Optimization of Escherichia coli expression, purification and characterization of antibody fragments

Nakić, Mirna

Master's thesis / Diplomski rad

2023

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, Faculty of Food Technology and Biotechnology / Sveučilište u Zagrebu, Prehrambeno-biotehnološki fakultet**

Permanent link / Trajna poveznica: https://urn.nsk.hr/urn:nbn:hr:159:191157

Sveučilište u Zagrebu

Prehrambeno-biotehnološki fakultet

Molekularna biotehnologija

Diplomski rad

Optimizacija ekspresije, pročišćavanja i karakterizacije fragmenata protutijela iz bakterije *Escherichia coli*

Mirna Nakić, 0217 MB

Zagreb, lipanj 2023.

University of Zagreb

Faculty of Food Technology and Biotechnology

Molecular biotechnology

Master Thesis

Optimization of *Escherichia coli* expression, purification and characterization of antibody fragments

Mirna Nakić, 0217 MB

Zagreb, June 2023.

Université d'Orléans-Université de Zagreb

UFR Sciences et Techniques - Faculté de Nutrition et Biotechnologie - Faculté des Sciences

MASTER SCIENCES DU VIVANT

Spécialité : **Techniques Bio-Industrielles**

INTERNSHIP REPORT

Optimization of Escherichia coli expression, purification and characterization of antibody fragments

BY

Mirna NAKIĆ

1st February – 16th June, 2023

Zagreb, June 2023

ACKNOWLEDGEMENTS

I would like to express my gratitude to my mentor and group leader Vincent Aucagne for accepting me into his team for my internship, and for his help and advice.

Also, I would like to thank my supervisors, Stéphane Charpentier and Mélanie Chenon, for their patience, encouragement, and advice they provided during my internship.

I would like to thank everyone in the "Synthetic protein and bioorthogonal chemistry" group, especially Lylia Azzoug for her help with the LC-MS analyses and Carlo Pifferi for providing the biotinylated peptide for my ELISA test.

I also would like to thank my friends who made my stay in France an unforgettable experience.

But most of all, a huge thanks goes to my family for eternally believing in me and supporting me.

TEMELJNA DOKUMENTACIJSKA KARTICA

Diplomski rad

Sveučilište u Zagrebu
Prehrambeno-biotehnološki fakultet
Centar za molekularnu biofiziku, CNRS, Orléans, Francuska
Znanstveno područje: Biotehničke znanosti
Znanstveno polje: Biotehnologija

Optimizacija ekspresije, pročišćavanja i karakterizacije fragmenata protutijela iz bakterije *Escherichia coli*

Mirna Nakić, 0058213093

Sažetak: Kamelidna nanotijela privukla su značajnu pozornost među istraživačima zbog svojih prednosti u odnosu na tradicionalna antitijela. Ove prednosti uključuju njihovu malu veličinu i prisutnost samo jedne disulfidne veze kao jedine post-translacijske modifikacije, što ih čini vrlo podložnim proizvodnji u mikrobnim sustavima. Međutim, proizvodnja proteina povezanih disulfidnom vezom u reducirajućoj citoplazmi *Escherichiae coli* izazovan je pothvat. Dakle, korišteno je nekoliko tehnika s ciljem optimiziranja proizvodnje nanotijela, kao što je fuzija s redoks aktivnim i/ili proteinom šaperonom kao i korištenje Origami^{TM} soja za kojeg je redoks inženjerstvom postignuto da posjeduje citoplazmu oksidativnih svojstava. Najveći prinos topljivog fuzijskog proteina dobiven je s tioredoksinom (Trx) kao fuzijskim partnerom u BL21 (DE3) soju *E. coli*. Ovaj rezultat je pokazatelj snažne šaperonske aktivnosti tioredoksina koji pomaže nanotijelu da postigne svoju nativnu konformaciju. Uklanjanje fuzijske oznake provedeno je cijepanjem TEV (engl. *tobacco etch virus*) proteazom koja cilja visoko specifičnu sekvencu unutar fuzijskog proteina i nakon što je pocijepano nanotijelo uspješno dobiveno, testovi karakterizacije mogli su započeti. Otkriveno je da proizvedeno nanotijelo ne sadrži očekivani disulfidni most, što je rezultat koji bi potencijalno mogao objasniti neka od opaženih svojstava nanotijela, kao što su neučinkovita renaturacija i stvaranje dimera.
Ključne riječi: nanotijelo, topljiv, disulfidna veza, *Escherichia coli*

Rad sadrži: 25 stranica, 15 slika, 2 tablice, 43 literaturna navoda, 2 priloga
Jezik izvornika: engleski

Rad je u tiskanom i elektroničkom obliku (pdf format) pohranjen u: Knjižnica Prehrambeno-biotehnološkog fakulteta, Kačićeva 23, Zagreb

Mentor: dr.sc. Vincent Aucagne
Nadzornici: prof.dr.sc. Stéphane Charpentier, Mélanie Chenon
Stručno povjerenstvo za ocjenu i obranu:

1. Prof. dr.sc. Igor Stuparević
2. Prof.dr.sc. Višnja Besendorfer
3. Prof.dr.sc. Josef Hamacek
4. Dr.sc. Vincent Aucagne
5. Dr.sc. Marcin Suskiewicz
6. Izv.prof.dr.sc. Ivana Ivančić Baće

Datum obrane: 3.7.2023.

BASIC DOCUMENTATION CARD

Master thesis

University of Zagreb
Faculty of Food Technology and Biotechnology
Centre for Molecular Biophysics, Orléans, France
Scientific area: Biotechnical sciences
Scientific field: Biotechnology

Optimization of Escherichia coli expression, purification and characterization of antibody fragments

Mirna Nakić, 0058213093

Abstract: Camelid VHHs, commonly known as nanobodies, have gained significant attention among researchers due to their distinct advantages over traditional antibodies. These advantages include their small size and the presence of only one disulfide bond as the sole post-translational modification, which makes them highly amenable to production in microbial systems. However, producing disulfide bonded proteins in a reducing cytoplasm of *Escherichia coli* is a challenging endeavor. So, several techniques were employed with the goal of optimizing the production of VHH, such as fusion with a redox active and/or chaperone protein as well as utilizing a redox engineered (Origami™) strain that possesses a more oxidative cytoplasm. The highest yield of soluble VHH fusion protein was obtained with thioredoxin (Trx) as the fusion partner in BL21 (DE3) strain of *E. coli*. This result is indicative of the robust chaperone activity of thioredoxin which assists the VHH to reach its native conformation. Fusion tag removal was performed by cleavage with tobacco etch virus (TEV) protease which targets a highly specific sequence within the fusion protein and once cleaved VHH was successfully obtained, characterization tests could proceed. It was revealed that the produced VHH does not contain the expected disulfide bridge which could potentially explain some of the observed properties of the VHH, such as inefficient renaturation and dimer formation.
Key words: nanobody, soluble, disulfide bond, *Escherichia coli*

Thesis contains: 25 pages, 15 figures, 2 tables, 43 references, 2 annexes
Original in: English
Graduate thesis in printed and electronic (pdf format) version deposited in: Library of the Faculty of Food Technology and Biotechnology, Kačićeva 23, Zagreb

Mentor: Vincent Aucagne, PhD, Centre for Molecular Biophysics, Orléans
Supervisors: Stéphane Charpentier, PhD; Mélanie Chenon
Rewievers:

1. Prof.dr.sc. Igor Stuparević
2. Prof.dr.sc. Višnja Besendorfer
3. Prof.dr.sc. Josef Hamacek
4. Dr.sc. Vincent Aucagne
5. Dr.sc. Marcin Suskiewicz
6. Izv.prof.dr.sc. Ivana Ivančić Baće

Thesis defended: 3.7.2023.

ABSTRACT

Camelid VHHs, commonly known as nanobodies, have gained significant attention among researchers due to their distinct advantages over traditional antibodies. These advantages include their small size and the presence of only one disulfide bond as the sole post-translational modification, which makes them highly amenable to production in microbial systems. However, producing disulfide bonded proteins in a reducing cytoplasm of *Escherichia coli* is a challenging endeavor. So, several techniques were employed with the goal of optimizing the production of VHH, such as fusion with a redox active and/or chaperone protein as well as utilizing a redox engineered (Origami™) strain that possesses a more oxidative cytoplasm. The highest yield of soluble VHH fusion protein was obtained with thioredoxin (Trx) as the fusion partner in BL21 (DE3) strain of *E. coli*. This result is indicative of the robust chaperone activity of thioredoxin which assists the VHH to reach its native conformation. Fusion tag removal was performed by cleavage with tobacco etch virus (TEV) protease which targets a highly specific sequence within the fusion protein and once cleaved VHH was successfully obtained, characterization tests could proceed. It was revealed that the produced VHH does not contain the expected disulfide bridge which could potentially explain some of the observed properties of the VHH, such as inefficient renaturation and dimer formation.

Key words: nanobody, soluble, disulfide bond, *Escherichia coli*

RÉSUMÉ

Les VHH de camélidés, communément appelés nanocorps, ont attiré l'attention des chercheurs en raison de certains avantages qu'ils possèdent par rapport aux anticorps traditionnels. Ces avantages comprennent leur petite taille et la présence d'un seul pont disulfure comme seule modification post-traductionnelle, ce qui les rend très propices à la production dans des systèmes microbiens. Cependant, la production de protéines à pont disulfure dans un cytoplasme réducteur d' *Escherichia coli* est une entreprise difficile. Ainsi, plusieurs techniques ont été employées dans le but d'optimiser la production de VHH, telles que la fusion avec une protéine redox active et/ou chaperonne ainsi que l'utilisation d'une souche génétiquement modifiée (Origami™) qui possède un cytoplasme plus oxydatif. Le rendement le plus élevé de protéine de fusion VHH soluble a été obtenu avec la thirédoxine (Trx) comme partenaire de fusion dans la souche BL21 (DE3) d'*E. coli*. Ce résultat suggère que l'activité chaperonne de la thiorédoxine aide le VHH à atteindre sa conformation native. L'élimination de l'étiquette de fusion a été réalisée par clivage avec une protéase du virus de la gravure du tabac (TEV) qui cible une séquence hautement spécifique dans la protéine de fusion et une fois que le VHH clivé a été obtenu avec succès, les tests de caractérisation ont pu se poursuivre. Il a été révélé que le VHH produit ne contient pas le pont disulfure attendu, ce qui pourrait potentiellement expliquer certaines des propriétés observées du VHH, telles qu'une renaturation inefficace et la formation de dimères.

Mots clés: nanocorps, soluble, pont disulfure, *Escherichia coli*

PRESENTATION OF THE LABORATORY

The experimental part of this master's thesis has taken place at CBM (Centre de Biophysique Moléculaire), UPR (unité propre de recherche) 4301, located in Orléans, France within the group "Synthetic protein and biorthogonal chemistry".

The Center for Molecular Biophysics (CBM) is a research unit within the National Center for Scientific Research (CNRS), a fundamental scientific organization in France that is among the world's leading scientific institutions. CBM collaborates with many companies and is affiliated with the University of Orléans. Sixteen thematic groups organized into four research teams make up the basic organizational structure of CBM. CBM supports an interdisciplinary approach by combining biology, chemistry and physics in investigating the structure and interactions of biomolecules at molecular and cellular levels. One of the main aims is to study the mechanisms that trigger the onset of certain diseases.

