Directed Evolution of HlyA Type I Secretion System from *Escherichia Coli*

Inaugural-Dissertation

For the attainment of the title of doctor
in the Faculty of Mathematics and Natural Sciences
at the Heinrich Heine University Düsseldorf

Presented by

Zohreh Pourhassan Najafabadi
from Iran

Düsseldorf, 11.05.2023

From the Institute of Biochemistry
At the Heinrich Heine University Düsseldorf

Published by permission of the
Faculty of Mathematics and Natural Sciences at
Heinrich Heine University Düsseldorf

Supervisor: Prof. Dr. Lutz Schmitt
Co-supervisor: Prof. Dr. Karl-Erich Jaeger

Date of the oral examination: 11.05.2023

"Knowledge is better than wealth because it preserves you, while you have to preserve wealth. Wealth decreases through spending, while knowledge increases by sharing…"

Ali ebne Abitaleb

Table of Contents

List of Figures

List of Tables

Abstract

Type 1 secretion systems (T1SS) are transport machineries wide spread among Gram-negative bacteria. T1SS translocon comprises an ABC transporter, a membrane fusion protein (MFP), and an outer membrane protein (OMP). The ABC transporter and the MFP form a complex in the inner membrane, while the OMP is recruited upon substrate recognition in the cytosol. The substrate is then translocated from the cytosol to the extracellular space through the translocon in one step, without a periplasmic intermediate.

T1SS are attractive candidates to be used as secretion hosts in different biotechnological applications due to their relatively simple architecture and easy plasmid-based expression compared to other secretion systems. However, their potential application as a secretion platform, particularly at large scales, is still limited by low secretion titers for some heterologous proteins.

In this thesis, the focus was on the HlyA T1SS from *Escherichia coli* to improve the secretion efficiency of the system. For this purpose, three KnowVolution campaigns were separately applied to the Hly enhancer fragment, the ABC transporter HlyB, and the MFP HlyD. This strategy is an integrative protein engineering approach, which uses a combination of directed evolution and computational analysis to achieve maximum improvement with minimal experimental efforts. The study also used a high-throughput screening system for directed evolution of a T1SS (the HlyA system) for the first time. In addition, it was shown that involving a native terminator region further enhances the secretion level. Researchers have already shown that the amount of TolC is constant under the overexpression condition of the inner membrane components (HlyB and HlyD), leading to the question of whether the amount of TolC is the limiting factor for the HlyA secretion. To address this question, the secretion efficiency was compared under condition of either the endogenous level or the overexpression of TolC protein. It was shown that TolC overexpression impairs the secretion efficiency.

Finally, this thesis dealt with the interaction between the nucleotide binding domain (NBD) of HlyB and the substrate HlyA. The interaction regions on the NBD for the secretion signal of HlyA were mapped, and two binding pockets were identified. The presence of these two binding pockets was supported with *in silico*, *in vivo*, and *in vitro* analysis. Noteworthy, it was proposed that the presence of a consensus length between the GG repeat and the amphiphilic helix affects on the secretion efficiency of heterologous proteins secreted by the HlyA T1SS.

Zusammenfassung

Typ-1-Sekretionssysteme (T1SS) sind unter gramnegativen Bakterien weit verbreitete Transportmaschinen. Das T1SS-Translokon umfasst einen ABC-Transporter, ein Membranfusionsprotein (MFP) und ein äußeres Membranprotein (OMP). Der ABC-Transporter und das MFP bilden einen Komplex in der inneren Membran, während das OMP bei der Substraterkennung im Cytosol rekrutiert wird. Das Substrat wird dann durch das Translokon in einem Schritt aus dem Zytosol in den extrazellulären Raum transportiert, ohne ein periplasmatisches Intermediat zu bilden.

Aufgrund ihrer relativ einfachen Architektur und Plasmid-basierten Expression sind T1SS in verschiedenen biotechnologischen Anwendungen als Sekretionswirte attraktive Kandidaten im Vergleich zu anderen Sekretionssystemen. Ihre potenzielle Anwendung als Sekretionsplattform, insbesondere im großen Maßstab, ist jedoch immer noch durch niedrige Sekretionstiter auf einige heterologe Proteine begrenzt.

In dieser Arbeit lag der Fokus auf dem HlyA T1SS aus *Escherichia coli* um die Sekretionseffizienz des Systems zu verbessern. Zu diesem Zweck, drei *KnowVolution*-Kampagnen wurden separat auf das Hly-*Enhancer*-Fragment, den ABC-Transporter HlyB und das MFP-HlyD angewendet. Diese Strategie ist ein integrativer Protein-*Engineering*-Ansatz, der eine Kombination aus gerichteter Evolution und Computeranalyse verwendet, um mit minimalem experimentellen Aufwand eine maximale Verbesserung zu erreichen. Die Studie verwendete zum ersten Mal auch ein Hochdurchsatz-Screening-System für die gerichtete Evolution eines T1SS (das HlyA-System). Zusätzlich wurde gezeigt, dass die Einbeziehung einer nativen Terminatorregion das Sekretionsniveau weiter erhöht.
Forscher haben bereits gezeigt, dass die TolC-Menge unter der Überexpressionsbedingung der inneren Membrankomponenten (HlyB und HlyD) konstant ist, was zu der Frage führt, ob die Menge an TolC der limitierende Faktor für die HlyA-Sekretion ist. Um diese Frage zu beantworten, wurde die Sekretionseffizienz entweder unter der Bedingung des endogenen Spiegels oder der Überexpression des TolC-Proteins verglichen. Es konnte gezeigt werden, dass die Überexpression von TolC die Sekretionseffizienz beeinträchtigt.
Abschließend befasste sich diese Arbeit mit der Wechselwirkung zwischen der Nukleotidbindedomäne (NBD) von HlyB und dem Substrat HlyA. Die Interaktionsregionen auf der NBD für das Sekretionssignal von HlyA wurden erfasst und zwei Bindungstaschen wurden identifiziert. Das Vorhandensein dieser beiden Bindungstaschen wurde durch *In-*

silico-, *In-vivo*- und *In-vitro*-Analysen gestützt. Bemerkenswerterweise wurde vorgeschlagen, dass das Vorhandensein einer Konsensuslänge zwischen der GG-Wiederholung und der amphiphilen Helix die Sekretionseffizienz von heterologen Proteinen, die vom HlyA T1SS sekretiert werden, beeinflusst.

Abbreviations

Å	Ångstrom
aa	Amino acid
ABC	ATP binding cassette transporter
ACP	Acyl carrier protein-dependent fatty acylation
ADP	Adenosine diphosphate
AH	Amphipathic helix
AMP	Adenosine monophosphate
A. aeolicus	*Aquifex aeolicus*
ATP	Adenosine triphosphate
B. subtilis	*Bacillus subtilis*
B. pertussis	*Bordetella pertussis*
CBT	ß-subunit of cholera toxin
CHO	Chinese hamster ovary
CLD	C39-peptidase-like domain
Ca^{2+}	Calcium ion
CD	Cytoplasmic domain
Cryo-EM	Cryogenic electron microscopy
Cryo-ET	Cryogenic electron tomography
Da	Dalton
3D	Three dimensional
DNA	Deoxyribonucleic acid
dNTPs	Deoxynucleotide triphosphates
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
e.g.	latin: exempli gratia
EPEC	Enteropathogenic *E. coli*
epPCR	Error-prone PCR
et al.	latin: et alii
FDA	Food and Drug Administration
GFP	Green fluorescent protein
His-tag	Polyhistidin-tag

Hly	Hemolysin
HlyA T1SS	HlyA type 1 secretion system
HTS	High-throughput screening
i.e.	Latin: “id est” or English “that is”
IFABP	Intestinal fatty acid-binding protein
IM	Inner membrane
IMAC	Immobilised metal ion chromatography
IMP	Inner membrane Protein
K_m	Michaelis-Menten constant
KISS	Keep it simple and smart
LPS	Lipopolysaccharide
M	Molar
MalE	Maltose-binding protein
MARTX	Multifunctional auto-processing repeat in toxin
MDR	Multidrug resistance
mg	Milligram
min	Minute
mL	Milliliter
MPF	Membrane fusion protein
NBD	Nucleotide binding domain
nm	Nanometer
nM	Nanomolar
NMR	Nuclear magnetic resonance
OM	Outer membrane
OMP	Outer membrane protein
PAGE	Polyacrylamide gel electrophoresis
pbp	Possible binding pocket
PDB	Protein data base
PI	Isoelectric point of a protein
PMF	Proton motive force
POTRA	Polypeptide transport-associated
PP	Periplasm
P. aeruginosa	*Pseudomonas aeruginosa*

PTM	Post translational modifications
RMSD	Root mean square deviation
RNC	Ribosome nascent chain
RND	Resistance-nodulation-division
RTX	Repeat In Toxin
Sec pathway	The general secretory pathway
Sec	Second
SEC	Size exclusion chromatography
S. marcescens	*Serratia marcescens*
sRNA	Soluble ribonucleic acid
SRP	Signal recognition particle
TBS	Tris-buffered saline
TEV	Tabacco etch virus
T1SS	Type 1 secretion system(s)
T2SS	Type 2 secretion system(s)
T3SS	Type 3 secretion system(s)
T4SS	Type 4 secretion system(s)
T5SS	Type 5 secretion system(s)
T6SS	Type 6 secretion system(s)
T8SS	Type 8 secretion system(s)
Tat pathway	Twin arginine transporter
TMD	Transmembrane domain
TMHs	Transmembrane helixes
V. cholerae	*Vibrio cholerae*
µl	Microliter
µm	Micromolar

List of 20 amino acids

Full name	Three-letter code	One-letter code	Amino acid type
Phenylalanine	Phe	F	Aromatic
Tyrosine	Tyr	Y	
Tryptophan	Trp	W	
Alanine	Ala	A	Aliphatic
Valine	Val	V	
Leucine	Leu	L	
Isoleucine	Ile	I	
Glycine	Gly	G	
Cysteine	Cys	C	Polar
Methionine	Met	M	
Proline	Pro	P	
Serine	Ser	S	
Threonine	Thr	T	
Asparagine	Asn	N	
Glutamine	Gln	Q	
Aspartic Acid	Asp	D	Charged acidic
Glutamic Acid	Glu	E	
Histidine	His	H	Charged basic
Lysine	Lys	K	
Arginine	Arg	R	

1 Introduction

1.1 Recombinant protein production

Protein production has quickly become a crucial step in studies and vast applications of proteins, from biological and biomedical sciences to industrial and environmental applications. Protein production has a long history. Many years ago, proteins for pharmaceutical purposes were isolated from natural resources. *i.e.*, animals or plants. Reports show that in the 1920s, pigs' pancreas was used for the production of insulin for diabetic therapy (Owczarek et al., 2019, Puetz and Wurm, 2019).

The technology behind protein production has changed since the 1980s when recombinant protein production was introduced to the world of biosciences for the first time (Johnson, 1983). Human insulin, produced by *Escherichia coli,* was the first product derived from recombinant DNA technology, which was approved by the Food and Drug Administration (FDA) (Johnson, 1983). Since then, recombinant protein technology has made remarkable progress towards biopharmaceutical production. To date, more than 170 unique products produced by recombinant technology have been approved by the FDA (Owczarek et al., 2019). Thus far, the demand for recombinant proteins has become greater and greater. In 2019 only, the market for biological protein drugs reached approximately 270 billion USD (https://www.prnewswire.com/).

Nevertheless, recombinant protein technology faces challenges concerning the production of proteins with high quality and yields, while at same time reducing the costs (Pourhassan et al., 2021). Although different producer hosts and many molecular tools are available, failed production of a target protein is not unusual due to the fact that each protein has its unique features (Rosano et al., 2019). This is why, there is a continuous interest in developing and improving protein production strategies which can be broadly applied.

Platforms for protein production originate from prokaryotic or eukaryotic cells, including bacteria, fungi, yeasts, insect cells, recombinant plants, transgenic animals, and mammalian cells. Each platform offers advantages for the production of a specific group of proteins (see **Table 1-1**) (Owczarek et al., 2019). For instance, mammalian cells are commonly employed for the production of "complicated" proteins that contain post-translational modifications (PTM), like human hormones (Khan, 2013). Since culturing of mammalian cells is very costly and time-consuming, these platforms are not the preferred choice for the production

of industrial enzymes (Puetz and Wurm, 2019). On the other hand, yeasts are simple eukaryotic cells that perform some PTMs, such as O-linked phosphorylation, glycosylation, and disulfide bond formation (Owczarek et al., 2019). However, genetic instability and clonal variation are drawbacks of using yeast cells (Pourhassan et al., 2021).

Among all the available platforms, bacterial hosts are more competitive in many cases, for example for the production of industrial enzymes. Most bacterial hosts grow quickly on inexpensive carbon sources with high productivity. The model organism of bacterial platforms is *E. coli* (Owczarek et al., 2019).

E. coli, a Gram-negative bacteria, is a cell factory of recombinant protein production on a laboratory scale (Rosano et al., 2019). A wide variety of molecular tools and biochemical assays (from gene cloning to protein purification) are available for handling this bacterium. The genetics and biochemistry of this microorganism have been studied in great detail. With a large set of existing vectors, genetic engineering of this bacterial cell is very affordable. Reduced cost of production, a short doubling time, and high yields of production make this bacterium a preferred host for protein production on laboratory scale (Rosano et al., 2019, Rosano and Ceccarelli, 2014).

Table 1-1: Different recombinant protein production platforms. Adopted and modified from (Owczarek et al., 2019).

Platform	Overall Costs	Production time	Scale-up capacity	Product yield	Product quality	Contamination risk	Purification cost
Bacteria	low	low	high	medium	low	medium	high
Yeast	medium	medium	high	high	medium	low	medium
Microalgae	low	high	high	high	high	very low	medium
Filamentous fungi	low	high	high	high	medium	low	low
Mammalian cell culture	high	high	very low	medium -high	high	very high	high
Transgenic plants	very low	medium	very high	high	high	low	high
Transgenic animals	high	high	low	high	high	very high	high
Insect cell culture	medium	medium	high	high	medium	very low	medium
Plant cell culture	medium	medium	medium	high	high	very low	high

Despite the benefits of *E. coli* as a producer host, biotechnological applications of this microorganism, especially on a large scale, still faces some challenges. In general, protein degradation by the pool of cytosolic proteases is a risk when using bacterial platforms. Lack of PTMs for a bacterial-derived protein might lead to the production of unstable, inactive, or improperly folded proteins (Owczarek et al., 2019, Pourhassan et al., 2021). In addition, a high expression level of a recombinant protein might lead to toxicity of the target protein for the producing host (Wingfield, 2015). Furthermore, a high level of expression might lead to the formation of inclusion bodies (Fahnert et al., 2004). Although inclusion body formation had been considered to be a limitation of bacterial platforms for a long time; in recent years, research has highlighted the benefits of efficient isolation of active proteins from inclusion bodies (Krauss et al., 2017).

1.1.1 Secretion of recombinant proteins

It is now well-established that protein secretion can provide solutions to the difficulties related to the cytosolic expression of recombinant proteins. Protein secretion is the process of transporting proteins from the cytosol into either other cell components, the extracellular space, or into other cells. For this purpose, bacterial cells possess a variety of secretion systems to transport proteins from one location to the other (Costa et al., 2015). These secretion systems are involved in numerous cell functions, such as nutrient acquisition, cell mobility, cell communication, toxicity, biofilm formation, respiration, etc. (Burdette et al., 2018, Costa et al., 2015).

Bacterial secretion systems can be employed for the secretion of heterologous recombinant proteins. The benefits of using secretion systems for recombinant protein production are listed below:

- Lower protein content of the extracellular space leads to higher solubility and stability of the secreted protein.
- Reduced risk of protein degradation by cytosolic proteases.
- Decreased risk of toxicity of a target protein for the producer host.
- Target proteins can be isolated in fewer steps, resulting in a significant reduction of the production costs (Kleiner-Grote et al., 2018).

1.2 Secretion systems of Gram-negative bacteria

Gram-negative bacteria are attractive hosts for the production of recombinant proteins. They grow quickly and are easier to handle. Moreover, many different tools and assays are available for their manipulation, and they are more stable at large-scale cultivations (Kleiner-Grote et al., 2018, Burdette et al., 2018).

Bacterial secretion systems are classified into two groups, which is based on whether a cargo bypassing (one-step secretion system) or entering the periplasmic space (two-step secretion system) (**Figure 1-1**). In a one-step secretion systems, a cargo is secreted directly from the cytosol to the extracellular space through a dedicated channel. Type 1 secretion systems (T1SS), T3SS, T4SS, and T6SS belong to this particular class.

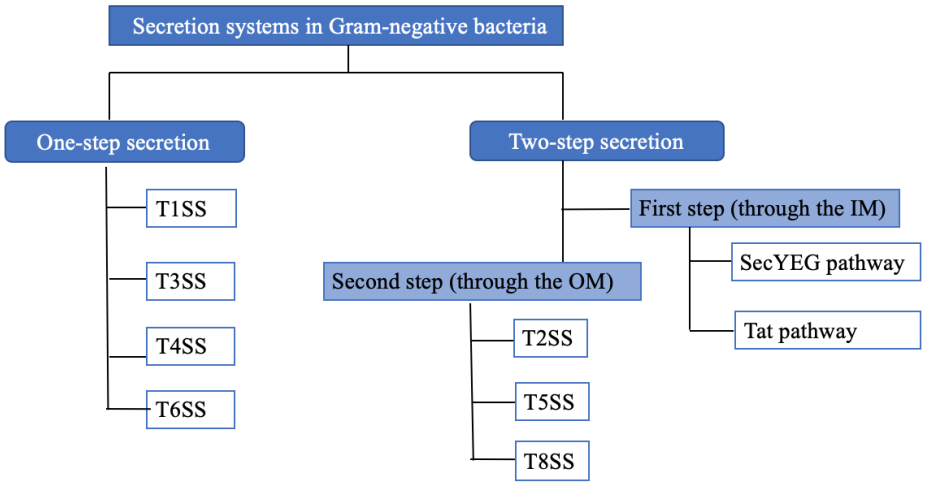


Figure 1-1: Classification of secretion systems in Gram-negative bacteria

In a two-step secretion system, cargo is first translocated into the periplasmic space by one of the general secretion systems: the secretion translocation system (Sec)-pathway or the twin-arginine transporter (Tat)-pathway. Subsequently, cargo is transported through a channel located in the outer membrane. Type 2 secretion systems (T2SS), T5SS, and T8SS belong to this class of secretion systems (Pourhassan et al., 2021, Burdette et al., 2018, Costa et al., 2015). **Table 1-2** summarizes some of the features of Gram-negative secretion systems. In the following sections, these systems will be explained and discussed in more detail.

Table 1-2: Gram-negative bacterial secretion systems.

	S*signal/ cleavage	Release into	Complexity	Model organism	Model substrate
One-step Secretion systems					
T1SS	C-terminal/ No	culture	low	*Pseudomonas fluorescens* *E. coli*	TliA HlyA
T3SS	N-terminal/ No	target cell	high	*Salmonella enterica* *S. enterica* *Enteropathogenic E. coli* *Shigella flexneri*	SptP FlgM EspA OspB
T4SS	Diverse/ Diverse	target cell	high	*Agrobacterium tumefaciens*	DNA
T6SS	Diverse/ Yes	target cell	high	*Vibrio cholerae*	VgrG
Two-step secretion systems					
T2SS	N-terminal/ Yes	Periplasm and then medium	high	*Pseudomonas aeruginosa*	Unknown
T5SS	N-terminal/ Yes	Periplasm and then medium	low	*E. coli*	Pet
T8SS	N-terminal/ Yes	Anchored to the OM	low	*E. coli*	CsgA

S* : secretion

1.3 One-Step secretion systems

1.3.1 Type I secretion system

T1SS is a double-membrane-spanning secretion system that is widespread in Gram-negative bacteria (Costa et al., 2015). It secretes a variety of proteins involved in different cell functions, ranging from pathogenicity (such as the pore-forming toxin HlyA from *E. coli*) to nutrient acquisition (such as the hemophore HasA from *Serratia marcescens*). Thomas *et al.* classified substrates of T1SS into six groups: Heme-binding proteins, RTX toxins, proteases, lipases, adhesion proteins, and S-layer proteins (Thomas et al., 2014b). These substrates vary in size from 5.8 kDa (such as ColicinV from *E. coli*) to 520 kDa (such as the huge surface protein LapA from *Pseudomonas fluorescens*) (Kanonenberg et al., 2013).

Despite such differences, T1SS substrates have some properties in common, as listed below; although, some exceptions may occur:

- Substrates of T1SS are transported in one single step from the cytosol to the extracellular space through a tripartite channel.
- T1SS substrates are transported unfolded, and they fold after secretion in the extracellular space.
- T1SS secretion signals are located at the C-terminus and are not cleaved off during, prior to, or after secretion (Holland et al., 2016, Pourhassan et al., 2021).
- Almost all T1SS substrates harbor a repetitive glycine- and aspartate-rich sequence upstream of their secretion signals.

The glycine- and aspartate-rich sequence, known as GG repeats or nonapeptide, contains a consensus sequence of GGxGxDxUx, where x refers to any amino acid (aa) and U to a large hydrophobic aa. The presence of this sequence gives the RTX (Repeats in Toxin) family its name. The number of GG repeats varies from only a few to 40 repeats, depending on the specific RTX protein (Linhartová et al., 2010). The RTX domain is a site for Ca^{2+} ion binding (Spitz et al., 2019). The binding stimulates the folding of an RTX substrate in the extracellular space, where the Ca^{2+} concentration is higher (almost 2 mM) than the K_d of the RTX domain for Ca^{2+} ions (150 µM). Accordingly, an RTX substrate stays unfolded in the cytosol where the Ca^{2+} concentration is only around 300 nM (Spitz et al., 2022, Pourhassan et al., 2021).

In terms of architectural composition, T1SS is simpler compared to other secretion systems because the secretion machinery consists of only three membrane proteins: an ATP-binding cassette (ABC) transporter, a membrane-fusion protein (MPF), and an outer membrane protein (OMP). These three membrane proteins form a channel that directs the substrate in a single step from the cytosol into the extracellular space (Kanonenberg et al., 2018). The energy of transport is provided by the ABC transporters through ATP hydrolysis. Moreover, research has shown that the proton motive force (PMF) is also providing the energy for the transport, as the negatively-charged RTX domain is translocated through the negative gradient of the membrane (Bumba et al., 2016, Koronakis et al., 1991).

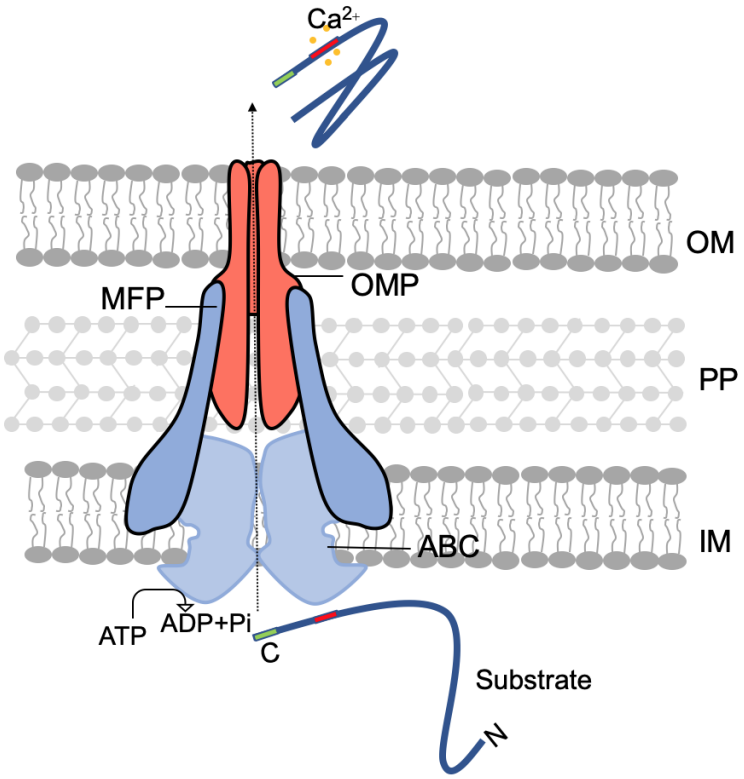


Figure 1-2: Schematic representation of protein export via T1SS
The channel of T1SS consists of an ABC transporter plus a MFP in the inner membrane, and an OMP. The translocation of the substrate occurs in one step from the cytosol to the extracellular space. The secretion signal is at the extreme C-terminus (highlighted as an ellipse in green), and the RTX domain is upstream of the secretion signal (highlighted as an ellipse in red). The secretion signal is not cleaved off during, while, or after the secretion. The protein substrate remains unfolded in the cytosol and folds only in the extracellular space, where the Ca2+ concentration is higher. IM: inner membrane, OM: outer membrane, PP: periplasm.

ABC transporters are assigned to a superfamily of proteins present in all domains of life. ABC transporters have a transmembrane domain (TMD), a nucleotide-binding domain (NBD), and an N-terminal cytosolic domain (Schmitt et al., 2003). A crystal structure of PCAT1 (PDB: 4RY2) is shown in **Figure 1-3a,** representing the aforementioned domains. PCAT1 is an ABC transporter from the Gram-positive bacterium *Clostridium thermocellum*. PCAT1 contains an active peptidase in its N-terminus which processes antimicrobial peptidase before secretion (Lin et al., 2015).

The ABC transporters of T1SS are classified into three groups based on their N-terminal domains. ABC transporters of groups 1 and 2 both contain this domain in their N-termini, which can be either an active peptidase (group 1) or an inactive one (group 2). The third group of ABC transporters do not contain this domain at all (Kanonenberg et al., 2013). The crystal structure of a homolog of the ABC component (PrtD) from *Aquifex aeolicus* (AaPrtD) was solved and visualized (**Figure 1-3b)**. PrtD is grouped into the third group of ABC transporters (Morgan et al., 2017).

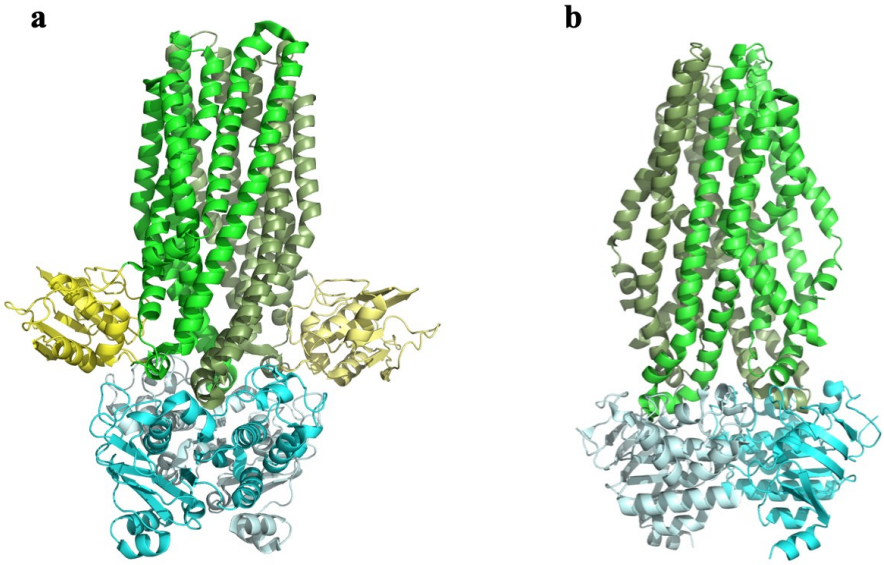


Figure 1-3: Cartoon representations of two examples of ABC transporters
Both structures were obtained by X-ray crystallography and visualized using Pymol software (www.pymol.org). The N-terminal cytosolic domain is shown in yellow, the TMDs in green, and the NBD in cyan; pale colors correspond to the same regions in the second monomer. **a**) The structure of PCAT1 (PDB: 4RY2), an ABC transporter from the Gram-positive bacterium *Clostridium thermocellum* (Lin et al., 2015), is shown. PCAT1 contains an N-terminal C39-peptidase domain which is an active peptidase, but it is not grouped within the aforementioned T1SS groups, as it belongs to a Gram-positive bacterium. **b**) The structure of AaPrtD (PDB: 5L22), an ABC transporter from *Dickeya dadantii*, is shown. AaPrtD is a homolog of the ABC component from the Gram-negative bacterium *Aquifex aeolicus*. PrtD does not harbor an N-terminal accessory domain in the cytosol, therefore, it is grouped under the third group of ABC transporters (Morgan et al., 2017).

Concerning the secretion mechanism of T1SS, two models of secretion have so far been suggested. However, it is still ambiguous which model is valid for the translocation through T1SS.

- The Brownian ratchet model suggests that the fuel for transport is provided primarily by the Ca^{2+}-ion-dependent folding of the substrate, which is promoted on the C-terminus of a substrate at the RTX domain. A folded RTX domain then prevents any tracking of the substrate back into the channel (Bumba et al., 2016).
- The Power stroke model suggests that conformational changes caused by ATP hydrolysis are the promoting factor, forcing the substrate to pass through the T1SS channel. First, ATP binds to the ABC transporter, which turns the ABC transporter from an inward-facing conformation form to an outward-facing form. Then, these conformational changes move the substrate from the cytosol to the trans-side of the membrane (Rempel et al., 2019, Stefan et al., 2020).

The biotechnological applications of T1SS have been taken into consideration mainly due to the two advantages of these systems. First of all, a wide variety of substrates can be secreted using T1SS. Secondly, the plasmid-based expression of its components is rather easy, as only three components form the entire T1SS machinery (Burdette et al., 2018). A long list of heterologous proteins with different origins have been successfully secreted via T1SS (Pourhassan et al., 2021).
So far, three different T1SS have been employed for the secretion of heterologous proteins:

- The TliDEF system from *Pseudomonas fluorescens*
- The PrtDEF system from *Erwinia chrysanthemi*
- The HlyABD secretion system from *E. coli*

However, the main drawback of using T1SS in biotechnological applications is formed by the limited secretion titers (Pourhassan et al., 2021).

1.3.2 Type III secretion system

T3SS, also known as injectosomes, deliver effector proteins into eukaryotic host cells. These proteins are mostly virulence-associated factors (Wagner et al., 2018). T3SS induction occurs upon sensing host-cell contact. Research has shown that a low environmental calcium ion concentration (*Yersinia*) (Heesemann et al., 1986) or an altering pH can also initiate T3SS induction (*Salmonella* SPI-2)(Yu et al., 2010).
T3SS are found in various pathogenic bacteria, including the enteropathogenic *E. coli* (EPEC), *Shigella*, *Yersinia*, *Chlamydia*, and *Salmonella* (Costa et al., 2015). The basic function of T3SS is to assemble bacterial flagella (Abby and Rocha, 2012). However, T3SS are also involved in the secretion of different substrates. These substrates have roles in different bacterial functions, such as nutrient acquisition (in *Chlamydia*), adhesion mechanism (in EPEC), and killing human macrophages (in *Salmonella*) (Wagner et al., 2018). Substrates of T3SS differ in their sizes, structures, and functions, but all are secreted unfolded in one step. The secretion signal is located at the N-terminus and remains intact during and after the secretion process (Burdette et al., 2018).
In terms of structure, T3SS are highly conserved among different Gram-negative bacteria. These secretion systems are among the most complex membrane machineries, with a size of around 6 MDa. The T3SS machine consists of around 20 different proteins with different copy numbers, ranging from 1 to 100. The involved subunits form a membrane-spanning

multi-ring complex, a needle-like filament complex, an inner membrane complex, and a set of cytosolic proteins (**Figure 1-4)**. The cytosolic subunits have roles in either the regulation of the system, or in providing the energy for transport (Wagner et al., 2018, Burdette et al., 2018, Costa et al., 2015). Note, the energy for transport is provided by ATP hydrolysis (Woestyn et al., 1994) as well as the PMF (Wagner et al., 2018).

A fascinating application of T3SS is its use in many biomedical studies and applications. T3SS is able to deliver a specific protein, such as proteins involved in gene editing, signaling, and transcriptional regulation, into a eukaryotic cell (Wagner et al., 2018). Secretion of a wide variety of proteins using T3SS has already been reported, such as scFvs (Metcalf et al., 2016), spider silk monomers (Widmaier et al., 2009), and alkaline phosphates (Burdette et al., 2018).

Besides the structural and regulatory proteins, sRNA is required for the proper functioning of T3SS. Therefore, transferring the related encoding genes to a non-native host is a complex procedure (Burdette et al., 2018). As the native hosts are pathogens, they are not suitable to use in many standard laboratories, especially on industrial scales. The attenuation of the native host to reduce the toxicity capacity of the host cells is a prerequisite. The most studied and applied T3SS originated from strains of *Salmonella spp*., *Yersinia spp*., and *Pseudomonas aeruginosa*. These strains have already been altered genetically for bacterial attenuation (Bai et al., 2018).

Researchers are facing some limitations upon using T3SS for biotechnological applications. The secretion through T3SS is tightly controlled by environmental changes and growth phases. Although high titers of secretion for some heterologous have been achieved, the low secretion yield for some heterologous proteins is still a challenge. Attenuation of the native organism is a time-consuming and costly process (Pourhassan et al., 2021, Bai et al., 2018, Burdette et al., 2018).

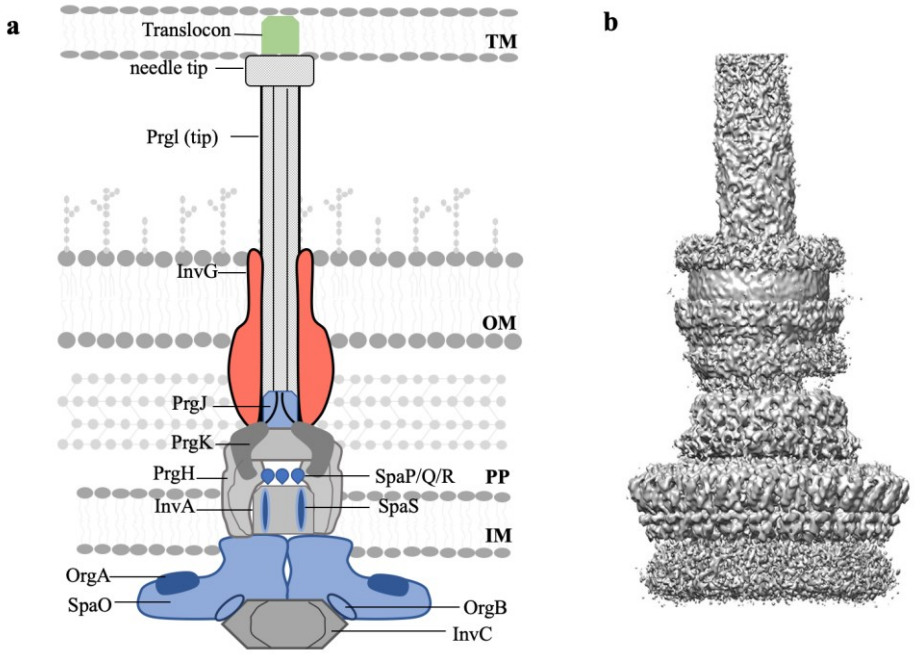


Figure 1-4: Representation of T3SS machinery in the membrane
a) The T3SS is a double-membrane-embedded nanomachine that translocates effector proteins in a single step from the cytosol to the transmembrane space of the target cell. It consists of around 25 different proteins and sums up to a size of 3.5 MDa to 6 MDa. T3SS, also known as injectosome, has a needle-like filament plus a set of rings forming the base structure. InvC is a hexameric ATPase located in the cytosol. The secretin InvG extends from the IM to the OM complex. The IM complex consists of PrgK, PrgH, SpaP, SpaQ, SpaR, SpaS, and InvA. This IM complex acts as a base for the protection and support of the needle (Prgl). The needle tip contacts the target membrane and creates a pore in it. The schematic representation in panel a was drawn based on (Costa et al., 2015). **b**) A side view of the Shigella needle complex adopted from (Lunelli et al., 2020). OM: outer membrane; PP: periplasm; IM: inner membrane.

1.3.3 Type IV secretion system

The ability to translocate DNA, as well as proteins, into bacterial or eukaryotic cells makes T4SS unique among other secretion systems. T4SS involves the export of DNA into the environment, or the import of exogenous DNA into a cell (Lacerda et al., 2013). Recent studies have revealed that T4SS also contributes to the transfer of other macromolecules, such as peptidoglycan, into other cells (Grohmann et al., 2018). T4SS are classified into two groups based on whether the substrate is a protein or DNA (Costa et al., 2021).

T4SS has a role in spreading antibiotic resistance genes through the conjugation of plasmid DNA between species. T4SS contributes to the pathogenicity of certain bacteria by translocating the related effector protein. For instance, the pathogenesis of *Bordetella pertussis* is related to T4SS (Costa et al., 2015).

In terms of structure, T4SS harbors a complex machine consisting of 6 to 12 proteins (**Figure 1-5**). The T4SS has a core subunit plus some accessory proteins, which are system-specific,

and may vary between different species (Costa et al., 2021). Generally, the entire system consists of a scaffold and translocation apparatus, as well as a pilus that is extended into the extracellular space (Terradot and Waksman, 2011).

The energy for transport is provided by three ATPases: VirB11, VirD4, and VirB4 (**Figure 1-5**). VirD4 plays an additional role in binding to either DNA or protein substrates (Llosa and Alkorta, 2017). The channel of T4SS consists of two main parts: the first part is located in the IM and the second part is located in the periplasmic space and OM. VirB3, VirB6, and VirB8 are proteins forming the IM part, and VirB4 (an ATPase) is attached to the IM complex. VirB7 and VirB9 are the proteins that form the OM complex, and protein VirB10 associates in the formation of both IM and OM complexes (Low et al., 2014). The T-pilus of *Agrobacterium tumefaciens* consists of the major component VirB2 and the minor component VirB5 (Aly and Baron, 2007).

The T4SS is almost never considered for the secretion of heterologous proteins in comparison to the other secretion systems discussed herein, possibly caused by its complicated structure.

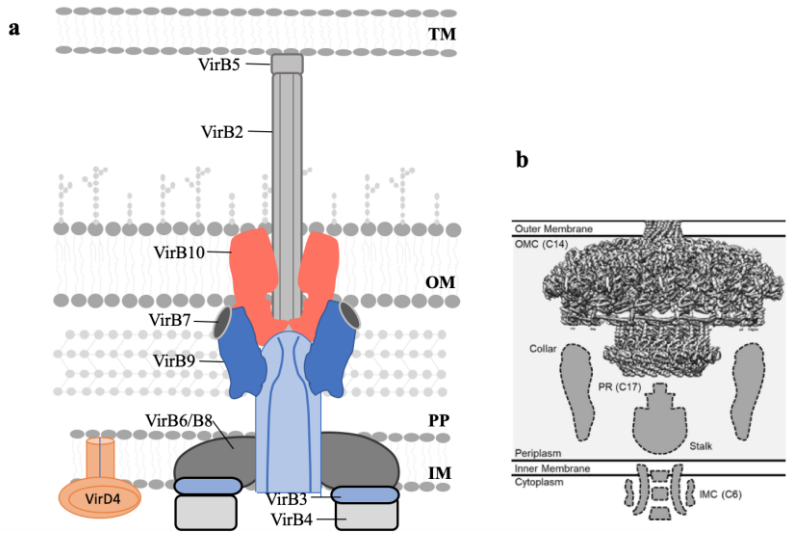


Figure 1-5: Structural organization of T4SS
a) The nanomachine of T4SS spans both membranes. Three ATPases fuel the substrate translocation: VirB11, a cytosolic soluble ATPase; VirD4, located in the IM; and VirB4, associated with the IM complex (VirB3, VirB6, and VirB8). Vir10 extends from the IM complex to the core-outer membrane complex (VirB7 and VirB9). VirB2 and VirB5 together form the pilus. The schematic is based on (Costa et al., 2015). **b**) Cryo-EM structure of the core complex of T4SS from the Helicobacter pylori, adopted from (Sheedlo et al., 2020).

1.3.4 Type VI secretion system

T6SS are double-membrane-spanning systems that transport toxic effector proteins from the cytosol into prokaryotic and eukaryotic cells in a single step. The substrate is either in the folded or partially folded state. T6SS has an important role in bacterial pathogenesis, specifically for Proteobacteria, and consists of 13 main protein components plus a few accessory proteins. The T6SS nanomachine consists of two main complexes, including a membrane complex and a tail complex (**Figure 1-6**). The tail part is anchored to the membrane by mediation of the membrane complex (Sana et al., 2017, Ho et al., 2014).

Only recently, the T6SS from *Aeromonas dhakensis* was reconstituted in an *E. coli* laboratory strain. The heterologous assembly of T6SS empowered *E. coli* to deliver a range of proteins into the cytosol of other cells, thereby creating an antimicrobial tool. The complexity of T6SS and its complicated gene regulation, made heterologous expression difficult. Furthermore, the native hosts are among pathogenetic bacteria (Jana et al., 2021), resulting in less or no application of T6SS in biotechnology.

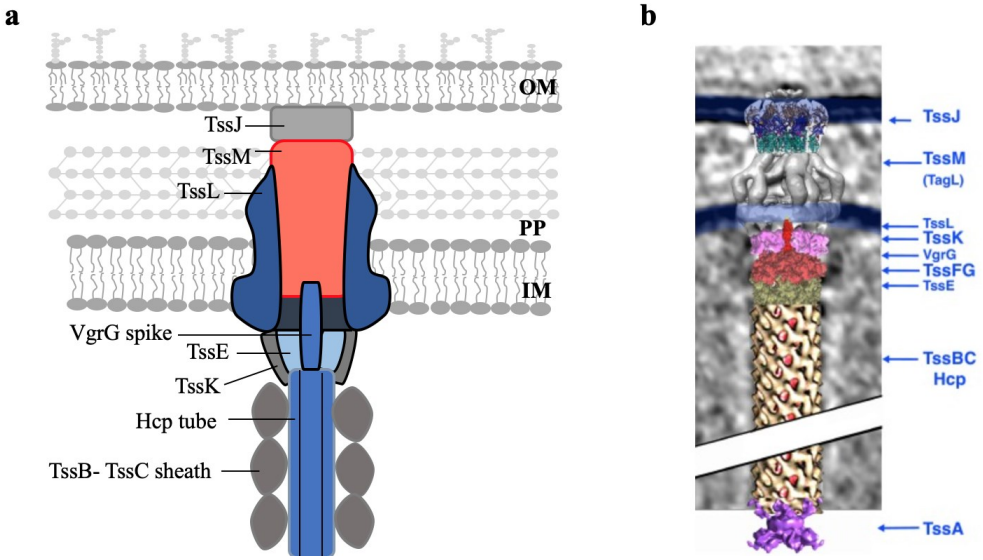


Figure 1-6: A schematic structural view of T6SS
T6SS consists of two main parts, the membrane complex and the tail complex, with a baseplate in between of them. T6SS translocates effector proteins from the cytosol to the eukaryotic or prokaryotic cells. Thirteen conserved proteins, plus some accessory proteins, are involved in the formation of T6SS nanomachine. **a**) A schematic view of T6SS based on (Costa et al., 2015). The membrane complex (TssJ, TssL, TssM) keeps the tail complex anchored to the membrane. A baseplate complex is involved in the association of the tail tube and the sheath. TssK might act as a connector of the membrane complex to the tail. VgrG is in a trimeric state and forms the spike part. Hcp is in a hexameric state and forms a tubular structure of the tail. TssB-TssC forms a boundary around the Hcp tube. **b**) A view of T6SS adopted from (Douzi et al., 2018).

1.4 Two-step secretion systems

Two-step secretion systems involve the translocation of substrates into the periplasmic space first, using either the Sec- or Tat-pathway. The second step is the translocation across the outer membrane, with the substrate either being released into the extracellular space or remaining attached to the membrane.

1.4.1 Translocation across the inner membrane

1.4.1.1 Sec pathway

The general secretory (Sec) pathway is an universal and ubiquitous secretion machinery found in all kingdoms of life (Tsirigotaki et al., 2017). Among all bacterial secretion systems, only the Sec pathway is essential for cell viability, which emphasizes the importance of this secretion pathway (Berg et al., 2004). The Sec pathway translocates a large number of different proteins with different functionality. For instance, around 96% of proteins crossing across the inner membrane of *E. coli* employ the Sec pathway (Orfanoudaki and Economou, 2014).

The core of the Sec pathway is the SecYEG channel, also known as SecYEG translocon, which mediates the translocation of proteins into the periplasm space or inserts proteins into the inner membrane. SecA, the motor molecule of the Sec pathway, assembles with the SecYEG translocon to form the SecYEG holoenzyme (Tsirigotaki et al., 2017).

Polypeptides destined for the Sec pathway are synthesized first as precursor proteins (pre-proteins) containing an N-terminal secretion signal. The secretion signal has a length of around 20 residues and is cleaved off during translocation across the inner membrane by a peptidase known as SPase. After the cleavage, the mature protein either enters into the periplasmic space or is integrated into the inner membrane. Although the sequences of different secretion signals are not identical, they all harbor three conserved parts: a positively-charged region at the N-terminus, a central hydrophobic core, and a polar terminal domain. The Sec pathway translocates respective substrates in an unfolded state (Freudl, 2018).

The Sec pathway mediated translocation takes place in two modes, as explained below. The first mode is the co-translational mode, in which precursor proteins are recognized by a signal recognition particle (SRP) during their synthesis at the ribosome. The ribosome-nascent chain (RNC) attached to SRP is then redirected to the membrane via an interaction of SRP with its membrane receptor, FtsY. The interaction brings the RNC close to the

SecYEG translocon. The remaining part of the nascent chain is then synthesized directly into the pore of the SecYEG channel. The energy for translocation in this mode is provided by the continued synthesis of the nascent chain (**Figure 1-7a**) (Tsirigotaki et al., 2017, Peng et al., 2019).

The second mode is the post-translational mode, in which precursor proteins are first completely synthesized into the cytosol. They are then kept in an unfolded state by cytosolic chaperones, most often the SecB protein. Subsequently, precursor proteins are translocated through the SecYEG channel with the help of a secretion signal and SecA. Afterwards, precursor proteins are translocated through the SecYEG channel through the energy generated by SecA. Note, the PMF is also providing the energy for the transport (**Figure 1-7b**) (Tsirigotaki et al., 2017, Green and Mecsas, 2016).

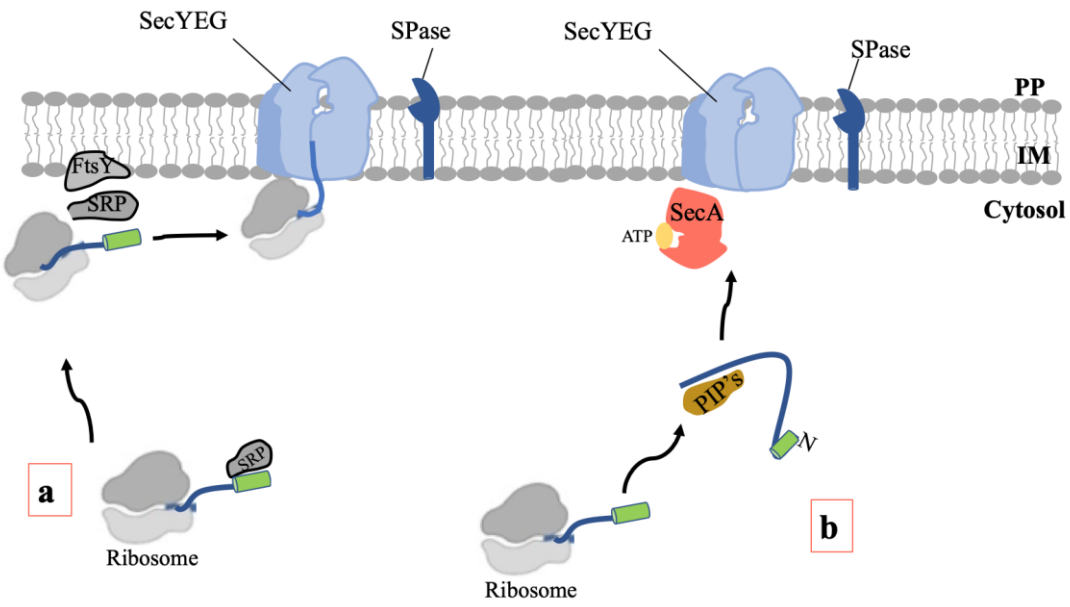


Figure 1-7: The mode of protein export by the Sec pathway
a) The co-translational export model: The signal recognition particle (SRP) binds to the nascent chain during synthesis at the ribosome. Then, the ribosome-nascent chain (RNC) is directed to the SRP receptor, FtsY. The ribosome is then close proximal to the SecYEG channel and the rest of the nascent chain is synthesized and exported through the SecYEG translocon. **b**) Post-translational export model: A polypeptide is first synthesized completely inside the cytosol. Sec-dependent substrates remain attached to post-translationally interacting proteins such as SecB. Then, it is handed over to the ATPase SecA. The whole complex is redirected to the SecYEG translocon. The secretion signal (highlighted as a green ellipse) is cleaved off by SPase, located in the inner membrane.

The translocation capacity of the Sec pathway is limited, and overexpression of target proteins may lead to an intracellular protein accumulation. In turn, this can result in inclusion body formation, protein aggregation, and/or proteolysis. To address this issue,

overexpression of proteins involved in the Sec pathway, such as SecYEG and the SecB protein, can be beneficial (Kleiner-Grote et al., 2018). It should be noted that the target protein and dedicated secretion signal have significant influence on the secretion efficiency of the Sec pathway. Studies on biotechnological applications have revealed that there is no universal secretion signal that can be used on different target proteins that promotes the best secretion efficiency (Freudl, 2018). In other word, the best-performing secretion signal should be experimentally identified for each target protein.

1.4.1.2 Tat pathway

The second major general secretion system, present in many bacteria, is the twin-arginine (Tat) pathway. It has been reported that around 80% of all bacteria contain the Tat pathway (Anné et al., 2016). The pathway gets its name from the presence of two conserved arginine residues in the secretion signals of target proteins, which are located at the N-terminus. After secretion, these signals are cleaved off by a membrane signal peptidase, and the mature protein is then released into the trans-side of the membrane (Freudl, 2018).

A special feature of this pathway is that proteins are transported correctly folded and in their final oligomeric state across the inner membrane. If needed, a dedicated cofactor is bound tightly, or even covalently, to the protein before secretion (Walker et al., 2015). This property of the system is especially important for proteins that require cytosolic conditions and/or cytosolic enzymes for appropriate folding. It appears that there is some form of quality control that only allows the secretion of proteins in their correct folding. This feature results in a higher quality of recombinant protein in the periplasm compared to the Sec pathway (Walker et al., 2015, Kleiner-Grote et al., 2018).

The Tat pathway consists of mainly three proteins: TatA, TatB, and TatC (**Figure 1-8**). TatB and TatC form an oligomeric receptor that binds to the secretion signal. After interacting with the secretion signal, TatA is recruited by the energy provided by the PMF. TatA forms a multimeric protomer translocation channel in the inner membrane. The number of TatA subunits depends on the size of folded protein (Kleiner-Grote et al., 2018, Walker et al., 2015). Interestingly, a target protein is either translocated through a pore, created by a variable number of TatA subunits (Gohlke et al., 2005), or passed through a TatA-induced weak region of the membrane, closely located to the TatBC receptor (Brüser and Sanders, 2003).

The Tat pathway is less efficient compared to the Sec pathway. In addition, secretion through the Tat pathway is slower due to the lower number of translocons within the inner membrane. However, the quality of secreted proteins in the Tat pathway is higher than the Sec pathway. To overcome the limitations of the Tat pathway, overexpression of Tat translocation has been investigated in *E. coli* (Matos et al., 2012). The overexpression of involved chaperones, such as DnaK, showed improved secretion through the Tat pathway. Note that the translocation efficiency of the Tat pathway is anyway limited by the source of energy, the PMF (Bergès et al., 1996, Kleiner-Grote et al., 2018).

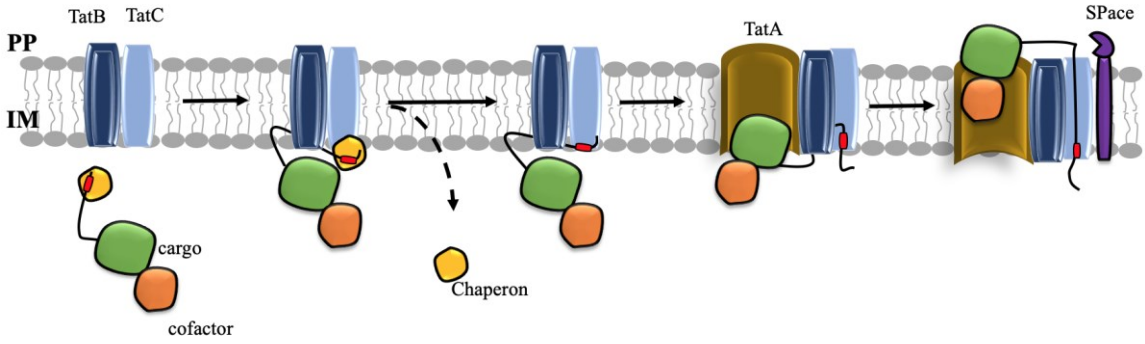


Figure 1-8: Schematic representation of the translocation through the Tat pathway
The cargo is synthesized and folded completely in the cytosol, and a cofactor binds to the protein if needed. A chaperone accompanies the protein by binding to the secretion signal (shown in red ellipse). The chaperon-protein complex is recognized by the TatB protein in the inner membrane. The inner membrane complex of the Tat pathway consists of TatB and TatC, and TatA added thereafter releasing the chaperon from the cargo. The secretion signal is removed by the inner membrane peptidase, SPase, and the cargo is released to the periplasmic space.

1.4.2 Translocation across the outer membrane

1.4.3 Type II secretion system

T2SS are widespread in both pathogenic and non-pathogenic Gram-negative bacteria. The T2SS apparatus spans both membranes, secreting folded substrates in two steps from the cytosol to the periplasmic space, and then into the extracellular space. The first step of the secretion depends on either the Sec or Tat pathway. T2SS substrates contain an N-terminal secretion signal which is cleaved off during passage through the inner membrane. Reports show that the transport through T2SS is fueled by ATP hydrolysis (Green and Mecsas, 2016, Nivaskumar and Francetic, 2014).

T2SS has a complex structure consisting of 12-15 components that together form different parts of the T2SS system, including an outer membrane complex, an inner membrane platform, a cytoplasmic ATPase, and a periplasmic pseudopilus (**Figure 1-9**). The structure

and biochemical function of some of these subunits have been determined, but the structure of the entire system remains to be elucidated (Costa et al., 2015, Cianciotto, 2005).

T2SS secretes a broad range of different substrates, which mostly contribute to the pathogenicity of harboring microorganisms, including the protease pseudolysin from *Pseudomonas aeruginosa*, the cholera toxin from *Vibrio cholera*, and the heat-labile enterotoxin from enterotoxigenic *E. coli* (Pourhassan et al., 2021). Hydrolyzing enzymes, such as proteases, lipases, and phosphatases, are also among T2SS substrates (Cianciotto, 2005). Interestingly, some bacterial species that live in harsh conditions, employ T2SS to secrete stable proteins, including hydrolyzing enzymes for nutrient acquisition, mobility, respiration, and biofilm formation (Nivaskumar and Francetic, 2014).

The secretion of the substrate across the OM specifically depends on the substrate's structure. However, the mechanism of substrate recognition has not yet been characterized, as the structure of the system is complicated. These are the reasons why the usage of T2SS in biotechnology has been limited (Burdette et al., 2018, Reed and Chen, 2013). The expression of T2SS from *Erwinia chrysanthemi* in a non-native organism (*E. coli*) has been performed successfully. However, this heterologous T2SS was able to secrete only its native substrate (He et al., 1991).

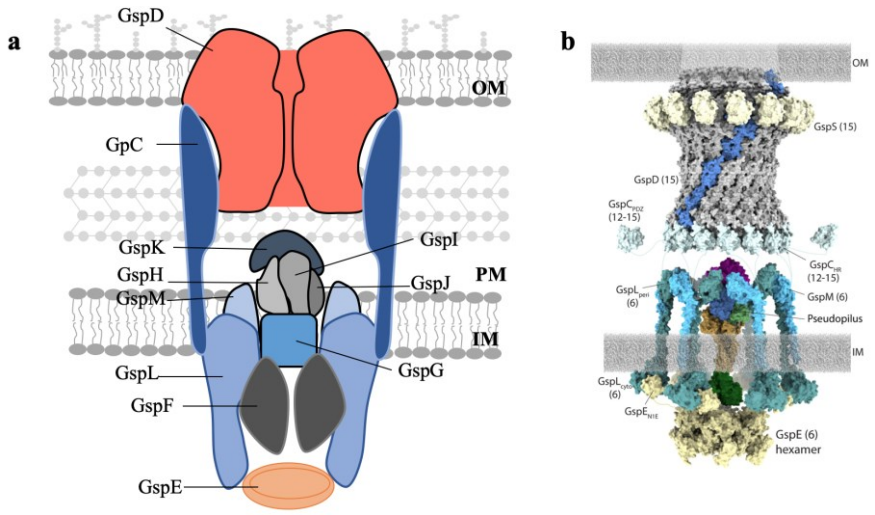


Figure 1-9: Structural organization of T2SS in the membranes
The T2SS is a double-membrane spanning system that translocates its substrate in two steps. The first step of translocation is across the IM, which depends on either the Sec or Tat pathway. The second step is translocation from the periplasm to the extracellular space. Around 12 kinds of protein components form the T2SS apparatus. **a**) Schematic representation of T2SS apparatus. The T2SS apparatus consists of a cytosolic ATPase (GspE), a pseudopilin (GspG, H, I, J, K), an IM complex (GspC, GspF, GspL, and GspM), and an OM complex (GspD). Some other proteins might be involved in the formation of T2SS apparatus in different species. The general secretion pathway (Gsp) shown here specifically belongs to *E. coli*. **b**) The T2SS apparatus was driven from the cryo-ET map of *Legionella pneumophila*, adopted from (Naskar et al., 2021).

1.4.4 Type V secretion system

T5SS, also known as autotransporter, secrete proteins into the extracellular space in a two-step mechanism. The transport across the inner membrane depends on the Sec pathway. Dedicated substrates harbor an N-terminal secretion signal, which is cleaved off during its passage over the inner membrane. Substrates of T5SS are mostly virulence factors, however, proteins involved in adhesion and biofilm formation can also be secreted through this system (van Ulsen et al., 2014). A polypeptide substrate of T5SS consists of three parts, including an N-terminal secretion signal for recognition by the Sec pathway, a passenger domain, and a C-terminal ß-barrel (translocator) domain (Reed and Chen, 2013). The five different groups of T5SS are schematically visualized in the **Figure 1-10**.

The mechanism of T5SS is unique as the substrate and its dedicated ß-barrel domain are fused. Firstly, a substrate is transported into the periplasm via the Sec pathway. Then, the ß-barrel domain is inserted into the outer membrane and creates a pore for the translocation of the passenger domain into the extracellular space (Burdette et al., 2018). The energy needed for the translocation across the outer membrane is provided by the sequential winding of the ß-strands. These ß-strands belong to the stacked ß-helical structures available in most autotransporters. Note that no ATP is present in the periplasm, and no proton concentration gradient is formed across the outer membrane (van Ulsen et al., 2014).

After the secretion process, the passenger domain either remains anchored to the translocator, or is cleaved off and released into the extracellular space (Costa et al., 2015). The property that a substrate might remain attached to the membrane can be used in *e.g.* live vaccine development, whole-cell biocatalysis, protein screening, epitope mapping, etc. (van Ulsen et al., 2014). Ulsen *et al.* reviewed a list of secreted heterologous proteins by T5SS including lipases for whole-cell biocatalysis, metallothionein for bioremediation, and the ß-subunit of cholera toxin (CBT) for translocation studies (van Ulsen et al., 2014).

Although different interesting applications of T5SS have been explored, the biotechnological applications of T5SS have been limited due to a low secretion titer and the narrow range of possible heterologous substrates. It should be noted that the size, complexity, and folding features of cargo have a great influence on the secretion via this autotransporter system (Reed and Chen, 2013).

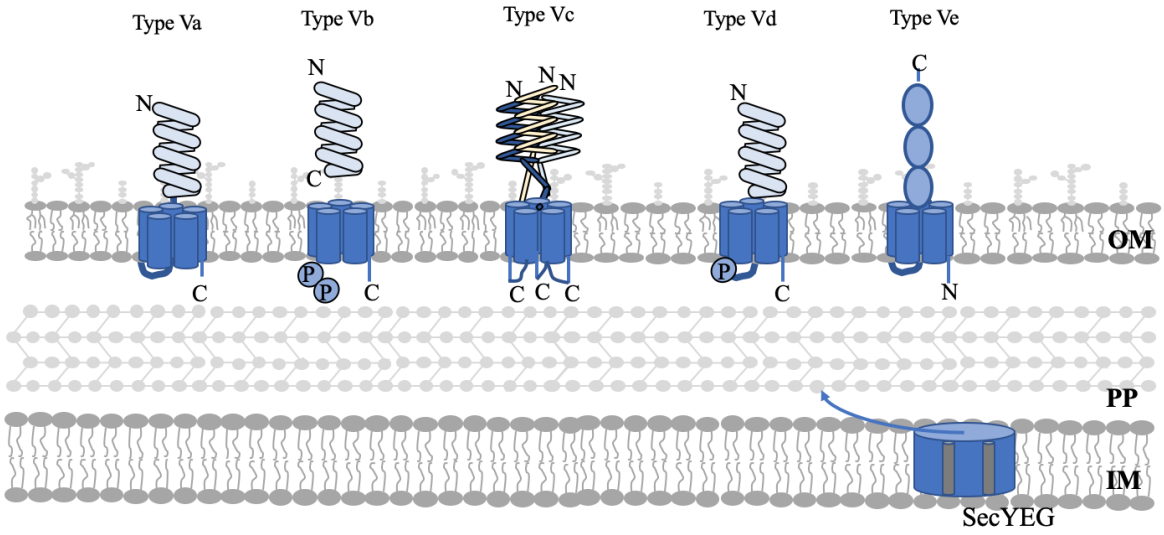


Figure 1-10: A schematic representation of five different models of T5SS
The domain located in the OM, was shown in dark blue, is the translocation domain (or ß-domain). The linker that associates the translocation domain with the passenger was shown as a dark blue line. The passenger domains were shown as coiled/circular shapes in the extracellular space. The N- and C- termini were identified. Type **Va** (classical autotransporter): The ß-barrel is inserted into the OM and the passenger domain remains anchored or cleaved off. Type **Vb** (two-partner secretion system): the passenger and translocator domains are synthesized in separate polypeptide chains. The translocation domain is inserted into the membrane with two POTRA domains (shown in two P-labeled circular shape). Type **Vc** (trimeric autotransporters). This group remains anchored to the membrane. The difference between these groups to the classical ones is their ability to trimerize. Type **Vd** (Fused-two partner secretion system): Here only one POTRA domain is involved. Type **Ve** (inverted-classical autotransporter): The only difference with the classical one is that they proceed from the N- to the C- terminus. The schematics were drawn based on (Leo et al., 2012).

1.4.5 Type VIII secretion system

Type 8 secretion system, also known as the curli biogenesis pathway, is used for the secretion of amyloid fibers (curli) onto the bacterial surface of Enterobacteriaceae. The main protein component of biofilms is the curli protein. Each curli consists of a Sec secretion signal at the N-terminal, an amyloid core, and a curli secretion signal (Van Gerven et al., 2015, Yan et al., 2020). The curli subunits are translocated in an unfolded state into the periplasm and kept unfolded by the periplasmic chaperone, CsgC. The assembly of curli fibers initiates upon reaching a specific nucleation concentration of these proteins. The assembly of curli subunits, CsgA, is mediated by the CsgB protein. A special transporter of this system, CsgG, with two accessory proteins, CsgE and CsgF, secretes curli subunits to the extracellular space (**Figure 1-11**). It is assumed that the translocation across the outer membrane is driven by an entropy gradient (Costa et al., 2015, Burdette et al., 2018, Yan et al., 2020).

Currently, the only host employed for T8SS engineering is *E. coli*. In a study conducted on T8SS, Gerven *et al.* employed the T8SS of *E. coli* to secrete non-native polypeptides up to 260 residues fused to CsgA subunits. They reported that the secretion efficiency depends on the folding properties of the fused passenger. The curli biogenesis pathway can be used for the surface display of small heterologous proteins (Van Gerven et al., 2014). Thus, it seems that the size of the passenger protein is the limiting factor for the secretion of heterologous proteins via T8SS (Burdette et al., 2018). An interesting application of the curli biogenesis pathway is the creation of a nanofibrous mesh with different functional properties on the surface of the bacterial host (Nussbaumer et al., 2017).

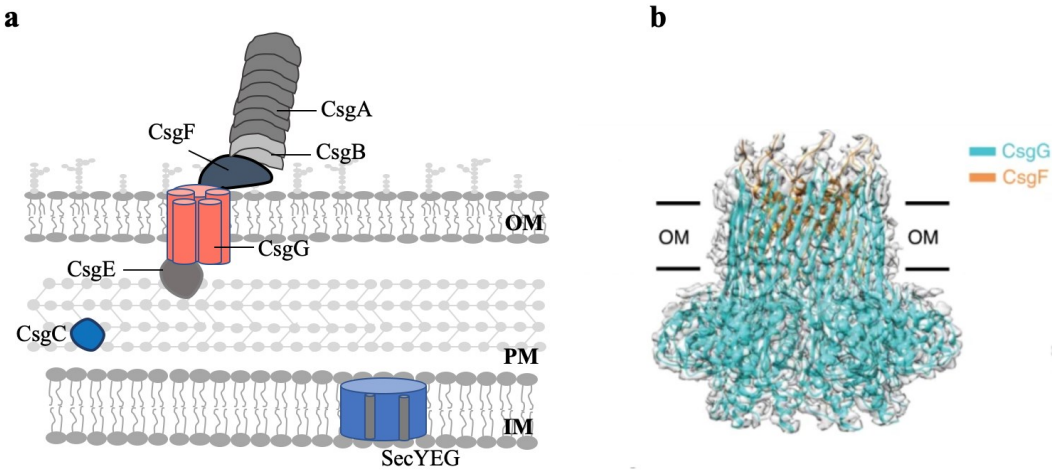


Figure 1-11: Structural view of T8SS (curli biosynthesis system)
Curli are extracellular fibers that function as protective layers for bacteria. The T8SS is known as the curli biogenesis system and is a single membrane-spanning system. **a**) Each curli consists of several CsgA (major subunits) and CsgB (minor subunits). The T8SS apparatus consists of CsgG and CsgE as main components, and two accessory proteins (CsgF and CsgC). **b**) Atomic model of the OM complex of the curli biogenesis system taken from (Yan et al., 2020).

