Master's thesis in biotechnology, BIO-3941

Purification of the recombinantly produced *Helicobacter pylori* antigens NAP and Flagellin A

Performed at University of Gothenburg

June 2008

By Ingrid Lea

Institute of Medical Biology, Faculty of Medicine
University of Tromsø

Acknowledgements

This study was an external study performed at the laboratory of Prof. Jan Holmgren at the Department of Microbiology and Immunology, Sahlgrenska Academy, University of Gothenburg in Sweden. The work was conducted from January to May 2008.

First of all I would like to thank my excellent supervisor Dr Carl-Fredrik Flach for all his inspiration and encouragement, his time and engagement, and for sharing his wide knowledge with me. I also want to thank all the students, researchers and employers at the department for making me feel welcome and providing guidance when needed. The members of the *H. pylori* group and my colleagues on “våning 5” have given me insight into many interesting scientific (and non-scientific ☺) subjects, creating a nice working environment.

Secondly, I am grateful for the support and advice from my supervisor at the University of Tromsø, Ass. Prof. Johanna Ericsson Sollid.

At last I wish to thank my family and friends for their help and support throughout five years of studies. Thank you Anders, for your encouragement, patience, and love. ♥

Gothenburg June 2008, Ingrid Lea

Summary

Helicobacter pylori is a Gram-negative flagellated bacterium that infects the stomach of around 50 % of the worlds population, and can give rise to serious disease such as chronic gastritis, duodenal ulcer and stomach cancer. The current treatment against *H. pylori,* consisting of a combination of a proton pump inhibitor and two different antibiotics, is effective but has several drawbacks such as increased risk of antibiotic resistance and no protection against re-infection. Although extensive research is going on to develop a vaccine that would be an attractive alternative or complement to the current treatment, there is no vaccine available against *H. pylori* today. However, in experimental models, oral vaccination with selected *H. pylori* antigens given together with cholera toxin as adjuvant can give rise to protective immunity, both when the vaccine is given before infection and when administered to already infected animals. Antibodies to a number of antigens have been detected in the sera of infected patients, and many of these are known to be virulence factors, and are considered as candidate *H. pylori* vaccine antigens, alone or in a combination. Two such antigens are Neutrophil-activating protein (NAP) and Flagellin A (FlaA). The *H. pylori* NAP, localised in the bacterial cytosol, is known to stimulate human neutrophils. *H. pylori* carry 5-7 flagella that provide motility required for colonisation and infection. *H. pylori* FlaA is the major protein and structural subunit of the flagellar filament. Both NAP and flagella preparations induce protection against *H. pylori* infection. Hence, the aim of this thesis work was to purify NAP and FlaA to be used for vaccine development purposes. Since preparations enriched by a specific protein are not easily obtained from natural host cells, recombinant protein production is a frequently used procedure. *Escherichia coli* facilitates protein expression by its inexpensive and fast high-density cultivation, its relative simplicity and the well-known genetics.

In this project the *H. pylori* antigens NAP and FlaA were cloned and expressed in *E. coli,* and a purification strategy were set up for the two proteins. The *flaA* and *napA* genes from *H. pylori* strain SS1 were amplified by high fidelity PCR. After TA cloning, the two genes were sequenced, and inserted into the expression vector pML-λP_L, relying on the heat-induced promoter λP_L. The recombinant expression vectors were transformed into *E. coli* N4830I host cells, and the induced expression of NAP and FlaA protein was detected on SDS-gels and immunoblots. NAP protein produced in heat-induced N4830I-pML-λP_L-*napA* was purified through several steps, including ammonium sulphate precipitation, anion exchange chromatography and gel filtration. The FlaA protein was produced as inclusion bodies in heat-induced N4830I-pML-λP_L-*flaA*. In conclusion, NAP was purified with good quality (high purity), although the yield was limited. Despite that several strategies were tested in the attempt to purify FlaA, the protein could not be purified and solubilized in its native form. Problems with polymerisation and precipitation suggested that alternative strategies should be considered for the future expression and purification of FlaA. Further characterisation and optimisation of the protein expression are required to develop highly effective protein purification protocols for NAP and FlaA.

Contents

Acknowledgements **3**

Summary **5**

Abbreviations **9**

Introduduction **10**

History *10*

Helicobacter pylori *10*

Neutrophil-activating protein *11*

Flagellin A *11*

Recombinant protein expression *12*

Escherichia coli expression systems *12*

Inclusion bodies *12*

Protein purification *13*

Ion exchange (IEX) chromatography *13*

Gel filtration (GF) *13*

Background for this study *14*

Aims of the thesis **15**

Materials and methods **16**

Materials *16*

Buffers and solutions *16*

Cells culture, vector, and growth media *16*

Methods *16*

Cloning of the H. pylori napA and flaA genes *16*

Polymerase chain reaction *16*

Agarose gel electrophoresis *16*

TOPO TA cloning *16*

Transformation of competent bacteria *16*

Plasmid Miniprep and Determination of DNA concentration..........................*16*

DNA Sequencing..*16*

DNA digestion by restriction enzymes..*16*

Ligation of flaA /napA fragments into pML-λPL vector...................................*17*

Purification of PCR product from agarose gel...*17*

Expression of NAP and FlaA..*17*

Cell cultures: growth, induction and fractioning..*17*

SDS-PAGE...*17*

Western blotting...*17*

Periplasmic preparation...*18*

Protein purification...*18*

Preparation of inclusion bodies... *18*

Ammonium sulphate precipitation...*18*

Anion exchange chromatography.. *18*

Gel filtration..*19*

Determination of protein concentration.. *19*

Results..**22**

Cloning of NAP and FlaA..*22*

Characterisation of NAP and FlaA expression..*23*

Attempts to purify FlaA.. *24*

Discussion.. **28**

Expression and purification of NAP..*28*

Expression and purification of FlaA...*29*

Conclusion..**31**

References..**32**

Appendix - ***A1:*** *Buffers and solutions*
A2: *SDS-PAGE*
A3: *agarose gel elecrophoresis / expression vector*

Abbreviations

Amp	Ampicillin
bp	Base pairs
BSA	Bovine Serum Albumin
Conc.	Concentration
dH_2O	destilled water
DMI	Department of Microbiology and Immunology
dNTP	deoxyribonucleoside triphosphate
E. coli	*Escherichia coli*
EDTA	ethylene diamine tetraacetic acid
EtBr	Etidium bromide
EtOH	Ethanol
FlaA	Flagellin subunit A
h	hour(s)
HAC	Acetic Acid
H. pylori	*Helicobacter pylori*
Ig	Immunoglobulin
IPTG	Isopropylthiogalactoside
Kb	kilo base
kDa	kilo Dalton
LB	Luria-Bertani broth
mA	milli-Ampere
mAb	monoclonal antibody
mM	milli-Molar
NAP	Neutrophile-activating protein
nm	nano meter
OD	Optical density
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
SDS-PAGE	Sodium Dodecyl Sulphate – polyacrylamide gel electrophoresis
SS1	Sydney strain 1
TBE	Tris-borate-EDTA
TBS	Tris Buffered Saline
Å	Ångström, 1 Å=0.1 nm

Introduduction

History

For a long time, the human stomach was considered to be a sterile organ where no microorganisms could live due to the harsh acidic conditions. This changed in 1982, when Warren and Marshall were able to culture *Helicobacter pylori* bacteria from gastric biopsies, and found that the bacteria were present in patients with active chronic gastritis, duodenal ulcer or gastric ulcer (Marshall and Warren, 1984). They were later awarded the Nobel Prize (in 2005) for their discovery of *H. pylori* and its role in gastritis and peptic ulcer disease. The possibility that peptic ulcers could be caused by a bacterium, and not stress, spicy food or other factors was a surprise to the scientific community, and it was difficult for Warren and Marshall to change the prevailing dogma. To convince colleagues and the public, Barry Marshall drank a suspension of bacterium and proved Koch´s postulate for gastritis. (Marshall, *et al.,* 1985) However, it was not until the early 1990s that is was recognized that *H. pylori* causes peptic ulcer disease. During that period, several researchers confirmed that *H. pylori* eradication cured peptic ulcer (Coghlan *et al*. 1987, Rauws and Tytgat, 1990, Graham *et al*. 1991).

In retrospect, it is interesting to notice that there were many references to the presence of *H. pylori* in gastric mucosa before its culture in 1982. Spiral-shaped bacteria were noted many times in the literature, but their presence was not properly correlated with gastro-duodenal disease (Marshall, 2002, Mobley *et al*. 2001). One of the first well known reports of gastric helicobacters was done by Bizzozero in Turin in 1893 (Bizzozero, 1893).

The bacterium was initially named *Campylobacter pyloridis*, then *C. pylori* (after correction to Latin grammar). In 1989, after DNA sequencing and other data had shown that the bacterium did not belong to the *Campylobacter* genus, it was placed in its own genus, *Helicobacter*. The name *pylōri* means "of the pylorus" or pyloric valve (the circular opening leading from the stomach into the duodenum), from the Greek word gatekeeper.(Mobely et al. 2001). Presently, over 20.000 articles have been published on "helicobacter", not counting the articles under the previous classification of "Campylobacter".

Helicobacter pylori

Helicobacter pylori is a spiral-shaped Gram-negative flagellated bacterium that infects the gastric mucosa of more than half of the world's population, making it the most prevalent of all bacterial infections. The prevalence in developing countries can be as high as 80-90%, whereas it is lower in industrialized countries, ranging between 10-50 % (Rothenbacker and Brenner, 2003). *H. pylori* is an important cause of chronic gastritis, peptic ulcer disease and gastric cancer (Goodwin, 1997). The natural progression of *H. pylori* infection is presented in fig. 1. Infection usually occurs during childhood and causes symptomatic acute gastritis in most patients and persists for decades or life-long. The infection can take multiple courses. Most people infected with *H. pylori* will never develop symptomatic disease. 10-15 % will develop peptic ulcer disease (gastric or duodenal ulcers), approximately 1% will develop gastric adenocarsinoma, and a small group of patients will develop gastric MALT lymphoma (Sauerbaum and Josenhans, 2007).

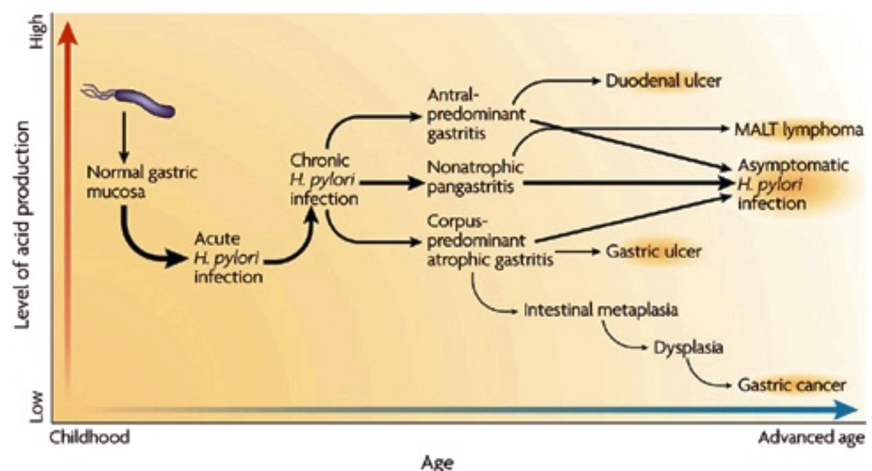

Fig. 1. Natural progression of *H. pylori* infection (adapted from figure 1, Sauerbaum and Josenhans, 2007).

The current treatment against *Helicobacter pylori* infection consists of a combination of two antibiotics and a proton pump inhibitor. The major drawbacks to this therapy are high cost, poor patient compliance and risk of developing antibiotic resistance. Furthermore, such treatment does not protect against reinfection. Extensive research is going on to develop a vaccine against *H. pylori* that will have (a) prophylactic use to prevent infection and/or (b) therapeutic use to eradicate an ongoing infection. A prophylactic vaccine would primarily be useful in young children in high endemic areas, whereas a therapeutic vaccine may be the most relevant one for treatment for those that are already infected (Svennerholm and Lundgren, 2006). Antibodies to a number of antigens such as UreB, VacA, CagA, HpaA, NAP, FlaA and FlaB havebeen detected in the sera of infected patiens, and many of these virulence factors are considered as candidate vaccine antigens, alone or in a combination. In experimental models, it has been shown that oral vaccination with selected *H. pylori* antigens given together with cholera toxin as adjuvant can give rise to protective immunity, both when the vaccine is given before infection and when administered to already infected animals (Nyström and Svennerholm 2007, Kabir., 2007).

