} ATC CGC TGT TGT TTA TAA CAA C 3' and HIP2B, 5'C\textsc{CTG} CGA GTT AGA AGT AAA CGC GTA AAC 3'. Synthetic \textit{Bam}HI and \textit{XhoI} restriction enzyme sites (underlined in primers) were engineered into the upstream and downstream primers, respectively. PCR amplification of \textit{ompP2} was carried out as described above. The amplified product was ligated into a TA cloning vector, pCR\textsc{TM}. The resulting plasmid was designated as pTAompP2. The plasmid pTAompP2 was digested with the restriction enzymes \textit{Bam}HI and \textit{XhoI} and ligated into the complementary restriction enzyme sites of the expression vector pGEX-2T. The resulting construct was designated pGST-P2. Recombinant plasmid pGST-P2 was transformed into competent cells of \textit{E. coli} strain TOP10.

\textit{E. coli} strain TOP10 transformed with pGST-P2 was grown overnight in 2 ml of BHI or LB broth. The aliquots of the overnight cultures were diluted 1:100 into 100 ml BHI or LB broth. When growth of the cultures reached an OD600 of about 0.5, expression of GST-P2 was induced by adding IPTG to a final concentration of 0.1 mM. Cells were grown in the presence of IPTG for 5 hours at 25\textdegree C. Cells were then harvested by centrifugation at 7,700 rpm for 10 min and taken up in 10 ml PBS. Bacteria were disrupted by sonication, and followed by the addition of Triton X-100 to a final concentration of 1%. After shaking gently in room temperature for 1 hour, the lysate was centrifuged at 12,000 rpm for 20 min at 4\textdegree C. GST-P2 fusion proteins were purified by glutathione-Sepharose 4B. The fusion proteins were eluted by 20 mM reduced glutathione in 50 mM Tris-HCl (pH 8.0). The elution was repeated twice.
RESULTS

1. Expression and characterization of the putative PE-binding adhesin of *H. influenzae*

1.1. Cloning and sequence of the putative PE-binding adhesin

The experimentally determined N-terminal 35 amino acid sequence of the putative PE-binding adhesin (Busse, 1997) starts at residue 25, indicating the first 24 amino acids form a signal sequence which closely resembles typical signal peptides, in which 3 out of the first 6 amino terminal residues are positively charged, followed by a hydrophobic domain (5' ISAISAALLSAPMMANA 3') and end at a signal peptide cleavage site (between Asp24 and Val25) (Heijne, 1990). The DNA sequence corresponding to the mature protein was obtained from the genomic DNA of nontypeable *H. influenzae* strain 6564 by PCR (Fig. 3). An approximate 950 bp DNA fragment was amplified (Fig. 3). The amplified product was ligated into the TA cloning vector, generating pTA119. The insert could be cleaved with the restriction enzymes BamHI and EcoRI, whose recognition sites were synthetically engineered into the upstream and downstream primers (Fig. 3). Sequence analysis revealed a coding region of 936 bp encoding a 312 amino acid protein with molecular mass of 34,863 (Fig. 4) and predicted pI of 6.74. The nucleotide and protein sequences have 98.7% and 97.7% identity, respectively, to the sequence of *H. influenzae* strain Rd for which the genomic DNA and predicted protein sequences have been determined (Fleischmann, et al., 1995). The preliminary classification of all *H. influenzae* Rd genes indicated that the nearest relative of HI0119 is adhesin B (FimA) of *Streptococcus parasanguis* (Fleischmann et al., 1995). The homology search showed that the putative PE-binding adhesin (HI0119N) from NTHI strain 6564 has 23.7%
identity and 47.8% similarity to the adhesin B (FimA) of *S. parasanguis* (Fig. 5). FimA of *S. parasanguis* is part of a larger family of related streptococcal adhesins, recently called lipoprotein antigen I family (jenkinson, 1994). However, the putative PE-binding adhesin is distinguished from the FimA family by the absence of the N-terminal lipid anchor consensus sequence LXXC, and the presence of a central histidine-rich domain (residues 90-137) (Fig. 4). The hydrophathy plot indicated a central hydrophilic region, corresponding to the histidine-rich domain but no obvious membrane spanning domain (Fig. 6). In the hydrophilic histidine-rich domain every other residue is histidine alternating with either aspartate, lysine or glutamate (Fig. 6A). Secondary structure predictions revealed no unusual features (Fig. 6B). The G + C content of the HI0119 gene is 36.2%, corresponding to the overall value of 38% for the *H. influenzae* genome (Fleischmann et al., 1995).

In a recent BLAST search we found that the putative PE-binding adhesin from NTHI6564 has 49.2% identity and 59.4% similarity to an unidentified protein (YebL) in *E. coli*, a 31.1 kDa protein in the mabB-ruvB intergenic region precursor (Fig. 5). Particularly, YebL also contained a histidine-rich region located at the similar position to that in HI0119 (Fig. 5). The high homology between the HI0119 protein and *E. coli* YebL strongly suggests that both proteins may have similar function, although the functions of both proteins remain to be established.

In addition, the putative PE-binding adhesin and YebL of *E. coli* contains two cysteine residues near the carboxyl terminus (Fig. 5), which are absent in FimA of *S. parasanguis*. 
Figure 3. Cloning of the DNA sequence corresponding the mature putative PE-binding adhesin (HI0119). (A) A schematic diagram of the cloning strategy. The DNA fragment of the mature HI0119 protein was obtained from the genomic DNA of NTHI6564 by PCR. The amplified product was ligated into the TA cloning vector, generating pTA119. (B) The amplified PCR product. (C) BamHI/EcoRI restriction enzyme digest of plasmid pTA119. The DNA fragments were resolved in 0.7% agarose gel. Lane M: 1 kb DNA marker. Sizes of DNA marker are indicated on left side.
Figure 4. Nucleotide and deduced amino acid sequences corresponding to the mature putative PE-binding adhesin. N-terminal amino acid sequence determined from purified adhesin is underlined. The histidine rich region is boxed.
Figure 5. Alignment of HI0119 and other bacterial proteins. The deduced amino acid sequences of the putative PE-binding adhesins (HI0119) from NTH16564 (HI0119N) and H. influenzae Rd strain (HI0119R), YebL from E. coli. and FimA from Streptococcus parasanguis were aligned with the aid of the PILEUF program.
Figure 6. Predicted structural properties of the putative PE-binding adhesin of *H. influenzae*. (A) Hydropathy profile. Hydropathy was determined by the methods of Kyte and Doolittle (Kyte & Doolittle, 1982). The x axis corresponds to the amino acid residues, while y axis corresponds to the relative hydrophathy index. (B) and (C) Predicted secondary structure of turn, strand and helix based on its deduced amino acid sequence. This analysis was performed by the PSA server (Protein Sequence Analysis System), the BioMolecular Engineering Research Center, University of Boston, USA (Stulz et al., 1993; White et al., 1994).
C

STRAND

4
3
TURN
1

HELIX

LOOP

Residue
1.2. Expression of recombinant fusion proteins

A BamHI-EcoRI fragment was cleaved from the plasmid pTA119 and ligated into the complementary restriction enzyme sites of the E. coli expression vector pGEX-2T (Fig. 7). A recombinant fusion protein with glutathione S-transferase (GST, 29 kDa) was thus prepared. The GST fusion protein system was chosen because its high level expression and rapid purification. Moreover, the protein of interest can be cleaved away from GST. Approximately 25 mg of the recombinant protein (GST-0119) was purified from 1 liter of culture by column chromatography on glutathione-Sepharose 4B. The apparent molecular weight of purified GST-0119 is 67 kDa by SDS-PAGE (Fig. 8A). Cleavage of the 37 kDa putative PE-binding adhesin from GST using thrombin was then attempted. However, only partial digestion (5-10%) was obtained, in spite of higher thrombin concentration, longer digestion time, or presence of detergents or denaturants, such as 0.5% SDS and 1 M urea (Fig. 8A).

Since GST-0119 was resistant to thrombin digestion, the coding sequence was subcloned into the pTrcHisA expression vector (Fig. 7). The fusion protein, His-0119, was purified from the bacterial lysate by Ni-NTA agarose chromatography (Fig. 8B). Although His-0119 was obtained in the soluble fraction of bacterial sonicates, it could not be isolated using Ni-NTA agarose chromatography under native conditions. The fusion protein was subsequently purified in the presence of 6 M guanidine-HCl, suggesting that the 6xHis tag is exposed for Ni²⁺ binding only in the denatured state. The molecular mass of the purified His-0119 was 40 kDa, due to its 3 kDa extra tag (Fig. 8B). A yield of 5 mg of His-0119 was typically obtained from one liter bacterial culture.
Figure 7. Construction of pGHA1 and pTHA1. The coding region of the putative adhesin was removed from plasmid pTA119 using the restriction enzymes BamHI and EcoRI. The 936 bp fragment was subsequently cloned into the expression vectors pGEX-2T and pTrc-A. The resulting constructs were designated pGHA1 and pTHA1, respectively.
Figure 8. Analysis of the fusion proteins by SDS-PAGE. Proteins were separated on 12% acrylamide gels and stained with Coomassie blue as described in the Methods. Approximately the same amount of protein (10 µg) was loaded in each lane. Molecular mass marker (in kilodaltons) are indicated. (A) Analysis of purified GST-0119. Lanes: M, marker proteins; 1 & 5, purified GST-0119; 2, typical thrombin digestion pattern of GST-0119, the 29 kDa band represents GST; 3 & 4, thrombin digestion pattern of GST-0119 in the present 0.5% SDS and 1 M Urea, respectively. (B) Analysis of purified His-0119. His-0119 was purified using Ni-NTA agarose under denatured conditions. Lanes: M, marker proteins; 1, purified His-0119.
1.3. The putative PE-binding adhesin has a disulfide bond in the C-terminus

Sequence analysis showed that the putative PE-binding adhesin contains two cysteine residues near the C-terminus. To determine whether these two cysteine residues form a disulfide bond, both GST-0119 and His-0119 fusion proteins were analyzed on SDS-PAGE under "native", reducing and nonreducing conditions (Fig. 9). The shift in molecular weight between reducing and nonreducing conditions in both fusion proteins, is indicative of the presence of a disulfide bond in the native protein (GST contains no disulfides). The disulfide-containing proteins migrate faster in SDS-PAGE than do their reduced counterparts, presumably because the form containing disulfide is more compact (Scheele & Jacoby, 1982). We did not observe oligomers formed from either fusion protein when samples were treated with only 0.1% SDS, no mercaptoethanol, and not boiled prior to electrophoresis ("native" condition).