1.5 The HlyA type 1 secretion system

In 1951, the HlyA was discovered in uropathogenic *E. coli* strains as a heat-labile and cell-associated toxin (Robinson, 1951). A paradigm of T1SS is the HlyA T1SS. HlyB and HlyD, the transport components involved in the HlyA secretion, were discovered later in the 1970s and 1980s, respectively (Noegel et al., 1979, Mackman et al., 1985). The description of TolC as the third membrane component of the HlyA T1SS, dates back to 1990 (Wandersman and Delepelaire, 1990).

HlyA is translocated through a channel, formed by HlyB, HlyD, and TolC, from the cytosol to the extracellular space in one step. The mechanism, functionality, and structure of this

system have already been studied for decades (Holland et al., 2016). Recently, in 2022, the stoichiometry and structure of the HlyB/D complex were determined by cryo-EM (Zhao et al., 2022). The HlyA T1SS is the main focus of this study, and the following sections will describe the components of this system in more detail.

1.5.1 HlyA

The substrate of HlyA T1SS, hemolysin A, is a pore-forming toxin belonging to the RTX (repeats in toxin) family. It is synthesized as an inactive protein, known as pro-HlyA that has a molecular weight of 110 kDa (1024 residues). The acyltransferase HlyC, a cytosolic protein, acylates the pro-HlyA at two lysine residues (K564 and K690) (Thomas et al., 2014c). Acylation of these lysine residues occurs prior to its secretion and turns pro-HlyA into an active toxin (a process that it is not required for secretion). Active HlyA can create a pore in the membrane of, for example, human erythrocytes and macrophages (Nicaud et al., 1985).

HlyA has three domains, including a functional N-terminal domain, an RTX domains consisting of six conserved GG repeats, and a secretion signal at the C-terminus (**Figure 1-12**) (Holland et al., 2016). All the information required for secretion is located within the last 50 to 60 C-terminal residues (Koronakis et al., 1989), and the presence of three RTX domains results in an enhanced secretion efficiency. Secretion of HlyA is directional, meaning the C-terminus reaches the extracellular space first (Lenders et al., 2015). The secretion signal is not cleaved off before, while, or after secretion.

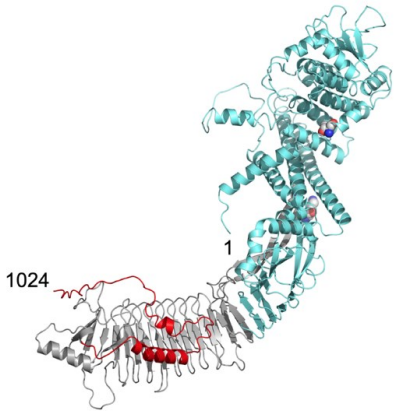


Figure 1-12: A model of pro-Hly predicted by AlphaFold2 (Jumper et al., 2021)
The secretion signal and the RTX in the C-terminus are colored red and gray, respectively. The pore-forming domain in the N-terminus is colored cyan. The two lysine residues (K564 and K690) which are acylated for activation of HlyA shown in spheres. The structure was adopted from (Spitz et al., 2022).

It has been shown that HlyA interacts with the transporter components before secretion (Lecher et al., 2012, Benabdelhak et al., 2003). For a long time, it was enigmatic which property of the secretion signal promoted this interaction, as sequence analysis of secretion signals of different RTX proteins did not reveal any defined sequence. Prior to the work of Spitz *et al*. (Spitz et al., 2022), the presence of an amphipathic helix as a secondary structure in the secretion signal was under debate (Koronakis et al., 1989, Stanley et al., 1991). Only recently, Spitz *et al.* provided evidence supporting the idea that the presence of an amphipathic helix between residues 975 and 987 is essential for the secretion process. Thus, a secondary structure in the secretion signal of RTX proteins promotes the secretion rather than a specific amino acid sequence (Spitz et al., 2022).

For studies on HlyA T1SS, two truncated forms of HlyA have been used. HlyA1, also termed HlyAc, has special importance, as it has been used frequently as a secretion carrier. HlyA1, with a molecular weight of 23 KDa, contains the secretion signal as well as three conserved RTX domains (**Figure 1-13b**). The first description of HlyA1 as a secretion carrier dates back to 1986 (Nicaud et al., 1986), and represents a turning point in the usage of HlyA T1SS for the secretion of heterologous proteins. Since then, a long list of heterologous proteins originating from prokaryotes or eukaryotes have been successfully secreted using HlyA1 (Pourhassan et al., 2021). However, it is acknowledged that the potential use of HlyA T1SS has not been fully exploited, due to its limitations concerning the narrow range of possible target proteins and the low yields of secretion (Khosa et al., 2018, Pourhassan et al., 2021).

The other most commonly used HlyA truncated version is HlyA2, which was derived from HlyA1, but lacks the secretion signal (**Figure 1-13b**). HlyA2 has been used in the work of Lecher *et al.*, in which a series of pull-down assays were conducted to map the interaction sites between HlyA and the CLD domain of HlyB (Lecher et al., 2012).

The encoding genes of HlyA T1SS are located in the *hly* operon, except for TolC. It was revealed that in *E. coli*, the transcriptional factor RfaH is also involved, as its binding to the *hly* operon causes the readout of a stem-loop termination region between the *hlyA* and *hlyB* genes (**Figure 1-13b**).

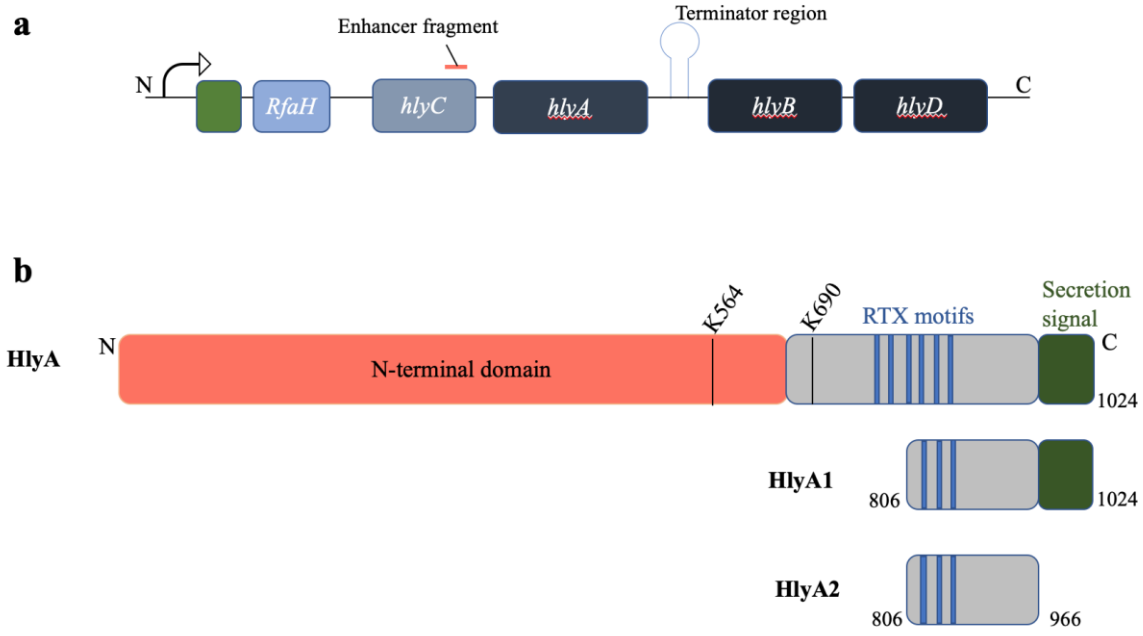


Figure 1-13: Schematic representation of the *hly* operon
a) The promoter and the transcriptional regulator (RfaH) were indicated, and a terminator region was indicated between the *hlyA* and *hlyB* genes. The *tolC* gene is not encoded by the *hly* operon. The C-terminus of the *hlyC* gene acted as an enhancer fragment, indicated with a red line. The Jumpstart (just upstream of many polysaccharide-associated starts) is shown in green box, which is a site that the RNA polymerase stops allowing binding of RfaH. **b**) Schematic representation of HlyA (full length) and two HlyA truncated versions (HlyA1 and HlyA2). The N-terminal functional domain (red) is followed by six RTX motifs and the C-terminal secretion signal. Two lysine residues (K564 and K690) that are acylated by HlyC were indicated. HlyA1 lacks three conserved GG repeats, while HlyA2 lacks three GG repeats plus the secretion signal.

1.5.2 HlyC

RTX toxins are expressed with a cognate acyltransferase, which is required for their activation. Two examples of these acyltransferases are: CyaC from *Bordetella pertussis* and RtxC from *Kingella Kingae* (Šebo et al., 1991, Osickova et al., 2018). HlyC is an acyltransferase dedicated to the HlyA T1SS, which is required for turning pro-HlyA into an active toxin before secretion. A mutation analysis showed that acylation of both residues is essential for the toxicity of HlyA (Thomas et al., 2014c).

HlyC transfers the acyl groups on two internal lysin residues of HlyA in collaboration with the acyl carrier protein-dependent fatty acylation (acyl-ACP). HlyC is synthesized in a 1:1 equimolar ratio to HlyA, and these two proteins form a stoichiometric complex (Holland et al., 2016). The structure of HlyC has not been yet solved, but a crystal structure of a HlyC homology model, ApxC, from *Actinobacillus pleuropneumoniae*, shows that this protein forms a dimer in the cytosol. ApxC protein has 70% amino acid identity with HlyC, and can perform the same functions in *E. coli*. The structure of ApxC consists of five β-sheets. A

cleft is formed by the contribution of two strands. This cleft appears to be conserved among acyltransferases of the Gcn5-like N-acyltransferase (GNAT) family, and harbors the active site residues involved in the catalysis (Greene et al., 2015, Holland et al., 2016).

Surprisingly, it has been reported that the C-terminus of the *hlyC* gene acts as an enhancer fragment for HlyA expression (**Figure 1-13a**). The presence of this region, upstream of the *hlyA* gene, led to an increased secretion efficiency. This region is rich in adenosine and uracil nucleotides, and it has an affinity for the ribosomal protein S1 (Khosa et al., 2018). The second chapter of this thesis describes the engineering of this enhancer fragment that resulted in a further improved secretion efficiency of the HlyA T1SS.

1.5.3 HlyB

HlyB belongs to the superfamily of ATP-binding cassette (ABC) transporter. Early studies on the topology of HlyB, mostly provided by co-constructional experiments, revealed that this ABC transporter harbors a cytosolic N-terminal domain C39-peptidase-like domain (CLD), a transmembrane domain (TMD) composed of hydrophobic transmembrane helices (TMHs), and a cytosolic C-terminal nucleotide-binding domain (NBD) (**Figure 1-14a**) (Wang et al., 1991, Holland et al., 2005).

For a long time, the available structures of HlyA T1SS were limited to a NMR structure of the CLD (Lecher et al., 2012), X-ray crystallography structure of the NBD (Schmitt et al., 2003), a crystal structure of truncated HlyD (lacking both N-terminal and C-terminal residues) (Kim et al., 2016), and a crystal structure of TolC (Koronakis et al., 2000). Only recently, structural information of the HlyA T1SS from single particle cryo-EM became available. The cryo-EM data represented an unusual structure of a heterododecameric assembly of three dimeric HlyB and dimeric HlyD units. Three protomers, each consisting of a HlyB dimer and two HlyD molecules, form the inner membrane complex of the HlyA T1SS (**Figure 1-14b**). In addition, Zhao *et al.* reported that one of the HlyB dimers provides the channel for the translocation of HlyA, while all three contribute to ATP hydrolysis to provide the energy for transport (Zhao et al., 2022).

The N-terminal domain (the CLD) resembles C39-peptidases, which are cysteine proteases from bacteriocin exporters that cleave off the C-terminal secretion signal from a protein or peptide prior to, or during the secretion process. Despite almost 40% homology between the CLD and C39-peptidases, no proteolytic activity is known for the CLD of HlyB, due to an

inactivated catalytic triad (containing a tyrosine instead of a cysteine in the catalytic triad). Nevertheless, the CLD of HlyB is still essential for secretion (Lecher et al., 2012).
In vitro analysis of the isolated CLD revealed an interaction between this domain and HlyA1 and/or HlyA2. This interaction is independent of the secretion signal, as the HlyA2, a HlyA truncated version without the secretion signal, also interacts with the CLD. This interaction likely depends on the RTX domain (Lecher et al., 2012). Interestingly, the cleavage site of an active C39-peptidase is located directly after a double-glycine motif (Wu and Tai, 2004). Pull-down assays, along with NMR structure, precisely mapped substrate-binding regions of the CLD (Lecher et al., 2011, Lecher et al., 2012).

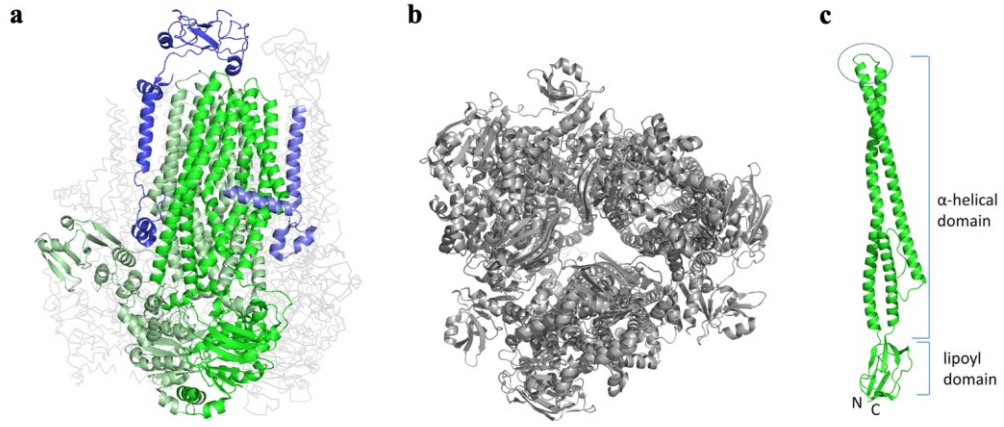


Figure 1-14: Overall structure of the HlyB /HlyD complex
a) Structure of HlyB dimer (green), the ABC transporter of the HlyA T1SS from *E. coli* associated with two molecules of HlyD (blue). **b**) Cartoon representation of the trimer of dimers (PDB 7SGR) of the HlyB NBDs (Zhao et al., 2022). The view is from the cytosol toward the membrane. The structures a and b were drawn using Pymol software. **c**) The structure of a HlyD protomer. The α-helical domain in the tip region indicated by a blue circle. The panel c adopted from (Kim et al., 2016).

It was suggested that the CLD domain performs a chaperone-like activity for HlyA (Lecher et al., 2012). This suggestion was established from different observations, including:

- Secretion of HlyA, the large polypeptide with a 110 kDa molecular weight, is C-directional, so it must be completely translated into the cytosol before secretion, and remain unfolded. (Bakkes et al., 2010, Lenders et al., 2015). Thus, the role of a chaperon might be required to protect the unfolded HlyA from the cytosolic proteases.
- Secretion of HlyA is independent of the general chaperones, such as SecB. In addition, no chaperone is encoded by the *hly* operon (Holland et al., 2016, Bakkes et al., 2010).

- The aforementioned interaction between HlyA and the CLD occurs only with the unfolded substrate (Lecher et al., 2012).

The NBD provides the energy for transport by hydrolyzing ATP (Holland et al., 2005). The NBD of HlyB plays a role in one of the early steps of secretion, which involves substrate recognition through a specific interaction between the NBD and the C-terminus of HlyA (Benabdelhak et al., 2003). Unlike the CLD, the NBD does not interact with HlyA2, indicating that this interaction is depending on the secretion signal. The presence of ATP causes the dissociation of the NBD/HlyA complex (Benabdelhak et al., 2003). The interaction between HlyA and the NBD is the main focus of chapter 3 of this study.
The NBD is highly conserved among ABC transporters. Three motifs are characteristics for all ABC transporters: the walker A motif with the sequence GXXGXGKS/T (where X refers to any aa), the walker B motif with the sequence ΦΦΦΦD (where Φ refers to any hydrophobic aa), and the C-loop, known as the signature motif, with the sequence LSGGQ. Besides that, the Q-loop, D-loop, and the H-loop are other conserved motifs present in NBDs (Zaitseva et al., 2005b, Holland et al., 2005).
The NBDs of HlyB are symmetric. Dimerization of both HlyB monomers is induced by binding of two ATP molecules. Each site is formed by coordination of the Walker A and Walker B of one monomer, and the C-loop of the second monomer, and vice versa. The D-loop is involved in the formation of the dimer interface, while the H-loop is involved in monomer-monomer communication. ATP hydrolysis requires Mg^{2+} ions as cofactor (Kanonenberg et al., 2019b). Reconstitution of the HlyB was performed successfully in lipid-nanoparticles, showing that interactions between HlyB and the substrate is still possible in the absence of other HlyA system components. However, we know that the presence of HlyD and TolC are essential for secretion (Kanonenberg et al., 2019b).

1.5.4 HlyD

HlyD belongs to the family of membrane fusion proteins and is known as the adaptor protein of the HlyA T1SS. It stabilizes the tunnel channel of the HlyA T1SS, by connecting HlyB and TolC (Holland et al., 2016). HlyD consists of 478 residues and has a molecular weight of 55 kDa. Data on the topology of the HlyD shows the presence of an N-terminal cytosolic region (residues 1-59), followed by a single transmembrane helix (residues 60-80), and a large periplasmic region (residues 81-487) (Wang et al., 1991). The first solved structure of HlyD covered residues 96-361 and showed a long α-helical domain, consisting of an α-

helical hairpin along with a lipoyl domain located in the periplasm. It is assumed that the tip of the α-hairpin interacts with the TolC protein (**Figure 1-14c**) (Kim et al., 2016).

Sequence analysis revealed three features of the HlyD cytoplasmic domain (CD): a putative amphipathic-helix (2-23), a box of five charges amino acids (34-38), and three proline residues towards the C-terminus of the CD (Holland et al., 2016, Balakrishnan et al., 2001). The first two features are essential for secretion, but only the charged box is a prerequisite for the TolC recruitment (Balakrishnan et al., 2001, Thanabalu et al., 1998). The presence of the charged box of HlyD (5 residues) was first demonstrated by Balakrishnan *et al.*, and later supported by Zhao *et al.*. It was shown that the residues of the charged box (E35, D37, and E38) are located in the main interface between two protomers (Zhao et al., 2022). Additionally, a network of electrostatic interactions between the charged box of HlyD and the CLD of HlyB is necessary for substrate secretion (Zhao et al., 2022).

A conserved region, covering the last 35 residues of HlyD-like proteins, was identified. However, its functionality is still unknown (Schülein et al., 1994). Residues 127-170 show 47% identity with residues 233-274 from the β-barrel region of TolC, but the functionality of this region remains unclear (Schülein et al., 1994).

Deletion of the first 26 HlyD residues reduces the secretion by around 90% compared to the wild-type HlyD. Also, deletion of either the first 45 residues, residues 26-45 or the last 10 C-terminal residues of HlyD, completely abolishes secretion (Balakrishnan et al., 2001, Thanabalu et al., 1998). A mutant lacking the first 45 residues can still insert into the membrane, but this mutant is unable to recruit TolC, even in the presence of HlyA and HlyB (Pimenta et al., 1999). The results of a mutagenesis study on HlyD confirmed that the integration of HlyD in the HlyA T1SS affects the correct folding of HlyA in the extracellular space, via an unknown mechanism (Pimenta et al., 2005).

Balakrishnan *et al.* reported that the HlyD CD is highly conserved in T1SS. They proposed that an interaction between the HlyD CD and HlyA is important for substrate recognition. Upon this interaction, HlyB and HlyD form a complex together. It was further suggested that this region plays a role in TolC recruitment (Balakrishnan et al., 2001).

Early studies, conducted by the Koronakis laboratory, suggested a trimeric state for HlyD (Balakrishnan et al., 2001), however after analysis of other MFP analogs, a hexameric state seemed more valid. Information obtained from the Cryo-EM confirmed this hexameric state of HlyD, in which two monomers of HlyD are associated with a HlyB dimer (**Figure 1-14a**) (Zhao et al., 2022).

Each HlyD subunit is located on the opposite side of the HlyB dimer. The transmembrane region of each HlyD is packed against a HlyB dimer. The cryo-EM information also revealed an amphipathic helix (AH) between residues 9 to 24. This AH is located parallel to the inner membrane and connects hydrophobically to HlyB (Zhao et al., 2022). The very last C-terminal residues of HlyD (residues 355 to 471) form a β-barrel domain and α-helix, which both lie at the trans-side of the inner membrane. The periplasmic region of HlyD is not well visible in the absence of TolC and substrate, suggesting that HlyD has a very flexible structure. Furthermore, the data also confirms that even in the absence of a substrate, HlyB and HlyD form an inner membrane complex together (Zhao et al., 2022).

1.5.5 TolC

TolC is the outer membrane protein of the HlyA T1SS with a molecular weight of 53 kDa. Insertion of TolC into the outer membrane is Sec-dependent. The 20-residues long secretion signal of TolC is cleaved off, after which the mature protein is translocated into the outer membrane (Wandersman and Delepelaire, 1990). Unlike the HlyB and HlyD, TolC does not only dedicate itself to the HlyA T1SS, but also associates with the tripartite pumps AcrAB-TolC (Du et al., 2014) and MacAB-TolC (Fitzpatrick et al., 2017). TolC protects *E. coli* from different cationic dyes, a range of antibiotics, and some detergents. Furthermore, TolC is an uptake site for, for example, colicins (Holland et al., 2016). In the AcrAB pump, TolC interacts with a resistance-nodulation-division (RND) secondary active transporter, instead of with an ABC transporter, to extrude small molecules from the cell (Du et al., 2014).

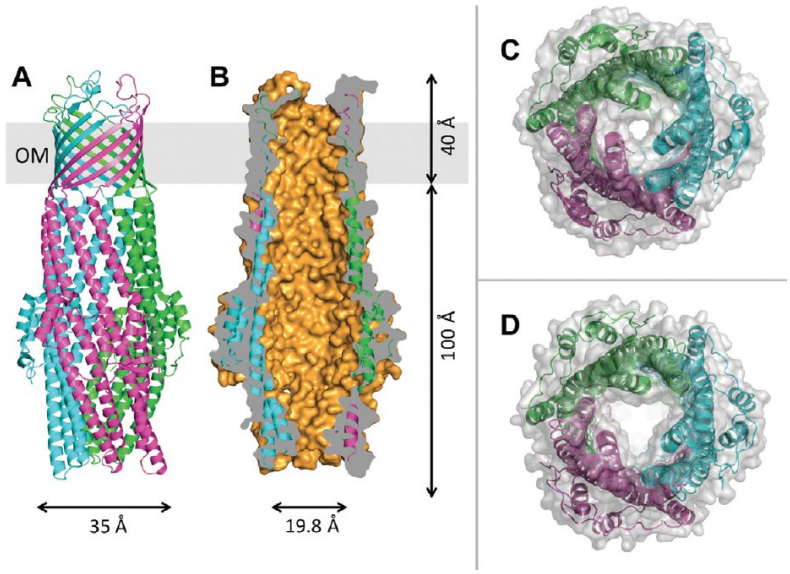


Figure 1-15: Crystal structure of TolC from *E. coli* (pdb: IEK7)
A) The close conformation of TolC. TolC extended 100 Å into the periplasmic space and 40 Å into the outer membrane. Each monomer was colored in a different color. **B**) The cross-sectional view of TolC. The view of TolC in close (pdb: 1EK9) (**C**) and open (pdb: 2XMN) (**D**) conformations. The view is from the periplasm toward the outer membrane. The figure adopted from (Lenders et al., 2013).

Note that many trimeric OMPs form a pore, each with their dedicated monomers. However, in the case of TolC, all three monomers contribute to the formation of a central pore (Koronakis et al., 2000). The crystal structure of TolC, solved by Koronakis in 2000 (resolution of 2.1 Å), revealed that TolC form a homotrimer into the outer membrane. Overall, TolC has a length of 140 Å, with an extension into the periplasmic space of 100 Å (**Figure 1-15 a, b**). TolC forms a β-barrel inside the outer membrane, consisting of 12 β-stranded β-sheets, and an α-helical domain that residues in the periplasmic space, consisting of 12 α-helices (Koronakis et al., 2000).

Unlike most other OMPs, TolC has no lid that covets the barrel opening. However, the periplasmic entrance of TolC can range from a diameter of only 3.9 Å in its closed state, and up to a maximum diameter of 22 Å in its open state, illustrating that a lid is not necessary. The opening of the entrance towards the periplasm follows an iris-like motion, which is triggered by the substrate (**Figure 1-15 c, d**). The diameter of open TolC allows the translocation of small molecules from the tripartite efflux system, or α-helixes belonging to the T1SS substrates, indicating that the T1SS only supports the transport of unfolded proteins (Koronakis et al., 2000, Masi et al., 2009).

1.5.6 Secretion mechanism of HlyA T1SS

Laboratory strains, such as *E. coli* K-12 and *E. coli* BL21, do not harbor the genes encoding HlyA T1SS. So, the expression of HlyA T1SS components is plasmids-based (Khosa et al., 2018) but TolC protein is encoded endogenously (Wandersman and Delepelaire, 1990).

HlyA secretion is a stepwise process. First, the very long polypeptide, is completely synthesized in the cytosol, to our knowledge without the mediation of chaperones. Then, the substrate is recognized by the secretion machinery, which happens through interactions between the substrate and the cytosolic domains of the inner membrane complex HlyB and HlyD. The CLD and substrate do not interact by means of the secretion signal, but most likely through the RTX domain (Lecher et al., 2012). However, an interaction between the NBD and substrate is suggested to involve the secretion signal region (Benabdelhak et al., 2003). The secretion signal is mapped to the last 50 or 60 C-terminal residues. However, the presence of three RTX domains guarantees efficient secretion (Nicaud et al., 1986). The secretion signal remains intact during and after the secretion (Koronakis et al., 1989).

Interaction between the HlyD CD and HlyA have been suggested to involve a putative amphipathic helix. The interaction between HlyD and substrate is independent of HlyB. All the above-mentioned interactions are necessary for successful secretion. Deletion of either of the cytosolic domains completely blocks the secretion process (Thanabalu et al., 1998, Balakrishnan et al., 2001).

After the recognition step, TolC is recruited through a yet unclear mechanism by HlyD. Then, together with HlyB, the three proteins form a tunnel-channel, redirecting HlyA in one step from the cytosol to the extracellular space, with the C-terminus of HlyA reaching the extracellular space first. ATP hydrolysis, together with the PMF provides the energy for transport (Schmitt et al., 2003). The PMF is also involved in powering the system (Rempel et al., 2019). As soon as the C-terminus exits from the tunnel channel, Ca^{2+} ions bind to the RTX domain, promoting the folding process (Lenders et al., 2015). This folding prevents the backsliding of the substrate inside the cytosol (Bumba et al., 2016). A schematic visualization of the secretion process is presented in **Figure 1-16**.

Secretion begins in the early logarithmic phase of the growth cycle, and it switches off before the stationary phase. HlyA secretion is secreted with a rate of 16 residues per transporter per sec. Thus, only 1 minute is required for the secretion of a single HlyA molecule (Lenders et al., 2016).

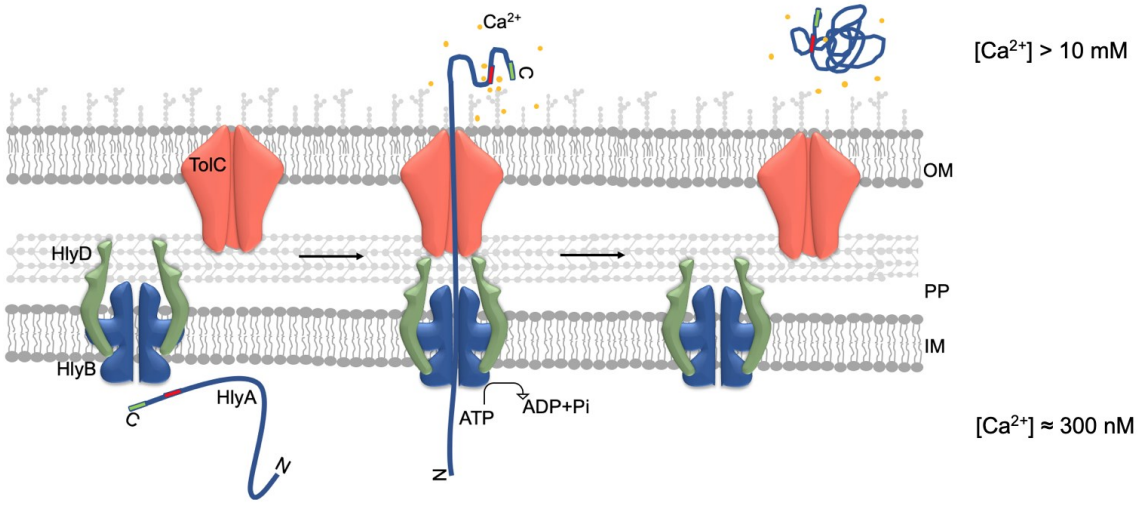


Figure 1-16: The mechanism of the secretion through the HlyA T1SS
The secretion of proteins through the HlyA T1SS is a process that involves the interaction of different components. The inner membrane complex consists of HlyB and HlyD, which interact with the substrate HlyA to be secreted. This interaction recruits TolC, which forms a tunnel channel with the inner membrane complex. The substrate is then transported through the tunnel channel and secreted outside the cell. After the secretion cycle is complete, TolC leaves the complex, although the inner membrane complex remains assembled.

1.6 Protein engineering for improving the secretion efficiency

Protein engineering is a powerful approach to enhance the properties of proteins/enzymes, with the aim to improve their specific function. Protein engineering is based on gene modifications, hence, advantages in the molecular sciences can be beneficial as well (Bilal et al., 2018). Protein engineering includes directed protein evolution and (semi)rational design. Rational design is a method of improving protein functions based on knowledge of the structure-function relationship of proteins. Directed Evolution, on the other hand, is an iterative process where researchers create random variations in a protein's genetic code and use selection techniques to identify the variants with improved function (Cheng et al., 2015). Engineering bacterial secretion systems is difficult, as these nanomachines are in many cases multicomponent systems that have complicated structures (Pourhassan et al., 2021). So far, different approaches have been applied to improve the secretion titers of bacterial secretion systems for heterologous recombinant proteins. **Table 1-3** summarizes some of these efforts.

Table 1-3: Examples of efforts to improve secretion efficiency of secretion systems

Organism	Strategy for engineering	cargo	Improved properties	Reference
E. coli	Applying an enhancer fragment	Lipase	Secretion efficiency of the HlyA T1SS	(Khosa et al., 2018)
Bacillus subtilis	Inactivation of six extracellular proteases and setting of a novel secretion cassette	ß-lactamase and other proteins	Quality and quantity (up to 18-fold) of the secreted proteins	(Wu et al., 1991)
Streptomyces lividans	Overexpression of phage-shock protein A (involved in maintenance of the PMF)	xylanaseC and GFP	Secretion efficiency of the Sec- and Tat-pathways	(Vrancken et al., 2007)
E. coli	Reducing metabolic burden, removal of protein degradation system, setting of a novel secretion cassette	heterologous enzyme (cutinase)	Secretion efficiency of T3SS Up to 24-fold	(Green et al., 2019)
Corynebacterium glutamicum	Use of a Sec signal library from *Bacillus subtilis*	cutinase	Secretion efficiency of Gram-positive bacteria	(Hemmerich et al., 2016)

1.6.1 Directed evolution

The most versatile method suitable to improve protein properties is directed protein evolution (Wong et al., 2006). Directed evolution was first described by Spiegelman as a concept of Darwinian experiments in a test tube (Mills et al., 1967). A major advantages of directed protein evolution is that it does not require knowledge of the 3D structure of a target protein, or even the related mechanism-structure relationship. Therefore, it has been applied to many different proteins and enzymes. There are thousands of successful case studies that employed directed evolution for improving the properties of proteins/enzymes. In 2018, the Nobel prize in chemistry for directed evolution highlighted once again the importance of this technique for different applications (Bornscheuer et al., 2019).

To achieve successful directed protein evolution, the following steps should be performed:

- A library generation method that generates unbiased diversity within the gene of interest.

- A selection strategy to ensure that the desired mutations are maintained in the population.
- A recombination strategy to combine different mutations and gain further improvements.

The KISS (Keep it simple and smart) principle was introduced by Schwaneberg and coworkers, and covers the design of different steps that lead to a successful protein evolution experiment (Wong et al., 2006). The basic requisites for a successful evolution are listed below:

- Adequate number of mutations in the gene of interest.
- Functional expression of the target protein in a lab-suitable host strain
- Reliable screening system to select the best performing-variants
- The use of computational tools to predict the effects of mutations on protein structure and function can be beneficial as well. In particular for designing the library and selecting the best-performing variants (Wong et al., 2006, Packer and Liu, 2015).

To sum up, directed evolution involves the generation of a library of mutants, screens for the desired trait, and optimization of the mutant. Once the optimized mutant has been identified, it can be tested *in vivo* to confirm its improved performance. Once the desired trait has been identified, the mutant can be characterized and further optimized. This can be achieved by applying additional mutagenesis methods by directed evolution and/or rational design. Usually, the directed evolution procedure is performed in iterative cycles of mutant library generation and screening, until the desired improvement is obtained (Packer and Liu, 2015).

Error-prone PCR (epPCR) is one of the most successfully applied techniques to introduce mutations in the gene of interest. epPCR is simple, robust, and adaptable (Wong et al., 2004). This technique is based on inaccurate polymerization of the gene of interest. It can be done either by using a low-fidelity DNA polymerase, increasing the concentration of Mn^{2+} in the PCR reaction buffer, by using unbalanced dNTPs (Wong et al., 2006). Despite the advantages of using epPCR for library generation, there are some limitations to using this method as well. First of all, epPCR is not completely random, due the degeneracy of genetic codes. Secondly, the biased mutational spectra of DNA polymerases are also a limiting factor (Vanhercke et al., 2005). Finally, reports show that epPCR is unable to identify around 70% of the beneficial positions of a protein (Yang et al., 2017).

To address the latter issue, Schawneberg and coworkers developed a novel library generation strategy, named casting error-prone PCR (cepPCR). In this strategy, the gene of interest is divided into smaller fragments, but only one fragment is subjected to library generation and screening each time. In the end, all identified positions related to each fragment are recombined to gain further improvements (Yang et al., 2017).

Another successful method for library generation is site saturation mutagenesis, which usually is applied to already identify beneficial positions (either determined by epPCR or cepPCR method). This allows researchers to generate complete diversity for each position, resulting in further exploration of the sequence. SSM generates substitutions that might be never generated via other library generation methods, such as epPCR (Wong et al., 2004). Methods like DNA shuffling (Stemmer, 1994) and StEP PCR (Zhao and Zha, 2006) are among other recombination methods that cause accumulation of beneficial positions in a way that enhances the desired property of the target protein.

The KnowVolution strategy, published in 2015, is a protein engineering approach that integrates directed evolution and computational analysis to maximize improvements while minimizing experimental efforts (Cheng et al., 2015). It consists of four phases: phase I involves identifying beneficial positions through random mutagenesis and screening; phase II involves subjecting potentially beneficial positions to site-saturation mutagenesis to explore full diversity and eliminate non-contributing positions. Phase III involves a computationally assisted analysis of beneficial amino acid substitutions, grouping those that might interact with each other and those that are independent. This analysis helps identify further beneficial amino acid positions. In the final phase, clustered or interacting positions are recombined to yield final variants with optimized performance (Islam et al., 2018).

Aims of the thesis

The HlyA T1SS from *E. coli* is the best-studied type 1 secretion system in terms of structure and secretion mechanism (Kanonenberg et al., 2018). The simple structure and easy plasmid-based expression make it attractive candidate for different applications (Burdette et al., 2018). Since the 1990s, many studies have explored this system for secreting a wide range of proteins originated from both prokaryotic and eukaryotic cells (Pourhassan et al., 2021, Koronakis et al., 1989, Nicaud et al., 1986, Schwarz et al., 2012). These studies mainly focused on two main areas: production of heterologous proteins and production of antigens for vaccination.

The ability of the HlyA T1SS to correctly folding of proteins, even those with disulfide bounds, is an important feature of the HlyA T1SS that makes it a favorable system for biotechnological applications (Schwarz et al., 2012, Ruano-Gallego et al., 2019). Large-scale secretion using the HlyA T1SS has already demonstrated the potential of the system for commercial uses (Ihling et al., 2019). The successful application of the system in producing peptides as active pharmaceutical ingredients highlighted the importance of further developing the system (Nguyen et al., 2021).

Although many heterologous proteins have been successfully secreted using the HlyA T1SS, the exploitation of the system is still facing two challenges: not all heterologous proteins can be secreted through the system, and many proteins can be secreted only in a low secretion titer (Khosa et al., 2018, Burdette et al., 2018, Kanonenberg et al., 2018, Pourhassan et al., 2021). To overcome these challenge, further optimization of the system is necessary. However, efforts towards improving the secretion efficiency of the HlyA T1SS have been limited.

The overall aim of the thesis was to enhance the secretion efficiency of the HlyA T1SS. To achieve this, the following objectives were set: 1) To increase secretion efficiency through directed evolution of the enhancer fragment, the ABC transporter HlyB, and the membrane fusion protein HlyD. 2) To evaluate the potential for further improvement by investigating the impact of the terminator region of the system. 3) Recently, it was discovered that the number of TolC proteins remains constant even with overexpression of HlyB and HlyD (Beer et al., 2022). This raised the question of whether the number of TolC is a limiting factor for secretion through the HlyA T1SS. Thus, this became another objective of the study to be addressed.

A deeper understanding of the system could contribute to increase the secretion efficiency. To achieve this objective, a functional study was conducted on the interaction between the

substrate HlyA and HlyB. The aim was to map areas the regions on the HlyB NBD that are involved substrate binding.

2 Publications

Chapter I: T1SS – an alternative platform for protein production

Chapter II: Directed evolution of the Hly enhancer fragment

Chapter III: Substrate binding sites of hemolysinB

Chapter IV: Engineering of the inner membrane complex of the HlyA T1SS

Chapter V: TolC overexpression impairs the hemolysin secretion

2.1 Chapter I: T1SS – an alternative platform for protein production

Title: Biotechnological applications of type 1 secretion systems

Authors: Zohreh Pourhassan N., Sander H.J. Smits, Jung Hoon Ahn, Lutz Schmitt

Published in Biotechnology Advances

Impact score (2021): 16.85

Own proportion of this work: 75 %

Biotechnology Advances 53 (2021) 107864

Contents lists available at ScienceDirect

Biotechnology Advances

journal homepage: www.elsevier.com/locate/biotechadv

Research review paper

Biotechnological applications of type 1 secretion systems

Zohreh Pourhassan N. [a], Sander H.J. Smits [a], Jung Hoon Ahn [b], Lutz Schmitt [a,*]

[a] Institute of Biochemistry, Heinrich Heine University Düsseldorf, Universitätsstr. 1, 40225 Düsseldorf, Germany
[b] Department of Chemistry and Biology, Korea Science Academy of Korea Advanced Institute of Science and Technology, Busan 47162, South Korea

ARTICLE INFO

Keywords:
Protein secretion
Gram-negative bacteria
Fusion proteins
Vaccination
Protein engineering

ABSTRACT

Bacteria have evolved a diverse range of secretion systems to export different substrates across their cell envelope. Although secretion of proteins into the extracellular space could offer advantages for recombinant protein production, the low secretion titers of the secretion systems for some heterologous proteins remain a clear drawback of their utility at commercial scales. Therefore, a potential use of most of secretion systems as production platforms at large scales are still limited. To overcome this limitation, remarkable efforts have been made toward improving the secretion efficiency of different bacterial secretion systems in recent years.

Here, we review the progress with respect to biotechnological applications of type I secretion system (T1SS) of Gram-negative bacteria. We will also focus on the applicability of T1SS for the secretion of heterologous proteins as well as vaccine development. Last but not least, we explore the employed engineering strategies that have enhanced the secretion efficiencies of T1SS. Attention is also paid to directed evolution approaches that may offer a more versatile approach to optimize secretion efficiency of T1SS.

1. Introduction

The production of proteins is a pivotal step in their study and in their diverse applications that was initially based on the isolation from natural resources such as animal and plant resources (Rosano et al., 2019; Puetz and Wurm, 2019). For instance, in the 1920s insulin for diabetes therapy was provided by isolation from pig's pancreas (Owczarek et al., 2019). Historically, in 1982, the production of a recombinant biopharmaceutical, human insulin, for the first time met all the required medical standards and quality constraints and was approved by the Food and Drug Administration (FDA) (Johnson, 1983; Owczarek et al., 2019). This approval is considered as a breakthrough in the introduction of a novel technology for protein synthesis to biosciences (Owczarek et al., 2019). Since then, recombinant protein technologies have made significant progress toward providing a variety of proteins for different applications (Owczarek et al., 2019), and thus far the market of the recombinant biological protein drugs has reached approximately 270 billion USD in 2019 (https://www.prnewswire.com/), representing the vast demands on the today recombinant proteins.

For the production of recombinant proteins, a wide range of platforms have been developed that originated from either prokaryotic or eukaryotic cells, including bacteria, fungi, yeasts, insect cells, microalgae, mammalian cells, transgenic animals, and plants (Puetz and Wurm, 2019). There are limits to how efficient these production platforms can be applied for a specific target protein to be produced in optimum quality and yields yet at a reduced cost (Rosano and Ceccarelli, 2014). For instance, mammalian cells are employed for complex proteins, but they grow slowly even on expensive media. In contrast, yeasts grow faster, but the genetic instability and clonal variation are weaknesses of yeast platforms (Burdette et al., 2018). In principle, *Escherichia coli*, a bacterial host, is often one of the first organism of choice for producing recombinant proteins in research laboratories (Cronan, 2014) and used in many cases in industrial scales (Huang et al., 2012). The main reasons for this preference seem to include: (i) the short and rapid growth cycle of this microorganism on the relatively cheap carbon resources, (ii) the ability to produce a broad range of proteins in high titers and yields, (iii) the availability of biochemical and genetical methods, and (iv) a reduced cost of production (Jia and Jeon, 2016; Huang et al., 2012).

Today, almost one-third of already FDA-approved recombinant proteins - in total more than 170 unique recombinant proteins were approved so far - are produced by *E. coli*, a Gram-negative bacteria, hinting at the importance of this workhorse for recombinant protein production (Puetz and Wurm, 2019; Sharma and Chaudhuri, 2017). A human growth hormone, interferon-α2b, β-glucosidase, a tumor necrosis factor, and interleukins are just a few examples of an already long list of

* Corresponding author.
E-mail address: lutz.schmitt@hhu.de (L. Schmitt).

https://doi.org/10.1016/j.biotechadv.2021.107864
Received 18 June 2021; Received in revised form 30 October 2021; Accepted 31 October 2021
Available online 10 November 2021

recombinant proteins approved for production in *E. coli* (Owczarek et al., 2019).

Despite the many and obvious benefits of using *E. coli* as a bacterial host, its biotechnological application still suffers from some limitations, for example: First, a bacteria-derived recombinant protein might be unstable, inactive, or not properly folded due to lack of the proper posttranslational modifications in bacteria, although a synthetic post-translational step could overcome this weakness (Owczarek et al., 2019). In addition, a high expression level might result in toxicity of the target protein for the producer host (Wingfield, 2015). High expression level of a target protein might also lead to inclusion body formation which has been considered for a long time as a drawback of using bacterial hosts for protein expression (Baneyx and Mujacic, 2004). However, there is a growing body of literature on the benefits of isolating active proteins from inclusion body in decreased number of operational steps and in more cost-effective approaches (for excellent reviews see (Singhvi et al., 2020, Krauss et al., 2017) and references therein).

And last but not least, the risk of protein degradation by cytosolic proteases, posed in general by the cytosolic expression of recombinant proteins, should be taken into account while applying *E. coli* as a bacterial host (Baneyx and Mujacic, 2004).

Protein secretion provides solutions to overcome difficulties associated with the cytosolic expression of the recombinant proteins. Due to the diluted total protein content outside of the cells, secretion leads to higher solubility and stability of the target protein. Also, it results in a reduced risk of protein degradation by cytosolic proteases (Kleiner-Grote et al., 2018). Notably, secretion of the recombinant proteins could be coupled with a significant reduction in the production costs owing to simplified downstream processes (Azam et al., 2016), especially because of no need of cell lysis and complicated refolding processes. This can be seen in the case of long historical use of Bacillus species for producing naturally secreted enzymes such as lipases, proteases and amylases in high production yields- even more than 20 g/l of supernatant. However, one should bear in mind that in contrast to the above-mentioned examples, the yields of secretion of heterologous proteins by Bacillus species are often very low (Harwood, 1992; Pohl and Harwood, 2010; Freudl, 2018).

All in all, at the outset, we acknowledge that the potential use of many secretion systems is still far away from being a potential platform for recombinant protein production in industrial scales mostly because of low secretion titers of those systems. In industrial scales, reaching high production yields for secreted proteins requires large volume of bacterial culture and concentrating culture supernatant which indeed make the production not cost-efficient. Therefore, improving secretion systems, especially in regard of the secretion titer, will require continued studies. The evidence reported thus far for production of secretory proteins by cell lines supports the idea that putting much more efforts on secretion systems is worth and could address the current drawbacks of using them. Supporting this view, after three decades of research on protein production by producter cell lines, the initial production yield of 100 mg/l monoclonal antibody has gained a yield of 10 g/l, which was based on improving basal media, using high cell density cultures, development of feed concentration (Kunert and Reinhart, 2016).

This review aims to provide a better understanding of biotechnological applications of type 1 secretion system (T1SS). Although reviewing different secretion systems is not the scope of this review, a brief overview of those systems that have been already used in biotechnological applications is provided to gain a better view of protein secretion. For more details on those secretion systems, the reader is directed to comprehensive reviews written recently on secretion systems (Costa et al., 2015; Burdette et al., 2018; Anné et al., 2016).

2. Bacterial secretion systems

In the case of Gram-positive bacterial cells, the secreted protein should be translocated across only one membrane, the plasma membrane, which is surrounded by a thick peptidoglycan layer (Green and Mecsas, 2016), and in the case of Gram-negative bacteria it should be transported across two membranes, the cytoplasmic inner membrane (IM) and the outer membrane (OM). A peptidoglycan (murein) layer is also located in between of two membranes (Kleiner-Grote et al., 2018).

2.1. Gram-positive bacterial cell secretion systems

Different Gram-positive bacteria such as *Lactococcus lactis*, *Streptomyces lividans*, *Corynebacterium glutamicum*, and *Bacillus subtilis* have been evaluated as hosts for the secretion of recombinant proteins (Anné et al., 2016; Freudl, 2018). Bacillus species bear great potential for the industrial-scale production of different enzymes including lipases, proteases, amylases in a product yield of up to 20 g/l, but only for their naturally secreted proteins or proteins from close relative hosts (Freudl, 2018). In many cases, not only insufficient yields but also the risk of extracellular proteolytic degradation impose limitations for applying Gram-positive bacteria as a secreting host. Furthermore, fewer genetic tools, in comparison to Gram-negative bacteria, especially *E. coli*, are available for handling Gram-positive bacteria (Anné et al., 2016).

At least six different secretion systems have so far been identified in Gram-positive bacteria including (i) Sec (secretion translocation system)-pathway, (ii) twin arginine transporter (Tat)-pathway, (iii) fimbrillin-protein exporter (FPE), (iv) flagella export apparatus (FEA), (v) Holins, and (vi) ESAT-6/WXG100 secretion system (WSS) (Forster and Marquis, 2012; Anné et al., 2016). Although protein translocation of Gram-positive bacteria is not the focus of this review, a brief overview of two translocation systems, the Sec- and Tat- pathway, that have been commonly applied for biotechnological application is provided. The Sec-pathway and the Tat-pathway are found in both, Gram-positive and -negative bacteria (Green and Mecsas, 2016).

Sec-pathway: this pathway exports its cognate proteins in an unfolded state either co-translationally or post-translationally. In the co-translational export pathway, precursor proteins are recognized earlier during their synthesis at ribosomes by a signal recognition particle (SRP). Interacting SRP and its dedicated membrane receptor, called FtsY receptor, brings the ribosome-nascent chain complex close to the membrane. Thereafter, the complex is redirected to the transmembrane SecYEG channel, and the remaining part of the nascent chain is synthesized and exported through the channel (Fig. 1-A) (Peng et al., 2019).

In the post-translational mode, the fully synthesized precursor protein maintains in an unfolded state by interacting with cytosolic proteins such as chaperon proteins, mainly SecB (Fig. 1-B). The precursor protein

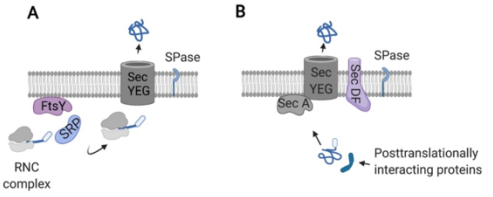


Fig. 1. Simplified protein export pathways by the Sec system. A) The co-translational export mode of the Sec pathway: In this route, the ribosome plus nascent chain (RNC) complex is recognized by the SRP. Thereafter, the RNC-SRP complex is directed to the SRP-receptor, FtsY, in the membrane, and then while the ribosome is close to the SecYEG channel, the RNC complex is exported and the secretion signal is cleaved during or after transport. B) Post-translational export mode of the Sec pathway: Sec-dependent substrates stay associated with post-translationally interacting proteins such as common chaperone, SecB, and handed over to the ATPase SecA. Thereafter, this complex is redirected to the SecYEG channel while the SecDF provides a pulling force. The secretion signal is cleaved during or after transport. All figures used in this review were developed by Biorender (BioRender.com/).

is then translocated through the SecYEG channel with the motor protein, SecA powered by ATP hydrolysis (Peng et al., 2019; Freudl, 2018). Due to the limited translocation capacity of the Sec-pathway, overexpression of the target proteins leads often to the formation of inclusion bodies and accordingly reduced yield of secreted protein. To avoid this problem, overexpression of translocation-involved proteins was reported as a solution (Kleiner-Grote et al., 2018). Furthermore, the Sec-pathway has limitation to export proteins that fold rapidly in the cytosol (Walker et al., 2015).

Tat-pathway: This pathway is the second major secretion route used by many bacteria. It exports proteins across the plasma membrane using the energy derived from the proton motive force (PMF). The exported proteins are completely folded and in their native oligomeric state and along with their cofactor. The Tat-pathway is capable of exporting a wide range of substrates including relatively complex substrates.

The source of obtaining energy, PMF, as well as the secretion capacity of the Tat-pathway, which is lower than the Sec-pathway due to the lower number of Tat translocons in the membrane, are limiting factors for Tat-pathway dependent secretion (Fig. 2) (Kleiner-Grote et al., 2018; Anné et al., 2016; Walker et al., 2015).

2.2. Gram-negative bacterial cell secretion systems

Extensive knowledge of the genetics of Gram-negative bacteria, as well as the availability of many genetic tools for handling these bacteria made them in principle more attractive for biotechnological purposes (Anné et al., 2016). The presence of lipopolysaccharides (LPS), however, can cause an immune response and need to be removed (Mack et al., 2014).

The secretion systems of Gram-negative bacteria are classified into two different classes based on the fact that the substrates either bypass or enter the periplasmic space and are consequently named one-step and two-steps secretion systems, respectively. Among the seven different secretion systems identified so far, six have been applied for the secretion of heterologous proteins listed as T1SS, T2SS, T3SS, T4SS, T5SS, and T8SS (Burdette et al., 2018). Table 1 summarizes some features of these systems.

2.2.1. Two-step secretion systems

For these two-step systems, T2SS and T5SS, substrates are exported first into the periplasmic space by using either the Sec- or/and Tat-pathway, and thereafter across the outer-membrane through a dedicated channel (Costa et al., 2015). Cognate substrates of the two-steps systems have two fates, either releasing to the supernatant or displaying on the bacterial surface. Please note that the T2SS is the only export route for exported Tat substrates (Nivaskumar and Francetic, 2014).

In addition to two mentioned secretion systems, some microorganism, in particular pathogenic bacteria, have developed another way of secretion which also operates in two steps but only with two involved proteins: a transporter protein and an exoprotein. Noted that each transporter is dedicated to a specific exoprotein. The way of secretion is known as two-partner-secretion (TPS) and can be found in different bacteria such as *Bordetella pertussis* and *Serratia marcescens* for secretion of the filamentous haemagglutinin (FHA) and Ca^{2+}-independent hemolysin, respectively (Jacob-Dubuisson et al., 2001). It should be mentioned that FHA has been used as a heterologous antigen delivery system to induce the mucosal immune response (Renauld-Mongenie et al., 1996).

The T2SS has a complicated architecture consisting of 15 components, the structural and biochemical features of some of these components are still unknown (Fig. 3) (Green and Mecsas, 2016; Costa et al., 2015). The functionalities of T2SSs in the related bacteria are nutrient acquisition by secreting hydrolytic enzymes, toxicity, mobility, respiration, and biofilm formation. Bacterial species living in extreme environments secrete stable proteins using T2SS (Nivaskumar and Francetic,

Table 1
Comparison of the applied bacterial secretion systems for the biotechnological applications.

	Secretion signal	Cleavage of the SS	Release location	Structural complexity
One-step system				
T1SS	C-terminal	No	Medium	Low
T3SS	N-terminal	No	Into a target cell	High
T4SS	Diverse	Diverse	Into a target cell (Diverse)	High
Two-step system				
T2SS	N-terminal	Yes (most cases)	Periplasm and then medium	High
T5SS	N-terminal	Partial	Periplasm and then medium	Low
T8SS	N-terminal	No	Anchored to the outer membrane	Low

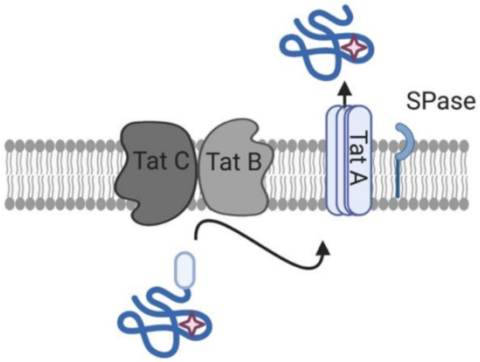


Fig. 2. Protein export by the Tat system. Tat substrates harbor an N-terminal secretion signal and fold beforehand in the cytosol and if required, a cofactor is also added prior to secretion. The secretion signal is cleaved off during transport through the membrane by the SPase. The TatB-TatC complex recognizes the secretion signal. The TatA complex translocases the substrate through the membrane.

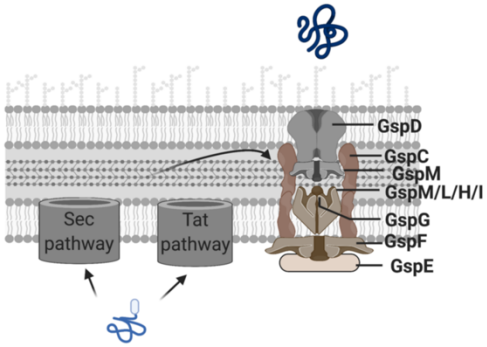


Fig. 3. Protein export via T2SS. In the first step, proteins are exported to the periplasmic space either by the Sec pathway or the Tat pathway. In the second step, the substrates enter the T2SS apparatus and are exported across the outer membrane.

2014).

Remarkably, T2SS substrates play important roles in pathogenicity of the harboring bacterial cells. This have been seen in many cases, including: the protease pseudolysin, cholera toxin, and heat-labile enterotoxin secreted by *Pseudomonas aeruginosa*, *Vibrio cholerae, and enterotoxigenic E. coli*, respectively (Yang et al., 2015; Reichow et al., 2010).

The T5SS exports proteins to the extracellular media preceded by secretion through the IM via the Sec-pathway in which the proteins' N-terminal secretion signal is cleaved off after passing the membrane. The substrates of this system have a unique property that, unlike other substrates of other secretion systems, the substrate and dedicated β-barrel domain are fused, subsequently the β-barrel inserts into the outer membrane and forms a channel, and thereafter substrate can be exported through it (Fig. 4). The passenger domain remains attached to the translocation domain or is cleaved off after secretion (Van Ulsen et al., 2014; Gawarzewski et al., 2013).

The surface display property of this system offers advantages for applying it in live vaccine development, whole-cell biocatalysis, bio-absorbent, biosensor development, epitope mapping, and protein library screening. The biotechnological application of this system has received less attention because of its limitations of heterologous substrates which can be included and the low secretion yields (Van Ulsen et al., 2014).

2.2.2. One-step secretion systems

One-step secretion systems (T1SS, T3SS, T4SS, and T6SS) secrete substrates across the bacterial envelope in one step into the extracellular space or into a target cell.

T1SS and T3SS translocate proteins in a partially or completely unfolded state from the cytosol to the extracellular space (T1SS) or cytosol of the host (T3SS). Subsequently, the secreted protein can fold properly in the supernatant even if disulfide bridges are required (Schwarz et al., 2012a; Burdette et al., 2018).

The structure of the T3SS system, which is highly conserved among Gram-negative bacteria, such as *Salmonella*, *Shigella*, pathogenic *E. coli*, and *Pseudomonas*, consists of a multi-ring needle complex, inner membrane complex, and a set of cytosolic proteins that provide the energy of secretion (Costa et al., 2015)(Fig. 5- A). For a high-resolution structure of the T3SS nanomachine, injectosome, of *Salmonella* see the recently

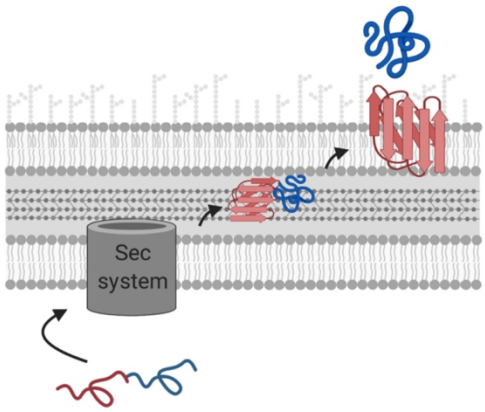


Fig. 4. Protein export via T5SS. The translocation domain and the substrate remain fused during secretion across the inner membrane. The translocation domain inserts in the outer membrane, then the substrate is either cleaved off autonomously or remains attached.

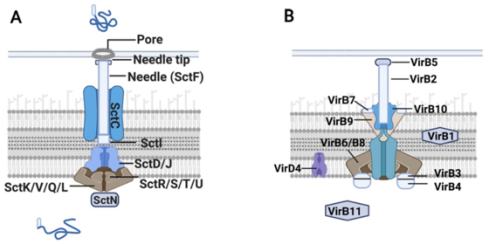


Fig. 5. Protein export via A) T3SS: the transport through T3SS occurs in one step from bacterial cytosol to eukaryotic host cytoplasm. The secretion sequence is located at the N-terminus of the substrate and is not cleaved off during or after secretion. B) T4SS has ability to translocate DNA and protein into target cells. This system contains three ATPases (VirB11, VirD4, and VirB4). The pilus of this system consists of VirB2 and VirB5 which extends to the extracellular space.

published cryo-electron tomography structure (Hu et al., 2017).

T3SS shows a unique delivery behavior that makes it attractive for many biomedical applications. The substrates of T3SS are exported directly into a target cell by passing through the multimeric needle-like structure of this secretion system (Bai et al., 2018). Induction of T3SS initiated by cell-cell contact between the host and bacterial cell results in protein delivery into the host-cytosol. However, low calcium ion concentration of the environment (*Yersinia*) (Heesemann et al., 1986) or altering pH (Salmonella SPI-2) (Yu et al., 2010) promotes the secretion of T3SS effectors into the surrounding environment. The mechanism of secretion of this system has been investigated in great details, but will not be covered in this review. Rather we refer to other excellent reviews, for example (Bai et al., 2018; Abrusci et al., 2014).

T3SS acts as a virulence factor in many pathogenic bacteria causing different plant or animal infections. Therefore, making use of this secretion system in biotechnological purposes, requires an attenuation of the bacterial pathogen. This can be achieved through extensive genetic alterations. Some of the already recombinant attenuated strains for using this system include: *Salmonella typhimurium*, *Yersinia species*, and *P. aeruginosa* (Bai et al., 2018). There are many examples of applying T3SS for a delivery of antigenic peptides and proteins into diverge target cells for different aims, e.g. vaccination, induction of immune response, genome editing, cancer immunotherapy, etc. (Bai et al., 2018). Although extensive research has been carried out on T3SS, still some of the limitations of using this system in biotechnology should be addressed such as low translocation efficiencies of some foreign proteins and toxicity of the delivery system (Burdette et al., 2018).

The ability to translocate DNA, in addition to proteins and other macromolecules, makes the superfamily of T4SS unique among other bacterial secretion systems (Fig. 5-B). T4SSs play an important role in conjugation of DNA and spreading of DNA-plasmids between bacteria. Although the conjugation is promoted by cell-to-cell contact, but there are also examples representing that a few T4SSs import or export DNA independent of such interactions which causes the release of macromolecule into the extracellular medium. The functionality of these bacterial systems mediate pathogenicity of some microorganisms such as *Helicobacter pylori* and *Legionella pneumophila* (Costa et al., 2021; Costa et al., 2015).

Interestingly, while T4SSs commonly translocate their dedicated substrate in one step from cytosol, the secretion of pertussis toxin across T4SS present a two-step mechanism (Verma and Burns, 2007). Research in recent years has provided a deeper insight into the functionality, mechanism, and structure of T4SS and comprehensive reviews have been published highlighting the recent progresses on this system (Costa et al., 2021; Grohmann et al., 2018).

3. Type 1 secretion system

Type 1 secretion systems are widespread in Gram-negative bacterial cells and secrete a wide range of substrates with different size and functionality that have different roles in bacteria ranging from nutrient acquisition to bacterial pathogenicity (Kanonenberg et al., 2018). In terms of structure, T1SS has a relatively simple architecture compared to other secretion systems. T1SS spans both membranes and periplasmic space with only three compartments listed as follows: an ATP-binding cassette (ABC) transporter, a membrane fusion protein (MFP), and an outer membrane protein (OMP) (Fig. 6) (Holland et al., 2016). These proteins present a multimeric structure and all together form a transmembrane tunnel spanning both inner and outer membranes for transporting an unfolded substrate from the cytosol to the extracellular space. ABC transporters belong to a superfamily of membrane proteins found in all domains of life with the ability to hydrolyze ATP providing the energy for transport; T1SS is assumed to involve the PMF as the energy source (Bumba et al., 2016; Koronakis et al., 1991). The secreted proteins are transported by electrophoretic movement of negatively charged RTX polypeptide along the gradient of negative membrane potential (Bumba et al., 2016).

3.1. Protein charge and secretion of T1SS

Interestingly, all known natural substrates of ABC transporter have very acidic isoelectric points (pI) taking negative charge at neutral pH (Delepelaire, 2004). Byun et al. also reported that proteins that are highly negatively charged in overall are secreted efficiently (Byun et al., 2017). One obvious explanation might be the PMF, which is the electrochemical gradient of protons and charges across the phospholipid bilayer membranes. The inside negative potential of the membrane might favor the secretion of negatively charged proteins but positively charged proteins might receive an extensive electrostatic force from the membrane potential (Rempel et al., 2019). There are two proposed models for ABC transporter's mechanism. One is the Brownian ratchet model, primarily powered by the calcium ion-dependent folding of the β-helix of the RTX domain (Bumba et al., 2016; Kanonenberg et al., 2018; Hepp and Maier, 2017). In the Brownian ratchet mechanism, the free Brownian motion of the polypeptide substrate is biased toward secretion by a mechanism that prevents backtracking of the substrate, which is the Ca^{2+}-dependent folding of the C-terminal secretion signal (Bumba et al., 2016). In contrast to the Brownian ratchet mechanism, in HlyA, after two hours of secretion, the secretion rate is independent on the extracellular Ca^{2+} concentration. Therefore, the Brownian ratchet mechanism is not operational for HlyA. However, Ca^{2+} is important for proper folding and stability of HlyA (Lenders et al., 2016).

The transporter's input energy, which is ATP for the T1SS, is used to reset the transporter into the initial state whenever the transporter gets stuck. It is noteworthy, however, that there was some evidence that pointed against the concept of membrane potential taking a role in the ABC transporter's secretory mechanism, as the membrane potential was not affecting the rate of secretion in some cases (Kanonenberg et al., 2018). The other one is the "power stroke" model, which forces the substrate to move from one side of the membrane to the other via a series of conformational changes by ATP (Rempel et al., 2019). In this model, ATP binding, which initiates the inward-facing to outward-facing transition, is stoichiometrically coupled to substrate translocation (Stefan et al., 2020). It is still elusive which model is correct.