Neutrophil-activating protein

The *H. pylori* neutrophil-activating protein (NAP) has been shown to be highly immunogenic in mice and humans (Satin *et al.*, 2000), and is suggested to play a central role in the accumulation of neutrophils at the site of infection. NAP has been shown to be capable of binding iron *in vitro*, increase the adhesion of neutrophils to endothelial cells, and induce migration and activation of human neutrophils and monocytes (summarized by Zanotti *et al.*, 2002). In addition, *H. pylori* NAP is described as a key factor driving T helper (Th) 1 inflammation in the *H. pylori* infection (D'Elios *et al.,* 2007), and has been shown to be a protective antigen in an *H. pylori* infection model (Satin *et al.,* 2000). Previous spectroscopic and electron microscope studies suggested that NAP was comparable with a four-helix bundle protein oligomerizing to form a dodecamer structurally similar to the *E. coli* DNA-binding protein Dps (Grant *et al.*, 1998, Tonello *et a*l., 1999). In 2002, the three-dimensional atomic structure of NAP was resolved, confirming that NAP is a dodecameric protein, about 90 Å in diameter, with 32 symmetry. The arrangement of the twelve 17 kDa monomers give rise to a nearly spherical shell, with an internal cavity for iron accumulation. (Zanotti *et al.*, 2002). The structure of *H. pylori* NAP is shown in fig. 2.

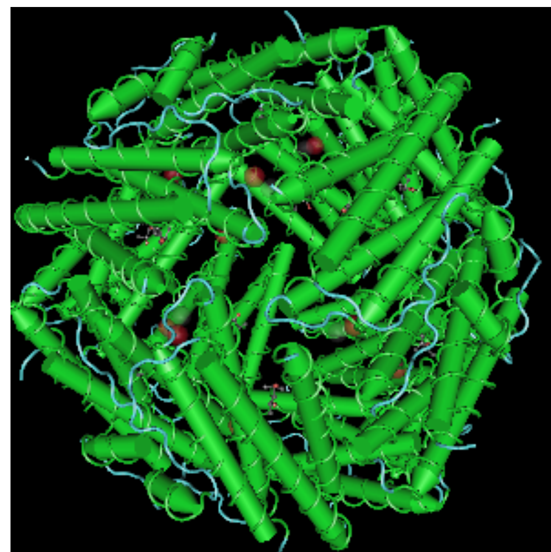

Fig. 2. Structure of H. pylori Neutrophil-activating protein (Zanotti *et al.,* 2002).

Flagellin A

The bacterial flagellum is both a motor organelle and a protein export/assembly apparatus that extends from the cytoplasm to the cell exterior. The flagellum contains three structural elements; the export/switch ring complex, the basal body (rod, rings and hook), and the filament. The filament is a long, thin cylindrical structure that is helical in shape, and therefore when rotated functions like a propeller (Macnab., 2003). *H. pylori* carries 5-7 flagella that provide motility required for colonization and infection. The flagellum is covered with a sheath continuous with the bacteria outer membrane, and contains a complex filament that is composed of two distinctly different flagellin subunits; FlaA (53 kDa) and FlaB (54 kDa). Geis *et al.* first

published a report on the purification of FlaA, which is the major subunit of the flagellar filament, building up the central and distal parts of the *H. pylori* flagellum. Kostrzynska *et al.* demonstrated that the filament also contains a small amount of a second flagellin subunit (FlaB) that seems to be located mainly at the proximal parts of the filament (Geis *et al.*, 1989, Kostrzynska *et al.,* 1991). Since the motility of *H. pylori* is an important virulence factor, it has been suggested that FlaA can be a potential antigen candidate for a *H. pylori* vaccine (Kabir., 2007). The structure of H. pylori has not been resolved, but the structure of Salmonella FlaA is presented in figure 3.

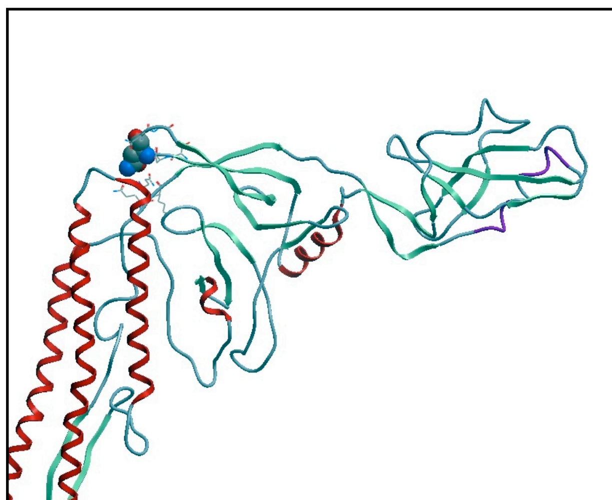

Fig. 3. Structure of Salmonella Flagellin A (Lars Brive, unpublished)

Recombinant protein expression

Escherichia coli expression systems

Because of the vast fund of knowledge about its genetics, biochemistry, and molecular biology, *Escherichia coli* is the system of first choice for expression of many heterologous proteins. Genetic manipulations are straight-forward, cultures of *E. coli* are easily and inexpensively grown, and many foreign proteins are well-tolerated and may be expressed at high levels (Sambrook and Russel, 2001). *E. coli* expression systems can be used for the production of recombinant proteins either intracellularly or extra-cellularly. Recombinantly expressed proteins can in principle be directed to four different locations namely the cytoplasm, the periplasm, the membrane fraction or the cultivation medium. Recombinant expression plasmids require a strong transcriptional promoter to control high-level gene expression. Promoter induction is either thermal or chemical and the most common inducer is the sugar molecule IPTG. T7 promoter based expression systems are often used in recombinant protein preparation, but systems using the λP_L promoter/cI repressor or the *tac* promoter are also common (Sørensen *et al.,* 2005, Terpe., 2006). The *tac* promoter is a hybrid between the *trp* and *lac* promoters. It is stronger than either, but still induced by IPTG (De Boer *et al.,* 1983, Brosius *et al.,* 1985). The λP_L promoter is also a very strong promoter. Expression vectors that carry the λP_L promoter are used together with a mutant *E. coli* host that synthesizes a temperature-sensitive form of the *c*I protein. At low temperature (less than 30°C) this mutant *c*I protein is able to repress the λP_L promoter, but at 42°C the protein is inactivated, resulting in transcription from the promoter (Elvin *et al.,* 1990). The T7 promoter is specific for the RNA polymerase coded by T7 bacteriophage, and a gene inserted downstream of this promoter will be expressed at a high level. Expression requires a host strain lysogenized by a DE3 phage fragment, encoding the T7 RNA polymerase, under the control of the IPTG-inducible *lacUV5* promoter. Addition of ITPG to the growth medium switches on synthesis of the T7 RNA polymerase by triggering the release of tetrameric LacI from the *lac* operator. This in turn leads to the transcription of the target gene from the T7 promoter, which is initiated by T7 RNA polymerase (Sørensen *et al.,* 2005, Studier and Moffat., 1986).

Inclusion bodies

The expression of foreign proteins at high level in *E. coli* often results in the formation of inclusion bodies composed of insoluble aggregates of the expressed protein. These cytoplasmic granules can be seen with a phase-contrast microscope and isolated from most soluble and membrane-bound bacterial proteins. Cells expressing high levels of foreign proteins can be concentrated by

centrifugation and lysed by mechanical techniques, sonication, or lysozyme plus detergents. It is crucial to obtain maximal cell lysis in order to obtain inclusion bodies in high yields. The inclusion bodies are recovered by centrifugation and extensively washed. The purpose of washing steps is to remove as much soluble, adherent bacterial protein as possible from the aggregated foreign protein. In most cases, adjusting the washing conditions allows the isolation of inclusion bodies that contain more than 90% pure foreign protein. (Sambrook and Russel, 2001) The material extracted from the purified inclusion body can potentially be used directly as an antigen (Harlow and Lane, 1988).

Protein purification

Protein purification is a series of processes that aims to isolate a single type of protein from a complex mixture, which is vital for the characterisation of the function, structure and interactions of the protein of interest. Biological tissue or a microbial culture are often used as starting material. The various steps in the purification process may free the protein from a matrix that confines it, separate the protein and non-protein parts of the mixture, and finally separate the target protein from other proteins, which often is the most labour-intensive step of protein purification. The separation steps exploits differences in protein size, physico-chemical properties and binding affinity. Initial planning, gathering of information and the design of a purification strategy are all important prerequisites for a successful protein purification. The purpose of the protein purification will indicate what is of most importance, whether one wants to achieve the highest possible quality (protein purity), quantity (protein yield) or cost-effectiveness. A purified protein can for example be used for sequencing and identification, antibody production, functional (kinetics) or structural (crystallography) studies, physico-chemical characterisation methods, or *in vivo* studies.

If possible, all known properties about the target protein are collected, such as function, stability, localisation, solubility, size, charge and known impurities. Then, the purification steps are mapped out in detail. Chromatography techniques takes advantage of the different properties of the proteins in the total mixture, and can separate proteins based on their hydrophobicity (hydrophobic interaction chromatography), binding affinity for a specific ligand (affinity chromatography), isoelectric point (ion exchange chromatography), or their size and shape (gel filtration). The two latter techniques are used in this study, and are therefore described in more detail.

Ion exchange (IEX) chromatography

An ion exchange column separates proteins based on their isoelectric point (pI). Most biomolecules are charged due to presence of ionic groups within their structure. The pH value at which a biomolecule carries no net charge is called the isoelectric point (pI), which is specific for each biomolecule. When exposed to a pH below its pI, the biomolecule will carry a positive charge, and above its pI, a negative charge. At a given pH value, a typical sample contains a mixture of molecules carrying different net charge of varying strength. To separate the components, the sample can be applied to an ion exchange column packed with gel bearing either negative charges (cation exchanger) or positive charges (anion exchanger). Molecules that carry a net charge opposite to that of the gel will bind to the gel by electrostatic forces, molecules that carry the same or no net charge will pass through. The binding of molecules to the column is reversible, as each molecule is normally displaced and selectively eluted from the column by an increasing salt gradient or by changing the pH. There are several advantages with ion exchange chromatography; it is easy to use and results can be predicted. Also, since the method involves high loading capacity, and concentrates the sample, it can be used as an early purification. It is however important to note that the choice of eluent and gradient is critical. In this study, the anion exchange column Resource Q, is used.

Gel filtration (GF)

A gel filtration column separates proteins based on their size and shape. A gel filtration column is packed with a gel which comprises porous beads, *e.g.* highly cross-linked agarose. When a

sample is passed down the column, separation depends on the different abilities of the sample components to enter the pores within the gel beads. Larger molecules, which cannot enter even the largest pores, pass through the column fastest. Smaller molecules, which can enter the pores freely, are delayed to different degree during the passage through the gel, depending on their size and shape. Proteins are therefore eluted in order of decreasing size. The advantages of gel filtration are that the method is easy to use, it provides free choice of eluent, works almost always, and the result is easy to predict. Due to limitations such as low resolution and small sample volume, this chromatography technique is preferably used as an intermediate, or final purification step. Also, it should be noted that influence of flow rate, and column efficiency, plays important roles. In this study, the gel filtration column Superdex 200, is used.

Background for this study

In 2007 the project "Cloning and expression of the *Helicobacter pylori* antigens NAP and FlaA in *E. coli*" was performed (Ingrid Lea, unpublished). One aim of the project was to compare the effects of three different expression systems based on either the *tac* promoter, the T7 promoter or the λP_L promoter. As a result of the project work, five recombinant *E. coli* strains, in which expression of recombinant *H. pylori* proteins could be induced, were successfully constructed (two for the expression of NAP and three for the expression of FlaA). NAP was expressed from the pRSET (T7 promoter) vector in *E. coli* BL21DE3 and from the pML-λP_L vector (λP_L promoter) in *E. coli* N4830I, while FlaA, in addition to the two abovementioned expression systems, also was expressed from the pML vector (*tac* promoter) in *E. coli* BL21. We concluded that the heat-induced N4830I-pML-λP_L expression system was superior to the other systems in the case of both NAP and FlaA expression, because it provided a high-level expression of NAP, the best FlaA expression observed, and the best target protein/total protein ratio. Bacteriophage λP_L promoter based vector systems have also previously been shown to be very efficient and convenient for the expression of foreign genes, and it has been suggested that the heat-inducibility of the λP_L promoter makes it especially attractive for large-scale production and purification of gene products (Cheng and Patterson, 1992). This, together with the observations made in the described project, suggested that the expression system based on the heat-induced λP_L promoter should be the first choice for production and purification of the *H. pylori* antigens NAP and FlaA.

In this study, the H. pylori antigens NAP and FlaA were expressed in the N4830I-pML-λP_L system. The purification strategies to develop protein purification protocols for the two proteins, included optimisation of protein expression and cell disruption, separation by precipitation and/or centrifugation techniques, and chromatography as final purification steps. NAP was purified by a combination of ammonium sulphate precipitation, anion exchange chromatography and gel filtration. FlaA was isolated as inclusion bodies. However, no downstream purification was possible due to problems with re-precipitation.