1.4. Immunoprecipitation of the putative PE-binding adhesin in H. influenzae

Antisera against GST-0119 and His-0119 were made in rabbits. Immunoblot analysis after the first boost immunization demonstrated that both antisera could specifically recognize the fusion proteins GST-0119, His-0119 and a single band of 37 kDa in a whole cell lysate from H. influenzae. Since both antisera have similar sensitivity and specificity on Western blot, antibody against His-0119 was used for subsequent studies. It was also felt that the small His tag of this fusion protein would make this antiserum more specific than the anti-GST-0119 antiserum since GST is quite immunogenic.
Figure 9. Determination of the oligometric and intramolecular disulfide-bond status of purified GST-0119 and His-0119 using SDS-PAGE. Proteins were separated on 12% acrylamide gels and stained with Coomassie blue. Protein samples were treated under minimally denaturing conditions (0.1% SDS, no heating) (lanes 1), or denatured in sample buffer in the presence (lanes 2), or absence (lanes 3) of reducing agent prior to loading and running the gel.
To determine whether antisera against the fusion proteins could recognize the native *Haemophilus HI0119* protein, antibody against His-0119 was used to immunoprecipitate the *Haemophilus* native protein. Compared with preimmune serum, antibody against His-0119 immunoprecipitated the 37 kDa protein band. Western blotting confirmed that the 37 kDa protein is the putative PE-binding adhesin, indicating that antiserum to His-0119 could specifically recognize the native *Haemophilus HI0119* protein (Fig. 10).

1.5. The putative PE-binding adhesin cannot bind to Hep2 cells

As a putative PE-binding adhesin, purified fusion proteins should bind to mammalian cells which bind *H. influenzae*. To confirm this, FITC labeled GST-0119 and His-0119 were used in a Hep2 cell binding study. Earlier work showed that antiserum against the PE-binding adhesin preparation could partially inhibit the binding of *H. influenzae* to Hep2 cells (Busse *et al.*, 1997). FITC labeled proteins were incubated with Hep2 cells for different times at 4°C as described in the methods section. Unexpectedly, both fusion proteins were unable to bind to the surface of Hep2 cells (results not shown).
Figure 10. Immunoprecipitation of the putative PE-binding adhesin from H. influenzae. Whole cell extracts of NTHi6564 were immunoprecipitated with antiserum to His-0119 (lane 1) and preimmune serum (lanes 2) as described in the method, and followed by SDS-PAGE (A) and immunoblotting (B) with antiserum against His-0119. Lane M, marker proteins. Molecular mass marker (in kilodaltons) are indicated on the right.
1.6. The putative PE-binding adhesin cannot bind to PE by TLC overlay

Since the putative PE-binding adhesin was originally purified from a PE affinity matrix, this protein should bind to PE. To confirm this, both purified fusion proteins GST-0119 and His-0119 were used in a TLC overlay assay. Unexpectedly, neither fusion protein bound to PE by TLC overlay (Fig. 11). Why couldn't the fusion proteins bind to PE in vitro, while this protein could be purified from a PE affinity matrix? One possibility was that the GST or 6Xhistidine tag in the fusion proteins interfered with binding to PE on cells or on TLC plates. It was therefore necessary to express the putative PE-binding adhesin without a fusion tag.

1.7. Expression of the putative PE-binding adhesin in E. coli

The putative PE-binding adhesin gene and its upstream 590 bp region was amplified as a 1.6 kb fragment from NTHI6564 chromosomal DNA, and was cloned into the TA cloning vector to generate plasmids pTAp119a and pTAp119b (Fig. 12). In pTAp119a, the HI0119 gene is in the same orientation as the lac promoter but is inverted with respect to this promoter in plasmid pTAp119b. Recombinant E. coli strains harboring either plasmid were found to overexpress the 37 kDa protein at a similar level, suggesting that transcription of this protein is initiated at a H. influenzae promoter, located in the 590 bp upstream region, which is well recognized by E. coli RNA polymerase. Recombinant protein was found to localize in the periplasmic space of E. coli and constituted about 35% of the total protein in the periplasmic shock fluid (Fig. 13). Interestingly, large amounts of recombinant
Figure 11. TLC overlay of the purified GST-0119 and His-0119 fusion proteins. Purified fusion proteins (10 μg/ml) were tested for binding of PE and PC (2 μg) separated by TLC as described in the method. Panel A was visualized by iodine spray, and panel B (binding of GST-0119) and panel C (binding of His-0119) were visualized with antiserum against His-0119. Lanes: 1, PE; 2, PC. This experiment was repeated at least four times with identical results.
Figure 12. Construction of plasmid pTAp119. The putative PE-binding adhesin gene and its upstream 590 bp region from NTHI6564 chromosomal DNA was amplified by PCR. The amplified 1.6 kb fragment was cloned into the TA cloning vector to generate plasmid pTAp119. (A) A schematic diagram of cloning. (B) The amplified 1.6 kb fragment (lane 1). (C) BamHI/EcoRI restriction enzyme digest of plasmid pTAp119 (lane 1). The DNA fragments were resolved in 0.7% agarose gel. Lane M: 1 kb DNA marker. Molecular mass marker are indicated.
Figure 13. Expression of recombinant putative adhesin in *E. coli*. *E. coli* strain TOP10 was transformed with pTAp119 generating the derivative *E. coli* strain Y119. Control was TOP10 strain transformed with vector pCRHTM. Y119 and control were grown overnight at 4°C. Whole cell extracts and periplasmic extracts were prepared as described in the methods. Both extracts were analyzed on SDS-PAGE (A) and Western blotting (B) using antisera to His-0119. Lanes: M, marker proteins, Molecular mass marker (in kilodaltons) are indicated; 1, control; 2, Y119.
protein was also found to exist in the supernatant of overnight cultures, suggesting that autolysis of bacteria had occurred during cultivation, presumably because of high expression of recombinant protein which resulted in an exhaustion of the components of the protein secretory pathway. However, the existence of recombinant protein in the culture supernatant provided a convenient protein source for testing PE-binding ability of recombinant putative PE-binding protein.

1.8. The putative PE-binding adhesin is not a PE-binding protein

The culture supernatant containing recombinant protein was applied on a PE-celite affinity column. After the column was washed to remove unbound material, the column was eluted with a 1 M Tris solution (pH 11.2). SDS-PAGE analysis demonstrated that recombinant protein could be purified from the PE column, but large amounts of this protein still existed in the unbound fraction. There are several possibilities to explain this phenomenon. Firstly, there could be at least two different conformations of expressed adhesin; one that binds to PE, another that does not. Secondly, the PE-binding ability of recombinant protein depends on other accessory factor(s), which are not present in quantities which support PE binding of the entire adhesin sample. Finally, recombinant protein may not bind to PE, but bind to a PE-binding protein. If one of these probabilities is true, recombinant protein in the unbound fraction should not bind to PE. To confirm this hypothesis, the unbound fraction was re-applied to a fresh PE column. After washing and elution, the Tris eluate was examined by SDS-PAGE. Similar amounts of recombinant protein were found to exist in the eluate (Fig. 14A), suggesting that another mechanism may be responsible for binding of this protein to the PE column.
Figure 14. Purification of recombinant adhesin by PE-celite or celite chromatography. The culture supernatant of *E. coli* was separated on a PE-celite or celite column, and aliquots of fractions were separated by SDS-PAGE and detected by Coomassie blue staining. (A) PE-affinity chromatography. The culture supernatant (lane 1) of Y119 was applied to a PE-celite column. The bound proteins were eluted with 1 M Tris solution (pH 11.2) (lane 3). Then unbound fraction (lane 2) was re-applied to a fresh PE-celite column. After washing, the column was also eluted with 1 M Tris solution (pH 11.2)(lane 4). (B) Celite chromatography. Recombinant adhesin (lane 1) was purified by celite chromatography as described in the methods. Lane M, marker proteins. Molecular mass marker (in kilodaltons) are indicated on the left.
In preparation of PE affinity matrix, PE was adsorbed onto celite beads (Boulanger et al., 1994). To determine whether recombinant protein could bind celite only, the culture supernatant which contained recombinant protein was applied to a celite column and eluted as before. SDS-PAGE demonstrated large amounts of recombinant protein in the Tris eluate (Fig. 14B). This result clearly indicates that recombinant protein could bind celite in the absence of PE, and the existence of PE actually partially blocked binding of the protein to the PE column. Thus, the putative PE-binding adhesin should be called the celite binding protein.

1.9. Antiserum against His-0119 cannot inhibit *H. influenzae* binding to PE by TLC overlay

Previous results demonstrated that antiserum raised against the PE-binding adhesin preparation could block binding of *H. influenzae* to PE and Gg3 by TLC overlay (Busse et al., 1997). To determine if this result was due to anti-HI0119 antibody, the experiment was repeated using anti-His-0119 antiserum. After NTH16564 was preincubated with antiserum to His-0119, the bacterium retained the ability to bind to PE or Gg3 as monitored by TLC overlay (Fig. 15). Nonimmune serum was used as control. Thus another component of the original preparation was likely responsible for the inhibitory activities observed.
Figure 15. Inhibition of NTHI6564 PE binding by antiserum against His-0119. *H. influenzae* was preincubated with the indicated antibody prior to assay of glycolipid binding specificity by TLC overlay. Inhibition of lipid binding after preincubation of *H. influenzae* with preimmune serum (A), antiserum against His-0119 (B) was assessed. Lanes: 1, PE; 2, Gg3 and Gg4; 3, SGC; 4, SGG.
The unexpected discoveries that the putative PE-binding adhesin is unable to bind to PE and Hep2 cells and PE binding of *H. influenzae* could not be inhibited by antiserum to His-0119 strongly suggested that this protein may not be involved in attachment of *H. influenzae* to PE. However, the progress which had been made at that time provided us an excellent opportunity to identify the function of this unknown protein. As a first step, the unique celite binding ability was characterized.

2. Characterization of the celite binding protein

Although the celite binding ability of recombinant HI0119 protein has been demonstrated, it was necessary to determine whether GST-0119, His-0119 and the native *Haemophilus* protein have similar binding ability.