My internship was done within the group "Synthetic protein and biorthogonal chemistry", led by Vincent Aucagne. The group's main research themes include new chemical tool for protein synthesis and synthesis of disulfide rich miniproteins.

RESEARCH OBJECTIVES

The objective of this master's thesis was to optimize the recombinant production method of soluble, disulfide bonded VHH (nanobody) targeting a peptide derived from the C-terminal part of the $5HT_7$ receptor using *Escherichia coli* as the expression system. The initial step of the study aimed to evaluate and compare the expression levels of six different VHH fusion constructs to determine the expression of which construct resulted in the highest yield of soluble protein. The next step involved optimizing the reaction conditions of cleavage by tobacco-etch virus (TEV) protease to remove the fusion tag and obtain the cleaved VHH. When significant quantity of VHH was acquired, characterization tests, such as HRMS, circular dichroism and ELISA, were performed.

TABLE OF CONTENTS

1. INTRODUCTION

1.1. Nanobodies

Antibodies and antibody-derived biomolecules have emerged as very prominent tools for a myriad of applications, not only as therapeutics, but their valuable properties also make them crucial for research and diagnostic purposes. Firstly, there is a wide range of potential targets to which they can be directed. Secondly, they exhibit both high affinity and high specificity towards their respective antigens, similar to conventional IgGs (Muyldermans, 2020).

In 1989 the conventional view of the antibodies changed with the discovery of heavy chain-only antibodies (HCAbs) from the sera of camelids (Figure 1). What made this discovery even more interesting was the fact that HCAbs showed robust antigen-binding capabilities, which stood in contrast by the fact that the pathology of the human « heavy chain disease » was associated with truncated antibodies that lack functionality (Arbabi-Ghahroudi, 2017). The smallest antibody fragment with the same antigen-binding activity as the original Ab is the antigen binding domain of HCAbs also known as nanobody (Nb) or VHH (Asaadi et al., 2021). Furthermore, HCAbs lack the first constant CH1 domain within the heavy chain (Jovcevska and Muyldermans, 2019).

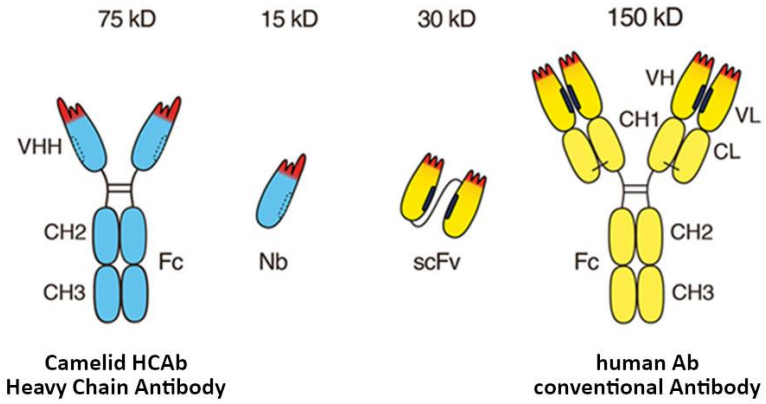

Figure 1. Conventional antibodies consist of 2 heavy and 2 light chains with variable domains of the heavy (VH) and light chain (VL) making up the antigen-binding site. Fc (fragment crystallizable) is made up of second (CH2) and third (CH3) constant domains of the antibody's two heavy chains. A pair of VH and VL domains can be genetically engineered to yield a single-chain variable fragment (scFv). VHH (Nb) is the antigen-binding site of camelid heavy chain-only antibodies. (Bannas et al., 2017)

The nanobodies are arranged in the same manner as the VH domain of human Abs, meaning that the three hypervariable antigen-binding loops (complementarity determining regions, CDR1/2/3) are accompanied by four conserved sequence regions (framework region, FR1/2/3/4) (Figure 2). The three CDRs constitute the paratope of the antibody, but the FRs also participate in antigen

binding by providing a scaffold for the loops. However, the role of CDR3 has been identified as the most decisive for the specificity of antigen binding, whereas CDR1 and CDR2 have a supportive role in strengthening the binding process (Mitchell and Colwell, 2018; Jovcevska and Muyldermans, 2019).

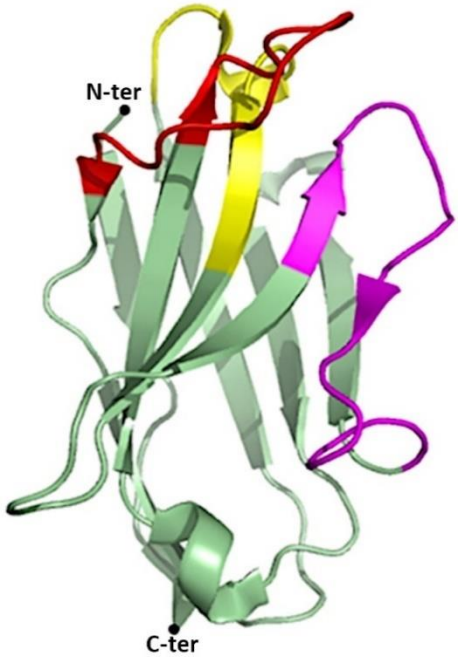

Figure 2. 3D structure of a nanobody. CDR1/2/3 are in yellow, magenta and red, respectively, and the framework regions are in green. N-ter and C-ter of the protein are also indicated. (Beghein and Gettemans, 2017)

1.1.1. Advantages of nanobodies over conventional antibodies

Due to their large size (14,2 nm × 8,2 nm × 3,8 nm) and consequently their large molecular weight (~ 150 kDa) and given their complex structure, the production of antibodies in bacteria and eucaryotic cells remains challenging (Muyldermans, 2020). On the other hand, typical dimensions of a Nb are 4 nm × 2,5 nm × 3 nm with a molecular weight of 12 -14 kDa, and those factors significantly contribute to developing of cost-effective microbial production systems (Jovcevska and Muyldermans, 2019; Wang et al, 2021). Additionally, nanobodies usually have one conserved disulfide bond between Cys22 and Cys92 which is required for their stability, but not necessarily

for their antigen-binding affinity (Kunz et al., 2018). In contrast to that, a typical molecule of IgG1 contains 16 disulfide bonds (Liu et al., 2010).

The overall amino acid composition of the previously mentioned FR2 region of the nanobody is more hydrophilic than that of the conventional antibody which is why nanobodies have better solubility in polar solvents and tend to aggregate less (Kolkman and Law, 2010). The CDR3 loop of nanobodies is usually longer (18 amino acids) than the CDR3 of human Abs because Nbs recognize the antigen with 3 loops instead of 6 loops, however, that finding applies only to VHHs isolated from dromedaries and not from llamas (Kolkman and Law, 2010, Muyldermans, 2013). The elongated CDR3 loop of Nbs is capable of forming an atypical convex paratope or a finger like structure that allows them to access antigens that are often unattainable to canonical Abs, such as catalytic grooves of enzymes or receptor domains (Könning et al., 2017; Jovcevska and Muyldermans, 2019).

Many of the robust physical qualities of nanobodies, such as resistance to proteolytic degradation, high pressures and high temperatures, resistance to denaturing chemicals, stem from their high refolding efficiency (Jovcevska and Muyldermans, 2019). This unconventional stability of Nbs opens up new possibilities of using them as therapeutics for the next generation biodrugs. Nbs also show better tissue penetration that classical Abs and can even be designed to interfere with tumor angiogenesis (Kolkman and Law, 2010; Jovcevska and Muyldermans, 2019). The size of the nanobodies makes them useful for *in vivo* imaging, for which they must be labeled with a radioisotope, but because their molecular size is below the glomerular filtration cutoff (65 kDa), they are cleared from the circulation fast, resulting in their accumulation in kidneys (Jovcevska and Muyldermans, 2019; Asaadi et al., 2021).

1.2. Recombinant production of nanobodies

1.2.1. Choice of the host organism

Goals of recombinant protein production are to achieve high production yields along with proper folding and solubility with a straightforward purification method (Costa et al., 2014). The choice of a host cell in which the heterologous protein will be expressed is one of the first and most important decision a researcher makes, and all the experiments afterwards are tailored towards the chosen expression platform. Advantages of using *E. coli* as an expression organism are readily known: 1) It grows fast; under optimal conditions the generation time of *E. coli* is 20 minutes. 2)

It is easy to achieve high density culture with maximum theoretical density of 1×10^{13} cells/mL. 3) The culture media can be made easily and cheaply. 4) The uptake of exogenous DNA is fast and easy (Rossano and Ceccarelli, 2014). Another factor determining the widespread implementation of *E. coli* in recombinant protein production lays in its well described genome which facilitates genetic manipulation immensely. Still, researchers can have trouble obtaining functional recombinant protein in high yield from *E. coli*. That can occur for a few reasons: toxicity of the exogenous protein, lack of appropriate posttranslational modifications, protein insolubility leading, for example, to the formation of inclusion bodies.

1.2.2. Inclusion body formation

Because of the overexpression of the recombinant protein and/or the innate inability of *E. coli* to carry out proper posttranslational modifications, some proteins tend to form insoluble aggregates called inclusion bodies (IBs) (Bhatwa et al., 2021).

Even though bacterial inclusion bodies have been considered as an obstruction to the recombinant protein production process, their presence has some perks, e.g. they are easy to isolate because of their difference in size and density compared to host cell proteins, they are resistant to proteolysis and mechanically stable (Singhvi et al., 2020). However, it remains a challenge to solubilize the IBs and then to refold the solubilized proteins into their native, active form (Singh et al., 2015). The conventional strategy for obtaining bioactive proteins from IBs includes four steps: isolation of inclusion bodies, their solubilization, refolding of the solubilized proteins and purification of refolded proteins by exploiting various methods (Singhvi et al., 2020).

1.2.3. Fusion of the recombinant protein with solubility-enhancing partners

Among the many tags enhancing the solubility of recombinant proteins, the oxidoreductase thioredoxin (Trx) is one of the most commonly employed fusion partners that is very soluble in the cytoplasm of *E. Coli* (Costa et al., 2014). It is worth noting that bacterial thiol disulfide oxidoreductase proteins DsbA and DsbC were also found to be of interest as solubilizing fusion partners (Nozach et al., 2013).

Fusion tags can enhance solubility and folding of the target protein through two mechanisms. Firstly, the fusion tag protein in itself is inherently very soluble and secondly, the fusion protein employed can be an enzyme that participates in disulfide bond formation (Kim and Lee, 2008).

1.2.4. Fusion of the recombinant protein with partners enabling purification

Since the imidazole ring binds to nickel ions, poly-His tags are the most widely used fusion tags enabling a one-step purification process on Ni^{2+}-loaded resins. (Rosano and Ceccarelli, 2014).

The fusion partner must be cleaved from the recombinant protein if the recombinant protein is to be analyzed for its biochemical or structural properties. This can be achieved by enzyme digestion. Enzymatic digestion of tags necessitates that the expression vector encodes for a protease cleavage site located between the sequence encoding for the tag and the sequence of the protein of interest. Widespread usage of His-tagged tobacco etch virus (TEV) protease for enzymatic cleavage can be attributed to its high specificity and facile production of large quantities of the enzyme (Rosano and Ceccarelli, 2014).