Aims of the thesis

The overall aim of this project was to purify the *Helicobacter pylori* antigens NAP (Neutrophil-activating protein) and FlaA (Flagellin A) obtained from recombinant expression in *E. coli*.

The specific aims were:

- To characterize and optimise the expression of recombinant NAP and FlaA in an *E. coli* expression system based on the λP_L promoter.
- To establish protein purification protocols for NAP and FlaA that lead to purified proteins of good quantity and quality to be used for vaccine development purposes.

Materials and methods

Materials

Buffers and solutions

Buffers and solutions used in this study are described in the appendix, table A1.

Cells culture, vector, and growth media

The expression vector pML-λP_L was kindly provided by Mike Lebens at the Department of Microbiology and Immunology. NAP and Flagellin A proteins were expressed from this vector under the control of the heat-induced λP_L promoter, and *E. coli* N4830I was used as host. Cells were grown at 30 ºC on LB liquid medium and LB plates containing 100 μg/ml Amp.

Methods

Cloning of the H. pylori napA and flaA genes

Polymerase chain reaction

The *flaA* and *napA* genes from *H. pylori* strain SS1 were amplified by high fidelity PCR (2 min at 94 ºC, 5 cycles of 30 s at 94 ºC, 30 s at 50 ºC and 1 min at 72 ºC, 20 cycles of 30 s at 94 ºC, 30 s at 62 ºC and 1 min at 72 ºC, and finally 7 min at 72 ºC) using gene specific primer pairs where sites for the restriction enzyme BsaI were added. The *flaA* gene was amplified using the primers BsaI-FlaA-f (5'-CGGTCTCGAATTCTATAACAAGGA GTTACAACA**ATG**GCT-3') and BsaI-FlaA-r (5'-CGG TCTCAAGCTTACAAACACCTTTCTCAAAA**CTA**A GT3-'). The *napA* gene was amplified using the primers BsaI-NAP-f (5'-C GGTCTCGAATTCAAAAGGACTT TTG**ATG**AAAACATTTGA-3') and BsaI-NAP-r (5'-C GGTCTCAAGCTT**TTA**AGCTAAATGGGCTTCTAGC AT-3'). The start and stop codons are marked in bold. The Expand High Fidelity PCR System (Roche) was used in a 20 μl PCR reaction that consisted of 0.5 μl dNTP, 2.5 μl $MgCl_2$, 2.5 μl buffer (with $MgCl_2$), 0.75 μl of each primer, 0.5 μl enzyme, and dH_2O. 250 ng genomic DNA from *H. pylori* SS1 strain was added as template. The TrueStart™ Taq PCR System (Fermentas) was used to verify the presence of correct inserts in all cloning and transformation steps. The 25 μl PCR reaction consisted of 0.5 μl dNTP, 2.5 μl buffer, 4 μl $MgCl_2$, 0.75 μl of each primer, 0.15 μl enzyme, and dH_2O. The bacterial colony to be screened was added as template.

Agarose gel electrophoresis

Agarose gel electrophoresis was used to check PCR-products, ligation products or DNA cut with restriction enzymes. 9 g DNA grade agarose was dissolved in 60 ml 0.5 % TBE buffer to make a 1.5 % agarose gel and the gel was left to polymerise for 30 minutes. The samples to be run on the gel were mixed with gel loading buffer, and 10 µl was loaded into each of the wells. 6 µl of 1 Kb DNA ladder or 100 bp DNA ladder from Fermentas (Appendix, fig. A2 / A3) was used as standard. The gel was run in 0.5 % TBE buffer at 100 V for 0.5 - 1.0 h, depending on the degree of separation needed. The gel was stained with EtBr, and the DNA bands were visualised by exposing the gel to UV light.

TOPO TA cloning

The PCR products amplified by High Fidelity PCR were ligated into a cloning vector using the TOPO TA cloning Kit from Invitrogen. A cloning reaction of 1 μl fresh PCR product, 1 μl diluted salt solution, 1 μl pCR 2.1-TOPO vector and 3 μl H_2O was mixed gently, incubated for 20 minutes at room temperature and placed on ice until transformation into bacteria.

Transformation of competent bacteria

The competent *E. coli* cells were thawed on ice. 1 μl ligation reaction and 40 μl competent cells were mixed in a cuvette on ice. The samples were electroporated at 2.5 V, and 250 μl of LB medium was immediately added to the cells. The mixture was transferred to a tube, and incubated at 30 °C on a shaker for 1 h. The transformation mixture was spread on pre-warmed LB-Amp plates. The plates were incubated overnight at 30 °C, and analysed by PCR screening.

Plasmid Miniprep and Determination of DNA concentration

Plasmids were isolated from bacterial culture using the GeneJET™ Plasmid Miniprep Kit from Fermentas. The over-night culture was centrifuged at 5000 x g for 10 minutes, and the resulting pellet was used to extract plasmid DNA according to the GeneJET™ Plasmid Miniprep Protocol. The concentration of the DNA extracted from the Plasmid Mini-prep procedures was measured by using the NanoDrop ND-100 Spectrophotometer.

DNA Sequencing

Target sequences were sequenced after cloning into pCR 2.1 TOPO / XL1. The purified plasmid DNA samples were diluted in elution buffer (10 mM Tris-HCl) and 10 μl H_2O to reach a concentration of 2.5 μg DNA per 20 μl. The prepared samples were sequenced at Eurofins MWG Biotech in Germany, and the resulting sequences were analysed by using Blast2.

DNA digestion by restriction enzymes

Plasmid DNA was digested with restriction enzymes in buffer G or Tango (Fermentas) in order to isolate *napA*

and *flaA* inserts from pCR 2.1 TOPO, and to prepare the expression vector for ligation with the *napA* and *flaA* inserts. The pCR-*napA* and pCR-*flaA* plasmids with inserts generated from PCR amplification using gene specific primers were cut with BsaI in buffer G in order to create *napA* and *flaA* inserts with EcoRI/HindIII overhangs. The expression vector pML-λP_L (Appendix, fig. A4) was opened with EcoRI (Fermentas) and HindIII (Boehringer Mannheim) in buffer Tango. (For details on digestion reactions, see table A2 in the Appendix).All digestion reactions were incubated at 37 ºC overnight.

Ligation of flaA /napA fragments into pML-λPL vector

The expression vector pML-λP_L was ligated with the flaA/nap inserts using T4 Ligase and ligation buffer from Fermentas. Reaction I (NAP) consisted of 8 μl vector (pML-λP_L, 6.2 ng/μl), 22 μl insert (*napA*, 1.5 ng/μl), 10x ligation buffer and 1 μl Ligase. Reaction II (FlaA) consisted of 8 μl vector (pML-λP_L, 6.2 ng/μl), 13 μl insert (*flaA*, 7.5 ng/μl), 10x ligation buffer and 1 μl Ligase. The ligation reactions were incubated at 22ºC overnight.

Purification of PCR product from agarose gel

Bands from agarose gels were purified using the E.Z.N.A™ Cycle-Pure Kit from Omega Bio-Tek. The gel piece was melted at 50 ºC in QX1 Buffer from the QIEAII Gel Extraction Kit (Qiagen), and this sample was used in DNA purification according to the E.Z.N.A™ Cycle-Pure Protocol. The concentration of the eluted DNA was determined by NanoDrop as described above.

Expression of NAP and FlaA

Cell cultures: growth, induction and fractioning.

All incubations were performed with shaking (180 rpm). Cells were grown in 1.5x LB-Amp at 30 °C overnight to make pre-cultures. Pre-cultures were diluted 1:100 in 1.5 x LB-Amp, and then incubated for 4 h before dividing it in two separate cultures. Recombinant protein expression was induced in one of the cultures by raising the temperature to 42 °C. The induction time was 3 h for all cultures. Both induced and non-induced cultures were centrifuged at 10 000 x g at 4 °C for 20 minutes, and the bacterial pellets were dissolved in PBS. The bacterial suspension was mixed with lysozyme (1 mg/ml) and EDTA (to a final concentration of 5 mM), and incubated for 10 minutes at room temperature (in which time the suspension should become viscous due to the presence of lysed cells that release their DNA). $MgCl_2$ (to a final concentration of 20 mM) and a small amount of solid DNase were added, and left to stand for 10 min at r.t. (in which time the viscosity should disappear). The lysate was ultra-sonically treated (60 Amp, 2 x 4 min, pulser 2 sec) and then checked under the microscope (to make sure that there were no intact cells left). The sonicate was centrifuged at 600 x g at 4 °C for 10 minutes, and the pellet (cellular debris) was dissolved in PBS. The supernatant was centrifuged at 5000 x g at 4 °C for 10 minutes, and the pellet (inclusion bodies) was dissolved in PBS. The resulting supernatant was ultra-centrifuged at 76 000 x g for 30 minutes to separate the membrane fraction (pellet) from the cytosolic sample (supernatant). The total protein samples, sub-cellular fractions and cultivation medium were analysed using SDS-PAGE and Western blotting.

SDS-PAGE

The NuPAGE® electrophoresis system from Invitrogen were used for SDS-PAGE according to the NuPAGE® Technical Guide. Each sample was mixed with an equal volume of 2x sample buffer, and 10 μl were loaded into each well of a NuPAGE® 12 % Bis-Tris gel. 8μl All Blue protein ladder from BIO-RAD (Appendix, fig. A1) was used as standard. The gel was run in 1x MOPS buffer (50 ml 20 x NuPAGE MOPS SDS Running Buffer, 950 ml dH_2O) at 200 V for 50 min. The gel was either used in western blotting or Coomassie staining. The gel was stained in Coomassie 0.25 % Brilliant BlueR-250 Staining Solution for 2 h (or overnight), and incubated in Destaining Solution (10 % HAC, 30 % EtOH, 60 % H_2O) for 3 h. The gel was washed in H_2O for 3 x 10 minutes, and incubated in NuPAGE® Gel Drying solution (available from Invitrogen) at gentle agitation for 20 minutes. The gel was dried using the NuPAGE® DryEase Mini-Gel Drying System from Invitrogen.

Western blotting

Western Transfer of the gels were performed according to the NuPAGE® Technical Guide. The gels were transferred in 1x NuPAGE® Transfer Buffer (50 ml 20x NuPAGE transfer buffer, 100 ml EtOH, 850 ml dH_2O) at 30 V for 1 h. All incubation and wash steps were performed with gentle agitation. The transfer membrane was put in protein blocking solution (1 % BSA in PBS) for 30 min (or overnight), and rinsed briefly in PBS. The membrane was incubated with primary antibody solution (10 ml 0.1% BSA in PBS, 20 μl Tween-20, 500 μl mAb FlaA (75 mg/ml) or mAb NAP (31 mg/ml) for 2 h (or overnight), and washed for 3 x 5 min in wash solution (0.05 % Tween-20, PBS). The blot was incubated with enzyme conjugated second antibodies (10 ml 0.1% BSA in PBS, 20 μl Tween-20, 100 μl Jackson HRP Goat anti-mouse IgG) for 2 h, and washed 2 x 5 min in wash solution containing Tween-20, and 2 x 5 min in PBS. The blot was developed in substrate-chromogen solution (2 ml alfa-chloro-naphthol in cold methanol, 10 ml TBS

Buffer (0.02 M phosphate, 0.15 M NaCl, pH7.5), 6 μl 30 % H_2O_2) for 5-15 min. The membrane was rinsed in tap water to stop the reaction, and left to dry between filter papers.

Periplasmic preparation

E. coli N4830-I cells harbouring the recombinant plasmid (pML-λP_L-*napA*) were used. Starting with 5 ml of induced culture, cells were harvested by centrifugation (10 000 x g for 20 min) and suspended in 1 ml of 20 % (w/v) sucrose, 0.3 M Tris-HCl, pH 8.0, 1 mM EDTA, and incubated for 20 min at r.t. The lysate was centrifuged at 10 000 for 20 minutes. The resulting pellet was suspended in 0.5 mM cold $MgCl_2$ and incubated on ice for 15 min, and then centrifuged at 20 000 x g for 10 min. An aliquot of the periplasmic fraction (supernatant) was diluted 10 x and checked for NAP with SDS-PAGE.