2.1. GST-0119 and His-0119 bind to celite

FITC-conjugated GST-0119 was incubated with celite 545. After washing, extensive GST-0119 binding to celite was observed. This binding was specifically inhibited by unlabeled GST-0119, but not by unlabeled GST (Fig. 16ABC). FITC-BSA did not bind to celite. In addition, FITC-labeled GST-0119 did not bind to silica, the major component of celite (results not shown).
Figure 16. Celite binding of FITC-conjugated GST-0119. FITC-GST-0119 (0.2 μg/ml) was incubated with celite in the present or absence of a 20-fold molar excess of unlabeled inhibitors. After washing, the celite was observed under incident uv illumination. (A) FITC-GST-0119 only. (B) FITC-GST-0119 plus unlabeled GST. (C) FITC-GST-0119 plus unlabeled GST-0119. (D) FITC-GST-0119 plus unlabeled GST-0119SII.
To confirm the celite binding ability of His-0119, the sonic lysate of an *E. coli* strain containing pTHAl was applied to a celite matrix. After the column was washed to remove unbound material, the column was specifically eluted with EDTA. The His-0119, identified by Western blotting, was eluted as a single protein band (Fig. 17).

2.2. Purification of the native *Haemophilus* HI0119 protein by celite chromatography

To determine whether the native *Haemophilus* HI0119 protein could bind celite, the sonic lysate of NTHI6564 was fractionated using celite chromatography. The eluted fraction was analyzed on SDS-PAGE and Western blotting. Both analyses demonstrated the native HI0119 protein in the eluate as the primary protein band (Fig. 18). These results are consistent with the previous observation that HI0119 was purified from a PE-celite column using EDTA or Tris solution (pH 11.2) (Busse *et al.*, 1997).

2.3. Mutational analysis

The celite binding protein contains two cysteine residues at amino acid positions 254 and 308. To gain insight into the possible function of this putative disulfide bond, a Cys254 to Ser254 mutant was constructed and expressed as a GST fusion protein (GST-0119SII) (Fig. 19AB). Purified GST-0119SII had the same migration by SDS-PAGE under reducing and non-reducing conditions (Fig. 19C), suggesting that the reduced migration of both His-0119 and GST-0119 under non-reducing conditions is due to disulfide bond formation. However, this mutant was able to inhibit binding of the FITC-conjugated GST-0119 to celite beads (Fig. 16D).
Figure 17. Purification of His-0119 by celite chromatography. Whole cell extract of *E. coli* TOP10 strain transformed with pTHAl was prepared and applied to a celite matrix as described in the methods. Bound proteins were eluted with 10 mM EDTA. The eluted proteins were analyzed on SDS-PAGE and Western blotting. Lane M, protein markers. Molecular mass marker (in kilodaltons) are indicated on the left. Lane 1, His-0119 (10 µg) purified using Ni-NTA under denatured conditions. Lane 2, His-0119 (5 µg) eluted from celite matrix. (A) SDS-PAGE (12% gel). Protein was stained with Coomassie blue. (B) Immunoblot with polyclonal antiserum against His-0119.
Figure 18. Purification of the putative adhesin of *H. influenzae* by celite chromatography. The sonicate lysate of NTHI6564 was separated on a celite column, and aliquots of fractions were analyzed on SDS-PAGE (12% gel) and Western blotting. Lanes: M, protein markers; 1, sonicated lysate (20 µg) of NTHI6564; 2, unbound fraction (20 µg) after celite column; 3, the Tris eluate (5 µg) from celite column. Molecular mass marker (in kilodaltons) are indicated on the left. (A) Protein stained with Coomassie blue. (B) Immunoblot with polyclonal antiserum against His-0119.
Figure 19. Construction and expression of the cysteine mutant. The strategy used is described in Methods. (A) The amino acid sequence in the C-terminal region. The corresponding amino acid sequences are indicated below the nucleotide sequence, the numbers below the amino acids denote their position in the mature protein. (B) The primer HICM used to mutate cysteine 308. The codon for Cys308 was altered to encode serine (nucleotide change indicated in box, EcoRI site is underlined). (C) the cysteine mutant was expressed as a GST fusion protein in *E. coli*. The purified cysteine mutant was resolved by SDS-PAGE (12% gel) under reducing (lane 1) or nonreducing (lane 2) conditions. Lane M, marker proteins. Molecular mass marker (in kilodaltons) are indicated on the left.
A.  

HI0119

5' AGC TAC ATG GAA TGT TTA GCT AAA TAA 3'  
Ser Tyr Met Glu Cys Leu Ala Lys OCH  
304 305 306 307 308 309 310 311 312

B.  

Primer (HIcm): 5' GAA TTC TTA TTT AGC TAA AGA TTC CAT GTA GC 3'

C.
Fortunately, a unique DraIII recognition site was found to exist in the insert of plasmid pTHA1, not in the vector itself. The C-terminal 235 bp of the celite binding protein was therefore truncated by DraIII and EcoRI digestion of pTHA1. The resulting plasmid was denoted pTHAc (Fig. 20). The mutant product (His-0119cm) was purified using Ni-NTA agarose under denatured conditions. SDS-PAGE analysis under reducing and non-reducing conditions confirmed that the disulfide-bonded domain was absent from the purified mutant (Mr of 34 kDa) (Fig. 21). Furthermore, this mutant could be isolated from a celite matrix (Fig. 22). The results demonstrated that the disulfide-bonded C-terminal domain did not contribute to the celite binding of HI0119.

2.4. Conservation of the celite binding protein in H. influenzae

To examine the distribution of the HI0119 gene and its protein product in H. influenzae, 23 H. influenzae clinical strains, including 9 Hib and 14 NTHI strains, were analyzed by Western blotting and PCR (Fig. 23). These analyses revealed that the HI0119 gene and its expressed product were present in all nontypeable and type b H. influenzae strains tested. Some heterogeneity in protein size was observed from strain to strain, which reflected the variable size of the PCR products (Fig. 23).
Figure 20. Construction of the C-terminal truncated mutant. (A) schematic diagram for constructing C-terminal truncated mutant. C-terminal 235 bp of the celite binding protein was removed by DraIII and EcoRI digestion of pTHA1. (B) BamHI/HindIII restriction enzyme digest of plasmids pTHA1 and pTHA1cm. Lanes: M, 1 kb DNA marker; 1-9, different clones of plasmid pTHA1cm.
Figure. 21. SDS-PAGE analysis of purified His-0119cm. Purified His-0119cm and His-0119 were resolved by SDS-PAGE (12% gel) under reducing or nonreducing conditions. Lanes: M, marker proteins. Molecular mass marker (in kilodaltons) are indicated on the left: 1, reducing conditions; 2, nonreducing conditions.
Figure 22. Purification of His-0119cm by celite chromatography. Purification of the C-terminal truncation mutant, His-0119cm, was monitored by SDS-PAGE (12% gel). Lanes: M, molecular mass markers; 1, sonicated lysate (20 μg) of bacteria; 2, unbound fraction (20 μg) after celite chromatography; 3, His-0119cm (5 μg) eluted from celite column. (A) Protein stained with Coomassie blue. (B) Immunoblot with polyclonal antiserum against His-0119.
Figure 23. Determination of the conservation and expression of the celite binding protein in various clinical isolates *H. influenzae*. (A) SDS-PAGE pattern of whole cell preparation on 12% gel. (B) Western blotting with polyclonal antiserum against His-0119. Lanes: M, protein markers; 1 to 8 & 12, Hib clinical isolates; 9-11 & 13-23, NTHI clinical isolates. (C) Agarose gel electrophoresis of PCR products stained with ethidium bromide. DNAs were amplified with primers HIMA1 and HIMA2 as described in the methods. Lane M, 1 kb DNA ladder; Lanes 1-23 same as in (A) and (B).
3. The celite binding protein is a periplasmic zinc-binding protein (Pzp1)

3.1. Localization of the celite binding protein

The previous study has shown that the N-terminal 24-amino-acid sequence of the celite binding protein forms a signal peptide (Busse, 1997), which suggests that this protein is extracytoplasmic. Since the celite binding protein was originally purified from a surface protein (water) extract from a NTHI stain, immunofluorescence with anti-His-0119 was used to determine the possible surface localization of this protein. This study revealed that this protein is not on the H. influenzae cell surface (results not shown). Furthermore, Western blotting with antisera raised against His-0119 showed that, following overnight growth of H. influenzae in liquid culture, the celite binding protein is cell-associated and is not secreted into the culture medium (Fig. 24). Other members of the FimA family possess an N-terminal lipid linkage consensus sequence LXXC. Although the celite binding protein contains no such sequence, and does not contain obvious membrane spanning hydrophobic regions, the possibility of membrane anchorage was investigated.

A characteristic of integral membrane proteins, including outer membrane proteins from gram-negative bacteria, is their selective partitioning into the detergent phase after Triton X-114 extraction (Bordier, 1981). NTH6564 was extracted with Triton X-114; most proteins were insoluble. Triton X-114 soluble proteins were partitioned into a detergent phase and an aqueous phase. The celite binding protein was exclusively found in the aqueous phase (Fig. 25), indicating that this protein is not a membrane protein.
Figure 24. SDS-PAGE and Western blotting of the culture supernatant of NTHI6564. The supernatant of NTHI6564 overnight growth culture was collected and concentrated 100 times as described in the methods. The concentrated supernatant was analyzed by SDS-PAGE and Western blotting. Lanes: M, protein markers; 1 & 3, the concentrated supernatant (15 μg); 2 & 4, bacterial pellet (15 μg). (A) Protein stained with Coomassie blue. (B) Immunoblot with polyclonal antiserum against His-0119.
Figure 25. Localization of the celite binding protein in NTHI6564. Triton X-114 extract, periplasmic and whole cell extracts of NTHI6564 were suspended in sample buffer and boiled for 3 min under reducing conditions prior to loading. (A) SDS-PAGE analysis (12% gel). (B) Western blotting with antiserum against fusion protein His-0119. Lane 1, Triton X-114 insoluble protein of NTHI6564. Lane 2, Triton X-114 detergent fraction. Lane 3, Triton X-114 aqueous fraction. Lane 4, periplasmic proteins of NTHI6564. Lane 5, whole cell extract of NTHI6564.
The periplasmic proteins of gram-negative bacteria can be selectively released by osmotic shock (Ames, 1994). The celite binding protein was found in the periplasmic extract of NTHI6564 (Fig. 25), indicating that this protein is localized in the periplasmic space in H. influenzae.