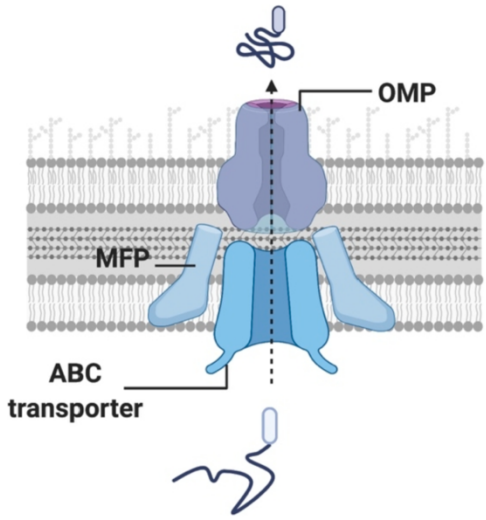


Fig. 6. Protein export via T1SS. The apparatus of T1SS consists of the ABC transporter plus MFP in the inner membrane and an OMP. The export is in one step from the cytosol to the extracellular space. The secretion signal is at the extreme C-terminus (highlighted as an ellipse in the cartoon) of the substrate and is not cleaved off during or after the secretion. Note that the oligomeric state of the MFP (drawn here as a dimer) is unknown, but likely a hexamer although experimental evidence points to a trimer (Balakrishnan et al., 2001).

3.2. Characteristics of T1SS substrates

The impressive variety of T1SS substrates results not only from their role in different bacterial functions, from nutrient acquisition to the bacterial pathogenicity (Kanonenberg et al., 2018), but also from their diverse sizes, for instances the hemophore HasA from *S. marcescens* (Letoffe et al., 1994) with a size of only 19 kDa to the adhesion factor LapA from *Pseudomonas fluorescence* with a size of 520 kDa (Boyd et al., 2014).

Thomas et al. classified the native substrates of T1SS into 6 classes based on their functionality as follows: Heme-binding, RTX toxins, proteases, lipases, adhesions, and S-layer proteins (Thomas et al., 2014b). Notwithstanding such differences, T1SS substrates -exceptions may occur- show some properties in common as follows:

1) Substrates are transported in a single-step secretion strategy from the cytosol to the medium (Thomas et al., 2014a; Thomas et al., 2014b), however, recent research on the giant RTX adhesin LapA of *P. fluorescens* Pf0-1 introduced a distinct subgroup of T1SS. Secretion of LapA occurs in a two-step secretion strategy since LapA first remains on the cell surface via a cognate outer membrane protein and can be released from cell surface at a later stage (Smith et al., 2018).
2) The next feature of almost all T1SS substrates is related to the secretion signal, which is detected by the dedicated ABC transporter. The majority of T1SS-dependent substrates harbor an uncleavable C-terminal secretion signal, which is not cleaved while or after secretion. One of the stated exception for having C-terminal secretion signal is the bacteriocin colicin V with a secretion signal at the N-terminal end, which is cleaved during secretion although the secretion is dependent on a T1SS (Holland et al., 2016).
3) T1SS substrates are secreted in an unfolded state to pass through a narrow tunnel. One has to note that HasA protein requires a cytoplasmic chaperone, SecB, for its efficient secretion (Holland et al., 2016).
4) Another common property of T1SS substrates is the presence of a repetitive Glycine- and Aspartate-rich sequence, also known as either nonapeptide sequence or GG repeats, upstream of the secretion signal having the GGxGxDxUx consensus sequence in which x can be any amino acid and U refers to a large hydrophobic amino acid. The

presence of this sequence brings about the name of a group of T1SS substrates called RTX (repeats in toxins) family, a superfamily of proteins produced by Gram-negative bacteria. The RTX is a site for the binding of Ca^{2+} ions triggering the folding of a substrate in the extracellular space where the Ca^{2+} concentration is much higher than the cytoplasm. The number of RTX repeats can be less than eight to more than forty repeats in different proteins and exceptionally none for HasA protein and the bacteriocin colcinV (Holland et al., 2016; Kanonenberg et al., 2018; Linhartová et al., 2010).

A subfamily of large repetitive RTX proteins are known as multifunctional-autoprocessing RTX (MARTX family) toxins that includes extremely large polypeptides with size of 5000 to 9000 kDa. The members of MARTX family are encoded in chromosomal islands in human pathogens such as *Vibrio cholerae*. It has been predicted that most of MARTX members are secreted by a T1SS that consists of a heterodimeric ATPase complex. For instance the MARTX from *V. cholerae* requires the ATP hydrolysis by two ABC transporters, RtxB and RtxE (Boardman and Fullner Satchell, 2004; Satchell, 2011).

3.3. Molecular model of the secretion through T1SS

Since the mechanism of secretion has been studied in more detail for the Hly system rather than other type 1 secretion systems, the model of secretion through Hly system is discussed herein. Upon interaction of HlyA with its cognate inner membrane complex, consisting of HlyB and HlyD, the recruitment of the outer membrane protein, TolC, occurs. Consequently, all three membrane proteins form a channel tunnel redirecting HlyA, C-terminus first, in an unfolded state to the exterior. After completing one cycle of secretion, TolC leaves the complex -yet the inner membrane complex remains assembled waiting for the next cycle of secretion (Holland et al., 2016). First, Letoffe et al. (1996) provided evidence suggesting that the assembly of T1SS components for the secretion of T1SS substrates, PrtC and HasA from *Erwinia chrysanthemi*, occurs within an ordered reaction, and overall secretion process is promoted by the interaction of the substrates and their cognate ABC transporters (Letoffe et al., 1996). Later on, Thanabalu et al. proved that HlyB forms a complex with HlyD then assembles with TolC upon substrate binding (Thanabalu et al., 1998).

To investigate the direction and rate of secretion through HlyA secretion system, a stalled system was used. In the stalled system, the C-terminal end of HlyA can be detected on the cell surface by using an HlyA specific polyclonal antibody and a Cy3-fluorophore labeled secondary antibody confirming this concept that the secretion through T1SS is directional and the C-terminus exits first from the T1SS channel tunnel (Lenders et al., 2015). To determine the secretion rate, the number of active HlyA secretion system in *E. coli* was estimated by the application of stalled T1SS which resulted in an average number of 4554 ± 1616 active transporter per cell (Lenders et al., 2016). The rate of secretion through T1SS is approximately 16 amino acids per second per transporter. Pieces of evidence also were provided by Lenders et al. showing that the secretion rate in the case of HlyA secretion system is independent of the size of the substrate, the absolute number of GG repeats, as well as Ca^{2+} concentration of medium (Lenders et al., 2016).

3.4. Folding rate of protein is an important factor of secretion by T1SS

One point which should be taken into consideration while planning for heterologous protein secretion through the T1SS is that, although, the size of heterologous protein has no impact on the secretion efficiency, the folding rate of the fused protein is the determinant index for successful secretion. Bakkes et al. fused HlyA1 to three versions of maltose-binding protein (MalE) with different folding rates, resulting from various mutations. As their data showed, a significant different secretion level between the wild-type and mutant MalEs was observed. These differences were related to the folding rate. According to Bakkes et al., a MalE protein with the lowest folding rate showed the strongest secretion level compared to other MalEs (Bakkes et al., 2010). Another example of how the folding rate affects secretion level was illustrated for the intestinal fatty acid-binding protein (IFABP). The IFABP containing a mutation causing slow-folding kinetics (G121V mutant) could be secreted through T1SS when it was fused to HlyA1, although the wild-type of this protein could not be secreted (Schwarz et al., 2012b; Schwarz et al., 2012a). Overall, these case studies support the view that the fast-folding rate of a substrate in the cytosol is a limiting factor for secretion through T1SS since a fast-folded protein could block the translocation channel and could not be secreted through it. Although, it is apparent from the performed experiments that the cytosolic folding rate of a protein dictates the possibility of secretion but lack of either a laboratory assay or computational approach for determining the folding rate of a protein does not allow any priori prediction of successful secretion.

For those proteins for which a slow-folding version is not available, a directed evolution approach could be an option for identifying mutations that decrease the folding rate. However, these methods are time-consuming and costly. One other suggested solution is using a destabilizing domain in front of the target protein sequence. This solution was then assayed for IFABP protein using a ddFKBP, an eukaryotic destabilizer in front of IFABP gene, nevertheless, no significant correlation was found between applying the destabilizer and the secretion efficiency of T1SS in this study (Schwarz et al., 2012a; Schwarz et al., 2012b).

3.5. Biotechnological application of T1SS

The exploitation of a T1SS for biotechnological applications has been taken into consideration, because of the ability of the T1SS to secrete a wide variety of proteins as well as easy plasmid-based expression of the T1SS components (Burdette et al., 2018). The main T1SS systems that have been developed and applied frequently for biotechnological purposes are discussed below. The biotechnological application of T1SS has been focused mainly on two general areas including heterologous protein secretion and vaccine development (Kanonenberg et al., 2018) but it will extend more to useful biologics (Nguyen et al., 2021) including hormones, and growth factors (Chung et al., 2009). Biocatalysts using enzymes (Jaeger et al., 1999) or whole cell biocatalysts (Son et al., 2012) and biosensors using receptors or adhesion molecules (Barlag and Hensel, 2015) will also be made possible using the proteins produced by T1SS.

3.5.1. Main T1SS machineries for the biotechnological applications

Main T1SS machineries that have been applied so far to achieve secretion of recombinant proteins are the HlyA system from *E. coli* (Khosa et al., 2018; Schwarz et al., 2012a; Schwarz et al., 2012b), the TliDEF system from *P. fluorescens* (Ahn et al., 1999; Park et al., 2020; Park et al., 2012), and the PrtDEF system of *E. chrysanthemi* (Ghigo and Wandersman, 1994; Palacios et al., 2001).

3.5.1.1. TliDEF system from P. fluorescens. TliDEF was identified by Ahn in 1999 (Ahn et al., 1999) and consists of three proteins: TliD (ABC transporter), TliE (MFP), and TliF (outer membrane factor) (Park et al., 2020). Natively, the TliDEF system secretes a thermostable lipase, TliA, in a temperature range from 20 °C to 25 °C (Ahn et al., 1999), optimal for cell growth of *P. fluorescens* (Park et al., 2012). A study by Eom reported that the temperature-dependent secretion of *P. fluorescens* at 30 °C recovered by a single mutation in one of the transmembrane domain of the TliD (Eom et al., 2016b). *P. fluorescens* is classified as a biosafety level 1 bacteria and several reliable platforms were developed for expressing recombinant proteins in this microorganism (Retallack et al., 2012). Another advantageous of using this bacteria is that unlike *E. coli*, *P. fluorescens* does not produce acetate during fermentation (Retallack et al., 2012).

Concerning the recognition of the passenger protein by the ABC protein, a C-terminal domain of TliA composed of roughly 103 residues harboring 6 RTX repeats, known as LARD3, was reported to be sufficient for an efficient secretion not only in *P. fluorescens* but also in *E. coli* (Park et al., 2012). In another case study, the C-terminal region of another *Pseudomonas* lipase, MIS38 lipase (PML) was successfully used for redirecting the *E. coli* alkaline phosphatase by a heterologous T1SS in *E. coli* even when the C-terminal domain was lacking all its 12 RTX repeats (Angkawidjaja et al., 2006).

In 2015, Ryu et al. established a new vector for *P. fluorescens* encoding genes for the ABC transporter, which was applied successfully for the secretion of GFP and alkaline phosphatase in this bacterial host (Ryu et al., 2015). The TliDEF system is used as an alternative for producing bacteriotoxic proteins such as bacterial phospholipase A1, which was previously verified to hamper the expression in other bacterial cells (Park et al., 2020). These examples verify efforts to achieve improvements of the application of the TliDEF system for recombinant protein production.

3.5.1.2. PrtDEF system of Erwinia chrysanthemi. *E. chrysanthemi* natively secretes a degradative metalloprotease by a T1SS, named PrtDEF system. It consists of two associated inner membrane proteins, PrtD and PrtE, and one outer membrane protein, PrtF. Studies on the proteases from *E. chrysanthemi* show that the secretion signal localizes to the C-terminus and is located within the last 50 amino acids, but in the case of PrtG protease the last 29 amino acids are recognized as the secretion signal for efficient secretion. In addition to the secretion signal, the presence of the RTX domain is critical for the secretion. The PrtDEF system was employed successfully for the secretion of a subset of hybrid eukaryotic proteins but interestingly just in the case of proteins lacking disulfide bonds (Palacios et al., 2001; Ghigo and Wandersman, 1994).

3.5.1.3. HlyA secretion system of E. coli. The most prominent T1SS of uropathogenic *E. coli* is the HlyA secretion system, which was the first described in 1979 (Noegel et al., 1979) and is likely the most studied T1SS. The apparatus of this system consists of HlyB (ABC transporter), HlyD (MFP), and TolC (outer membrane protein). The substrate of this system is HlyA, a member of the RTX family with a size of 110 kDa (Spitz et al., 2019) and with the ability to lyse for example erythrocytes among other cell types (Noegel et al., 1979). The encoding genes of HlyA secretion system, except for TolC, are located in an operon called the *hly* operon where the fourth gene, encoding for HlyC, is also located. HlyC is not essential for the secretion process, but required for activation of HlyA by acylation of two internal lysine residues of the unfolded HlyA prior to secretion (Khosa et al., 2018).

HlyB, the ABC transporter, resides in a dimeric state in the inner membrane. Data on the topology of HlyB, provided mostly by co-construction studies, shows the presence of a transmembrane domain (TMD) composed of six hydrophobic transmembrane helix (TMHs), a hydrophilic nucleotide-binding domain (NBD), plus a cytosolic N-terminal domain, known as C39-peptidase-like domain (CLD) per monomer (Kanonenberg et al., 2019; Thanabalu et al., 1998; Lecher et al., 2012). Concerning the latter one, unlike other canonical ABC transporters, HlyB contains a CLD domain that shows about 40% homology to C39 peptidase (a cysteine protease), despite no known proteolytic activity (Lecher et al., 2012). Lecher et al. observed a specific interaction between the CLD domain and unfolded substrate notably not within its secretion signal. This observation led to the speculation that CLD may play a role in the positioning of HlyA into the secretion complex (Lecher et al., 2012). The NBD not only plays an important role in providing energy of the transport by hydrolyzing ATP but it was also proposed that NBD takes part in one of the early steps of the secretion process by recognition of the substrate, which occurs upon a specific interaction between NBD and C-terminus of HlyA (Benabdelhak et al., 2003).

HlyD together with HlyB forms the IM complex of HlyA secretion system. Although a trimeric state for HlyD was suggested by the Koronakis laboratory in 2001 (Balakrishnan et al., 2001), a hexameric state for HlyD in the membrane was suggested later based on several lines of evidence from other MFP analogues (Holland et al., 2016). However, the oligomeric state of this protein is still an open question.

HlyD has a short cytoplasmic region by a size of 60 residues that initializes the assembly of secretion complex upon interaction with HlyA (Pimenta et al., 1999; Schtilein et al., 1992). By deletion of this cytoplasmic region, however, HlyD can still be inserted into the inner membrane (Balakrishnan et al., 2001). A single α-helical domain spanning the inner membrane, an α-helix-periplasmic domain, plus a large β-strand C-terminal domain comprise the other part for the MFP. Based on data from beta-lactamase fusion as well as cellular fractionation analysis, it revealed that HlyD has only one single transmembrane domain (Holland et al., 2016).

The outer membrane component of the HlyA secretion system is TolC that also evolves in tripartite drug efflux complexes of *E. coli* including AcrAB-TolC (Du et al., 2014) and MacAB-TolC (Fitzpatrick et al., 2017). The crystal structure of TolC represents a homotrimer state that forms a channel with a length of 140 Å residing in the periplasmic space. Periplasmic part of each protomer of TolC protein consists of two internal repeats of α-hairpin forming a very long α-barrel and then all three protomers form a 12 β-strands in the transmembrane part, and thereby the exit route to complete the secretion (Koronakis et al., 1997; Jo et al., 2019; Anderson et al., 2000).

A 23-kDa C-terminus fragment of HlyA, known as either HlyA1 or HlyAc can be employed for secretion purposes. This observation by Nicaud et al., 1986 was a turning point stressing the possibility of T1SS application for the secretion of heterologous proteins (Nicaud et al., 1986). In 1987, Mackman reported that the *E. coli* HlyA secretion system can secrete OmpF protein fused to HlyA1 highlighting HlyA T1SS as an efficient system for protein secretion applications (Mackman et al., 1987). Since then, this system has been used for secreting not only different prokaryotic proteins but also proteins from eukaryotic origins (Schwarz et al., 2012b).

3.5.2. Heterologous protein secretion

T1SS can be used as an interesting platform for the secretion of a wide variety of proteins with different importance for industrial and pharmaceutical applications. Up to now, several studies have reported utilization of T1SS for expressing enzymes with industrial importance such as cutinases (Su et al., 2012), lipases (Eom et al., 2014a), metalloproteases (Eom et al., 2014b), and alkaline phosphatases (Angkawidjaja et al., 2006).

Lipases are natural substrates of T1SS with high importance in industrial applications. Lipase family I.3 is an important class of lipolytic enzymes secreted natively through T1SS of so far only two genera of bacteria, *Pseudomonas* and *Serratia*. Some of lipase family I.3 have been secreted successfully through a heterologous T1SS. For instance, the lipase from *Pseudomonas* sp. MIS38 can be secreted through reconstituted T1SS from *S. marcescens* Sr41 in *E. coli* (Kwon et al., 2002). A number of case studies on secretion of different proteins through T1SS are listed in the Table 2.

Although, comparison of the secretion level of heterologous proteins secreted through T1SS with of some other secretion system, e.g. 20 g/l for Bacillus species (Pohl and Harwood, 2010) confirms a low yield of secretion for T1SS, one should stress that an efficient high-titer expression of peptides and small proteins, more than 2 g/l, was reported only recently by using T1SS in a cost-efficient approach (Nguyen et al., 2021). This suggests that further optimization might be possible.

3.5.3. The potential use of T1SS for vaccine development

Secretion and screening of antibodies: There are relatively few published studies on the usage of T1SS for antibody secretion, although the early example on this application dates back to 1990s study of Holland et al. for the secretion of both antibody domains, Fv and Fab

Table 2
A list of heterologous proteins secreted by T1SS, n.r.: not reported.

Substrate	Secretion signal	Host cell	Secretion titer	Application	Ref.
Therapeutic application					
Nanobodies	$HlyA_s$	*E. coli*	0.3–1.3 mg/l	Screening and easy purification of Nbs	(Ruano-Gallego et al., 2019)
Prostate specific antigen	$HlyA_s$	*Salmonella enterica serovar*	n.r.	Prevention of tumor growth in mice	(Fensterle et al., 2008)
HlyA	$HlyA_s$	*Salmonella typhi*	1–4 mg/l	Live bacterial vaccine	(Gentschev et al., 2004)
C-Raf	$HlyA_s$	*S. enterica serovar*	2–3 mg/l	Cancer therapy in mice	(Gentschev et al., 2005)
Immunoglobulin VHH	$HlyA_s$	*E. coli*	~ 2 mg/l	Correct protein folding and dimerization	(Fraile et al., 2004)
Human interleukin −6	$HlyA_s$	*Salmonella typhimurium*	n.r.	Mice immunization against bacterial LPS	(Li et al., 2003)
Helicobacter pylori antigen	$HlyA_s$	*E. coli*	~ 1 mg/l	Generation and detection of immunogenic antigens	(Spreng et al., 2000a)
Heterologous microbial antigens	$HlyA_s$	*E. coli* and *Salmonella aroA*	n.r.	Antigen-delivery systems	(Gentschev et al., 1996)
PagC, of *S. typhimurium*	$HlyA_s$	*E. coli*	0.75 mg/l	Generation and identification of Salmonella specific antibodies	(Mollenkopf et al., 1996)
Shiga toxin B-subunit	$HlyA_s$	*Salmonella typhimurium*	n.r.	antigen-specific immune responses in mice	(Su et al., 1992)
Industrial application					
Cutinase	$HlyA_s$	*E. coli*	1.5 mg/ml	2.5 folds higher than T2SS	(Su et al., 2012)
Alkaline phosphatase	C-terminal of PML	*E. coli* carrying the *S. marcescens* lipase system	10 mg/l	Secretion through a heterologous T1SS	(Angkawidjaja et al., 2006)
Serratia proteamaculans Metalloprotease	Cognate signal	*P. fluorescens*	789 mg/l	Secretion through a heterologous T1SS	(Eom et al., 2014b)
Cyclodextrin glucanotransferase	$HlyA_s$	*E. coli*	0.58 mg/l	Optimizing secretion	(Low et al., 2011)
endo-β-1, 4 mannanase	From lipase (TliA)	*E. coli*	4.65 mg/l	Secretion through a heterologous T1SS	(Eom et al., 2016a)
Different proteins/peptides of interest of interest					
Teriparatide	$HlyA_s$	*E. coli*	>2 g/l	Cost-efficient industrial production of peptides	(Nguyen et al., 2021)
Intestinal fatty acid binding protein (IFABP)	$HlyA_s$	*E. coli*	1 mg/l	Secretion of eukaryotic protein	(Schwarz et al., 2012a)
A major capsid protein of a fish pathogen	$HlyA_s$	*E. coli*	0.4 mg/l	Live vaccine against viral pathogens in Fish	(Yang et al., 2008)
β-Lactamase	$HlyA_s$	*E. coli*	5 mg/l	Secretion of mature folded protein	(Chervaux et al., 1995)
S. marcescens HasA	Cognate signal	*E. coli*	n.r.	Secretion through a heterologous T1SS	(Letoffe et al., 1994)
Glycoprotease	$HlyA_s$	*E. coli*	n.r.	Suitable for use in vaccine trials	(Lo et al., 1994)
Streptokinase (Skc) from *Streptococcus equisimilis*	$HlyA_s$	*E. coli*	n.r.		(Kern and Ceglowski, 1995)
Cytoplasmic protein dihydrofolate reductase (DHFR)	$HlyA_s$	*E. coli*			(Nakano et al., 1992)
Antibody production					
IgA VH and IgA VL	$HlyA_s$	*E. coli*	0.1 mg/l		(Holland et al., 1990)
Single-chain Fv antibody	$HlyA_s$	*E. coli*	1–2 mg/l		(Fernández et al., 2000)

fragments, of a murine (*Mus musculus*) immunoglobulin using the HlyA secretion system (Holland et al., 1990). According to another study performed by Fernández et al., the active single-chain Fv antibodies (scFvs) secreted through the HlyA secretion system into culture supernatant can be stored for some while or used directly in a variety of immunoassays without further purification steps (Fernández et al., 2000). It has been shown that T1SS can be applied also for the generation of monoclonal and polyclonal antisera against specific proteins. This was exemplified by the fusion of an outer membrane protein of *Salmonella typhimurium*, PagC, to HlyA1. The secreted PagC-HlyA1 was easily isolated and used to generate either polyclonal or monoclonal anti-PagC that could specifically detect *Salmonella* in immunological assays (Mollenkopf et al., 1996).

Nanobodies are single-domain antibodies derives from the heavy-chain-only antibodies from the serum of camelids. The production of nanobodies has attracted significant scientific attention because of their usage in wide variety of applications such as protein crystallography, cell biology, diagnostics or therapy of human diseases. Successful secretion of nanobodies is an interesting approach facilitating not only the purification process but also providing a platform for screening of nanobodies against a specific antigen (Ruano-Gallego et al., 2019). Only recently, it was reported that the HlyA secretion system can be used as a secretion platform for the secretion and screening of a large diversity of nanobodies. The culture supernatant can be used without any further purification in a screening system for the detection of antigen-binding nanobodies. Interestingly, the secreted Nb-HlyAs contain the disulfide bonds between the cysteine pairs although the reduced forms of cysteines are expected in the cytoplasm of bacteria. This finding also indicates that the nanobodies were oxidized during or after secretion by an unknown mechanism (Ruano-Gallego et al., 2019).

Live antigen-delivery systems based on T1SS: Among different secretion systems, the T1SS could stand as a perfect carrier for the delivery of antigens in different Gram-negative attenuated bacteria. T1SS offer a very interesting area of application for antigen production and antigen delivery in the form of live vaccines. Attenuated live vaccines secreting a specific antigen are highly effective in the immunization of animals and humans, because secreted antigens are accessible to the host immune system without any need of disintegration of the bacterial cells. Furthermore, secreting the antigen avoids the insolubility problems occurring in the reducing condition of bacterial cytosol. The early

successful examples of in vitro functionality of Hly secretion system in different Gram-negative vaccine microorganisms include *Salmonella serotypes*, *Shigella spp*, *V. cholerae*, enterohaemorrhagic *E. coli* (EHEC), enteroinvasive *E. coli* (EIEC), *S. marcescens*, and *P. aeruginosa* (Spreng et al., 1999; Hahn and von Specht, 2003).

Various studies described the successful application of Hly secretion system in secretion of, delivery of, and immunization against different pathogens including bacteria (Gentschev et al., 1996; Spreng et al., 2000a), viruses (Spreng et al., 2000b) and parasites (Gómez-Duarte et al., 2001) using a carrier strain, mostly attenuated Salmonella strains in several model animals including mice, rabbit, and cattle.

One of the most encouraging examples of vaccine development using T1SS was reported by Gentschev et al. in 2004. This group applied *Salmonella typhi* Ty21a, the only licensed live oral Salmonella vaccine for human use, which harbored a heterologous expressed hemolysin apparatus for delivering hemolysin antigen into mice on a pilot scale. All applied conditions agreed well with the condition used for the production of the previously approved attenuated oral vaccine against typhoid fever, Vivotif. Stable replication of the vector, adequate episomal expression, antigen expression even after a short stress condition, and the immunization of applied mice model were all tested and point out that the HlyA secretion system could be applied for the development of an oral vaccine (Gentschev et al., 2004; Hotz et al., 2009; Gentschev et al., 2007).

Suggested application of T1SS for cancer therapy: Using the features of the HlyA secretion system, researchers have been able to develop live vaccines for cancer therapy of animal models (Gentschev et al., 2005; Fensterle et al., 2008). A live bacterial vaccine was developed for the treatment of C-Raf dependent human tumors in animal model (Gentschev et al., 2005). Injection of an attenuated *S. enterica serovar* Typhimurium strain harboring the hemolysin apparatus secreting Raf antigen into two mice models suffering from oncogene-induced lung adenomas significantly reduced the C-Raf tumor growth. It revealed that the injection of the afore mentioned live vaccine elicits specific C-Raf antibody production and T-cell response in both animal models (Gentschev et al., 2005).

Another successful example of cancer immunotherapy using T1SS was illustrated with *S. enterica serovar* Typhimurium aroA against Prostate cancer. In this study, Prostate specific-antigen was secreted through the heterologous hemolysin secretion system of the mentioned strain. The injection of this attenuated live vaccine caused the induction of cytotoxic $CD8^+$ T-cell response in mice having tumor cells (Fensterle et al., 2008). The versatility of T1SS for the generation of live bacterial vaccines for delivery antigens from bacteria, viruses, and parasites to human cells could be seen by the given examples, however, its application for human cells needs further investigation.

Although aforementioned case studies are attractive examples of using T1SS for vaccine development, we are not aware of any use of this system in biopharmaceutical industry. The therapeutic proteins not only should have full protein functionality but also their production need to be cost-effective to receive attention from biopharmaceutical industry (Sanchez-Garcia et al., 2016). The limited secretion level of proteins by T1SS on one side and all restricted constraints of the biopharmaceutical industry on the other side resulted in a non-use of T1SS in biopharmaceutical drug production. However, recent research has generated hope in application of T1SS for producing peptides as active pharmaceutical ingredients.

Production of peptides using T1SS: Recombinant production of peptides is costly and time consuming because of their characteristics such as proteolytic sensitivity, aggregation and cytotoxicity. Lack of a decisive expression platform for peptide production made the chemical synthetics the strategy for peptides production, however, the recombinant production of peptides has also attracted attentions, especially for peptides above a certain length that cannot be easily accessed by classic solid phase synthesis. Recently, the secretion signal of the HlyA secretion system was applied as a protein tag to produce peptides and small proteins. By using this protein tag, the peptides expressed as inclusion bodies. Therefore, beside the high yield of expression, the fused peptide was protected from proteolytic aggregation. Also, the producer host was protected from cytotoxicity feature of the peptide. Renaturation and purification were performed in few steps and with relatively less efforts. With this T1SS based approach, named Numaswitch, more than 2 g/l of Teriparatide was produced which is almost 20-folds higher yield compared to other production strategies for peptide production (Nguyen et al., 2021).

3.6. Applied engineering strategies for T1SS

Despite long-term studies on T1SS since its discovery in 1979, a general platform based on T1SS for the secretion of a wide range of recombinant proteins in reasonable titers has not been provided so far, even though a growing body of scientific research opens up hope by providing practical supports to overcome the limitations of this system in particular concerning the yield of secretion. The efforts given on the engineering of T1SS for improving secretion efficiency can be summarized in four different areas, as follows: culture condition, engineering on the molecular level, changing ABC transporters and host, and protein directed evolution.

3.6.1. Culture condition

The impact of growth conditions on the secretion level of T1SS has been one of the early and frequently discussed strategies to improve the secretion efficiency of this system. The first reported case on this strategy backs to 1980 when Springer and Gobbel reported that the composition of the medium and the growth temperature are two altering factors on the secretion level of HlyA. Also, it was observed that the extracellular and intracellular amounts of HlyA accumulate up to the mid-logarithmic phase (Springer and Goebel, 1980). The concept that efficient protein secretion through T1SS proceeds only by actively growing cells was confirmed also by Eom et al. investigating the TliDEF system of *P. fluorescens* in recombinant *E. coli* (Eom et al., 2006). It also pointed out that the concentration of inducer is an influencing factor on the secretion level probably by controlling the amount of expressed ABC transporter and substrate (Eom et al., 2006). A statistical model developed by Low et al. showed that there is a synergistic relationship between the inducer concentration and post-induction temperature to improve the secretion level of CGTase through Hly system of *E. coli* (Low et al., 2011).

Only recently, Ihling et al. reported culturing *E. coli* in a defined mineral medium to secrete HlyA1. The use of mineral medium results in less impurities of the culture supernatant, which enhances the advantages of recombinant proteins secretion of *E. coli*. They also scaled up the secretion of T1SS for the first time in a stirred tank fermentation yielding 0.54 g/l HlyA1 in the supernatant and showed the possibility toward using T1SS for the industrial scales (Ihling et al., 2019).

3.6.2. Engineering on the molecular level

Genetic engineering of T1SS can be considered as a highly efficient strategy to deal with T1SS limitations. In support of this idea, studies on two genetically relevant elements, JUMPStart DNA sequence and anti-termination factor element RfaH, illustrated the significant influences of these two elements on the secretion level of hemolysin by enhancing transcript elongation. RfaH-enhanced transcript elongation occurs after evolving RfAH with JUMPStart sequence (for just upstream of many polysaccharide-associated genes starts). Deletion of each of these elements causes a decrease in the secretion level (Leeds and Welch, 1997; Leeds and Welch, 1996). Another case study on the T1SS suggests that co-expression of HlyA with RfaH, anti-terminator regulatory factor, leads to a high improvement of the secretion efficiency of HlyA in the attenuated *S. enterica serovar* Typhi strain Ty21a (Hotz et al., 2009).

Khosa et al. determined that a genetic element on the expression vector of HlyA increases significantly the secretion level of HlyA system.

This element, which is a part of *hly*C gene, is rich of adenosine and uracil nucleotides and acts as an enhancer fragment. What is striking about this enhancer fragment is that the established vector containing this enhancer fragment enabled the secretion of some fast-folding proteins that could not be secreted by T1SS system so far, possibly by attracting ribosomal protein S1 which results in a faster and stronger translation of *hly* mRNA (Khosa et al., 2018).

Son et al. demonstrated a further development on the genetic level for the secretion of T1SS from *P. fluorescens* by deleting two genes encoding the protease (PrtA) and lipase (TilA) of this strain. In that study, secretion level increased by protection from degradative protease and removal of competitive substrate lipase (Son et al., 2012).

Interestingly, Byun et al. conducted a series of the secretion experiments through TliDEF system in which the pI of the fused recombinant protein was the focus of the study. They suggested that proteins with acidic pI and thus negatively charge at neutral pH are exported more efficient through TliDEF system probably due to easy entry into the positively charged pore of the transporter and the existing membrane potential of the inner membrane (negative on the cytosolic side) (Byun et al., 2017). Considering all the above-mentioned case studies, more research should be performed on genetic modifications of T1SS to hopefully result in further improvement of secretion titers.

3.6.3. Changing ABC transporters and host

It has been shown by Eom et al. that the ABC transporters could have an impact on the secretion level of target proteins. Eom et al. tested the secretion of *P. fluorescens* lipase (TliA) in recombinant *E. coli* harboring one of the four following ABC transporters: *P. fluorescens* TliDEF, *P. aeruginosa* AprDEF, *Erwinia chrysanthemi* PrtDEF, and *S. marcescens* LipBCD. The latter one showed the highest lipase secretion in recombinant *E. coli* (Eom et al., 2014a). However, Eom et al. did not attempt to provide an explanation. Nevertheless, these results suggest that the type of ABC transporter should be taken into account as an influencing factor on the secretion level. Eom et al. showed that *Serratia proteamaculans* metalloprotase could be secreted in *E. coli* through a heterologous TliDEF exporter of *P. fluorescens* and switching the microbial host from *E. coli* to *P. fluorescens* resulted in a 34.8-fold increase of secretion (Eom et al., 2014b).

3.6.4. Directed protein evolution

Due to the enormous progress of molecular biology techniques over the past decades, directed protein evolution has become a powerful method for protein engineering and evolving protein properties in an artificial and practical manner (Wong et al., 2006). Until now, just a few studies have published on direct evolution of T1SS.

The first reported case on a directed protein evolution of a T1SS was performed on the TliDEF system from *P. fluorescence*, secreting thermostable lipase (TliA) in recombinant *E. coli*. Eom et al. could isolate four mutants with the highest secretion level up to 3.2-fold increase. Interestingly, all beneficial positions of the isolated mutants were located in the predicted cytoplasmic region of the TliD, the ABC transporter of this system. They also provided evidence showing that the observed higher secretion level is contributed by the higher efficiency of the ABC transporter and not due to the higher expression level of this transporter protein. The novel ABC transporter harboring those mutations did not enhance secretion level of an alternative target protein, the protease PrtA from *P. fluorescens* SIK_W1 (Eom et al., 2005).

The second published study on the directed evolution of a T1SS was performed on HlyA system from *E. coli*. By screening two generated libraries of the *hly*B and *hly*D genes with a size of 1.5×10^4, two mutants were isolated to have significantly enhanced secretion level of subtilisin by 27 and 15 folds. However, the discovered mutants showed different secretion profiles according to different target proteins or temperatures (Sugamata and Shiba, 2005).

The HlyA system of *E. coli* was subjected once again to a directed protein evolution by Kheng et al. and resulted a 35–217% increase in secretion level of cyclodextrin glucanotransferase compared to the wild-type strain. The detected beneficial positions were mostly spread over the entire transmembrane domains of HlyB (Low et al., 2010).

All three published studies on the directed evolution of T1SS applied error-prone PCR (epPCR) with a low mutation load for generating libraries. The epPCR is a simple and robust method, but one has to stress that this method fails to generate more than 70% of the beneficial positions of the target protein(s) (Yang et al., 2017). Another weakness of the reported studies is the possible high standard deviations of applied screening methods since all the studies screened the mutants via agar-plate activity assays. Research on protein evolution of T1SS has been mostly restricted to the traditional directed evolution strategies, and now the strategic direct evolution of T1SS is needed for significantly enhanced secretion properties.

4. Conclusion and future outlook

T1SS has been successfully applied for the secretion of a long list of heterologous recombinant proteins, small proteins, and peptides. This system is also able to export a range of biopharmaceutical relevant products including antibody fragments, cytokines, hormone proteins, growth factors and different antigens. Notably, the correct folding of proteins containing disulfide bonds exported by T1SS in the culture supernatant opens another encouraging property of using this secretion system. Large-scale secretion of T1SS has started to demonstrate its applicability for commercial scales, although the secretion yields are still far away of what is termed industrial standard as exemplified by lipases secreted in the 20 g / l range from Bacillus strains. However, the described developments and properties should encourage researcher to exploit this system, despite the fact that there is a clear need for improving the secretion efficiency of this system. Providing a general secretion platform based on T1SS for different recombinant proteins can address the long-standing issue in biotechnology that is protein production in optimum quality and yield yet at reduced cost.

Declaration of funding source

Zohreh Pourhassan N. was financially supported by the CLIB Competence Center Biotechnology (CKB) funded by the European Regional Development Fund ERDF (34.EFRE-0300096).

Submission declaration

This article has not been submitted to, nor is under review at, other journals. The publication of this work is approved by all authors. All authors declare that the article, in the case of acceptance, will not be published in same form to other journals, neither in English nor in other languages, including electronical form without the written acquiescence of the copyright holders.

Author contributions

Zohreh Pourhassan N.: Conceptualization, Literature search, Writing and editing the draft.

Sander H.J. Smits: Conceptualization, Reviewing and revising the draft for the content. Jung.

Hoon Ahn: Conceptualization, Writing one section of the initial draft, Revising the draft for.

the content. Lutz Schmitt: Supervision, Conceptualization, Revising the draft for the content.

Declaration of Competing Interest

The authors declare that they have no conflict of interest.

Acknowledgements

We thank all current and former members of the Institute of Biochemistry for stimulating discussions. We also thank Prof. Karl-Erich Jaeger for suggestions, discussions and support.

References

Abrusci, P., Mcdowell, M.A., Lea, S.M., Johnson, S., 2014. Building a secreting nanomachine: a structural overview of the T3SS. Curr. Opin. Struct. Biol. 25, 111–117.

Ahn, J.H., Pan, J.G., Rhee, J.S., 1999. Identification of the tliDEF ABC transporter specific for lipase in *Pseudomonas fluorescens* SIK W1. J. Bacteriol. 181, 1847–1852.

Anderson, C., Hughes, C., Koronakis, V., 2000. Chunnel vision: export and efflux through bacterial channel-tunnels. EMBO Rep. 1, 313–318.

Angkawidjaja, C., Kuwahara, K., Omori, K., Koga, Y., Takano, K., Kanaya, S., 2006. Extracellular secretion of *Escherichia coli* alkaline phosphatase with a C-terminal tag by type I secretion system: purification and biochemical characterization. Protein Eng. Des. Sel. 19, 337–343.

Anné, J., Economou, A., Bernaerts, K., 2016. Protein Secretion in Gram-positive Bacteria: From Multiple Pathways to Biotechnology. Protein and Sugar Export and Assembly in Gram-positive Bacteria. Springer.

Azam, A., Li, C., Metcalf, K.J., Tullman-Ercek, D., 2016. Type III secretion as a generalizable strategy for the production of full-length biopolymer-forming proteins. Biotechnol. Bioeng. 113, 2313–2320.

Bai, F., Li, Z., Umezawa, A., Terada, N., Jin, S., 2018. Bacterial type III secretion system as a protein delivery tool for a broad range of biomedical applications. Biotechnol. Adv. 36, 482–493.

Bakkes, P.J., Jenewein, S., Smits, S.H., Holland, I.B., Schmitt, L., 2010. The rate of folding dictates substrate secretion by the *Escherichia coli* hemolysin type 1 secretion system. J. Biol. Chem. 285, 40573–40580.

Balakrishnan, L., Hughes, C., Koronakis, V., 2001. Substrate-triggered recruitment of the TolC channel-tunnel during type I export of hemolysin by *Escherichia coli*. J. Mol. Biol. 313, 501–510.

Baneyx, F., Mujacic, M., 2004. Recombinant protein folding and misfolding in *Escherichia coli*. Nat. Biotechnol. 22, 1399–1408.

Barlag, B., Hensel, M., 2015. The giant adhesin SiiE of Salmonella enterica. Molecules 20, 1134–1150.

Benabdelhak, H., Kiontke, S., Horn, C., Ernst, R., Blight, M.A., Holland, I.B., Schmitt, L., 2003. A specific interaction between the NBD of the ABC-transporter HlyB and a C-terminal fragment of its transport substrate haemolysin A. J. Mol. Biol. 327, 1169–1179.

Boardman, B.K., Fullner Satchell, K.J., 2004. Vibrio cholerae strains with mutations in an atypical type I secretion system accumulate RTX toxin intracellularly. J. Bacteriol. 186, 8137–8143.

Boyd, C.D., Smith, T.J., El-Kirat-Chatel, S., Newell, P.D., Dufrêne, Y.F., O'toole, G.A., 2014. Structural features of the *Pseudomonas fluorescens* biofilm adhesin LapA required for LapG-dependent cleavage, biofilm formation, and cell surface localization. J. Bacteriol. 196, 2775–2788.

Bumba, L., Masin, J., Macek, P., Wald, T., Motlova, L., Bibova, I., Klimova, N., Bednarova, L., Veverka, V., Kachala, M., 2016. Calcium-driven folding of RTX domain β-rolls ratchets translocation of RTX proteins through type I secretion ducts. Mol. Cell 62, 47–62.

Burdette, L.A., Leach, S.A., Wong, H.T., Tullman-Ercek, D., 2018. Developing Gram-negative bacteria for the secretion of heterologous proteins. Microb. Cell Factories 17, 196.

Byun, H., Park, J., Kim, S.C., Ahn, J.H., 2017. A lower isoelectric point increases signal sequence–mediated secretion of recombinant proteins through a bacterial ABC transporter. J. Biol. Chem. 292, 19782–19791.

Chervaux, C., Sauvonnet, N., LE Clainche, A., Kenny, B., Hunt, A.L., Broome-Smith, J.K., Holland, I.B., 1995. Secretion of active β-lactamase to the medium mediated by the *Escherichia coli* haemolysin transport pathway. Mol. Gen. Genet. MGG 249, 237–245.

Chung, C.W., You, J., Kim, K., Moon, Y., Kim, H., Ahn, J.H., 2009. Export of recombinant proteins in *Escherichia coli* using ABC transporter with an attached lipase ABC transporter recognition domain (LARD). Microb. Cell Factories 8, 1–12.

Costa, T.R., Felisberto-Rodrigues, C., Meir, A., Prevost, M.S., Redzej, A., Trokter, M., Waksman, G., 2015. Secretion systems in gram-negative bacteria: structural and mechanistic insights. Nat. Rev. Microbiol. 13, 343–359.

Costa, T.R., Harb, L., Khara, P., Zeng, L., Hu, B., Christie, P.J., 2021. Type IV secretion systems: advances in structure, function, and activation. Mol. Microbiol. 115, 436.

Cronan, J.E., 2014. *Escherichia coli* as an Experimental Organism. eLS.

Delepelaire, P., 2004. Type I secretion in gram-negative bacteria. Biochim. Biophys. Acta 1694, 149–161.

Du, D., Wang, Z., James, N.R., Voss, J.E., Klimont, E., Ohene-Agyei, T., Venter, H., Chiu, W., Luisi, B.F., 2014. Structure of the AcrAB–TolC multidrug efflux pump. Nature 509, 512–515.

Eom, G.T., Song, J.K., Ahn, J.H., Seo, Y.S., Rhee, J.S., 2005. Enhancement of the efficiency of secretion of heterologous lipase in *Escherichia coli* by directed evolution of the ABC transporter system. Appl. Environ. Microbiol. 71, 3468–3474.

Eom, G.-T., Rhee, J.-S., Song, J.-K., 2006. An efficient secretion of type I secretion pathway-dependent lipase, TliA, in *Escherichia coli:* effect of relative expression levels and timing of passenger protein and ABC transporter. J. Microbiol. Biotechnol. 16, 1422–1428.

Eom, G.T., Lee, S.H., Oh, Y.H., Choi, J.E., Park, S.J., Song, J.K., 2014a. Efficient extracellular production of type I secretion pathway-dependent *Pseudomonas fluorescens* lipase in recombinant *Escherichia coli* by heterologous ABC protein exporters. Biotechnol. Lett. 36, 2037–2042.

Eom, G.T., Oh, J.Y., Song, J.K., 2014b. High-level production of *Serratia proteamaculans* metalloprotease using a recombinant ABC protein exporter-mediated secretion system in *Pseudomonas fluorescens*. Process Biochem. 49, 1718–1722.

Eom, G.T., Oh, J.Y., Park, J.H., Jegal, J., Song, J.K., 2016a. Secretory production of enzymatically active endo-β-1, 4-mannanase from *Bacillus subtilis* by ABC exporter in *Escherichia coli*. Process Biochem. 51, 999–1005.

Eom, G.T., Oh, J.Y., Park, J.H., Lim, H.J., Lee, S.J., Kim, E.Y., Choi, J.-E., Jegal, J., Song, B.K., Yu, J.-H., 2016b. Alleviation of temperature-sensitive secretion defect of *Pseudomonas fluorescens* ATP-binding cassette (ABC) transporter, TliDEF, by a change of single amino acid in the ABC protein, TliD. J. Biosci. Bioeng. 122, 283–286.

Fensterle, J., Bergmann, B., Yone, C., Hotz, C., Meyer, S., Spreng, S., Goebel, W., Rapp, U., Gentschev, I., 2008. Cancer immunotherapy based on recombinant *Salmonella enterica* serovar Typhimurium aroA strains secreting prostate-specific antigen and cholera toxin subunit B. Cancer Gene Ther. 15, 85–93.

Fernández, L.A., Sola, I., Enjuanes, L., DE Lorenzo, V., 2000. Specific secretion of active single-chain Fv antibodies into the supernatants of *Escherichia coli* cultures by use of the hemolysin system. Appl. Environ. Microbiol. 66, 5024–5029.

Fitzpatrick, A.W., Llabrés, S., Neuberger, A., Blaza, J.N., Bai, X.-C., Okada, U., Murakami, S., van Veen, H.W., Zachariae, U., Scheres, S.H., 2017. Structure of the MacAB–TolC ABC-type tripartite multidrug efflux pump. Nat. Microbiol. 2, 1–8.

Forster, B.M., Marquis, H., 2012. Protein transport across the cell wall of monoderm gram-positive bacteria. Mol. Microbiol. 84, 405–413.

Fraile, S., Muñoz, A., de Lorenzo, V., Fernández, L.A., 2004. Secretion of proteins with dimerization capacity by the haemolysin type I transport system of *Escherichia coli*. Mol. Microbiol. 53, 1109–1121.

Freudl, R., 2018. Signal peptides for recombinant protein secretion in bacterial expression systems. Microb. Cell Factories 17, 52.

Gawarzewski, I., Smits, S.H., Schmitt, L., Jose, J., 2013. Structural comparison of the transport units of type V secretion systems. Biol. Chem. 394, 1385–1398.

Gentschev, I., Mollenkopf, H., Sokolovic, Z., Hess, J., Kaufmann, S.H., Goebel, W., 1996. Development of antigen-delivery systems, based on the *Escherichia coli* hemolysin secretion pathway. Gene 179, 133–140.

Gentschev, I., Dietrich, G., Spreng, S., Neuhaus, B., Maier, E., Benz, R., Goebel, W., Fensterle, J., Rapp, U.R., 2004. Use of the α-hemolysin secretion system of *Escherichia coli* for antigen delivery in the *Salmonella typhi* Ty21a vaccine strain. Int. J. Med. Microbiol. 294, 363–371.

Gentschev, I., Fensterle, J., Schmidt, A., Potapenko, T., Troppmair, J., Goebel, W., Rapp, U.R., 2005. Use of a recombinant *Salmonella enterica* serovar Typhimurium strain expressing C-Raf for protection against C-Raf induced lung adenoma in mice. BMC Cancer 5, 15.

Gentschev, I., Spreng, S., Sieber, H., Ures, J., Mollet, F., Collioud, A., Pearman, J., Griot-Wenk, M.E., Fensterle, J., Rapp, U.R., 2007. Vivotif®–a 'magic shield'for protection against typhoid fever and delivery of heterologous antigens. Chemotherapy 53, 177–180.

Ghigo, J.-M., Wandersman, C., 1994. A carboxyl-terminal four-amino acid motif is required for secretion of the metalloprotease PrtG through the *Erwinia chrysanthemi* protease secretion pathway. J. Biol. Chem. 269, 8979–8985.

Gómez-Duarte, O.G., Pasetti, M.F., Santiago, A., Sztein, M.B., Hoffman, S.L., Levine, M. M., 2001. Expression, extracellular secretion, and immunogenicity of the *Plasmodium falciparum* sporozoite surface protein 2 in Salmonella vaccine strains. Infect. Immun. 69, 1192–1198.

Green, E.R., Mecsas, J., 2016. Bacterial secretion systems: an overview. In: Virulence Mechanisms of Bacterial Pathogens, pp. 213–239.

Grohmann, E., Christie, P.J., Waksman, G., Backert, S., 2018. Type IV secretion in Gram-negative and Gram-positive bacteria. Mol. Microbiol. 107, 455–471.

Hahn, H.P., von Specht, B.-U., 2003. Secretory delivery of recombinant proteins in attenuated Salmonella strains: potential and limitations of type I protein transporters. FEMS Immunol. Med. Microbiol. 37, 87–98.

Harwood, C.R., 1992. Bacillus subtilis and its relatives: molecular biological and industrial workhorses. Trends Biotechnol. 10, 247–256.

Heesemann, J., Gross, U., Schmidt, N., Laufs, R., 1986. Immunochemical analysis of plasmid-encoded proteins released by enteropathogenic Yersinia sp. grown in calcium-deficient media. Infect. Immun. 54, 561–567.

Hepp, C., Maier, B., 2017. Bacterial translocation ratchets: shared physical principles with different molecular implementations: how bacterial secretion systems bias Brownian motion for efficient translocation of macromolecules. Bioessays 39, 1700099.

Holland, I.B., Kenny, B., Steipe, B., Plückthun, A., 1990. [11] Secretion of Heterologous Proteins in *Escherichia coli*. Methods in Enzymology. Elsevier.

Holland, I.B., Peherstorfer, S., Kanonenberg, K., Lenders, M., Reimann, S., Schmitt, L., 2016. Type I protein secretion-deceptively simple yet with a wide range of mechanistic variability across the family. EcoSal Plus 7.

Hotz, C., Fensterle, J., Goebel, W., Meyer, S.R., Kirchgraber, G., Heisig, M., Fürer, A., Dietrich, G., Rapp, U.R., Gentschev, I., 2009. Improvement of the live vaccine strain *Salmonella enterica* serovar Typhi Ty21a for antigen delivery via the hemolysin secretion system of *Escherichia coli*. Int. J. Med. Microbiol. 299, 109–119.

Hu, B., Lara-Tejero, M., Kong, Q., Galán, J.E., Liu, J., 2017. In situ molecular architecture of the Salmonella type III secretion machine. Cell 168, 1065–1074 (e10).

Huang, C.-J., Lin, H., Yang, X., 2012. Industrial production of recombinant therapeutics in *Escherichia coli* and its recent advancements. J. Ind. Microbiol. Biotechnol. 39, 383–399.

Ihling, N., Uhde, A., Scholz, R., Schwarz, C., Schmitt, L., Büchs, J., 2019. Scale-up of a type I secretion system in *E. coli* using a defined mineral medium. Biotechnol. Progr. 36 (2), e2911.

Jacob-Dubuisson, F., Locht, C., Antoine, R., 2001. Two-partner secretion in Gram-negative bacteria: a thrifty, specific pathway for large virulence proteins. Mol. Microbiol. 40, 306–313.

Jaeger, K., Dijkstra, B., Reetz, M., 1999. Bacterial biocatalysts: molecular biology, three-dimensional structures, and biotechnological applications of lipases. Ann. Rev. Microbiol. 53, 315–351.

Jia, B., Jeon, C.O., 2016. High-throughput recombinant protein expression in *Escherichia coli*: current status and future perspectives. Open Biol. 6.

Jo, I., Kim, J.-S., Xu, Y., Hyun, J., Lee, K., Ha, N.-C., 2019. Recent paradigm shift in the assembly of bacterial tripartite efflux pumps and the type I secretion system. J. Microbiol. 57, 185–194.

Johnson, I.S., 1983. Human insulin from recombinant DNA technology. Science 219, 632–637.

Kanonenberg, K., Spitz, O., Erenburg, I.N., Beer, T., Schmitt, L., 2018. Type I secretion system—it takes three and a substrate. FEMS Microbiol. Lett. 365, fny094.

Kanonenberg, K., Smits, S.H., Schmitt, L., 2019. Functional reconstitution of HlyB, a type I secretion ABC transporter, in saposin-A nanoparticles. Sci. Rep. 9, 1–12.

Kern, I., Ceglowski, P., 1995. Secretion of streptokinase fusion proteins from *Escherichia coli* cells through the hemolysin transporter. Gene 163, 53–57.

Khosa, S., Scholz, R., Schwarz, C., Trilling, M., Hengel, H., Jaeger, K.-E., Smits, S.H., Schmitt, L., 2018. An A/U-rich enhancer region is required for high-level protein secretion through the hlya type i secretion system. Appl. Environ. Microbiol. 84, e01163-17.

Kleiner-Grote, G.R., Risse, J.M., Friehs, K., 2018. Secretion of recombinant proteins from *E. coli*. Eng. Life Sci. 18, 532–550.

Koronakis, V., Hughes, C., Koronakis, E., 1991. Energetically distinct early and late stages of HlyB/HlyD-dependent secretion across both *Escherichia coli* membranes. EMBO J. 10, 3263–3272.

Koronakis, V., Li, J., Koronakis, E., Stauffer, K., 1997. Structure of TolC, the outer membrane component of the bacterial type I efflux system, derived from two-dimensional crystals. Mol. Microbiol. 23, 617–626.

Krauss, U., Jäger, V.D., Diener, M., Pohl, M., Jaeger, K.-E., 2017. Catalytically-active inclusion bodies—carrier-free protein immobilizates for application in biotechnology and biomedicine. J. Biotechnol. 258, 136–147.

Kunert, R., Reinhart, D., 2016. Advances in recombinant antibody manufacturing. Appl. Microbiol. Biotechnol. 100, 3451–3461.

Kwon, H.-J., Haruki, M., Morikawa, M., Omori, K., Kanaya, S., 2002. Role of repetitive nine-residue sequence motifs in secretion, enzymatic activity, and protein conformation of a family I. 3 lipase. J. Biosci. Bioeng. 93, 157–164.

Lecher, J., Schwarz, C.K., Stoldt, M., Smits, S.H., Willbold, D., Schmitt, L., 2012. An RTX transporter tethers its unfolded substrate during secretion via a unique N-terminal domain. Structure 20, 1778–1787.

Leeds, J.A., Welch, R.A., 1996. RfaH enhances elongation of *Escherichia coli* hlyCABD mRNA. J. Bacteriol. 178, 1850–1857.

Leeds, J.A., Welch, R.A., 1997. Enhancing transcription through the *Escherichia coli* hemolysin operon, hlyCABD: RfaH and upstream JUMPStart DNA sequences function together via a postinitiation mechanism. J. Bacteriol. 179, 3519–3527.

Lenders, M.H., Weidtkamp-Peters, S., Kleinschrodt, D., Jaeger, K.-E., Smits, S.H., Schmitt, L., 2015. Directionality of substrate translocation of the hemolysin A Type I secretion system. Sci. Rep. 5, 12470.

Lenders, M.H., Beer, T., Smits, S.H., Schmitt, L., 2016. In vivo quantification of the secretion rates of the hemolysin A Type I secretion system. Sci. Rep. 6, 33275.

Letoffe, S., Ghigo, J., Wandersman, C., 1994. Secretion of the Serratia marcescens HasA protein by an ABC transporter. J. Bacteriol. 176, 5372–5377.

Letoffe, S., Delepelaire, P., Wandersman, C., 1996. Protein secretion in gram-negative bacteria: assembly of the three components of ABC protein-mediated exporters is ordered and promoted by substrate binding. EMBO J. 15, 5804–5811.

Li, Y., Reichenstein, K., Ullrich, R., Danner, T., Von Specht, B.-U., Hahn, H.P., 2003. Effect of in situ expression of human interleukin-6 on antibody responses against *Salmonella typhimurium* antigens. FEMS Immunol. Med. Microbiol. 37, 135–145.

Linhartová, I., Bumba, L., Masín, J., Basler, M., Osička, R., Kamanová, J., Procházková, K., Adkins, I., Hejnová-Holubová, J., Sadílková, L., Morová, J., Sebo, P., 2010. RTX proteins: a highly diverse family secreted by a common mechanism. FEMS Microbiol. Rev. 34, 1076–1112.

Lo, R.Y., Watt, M.-A., Gyroffy, S., Mellors, A., 1994. Preparation of recombinant glycoprotease of *Pasteurella haemolytica* A1 utilizing the *Escherichia coli* α-hemolysin secretion system. FEMS Microbiol. Lett. 116, 225–230.

Low, K.O., Mahadi, N.M., Rahim, R.A., Rabu, A., Bakar, F.D.A., Murad, A.M.A., Illias, R. M., 2010. Enhanced secretory production of hemolysin-mediated cyclodextrin glucanotransferase in *Escherichia coli* by random mutagenesis of the ABC transporter system. J. Biotechnol. 150, 453–459.

Low, K.O., Mahadi, N.M., Rahim, R.A., Rabu, A., Bakar, F.D.A., Murad, A.M.A., Illias, R. M., 2011. An effective extracellular protein secretion by an ABC transporter system in *Escherichia coli*: statistical modeling and optimization of cyclodextrin glucanotransferase secretory production. J. Ind. Microbiol. Biotechnol. 38, 1587–1597.

Mack, L., Brill, B., Delis, N., Groner, B., 2014. Endotoxin depletion of recombinant protein preparations through their preferential binding to histidine tags. Anal. Biochem. 466, 83–88.

Mackman, N., Baker, K., Gray, L., Haigh, R., Nicaud, J.M., Holland, I.B., 1987. Release of a chimeric protein into the medium from *Escherichia coli* using the C-terminal secretion signal of haemolysin. EMBO J. 6, 2835–2841.

Mollenkopf, H.-J., Gentschev, I., Bubert, A., Schubert, P., Goebelc, W., 1996. Extracellular PagC-HlyA S fusion protein for the generation and identification of Salmonella-specific antibodies. Appl. Microbiol. Biotechnol. 45, 629–637.

Nakano, H., Kawakami, Y., Nishimura, H., 1992. Secretion of genetically-engineered dihydrofolate reductase from *Escherichia coli* using an *E. coli* α-hemolysin membrane translocation system. Appl. Microbiol. Biotechnol. 37, 765–771.

Nguyen, B.-N., Tieves, F., Rohr, T., Wobst, H., Schöpf, F.S., Solano, J.D.M., Schneider, J., Stock, J., Uhde, A., Kalthoff, T., 2021. Numaswitch: an efficient high-titer expression platform to produce peptides and small proteins. AMB Express 11, 1–7.

Nicaud, J.-M., Mackman, N., Gray, L., Holland, I., 1986. The C-terminal, 23 kDa peptide of *E. coli* haemolysin 2001 contains all the information necessary for its secretion by the haemolysin (Hly) export machinery. FEBS Lett. 204, 331–335.

Nivaskumar, M., Francetic, O., 2014. Type II secretion system: a magic beanstalk or a protein escalator. Biochim. Biophys. Acta 1843, 1568–1577.

Noegel, A., Rdest, U., Springer, W., Goebel, W., 1979. Plasmid cistrons controlling synthesis and excretion of the exotoxin α-haemolysin of *Escherichia coli*. Mol. Gen. Genet. MGG 175, 343–350.

Owczarek, B., Gerszberg, A., Hnatuszko-Konka, K., 2019. A brief reminder of systems of production and chromatography-based recovery of recombinant protein biopharmaceuticals. BioMed Res. Int. 2019.

Palacios, J.L., Zaror, I., Martínez, P., Uribe, F., Opazo, P., Socías, T., Gidekel, M., Venegas, A., 2001. Subset of hybrid eukaryotic proteins is exported by the type I secretion system of *Erwinia chrysanthemi*. J. Bacteriol. 183, 1346–1358.

Park, Y., Moon, Y., Ryoo, J., Kim, N., Cho, H., Ahn, J.H., 2012. Identification of the minimal region in lipase ABC transporter recognition domain of *Pseudomonas fluorescens* for secretion and fluorescence of green fluorescent protein. Microb. Cell Factories 11, 60.

Park, J., Eom, G.T., Young, O.H., Hyun Park, J., Kim, S.C., Kwang Song, J., Hoon Ahn, J., 2020. High-level production of bacteriotoxic phospholipase A1 in bacterial host *Pseudomonas fluorescens* via ABC transporter-mediated secretion and inducible expression. Microorganisms 8, 239.

Peng, C., Shi, C., Cao, X., Li, Y., Liu, F., Lu, F., 2019. Factors influencing recombinant protein secretion efficiency in gram-positive bacteria: signal peptide and beyond. Front. Bioeng. Biotechnol. 7, 139.

Pimenta, A., Young, J., Holland, I., Blight, M., 1999. Antibody analysis of the localisation, expression and stability of HlyD, the MFP component of the *E. coli* haemolysin translocator. Mol. Gen. Genet. MGG 261, 122–132.

Pohl, S., Harwood, C.R., 2010. Heterologous protein secretion by Bacillus species: from the cradle to the grave. Adv. Appl. Microbiol. 73, 1–25.

Puetz, J., Wurm, F.M., 2019. Recombinant proteins for industrial versus pharmaceutical purposes: a review of process and pricing. Processes 7, 476.

Reichow, S.L., Korotkov, K.V., Hol, W.G.J., Gonen, T., 2010. Structure of the cholera toxin secretion channel in its closed state. Nat. Struct. Mol. Biol. 17, 1226–1232.

Rempel, S., Stanek, W., Slotboom, D., 2019. ECF-type ATP-binding cassette transporters. Annu. Rev. Biochem. 88, 551–576.

Renauld-Mongenie, G., Mielcarek, N., Cornette, J., Schacht, A.-M., Capron, A., Riveau, G., Locht, C., 1996. Induction of mucosal immune responses against a heterologous antigen fused to filamentous hemagglutinin after intranasal immunization with recombinant *Bordetella pertussis*. Proc. Natl. Acad. Sci. 93, 7944–7949.

Retallack, D.M., Jin, H., Chew, L., 2012. Reliable protein production in a *Pseudomonas fluorescens* expression system. Protein Expr. Purif. 81, 157–165.

Rosano, G.L., Ceccarelli, E.A., 2014. Recombinant protein expression in *Escherichia coli*: advances and challenges. Front. Microbiol. 5, 172.

Rosano, G.L., Morales, E.S., Ceccarelli, E.A., 2019. New tools for recombinant protein production in Escherichia coli: a 5-year update. Protein Sci. 28, 1412–1422.

Ruano-Gallego, D., Fraile, S., Gutierrez, C., Fernández, L.Á., 2019. Screening and purification of nanobodies from *E. coli* culture supernatants using the hemolysin secretion system. Microb. Cell Factor. 18, 47.

Ryu, J., Lee, U., Park, J., Yoo, D.-H., Ahn, J.H., 2015. A vector system for ABC transporter-mediated secretion and purification of recombinant proteins in Pseudomonas species. Appl. Environ. Microbiol. 81, 1744–1753.

Sanchez-Garcia, L., Martín, L., Mangues, R., Ferrer-Miralles, N., Vázquez, E., Villaverde, A., 2016. Recombinant pharmaceuticals from microbial cells: a 2015 update. Microb. Cell Factories 15, 33.

Satchell, K.J., 2011. Structure and function of MARTX toxins and other large repetitive RTX proteins. Annu. Rev. Microbiol. 65, 71–90.

Schtilein, R., Gentschev, I., Mollenkopf, H.-J., Goebel, W., 1992. A topological model for the haemolysin translocator protein HlyD. Mol. Gen. Genet. 234, 155–163.

Schwarz, C.K., Landsberg, C.D., Lenders, M.H., Smits, S.H., Schmitt, L., 2012a. Using an *E. coli* Type 1 secretion system to secrete the mammalian, intracellular protein IFABP in its active form. J. Biotechnol. 159, 155–161.

Schwarz, C.K., Lenders, M.H., Smits, S.H., Schmitt, L., 2012b. Secretion of slow-folding proteins by a type 1 secretion system. Bioengineered 3, 289–292.

Sharma, A., Chaudhuri, T.K., 2017. Revisiting *Escherichia coli* as microbial factory for enhanced production of human serum albumin. Microb. Cell Factories 16, 1–19.