Protein purification

Preparation of inclusion bodies

E. coli N4830-I cells harbouring the recombinant plasmid (pML-λP_L-*flaA*) were grown in LB-Amp and induced as described above. Samples were collected after 2, 3 and 4 h as well as after over night induction, for microscopic and SDS-PAGE analyses. Starting with one l of 4 h induced culture, cells were harvested by centrifugation (10 000 x g for 20 min) and lysed as described above. The inclusion bodies were pelleted by centrifugation (3000 x g for 10 min) and washed in 10 ml of cold PBS, 10 mM EDTA a total of 3 times. After a final centrifugation (6000 x g for 10 min), the inclusion body pellet was dissolved in 1 ml dH2O, and mixed quickly with 20 ml of 6.5 M Urea, 1 mM EDTA. The solution was left at r.t over night at gently rotation to dissolve. The next day, the insoluble material was removed by centrifugation (27 000 x g for 20 min), and the supernatant was dialysed against 10mM Tris-HCl, pH 7.5, 50 mM NaCl. The buffer was changed two times before and one time after over night dialysis. The dialysate was centrifuged at 27 000 x g for 30 min to remove any particles. The supernatant and the pellet were checked for FlaA protein with SDS-PAGE. An alternative strategy was also applied to avoid re-precipitation of FlaA. Inclusion bodies were pelleted by centrifugation (6000 x g for 10 min) and dissolved in dH2O and Urea, as described above. After removal of insoluble material, the supernatant was dialysed as above, but this time with a stepwise removal of urea. Urea concentrations in the buffer were divided by two a total of three times (from 1 M to 0.125 M Urea), before over night dialysis against 10mM Tris-HCl, pH 7.5, 50 mM NaCl. After centrifugation of the dialysate (27 000 x g for 30 min), the supernatant and pellet were checked by SDS-PAGE. The supernatant was stored at -20 °C, and after thawing the sample, a new centrifugation (48 000 x g for 30 min) was performed, to confirm the localisation of FlaA (soluble or non-soluble fraction).

Ammonium sulphate precipitation

All incubation steps were performed at room temperature. Starting with 2 l of induced culture, cells were harvested and dissolved in 10 ml of 500 mM Tris-HCL pH 7.5, 1 mM EDTA, 500 mM NaCl. The cells were lysed as described above. The lysate was centrifuged at 15 000 x g for 20 min, and the resulting supernatant was centrifuged again at 15 000 x g for 20 min. For NAP, the supernatant was made 60 % (w/v) saturated with respect to ammonium sulphate. After centrifugation (15 000 x g for 45 min) the supernatant was dialysed against 10 mM Tris-HCl, pH 8.0, 20 mM NaCl, with several buffer changes before and after overnight dialysis. The dialysate was concentrated by centrifugation (5000 x g for 3 x 10 min) in a Vivaspin (5000 MWCO) spin column. An attempt to purify FlaA by ammonium sulphate precipitation was also performed. For FlaA, the dialysed supernatant obtained from preparation of inclusion bodies was made 30 % (w/v) saturated with respect to ammonium sulphate. After centrifugation (15 000 x g for 45 min), the resulting supernatant was made 50 % (w/v) saturated with ammonium sulphate. Precipitated material was again pelleted by centrifugation (20 000 x g for 45 min). The resulting pellets from 30 % and 50% saturation were resuspended in 10 ml of 10 mM Tris-HCl, pH 8.0, 20 mM NaCl. Both of the pellets, and the 50 % supernatant, were checked for FlaA protein by SDS-PAGE.

Anion exchange chromatography

Protein purification was performed using the UNICORN™ controlled ÄKTA*explorer*™ system. The dialysed NAP supernatant obtained from ammonium sulphate precipitation was applied to a Resource Q column that had been equilibrated with 10 mM Tris-HCl, pH 8.0, 20 mM NaCl. Proteins were eluted with a NaCl gradient generated by mixing the abovementioned buffer with 10 mM Tris-HCl, pH 8.0, 1 M NaCl. The eluted fractions were resolved by SDS-PAGE. The fractions containing NAP were concentrated by use of a vivaspin column, and applied in the next purification step (Gel filtration). The dialysed FlaA supernatant obtained from preparation of inclusion bodies was applied to a Resource Q column that had been equilibrated with 10 mM Tris-HCl, pH 7.5, 50 mM NaCl. Proteins were eluted with a NaCl gradient generated by mixing the equilibration buffer with 10 mM Tris-HCl, pH 7.5, 1 M NaCl. The eluted fractions were analysed with SDS-PAGE.

Gel filtration

The fractions containing partially purified NAP, obtained from anion exchange chromatography, were pooled and further purified by gel filtration chromatography using a Superdex 200 10/30 column, and the UNICORN™ controlled ÄKTA*explorer*™ system. Proteins were eluted with 10 mM Tris-HCl, pH 8.0, 0.5 M NaCl, and the fractions containing purified NAP were resolved by SDS-PAGE, and pooled.

Determination of protein concentration

The concentration of the target protein in the pooled fractions after gel filtration chromatography was measured by use of the DC Protein Assay (BIO-RAD) and the NanoDrop ND-100 Spectrophotometer. The DC Protein Assay was used according to the manufacturer's instructions. BSA dissolved in the same buffer as the NAP protein was used for preparation of a standard curve.

An outline are drawn for the cloning and expression of *H. pylori* NAP and FlaA in *E. coli*, in fig. 4.

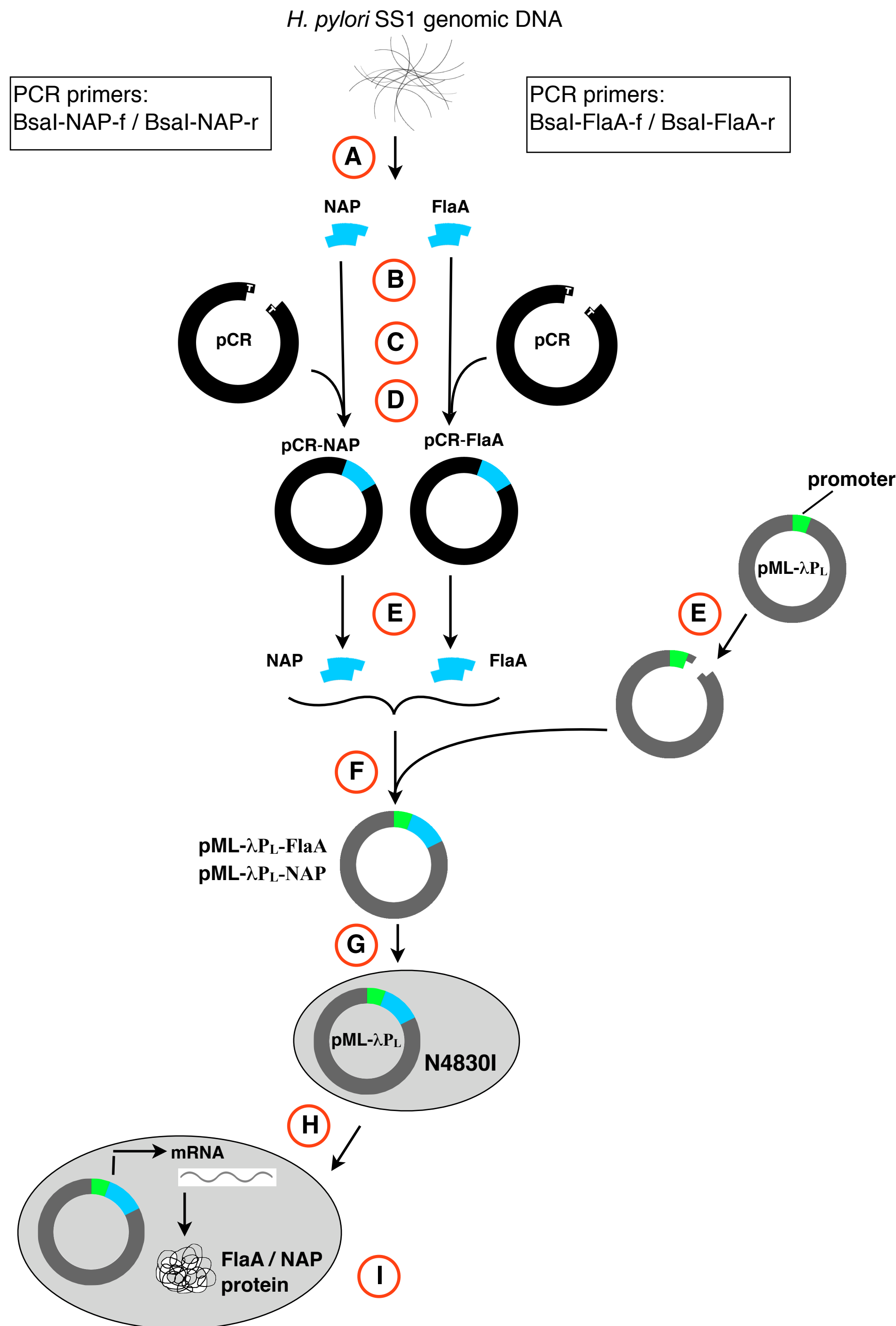

Fig. 4. Cloning and expression of *H. pylori* NAP and Flagellin A in *E. coli*. **Cloning and Sequencing:** (A) Primers were designed to amplify flaA and napA genes from *H. pylori* strain SS1. (B) The napA and flaA genes were cloned into the pCR vector by TA cloning. (C) The recombinant plasmids were amplified in *E. coli* XL-1 strain, and extracted by plasmid miniprep. (D) Sequencing analysis of napA and flaA fragments. (E) Plasmid DNA were digested by the use of restriction enzymes. (F) The target fragments and expression vectors were recovered and ligated. (G) The recombinant expression vectors were transformed into *E. coli* N4830I host cells. **Expression and identification of recombinant protein:** H) Protein expression was induced in N4830I by heat (42 °C). (I) The molecular weight and output of Flagellin A and NAP protein were examined by SDS-PAGE and western blotting. Monoclonal antibodies (mAbs) against FlaA and NAP were used as primary antibodies. Jackson HRP-conjugated Goat anti-mouse IgG was used in the secondary step.

The purification protocol for *H. pylori* NAP is outlined in fig. 5. Attempts to purify *H. pylori* FlaA are shown in fig.6.

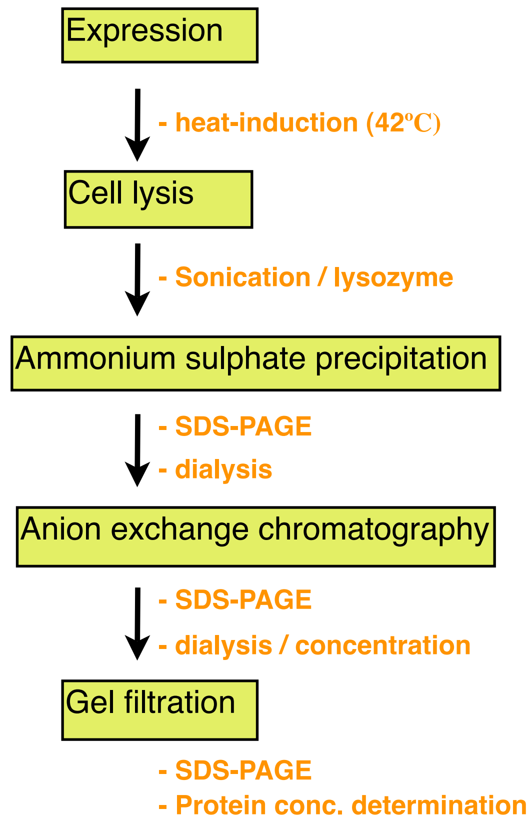

Fig. 5. Flow scheme, purification of *H. pylori* NAP.

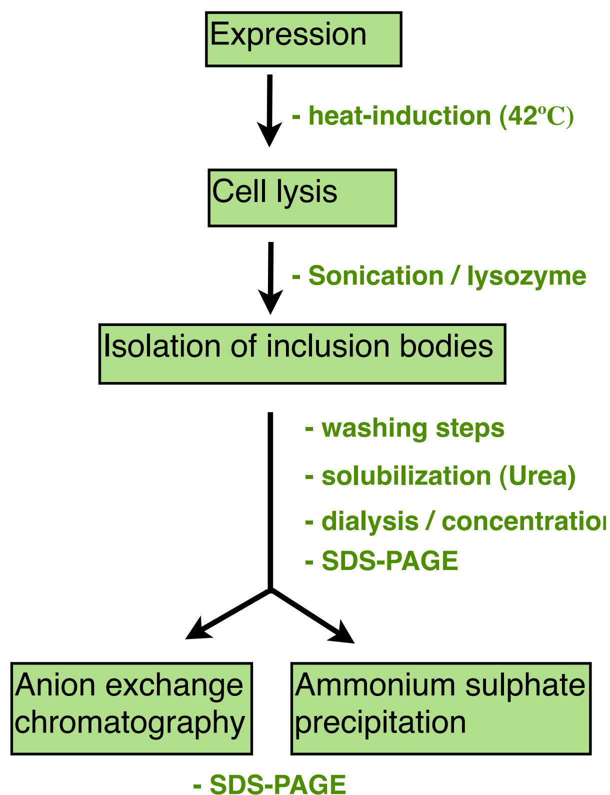

Fig. 6. Flow scheme, attempt to purify *H. pylori* FlaA.