3.2. Purification of recombinant celite binding protein in E. coli

The periplasmic extracts of an E. coli strain harboring plasmid pTAp119 were prepared. Recombinant protein was purified by chromatofocusing (Fig. 26) and gel filtration chromatography (Fig. 27). The celite binding protein thus obtained, has been purified to >98% homogeneity as determined by SDS-PAGE (Fig. 27B) with a yield of approximately 7 mg pure protein from 1 liter of bacterial culture. The observed isoelectric point (pI) of this protein, as estimated by chromatofocusing, was 6.4, which is close to the theoretical pI (pH 6.7) (Fig. 26). Like the fusion proteins GST-0119 and His-0119, the recombinant protein contained a disulfide bond as determined by differential migration by SDS-PAGE under reducing and nonreducing conditions. Amino acid analysis of the purified recombinant protein gave the expected amino acid composition for the mature, processed form of the protein (Table 2).

3.3. Circular dichroism (CD) analysis of recombinant celite binding protein

In order to obtain initial experimental information about the secondary structure of the purified protein, a CD spectrum was recorded. The strong molar ellipticity at 222 nm suggested that the major secondary structure of the recombinant protein is α-helix (25%) (Fig. 25).
Table 2. Amino acid analysis of recombinant celite binding protein

<table>
<thead>
<tr>
<th>Amino acids</th>
<th>expected number</th>
<th>observed number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ala</td>
<td>27</td>
<td>28.1</td>
</tr>
<tr>
<td>Cys</td>
<td>2</td>
<td>0.16</td>
</tr>
<tr>
<td>Asp &amp; Asn</td>
<td>41</td>
<td>38.3</td>
</tr>
<tr>
<td>Glu &amp; Gin</td>
<td>28</td>
<td>29.9</td>
</tr>
<tr>
<td>Gly</td>
<td>16</td>
<td>15.9</td>
</tr>
<tr>
<td>Ser</td>
<td>17</td>
<td>11.5</td>
</tr>
<tr>
<td>His</td>
<td>28</td>
<td>29.5</td>
</tr>
<tr>
<td>Arg</td>
<td>2</td>
<td>2.4</td>
</tr>
<tr>
<td>Thr</td>
<td>16</td>
<td>13.1</td>
</tr>
<tr>
<td>Pro</td>
<td>12</td>
<td>13.5</td>
</tr>
<tr>
<td>Tyr</td>
<td>9</td>
<td>9.2</td>
</tr>
<tr>
<td>Val</td>
<td>21</td>
<td>22.0</td>
</tr>
<tr>
<td>Met</td>
<td>1</td>
<td>1.2</td>
</tr>
<tr>
<td>Ile</td>
<td>18</td>
<td>18.0</td>
</tr>
<tr>
<td>Leu</td>
<td>25</td>
<td>25.4</td>
</tr>
<tr>
<td>Phe</td>
<td>12</td>
<td>12.3</td>
</tr>
<tr>
<td>Lys</td>
<td>33</td>
<td>35.4</td>
</tr>
</tbody>
</table>
Figure 26. Chromatofocusing of recombinant celite binding protein. The celite binding protein was purified from a periplasmic extract of TApep1 using chromatofocusing as described in the methods. Fractions were collected, and \( \text{OD}_{280\text{nm}} \) (○) and pH (●) were determined. The celite binding protein peak (identified by Western blotting) is indicated with an arrow.
Figure 27. Gel filtration of recombinant celite binding protein. The peak fractions from chromatofocusing were pooled and applied on gel filtration as described in the methods. The peak fractions obtained were analyzed on SDS-PAGE (12% gel). (A) Gel filtration elution profile. (B) Protein stained with Coomassie blue. The fraction number is given above each lane. M denotes the molecular weight marker.
Figure 28. CD spectrum of recombinant celite binding protein. CD measurement was carried out as described in the methods. The protein concentration was 0.56 mg/ml.
3.4. The celite binding protein can bind zinc specifically

Celite is a diatomaceous earth product, the shell of ancient diatoms. It consists of 90% silica and minerals (Table 3). Several lines of evidence suggest that the celite binding protein may have a capability to bind metal(s). First, this protein bound to celite but not to silica, which suggested that this protein may bind to a mineral substance on the celite surface. Second, this protein was repeatedly eluted with EDTA, suggesting that the celite binding was ion-dependent. Third, this protein had a distinct histidine-rich domain which was suggestive of a metal-binding function. Since typical chemical analysis of celite (Table 3) showed that minerals in the celite exist as oxidized form which may compromise coordination to this protein, but still allow charge interaction. However, it is possible that hydrophobic interaction may also play a role in the celite binding.

Table 3. Typical chemical analysis of celite 545*

<table>
<thead>
<tr>
<th>Chemical compositions</th>
<th>Percentage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SiO₂</td>
<td>89.6</td>
</tr>
<tr>
<td>Al₂O₃</td>
<td>4.0</td>
</tr>
<tr>
<td>Fe₂O₃</td>
<td>1.3</td>
</tr>
<tr>
<td>P₂O₅</td>
<td>0.2</td>
</tr>
<tr>
<td>TiO₂</td>
<td>0.2</td>
</tr>
<tr>
<td>CaO</td>
<td>0.5</td>
</tr>
<tr>
<td>MgO</td>
<td>0.6</td>
</tr>
<tr>
<td>Na₂O + K₂O</td>
<td>3.3</td>
</tr>
</tbody>
</table>

* from Fisher Scientific Inc.
To determine the nature and amount of metals bound to the celite binding protein, the purified protein was analyzed by Neutron Activation (NA). NA analysis revealed that this protein contained about two zinc atoms per protein molecule, but did not contain measurable levels of Cr, Ni, Sc, Fe, Co or Mn. The zinc content of the purified protein was also confirmed by atomic absorption spectroscopy (Table 4). In addition, $^{65}\text{Zinc}$ binding to the celite binding protein was demonstrated using a $^{65}\text{Zinc}$ blot technique (Fig. 29).

To assess the zinc binding capacity of the celite binding protein, the purified protein was incubated in the presence of 5 mM EDTA or 1 mM Zn(NO$_3$)$_2$ at 40°C overnight. Following exhaustive dialysis, the zinc content was determined by atomic absorption spectroscopy. The untreated protein contained 1.58 zinc atoms per protein molecule while EDTA treatment reduced bound zinc to an undetectable level. On incubation with zinc, a total of 5 zinc atoms per protein molecule could be bound (Table 4).

Table 4. Analysis of Zinc Content by Atomic Absorption Spectroscopy

<table>
<thead>
<tr>
<th>Sample</th>
<th>Zinc (µM)</th>
<th>Protein (µM)</th>
<th>No. of zinc/protein molecule</th>
</tr>
</thead>
<tbody>
<tr>
<td>PZP1+Zn</td>
<td>4.18</td>
<td>0.84</td>
<td>4.96</td>
</tr>
<tr>
<td>Dialysate</td>
<td>&lt;0.1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>PZP1</td>
<td>1.61</td>
<td>1.02</td>
<td>1.58</td>
</tr>
<tr>
<td>Dialysate</td>
<td>&lt;0.1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>PZP1+EDTA</td>
<td>&lt;0.1</td>
<td>0.772</td>
<td>&lt;0.13</td>
</tr>
<tr>
<td>Dialysate</td>
<td>&lt;0.1</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

*Purified PZP1 was treated with 1 mM zinc or 5 mM EDTA. After exhaustive dialysis, the zinc content of samples and dialysates was measured by atomic absorption spectroscopy.*
Figure 29. $^{65}$Zinc-blotting with the celite binding protein. The periplasmic extracts from *E. coli* strains which contained plasmid pTApzp1 (lane 1) or pCriTM (lane 2), and purified celite binding protein (lane 3) were analyzed by 12% SDS-PAGE followed by Coomassie blue staining (A), Western blotting with anti-His-0119 (B), or with $^{65}$Zinc followed by autoradiography (C). $^{65}$Zinc blotting was completed as described in the methods.
It is clear that the putative PE-binding adhesin, subsequently identified as the celite binding protein is in fact, a periplasmic zinc binding protein (Pzp1).

3.5. Regulation of the \(pzp1\) gene of \(H.\ influenzae\)

To gain insight into the regulation of the \(pzp1\) gene, NTHI6564 was grown under different conditions, and the synthesis and the cellular content of Pzp1 were assessed by SDS-PAGE and Western blotting using antiserum to His-0119. The tested conditions included temperature (37°C or 42°C), atmosphere (aerobic or anaerobic), salt concentration (1% NaCl or 1.5% NaCl), supplemental zinc (0 μm, 5 μm, 100 μm, 200 μm), iron-limitation (EDDHA: 0 μm, 100 μm, 150 μm), and metal chelation (EDTA: 0 μm, 100 μm). Samples also were taken at different stages of growth (log, stationary). No appreciable change in expression levels of Pzp1 due to the above conditions was observed (results not shown).

4. Pzp1 plays a key role for zinc uptake in \(H.\ influenzae\)

4.1. Construction of the \(pzp1\) negative mutant

A \(pzp1\)-deficient mutant was constructed as described in the methods section. Initial attempts to select transformants under aerobic conditions were unsuccessful; however, under anaerobic conditions, small colonies were observed after overnight incubation. Transformant colonies were screened by direct amplification of chromosomal DNA from single colonies by PCR. Oligonucleotide primers used in the screening PCR were primer P1 and P2 to allow amplification of the coding region of the \(pzp1\) gene (Fig. 30A). A comparison of PCR in Fig. 30B illustrates that an insertion of the 1.2 kb kanamycin cassette into the \(pzp1\) gene had occurred in
transformants of Ypzp1::kan, as expected. Western blotting with antiserum against fusion protein His-0119 demonstrated that Pzp1 was absent in the whole cell extract of the pzp1 mutant (Fig. 31). Further study demonstrated that the pzp1 mutant did not grow on chocolate agar plates or in 20% Levinthal broth under aerobic conditions, but did grow poorly under anaerobic conditions (Fig. 32). This result indicates that the pzp1 gene is essential for growth under aerobic conditions, and is also important for anaerobic growth of H. influenzae.