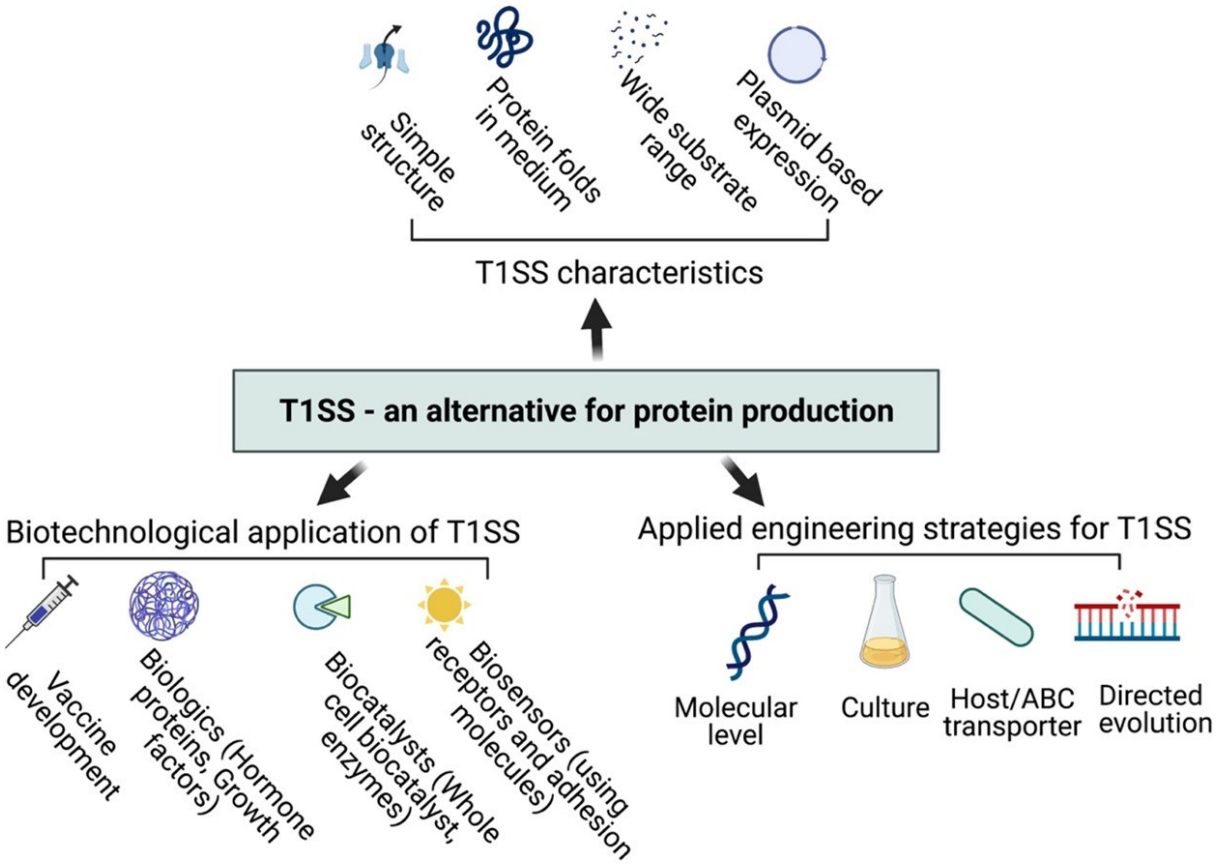
Singhvi, P., Saneja, A., Srichandan, S., Panda, A.K., 2020. Bacterial inclusion bodies: a treasure trove of bioactive proteins. Trends Biotechnol. 38, 474–486.

Smith, T.J., Sondermann, H., O'toole, G.A., 2018. Type 1 does the two-step: type 1 secretion substrates with a functional periplasmic intermediate. J. Bacteriol. 200, e00168-18.

Son, M., Moon, Y., Oh, M.J., Han, S.B., Park, K.H., Kim, J.-G., Ahn, J.H., 2012. Lipase and protease double-deletion mutant of *Pseudomonas fluorescens* suitable for extracellular protein production. Appl. Environ. Microbiol. 78, 8454–8462.

Spitz, O., Erenburg, I.N., Beer, T., Kanonenberg, K., Holland, I.B., Schmitt, L., 2019. Type I Secretion Systems—One Mechanism for All? Protein Secretion in Bacteria, pp. 215–225.
Spreng, S., Dietrich, G., Goebel, W., Gentschev, I., 1999. The *Escherichia coli* haemolysin secretion apparatus: a potential universal antigen delivery system in Gram-negative bacterial vaccine carriers. Mol. Microbiol. 31, 1596–1598.
Spreng, S., Gentschev, I., Goebel, W., Mollenkopf, H.-J., Eck, M., Müller-Hermelink, H.K., Schmausser, B., 2000a. Identification of immunogenic antigens of *Helicobacter pylori* via the *Escherichia coli* hemolysin secretion system. FEMS Microbiol. Lett. 186, 251–256.
Spreng, S., Gentschev, I., Goebel, W., Weidinger, G., TER Meulen, V., Niewiesk, S., 2000b. Salmonella vaccines secreting measles virus epitopes induce protective immune responses against measles virus encephalitis. Microbes Infect. 2, 1687–1692.
Springer, W., Goebel, W., 1980. Synthesis and secretion of hemolysin by *Escherichia coli*. J. Bacteriol. 144, 53–59.
Stefan, E., Hofmann, S., Tampé, R., 2020. A single power stroke by ATP binding drives substrate translocation in a heterodimeric ABC transporter. Elife 9, e55943.
Su, G.-F., Brahmbhatt, H.N., De Lorenzo, V., Wehland, J., Timmis, K.N., 1992. Extracellular export of Shiga toxin B-subunit/haemolysin A (C-terminus) fusion protein expressed in *Salmonella typhimurium* aroA-mutant and stimulation of B-subunit specific antibody responses in mice. Microb. Pathog. 13, 465–476.
Su, L., Chen, S., Yi, L., Woodard, R.W., Chen, J., Wu, J., 2012. Extracellular overexpression of recombinant Thermobifida fusca cutinase by alpha-hemolysin secretion system in *E. coli* BL21 (DE3). Microb. Cell Factories 11, 8.
Sugamata, Y., Shiba, T., 2005. Improved secretory production of recombinant proteins by random mutagenesis of hlyB, an alpha-hemolysin transporter from *Escherichia coli*. Appl. Environ. Microbiol. 71, 656–662.
Thanabalu, T., Koronakis, E., Hughes, C., Koronakis, V., 1998. Substrate-induced assembly of a contiguous channel for protein export from *E. coli*: reversible bridging of an inner-membrane translocase to an outer membrane exit pore. EMBO J. 17, 6487–6496.
Thomas, S., Bakkes, P.J., Smits, S.H., Schmitt, L., 2014a. Equilibrium folding of pro-HlyA from *Escherichia coli* reveals a stable calcium ion dependent folding intermediate. Biochim. Biophys. Acta 1844, 1500–1510.
Thomas, S., Holland, I.B., Schmitt, L., 2014b. The type 1 secretion pathway—the hemolysin system and beyond. Biochim. Biophys. Acta 1843, 1629–1641.
Van Ulsen, P., Ur Rahman, S., Jong, W.S., Daleke-Schermerhorn, M.H., Luirink, J., 2014. Type V secretion: from biogenesis to biotechnology. Biochim. Biophys. Acta 1843, 1592–1611.
Verma, A., Burns, D.L., 2007. Requirements for assembly of PtlH with the pertussis toxin transporter apparatus of *Bordetella pertussis*. Infect. Immun. 75, 2297–2306.
Walker, K.L., Jones, A.S., Robinson, C., 2015. The Tat pathway as a biotechnological tool for the expression and export of heterologous proteins in *Escherichia coli*. Pharm. Bioprocess. 3, 387–396.
Wingfield, P.T., 2015. Overview of the purification of recombinant proteins. Curr. Protoc. Protein Sci. 80, 6.1.1–6.1.35.
Wong, T.S., Zhurina, D., Schwaneberg, U., 2006. The diversity challenge in directed protein evolution. Comb. Chem. High Throughput Screen. 9, 271–288.
Yang, Z., Liu, Q., Wang, Q., Zhang, Y., 2008. Gene cloning and expression of major capsid protein of large yellow croaker iridovirus and its secretion based on α-hemolysin transport system. J. World Aquacult. Soc. 39, 774–780.
Yang, J., Zhao, H.-L., Ran, L.-Y., Li, C.-Y., Zhang, X.-Y., Su, H.-N., Shi, M., Zhou, B.-C., Chen, X.-L., Zhang, Y.-Z., 2015. Mechanistic insights into elastin degradation by Pseudolysin, the major virulence factor of the opportunistic pathogen *Pseudomonas aeruginosa*. Sci. Rep. 5, 9936.
Yang, J., Ruff, A.J., Arlt, M., Schwaneberg, U., 2017. Casting epPCR (cepPCR): a simple random mutagenesis method to generate high quality mutant libraries. Biotechnol. Bioeng. 114, 1921–1927.
Yu, X.-J., Mcgourty, K., Liu, M., Unsworth, K.E., Holden, D.W., 2010. pH sensing by intracellular Salmonella induces effector translocation. Science 328, 1040–1043.

Graphical Abstract



2.2 Chapter II: Directed evolution of the Hly enhancer fragment

Title: Optimized Hemolysin Type 1 Secretion System in *Escherichia coli* by Directed Evolution of the Hly Enhancer Fragment and Including a Terminator Region

Authors: Zohreh Pourhassan N., Haiyang Cui, Sakshi Khosa, Mehdi D. Davari, Karl-Erich Jaeger, Sander H. J. Smits, Ulrich Schwaneberg, and Lutz Schmitt

Published in ChemBioChem

Impact score: 3.44

Own proportion of this work: 60 %

ChemBioChem
Research Article
doi.org/10.1002/cbic.202100702
Chemistry Europe
European Chemical Societies Publishing
www.chembiochem.org

Optimized Hemolysin Type 1 Secretion System in *Escherichia coli* by Directed Evolution of the Hly Enhancer Fragment and Including a Terminator Region

Zohreh Pourhassan N.,[a] Haiyang Cui,[b, c, f] Sakshi Khosa,[a] Mehdi D. Davari,[d] Karl-Erich Jaeger,[e] Sander H. J. Smits,[a] Ulrich Schwaneberg,[b, c] and Lutz Schmitt*[a]

Type 1 secretion systems (T1SS) have a relatively simple architecture compared to other classes of secretion systems and therefore, are attractive to be optimized by protein engineering. Here, we report a KnowVolution campaign for the hemolysin (Hly) enhancer fragment, an untranslated region upstream of the *hly*A gene, of the hemolysin T1SS of *Escherichia coli* to enhance its secretion efficiency. The best performing variant of the Hly enhancer fragment contained five nucleotide mutations at five positions (A30U, A36U, A54G, A81U, and A116U) resulted in a 2-fold increase in the secretion level of a model lipase fused to the secretion carrier HlyA1. Computational analysis suggested that altered affinity to the generated enhancer fragment towards the S1 ribosomal protein contributes to the enhanced secretion levels. Furthermore, we demonstrate that involving a native terminator region along with the generated Hly enhancer fragment increased the secretion levels of the Hly system up to 5-fold.

Introduction

Recombinant protein production at high titers in functional form and at reduced costs has been one of the central issues in biotechnology for decades.[1] *Escherichia coli* (*E. coli*) has been continuously implemented in recombinant protein production[2] and is often still the preferred production host in research laboratories[3] and used in industrial scales.[4,5] Recombinant protein secretion is beneficial over intracellular production strategies to meet biotechnological purposes because of various reasons such as simplifying downstream processes, reducing risk of cytosolic degradation, higher quality and solubility of secreted protein, avoidance of inclusion body formation, and reduced costs of production processes.[5–7]

So far, at least eight different secretion systems for Gram-negative bacteria, including *E. coli*, have been elucidated that secrete a wide range of transport-substrates.[7,8] Successful secretion of recombinant proteins by five of these secretion systems has been reported so far, but often with low titers and for a narrow range of substrates. Attempts to achieve reasonable secretion levels for recombinant protein secretion have been continuously made commonly with limited success, presumably due to the complexity of the multi-component secretion machineries.[6,7,9]

Type 1 secretion systems (T1SS) possess a relatively simple architecture consisting of only three membrane proteins, providing a secretory pathway for Gram-negative bacteria to secrete substrates in a single step and in an unfolded state.[10,11] The first and probably the best studied T1SS is the HlyA secretion system of uropathogenic *E. coli*.[10,12] The substrate of this system is HlyA, an enzyme of 110 kDa molecular weight which belongs to the RTX (repeats-in-toxins) family. Members of this family exhibit as a common feature a consensus glycine-rich motif (GGxGxDxUx), where x can be any amino acid and U refers to a large hydrophobic amino acid.[13]

The secretion signal of HlyA is located within the 50–60 C-terminal amino acid residues.[10] The secretion signal of HlyA along with three conserved glycine-rich motifs is known as HlyA1 which has a molecular weight of 24 kDa. HlyA1 has been employed as a carrier for secretion of heterologous proteins. Despite a long list of successfully secreted heterologous proteins, HlyA1 is still far away from being an universal secretion platform, due to a narrow range of possible substrates

[a] Z. Pourhassan N., S. Khosa, S. H. J. Smits, L. Schmitt
Institute of Biochemistry,
Heinrich Heine University
Universitätsstrasse 1, 40225 Düsseldorf (Germany)
E-mail: lutz.schmitt@hhu.de
[b] H. Cui, U. Schwaneberg
Institute of Biotechnology,
RWTH Aachen University
Worringerweg 3, 52074, Aachen (Germany)
[c] H. Cui, U. Schwaneberg
DWI-Leibniz Institute for Interactive Materials
Forckenbeckstraße 50, 52056 Aachen (Germany)
[d] M. D. Davari
Department of Bioorganic Chemistry
Leibniz Institute of Plant Biochemistry
Weinberg 3, 06120 Halle (Germany)
[e] K.-E. Jaeger
Institute of Molecular Enzyme Technology
Heinrich Heine University and Institute of Bio- and Geosciences IBG-1:
Biotechnology, Forschungszentrum Jülich GmbH
52426 Jülich (Germany)
[f] H. Cui
Present address: Carl R. Woese Institute for Genomic Biology
University of Illinois at Urbana-Champaign
1206 West Gregory Drive, Urbana, Il 61801 (USA)

ChemBioChem
Research Article
doi.org/10.1002/cbic.202100702
Chemistry Europe
European Chemical Societies Publishing

as well as low yields of secretion.[14–17] It should be noted that an engineered T1SS has recently been developed to allow efficient production of peptides and small proteins.[18]

The HlyA secretion system forms a tripartite double-membrane-spanning channel trough association of an inner membrane ATP-binding cassette transporter (HlyB), a membrane fusion protein (HlyD), and an outer membrane protein (TolC).[19] The next component of the HlyA system is HlyC. Though not essential for secretion, it acylates HlyA at two internal lysine residues (Lys564 and Lys690) prior to secretion. This modification turns the substrate into an active toxin which is able to lyse the membranes, for example, of human erythrocytes.[11,20]

The encoding genes of the Hly system, except of the *tol*C gene, are located within the *hly* operon (Figure 1).[10] Interestingly, a terminator region is located between the genes *hly*A and *hly*B, and therefore, the *hly* operon requires an anti-terminator factor RfaH to read through the terminator region. It was reported that the expression of *hly* operon is reduced in the absence of the RfaH factor and the expression of *hly*C and *hly*A are uncoupled from *hly*B and *hly*D genes.[21]

Recently, it was demonstrated that the presence of a 5′-untranslated region upstream of the *hly*A gene resulted in a multiple fold enhanced secretion level of heterologous proteins through the Hly system. This 5′-untranslated region, known as Hly enhancer fragment, is rich in uracil and adenine bases, and represents the C-terminal part of the *hly*C gene. Interestingly, the established vector harboring the Hly enhancer fragment enabled secretion of fast-folding fusion proteins to HlyA1 that could not be previously secreted.[20] We thus envisaged to subject this enhancer fragment to directed evolution attempting to optimize the secretion efficiency of the Hly system. It should be noted here that other parts of the Hly system might also be targeted by directed evolution when attempting to optimize this system.

Enhancer fragments, known to be involved in the translation initiation process, have an influence on the amount of proteins translated.[22–24] Protein translation is initiated by association of the ribosome to mRNA, which is considered the rate-limiting step in protein synthesis. This step of translation is facilitated by interaction of Shine-Dalgarno (SD)/anti-SD sequences as well as the S1 ribosomal protein/enhancer fragment.[23] Subsequently, the ribosome should dissociate from the initiation complex to slide over the mRNA for translating the downstream codons. It has been demonstrated that the presence of an enhancer fragment upstream of the SD sequence causes an increased dissociation rate and destabilizes the initiation complex.[23]

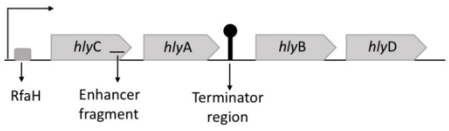


Figure 1. The *E. coli hly* operon. RfaH encodes the anti-terminator factor that is required for reading through the terminator region between *hly*A and *hly*B genes.

Until now only little attention has been paid to engineering of the Hly secretion system that may stem from lack of an elucidative structure and derived secretion mechanism. Directed evolution may represent a promising approach to engineer this system as it does not require any knowledge of the molecular structures to identify protein variants with optimized performance.[25–27]

The KnowVolution strategy was published in 2015 as an integrative protein engineering approach.[28,29] It combines directed evolution with computational analysis in order to obtain maximized improvements with minimized experimental efforts. A KnowVolution campaign consists of four phases: Phase (I) beneficial positions are identified through random mutagenesis and screening; Phase (II) potentially beneficial positions are subjected to site-saturation mutagenesis (SSM) in the wildtype enzyme to explore the full natural diversity and eliminate non-contributing positions. Sequencing of beneficial SSM libraries generates a molecular understanding of each amino acid exchange; Phase (III) comprises of a computationally assisted analysis in which beneficial amino acid substitutions are analyzed and grouped (amino acid substitutions that might interact with each other and independent ones). The latter analysis enables to further identify beneficial amino acid positions that interact with the identified ones. Clustered/interacting positions are recombined in the final phase to yield final variants with optimized performance. Various examples of successful KnowVolution campaigns have been published, *e.g.*, with glucose oxidases, phytases, proteases, cellulases, laccases, aryl sulfotransferases, or polymer-binding peptides.[7,30–34]

Here, we report on a KnowVolution campaign, according to the published procedure[35] with slight modifications, to optimize the 5′-untranslated region of the HlyA enhancer fragment. To this end, we have applied a high-throughput screening system, for the first time for directed evolution of the HlyA system, based on a lipase (lip).[36] Notably, all the previously published studies that have applied KnowVolution campaign have dealt with catalytic or adhesion promoting proteins, whereas in the current study an untranslated region, the Hly enhancer fragment, was subjected to a KnowVolution campaign.

Results

Secretion of lipase through the Hly secretion system using a one-plasmid system

A plasmid containing multiple genes essential for secretion via the Hly system was successfully constructed that included the regions encoding for the Hly enhancer fragment (untranslated region consisting of a 159 base pairs), the lipase fused to HlyA1, and the proteins HlyB and HlyD. The gene encoding TolC, the OM protein of the Hly secretion system, is endogenous and thus not present on the plasmid. Instead of a two-plasmid system, a one-plasmid system as constructed in this study was

ChemBioChem
Research Article
doi.org/10.1002/cbic.202100702
Chemistry Europe
European Chemical Societies Publishing

used because it results in a lower deviation between clones which generally should be avoided for directed evolution studies.

The constructed plasmid was introduced into chemically competent *E. coli* BL21-Gold (DE3) and the cells were grown on tributyrin agar plates. Halo formation around single cell clones indicated successful secretion of functional lipase-HlyA1 (Figure 2a). Cell clones harboring single plasmid either pSU2627-*EF*(159bp)-*lip*A-*hly*A1 or pK184-*hly*BD did not form halos (Figure 2b and 2c), indicating that lipase-HlyA1 secretion is dependent on the Hly secretion system. Test expression of clones secreting lipase-HlyA1 was performed in shaking flasks directly employing the supernatant of cultures (see SDS-PAGE analysis in Figure 2d).

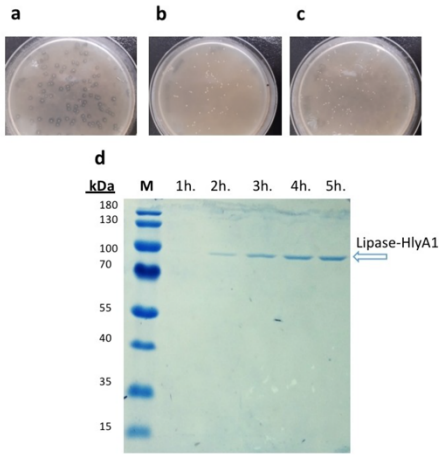


Figure 2. Secretion of lipase-HlyA1 through the Hly secretion system. Lipolytic activity of secreted lipase-HlyA1 was analyzed on tributyrin agar plates. Cells harboring the plasmid a) pK184_*EF-lip*A-*hly*A1BD, b) pSU2627_*EF*(159bp)-*lipA-hly*A1, and c) pK184_*hly*BD. d) SDS-PAGE analysis of supernatant (unconcentrated) of clones secreting lipase-HlyA1 through the Hly secretion system. M, marker proteins; the molecular weight of the marker proteins is given on the left; (xh.) unconcentrated supernatant of culture, where x denotes the number of hours after induction.

MTP-based screening system for secreted lipase

Lipase activity was determined through hydrolysis of *p*-Nitrophenyl butyrate (*p*NPB) to butyric acid and yellow *p*-nitrophenolate (Figure 3). The production of *p*-nitrophenolate was continuously monitored spectrophotometrically at 410 nm. The *p*NPB-based screening system in 96-well microtiter plates was reported previously.[37,38] Here, the amount of the secreted lipase was monitored by determination of lipase activity in supernatants of cultures grown for 16 h. The amount of *p*-nitrophenolate formed within 8 minutes of *p*NPB-lipase assay is correlated to the amount of secreted lipase in the supernatant.

Directed evolution of the Hly enhancer fragment through the KnowVolution strategy

Phase I (identification of beneficial positions): A random mutagenesis library of the Hly enhancer fragment was generated through random mutagenesis *in vitro*, employing error-prone PCR and a low-fidelity DNA polymerase in a buffer with 0.9 mM $MnCl_2$ resulting in an average of 8.5 mutations per enhancer fragment (159 bp). Hence, the calculated mutation load was 40.28 mutations per kb. The generated library with a size of 1408 clones was screened in a 96-well MTP format using the *p*NPB screening assay to identify possible beneficial variants. After screening of the random mutant library, rescreening of promising variants was repeated for five times. Promising variants showing an increased secretion level compared to the WT clone were selected and sequenced. Alignment of the enhancer fragment of the promising variants indicated that those variants harbored mutations repeatedly in 5 positions, including: +30, +36, +54, +81, +116.

Phase II and Phase III (determination and recombination): All five positions were subjected to site saturation mutagenesis to explore the natural full diversity. Notably, KnowVolution campaigns have been used so far for protein evolution only, whereas we have now used this evolution strategy to optimize the function of a DNA sequence, which in our case is not translated into a protein sequence. Saturation mutagenesis and recombination of possible beneficial positions were performed in parallel. Thereby, a library with a size of 1232 clones was generated at these 5 positions and screened using the *p*NPB screening system. The rescreening of promising variants was performed and led to the selection of seven clones that showed a 1.5 to 2-fold increase of secreted lipase-HlyA1 protein

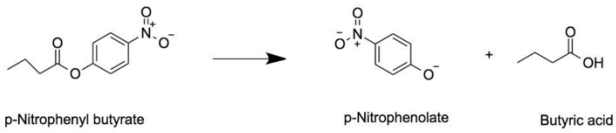


Figure 3. Hydrolysis of the substrate *p*NPB by secreted lipase. The production of *p*-nitrophenolate was monitored spectrophotometrically at 410 nm for 8 min. The figure was generated using the ChemDraw software.

ChemBioChem
Research Article
doi.org/10.1002/cbic.202100702
Chemistry Europe
European Chemical Societies Publishing

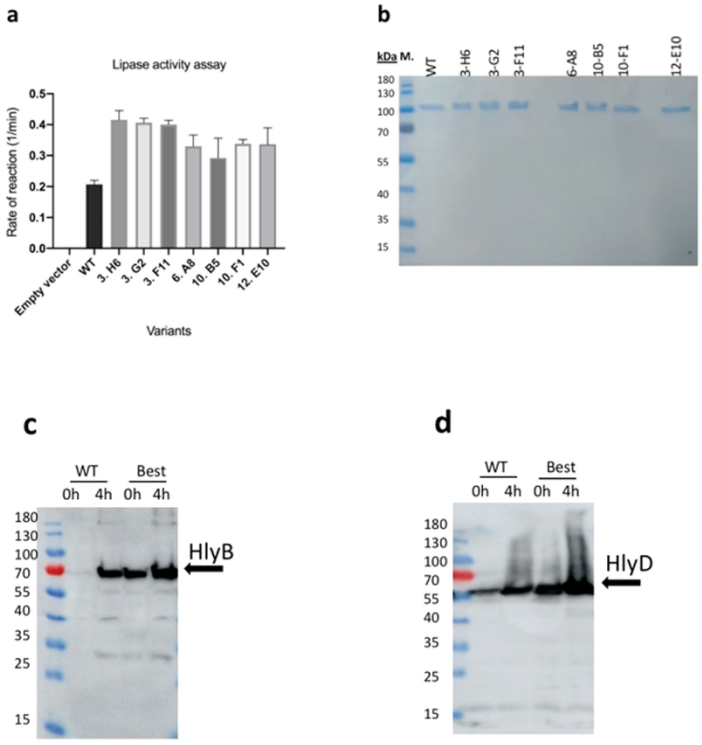


Figure 4. Secretion of lipase-HlyA1 through the Hly secretion system using enhancer fragment variants. a) Rate of lipase activity of the wildtype and the improved variants. The reported values and calculated error bars are based on at least three independent replicates. b) SDS-PAGE of unconcentrated supernatant of wildtype and the improved variants. Western blot analysis of *E. coli* cells demonstrated that the expression level of c) HlyB and d) HlyD were higher for cells having the "Best" enhancer fragment in comparison to the cells having WT enhancer fragment. The pLac promoter on plasmid pK184 is known to be leaky, therefore, HlyB and HlyD are expressed even before induction at 0 h.

compared to the WT (Figure 4a). The nucleotide mutations of each variant in the enhancer fragment are presented in Table 1.

Next, the secretion levels of lipase-HlyA1 of the seven improved variants and the WT clone were compared via test expressions in 100 mL Erlenmeyer flasks. The supernatant of the cultures (unconcentrated) was analyzed by SDS-PAGE (Figure 4b). Among the improved variants, variant 3-H6 seems to be the best as determined by either the *p*NPB lipase activity assay or test expression. The respective enhancer fragment was termed "Best" enhancer fragment and was used for further experiments. Furthermore, the level of secretion of HlyB (Figure 4c) and HlyD (Figure 4d) was analyzed by western blot for *E. coli* cells having either the WT or "Best" enhancer fragment. The analysis of western blot of three biological replicates confirmed that the expression level of both proteins was higher in cells having "Best" enhancer fragment as compared to cells having the WT one.

Table 1. Nucleotide mutations in enhancer fragment – untranslated region – of the improved variants.

Variants	Nucleotide and mutations
WT	A30, A36, A54, A81, A116
3-H6	A30U, A36U, A54G, A81U, A116U
10-F1	A30G, A36A, A54G, A81U, A116U
3-F11	A30U, A36A, A54G, A81U, A116C
6-A8	A30A, A36A, A54G, A81U, A116U
12-E10	A30U, A36U, A54G, A81U, A116U
3-G2	A30G, A36U, A54G, A81G, A116A
10-B5	A30C, A36C, A54G, A81U, A116A

RNA folding simulation and 3D structure prediction of the enhancer fragments

It is known that the enhancer fragment, upstream of the SD sequence, plays an important role in promoting protein trans-

ChemBioChem
Research Article
doi.org/10.1002/cbic.202100702
Chemistry Europe
European Chemical Societies Publishing

lation via interaction with the S1 ribosomal protein,[22,23] albeit not being translated into a protein sequence. We have previously reported that the HlyA enhancer fragment interacts with the S1 ribosomal protein.[20]

In order to explain the structural effects of the nucleotide mutations of the "Best" enhancer fragment and their consequences on the increased secretion level of lipase-HlyA1, the following computational analysis was performed. First, RNA folding simulation studies on both enhancer fragment variants (WT and "Best") were performed. As shown in Figure 5a, the "Best" enhancer showed a significant conformational change compared to WT, indicating that the introduced mutations affected mRNA folding. The structural angle change of the "Best" enhancer compared to WT is shown in Figure 5a. Subsequently, the modified mRNA's structure resulted in changes in docking pose upon interaction between S1 ribosomal protein and the enhancer fragment (Figure 5b and 5c). The WT/S1 ribosomal protein complex had a docking score of −7858708.37 while "Best"-S1 complex showed a reduced docking score of −6379871.3 suggesting that the binding interactions between the "Best" enhancer and S1 is not as strong as in the WT. Interestingly, three out of the five determined beneficial positions were located on the interface between ribosomal protein S1 and mRNA (*i.e.*, positions +30, +36, +81).

Ribosomal protein S1 binds with different affinities to the enhancer fragment variants

The above computational analysis indicated differences in interactions of the "Best" enhancer fragment with the S1 ribosomal protein compared to the WT fragment. This observation was experimentally tested by a RNA pull-down assay. RNA molecules of both "Best" and WT enhancer fragments were synthesized and labeled with 3′-Biotin TEG.

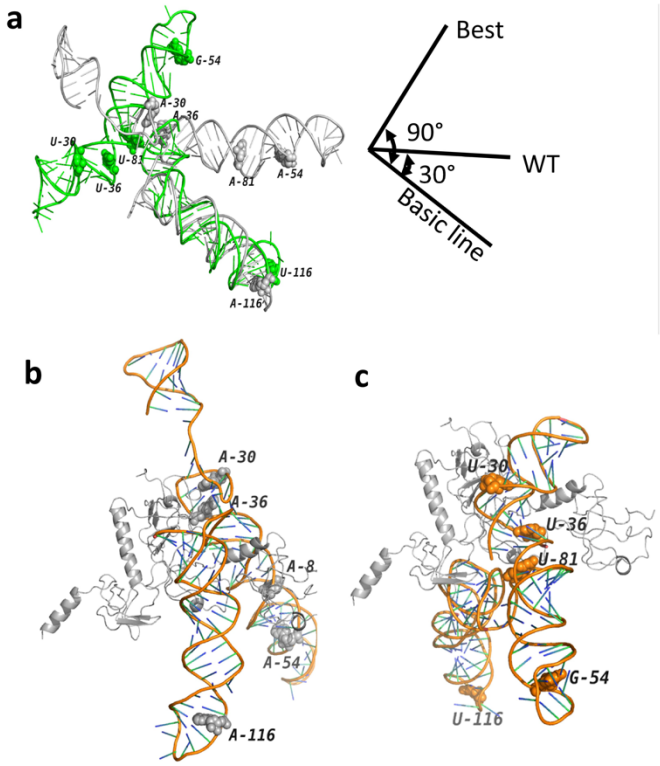


Figure 5. RNA folding simulation and docking analysis. a) The comparison between WT and the "Best" mRNA enhancer fragments. The "Best" enhancer is shown in green and WT enhancer is depicted in grey. The right panel shows a schematic diagram of the structural angle change of the "Best" enhancer compared to WT. 3D model of the interaction of S1 ribosomal protein with enhancer fragment b) WT and c) the "Best" variant. The S1 ribosomal protein is displayed as grey cartoon. The substituted nucleotides are highlighted with spheres in grey for WT and in green/orange for the "Best" enhancer. The docking pose was obtained from the NP-dock molecular docking server.

ChemBioChem
Research Article
doi.org/10.1002/cbic.202100702
Chemistry Europe
European Chemical Societies Publishing

Next, both enhancer fragment variants were immobilized on magnetic beads and thereafter, incubated with *E. coli* cell extract. The amount of bound protein on the beads carrying either the "Best" enhancer fragment or the WT enhancer fragment was analyzed by SDS-PAGE (Figure 6a). The two strong pull-down bands around 70 and 100 kDa on the SDS-gel were cut out and the proteins were identified by liquid chromatography-tandem mass spectrometry. The dedicated band around 100 kDa was identified as aldehyde-alcohol dehydrogenase from *E. coli*, while the band around 70 kDa with 58% coverage of peptide sequences was identified as *E. coli* S1 ribosomal protein. The semi-quantification of SDS-PAGE using the ImageJ indicated that in comparison to the RNA molecule of WT enhancer fragment, about 50% less S1 ribosomal protein was pulled down by the RNA molecule of the "Best" enhancer fragment (Figure 6b). This data supports the interpretation of the computational analysis suggesting that the interaction of S1 ribosomal protein with the "Best" enhancer fragment is weaker than with the WT fragment. These findings suggest that in the presence of the "Best" enhancer fragment, the translation of lipase-HlyA1 might increase which leads to a higher level of lipase secretion.

Enhanced secretion of cutinase-HlyA1

To test whether the "Best" Hly enhancer fragment could also improve the secretion of other fusion proteins, a fusion between the gene encoding cutinase from *Fusarium solani pisi*[39] was constructed either with the "Best" or WT enhancer fragment. The test expression for both constructs was conducted and the supernatant was used for western blot analysis as well as *p*NPB hydrolysis assay. Western blot analysis of the supernatant revealed that the "Best" enhancer fragment was also able to promote the secretion of the cutinase-HlyA1 (Figure 7a). Furthermore, the *p*NPB assay confirmed more than 2-fold secretion of cutinase as compared to the WT (Figure 7b).

This observation indicates that the "Best" enhancer fragment variant increases the secretion efficiency of the Hly secretion system not only for lipase-HlyA1 but also for the target protein, cutinase-HlyA1.

Introduction of a terminator region between *hly*A1 and *hly*B genes

Given the ability of the "Best" enhancer fragment to improve the secretion of two tested HlyA1 fusion proteins, *i.e.*, lipase and cutinase, we also investigated whether employing the Hly terminator region between *hly*A1 and *hly*B genes would further improve the secretion of the Hly secretion system. Previous studies have demonstrated that the anti-terminator factor RfaH is required for reading through the terminator signal located between *hly*A and *hly*B genes in the *hly* operon and increases the transcriptional elongation as well as the steady-state level of *hly* mRNA.[21]

In this study, the *hly* terminator, a stem loop between *hly*A1 and *hly*B gene, was introduced on the plasmids encoding either lipase or cutinase and along with the "Best" enhancer fragment. The presence of the terminator resulted in enhanced secretion of both target hydrolases. A higher level of lipase-HlyA1 secretion was achieved in the strain carrying the plasmid containing the "Best" enhancer fragment along with the terminator region compared to the strain having either the WT plasmid or the "Best" plasmid only (Figure 8a). Estimation of the amount of secreted protein by SDS-PAGE showed an increase of more than 3.5-fold for lipase-HlyA1 with the newly constructed plasmid. In the presence of the "Best" enhancer fragment and terminator region, successive increase in the amount of secreted lipase-HlyA1 in the supernatant was observed. This increase only after four hours of induction led to formation of enzyme oligomers (observed as white particles) in the supernatant culture, which mostly contained lipase-HlyA1. Based on the known concentration of purified HlyA, concentration of secreted lipase-HlyA1 in the supernatant was

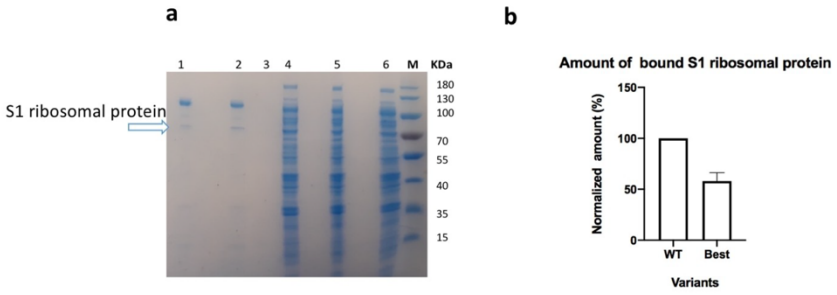


Figure 6. RNA pull-down assay. **a**) SDS-PAGE of eluted proteins by immobilized RNA molecules of the "Best" enhancer fragment (1), WT enhancer fragment (2), without RNA (3) and washing flow through of washing of the beads "Best" (4), WT (5), and without RNA (6). The position of the S1 ribosomal protein is indicated by an arrow. **b**) Semi quantification of SDS-PAGE via ImageJ software. The reported values and calculated error bars are based on three independent replicates. In each replicate, the reported value of the "Best" enhancer fragment was normalized based on the value of the WT.

ChemBioChem
Research Article
doi.org/10.1002/cbic.202100702
Chemistry Europe
European Chemical Societies Publishing

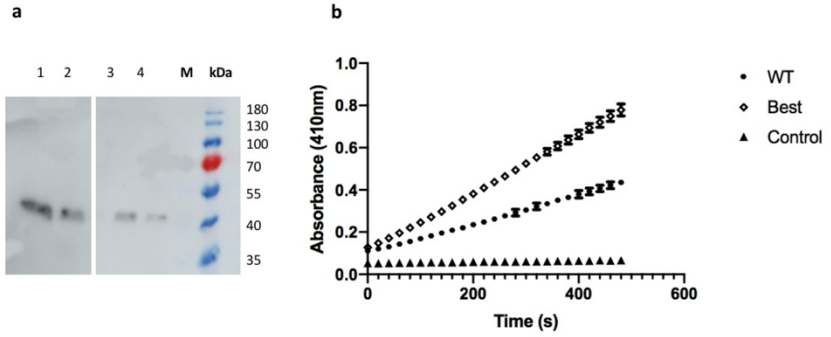


Figure 7. Enhanced secretion of cutinase-HlyA1 by the "Best" enhancer fragment. a) Western blot analysis of supernatant fractions from the "Best" enhancer fragment (1 and 2) and the WT (3 and 4) cultures. b) Activity of the cutinase-HlyA1 in the supernatant of the culture dedicated to either the "Best" or WT enhancer fragment using the *p*NPB assay. The reported absorbance at 410 nm and the error bars correspond to average values of three replicates. The control is the absorbance of the buffer and substrate without addition of culture supernatant.

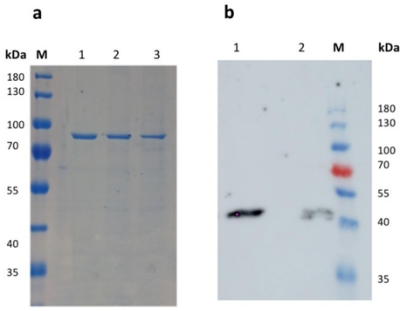


Figure 8. Influence of the HlyA terminator on the amount of secreted proteins. a) Secreted lipase-HlyA1; SDS-PAGE of supernatant fractions from terminator with the "Best" enhancer fragment (1), the "Best" enhancer fragment only (2), and the WT (3) strains; stained with Coomassie Brilliant Blue G-250. b) Secreted cutinase-HlyA1; western blot analysis of supernatant fractions from terminator along with the "Best" enhancer fragment (1) and the WT (2) strains.

quantified as about 180 mg per litter (corresponding to 60 mg/OD_{600}). Furthermore, the amount of lipase-HlyA1 (wet weight) in collected white particles from 25 mL culture was estimated to be above 1.5 mg (~60 mg/L).

In case of cutinase, Western blot analysis revealed that the presence of the terminator region along with the "Best" enhancer fragment secreted significantly higher amounts of cutinase-HlyA1 in the culture medium compared to the WT plasmid (Figure 8b). Semi quantification of the Western blot analysis with ImageJ software estimated an increase of around 5-fold increase for the cutinase-HlyA1 when expressed using the newly constructed plasmid.

Discussion

For heterologous protein production, secretion of proteins is beneficial compared to cytosolic expression, especially in terms of costs and the quality of produced proteins. Secretion offers various advantages for protein production, such as higher solubility and stability of target protein, reduced risk of protein degradation by cytosolic proteases, and easy purification process.

Although, *E. coli* is the most preferred producer host for research laboratories,[3] its potential uses as a secretory host is limited since the expression of recombinant proteins mostly takes place either in the cytoplasm or periplasm and if in the supernatant, in low titers. For secreted proteins, *Bacillus* species are the expression hosts of choice with a secreted product yield of 20 g/L.[4,6,9,40] In this regard, investigating the secretion of recombinant proteins by *E. coli* is a continuous topic of interest and therefore, tremendous efforts have been made to investigate further improvements in the secretion efficiencies of *E. coli* to make it more competitive, as both laboratory and industrial strain.[6,9,41,42]

Among the secretion systems of Gram-negative bacteria including *E. coli*, T1SS, T2SS, T3SS, T5SS, and T8SS have been used for biotechnological purposes.[6] Among them, T1SS has the most minimalistic architecture and is therefore an attractive target for further improvements.

This study set out with the aim of improving the secretion efficiency of the Hly secretion system of *E. coli*. As a member of the T1SS, the Hly system consists of only three compartment proteins[10] that makes the plasmid-expression of this system in the production hosts easy.[6] Recently, a 5′-untranslated region upstream of the *hly*A gene, known as the Hly enhancer fragment, was identified that was able to increase the secretion level of HlyA1 fusion proteins by several folds.[20] In the current study, this Hly enhancer fragment was subjected to optimization by directed evolution using a KnowVolution campaign.[28,31]

ChemBioChem
Research Article
doi.org/10.1002/cbic.202100702
Chemistry Europe
European Chemical Societies Publishing

Noteworthy, this study is the first case study that applies a KnowVolution campaign for evolution of an untranslated DNA region for further improvement of the HlyA secretion system of *E. coli*.

To initiate the translation in prokaryotic cells, the first and key step is binding the ribosome to the 5′-untranslated region of mRNA, which is followed by ribosome dissociation from the initiation complex to move to downstream codons of the mRNA for completion of translation. It has been suggested that the S1 ribosomal protein, the largest ribosomal protein of *E. coli* with a molecular weight of 68 kDa, interacts with mRNA within the enhancer region, upstream of the SD sequence, by an unknown mechanism.[24,43,44] In an elegant set of experiments, it was demonstrated that the transition step from translation initiation to elongation takes effect by both interaction of SD/anti-SD sequences and enhancer/S1 by a negative allosteric effect. It is notable that the ribosome dissociation from initiation complex is the determining factor for translation efficiency.[23]

Through KnowVolution campaign, two libraries with sizes of 1408 and 1232 clones, constructed by epPCR and site saturation mutagenesis, respectively, were screened. Screening a library size of 1000 to 2000 clones in which 40–60% of the generated enzymes variants are still active are commonly and successfully used in standard directed evolution campaigns.[45] Iterative rounds of screening of a few thousand clones have proved to yield superior results than screening twenty to fifty thousand of clones since the number of "new" substitutions decreases rapidly with increased library size.[46] The latter can be attributed to the strong mutational bias of polymerases.[47] Therefore, superior strategies have been developed, *e.g.* iterative rounds of screening of small libraries (1000–2000 clones) with varied mutational biases, flow cytometry, and microfluidic screening systems that enable to screen millions of variants, and combined computational /experimental strategies which analyze and identify further beneficial positions based on the identified ones in the initial round of directed evolution (implemented if required in Phase III of a KnowVolution campaign).

All the mentioned finding are on a solid ground for enzyme properties such as activity, organic solvent resistance, pH or ionic liquid resistance, *etc.* In respect to expression/translocation improvement no datasets are available to link potential improvement possibilities to the screened library size. Therefore standard conditions of screening (number of clones) have been applied and combined in a KnowVolution campaign to identify key residues and to generate a molecular understanding and nature of mutations at identified key positions. In summary, the molecular understanding will be broadened through iterative rounds of screening of small libraries in the future, since alternatives, such as a beneficial screening that enables to screen millions of variants, are not available.

After screening of libraries, five beneficial positions (+30, +36, +54, +81, and +116) were identified in the Hly enhancer fragment. Mutations at these positions boosted the secretion levels of lipase and cutinase through the Hly system by 100%. It was somewhat surprising that adenine nucleotides were present at all the five identified positions and were randomly substituted to uracil in the best performing enhancer variant, named as the "Best" enhancer fragment. It was already reported that uracil nucleotides of the enhancer fragment are beneficial for its interaction with the S1 ribosomal protein.[23] Therefore, we reasoned that the flexibility of the interaction between S1 ribosomal protein and the enhancer fragment is increased in the "Best" enhancer fragment. Most likely, the presence of uracil at these positions may destabilize the initiation complex and increase the dissociation rate between S1 and mRNA. This was further supported by molecular docking studies and an RNA pull-down assay.

Furthermore, the results obtained from the secretion of cutinase from *Fusarium solani pisi* in the presence of the "Best" enhancer fragment indicated that this novel enhancer fragment is compliant with other secretory proteins in addition to lipase-HlyA1. Notably, it is the first time that cutinase protein from *Fusarium solani pisi* was tested as a secretory protein in the Hly secretion system, while lipase had been tested before.[20] It also suggested a general applicability of the "Best" enhancer fragment.

Kondo *et al.* six described six genetics elements of *E. coli* which contribute to an increased gene expression levels and result in higher yields of the desired protein, namely: promoter, translation enhancer, Shine-Dalgarno sequence, spacer, encoding gene, and terminator of the gene.[48] Our results additionally demonstrated that the presence of a terminator region along with the "Best" enhancer fragment can result in further improvement of secretion through the Hly secretion system. In *E. coli* cells, the anti-terminator factor RfaH reads through terminator regions. It was reported that RNA polymerase stops at so-called JUMPstart (just upstream of many polysaccharide-associated starts) sequences, where RfaH binds to RNA polymerase forming a complex able to read through the terminator region.[20]

Conclusions

This study demonstrated that the secretion through the Hly secretion system can be improved significantly by engineering of genetic elements. *E. coli* cells expressing the optimized HlyA secretion system constructed in this study showed a secretion level of around 180 mg per liter for soluble lipase-HlyA1. Furthermore, a high protein content of above 1.5 mg of lipase-HlyA1 was estimated for the white particles in only 25 mL of culture. Additionally, the one-plasmid system developed in this study might offer an interesting alternative to two-vector systems, which have been used so far. This would also significantly reduce the costs of antibiotics when cells are cultivated at larger scale.

In addition, the employed evolutionary and genetic engineering strategies can generally be applied to all types of fusion proteins if required and provide an opportunity for optimizing the secretion titers by evolutionary approaches. Eventually, our approach is also applicable to those secretion systems that have been used in the past for biotechnological approaches

ChemBioChem
Research Article
doi.org/10.1002/cbic.202100702
Chemistry Europe
European Chemical Societies Publishing

such as T2SS, T3SS, T5SS and T8SS, to significantly improve the *E. coli* secretion performance, thereby rendering *E. coli* a more competitive enzyme secreting host that might challenge the industrial applied host systems in the future.

Experimental Section

Materials

All chemicals used in this study were purchased from Sigma-Aldrich, Roche Diagnostics GmbH, or Applichem GmbH if not stated otherwise. Enzymes were purchased from New England Biolabs, except of polymutarase polymerase, which was provided from the SeSaM-Biotech (Aachen, Germany). All oligonucleotides were purchased from Eurofins MWG Operon and diluted in MilliQ-water. The commercially available kits NucleoSpin plasmid miniprep kit and PCR clean-up kit were purchased from Macherey Nagel (Dueren, Germany). DNA sequencing was done at either Eurofins MWG-Operon (Ebersberg, Germany) or Microsynth Seqlab (Göttingen, Germany). The Gibson assembly master mix kit as well as the Q5 site-directed mutagenesis kit were purchased from New England Biolabs.

Construction of a one-plasmid system for secretion

For cloning and library generation either *E. coli* DH5α or *E. coli* BL21-Gold (DE3) were used. Plasmids and oligonucleotides used in this work are listed in Table 2 and Table 3, respectively. For all plasmids constructed in this study, pK184_*hly*BD was used as the backbone vector. The *lac* promoter of plasmid pK184 is inducible with isopropyl β-D-1-thiogalactopyranoside (IPTG).

Table 2. Plasmids used in this study.

Backbone plasmid	Contained genes/fragments	Source
pK184	*hly*BD	[20]
pSU2726	*EF** (159bp)-*lipA-hly*A1	[20]
pK184	*EF-lipA-hly*A1BD	this study
pK184	*EF*(Best)-*lipA-hly*A1BD	this study
pK184	*EF*(Best)-*lipA-hly*A1-*Ter-hly*BD	this study
pEKEx2	NprE-*cutinase*	[39]
pK184	*EF-cutinase-hly*A1BD	this study
pK184	*EF*(Best)-*cutinase-hly*A1BD	this study
pK184	*EF*(Best)-*cutinase-hly*A1-*Ter*^-*hly*BD	this study

The one-plasmid system for the secretion of the *lip*A gene from *Serratia marcescens*[36] was constructed as follows: plasmid pSU2726_*hly*C(159bp)-*lipase-hly*A1 was used for the PCR amplification of *EF* (enhancer fragment 159bp)-*lipA-hly*A1 fragment using the primer sets P1 and P2, both containing an overhang region towards the backbone plasmid. The backbone plasmid, pK184_*hly*BD, was linearized using the primer sets P3 and P4. Subsequently, both PCR products were digested with DpnI enzyme (20 U, 37 °C, overnight). Then, PCR products were purified with the PCR clean-up kit. The *EF-lipA-hly*A1 fragment was assembled into the pK184_*hly*BD vector using the Gibson assembly kit, according to the instructions of the manufacturer. Subsequently, 2 µl of the assembled product was transformed into chemically competent cells of *E. coli* DH5α. The sequence of the constructed plasmid pK184_*EF-lipA-hly*A1BD was confirmed via plasmid sequencing.

Generation of enhancer fragment mutant library by epPCR

Random mutagenesis of the enhancer fragment, with a size of 159 base pairs, was performed via error prone PCR with a high mutational load.[46] To obtain the mutations in the first nucleotides of the enhancer fragment, the forward primer was designed to bind behind the *lac* promoter but upstream of the starting nucleotide of the enhancer fragment. The enhancer fragment was amplified using the primer sets P9 and P10. The epPCR mixture (50 µL) contained: 50 ng of plasmid template pK184_*EF-lipA-hly*A1BD, 1× ThermoPol buffer, 0.2 mM dNTP mix, 400 pM of each primer, 10 U Polymutarase DNA polymerase, and 0.9 mM $MnCl_2$.

The vector template pK184_*EF-lipA-hly*A1BD lacking the *EF* region was amplified using a high-fidelity polymerase, Q5 polymerase, and primer sets P11 and P12. The PCR mixture (50 µL) contained: 50 ng of plasmid pK184_*EF-lipA-hly*A1BD plasmid, 1×Q5 standard reaction buffer, 0.2 mM dNTP mix, 400 pM of each primer, and 1 U Q5 DNA Polymerase. Subsequently, the ep-PCR product of the enhancer fragment and the PCR product of the backbone vector were digested with DpnI enzyme, overnight, at 37 °C to remove any wildtype vector. The PCR products were purified using PCR clean-up kit. The whole plasmid construction of the epPCR product and the linear plasmid was performed via Gibson assembly according to the instructions of the manufacturer. Subsequently, 2 µL of the assembled product was transformed into chemically competent cells of *E. coli* BL21-Gold (DE3). The transformed cells were cultured directly on tributyrin lysogeny broth (LB) agar plates (tryptone 10 g/L, yeast extract 15 g/L, NaCl 5 g/L, Agar 15 g/L, Gum Arabic 1.5 g/L, tributyrin 15 g/L) containing kanamycin 50 µg/mL. Ten single cell clones of epPCR library were picked and sent for sequencing.

Table 3. Oligonucleotides used in this study.

Oligonucleotides	Sequence
P1	GTGAGCGGATAACAATTTGATTACGAATTCGAGC
P2	GACTGTTTCCTGTGTGATTATGCTGATGTGGTCAG
P3	TCACACAGGAAACAGTCATG
P4	AATTGTTATCCGCTCACAATTC
P5	GGAAATTCTCTTGCAAAAAATG
P6	CATTTAATTACCTCTTAACCAG
P7	GGTTAAGAGGTAATTAAATGGCGCCTACTAGTAAC
P8	TTTTTTGCAAGAGAATTTCCAGCAGAACCACGGAC
P9	AATTGTGAGCGGATAACAATTTG
P10	CTAAAGATGCCCATTTAATTACC
P11	TAAATGGGCATCTTTAGCTATAAG
P12	GTTATCCGCTCACAATTCCAC
P13	GTGCCACATAAGATTGCTATTTTTTTGGAGTCATAATGGCGAATTCTGATTCTTG
P14	AGCCCAGTAAGATTGCTATCATTTAAATTAATATATTATGCTGATGTGGTCAG

ChemBioChem
Research Article
doi.org/10.1002/cbic.202100702
Chemistry Europe
European Chemical Societies Publishing

Cultivation in 96-well microtiter plates

For the cultivation of mutants from libraries, single clones showing lipase activity on tributyrin LB agar plates were cultured into 150 μL LB_{Kan} in 96-well MTP (Microtiter plate, PS-F-bottom, Sarstedt, Germany). In each 96-well MTP plate, 4 wildtype clones and 4 empty vector clones were included as positive and negative clones, respectively. The plates were incubated for 24 hours (37 °C, 900 rpm, and 80% humidity) in a 96-well MTP shaker (SI505 MTP shaker incubator, Avantor). After cultivation, 50 μL of 50% glycerol (v/v) was added to each well and the plates, so called master plates, were stored at −80 °C until further experiments. For the expression of lipase-HlyA1, the master plates were used to inoculate the pre-culture in MTPs containing 150 μL LB_{Kan} and cells were cultivated overnight (37 °C, 900 rpm, and 80% humidity). Subsequently, 2 μL of the pre-culture were used to inoculated the main culture MTPs (150 μL LB_{Kan}) and the cells were grown for 3 hours (37 °C, 900 rpm, and 80% humidity) before induction with 1 mM of IPTG and 5 mM $CaCl_2$. The cultures were further grown for 16 hours (37 °C, 900 rpm, and 80% humidity) in 96-well MTP shaker incubator. Afterwards, the MTPs were centrifuged (4 °C, 20 min, 3500 rpm) and the supernatants were harvested and used for the screening assay.

Site-saturation mutagenesis

Individual site-saturation mutagenesis at positions 30, 36, 54, 81, and 116 of the Hly enhancer fragment was performed to obtain all the possible mutations at these positions. Two reverse complementary oligonucleotides carrying degenerative nucleotides in all individual positions were designed and synthesized. Annealing of the oligos was performed according to the protocol mentioned by Sigma-Aldrich as follows. In short, oligos were dissolved in the annealing buffer containing 10 mM Tris pH 7.5, 50 mM NaCl, and 1 mM EDTA. Equal volumes of the equimolar oligonucleotides of both forward and reverse oligos were mixed and heated up to 95 °C for 2 min. Subsequently, the mixture was gradually cooled to 25 °C over a period of 45 min. The annealed oligos were used as mega primers in a Megawhop PCR in order to amplify the whole plasmid. Megawhop PCR mix (100 μL) contained: 0.3 mM dNTP, 500 ng annealed oligos, 2 U Q5 polymerase, 70 ng of the plasmid template pK184_*EF-lip*A-*hly*A1BD, and 1× Q5 buffer. The PCR product was digested with DpnI enzyme, overnight, at 37 °C and purified using PCR clean-up kit. Subsequently, 2 μL of purified Megawhop PCR product was transformed into chemically competent cells of *E. coli* BL21-Gold (DE3). The transformed cells were cultured on tributyrin LB agar plates. Single cell clones were picked up and grown in 96-well MTP plates.

Colorimetric screening system

The substrate used in this study to detect the amount of lipase-HlyA1 secretion was *p*-nitrophenyl butyrate (*p*NPB).[49] Upon hydrolysis by lipase, *p*-nitrophenolate is formed, which can be detected spectrophotometrically at 410 nm. After cultivation and expression of the libraries, 10 μL supernatant of each single culture was transferred into a new 96-well MTP containing 90 μL of 50 mM TEA buffer (pH 7.4). The reaction was started by supplementing 100 μL of freshly prepared substrate solution containing 0.5 mM *p*NPB (dissolved in 10% v/v acetonitrile) in 50 mM TEA buffer (pH 7.4) in a final reaction volume of 200 μL. The amount of released *p*-nitrophenolate was recorded by measuring A_{410nm} at 30 °C for 8 min with a kinetic interval of 20 seconds in a microtiter plate reader (Tecan infinite® M200 Pro Austria GmbH, Männedorf, Switzerland). The rate of conversion of *p*NPB to *p*-nitrophenolate (the differences of A_{410nm} within 8 minutes) was calculated for each reaction and used to identify clones with enhanced secretion of lipase-HlyA1 into the supernatant.

Construction of one-plasmid system for secretion

The one-plasmid system for secretion of cutinase from *Fusarium solani pisi* was constructed as follows: the plasmid vector pK184_*EF-lip*A-*hly*A1BD was linearized, excepting the *lip*A gene, using the primer sets P5 and P6. The *cutinase* gene from plasmid pEKEx2_NprE-*cutinase* was amplified using the primer sets P7 and P8. Both PCR products were digested with DpnI enzyme overnight at 37 °C and purified using PCR clean-up kit. The *cutinase* gene was assembled into the linearized plasmid pK184_*EF-hly*A1BD via Gibson assembly, according to the manufacture's protocol. 2μL of the assembled reaction was transformed into *E. coli* DH5α competent cells and cultured on LB_{Kan} agar plates. Cloning of the plasmid pK184_*EF-cutinase-hly*A1BD was verified via sequence analysis.

Plasmid pK184_*EF*(Best)-*cutinase-hly*A1BD was constructed in similar manner as it was already mentioned above, using plasmid pK184_*EF*(Best)-*lip*A-*hly*A1BD as template.

To insert a terminator element between the *hly*A1 and *hly*B genes on the above constructed one-plasmid systems for secretion, the sequence of this terminator element for the *hly* operon of uropathogenic *E. coli* UTI89 was obtained from the National Center for Biotechnology Information (NCBI). Two non-overlapping primers P13 and P14 were designed based on the sequence of the terminator. The plasmids pK184_*EF*(Best)-*lipase-hly*A1BD and pK184_*EF*(Best)-*cutinase-hly*A1BD were used as backbone templates. The PCR amplification of the plasmids was performed using primer sets P13 and P14. The PCR products were separately incubated (5 min, room temperature) in the reaction buffer of the Q5-site directed mutagenesis kit based on the manufacture's instruction (New England Labs, England). Subsequently, 2 μL of the resultant mixture was transformed into 50 μL of chemically competent cells of *E. coli* DH5α. Sequencing data of single cell clones confirmed the insertion of the terminator sequence between the *hly*A1 and *hly*B gene on the plasmids. The sequencing data of single cell clones confirmed the insertion of the terminator sequence between *hly*A1 and *hly*B genes on the plasmids and thereby, successful construction of plasmids pK184_*EF*(Best)-*lip*A-*hly*A1-*Ter-hly*BD and pK184_*EF*(Best)-*cutinase-hly*A1-*Ter-hly*BD.

RNA pull-down assay

The 3'-Biotin TEG (triethylene glycol)-labeled RNA of wildtype enhancer fragment (5'-GAAUUUCAUG GAGGUAAAAU UGAUAAA-CAG UUAGCGAAUA AAAUUUUUAA ACAAUAUCAC CACGAGUUAA UAACUGAAGU AAAAAGAAAG-3') and the "Best" enhancer fragment (5'-GUAUUUCUUG GAGGUAAAAU UGAUAGACAG UUAGCGAAUA AAAUUUUUAA ACUAUAUCAC CACGAGUUAA UAACUGAAGU AAAAAGAUAG-3') were synthesized by BioSpring (Germany). The RNA pull-down assay was performed using the Thermo Scientific Pierce magnetic RNA-protein pull-down kit (ThermoFisher Scientific, US) according to the manufacturer's instructions. Eluted proteins from the magnetic beads were run on a SDS-PAGE. Two SDS-PAGE bands depicting the bound proteins were extracted from the gel and analyzed by liquid chromatography-tandem mass spectrometry.

Expression and secretion experiments in shaker flasks

E. coli BL21-Gold (DE3) competent cells were transformed with the desired plasmids and grown on LB agar plates supplemented with

ChemBioChem | Research Article doi.org/10.1002/cbic.202100702 | Chemistry Europe European Chemical Societies Publishing

50 μg/mL kanamycin. A single colony was used to prepare a pre-culture and cultivated overnight (37 °C and 180 rpm). The overnight culture was used to inoculate 25 mL of 2YT medium supplemented with 50 μg/mL kanamycin at an OD_{600} of 0.1 in 100 mL Erlenmeyer shaking flask. The cultures were cultivated at 37 °C and 180 rpm to an OD_{600} of 0.4 to 0.6. Subsequently, the expression was induced with 1 mM IPTG and 5 mM $CaCl_2$. The culture was harvested after specific hours of expression by centrifugation and was analyzed either by SDS-PAGE[50] or Western blot analysis. Staining of the SDS-gels was performed via Colloidal Coomassie G-250 Staining protocol.[51] The intensity of the protein bands on the SDS-gels were quantified using ImageJ software (Image Processing and Analysis in Java). To determine the amount of secreted lipase-HlyA1, a series of purified HlyA solutions with known concentrations were also loaded on same SDS-PAGE. The amount of secreted lipase-HlyA1 of the supernatant was then compared to the HlyA solution through a calibration curve. Furthermore, to determine the amount of lipase-HlyA1 protein in the white particles of the 25 mL culture, those particles were isolated after four hours of induction from the supernatant and resuspended in water and SDS-sample buffer. The amount of protein was estimated as mentioned for lipase-HlyA1 of the supernatant.

Computational methods

The crystal structure of the S1 ribosomal protein (PDB ID: 6h4n, chain Y) was taken from PDB bank (http://www.rcsb.org).[52] The online webserver SimRNA (https://genesilico.pl/SimRNAweb) was used to construct the 3D structure of the wild-type (WT) and "Best" enhancer fragments.[53] The RNA sequence of enhancer fragments in a version with the residues A, C, G, and U only was used in Fasta format for the SimRNA modeling as described in the SimRNA user manual. The number of iterations was set to 500 steps of simulations with default parameters. After the simulation steps, 1% of the frames with the lowest energy were selected for clustering while the remaining frames were discarded. To compare RNA 3D structure of both fragment variants, their predicted structures were aligned using the RNA-align open-source algorithm (https://zhanglab.ccmb.med.umich.edu/RNA-align).[54]

To model the plausible S1 ribosomal protein-Enhancer fragment interaction, docking of the enhancer fragments (WT, and the "Best") with the S1 ribosomal protein were performed using Nucleic Acid-Protein Dock (NPDock) web server (http://genesilico.pl/NPDock).[55] Docking was carried out using the default parameters (1000 steps of simulation, initial temperature of 15,000 K). In total, 100 best scored models were used for clustering similar structures to obtain the final optimized structure. RMSD cut-off of 5 Å was used for clustering.

Acknowledgment

This study was funded by the CLIB Competence Center Biotechnology (CKB) funded by the European Regional Development Fund ERDF (34.EFRE-0300096) to L.S. and in part by Deutsche Forschungsgemeinschaft (CRC 1208 project A01 to L.S.). We thank all members of the Institute of Biochemistry for fruitful discussions. Open Access funding enabled and organized by Projekt DEAL.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Keywords: directed evolution · Hly enhancer fragment · improved secretion efficiency · protein · protein secretion

[1] J. D. Linton, W. Xu, *Trends Biotechnol.* **2021**, *39*, 107–110.
[2] J. Puetz, F. M. Wurm, *Processes* **2019**, *7*, 476.
[3] J. E. Cronan, *eLS* **2014**, https://doi.org/10.1002/9780470015902.a0002026.pub2.
[4] G. L. Rosano, E. A. Ceccarelli, *Front. Microbiol.* **2014**, *5*, 172–172.
[5] M. Theisen, J. C. Liao, *Industrial Biotechnology: Microorganisms* **2017**, *1*, 149–181.
[6] L. A. Burdette, S. A. Leach, H. T. Wong, D. Tullman-Ercek, *Microbial Cell Factories* **2018**, *17*, 196.
[7] Z. Pourhassan, S. H. Smits, J. H. Ahn, L. Schmitt, *Biotechnol. Adv.* **2021**, *53*, 107864.
[8] T. R. Costa, C. Felisberto-Rodrigues, A. Meir, M. S. Prevost, A. Redzej, M. Trokter, G. Waksman, *Nature Rev. Microbiol.* **2015**, *13*, 343–359.
[9] G. R. Kleiner-Grote, J. M. Risse, K. Friehs, *Eng. Life Sci.* **2018**, *18*, 532–550.
[10] I. B. Holland, S. Peherstorfer, K. Kanonenberg, M. Lenders, S. Reimann, L. Schmitt, *EcoSal Plus* **2016**, *7*.
[11] K. Kanonenberg, O. Spitz, I. N. Erenburg, T. Beer, L. Schmitt, *FEMS Microbiol. Lett.* **2018**, *365*, fny094.
[12] A. Noegel, U. Rdest, W. Springer, W. Goebel, *Mol. Gen. Genet.* **1979**, *175*, 343–350.
[13] O. Spitz, I. N. Erenburg, T. Beer, K. Kanonenberg, I. B. Holland, L. Schmitt, *Prot. Secr. Bact.* **2019**, *7*, 215–225.
[14] C. K. Schwarz, C. D. Landsberg, M. H. Lenders, S. H. Smits, L. Schmitt, *J. Biotechnol.* **2012**, *159*, 155–161.
[15] D. Ruano-Gallego, S. Fraile, C. Gutierrez, L. Á. Fernández, *Microb. Cell Fact.* **2019**, *18*, 47.
[16] I. Gentschev, G. Dietrich, W. Goebel, *Trends Microbiol.* **2002**, *10*, 39–45.
[17] L. A. Fernández, I. Sola, L. Enjuanes, V. De Lorenzo, *Appl. Environ. Microbiol.* **2000**, *66*, 5024–5029.
[18] B.-N. Nguyen, F. Tieves, T. Rohr, H. Wobst, F. S. Schöpf, J. D. M. Solano, J. Schneider, J. Stock, A. Uhde, T. Kalthoff, *AMB Express* **2021**, *11*, 1–7.
[19] M. H. Lenders, S. Reimann, S. H. Smits, L. Schmitt, *Biol. Chem.* **2013**, *394*, 1371–1384.
[20] S. Khosa, R. Scholz, C. Schwarz, M. Trilling, H. Hengel, K.-E. Jaeger, S. H. Smits, L. Schmitt, *Appl. Environ. Microbiol.* **2018**, *84*, e01163–01117.
[21] J. A. Leeds, R. A. Welch, *J. Bacteriol.* **1997**, *179*, 3519–3527.
[22] J. Fricke, S. Pedersen, *J. Mol. Biol.* **1998**, *280*, 561–569.
[23] S. Takahashi, H. Furusawa, T. Ueda, Y. Okahata, *J. Am. Chem. Soc.* **2013**, *135*, 13096–13106.
[24] J. Sengupta, R. K. Agrawal, J. Frank, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 11991–11996.
[25] T. S. Wong, D. Zhurina, U. Schwaneberg, *Comb. Chem. High Throughput Screening* **2006**, *9*, 271–288.
[26] M. S. Packer, D. R. Liu, *Nature Rev. Genetics* **2015**, *16*, 379–394.
[27] F. H. Arnold, *Acc. Chem. Res.* **1998**, *31*, 125–131.
[28] F. Cheng, L. Zhu, U. Schwaneberg, *Chem. Commun.* **2015**, *51*, 9760–9772.
[29] H. A. Bunzel, J. R. Anderson, A. J. Mulholland, *Curr. Opin. Struct. Biol.* **2021**, *67*, 212–218.
[30] C. Novoa, G. V. Dhoke, D. M. Mate, R. Martínez, T. Haarmann, M. Schreiter, J. Eidner, R. Schwerdtfeger, P. Lorenz, M. D. Davari, *ChemBioChem* **2019**, *20*, 1458–1466.
[31] S. Islam, D. Laaf, B. Infanzón, H. Pelantová, M. D. Davari, F. Jakob, V. Křen, L. Elling, U. Schwaneberg, *Chem. Eur. J.* **2018**, *24*, 17117–17124.
[32] K. Rübsam, M. D. Davari, F. Jakob, U. Schwaneberg, *Polymer* **2018**, *10*, 423.
[33] F. Contreras, M. J. Thiele, S. Pramanik, A. M. Rozhkova, A. S. Dotsenko, I. N. Zorov, A. P. Sinitsyn, M. D. Davari, U. Schwaneberg, *ACS Sustainable Chem. Eng.* **2020**, *8*, 12388–12399.
[34] Y. Wang, P. Xue, M. Cao, T. Yu, S. T. Lane, H. Zhao, *Chem. Rev.* **2021**, *121*, 12384–12444.
[35] S. Brands, H. U. Brass, A. S. Klein, J. G. Sikkens, M. D. Davari, J. Pietruszka, A. J. Ruff, U. Schwaneberg, *Catal. Sci. Technol.* **2021**, *11*, 2805–2815.

