Surface expression using the AIDA autotransporter: Towards live vaccines and whole-cell biocatalysis

Martin Gustavsson
M. Sc.

Royal Institute of Technology
Stockholm 2011
Martin Gustavsson (2011): Surface expression using the AIDA auto-transporter: 
Towards live vaccines and whole-cell biocatalysis. School of Biotechnology, Royal 
Institute of Technology (KTH), AlbaNova University Center, Stockholm, Sweden.

Abstract

The area of surface expression has gathered a lot of interest from research groups 
all over the world and much work is performed in the area. Autotransporters have 
been used for surface expression in Gram-negative bacteria. One of the more 
commonly used autotransporters is the Adhesin Involved in Diffuse Adherence 
(AIDA) of pathogenic Escherichia coli. The surface expression of enzymes and 
vaccine epitopes offer several advantages. Surface expressed enzymes gain similar 
properties to immobilised enzymes, mainly simplified handling and separation using 
centrifugation. Surface expressed vaccine epitopes can have longer half-lives inside 
the animal that is to be immunized and surface groups on the host cell can act as 
adjuvants, increasing the immune response and leading to a better immunisation.

However, while much basic research is directed towards mechanisms of surface 
expression using autotransporters there are few reports regarding production of 
surface expressed protein. Thus the aim of this work was the optimisation of the 
yield and productivity of surface expressed protein. Protein Z, an IgG-binding 
domain of Staphylococcal protein A, was used as a model protein for the investigation 
of which cultivation parameters influenced surface expression. The choice of 
cultivation medium gave the largest impact on expression, which was attributed to 
effects based on the induction of the native promoter of AIDA. The AIDA system 
was then used for the expression of two Salmonella surface proteins, SefA and H:gm, 
with potential for use as vaccine epitopes. SefA was verified located on the cell 
surface, and H:gm was found in the outer membrane of the host cell, though only 
in proteolytically truncated forms lacking the His6-tag used for detection. This 
proteolysis persisted in E. coli strains deficient for the outer membrane protease 
OmpT and was concluded to be dependent on other proteases. The removal of 
proteolysis and further optimisation of the yield of surface-expressed protein are 
important goals of further work.

Keywords: AIDA-autotransporter, Escherichia coli, live vaccines, surface expression
List of publications

This thesis is based on the following publications, which are referred to by their roman numerals:


Table of contents

1 Introduction 7

2 Surface expression 3
  2.1 Applications for surface expression 3
  2.2 Bacterial surface expression 6

3 Autotransport 10
  3.1 Autotransporter passenger proteins 11
  3.2 Secretion mechanism of autotransporters 12
  3.3 The autotransporter Adhesin Involved in Diffuse Adherence (AIDA) 18

4 Present investigation 21
  4.1 Strategy and model system 21
  4.2 Optimisation of surface expression (II) 24
  4.3 Surface expression of Salmonella epitopes 32
  4.4 Effects of proteolysis on surface expression (I, II) 34

5 Concluding remarks 38

6 Abbreviations 39

7 Acknowledgements 40

8 References 41
# Introduction

Recombinant protein production is today a large, multibillion industry. The proteins produced can be divided into industrial enzymes, with a revenue of approximately 3.3 billion dollars in 2010 [1], and therapeutic proteins reaching a revenue of 93 billions in 2010 [2]. This was all made possible by a discovery of Smith and co-workers in 1970. They discovered the DNA-modifying type II restriction enzymes, which led to big advances in the area of biotechnology [3]. Restriction enzymes provided the tool needed for manipulation of DNA and enabled the construction of recombinant DNA. Recombinant DNA technology has enabled the cloning and production of proteins of interest using well-studied production organisms, such as the *Escherichia coli*, *Saccharomyces cerevisiae*, *Bacillus* and *Aspergillus*.

Industrial enzymes can be broadly divided into food enzymes, enzymes used in the textile industry, enzymes for wood, pulp and paper processing and washing powder enzymes. An example of a large-scale enzymatic processes is the production of high-fructose syrup [4]. The use of enzymes for biocatalysis of the synthesis of fine chemicals is an important industrial area [5]. Enzymes bring many advantages compared to conventional, chemical catalysts, such as high catalytic efficiency and high enantiomeric selectivity. Furthermore, the use of organic solvents can often be avoided in favour of water, leading to more environmentally friendly processes. However, enzyme catalysis has its disadvantages, one being the cost of the enzyme production and purification. Another obstacle is that it is generally not possible to separate the enzyme from the product for reusing. This can be circumvented through the immobilisation of the enzyme on a solid support. However, this adds more steps to the production process and increases the cost. Another option is the use of whole cells, often yeast or bacteria, as biocatalysts. This removes the purification and immobilisation of the enzyme and enables easy separation of the catalyst from the reaction medium through for instance centrifugation. However, this adds a mass transfer barrier between the reactants in the reaction medium and the enzymes inside the cells.
Therapeutic recombinant proteins include for instance insulin [6] and recombinant factor VIII [7]. Vaccines are another large market, which is shifting from the use of traditional, attenuated pathogens towards the use of recombinant subunit vaccines. The use of subunit vaccines is advantageous compared to the use of attenuated bacteria or viruses, since the risk of the attenuated vaccine reverting to virulent forms is removed [8]. However, the half-life of a subunit vaccine inside animals and humans is low, leading to a weak immune response and inefficient immunisation.

Surface expression technology may be used to circumvent these problems. Surface expression is the process of producing a protein anchored to the surface of cells, for instance bacteria or yeast, or bacteriophages. By tethering the subunit vaccine to the surface of a non-pathogenic cell or virus particle it is readily accessible for recognition by the host immune system. Additionally, the surface of the cell or virus used as a host for surface expression can act as an adjuvant for the vaccine, thereby increasing the strength of the immunogenic response [8][9]. Finally, if the cell used for surface expression is a part of the natural flora of the host it is possible for it to colonize the host and persist for a longer time, further increasing the immune response to the presented subunit vaccine.

Surface expressed enzymes could bring the ease of separation of immobilized enzymes and whole-cell biocatalysts while avoiding some of the disadvantages. The enzyme is produced and ready for use with minimal downstream operations compared to conventional immobilisation techniques, while avoiding the diffusion barrier associated with whole cell catalysts with intracellular expression [10].

Much work has been put into the area of surface expression of both enzymes and vaccines. However, for the realisation of large-scale use of surface expression technology there are some requirements that have to be met. The system used for achieving surface expression must be flexible with respect to the structure and size of proteins that can be expressed, in order to enable the use in different applications. Furthermore, it must be possible to express large amounts of protein on the surface on the chosen host cell, and this high level of expression must be possible to maintain at high cell densities in order to achieve economical processes. While much work has been reported regarding surface expression there are few studies on the production aspect.
The aim of this work was to address these problems using autotransport [11] to achieve surface expression, and to optimise the yield of surface expressed protein as well as the productivity of surface expressing cells. Autotransporters have successfully been introduced into non-pathogenic laboratory strains lacking the export systems of pathogenic strains, and has since gathered much interest for use in recombinant surface expression. This is due to autotransporters having been found to readily work for the expression of large proteins and to high protein concentrations ($10^5$ copies per cell) [12]. *E. coli* has been employed frequently due to being a good expression host since it is well studied and grows readily with a high rate on cheap, defined media. Furthermore, as laboratory *E. coli* export few proteins due to the mentioned lack of export systems associated with virulence, leading to recombinant proteins expressed on the surface being relatively pure.

2 **Surface expression**

Smith and co-workers were the first to report surface expression back in 1985. They successfully surface displayed peptides by fusing them to coat proteins of filamentous bacteriophages [13]. This discovery was followed shortly by the first reports of the use of bacteria for surface expression in 1986 [14, 15], and later the use of yeast [16]. The area of surface expression has since gathered much interest and surface display of peptides and proteins has been suggested for many different applications. Some applications for surface expression are described in more detail below.

2.1 **Applications for surface expression**

*Library screening*

Surface display technology has been used extensively for the screening of randomised protein libraries for binders to target molecules of interest [9, 10]. Library screening using surface display is based on the fundamental principle that each cell contains the genetic information about the peptide displayed on its surface. Initially phage display was the go-to technique for library screening, but as
cell surface display systems have become more sophisticated they have become more common [17]. The major advantage of cell surface displayed libraries compared to phage display is the possibility of using optical methods such as fluorescence activated cell sorting (FACS) for high-throughput screening [9]. This is possible due to the large size of bacteria when compared to phages. Another contributing factor is the recent improvements in flow cytometry technology with reduced background noise, allowing the detection of particles in the size of bacteria.