Results

Cloning of NAP and FlaA

PCR followed by agarose gel electrophoresis confirmed both the generation of *napA* and *flaA* target DNA amplification products, and the successful ligation of these into the cloning vector to make pCR-*napA* and pCR-*flaA*. Target fragments of *napA* and *flaA* genes with expected sizes (*napA* ~450 bp and *flaA* ~1530 bp) amplified from DNA template of *Helicobacter pylori* SS1 strain are shown in figure 7. Sequence analysis of the inserts identified the correct *napA* and *flaA* clones. After amplification in the XL-1 strain and plasmid miniprep, the *napA* and *flaA* genes were cut out from the cloning vector with rectriction enzymes, and ligated into the expression vector pML-λPL to make the constructs pML-λPL-*napA* and pML-λPL-*flaA*. PCR followed by agarose gel electrophoresis confirmed the successful restriction enzyme cutting and ligation of correct DNA fragments.

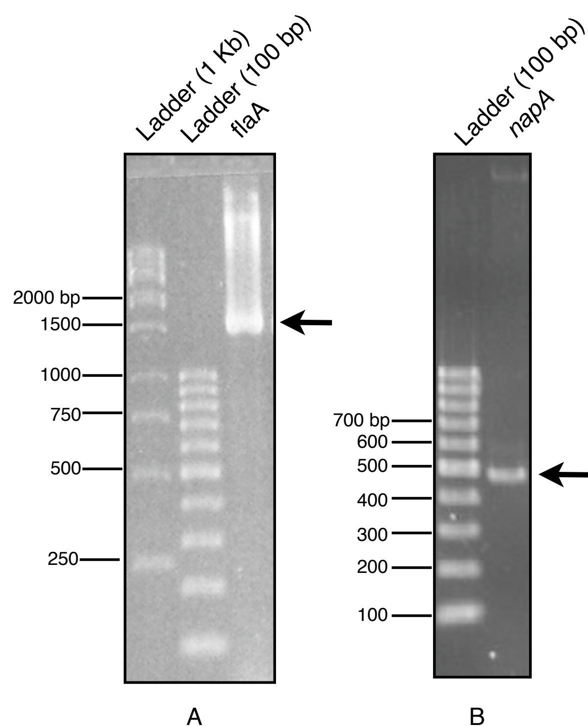

Fig. 7 Agarose gel electrophoresis showing target fragments of *flaA* and *napA* genes amplified from *H. pylori* SS1 strain using gene specific primers. From left to right: (A) 1 Kb ladder, 100 bp ladder, *flaA* (~1530 bp), (B) ladder (100 bp), *napA* (~450 bp). Arrows indicate the *flaA* and *napA* fragments. The sizes of DNA standards are labelled on the left.

Expression of NAP and FlaA

The expression vector constructs pML-λP_L-*napA* and pML-λP_L-*flaA* were successfully electroporated into the *E. coli* N4830I strain. PCR screening and agarose gel electrophoresis of clones selected on growth media with amp confirmed the presence of *napA* and *flaA* inserts. The NAP and FlaA proteins produced in N4830I-pML-λP_L-*napA* and N4830I-pML-λP_L-*flaA* induced by heat (42ºC), were detected on SDS-gels and immunoblots. Figure 8 and 9 (see next page) shows the expression of *H. pylori* FlaA and NAP in *E. coli*, respectively. Arrows indicate the 53 kDa FlaA protein band (fig. 8) and 17 kDa NAP protein band (fig. 9).

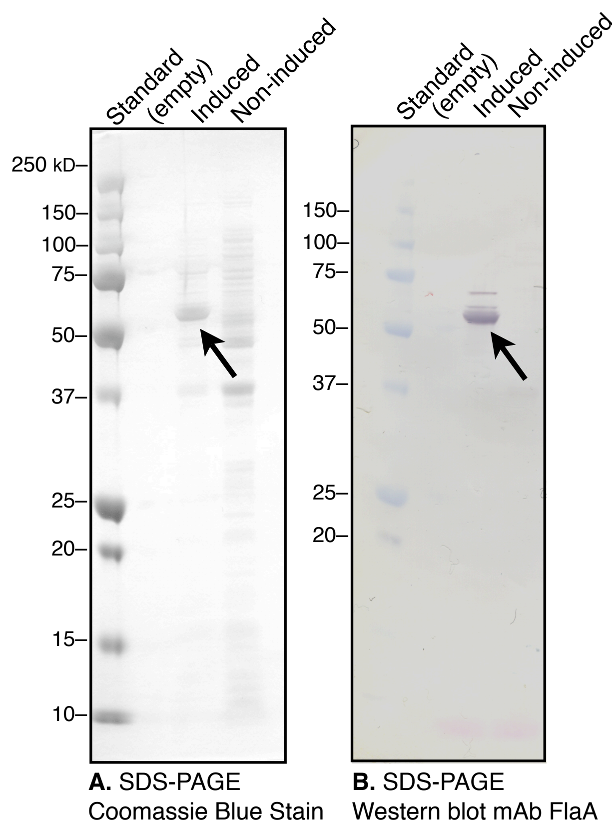

Fig. 8. Expression of recombinant H. pylori Flagellin A protein. (A) SDS-PAGE (B) Immunoblot. From left to right (both A and B): Protein Standard, cell extract of heat-induced (42ºC) *E. coli* N4830I containing pML-λPL-*flaA,* cell extract of non-induced *E. coli* N4830I containing pML-λPL-*flaA*. Arrows indicate the 53 kDa FlaA band, which was present in induced cells but lacking in non-induced cells. The sizes of protein standards are labelled on the left.

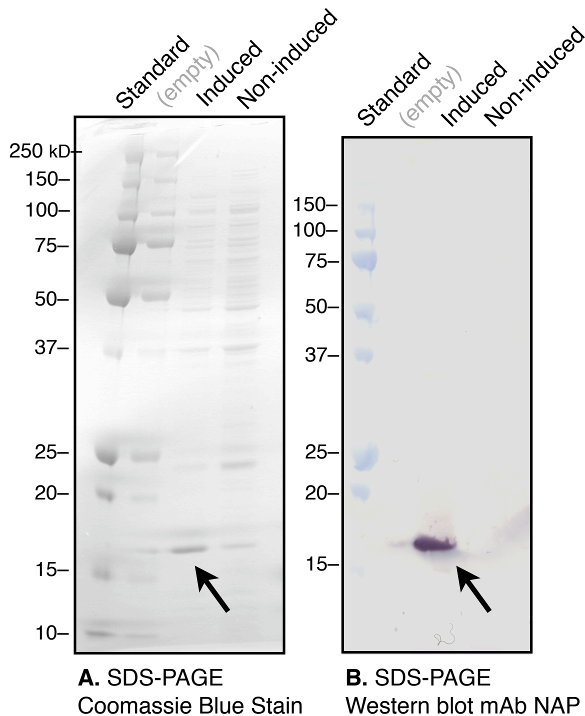

Fig. 9. Expression of *H. pylori* NAP protein. SDS-PAGE (12 % Bis-Tris gels) analysis of soluble cell extracts of N4830I-pML-λPL-*napA* by Coomassie Staining (A) and Western Immunoblotting (B) using mAb NAP. Detection was performed using . From left to right (both A and B): Protein Standard, cell extract of heat-induced (42ºC) *E. coli* N4830I containing pML-λPL-*napA,* cell extract of non-induced *E. coli* N4830I containing pML-λPL-*napA*. Arrows indicate the 17 kDa NAP band, which is mainly present in the induced cells. The sizes of protein standards are labelled on the left.

Characterisation of NAP and FlaA expression

To investigate where NAP and FlaA reside in the host cells after induced expression, the cultivation media and different cellular fractions (obtained according to the fractionation by centrifugation described in the materials and methods section) were analysed by SDS-PAGE and western blot. The outcome of this fractionation showed that the majority of the FlaA protein ends up in inclusion bodies (data not shown), whereas the majority of the NAP protein was detected in the soluble cytosolic fraction (fig. 10). A periplasmic preparation (see materials and methods section) was also performed, to find out whether NAP is produced as a periplasmic protein or not. The resulting fractions were controlled by SDS-PAGE, and shown in fig. 11. The NAP observed in the total cell extract was also present in the pellet, but not in the periplasmic preparation (supernatant).

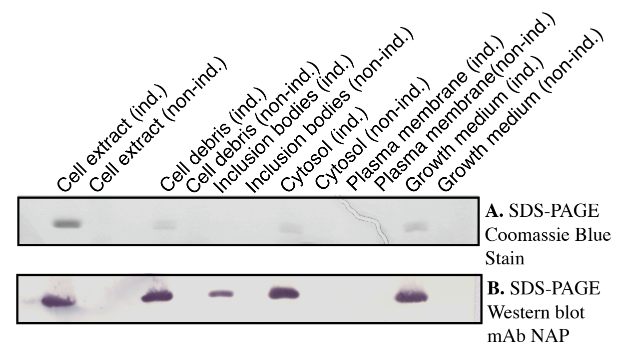

Fig. 10. SDS-PAGE analysis of induced and non-induced N4830I-pML-λPL-*napA* cell extracts, cellular fractions, (cell debris pellet, inclusion body pellet, cytosolic fraction /supernatant and plasma membrane pellet) and growth medium, by Coomassie Staining (A) and Western Immunoblotting (B) using mAb NAP (see material and methods, and discussion section for more details).

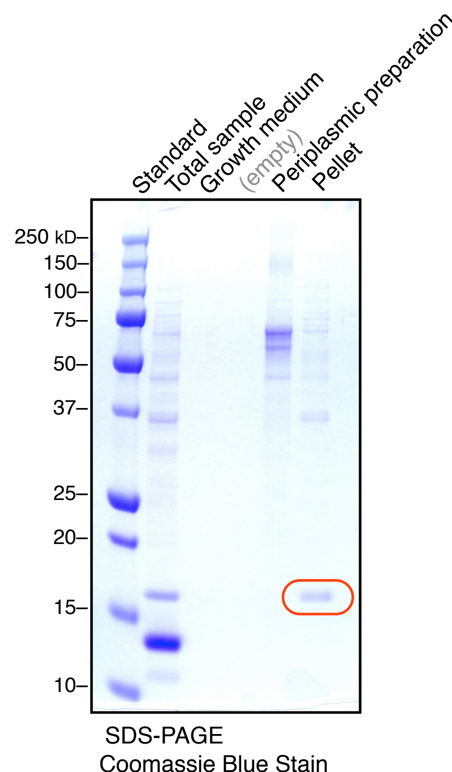

Fig. 11. The NAP protein is not found in the periplasmic preparation. From left to right: Protein standard, cell extract of heat-induced (42ºC) *E. coli* N4830I containing pML-λPL-*napA*, growth medium, supernatant (periplasmic proteins) and pellet after periplasmic preparation. The red circle indicates the 17 kDa NAP band, which is present in the pellet after periplasmic preparation but not in the supernatant (periplasmic proteins). The sizes of protein standards are labelled on the left.

To optimise the expression of FlaA, the effect of induction time on FlaA expression was explored. Samples of cell culture with *E. coli* N4830I harbouring the pML-λPL-*flaA* plasmid were collected after 2, 3, or 4 h induction at 42 ºC and 180 rpm agitation, and the rest of the cell culture was left overnight. Inclusion body count (microscope) and SDS-PAGE both showed an increased number of inclusion bodies and FlaA expression from 2 to 4 h. Over night induction did not improve the FlaA expression further. The SDS-PAGE results showed an increase in the total amount of proteins after over night induction. However, the best target protein/total protein ratio was observed after 4 h induction.

Attempts to purify FlaA

Inclusion bodies containing FlaA were isolated from the rest of the lysed host cell by centrifugation, and dissolved in urea. The inclusion body fraction and total cell extract were analysed by SDS-PAGE. The results showed that the inclusion bodies consisted of several proteins, in addition to FlaA. After pelleting insoluble material, the solubilized inclusion bodies were dialysed against a buffer without urea. Since the preparation of inclusion bodies only generated partially purified FlaA, additional downstream purification was needed. Therefore, the dialysed supernatant was precipitated using ammonium sulphate cuts at 30% and 50 % saturation (w/v), followed by centrifugation. SDS-PAGE analyses of the resulting 30 % and 50 % pellets and the 50 % supernatant, showed that all proteins, including FlaA, precipitated at 30%, and therefore were lost in the insoluble pellet.