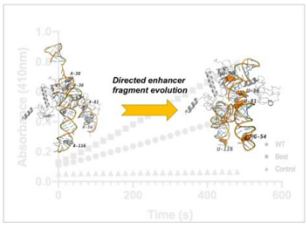
ChemBioChem
Research Article
doi.org/10.1002/cbic.202100702
Chemistry Europe
European Chemical Societies Publishing

[36] X. Li, S. Tetling, U. K. Winkler, K.-E. Jaeger, M. J. Benedik, *Appl. Environ. Microbiol.* **1995**, *61*, 2674.
[37] V. J. Frauenkron-Machedjou, A. Fulton, L. Zhu, C. Anker, M. Bocola, K. E. Jaeger, U. Schwaneberg, *ChemBioChem* **2015**, *16*, 937–945.
[38] J. Zhao, V. J. Frauenkron-Machedjou, A. Fulton, L. Zhu, M. D. Davari, K.-E. Jaeger, U. Schwaneberg, M. Bocola, *Phys. Chem. Chem. Phys.* **2018**, *20*, 9600–9609.
[39] J. Hemmerich, M. Moch, S. Jurischka, W. Wiechert, R. Freudl, M. Oldiges, *Biotechnol. Bioeng.* **2019**, *116*, 644–655.
[40] R. Freudl, *Microb. Cell Fact.* **2018**, *17*, 52.
[41] T. Baumgarten, A. J. Ytterberg, R. A. Zubarev, J.-W. de Gier, *Appl. Environ. Microbiol.* **2018**, *84*.
[42] F. Bai, Z. Li, A. Umezawa, N. Terada, S. Jin, *Biotechnol. Adv.* **2018**, *36*, 482–493.
[43] T. Suryanarayana, A. R. Subramanian, *Biochemistry* **1984**, *23*, 1047–1051.
[44] A. R. Subramanian, *Trends Biochem. Sci.* **1984**, *9*, 491–494.
[45] K. Lan Tee, U. Schwaneberg, *Comb. Chem. High Throughput Screening* **2007**, *10*, 197–217.
[46] J. Yang, A. J. Ruff, M. Arlt, U. Schwaneberg, *Biotechnol. Bioeng.* **2017**, *114*, 1921–1927.
[47] J. Zhao, V. J. Frauenkron-Machedjou, T. Kardashliev, A. J. Ruff, L. Zhu, M. Bocola, U. Schwaneberg, *Appl. Microbiol. Biotechnol.* **2017**, *101*, 3177–3187.
[48] T. Kondo, S. Yumura, *Appl. Microbiol. Biotechnol.* **2020**, *104*, 3825–3834.
[49] K. Shirai, R. Jackson, D. Quinn, *J. Biol. Chem.* **1982**, *257*, 10200–10203.
[50] U. K. Laemmli, *Nature* **1970**, *227*, 680–685.
[51] N. Dyballa, S. Metzger, *J. Vis. Exper.* **2009**, *30*, e1431.
[52] B. Beckert, M. Turk, A. Czech, O. Berninghausen, R. Beckmann, Z. Ignatova, J. M. Plitzko, D. N. Wilson, *Nature Microbiol.* **2018**, *3*, 1115–1121.
[53] M. J. Boniecki, G. Lach, W. K. Dawson, K. Tomala, P. Lukasz, T. Soltysinski, K. M. Rother, J. M. Bujnicki, *Nucleic Acids Res.* **2016**, *44*, e63-e63.
[54] S. Gong, C. Zhang, Y. Zhang, *Bioinformatics* **2019**, *35*, 4459–4461.
[55] I. Tuszynska, M. Magnus, K. Jonak, W. Dawson, J. M. Bujnicki, *Nucleic Acids Res.* **2015**, *43*, W425-W430.

Manuscript received: December 22, 2021
Revised manuscript received: January 21, 2022
Accepted manuscript online: January 21, 2022
Version of record online: ■■, ■■

RESEARCH ARTICLE

Secretion efficiency of the Hly secretion system was optimized through directed evolution of the Hly enhancer fragment using KnowVolution strategy. The structural analysis of wildtype and the developed enhancer fragment suggested that the improved secretion is due to a different affinity of the novel enhancer fragment towards the S1 ribosomal protein. This was further supported by a RNA pulldown assay.



*Z. Pourhassan N., H. Cui, S. Khosa, M. D. Davari, K.-E. Jaeger, S. H. J. Smits, U. Schwaneberg, L. Schmitt**

1 – 13

Optimized Hemolysin Type 1 Secretion System in *Escherichia coli* by Directed Evolution of the Hly Enhancer Fragment and Including a Terminator Region

2.3 Chapter III: Substrate binding sites of hemolysinB

Title: Investigations on the substrate binding sites of hemolysin B, an ABC transporter, of a type 1 secretion system

Authors: Zohreh Pourhassan N., Eymen Hachani, Olivia Spitz, Sander H. J. Smits, and Lutz Schmitt

Published in: Frontiers in Microbiology

Impact factor (2022): 6.06

Own proportion of this work: 60%

frontiers | Frontiers in Microbiology

TYPE Original Research
PUBLISHED 01 December 2022
DOI 10.3389/fmicb.2022.1055032

OPEN ACCESS

EDITED BY
Hans-Georg Koch,
University of Freiburg, Germany

REVIEWED BY
Thomas Brüser,
Leibniz University Hannover, Germany
Vassiliy Bavro,
University of Essex, United Kingdom

*CORRESPONDENCE
Lutz Schmitt
lutz.schmitt@hhu.de

SPECIALTY SECTION
This article was submitted to Microbial Physiology and Metabolism, a section of the journal Frontiers in Microbiology

RECEIVED 27 September 2022
ACCEPTED 09 November 2022
PUBLISHED 01 December 2022

CITATION
Pourhassan N Z, Hachani E, Spitz O, Smits SHJ and Schmitt L (2022) Investigations on the substrate binding sites of hemolysin B, an ABC transporter, of a type 1 secretion system. *Front. Microbiol.* 13:1055032. doi: 10.3389/fmicb.2022.1055032

Investigations on the substrate binding sites of hemolysin B, an ABC transporter, of a type 1 secretion system

Zohreh Pourhassan N.[1], Eymen Hachani[1], Olivia Spitz[1], Sander H. J. Smits[1,2] and Lutz Schmitt[1]*

[1]Institute of Biochemistry, Heinrich Heine University, Düsseldorf, Germany, [2]Center for Structural Studies, Heinrich Heine University, Düsseldorf, Germany

The ABC transporter hemolysin B (HlyB) is the key protein of the HlyA secretion system, a paradigm of type 1 secretion systems (T1SS). T1SS catalyze the one-step substrate transport across both membranes of Gram-negative bacteria. The HlyA T1SS is composed of the ABC transporter (HlyB), the membrane fusion protein (HlyD), and the outer membrane protein TolC. HlyA is a member of the RTX (repeats in toxins) family harboring GG repeats that bind Ca^{2+} in the C-terminus upstream of the secretion signal. Beside the GG repeats, the presence of an amphipathic helix (AH) in the C-terminus of HlyA is essential for secretion. Here, we propose that a consensus length between the GG repeats and the AH affects the secretion efficiency of the heterologous RTX secreted by the HlyA T1SS. Our *in silico* studies along with mutagenesis and biochemical analysis demonstrate that there are two binding pockets in the nucleotide binding domain of HlyB for HlyA. The distances between the domains of HlyB implied to interact with HlyA indicated that simultaneous binding of the substrate to both cytosolic domains of HlyB, the NBD and CLD, is possible and required for efficient substrate secretion.

KEYWORDS

ABC transporter, protein secretion, putative binding pockets, substrate interaction, bacterial secretion systems

Introduction

Gram-negative bacteria have evolved type 1 secretion systems (T1SS) to translocate a diverse variety of substrates, mainly virulence factors, in one step to the exterior space (Holland et al., 2016). The substrates of T1SS are transported in an unfolded state and classified based on their functionality, as follows: toxins, lipases, heme-binding proteins, adhesion proteins, proteases, S-layer binding proteins, and many of unknown functions (Linhartová et al., 2010; Thomas et al., 2014).

In terms of architecture, T1SS nanomachineries consist of three components: an ATP binding cassette (ABC) transporter, a membrane fusion protein (MFP), and an outer membrane protein (OMP). These three components form a continuous channel, spanning

both, inner and outer, membrane for redirecting substrates in one step from the cytosol to the extracellular space (Pourhassan et al., 2021).

A common feature of T1SS substrates is the presence of a secretion signal in the extreme C-terminus, which is not cleaved prior, while, or after secretion (Pourhassan et al., 2021). Additionally, these substrates harbor repetitive Glycine- and Aspartate-rich sequences, named as either nonapeptide sequence or GG repeats, having the consensus sequence GGxGxDxUx (x refers to any amino acid; U refers to a large and hydrophobic amino acid) upstream of the secretion signal (Bumba et al., 2016).

The GG repeats are characteristic for a group of T1SS substrates known as RTX (repeats in toxins) family, a superfamily of mostly toxic proteins secreted by Gram-negative bacteria. It has been demonstrated that the RTX is a site for binding of Ca^{2+} ions stimulating the folding of a substrate in the extracellular space where the Ca^{2+} concentration (around 2 mM) is higher than the K_D of the RTX (around 150 μM) to Ca^{2+}. Obviously, substrates remain unfolded in the cytosol where the Ca^{2+} concentration is only around 300 nM (Bumba et al., 2016; Spitz et al., 2022).

The HlyA secretion system is a paradigm of the T1SS, first identified in an uropathogenic *Escherichia coli* strain (Noegel et al., 1979). The secretion machinery of this system consists of three membrane components, as follows: the ABC transporter HlyB, the membrane fusion protein HlyD, and the outer membrane TolC (Holland et al., 2016; Pourhassan et al., 2022). Only very recently, structural information derived from single particle cryo-EM was reported and revealed an unusual six stoichiometry of the inner membrane components of the HlyA T1SS, HlyB, and HlyD, which form the so-called inner membrane complex (Zhao et al., 2022). This structure revealed that a trimer of HlyB dimers incorporated with in sum six HlyD monomers. This complex formation of HlyB and HlyD is essential for secretion. They also reported a novel organization of HlyB trimers in which three perform ATP hydrolysis, but only one acts as the protein channel for the HlyA secretion (Zhao et al., 2022).

Another protein of this system is the acyl carrier HlyC that acylates two specific lysin residues of HlyA and converts it from an inactive form into an active toxin. However, this posttranslational modification is not required for secretion (Hardie et al., 1991).

HlyA is a member of RTX family with a size of 110 kDa and has the ability to create a pore in the membrane of, for example, human erythrocytes. The required information for the secretion of HlyA is encoded within the last 50–60C-terminal amino acids (Nicaud et al., 1986). A C-terminal fragment of HlyA, termed HlyA1, has been used frequently as a transport-carrier for secretion of different heterologous proteins. HlyA1 consists of the C-terminal secretion signal along with three of the six conserved GG repeats (Pourhassan et al., 2021).

Extensive research on the T1SS substrates demonstrated that a secondary structure might be encoded by the secretion signals, since no significant conservation on the primary structure of the secretion sequences was evident. In this regard, the presence of an amphipathic helix (AH) located between residues 973 and 987 of HlyA was proposed (Koronakis et al., 1989), and only recently further data was provided by Spitz et al. supporting the idea that the presence of this AH in the C-terminus of HlyA plays an essential role in the early steps of the secretion process (Spitz et al., 2022). In the current study, we identified a consensus length between the AH and the GG repeats that influences the secretion efficiency of heterologous RTX proteins secreted by the HlyA T1SS. We observed that a reduction in the secretion rate of HlyA occurs by shortening this consensus length.

It is now understood that secretion through HlyA T1SS occurs within an ordered reaction. This reaction initiates upon interaction of HlyA with the inner membrane complex, consisting of HlyB and HlyD. This interaction recruits TolC protein, and only then all these membrane proteins form a channel bypassing HlyA in one step, the C-terminus first (Lenders et al., 2015), to the extracellular space (Letoffe et al., 1996; Thanabalu et al., 1998).

The central protein of the HlyA T1SS is the ABC transporter HlyB. In terms of topology, HlyB harbors six transmembrane domains (TMD), a nucleotide-binding domain (NBD), and a C39-peptidase-like domain (CLD) (Holland et al., 2016). Concerning the CLD, despite 40% homology to other C39-peptidases, no protease activity has been reported for this domain, because of a defective catalytic site. Nevertheless, it is essential for the secretion. Importantly, Lecher et al. reported an interaction between the CLD and the unfolded substrate outside of the secretion signal, suggesting a receptor or a chaperone-like activity for this domain, but the exact function of this domain remains unclear (Lecher et al., 2012).

The NBD domain not only provides the energy of the transport by hydrolyzing ATP (Schmitt et al., 2003; Zhao et al., 2022), but also interacts with HlyA prior to the secretion, revealed by surface plasmon resonance analysis (Benabdelhak et al., 2003). This interaction vanishes in the absence of the signal sequence of HlyA, and is highly dependent on the presence of nucleotides either ATP or ADP (Benabdelhak et al., 2003). Like other ABC transporters, the ATP binding site of HlyB is highly conserved and sequence analysis reveals the presence of three characteristic motifs, the Walker A motif with the sequence of GXXGXGKS/T (x refers to any amino acid), the Walker B motif with the sequence of ΦΦΦΦD (Φ refers to any hydrophobic amino acid), and the C-loop with the sequence of LSGGQ (also called as signature motif) (Oswald et al., 2006).

Lecher et al. already reported the substrate binding region within the CLD by conducting structural studies and pulldown assays (Lecher et al., 2012). Regarding the binding sites of NBD for HlyA no study has been conducted so far. Here, we were aiming to address this question by performing *in silico* studies. Furthermore, we analyzed the available structures of the NBD of HlyB and found two possible interaction sites for the AH of HlyA. Our mutagenesis and functional studies also support the presence of these two interaction sites. In the following, these interaction sites as well as the previously identified HlyA interaction site of the CLD were mapped onto a model of HlyB suggesting that simultaneous binding of the substrate to both cytosolic domains of the transporter is necessary for efficient substrate secretion.

Materials and methods

All chemicals used in this study were purchased from Sigma-Aldrich, Applichem GmbH, and Roche Diagnostics GmbH if not stated otherwise. Enzymes were purchased from New England Biolabs. All oligonucleotides were purchased from Eurofins MWG Operon and diluted in MilliQ-water. The NucleoSpin plasmid miniprep kit and the Q5 site-directed mutagenesis kit were purchased from Macherey Nagel (Dueren, Germany) and New England Biolabs, respectively. DNA sequencing was performed at Microsynth Seqlab (Göttingen, Germany).

Construction of plasmids used in this study

In this study *E. coli* DH5α, *E. coli* BL21-Gold (DE3), and *E. coli* BL21 (DE3) were used for cloning, overexpression, and purification purposes. Plasmids and oligonucleotides used in this study are listed in Tables 1, 2, respectively.

Plasmids containing truncated variants of HlyA were constructed as follows

Plasmid pSU2726_*hlyC* (159 bp)-*hlyA* was used as the backbone plasmid. This plasmid contains an ampicillin resistance gene. The *lac* promoter of this plasmid is inducible with isopropyl β-D-1-thiogalactopyranoside (IPTG). To delete the encoding region of 30 and 50 residues of HlyA, plasmid pSU2726_*hlyC* (159 bp)-*hlyA* was linearized *via* PCR reaction using primer sets of P9/P10 and P9/P11, respectively. Subsequently, both PCR products were incubated with the KLD (kinase-ligase-DpnI) reaction mix of the Q5 site-directed mutagenesis kit (5 min, room temperature), according to the instructions of the manufacturer. Subsequently, 2 μL of the KLD mixture was transformed into chemically competent cells of *E. coli* DH5α. The sequence of the constructed plasmids pSU2726_*hlyC* (159 bp)-*hlyA* (Δ939-968) and pSU2726_*hlyC* (159 bp)-*hlyA* (Δ919-968) were confirmed *via* plasmid sequencing.

TABLE 1 Plasmids used in this study.

Plasmids name	Source
pK184_*hly*BD	Khosa et al. (2018)
pK184_*hly*B (F518D, Y519D)-*hly*D	This study
pK184_*hly*B (Y477D, F518D, Y519D)-*hly*D	This study
pK184_*hly*B (V682D)-*hly*D	This study
pK184_*hly*B (V682D, L697D, Y700D)-*hly*D	This study
pPSG122_His::NBD	Zaitseva et al. (2004)
pPSG122_His::NBD (Y477D)	This study
pPSG122_His::NBD (F518D, Y519D)	This study
pPSG122_His::NBD (Y477D, F518D, Y519D)	This study
pPSG122_His::NBD (V682D)	This study
pPSG122_His::NBD (L697D, Y700D)	This study
pPSG122_His::NBD (V682D, L697D, Y700D)	This study
pSU2726_*hly*C (159 bp)-*hly*A	Khosa et al. (2018)
pSU2726_*hly*C (159 bp)-*hly*A (Δ939-968)	This study
pSU2726_*hly*C (159 bp)-*hly*A (Δ919-968)	This study

Mutations in hlyB gene were introduced as follows

In general, plasmid pK184_*hlyBD* was used as the backbone plasmid. The *lac* promoter of plasmid pK184 is inducible with IPTG. This plasmid has a kanamycin resistance gene. The mutations were introduced in the *hlyB* gene using the Q5 site-directed mutagenesis kit, according to the instructions of the manufacturer. First the PCR for introducing each mutation was performed with the proper primer sets. Then, the PCR product was incubated with the KLD mix (5 min, room temperature). Subsequently, 2 μL of the KLD mixture was transformed into chemically competent cells of *E. coli* DH5α.

In detail, to introduce the mutations concerning the putative binding pocket inside (pbp-in), first plasmid pK184_*hlyBD* (F518D, Y519D)-*hlyD* was constructed using primer sets of P3/P4. To introduce the third mutation of pbp-in, plasmid pK184_*hly*B (F518D, Y519D)-*hly*D was amplified using primer sets of P1/P2. The successful construction of plasmid pK184_*hlyB* (Y477D, F518D, Y519D)-*hlyD*, containing the three mutations in the *hlyB* gene, was verified *via* sequencing.

To introduce the mutations concerning the putative binding pocket outside (pbp-out), first plasmid pK184_*hly*B (V682)-*hly*D was constructed using primer sets of P5/P6, and then used as the backbone plasmid for the introduction of the second and third mutations. The PCR amplification of plasmid pK184_*hly*B (V682)-*hly*D was performed using the primer sets of P7/P8. The successful construction of plasmid pK184_*hly*B (V682D, L697D, Y700D)-*hly*D, containing three mutations in the *hly*B gene, was verified *via* sequencing.

Mutations in the coding region of the nucleotide-binding domain of HlyB were introduced as follows

Plasmid pPSG122_*His::NBD* was used as the backbone plasmid. The *pBAD* promoter of this plasmid is inducible with L-arabinose. This plasmid has an ampicillin resistance gene. All the mutations in the *NBD* were introduced as described for the pK184_*hlyBD* backbone plasmid. The only exception was that primer P2' was used instead of primer P2. Sequencing was performed to verify the successful construction of the desired plasmids, listed in Table 1.

TABLE 2 Oligonucleotides used in this study.

Name	Details	Sequence
P1	Fp, pbp-in, Y477D	CCGGTTTCGCGATAAGCCTGACT
P2	Rp, pbp-in, Y477D	ATATTACGAAAAGTGATATCACCATTAATTTC
P2'	Rp, pbp-in, Y477D	ATATTACGAAAAGTGATATCGTG
P3	Fp, pbp-in, F518D, Y519D	AATTCAACGTGATGATATTCCTGAAAATGGC
P4	Rp, pbp-in, F518D, Y519D	AATTTAGTTAATGTGCTTTTTCC
P5	Fp, pbp-out, V682D	AGGGAAAATTGATGAACAGGGTAAAC
P6	Rp, pbp-out, V682D	TTTTCCATGACAATAATGCG
P7	Fp, pbp-out, L697D, Y700D	AGTGATTTATATCAGTTACAGTCAGAC
P8	Rp, pbp-out, L697D, Y700D	GTAATCACTTTCCGGTTCAGAAAG
P9	Fp, deletion of 30 or 50 residues of HlyA	AGTCAGGGTGATCTTAATCC
P10	Rp, deletion of 30 residues of HlyA	GATTATCCGGCCACTTTTATC
P11	Rp, deletion of 50 residues of HlyA	CTCTTTTTCAAACCAGTTCCTG

Fp, forward primer; Rp, reverse primer.

Expression and secretion experiments in shaker flasks

Escherichia coli BL21 (DE3) chemically competent cells were transformed with the desired plasmids and grown on LB agar plates supplemented with 50 µg/mL kanamycin and/or 100 µg/mL ampicillin.

In all of the cultures, first a mixture of transformed clones was used to prepare a pre-culture and cultivated overnight (37°C and 180 rpm). The overnight culture was used to inoculate 25 ml of 2YT (16 g/l Tryptone, 10 g/l yeast extract, 5 g/l NaCl) medium supplemented with 50 µg/mL kanamycin and/or 100 µg/mL ampicillin at an OD_{600} of 0.1 in 100 ml Erlenmeyer shaking flask. The cultures were cultivated at 37°C and 180 rpm to an OD_{600} of 0.7–0.8. Subsequently, the expression was induced with 1 mM IPTG and 5 mM $CaCl_2$. Cells were grown for 4 h and samples were taken every hour. The samples were centrifuged (5,200 x g, 5 min). Then, supernatant and cell pellet were collected in separate Eppendorf tubes and mixed with an amount of SDS sample buffer to normalized the samples in the respect of the OD of cultures. Cell growth was also monitored each hour by measuring the OD_{600}. The samples were analyzed either by SDS-PAGE or Western blot analysis. Staining of the SDS-PAGE gels was performed *via* the Colloidal Coomassie G-250 Staining protocol (Dyballa and Metzger, 2009). The intensity of the protein bands on the SDS-PAGE gels were semi-quantified using the ImageJ software (Image Processing and Analysis in Java) if necessary (Abràmoff et al., 2004).

To determine the amount of secreted target protein, a series of purified HlyA solutions with known concentrations were also loaded on the same SDS-PAGE. The amount of secreted target protein in the supernatant was then compared to the HlyA solution through a calibration line. The rate of secretion was calculated according to the published protocol (Lenders et al., 2016).

The supernatants samples were mixed with SDS-sample buffer. The pellet samples were first resuspended in resuspension buffer (50 mM Na_2HPO_4, pH 8, 300 mM NaCl) and then mixed with SDS-sample buffer. The supernatants and pellets samples were analyzed using SDS-PAGE analysis and Western-blotting.

Expression and purification of HlyB nucleotide-binding domain variants

Expression and purification of the NBD variants were performed according to the published procedure (Zaitseva et al., 2004) with slight modifications, as described below.

The NBD plasmids (WT or mutant variants) were introduced into chemically competent *E. coli* BL21 (DE3) or *E. coli* BL21-Gold (DE3) cells. Then, the cells were grown on LB agar plates supplemented with 100 µg/mL ampicillin. A mixture of clones was used to prepare a pre-culture (50 ml 2YT) and cultivated overnight (37°C and 180 rpm). The overnight culture was used to inoculate 2 L of 2YT medium supplemented with 100 µg/mL ampicillin in a ratio of 1:100. The cultures were cultivated (37°C and 180 rpm) to an OD_{600} of 0.8. The cultures were cooled down to 20°C and cultured again to an OD_{600} of 1. Then, the cultures were induced with 0.002% L-arabinose and continued to grow for 3 h (20°C and 180 rpm). All following steps were performed at 4°C. Cells were harvested by centrifugation (20 min, 4,500 × g, 4°C). Cells were resuspended in buffer A (25 mM sodium phosphate, 100 mM potassium chloride, 10 mM imidazole, 20% Glycerol, pH 8) supplemented with protease inhibitor cocktails and DNAase. Cells were disrupted by passing three times through a cell disruptor (Microfluidizer M-110P, Microfluidics) at 1.5 kbar. Cell debris and undisrupted cells were removed by centrifugation (60 min, 125,000 × g, 4°C). The supernatant was loaded onto a 5 ml Zn^{2+}-charged chelating HiTrap™ HP column that was already equilibrated with buffer A. The column washed with 25 ml of

buffer A and the elution step was performed with a linear gradient of buffer A2 (25 mM sodium phosphate, 100 mM potassium chloride, 300 mM imidazole, 20% glycerol, pH 8). Based on SDS-PAGE analysis, fractions containing the HlyB NBD variants were mixed and concentrated to a final volume of 500 µL using an Amicon Ultra filter (15 ml, MWCO = 10,000 Da Merck/Milipore). The concentrated protein was centrifuged (10 min, 100,000 x g, 4°C) and subjected to a size-exclusion chromatography on the Superdex 200 increased 10/300 Gl column (GE Healthcare) in buffer B (10 mM CAPS-NaOH, 20% glycerol, pH 10.4). The purified HlyB NBD variants were concentrated using an Amicon Ultra filter (5 ml, MWCO = 10,000 Da Merck/Milipore), and used for further analysis.

ATPase activity assays

ATPase assays were performed according to the published procedure (Zaitseva et al., 2005b) with slight modifications. In this assay, the amount of released free phosphate was determined *via* a colorimetric assay. The concentrated purified NBD was diluted in HEPES buffer (100 mM HEPES, 20% Glycerol, pH 7) only before starting the assay. Subsequently, 30 µL of NBD solution in HEPES buffer was added to 10 µL of 50 mM $MgCl_2$ and 10 µL of ATP solution (to a final concentration of 0 to 6 mM, pH 8). Instead of $MgCl_2$, HEPES buffer was used in negative control samples. The ATPase assay was started by supplementing protein solution to the reaction.

The reaction mixture in a final volume of 50 µL was incubated for 120 min at 22° C. Then, 25 µL of sample reactions was added to a 96 well pellet and reaction stopped by adding 175 µL of 10 mM H_2SO_4. The amount of released free phosphate was monitored by adding 50 µL of staining solution (0.096% (w/v) malachite green, 1.48% (w/v) ammonium molybdate, 0.173% (w/v) Tween-20 in 2.36 M H_2SO_4). The mixture was incubated 8 min at room temperature. The amount of released free phosphate was recorded by measuring $A_{595\ nm}$ with a micro plate reader (iMark Microplate Reader, Bio Rad) and using phosphate buffer as standard. Raw data were fitted using GraphPad Prism 8 Software (GraphPad).

Secondary structure prediction

Quick2D (Zimmermann et al., 2018) and AmphipaSeeK (Combet et al., 2000; Sapay et al., 2006) were used to predict the secondary structures. Quick2D is used for prediction of α-, π- and TM-helices, β-strands, coiled coils, as well as disordered regions (Zimmermann et al., 2018). AmphipaSeeK is specifically designed to identify amphipathic helices (Sapay et al., 2006). The AmphipaSeeK predicts a secondary structure, a membrane topology (in-plane or not-in-plane), a score for the proposed membrane topology, and an amphipathy score for each residue in dependence to the neighboring residues.

Structure prediction of HlyB

The structure of HlyB was modeled based on the structure of PCAT1 with TopModel tool (Lin et al., 2015; Mulnaes et al., 2020). The monomers of our HlyB model superimposed with a RMSD of 3.3 Å (558 Cα atoms) and 3.6 Å (558 Cα atoms), respectively, on the monomers of the inward facing conformation of HlyB in the structure of the HlyB / HlyD complex (Zhao et al., 2022). These differences are due to a different orientation of TMD and NBD in our model compared to the single particle structure. Superimposition of the isolated TMDs revealed a RMSD of 1.8 Å (311 Cα atoms) and 1.3 Å (273 Cα atoms), respectively for the monomers of our HlyB model and the monomers of the inward facing conformation of HlyB in the HlyB / HlyD structure. The crystal structures of the isolated NBD of HlyB in the apo form (1MT0; Schmitt et al., 2003) and the ATP / Mg^{2+} bound state (1XEF; Zaitseva et al., 2005a) superimposed with RMSD values of 0.7 Å (191 Cα atoms, 1MT0) and 1.2 Å (449 Cα atoms, 1XEF), respectively. Due to the high structural identity between the crystal structures of the isolated NBD and the NBD in the single particle cryo-EM structure, we primarily used these crystal structures for our analysis.

Illustration and visualization

We used Netwheel to visualize the amphipathic characteristics of a helix (Mol et al., 2018). Protein and peptide structures were processed in PyMOL (Delano, 2002; the PyMOL molecular Graphics System, Version 1.8.6.0 Enhanced for Mac OS X). In order to illustrate and identify hydrophobic surfaces the YRB-script was applied in PyMOL (Hagemans et al., 2015).

Sequence alignments

Alignments were performed using Clustal Omega (Madeira et al., 2019). Sequences of the homologous proteins were taken from Uniprot[1] or NCBI[2].

Results

In silico studies on the putative HlyB binding pockets

It was proposed that in one of the early steps of the secretion, the secretion signal of HlyA interacts with the NBD of HlyB. Here, it was also reported that the presence of ATP accelerates the dissociation of this complex (Benabdelhak et al., 2003). Also, it

1 https://www.uniprot.org/
2 https://www.ncbi.nlm.nih.gov/

was revealed that an amphipathic helix (AH) in the extreme C-terminus of HlyA plays an essential role in the secretion (Spitz et al., 2022). In this study, we assumed that the interaction between HlyA and NBD takes place within this essential AH. Based on this assumption, the binding region of the AH is expected to be a hydrophobic site. Accordingly, *in silico* studies explained in the following were performed.

It should be noted that different structures of the isolated NBD from different stages of ATP hydrolysis cycle are available, for instances: structures of the wild type NBD either as a nucleotide free monomer (PDB entry 1MT0; Benabdelhak et al., 2003), as an ADP bound monomer (PDB entry 2FF7; Zaitseva et al., 2005a), or as a TNP-ADP bound monomer (PDB entry 2PMK; Oswald et al., 2008). Additionally, the structures of the dimeric forms are available for the hydrolysis-impaired mutant E631Q with two ATP bound (PDB entry 2FGK; Zaitseva et al., 2005a), for the hydrolytic inactive mutant H662A (activity <0.1%; Zaitseva et al., 2005a) with two ATP bound (PDB entry 2FGJ; Zaitseva et al., 2005a), and for the mutant H662A with two ATP and two Mg^{2+} ions bound (PDB entry 1XEF; Zaitseva et al., 2005b). As outlined in Materials and Methods, the differences between the crystal structures of the isolated NBD of HlyB and the NBD of the HlyB structure in the cryo-EM structure of the HlyB / HlyD complex (Zhao et al., 2022) are marginal. Therefore, we focused on the structures of the isolated NBD as different conformations of the catalytic cycle are available.

Based on Benabdelhak et al., the interaction between HlyA and the NBD domain occurs in the absence of ATP, or on the other hand, the interaction occurs with the nucleotide free form of the NBD (Benabdelhak et al., 2003). Therefore, the structure of the nucleotide free monomer (PDB entry 1MT0) was used to search and identify putative hydrophobic binding pockets using the YRB-script (Hagemans et al., 2015). Furthermore, the dimeric ATP-bound structure (1XEF) was also analyzed to examine if the interaction is disrupted in the presence of ATP. Note that the positional changes of the residues can be analyzed by superimposing both structures and calculating the RMSD value only for the residues involved in forming the putative binding site. By performing this, we identified two regions of high hydrophobicity, which we called putative binding pocket inside (pbp-in) and putative binding pocket outside (pbp-out).

The pbp-in is located closely to the dimer interface and the ATP binding site (Figure 1A). While some residues from this pocket are polar amino acids, they contribute to the nonpolar character of the binding pocket with the carbon atoms of their side chains. Furthermore, charges at the side of a binding pocket for an AH may interact with the polar side of the AH and help its orientation. The residues that form pbp-in are: F475, Y477, K478, I484, T510, K513, Q516, F518, and Y519 (Figure 1C).

The pbp-in shares at least three residues with the ATP binding site (Y477, I484, T510): Y477, located in the A-loop (Ambudkar et al., 2006), interacts with the adenine base, I484 with the ribose moiety, and T510 with the Pα of ATP (Zaitseva et al., 2005a). When superimposing the ATP free monomer (1MT0) with the ATP bound dimer (1XEF) the residues of the pbp-in show a change in position of approximately 2 Å as reflected by the RMSD value. Both Tyr (Y477 and Y519) and both Lys (K487 and K513) display the largest change in position. The RMSD value is 1.3 Å when compared to the ADP-bound state (2FF7).

The pbp-out is located opposite to the membrane and is exposed to the cytosol (Figures 1B,F). It is made up by the following residues: V675, E677, K680, V682, E683, L697, Y700, L701 and L704 (Figure 1D). Y700, L701, and L704 point toward the dimer interface and are involved in monomer-monomer contacts (Zaitseva et al., 2005a) in the ATP-bound state. The pbp-out changes less than the pbp-in upon ATP binding and dimerization as shown by the RMSD value of 1.1 Å. The difference to the ADP-bound monomer is only 0.9 Å.

Both suggested pbp's hold the potential to be the interaction sites of the C-terminal part of HlyA, as: (i) Both pbp's display a hydrophobic region that matches the length of the AH in HlyA. (ii) They hold residues that are able to form π-π stacking interactions with their side chains and could, therefore, act as an interaction partner to F990 of HlyA. F990 is an essential amino acid of HlyA for the efficient secretion as revealed by mutagenesis studies (Chervaux and Holland, 1996). (iii) Both pbp's are changed upon ATP binding and dimerization, supporting the already published observation of Benabdelhak et al. (2003).

In vivo studies on the putative HlyB binding pockets

We applied directed mutagenesis to the proposed binding pockets to perform *in vivo* studies on the putative binding pockets. The first sets of mutagenesis experiments investigated the pbp-in. Two different variants of HlyB were cloned by introducing mutations at either two or three positions in the pbp-in site. One of the mutants harbored two point-mutations, F518D and Y519D, whereas the other one harbored three point-mutations, Y477D, F518D, and Y519D.

Additionally, the proposed pbp-out was investigated by mutagenesis studies. A pbp-out triple mutant was cloned that harbored the mutations V682D, L697D, and Y700D.

The constructed plasmids harboring different variants of HlyB protein were co-transformed into chemically competent *E. coli* BL21(DE3) cells along with pSU2627_*hlyC*(159 bp)-*hlyA* plasmid. Cell colonies harboring plasmids pK184_*hlyBD* and pSU2627_*hlyC*(159 bp)-*hlyA* were used as positive controls. Test expression of clones secreting HlyA was performed in shaking flasks employing directly the supernatant of cultures for analysis by SDS-PAGE. The test expression experiments were performed with three biological replicates.

Interestingly, we observed that neither the double pbp-in mutant (Figure 2A), triple pbp-in (Figure 2B), nor triple pbp-out mutant (Figure 2B) were able to secrete HlyA. However, in the same condition the positive controls were able to secrete HlyA (Figures 2A,B).

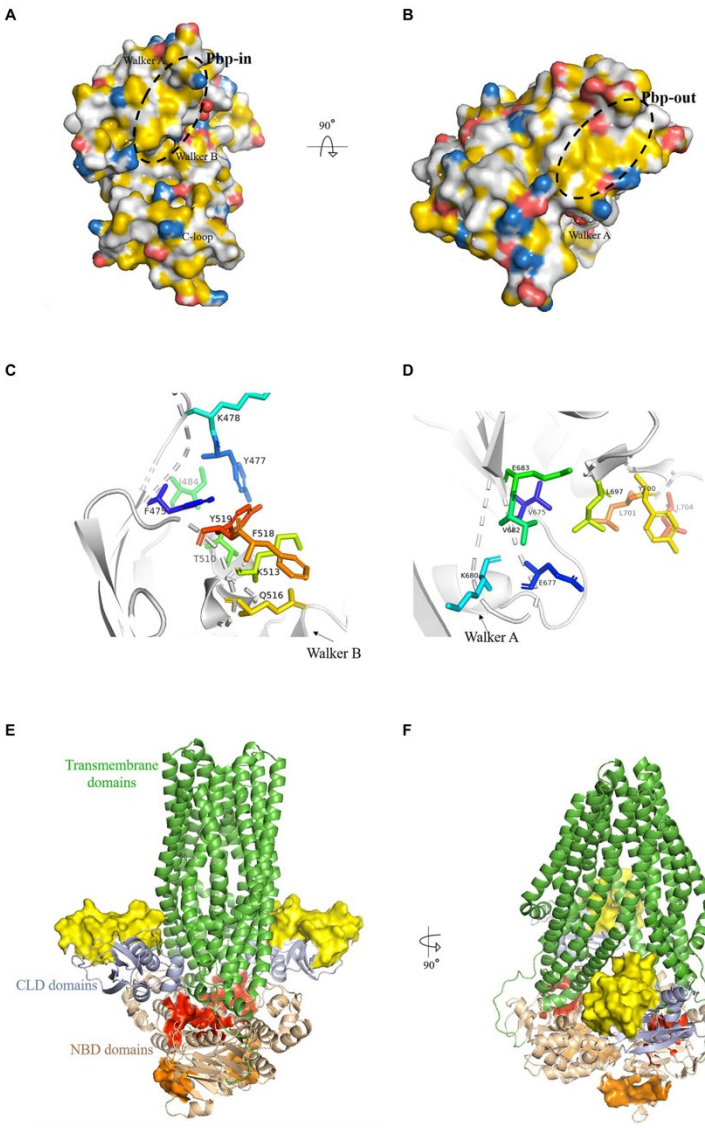


FIGURE 1
Putative binding pockets (pbp) of HlyB NBD. **(A,B)** The HlyB NBD monomer (1MT0) is shown in surface representation. The surface was colored with the YRB-script (Hagemans et al., 2015), which highlights carbon atoms that are not bound to oxygen or nitrogen in yellow, the charged oxygens of Glu and Asp in red, the charged nitrogen of Lys and Arg in blue while all other atoms in white. The structural elements of the NBD were labeled. The pbp's are circled with a black dashed line. From panel **A**, which shows the pbp-in, to panel **B**, which shows the pbp-out, the molecules have been rotated toward the reader. **(C,D)** The residues of pbp-in **(C)** and pbp-out **(D)** from the monomeric structure (1MT0) are shown as stick representations. **(E,F)** Cartoon representation of a model of HlyB based on PCAT1 shown in green (Lin et al., 2015). The identified pbp's are shown as surface representation with pbp-in in red and pbp-out in orange. The yellow surface maps the HlyA-interaction region on the CLD (Lecher et al., 2012). Pale colors correspond to the same regions in the second monomer. The overall domains of HlyB were labeled.

The expression level of HlyB and HlyD of the different mutants was analyzed using Western blotting demonstrating that HlyB (Figures 2C,E) and HlyD (Figures 2D,F) were successfully expressed in the different mutants. Thus, the lack of secretion of the mutants observed in our test secretion experiments cannot be attributed to the lack of the HlyA secretion system apparatus or individual components.

In vitro studies on the putative HlyB binding pockets

To express and purify the HlyB-NBD mutants, a construct containing the C-terminal residues 467 to 707 of HlyB was overexpressed and purified according to the published protocol (Zaitseva et al., 2004). The expression and purification of wildtype HlyB-NBD, double-pbp-in mutant, and triple-pbp-in mutant were performed in high yield and purity (Figure 3). Gel shifting, which is uncorrelated migration with formula molecular weight of a protein on SDS-PAGE, was observed for the purified mutated NBDs (Figure 3). Amino acid substitutions can cause mobility changes during electrophoresis on SDS-PAGE. To date, several studies have investigated the correlation between the mobility changes and biochemical features of proteins, but still this relationship has to be explained (Rath et al., 2009; Shi et al., 2012). Shi et al. (2012) observed that net charge of a protein, binding ability to SDS-molecules are factors which affect protein migration on the SDS-PAGE.

Although, different expression strains and different expression conditions were tested, the purification of the pbp-out mutant was not successful.

ATPase activity of the HlyB nucleotide-binding domain mutants

Isolated NBD variants were subsequently used to characterize the ATPase activity of each NBD. ATPase activity of the HlyB-NBD mutants was measured by a colorimetric assay based on the amount of released inorganic phosphate from the ATP hydrolysis.

The ATPase activity of the NBD pbp-in triple mutant was completely abolished (range of ATP concentrations 0 to 6 mM).

The maximum velocity (V_{max}) of 253.8 nmol/mg.min was determined for the wild type NBD under the conditions of the assay. The V_{max} of the NBD pbp-in double mutant was reduced two-fold in comparison to the wild type and determined to be 133.4 nmol/mg.min (Table 3; Figure 4). Despite of ATPase activity of the NBD pbp-in double mutant, no secretion was observed for this mutant.

Note that the double and triple mutants of pbp-in differ in one residue (Y477), but the double mutant still is able to hydrolyze ATP, nevertheless, these mutations led to the lack of secretion. The kinetic parameters of the different ATPase activity are summarized in Table 3.

Y477 is a residue of the A-loop, which coordinates the adenine ring. Interestingly, Pdr5, the ABC transporter involved in pleiotropic resistance from baker's yeast, have no aromatic A-loop residue but still able to hydrolyses ATP with high efficiency (Raschka et al., 2022).

Investigations on the HlyB C39-peptidase-like domain

It is now understood that the CLD domain of HlyB is essential for secretion. Additionally, an interaction between HlyA and the CLD domain has been confirmed (Lecher et al., 2012). It seems that the AH at the C-terminus of HlyA has no impact on this interaction which was revealed by pull-down assay of the isolated CLD with unfolded/folded HlyA, HlyA1 (C-terminal 217 residues), and HlyA2 (HlyA1 lacking the C-terminal 60 residues that make up the secretion signal). Despite lacking the AH in HlyA2, no change in the interaction between the CLD and HlyA was observed. Instead, the interaction was disrupted when the experiment was performed with folded versions of the substrate (Lecher et al., 2012). In this regard, the authors assigned the interaction to the conserved GG repeats of the RTX domain of HlyA, which induce folding of RTX proteins upon Ca^{2+} binding (Baumann et al., 1993; Lecher et al., 2012).

By a combination of chemical shift experiments and mutational studies Lecher et al. were able to map the region in the CLD that interacts with HlyA (Figures 1E,F). Since HlyB seems to interact with different regions of the substrate, we were interested if a simultaneous binding would be possible.

The last C-terminal GG repeat of HlyA2 is located 122 amino acids upstream of the AH. Assuming 3.8 Å per amino acid in an unfolded protein (Carrion-Vazquez et al., 1999; Crecca and Roitberg, 2008), the distance between the last C-terminal GG repeat and the AH (NBD-interaction site) is approximately 450 Å. For a visualization of these distances see Figure 5.

In order to measure the distances between the pbp's of the NBD and the interaction site on the CLD the structure of full-length HlyB is needed. However, the HlyB structure became available only recently and we modeled it based on another ABC transporters, PCAT1 (Lin et al., 2015; Figures 1E,F).

In this model the outmost residue of the HlyA interaction site in the CLD is E92. The minimal distance between E92 and pbp-in measures ~40 Å and the maximal distance from E92 to pbp-in is ~64 Å. The pbp-out shows a distance to E92 between 54 Å and 68 Å, respectively. Based on these calculations, HlyA can easily bridge these distances and simultaneous binding of the substrate HlyA to the CLD and NBD of HlyB is possible. In the recently published structure of the HlyB / HlyD complex (Zhao et al., 2022) only one of the three HlyB dimers is transport competent. Furthermore, only one of the two CLDs is visible in the electron density of the transport competent dimer. Here, the distance

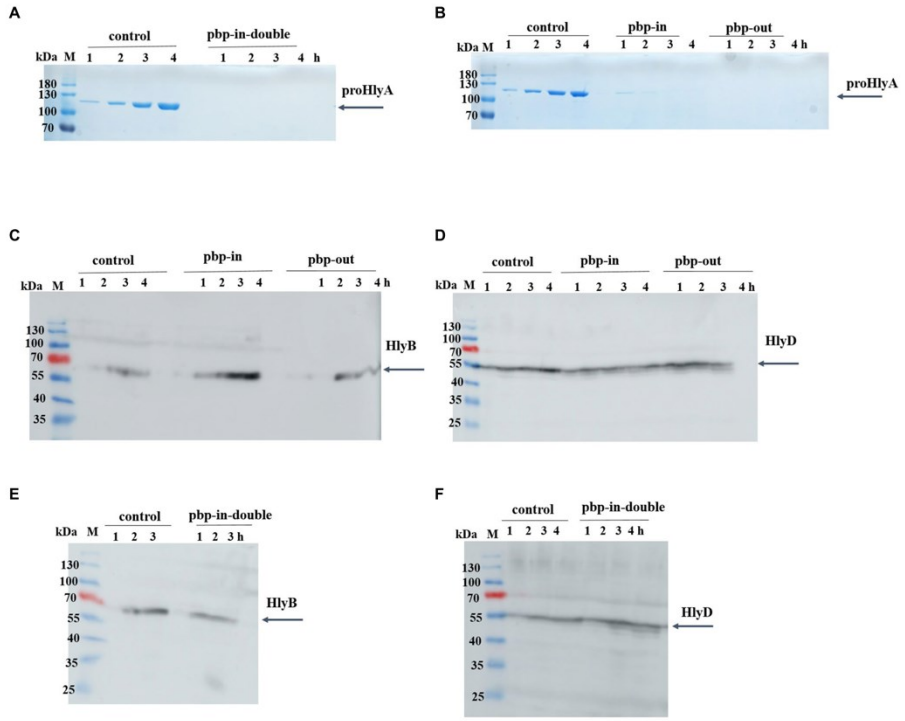


FIGURE 2
Secretion of proHlyA through the HlyA T1SS, in which HlyB harbors mutation. **(A)** SDS-PAGE analysis of the supernatant (unconcentrated) of clones secreting HlyA, in which HlyB is the wild type (control) or contains double mutations in pbp-in. **(B)** SDS-PAGE analysis of the supernatant (unconcentrated) of clones secreting HlyA, in which HlyB is the wild type (control), triple mutations in the pbp-in, or triple mutations in the pbp-out. Western blotting of *E. coli* cells (whole cells) demonstrated that **(C,E)** HlyB and **(D,F)** HlyD were expressed for all three mutants at levels comparable to the control cells. M, marker proteins; the molecular weight of the marker proteins is given on the left; xh, unconcentrated supernatant of culture, where x denotes the number of hours after induction.

between E92 and pbp-in is approximately 60 Å, while the distance to pbp-out is with slight increased, approximately 75 Å. This also supports our conclusion that HlyA can bridge this distance and that the proposed simultaneous binding is indeed possible.

If multiple interaction sites on HlyB are occupied simultaneously by one substrate molecule, this would result in a strictly ordered substrate arrangement, which could confer specificity between substrate and transport machinery. The interaction between substrate and the secretion machinery has also essential impact on inducing the assembly of the translocation channel (Thanabalu et al., 1998; Zhao et al., 2022).

The transport components HlyB and HlyD can also secrete heterologously expressed RTX toxins such as FrpA from *N. meningitidis* and HlyIA from *A. pleuropneumoniae* serotype 1 (Gygi et al., 1990; Thompson and Sparling, 1993). Both are predicted to have an AH in their C-terminus and a GG repeat can be found 122 residues (FrpA) and 121 residues (HlyIA) upstream of this AH.

Investigations on the conserved distance between the predicted amphipathic helix domain and the GG-repeat of HlyA

The secretion signal of HlyA is located in the C-terminal 60 residues (Holland et al., 2016). Interestingly, heterologous substrates that can be secreted by HlyBD show AHs in their C-terminal secretion signal when analyzed with the prediction tool AmphipaSeeK. Furthermore, these heterologous substrates display the same distance between the predicted AH and the GG-repeat of the RTX domain as HlyA (Table 4). A linker of a conserved length between the two interaction sites further

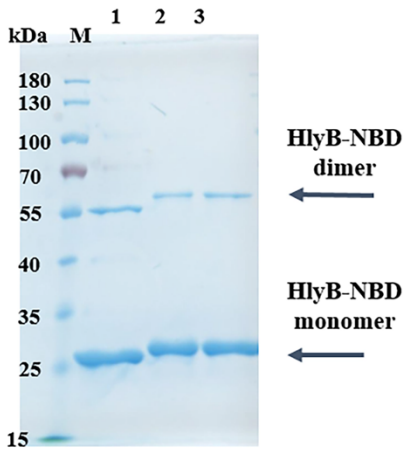


FIGURE 3
Purification of HlyB NBD mutants. SDS-PAGE of purified HlyB NBD of wild type (1), double-pbp-in (2), and (3) triple-pbp-in (3). M, marker proteins; the molecular weight of the marker proteins is given on the left.

TABLE 3 Summary of the kinetic parameters of HlyB NBD wild type and the mutated variants.

	Kinetic parameters	WT NBD	Pbp-in-double
Michaelis-Menten kinetic	V_{max} (nmol/mg.min)	253.8 ± 7.6	133.4 ± 7.9
	K_m (mM)	0.65 ± 0.06	0.73 ± 0.12
Hill equation	h	1.66 ± 0.154	2.37 ± 0.2
	$K_{0.5}$ (1/min)	0.49 ± 0.026	0.5 ± 0.018

strengthens the theory that the RTX domain and the secretion signal interact simultaneously with the ABC transporter, which might be important for the correct orientation of the substrate. Additionally, an aromatic residue (F990) close to the AH was shown to be important for HlyA secretion and a Phe residue close to the predicted AHs can be found in the heterologous substrates as well (Table 4).

The following mutagenesis experiments were performed to modify the length between the two docking sites, AH and GG repeats. Two different truncated versions of HlyA were cloned by deleting either 30 or 50 residues of the 130 residues that are encoded between the AH and the GG-repeats of HlyA. For the 30-residues-truncated HlyA, residues from 939 to 968 were deleted, and for the 50-residues-truncated HlyA, residues from 919 to 968 were deleted. The predicted AH domain is located between residues 975 to 987. The strain containing the plasmid encoding full-length HlyA was used as the control strain. Expression experiments of clones secreting wild type HlyA and both truncated variants of HlyA were performed in shaking flasks employing directly the supernatant of cultures for analysis by SDS-PAGE. The analysis confirmed that the secretion efficiency of both truncated HlyA variants were strongly reduced compared to the wild type HlyA (Figure 6A). The secretion rates of two HlyA variants were quantified according to the published protocol (Lenders et al., 2016). The secretion rates for the HlyA variants were displayed in Figure 6B indicating that the secretion rate of wild type HlyA is almost two-fold higher than the 30-residues-truncated HlyA variant. Additionally, the 50-residues-truncated HlyA variant showed the lowest secretion rate compared to the other two tested variants. Consequently, these observations demonstrated that by shortening the distance between the predicted AH domain and the GG-repeats of HlyA, the secretion level of this substrate is reduced.

The expression level of HlyB and HlyD of the different variants as well as the amount of HlyA was analyzed using Western blotting demonstrating that HlyB (Figure 6C), HlyD (Figure 6D), and pro HlyA (Figure 6E) were expressed in comparable levels in the different variants. Thus, the lower secretion of proHlyA observed in our test secretion experiments cannot be attributed to the lower amount of the secretion system apparatus or the secreted substrate.

Discussion

The ABC transporter HlyB plays a central role in the HlyA T1SS, because it not only provides the energy of the transport by hydrolyzing ATP, but also plays a role in the early steps of the secretion by interacting with the substrate prior to the secretion. Previous studies have explored the relationship between HlyB and the substrate, and reported two interactions: An interaction site within the CLD domain (Lecher et al., 2012) and an interaction site within the NBD domain (Benabdelhak et al., 2003). Lecher et al. have been able to map the binding site of the CLD for HlyA (Lecher et al., 2012), but the binding site of the NBD for HlyA is still an open question. Here, we aimed to address this question by performing *in silico*, mutagenesis, and biochemical studies.

Recent studies demonstrated that the presence of an AH, covering the residues 970–987, in the C-terminus of HlyA is important for secretion. More important, this secondary structure rather than the primary structure plays a crucial role in the recognition of HlyA (Spitz et al., 2022). Furthermore, based on studies conducted by Benabdelhak et al. (2003), the interaction between the NBD and HlyA occurs within the C-terminus of the substrate, most probably the signal sequence is involved in this interaction, since HlyA2 lacking the C-terminal 57 amino acids does not interact with the NBD. Interestingly, the formed complex HlyB-NBD / C-terminus of HlyA dissociates when ATP is present (Benabdelhak et al., 2003).

Putting the pieces of this puzzle together, we assumed that the C-terminal AH is the site of HlyA that interacts with HlyB-NBD. To identify putative binding pocket(s) in HlyB, *in silico* studies were conducted and based on that, two putative binding

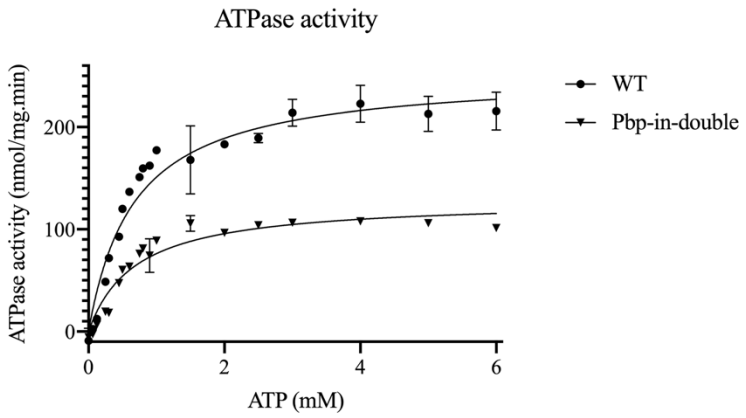


FIGURE 4
Kinetic measurements of the ATPase activity of HlyB NBD mutants, WT HlyB NBD (circle) and double-pbp-in (triangle). Data are from minimum two independent replicates, error bars were calculated by SEM.

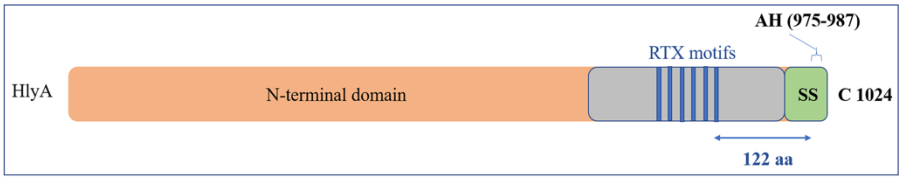


FIGURE 5
Schematic view of the HlyA and HlyA2. The distance the last C-terminal RTX and the beginning of the AH is 122 residues (122*3.8Å=463.6Å).

pockets were identified, pbp-in and pbp-out. Interestingly, both binding pockets are located in hydrophobic regions with a size capable to accommodate the C-terminal AH of HlyA. Also, there are residues in both pockets that are able to form π-π interactions, thus those sites could be sites for binding of F990. This residue is in the C-terminus of HlyA, however, not part of the AH, but proven to be essential for secretion (Chervaux and Holland, 1996; Spitz et al., 2022). Both pockets showed changes upon ATP binding and dimerization that could go along with published observations (Benabdelhak et al., 2003) that ATP binding accelerates the dissociation of HlyB and HlyA.

A set of mutagenesis experiments were conducted on these putative binding pockets. First, for pbp-in two mutants were cloned, one mutant with three mutated residues (Y477D, F518D, and Y519D), and a mutant with two mutated residues (F518D, and Y519D). Since Y477 is a part of the ATP binding site, more precisely the A-loop, we decided for the variant with only two mutations, as well. For the pbp-out only a variant with three mutations was constructed (V682D, L697D, and Y700D).

Interestingly, neither of the variants were able to secrete HlyA anymore. It seems that HlyA cannot bind to the mutated binding sites, therefore, no secretion occurs. To examine ATPase activity of HlyB NBD variants, the corresponding mutants were generated, and purifications of all HlyB NBD variants were performed. Purification of the pbp-out NBD was not successful, although different expression strains and expression conditions were tested. The protein was in all cases expressed as inclusion bodies. This might indicate misfolding of the HlyB NBD resulting from the mutations in pbp-out.

Purification of the wild type NBD and both variants of pbp-in were performed in high yield and quality. It was not surprising that the pbp-in with three mutated residues showed no ATPase activity at all, because residue 477 is a part of the ATP binding site (A-loop) and mutation of this residue might affect the ATPase activity of the NBD. The results obtained from the ATPase activity experiment of pbp-in double mutations demonstrated an active NBD, despite lack of secretion for this mutant. What stands out from the secretion experiments and the ATPase activity assays,

TABLE 4 A list of heterologous proteins that can be secreted by the HlyA T1SS.

RTX protein	Host	Accession number	Uniprot entry link	Position of GG repeat	Position of AH	Distance GG-AH	Aromatic residue	Reference
HlyA	*E. coli*	P08715	https://www.uniprot.org/uniprotkb/P08715/entry	844	974	130	F990	
FrpA	*Neisseria meningitidis*	Q9K0K9	https://www.uniprot.org/uniprotkb/Q9K0K9/entry	1,125	1,257	132	F1274	Thompson and Sparling (1993)
FrpA	*Kingella kingae*	F5S7Z9	https://www.uniprot.org/uniprotkb/F5S7Z9/entry	600	720	120	F733	Erenburg (2022)
MbxA	*Mycobacterium bovis*	A7XER5	https://www.uniprot.org/uniprotkb/A7XER5/entry	756	886	130	F901	Erenburg (2022)
HlylA	*Actinobacillus pleuropneumoniae*	P55128	https://www.uniprot.org/uniprotkb/P55128/entry	840	969	129	F987	Gygi et al. (1990)
LktA	*Mannheimia haemolytica*	P0C083	https://www.uniprot.org/uniprotkb/P0C083/entry	779	909	130	F923	Highlander et al. (1990)
PaxA	*Pseudomonas aeruginosa*	Q9RCG8	https://www.uniprot.org/uniprotkb/Q9RCG8/entry	845	974	129	F991	Kuhnert et al. (2000)

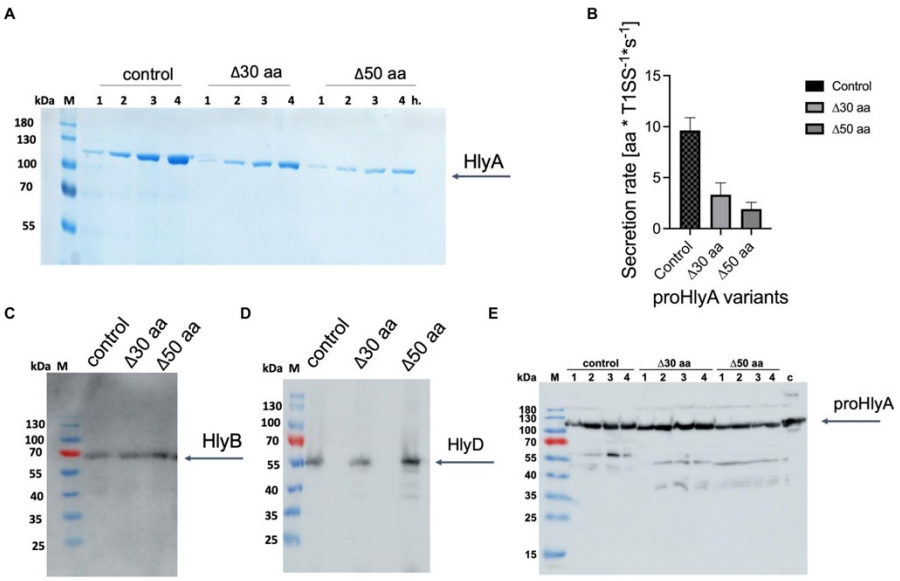


FIGURE 6
Secretion of proHlyA variants through the HlyA T1SS. **(A)** SDS-PAGE analysis of the supernatant (unconcentrated) of clones secreting wild type proHlyA (control), 30-residues-truncated proHlyA (Δ30 aa), and 50-residues-truncated proHlyA (Δ50 aa) through the HlyA T1SS. **(B)** Secretion rates of truncated proHlyA. Error bars represent the standard deviation of at least three biological replicates. Western blot analysis of *E. coli* cells demonstrated that the amounts of **(C)** HlyB, **(D)** HlyD or **(E)** proHlyA were expressed for cells secreting either WT proHlyA, 30-residues-truncated proHlyA, or 50-residues-truncated proHlyA in comparable levels. M, marker proteins; the molecular weight of the marker proteins is given on the left; x h., unconcentrated supernatant of culture, where x denotes the number of hours after induction.

the pbp-in and pbp-out are two binding sites on NBD for HlyA. In other word, the results of *in silico*, *in vitro* and *in vivo* studies are in accordance with each other and support the proposal that the NBD of HlyB harbors two pbp's in.

As discussed above, the second interaction between HlyB and HlyA occurs within the CLD. To assess if simultaneous HlyA binding to both areas, the NBD and the CLD, is possible, the distances between the domains known to interact with the C-terminal part of

HlyA were determined in a model of the full-length ABC transporter HlyB. These measurements suggest that a simultaneous binding of the substrate to two domains of the transporter is possible.

Interestingly, heterologous substrates that can be secreted by HlyA T1SS, also show AHs in their C-terminal secretion signal when analyzed with the prediction tool AmphipaSeeK. These heterologous substrates display the same distance between the predicted AH and a GG-repeat of the RTX domain as HlyA. A linker of conserved length between the two interaction sites further strengthens the hypothesis that the RTX domain and the secretion signal interact simultaneously with the ABC transporter, which might be important for the correct orientation of the substrate. Additionally, an aromatic residue (F990) close to the AH was shown to be important for HlyA secretion and a Phe residue close to the predicted AHs can be found in the heterologous substrates as well. Further mutagenesis studies showed the importance of this consensus length, as shortening this distance significantly reduced the rate of secretion. Albeit the remarkable alteration of the conserved linker length (by deletion of 30 or 50 residues), still secretion of truncated HlyA occurred in a reasonable titers. This clearly demonstrates that a simultaneous binding of the two binding sites is not strictly necessary, although it significantly affects the efficiency of secretion and thus the amount of secreted HlyA.

Conclusion

This study demonstrated that NBD of the ABC transporter HlyB harbors two binding sites for the substrate. The presence of these two putative binding pockets could be validated by *in silico*, *in vitro*, and *in vivo* investigations. The modeling of the HlyB structure allowed measuring distances between the domains known to interact with HlyA's C-terminal fragment. This measurement indicated that simultaneous binding of the substrate to two cytosolic domains (the CLD and the NBD) of the transporter is possible (Figure 7).

Last but not the least, we proposed that a consensus length between the GG repeats and the AH of the C-terminal HlyA is required for efficient secretion. The sequence analysis

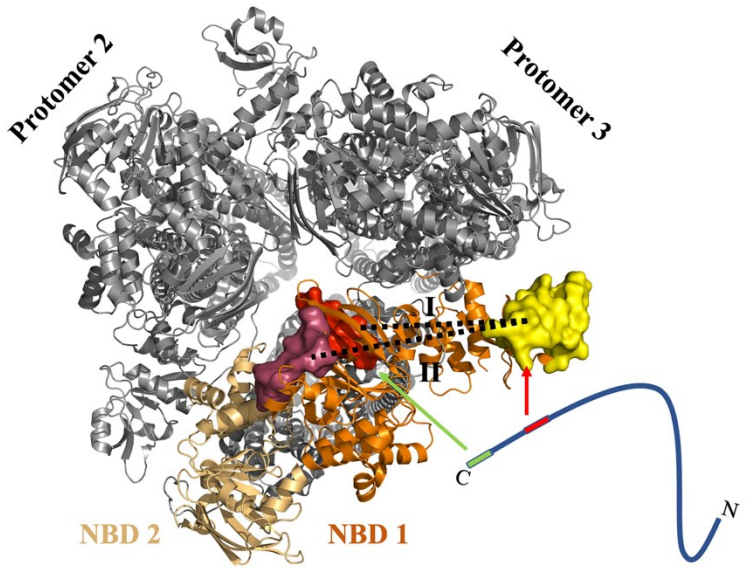


FIGURE 7
Proposed model of the interaction of HlyA with the NBD of HlyB. Cartoon representation of the trimer of dimers (PDB 7SGR) of the HlyB NBDs (Zhao et al., 2022). The view is from the cytosol toward the membrane. The NBDs of the transport-competent HlyB dimer are colored in orange and light orange, respectively, while the two non-transport competent HlyB dimers are shown in gray. The identified pbp's are shown as surface representation with pbp-in in red and pbp-out in magenta. The yellow surface maps the HlyA-interaction region of the CLD. The substrate, HlyA, is drawn schematically as a line. The two interacting regions of HlyA are highlighted in green (secretion signal) and red (RTX domain). The arrows emphasis the interaction regions within the NBD of HlyB. Please note that the CLD of NBD 2 was not traced in the single particle cryo EM structure (Zhao et al., 2022). For most efficient secretion, HlyA has to interact simultaneously with both regions of the NBD. In the proposed model, a 1:1 stoichiometry of HlyA and HlyB is assumed. However, one cannot exclude a 2:1 stoichiometry as demonstrated for PCAT1 (Kieuvongngam et al., 2020).

revealed that this consensus length is present in the heterologous RTX proteins secreted by the HlyA T1SS, while mutagenesis studies showed that shortening this distance reduces the secretion efficiency.