**Live vaccines**

Bacterial cells with surface expressed vaccine epitopes offer several advantages. First, the usage of non-pathogenic host cells leads to a safer vaccine compared to using attenuated pathogens. When using inactivated cells as vaccines there is a risk of reversion back to the pathogenic form. In addition the surface of the bacterium can act as an adjuvant, enhancing the immune response to the vaccine [8, 9]. Furthermore, if the bacterium used can colonise the recipient it is possible to achieve exposure to the antigen for a longer time span than is possible using attenuated or subunit vaccines, giving a stronger immunisation [8].

Examples of surface expression used for development of live vaccines include surface expression of M6 protein of *Streptococcus pyogenes* in *Streptococcus gordonii* with successful immune response in mice [18, 19]. *E. coli* has been used for display of antigens from hepatitis B and foot and mouth disease [20, 21]. Other examples of bacteria suggested for vaccine development are the food grade *Staphylococcus carnosus* [22], *Staphylococcus xylosus* [23] and *Lactococcus lactis* [24].

**Whole-cell biocatalysis**

Synthesis of enantiomerically pure chemicals using enzymes as biocatalysts is an important area of biotechnology [5]. Traditionally there are two main ways to approach a biocatalytic process. One is the use of purified enzymes, in solution or immobilized. The other is the use of whole cells expressing enzymes as intracellular biocatalysts [10]. Both approaches have advantages and disadvantages. When working with pure enzymes the main disadvantage is the cost associated with
purification of the enzyme, which can be a major issue in a cost sensitive process [25]. Furthermore, if the enzyme is not immobilized it is more difficult to separate from the final product and the enzyme is not reusable for more than one reaction cycle. While enzyme immobilization may overcome this drawback it further adds to the cost of processing of the biocatalyst. Whole-cell processes overcome the difficulties of separation of the catalyst, since the cell that contains the enzyme is large enough to enable easy separation using centrifugation or filtration. This means that the catalyst can be reused as long as the enzyme and cell used is stable under the reaction conditions of the chemical synthesis. However, whole-cell systems have a disadvantage. The substrates of the reaction have to pass into the cell for reaction to occur. This may lead to mass transport becoming a limitation, leading to the inability to perform reactions of interest [5, 10]. Using surface expressed enzymes bring the advantages of both of the approaches. No downstream, other than cell harvest, is needed. Just as in whole cell systems with intracellular enzyme the cells are easily separated from the reaction medium, but the mass transfer limitation is avoided since the enzymes are exposed on the cell surface.

**Bioresomediation**

Contamination with organic and inorganic compounds, for instance pesticides and heavy metals, is a growing environmental problem. Microbial cells with surface-expressed enzymes may be used to degrade organic pollutants in the environment [26]. Surface expression brings the advantage of better accessibility for the organic pollutants to bind and react in the enzyme compared to cytoplasmic enzyme expression. Several studies have been reported using surface expressed enzymes for this purpose, including surface expression of organophosphorous hydrolase for degradation of organophosphate pesticides [27].

Contrary to organic pollutants it is not possible to degrade heavy metals. They must instead be removed from the environment by other methods such as adsorption. One method that has been studied is the use of microorganisms for adsorption of metal ions [28, 29]. Most studies have focused on the intracellular uptake of the metal ions into the microorganism combined with cytoplasmic overexpression of detoxifying proteins, mainly metal chelators, for protection [30]. A main disadvantage of this is that the accumulated metal ions are difficult to
recover without breaking the cells, rendering it impossible to reuse the cell [29, 30]. This drawback may be circumvented by surface expression of metal-binding peptides, which enables recovery of the metal ions without cell breakage. This strategy has been evaluated for surface display using bacteria [31-33] and yeast [34]. Even bacteriophages have been suggested for use in the related application of mineral separation in ore mining [35].

**Biosensors**

Cells with surface expressed protein have the potential to be used as biosensors for detection of for instance pollutants. This can be achieved through the expression of hydrolytic enzymes that degrade the target compounds into molecules that can be detected, for instance through a pH or colour change. One such example is the use of the aforementioned organophosphorous hydrolase for rapid detection of pesticides [36]. José et al have used another approach. They expressed an antibody-binding domain of staphylococcal protein A on the surface of *E. coli* and used the resulting cell for amplifying the signal of a surface plasmon resonance based detector [12].

**Biofuel production**

Production of ethanol from cellulosic waste is an attractive solution to the impending shortage of oil. Yeast readily converts glucose to ethanol, but is unable to process cellulose. Therefore pre-processing is required to break down the cellulose to simple sugars. One solution that has been proposed is the expression of cellulolytic enzymes on the surface of *Saccharomyces cerevisiae*, and this approach has been demonstrated to work in laboratory scale [37].

### 2.2 Bacterial surface expression

This study has been focused on the surface display on bacteria. Yeast based surface expression systems have similar properties as bacterial, in that they offer easier handling and purification of the protein in question due to their size.
Additionally yeast systems have the advantage of providing glycosylation, which may be an advantage when working with non-bacterial proteins [16]. The commonly used *Saccharomyces cerevisiae* is also a GRAS (generally regarded as safe) organism. However, yeast cells grow slower than bacteria such as *E. coli*, which is an obvious processing disadvantage. Furthermore, yeast cells are less suitable for vaccine development due to having weaker antigenic properties than bacteria. Many different strategies have been used for recombinant surface expression in bacteria, but they all share a common feature: the fusion of the protein that is to be displayed on the surface with a protein that is naturally incorporated into the cell wall or outer membrane. The choice of the anchoring protein is where the strategies differ. The first thing deciding which strategy can be used is the choice of the host cell. Gram-negative and Gram-positive surface expression requires different strategies due to the significant differences in structure of the cell surface. The different classes of proteins that have been used for surface expression are discussed below. Table 1 shows a summary of these proteins and their advantages and disadvantages for surface expression.

**Table 1:** Fusion proteins used to achieve surface expression.

<table>
<thead>
<tr>
<th>Fusion protein</th>
<th>Example</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gram-positive</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cell-wall bound</td>
<td>SpA</td>
<td>Tightly anchored</td>
<td>Needs cell-wall removal</td>
</tr>
<tr>
<td>Membrane bound</td>
<td>DppE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gram-negative</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Porins</td>
<td>OmpA, LamB, OmpC</td>
<td>High expression</td>
<td>Small proteins</td>
</tr>
<tr>
<td>Lipoprotein</td>
<td>Ice nucleation protein</td>
<td>Large proteins</td>
<td></td>
</tr>
<tr>
<td>Autotransporters</td>
<td>AIDA, IgA1 protease</td>
<td>Large proteins, Disulphide bridges</td>
<td>High expression difficult</td>
</tr>
<tr>
<td>Flagellin</td>
<td>FlicC, FlicD</td>
<td>High expression</td>
<td>Small peptides</td>
</tr>
<tr>
<td>Fimbriae</td>
<td>FimA, FimH</td>
<td>High expression</td>
<td>Small peptides</td>
</tr>
</tbody>
</table>
The main difference in Gram-negative surface expression compared to Gram-positive is the existence of the second, outer cell membrane in Gram-negative bacteria. This adds an additional obstacle that needs to be overcome, since the protein that is to be displayed needs to be translocated across two membranes instead of just one. However, there is a major advantage in Gram-negative systems: the existence of a very well studied host in *E. coli*. *E. coli* makes for an excellent host due to its rapid growth and ease of transformation [38]. Additionally, laboratory *E. coli* strain such as K12 lack the export systems employed by pathogenic strains or secretion of various proteins involved in virulence. This provides an advantage since any recombinant proteins that are surface expressed will be relatively pure due to a lack of other secreted proteins. However, this lack of export systems also pose an obstacle for recombinant surface expression, since the machinery required is missing. Several different strategies have been devised to circumvent this, mainly based on fusion to different natural surface proteins.