Next, the dialysed supernatant obtained after the preparation of inclusion bodies was applied directly to a Resource Q column equilibrated with running buffer (10 mM Tris-HCl, pH 7.5, 50 mM NaCl). Proteins were eluted with a NaCl gradient. The eluted fractions were analysed by SDS-PAGE, but no FlaA band corresponded to the peaks on the chromatogram (data not shown). In another run with inclusion bodies on a Resource Q column, the pressure in the system became too high. This indicated that FlaA tends to re-precipitate and thus cause clogging of the column/system. Therefore an alternative dialysis approach was applied where the dissolved inclusion bodies were dialysed against running buffer with stepwise removal of urea. This strategy kept the FlaA protein in solution, as shown in figure 12 (A). This result was very promising, and it was planned to continue with this sample in anion exchange chromatography. However, when the FlaA supernatant (from 27 000 x g centrifugation) was stored at -20 ºC overnight, and then thawed, the FlaA protein re-precipitated, and was lost in the insoluble fraction. This was confirmed by a new centrifugation at 48 000 x g for 30 min, followed by SDS-PAGE of the resulting pellet and supernatant, as shown in figure 12 (B).

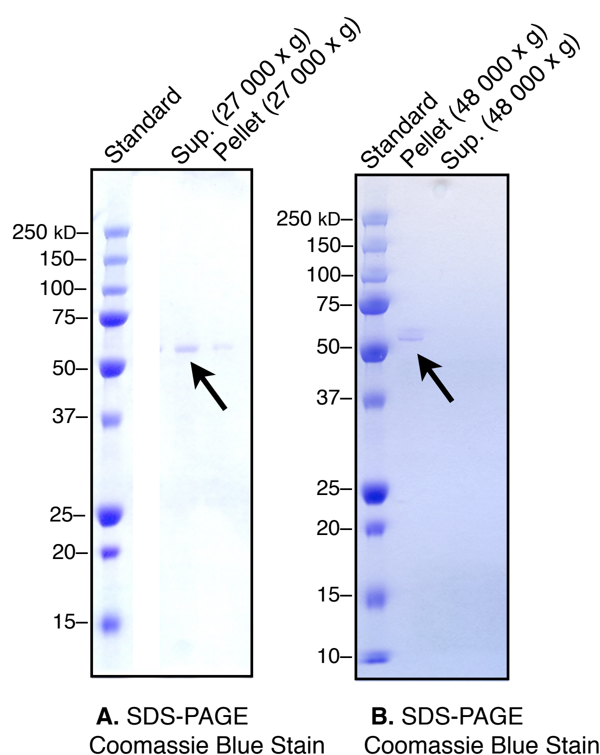

Fig. 12. FlaA inclusion body solubilisation (A) and re-precipitation after freeze-thaw cycle(B). From left to right: (A) Protein Standard, supernatant and pellet after 15000 rpm, (B) Protein Standard, pellet and supernatant after 20000 rpm. Arrows indicate the 53 kDa Flagellin A protein. The sizes of protein standards are labelled on the left. The pellet was dissolved in urea and dialysed against running buffer, and the urea concentration was lowered stepwise, before centrifugation at 27 000 x g for 30 min. This procedure kept FlaA in the soluble fraction (indicated by the arrow in 10A). When this supernatant was kept at -20 ºC, the FlaA protein re-precipitated, and FlaA was lost in the insoluble pellet after centrifugation at 48 000 x g for 30 min (indicated by the arrow in 10B).

Purification of NAP

To separate NAP from other contaminating proteins in the cell extract, an ammonium sulphate precipitation was performed (see materials and methods section). The supernatant was made 60% saturated with respect to ammonium sulphate, and SDS-PAGE analysis of the cell debris pellet, 60% supernatant and 60% pellet showed that at this saturation, NAP remained in the soluble fraction (supernatant), while most of the other contaminating proteins precipitated and ended up in pellet after centrifugation. Thus, the SDS-PAGE results in figure 13 shows that many proteins were effectively removed by the procedure.

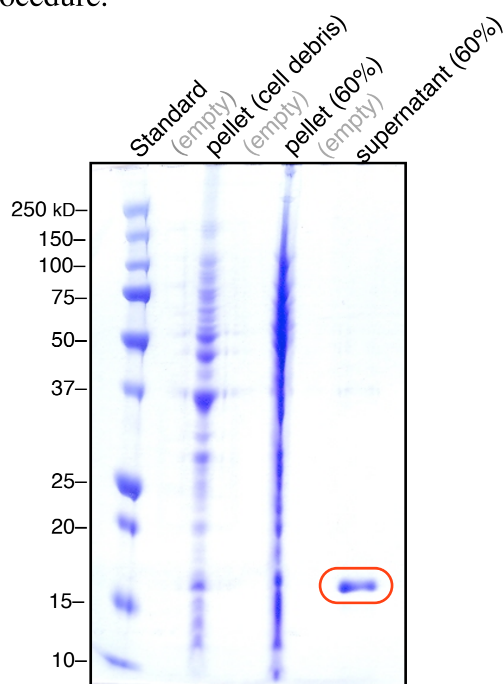

Fig. 13. SDS-PAGE analysis of ammonium sulphate precipitation. From left to right: Protein standard, cell debris pellet, 60 % ammonium sulphate pellet, 60 % ammonium sulphate supernatant. The red circle indicates the 17 kDa NAP band, which is present in the supernatant after ammonium sulphate precipitation. The sizes of protein standards are labelled on the left.

Next, the dialysed NAP supernatant obtained from ammonium sulphate precipitation was applied to a Resource Q column equilibrated with running buffer (10 mM Tris-HCl, pH 8.0, 20 mM NaCl). Proteins were eluted by a NaCl gradient. The eluted fractions were analysed by SDS-PAGE and NAP was detected in fractions 50 to 59. The eluted fractions containing NAP were pooled, concentrated and resolved by SDS-PAGE, shown in fig. 14 (A).

Then, the pooled fractions obtained from anion exchange chromatography were applied to a Superdex 200 10/30 gel filtration column. Proteins were eluted with 10 mM Tris-HCl, pH 8.0, 0.5 M NaCl. The eluted fractions were analysed by SDS-PAGE, and NAP were detected in fractions 35 to 44. The fractions containing purified NAP were pooled and resolved by SDS-PAGE, shown in fig. 14 (B).

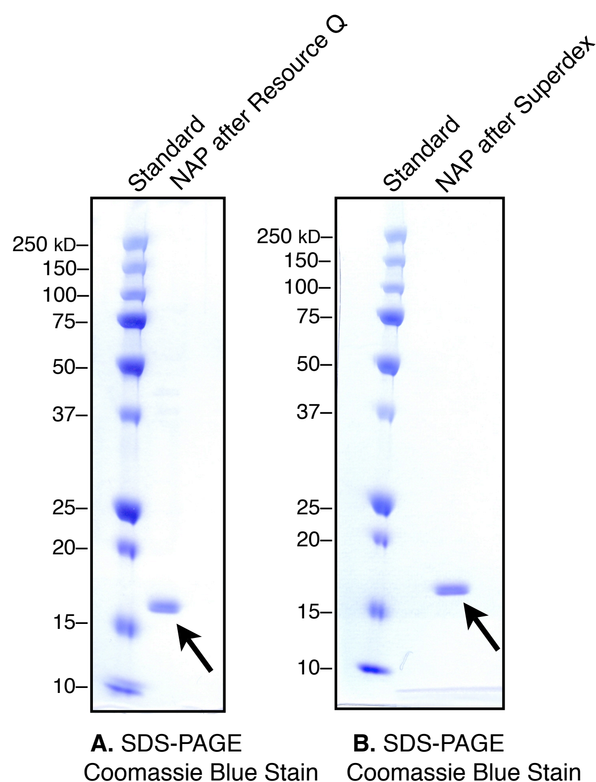

Fig. 14. Purification of *H. pylori* NAP protein. (A) Result Resource Q column, (B) Result after Superdex 200 10/30 column. From left to right: (A) Protein Standard, concentrated NAP fractions after anion exchange chromatography (Resource Q column), (B) Protein Standard, concentrated NAP fractions after gel filtration (Superdex column). Arrows indicates the 17 kDa NAP band. The sizes of protein standards are labelled on the left.

The concentration of the NAP protein in the pooled fractions after gel filtration chromatography was measured by using the NanoDrop ND-100 Spectrophotometer, and found to be 0.2 mg/ml. The DC Protein Assay results from using the same NAP sample, read a concentration of 0.18 mg/ml.

SDS-PAGE analysis after anion exchange chromatography showed that NAP was eluted in fractions 50 to 59, which corresponds to one of the six peaks on the chromatogram in fig. 15. Thus, the pooled fractions containing NAP were effectively separated from other contaminants by this procedure. The SDS-PAGE analysis after gel filtration showed that NAP was eluted in fractions 35 to 44, which corresponds to the first and highest peak on the chromatogram in fig. 16. The two other peaks represent proteins that also were observed as bands on the SDS-gel (data not shown), and thus were effectively removed by this final purification step. The values on the Y-axis in fig. 16 are not in the same scale as in fig. 15.

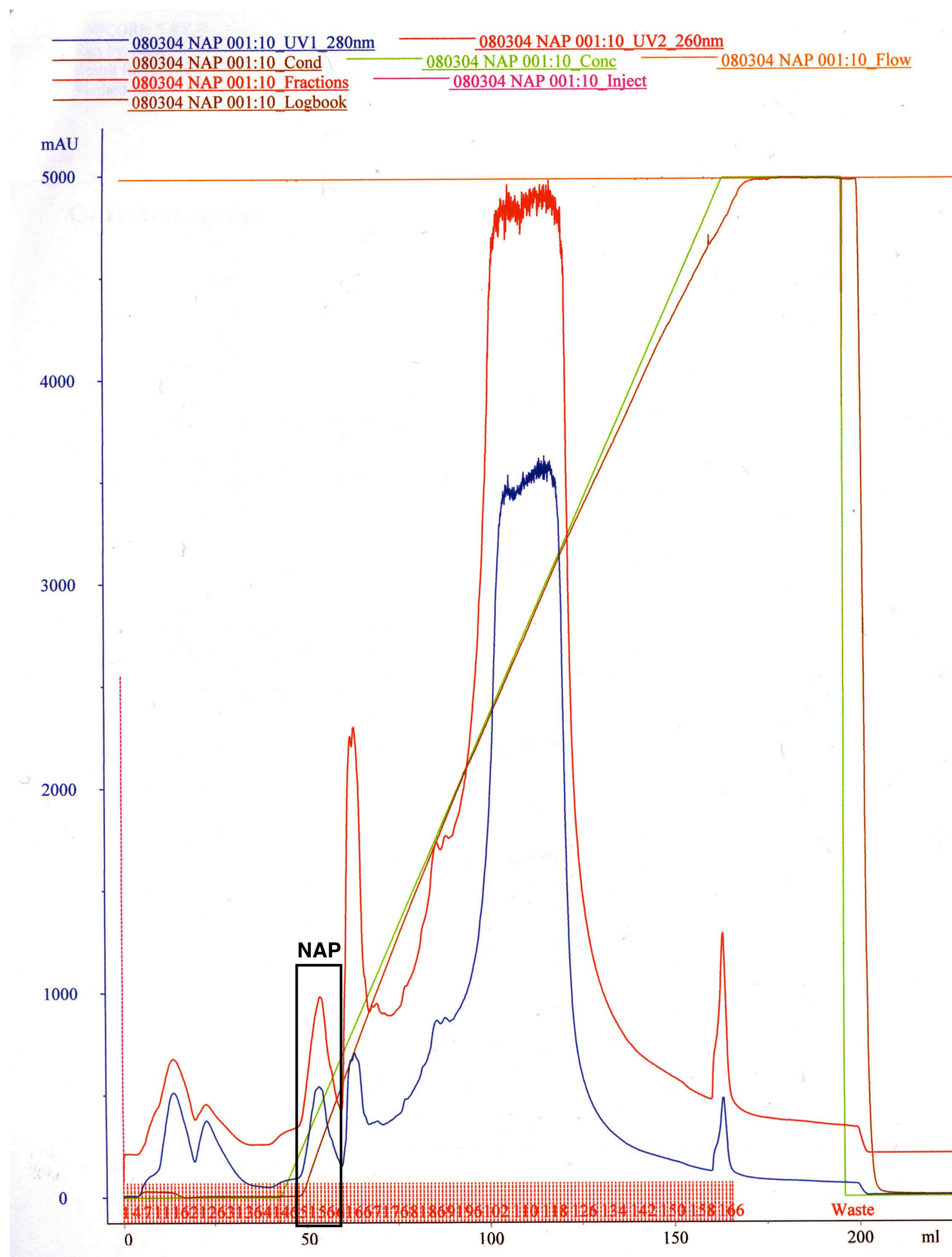

Fig. 15. Chromatogram from purification of *H. pylori* NAP using anion exchange chromatography. The dialysed NAP supernatant obtained from ammonium sulphate precipitation was loaded onto a Resource Q column equilibrated with running buffer (10 mM Tris-HCl, pH 8.0, 20 mM NaCl). Proteins were eluted with a NaCl gradient generated by mixing running buffer (see above) with elution buffer (10 mM Tris-HCl, pH 8.0, 1 M NaCl). The eluted fractions were resolved by SDS-PAGE, and NAP was detected in fractions 50 to 59, which were pooled and applied in gel filtration. These fractions represent the NAP peak (blue line) marked by a rectangle in the chromatogram. The Y-axis indicates the absorbance given in mAU (milli absorbance unit). The X-axis indicates the elution volume in ml.