Data availability statement

All data are avaliable upon request from the corresponding author.

Author contributions

ZN: performing experiments and writing and editing the draft. EH: performing experiments and editing the draft. OS: performing experiments and writing the draft. SS: conceptualization and revising the draft for the content. LS: supervision, conceptualization, and revising the draft for the content. All authors contributed to the article and approved the submitted version.

Funding

This research was funded by Deutsche Forschungsgemeinschaft through CRC 1208 project A01 to LS and in part by CLIB Competence Center Biotechnology (CKB) funded by the European Regional Development Fund ERDF (34.EFRE-0300096) to ZN and LS. The Center for Structural Studies was funded by the Deutsche Forschungsgemeinschaft (DFG Grant number 417919780) to SS.

Acknowledgments

We thank all current and former members of the Institute of Biochemistry for support and fruitful discussions. We dedicate this manuscript to Karl-Erich Jaeger for his pioneering contributions in the field of molecular enzyme technology.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Abràmoff, M. D., Magalhães, P. J., and Ram, S. J. (2004). Image processing with ImageJ. *Biophoton. Int.* 11, 36–42.

Ambudkar, S. V., Kim, I.-W., Xia, D., and Sauna, Z. E. (2006). The A-loop, a novel conserved aromatic acid subdomain upstream of the Walker a motif in ABC transporters, is critical for ATP binding. *FEBS Lett.* 580, 1049–1055. doi: 10.1016/j.febslet.2005.12.051

Baumann, U., Wu, S., Flaherty, K. M., and Mckay, D. B. (1993). Three-dimensional structure of the alkaline protease of *Pseudomonas aeruginosa*: a two-domain protein with a calcium binding parallel beta roll motif. *EMBO J.* 12, 3357–3364. doi: 10.1002/j.1460-2075.1993.tb06009.x

Benabdelhak, H., Kiontke, S., Horn, C., Ernst, R., Blight, M. A., Holland, I. B., et al. (2003). A specific interaction between the NBD of the ABC-transporter HlyB and a C-terminal fragment of its transport substrate haemolysin a. *J. Mol. Biol.* 327, 1169–1179. doi: 10.1016/S0022-2836(03)00204-3

Bumba, L., Masin, J., Macek, P., Wald, T., Motlova, L., Bibova, I., et al. (2016). Calcium-driven folding of RTX domain β-rolls ratchets translocation of RTX proteins through type I secretion ducts. *Mol. Cell* 62, 47–62. doi: 10.1016/j.molcel.2016.03.018

Carrion-Vazquez, M., Marszalek, P. E., Oberhauser, A. F., and Fernandez, J. M. (1999). Atomic force microscopy captures length phenotypes in single proteins. *Proc. Natl. Acad. Sci.* 96, 11288–11292. doi: 10.1073/pnas.96.20.11288

Chervaux, C., and Holland, I. (1996). Random and directed mutagenesis to elucidate the functional importance of helix II and F-989 in the C-terminal secretion signal of *Escherichia coli* hemolysin. *J. Bacteriol.* 178, 1232–1236. doi: 10.1128/jb.178.4.1232-1236.1996

Combet, C., Blanchet, C., Geourjon, C., and Deleage, G. (2000). NPS@: network protein sequence analysis. *Trends Biochem. Sci.* 25, 147–150. doi: 10.1016/S0968-0004(99)01540-6

Crecca, C. R., and Roitberg, A. E. (2008). Using distances between α-carbons to predict protein structure. *Int. J. Quantum Chem.* 108, 2782–2792. doi: 10.1002/qua.21769

Delano, W. L. (2002). Pymol: an open-source molecular graphics tool. CCP4 Newsl. *Prot. Crystall.* 40, 82–92.

Dyballa, N., and Metzger, S. (2009). Fast and sensitive colloidal coomassie G-250 staining for proteins in polyacrylamide gels. *JoVE* e1431. doi: 10.3791/1431

Erenburg, I. N., Hansch, S., Chacko, F. M., Hamacher, A., Wintgens, S., Stuhldreier, F., et al. (2022). Heterologously secreted MbxA from Moraxella bovis induces a membrane blebbing response of the human host cell. *Sci. Rep.* 12:17825. doi: 10.1038/s41598-022-22480-x

Gygi, D., Nicolet, J., Frey, J., Cross, M., Koronakis, V., and Hughes, C. (1990). Isolation of the Actinobacillus pleuropneumoniae haemolysin gene and the activation and secretion of the prohaemolysin by the HlyC, HlyB and HlyD proteins of *Escherichia coli*. *Mol. Microbiol.* 4, 123–128. doi: 10.1111/j.1365-2958.1990.tb02021.x

Hagemans, D., Van Belzen, I. A., Morán Luengo, T., and Rüdiger, S. G. (2015). A script to highlight hydrophobicity and charge on protein surfaces. *Front. Mol. Biosci.* 2:56. doi: 10.3389/fmolb.2015.00056

Hardie, K., Issartel, J. P., Koronakis, E., Hughes, C., and Koronakis, V. (1991). In vitro activation of *Escherichia coli* prohaemolysin to the mature membrane-targeted toxin requires HlyC and a low molecular-weight cytosolic polypeptide. *Mol. Microbiol.* 5, 1669–1679. doi: 10.1111/j.1365-2958.1991.tb01914.x

Highlander, S. K., Engler, M. J., and Weinstock, G. M. (1990). Secretion and expression of the Pasteurella haemolytica leukotoxin. *J. Bacteriol.* 172, 2343–2350. doi: 10.1128/jb.172.5.2343-2350.1990

Holland, I. B., Peherstorfer, S., Kanonenberg, K., Lenders, M., Reimann, S., and Schmitt, L. (2016). Type I protein secretion-deceptively simple yet with a wide range of mechanistic variability across the family. *EcoSal Plus* 7, 1–46. doi: 10.1128/ecosalplus.ESP-0019-2015

Khosa, S., Scholz, R., Schwarz, C., Trilling, M., Hengel, H., Jaeger, K.-E., et al. (2018). An a/U-rich enhancer region is required for high-level protein secretion through the

hlya type i secretion system. *Appl. Environ. Microbiol.* 84, e01163–e01117. doi: 10.1128/AEM.01163-17

Kieuvongngam, V., Olinares, P. D. B., Palillo, A., Oldham, M. L., Chait, B. T., and Chen, J. (2020). Structural basis of substrate recognition by a polypeptide processing and secretion transporter. *elife* 9:e51492. doi: 10.7554/eLife.51492

Koronakis, V., Koronakis, E., and Hughes, C. (1989). Isolation and analysis of the C-terminal signal directing export of *Escherichia coli* hemolysin protein across both bacterial membranes. *EMBO J.* 8, 595–605. doi: 10.1002/j.1460-2075.1989.tb03414.x

Kuhnert, P., Heyberger-Meyer, B., Nicolet, J., and Frey, J. (2000). Characterization of PaxA and its operon: a cohemolytic RTX toxin determinant from pathogenic Pasteurella aerogenes. *Infect. Immun.* 68, 6–12. doi: 10.1128/IAI.68.1.6-12.2000

Lecher, J., Schwarz, C. K., Stoldt, M., Smits, S. H., Willbold, D., and Schmitt, L. (2012). An RTX transporter tethers its unfolded substrate during secretion via a unique N-terminal domain. *Structure* 20, 1778–1787. doi: 10.1016/j.str.2012.08.005

Lenders, M. H., Beer, T., Smits, S. H., and Schmitt, L. (2016). In vivo quantification of the secretion rates of the hemolysin a type I secretion system. *Sci. Rep.* 6:33275. doi: 10.1038/srep33275

Lenders, M. H., Weidtkamp-Peters, S., Kleinschrodt, D., Jaeger, K.-E., Smits, S. H., and Schmitt, L. (2015). Directionality of substrate translocation of the hemolysin a type I secretion system. *Sci. Rep.* 5:12470. doi: 10.1038/srep12470

Letoffe, S., Delepelaire, P., and Wandersman, C. (1996). Protein secretion in gram-negative bacteria: assembly of the three components of ABC protein-mediated exporters is ordered and promoted by substrate binding. *EMBO J.* 15, 5804–5811. doi: 10.1002/j.1460-2075.1996.tb00967.x

Lin, D. Y.-W., Huang, S., and Chen, J. (2015). Crystal structures of a polypeptide processing and secretion transporter. *Nature* 523, 425–430. doi: 10.1038/nature14623

Linhartová, I., Bumba, L., Mašín, J., Basler, M., Osička, R., Kamanová, J., et al. (2010). RTX proteins: a highly diverse family secreted by a common mechanism. *FEMS Microbiol. Rev.* 34, 1076–1112. doi: 10.1111/j.1574-6976.2010.00231.x

Madeira, F., Park, Y. M., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., et al. (2019). The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* 47, W636–W641. doi: 10.1093/nar/gkz268

Mol, A. R., Castro, M. S., and Fontes, W. (2018). NetWheels: a web application to create high quality peptide helical wheel and net projections. *BioRxiv* 416347. doi: 10.1101/416347

Mulnaes, D., Porta, N., Clemens, R., Apanasenko, I., Reiners, J., Gremer, L., et al. (2020). TopModel: template-based protein structure prediction at low sequence identity using top-down consensus and deep neural networks. *J. Chem. Theory Comput.* 16, 1953–1967. doi: 10.1021/acs.jctc.9b00825

Nicaud, J.-M., Mackman, N., Gray, L., and Holland, I. (1986). The C-terminal, 23 kDa peptide of *E. coli* haemolysin 2001 contains all the information necessary for its secretion by the haemolysin (Hly) export machinery. *FEBS Lett.* 204, 331–335. doi: 10.1016/0014-5793(86)80838-9

Noegel, A., Rdest, U., Springer, W., and Goebel, W. (1979). Plasmid cistrons controlling synthesis and excretion of the exotoxin α-haemolysin of *Escherichia coli*. *Mol. Gen. Genet. MGG* 175, 343–350. doi: 10.1007/BF00397234

Oswald, C., Holland, I. B., and Schmitt, L. (2006). The motor domains of ABC-transporters. *Naunyn Schmiedeberg's Arch. Pharmacol.* 372, 385–399. doi: 10.1007/s00210-005-0031-4

Oswald, C., Jenewein, S., Smits, S. H., Holland, I. B., and Schmitt, L. (2008). Water-mediated protein-fluorophore interactions modulate the affinity of an ABC-ATPase/TNP-ADP complex. *J. Struct. Biol.* 162, 85–93. doi: 10.1016/j.jsb.2007.11.006

Pourhassan, N. Z., Cui, H., Khosa, S., Davari, M. D., Jaeger, K. E., Smits, S. H., et al. (2022). Optimized Hemolysin type 1 secretion system in *Escherichia coli* by directed evolution of the Hly enhancer fragment and including a terminator region. *Chembiochem* 23:e202100702. doi: 10.1002/cbic.202100702

Pourhassan, Z., Smits, S. H., Ahn, J. H., and Schmitt, L. (2021). Biotechnological applications of type 1 secretion systems. *Biotechnol. Adv.* 53:107864. doi: 10.1016/j.biotechadv.2021.107864

Raschka, S. L., Harris, A., Luisi, B. F., and Schmitt, L. (2022). Flipping and other astonishing transporter dance moves in fungal drug resistance. *BioEssays* 44:2200035. doi: 10.1002/bies.202200035

Rath, A., Glibowicka, M., Nadeau, V. G., Chen, G., and Deber, C. M. (2009). Detergent binding explains anomalous SDS-PAGE migration of membrane proteins. *Proc. Natl. Acad. Sci.* 106, 1760–1765. doi: 10.1073/pnas.0813167106

Sapay, N., Guermeur, Y., and Deléage, G. (2006). Prediction of amphipathic in-plane membrane anchors in monotopic proteins using a SVM classifier. *BMC Bioinformatics* 7, 1–11. doi: 10.1186/1471-2105-7-255

Schmitt, L., Benabdelhak, H., Blight, M. A., Holland, I. B., and Stubbs, M. T. (2003). Crystal structure of the nucleotide-binding domain of the ABC-transporter haemolysin B: identification of a variable region within ABC helical domains. *J. Mol. Biol.* 330, 333–342. doi: 10.1016/S0022-2836(03)00592-8

Shi, Y., Mowery, R. A., Ashley, J., Hentz, M., Ramirez, A. J., Bilgicer, B., et al. (2012). Abnormal SDS-PAGE migration of cytosolic proteins can identify domains and mechanisms that control surfactant binding. *Protein Sci.* 21, 1197–1209. doi: 10.1002/pro.2107

Spitz, O., Erenburg, I. N., Kanonenberg, K., Peherstorfer, S., Lenders, M. H., Reiners, J., et al. (2022). Identity determinants of the translocation signal for a type 1 secretion system. *Front. Physiol.* 12:2585–2598. doi: 10.3389/fphys.2021.804646

Thanabalu, T., Koronakis, E., Hughes, C., and Koronakis, V. (1998). Substrate-induced assembly of a contiguous channel for protein export from E. coli: reversible bridging of an inner-membrane translocase to an outer membrane exit pore. *EMBO J.* 17, 6487–6496. doi: 10.1093/emboj/17.22.6487

Thomas, S., Holland, I. B., and Schmitt, L. (2014). The type 1 secretion pathway—the hemolysin system and beyond. Biochimica et Biophysica Acta (BBA)-molecular. *Cell Res.* 1843, 1629–1641. doi: 10.1016/j.bbamcr.2013.09.017

Thompson, S. A., and Sparling, P. (1993). The RTX cytotoxin-related FrpA protein of *Neisseria meningitidis* is secreted extracellularly by meningococci and by HlyBD+ *Escherichia coli*. *Infect. Immun.* 61, 2906–2911. doi: 10.1128/iai.61.7.2906-2911.1993

Zaitseva, J., Holland, I. B., and Schmitt, L. (2004). The role of CAPS buffer in expanding the crystallization space of the nucleotide-binding domain of the ABC transporter haemolysin B from *Escherichia coli*. *Acta Crystallogr. D Biol. Crystallogr.* 60, 1076–1084. doi: 10.1107/S0907444904007437

Zaitseva, J., Jenewein, S., Jumpertz, T., Holland, I. B., and Schmitt, L. (2005a). H662 is the linchpin of ATP hydrolysis in the nucleotide-binding domain of the ABC transporter HlyB. *EMBO J.* 24, 1901–1910. doi: 10.1038/sj.emboj.7600657

Zaitseva, J., Jenewein, S., Wiedenmann, A., Benabdelhak, H., Holland, I. B., and Schmitt, L. (2005b). Functional characterization and ATP-induced dimerization of the isolated ABC-domain of the haemolysin B transporter. *Biochemistry* 44, 9680–9690. doi: 10.1021/bi0506122

Zhao, H., Lee, J., and Chen, J. (2022). The hemolysin a secretion system is a multi-engine pump containing three ABC transporters. *Cells* 185, 3329–3340 e13. doi: 10.1016/j.cell.2022.07.017

Zimmermann, L., Stephens, A., Nam, S.-Z., Rau, D., Kübler, J., Lozajic, M., et al. (2018). A completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its core. *J. Mol. Biol.* 430, 2237–2243. doi: 10.1016/j.jmb.2017.12.007

2.4 Chapter IV: Engineering of the inner membrane complex of the HlyA T1SS

Title: KnowVolution Campaign of the ABC Transporter Hemolysin B Increases the Secretion through HlyA type 1 Secretion System

Authors: Zohreh Pourhassan N., Haiyang Cui, Neele Muckhoff, Mehdi D. Davari, Sander H. J. Smits, Ulrich Schwaneberg, and Lutz Schmitt

Submitted in: Applied Microbiology and Biotechnology

Impact factor (2023): 5.56

Own proportion of this work: 60%

KnowVolution Campaign of the ABC Transporter Hemolysin B Increases the Secretion through HlyA type 1 Secretion System

Zohreh Pourhassan N.[a], Haiyang Cui[b,c,§], Neele Muckhoff[a], Mehdi D. Davari[d], Sander H. J. Smits[a], Ulrich Schwaneberg[b,c], and Lutz Schmitt[a,*]

[a:] Institute of Biochemistry, Heinrich Heine University, Universitätsstr. 1, 40225 Düsseldorf

[b:] Institute of Biotechnology, RWTH Aachen University, Worringerweg 3, 52074, Aachen (Germany)

[c:] DWI-Leibniz Institute for Interactive Materials, Forckenbeckstraße 50, 52056, Aachen (Germany)

[d:] Department of Bioorganic Chemistry, Leibniz Institute of Plant Biochemistry, Weinberg 3, 06120, Halle (Germany)

[§:] present address; Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign, 1206 West Gregory Drive, Urbana, IL 61801 (USA)

*: to whom correspondence should be addressed
Lutz Schmitt
Phone: +49(0)211-81-10773
Email: lutz.schmitt@hhu.de

Abstract

Secretion of proteins into the extracellular space has great advantages for the production of recombinant proteins. Type 1 secretion systems (T1SS) are attractive candidates to be optimized for biotechnological applications, as they have a relatively simple architecture compared to other classes of secretion systems. In this study, we report a KnowVolution campaign on improving the secretion efficiency of the hemolysin A type 1 secretion system (HlyA T1SS) from *Escherichia coli*. After the screening of around 5760 variants in four KnowVolution phases, an enhanced HlyB variant containing four substitutions (T36L/F216W/S290C/V421I) was obtained with an up to 2.5-fold improved secretion for two hydrolases, a lipase and a cutinase. This study represents a step forward to advance protein secretion using a T1SS to make *E. coli* more competitive for its applications in industrial biotechnology.

Key words: Protein secretion, KnowVolution, Gram-negative bacteria, protein engineering, directed evolution, bacterial secretion system

Introduction

Secretion of proteins is important for recombinant protein production as it can lead to protein production in high quality and yield, while at same time reducing the production costs (Kleiner-Grote et al., 2018). Protein secretion in Gram-negative bacteria, such as *Escherichia coli*, often results in low titers. This limitation is of particular concern and reduces the general utilities of biotechnological hosts, such as *E. coli*, as secreting hosts (Kleiner-Grote et al., 2018, Burdette et al., 2018).

To enhance the secretion efficiency of different classes of secretion systems, different approaches have been applied. For instance, this approaches include modification of secretion signal (Freudl, 2018), transferring the secretion system to a more efficient secretion host (Eom et al., 2014), modification of genetical elements such as enhancer fragment (Pourhassan et al., 2022a). In addition to these approaches, protein engineering has proved to be a powerful tool for tailoring protein properties (Wong et al., 2006, Pramanik et al., 2021). Until now, different secretion systems have been subjected to various engineering approaches, but these efforts have commonly had little success due to the complexity of the secretion systems (Burdette et al., 2018).

Among different secretion systems of Gram-negative bacteria, type 1 secretion systems (T1SS) have relatively simple architecture, and this makes them attractive targets to be optimized (Burdette et al., 2018). T1SS translocates proteins in an unfolded state across both membranes of Gram-negative bacteria in a single step. Subsequently, proteins fold in their functional state in the extracellular space (Holland et al., 2016).

The T1SS paradigm is the hemolysin A type 1 secretion system (HlyA T1SS). The HlyA T1SS machinery consists of three membrane proteins: the ABC transporter HlyB, the membrane fusion protein HlyD, and the outer membrane protein TolC. TolC is recruited upon an interaction between the substrate (HlyA) and the inner-membrane complex. Then, these membrane proteins form a tunnel-channel redirecting HlyA from the cytosol to the extracellular space (Holland et al., 2016, Pourhassan et al., 2021).

Recently, the single particle cryo-EM structure of HlyB/HlyD complex revealed that a trimer of HlyB dimers, each associated with two molecules of HlyD (per HlyB dimer), forms the inner membrane of the system complex (Zhao et al., 2022).

HlyB is an ATP-binding cassette (ABC) transporter that powers the secretion of HlyA by hydrolyzing ATP. HlyB has a domain organization consists of a transmembrane domain (TMD) composed of six transmembrane helices (TMH), a hydrophilic cytoplasmic

nucleotide-binding domain (NBD), and an N-terminal cytosolic domain called the C39-peptidase like domain (CLD) (Schmitt et al., 2003, Lecher et al., 2012). In addition to its ATPase activity, NBD has a supplementary function. Specifically, it interacts with HlyA prior to secretion, which is proposed to represent the substrate recognition step (Benabdelhak et al., 2003, Pourhassan et al., 2022b). Based on a series of pull-down assays, Lecher *et al.* suggested that the CLD is involved in positioning HlyA into the secretion pathway, indicating a chaperone-like activity for this domain (Lecher et al., 2012).

HlyD is a membrane fusion protein that interacts with HlyB and TolC to stabilize and seal the tunnel-channel of HlyA T1SS (Holland et al., 2016). In terms of topology, HlyD has an N-terminal cytosolic domain (60 residues) that follows by a single transmembrane helix (20 residues), and a large periplasmic region (Schtilein et al., 1992). Less is known about HlyD compared to other membrane proteins of the system.

Within various protein engineering strategies, KnowVolution is a general applicable protein engineering strategy to improve protein properties (Cheng et al., 2015). KnowVolution combines directed evolution and computational analysis in order to maximize improvements with a minimal experimental effort. A KnowVolution campaign is divided into four phases: I) Identification; II) Determination; III) Computational analysis; IV) Recombination (Pramanik et al., 2021, Yang et al., 2015). KnowVolution has been reported for the successful engineering of *e.g.* the ligase PigC (Brands et al., 2021), laccases (Novoa et al., 2019), aryl sulfotransferases (Islam et al., 2018), and polymer-binding peptides (Rübsam et al., 2018).

HlyA1 or HlyAc, a 23-kDa C-terminal fragment of HlyA containing the secretion signal plus three GG repeats, is a carrier for the secretion of heterologous proteins (Pourhassan et al., 2021). While it has been successful in secreting various proteins, its potential as a universal secretion platform is limited by low secretion titers. Improving the efficiency of HlyA T1SS for recombinant proteins remains a challenge (Khosa et al., 2018).

To optimize the secretion efficiency of the system, we recently worked on the HlyA enhancer fragment and have achieved great success, which promoted us to a further improvement of the system (Pourhassan et al., 2022a). In the current study, we report a two-round based KnowVolution campaigns on the inner membrane complex of HlyA T1SS (HlyB and HlyD) to achieve further improved HlyA T1SS secretion. The KnowVolution on HlyB yielded a variant with four beneficial substitutions that increased the secretion efficiency of lipase for

2.5-folds. In addition, the general applicability of the approach was also investigated with a cutinase as a second model system.

Results

Overall, two KnowVolution campaigns were employed on the HlyA T1SS of *E. coli*, which resulted in an increased secretion. The first part of the results section describes the four phases of the KnowVolution campaign on HlyB. Next, we explain the results of the secretion experiments using the improved variant in shaking flasks. Then, we report the data on the use of the novel HlyB for the secretion of another targeted hydrolase, a cutinase. Encouraged by these results, we also conducted a KnowVolution campaign on HlyD to further improve the secretion efficiency (see the final results section).

KnowVolution of HlyB toward increasing secretion efficiency

To identify amino acid positions of the HlyB transporter that increase the secretion efficiency of HlyA T1SS, a KnowVolution campaign, in 4 phases was conducted (see **Figure 1)**:

Phase I (Identification): In the first phase of KnowVolution campaign on HlyB, beneficial positions of the *hly*B gene were identified through a random mutagenesis approach known as casting error-prone PCR (cepPCR) according to the published protocol (Yang et al., 2017). In cepPCR approach, the gene of interest is divided into smaller fragments, and subsequently, fragments were subjected to epPCR individually. To generate a library on HlyB using cepPCR approach the gene of *hlyB* was divided into six fragments (F_{1B}: 275 bp, F_{2B}: 251 bp, F_{3B}: 283 bp, F_{4B}: 256 bp, F_{5B}:279 bp, F_{6B}: 291 bp). The region of *hlyB* gene encoding the NBD was excluded. Excluding the region of the *hlyB* gene that encodes the NBD was expected to reduce the number of inactive clones in the resulting library, as this region contains highly conserved motifs that are essential for transporter function.

Random mutagenesis libraries of those six fragments were generated using a low-fidelity DNA polymerase in buffer conditions with different $MnCl_2$ concentrations (ranging from 0.1 to 0.9 mM). The mutation frequency was adjusted by altering the $MnCl_2$ concentration to obtain libraries with an inactive ratio of 40-50%. The inactive ratio of libraries was determined via prescreening clones on tributyrin agar plates, as active clones were able to generate halos because of the secreted lipase-HlyA1, which was used as model substrate. The mutation load of libraries was in the range of 3.6 to 5.18 mutations per fragment (on average almost 15 mutations per kbp).

In the first phase, ~ 4576 clones were screened in a 96-well MTP format using the *para*-nitrophenyl- butyrate (*p*NPB) assay to identify *E. coli* clones that secrete the lipase with

higher titers. The promising variants were rescreenind four times, and those with increased secretion levels were sent for sequencing. Despite generating and screening of a larger library for the fifth fragment, we did not identified any beneficial position. Upon aligning the sequence of the promising variants, it was observed that twelve positions were repeatedly mutated, which were considered as "potential beneficial positions". These positions include: Thr 36, Ser 71, Val 108, Leu 164, Phe 216, Glu 226, Ser 272, Ser 290, Ile 407, Val 421, Phe 434, and Asn 465.

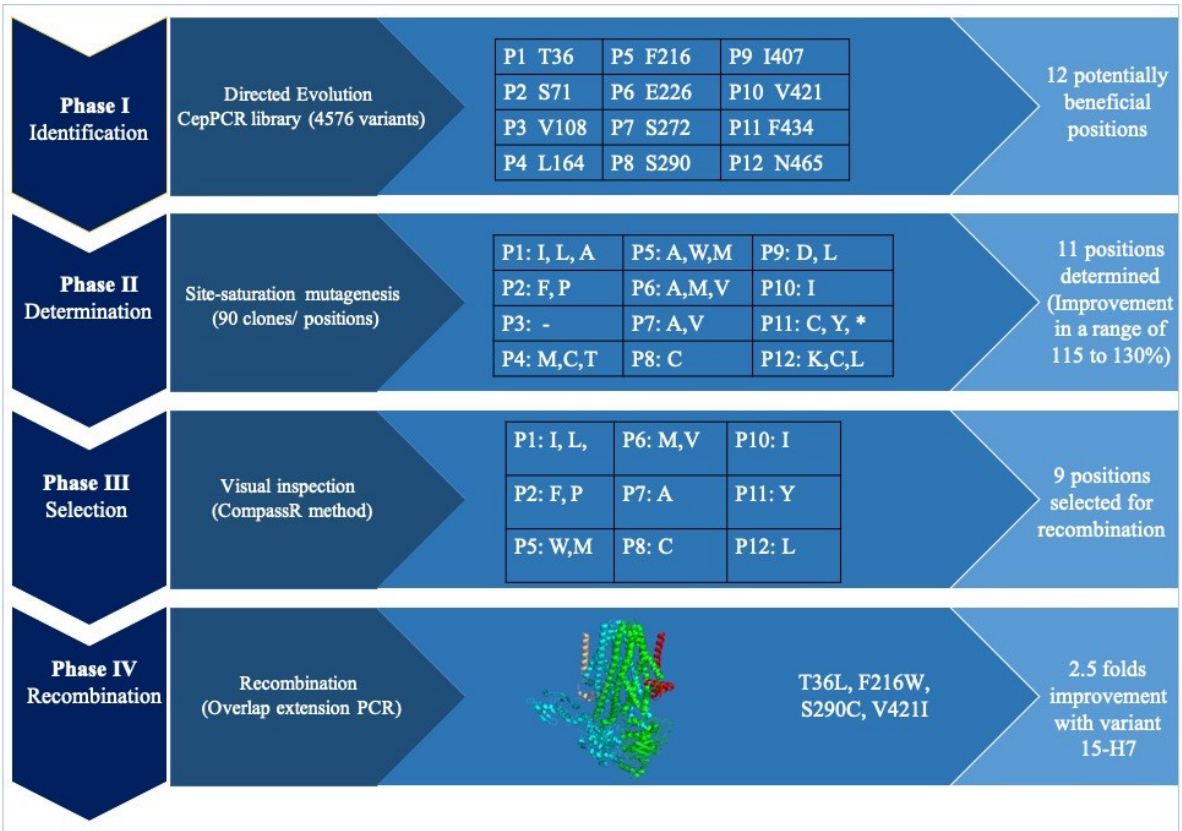


Figure 1: An overview of the HlyB KnowVolution campaign aimed at increasing the secretion efficiency of the HlyA T1SS. In phase I, the hlyB gene underwent random mutagenesis, resulting in the identification of 12 potential beneficial positions. In phase II, all 12 positions were subjected to site saturation mutagenesis. n phase III, the computer-assisted recombination (CompassR) strategy was used to select substitutions for recombination based on the HlyB homology model. Finally, in phase IV, recombination was performed resulting in a variant harboring four substitutions (T36L, F216W, S290C, V421I). This variant showed a 2.5-fold improvement in secretion efficiency compared to the wildtype.

Phase II (Determination): To obtain the full diversity of those 12 identified positions, individual site-saturation mutagenesis were performed on each position to identify variants with increased secretion levels. A total of 96 variants were screened for each position to ensure that all possible amino acids were represented. The promising variants were rescreened four times and the truly improved ones were sent for sequencing. **Table 1**

summarizes the substitutions that exhibited higher secretion efficiency compared to the wildtype.

Table 1: Sequence analysis of saturated positions in HlyB variants

Positions	Substitutions with improved secretion of targeted lipases	Compatible substitutions suggested by CompassR
Position 1: T36	**Ile***, Leu, Ala	Ile and Leu
Position 2: S71	**Phe***, Pro	Pro and Phe
Position 3: V108	**-**	-
Position 4: L164	Met, Cys, **Thr***	-
Position 5: F216	Ala, Trp, **Met***	Met and Trp
Position 6: E226	Ala, Met, **Val***	Met and Val
Position 7: S272	**Ala***, Val	Ala
Position 8: S290	Cys	Cys
Position 9: I407	Asp, **Leu***	-
Position 10: V421	Ile	Ile
Position 11: F434	Cys, Tyr, **a silent mutation***	Tyr
Position 12: N465	**Lys***, Cys, Leu	Leu

Amino acids highlighted in bold showed higher secretion efficiency compared to other substitutions.

Phase III (Selection): No three-dimensional structure of HlyB was available during the experimental phase of the study. Therefore, we created a homology model based on the crystal structure of the ABC transporter PCAT1 (Lin et al., 2015) to performed the calculations. However, a single particle cryo-EM structure of HlyB/HlyD was recently published (Zhao et al., 2022). A comparison between the homology model and the structure used for visualization of the identified beneficial positions (**Figure 2 and 3**) is provided in the Materials and Method section.

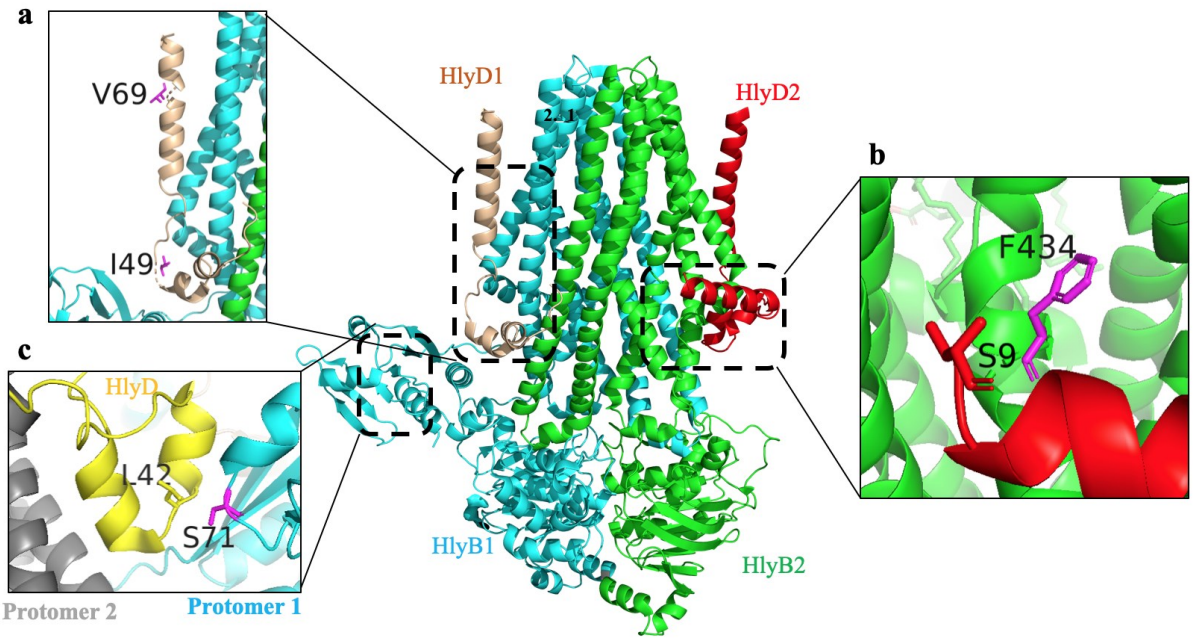


Figure 2: An illustration of the potential beneficial positions in the three-dimensional structure of HlyB wildtype (PDB: 7SGR). HlyB dimer is in cyan and green. **a**) Ile49 of HlyD is among the residues involved in the interaction with the positively charged region of the CLD. **b)** F434 of HlyB is located in close proximity of the amphipathic helix of HlyD. The amphipathic helix is oriented parallel to the membrane. **c)** Residue S71 is located in the protomer-protomer interface. The figure was generated using the PyMol software (www.pymol.org).

To investigate whether the identified beneficial positions exhibit cooperative effects, we explored the recombination of these residues. However, recombining multiple beneficial positions is a major challenge in protein engineering, as it often leads to variants with reduced or even no activity after recombining three or four substitutions.
(Liebeton et al., 2000). To address this challenge, we employed a computer-assisted recombination (CompassR) strategy, which has been previously validated for improving the performance of the BSLA lipase (Cui et al., 2020). The CompassR strategy relies on the calculation of the relative free energy of folding ($\Delta\Delta G_{fold}$) of different substitutions, and enables the selection of substitutions that can be combined for improved protein function. **Table 2** summarizes the calculated $\Delta\Delta G_{fold}$ values for HlyB substitutions. Based on the CompassR selection criteria ($\Delta\Delta G_{fold} < 0.36$), we identified nine beneficial substitutions (**Table 1**) that were recombined in phase IV.

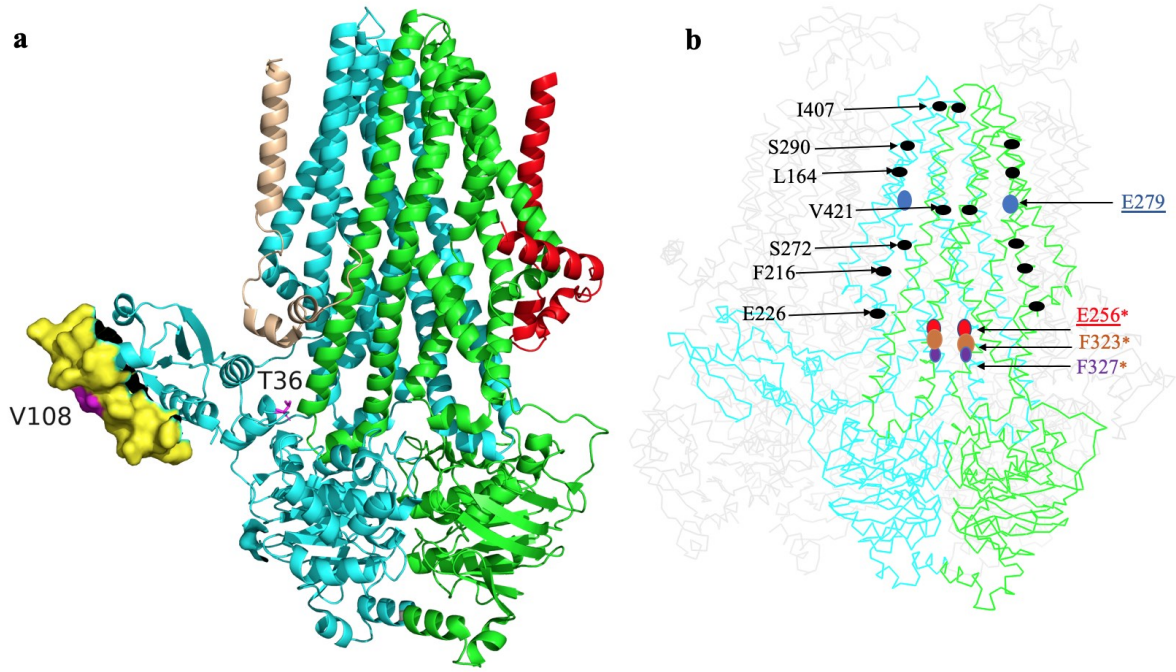


Figure 3: An illustration of the potential beneficial positions in the three-dimensional structure of the HlyB wildtype (PDB: 7SGR). **a**) carton representation of HlyB dimer (cyan and green) associated with two HlyD molecules (light brown and red), with the yellow surface mapping the HlyA-interaction region on the CLD (Lecher et al., 2012). The magenta color highlights residue 108, which is located within this interaction region. **b**) An illustration of the identified beneficial positions determined in the current study as black dots. The positions suggested by Holand *et al.* involved in oligomer formation are underlined (Holland et al., 2016), and the positions at the opening of HlyB suggested by Zhao *et al.* (Zhao et al., 2022) are highlighted by an asterisk. The figure was generated using PyMol software (www.pymol.org).

Phase IV (Recombination): Recombination between selected substitutions from phase III was accomplished through an overlap extension PCR according to a published protocol (Hilgarth and Lanigan, 2020). The resulting library of recombinant variants estimated to have 1500 different combinations. Variants showing higher secretion levels of lipase were isolated and rescreened four times. The screening of the recombinant library yielded a variant, termed as 15-H7, with an improvement of around 2.5-folds in the secretion of lipase-HlyA1 (**Figure 4a**). This variant contained 4 substitutions: T36L, F216W, S290C, and V421I.

Table 2: The $\Delta\Delta G_{fold}$ of substitutions and their effect on HlyB. The selected substitutions based on the CompassR criteria ($\Delta\Delta G_{fold}$<0.36) are highlighted in green. The larger the $\Delta\Delta G_{fold}$ negative values, the higher the stability.

Substitution	$\Delta\Delta G_{fold}$ (kcal/mol)	Effect of the substitution
T36A	0.68	Slightly destabilizing
T36I	-4.54	Stabilizing
T36L	-2.47	Stabilizing
S71P	-0.89	Stabilizing
S71F	-3.62	Stabilizing
L164M	1.55	Slightly destabilizing
L164C	3.67	highly destabilizing
L164T	3.95	highly destabilizing
F216A	0.82	Slightly destabilizing
F216M	-1.68	Stabilizing
E226A	1.24	Slightly destabilizing
E226M	-2.31	Stabilizing
E226V	0.12	neutral effect
S272A	-0.35	neutral effect
S272V	0.6	Slightly destabilizing
S290C	-2.03	Stabilizing
I407D	4.21	highly destabilizing
I407L	0.6	Slightly destabilizing
V421I	-1.84	Stabilizing
F434C	2.54	highly destabilizing
F434Y	0.01	neutral effect
N465K	0.56	Slightly destabilizing
N465C	1.16	Slightly destabilizing
N465L	-0.95	Stabilizing

Comparing the secretion level of HlyB wildtype and 15-H7 variant

A test secretion experiment was conducted in 100 mL Erlenmeyer flasks to compare the secretion levels of the 15-H7 variant (the best variant obtained from the HlyB KnowVolution campaign) and the wildtype. The unconcentrated supernatant of both cultures was analyzed by SDS-PAGE (**Figure 4**). Prior to loading on the SDS-PAGE, the samples were normalized based on OD_{600} of the cultures. A semi-quantification of the SDS-PAGE gel using the ImageJ software (Collins, 2007) confirmed approximately a 2.5-fold increase in the secretion efficiency of the 15-H7 variant of compared to the wildtype at one hour after induction. However, the improvement decreased to 1.5-fold after the second hour due to

aggregate formation, as evident from the formation of white foam in the supernatant of the 15-H7 culture (**Figure 4b**).

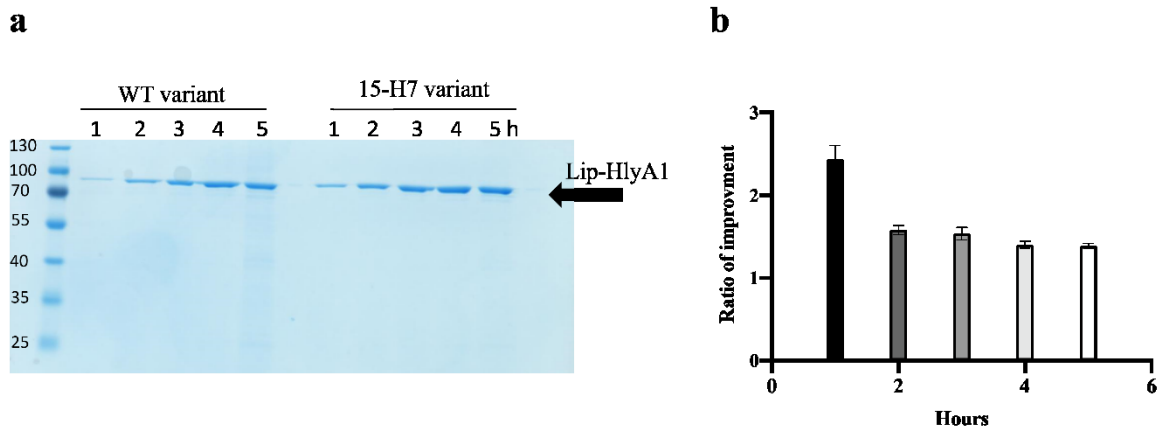


Figure 4: Secretion of lipase-HlyA1 through the HlyA T1SS. a) SDS-PAGE of the unconcentrated supernatant of wildtype and the 15-H7 variant secreting lipase-HlyA1. b) Semi quantification of SDS-PAGE gel via the ImageJ software [44]. Ratio of the improvements of 15-H7 variant compared to the wildtype. The error bars are based on three independent replicates with the error bars reporting S.D. The molecular weight of the marker proteins is given on the left (kDa). xh., x hours after induction.

In addition, the supernatant samples collected every hour after induction were used to perform the pNPB assay to determine lipase activity. The initial activities were measured spectrophotometrically at 410 nm. The pNPB assay results confirmed the SDS-PAGE analysis and showed that the secretion level of 15-H7 was higher than the wildtype (**Figure 5a, b**). The improvement in secretion was higher in the early hours after induction, with the amount of soluble enzyme in the 15-H7 culture being 2.5-fold higher than in the wildtype during the first hour after induction. However, the amount of secreted Lipase-HlyA1 in the 15-H7 culture reached a concentration that promoted aggregation, visible as white particles.

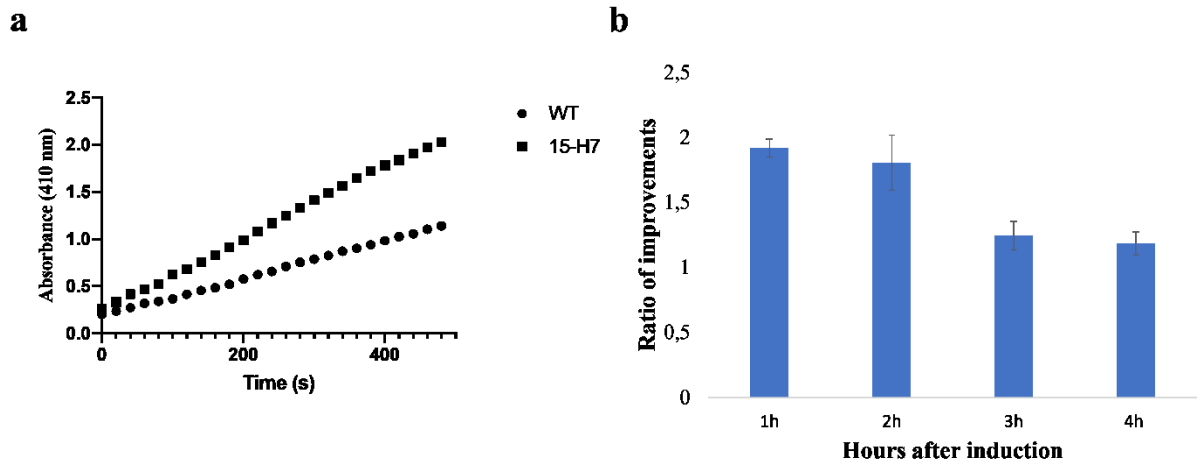


Figure 5: Enhanced secretion of lipaase-HlyA1 by the 15-H7 variant. a) Activity of the lipase-HlyA1 in the supernatant of the culture dedicated to either the "15-H7" or wildtype HlyB using the *p*NPB assay. b) Ratio of improved lipase activity of the wildtype and the improved variants using the *p*NPB assay. The reported values and calculated error bars are based on at least three independent replicates.

The level of expression of HlyB (**Figure 6a**) and HlyD (**Figure 6b**) was analyzed by Western blotting for *E. coli* cells having either the wildtype or the 15-H7 variants. This analysis performed in three biological replicates, and confirmed that HlyB and HlyD were expressed at comparable levels in both variants. Thus, the improved secretion efficiency of the 15-H7 variant cannot be attributed to a higher expression of the HlyA secretion system apparatus.

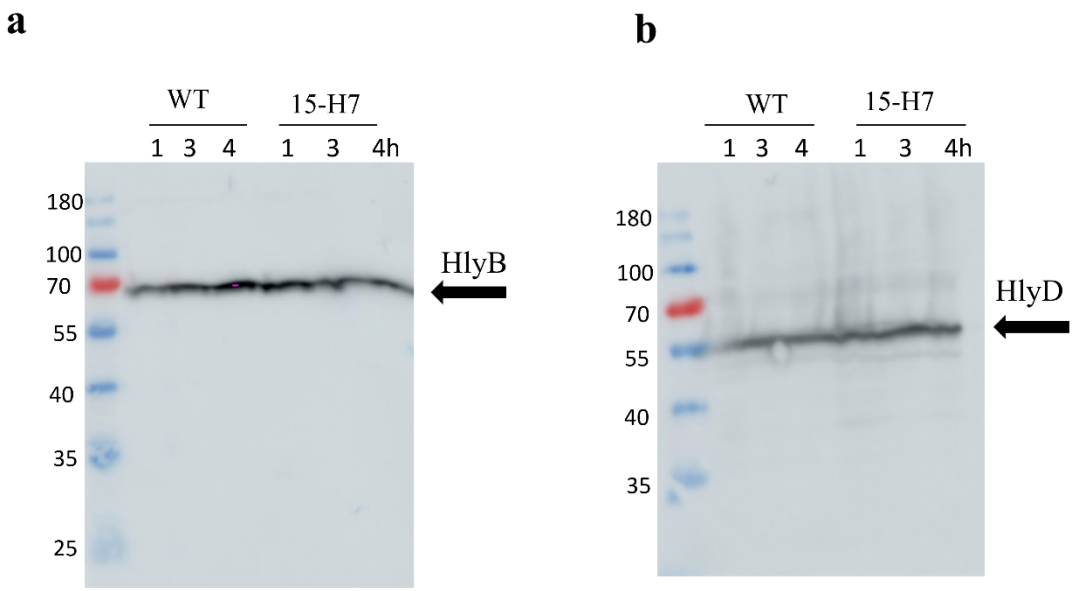


Figure 6: Western blot analysis of *E. coli* cells secreting lipaase-HlyA1 through the HlyA T1SS. **a**) Western blot of HlyB and **b**) Western blot of HlyD using polyclonal antibodies against HlyB and HlyD, respectively. The blots indicate that both membrane proteins were expressed for both variants at comparable levels. The molecular weight of the marker proteins is given on the left; xh: cell pellets of culture, where x denotes the number of hours after induction.

To determine the amount of secreted lipase-HlyA1 in the 15-H7 culture, a serial dilution of purified HlyA with known concentration was prepared and a calibration curve was generated. The concentration of secreted lipase-HlyA1 in the supernatant was quantified to approximately 400 mg per liter. The amount of lipase-HlyA1 (wet weight) collected in the white particles was estimated to 5 mg from 25 mL culture (200 mg/L).

Improved secretion of cutinase-HlyA1 through HlyA T1SS

We were eager to see if the 15-H7 HlyB variant could promote the secretion of a different target protein rather than lipase. To do this, a fusion gene encoding the cutinase protein from

Fusarium solani pisi was cloned instead of the lipase gene in the plasmid harboring the 15-H7 HlyB variant. The test expression was conducted for wildtype and 15-H7 variant, and the supernatant of the cultures was analyzed via Western blotting. This results showed that the secretion of cutinase-HlyA1 in the presence of the 15-H7 HlyB variant was 2.5-fold higher than the wildtype one. This observation demonstrated that the novel HlyB is capable of enhancing the secretion of another target protein, cutinase-HlyA1 (**Figure S1**).

KnowVolution of HlyD toward increasing secretion efficiency

Following the success in improving HlyB secretion efficiency, we applied the KnowVolution strategy to HlyD in an attempt to further enhance the secretion efficiency of the HlyA T1SS system. The KnowVolution strategy for directed evolution of HlyD was the same as already explained for HlyB.

In the Identification step (Phase I), we identified three potential beneficial positions located in the cytoplasmic domain of HlyD: Met6, Ile49, Val69. However, despite screening site-saturation libraries for these positions, no promising variant was found.

We attempted to regenerate the mutation libraries on fragments 1-2 to cover more beneficial variants, but still no positions were determined. This led us to conclude that either HlyD may require multiple substitutions simultaneously for better secretion or that the wild-type HlyD represents the best naturally evolved variant for interactions with HlyB and TolC, which cannot be further optimized through protein engineering approaches.

Discussion

Secretion, as a strategy of recombinant protein production, could offers several advantages such as increased stability and solubility of the target protein, as well as a significant reduction in the production costs due to simplified downstream processes (Kleiner-Grote et al., 2018, Burdette et al., 2018).

Note almost one-third of all FDA-approved recombinant proteins are produced by *E. coli* (Puetz and Wurm, 2019, Sharma and Chaudhuri, 2017), but to our knowledge, there is no example of usage of *E. coli* as a secretion host in industrial scales. Despite the benefits of using *E. coli* in large-scale protein production, the low secretion levels have posed a challenge. To address this issue, researchers have attempted to engineer different secretion systems of *E. coli,* but with little success so far (Pourhassan et al., 2021, Burdette et al., 2018).

T1SS have a relatively simple architecture among other secretion systems of Gram-negative bacteria (Costa et al., 2015), making this class of secretion systems attractive targets for engineering efforts. The best achieved secretion titer for T1SS in different Gram-negative bacteria has been reported for metalloprotease secreted by TliDEF, a T1SS from *Pseudomonas Fluorescens*, with a yield of 789 mg/mL (Burdette et al., 2018, Pourhassan et al., 2021).

Previously, we reported on a successful improvement of the secretion efficiency of the HlyA T1SS through a KnowVolution approach on the HlyA enhancer fragment, as well as including a terminator region. Our previous work resulted in a secretion level of approximately 180 mg/L for soluble lipase-HlyA1. Additionally, the white particles in a small culture volume of only 25 mL had a high protein content estimated to be above 1.5 mg of lipase-HlyA1.

In the current work, we applied KnowVolution strategy on the inner membrane complex of the system to further optimize secretion of target proteins, resulting in significant improvements in secretion efficiency. Our T1SS system achieved now more than 400 mg/L soluble lipase plus around 5 mg in form of aggregated enzyme in a culture volume of only 25 mL through a KnowVolution strategy. This indicates that putting more efforts on the HlyA system is worth and could address the current drawbacks of using them, as there might be still room for improvement in efficiency of the HlyA system.

In contrast to previous KnowVolution campaigns that used epPCR or SeSAM (Ji et al., 2020) for diversity generation, we employed a cepPCR method (Yang et al., 2017) that enables us

to adjust mutation frequencies to small peptide stretches. This approach led to the identification of several beneficial amino acid positions within the cepPCR library for HlyB, with 12 positions distributed in the CLD and TMD. Furthermore, the CompassR strategy (Cui et al., 2020) was performed in the phase III of the HlyB KnowVolution compaign. The best variant after all four phases contained four mutations: T36L/F216W/S290C,/V421I.

During the KnowVolution campaign, three positions (T36, S71, and V108) were identified located in the CLD, which is the domain that interacts with HlyA before secretion (Lecher et al., 2012). Recently, Zhao et al. showed that an interaction between HlyD and the CLD is essential for secretion. They also defined the main interface between the three protomer, which is a projection between the cytosolic domain of HlyD, residues 35-50, with a positively charged surface of the CLD. Interestingly, one of the positions we identified as potentially beneficial for HlyD, Il49, is located within this region of HlyD. In addition, another identified possible beneficial of HlyD, V69, is involved in making contact with TM1/2 of HlyB. However, site saturation of these positions did not result in any improved variants.

One potential position of HlyB, S71, is located in the inter-protomer interface in close proximity to the residues involved in the HlyD/CLD interaction. In the work by Zhao *et al.*, the S71C mutation was tested and led to the conclusion that oligomerization between HlyB and HlyD occurs in the native environment, which is a prerequisite for secretion (Zhao et al., 2022).

Previous reports have indicated that the central part of the CLD interaction site is composed of hydrophobic residues, and upon interaction with substrate, chemical shifts were observed in the first α-helix and β-sheets number 3 to 5 of the CLD (Lecher et al., 2012). Notably, the identified potential beneficial position V108 is located within the β-sheet 5 of the CLD.

Experiments conducted by Holland laboratory using cross-linking reported that two HlyB mutants, E256K and S279L, were unable to form oligomers with HlyD (Holland et al., 2016). In the work by Zhao *et al.*, residues R256, F323, and A327 were identified as being located at the lateral opening of the protomer involved in HlyA translocation. Furthermore, they identified several positions of HlyB that were in close proximity to the substrate eGFP-HlyA1, including positions 256, 263, 319, 323, 327, and 422 (Zhao et al., 2022).

Our suggested possible beneficial positions of HlyB, including F216, E226, S272, and S290, appear to be in the vicinity of the substrate interaction region of the TMD of HlyB.

Especially, V421, which was identified as a beneficial positions in the current study, is located within the translocation pathway (Zhao et al., 2022).

Interestingly, the potential beneficial position I407 is located in the periplasmic loop P3 which is between TMHs 5 and 6. This loop with a size of 8 residues in HlyB is conserved in a wide variety of ABC transporters. A site saturation mutagenesis study on the residues of this loop demonstrated that at least four residues of this loop are essential for secretion, as for instance, I401T and D404G mutants showed only 20% of hemolytic activity and a double mutant, S402P and D404K, showed no hemolytic activity at all (Holland et al., 2016).

Structural analysis shows that our beneficial position, F434, is located exactly at a location in which the amphipathic helix of HlyD (residues 9-24) contacts the elbow helix of HlyB (Zhao et al., 2022).

The KnowVolution campaign in HlyD did not identify any single residue for improving the secretion efficiency of HlyA T1SS. This is consistent with two previous directed evolution studies on HlyA T1SS, which also did not identify any improved mutants harboring HlyD mutations (Sugamata and Shiba, 2005, Low et al., 2010). One possible explanation for this is that only 60 residues of HlyD are located in the cytosol, limiting its direct contact with the substrate to these residues. The rest of the protein is either located in the membrane (20 residues) or in the periplasmic space (approximately 400 residues). Another possibility is that HlyD is already optimized by evolution and is in its best functional state.

Conclusion

This study demonstrated that the secretion efficiency of the HlyA T1SS can be improved further by the engineering its ABC transporter HlyB. *E. coli* cells harboring the novel HlyB variant showed a secretion level of 400 mg/L of soluble lipase-HlyA1, plus 5 mg estimated to be in the form of enzyme oligomers in the supernatant of only 25 mL of culture. The applied evolutionary approach of this study can be generally applied to other section systems to improve the secretion performance of bacterial secretion systems. An improved secretion system opens up an avenue for the usage of Gram-negative bacteria such as *E. coli* as a secreting host for industrial protein production.

Experimental Section
Material

All chemicals used in this study were purchased from Sigma-Aldrich, Roche Diagnostics GmbH, and Applichem GmbH if not stated otherwise. Enzymes were purchased from New England Biolabs (England), Promega GmbH (Germany), SeSam-Biotech GmbH (Germany). All oligonucleotides were purchased from either Eurofins MWG Operon or GeneScript. The commercial available kits NucleoSpin plasmid miniprep kit, and PCR clean-up kit were purchased from Macherey Nagel (Düren, Germany). Gibson assembly master mix was purchased from New England Biolabs (England). DNA sequencing was performed by Microsynth Seqlab (Göttingen, Germany).

Generation of *hlyB* and *hlyD* casting error-prone PCR mutant library

For cloning and library generation *E. coli* BL21-Gold (DE3) was used. Oligonucleotides used in this work are listed in **Table 1S**. Plasmid pK184_*EF*(Best)-*lipA-hlyA1-Ter-hlyBD* (Pourhassan et al., 2022a) was used as the backbone vector for the library generation. Note that the gene encoding the TolC protein is endogenous, and the genes encoding HlyB and HlyD are plasmid-based. The abbreviations in the name of the plasmid stand for: *EF* for enhancer fragment, *lip*A for *lipase* gene from *Serratia marcescens, hly* for hemolysin, and *Ter* for a terminator region. The *lac* promoter of plasmid pK184 is inducible with isopropyl β-D-1-thiogalactopyranoside (IPTG). This plasmid contains a kanamycin resistance gene.

Casting error-prone PCR (cepPCR) was used for generation of random mutagenesis libraries for HlyB and HlyD according to the published cepPCR protocol (Yang et al., 2017). Accordingly, the *hlyB* gene from *E. coli* UTI89, except of the region encoding for the NBD, was divided into six fragments, listed in **Table 1S**.

The *hlyD* gene from *E. coli* UTI89 was divided into six fragments, listed in **Table 1S**. For each fragment a primer set, forward and reverse, was designed (**Table 1S**) and employed for amplification of the fragment via epPCR.

The parameters and reaction conditions of cepPCR (50 µL) were as follows: 1X ThermoPol buffer, 0.2 mM dNTP mix, 400 pM of each primer, 10U polymutarase polymerase, 60 ng DNA template, 0.1-0.9 mM $MnCl_2$. The PCR was performed as: 94°C for 2min (1 cycle), 94°C for 30 s, primer melting temperature (Tm) for 30 s, 68°C for 1 min (for 30 cycles), and 68°C for 10 min (1 cycle). To obtain mutations of the first amino acids of each fragment, the forward primers annealed upstream of reverse primer of the previous fragment. PCR

products were checked on agarose gel and purified with the PCR clean-up kit according to the instructions of the manufacturer.

For cloning of the individual mutant fragments into plasmid pK184_ *EF*(Best)-*lipA-hlyA1-Ter-hlyBD* either the megawhop PCR or Gibson assembly was used, with the following details:

For each individual fragment, nine mutant libraries applying different $MnCl_2$ concentrations (0.1 mM to 0.9 mM) were generated. The inactive ratio of libraries were determined by prescreening on tributyrin agar plates (trypton 10 g/L, yeast extract 5 g/L, NaCl 10 g/L, agar 15 g/L, gum arabic 1.5 g/L, tributyrin 15 g/L, kanamycin 50 μg/mL). The condition of library generation for each fragment was adjusted to obtain libraries with an inactive ratio of 40-50%. To construct the entire vector, the amplified fragment containing mutations was used as either megaprimer for megawhop PCR or insert for Gibson assembly.

The megawhop PCR was performed based on the published protocol (Miyazaki, 2011). The megawhop PCR products were digested with DpnI (20U, 37°C, overnight) and purified using the PCR clean-up kit.

For Gibson assembly cloning, the vector template lacking the related fragment was amplified using a high-fidelity polymerase, Q5 polymerase. The PCR mixture (50 μL) contained 50 ng of template plasmid, 1×Q5 standard reaction buffer, 0.2 mM dNTP mix, 200 pM of each primer, and 1 U Q5 DNA polymerase. Subsequently, the epPCR product of the related fragments and the PCR product of the backbone vector were digested with DpnI enzyme (overnight, at 37°C) to remove any undigested wildtype vector. The PCR products were purified using a PCR clean-up kit. The whole plasmid construction of the epPCR product and the linear plasmid was performed via Gibson assembly according to the instructions of the manufacturer.

Subsequently, 2 μL of the assembled product of Gibson assembly reaction or the megawhop PCR product was transformed into chemically competent cells of *E. coli* BL21-Gold (DE3). The transformed cells were cultured directly on tributyrin LB agar plates containing kanamycin 50 μg/mL. Ten single cell clones of each library were picked and sent for sequencing. For each fragment, single active clones were picked up and cultured in 96-well plates (PS-F-bottom, Sarstedt, Germany).

Site saturation mutation of potential beneficial positions

Site-saturation mutagenesis of potential beneficial positions was performed by two-steps PCR based on the published protocol (Wang and Malcolm, 1999). The primers having degenerative nucleotides used for site-saturation mutagenesis are listed in **Table 2S**. The PCR reaction was performed in two steps. In the first step, two separate reactions in separate tubes were performed; one tube containing the forward primer and the another one containing the reverse primer. In the second step, both tubes were mixed and amplification was continued.

The PCR mixture (50 µL) of the first step contained 50 ng backbone plasmid, 1X Q5 standard reaction buffer, 0.2 mM dNTP mix, 200 pM of either forward or reverse primer, 1 U Q5 DNA polymerase. The PCR was performed as: 98°C for 2 min (1 cycle); 98°C for 20 s, Tm for 20 s, 72°C for 8 min (for 3 cycles), and 72°C for 10 min (1 cycle). After mixing two tubes, another PCR reaction was performed as: 98°C for 2 min (1 cycle); 98°C for 20 s, primer melting temperature (Tm) for 20 s, 72°C for 8 min (for 15 cycles), and 72°C for 10 min (1 cycle). The PCR products were checked on the agarose gel, and subsequently were subjected to DpnI digestion (20U, 37°C, overnight). DNA purification of the PCR products was performed using PCR purification kit. Purified PCR products were transformed directly into chemically competent *E. coli* BL21 Gold (DE3) for expression and screening.

Recombination of compatible substitutions experimentally

The recombination library of the HlyB was generated using an overlap extension PCR method based on the published protocol (Hilgarth and Lanigan, 2020). Two oligonucleotides of the *hly*B gene with 6 and 7 substitutions were synthesized by GenScript. Both synthesized oligonucleotides, the *hlyB* wildtype gene, plus the plasmids with improved positions were used as templates for the overlap extension PCR. Subsequently, the HlyB overlap extension library was cloned into the pK184 plasmid using Gibson assembly method and transformed into *E. coli* BL21 Gold (DE3) chemical competent cells.

Cultivation and expression of libraries in 96-well plates

For each library, single active clones were picked and placed in 150 µL LB_{kan} medium in 96-well plates (PS-F-bottom, Sarstedt, Germany) along with four wildtype clones and four empty vector clones as positive and negative controls, respectively. Active clones were identified by generating halo on the tributyrin-agar plates. The cultures were incubated overnight (24 hours, 37°C, 900 rpm, and 80% humidity) in a 96-well MTP shaker (SI505

MTP shaker incubator, Avantor). Then, 50 μL of glycerol 50% (w/w) were added to each wells and the plates, named as master plates, were stored at -80°C for further experiments. For expression, each master plate was sub-cultured into a new 96-well MTP plate containing 150 µL LB_{kan} medium and cultivated (24 hours, 37°C, 900 rpm, and 80% humidity). The cultures were used to inoculate new 96-well MTP plates containing 150 µL LB_{kan} medium and cultivated for 3 to 4 hours (37°C, 900 rpm, and 80% humidity). Subsequently, main cultures were induced with 1 mM of IPTG and 5 mM $CaCl_2$. The cultures were further cultivated for 16 hours (37°C, 900rpm, and 80% humidity) in 96-well MTP shaker incubator. Afterwards, the MTPs were centrifuged (4°C, 20 min, 3500 rpm) and supernatants were harvested and used for the further investigation.

Colorimetric screening system

The screening of the libraries with the p-nitrophenyl butyrate (*p*NPB) assay was performed as previously reported in 96-well MTPs (Pourhassan et al., 2022a). Screening of the recombinant library via lipase colorimetric assay proved to be a challenging task as the amount of lipase in the supernatant of clones was high and reached the saturation level of the assay. To solve this issue, the supernatant was further diluted in the assay reaction buffer (20 fold dilution) and the incubation time before inducing the cultures was reduced to 3 hours.