Bacterial surface expression was first achieved using Gram-negative outer membrane protein fusions back in 1986 [14, 15]. These outer membrane proteins belong to the porin family and have β-strands spanning the outer membrane of the bacterium. The β-strands are connected by unstructured loop regions and form a β-barrel structure. The strategy used for surface expression of recombinant proteins has been insertion inside these loop regions on the surface-exposed side, and a variety of different membrane proteins have been used as carriers [10, 39]. Two examples are the usage of OmpA and LamB for surface display [14, 15]. This strategy has proven successful for a variety of recombinant proteins, but has been found to be restrictive regarding the size of the surface-expressed proteins due to disruption of the structure of the membrane protein when the insert is too large. Only smaller proteins have been possible to surface-express, typically up to around 70 aa with the exception of OmpC, which has been shown to enable expression of a polypeptide of 162 aa [9, 40].

A different class of surface proteins that have been widely studied for surface expression is the type Va autotransporters. The use of autotransporters has proven advantageous due to the possibility to surface express large proteins with high copy numbers per cell, as well as the possibility for the surface expressed protein to dimerize on the cell surface [41]. As the use of autotransporters for surface expression is the focus of the present study the use and mechanism of
autotransporters will be discussed in more detail below.

Lipoproteins differ from the β-barrel outer membrane proteins in that they are anchored to the cell membrane through covalent linkage to a lipid moiety [10]. Several different lipoproteins have been evaluated as vehicles for surface expression. One example of a lipoprotein that has been used extensively for surface expression is the *Pseudomonas syringae* ice nucleation protein [42]. It has been used for display of for instance enzymes and antigens [17], and is considered as one of the more promising vehicles for surface expression in Gram-negative bacteria [10]. TraT, an *E. coli* surface protein involved conjugation [43], is another example of a lipoprotein that has been used for surface expression [44].

The final class of proteins that has been used for surface expression in Gram-negative bacteria is the subunits of surface appendages such as flagella and fimbriae [10, 45]. Fimbriae are interesting for surface display for several reasons, including being present at high copy-numbers per cell (there can be around 500 fimbriae per cell, each made up of approximately 1000 identical subunits) and also being strongly immunogenic which could be advantageous for use as vaccine adjuvants [45]. However, the size of the expressed peptide is rather limited, with reports of 10-30 aa [45]. Flagellin subunits FlicC and FlicD that make up flagella have also been used for surface expression by replacing a non-conserved, central part of the flagellin sequence with the peptide of interest. This was first demonstrated by the presentation of egg-white lysozyme [46]. Again a major drawback with flagellar display is that the acceptable size of the presented peptide is rather small at around 60 aa [9].

The surface of Gram-positive bacteria is characterized by a thick peptidoglycan cell wall and a lack of outer membrane. There are many known surface proteins in Gram-positive bacteria [47] that could have potential for surface expression of recombinant fusion proteins. These proteins can be divided in two main classes: cell wall bound and cell membrane bound proteins [10]. Of the known cell wall anchored proteins staphylococcal protein A is the best studied. It has an N-terminal signal peptide targeting it for translocation over the cell membrane, four to five IgG-binding domains, a cell wall anchoring domain and a cell wall sorting signal targeting the protein for incorporation in the cell wall [10]. Protein A has been exploited for surface expression by fusing the cell wall binding domain X and the
sorting signal M with a protein of interest. Fusion proteins have been successfully expressed in several Gram-positive strains using this system, including *Staphylococcus eburneus* [48] and *Staphylococcus xylosus* [49]. This system has been used for the display of vaccines, but also for expression of metal binding peptides and for library screening.

The second class of Gram-positive proteins used is membrane-anchored proteins. Examples include the lipoprotein DppE of *Bacillus subtilis* [50]. This system had the disadvantage that the cell wall had to be stripped from the bacterium in order for the protein to become accessible on the cell surface. Thus, it is debatable if this should really be considered surface expression since the actual cell surface needs to be removed to achieve surface exposure.

## 3 Autotransport

Autotransport, also known as type Va secretion, is used by Gram-negative pathogens to secrete a variety of virulence factors [51]. The first autotransporter discovered was IgA1 protease from *Neisseria gonorrhoeae*, which was discovered in 1987 [52]. Many additional autotransporters have been identified since then and the autotransporter family is now known to contain more than 700 members, making it the largest family of secreted proteins in Gram-negative bacteria [53]. In their functional form autotransporters consist of two domains: a β-barrel that integrates into the outer membrane and a surface-exposed “passenger” protein domain [54]. However, autotransporters are synthesised as pro-proteins with an N-terminal signal peptide for transport over the inner membrane [55]. Additionally, the C-terminal part of the passenger domain, termed the “junction” region, has been found to play an important part in translocation and folding, as will be discussed below. An overview of the domain organisation can be seen in figure 1.
3.1 Autotransporter passenger proteins

The passengers of autotransporters fulfil diverse functions, most of them related to host virulence [51]. These functions include cell adhesion, proteolysis, toxins and many more [51, 56]. The passenger proteins of autotransporters are generally large, frequently with molecular weights of >100 kDa [56]. This makes autotransporters interesting candidates for recombinant surface expression, since many of the other systems used for surface expression are very restrictive regarding the size of the surface-expressed protein, as previously mentioned. Table 2 shows a summary of a few autotransporters, their function and the size of the passenger protein.

<table>
<thead>
<tr>
<th>Autotransporter</th>
<th>Bacterial species</th>
<th>Function</th>
<th>Passenger size kDa</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIDA-I</td>
<td><em>E. coli</em></td>
<td>Adhesion</td>
<td>79.5</td>
</tr>
<tr>
<td>Ag43</td>
<td><em>E. coli</em></td>
<td>Biofilm formation</td>
<td>49.8</td>
</tr>
<tr>
<td>IgA1 protease</td>
<td><em>N. gonorrhoea</em></td>
<td>Protease</td>
<td>106</td>
</tr>
<tr>
<td>EspP</td>
<td><em>E. coli</em></td>
<td>Cytotoxic protease</td>
<td>105.4</td>
</tr>
<tr>
<td>Hbp</td>
<td><em>E. coli</em></td>
<td>Hemoglobin protease</td>
<td>111.8</td>
</tr>
<tr>
<td>BrkA</td>
<td><em>B. pertussis</em></td>
<td>Serum resistance</td>
<td>68.4</td>
</tr>
<tr>
<td>Pertactin</td>
<td><em>B. pertussis</em></td>
<td>Adhesion</td>
<td>60.2</td>
</tr>
</tbody>
</table>
Even though the function, amino acid sequence and length of the passenger proteins are diverse, many passengers show a common structural theme: a right-handed $\beta$-helix (figure 2). In fact a majority of the crystal structures of autotransporter passengers solved to date share this basic scaffold [57-59]. Furthermore, it has been predicted that a vast majority of the known autotransporters have passengers that form this kind of $\beta$-helical structure [60]. This appears to not be a requirement for surface expression, however, based on the successful display of a variety of recombinant passenger proteins of different folds using autotransporters.

**Figure 2:** Passenger proteins of the autotransporters Hbp (left, PDB id 1WXR) and pertactin (right, PDB id 1DAB)

### 3.2 Secretion mechanism of autotransporters

The mechanism used by autotransporters to achieve expression in the extracellular environment is beautiful in its simplicity. Both the protein that is to be exported, called the 'passenger' as previously mentioned, and the transporter itself are synthesized as a single protein containing all the functionality needed for transport. The protein is made up of three major parts: an N-terminal signal peptide targeting the protein for transport over the inner membrane through the Sec-dependent secretion pathway, the passenger protein and a C-terminal $\beta$-barrel [54]. An overview of the secretion mechanism of autotransporters is shown in figure 3.
Figure 3: Mechanism of autotransport. 1: The autotransporter is synthesised in the cytoplasm as a pro-protein with an N-terminal signal peptide for Sec transport. 2: The unfolded autotransporter is translocated to the periplasm and the signal peptide is cleaved off. 3: The β-barrel is inserted into the OM. 4: The passenger is translocated through the pore formed by the β-barrel and folds on the cell surface. 5: Depending on the autotransporter used the passenger may be autoproteolytically cleaved off and released to the medium.

Transport to the periplasm

Autotransporters are transported over the inner membrane via the Sec system [61]. An N-terminal signal peptide is required for targeting the autotransporter for export through Sec, and is proteolytically removed after translocation. A majority of the known autotransporters have signal peptides of 20-30 aa residues in length, which is the usual size for export through Sec. However, around 10% of the autotransporters have unusually long signal peptides due to an extension in the N-terminal part. These long signal peptides are at least 42 aa long [56]. The C-terminal part of these signal peptides is similar to regular signal peptides consisting of a charged region and a hydrophobic region, as well as a recognition site for a signal peptidase. The N-terminal extension shows a high homology, and contains a further
hydrophobic region as well as a charged region [56]. It is unclear what the function of the extended signal peptide is, but it has been suggested that it plays a role in the transport through the periplasm by temporary tethering the autotransporter to the inner membrane, thereby preventing misfolding [62]. This could assist in keeping the passenger in the unfolded conformation required for translocation over the outer membrane. Examples of autotransporters with long signal peptides include the *Bordetella pertussis* autotransporter BrkA [63], the *E. coli* autotransporter AIDA as well as in all autotransporters of the SPATE (serine protease autotransporters of *Enterobacteriaceae*) family [51].