1.2.5. Periplasmic expression as a method for obtaining disulfide-linked proteins

Expression of disulfide-bonded recombinant proteins in the bacterial periplasm offers some advantages compared to cytoplasmic expression. Namely, the bacterial periplasm is an oxidative environment in which the disulfide bonds can be formed (Schimek et al., 2020). In *E. coli* wild types, disulfide bonds (DSBs) are formed in the periplasm because it is a more oxidizing environment with its redox potential around – 165 mV compared to the cytoplasm where the redox potential is between – 260 and – 280 mV (Depuydt et al., 2011).

If the final destination of a protein is the periplasm, such proteins are synthesized in the cytoplasm with an N-terminal signal peptide sequence (SP), which prompts the translocation of the protein after which the signal peptide is cleaved by a signal peptidase. (Schimek et al., 2020) In *E. Coli*, 8 different systems for export of proteins have been found, with the *sec*-pathway being responsible for the transport of majority of proteins from the cytoplasm to into the periplasm (Crane and Randall, 2017). Within the *sec*-pathway, unfolded proteins are translocated across the inner membrane (IM) either post-translationally or co-translationally in adenosine triphosphate (ATP)- and proton-motive force (PMF)-dependent manner (Kleiner-Grote et al., 2018).

Periplasmic expression of recombinant proteins can ease their purification as only the bacterial outer membrane (OM) must be broken down to retrieve the protein and, in theory, there should be less cytoplasmic and cell wall debris in periplasmic extracts (Schimek et al., 2020).

1.2.5.1. Disulfide bond formation

The set of enzymes that participate in generating DSBs, located in the periplasmic compartment, are named Dsb proteins (Berkmen, 2012) (Figure 3). The enzyme that is responsible for the formation of DSBs, DsbA, is a 21 kDa monomeric protein with a CXXC motif in its active site (X is any amino acid) (Berkmen 2012; Landeta et al., 2018). The two cysteines are disulfide linked and catalyze the DSB formation by thiol-disulfide exchange reaction (Landeta et al, 2018). The now reduced DsbA is re-oxidixed into its active oxidized state by a small membrane protein called DsbB which then transfers the electrons to either ubiquinone during aerobic growth, or to menaquinone under anaerobic conditions. (Depuydt et al., 2011; Berkmen, 2012; Landeta et al., 2018). Not all proteins that are excreted in the periplasm demonstrate appropriate folding as DsbA may not have oxidized them correctly. Since DsbA catalyzes the formation of DSBs in the protein as it's being translocated into the periplasm by the *sec* pathway, the disulfide bonds are formed consecutively. That leads to misfolding of proteins whose cysteines are not intended to be linked consecutively. Such misfolded proteins can either be isomerized properly by the protein DsbC or they can be degraded by periplasmic proteases (Berkmen, 2012). DsbC is a homodimeric protein composed of thioredoxin and dimerization domains, with the thioredoxin domain being characterized by four cysteine residues. When the two monomers of DsbC dimerize, a hydrophobic cleft is formed in the middle of the protein which is why DsbC can recognize the hydrophobic surface of misfolded proteins (Berkmen, 2012). The C-terminal cysteine pair forms a structural disulfide link, whereas the N-terminal cysteine pair is a part of the CXXC motif in the active site. DsbD, an inner-membrane protein, is keeping the cysteines in the active site of DsbC in their reduced state and DsbD is reduced by cytoplasmic thioredoxins (Berkmen, 2012; Landeta et al., 2018).

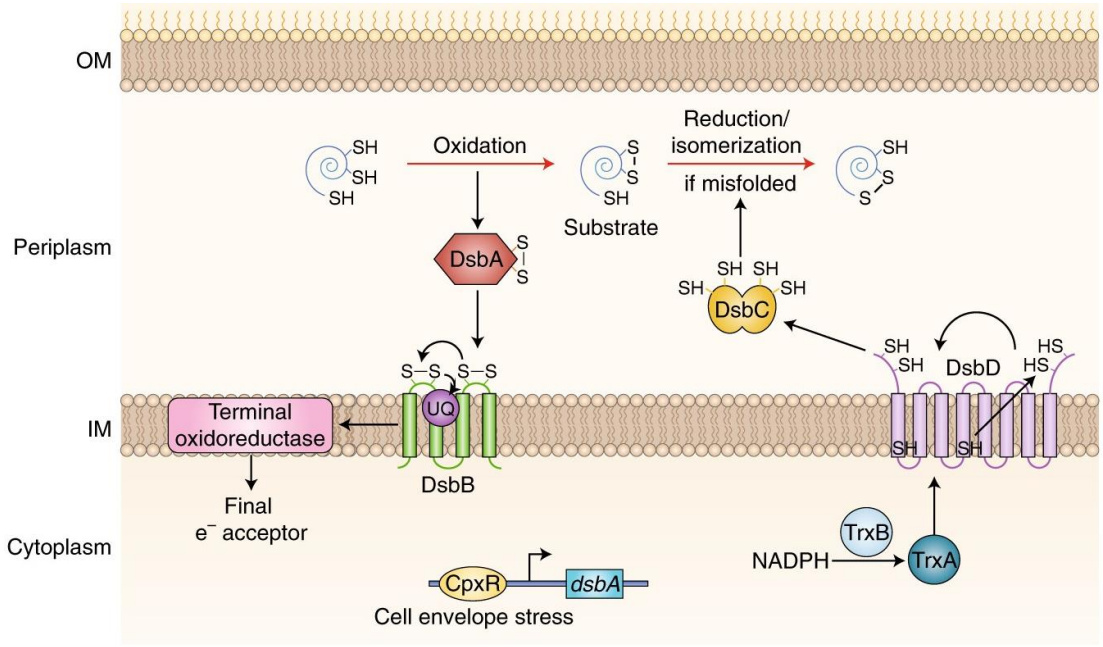

Figure 3. Dsb pathway of disulfide bond formation in *E. coli* (Landeta et al., 2018).

2. MATERIALS AND METHODS

2.1. Transformation of bacteria

2.1.1. Bacterial strains

For the expression of recombinant fusion proteins two competent *Escherichia coli* strains were used: BL21(DE3) (Fisher) and Origami™ B(DE3) pLysS (Novagen, Sigma Aldrich). BL21 (DE3) strain contains DE3 prophage which carries a gene for T7 RNA polymerase under the control of lacUV5 promoter. This strain also lacks 2 proteases: lon protease and outer membrane protease OmpT, which reduces proteolytic degradation of recombinant proteins (Thermo Fisher website, 2023). In Origami™ strain, thioredoxin reductase (*trxB*) and glutathione reductase (*gor*) genes are deleted, resulting in a more oxidative cytoplasm which should enable disulfide bridge formation (Sigma Aldrich website, 2023).

2.1.2. Constructs

Expression plasmids were prepared by Mélanie Chenon using standard molecular biology cloning techniques. The fusion constructs are schematically represented in Figure 4 and their amino acid sequence can be found in Annex (Table 2). The vector that was used for the construction of

expression plasmids was pET which is most widely used for expression of heterologous proteins and is under control of a strong bacteriophage T7 promoter (Sigma Aldrich, 2023).

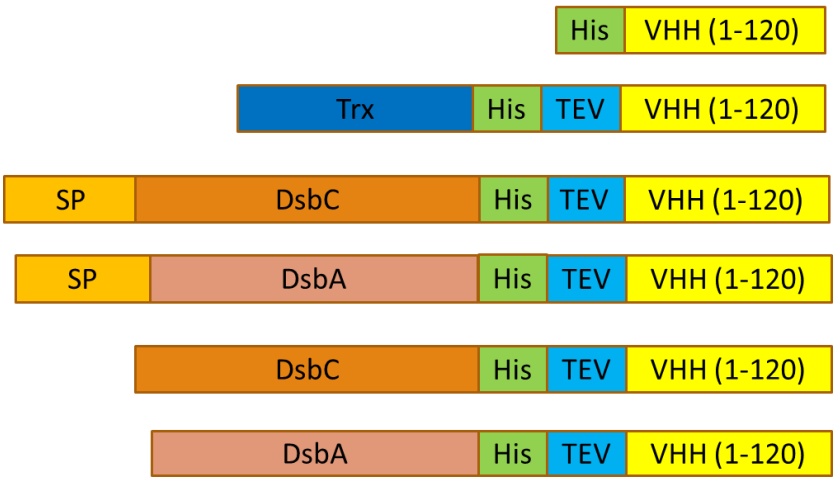

Figure 4. Schematic representation of used fusion constructs. The green block represents 6xHis affinity tag and the blue block represents specific TEV cleavage site (ENLYFQ|A) SP: periplasm-addressing signal peptide, TEV: tobacco etch virus protease site. The yellow block represents the sequence of the VHH directed towards peptide derived from the C-terminal part of the 5HT7 receptor.

Competent bacterial cells were transformed with 1 µL of plasmid DNA (10-40 ng/µL) followed by incubation on ice for 20 minutes, then 1 minute on 42 ˚C and then again 2 minutes on ice. 900 µL of LB media was added to the transformation mixture. That step was followed by incubation for 1 hour at 37 ˚C (180 rpm). Then the cells were centrifuged for 4 minutes at 4000 rpm (centrifuge 5430 R, FA-45-30-11 rotor, Eppendorf) and after removing 900 µL of the LB media, the bacterial pellet was resuspended in the remaining media and used to inoculate LB agar plates supplemented with ampicillin (final concentration of 100 µg/mL) which were incubated on 37 ˚C overnight.

2.2. Recombinant protein expression

2.2.1. Expression in autoinduction media

The next day, one of the colonies that grew on the LB agar plate was transferred into a tube with 2 mL of LB media supplemented with 250 µL of 40 % glucose, 40 µL $MgSO_4$ and 2 µL of ampicillin (final concentration of 100 µg/mL). The bacterial pre-culture was incubated overnight (37 ˚C, 180 rpm) and the next day 1 L of autoinduction media, prepared according to Studier (2005), supplemented with ampicillin was inoculated with 1 mL of the bacterial pre-culture. The inoculated

autoinduction media was incubated for 6 hours (37 ˚C, 180 rpm) and then overnight at 20 ˚C (180 rpm).

2.2.2. Expression in LB media

For the bacterial cultures cultivated in LB broth, one of the colonies was picked from the LB agar plate and transferred into a flask with 50 mL of LB media to which 50 µL of ampicillin was added. The bacterial pre-culture was incubated in the same manner and the next day 1 L of LB media was innoculated with bacterial pre-culture to obtain an inital OD between 0.1 and 0.2. Incubation at 37 ˚C, 180 rpm until the OD reached 0.6 proceeded after which the expression of recombinant protein was induced by adding 1 mM isopropyl β-d-1-thiogalactopyranoside (IPTG). The induced culture was then incubated overnight at 20 ˚C (180 rpm).