4.2. Growth defect of the pzp1 mutant can be specifically rescued by zinc

Considering the periplasmic location and potent zinc binding capability of Pzp1, it seemed reasonable that the growth defect was caused by zinc deficiency. Accordingly, chocolate agar plates were supplemented with ZnCl2 at 100 μM, and the pzp1 mutant was grown on the plates with zinc under aerobic and anaerobic conditions. The growth defect of the pzp1 mutant under both aerobic and anaerobic conditions was rescued by zinc addition (Fig. 32), suggesting that the pzp1 mutant is defective in zinc uptake. To assess the substrate specificity of suppression, 100 μM CaCl2, MgCl2, CuCl2, NiCl2, Cd(NO3)2, MnCl2, FeCl3, ZnCl2 or Zn(NO3)2 were added into 20% Levinthal broth. Only ZnCl2 and Zn(NO3)2 rescued the growth defect of the pzp1 mutant under aerobic conditions (Fig. 33A), indicating that the suppression of the growth defect is zinc specific.
Figure 30. Construction of the pzpl deficient mutant. (A) Schematic of the *H. influenzae* chromosomal region encompassing the *pzpl* gene. ORFs are indicated by boxes. The gene *chlN* encodes a putative molybdopterin biosynthesis protein. The arrows indicate the direction of transcription. The localization of oligonucleotide primers (P1 and P2) used in screening reactions are shown above. The deleted region was replaced by a 1.2 kb Kan^R^ gene. (B) PCR analysis of chromosomal DNA from NTHI6564 wt and the *pzpl* mutant. The oligonucleotide primers P1 and P2 were used to amplify the *pzpl* gene.
Figure 31. SDS-PAGE and Western blotting of wild type and the pzp1 mutants. The whole cell lysates from NTHI6564 and the two separate pzp1 mutants (c1 and c2) were suspended in sample buffer and boiled 3 min under reducing conditions prior to loading the gel. (A) 12% SDS-PAGE. (B) Western blotting with antiserum against fusion protein His-0119.
Figure 32. Zinc suppression of the growth defect of the \textit{pzpl} mutant. NTHI6564 \textit{wt} and the \textit{pzpl} mutant were spread onto chocolate agar plates lacking or containing 100 μM ZnCl$_2$ supplementation under aerobic or anaerobic conditions. The plates were incubated at 37°C for 18 hours.
Anaerobic

Aerobic

Zinc (μM)

0 100

wt

pzp1 mutant
Figure 33. Recovery of growth defect of the pzp1 mutant is zinc specific. (A) Effect of various metals on pzp1 mutant growth. The pzp1 mutant was grown aerobically in 5 ml 20% Levinthal broth which was supplemented with different metals at 100 μM at 37°C with shaking at 180 rpm. After overnight growth optical densities at 600 nm were determined. (B) Growth curve of the pzp1 mutant at different concentrations of zinc. NTHI6564 wt and the pzp1 mutant were grown in 50 ml 20% Levinthal broth which was supplemented with ZnCl2 at 5 μM and 100 μM at 37°C under aerobic conditions. Growth was monitored by optical density at 600 nm.
A

B

0.0 0.2 0.4 0.6 0.8

OD at 600 nm

WT npl mutant ZnCl₂ Zn(OC₃)₂ CaCl₂ CoCl₂ FeCl₂ MgCl₂ MnCl₂ Mn(OC₃)₂ NC₃

0 2 4 6 8 10 12

Time (hours)

OD at 600 nm

wild type pepF mutant 5 μM ZnCl₂ 100 μM ZnCl₂
The growth characteristics of the \textit{pzpl} mutant at different zinc concentrations under aerobic conditions were compared with that of the wild-type parent strain (Fig. 33B). In the presence of 5 \( \mu M \) ZnCl\(_2\), the growth defect of the mutant cannot be rescued. At 100 \( \mu M \) ZnCl\(_2\), the growth rate of the mutant was similar to that of the \textit{wt} strain, although the cell density at stationary phase was lower than that of the \textit{wt} strain (Fig. 33B).

5. Outer membrane protein P2 is a PE-binding protein

Although \textit{Pzp1} is probably not involved in adhesion of \textit{H. influenzae} to PE, it is not clear why the original antiserum raised against the PE-binding preparation could inhibit \textit{H. influenzae} binding to PE and Hep2 cells. We believed that this original protein preparation must have contained the real PE-binding component which was recognized by this antiserum. The previous experimental records were rechecked. It was found that OmpP2 in fact existed in the original PE-binding adhesin preparation as a minor component, even though a single protein band was subjected to N-terminal sequencing (Busse, 1997). In order to identify the adhesin from \textit{H. influenzae}, whole cell extracts of NTHI6564 were prepared and applied to a PE affinity column. Bound proteins were eluted from the column with 1 M Tris (pH 11.2) and then subjected to SDS-PAGE. Several proteins were eluted from the PE column, the most prominent band migrating at about 42 kDa (Fig. 34A). By Western blotting, three protein bands, including 42 kDa, 32 kDa and 18 kDa, were recognized by the antiserum raised previously against the PE-binding preparation (Fig. 34B), but both the 42 kDa and the 32 kDa bands did not react with the antiserum to His-0119.
Figure 34. Purification of *H. influenzae* adhesin(s) from a PE affinity matrix. (A) Protein staining. The *H. influenzae* sonicate lysate was separated on a PE affinity matrix, and the Tris eluate was resolved on 12% SDS-PAGE and stained by Coomassie blue. Lane M, molecular mass markers. Lane 1, the Tris eluate (15 μg). (B) Western blotting with antiserum raised previously against the PE-binding preparation. (C) Western blotting with antiserum against His-0119.
These results suggested that the 42 kDa and 32 kDa proteins were PE-binding adhesin candidates. N-terminal sequences of the 42 kDa and 32 kDa proteins were determined after SDS-PAGE and transfer to a PVDF membrane. The results showed that the amino acid sequences were AVVYNNEGTVNE, and APQENTFYAG, corresponding to the 42 kDa and 32 kDa protein bands, respectively. The 18 kDa band was also sent for sequencing. However, for unknown reasons, the 18 kDa band could not be sequenced. N-terminal sequences were used to search the TIGR H. influenzae genome database (Fleischmann et al., 1995), and it was found that the 42 kDa and 32 kDa proteins corresponded to outer membrane protein P2 and outer membrane protein P5, respectively.

OmpP2 is the most abundant protein in the H. influenzae outer membrane and a target for protective antibodies. It functions as a porin protein. To confirm that OmpP2 was a PE-binding protein, PCR was employed to amplify the about 1 kb coding region of ompP2 from NTHI6564 chromosomal DNA (Fig. 35A). The amplified product was ligated into the TA cloning vector, generating plasmid pTAompP2. The plasmid pTAompP2 was digested with BamHI and XhoI, and the insert was subcloned into the complementary sticky ends of the pGEX-2T expression vector. The resulting plasmid was denoted as pGHP2 (Fig. 35B). GST-P2 was expressed in E. coli. The kinetics of IPTG-induced protein expression were demonstrated in SDS-PAGE (Fig. 36A). Western blotting with an antiserum against OmpP2 detected the expressed GST-P2 (Fig. 36B). However, GST-P2 was found to be highly insoluble during the purification (Fig. 37). Lowering the growth temperature to 25°C appeared to increase the solubility of GST-P2. Small amounts of GST-P2
Figure 35. Separation of amplified mature OmpP2 gene and restriction enzyme digestions of pGHP2 by 0.7% agarose gel electrophoresis. (A) The PCR product appearing at 1050 bp. (B) BamHI/XhoI restriction enzyme digest of plasmid pGHP2 from two separate clones (c1 & c2). Lane M contain a DNA marker.
Figure 36. SDS-PAGE and Western blotting of whole cell lysates of induced and noninduced *E. coli* TOP10 harboring pGHP2. Lanes: M, molecular mass markers; 1, bacteria at an OD$_{600nm}$ of 0.5; 2-4, induced cells 1, 2, 3, 4 hour after induction, respectively, showing the fusion protein GST-P2 at the expected molecular mass of 71 kDa. (A) Protein stained with Coomassie blue. (B) Immunoblot with polyclonal antiserum against OmpP2.
Figure 37. Purification of the fusion protein GST-P2. Cultures of *E. coli* expressing GST-P2 were harvested, and the sonicate lysate was prepared. Purified GST-P2 was resolved on 12% SDS-PAGE and detected by Coomassie blue. Lanes: M, molecular mass markers. 1 & 2, whole cell lysates of noninduced and induced bacteria, respectively; 3 & 4, the post-sonicate pellet of induced bacteria. 5 & 6, purified GST-P2.
Figure 38. TLC overlay of the fusion protein GST-P2. Purified fusion protein GST-P2 (10 μg/ml) was tested for binding of glycolipids (2 μg) separated by TLC as described in the methods. Panel A was visualized by orcinol spray, and panel B (binding of GST) and panel C (binding of GST-P2) were visualized with antiserum against GST. Lanes: 1, from bottom, GMI, Gb4 and Gg3; 2, PE from E. coli; 3, PE from soybean; 4, PC; 5, PG.
<table>
<thead>
<tr>
<th></th>
<th>OmpP2:</th>
<th>OmpP5:</th>
</tr>
</thead>
<tbody>
<tr>
<td>220</td>
<td>KIANVRNXYNE -DESHTQLNVLATLGYRFSDGLVLSELDSOYAKWIYKX</td>
<td>KVGYHRNSFTFGGGYQTLQNNNLGAVELGYDFQAKREKGTYVXK--TN</td>
</tr>
<tr>
<td>230</td>
<td></td>
<td></td>
</tr>
<tr>
<td>240</td>
<td></td>
<td></td>
</tr>
<tr>
<td>250</td>
<td></td>
<td></td>
</tr>
<tr>
<td>260</td>
<td></td>
<td></td>
</tr>
<tr>
<td>270</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 39. Comparison of OmpP2 and OmpP5 amino acid sequences. In a 79 amino acid region, both proteins have 21% identity. The numbers corresponding to the amino acids denote their position.
could be purified by glutathione-Sepharose 4B under non-denatured conditions (Fig. 37). Purified GST-P2 was assayed by TLC overlay for binding to lipids. As expected, GST-P2 specifically bound to PE from *E. coli* and PE from soybean, but not Gg4, GM1 PC and PG (Fig. 38). GST-P2 also did not bind Gg4 (results not shown). Furthermore, GST-P2 was found to bind to the upper edge of the Gg3 band (Fig. 38), which suggested that GST-P2 might bind to a fast migrating Gg3 species.