**Surface translocation**

Once transported over the inner membrane the signal peptide is proteolytically cleaved off, leaving the passenger protein and β-barrel. The protein is then faced with the main obstacle in Gram-negative secretion: transport over the outer membrane. The way autotransporters overcome this obstacle is a subject of some controversy, and several different mechanisms have been suggested. The one presented here is the “classical” autotransport mechanism initially proposed by Pohlner *et al.* [52]. Once the autotransporter has been translocated over the inner membrane the β-barrel is thought to fold and insert into the outer membrane, forming a pore. Following insertion, the passenger protein is translocated through the pore to the cellular surface [64]. This translocation can proceed in two ways, either with the N-terminal of the passenger “threading” through the pore or the C-terminal being pulled through in a hairpin conformation. Recent findings suggest that the latter is the case [65]. It is believed that the passenger must remain in an unfolded state during translocation due to space limitation in the narrow β-barrel pore, as shown by crystal structures [66, 67]. Finally the passenger must fold into its native conformation once outside the outer membrane. In fact, continuous folding as the passenger exits the pore has been suggested to be the driving force of the translocation [60]. Once folded on the cell surface the fate of the passenger diverges depending on which autotransporter is studied. Some passengers remain covalently linked to their β-barrel domains and are presented on the cell surface. Other autotransporters possess autoproteolytic activity with a cleavage site between the β-barrel and passenger domains. The passengers of these autotransporters are cleaved from their β-barrel and are released to the surrounding environment.
While beautiful, this proposed mechanism leads to some limitations. Mainly due to the requirement for the passenger to remain in an unfolded state until translocation is complete. This can be problematic when working with certain passenger proteins, for instance proteins containing cysteines. As the periplasm is a reducing environment there is a risk that the cysteines will form disulphide bridges after transport over the inner membrane, thereby fixing the passenger in a translocation-incompetent conformation that cannot pass through the pore formed by the β-barrel, leading to the protein stalling on the periplasmic side of the outer membrane. Proteins stalled in this manner may then be subject to proteolysis catalysed by periplasmic proteases. However, premature folding due to disulphide bond formation has been reported possible to overcome by the addition of reducing agents such as mercaptoethanol to the growth medium, which prevents oxidation of cysteines in the periplasm [11].

The junction region

The C-terminal part of the passenger domain, termed “junction” region [68], has been shown to play an important part in translocation in many autotransporters. An increased sequence conservation can bee seen in this part of the passenger proteins of the SPATE family, and mutants with insertion in the conserved positions has shown a higher degree of translocation deficiency compared to mutants with insertions in other positions for the serine protease EspP [68]. Observations that the junction region is required for passenger folding in some autotransporters lead to the suggestion that it acts as an intramolecular chaperone [69]. It has been shown that the C-terminal part of the passenger domains of Bordetella pertussis pertactin and E. coli Pet is more stable than the rest of these passenger domains [60, 70]. It was thus suggested by Junker et al. that the folding of this stable part of the passenger is part of the driving force for passenger translocation over the outer membrane [65]. The folded C-terminal can then act as a scaffold for the folding of the rest of the passenger protein, which may explain the observed autochaperone properties of the junction region.
Influence of periplasmic chaperones

Folding and insertion of outer membrane proteins has been found to be influenced by periplasmic chaperones. The main chaperones involved in assisting outer membrane proteins are Skp and SurA [71]. Skp is a generic chaperone that captures unfolded proteins after translocation through the Sec system and assists in folding and outer membrane insertion. SurA is a cis/trans peptidyl prolyl isomerasers mainly involved in the folding and insertion of trimeric β-barrel proteins into the outer membrane. Both SurA and Skp has been reported by Ruiz-Perez et al. [72] to interact with the serine protease autotransporter EspP during binding assays and Qu et al. also showed that Skp can interact with the C-terminal domain of the autotransporter NalP [73]. Decreased expression levels of EspP was also found by Ruiz-Perez et al in SurA and Skp mutants, confirming the influence of these chaperones. Sauri et al. showed that the autotransporter Hbp accumulates intracellularly in a SurA negative strain with reduced levels of surface translocated protein, further showing the importance of this chaperone [74]. Together these studies highlight the importance of SurA and Skp for successful processing of autotransporters. In addition to Skp and SurA it has been shown that DsbA can influence the translocation of passenger proteins over the outer membrane. DsbA is responsible for the correct formation of disulphide bonds in proteins exported to the periplasm. The formation of disulphide bonds in passenger proteins inside the periplasm is a problem, since it can cause the folding of the passenger into translocation incompetent forms. It has been shown that DsbA mutant strains can be used to circumvent this [75, 76].

Influence of the β-Barrier Assembly Machinery (BAM) complex

This insertion of autotransporter β-barrels into the outer membrane was initially thought to be autonomous, but has recently been shown to be influenced by the β-Barrier Assembly Machinery (BAM) complex [74, 77, 78]. The BAM complex is a hetero oligomer consisting of five subunits, BamA-BamE, and has been implicated to be involved in the folding and insertion of most of the known outer membrane proteins [77]. BamA (previously called Omp85 or YaeT) is the central component of the complex and forms a β-barrel pore in the outer membrane with a predicted diameter of approximately 2.5 nm [79].
Rossiter et al. found that both BamA and BamD are required for secretion of the passengers of the autotransporters Pet and Ag43, while no effect on surface expression levels was seen in strains lacking BamB, BamC or BamE [77]. This was consistent with the previous findings that BamA, but not BamB, is required for surface expression of the Hbp autotransporter, reported by Sauri et al. [74]. Going by these results it is tempting to suggest a modified mechanism for autotransporter translocation, where the autotransporter β-barrel is folded and inserted into the outer membrane by the BAM complex, followed by passenger translocation through the autotransporter pore. However, it has been suggested that the BAM complex may not only be involved in insertion of the β-barrel into the outer membrane, but may also play a role in passenger translocation [80, 81]. This model would enable the whole autotransporter to fold inside the cell and then be translocated through the BamA pore. The proponents of this mechanism find support from several reports of the successful translocation of passengers that have been shown to exist in folded conformations in the periplasm [75, 82], which should not be possible according to the classical translocation mechanism.

*Use of autotransporters for recombinant surface expression*

The potential for recombinant protein secretion using autotransporters was realised soon after the discovery of the first autotransporter. The surface display of cholera toxin B was reported by Klauser and co-workers only three years after the discovery of *N. gonorrhoeae* IgA1 protease [11]. Since then much research has been directed towards the use of autotransporters for surface expression of a variety of recombinant proteins. The main advantage of autotransporters compared to many of the other expression systems is that they are relatively flexible regarding the size of the recombinant passenger. As previously mentioned many natural passenger proteins are over 100 kDa in size, which shows that autotransporters have the potential for transporting large proteins. Furthermore, it has been shown that it is possible for passenger proteins to dimerize on the cell surface, enabling the expression of active, dimeric proteins [41]. The two main autotransporters that have been used for recombinant expression are the previously mentioned IgA1 protease and Adhesin Involved in Diffuse Adherence (AIDA) from pathogenic *E. coli* [83]. Some examples proteins include functional enzymes [41, 84, 85], enzyme inhibitors [86], and epitopes [87].
3.3 The autotransporter Adhesin Involved in Diffuse Adherence (AIDA)

The Adhesin Involved in Diffuse Adherence (AIDA) [83] is an autotransporter that has been used frequently for recombinant surface expression. The wild-type function of AIDA is to enable enteropathogenic *E. coli* to adhere to epithelial cells in the intestines [83]. AIDA has been considered advantageous for recombinant surface expression compared to other autotransporters, such as IgA1 protease, due to being a native protein of *E. coli* [88]. Furthermore, AIDA has been shown to be more effective than IgA1 protease for recombinant surface expression [89]. These factors have made AIDA the autotransporter of choice for many research groups.