2.3. Purification of recombinant protein

2.3.1. Extraction of protein from bacteria

2.3.1.1. Cytoplasmic extraction

Bacterial pellet was obtained by centrifuging the culture for 20 minutes at 4000 rpm (4 ˚C, F12-6x500 rotor) in the Thermo Fisher scientific Sorvall Lynx 400 centrifuge. After an additional wash with PBS, the bacterial pellet was resuspended in buffer A (20 mM Tris-HCl pH 7.5, 1 mM EDTA, 500 mM NaCl, 25 mM imidazole pH 8) and 1 tablet of inhibitors of proteases was added (cOmplete Mini, EDTA-free, Roche). The resuspended bacterial cells were frozen in liquid nitrogen and thawed at room temperature (1 cycle). Lysozyme was added (0.5-0.8 mg/mL final concentration) and the bacterial lysate was agitated for 30 minutes in the cold room. Sonication (40 % intensity, 6 cycles of 30 seconds ON/30 seconds OFF) using the Bioblock scientific, Vibra cell 75115 sonicator was performed followed by centrifugation (F21-8x50y rotor, 22 000 g, 35 minutes, 4 ˚C) in the same previously mentioned centrifuge.

2.3.1.2. Periplasmic extraction

For periplasmic purification, the bacterial pellet ought not to be stored at -20 ˚C because freezing and defreezing prompts the beginning of lysis of bacterial cells, so it is important to have a fresh bacterial pellet which is then resuspended in buffer 1 (100 mM Tris-HCl pH 7.5, 0,5 mM EDTA, 500 mM sucrose, 1 protease inhibitor tablet) by pipetting, as it is important not to mix periplasmic and cytoplasmic fractions. After centrifugation (F21 rotor, 20 000 g, 3 minutes, 4 ˚C) the pellet is

resuspended in Mg-water (1 mM $MgCl_2$) by pipetting and centrifuged again (F21 rotor, 20 000 g, 3 minutes, 4 ˚C). The supernatant is kept for the purification.

2.3.2. Purification on the His-Trap column

For both the cytoplasmic and periplasmic extraction, the protocol for the purification using the ÄKTA pure 25 M system (Cytiva) and the His-Trap column (Cytiva) (V_{column}= 5 mL) was the same. Loading of the soluble fraction (after last centrifugation) on the column (F= 3 mL/min) is followed by a washing of the column with the buffer A. Elution of the recombinant His-tagged protein is carried out stepwise, by increasing the % of buffer B (20 mM Tris pH 7.5, 1 mM EDTA, 500 mM NaCl, 500 mM imidazole pH 8). The eluate is collected in 1 mL fractions which are then analysed on SDS polyacrylamide gel electrophoresis (SDS-PAGE) (15 %, 200 V, 50 min). Protein samples were also analyzed by liquid chromatography-mass spectrometry (LC-MS) for confirmation of identity and integrity.

2.3.3. Cleavage by TEV protease

To optimize the cleavage of purified protein it is important to perform a buffer exchange overnight (ON) at 4 °C. The composition of the dialysis buffer was: 20 mM Tris-HCl pH 7.5, 1 mM EDTA, 200 mM NaCl, 25 mM imidazole pH 8, and the dialysis was performed using Spectra/Por dialysis membrane (7 MWCO 3500, Repligen). Dialyzed construct was incubated with recombinantly expressed TEV protease (produced by Mélanie Chenon). TEV that was used is His-tagged, so when the reaction mixture is purified on the HisTrap column, the TEV protease is retained and thereby separated from the wanted product. TEV protein was added in ratio of 1/20 to the mass of Trx-His-VHH and the concentration of DTT was adjusted to 1 mM final. After incubating the reaction on 25 ˚C for 24 h, additional 1/20 of TEV protease was added to the reaction, thus making the final ratio of TEV 1/10 compared to the fusion protein, and incubated for 24 more hours on 25 ˚C. Afterwards, the sample was purified on the His-Trap column and the cleaved VHH is expected to be in the flow-through, as after cleavage it is not His-tagged anymore and therefore is not retained on the column. Both the flow-through and the eluate were collected in 1 mL fractions and analyzed by SDS-PAGE.

2.4. Characterization of recombinant proteins

2.4.1. Sodium dodecyl-sulfate polyacrylamide gel electrophoresis (SDS-PAGE)

To observe the proteins present in each step of the purification, the protein samples were denatured in the presence of loading sample buffer 5x (0.06 M Tris-HCl pH 6.8, 4 % SDS, 25 % glycerol, 5 % 2-mercaptoethanol, 0,1 % bromophenol blue) at 100 ˚C for 3 to 5 minutes.

The proteins in the sample were then separated by SDS-PAGE (4 % stacking gel and 15 % resolving gel). The separation was performed under 200 V for 50 minutes in 1 X SDS-Glycine migration buffer. The molecular weight marker was PageRuler Prestained Protein Ladder by Thermo Fisher. Following the end of migration, the gel was stained using ReadyBlue Protein Gel Stain (Sigma-Aldrich) and afterwards decolorized in water before taking a picture with Molecular Imager Gel Doc XR+ system (Bio-Rad).

2.4.2. Liquid chromatography-mass spectrometry (LC-MS)

LC-MS analyses were carried out on an Agilent 1260 Infinity HPLC system, coupled with an Agilent 6120 mass spectrometer (ESI + mode), and fitted with an Aeris Widepore XB-C18 2 (3.6 µm, 150 × 2.1 mm, 0.5 mL/min flow rate, 60 °C) column. Solvents A and B were 0.1 % formic acid in H_2O and 0.1 % formic acid in MeCN, respectively. Gradient: 3% B for 0.6 min, then 3 to 50% B over 10.8 min. The multiply charged envelope was deconvoluted using the charge deconvolution tool in Agilent OpenLab CDS ChemStation software to obtain the average [M] value.

2.4.3. Characterization of the disulfide bond with liquid chromatography-high resolution mass spectrometry (LC-HRMS)

The molecular weight observed by the LC-MS method has a somewhat limited precision, typically with an error range of ± 2 Da. This level of precision unfortunately corresponds exactly to the shift in molecular weight associated with the absence or presence of a single disulfide bond (2 Da). To obtain more accurate results and higher resolution, a more precise method, LC-HRMS was employed. The increased resolution of LC-HRMS allows for a more precise determination of molecular weights and can provide more reliable information regarding the presence or absence of a disulfide bond.

HRMS analysis was performed by Guillame Gabant on a purified VHH sample with the concentration of 0.4 mg/mL with the sample previously desalted by LC and injected into the HRMS instrument (maXis II, Bruker) by electrospray ionization (ESI).

2.4.4. Circular dichroism (CD)

Circular dichroism (CD) analyses were performed (by Vincent Aucagne) on a Jasco J-810 spectropolarimeter equipped with a cell with an optical path length of 1 mm. Samples of VHH were dialyzed overnight (4 ˚C) against a buffer (10 mM phosphate buffer pH 7.4, 300 mM NaF) prior to CD analysis. Nanobody molecules were diluted in the dialysis buffer to a final concentration of 30 µg/mL.

CD spectra were taken between 190 and 250 nm. In the variable temperature measurements, spectra (12 scans, 50 nm/min) were recorded from 20 to 90 °C (every 2 °C), with a slope of 5 °C/min between each temperatures. The melting temperature (Tm) was calculated by determining the inflection point of the transition in the measured ellipticity at $\lambda = 216$ nm versus temperature curve, which was fitted using the Origin 9 software according to a Boltzmann function.

2.4.5. Enzyme-linked immunosorbent assay (ELISA)

VHH samples obtained in the previously described manner were prepared for ELISA by performing dialysis overnight (4 ˚C) against a PBS buffer with 100 mM NaCl. The sample was then diluted with PBS to obtain the concentration of 10 µg/mL. 100 µL (1 µg) of VHH was used to coat 33 wells on a 96-well Maxisorp (Thermo Fisher Scientific) plate, overnight at 4 ˚C. After the adsorption of VHH, the wells were washed 3 times with PBS-T (PBS + 0.05 % Tween 20). Non-specific binding sites were saturated with PBS-T+1 % BSA for 1 h at room temperature (with agitation), after which the wells were washed 3 times with PBS-T. In each first well 1 mM of the antigen was put (anti-biotinylated-CPII-072 D-peptide, prepared by Carlo Pifferi) and then the peptide was diluted by 1/10 in each next well (dilutions were done in PBS-T; the range of concentration was 1 mM-1 pM). The biotinylated peptide was incubated for 1 h at 37 ˚C, followed by 3 more washings with PBS-T. Then horseradish peroxidase (HRP)-conjugated streptavidin (Ultra-Streptavidin-HRP, Thermo Fisher Scientific) diluted at 1/1000 in PBS-T was added in each well, incubated for 1 h at 37 ˚C and washed 3 times with PBS-T. The substrate 3,3',5,5'-tetramethylbenzidine (TMB; 1-Step Ultra TMB-ELISA, Thermo Fisher Scientific) was added in

each well to reveal the peroxidase activity and incubated until the desired color developed (15 minutes), which was followed by addition of 2M sulfuric acid to stop the reaction. The absorbance was read at 450 nm using the CLARIOstar Plus microplate reader (BMG Labtech).
The ELISA curve was fitted using the Origin 9 software according to a Boltzmann function.

3. RESULTS

3.1. Overview

With the aim of optimizing the recombinant production and characterization of nanobodies, the primary focus of this master's thesis was to obtain the fusion constructs mentioned earlier by expressing them in *Escherichia coli* and subsequently purifying them using affinity chromatography. Initial experiments involved evaluating two different growth media (LB and autoinduction media) under two conditions (one for LB media and one for autoinduction media) using the same fusion construct. Since higher yield was observed in autoinduction media, it was selected as the preferred medium for expressing the remaining fusion constructs. Following each purification, the fractions that were presumed to contain the desired protein were subjected to analysis using SDS-PAGE and LC-MS to confirm the correct molecular weight and integrity. The highest yield was achieved when bacteria were transformed with a plasmid encoding the Trx-His-VHH construct, prompting the repetition of its expression to recover a substantial amount of the fusion protein for subsequent removal of the fusion tag and generation of active recombinant VHH. Reaction conditions of cleavage by TEV protease were optimized (results shown in Annex) and applied in the final reaction where a sufficient quantity of the VHH was recovered to perform characterization assays such as circular dichroism, ELISA assay and HRMS to precisely determine the molecular weight of the VHH to assess the presence/absence of the disulfide bond.

3.2. Protein expression and purification

For every purification that was carried out on the His-Trap column (as described in Materials and Methods), similar chromatograms were obtained, and on Figure 5 one exemplary chromatogram was shown.