OmpP5 is the *Haemophilus* analog of the *E. coli* OmpA. This protein was recently identified as a pilin component of a nonhemagglutinating filament in the NTH1 strains. Since both OmpP2 and P5 were purified simultaneously from the PE column, it is possible that the proteins share some sequence similarity. By comparing protein sequences, a region of 79 amino acids was found in which both proteins have 21.5% identity. This region may be involved in PE binding.
DISCUSSION

Zinc is essential for all organisms because it plays a critical role in the catalytic activity and/or structural stability of many proteins. Despite this importance, very little is known about the mechanisms and regulation of zinc transport in bacteria (Silver & Walderhaug, 1993). Although the few reports suggest that bacteria appear to possess a specific energy-dependent zinc transport system, these studies have been confounded by both the nonspecific binding of zinc to the bacterial surface and the rapid exchange of cellular zinc with zinc in the medium (Bucheder & Broda, 1974; Chipley, 1972). So far, a specific zinc transport mechanism has not been properly demonstrated and thus there has been no means to study the regulation of this metal's transport in any prokaryotes (Silver & Walderhaug, 1992). In this study, it has been demonstrated that an originally designated phosphate-binding protein, the HI0119 gene product, is in fact a periplasmic zinc binding protein (Pzp1) which plays a key role in the zinc uptake of H. influenzae. This is the first description of a protein involved in zinc uptake in prokaryotes.

The *pzp1* isogenic mutant described in this report provides an opportunity to study zinc transport in bacteria using biochemical and genetic methods. Furthermore, purified, functional Pzp1 has been readily isolated in high yield. A CD spectrum of purified Pzp1 suggested that the major secondary structure of the recombinant protein is α-helix. Further structural studies on this protein will give insights as to the precise role Pzp1 plays in zinc processing in *H. influenzae*. The high solubility of Pzp1 means it may be a good candidate for X-ray crystallography or NMR studies to determine the structure of the zinc binding domain.
Since GST-0119 (Pzp1) could not be cleaved by thrombin completely, it is possible that the local secondary structure of the recombinant protein may prevent effective thrombin binding. This may also explain why the 6XHis-0119 (Pzp1) did not bind to the Ni$^{2+}$-NTA agarose column under native conditions. Interestingly, although the protein contains an inherent histidine-rich domain, which might be expected to have a nickel binding function, this did not confer binding of either fusion protein to Ni$^{2+}$-NTA agarose under native conditions. Perhaps, the Asp and Lys residues interspersed within the His residues confer a specificity for zinc binding. However, the ability of Pzp1 to bind other metals in vitro was not assessed.

The ATP binding-protein cassette (ABC) system is involved in the transport of a diverse array of macromolecules across the cytoplasmic membranes of bacteria and eukaryotes (Higgins, 1992). This system consists of three basic parts: one or two ATPases, one or two integral membrane proteins and one substrate-specific binding protein. In gram-negative bacteria the binding protein is soluble and periplasmic, but in gram-positive bacteria the binding protein is lipid-linked to the cytoplasmic membrane. Usually these three components are encoded together in one operon in bacteria (Ames, 1986; Tam & Seier, 1993). Pzp1 of H. influenzae, a periplasmic zinc binding protein, is 23.7% identical and 47.8% similar to FimA of Streptococcus parasanguis. FimA of S. parasanguis is a lipoprotein which is involved in adherence of these bacteria to the salivary pellicle of dental surfaces (Fenno et al., 1989; Oligino & Eves-Taylor, 1993; Burnette-Curley et al., 1995). DNA sequence data showed that the S. parasanguis fimA locus encodes an ATP-binding membrane transport system (Fenno et al., 1995). Interestingly, a fimA isogenic mutant did not display an obvious growth defect in vitro but was found to be less virulent in animal models (Burnette-Curley et al., 1995). The present results suggest that there is functional heterogeneity between Pzp1 of H. influenzae and FimA of S. parasanguis, since Pzp1 has a central
histidine rich domain and a C-terminal disulfide bonded domain which are absent in the FimA protein of *S. parasanguis*. One can speculate that FimA of *S. parasanguis* may be involved in an ATP transport system similar to Pzp1, but has different substrate binding specificity. Furthermore, a recently performed BLAST search showed that Pzp1 has 49.2% identity and 59.4% similarity to an unidentified protein (YebL) in *E. coli*, a 31.1 kDa protein in the msbB-ruvB intergenic region precursor. Compared with the sequence of Pzp1, YebL in *E. coli* seems to have similar domain structure including a central potential metal binding domain of 21 amino acid residues and two conserved cysteines in the C-terminal region which may form a disulfide bond. The *yebL* locus also appears to be organized in an operon. YebL of *E. coli* may thus have similar function to Pzp1 of *H. influenzae*, which is involved in the transport of zinc.

FimA is a member of the lipoprotein receptor antigen groups (Lra1) (Jenkinson, 1994). This family currently consists of six lipoproteins, including FimA from *Streptococcus parasanguis* FW213, ScaA from *S. gordonii* PK448, PsaA from *S. pneumoniae* B36A, SsaB from *S. sanguis*, EfaA from *Enterococcus faecalis* and ScbA from *S. cristal* CC5A. These six members may constitute a group of FimA-related adhesins with functional heterogeneity. ScaA may function as a mediator of coaggregation with human oral actinomyces. SsaB binds both to salivary components and to the actinomyces cells (Ganeshkumar *et al.*, 1993). Mutagenesis of *psaA* suggested that a functional *psaA* gene is essential for full virulence of *S. pneumoniae* (Berry & Paton, 1996). No function has, so far, been assigned to *efaA* and *scbA*. 
Although there is obvious functional heterogeneity, all members of the LraI family have considerable amino acid sequence homology. Sequencing analysis suggested that each member of this family appears to be part of an ABC protein complex. In this study, Pzp1, one of two homologues of FimA in *H. influenzae*, was identified as a key protein for zinc uptake. In addition, a manganese transport protein (MntC) of the cyanobacterium *Synechocystis* 6803 is 30% identical in sequence to ScaA (Bartsevich & Pakrasi, 1995). Based on all these data, it is postulated that there is a large family (Bmt) which is involved in bacterial metal transport. In gram-positive bacteria, this family includes all six members of LraI. They are lipoproteins on the surface of bacteria and may additionally function as adhesins. In gram-negative bacteria, the Bmt family may include Pzp1 and HI0362 protein from *H. influenzae*, MntC from *Synechocystis*, and YebL from *E. coli*. They are localized in the periplasmic region and may be involved in metal binding. It is expected that many more members of the Bmt family will be identified in future.

The upstream region of *pzpl*, HI0117 encodes a small hypothetical protein of 82 amino acids. The deduced product of HI0118 is a putative molybdopterin biosynthesis protein (chLN). The nucleotide sequence downstream of *pzpl* contains two open reading frames with different transcriptional direction to *pzpl*, HI0120 and HI0121, which encode hypothetical proteins (Fleischmann *et al.*, 1995). When the *pzpl* gene and its upstream 590 bp fragment were cloned into vector pCRYI™ in either orientation, recombinant *E. coli* strains harboring either plasmid pTA119a or pTA119b were found to overexpress Pzp1 at a similar level, suggesting that transcription of this protein is initiated at a *H. influenzae* promoter. Overall, these data suggest that the *pzpl* locus does not appear to be part of an operon unlike other members of the Bmt family described above.
The elucidation of the complete DNA sequence of the *H. influenzae* strain Rd genome has greatly facilitated direct comparisons of genes and protein sequences with those of other species (Fleischmann *et al.*, 1995). In the *H. influenzae* Rd genome, two genes, HI0119 and HI0362, were identified as putative adhesin B precursors on the basis of homology to the adhesin B (FimA) of *S. parasanguis* (Fleischmann *et al.*, 1995). In this study, the HI0119 gene product was demonstrated to be a periplasmic zinc binding protein (Pzp1). Harkness *et al.* identified two iron-repressed periplasmic proteins in *H. influenzae*, with molecular weight of 40 and 31 kDa respectively (Harkness *et al.*, 1992). The N-terminal sequence of the 40 kDa protein had 81% similarity to the N terminus of Fbp, the major iron-binding protein of *N. gonorrhoeae* and *N. meningitidis*. The N-terminal sequence of the 31 kDa protein was also determined. This sequence showed no homology with any known protein sequence (Harkness *et al.*, 1992). Recently, it was noticed that the N-terminal sequence of the 31 kDa protein had 100% identity to the N terminus of the HI0362 gene product. Additionally, the gene HI0362 locus seems to have similar genetic organization to typical ABC systems, in which two upstream genes, HI0360 and HI0361, appear to encode a hydrophobic membrane protein and an iron(III) dicitrate transport ATP-binding protein, respectively (Fleischmann *et al.*, 1995). These data suggest that the HI0362 gene product appears to localize in the periplasmic space and may be involved in iron or iron compound transport in *H. influenzae*.

In general, there is little sequence conservation between the binding proteins for different substrates among ABC transport systems (Higgins, 1992). However, for some binding-protein-dependent transport systems, two distinct binding proteins with different substrate specificities interact with the same inner membrane transport complex. The genes encoding the two alternative binding proteins are usually related. For example, the *E. coli* thiosulfate- and sulfate-binding proteins are
highly similar, but encoded in different transcriptional units. They share common membrane components (Hryniewicz et al., 1990). In the *H. influenzae* genome, the HI0362 gene product has 21% identity to Pzp1, and the gene HI0362 locus seems to encode an ABC transport system (Fleischmann et al., 1995). Even though the *pzpl* locus does not appear to be part of an operon as is usually found for comparable transporters, we expect that Pzp1 and the product of the HI0362 gene may interact with the same core transmembrane complex.

Our results demonstrate that the *pzpl* mutant cannot grow under aerobic conditions and grows poorly under anaerobic conditions. Only zinc can suppress the growth defects of the *pzpl* mutant. This suggests that the metabolic process under aerobic conditions may be more dependent on zinc than that under anaerobic conditions in *H. influenzae*. Alternatively, there may be an additional, lower affinity zinc transport system operating when *H. influenzae* grows under anaerobic conditions. Western blotting has shown that the expression level of Pzp1 was not found to be decreased during anaerobic growth (results not shown), supporting the former explanation.