Like other autotransporters AIDA is synthesised with an N-terminal signal peptide followed by a passenger domain and finally a C-terminal domain containing the membrane-spanning β-barrel (figure 4a). The C-terminal domain (AIDA\(^C\)) consists of two separate parts, \(\beta_1\) and \(\beta_2\). \(\beta_2\) is the β-barrel domain consisting of 14 membrane-spanning β-strands, while \(\beta_1\) forms a short β-helix cap on the outside of the barrel and is the junction region of AIDA (figure 4b) [90]. The two β-domains are connected by a short α-helix spanning the pore formed by \(\beta_2\).

The AIDA passenger is proteolytically cleaved after translocation to the cell surface. This proteolysis is autocatalytic and is believed to be catalysed by two acidic amino acid residues (Glu\(^{897}\) and Asp\(^{878}\)) located in the junction region [91]. The mature AIDA passenger remains associated with the cell surface after the proteolysis, but can be released by brief heat treatment [92]. The cleavage site for autoproteolysis has been identified and marks the border between the passenger domain and AIDA\(^C\). It has been shown that the autoproteolytic release of the passenger domain can be removed by a single amino acid mutation in this cleavage site [93].
Figure 4: Structure of the AIDA autotransporter. A: Domain structure of AIDA. From N- to C-terminal there is a signal peptide for Sec transport (SP), a passenger domain with adhesin properties and a C-terminal translocation unit termed AIDA\textsuperscript{C}. AIDA\textsuperscript{C} is divided into two domains, \( \beta_1 \) and \( \beta_2 \), where \( \beta_1 \) is the junction region and \( \beta_2 \) is the C-terminal \( \beta \)-barrel. Autoproteolysis takes place between \( \beta_1 \) and the passenger domain. B: Theoretical structure of AIDA\textsuperscript{C} modelled using the Phyre2 online tool\cite{94}.

The functionality of mature AIDA-I is additionally dependent on modification by the protein autotransporter adhesin heptosyltransferase (AAH) \cite{95}. AAH is located in the cytoplasm and transfers heptosyl residues to the AIDA passenger. In other words, AIDA is an \( E. \text{coli} \) protein that undergoes glycosylation. It has been shown that this glycosylation takes place in the cytoplasm prior to export to the periplasm \cite{96}. The \( aah \) gene is located directly upstream of the \( aidA \) gene encoding for AIDA. It was recently shown by Benz and Schmidt that AIDA can be transcribed from
two promoters, one upstream of \textit{aah} producing a bis-cistronic mRNA coding for both AIDA and the heptosyltransferase, and one located between the two genes giving rise to mRNA coding for only AIDA \cite{97}. Factors influencing the induction of these promoters were studied in the report by Benz and Schmidt \cite{97}. They found that the growth medium and growth phase of the cells had a big impact on expression from both promoters. Interestingly, expression was significantly higher in rich LB medium compared to a defined salt medium, and both promoters showed a higher expression in the stationary phase. A consensus sequence for the stationary phase sigma factor coded by \textit{rpoS} was found in the \textit{aah} promoter which could explain the increased expression from this promoter in the stationary phase, but no such consensus was present in \textit{aidA}. Regardless of the mechanism this is not a surprising result. Many genes involved in adhesion are induced by nutrient availability and/or stress \cite{98}. In this way they enable the bacterial cells to colonise a host when nutrient availability is high or protect themselves from different stresses.

One stress that could be involved in regulation is the stringent response \cite{99}. The stringent response is triggered by either carbon starvation or a lack of charged aminoacyl-tRNA during protein synthesis and leads to rapid synthesis of guanosine tetraphosphate (ppGpp) by two enzymes produced by the genes \textit{spoT} and \textit{relA}. \textit{spoT} is responsible for ppGpp production in case of carbon starvation while \textit{relA} is activated when there is a lack of amino acids. ppGpp rapidly halts rRNA and protein synthesis, and shifts gene expression in the cell to biosynthetic genes through altering the affinity of RNA polymerase to different sigma factors \cite{99}. 


4 Present investigation

Much research is going into the area of surface expression and several applications have been suggested, including whole-cell biocatalysis and live vaccine development. Recently autotransporters have been widely employed as vehicles for achieving surface expression in Gram-negative bacteria, for instance the frequently used E. coli. However, even though there exists a large number of studies on the mechanism of autotransporters there is a lack of studies on factors affecting surface expression levels during cultivation. This information is crucial for the formulation of production processes of surface expressed protein, which in turn is a requirement for the industrial application of surface expression technology. Therefore the aim of this study was to achieve stable surface expression levels with a high yield of protein per cell and a high productivity of surface expressing cells.

4.1 Strategy and model system

The surface expression system that was chosen is based on the AIDA autotransporter, described in 3.3. The vector used was originally described by Casali et al. [100] and consists of the native AIDA signal peptide, a linker region composed of the C-terminal 54 amino acid residues of the AIDA passenger followed by AIDA.C. The recombinant passenger is cloned in frame between the signal peptide and the linker, creating a construct with the surface expressed passenger forming the N-terminal domain in the mature protein. The autoproteolytic cleavage site between AIDA.C and the passenger has been removed in this vector, enabling the recombinant passenger to remain tethered to the outer membrane. The whole construct is under the control of the native aidA promoter. A schematic representation of the surface expression construct can be seen in figure 5.

![Figure 5: Schematic representation of autotransport vector. SP: signal peptide, L: linker.](image-url)
The present study was divided in two main steps. First, the influence of external factors on surface expression was studied in order to develop a methodology for achieving optimised cultivation conditions that provide a high, stable level of surface expression (II). To achieve this we wanted to express a protein that was easy to produce and transport to the periplasm, in order to minimise the impact of protein chosen on expression and transport. This would allow us to isolate the effects related to surface expression. The protein chosen as a model was protein Z [101]. Z is a small (6.6 kDa), soluble protein consisting of the synthetic B domain of Staphylococcus aureus protein A (SpA). SpA is a natural surface protein in S. aureus, which gives it a good potential for successful export to the surface. Z is also a highly soluble protein, thus minimising the risk of inclusion body formation during production. The structure of Z is a compact three-helix bundle (figure 6a). In addition, Z lacks disulphide bridges which is an advantage since proteins with disulphide bridges has been found to be difficult to surface express using autotransport due to premature folding in the periplasm, as previously mentioned. Finally, Z binds to the Fc domain of IgG antibodies, enabling easy quantification of surface expression using flow cytometry.

The second part was the study of the expression of two more complex proteins of interest (I). The proteins chosen for expression were the fimbrial subunit protein SefA and the flagelling of Hgm from Salmonella enterica serotype Enteritidis (SE) [102][103]. SE is the most or second most common cause of Salmonella infections in the United States and a major cause of infection in Europe as well [104], which highlights the interest for developing efficient vaccines against this serotype. SefA, together with SefB and SefC make up the fimbriae of SE, and SefA is exclusive to S. enteritidis [105]. In addition SefA plays an important role in S. enterica infections [105]. These factors make SefA interesting as a vaccine epitope. Hgm has previously been proven to raise an immune response in mice [87] when administered nasally, highlighting its potential for use as a vaccine as well. Genes for SefA and Hgm was made available through a collaboration with the group of Professor Truong Nam Hai from Vietnam Institute of Biotechnology. A a vaccine against SE is of high interest in Vietnam due to a high portion of the chicken carrying SE.
SefA and H:gm were also of scientific interest in addition to their potential commercial use. SefA and H:gm are both surface proteins in *S. enterica*, and as such should have good secretion potential just like Z. However, in contrast to Z, they represent full proteins of larger size. SefA has a molecular weight of 14.4 kDa and H:gm has a size of 53 kDa [103, 105, 106], giving an opportunity to study the expression of two proteins of vastly different sizes. SefA and H:gm also have the advantage of not containing disulphide bridges, enabling the removal of this influence on their surface expression. The structure of H:gm is divided into two main domains. One domain forms a long stem formed by several adjacent α-helices, and the other part forms a compact domain composed of β-sheets (figure 6b). The α-helical domain interfaces with identical subunits and forms a coiled coil, building up the flagella. These structures can reach up to 15 µM in length [106]. Unfortunately there is no structure for SefA available at this date.