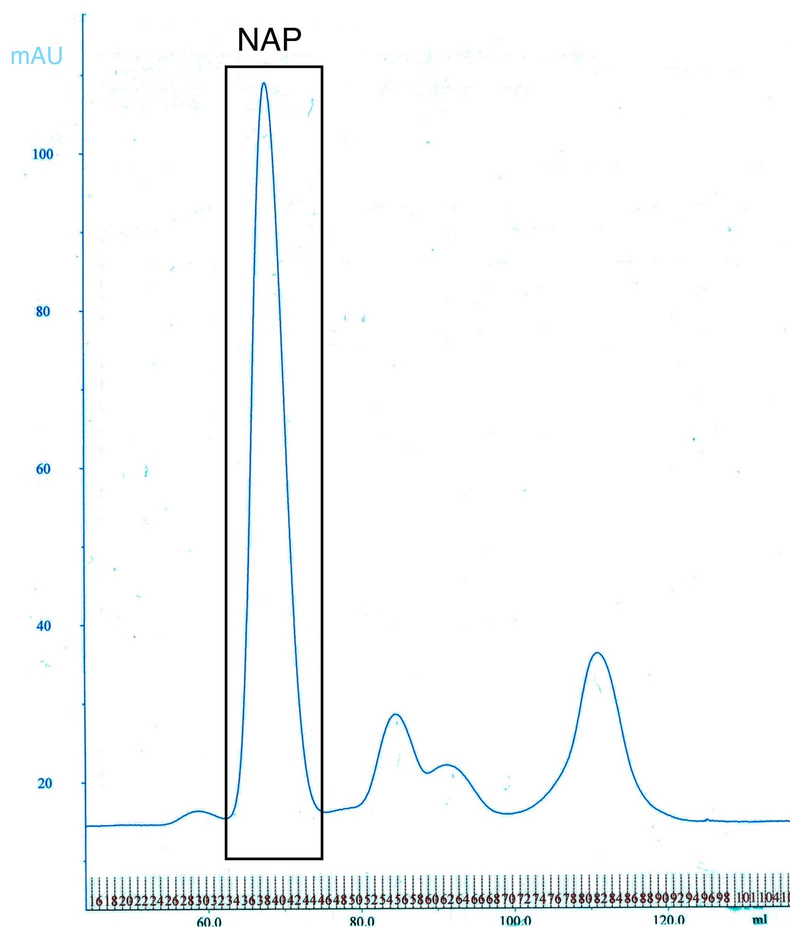

Fig. 16. Chromatogram from purification of *H. pylori* NAP using gel filtration. The partially purified NAP fractions obtained from anion exchange chromatography (fig. 13) were after concentration and buffer exchange by use of a vivaspin column, applied to a Superdex 200 10/30 column. Proteins were eluted with 10 mM Tris-HCl, pH 8.0, 0.5 M NaCl buffer. The eluted fractions were pooled and resolved by SDS-PAGE, NAP was detected in fractions 35 to 44, represented by the NAP peak marked by a rectangle in the chromatogram. The NAP fractions were pooled and tested with Nanodrop and DC protein assay. The two other peaks represent proteins that also were observed as bands on the SDS-gel, and thus were effectively removed by this purification step. No other contaminating proteins were observed by SDS-PAGE analysis. The Y-axis indicates the absorbance given in mAU (milli absorbance unit). The X-axis indicates the elution volume in ml.

Discussion

The aim of this study were to characterize and optimise the expression of recombinant NAP and FlaA in an *E. coli* expression system based on the λP_L promoter, in order to establish protein purification protocols for NAP and FlaA that led to purified proteins of good quantity and quality to be used for vaccine development purposes. Both proteins has been suggested as strong *H. pylori* vaccine candidate antigens (Dundon *et al.,* 2002, Kabir, 2007). In this study, the heat-induced N4830I-pML-λP_L expression system previously established (see introduction section), was used for the production of NAP and FlaA. A protocol for the purification of recombinantly produced *H. pylori* NAP of good quality, was successfully established, although the yield was limited.

Expression and purification of NAP

The total and sub-cellular expression of NAP was detected by SDS-PAGE, shown in fig. 11. The analysis of the N4830I-pML-λP_L-*napA* cultures shows expression of NAP in the induced total sample. The presence of NAP in the induced cell debris fraction shows that some of the protein is lost during the separation steps, which could be explained by insufficient cell lysis and/or adherence of released NAP to cell debris and non-lysed cells. The weak band that was seen in the inclusion body fraction is probably due to contamination from the cell debris fraction. Most importantly, large amounts of NAP are found in the cytosolic fraction and cultivating medium, the latter implies that the protein is released from the *E. coli* cells. However, when I repeated the experiment, limited amounts were seen in the cultivation medium, indicating that the high amounts seen in the initial experiment probably was due to release of NAP upon cell lysis. Anyhow, the observed localisation of NAP is beneficial in the subsequent protein purification procedure, since the target protein should preferably be in a soluble fraction to avoid the need for detergents.

The above-mentioned results confirms what has already been reported; that *H. pylori* NAP is localised in the bacterial cytosol and is released upon autolysis. NAP can bind to the external surface of the outer membrane (Namavar *et al.,* 1998). In such a location, NAP can mediate the binding of H. pylori to the cell surface via interactions with carbohydrate (Teneberg *et al.,* 1997).

Further characterisation of the cellular localisation of NAP after expression was performed by periplasmic preparation. The periplasmic space is the space seen between the plasma membrane and outer membrane in the Gram-negative bacteria. Periplasmic preparation provides a way to extract periplasmic proteins from bacterial cells. When the pelleted cells are placed in a hypertonic solution (*e.g.* sucrose), plasmolysis occurs. This causes the cell to loose water, the plasma membrane to shrink, and will also increase the permeability of the outer membrane. Subsequent adding of ice-cold $MgCl_2$ makes the plasma membrane swell, and press out the periplasmic proteins. The effect of periplasmic preparation can be exploited in the purification of periplasmic proteins. Thus, I wanted to check whether NAP was a periplasmic protein, since the cellular localisation of *H. pylori* NAP remains a subject of debate. The outcome of this experiment shows that NAP does not reside in the periplasmic preparation (fig. 10), hence the recombinant *H. pylori* NAP is not a periplasmic protein.

One of the earliest forms of protein purification was the use of ammonium sulphate precipitation. By increasing the concentration of ammonium sulphate in steps, differential preparation of the protein mixture will occur. Centrifugation of the precipitate results in “cuts” of protein populations, significantly enriching the purity of the starting material. The precipitate will normally resolubilize once the salt is removed, either by filtration, or by dialysis. Recent studies have reported the use of ammonium sulphate precipitation as an initial step in the purification of *H. pylori* NAP. However, the studies contradict each other in the case of whether NAP ends up in the precipitate, or remains in the soluble fraction. Kottakis *et al.* reported that NAP remained in solution after a 98% precipitation with respect to ammonium sulphate (Kottakis *et al.,* 2007).

Ceci *et al.,* on the other hand, reported that with two ammonium sulphate cuts at 30% and 60%, NAP precipitated at 60% concentration (Ceci *et al.,* 2007). The studies also differed with regard to the subsequent purification steps used to purify *H. pylori* NAP. In this study, the result of ammonium sulphate precipitation shows that NAP remains in solution (supernatant) after 60% saturation with respect to ammonium sulphate (fig. 13). The outcome of this purification step also shows that the majority of contaminating proteins ends up in the 60% precipitate (pellet), and are effectively removed by this procedure. With more time available to optimise this purification step, I would suggest a strategy of testing different ammonium sulphate cuts, in order to minimise the potential loss of NAP to the precipitate. Anyhow, the comparison of protein purity by SDS-PAGE, shows a remarkable difference before and after the use of ammonium sulphate precipitation. With NAP already in a soluble fraction, and no need for detergents, chromatography techniques was a natural choice for the next purification step. After dialysis against running buffer followed by concentration, the sample was run on a Resource Q anion exchange column. The eluted fractions containing NAP were pooled, dialysed and concentrated, and this sample were applied to a Superdex 200 gel filtration column. SDS-PAGE analyses after anion exchange chromatography (fig. 14 A) and gel filtration (fig.14 B) confirmed the presence of NAP in the eluted fractions corresponding to the NAP peaks marked in the chromatograms (fig.15 and 16). Overall, the results shows that the number of contaminating proteins is effectively reduced by the two chromatography steps. The final sample contains NAP of good quality (high purity), since no other proteins than NAP are detected in the pooled fractions after gel filtration. The protein concentration was determined by use of the Nanodrop ND-100 spectrophotometer, and the DC Protein Assay. The values obtained from the different methods were similar, hence the observed protein concentrations, although quite low, are reliable. Considering the initial expression level observed for NAP, which was very good, the limited yield (0.2 mg/ml) after the purification procedure, is surprising. From this overall result it is clear that a large amount of NAP is lost throughout the purification, possibly due to insufficient disruption of cells and/or too many buffer changes and centrifugation steps. With more time to improve the output of recombinant NAP, I would suggest further optimalisation experiments that explore the effect of changing the various parameters, *e.g.* expression (growth and induction time), cell disruption (sonication and enzymatic lysis), ammonium sulphate precipitation (different cuts), dialysis (less buffer changes and centrifugations/filtrations), chromatography (change buffers and columns).

Expression and purification of FlaA

To optimise the expression of FlaA, the effect of induction time on FlaA expression was explored. Cells were induced 2, 3 or 4 h, as well as over night, and the results were observed by microscopic and SDS-PAGE analysis, which clearly shows an increased FlaA expression level from 2 to 4 h. In the microscope, it was observed that both the size and number (one or two) of inclusion bodies per bacteria increased. The sizes of the FlaA protein bands on the SDS-gel increased correspondingly, and it was concluded that 4 h induction gives the best target protein/total protein ratio, since over night induction mainly resulted in larger amounts of the contaminating proteins, and not FlaA.

To further characterise the expression of FlaA, different cellular fractions were analysed by SDS-PAGE. The outcome of this fractioning shows that majority of the FlaA protein ends up in inclusion bodies. It should however be pointed out that the vast majority of the FlaA protein was pelleted already after 600 x g centrifugation and not, as expected, after the subsequent 5000 x g centrifugation. This was probably due to insufficient lysis of the cells and/or the formation of high molecular weight inclusion bodies.