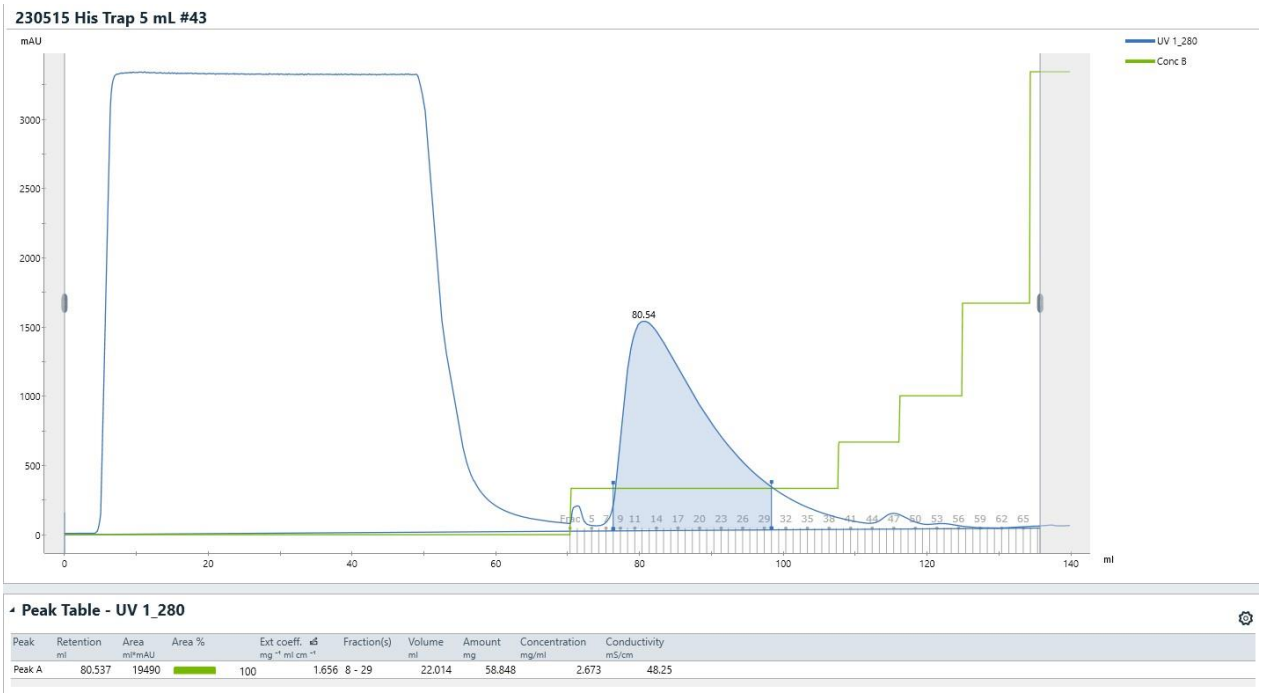

Figure 5. A standard chromatogram obtained from affinity chromatography is depicted (purification of Trx-His-VHH). The absorbance at 280 nm is represented by the blue line, indicating the protein elution profile. A peak in absorbance indicates presence of a protein. The green line corresponds to the concentration of buffer B (elution buffer) during the stepwise elution process.

As previously stated in Materials and Methods section, after every purification done by affinity chromatography, SDS-PAGE analysis was conducted to verify the presence of protein of interest in fractions corresponding to the peak in absorbance. An exemplar SDS-PAGE gel is displayed in Figure 6.

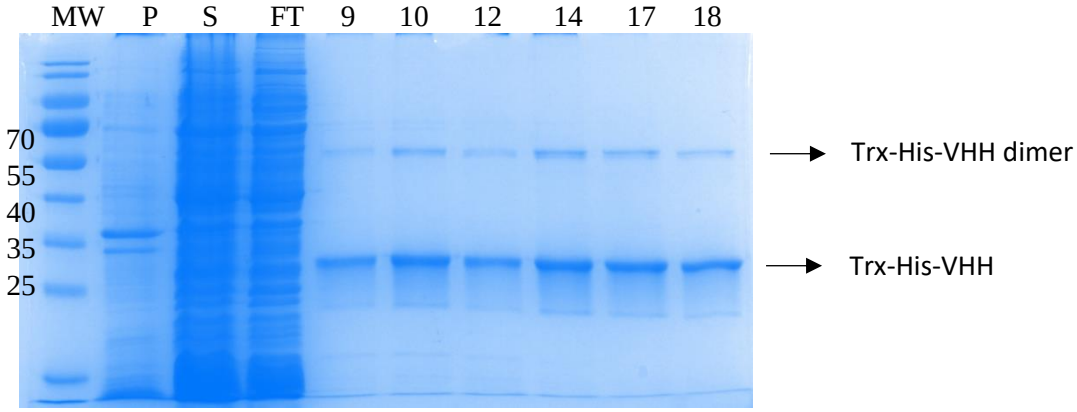

Figure 6. A representative SDS-PAGE gel performed after every purification step to validate the presence of target protein (MW is the molecular weight marker, P-pellet, S-supernatant, FT-flow through, 9, 10, 12, 14, 17, 18-samples from the collected fractions corresponding to the peak in absorbance). The most visible band is Trx-His-VHH (28.3 kDa), and also a dimeric form of the protein is present (56.6 kDa)

All purified constructs can be seen in Figure 7 and in Table 1 all the yields after purification were listed. SDS-PAGE in Figure 7 was carried out with 3 μg of each fusion protein.

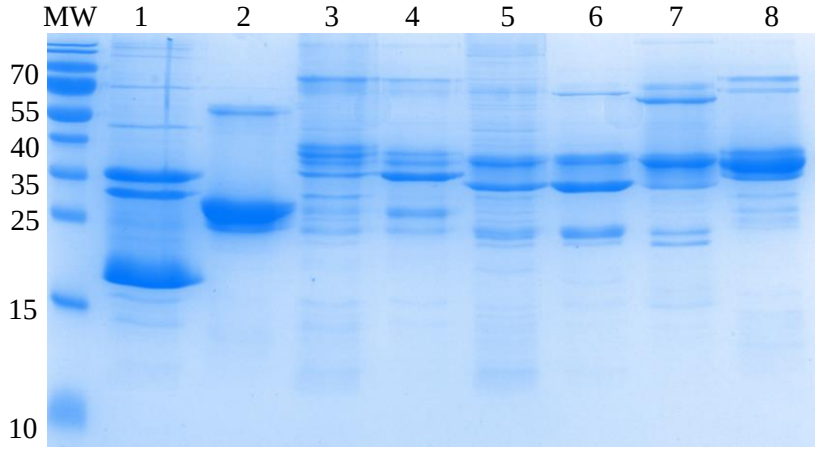

Figure 7. SDS-PAGE with all the purified constructs. MW is the molecular weight marker. 1: His-VHH (16.8 kDa), 2: Trx-His-VHH (28.3 kDa), 3: SP-DsbC-His-VHH (42.1 kDa), 4: SP-DsbC-His-VHH (periplasmic extraction, 40 kDa), 5: SP-DsbA-His-VHH (39.6 kDa), 6: SP-DsbA-His-VHH (periplasmic extraction, 37.6 kDa), 7: DsbA-His-VHH (37.9 kDa), 8: DsbC-His-VHH (40.3 kDa)

Table 1. All purified constructs with the yields.

Construct	Condition	Extraction protocol	Yield ($mg_{fusion\ protein}/L_{culture}$)	Theoretical MW (with DSB) [Da]	MW observed on LC-MS [Da]
Trx-His-VHH	LB ON at 20 ˚C after induction	Cytoplasmic	2.5	28304	28304
	autoinduction media 6 h at 37 ˚C then ON at 20 ˚C		84 (1st try), 120 (2nd try) 0 (in Origami strain)	28304 28304 /	28304 28322** /
His-VHH		Cytoplasmic	17.2	16824	16824
SP-DsbC-His-VHH		Periplasmic	1.4	40008	40006
		Cytoplasmic	6	40008* 42172	40007* 42171
DsbC-His-VHH		Cytoplasmic	60	40268	40267
SP-DsbA-His-VHH		Periplasmic	2.4	37680	37680
		Cytoplasmic	17.7	37680* 39654	37679* 39654
DsbA-His-VHH		Cytoplasmic	9.8	37940	37941

*Molecular weight with the cleaved signal peptide, as when the periplasmic proteins enter the periplasm, the signal peptide is cleaved by signal peptidases.
** The difference of 18 Da could be attributed to Met oxidation.

3.2.1. Enzymatic digestion with TEV protease

The highest yield was obtained with Trx-His-VHH construct (Table 1.) so it was expressed once more with the purpose of performing enzymatic cleavage to remove the fusion tag and to obtain cleaved VHH. The cleavage reaction of the Trx-His-VHH construct was terminated after 48 hours by subjecting the reaction mixture to purification on the His-Trap column. The resulting chromatogram is depicted in Figure 8, where a distinct peak is observed in the flow-through (colored blue). Fractions corresponding to this peak were subsequently analyzed using SDS-PAGE to confirm the presence of the desired product, as shown in Figure 9.

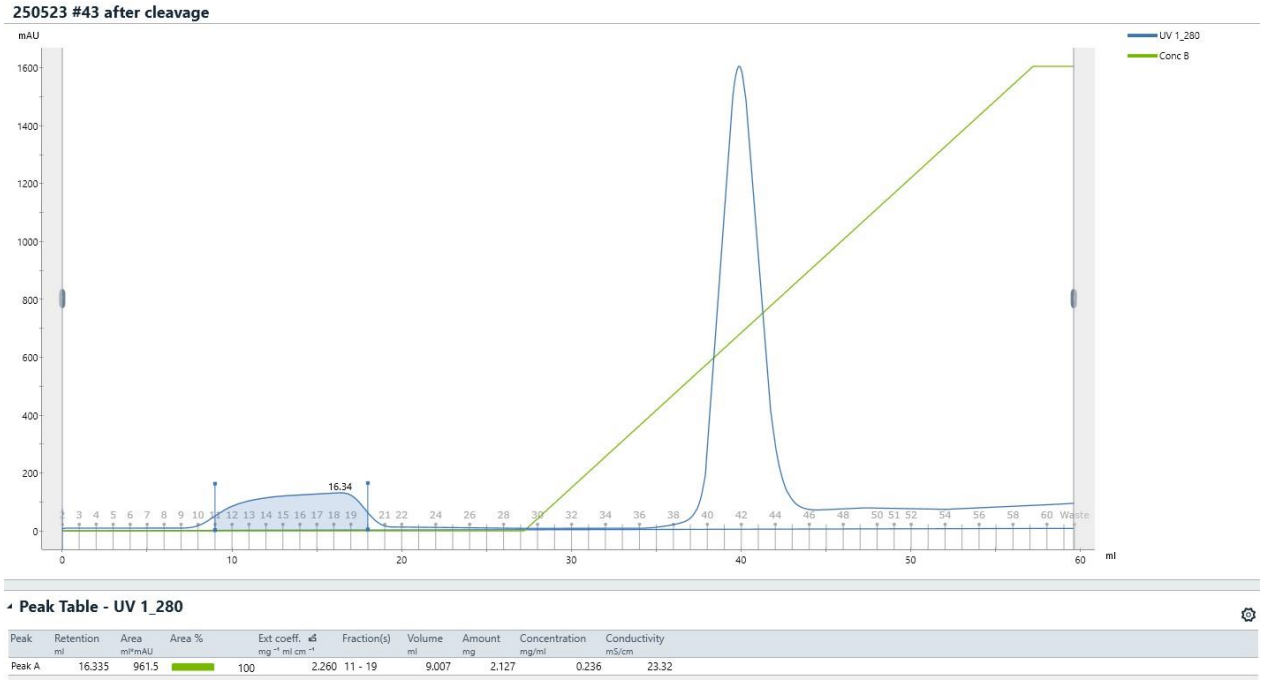

Figure 8. Chromatogram showing purification of the enzymatic cleavage reaction. The flow-through fractions (fractions 11-19) contain the VHH protein, which no longer possesses a His-tag. In contrast, the components retained on the column, including the uncleaved starting material (Trx-His-VHH), Trx-His, and His-tagged TEV protease, are eluted using a gradient of the elution buffer (buffer B).