**Figure 6:** Structure of proteins used in this study. A: protein Z (PDB id 2SPZ), B: H:gm (PDB id 1UCU).
Surface expression of H:gm and SefA in *E. coli* using the mentioned AIDA construct was compared to an existing, Gram-positive surface expression system based on *Staphylococcus carnosus* in (I) [107]. The *S. carnosus* model system used for comparison was based on *Staphylococcus aureus* protein A (SpA) as described in section 2.2. The vector contains the X and M domains of SpA, followed by an albumin-binding domain (ABP) used for detection, the recombinant surface expressed protein and finally a signal peptide for inner membrane translocation through the Sec system (figure 7). The *S. carnosus* system has previously been reported suitable for use in vaccination [108], and thus made a good reference for this work. In addition, the *S. carnosus* expression system was available in-house through collaboration with Professor Stefan Ståhl.

**Figure 7:** Surface expression construct for expression in *S. carnosus*. SP: signal peptide, ABP: albumin binding domain used as detection tag, XM: cell wall anchoring domains.

### 4.2 Optimisation of surface expression (II)

In (II) we studied the effect on different cultivation conditions and techniques, such as batch and fed-batch, on surface expression using AIDA. Based on previous findings by Benz et al. [97] on the regulation of the *aidA* promoter and the expected behaviour of promoters controlling cellular adhesion [98] it was hypothesized that the conditions during cultivation, such as medium use, cell growth rate and nutrient availability, would be important factors that influence the surface expression. Thus the factors investigated were the choice of cultivation medium, growth rate using fed-batch technology and effects due to nutrient depletion.
**Effects of carbon source depletion**

An increase of expression of intracellular β-galactosidase under the control of the aidA promoter was found during the stationary growth phase in a study by Benz *et al.* [97]. We wanted to investigate if this could be used to achieve an increase in surface expression at the end of a production process. As previously mentioned, Benz *et al.* showed that the stationary phase sigma factor encoded by rpoS was likely not the cause of this behaviour. Thus we hypothesized that the increase in expression observed by Benz *et al.* was rather due to a stringent response due to carbon source depletion, regulated by spoT [99]. To test if this was the case we performed a set of shake flask cultivations were some flasks had a low starting glucose concentration and some had a higher. The cultures supplemented with less glucose would enter stationary phase earlier due to carbon source depletion. The surface expression of Z was measured using flow cytometry and compared to the reference shake flasks with a higher glucose concentration. Furthermore, a second addition of glucose was made after 1.5 h in the stationary phase. Contrary to the hypothesis it was found that the surface expression in fact decreased when the carbon source was depleted (figure 8). Upon the second addition of glucose to one of the flasks the surface expression rapidly increased back to the original values. Thus it was concluded that spoT was likely not responsible for the increase in expression observed by Benz *et al.* It also has to be kept in mind that Benz *et al.* used a rich medium in their expression study, meaning that it is unclear what limitation lead to stationary phase entry in their case.

*Figure 8:* Growth and surface expression of protein Z in relation to stationary phase entry. Three cultivations are shown: one with 10 g l⁻¹ initial glucose (squares) and two with 3 g l⁻¹ initial glucose (circles and diamonds). Open symbols show OD and filled symbols show surface expression. Arrow: a second addition of glucose was made to one of the two flasks with low starting glucose (circles). All curves are normalized to the entry to stationary phase, which is shown in grey.
Effects of growth medium

Benz et al. [97] showed a significant difference in intracellular expression from the aidA promoter depending on the growth medium used. Thus, we wanted to use this info to optimise the surface expression from the same promoter, in order to choose the optimal medium for expression. The expression of cells growing in rich Lysogeny broth (LB) medium was compared to cells growing in minimal salt medium. A significantly higher surface expression was found in minimal medium with glucose (fig 9A), which is in accordance with the results obtained by Benz et al. It is thus likely that this medium effect on surface expression is due to the induction of the promoter and not due to factors influencing transport to the cell surface.

We hypothesised that the reduction of expression in LB could be due to the lack of a readily available carbon source. This leads to rapid depletion of energy amino acids, and the triggering of a relA-mediated stringent response. This stringent response could be responsible for the lower expression levels detected in LB medium. To check this hypothesis, glucose was added to the LB medium to avoid consumption of amino acids for energy metabolism. Samples were taken and analysed for surface expression using flow cytometry (fig 9A) and amino acid consumption was monitored using HPLC analysis (fig 9B). Energy amino acids [109] were rapidly consumed in LB medium without glucose, and the consumption was markedly reduced in cultures where glucose was added to the medium. The one exception was serine, which was rapidly exhausted in both cases. However, contrary to the hypothesis, the surface expression was even lower in LB medium with glucose, which indicates that the reduction of expression in LB compared with minimal medium was not due to effects of stringent response but rather to still unknown effects.
Figure 9: Effects of growth medium on surface expression. (A) Comparison of surface expression of protein Z in minimal medium with glucose (grey), LB medium (white) and LB medium with glucose (black). Samples were taken at OD600 = 0.5 and 3. (B) Consumption of amino acids and accumulation of NH$_3$ in LB medium (filled symbols) and LB medium with glucose (open symbols) as a function of cultivation time. Legend: Serine (circles), aspartate (squares), threonine (upwards pointing triangle), proline (downwards pointing triangle) and NH$_3$ (diamonds). Arrow: depletion of the first amino acid (serine) in the LB cultures.

**Surface expression during bioreactor cultivation**

After selecting minimal medium as cultivation medium was to test if it was possible to maintain a stable surface expression during process conditions. This was done by performing small-scale bioreactor cultivations. Commonly the limiting factor for how much cells can be obtained from a cultivation is determined by the oxygen transfer capacity of the reactor. Once a critical cell mass is reached the total oxygen consumption rate (OCR) of the cells becomes higher than the maximum oxygen transfer rate and the oxygen supply becomes limiting. A common way to achieve higher cell masses is to reduce the specific oxygen consumption of each cell, thereby allowing a higher cell mass before the OCR reaches critical levels. This is typically achieved by reducing the specific growth rate of the cell by limiting a substrate, such as the carbon source. However, our previous results showed that carbon source starvation gave a reduction of surface expression (figure 7). Thus we
wanted to test if the carbon limitation during fed-batch cultivation would negatively affect surface expression. Furthermore, as we observed progressively lower surface expression levels using richer mediums it was also possible that the lowering of surface expression might be due to a too high growth rate. If this was true then lowering the growth rate even further in minimal medium might then improve the surface expression of Z.

To test these two possibilities, fed-batch cultivation with exponential feeding was employed in a multiparallel bioreactor system. To achieve three separate growth rates performed in duplicate. Feed exponents were chosen to achieve specific growth rates, $\mu$, of 0.1, 0.2 and 0.4 h$^{-1}$, according to equation 1, where $F$ is the feed at time $t$ and $F_0$ is the starting feed rate.

$$F = F_0 e^{\mu t}$$  \hspace{1cm} (1)

The cultures were monitored by regular sampling of optical density (OD) and freezing of samples for later flow cytometric analysis. All six cultures showed similar values during the initial batch phase, as expected. At the start of the feed there was an initial drop of surface expression in the two lower feed rates. With time the surface expression recovered and showed similar values as the batch phase (figure 10). In other words, the yield of surface expressed protein per cell was more or less constant at the different growth rates. The yield is defined as the specific productivity divided by the specific growth rate. This means that the specific productivity must be higher in the faster growing cell if the yield is unaffected by the growth rate. Thus it appears that no limitations due to transport to the cell surface could be observed at the studied growth rates, and growing the cells faster has no negative impact at production of surface expressed protein. Also, it can be concluded that apart from the initial adaptation phase there was no negative effect on surface expression due to the carbon source limitation during the fed-batch.
Having seen no effects on surface expression based on growth rate we wanted to see if it would be possible to maximise the specific productivity of surface expressed protein by growing cells at their maximum growth rate, which is achieved in a batch culture. However, to have an interesting process it is not enough to have a high specific productivity. It is also necessary to achieve a high cell mass in order to maximise the total productivity. This is generally not possible in a batch culture since the water activity becomes growth inhibiting at high glucose concentrations. To circumvent this we used a strategy of repeated, batch-wise addition of glucose of non-limiting concentration to keep the cells growing at their maximum rate. In this kind of process the maximum cell mass is dictated by the oxygen transfer capacity of the bioreactor, as well as the accumulation of by-products such as acetic acid. Using this technique we were able to achieve a cell mass of 18 g l\(^{-1}\) dry weight before the oxygen transfer rate (OTR) of the reactor became limiting. This was indicated by the dissolved oxygen tension (DOT) reaching zero and an increase in acetic acid production. The culture kept growing until 30 g l\(^{-1}\), when the experiment was terminated. Surface expression per cell was kept approximately constant until the oxygen limitation was reached, at which point it started to drop together with the growth rate (figure 11).