SDS-PAGE analysis of the cell extract and isolated inclusion bodies, comparing the total protein amount of the two, shows that the purity did not improve much after preparation of inclusion bodies. This indicates that the inclusion bodies contain other proteins besides

FlaA, and not, as expected, consist mainly of the over-expressed target protein.
Ammonium sulphate precipitation worked very well for purifying NAP in a soluble fraction, and this technique was also tested with FlaA. However, SDS-PAGE analysis of the pellets and supernatant from 30% and 50% cuts with ammonium sulphate, shows that FlaA precipitates and ends up in the insoluble 30% pellet. Therefore, the dissolved inclusion bodies were dialysed against running buffer, and applied directly to a Resource Q anion exchange column, in an attempt to further purify FlaA. In one experiment, the concentration of the sample by a vivaspin column before the run, resulted in visible particles of precipitated protein. To prevent high pressure in the system, particles were pelleted by thorough centrifugation of the sample. Despite several attempts with different running buffers (pH, salt), and trying to avoid concentrating steps, FlaA proteins re-precipitated. This, in turn, created problems with particles clogging the column, and loss of significant amounts of FlaA in insoluble fractions during the dialysis, centrifugation and chromatography steps. The initial expression level of FlaA observed on SDS-gels was not as high as observed for NAP. This, in addition to the problems I experienced with precipitation throughout the purification procedure, contributed to a very small output of FlaA in the eluted fractions after anion exchange chromatography. Overall, the results of the FlaA experiments show that it is hard to purify this protein in its native form. The lack of recent publications concerning purification of recombinant *H. pylori* FlaA (few or none to my knowledge), might indicate that this is a challenge yet to be overcome by scientists. The purification problems with FlaA in this study, are possibly explained by the structural role of this protein. When building up the flagellar filament, flagellin polymerize (with itself, and possibly other proteins, the latter can also explain the contamination of other proteins in the inclusion bodies), and therefore tend to precipitate into insoluble complexes. Flagellin is known to be a protein that arrange itself in a hollow cylinder to form the filament of the bacterial flagellum. The filament consists of a long homopolymer of a single protein, flagellin, with a small cap protein at the end. Polymerization of flagellin occurs as a result of relatively conserved structures at the N and C termini, although the intervening regions of the protein are highly diverse (Donnelly and Steiner, 2002). In 2001, the crystal structure of a central proteolytic fragment of Salmonella flagellin was solved, contributing to an understanding of how these conserved structures are involved in filament formation (Samatey et al., 2001).
The purification strategy involving inclusion bodies was not very successful in this study, for reasons unknown. A lesson to be drawn from this might be that we should rethink the whole strategy of expression and purification of FlaA. Various alternative purification procedures can be thought of, *e.g.* directing the recombinant *H. pylori* FlaA to specific compartments, or to the flagellum of the *E. coli* host cell. What about the use of fusion proteins such as His-tags? Could other bacterial species be better hosts for expressing recombinant FlaA? Or maybe the best strategy to obtain FlaA in a pure, native form is to knock the flagellum directly of the *H. pylori*, and use the crude preparation for immunisation*?* In 1989, Geis *et al*. isolated the flagella of *C. jejuni* by mechanical shearing from the cell surface, sucrose density gradient centrifugation, and Sepharose CL-4B gel chromatography (Geis *et al*., 1989). The role of FlaA in *H. pylori* virulence and colonisation has already been mentioned(Kabir., 2007), but the potential of this antigen in vaccine development remains to be investigated. In a recent study, mice immunized with a preparation enriched for H. pylori flagella sheath proteins exhibited significantly reduced colonisation, equivalent to that observed in mice immunized with whole-cell lysate (Skene et al., 2007).

For more than ten years a vaccine against *Helicobacter pylori* har been the elusive goal of many investigators. Although attempts to produce a vaccine against *H. pylori* have failed in their ultimate goal, considerable knowledge has been developed on the pathogenesis and immunity of *Helicobacter* infections. Marshall and Schoep even suggests an alternative use of

this knowledge, namely, the use of *Helicobacter* species to deliver vaccines against other organisms (Marshall and Schoep, 2007).

Based on the very high prevalence diseases caused by *H. pylori,* and the emerge of antibiotic resistance among clinical isolates, there is a need for an effective vaccine against *H. pylori*. An important part of this vaccine development process will be to identify the main protective immune mechanisms against *H. pylori*. Several vaccine development strategies are currently being explored. One such strategy involves to make a cocktail of strong protective antigens, or a recombinant bacterial strain that express such antigens, that could be administered by a regimen that gives rise to effective immune responses in humans (Svennerholm and Lundgren, 2006). Potential *H. pylori* antigens to include in such a cocktail, would be NAP, *H. pylori* adhesin A (HpaA), Urease B subunit (UreB), and FlaA.

Conclusion

In this study, *H. pylori* Neutrophil-activating protein (NAP) and Flagellin A (FlaA) were cloned and expressed in *E. coli* N4830I-pML-λP_L-*napA* and N4830I-pML-λP_L-*flaA*, respectively. The majority of the produced FlaA protein ended up in inclusion bodies, whereas the majority of the NAP protein was detected in the soluble cytosolic fraction. NAP was purified by a combination of ammonium sulphate precipitation, anion exchange chromatography and gel filtration, resulting in recombinant NAP with good quality (high purity), although the yield was limited.

FlaA was isolated as inclusion bodies, however, no downstream purification was possible due to problems with re-precipitation. This suggests that alternative strategies should be considered for the future expression and purification of FlaA. Finally, further characterisation and optimisation of the protein expression are required to develop highly effective purification protocols for *H. pylori* NAP and FlaA.

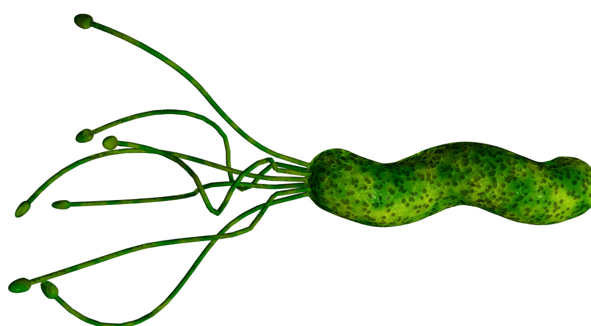

Fig. 17. *Helicobacter pylori* (www.hpylori.com/au).

References

Web pages

http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi
http://www.ncbi.nlm.nih.gov/books/bv.fcgi?rid=hp
http://www.hpylori.com/au
http://www.fermentas.com
http://www.invitrogen.com
http://www.mwg-biotech.com
http://www.omegabiotek.com
http://qiagen.com

Books

Marshall, B., 2002. Helicobacter pioneers: Firsthand Accounts from the Scientists who discovered Helicobacters 1892-1982. Wiley-Blackwell

Mobley, H.L.T., Mendz, G.L., Hazell, S.L., 2001. *Helicobacter pylori*: Physiology and Genetics. AMS Press (Available as e-book at http:/www.ncbi.nlm.nih.gov/books/bv.fcgi?rid=hp)

Sambrook, J., Russel, D., 2001. Molecular Cloning, A Laboratory Manual, 3rd edition.

Articles

Andersen-Nissen, E., Smith, K.D., Bonneau, R., Strong, R., Aderem, A., 2007. A conserved surface on Toll-like receptor 5 recognizes bacterial flagellin. J Exp Med 204 (2), 393-403.

Bizzozero, G., 1893. Ueber die schlauchformigen drusen des magendarmkanals und die bezienhungen ihres epithels zu dem oberflachenepithel der schleimhaut. Arch. Mikr. Anat. 42, 82.

Brosius, J., Erfle, M., Storella, J., 1985. Spacing of the -10 and -35 regions in the *tac* promoter. Effect on its *in vivo* activity. J Biol Chem. 260, 3539-3541.

Ceci, P., Mangiarotti, L., Rivetti, C., Chiancone, E., 2007. The neutrophil-activating Dps protein of *Helicobacter pylori*, HP-NAP, adopts a mechanism different from Escherichia coli Dps to bind and condense DNA. Nucleic acids Res. 35(7), 2247-2256

Cheng, X., Patterson, T.M., 1992. Construction and use of lambda PL promoter vectors for direct cloning and high-level expression of PCR amplified DNA coding sequences. Nucleic acids Res. 2 (17), 4591-4598.

Coghlan, J.G., Gilligan, D., Humphries, H., KcKenna, D., Dooley, C., Sweeney, E., Keane, C., O'Morain, C., 1987. *Campylobacter pylori* and recurrence of duodenal ulcers - a 12-month follow-up study. Lancet ii, 1109-1111.

De Boer H,.A., Comstock, L.J., Vasser, M., 1983. The tac promoter: a functional hybrid derived from the trp and lac promoters. Proc Natl Acad Sci USA. 80(1), 21-25

D´Elios, M.M., Amedei, A., Cappon, A., Del Prete, G., de Bernard, M., 2007. The neutrophil-activating protein of *Helicobacter pylori* (HP-NAP) as an immune modulating agent. Immunol Med Micriobiol. 50, 157-164

Donnelly, M.A., Steiner, T.S., 2002. The nonadjacent regions in enteroaggretative E. coli Flagellin are Required for Activation of Toll-like receptor 5. J.Biol. Chem. 277(43), 40456-40461.

Dundon, W.G., Nishioka, H., Polenghi, A., Papinutto, E., Zanotti, G., Montemurro, P., Del Giudice, G., Rappuoli, R., Montecucco, C., 2002, The neutrophil-activating protein of *Helicobacter pylori*. Int. J. Med. Microbiol. 291, 545-550.

Elvin, C.M., Thompson, P.R., Argall, M.E., Hendry, P., Stamford, N.P., Lilley, P.E., Dixon, N.E., 1990. Modified bacteriophage lambda promoter for overproduction of proteins in *E. coli*. Gene. 87, 123-126.

Geis, G., Leying, H., Sauerbaum, S., Mai, U., Opferkuch, W., 1989. Ultrastructure and chemical analysis of *Campylobacter pylori* flagella. J. Clin. Microbiol. 27, 434-441.

Goodwin, C.S., 1997. *Helicobacter pylori* gastritis, peptic ulcer and gastric cancer: clinical and molecular aspects. Clin. Infect. Dis. 25, 1017-1019.

Graham, D.Y., Lew, G.M., Evans, D.G., Evans, Jr D.J., Klein, P.D., 1991. Effect of triple therapy (antibiotics plus bismuth) on duodenal ulcer healing. A randomized controlled trial. Ann. Intern. Med. 115, 266-269.

Grant, R.A., Filman, D.J,. Finkel, S.E., Kolter, R., Hogle, J.M., 1998. The crystal structure of Dps, a ferritin homolog that binds and protects DNA. Nature Struct. Biol. 5, 294-303.

Harlow, E., Lane, D., 1988. Antibodies: A laboratory manual, Cold Spring Harbor , NY, Cold Spring Harbor Laboratory.

Kabir, S., 2007. The Current Status of *Helicobacter pylori* Vaccines: A Review. Helicobacter. 12, 89-102.

Kottakis, F., Papadopoulos, G., Pappa, E.V., Cordopatis, P., Pentas, S., Choli-Papadopoulou, T., 2008 (Epub. 2007). *Helicobacter pylori* neutrophil-activating protein activates neutrophils by its C-terminal region even without dodecamer-formation, which is a prerequisite for DNA protection - novel approaches against *Helicobacter pylori* inflammation. FEBS J. 275 (2), 302-317.

Macnab, R.M., 2003. How bacteria assemble flagella. Ann. Rev. Microbiol. 57, 77-100.

Marshall, B., Schoep, T., 2007. Helicobacter pylori as a vaccine delivery system. Helicobacter 12:s 2, 75-79.

Marshall, B.J., Warren, J.R., 1984. Unidentified curved bacilli in the stomach of patients with gastritis and peptic ulceration. Lancer. 1(8390), 1311-1315.

Mergulhao, F.J., Summers, D.K., Monteiro, G.A., 2005. Recombinant protein secretion in *Escherichia coli*. Biotechnol. Adv. 23, 177-202.

Namavar, F., Sparrius, M., Veerman, E.C., Appelmek, B.J., Vandenbroucke-Grauls, C.M., 1998. Neutrophil-activating protein mediates adhesion to sulfated carbohydrates on high-molecular-weight salivary mucin. Infect. Immun. 66, 444-447.

Nyström, J., Svennerholm, A.M., 2007. Oral immunization with HpaA affords protective immunity against *H. pylori* that is reflected by specific mucosal immune responses. Vaccine. 25(14), 2591-2598.

Rauws, E.A., Tytgat, G.N., 1990. Cure of dudenal ulcer associated with eradication of *Helicobacter pylori*. Lancet 335, 1233-1235.

Samatey, F.A., Imada, K., Nagashima, S., Vonderviszt, F., Kumasaka, T., Yamamoto, M., Namba, H., 2001. Structure of the bacterial flagellar protofilament and implications for a switch for supercoiling. Nature 410, 331-337.

Sauerbaum, S., Josenhans, C., *Helicobacter pylori* evolution and phenotypic diversification in a changing host. 2007 Nature rev. microbiol. 5, 441-452

Skene, C., Young, A., Every, A., Sutton, P., 2007. Heliocobacter pylori flagella: antigenic profile and protective immunity. FEMS Immunol. & Med. Microbiol. 50(2) 249-256.

Studier, F.W., Moffatt, B.A., 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. J Mol Biol. 189 (1), 113-130.

Svennerholm, A.M., Lundgren, A., 2006 Progress in vaccine development against *Helicobacter pylori*. FEMS Immunol Med Microbiol. 50, 146-156.

Sørensen, H.P., Mortensen, K.K., 2005. Advanced genetis strategies for recombinant protein expression in *Escherichia coli*. J Biotechnol. 115, 113-128.

Teneberg, S., Miller-Podraza, H., Lampert, H.C., Evans, D.J., Evans, D.G., Danielson, D., Karlsson, K.A., 1997. Carbohydrate binding specificity of the neutrophil-activating protein of Helicobacter pylori. J. Biol. Chem. 272, 19067-19071.

Terpe, K., 2006. Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems. Appl Microbiol. Biotechnol. 72, 211-222.

Zanotti, G., Papinutto, E., Dundon, W., Battistutta, R., Serveso, M., del Giudice, G., Rappuoli, R., Montecucco, C., 2002. Structure of the neutrophil-activating protein from *Helicobacter pylori*. J Mol Biol. 323, 125-130

Ingrid Lea

- 2008 -