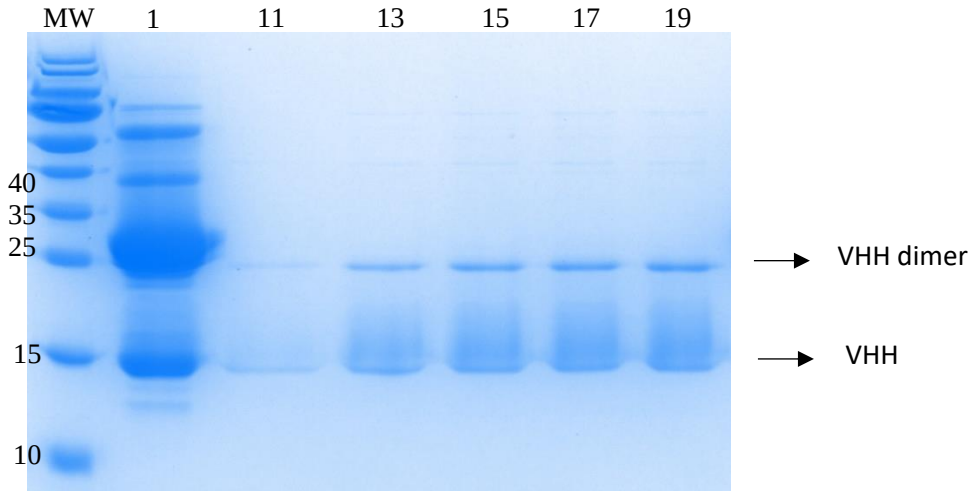

Figure 9. SDS-PAGE after purification of the enzymatic cleavage reaction. MW is the molecular weight marker. 1- Sample before purification; 11, 13, 15, 17, 19- fractions of the flow-through. Sample before purification contains: TEV protease, Trx-His-VHH (starting material, 28.30 kDa), products of the cleavage reaction: Trx-His (14.42 kDa) and VHH (13.88 kDa). In the fractions from the flow-through (not retained on the column), VHH (13.88 kDa) and its dimer (27.76 kDa) can be seen.

Correct molecular of the cleaved purified VHH weight was confirmed by LC-MS analysis (13887.8 Da, figure not shown).

Monitoring the cleavage reaction using SDS-PAGE poses certain challenges due to several reasons. Firstly, the difference in molecular weight of the products of the reaction (Trx-His and VHH) is too small for their effective separation on the gel. Secondly, the molecular weight of Trx-His-VHH and TEV protease is also similar, and in addition to that, the dimeric form of VHH has a similar molecular weight to those 2 proteins. So, a more insightful way of monitoring the reaction is by LC-MS (Figure shown in Annex).

3.3. Characterization of the VHH

3.3.1. Dimerization test

In each SDS-PAGE analysis conducted, the presence of a dimeric form of the purified protein was consistently observed. In the literature, Baral et al. (2012) characterized the dimeric structure of a human VHH by crystallography, whereas in the paper by Watanabe et al. (2019) a dimeric form of a VHH can be seen on SDS-PAGE.

As the dimer was not observed by LC-MS, to gain a deeper understanding of the nature of this dimer, a straightforward experiment was conducted using the His-VHH fusion protein.

The experimental results depicted in Figure 10 were obtained with 8 μg of the fusion protein (in sample buffer) in each well and they indicate that the band representing the monomeric state of the protein is most prominent when the sample is not subjected to any heating during the preparation for SDS-PAGE.

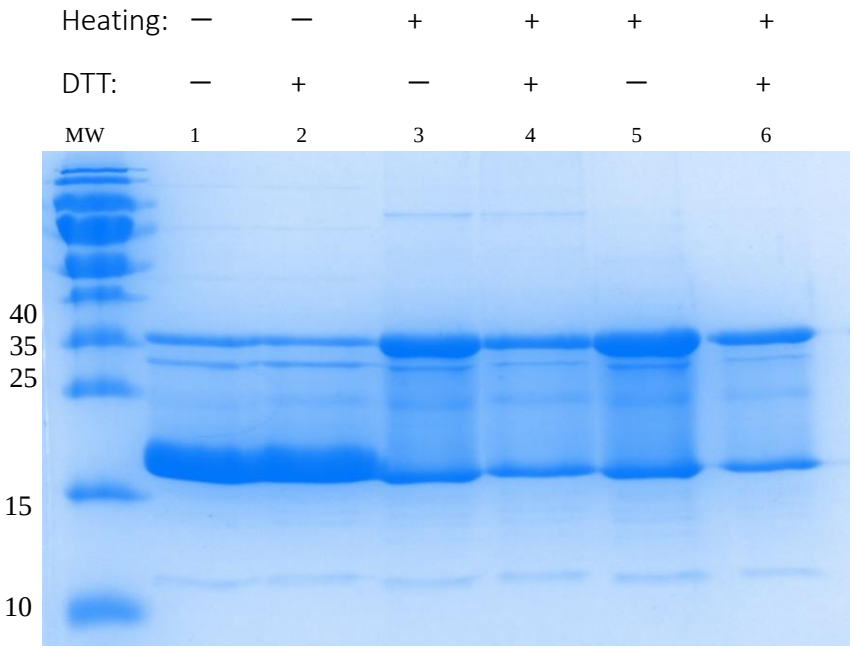

Figure 10. SDS-PAGE with the results of the dimerization experiment. MW is the molecular weight marker. Molecular weight of the tested fusion protein is 16.82 kDa, whereas the dimer's molecular weight is 33.64 kDa. The samples that were not heated (lanes 1 and 2) were kept at 4 °C for 30 minutes, samples from lanes 3 and 4 were kept at room temperature for 20 minutes after heating and samples from lanes 5 and 6 were kept at 4 °C for 20 minutes after heating. The heating of the samples lasted 10 minutes at 100 °C. The final concentration of DTT was 0.1 M.

3.3.2. Characterization of the disulfide bond with LC-HRMS

The LC-HRMS analysis of the purified VHH revealed an observed monoisotopic mass that was 2 Da larger (measured as 13880.7 Da) than the expected monoisotopic mass of the protein when it contains one disulfide bridge (18878.7 Da) (Figure 11). This significant difference in mass strongly suggests that the cysteine residues within the VHH are not disulfide linked.

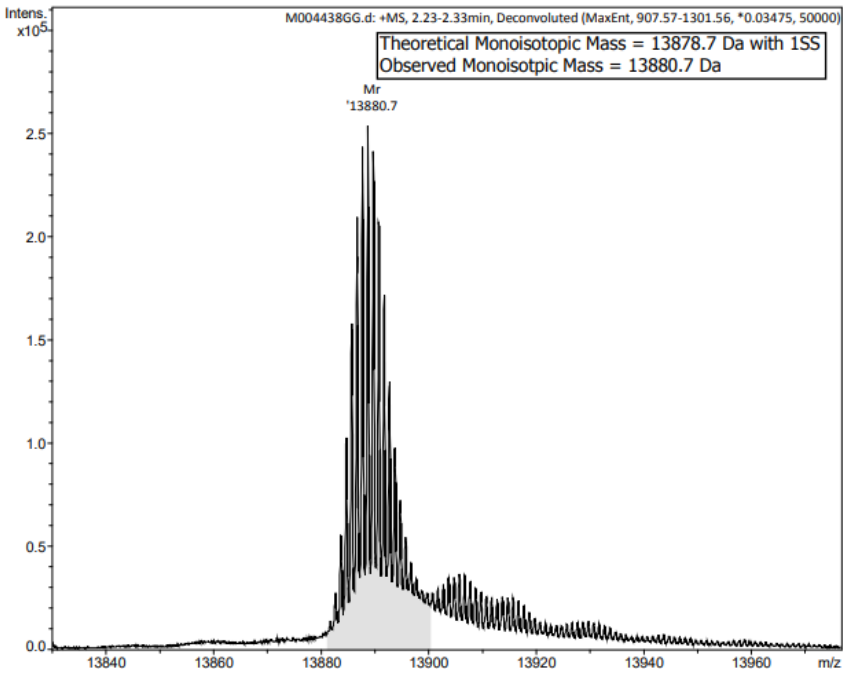

Figure 11. HRMS spectrum of the VHH.

3.3.3. Enzyme-linked immunosorbent assay (ELISA)

To assess the antigen-binding potential of the produced recombinant VHH, ELISA test was performed with the biotinylated peptide derived from the C-terminal part of the 5HT7 receptor as the antigen as described in the Material and methods section. The fitted ELISA curve is shown in Figure 12, and EC50 value was determined to be 8.3 μM.

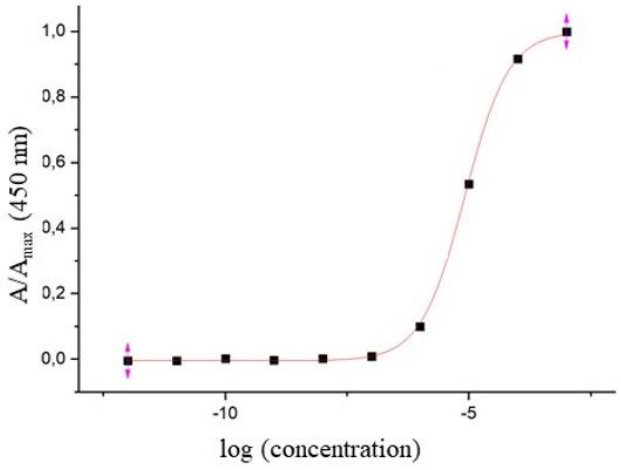

Figure 12. Fitted standard curve obtained from results after performing an ELISA test on purified VHH with biotinylated peptide derived from the C-terminal part of the 5HT7 receptor as the antigen.

3.3.4. Circular dichroism

Circular dichroism (CD) spectra analysis is a valuable tool for investigating the secondary structure of proteins. This method allows for the determination of important structural information, including the relative amounts of α-helix, β-sheet, and random coil structures present in a protein. CD spectra for VHH are shown in Figure 13.

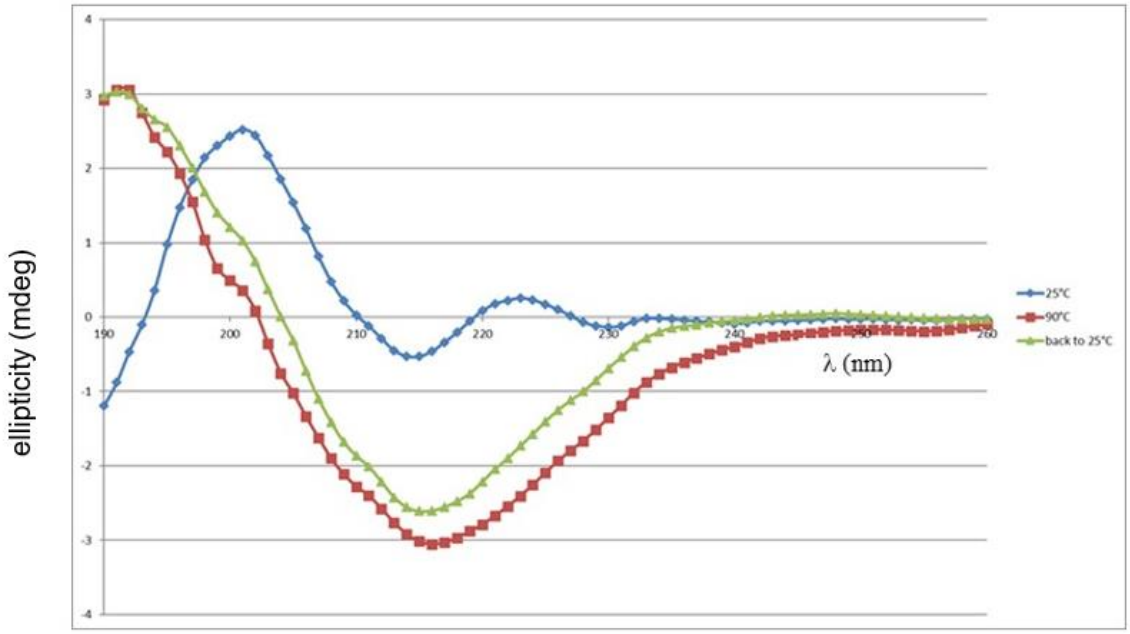

Figure 13. CD spectra of the purified VHH recorded at 25 ˚C, 90 ˚C (fully denatured protein) and spectra after cooling the protein to 25 ˚C.