Computational method to analyze stability

The structure of HlyB was constructed based on the structure of PCAT1 (PDB: 4RY2) (Lin et al., 2015) with TopModel tool (Mulnaes et al., 2020). The homology model of HlyB was used for the computational analysis. For visualization the HlyB /HlyD structure (PDB: 7SGR) driven from Cryo-EM from Zhao *et al*. work was used (Zhao et al., 2022). The monomers of the HlyB model superimposed with a RMSD of 3.3 Å (558 Cα atoms) and 3.6 Å (558 Cα atoms), respectively, on the monomers of the inward facing conformation of HlyB (Zhao et al., 2022). These differences are due to a different orientation of TMD and NBD in the model compared to the single particle structure. Superimposition of the isolated TMDs revealed a RMSD of 1.8 Å (311 Cα atoms) and 1.3 Å (273 Cα atoms). The crystal structures of the isolated NBD of HlyB in the apo form (1MT0 (Schmitt et al., 2003)) and the ATP / Mg^{2+} bound state (1XEF (Zaitseva et al., 2005a)) superimposed with RMSD

values of 0.7 Å (191 Cα atoms, 1MT0) and 1.2 Å (449 Cα atoms, 1XEF), respectively (Lin et al., 2015).

The $\Delta\Delta G_{fold}$ ($\Delta\Delta G_{fold}$= $\Delta\Delta G_{fold,Sub}$- $\Delta\Delta G_{fold,WT}$) values of individual substitutions were computed using FoldX method version 3b5.1 (Guerois et al., 2002), in YASARA structure version as described previously (Van Durme et al., 2011, Cui et al., 2020). Default FoldX parameters were used (temperature 298 K; ionic strength 0.05 m; pH 7) to generate the substitutions. Mutate residues commands of FoldX was used to compute the $\Delta\Delta G_{fold}$ values of substitutions. PyMOL molecular Graphics System, Version 1.8.6.0 Enhanced for Mac OS X (DeLano, 2002) was used for visualization.

Acknowledgment

This research was funded by CLIB Competence Center Biotechnology (CKB) funded by the European Regional Development Fund ERDF (34.EFRE-0300096) to Z.P and L.S. We thank all current and former members of the Institute of Biochemistry for support and fruitful discussions. We dedicate this manuscript to Karl-Erich Jaeger for his pioneering contributions in the field of molecular enzyme technology.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] G. R. Kleiner-Grote, J. M. Risse and K. Friehs, *Engineering in Life Sciences* **2018**, *18*, 532-550.

[2] L. A. Burdette, S. A. Leach, H. T. Wong and D. Tullman-Ercek, *Microbial cell factories* **2018**, *17*, 196.

[3] R. Freudl, *Microbial cell factories* **2018**, *17*, 52.

[4] G. T. Eom, S. H. Lee, Y. H. Oh, J. E. Choi, S. J. Park and J. K. Song, *Biotechnology letters* **2014**, *36*, 2037-2042.

[5] Z. N. Pourhassan, H. Cui, S. Khosa, M. D. Davari, K. E. Jaeger, S. H. Smits, U. Schwaneberg and L. Schmitt, *ChemBioChem* **2022**, *23*, e202100702.

[6] a) T. S. Wong, D. Zhurina and U. Schwaneberg, *Combinatorial chemistry & high throughput screening* **2006**, *9*, 271-288; b) S. Pramanik, F. Contreras, M. D. Davari and U. Schwaneberg, *Protein Engineering: Tools and Applications* **2021**, 153-176.

[7] I. B. Holland, S. Peherstorfer, K. Kanonenberg, M. Lenders, S. Reimann and L. Schmitt, *EcoSal Plus* **2016**, *7*.

[8] Z. N. Pourhassan, S. H. Smits, J. H. Ahn and L. Schmitt, *Biotechnology advances* **2021**, *53*, 107864.

[9] H. Zhao, J. Lee and J. Chen, *Cell* **2022**, *185*, 3329-3340. e3313.

[10] a) L. Schmitt, H. Benabdelhak, M. A. Blight, I. B. Holland and M. T. Stubbs, *Journal of molecular biology* **2003**, *330*, 333-342; b) J. Lecher, C. K. Schwarz, M. Stoldt, S. H. Smits, D. Willbold and L. Schmitt, *Structure* **2012**, *20*, 1778-1787.

[11] a) H. Benabdelhak, S. Kiontke, C. Horn, R. Ernst, M. A. Blight, I. B. Holland and L. Schmitt, *Journal of molecular biology* **2003**, *327*, 1169-1179; b) Z. N. Pourhassan, E. Hachani, O. Spitz, S. H. Smits and L. Schmitt, *Frontiers in Microbiology* **2022**, *13*.

[12] R. Schtilein, I. Gentschev, H.-J. Mollenkopf and W. Goebel, *Molecular and General Genetics MGG* **1992**, *234*, 155-163.

[13] F. Cheng, L. Zhu and U. Schwaneberg, *Chemical Communications* **2015**, *51*, 9760-9772.

[14] J. Yang, H.-L. Zhao, L.-Y. Ran, C.-Y. Li, X.-Y. Zhang, H.-N. Su, M. Shi, B.-C. Zhou, X.-L. Chen and Y.-Z. Zhang, *Scientific Reports* **2015**, *5*, 9936.

[15] S. Brands, H. U. Brass, A. S. Klein, J. G. Sikkens, M. D. Davari, J. Pietruszka, A. J. Ruff and U. Schwaneberg, *Catalysis Science & Technology* **2021**, *11*, 2805-2815.

[16] C. Novoa, G. V. Dhoke, D. M. Mate, R. Martínez, T. Haarmann, M. Schreiter, J. Eidner, R. Schwerdtfeger, P. Lorenz and M. D. Davari, *ChemBioChem* **2019**, *20*, 1458-1466.

[17] S. Islam, D. Laaf, B. Infanzón, H. Pelantová, M. D. Davari, F. Jakob, V. Křen, L. Elling and U. Schwaneberg, *Chemistry–A European Journal* **2018**, *24*, 17117-17124.

[18] K. Rübsam, M. D. Davari, F. Jakob and U. Schwaneberg, *Polymers* **2018**, *10*, 423.

[19] S. Khosa, R. Scholz, C. Schwarz, M. Trilling, H. Hengel, K.-E. Jaeger, S. H. Smits and L. Schmitt, *Applied Environmental Microbiology* **2018**, *84*, e01163-01117.

[20] J. Yang, A. J. Ruff, M. Arlt and U. Schwaneberg, *Biotechnology and bioengineering* **2017**, *114*, 1921-1927.

[21] D. Y.-w. Lin, S. Huang and J. Chen, *Nature* **2015**, *523*, 425-430.

[22] K. Liebeton, A. Zonta, K. Schimossek, M. Nardini, D. Lang, B. W. Dijkstra, M. T. Reetz and K.-E. Jaeger, *Chemistry & biology* **2000**, *7*, 709-718.

[23] H. Cui, H. Cao, H. Cai, K. E. Jaeger, M. D. Davari and U. Schwaneberg, *Chemistry–A European Journal* **2020**, *26*, 643-649.

[24] R. S. Hilgarth and T. M. Lanigan, *MethodsX* **2020**, *7*, 100759.

[25] T. J. Collins, *Biotechniques* **2007**, *43*, S25-S30.

[26] a) J. Puetz and F. M. Wurm, *Processes* **2019**, *7*, 476; b) A. Sharma and T. K. Chaudhuri, *Microbial cell factories* **2017**, *16*, 1-19.

[27] T. R. Costa, C. Felisberto-Rodrigues, A. Meir, M. S. Prevost, A. Redzej, M. Trokter and G. Waksman, *Nature Reviews Microbiology* **2015**, *13*, 343-359.

[28] Y. Ji, S. Islam, H. Cui, G. V. Dhoke, M. D. Davari, A. M. Mertens and U. Schwaneberg, *Catalysis Science & Technology* **2020**, *10*, 2369-2377.

[29] a) Y. Sugamata and T. Shiba, *Applied Environmental Microbiology* **2005**, *71*, 656-662; b) K. O. Low, N. M. Mahadi, R. A. Rahim, A. Rabu, F. D. A. Bakar, A. M. A. Murad and R. M. Illias, *Journal of biotechnology* **2010**, *150*, 453-459.

[30] K. Miyazaki in *MEGAWHOP cloning: a method of creating random mutagenesis libraries via megaprimer PCR of whole plasmids*, *Vol. 498* Elsevier, **2011**, pp. 399-406.

[31] W. Wang and B. A. Malcolm, *Biotechniques* **1999**, *26*, 680-682.

[32] D. Mulnaes, N. Porta, R. Clemens, I. Apanasenko, J. Reiners, L. Gremer, P. Neudecker, S. H. Smits and H. Gohlke, *Journal of chemical theory and computation* **2020**, *16*, 1953-1967.

[33] J. Zaitseva, S. Jenewein, T. Jumpertz, I. B. Holland and L. Schmitt, *The EMBO journal* **2005**, *24*, 1901-1910.

[34] R. Guerois, J. E. Nielsen and L. Serrano, *Journal of molecular biology* **2002**, *320*, 369-387.

[35] J. Van Durme, J. Delgado, F. Stricher, L. Serrano, J. Schymkowitz and F. Rousseau, *Bioinformatics* **2011**, *27*, 1711-1712.

[36] W. L. DeLano, *CCP4 Newsletter on Protein Crystallography* **2002**, *40*, 82-92.

Supporting information

Table S1: Primers used for epPCR amplification of fragments.

Fragments	Forward primer (5' to 3')	Reverse primer (5' to 3')
F_{1B} (275 bp)	GCCACATAAGATTGCTATTTTTTTG	CCATCCTCTCTCCAGACTAATG
F_{2B} (251 bp)	TCTGGAGAGAGGATGGACGTC	AACAAGGGTTTCAATAAATATTCTC
F_{3B} (283 bp)	GAATATTTATTGAAACCCTTGTTGTG	CGACTCTCAAAATAAGAGATCGG
F_{4B} (256 bp)	CTCTTATTTTGAGAGTCGTCGTG	CCGTGAAAACTTATCATCAAGGC
F_{5B} (279 bp)	TGATGATAAGTTTTCACGGAATGCG	GCAATTAACTGACCAATCGATAAATCC
F_{6B} (291 bp)	CTGGTTATTTCCGGGGATTTATCG	CCCCTGCTTAATACTGAGATTG
F_{1D} (279 bp)	CAGTTACAGTCAGACTAACAG	TTCCACCTGACCTAAAAC
F_{2D} (260 bp)	GGTTTCTGGTTATTGCTG	AATTCAATTGACCTGCTC
F_{3D} (289 bp)	CGGTATCAAATTCTGAGCAG	CCTGTTTATGCAATAAACTCC
F_{4D} (290 bp)	GGTTGAAAAAAGCCGTCTG	AAGCCTGTTGACGCTCTTC
F_{5D} (280 bp)	GCATTGAGTTATTAACTCTGGAG	CCCACCAGATAACCATATC
F_{6D} (287 bp)	CCTTTCCTTACACCCGATATG	CCGGCCGCTCTAGATTAC

Table S2: Primers used for site-saturation at potential beneficial positions

Beneficial positions	Forward primer	Reverse primer
SSM of potential beneficial positions of HlyB		
Ter 36	CATAGATTTGACNNKGACGG	CCGTCMNNGTCAAATCTATG
Ser 71	GATTAAACTTTATTNNKCTGCCCG	CGGGCAGMNNAATAAAGTTTAATC
Val 108	GAAATCCCCGTNNKCTCGAACAG	CTGTTCGAGMNNACGGGGATTTC
Leu 164	GTCTGTTTTTNNKCAATTATTTG	CAAATAATTGMNNAAAAACAGAC
Phe 216	GAACTTACATTNNKGCACATAGTACAAG	CTTGTACTATGTGCMNNAATGTAAGTTC
Glu 226	GATTGATGTTNNKTTGGGTG	CACCCAAMNNAACATCAATC
Ser 272	GACAGGCATTAACANNKGTTCTG	CAGAACMNNTGTTAATGCCTGTC
Ser 290	GTGGTATTACNNKCCAAAG	CTTTGGMNNGTAATACCAC
Ile 407	GATTTATCGNNKGGTCAGTTAATTG	CAATTAACTGACCMNNCGATAAATC
Val 421	GCTGGTCAGATTNNKGCACCG	CGGTGCMNNAATCTGACCAGC
Phe 434	GCACAAATCTGGCAGGATNNKCAGCAG	CTGCTGMNNATCCTGCCAGATTTGTGC
Asn 465	GGAAATTNNKGGTGATATCACTTTTCG	CGAAAAGTGATATCACCMNNAATTTCC
SSM of potential beneficial positions of HlyD		
Met 6	CATGGTTANNKGGGTTCAG	CTGAACCCMNNTAACCATG
Ile 49	GGAATTANNKGAAACGCCG	CGGCGTTTCMNNTAATTCC

Val 69	GTTTCTGNNKATTGCTGTC	GACAGCAATMNNCAGAAAC

Table S3: Internal-atomic distance between 12 beneficial positions of the HlyB wildtype measured by YASARA. The distance values were reported with Å unit.

	T36	S71	V108	L164	F216	E226	S272	S290	I407	V421	F434	N465
T36	0	26.3	27.6	53.1	40.9	30.1	40.2	63.0	70.9	49.3	29.2	37.5
S71		0	18.2	62.1	54.5	48.1	46.2	62.1	72.7	52.8	36.3	59.4
V108			0	52.3	46.2	44.1	38.5	50.6	59.4	40.5	27.7	65.0
L164				0	14.5	29.4	16.5	24.3	26.7	16.4	26.2	72.6
F216					0	15.5	15.0	36.4	39.4	22.6	19.1	58.5
E226						0	22.8	48.3	54.1	31.0	20.8	43.2
S272							0	26.4	34.6	15.4	12.3	64.1
S290								0	13.1	15.5	34.2	90.3
I407									0	21.7	42.3	96.9
V421										0	20.6	76.5
F434											0	57.7
N465												0

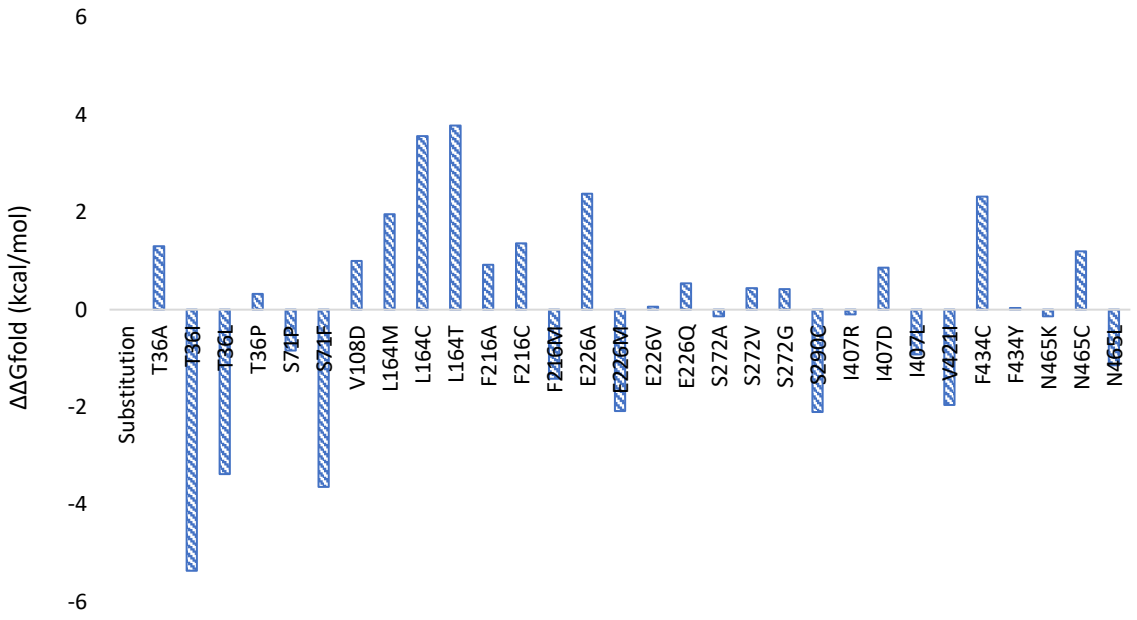


Table S4: $\Delta\Delta G_{fold}$ (kcal/mol) values of amino acid substitutions on the twelve beneficial positions of the HlyB calculated using FoldX method.

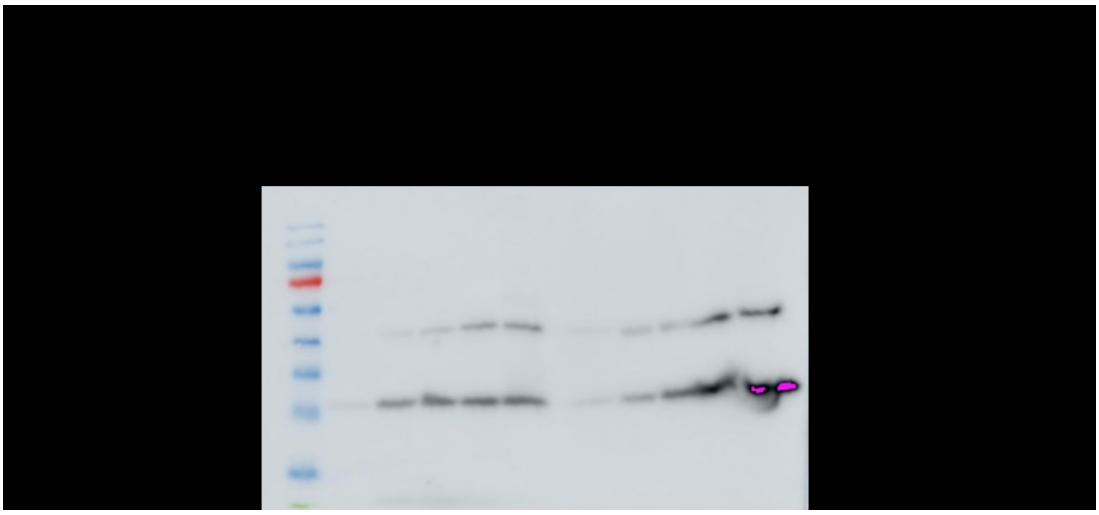


Figure S1. Improved secretion of cutinase-HlyA1 by the 15-H7 variant. a) Western blot analysis of supernatant fractions from the 15-H7 variant the wildtype cultures. The molecular weight of the marker proteins is given on the left; xh: cell pellets of culture, where x denotes the number of hours after induction. This experiment performed in triplicate.

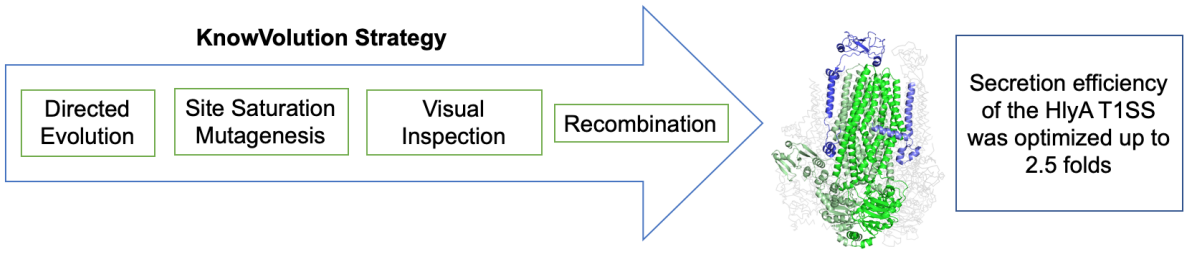
HlyD KnowVolution

Phase I (Identification): In the first phase, beneficial positions of HlyD were identified through the cepPCR as already explained above for the HlyB. Based on the cepPCR approcah, the gene of *hly*D was divided into six fragments (F_{1D}: 279 bp, F_{2D}: 260 bp, F_{3D}: 289 bp, F_{4D}: 290 bp, F_{5D}:280 bp, F_{6D}: 287 bp). All steps of library generation and screening were performed as already explained for the HlyB protein.

For each fragment, a library with a size of 786 was generated. Subsequently, the libraries were screened in a 96-well MTP format using the *p*NPB screening to identify clones with higher secreted lipase-HlyA1. None of the fragments, except of the first fragment, showed promising variants. The promising variants of the first fragment were sent for sequencing. Alignment of the sequence of this promising variants identified three potential beneficial positions, listed as: Met6, Ile49, Phe69.

Phase II (determination): To obtain full diversity on the identified beneficial positions, the two beneficial positions were subjected to individual site-saturation mutagenesis. In this step no promising variant were found after screening of the site-saturation libraries for none of the identified positions. To be sure about the outcome of the KnowVolution of HlyD, a new libraries were generated on the fragments 1-2. But still no positions were found.

Graphical Abstract



Isolated through the application of KnowVolution, a novel mutant of HlyB promotes the secretion efficiency of the HlyA T1SS of *E. coli* up to 2.5-fold for two hydrolases, a lipase and a cutinase. HlyB is an ABC transporter that plays the central role in the HlyA T1SS.

2.5 Chapter V: TolC overexpression impairs the hemolysin secretion

Title: TolC overexpression impairs the hemolysin secretion through the HlyA type 1 secretion system

Authors: Zohreh Pourhassan N. and Lutz Schmitt

Submitted in: to be submitted

Impact factor (2023): -

Own proportion of this work: 80 %

TolC overexpression impairs the hemolysin secretion through the HlyA type 1 secretion system

Zohreh N. Pourhassan [1] and Lutz Schmitt[1*]

[1] Institute of Biochemistry, Heinrich Heine University, Universitätsstr. 1, 40225 Düsseldorf (Germany)

[*] to whom correspondence should be addressed
Lutz Schmitt
Phone: +49(0)211-81-10773
Email: lutz.schmitt@hhu.de

Abstract

Secretion has advantages for recombinant protein production. Many secretion systems secrete heterologous proteins only in low titers if at all. This limits the applications of secretion systems in biotechnological applications and especially on industrial scales. Herein, we investigated if the limiting factor of secretion efficiency of the hemolysin A type 1 secretion system (HlyA T1SS) is the amount of TolC in the membrane.

HlyA T1SS of *Escherichia coli* is a paradigm of the type 1 secretion systems. This group of secretion systems of Gram-negative bacteria translocate a substrates across both membranes through a continuous tunnel-channel in one step. The tunnel of HlyA T1SS consists of the ATP binding cassette transporter HlyB, the membrane fusion protein HlyD, and the outer membrane protein TolC.

Recently, it was determined that overexpression of HlyBD did not influence on the amount of TolC. Herein, we observed that TolC overexpression rather reduces the secretion level of the HlyA T1SS.

Keywords:

Protein secretion, Gram-negative bacteria, bacterial secretion systems, TolC overexpression, secretion efficiency, the Hly secretion system.

Introduction

Secretion has great advantages for recombinant protein production. Secretion could result in a production of proteins in a higher quality and higher stability, meanwhile with a reduction in production costs because of a simplified downstream processes in compared to other production strategies (Kleiner-Grote et al., 2018, Burdette et al., 2018).

Gram-negative bacteria have evolved at least seven different secretion systems to translocate a wide variety of substrates across their membranes (Costa et al., 2015). So far, six out of seven Gram-negative bacterial secretion systems have been applied for the secretion of heterologous proteins, listed as T1SS, T2SS, T3SS, T4SS, T5SS, and T8SS (Burdette et al., 2018). Secretion titers of these systems for heterologous proteins are in most of the cases very low that limits their utilities in biotechnological applications (Khosa et al., 2018). To address this issue different approaches have been employed on different secretion systems *e.g.* secretion signal engineering (Freudl, 2018), KnowVolution of secretion system components (Pourhassan et al., 2022a), overexpression of a particular protein involved in a secretion system (Perez-Perez et al., 1994).

Herein, we focus on hemolysin A type 1 secretion system to answer this question if the overexpression of the outer membrane protein TolC could improve further the secretion titer of this system.

HlyA T1SS is one of the most prominent and probably the best studied T1SS (Holland et al., 2016). The relatively simple architecture of the HlyA T1SS makes it an interesting candidate for the further studies to improve its secretion efficiency. It has been elucidated that the secretion pathway of HlyA T1SS from *Escherichia coli* consists of three membrane proteins, listed as the ATP-binding cassette (ABC) transporter HlyB, the membrane fusion protein HlyD, and the outer membrane protein TolC (Holland et al., 2016, Balakrishnan et al., 2001).

Upon a specific interaction between the inner-membrane complex, consisting of HlyB and HlyD, TolC is recruited. Then, all three proteins form a tunnel-channel redirecting the substrate in one step from the cytosol to the extracellular space. After secretion, TolC dissociates from the complex, while HlyD and HlyB remain associated (Holland et al., 2016)Balakrishnan et al., 2001).

The substrate of this system is HlyA, a pore-forming toxin, that lyses cell membranes of, for examples, human erythrocytes (Noegel et al., 1979). HlyA has a size of 110 kDa and belongs to the RTX (repeats-in-toxins) family, a superfamily of proteins secreted by Gram-negative

bacteria harboring a number of conserved glycine- and aspartate-rich sequence (GGxGxDxUx, while U refers to a large hydrophobic amino acid and x could be any amino acids) (Spitz et al., 2019). The secretion signal of HlyA along with three conserved Glycine-rich motifs, termed HlyA1, is known as an interesting carrier for the secretion of the heterologous proteins (Nicaud et al., 1986).

HlyB resides in a dimeric state in the inner-membrane. HlyB provides the energy of the transport by ATP hydrolysis. Analysis on the HlyB structure shows the presence a transmembrane domain (TMD) consists of six transmembrane helixes, a nucleotide-binding domain (NBD), and a C39-peptidase-like domain (CLD) per monomer (Kanonenberg et al., 2018). The CLD presents around 40% of homology to the C39 peptidase (a cysteine protease), but without any peptidase activity. Researchers have reported a chaperon like activity for the CLD based on a series of conducted pull-down assays (Lecher et al., 2012).

HlyD resides in the inner-membrane. First a trimeric state was suggested for this protein (Balakrishnan et al., 2001), but later based on the structure of other membrane fusion proteins a hexametric state for this protein was suggested to be more valid (Pourhassan et al., 2021, Holland et al., 2016). HlyD harbor a short cytosolic domain with a size of 60 residues. This cytosolic domain continues with a single helix transmembrane domain and then a very long periplasmic domain (Pimenta et al., 1999, Pimenta et al., 2005).

Later on, it was found that the OM protein TolC is also engaged in HlyA T1SS machinery (Thanabalu et al., 1998). Beside that, TolC evolves in the tripartite drug efflux complexes of *E. coli* including AcrAB/TolC (Du et al., 2014) and MacAB/TolC (Fitzpatrick et al., 2017). The crystal structure of TolC was solved representing a homotrimeric state for this protein forming the exit route with a length of 140 Å in the periplasmic space and with a maximum diameter of 20 Å (Koronakis et al., 1997). The restricted diameter of TolC determines the allowed transit structure to an α-helix. Obviously, HlyA remains unfolded during the secretion and folds only when it is in the extracellular space. TolC forms a 12-stranded β-barrel in the OM, thought to be continually opened to the extracellular space, and a 12-α-helical helices with a length of 100 Å into the periplasmic space (Anderson et al., 2000, Jo et al., 2019, Koronakis et al., 1997).

Using a stalled Hly system, the number of active Hly system in *E. coli* BL21 (DE3) under the overexpression of HlyA, HlyB, and HlyD was estimated to be 4500 copies (Lenders et al., 2016). Only recently, the endogenous amount of TolC in *E. coli* BL21 (DE3)-BD quantified under two conditions of stalled system and actively secreting system (Beer et al.,

2022). Beer *et al.* reported that under both tested conditions, the number of trimeric TolC remains constant and around approximately 4500 trimeric TolC. Furthermore, Beer *et al.* also determined the number of active Hly system in *E. coli* strain UTI89 as well as number of TolC. Interestingly, they reported that although the number of active Hly secretion system in UTI89 is only 800, but still the number of TolC trimer is around 4500 (Beer et al., 2022). What is not yet clear is that whether the amount of TolC is the limiting factor of secretion through the Hly system. To address this question, the current study attempts to investigate the influence of TolC overexpression on the secretion efficiency of the Hly secretion system.

Materials and Methods

Materials

Chemicals used in this study were purchased from Sigma-Aldrich, Roche Diagnostics GmbH, or Applichem GmbH if not stated otherwise. Enzymes were purchased from New England biolabs.

Plasmids used in this study

For the overexpression studies *E. coli* BL21 (DE3) were used. Plasmid pK184_ *EF-lipA-hlyA1BD* (Pourhassan et al., 2022a) was used for the expression of Hly system apparatus as well as the expression of the lipase fuse to HlyA1. The pK184 plasmid is inducible with isopropyl β-D-1-thiogalactopyranoside (IPTG) under the P_{lac} and has kanamycin resistant. To overexpress TolC protein, plasmid pET28_*tolC* was used. This plasmid contains an ampicillin resistant gene. The TolC gene of this plasmid is under the control of the T_7 promoter.

Expression and secretion experiments in shaker flasks

E. coli BL21 (DE3) chemically competent cells were transformed with pK184 or/and pET28_*tolC* plasmids and grown on the tributyrin Lysogeny broth (LB) agar plates (Trypton 10 g/L, Yeast extract 15 g/L, NaCl 5 g/L, Agar 15 g/L, Gum arabic 1.5 g/L, Tributyrin 15 g/L, Kanamycin 50 μg/mL or/and ampicillin). The halo around single cell colons on the tributyrin agar plate showed the successful secretion of lipase. Single active cell colons were used to prepare pre-cultures (37°C and 180 rpm). The overnight pre-culture was used to inoculate 25 ml of 2YT (Trypton 16 g/L, Yeast extract 10 g/L, NaCl 5 g/L) medium supplemented with proper antibiotics to an OD_{600} of 0.1 in 100 mM Erlenmeyer shaking flask. The culture was cultivated (37°C and 180 rpm) before the culture was inducted at OD_{600} of 0.4 to 0.6 with 1 mM IPTG and 5 mM $CaCl_2$. The supernatant of the culture harvested after 0, 2, 4, and 6 hours after induction. The cell free supernatant harvested by centrifugation of the culture and used for further analysis, either for SDS-PAGE or *p*NPB lipase assay. The semi-quantifications of SDS-PAGE was performed using ImageJ software (Image Processing and Analysis in Java).

Colorimetric *p*NPB lipase assay

The lipase assay applied in the current study is based on the published protocol of *p*-Nitrophenyl butyrate (*p*NPB) assay (Pourhassan et al., 2022a). In the presence of lipase the

*p*NPB substrate is converted to *p*-nitrophenolate which is detectable spectrophotometrically at 410 nm. After having the cell free supernatant of the cultures, 10 µl of cell free supernatant was diluted in 90 µL of TEA buffer (50 mM, pH 7.4) into a 96-well MTP (PS-F-bottom, Sarstedt, Germany). The substrate solution was prepared freshly as 0.5 mM *p*NPB solved in 10% v/v of acetonitrile and diluted in TEA buffer (50 mM, pH 7.4). Upon supplementing 100 µL of freshly prepared substrate solution to the supernatant mixture in a final reaction volume of 200 µL the reaction was started. The amount of released *p*-nitrophenolate was recorded by measuring A_{410nm} (at 30°C) over 8 min with a kinetics interval of 20 seconds on a microtiter plate reader (Tecan infinite® M200 Pro Austria GmbH, Männedorf, Switzerland). The rate of conversion of *p*NPB to *p*-nitrophenolate was calculated for each reaction and used to compare the secretion level of different cultures.

Membrane isolation

A two-step centrifugation procedure was applied for membrane isolation based on a previously published protocol (Kanonenberg et al., 2019a). Cell pellet was resuspended in phosphate buffer (50 mM, pH 8) supplemented with DNAase and protease inhibitor. Then, cells were lysed by passing at-least three times through a cell disruptor (Microfluidizer M-110L, Microfluidics) at 1.5 kbar. Subsequently, cell lysate was centrifuged by low-spin centrifugation (18000 g, 30 min, at 4°C) to remove undisrupted cells and cell debris. Then, supernatant of low-spin centrifugation was applied to an ultra-centrifugation (150000 g, 90 min, at 4°C) to pellet the membranes. Membranes were homogenized in phosphate buffer and applied on western blot for the further analysis.

Results

TolC overexpression impairs the secretion level of HlyA T1SS

The first set of experiments aimed to address the question if overexpression of TolC improves the secretion efficiency through HlyA T1SS. To compare the secretion level of the substrate (a lipase from *Serratia marcescens* fused to HlyA1) under condition of either endogenous or overexpressed TolC, the following strains were cloned.

Chemically competent *E. coli* BL21 (DE3) cells were transformed with plasmid pK184_ *EF-lip*A-*hly*A1BD. This strain was termed "TolC endogenous" strain. Also, chemically competent *E. coli* BL21 (DE3) cells were co-transformed with two plasmids of pK184_ *EF-lip*A-*hly*A1BD and pET28_*tol*C. The expression of *tolC* gene of the pET28_*tol*C was under control of T7 promoter. This strain was termed "TolC overexpression" strain. The expression of TolC was under the strong T7 promoter of pET28_*tol*C plasmid, plus the expression of TolC was express continuously endogenous.

The amount of secreted lipase was determined either through lipase related assays such as tributyrin agar plate or *p*-Nitrophenyl butyrate (*p*NPB) assay or through SDS-PAGE analysis.

Tributyrin agar plate : After introducing the plasmids into *E. coli* BL21(DE3), cell were grown on tributyrin agar plate. The diameter of halos formed around colons represents the amount of secreted lipase. As seen in **Figure 1a-c**, the halos around single clones harboring plasmid pET28_*tolC* were smaller in sizes (in average of 2.1 mm) in compared to halos around cells having only plasmid pK184_ *EF-lip*A-*hly*A1BD (in average of 3.4 mm). Based on this observation, the amount of secreted lipase was lower in TolC overexpression strain. Cell clones harboring single plasmid pK184_*hly*BD used as negative clones.

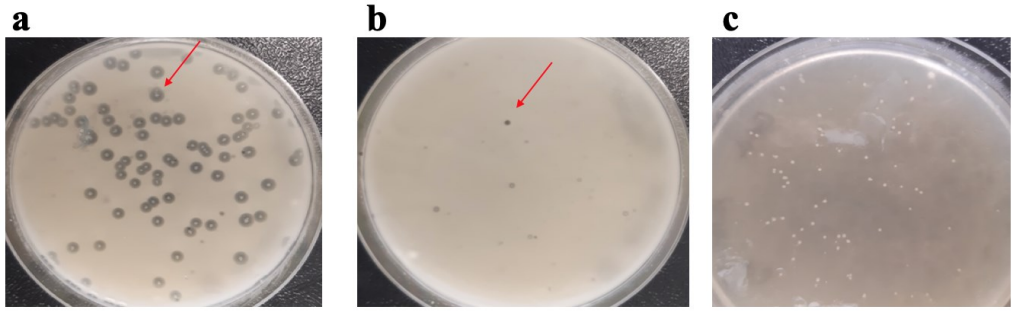


Figure 1: Secretion of lipase-HlyA1 through the HlyA T1SS. The amount of secreted lipase was detected on tributyrin agar plates. The size of halos around cells harboring the plasmids **a**) pK184_ *EF-lip*A-*hly*A1BD and **b**) pK184_ *EF-lip*A-*hly*A1BD plus pET28_tolC, **c**) pK184_*hly*BD. Arrows show the halo around colons.

SDS-PAGE analysis : Additionally, test expression of clones secreting lipase-HlyA1 was performed in shaking flasks and the supernatant of the cultures were employed directly on the SDS-PAGE. A significant reduction in the amount of secretion observed because of TolC overexpression (**Figure 2a**). The Western blot analysis confirmed that all expressed Lipase-HlyA1 in the cytosol of TolC endogenous strain was secreted as no signal for HlyA protein was detected in the cytosol of *E. coli* cells (**Figure 2b**). A strong band at around 60 kDa and some bands related to degraded products were observed in the cytosol of TolC overexpression strain (**Figure 2b**).

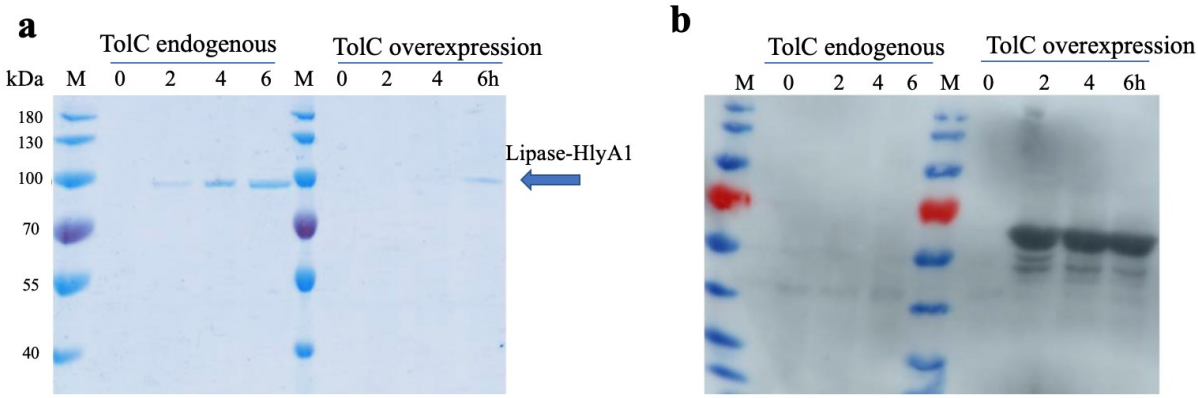


Figure 2: Secretion of Lipase-HlyA1 through the HlyA T1SS. **a**) SDS-PAGE analysis of supernatant (unconcentrated) of cultures of TolC overexpression strain and TolC endogenous strain. **b**) Western blot analysis of *E. coli* cells demonstrated that there was no HlyA in the cytosol of TolC endogenous strain, but for TolC overexpression strain a very strong band around 60 kDa and some degradation bands could be detected.

p-Nitrophenyl butyrate (pNPB) assay: This assay determined the lipase activity through hydrolysis of *p*NPB to butyric acid and p-nitrophenolate. The latter one is a yellow substrate and its production can be continuously monitored at 410 nm. The *p*NPB assay performed in 96-well microtiter plates based on the published protocol (Pourhassan et al., 2022a). Herein, the amount of the secreted lipase was monitored by employing the supernatant of the cultures grown overnight into the assay reaction buffer. The results of this assay confirmed that under condition of TolC overexpression, the amount of Lipase-HlyA1 was significantly reduced by around 3-folds (Figure 3).

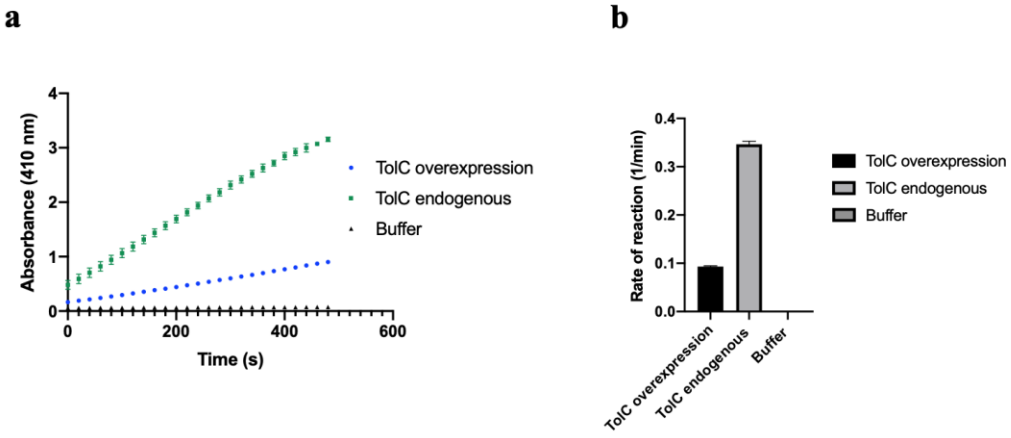


Figure 3: Reduced secretion of Lipase-HlyA1 by overexpression of TolC. a) Activity of the Lipase-HlyA1 in the supernatant of the cultures using the pNPB assay b) Rate of lipase activity dedicated to either the "TolC overexpression strain" or "TolC endogenous strain". The reported absorbance at 410 nm and the error bars correspond to average values of three replicates. The control is the absorbance of the buffer and substrate without addition of culture supernatant.

Analyzing if the overexpression led to an increased TolC in the membrane

To test if the cloned strain for the overexpression of TolC harboring two plasmids of pK184_*EF-lip*A-*hly*A1-*hly*BD and pET28_*tol*C does indeed express higher amount of TolC in the membrane, the amount of TolC in the membrane of two strains 'TolC overexpression' and 'TolC endogenous' was compared. To this end, first both strains were cultivated in shaking flasks in the exact same condition and membrane isolation were performed from *E. coli* cells. Samples from cell pellets were taken at different timepoints after induction and analyzed via western blot (**Figure 4a**). A drastic increase of the TolC was observed for the 'TolC overexpression' strain already at zero time of induction, note that the pET28 plasmid is a leaky plasmid. At three and 6 hours after induction, a band at around 150 kDa was observed on Western blot which represents the trimeric form of TolC, clearly indicating that the overexpression of TolC was successful. In other word, in the strain 'TolC overexpression' the amount of trimeric and monomeric TolC was significantly higher in compared to the membrane of 'TolC endogenous' strain (**Figure 4a and b**).

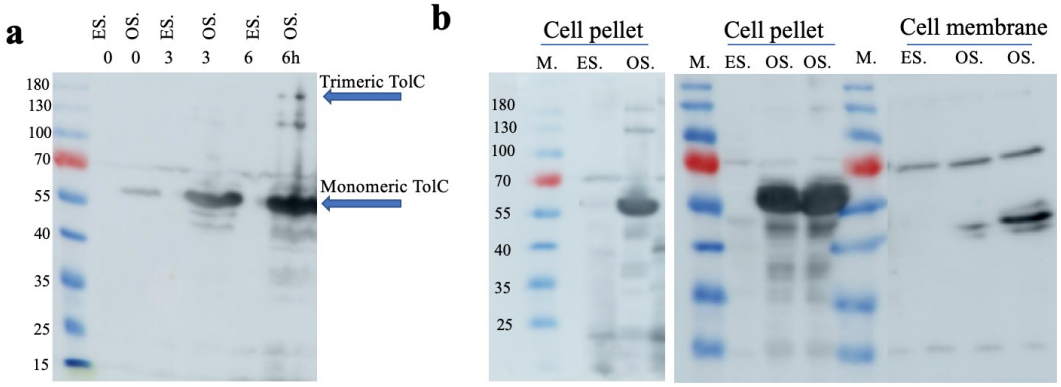


Figure 4: Western blot analysis of **a**) *E. coli* cells and **b**) isolated membrane demonstrated that the expression level of TolC was higher for cells having the pET28_tolC plasmid. ES. refers to endogenous strain and OS. refers to overexpression strain. The band around 55 kDa belongs to TolC protein.

Discussion

Secretion in comparison to cytosolic expression brings benefits to the production of heterologous proteins as it results in production of proteins with higher quality but at reduced costs (Pourhassan et al., 2021). Higher stability and quality of the target protein, reduced risk of cytosolic proteases, simplified downstream processes are advantages of protein secretion. Albeit these great advantages, the potential uses of secretion has been limited in laboratories and especially on industrial scales (Kleiner-Grote et al., 2018, Burdette et al., 2018). This can be seen in case of *E. coli*, which is the most preferred producer host in laboratories and used many times in industrial production, but there is no example of using this bacteria as a secretory host at industrial applications (Kleiner-Grote et al., 2018, Burdette et al., 2018). The main reason of this limitation is the low titers of secretion. In this regard, the secretion system of *E. coli* have been continuously subjected to engineering to gain further improvements (Pourhassan et al., 2021).

One of the secretion systems of *E. coli* is the HlyA T1SS which is interesting to be engineered because of its relatively simple structure. Substrates are transported through the dedicated channel of T1SSs in an unfolded state in one step from cytosol to the extracellular space and fold out of the cell upon secretion. Its secretion pathway consists of the ABC transporter HlyB, the membrane fusion protein HlyD, and the outer membrane protein TolC (Holland et al., 2016).

To improve the secretion efficiency of HlyA T1SS, different approaches have been applied, for instance: a) directed evolution of Hly apparatus (Sugamata and Shiba, 2005, Low et al., 2010), b) engineering on the enhancer fragment of this system (Khosa et al., 2018), c) KnowVolution of the enhancer fragment, HlyB, and HlyD (Pourhassan et al., 2022a).

It has been reported that number of overexpressed HlyA T1SS in *E. coli* BL21(DE3)-BD is around 4500 copies under the conditions of HlyB and HlyD overexpression (Lenders et al., 2016). Only recently, Beer *et al*. quantified the endogenous amount of TolC present in BL21(DE3)-BD. Based on their observation, the amount of trimeric TolC is also around 4500 either in actively secreting cells or cells harboring stalled system (Beer et al., 2022). This match arose a question if the number of active HlyA T1SS in *E. coli* cells is limited by the amount of TolC. To address this question, this study attempts to investigate the influence of TolC overexpression on the secretion efficiency of the Hly secretion system.

Herein, to improve further the system, we overexpressed the outer membrane TolC via plasmid based expression to see if this overexpression leads to further improvements of secretion titer of the HlyA T1SS.

Based on our observation, the plasmid based overexpression led to an increased level of TolC in the membrane of *E. coli*, but this overexpression indeed led to decrease secretion efficiency of the system.

It is possible that TolC overexpression exceeds the capacity of the membrane for this outer membrane protein, leading to intracellular TolC accumulation. Therefore, inclusion body formation and disturbance of the exciting balance between proteins involved in the system might occur. For this reason, an increase in expression of all three membrane protein at same levels might be a better approach for increasing the secretion yield of the system.

As it mentioned earlier, the secretion efficiency of the Hly system is not high enough for its application in industrial purposes. Investigating and improving the secretion yield of heterologous proteins using HlyA1 is a continuing concern for biotechnological purposes.

References

ANDERSON, C., HUGHES, C. & KORONAKIS, V. 2000. Chunnel vision: Export and efflux through bacterial channel-tunnels. *EMBO Reports,* 1**,** 313-318.

BALAKRISHNAN, L., HUGHES, C. & KORONAKIS, V. 2001. Substrate-triggered recruitment of the TolC channel-tunnel during type I export of hemolysin by Escherichia coli. *Journal of molecular biology,* 313**,** 501-510.

BEER, T., HÄNSCH, S., PFEFFER, K., SMITS, S. H., WEIDTKAMP-PETERS, S. & SCHMITT, L. 2022. Quantification and Surface Localization of the Hemolysin A Type I Secretion System at the Endogenous Level and under Conditions of Overexpression. *Applied and Environmental Microbiology,* 88**,** e01896-21.

BURDETTE, L. A., LEACH, S. A., WONG, H. T. & TULLMAN-ERCEK, D. 2018. Developing Gram-negative bacteria for the secretion of heterologous proteins. *Microbial cell factories,* 17**,** 196.

COSTA, T. R., FELISBERTO-RODRIGUES, C., MEIR, A., PREVOST, M. S., REDZEJ, A., TROKTER, M. & WAKSMAN, G. 2015. Secretion systems in Gram-negative bacteria: structural and mechanistic insights. *Nature Reviews Microbiology,* 13**,** 343-359.

DU, D., WANG, Z., JAMES, N. R., VOSS, J. E., KLIMONT, E., OHENE-AGYEI, T., VENTER, H., CHIU, W. & LUISI, B. F. 2014. Structure of the AcrAB–TolC multidrug efflux pump. *Nature,* 509**,** 512-515.

FITZPATRICK, A. W., LLABRÉS, S., NEUBERGER, A., BLAZA, J. N., BAI, X.-C., OKADA, U., MURAKAMI, S., VAN VEEN, H. W., ZACHARIAE, U. & SCHERES, S. H. 2017. Structure of the MacAB–TolC ABC-type tripartite multidrug efflux pump. *Nature microbiology,* 2**,** 1-8.

FREUDL, R. 2018. Signal peptides for recombinant protein secretion in bacterial expression systems. *Microbial cell factories,* 17**,** 52.

HOLLAND, I. B., PEHERSTORFER, S., KANONENBERG, K., LENDERS, M., REIMANN, S. & SCHMITT, L. 2016. Type I Protein Secretion-Deceptively Simple yet with a Wide Range of Mechanistic Variability across the Family. *EcoSal Plus,* 7.

JO, I., KIM, J.-S., XU, Y., HYUN, J., LEE, K. & HA, N.-C. 2019. Recent paradigm shift in the assembly of bacterial tripartite efflux pumps and the type I secretion system. *Journal of Microbiology,* 57**,** 185-194.

KANONENBERG, K., ROYES, J., KEDROV, A., POSCHMANN, G., ANGIUS, F., SOLGADI, A., SPITZ, O., KLEINSCHRODT, D., STÜHLER, K. & MIROUX, B. 2019. Shaping the lipid composition of bacterial membranes for membrane protein production. *Microbial cell factories,* 18**,** 131.

KANONENBERG, K., SPITZ, O., ERENBURG, I. N., BEER, T. & SCHMITT, L. 2018. Type I secretion system—it takes three and a substrate. *FEMS microbiology letters,* 365**,** fny094.

KHOSA, S., SCHOLZ, R., SCHWARZ, C., TRILLING, M., HENGEL, H., JAEGER, K.-E., SMITS, S. H. & SCHMITT, L. 2018. An A/U-rich enhancer region is required for high-level protein secretion through the hlya type i secretion system. *Appl. Environ. Microbiol.,* 84**,** e01163-17.

KLEINER-GROTE, G. R., RISSE, J. M. & FRIEHS, K. 2018. Secretion of recombinant proteins from *E. coli*. *Engineering in Life Sciences,* 18**,** 532-550.

KORONAKIS, V., LI, J., KORONAKIS, E. & STAUFFER, K. 1997. Structure of TolC, the outer membrane component of the bacterial type I efflux system, derived from two-dimensional crystals. *Molecular microbiology,* 23**,** 617-626.

LECHER, J., SCHWARZ, C. K., STOLDT, M., SMITS, S. H., WILLBOLD, D. & SCHMITT, L. 2012. An RTX transporter tethers its unfolded substrate during secretion via a unique N-terminal domain. *Structure,* 20**,** 1778-1787.

LENDERS, M. H., BEER, T., SMITS, S. H. & SCHMITT, L. 2016. *In vivo* quantification of the secretion rates of the hemolysin A Type I secretion system. *Scientific reports,* 6**,** 33275.

LOW, K. O., MAHADI, N. M., RAHIM, R. A., RABU, A., BAKAR, F. D. A., MURAD, A. M. A. & ILLIAS, R. M. 2010. Enhanced secretory production of hemolysin-mediated cyclodextrin glucanotransferase in Escherichia coli by random mutagenesis of the ABC transporter system. *Journal of biotechnology,* 150**,** 453-459.

NICAUD, J.-M., MACKMAN, N., GRAY, L. & HOLLAND, I. 1986. The C-terminal, 23 kDa peptide of *E. coli* haemolysin 2001 contains all the information necessary for its secretion by the haemolysin (Hly) export machinery. *FEBS letters,* 204**,** 331-335.

NOEGEL, A., RDEST, U., SPRINGER, W. & GOEBEL, W. 1979. Plasmid cistrons controlling synthesis and excretion of the exotoxin α-haemolysin of Escherichia coli. *Molecular and General Genetics MGG,* 175**,** 343-350.

PEREZ-PEREZ, J., MÁRQUEZ, G., BARBERO, J.-L. & GUTIÉRREZ, J. 1994. Increasing the efficiency of protein export in Escherichia coli. *Bio/technology,* 12**,** 178-180.

PIMENTA, A., RACHER, K., JAMIESON, L., BLIGHT, M. & HOLLAND, I. 2005. Mutations in HlyD, part of the type 1 translocator for hemolysin secretion, affect the folding of the secreted toxin. *Journal of bacteriology,* 187**,** 7471-7480.

PIMENTA, A., YOUNG, J., HOLLAND, I. & BLIGHT, M. 1999. Antibody analysis of the localisation, expression and stability of HlyD, the MFP component of the *E. coli* haemolysin translocator. *Molecular and General Genetics MGG,* 261**,** 122-132.

POURHASSAN, Z. N., CUI, H., KHOSA, S., DAVARI, M. D., JAEGER, K. E., SMITS, S. H., SCHWANEBERG, U. & SCHMITT, L. 2022. Optimized Hemolysin type 1 secretion system in Escherichia coli by directed evolution of the Hly enhancer fragment and including a terminator region. *ChemBioChem,* 23**,** e202100702.

POURHASSAN, Z. N., SMITS, S. H., AHN, J. H. & SCHMITT, L. 2021. Biotechnological applications of type 1 secretion systems. *Biotechnology advances,* 53**,** 107864.

SPITZ, O., ERENBURG, I. N., BEER, T., KANONENBERG, K., HOLLAND, I. B. & SCHMITT, L. 2019. Type I Secretion Systems—One Mechanism for All? *Protein Secretion in Bacteria***,** 215-225.

SUGAMATA, Y. & SHIBA, T. 2005. Improved secretory production of recombinant proteins by random mutagenesis of hlyB, an alpha-hemolysin transporter from Escherichia coli. *Appl. Environ. Microbiol.,* 71**,** 656-662.

THANABALU, T., KORONAKIS, E., HUGHES, C. & KORONAKIS, V. 1998. Substrate-induced assembly of a contiguous channel for protein export from *E. coli*: reversible bridging of an inner-membrane translocase to an outer membrane exit pore. *The EMBO journal,* 17**,** 6487-6496.

3 Discussion

3.1 Toward secretion of recombinant proteins

For recombinant protein production, secretion has advantages compared to the intracellular expression:

- Simplified purification: Secreted proteins from culture medium can be easily isolated using filtration, reducing the complexity of downstream processes.
- Higher stability of the target protein: As the protein content in the extracellular space is lower, secreted proteins are often more stable than those inside cells.
- Better folding: Due to the lower protein content of the culture medium, proteins fold better outside of the cell and therefore improve their functions.
- Increased yield: Secretion allows for high-level expression of the target protein.
- Lower cost: Secretion can simplify and lower the costs of downstream process (Kleiner-Grote et al., 2018, Azam et al., 2016).

In the outset, it should be acknowledged that not always the secretion offers benefits for recombinant protein production. The yield of secretion for some recombinant proteins can be low, especially when compared to bacterial expression system. Furthermore, secreting some recombinant proteins can be a complex process, requiring the use of a specific expression vector, host, and an optimized culture condition. Purification of secreted recombinant proteins can be challenging when they secrete at low concentration in the culture supernatant. All of these challenges can lead to a high cost of production.

Strategies such as improvement of gene expression rates, using an efficient secretion signal, co-expressing associated proteins such as chaperones, deleting native proteases of the host cell, and further mutagenesis and breeding can be applied to address the problems of a secretion process and exploit secretion systems (Schmidt, 2004).

E. coli is the most preferred host for recombinant protein production on laboratory scales and one of the most convenient on industrial scales for recombinant protein production (Rosano and Ceccarelli, 2014). It is a well-characterized bacterium, and its genetic manipulation is convenient due to many available tools and assays (Owczarek et al., 2019). Furthermore, *E. coli* has a fast growth rate and can quickly reach a high cell density in an appropriate culture condition (Sharma and Chaudhuri, 2017).

Secretion can be consider as a solution to tackle the problems of extracellular expression, such as toxicity of the target protein or degradation by the cytosolic proteases (Wingfield, 2015). Importantly, low secretion titers of recombinant proteins by *E. coli* are a particular concern and impede the utilization of this bacterium as a secretion host on large scales. To our best knowledge, there is no example of using *E. coli* as a secretion host on industrial scales, as its secretion titers for recombinant proteins are in range of mg/L (Burdette et al., 2018, Pourhassan et al., 2021).

Bacillus subtilis, a Gram-positive bacterium, is commonly used for the secretion of proteins on an industrial scale. *B. subtilis* is used to secrete technical enzymes such as lipases, proteases and amylases with product yields up to 20 g/L of supernatant. However, the yields of secretion of heterologous proteins by Bacillus species are often very low (Freudl, 2018, Pohl and Harwood, 2010). The genetically modified eukaryotic CHO cell lines is also used for the secretion of recombinant proteins on a large scale, such as the production of recombinant therapeutic monoclonal antibodies with a product yield of 5 g/L (5 grams per liter of serum free media) (Zhou et al., 2018).

Several manufacturers, such as Lonza at Slough (UK), can produce therapeutic monoclonal antibodies using CHO cell lines at a scale of 25,000 L and produce 120 kg of antibody per week per run (www.lonza.com) (Galbraith et al., 2006). Early example of research on the production of monoclonal antibodies using eukaryotic cell lines report yields of around 100 mg/L. After years of research and efforts into eukaryotic secretion systems, yields of 10 g/L for these cell lines have been achieved (Kunert and Reinhart, 2016). This supports the idea that putting more efforts into bacterial secretion system may also lead to such success.

Some of the efforts on engineering of *E. coli* secretion systems can be summarized as bellow:

- Engineering the Sec-dependent pathway to improve the efficiency of protein secretion (Lee and Jeong, 2013).
- Using chaperon proteins, such as DnaK and GroEL, have been shown to improve the folding and stability of recombinant proteins, resulting in improved secretion efficiency (de Marco et al., 2007).
- By optimizing signal peptides, researchers have been able to improve secretion efficiency (Shi et al., 2021).
- Designing expression vector containing strong promoters (Beshay et al., 2007) or enhancer fragments (Khosa et al., 2018) have been also used for an improved secretion efficiency.

- Optimizing the growth medium also has been consider for an improved secretion efficiency (Huang et al., 2016).

3.2 Biotechnological applications of T1SS

The relatively simple architecture of T1SS made them attractive candidates for biotechnological applications. The introduction of T1SS into a non-native host is rather simple, as it only requires three components of the secretion apparatus plus substrate to be transferred and overexpressed via plasmid (Burdette et al., 2018). Researchers have made significant progress in utilizing T1SS for biotechnology applications. One such application is the secretion of enzymes for biocatalysts. Additionally, T1SS has also been studied for its potential use in the production of vaccines and therapeutic proteins. However, two issues have limited the exploitation of T1SS in biotechnological applications, including a low titer of secretion and the possibility of successful secretion of a heterologous protein (Pourhassan et al., 2021).

▪ **A low secretion titer of T1SS is still a matter.**

A list of heterologous proteins secreted by T1SS was provided in chapter 2.1 of this thesis, along with their achieved secretion titers. Except for one case study, all secreted proteins showed a secretion titer of under 10 mg/L. Accordingly, researchers have tried to improve the secretion titer of T1SS with different approaches listed as:

- Engineering genetically elements (Khosa et al., 2018) (Pourhassan et al., 2022a),
- Transferring secretion system to non-native secretion hosts (Eom et al., 2014),
- Engineering the secretion apparatus (Sugamata and Shiba, 2005) (Low et al., 2010),
- Improving culture conditions (Eom et al., 2006).

Despite these efforts and their achieved successes, these efforts have not met the required concentration for the use on industrial scales due to the fact that the basic secretion of T1SS is generally very low. Nevertheless, large-scale secretion of T1SS has been tried reaching the secretion yield of 540 mg/L (Ihling et al., 2019). Ihling *et al*. reported on the scaling-up production of HlyA1 using a defined mineral medium and in a stirred tank bioreactor. The major advantage of their work was that there was no contamination with proteins and compounds from complex media (Ihling et al., 2019). This work indicated the applicability of the HlyA T1SS on an industrial scale, although the titer of the secretion needs to be further improved.

Nguyen and colleagues have opened up possibilities in using T1SS for producing peptides as active pharmaceutical ingredients. Their efforts resulted in an efficient and affordable method for high-titer expression of peptides and small proteins at levels of up to 2 g/L. In their work, they used the HlyA T1SS secretion signal as a tag, promoting expression of peptides and small proteins in the form of inclusion bodies. They named this system "Numaswitch". The main advantages of using "Numaswitch" include high-titer expression, fewer steps for peptide isolation, prevention of proteolytic aggregation, and protection of the host from cytotoxicity. Using "Numaswitch", they were able to successfully produce over 2 g/L of Teriparatide, with a reported titer 20-fold higher than other peptide production strategies (Nguyen et al., 2021).

- **Not all heterologous proteins can be secreted by T1SS**.

It is important to note that not all heterologous proteins can be successfully secreted through T1SS. According to a study by Bakkes *et al*. (2010), the folding rate of a heterologous protein plays a crucial role in determining its ability to be secreted (Bakkes et al., 2010). This theory was supported by two additional studies on maltose-binding protein (MalE) (Bakkes et al., 2010) and the intestinal fatty acid-binding protein (IFABP) (Schwarz et al., 2012). The data suggests that fast-folding proteins cannot be secreted through the HlyA T1SS channel, likely due to the fact that the folded proteins block the secretion apparatus within the cytosol. There is no available computational method for accurately determining the folding rate of a protein, therefore it is not so far possible to predict successful secretion beforehand.

Only recently, Ahn's group developed a tool to predict the secretion possibility of fused proteins through the ABC transporter. They proposed that the presence of a cationic supercharged region within the target protein hinders secretion. They also suggested that mutations resulting in the neutralization of the supercharged cationic regions rescues the secretion. They generalized their finding to all ABC transporters and mentioned that their work significantly makes progress in secretion through ABC transporters (Byun et al., 2022). However, their finding were not consistent with our observations of the secreted proteins from HlyA T1SS. When we used their provided tool, the result was somewhat counterintuitive as none of our already secreted proteins followed their provided rule of secretion. So, their provided predicting tool on the secretion through ABC transporters is inconsistent on what we have already observed on the HlyA T1SS.

3.3 Directed evolution of T1SS

So far, only a few studies paid attention to the engineering of this system to improve it secretion efficiency. This might be due to the lack of a comprehensive understanding of the mechanism and structure of the system. With the recent report of the cryo-EM structure of the HlyA T1SS (Zhao et al., 2022), it is expected that the engineering and improvement of this system will become much easier, leading to even more biotechnological applications in the future.

Directed protein evolution is one of the engineering approaches that has been employed by different groups to achieve enhanced secretion efficiency of T1SS listed in **Table 3-1**. This is a versatile method for engineering proteins without the need for knowledge of the target protein's molecular structure. To use directed evolution on a specific protein, two steps must be completed: creating a diverse set of mutants and then selecting the best among them (Wong et al., 2006).

Table 3-1: A list of conducted directed evolution on T1SS

Organism	Subject	Method	Folds of Improvement	Reference
E. coli	TliDEF system from *P. fluorescence*	epPCR	3.2-fold	(Eom et al., 2005)
E. coli	HlyA T1SS	epPCR	Up to 27 fold	(Sugamata and Shiba, 2005)
E. coli	HlyA T1SS	epPCR	Up to 2-fold	(Low et al., 2010)
E. coli	Enhancer fragment of the HlyA T1SS	KnowVolution plus including a terminator region	5-folds	(Pourhassan et al., 2022a)
E. coli	The inner membrane complex of the HlyA T1SS	KnowVolution	2.5-folds	Chapter 2.4 of the current thesis

- **Methods of generating diversity**

epPCR is considered a popular and powerful tool for generating diverse mutant libraries in laboratories. This is because epPCR is a fast, robust, versatile, and user-friendly method that offers many advantages over other methods that have been developed so far. The key benefits of epPCR include its ability to quickly generate large numbers of mutants, its ability

to work with a wide range of templates, and its ease of use, which makes it a popular choice (Wong et al., 2004).

But epPCR has its limitations, and this method is not completely random which is attributed to the degeneracy of genetic codes as well as the strong mutational bias of polymerases (Cline and Hogrefe, 2000). A solution to the latter is applying a site saturation mutagenesis approach on the identified beneficial positions determined by epPCR (Wong et al., 2004). In addition, research shows that epPCR is unable to determine around 70% of beneficial positions of a target protein. Because applying an epPCR with highly mutational load increases the risk of inactivation of a large number of populations. In other side, reducing the mutational load do not allow to identify some of the beneficial positions. Casting epPCR (cepPCR) can further increase the number of beneficial positions of a target protein (Yang et al., 2017).

The directed evolution method already applied on T1SS (**Table 3.1**) is limited to epPCR. However, the directed evolution methods employed in different courses of this thesis (chapter 2.2, and chapter 2.3) are the superior strategies, cepPCR and KnowVolution strategy, which usually leads to the identification of further beneficial positions.

- **Screening system**

Screening system is the heart of every directed evolution experiment (Wong et al., 2006). Advances in high-throughput technology fuel the progresses of directed evolution. Screening system are based on the linkage between phenotype and a genotype of the target protein. High throughput screening depends on optical properties such as color (Islam et al., 2018), turbidity (Moore and Arnold, 1996), fluorescent (Peck et al., 2011), and luminescent (Choi et al., 2013). It should be considered that many biomolecules targeted for directed evolution experiments have no directly observable properties. In this regard, a reporter with an observable phenotype is required, or alternatively the target can be associated with a known enzyme in a parallel reaction (Packer and Liu, 2015).

For having a reliable and high-throughput system (HTS), steps such as filtration for removal of cell debris should be avoided. A valuable screening system should be reproducible, reliable, and allow a high-throughput screening at an affordable price. To establish a protocol for an activity-based screening system, some parameters should be taken into account, including: the growth and expression conditions, limiting the amount of evaporation from multi-titer plates during different steps, and improving the activity of the target protein (Wong et al., 2006, Packer and Liu, 2015, Bornscheuer et al., 2019).

A pre-screening to eliminate inactive clones increases the chances of success. It is also important to consider doing a re-screening or double re-screening to avoid selecting false positives. Research has shown that iterative round of screening of a few thousands clones yielded a higher preferable improvement than screening of twenty or fifty thousand clones (Wong et al., 2006, Packer and Liu, 2015, Bornscheuer et al., 2019).