Pzp1 of *H. influenzae* contains an unusual histidine-rich domain of about 47 amino acids. This domain is extremely rich in potentially metal-binding amino acids, including 23 histidines, 10 aspartic acids and 6 glutamic acids. We expect that the zinc binding sites may be located at this domain. Some studies have demonstrated that HXXXH motif form a high affinity binding site for Cu$^{2+}$, Ni$^{2+}$ and Zn$^{2+}$ (Arnold & Haymore, 1991; Arnold, 1991). In fact there are 7 HXXXH motifs in this metal-binding domain. Our results showed that purified Pzp1 contained an average of 1.6-1.9 zinc atoms per protein molecule, although this protein has potential to bind up to five zinc atoms. This finding is consistent with a zinc
accumulation and transport role, since it would be expected that the Pzp1 population would contain molecules at various stages of substrate delivery to the membrane-bound component of the transporter. Future structural studies of Pzp1 in the fully zinc bound or apo-forms might reveal zinc-dependent conformational changes.

DsbA is a periplasmic E. coli protein which acts as a thiol-disulfide oxidase, involved in the formation of protein disulfide bonds (Bardwell et al., 1991). Rensing et al. isolated one mutant which was sensitive to Cd$^{2+}$ and Zn$^{2+}$ (Rensing et al., 1997a). This mutant had a single TnphoA insertion in the dsbA gene. The toxic effects of Cd$^{2+}$ and Zn$^{2+}$ in the mutant could be due to the binding of these metals to the free thiols of periplasmic proteins which are normally oxidized by the dsbA gene product. Another possibility is that the Cd$^{2+}$ or Zn$^{2+}$ specific transporter cannot fold properly without DsbA, thus making the cells hypersensitive to cadmium and zinc. The present results show that YebL of E. coli, a homolog of Pzp1, contains two conserved cysteine residues near the carboxyl terminus, which may form disulfide bond like Pzp1. They suggest that DsbA may be needed for the formation of the YebL disulfide bond. It is therefore possible that the disulfide bond of YebL cannot be formed in the dsbA mutant, which fails to regulate properly the zinc uptake rate, thus resulting in the toxic effect of zinc. If it is true, it suggests that the disulfide bond of YebL may be implicated in regulating the zinc uptake rate in E. coli.

A DsbA homologue, Por (45% identical to the E. coli DsbA protein), was identified in H. influenzae. Mutants in the Por-encoding gene impaired competence-induced DNA uptake in this bacterium presumably because this process involves outer membrane proteins with disulfide bonds in their active state (Tomb,
1992). It would be interesting to examine whether the por mutant of *H. influenzae* is more sensitive to zinc than the wild type strain.

Although Pzp1 was originally purified from a surface (water) extract, and inferred to be an adhesin, these results indicate that Pzp1 in *H. influenzae* is a periplasmic protein, and thus probably does not directly contribute to adhesion of *H. influenzae*. Water extraction of bacteria has proven to be an effective method of solubilizing surface proteins for some bacteria. However, Harkeness et al. have noticed that some nontypeable *H. influenzae* strains were lysed when suspended in water (Harkeness et al., 1992). This may explain how Pzp1 was isolated from a water extract of *H. influenzae*. Thus, caution should be taken when surface proteins are isolated using water extraction of bacteria; this technique may enrich for surface structures but is not proof of surface localization.

Some highly specific ion transport systems were known to be regulated at the gene level. That is, additional genes encode trans-acting proteins and cis-acting DNA sites which govern the level of transcriptional mRNA synthesis. For example, regulation of iron transport is frequently accomplished by the single iron-binding protein Fur. In order to explore the regulation of *pzp1* expression, *H. influenzae* was grown at different conditions, including temperature, atmosphere, salt concentration, supplemental zinc, iron-limitation, metal chelation (EDTA). However, no obvious change in the level of *pzp1* expression was observed. Due to the extreme difficulty in making a zinc-deficient medium, expression of *pzp1* that responds to zinc starvation would not be observed. It is possible that increased *pzp1* expression could be induced in response to zinc-limited growth conditions.
In the present study, sequence analysis of the mature Pzp1 from a nontypeable strain, HI6564, revealed a 97.7% identity to the sequence of *H. influenzae* Rd. Western blotting and PCR analysis indicated that the *pzpl* gene is highly conserved and its protein product is generally expressed in all 23 *H. influenzae* clinical strains tested. In addition, it was observed that no DNA fragment was amplified from *E. coli* TOP10 or HB101 strains and *Neisseria meningitidis* clinical isolates by primers (HIMA1 & HIMA2) which were used for PCR analysis of the *H. influenzae* *pzpl* gene, suggesting that *pzpl* is species specific. These data support the possibility for developing a quick technique for diagnosis of *H. influenzae* infection or species identification based on the *pzpl* sequence.

Celite is a natural product derived from fossil deposits of diatoms. The main chemical constituent in celite is silica (SiO₂), accounting for 90% by weight of the material. Other inorganic oxides have been found to be present in smaller amounts including Al₂O₃, Fe₂O₃, TiO₂, CaO, MgO (Table 3). The physiological significance of celite-binding is not clear. However, mutational analysis of the fusion proteins indicates that the C-terminus forms a disulfide-bonded domain which does not participate in celite binding. Though the purified FimA has ability to bind saliva-coated hydroxyapatite (SCHA) *in vitro*, it is not clear what the active binding portion of FimA in SCHA (Oligino & Fives-Taylor, 1993). We would speculate, that since dental calculus contains silicic acid, silica and minerals (Hidaka, 1993), the property of celite binding may be more relevant to the adhesin function of FimA of *Streptococcus parasanguis*, which is surface-expressed.

Recent studies showed that zinc uptake in the yeast *Saccharomyces cerevisiae* is transporter-mediated by at least two systems, one with high affinity and a second with lower affinity. The transporters which are responsible for both uptake systems
have been identified because of their significant similarity to Irt1, an Fe(II) transporter gene from the plant *Arabidopsis thaliana* (Eide et al., 1996). The zrt1 gene encodes the transporter protein of the high affinity system (Zhao, 1996a), whereas the zrt2 gene encodes the transporter of the low-affinity system (Zhao, 1996b). Based on protein sequences, Zrt1 and Zrt2 were predicted to be integral membrane proteins containing eight potential transmembrane domains. However, the zrt1/zrt2 double mutant is viable, indicating the existence of additional zinc uptake pathways (Zhao, 1996b). In this study, it is demonstrated that Pzp1 is a highly soluble periplasmic zinc binding protein. Our results suggest that, unlike in yeast, there may be only one zinc uptake system in *H. influenzae*.

Many bacterial pathogens, including *Pseudomonas aeruginosa*, *Chlamydia trachomatis*, *EPEC*, *H. pylori* and *H. influenzae* can specifically bind PE (Lingwood, 1991; Lingwood, 1992; Lingwood, 1993; Krivan, 1991; Busse et al; 1997). Adhesion of bacteria to epithelial cells correlated with the quantity of PE present, while binding to PE by TLC overlay was influenced by differences in the fatty acid composition of the phospholipid receptor (Gold et al., 1995). In various eukaryotic cells, aminophospholipids (PE and PS) are asymmetrically distributed in the plasma membrane; most PE and essentially all PS were found in the inner leaflet. In the early stages of apoptosis, PS is translocated from the inner side of the plasma membrane to the outer layer, which allows phagocytes to recognize and engulf the apoptotic cells (Fadoc et al., 1992). A recent study has shown that PE is also exposed on the cell surface in the early phase of apoptosis (Emoto et al., 1997). Apoptosis plays an important role in both respiratory and gastrointestinal epithelial turnover (Hall et al., 1994). Bacteria which bind PE may therefore preferentially bind apoptotic cells. Moreover, microbial infection might induce apoptosis which could thus amplify attachment. Preferential binding to apoptotic cells might improve bacterial...
nutrient access (Chen & Zychlinski, 1994).

*H. influenzae* is a strictly human pathogen which colonizes the respiratory tract, where it resides as a commensal or give rise to both localized and serious systemic infections (Moxon, 1989; Turk, 1984). The attachment of bacteria to host cells is considered to be the initial event in the infectious process. Adherence of *H. influenzae* to epithelial cells is promoted by pilus and nonpilus adhesins. In human and animal organ culture models, *H. influenzae* infection caused patchy and occasionally confluent damage to the epithelium, and the bacteria associated only with structurally damaged cells (Wilson et al., 1992). It is possible that *H. influenzae* infection and/or other causes may induce apoptosis of epithelial cells, resulting in the increased surface exposure of PE and epithelial damage. It is reasonable to hypothesize that the bacteria were closely associated with damaged epithelium by PE-binding adhesin(s). In this study, OmpP2 and P5 were identified as possible PE-binding adhesins, which may be involved in adherence of *H. influenzae* during initial colonization of damaged epithelium.

Several lines of evidence indicated that OmpP2 is a PE-binding protein. First, OmpP2 is the most prominent of the 1 M Tris (pH 11.2) PE column eluate. Second, OmpP2 was recognized by an antiserum against the PE-binding adhesin preparation, which inhibited the binding of *H. influenzae* to PE. Third, purified GST-P2 can specifically bind PE from *E. coli* and PE from soybean by TLC overlay. Finally, N terminal amino acid sequencing showed that both Pzpl and OmpP2, in fact, existed in the original PE-binding adhesin preparation, even though OmpP2 was a minor component.
OmpP2 has been shown to function as a porin in *H. influenzae*. Based on parameters of hydrophilicity and amphiphilicity a model for OmpP2 was generated. The model proposed sixteen membrane spanning β-strands connected by eight long loops on one side and eight short β-turns on the other side. Two surface-exposed regions have been mapped to loops L4 and L8 (Srikumar *et al.*, 1992). Duim *et al.* reported antigenic variation in OmpP2 from serial isolates of the same nontypeable strains obtained from patients with chronic obstructive pulmonary disease (Duim *et al.*, 1994). The sequence changes which result in these antigenic changes were localized to L6, suggesting that a portion of L6 is also at the cell surface. Recently, bactericidal epitopes of OmpP2 were identified on two different surface-exposed loops L5 and L8 (Haase *et al.*, 1994; Yi & Murphy, 1997). In this study, we noticed that OmpP2 and P5 had 21.5% identity in a region with 79 amino acids. Since both OmpP2 and P5 were simultaneously purified from a PE column, this region may be involved in PE binding or P5 may bind to P2. In the OmpP2 model, this region corresponds to extracellular loops L6 and L7. Further study is needed to determine whether loops L6 and L7 of OmpP2 are involved in PE binding.