Based on the data analysis by the BeStSel web server, the estimated secondary structure content for the VHH before denaturation showed that the VHH structure is predominantly made up of β-sheet, which, combined with the low content of α-helix, is expected.

Furthermore, CD spectroscopy can be utilized to determine the melting temperature (T_m) of a protein. The T_m is a critical parameter that provides insights into the folding and structural integrity of the protein and is determined by monitoring changes in the CD spectrum as a function of temperature. As can be seen in Figure 14, the T_m of VHH is 74.0 ˚C ± 0.3 ˚C.

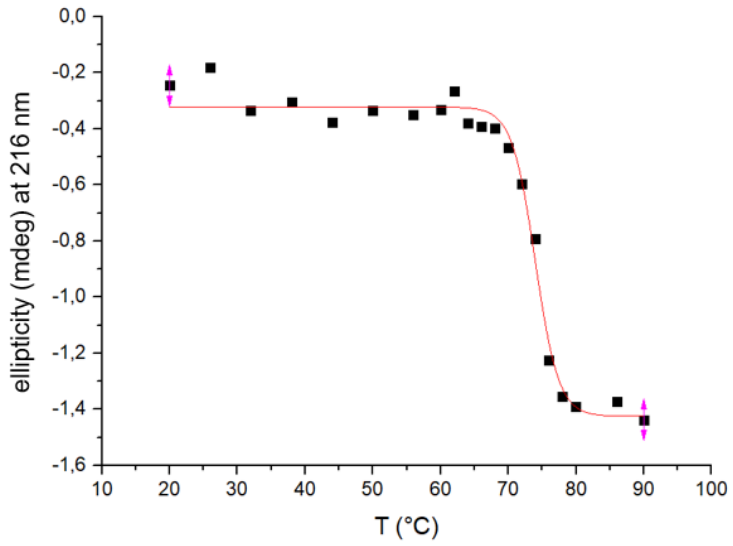

Figure 14. Fitted curve that corresponds to the measured ellipticity at λ = 216 nm versus temperature for the purified VHH.

4. DISCUSSION AND CONCLUSIONS

The goal of this master's thesis was to optimize the recombinant production method for the VHH targeting a peptide derived from the C-terminal part of the 5HT7 receptor. The focus was on producing a sufficient quantity of VHHs that specifically recognize and bind to the peptide.

Optimization of the recombinant nanobody production in *E. coli* included expressing 6 different fusion constructs (Figure 7) to determine which construct would yield the biggest quantity of the soluble protein. From Figure 7 it is visibly noticeable that for some fusion constructs, purification using immobilized metal-affinity chromatography (IMAC) is not enough to achieve a satisfactory purification and an additional purification step is needed, such as size exclusion chromatography (SEC). This occurs typically with samples that have a low concentration of the recombinant His-tagged protein, so other proteins that contain two or more consecutive histidines bind non-specifically to the His-Trap column. A way to circumvent that issue is to use a His-Trap column of a smaller volume (1 mL compared to 5 mL that was used in this protocol), as when less of the resin is used, the tagged protein will occupy most of the available binding sites, reducing the number of sites that other proteins could non-specifically bind to (Bornhorst and Falke, 2000).

From the preliminary experiments carried out with the Trx-His-VHH construct, it was shown that the obtained yield of the fusion protein is much higher in autoinduction media (84 mg/L) than in LB media (2.5 mg/L) (Table 1.). Autoinduction is generally considered as a more convenient

method than culturing bacteria in LB medium with IPTG induction, as the autoinduction media is simply inoculated with the expression strain which is then allowed to grow until saturation without the need for monitoring the optical density of the culture. In addition to that, with autoinduction media generally significantly higher cell culture densities and concentrations of protein per culture volume can be obtained (Studier, 2005).

In their high-throughput screening of fusion partners for soluble expression of small disulfide-rich proteins, Nozach et al (2013), identified disulfide isomerase DsbC as the most potent partner with DsbA being the second most potent partner. Results presented in this report also indicate that when the VHH was expressed with DsbC as a fusion partner, the yield of the protein was significantly improved as compared to the yield of fusion DsbA-His-VHH (60 mg/L for DsbC-His-VHH compared to 9.8 mg/L for DsbA-His-VHH, Table 1). However, results presented indicate that Trx is the most efficient partner for expression of VHH protein, in terms of acquired yield of the fusion protein (120 mg/L, Table 1, Figure 5 and 6).

Fusion protein yields obtained from periplasmic expression and extraction seem to be quite lower than those obtained from cytoplasmic extraction (Table 1), a phenomenon that has been described beforehand in literature. In order to express a high yield of recombinant protein, the gene that encodes for that protein must be expressed at the highest level possible which leads to the saturation of the Sec translocon, resulting in increased accumulation of the precursor of periplasmic protein in the cytoplasm, instead of them being exported into the periplasm (Baumgarten et al., 2018). Another limiting factor contributing to the lower yields of recombinant proteins in the periplasm is the size of the periplasm compartment, which accounts for 20 % of the total cell volume (Horga et al., 2018).

It is worth noting that there was no soluble Trx-His-VHH recombinant fusion protein detected when it was expressed in Origami™ strain of *E. Coli* (Table 1), even though the strain should have a more oxidizing cytoplasm which facilitates production of disulfide bonded proteins. Similar findings have been published by López-Cano et al. (2022). They found lower production yields of recombinant disulfide-bonded defensins coupled with reduced activity of the proteins in Origami™ strain compared to BL21 (DE3) strain. Lower production yields of soluble disulfide rich small proteins with Origami™ strain compared to BL21 (DE3) strain were also found by Nozach et al. (2013), even though it is expected that redox-active fusion partners such as thioredoxin or DsbC promote disulfide bond formation in the cytoplasm of redox engineered Origami™ strain.

Tobacco etch virus (TEV) protease is an enzyme employed to remove fusion tags from recombinant proteins as it targets a very specific sequence and it is easily produced (Sequeira et al., 2017; Kurussi et al., 2017). Even though the classical recognition site for the enzyme is ENLYFQ-G/S, the glycine or serine at the P1' position can be replaced by almost any other amino acid (except proline), so a desired amino acid can be inserted this way into the N-terminus of the recombinant protein (Renicke et al., 2013). The cleavage reaction is typically performed overnight (Kurussi et al., 2017), however our experimental findings (Figure 15) indicated that after 24 hours the reaction is still ongoing, so it was decided to go through with the reaction for 48 hours to maximize the obtained quantity of VHH. The optimal temperature for the reaction is 34 °C (Kurussi et al., 2017), but since the production described in this master's thesis is novel, the stability of the protein at 34 °C was unknown. That is the reason the reaction was performed at lower temperature (25 °C) and approximately 2 mg of cleaved VHH was obtained (Figure 8 and 9).

The results from LC-HRMS analysis provided concrete evidence that the VHH is not disulfide linked (Figure 11). Since the VHH was obtained from the Trx-His-VHH fusion protein, this result can be linked to the finding of Jurado et al. (2006). The findings of the study indicated that the fusion partner Trx (thioredoxin) does not serve as a catalyst for disulfide bond formation when used as a fusion partner with the single-chain variable fragment (scFv). Instead, Trx appears to have a chaperone effect, facilitating the correct folding of the scFv and increasing its solubility. It is possible that a similar scenario applies to the VHH protein being investigated, wherein Trx exhibits a robust chaperone activity that enhances the yield of soluble fusion protein. However, Trx does not play a direct role in promoting disulfide bond formation within the protein structure of the VHH.

When the cleaved VHH was acquired, characterization tests could proceed.
To inspect the secondary structure, thermal stability and renaturation efficiency of the produced VHH, CD experiments were carried out. Firstly, to assess the secondary structure, CD spectra were recorded at 25 °C and in Figure 13 (blue curve) a peak maximum can be seen at approximately 201 nm and a peak minimum at approximately 216 nm which is indicative of β-sheet secondary structure. Estimated secondary structure content (25 °C) correlates well with CD spectra performed for anti-Coronavirus VHHs (JASCO Global website, 2023). The protein was then subjected to heating to 90 °C and then cooled to 25 °C. CD spectra that were recorded after the protein has been denaturated and cooled to 25 °C (Figure 13) resemble more the denatured spectra which indicates

inefficient renaturation. Kunz et al. (2018) suggested that irreversible denaturation of nanobodies occurs due to aggregation. This fact could be supported by evidence presented in Figure 10, where it is demonstrated that heating the VHH causes it to dimerize more than if it was not heated to 100 °C. It is likely that the dimerization takes place through non-covalent hydrophobic interaction of the exposed hydrophobic core of the denatured protein. Additional evidence supporting the fact that inefficient renaturation is due to dimerization was also presented by Kunz et al. (2018). They found reduced aggregation in nanobodies with that have an additional non-canonical disulfide bond caused by the role of disulfide bonds in improving kinetic stability and prevention of aberrant association. Nanobody described in this thesis is expected to have only one disulfide bond, but is actually not disulfide bonded, and is thus lacking the mentioned protective effect which could be the determining reason for the observed dimer formation.

From the CD spectra analysis, it was determined that the T_m of VHH is 74.0 ˚C ± 0.3 ˚C (Figure 14) which is in the higher end of the range of what is expected for nanobodies (JASCO Global website, 2023), indicating reasonable thermal stability, in spite of the absence of the disulfide bond. ELISA test was another important assay performed in order to characterize the affinity of the produced nanobody for its antigen (biotinylated peptide derived from the C-terminal part of the 5HT7 receptor). The EC50 value that was obtained from the fitted standard ELISA curve (Figure 12) was 8.3 μM, which is significantly higher than the expected nM range (similar nanobody that was produced by the « Cellular microenvironment and receptor pharmacology » group at CBM had EC50 value of 10 nM, unpublished results), implying subpar affinity of the nanobody. The reasoning behind this low affinity of the VHH is still unclear. The produced VHH does not contain a disulfide bond, but that should not impair the antigen-binding affinity as the disulfide bind is not strictly required for the antigen binding ability (Liu et al., 2019). The conducted ELISA employed a method where the VHH was bound to the surface of a microtiter plate, followed by the addition of varying concentrations of the biotinylated peptide to evaluate the binding affinity of the VHH. However, it is worth noting that Butler et al. (1992) reported a potential issue with this approach. According to their findings, passive adsorption of antibodies onto polystyrene surfaces can lead to molecular alterations or denaturation, resulting in a loss of binding activity. It can be concluded that repeating the experiment would have been beneficial. However, due to the time constraints and the fact that the experiment was conducted towards the end of the internship, it may not have been feasible to repeat the experiment within the available timeframe.