In all three conducted directed evolutions on T1SS, mentioned above, the screening of libraries were performed on the solid phase agar plates. These studies have been limited to tedious and low-throughput assays such as screening based on agar plate assays. This work, for the first time, applied a high-throughput screening assay for directed evolution of T1SS. The applied colorimetric assay, which was based on lipase assay, was designed to measure the amount of release lipase into the supernatant.

3.4 Cross communication between the HlyA T1SS components

A deeper understanding of the HlyA T1SS secretion mechanism can help researchers to increase the secretion efficiency of this system. Further study of the system's components is necessary to gain detailed knowledge of the secretion mechanism. Investigating interactions between the protein components can also enhance understanding of the system.

HlyA is a long polypeptide with 1024 amino acids residues, and has a molecular weight of 110 kDa. It remains unfolded inside of the cytosol waiting to be secreted (Lenders et al., 2016). According to research by Lender *et al.*, the secretion of HlyA is C-directional, meaning that the entire polypeptide should be completely translated before it can be secreted (Lenders et al., 2015). It is still enigmatic how this polypeptide is kept protected from cytosolic proteases, as no known chaperon is involved in this system. However, the CLD domain of HlyB have been suggested to have a chaperon-like activity. This hypothesis is supported by an interaction between the substrate and the CLD. Interestingly, this interaction only occurs with unfolded HlyA and within the GG-repeats, as revealed in series of *in vitro* pull-down experiments. Structural information from NMR and the pull-down assays was used to map the interaction region of the CLD for HlyA (Lecher et al., 2012).

One of the earliest identified interactions between HlyA T1SS components was discovered by the work of Thanabalu *et al.* in 1998 through a series of *in vivo* cross-linking experiments. They reported that the formation of the IM complex is independent of the presence of HlyA and TolC. Both proteins of the IM complex, HlyB and HlyD, bind to HlyA independent of each other. Based on their report, TolC is recruited to form the channel only when HlyA is

engaged with the IM complex. Their data also supported that conformational changes of all three proteins during HlyA translocation occur, indicating a dynamic secretion process. After secretion, TolC is dissociated from the channel (Thanabalu et al., 1998).

Thanabalu *et al.* also found that a HlyB mutant lacking its ATP hydrolysis can still form a complex with HlyD, but cannot transport the substrate as it became trapped inside the channel. Furthermore, they observed that masking the C-terminus of HlyD blocks secretion, but the protein still is oligomerized within the complex (Thanabalu et al., 1998).

In 2003, Benabdelhak *et al.* specified the interaction between HlyB and HlyA to the NBD of HlyB. This interaction is suggested to be involved in the substrate recognition step and to take place in the secretin signal of HlyA. Since HlyA2 lacking the C-terminal 57 residues does not interact with the HlyB NBD. They also found that the interaction between HlyA and HlyB NBD is highly dependent on the presence of nucleotides, either ATP or ADP (Benabdelhak et al., 2003).

During one of the courses of this work, the main focus was placed on the identification of residues of HlyB NBD that are involved in the interaction with HlyA (Pourhassan et al., 2022b). Recent studies provided evidence that a secondary structure, an AH covering residues 970-987, plays a crucial role in recognizing HlyA and is crucial for secretion. In general, AHs are defined as helixes having hydrophobic and polar residues between two faces. It was supposed that the interaction between HlyB NBD and HlyA takes place within the AH of the C-terminal HlyA. Then, locations on HlyB NBD with hydrophobic features were searched. *In silico* experiments suggested for two regions on HlyB NBD as putative binding pockets, and these were characterized through mutagenesis experiments and biochemical assays, which further supported their roles in the interaction (Pourhassan et al., 2022b).

In addition to identifying the putative binding pockets, the study also reported that both pockets show changes upon ATP binding. This finding supports previous research that ATP binding accelerates the dissociation of HlyB and HlyA. It is known that secretion of HlyA is fueled by ATP hydrolysis and the PMF (Koronakis et al., 1991), but it is still not clear at which steps of the secretion process these energy sources are used.

The researcher observed that HlyD interacts with HlyA even in the absence of HlyB and TolC (Thanabalu et al., 1998). The deletion of the 45 N-terminal residues of HlyD blocks secretion, as the truncated version was unable to bind to the substrate, but it can still insert into the membrane and form complex with HlyB. The presence of an “AH” and a “charged

box" in the N-terminal of HlyD were found to play the central role of the cytoplasmic domain of HlyD. The deletion of the AH impaired the secretion, but substrate engagement and TolC recruitment still occurred. The charged box was shown to be necessary for interaction with HlyA and the subsequent recruitment of TolC (Balakrishnan et al., 2001). Subsequently, Pimenta *et al.* provided evidence supporting the idea that HlyD has effects on folding of HlyA (Pimenta et al., 2005).

An outstanding feature of HlyD is its integrity in the HlyA T1SS as a membrane fusion protein. It is assumed that the interaction between HlyD and TolC provides a seal channel spanning both membranes and the periplasmic space (Pimenta et al., 2005). A periplasmic region of HlyD, residues 127-170 with 47% homology to a TolC region, was found to be essential for secretion, but the specific function of this region is not known yet (Schülein et al., 1994). It was suggested that the α-helical tip region of HlyD interacts with TolC, and this region has an important role in the secretion (Kim et al., 2016). More studies is required to characterize the interaction between HlyD and TolC.

Letoffe *et al.* (1994) prior to all these studies demonstrated that the secretion of three T1SS substrates (metalloprotease B and C from *Erwinia chrysanthemi* and HasA from *Serratia marcescens*) requires an ordered association between the substrates and all three T1SS component proteins. They found that the first interaction is between the substrates and the cognate ABC transporter, which promotes the second interaction between the two inner membrane proteins. Finally, interaction between MFP and OMP completes the ordered process of secretion (Letoffe et al., 1994).

It is now well understood that the oligomerization between HlyB and HlyD is one of the earliest prerequisites of the secretion process. Six HlyD and three HlyB dimers associate to form of the inner membrane complex. The translocation pathway is built of a HlyB dimer, and HlyA is translocated through this pathway. However, only one of the HlyB dimers contributes to the formation the translocation pathway, while all three dimers are able to Hydrolyze ATP (Zhao et al., 2022).

The cryo-EM structure of the IM complex of the HlyA T1SS revealed more contacts between the components of the system. For instance, all six HlyD of the complex are located at the peripheral of the three protomers. It was also represented from the data that the AH of the HlyD CD locates close to the cytoplasmic membrane, having hydrophobic contacts with HlyB (Zhao et al., 2022).

Importantly, the interface between neighboring protomers is located in the cytoplasm and is composed of residues 35-50 of HlyD, which create two small helix structures. The acidic residues of these helices interact with a positively charged surface on the CLD of HlyB. The residues involved in oligomerization at the inter-protomer interface from both inner membranes are highly conserved, this concludes that the hetero-dodecamer complex (of the inner membrane component) is a common properties of T1SSs (Zhao et al., 2022).

3.5 The C-terminus of HlyA, a region with multi functions

A) The C-terminus of HlyA is involved in various stages of the secretion process by interacting with the HlyA T1SS machinery. This includes interaction with the HlyB CLD within the RTX domains (Lecher et al., 2012), interaction with the HlyB NBD, which is limited to the AH of HlyA (Pourhassan et al., 2022b). Additionally, it interacts with the CD of HlyD (Thanabalu et al., 1998).

It is not yet clear if these interactions occur in a stepwise or simultaneous manner. However, *in silico* analysis suggests that simultaneous interactions of the CLD and NBD with HlyA is possible, as the distance between the CLD and the putative binding pockets of NBD can be bridged with the RTX and the AH domains of HlyA (Pourhassan et al., 2022b).

B) Being a secretion carrier is the next remarkable feature of the C-terminus of HlyA. As the C-terminus of HlyA contains information required for secretion, it can be used for redirecting heterologous proteins into the milieu. This was first described in 1985 by Gray *et al.* who observed that the deletion of the last C-terminal 27 residues of HlyA abolishes secretion (Gray et al., 1986). The secretion signal for HlyA was found to be located within the 50-60 C-terminal residues of HlyA. However, a fragment of the C-terminus, with a weight of 23 KDa and harboring at least 3 RTX domains, known as HlyA1 lead to a more efficient secretion (Koronakis et al., 1989) (Jarchau et al., 1994) (Nicaud et al., 1986). The biotechnological application of the HlyA T1SS is attributed to this fragment, which is explained in chapter 2.1.

C) Representing a secondary structure within the C-terminus of HlyA is a recent characterized feature of the C-terminus of HlyA. The C-terminus of HlyA has been recently characterized as having a secondary structure playing the role of secretion signal. Spitz *et al*. confirmed the long-held speculation that this structure is involved in the secretion signal of HlyA. They found that an amphipathic alpha-helix between residues 975 and 987 is

necessary for secretion, and that the presence of an aromatic residue, F990, in close proximity to this helix is also important for secretion (Spitz et al., 2022).

D) Importantly, the C-terminus of HlyA has role in the folding of the substrate. This feature was dedicated to the RTX domain of HlyA (Bumba et al., 2016). The RTX binds to Ca^{2+} ions with an affinity of (K_D) approximately 100 μM. The Ca^{2+} concentration in the cytosol is around 300 nM, while this concentration is around 2 mM in the extracellular space. Obviously, RTX domain promotes the folding in the extracellular space (Spitz et al., 2022, Thomas et al., 2014a). This folding was suggested to be the driving force for pulling out the RTX substrates through the translocation pathway in a ratchet-like motion. The primary folding most probably prevents the back sliding of the substrate inside of the channel (Bumba et al., 2016). However, some other studies suggest that rather conformational changes due to ATP hydrolysis is the main driving force for T1SS (Rempel et al., 2019, Stefan et al., 2020).

E) Last but not least, the C-terminus of HlyA acts as a molecular ruler. It was found that there is a consensus length between the GG-repeat and the AH of heterologous proteins secreted by the HlyA T1SS. This consensus length covers 130 residues, and shortening this distance by 30 or 50 residues reduces significantly the amount of secretion.

4 References

ABBY, S. S. & ROCHA, E. P. 2012. The non-flagellar type III secretion system evolved from the bacterial flagellum and diversified into host-cell adapted systems. *Plos Genetics,* 8**,** e1002983.

ALY, K. A. & BARON, C. 2007. The VirB5 protein localizes to the T-pilus tips in *Agrobacterium tumefaciens*. *Microbiology,* 153**,** 3766-3775.

ANDERSON, C., HUGHES, C. & KORONAKIS, V. 2000. Chunnel vision: Export and efflux through bacterial channel-tunnels. *EMBO Reports,* 1**,** 313-318.

ANNÉ, J., ECONOMOU, A. & BERNAERTS, K. 2016. Protein secretion in gram-positive bacteria: from multiple pathways to biotechnology. *In:* BAGNOLI, F. & RAPPUOLI, R. (eds.) *Protein and Sugar Export and Assembly in Gram-positive Bacteria.* Springer.

AZAM, A., LI, C., METCALF, K. J. & TULLMAN-ERCEK, D. 2016. Type III secretion as a generalizable strategy for the production of full-length biopolymer-forming proteins. *Biotechnology and bioengineering,* 113**,** 2313-2320.

BAI, F., LI, Z., UMEZAWA, A., TERADA, N. & JIN, S. 2018. Bacterial type III secretion system as a protein delivery tool for a broad range of biomedical applications. *Biotechnology advances,* 36**,** 482-493.

BAKKES, P. J., JENEWEIN, S., SMITS, S. H., HOLLAND, I. B. & SCHMITT, L. 2010. The rate of folding dictates substrate secretion by the *Escherichia coli* hemolysin type 1 secretion system. *Journal of Biological Chemistry,* 285**,** 40573-40580.

BALAKRISHNAN, L., HUGHES, C. & KORONAKIS, V. 2001. Substrate-triggered recruitment of the TolC channel-tunnel during type I export of hemolysin by *Escherichia coli*. *Journal of molecular biology,* 313**,** 501-510.

BEER, T., HÄNSCH, S., PFEFFER, K., SMITS, S. H., WEIDTKAMP-PETERS, S. & SCHMITT, L. 2022. Quantification and surface localization of the Hemolysin A type I secretion system at the endogenous level and under conditions of overexpression. *Applied and Environmental Microbiology,* 88**,** e01896-21.

BENABDELHAK, H., KIONTKE, S., HORN, C., ERNST, R., BLIGHT, M. A., HOLLAND, I. B. & SCHMITT, L. 2003. A specific interaction between the NBD of the ABC-transporter HlyB and a C-terminal fragment of its transport substrate haemolysin A. *Journal of molecular biology,* 327**,** 1169-1179.

BERG, B. V. D., CLEMONS, W. M., COLLINSON, I., MODIS, Y., HARTMANN, E., HARRISON, S. C. & RAPOPORT, T. A. 2004. X-ray structure of a protein-conducting channel. *Nature,* 427**,** 36-44.

BERGÈS, H., JOSEPH-LIAUZUN, E. & FAYET, O. 1996. Combined effects of the signal sequence and the major chaperone proteins on the export of human cytokines in *Escherichia coli*. *Applied and environmental microbiology,* 62**,** 55-60.

BESHAY, U., MIKSCH, G., FRIEHS, K. & FLASCHEL, E. 2007. Increasing the secretion ability of the *kil* gene for recombinant proteins in *Escherichia coli* by using a strong stationary-phase promoter. *Biotechnology letters,* 29**,** 1893-1901.

BILAL, M., IQBAL, H. M., GUO, S., HU, H., WANG, W. & ZHANG, X. 2018. State-of-the-art protein engineering approaches using biological macromolecules: A review from immobilization to implementation view point. *International journal of biological macromolecules,* 108**,** 893-901.

BORNSCHEUER, U. T., HAUER, B., JAEGER, K. E. & SCHWANEBERG, U. 2019. Directed evolution empowered redesign of natural proteins for the sustainable production of chemicals and pharmaceuticals. *Angewandte Chemie (International ed. in English),* 58**,** 36-40.

BRANDS, S., BRASS, H. U., KLEIN, A. S., SIKKENS, J. G., DAVARI, M. D., PIETRUSZKA, J., RUFF, A. J. & SCHWANEBERG, U. 2021. KnowVolution of prodigiosin ligase PigC towards condensation of short-chain prodiginines. *Catalysis Science & Technology,* 11**,** 2805-2815.

BRÜSER, T. & SANDERS, C. 2003. An alternative model of the twin arginine translocation system. *Microbiological research,* 158**,** 7-17.

BUMBA, L., MASIN, J., MACEK, P., WALD, T., MOTLOVA, L., BIBOVA, I., KLIMOVA, N., BEDNAROVA, L., VEVERKA, V. & KACHALA, M. 2016. Calcium-driven folding of RTX domain β-rolls ratchets translocation of RTX proteins through type I secretion ducts. *Molecular cell,* 62**,** 47-62.

BURDETTE, L. A., LEACH, S. A., WONG, H. T. & TULLMAN-ERCEK, D. 2018. Developing Gram-negative bacteria for the secretion of heterologous proteins. *Microbial cell factories,* 17**,** 1-16.

BYUN, H., PARK, J., FABIA, B. U., BINGWA, J., NGUYEN, M. H., LEE, H. & AHN, J. H. 2022. Generalized Approach towards secretion-based protein production via neutralization of secretion-preventing cationic substrate residues. *International Journal of Molecular Sciences,* 23**,** 6700.

CHENG, F., ZHU, L. & SCHWANEBERG, U. 2015. Directed evolution 2.0: improving and deciphering enzyme properties. *Chemical Communications,* 51**,** 9760-9772.

CHOI, S., HAN, S., LEE, H., CHUN, Y.-J. & KIM, D. 2013. Evaluation of luminescent P450 analysis for directed evolution of human CYP4A11. *Biomolecules & Therapeutics,* 21**,** 487.

CIANCIOTTO, N. P. 2005. Type II secretion: a protein secretion system for all seasons. *Trends in microbiology,* 13**,** 581-588.

CLINE, J. & HOGREFE, H. 2000. Randomize gene sequences with new PCR mutagenesis kit. *Strategies,* 13**,** 157-161.

COLLINS, T. J. 2007. ImageJ for microscopy. *Biotechniques,* 43**,** 25-30.

COSTA, T. R., FELISBERTO-RODRIGUES, C., MEIR, A., PREVOST, M. S., REDZEJ, A., TROKTER, M. & WAKSMAN, G. 2015. Secretion systems in Gram-negative bacteria: structural and mechanistic insights. *Nature Reviews Microbiology,* 13**,** 343-359.

COSTA, T. R., HARB, L., KHARA, P., ZENG, L., HU, B. & CHRISTIE, P. J. 2021. Type IV secretion systems: Advances in structure, function, and activation. *Molecular microbiology,* 115**,** 436-452.

CUI, H., CAO, H., CAI, H., JAEGER, K. E., DAVARI, M. D. & SCHWANEBERG, U. 2020. Computer-assisted Recombination (CompassR) teaches us how to recombine beneficial substitutions from directed evolution campaigns. *Chemistry–A European Journal,* 26**,** 643-649.

DE MARCO, A., DEUERLING, E., MOGK, A., TOMOYASU, T. & BUKAU, B. 2007. Chaperone-based procedure to increase yields of soluble recombinant proteins produced in *Escherichia coli*. *BMC biotechnology,* 7**,** 1-9.

DELANO, W. L. 2002. Pymol: An open-source molecular graphics tool. *CCP4 Newsletter on Protein Crystallography,* 40**,** 82-92.

DOUZI, B., DURAND, E., ROUSSEL, A., CASCALES, E. & CAMBILLAU, C. 2018. Towards a complete structural deciphering of Type VI secretion system. *Current Opinion in Structural Biology,* 49**,** 77-84.

DU, D., WANG, Z., JAMES, N. R., VOSS, J. E., KLIMONT, E., OHENE-AGYEI, T., VENTER, H., CHIU, W. & LUISI, B. F. 2014. Structure of the AcrAB–TolC multidrug efflux pump. *Nature,* 509**,** 512-515.

EOM, G.-T., RHEE, J.-S. & SONG, J.-K. 2006. An efficient secretion of type I secretion pathway-dependent lipase, TliA, in *Escherichia coli*: effect of relative expression levels and timing of passenger protein and ABC transporter. *Journal of Microbiology and Biotechnology,* 16**,** 1422-1428.

EOM, G. T., LEE, S. H., OH, Y. H., CHOI, J. E., PARK, S. J. & SONG, J. K. 2014. Efficient extracellular production of type I secretion pathway-dependent *Pseudomonas fluorescens* lipase in recombinant *Escherichia coli* by heterologous ABC protein exporters. *Biotechnology letters,* 36**,** 2037-2042.

EOM, G. T., SONG, J. K., AHN, J. H., SEO, Y. S. & RHEE, J. S. 2005. Enhancement of the efficiency of secretion of heterologous lipase in *Escherichia coli* by directed evolution of the ABC transporter system. *Applied Environmental Microbiololgy,* 71**,** 3468-3474.

FAHNERT, B., LILIE, H. & NEUBAUER, P. 2004. Inclusion bodies: formation and utilisation. *In:* SCHEPER, T. (ed.) *Physiological Stress Responses in Bioprocesses.* Springer.

FITZPATRICK, A. W., LLABRÉS, S., NEUBERGER, A., BLAZA, J. N., BAI, X.-C., OKADA, U., MURAKAMI, S., VAN VEEN, H. W., ZACHARIAE, U. & SCHERES, S. H. 2017. Structure of the MacAB–TolC ABC-type tripartite multidrug efflux pump. *Nature microbiology,* 2**,** 1-8.

FREUDL, R. 2018. Signal peptides for recombinant protein secretion in bacterial expression systems. *Microbial cell factories,* 17**,** 52.

GALBRAITH, D. J., TAIT, A. S., RACHER, A. J., BIRCH, J. R. & JAMES, D. C. 2006. Control of culture environment for improved polyethylenimine-mediated transient production of recombinant monoclonal antibodies by CHO cells. *Biotechnology progress,* 22**,** 753-762.

GOHLKE, U., PULLAN, L., MCDEVITT, C. A., PORCELLI, I., DE LEEUW, E., PALMER, T., SAIBIL, H. R. & BERKS, B. C. 2005. The TatA component of the twin-arginine protein transport system forms channel complexes of variable diameter. *Proceedings of the National Academy of Sciences,* 102**,** 10482-10486.

GRAY, L., MACKMAN, N., NICAUD, J.-M. & HOLLAND, I. 1986. The carboxy-terminal region of haemolysin 2001 is required for secretion of the toxin from *Escherichia coli*. *Molecular and General Genetics,* 205**,** 127-133.

GREEN, C. A., KAMBLE, N. S., COURT, E. K., BRYANT, O. J., HICKS, M. G., LENNON, C., FRASER, G. M., WRIGHT, P. C. & STAFFORD, G. P. 2019. Engineering the flagellar type III secretion system: improving capacity for secretion of recombinant protein. *Microbial Cell Factory,* 18**,** 1-17.

GREEN, E. R. & MECSAS, J. 2016. Bacterial secretion systems: an overview. *In:* INDIRA T. KUDVA, N. A. C., PAUL J. PLUMMER, QIJING ZHANG, TRACY L. NICHOLSON, JOHN P. BANNANTINE, BRYAN H. BELLAIRE (ed.) *Virulence Mechanisms of Bacterial Pathogens.* American Society for Microbiology.

GREENE, N. P., CROW, A., HUGHES, C. & KORONAKIS, V. 2015. Structure of a bacterial toxin-activating acyltransferase. *Proceedings of the National Academy of Sciences,* 112**,** E3058-E3066.

GROHMANN, E., CHRISTIE, P. J., WAKSMAN, G. & BACKERT, S. 2018. Type IV secretion in Gram-negative and Gram-positive bacteria. *Molecular microbiology,* 107**,** 455-471.

GUEROIS, R., NIELSEN, J. E. & SERRANO, L. 2002. Predicting changes in the stability of proteins and protein complexes: a study of more than 1000 mutations. *Journal of molecular biology,* 320**,** 369-387.

HE, S. Y., LINDEBERG, M., CHATTERJEE, A. K. & COLLMER, A. 1991. Cloned *Erwinia chrysanthemi* out genes enable *Escherichia coli* to selectively secrete a diverse family of heterologous proteins to its milieu. *Proceedings of the National Academy of Sciences,* 88**,** 1079-1083.

HEESEMANN, J., GROSS, U., SCHMIDT, N. & LAUFS, R. 1986. Immunochemical analysis of plasmid-encoded proteins released by enteropathogenic Yersinia sp. grown in calcium-deficient media. *Infection and immunity,* 54**,** 561-567.

HEMMERICH, J., ROHE, P., KLEINE, B., JURISCHKA, S., WIECHERT, W., FREUDL, R. & OLDIGES, M. 2016. Use of a Sec signal peptide library from *Bacillus subtilis* for the optimization of cutinase secretion in *Corynebacterium glutamicum*. *Microbial cell factories,* 15**,** 1-11.

HILGARTH, R. S. & LANIGAN, T. M. 2020. Optimization of overlap extension PCR for efficient transgene construction. *MethodsX,* 7**,** 100759.

HO, B. T., DONG, T. G. & MEKALANOS, J. J. 2014. A view to a kill: the bacterial type VI secretion system. *Cell host & microbe,* 15**,** 9-21.

HOLLAND, I. B., PEHERSTORFER, S., KANONENBERG, K., LENDERS, M., REIMANN, S. & SCHMITT, L. 2016. Type I protein secretion-deceptively simple yet with a wide range of mechanistic variability across the family. *EcoSal Plus,* 7**,** 10-1128.

HOLLAND, I. B., SCHMITT, L. & YOUNG, J. 2005. Type 1 protein secretion in bacteria, the ABC-transporter dependent pathway. *Molecular membrane biology,* 22**,** 29-39.

HUANG, L., WANG, Q., JIANG, S., ZHOU, Y., ZHANG, G. & MA, Y. 2016. Improved extracellular expression and high-cell-density fed-batch fermentation of chitosanase from *Aspergillus Fumigatus* in *Escherichia coli*. *Bioprocess and biosystems engineering,* 39**,** 1679-1687.

IHLING, N., UHDE, A., SCHOLZ, R., SCHWARZ, C., SCHMITT, L. & BÜCHS, J. 2019. Scale-up of a type I secretion system in *Escherichia coli* using a defined mineral medium. *Biotechnology Progress,* 36**,** e2911.

ISLAM, S., LAAF, D., INFANZÓN, B., PELANTOVÁ, H., DAVARI, M. D., JAKOB, F., KŘEN, V., ELLING, L. & SCHWANEBERG, U. 2018. KnowVolution campaign of an aryl sulfotransferase increases activity toward cellobiose. *Chemistry–A European Journal,* 24**,** 17117-17124.

JANA, B., KEPPEL, K. & SALOMON, D. 2021. Engineering a customizable antibacterial T6SS-based platform in *Vibrio natriegens*. *EMBO reports,* 22**,** e53681-e53681.

JARCHAU, T., CHAKRABORTY, T., GARCIA, F. & GOEBEL, W. 1994. Selection for transport competence of C-terminal polypeptides derived from *Escherichia coli* hemolysin: the shortest peptide capable of autonomous HlyB/HlyD-dependent secretion comprises the C-terminal 62 amino acids of HlyA. *Molecular and General Genetics,* 245**,** 53-60.

JI, Y., ISLAM, S., CUI, H., DHOKE, G. V., DAVARI, M. D., MERTENS, A. M. & SCHWANEBERG, U. 2020. Loop engineering of aryl sulfotransferase B for improving catalytic performance in regioselective sulfation. *Catalysis Science & Technology,* 10**,** 2369-2377.

JO, I., KIM, J.-S., XU, Y., HYUN, J., LEE, K. & HA, N.-C. 2019. Recent paradigm shift in the assembly of bacterial tripartite efflux pumps and the type I secretion system. *Journal of Microbiology,* 57**,** 185-194.

JOHNSON, I. S. 1983. Human insulin from recombinant DNA technology. *Science,* 219**,** 632-637.

JUMPER, J., EVANS, R., PRITZEL, A., GREEN, T., FIGURNOV, M., RONNEBERGER, O., TUNYASUVUNAKOOL, K., BATES, R., ŽÍDEK, A. & POTAPENKO, A. 2021. Highly accurate protein structure prediction with AlphaFold. *Nature,* 596**,** 583-589.

KANONENBERG, K., ROYES, J., KEDROV, A., POSCHMANN, G., ANGIUS, F., SOLGADI, A., SPITZ, O., KLEINSCHRODT, D., STÜHLER, K. & MIROUX, B. 2019a. Shaping the lipid composition of bacterial membranes for membrane protein production. *Microbial cell factories,* 18**,** 1-12.

KANONENBERG, K., SCHWARZ, C. K. & SCHMITT, L. 2013. Type I secretion systems–a story of appendices. *Research in microbiology,* 164**,** 596-604.

KANONENBERG, K., SMITS, S. H. & SCHMITT, L. 2019b. Functional reconstitution of HlyB, a type I secretion ABC transporter, in saposin-A nanoparticles. *Scientific reports,* 9**,** 1-12.

KANONENBERG, K., SPITZ, O., ERENBURG, I. N., BEER, T. & SCHMITT, L. 2018. Type I secretion system—it takes three and a substrate. *FEMS microbiology letters,* 365**,** fny094.

KHAN, K. H. 2013. Gene expression in mammalian cells and its applications. *Advanced pharmaceutical bulletin,* 3**,** 257.

KHOSA, S., SCHOLZ, R., SCHWARZ, C., TRILLING, M., HENGEL, H., JAEGER, K.-E., SMITS, S. H. & SCHMITT, L. 2018. An A/U-rich enhancer region is required for high-level protein secretion through the HlyA type I secretion system. *Applied Environmental Microbiology,* 84**,** e01163-17.

KIM, J.-S., SONG, S., LEE, M., LEE, S., LEE, K. & HA, N.-C. 2016. Crystal structure of a soluble fragment of the membrane fusion protein HlyD in a type I secretion system of Gram-negative bacteria. *Structure,* 24**,** 477-485.

KLEINER-GROTE, G. R., RISSE, J. M. & FRIEHS, K. 2018. Secretion of recombinant proteins from *Escherichia coli*. *Engineering in Life Sciences,* 18**,** 532-550.

KORONAKIS, V., HUGHES, C. & KORONAKIS, E. 1991. Energetically distinct early and late stages of HlyB/HlyD-dependent secretion across both *Escherichia coli* membranes. *The EMBO journal,* 10**,** 3263-3272.

KORONAKIS, V., KORONAKIS, E. & HUGHES, C. 1989. Isolation and analysis of the C-terminal signal directing export of *Escherichia coli* hemolysin protein across both bacterial membranes. *The EMBO journal,* 8**,** 595-605.

KORONAKIS, V., LI, J., KORONAKIS, E. & STAUFFER, K. 1997. Structure of TolC, the outer membrane component of the bacterial type I efflux system, derived from two-dimensional crystals. *Molecular microbiology,* 23**,** 617-626.

KORONAKIS, V., SHARFF, A., KORONAKIS, E., LUISI, B. & HUGHES, C. 2000. Crystal structure of the bacterial membrane protein TolC central to multidrug efflux and protein export. *Nature,* 405**,** 914-919.

KRAUSS, U., JÄGER, V. D., DIENER, M., POHL, M. & JAEGER, K.-E. 2017. Catalytically-active inclusion bodies—carrier-free protein immobilizates for application in biotechnology and biomedicine. *Journal of biotechnology,* 258**,** 136-147.

KUNERT, R. & REINHART, D. 2016. Advances in recombinant antibody manufacturing. *Applied microbiology and biotechnology,* 100**,** 3451-3461.

LACERDA, T. L. S., SALCEDO, S. P. & GORVEL, J.-P. 2013. Brucella T4SS: the VIP pass inside host cells. *Current opinion in microbiology,* 16**,** 45-51.

LECHER, J., SCHWARZ, C. K., STOLDT, M., SMITS, S. H., WILLBOLD, D. & SCHMITT, L. 2012. An RTX transporter tethers its unfolded substrate during secretion via a unique N-terminal domain. *Structure,* 20**,** 1778-1787.

LECHER, J., STOLDT, M., SCHWARZ, C. K., SMITS, S. H., SCHMITT, L. & WILLBOLD, D. 2011. 1 H, 15 N and 13 C resonance assignment of the N-terminal C39 peptidase-like domain of the ABC transporter Haemolysin B (HlyB). *Biomolecular NMR assignments,* 5**,** 199-201.

LEE, Y. J. & JEONG, K. J. 2013. Enhanced production of antibody fragment via SRP pathway engineering in *Escherichia coli*. *Biotechnology and bioprocess engineering,* 18**,** 751-758.

LENDERS, M. H., BEER, T., SMITS, S. H. & SCHMITT, L. 2016. *In vivo* quantification of the secretion rates of the hemolysin A type I secretion system. *Scientific reports,* 6**,** 33275.

LENDERS, M. H., REIMANN, S., SMITS, S. H. & SCHMITT, L. 2013. Molecular insights into type I secretion systems. *Biological chemistry,* 394**,** 1371-1384.

LENDERS, M. H., WEIDTKAMP-PETERS, S., KLEINSCHRODT, D., JAEGER, K.-E., SMITS, S. H. & SCHMITT, L. 2015. Directionality of substrate translocation of the hemolysin A type I secretion system. *Scientific reports,* 5**,** 12470.

LEO, J. C., GRIN, I. & LINKE, D. 2012. Type V secretion: mechanism (s) of autotransport through the bacterial outer membrane. *Philosophical Transactions of the Royal Society B: Biological Sciences,* 367**,** 1088-1101.

LETOFFE, S., GHIGO, J. & WANDERSMAN, C. 1994. Secretion of the *Serratia marcescens* HasA protein by an ABC transporter. *Journal of bacteriology,* 176**,** 5372-5377.

LIEBETON, K., ZONTA, A., SCHIMOSSEK, K., NARDINI, M., LANG, D., DIJKSTRA, B. W., REETZ, M. T. & JAEGER, K.-E. 2000. Directed evolution of an enantioselective lipase. *Chemistry & biology,* 7**,** 709-718.

LIN, D. Y.-W., HUANG, S. & CHEN, J. 2015. Crystal structures of a polypeptide processing and secretion transporter. *Nature,* 523**,** 425-430.

LINHARTOVÁ, I., BUMBA, L., MAŠÍN, J., BASLER, M., OSIČKA, R., KAMANOVÁ, J., PROCHÁZKOVÁ, K., ADKINS, I., HEJNOVÁ-HOLUBOVÁ, J. & SADÍLKOVÁ, L. 2010. RTX proteins: a highly diverse family secreted by a common mechanism. *FEMS microbiology reviews,* 34**,** 1076-1112.

LLOSA, M. & ALKORTA, I. 2017. Coupling proteins in type IV secretion. *In:* BACKERT, S. & GROHMANN, E. (eds.) *Type IV Secretion in Gram-negative and Gram-positive Bacteria.* Springer.

LOW, H. H., GUBELLINI, F., RIVERA-CALZADA, A., BRAUN, N., CONNERY, S., DUJEANCOURT, A., LU, F., REDZEJ, A., FRONZES, R. & ORLOVA, E. V. 2014. Structure of a type IV secretion system. *Nature,* 508**,** 550-553.

LOW, K. O., MAHADI, N. M., RAHIM, R. A., RABU, A., BAKAR, F. D. A., MURAD, A. M. A. & ILLIAS, R. M. 2010. Enhanced secretory production of hemolysin-mediated cyclodextrin glucanotransferase in *Escherichia coli* by random mutagenesis of the ABC transporter system. *Journal of biotechnology,* 150**,** 453-459.

LUNELLI, M., KAMPRAD, A., BÜRGER, J., MIELKE, T., SPAHN, C. M. & KOLBE, M. 2020. Cryo-EM structure of the Shigella type III needle complex. *PLoS pathogens,* 16**,** e1008263.

MACKMAN, N., NICAUD, J.-M., GRAY, L. & HOLLAND, I. 1985. Genetical and functional organisation of the *Escherichia coli* hemolysin determinant 2001. *Molecular and General Genetics,* 201**,** 282-288.

MASI, M., DURET, G., DELCOUR, A. H. & MISRA, R. 2009. Folding and trimerization of signal sequence-less mature TolC in the cytoplasm of *Escherichia coli*. *Microbiology,* 155**,** 1847.

MATOS, C. F., BRANSTON, S. D., ALBINIAK, A., DHANOYA, A., FREEDMAN, R. B., KESHAVARZ-MOORE, E. & ROBINSON, C. 2012. High-yield export of a native heterologous protein to the periplasm by the tat translocation pathway in *Escherichia coli*. *Biotechnology and Bioengineering,* 109**,** 2533-2542.

METCALF, K. J., BEVINGTON, J. L., ROSALES, S. L., BURDETTE, L. A., VALDIVIA, E. & TULLMAN-ERCEK, D. 2016. Proteins adopt functionally active conformations after type III secretion. *Microbial cell factories,* 15**,** 1-10.

MILLS, D. R., PETERSON, R. & SPIEGELMAN, S. 1967. An extracellular Darwinian experiment with a self-duplicating nucleic acid molecule. *Proceedings of the National Academy of Sciences,* 58**,** 217-224.

MIYAZAKI, K. 2011. MEGAWHOP cloning: a method of creating random mutagenesis libraries via megaprimer PCR of whole plasmids. *In:* VOIGT, C. (ed.) *Methods in enzymology.* Elsevier.

MOORE, J. C. & ARNOLD, F. H. 1996. Directed evolution of a para-nitrobenzyl esterase for aqueous-organic solvents. *Nature biotechnology,* 14**,** 458-467.

MORGAN, J. L., ACHESON, J. F. & ZIMMER, J. 2017. Structure of a type-1 secretion system ABC transporter. *Structure,* 25**,** 522-529.

MULNAES, D., PORTA, N., CLEMENS, R., APANASENKO, I., REINERS, J., GREMER, L., NEUDECKER, P., SMITS, S. H. & GOHLKE, H. 2020. TopModel: template-based protein structure prediction at low sequence identity using top-down consensus and deep neural networks. *Journal of chemical theory and computation,* 16**,** 1953-1967.

NASKAR, S., HOHL, M., TASSINARI, M. & LOW, H. H. 2021. The structure and mechanism of the bacterial type II secretion system. *Molecular Microbiology,* 115**,** 412-424.

NGUYEN, B.-N., TIEVES, F., ROHR, T., WOBST, H., SCHÖPF, F. S., SOLANO, J. D. M., SCHNEIDER, J., STOCK, J., UHDE, A. & KALTHOFF, T. 2021. Numaswitch: an efficient high-titer expression platform to produce peptides and small proteins. *AMB Express,* 11**,** 1-7.

NICAUD, J.-M., MACKMAN, N., GRAY, L. & HOLLAND, I. 1985. Characterisation of HlyC and mechanism of activation and secretion of hemolysin from *Escherichia coli* 2001. *FEBS letters,* 187**,** 339-344.

NICAUD, J.-M., MACKMAN, N., GRAY, L. & HOLLAND, I. 1986. The C-terminal, 23 kDa peptide of *Escherichi coli* hemolysin 2001 contains all the information necessary for its secretion by the hemolysin (Hly) export machinery. *FEBS letters,* 204**,** 331-335.

NIVASKUMAR, M. & FRANCETIC, O. 2014. Type II secretion system: a magic beanstalk or a protein escalator. *Biochimica et Biophysica Acta - Molecular Cell Research,* 1843**,** 1568-1577.

NOEGEL, A., RDEST, U., SPRINGER, W. & GOEBEL, W. 1979. Plasmid cistrons controlling synthesis and excretion of the exotoxin α-hemolysin of *Escherichia coli*. *Molecular and General Genetics,* 175**,** 343-350.

NOVOA, C., DHOKE, G. V., MATE, D. M., MARTÍNEZ, R., HAARMANN, T., SCHREITER, M., EIDNER, J., SCHWERDTFEGER, R., LORENZ, P. & DAVARI, M. D. 2019. KnowVolution of a fungal laccase toward alkaline pH. *ChemBioChem,* 20**,** 1458-1466.

NUSSBAUMER, M. G., NGUYEN, P. Q., TAY, P. K., NAYDICH, A., HYSI, E., BOTYANSZKI, Z. & JOSHI, N. S. 2017. Bootstrapped biocatalysis: Biofilm-derived materials as reversibly functionalizable multienzyme surfaces. *ChemCatChem,* 9**,** 4328-4333.

ORFANOUDAKI, G. & ECONOMOU, A. 2014. Proteome-wide subcellular topologies of *E. coli* polypeptides database (STEPdb). *Molecular & Cellular Proteomics,* 13**,** 3674-3687.

OSICKOVA, A., BALASHOVA, N., MASIN, J., SULC, M., RODEROVA, J., WALD, T., BROWN, A. C., KOUFOS, E., CHANG, E. H. & GIANNAKAKIS, A. 2018. Cytotoxic activity of *Kingella kingae* RtxA toxin depends on post-translational acylation of lysine residues and cholesterol binding. *Emerging microbes & infections,* 7**,** 1-15.

OWCZAREK, B., GERSZBERG, A. & HNATUSZKO-KONKA, K. 2019. A brief reminder of systems of production and chromatography-based recovery of recombinant protein biopharmaceuticals. *BioMed Research International,* V2019**,** 1-13.

PACKER, M. S. & LIU, D. R. 2015. Methods for the directed evolution of proteins. *Nature Reviews Genetics,* 16**,** 379-394.

PECK, S. H., CHEN, I. & LIU, D. R. 2011. Directed evolution of a small-molecule-triggered intein with improved splicing properties in mammalian cells. *Chemistry & biology,* 18**,** 619-630.

PENG, C., SHI, C., CAO, X., LI, Y., LIU, F. & LU, F. 2019. Factors influencing recombinant protein secretion efficiency in gram-positive bacteria: signal peptide and beyond. *Frontiers in bioengineering and biotechnology,* 7**,** 139.

PEREZ-PEREZ, J., MÁRQUEZ, G., BARBERO, J.-L. & GUTIÉRREZ, J. 1994. Increasing the efficiency of protein export in *Escherichia coli*. *Bio/technology,* 12**,** 178-180.

PIMENTA, A., RACHER, K., JAMIESON, L., BLIGHT, M. & HOLLAND, I. 2005. Mutations in HlyD, part of the type 1 translocator for hemolysin secretion, affect the folding of the secreted toxin. *Journal of bacteriology,* 187**,** 7471-7480.

PIMENTA, A., YOUNG, J., HOLLAND, I. & BLIGHT, M. 1999. Antibody analysis of the localisation, expression and stability of HlyD, the MFP component of the *Escherichia coli* hemolysin translocator. *Molecular and General Genetics,* 261**,** 122-132.

POHL, S. & HARWOOD, C. R. 2010. Heterologous protein secretion by Bacillus species: from the cradle to the grave. *Advances in applied microbiology,* 73**,** 1-25.

POURHASSAN, Z. N., CUI, H., KHOSA, S., DAVARI, M. D., JAEGER, K. E., SMITS, S. H., SCHWANEBERG, U. & SCHMITT, L. 2022a. Optimized Hemolysin type 1 secretion system in *Escherichia coli* by directed evolution of the Hly enhancer fragment and including a terminator region. *ChemBioChem,* 23**,** e202100702.

POURHASSAN, Z. N., HACHANI, E., SPITZ, O., SMITS, S. H. & SCHMITT, L. 2022b. Investigations on the substrate binding sites of hemolysin B, an ABC transporter, of a type 1 secretion system. *Frontiers in Microbiology,* 13**,** e1055032.

POURHASSAN, Z. N., SMITS, S. H., AHN, J. H. & SCHMITT, L. 2021. Biotechnological applications of type 1 secretion systems. *Biotechnology advances,* 53**,** 107864.

PRAMANIK, S., CONTRERAS, F., DAVARI, M. D. & SCHWANEBERG, U. 2021. Protein engineering by efficient sequence space exploration through combination of directed evolution and computational design methodologies. *In:* ZHAO, H., LEE, S. Y., NIELSEN, J. & STEPHANOPOULOS, G. (eds.) *Protein Engineering: Tools and Applications.* WILEY-VCH.

PUETZ, J. & WURM, F. M. 2019. Recombinant proteins for industrial versus pharmaceutical purposes: a review of process and pricing. *Processes,* 7**,** 476.

REED, B. & CHEN, R. 2013. Biotechnological applications of bacterial protein secretion: from therapeutics to biofuel production. *Research in microbiology,* 164**,** 675-682.

REMPEL, S., STANEK, W. & SLOTBOOM, D. 2019. ECF-type ATP-binding cassette transporters. *Annual review of biochemistry,* 88**,** 551-576.

ROBINSON, G. 1951. The haemolysin of *Bacterium coli*. *Microbiology,* 5**,** 788-792.

ROSANO, G. L. & CECCARELLI, E. A. 2014. Recombinant protein expression in *Escherichia coli*: advances and challenges. *Frontiers in microbiology,* 5**,** 172-172.

ROSANO, G. L., MORALES, E. S. & CECCARELLI, E. A. 2019. New tools for recombinant protein production in *Escherichia coli*: a 5-year update. *Protein Science,* 28**,** 1412-1422.

RUANO-GALLEGO, D., FRAILE, S., GUTIERREZ, C. & FERNÁNDEZ, L. Á. 2019. Screening and purification of nanobodies from *Escherichia coli* culture supernatants using the hemolysin secretion system. *Microbial cell factories,* 18**,** 1-13.

RÜBSAM, K., DAVARI, M. D., JAKOB, F. & SCHWANEBERG, U. 2018. KnowVolution of the polymer-binding peptide LCI for improved polypropylene binding. *Polymers,* 10**,** 423.

SANA, T. G., LUGO, K. A. & MONACK, D. M. 2017. T6SS: The bacterial" fight club" in the host gut. *PLoS pathogens,* 13**,** e1006325.

SCHMIDT, F. 2004. Recombinant expression systems in the pharmaceutical industry. *Applied microbiology and biotechnology,* 65**,** 363-372.

SCHMITT, L., BENABDELHAK, H., BLIGHT, M. A., HOLLAND, I. B. & STUBBS, M. T. 2003. Crystal structure of the nucleotide-binding domain of the ABC-transporter haemolysin B: identification of a variable region within ABC helical domains. *Journal of molecular biology,* 330**,** 333-342.

SCHTILEIN, R., GENTSCHEV, I., MOLLENKOPF, H.-J. & GOEBEL, W. 1992. A topological model for the hemolysin translocator protein HlyD. *Molecular and General Genetic,* 234**,** 155-63.

SCHÜLEIN, R., GENTSCHEV, I., SCHLÖR, S., GROSS, R. & GOEBEL, W. 1994. Identification and characterization of two functional domains of the hemolysin translocator protein HlyD. *Molecular and General Genetics* 245**,** 203-211.

SCHWARZ, C. K., LANDSBERG, C. D., LENDERS, M. H., SMITS, S. H. & SCHMITT, L. 2012. Using an *E. coli* type 1 secretion system to secrete the mammalian, intracellular protein IFABP in its active form. *Journal of biotechnology,* 159**,** 155-161.

ŠEBO, P., GLASER, P., SAKAMOTO, H. & ULLMANN, A. 1991. High-level synthesis of active adenylate cyclase toxin of *Bordetella pertussis* in a reconstructed *Escherichia coli* system. *Gene,* 104**,** 19-24.

SHARMA, A. & CHAUDHURI, T. K. 2017. Revisiting *Escherichia coli* as microbial factory for enhanced production of human serum albumin. *Microbial cell factories,* 16**,** 1-19.

SHEEDLO, M. J., CHUNG, J. M., SAWHNEY, N., DURIE, C. L., COVER, T. L., OHI, M. D. & LACY, D. B. 2020. Cryo-EM reveals species-specific components within the *Helicobacter pylori* Cag type IV secretion system core complex. *Elife,* 9**,** e59495.

SHI, L., LIU, H., GAO, S., WENG, Y. & ZHU, L. 2021. Enhanced extracellular production of is PETase in *Escherichia coli* via engineering of the pelB signal peptide. *Journal of Agricultural and Food Chemistry,* 69**,** 2245-2252.

SPITZ, O., ERENBURG, I. N., BEER, T., KANONENBERG, K., HOLLAND, I. B. & SCHMITT, L. 2019. Type I Secretion Systems—One Mechanism for All? *Microbiology spectrum,* 7**,** 215-225.

SPITZ, O., ERENBURG, I. N., KANONENBERG, K., PEHERSTORFER, S., LENDERS, M. H., REINERS, J., MA, M., LUISI, B. F., SMITS, S. H. & SCHMITT, L. 2022. Identity determinants of the translocation signal for a type 1 secretion system. *Frontiers in Physiology,* 12**,** 2585.

STANLEY, P., KORONAKIS, V. & HUGHES, C. 1991. Mutational analysis supports a role for multiple structural features in the C-terminal secretion signal of *Escherichia coli* hemolysin. *Molecular microbiology,* 5**,** 2391-2403.

STEFAN, E., HOFMANN, S. & TAMPÉ, R. 2020. A single power stroke by ATP binding drives substrate translocation in a heterodimeric ABC transporter. *Elife,* 9**,** e55943.

STEMMER, W. P. 1994. DNA shuffling by random fragmentation and reassembly: in vitro recombination for molecular evolution. *Proceedings of the National Academy of Sciences,* 91**,** 10747-10751.

SUGAMATA, Y. & SHIBA, T. 2005. Improved secretory production of recombinant proteins by random mutagenesis of hlyB, an alpha-hemolysin transporter from *Escherichia coli*. *Applied Environmental Microbiology,* 71**,** 656-662.

TERRADOT, L. & WAKSMAN, G. 2011. Architecture of the *Helicobacter pylori* Cag-type IV secretion system. *The FEBS journal,* 278**,** 1213-1222.

THANABALU, T., KORONAKIS, E., HUGHES, C. & KORONAKIS, V. 1998. Substrate-induced assembly of a contiguous channel for protein export from *Escherichia coli*: reversible bridging of an inner-membrane translocase to an outer membrane exit pore. *The EMBO journal,* 17**,** 6487-6496.

THOMAS, S., BAKKES, P. J., SMITS, S. H. & SCHMITT, L. 2014a. Equilibrium folding of pro-HlyA from *Escherichia coli* reveals a stable calcium ion dependent folding intermediate. *Biochimica et Biophysica Acta - Proteins and Proteomics,* 1844**,** 1500-1510.

THOMAS, S., HOLLAND, I. B. & SCHMITT, L. 2014b. The type 1 secretion pathway—the hemolysin system and beyond. *Biochimica et Biophysica Acta -Molecular Cell Research,* 1843**,** 1629-1641.

THOMAS, S., SMITS, S. H. & SCHMITT, L. 2014c. A simple *in vitro* acylation assay based on optimized HlyA and HlyC purification. *Analytical biochemistry,* 464**,** 17-23.

TSIRIGOTAKI, A., DE GEYTER, J., ECONOMOU, A. & KARAMANOU, S. 2017. Protein export through the bacterial Sec pathway. *Nature Reviews Microbiology,* 15**,** 21-36.

VAN DURME, J., DELGADO, J., STRICHER, F., SERRANO, L., SCHYMKOWITZ, J. & ROUSSEAU, F. 2011. A graphical interface for the FoldX forcefield. *Bioinformatics,* 27**,** 1711-1712.

VAN GERVEN, N., GOYAL, P., VANDENBUSSCHE, G., DE KERPEL, M., JONCKHEERE, W., DE GREVE, H. & REMAUT, H. 2014. Secretion and functional display of fusion proteins through the curli biogenesis pathway. *Molecular microbiology,* 91**,** 1022-1035.

VAN GERVEN, N., KLEIN, R. D., HULTGREN, S. J. & REMAUT, H. 2015. Bacterial amyloid formation: structural insights into curli biogensis. *Trends in microbiology,* 23**,** 693-706.

VAN ULSEN, P., UR RAHMAN, S., JONG, W. S., DALEKE-SCHERMERHORN, M. H. & LUIRINK, J. 2014. Type V secretion: from biogenesis to biotechnology. *Biochimica et Biophysica Acta -Molecular Cell Research,* 1843**,** 1592-1611.

VANHERCKE, T., AMPE, C., TIRRY, L. & DENOLF, P. 2005. Reducing mutational bias in random protein libraries. *Analytical biochemistry,* 339**,** 9-14.

VRANCKEN, K., DE KEERSMAEKER, S., GEUKENS, N., LAMMERTYN, E., ANNÉ, J. & VAN MELLAERT, L. 2007. pspA overexpression in *Streptomyces lividans* improves both Sec-and Tat-dependent protein secretion. *Applied microbiology and biotechnology,* 73**,** 1150-1157.

WAGNER, S., GRIN, I., MALMSHEIMER, S., SINGH, N., TORRES-VARGAS, C. E. & WESTERHAUSEN, S. 2018. Bacterial type III secretion systems: a complex device for the delivery of bacterial effector proteins into eukaryotic host cells. *FEMS microbiology letters,* 365**,** fny201.

WALKER, K. L., JONES, A. S. & ROBINSON, C. 2015. The Tat pathway as a biotechnological tool for the expression and export of heterologous proteins in *Escherichia coli*. *Pharmaceutical Bioprocessing,* 3**,** 387-396.

WANDERSMAN, C. & DELEPELAIRE, P. 1990. TolC, an *Escherichia coli* outer membrane protein required for hemolysin secretion. *Proceedings of the National Academy of Sciences,* 87**,** 4776-4780.

WANG, R., SEROR, S. J., BLIGHT, M., PRATT, J. M., BROOME-SMITH, J. K. & HOLLAND, I. B. 1991. Analysis of the membrane organization of an *Escherichia coli* protein

translocator, HlyB, a member of a large family of prokaryote and eukaryote surface transport proteins. *Journal of molecular biology,* 217**,** 441-454.

WANG, W. & MALCOLM, B. A. 1999. Two-stage PCR protocol allowing introduction of multiple mutations, deletions and insertions using QuikChangeTM site-directed mutagenesis. *Biotechniques,* 26**,** 680-682.

WIDMAIER, D. M., TULLMAN-ERCEK, D., MIRSKY, E. A., HILL, R., GOVINDARAJAN, S., MINSHULL, J. & VOIGT, C. A. 2009. Engineering the Salmonella type III secretion system to export spider silk monomers. *Molecular systems biology,* 5**,** 309.

WINGFIELD, P. T. 2015. Overview of the purification of recombinant proteins. *Current protocols in protein science,* 80**,** 6.1.1-6.1.35.

WOESTYN, S., ALLAOUI, A., WATTIAU, P. & CORNELIS, G. R. 1994. YscN, the putative energizer of the Yersinia Yop secretion machinery. *Journal of bacteriology,* 176**,** 1561-1569.

WONG, T. S., TEE, K. L., HAUER, B. & SCHWANEBERG, U. 2004. Sequence saturation mutagenesis (SeSaM): a novel method for directed evolution. *Nucleic acids research,* 32**,** e26-e26.

WONG, T. S., ZHURINA, D. & SCHWANEBERG, U. 2006. The diversity challenge in directed protein evolution. *Combinatorial chemistry & high throughput screening,* 9**,** 271-288.

WU, K.-H. & TAI, P. C. 2004. Cys32 and His105 are the critical residues for the calcium-dependent cysteine proteolytic activity of CvaB, an ATP-binding cassette transporter. *Journal of Biological Chemistry,* 279**,** 901-909.

WU, X., LEE, W., TRAN, L. & WONG, S. 1991. Engineering a *Bacillus subtilis* expression-secretion system with a strain deficient in six extracellular proteases. *Journal of Bacteriology,* 173**,** 4952.

YAN, Z., YIN, M., CHEN, J. & LI, X. 2020. Assembly and substrate recognition of curli biogenesis system. *Nature communications,* 11**,** 1-10.

YANG, J., RUFF, A. J., ARLT, M. & SCHWANEBERG, U. 2017. Casting epPCR (cepPCR): A simple random mutagenesis method to generate high quality mutant libraries. *Biotechnology and bioengineering,* 114**,** 1921-1927.

YANG, J., ZHAO, H.-L., RAN, L.-Y., LI, C.-Y., ZHANG, X.-Y., SU, H.-N., SHI, M., ZHOU, B.-C., CHEN, X.-L. & ZHANG, Y.-Z. 2015. Mechanistic insights into elastin degradation by pseudolysin, the major virulence factor of the opportunistic pathogen *Pseudomonas aeruginosa*. *Scientific Reports,* 5**,** 9936.

YU, X.-J., MCGOURTY, K., LIU, M., UNSWORTH, K. E. & HOLDEN, D. W. 2010. pH sensing by intracellular Salmonella induces effector translocation. *Science,* 328**,** 1040-1043.

ZAITSEVA, J., JENEWEIN, S., JUMPERTZ, T., HOLLAND, I. B. & SCHMITT, L. 2005a. H662 is the linchpin of ATP hydrolysis in the nucleotide-binding domain of the ABC transporter HlyB. *The EMBO journal,* 24**,** 1901-1910.

ZAITSEVA, J., JENEWEIN, S., WIEDENMANN, A., BENABDELHAK, H., HOLLAND, I. B. & SCHMITT, L. 2005b. Functional characterization and ATP-induced dimerization of

the isolated ABC-domain of the haemolysin B transporter. *Biochemistry,* 44**,** 9680-9690.

ZHAO, H., LEE, J. & CHEN, J. 2022. The hemolysin A secretion system is a multi-engine pump containing three ABC transporters. *Cell,* 185**,** 3329-3340. e13.

ZHAO, H. & ZHA, W. 2006. *In vitro*'sexual'evolution through the PCR-based staggered extension process (StEP). *Nature Protocols,* 1**,** 1865-1871.

ZHOU, Y., RAJU, R., ALVES, C. & GILBERT, A. 2018. Debottlenecking protein secretion and reducing protein aggregation in the cellular host. *Current opinion in biotechnology,* 53**,** 151-157.

Acknowledgment

I wish to express my deep appreciation to **Prof. Dr. Lutz Schmitt**. I am truly grateful for his advice, motivation, guidance and support. He gave me all the freedom in the laboratory and his positive attitude and optimism about my experiments inspired me to continue my research even sometimes the screening results were not promising at all. I cannot thank him enough for always having an open ear for me and for my every new idea. The door of his office is always open to students which represents his dedication to them. I sincerely thank him for all encouragements and motivations. His patience and valuable time spent in many discussions have been invaluable. It was great time to be accompanying with him to the CKB meetings, CKB symposiums as well as the discussion talk in Aachen. I could always count on his support during all courses of this study, especially during my pregnancy and parental leave. I am truly fortunate to have had the opportunity to work under Professor Schmitt's supervision. I will miss my time at the institute and I am grateful for all that I have learned from him.

At this point I would like to thank **Prof. Dr. Karl-Erich Jaeger** for kindly taking over the co-supervisor and for the many tips during our follow-up meetings of the CKB. I am deeply appreciative for his valuable comments on the review and manuscripts of this thesis.

I express my gracious thank to **apl. Prof. Dr. Sander Smits**. His critical approach in scientific discussions has been a source of inspiration for me. He argues his points of view in scientific discussion in a clear and informative manner, making discussions educational. I am truly grateful by all the knowledge he has shared with us during the regular Type 1-subgroup meetings and weekly seminars. In particular, I am grateful for the opportunity he gave me to work on binding pocket project. Throughput my thesis, his support and motivation was invaluable. I also thank him for many tips on the review writing, which have greatly improved the quality of the review written in this thesis. I greatly thank him for the time he spent on editing the papers and this thesis.

I express my special thanks to **Dr. Ulrich Schulte** for all coordinating discussions concerning the protein biochemistry internship, which I have supervised in recent years. Despite the challenges that we faced, he always approached them with a positive attitude and found simple solutions to even the most complex problems. I sincerely thanks him for his constant

supports and for always answering my question when I was supervising students. I am also grateful to him for explaining me all the rules that I had to follow during my pregnancy in the lab. It is with great sadness that we have lost him, I always remember him as a friendly and competent collogue.

I take this privilege to express my gratitude to **Prof. Dr. Schwaneberg** for providing me all guidance and support. I am truly grateful to him for the opportunity to work in their lab in Aachen. Due to the friendly atmosphere in their group, I was felt at home-institute there. I would also like to express my sincere thanks to **Haiyang Cui (Ocean)** for sharing his knowledge of library generation and screening techniques with me as well as his support for publishing manuscripts. I would also like to extend my thanks to **Dr. Mahdi Davari** and **Dr. Anne Joelle Ruff** for their support and contributions to my research.

I would like to thanks CLIB Kompetenzzentrum Biotechnologie (CKB) for funding of this Ph.D. I am also grateful to the stuff of CKB for their assistance and supports throughout my study.

I would like to extend my sincere gratitude to **Martina Wesemann** for her efforts in the lab. With her warm and friendly attitude, she was always the first to address any issues, whether it was a malfunctioning machine or a shortage of chemicals. It was a pleasure to talk with her non-work-related topics. I would also like to express my gratitude to **Silke Mavaro** for her efforts in setting up our daily experiments and providing us with the laboratory supplies.

I would like to thank the Biochemistry secretaries, former secretary **Mathilde Blum** and **Tatjana Platz** and current secretary **Corrina Franzkowiak** for organizing all forms and arranging all official works. With their help, every bureaucratic problem was overcome.

I want to express my heartfelt gratitude for the fantastic colleagues (friends) **Dr. Sakshi Khosa**, **Julia Gottstein**, and **Feby Chako** for being great neighbor in lab as well as office. I am truly grateful for our own corner in the office talking about lab news, ourselves, Feby's marriage, Julia's travel etc. It was great time all the coffee breaks and lunches that we shared together. Julia brought with herself an international atmosphere into the lab (in the old building) even if she is a German girl. Special thanks for all night as well as weekend working sessions in the lab. I hope see always your face smiling. Special thanks to Sakshi

for her extreme care and support, answering my stupid questions concerning protein purification without getting annoyed. Sorry that I made her work sometimes more by asking her to correct my writings. Feby with her sincere feeling brought positivity and light to my days in the lab. I like to thank them all for making my time in the institute so enjoyable.

I am also thankful to the current members of the Hly group. It was great to have **Florestan Bilsing** and **Manuel Anlauf** as part of our T1SS team. I would like to thank them for their willingness to help and for their creative ideas. I'm also grateful for their help in solving all of my Akta column problems and for never becoming boring during those times. The long subgroup meetings, sometimes with cookies or pizza, are unforgettable. I would also like to express my thanks to **Eymen Hachani** for his contributions to the T1SS group, especially for his help with the binding pocket project. I wish him all the best in his future endeavors, particularly with his HlyD work. I also would like to welcome **Cigdem Guenes** to the group and wish her success with this really nice and challenging topic.

I would like to express my gratitude to the former Ph.D. students of type I secretion group, **Dr. Tobias Beer**, **Dr. Isabelle Erenburg**, **Dr. Olivia Spitz**, and **Dr. Kerstin Kanonenberg**. Their expertise and dedication to the Hly system greatly enhanced my knowledge about that amazing system. I am extremely grateful for all contributions and discussion I had during courses of this study. Their ready to help attitude allowed me to achieve my goals more effectively. I thank them all for answering my never ending list of questions.

I would also like to mention my deep gratitude to **Dr. Jens Reiners** for his great supports, invaluable advices and help. I would also like to thank him for all the discussions and the teaching time in the SAXS laboratory. Special thanks for him for solving the Akta related problems as well as answering all my questions.

I would like to mention my deep-hearted thanks to **Stefanie Rascka** for being a great support and friend during my promotion. It was great to talk about a long list of topics after work in the afternoons in the office which made me feel relax after work. It was a pleasure to have her as a colleague in supervising the protein biochemistry practical courses either in-person or on-line. It was a great time that we worked together to provide students educational materials during Covid-19 pandemic. I am looking forward to call her Mam Steffi soon. Special thanks for **Vivien Knospe** my first neighbor in the office and laboratory. I am also

grateful to her for all talks especially about our students in the lab. I wish her a good luck for the promotion.

I would like to mention my thanks to **Dr. Tim Kroll**. I am grateful by his cheerful manner, which make you laugh even if you had a bad day in the lab. It was funny to see someone is greeting you from other side of apartment (in the new building) by moving his hands and his body.

I would also like to express my thanks to **Dr. Martin Prescher** for his extreme care and support with the practical courses. Special thanks for sharing all tips and advices concerning supervising students. It was always nice to talk about unbelievable students.

I would like to mention my thanks to the group of **Jun. Prof. Dr. Alexej Kedrov**. They all increased the amount of cheerfulness in our group. I am also grateful to them for giving us company in the Biochemistry I events. Special thanks to **Maryna Loewe**, **Michael Kamel** and **Athanasios Padaopoulos** for the deep talks and bringing the international spirit alive in the institute. I also would like to thank Michael showing me how the gradient maker works. I wish the best luck for their coming-soon defenses.

I would like to express my gratitude to **Dr. Marten Exterkate** for being a great support for editing my introduction. I would also like to thanks **Alexandra Bork**, **Hans Klose**, and **Lea-Maria Nentwig** for being such a great collogues. Their supports, stimulus discussions and talks made my experience in the institute much more enjoyable.

I would also like to thank former Ph.D. of the institute **Dr. Marcel Lagedroste**, **Dr. Manuel Wagner**, **Dr. Julia Schumacher**, and **Dr. Diana Kleinschrodt** for their care and for being there to answer any questions.

I also express my gratitude to **Dr. Katharina Gräfe** und **Dr. Kalpana Shanmugarajah** for their warm welcome. Conversation with them in the beginning of my Ph.D. was helpful, as I was able to learn more about the institute and its unwritten rules. The one-day trip to paddle in the Rhur River in which Katrina, me and Lutz were in same boat is still a fresh memory. I still laugh when I recall that how our boat was spinning out of control before we were finally able to control it, and how we were the slowest team at the start.

During the doctoral thesis I have received active supports from the following list of students: **Florian Neusius**, **Bekdil Tabak**, **Lucas Seidler**, and **Neele Muckhoff**. I would like to thank them all and mention that without them the time in the lab would have been half as nice. I also wish Neele and Florian good luck and many successes by their promotions.

Last but not the least, I would like to thank **my parents** for all the endless love and supports that they have given me throughput my Ph.D. and in life in general. Sometimes a single phone call from my parents was a source of power and could give me the strength to get through a hard day. Special thanks for giving me courage and strength to continue my academic life. Whenever, I think about my father's calm attitude, it helps to reduce my stress. My mother has an incredible positive energy that she spreads around her. Thank them giving me this great opportunity to live in your house again, with my son, for writing my thesis. I feel incredibly lucky to have these amazing people as my parents, and I can never thank them enough for everything they have done for me. I hope they can forgive me for the pain of separation, and I am sorry that I cannot visit them more often, however, they are always in my heart.

I am also grateful of the love of all my four **siblings** as well as Forogh. Thanks for all supports and encouragement I have received from them. Mohammad, Razi and Marzi, they are the best people who can relieve my pressure. My special thanks to Elham for all the time she took care of Hanif allowing me to concentrate on my work. She is not only a bigger sister but also a good friend who has always open ears for me.

Amin words cannot describe my sincere appreciation for all of the support and encouragement you have given me throughout my Ph.D. I would like to thank you for your faith in me, which empowers me at all times and makes me feel stronger. I am grateful that you are always more eager for my academic progress than I am myself. Thank you for all of the times you took care of Hanif by yourself to give me the chance to work in the lab or on my thesis. Thank you for coming with me in the weekend to the lab and helping to refill the tip boxes for the next screening experiments, and carrying the big culture flasks. I am so grateful for having you as a mountain behind me. I promise to put more time for you and for our cute son, Hanif, from now on. I would also like to thank your family for their endless kindness to me. Thanks for all their supports during my writing phase in Iran.

Declaration

The dissertation presented here has been made independently and without unauthorized assistance. The dissertation has not been submitted in this or any similar form to any other institution. I have not taken any unsuccessful promotion exams so far.

_____ _____

Ort, Datum Unterschrift