**Figure 10:** Fed-batch cultivation of surface-expressing *E. coli*. Progress of cell growth (filled symbols) and surface expression (open symbols) before and after feed start (marked in grey). Legend: \(\mu = 0.1\) h\(^{-1}\) (circles), \(\mu = 0.2\) h\(^{-1}\) (squares) and \(\mu = 0.4\) h\(^{-1}\) (triangles).
Figure 11: Batch cultivation with repeated glucose addition. (A) Cell growth measured as optical density (filled squares) and dry weight (open squares), and surface expression (circles) as a function of cultivation time. Growth curves are fitted to exponential functions to approximate the growth rates (solid lines). (B) Growth rate (diamonds), acetic acid concentration (triangles) and dissolved oxygen tension (DOT, line) as a function of time. Downward pointing arrows indicate time points for the five additions of glucose. The leftmost upwards pointing arrow indicates the starting point for manual DOT regulation by increasing of the stirring or the air-flow, and the rightmost upwards pointing arrow indicates the point at which the maximum stirrer speed was reached.

An interesting effect was that the width of the fluorescence peaks obtained from the repeated batch cultivation was significantly more narrow than those obtained during fed-batch (figure 12). This indicates a more homogenous population in the batch culture. We attributed this to being an effect of to differential substrate uptake rates of the cells in a population due to different expression levels of the sugar permeases, leading to different production rates in different cells, as has been previously reported [110].
Conclusion

The choice of growth medium gave the strongest effect on surface expression of Z, even though the mechanism is still unknown. We showed that expression could be kept at high levels at cell concentrations up to 18 g l\(^{-1}\), and by further work an even higher cell mass could probably be achieved using fed-batch. This would result in a higher total yield of surface expressed protein. However, the population homogeneity of the surface expression observed in the batch cultivation with repeated glucose additions is an important advantage, especially if the product is to be used as a live vaccine were the dosage to the patient needs to be tightly regulated. If the higher diversity in the fed-batch is indeed an effect of sugar uptake as hypothesized, the situation will likely become worse during scale up since the slower mixing in larger reactors generally give rise to sugar gradients [111]. It can also be noted that the expression values obtained during the bioreactor cultivations were higher than those observed during shake-flask cultivation, highlighting the importance of having a controlled environment for maximised expression.

Figure 12: Comparison of fluorescence peak width between a selected but typical fed-batch experiment here from a cultivation at a feed corresponding to a growth rate of 0.4 h\(^{-1}\) (grey) and the batch experiment with repeated glucose additions (black).
4.3 Surface expression of *Salmonella* epitopes

The proteins H:gm and SefA of *S. enterica* were surface-expressed using. Correct localisation of the recombinant protein products in the *E. coli* outer membrane was verified using a membrane protein isolation procedure followed by Western blot analysis. Both proteins could be detected in the outer membrane fraction, with only minor traces in the inner membrane and soluble protein fractions, as seen in figure 13.

![Western blot](image)

**Figure 13:** Western blot developed with antibodies against AIDA<sup>C</sup> showing cellular distribution of AIDA<sup>C</sup>-SefA and AIDA<sup>C</sup>-H:gm. OM: Outer membrane protein fraction, IM: Inner membrane protein fraction, Soluble: soluble protein fraction.

After verification that both constructs were present in the outer membrane we verified the correct orientation of the constructs using flow cytometry. Fluorescent antibodies against the N-terminal histidine tags were used for detection. SefA gave a clear signal compared to the negative control (figure 14). Based on this result it is possible to conclude that SefA was expressed in the outer membrane facing the external environment. The translocation of SefA was efficient, as can be seen from the fact that very little protein is detected in the soluble and inner membrane protein fractions (figure 13). However, no H:gm was detected on the cell surface by flow cytometry.
Surface expression was compared to the *S. carnosus* reference system. *S. carnosus* cells were incubated with fluorescent human albumin to allow binding the ABP domain. As can be seen in figure 15, a positive signal was detected for both SefA and H:gm in *S. carnosus*. The fluorescence peaks were much broader compared to the ones obtained in *E. coli*, indicating that the population distribution of surface expression was larger when using the *Staphylococcus* system. It is also not possible to determine if SefA and H:gm were expressed in full-length forms on the surface in this system since the ABP domain used for detection is located on the C-terminal side of the passenger, i.e. inside of the passenger compared to the cell wall.

![Figure 14: Surface expression of SefA and H:gm using AIDA detected using flow cytometry.](image1)

![Figure 15: Surface expression of SefA and H:gm in *S. carnosus* detected using flow cytometry.](image2)
Disregarding the fact that H.gm was not detected on the cell surface there are several advantages to the \textit{E. coli} based system. \textit{E. coli} grows readily in minimal medium, which is an advantage from a processing perspective. In contrast, \textit{S. carnosus} did not show growth in minimal medium and we had to turn to rich medium to achieve growth. Expression of SefA in \textit{E. coli} also showed a more narrow distribution within the population, as is seen from the width of the peak obtained from flow cytometry (figure 14). In contrast, the peaks were broader in \textit{S. carnosus}, especially for H.gm where two populations could be seen.

\section*{4.4 Effects of proteolysis on surface expression (I, II)}

\textit{Effects of the outer membrane protease OmpT}

In (II) we evaluated the influence of the outer membrane protease OmpT on surface expression of the model protein Z. OmpT is not very specific, requiring only two consecutive basic amino acids, such as two arginines, as recognition sequence for its proteolytic activity [112]. Z contains a potential cleavage site (Lys\textsubscript{49}–Lys\textsubscript{50}) for OmpT and is thus likely to be affected by this protease. The evaluation of the effect of OmpT was done by comparing the surface expression in an OmpT-producing 0:17 wild type (wt) with the expression in 0:17\textsuperscript{ΔOmpT} used in the rest of this work.

Successful expression of the AIDA\textsuperscript{C}–Z fusion protein in the outer membrane was verified by isolating the outer membrane protein fraction followed by Western blotting using antibodies against AIDA\textsuperscript{C}. Full-length fusion protein could be detected in the outer membrane in both strains used, as seen in the resulting blot (figure 16A). However, a significant portion of the protein was degraded in the wt strain, as is seen by the appearance of a strong band at approximately the same size as AIDA\textsuperscript{C}. This corresponds well with expected size obtained from cleavage of Z at the C-terminal OmpT cleavage site.
After having verified the location of the AIDA\textsuperscript{C}-Z fusion in the outer membrane the correct orientation of Z towards the external side was investigated using flow cytometry. Cells expressing Z were labelled using biotinylated human IgG followed by labeling of the antibody using a streptavidin-conjugated fluorophore. Flow cytometric analysis showed a strong surface expression, with approximately a hundred times higher fluorescence in the OmpT-negative strain. However, the fluorescence was approximately five times lower in the WT strain (figure 16B) compared to the OmpT mutant, supporting the degradation observed on the Western blot. Thus it can be concluded that Z was produced facing the external environment in both strain but it is clear that OmpT can degrade surface expressed Z. We thus decided to use the OmpT negative strain for the optimisation study.

![Figure 16: Comparison of surface expression of Z in 0:17 and 0:17 ΔOmpT. A: Outer membrane Western blot developed using AIDA\textsuperscript{C} antiserum. Lane 1: protein ladder, lane 2: 0:17, lane 3: 0:17 ΔOmpT, lane 4+5: 0:17 and 0:17 ΔOmpT with vector containing AIDA\textsuperscript{C} without Z. Lane 6: 0:17 expressing Z-AIDA\textsuperscript{C}, lane 7: ΔOmpT expressing Z-AIDA\textsuperscript{C}. B: Surface expression measured using flow cytometry. Red: 0:17 ΔOmpT expressing AIDA\textsuperscript{C} (negative control), blue: 0:17 expressing AIDA\textsuperscript{C}-Z, green: 0:17 ΔOmpT expressing AIDA\textsuperscript{C}-Z.](image)
Other proteases involved after inner membrane translocation

Even though the work done in (I) employed the OmpT-negative mutant strain there was still proteolysis present. Outer membrane localisation of H:gm and SefA was evaluated using the same Western blot procedure used for Z in (II). Blotting with AIDA\(^c\) antibodies showed a band at the expected molecular weight for full-length AIDA\(^c\)-SefA (figure 13 above). However, AIDA\(^c\)-H:gm was only detected in digested forms, showing that indeed other proteases than OmpT plays a role in degradation of surface expressed proteins. Furthermore several bands appeared in the outer membrane fraction with molecular weights slightly higher and lower than that of AIDA\(^c\) in both the SefA and H:gm samples. This proteolysis pattern indicates cleavage of both proteins by the same protease or proteases, most likely in the linker region and AIDA\(^c\), since these are common in both proteins.