Why was Pzp1 purified from the PE affinity matrix as a major protein band? The current data show that Pzp1 and OmpP2 have similar molecular weight. When a water extract containing both proteins was applied to a PE-celite affinity column, both proteins were retained in the column, each binding to its ligand. Previous adhesion studies showed *H. pylori* binding to PE by TLC overlay was ion-dependent, being inhibited by EDTA. Therefore, initial attempts to isolate a PE-binding adhesin from *H. influenzae* by affinity chromatography employed an EDTA gradient (6 mM to 10 mM) to elute bound proteins. The EDTA eluate primarily consisted of Pzp1. When the column was eluted with 1 M Tris (pH 11.2), all bound proteins including Pzp1 and OmpP2 were denatured and separated with their ligand. OmpP2 was
observed as the most prominent band on SDS-PAGE. However, given the similar migration of Parp1 and OmpP2 the selective elution of these two proteins was not initially appreciated. The strong eluant, 1 M Tris, was used subsequently as it was felt that EDTA was a weak eluant of the "adhesin". Nevertheless, despite the initial confusion, the newly described functions of two proteins, Parp1 and OmpP2, have been clearly identified.
SUMMARY AND FUTURE DIRECTIONS

Summary

1. Although the zinc requirement for bacterial growth was demonstrated fifty years ago, a specific zinc transport mechanism has not been properly shown. In this study, Pzp1 was identified as a key protein for zinc uptake in *H. influenzae*. This is the first description of a protein involved in zinc uptake in prokaryotes.

2. In the *H. influenzae* Rd genome, two genes, HI0119 and HI0362, were identified as putative adhesin B precursors on the basis of homology to the adhesin B (FimA) of *Streptococcus parasanguis*. Moreover, our previous work suggested that HI0119 may be a PE-binding adhesin. In this study, the HI0119 gene was cloned and sequenced. The HI0119 protein was expressed as fusions; GST-0119, His-0119, and as recombinant protein in *E. coli*. The proteins were purified and further characterized. Our results indicated that HI0119 (Pzp1) is unlikely to function as an adhesin. This protein is unable to bind to PE, but specifically binds to celite matrix in the absence of PE.

3. Pzp1 mutants lacking one or both cysteine residues were used to confirm that Pzp1 has a disulfide bond in the C-terminal domain. The disulfide bond and C terminus of this protein are not involved in celite binding. A CD spectrum of recombinant Pzp1 suggested that Pzp1 is highly helical.

4. DNA sequence analysis of the mature Pzp1 from a nontypeable strain, HI6564, revealed a 97% identity to the sequence of *H. influenzae* Rd. PCR and Western blotting analysis of clinical isolates indicated that the *pzp1* gene is highly conserved.
and its protein product is generally expressed in *H. influenzae*.

5. Since Pzp1 was originally purified from a water extract, it was postulated to be localized on the surface of *H. influenzae*. This study demonstrated that Pzp1 is neither localized to the outer membrane nor secreted to growth culture. Pzp1 is in fact a periplasmic protein.

6. Using neutron activation analysis, atomic absorption spectroscopy and $^{65}$Zinc blotting, Pzp1 has been identified as a zinc binding protein. Recombinant Pzp1 contained about 2 zinc atoms per protein molecule. The zinc atoms could be removed *by incubation with EDTA and by further addition of zinc*, a total of 5 zinc atoms per Pzp1 could be bound.

7. An isogenic *pzp1* deficient mutant has been constructed. This mutant cannot grow under aerobic conditions and grows poorly under anaerobic conditions, revealing an absolute requirement for the aerobic growth, and some importance for anaerobic growth of *H. influenzae*.

8. Using PE affinity chromatography, OmpP2 and OmpP5 of *H. influenzae* were identified as PE binding proteins, and possible adhesins. The proteins share a 21% identity in a 79 amino acid region. Furthermore, OmpP2 was expressed as a GST fusion protein. Purified GST-P2 can specifically bind PE, as monitored by TLC overlay.

Future Directions

Our pioneer study has produced a series of important discoveries and resulted
in identification of the function of an unknown protein, HI0119. This study opened an avenue of research in further understanding the function and structure of Pzp1 and its related protein family. The following direction could be considered.

1. Pzp1 is a periplasmic zinc binding protein which plays a key role for zinc uptake in H. influenzae. Although the present data strongly suggests that Pzp1 is involved in zinc transport, direct evidence is required to confirm its zinc transport function. A highly sensitive zinc uptake assay should be established to detect the role of Pzp1 and the other related proteins in zinc uptake.

2. The present study suggests that YebL, an E. coli homolog of Pzp1, may have a similar function to Pzp1. Thus, the cloning and expression of YebL, and construction of a yebL isogenic mutant will be useful to confirm the function of E. coli YebL.

3. Based on the present study, we postulated that there is a large protein family, termed Bmt, which is involved in bacterial metal transport. In order to examine if the members of the putative Bmt family are involved in metal transport in bacteria, it might be useful to measure the metal uptake in isogenic mutants of each member and wild type strain.

4. Although the pzp1 gene does not appear to be part of an operon, we expect that Pzp1 and the product of HI0362 gene may interact with the same core transmembrane complex. It is obviously difficult to obtain direct evidence of this interaction since the membrane-bound components of periplasmic permeases are generally present in small amounts, they are more difficult to isolate (Kerppola et al., 1991). Nevertheless, cross linking and immunoprecipitation analysis might provide useful information if the HI0362 operon were expressed and antibodies
specific for each of these proteins were developed.

5. It is likely that the zinc binding site of Pzp1 may reside in the histidine rich region. Mutational analysis of this region will be important to clarify the role of this region in zinc binding.

6. This study showed that Pzp1 contains a disulfide bond in the C terminal domain. Since a dsbA mutant became sensitive to zinc, we speculate that the disulfide bond of periplasmic binding proteins may be implicated in regulating zinc uptake. A construction of a Pzp1 disulfide bond mutant strain will provide a direct evidence for the role of the disulfide bond of Pzp1.

7. This study has indicated that OmpP2 and possibly OmpP5 are PE binding proteins. However, it is not clear whether their PE binding abilities are involved in H. influenzae adherence to cells. Further study should focus on examining whether OmpP2 and OmpP5 are implicated in bacterial adhesion, and if this is true, whether their functions in adherence are associated with PE binding abilities.
REFERENCES


Adams WG, Deaver KA, Cochi SL, Plikaytis BD, Zell ER, Broome CV, Wenger JD. Decline of childhood Haemophilus influenzae type b (Hib) disease in the Hib vaccine era. JAMA 1993; 269:221-226.


Anderson RN, Ganeshkumar N, Kolenbrander PE. Cloning of the Streptococcus gordonii PK488 gene, encoding an adhesin which mediates coaggregation with


Barenkamp SJ, Bodor FF. Development of serum bactericidal activity following


Boulanger J, Huesca M, Arab S, Lingwood CA. Universal method for the facile production of glycolipid/lipid matrices for the affinity purification of binding


Burnette-Curley D, Wells V, Viscount H, Munro CL, Fenno JC, Fives-Taylor P,


Conklin DS, McMaster JA, Culbertson MR, Kung C. COT1, a gene involved in cobalt


Cunnane SC. Zinc: clinical and biochemical significance. CRC Press, Boca Raton, FL.


Gray-Owen SD, Loosmore S, Schryvers AB. Identification and characterization of genes encoding the human transferrin-binding proteins from *Haemophilus*


Hansson GC, Karlsson KA, Larson G, Stromberg N, Thurin J. Carbohydrate-specific


Hidaka S, Okamoto Y, Abe K. Possible regulatory role of silicic acid, silica and clay minerals in the formation of calcium phosphate precipitates. Archs Oral Biol 1993; 38:405-413.


Krivan HC, Ginsburg V, Roberts DD. Pseudomonas aeruginosa and Pseudomonas 
*cepacia* isolated from cystic fibrosis patients bind specifically to gangliotetraosylceramide (asialo GM1) and gangliotriosylceramide (asialo GM2). Arch Biochem Biophys 1988a; 260:493-496.


Krivan HC, Roberts DD, Ginsburg V. Many pulmonary pathogenic bacteria bind specifically to the carbohydrate sequence GalNAcβ1-4Gal found in some glycolipids. Pro Natl Acad Sci USA 1988c; 85:6157-6161.


Kroll JS, Loynds BM, Moxon ER. The *Haemophilus influenzae* capsulation gene...


Lee KY, Weinberg ED. Sporulation of Bacillus megaterium: roles of metals ions.


Lingwood CA, Nutikka A. A novel chemical procedure for the selective removal of nonreducing terminal N-acetyhexosamine residues from glycolipids. Anal
Biochem 1994; 217:119-123.


Uropathogenic *Escherichia coli* can express serologically identical pili of different receptor binding specificities. Mol Microbiol 1988; 2:255-263.


Munson R. Jr., Brodeur B, Chong P, Grass S, Martin D, Proulx C. Outer membrane


Reddy MS, Berenstein JM, Murphy TF, Faden HS. Binding between outer membrane proteins of nontypeable *Haemophilus influenzae* and human


Rubin LG, St. Geme III JW. Role of lipooligosaccharide in virulence of the Brazilian


Sharon N, Lis H. Carbohydrates in cell recognition. Scientific American 1993; 268:82-


Smirnov MD, Esmon CT. Phosphatidylethanolamine incorporation into vesicles


Steinhoff MC. Haemophilus influenzae type b infections are preventable everywhere. Lancet 1997; 349:1186-1187.

St. Geme JW, III, Falkow S. Capsule loss by Haemophilus influenzae type b results in enhanced adherence to and entry into human cells. J Infect Dis 1992; 165(suppl 1):


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


Weber A, Harris, Lohrke S, Forney L, Smith AL. Inability to express fimbriae results


White JV, Stultz CM, Smith TF. Protein classification by stochastic modeling and


Zwahlen AL, Rubin G, Moxon ER. Contribution of lipopolysaccharide to pathogenicity of Haemophilus influenzae: comparative virulence of genetically-