It was concluded that the observed proteolysis of H:gm was likely in the N-terminal of AIDA\(^c\)-H:gm, since the AIDA\(^c\) antibodies could still bind to the fusion protein, something that was not expected if AIDA\(^c\) was cleaved. This could easily be tested, since AIDA\(^c\)-H:gm and AIDA\(^c\)-SefA carried N-terminal His\(_6\)-tags for flow cytometric detection. A second Western blot was performed on the same outer membrane protein fractions and developed using a His-tag antibody (figure 17). This blot again produced a clear band at the expected size of AIDA\(^c\)-SefA and, as expected, showed no band in the lane containing the sample for AIDA\(^c\)-H:gm. This confirmed that the protein was indeed cleaved in the N-terminal. Judging by the molecular weight observed the cleavage site was most likely located in the N-terminal part of the H:gm passenger. This also explains the lack of signal for H:gm during the flow cytometric analysis (figure 14 above), since His-tag antibodies were used for the fluorescence labelling of the cells. However, no H:gm was found in detectable amounts other than in the outer membrane fraction (figure 13), showing that the transport to the outer membrane was successful. Thus it is likely that the fragment of H:gm that is not cleaved is successfully translocated to the outside, since otherwise it would likely have been completely degraded in the periplasm. Due to the lack of a suitable detection tag other than the proteolytically cleaved His-tag it was not possible to conclusively determine that this is the case.
Figure 17: Western blot of outer membrane fractions of cells producing H:gm and SefA. Lane 1: Purified His-tagged lipase A (positive control), lane 2: H:gm, lane 3: SefA, lane 4: His-tagged AIDA\(^\text{C}\) (positive control). The blot was developed using anti-His\(_6\) antibodies.

Since AIDA\(^\text{C}\)-H:gm was verified to be located in the outer membrane using Western blots it must have been successfully transported over the inner membrane. Since the signal peptide is located in the N-terminal, and is needed for transport through the Sec system, the proteolysis must have taken place after translocation to the periplasm. Translocation over the outer membrane is thought to proceed in a C-to N-terminal direction with the formation of a hairpin inside the β-barrel, as previously mentioned. This means that the N-terminal of the passenger is the last part of the protein to leave the periplasm. It is therefore not surprising that it is the N-terminal that is targeted by proteolysis, assuming that the proteases responsible are located in the periplasm. For instance, the periplasmic protease DegP could be involved in this proteolysis. Especially considering that H:gm is significantly larger than SefA, which should mean that the translocation takes longer, resulting in a longer exposure to periplasmic proteases. Another possibility is that some property of H:gm causes it to partly fold in the periplasm, resulting in stalling of the translocation until proteolysis has occurred. The third possibility is the presence of another outer membrane protease than OmpT that can cause proteolysis after outer membrane translocation. One such protease is OmpP, which is a close homologue to OmpT with similar specificity [113]. However, the dramatic reduction of proteolysis of Z in the OmpT-negative strain suggests that OmpP is not responsible for this degradation.
5 Concluding remarks

The usage of an OmpT-negative strain as well as the correct cultivation medium was shown to have the largest impact on surface expression of Z using the AIDA autotransporter. As the difference in surface expression obtained between minimal salt medium and LB medium closely mimics the results obtained by Benz et al. for intracellular expression using the same aidA promoter it is likely that the difference is a result of the induction of the promoter. Bioreactor cultivations were designed based on the knowledge gained, and a high, stable surface expression was achieved in both batch and fed-batch cultivations. Reducing the growth rate in fed-batch had little influence on surface expression. However, cells growing in batch showed a more homogenous surface expression. High levels of surface expression were maintained to a cell mass of 18 g l\(^{-1}\). This shows the feasibility of using AIDA for surface expression even in larger scale, and the information obtained could be used as a starting point for the design of processes for surface expression of specific proteins of interest, which was the aim of the study.

It is clear that proteolysis remains a problem for some passenger proteins even in OmpT-negative E. coli strains. Further investigation should identify the positions were the proteins in our study are cleaved and might allow identification of which protease or proteases are responsible for this cleavage. It may then be possible to remove the proteolytic sites in the recombinant proteins or engineer a strain deficient for the protease in question. The degradation of Hgm shows the downside of relying on a single tag for detection of the surface expression of the protein in question. Of course, in the ideal world one would use specific antibodies for the protein in question. However, specific antibodies are not available for the vast majority of proteins, and it may not be feasible to produce antibodies for each protein when working with a range of different proteins. Thus, an ideal surface expression vector should provide ample detection possibilities independent of the passenger protein in question. To achieve this while minimising the impact of for instance proteolysis as was the case in this study it is recommended that at least two detection tags, placed on either side of the passenger protein, is used. Thus it is possible to detect the protein on the cell surface even if one tag is cleaved.
6 Abbreviations

AIDA  Adhesin Involved in Diffuse Adherence
BAM  β-barrel Assembly Machine
DNA  Deoxyribonucleic acid
DOT  Dissolved oxygen tension
DW  Cell dry weight (g l\(^{-1}\))
E. coli  *Escherichia coli*
IgG  Immunoglobulin G
IM  Inner membrane
kDa  kilo Dalton
LB  Luria broth
nm  nanometer
OCR  Oxygen consumption rate
OD  Optical density
OM  Outer membrane
OMP  Outer membrane protein
OTR  Oxygen transfer rate
ppGpp  Guanosine tetraphosphate
RNA  Ribonucleic acid
*S. carnosus*  *Staphylococcus carnosus*
wt  Wild type
X  Cell mass (g l\(^{-1}\))
\(\mu\)  Specific growth rate (h\(^{-1}\))
7 Acknowledgements

I would like to thank the Swedish Agency for Innovation Systems (Vinnova) and the Swedish International Development Agency (SIDA) for funding of the work presented in this thesis.

Furthermore I would like to thank everyone involved in the work presented in this thesis. First and foremost my supervisor professor Gen Larsson for accepting me as a PhD student and for all guidance and help during the work.

I would like to thank our Vietnamese collaborators for the work they have put in to the project. Professor Hai for providing the Salmonella proteins for the study in the first paper, Nhan for all the laboratory work and Huyen and Tan for their conversations and assistance during the last year.

All the partners in the BIO-AMINES project are acknowledged for the collaboration during these years.

Thanks to professor Stefan Ståhl for supplying the surface expression system for S. carnosus and John for all the help with the flow cytometry assay.

Emma, thanks for all your help, all the conversations we have had during the years and your contribution to the second paper.

Thanks to the diplomaworkers Cecilia, Alberto, Mattia, Rafaela and Madhu for your contributions to the work.

Thanks to all co-workers at the department for discussions during lunches and “fika”, providing important advice and for making the department a nice working environment.

Finally I would like to thank my family for their constant support. Especially Ella for your support and putting up with disappearance to the lab for long hours, and for being who you are. I love you.
8 References

1. Enzymes in Industrial Applications: Global Markets

2. Therapeutic Proteins Market to 2017 - High Demand for Monoclonal Antibodies will Drive the Market


44. Chang HJ, Sheu SY, Lo SJ: Expression of foreign antigens on the surface of *Escherichia coli* by fusion to the outer membrane protein TraT. *Journal of Biomedical Science* 1999, 6:64-70.

45. Klemm P, Scheinberger MA: Fimbrial surface display systems in bacteria: from vaccines to random libraries. *Microbiology* 2000,


69. Oliver DC, Huang G, Nodel E, Pleasance S, Fernandez RC: A conserved region within the *Bordetella pertussis* autotransporter BrkA is necessary for folding of its passenger domain. *Molecular Microbiology* 2003, **47**:1367-1383.


93. Charbonneau M, Berthiaume F, Mourez M: **Proteolytic processing is not essential for multiple functions of the *Escherichia coli* autotransporter adhesin involved in diffuse adherence (AIDA-I).** *Journal of Bacteriology* 2006, 188:8504-8512.