5. FUTURE PERSPECTIVES

To optimize the cleavage of VHH, it is recommended to conduct the cleavage reaction at a higher temperature (closer to the optimal temperature for TEV protease activity). Additionally, longer incubation period should also be considered.

It is crucial to obtain a disulfide bond within the VHH structure because of its role in protein stability. Ban et al. (2020) proposed a method for obtaining a disulfide bond in the VHH firstly by denaturing the protein with either guanidinium chloride or urea, in the presence of reducing agents (DTT or 2-mercaptoethanol). The renaturation could be done by gradually removing the denaturing agents by e.g., dialysis. However, this approach does not always yield properly formed disulfide bonds, so adding a mixture of oxidized and reduced thiol reagents such as glutathione helps with the recovery of correct disulfide bonds in the protein structure.

Further ELISA assays should be conducted, wherein the antigen is adsorbed on the plate to prevent any conformational changes to the nanobody's paratope.

6. REFERENCES

Arbabi-Ghahroudi, M. (2017). Camelid Single-Domain Antibodies: Historical Perspective and Future Outlook. *Frontiers in Immunology*, *8*(NOV). https://doi.org/10.3389/FIMMU.2017.01589

Asaadi, Y., Jouneghani, F. F., Janani, S., & Rahbarizadeh, F. (2021). A comprehensive comparison between camelid nanobodies and single chain variable fragments. *Biomarker Research 2021 9:1*, *9*(1), 1–20. https://doi.org/10.1186/S40364-021-00332-6

Ban, B., Sharma, M., & Shetty, J. (2020). Optimization of Methods for the Production and Refolding of Biologically Active Disulfide Bond-Rich Antibody Fragments in Microbial Hosts. *Antibodies*, *9*(3), 1–18. https://doi.org/10.3390/ANTIB9030039

Bannas, P., Hambach, J., & Koch-Nolte, F. (2017). Nanobodies and Nanobody-Based Human Heavy Chain Antibodies As Antitumor Therapeutics. *Frontiers in Immunology*, *8*(NOV). https://doi.org/10.3389/FIMMU.2017.01603

Baral, T. N., Chao, S. Y., Li, S., Tanha, J., Arbabi-Ghahroudi, M., Zhang, J., & Wang, S. (2012). Crystal Structure of a Human Single Domain Antibody Dimer Formed through VH-VH Non-Covalent Interactions. *PLOS ONE*, *7*(1), e30149. https://doi.org/10.1371/JOURNAL.PONE.0030149

Baumgarten, T., Ytterberg, A. J., Zubarev, R. A., & de Gier, J. W. (2018). Optimizing Recombinant Protein Production in the Escherichia coli Periplasm Alleviates Stress. *Applied and Environmental Microbiology*, *84*(12). https://doi.org/10.1128/AEM.00270-18

Beghein, E., & Gettemans, J. (2017). Nanobody technology: A versatile toolkit for microscopic imaging, protein-protein interaction analysis, and protein function exploration. *Frontiers in Immunology*, *8*(JUL), 276923. https://doi.org/10.3389/FIMMU.2017.00771/BIBTEX

Berkmen, M. (2012). Production of disulfide-bonded proteins in Escherichia coli. *Protein Expression and Purification*, *82*(1), 240–251. https://doi.org/10.1016/J.PEP.2011.10.009

BeStSel - Protein Circular Dichroism Spectra Analysis. (n.d.). Retrieved June 02, 2023, from https://bestsel.elte.hu/index.php

Bhatwa, A., Wang, W., Hassan, Y. I., Abraham, N., Li, X. Z., & Zhou, T. (2021). Challenges Associated With the Formation of Recombinant Protein Inclusion Bodies in Escherichia coli and Strategies to Address Them for Industrial Applications. *Frontiers in Bioengineering and Biotechnology*, *9*, 630551. https://doi.org/10.3389/FBIOE.2021.630551/BIBTEX

BL21(DE3) Competent Cells. (n.d.). Retrieved June 07, 2023, from https://www.thermofisher.com/order/catalog/product/EC0114

Bornhorst, J. A., & Falke, J. J. (2000). [16] Purification of Proteins Using Polyhistidine Affinity Tags. *Methods in Enzymology*, *326*, 245. https://doi.org/10.1016/S0076-6879(00)26058-8

Butler, J. E., Ni, L., Nessler, R., Joshi, K. S., Suter, M., Rosenberg, B., Chang, J., Brown, W. R., & Cantarero, L. A. (1992). The physical and functional behavior of capture antibodies adsorbed on polystyrene. *Journal of Immunological Methods*, *150*(1–2), 77–90. https://doi.org/10.1016/0022-1759(92)90066-3

Costa, S., Almeida, A., Castro, A., & Domingues, L. (2014). Fusion tags for protein solubility, purification and immunogenicity in Escherichia coli: the novel Fh8 system. *Frontiers in Microbiology*, *5*(FEB). https://doi.org/10.3389/FMICB.2014.00063

Crane, J. M., & Randall, L. L. (2017). The Sec System: Protein Export in Escherichia coli. *EcoSal Plus*, *7*(2). https://doi.org/10.1128/ECOSALPLUS.ESP-0002-2017

Depuydt, M., Messens, J., & Collet, J. F. (2011). How proteins form disulfide bonds. *Antioxidants & Redox Signaling*, *15*(1), 49–66. https://doi.org/10.1089/ARS.2010.3575

Horga, L. G., Halliwell, S., Castiñeiras, T. S., Wyre, C., Matos, C. F. R. O., Yovcheva, D. S., Kent, R., Morra, R., Williams, S. G., Smith, D. C., & Dixon, N. (2018). Tuning recombinant protein expression to match secretion capacity. *Microbial Cell Factories*, *17*(1), 1–18. https://doi.org/10.1186/S12934-018-1047-Z/FIGURES/6

Jovčevska, I., & Muyldermans, S. (2020). The Therapeutic Potential of Nanobodies. *BioDrugs : Clinical Immunotherapeutics, Biopharmaceuticals and Gene Therapy*, *34*(1), 11–26. https://doi.org/10.1007/S40259-019-00392-Z

Jurado, P., De Lorenzo, V., & Fernández, L. A. (2006). Thioredoxin fusions increase folding of single chain Fv antibodies in the cytoplasm of Escherichia coli: evidence that chaperone activity is the prime effect of thioredoxin. *Journal of Molecular Biology*, *357*(1), 49–61. https://doi.org/10.1016/J.JMB.2005.12.058

Kim, S., & Lee, S. B. (2008). Soluble expression of archaeal proteins in Escherichia coli by using fusion-partners. *Protein Expression and Purification*, *62*(1), 116–119. https://doi.org/10.1016/J.PEP.2008.06.015

Kleiner-Grote, G. R. M., Risse, J. M., & Friehs, K. (2018). Secretion of recombinant proteins from E. coli. *Engineering in Life Sciences*, *18*(8), 532. https://doi.org/10.1002/ELSC.201700200

Kolkman, J. A., & Law, D. A. (2010). Nanobodies – from llamas to therapeutic proteins. *Drug Discovery Today: Technologies*, *7*(2), e139–e146. https://doi.org/10.1016/J.DDTEC.2010.03.002

Könning, D., Zielonka, S., Grzeschik, J., Empting, M., Valldorf, B., Krah, S., Schröter, C., Sellmann, C., Hock, B., & Kolmar, H. (2017). Camelid and shark single domain antibodies: structural features and therapeutic potential. *Current Opinion in Structural Biology*, *45*, 10–16. https://doi.org/10.1016/J.SBI.2016.10.019

Kunz, P., Zinner, K., Mücke, N., Bartoschik, T., Muyldermans, S., & Hoheisel, J. D. (2018). The structural basis of nanobody unfolding reversibility and thermoresistance. *Scientific Reports 2018 8:1*, *8*(1), 1–10. https://doi.org/10.1038/s41598-018-26338-z

Landeta, C., Boyd, D., & Beckwith, J. (2018). Disulfide bond formation in prokaryotes. *Nature Microbiology*, *3*(3), 270–280. https://doi.org/10.1038/S41564-017-0106-2

Liu, H., Chumsae, C., Gaza-Bulseco, G., Hurkmans, K., & Radziejewski, C. H. (2010). Ranking the susceptibility of disulfide bonds in human IgG1 antibodies by reduction, differential alkylation, and LC-MS analysis. *Analytical Chemistry*, *82*(12), 5219–5226. https://doi.org/10.1021/AC100575N/SUPPL_FILE/AC100575N_SI_002.PDF

Liu, H., Schittny, V., & Nash, M. A. (2019). Removal of a Conserved Disulfide Bond Does Not Compromise Mechanical Stability of a VHH Antibody Complex. *Nano Letters*, *19*(8), 5524–5529. https://doi.org/10.1021/ACS.NANOLETT.9B02062/SUPPL_FILE/NL9B02062_SI_001.PDF

López-Cano, A., Martínez-Miguel, M., Guasch, J., Ratera, I., Arís, A., & Garcia-Fruitós, E. (2022). Exploring the impact of the recombinant Escherichia coli strain on defensins antimicrobial activity: BL21 versus Origami strain. *Microbial Cell Factories*, *21*(1), 1–10. https://doi.org/10.1186/S12934-022-01803-7/FIGURES/5

Mitchell, L. S., & Colwell, L. J. (2018). Comparative analysis of nanobody sequence and structure data. *Proteins*, *86*(7), 697–706. https://doi.org/10.1002/PROT.25497

Muyldermans, S. (2021). A guide to: generation and design of nanobodies. *The FEBS Journal*, *288*(7), 2084–2102. https://doi.org/10.1111/FEBS.15515

Nozach, H., Fruchart-Gaillard, C., Fenaille, F., Beau, F., Ramos, O. H. P., Douzi, B., Saez, N. J., Moutiez, M., Servent, D., Gondry, M., Thaï, R., Cuniasse, P., Vincentelli, R., & Dive, V. (2013). High throughput screening identifies disulfide isomerase DsbC as a very efficient partner for recombinant expression of small disulfide-rich proteins in E. coli. *Microbial Cell Factories*, *12*(1), 1–16. https://doi.org/10.1186/1475-2859-12-37/TABLES/1

Origami B(DE3) Competent Cells - Novagen Origami B host strains carry the same mutations as the original Origami strain, except that they are derived from a lacZY mutant of BL21 to enable precise control of expression levels using IPTG. | Sigma-Aldrich. (n.d.). Retrieved June 07, 2023, from https://www.sigmaaldrich.com/HR/en/product/mm/70837

Raran-Kurussi, S., Cherry, S., Zhang, D., & Waugh, D. S. (2017). Removal of Affinity Tags with TEV Protease. *Methods in Molecular Biology (Clifton, N.J.)*, *1586*, 221. https://doi.org/10.1007/978-1-4939-6887-9_14

Renicke, C., Spadaccini, R., & Taxis, C. (2013). A Tobacco Etch Virus Protease with Increased Substrate Tolerance at the P1' position. *PLoS ONE*, *8*(6). https://doi.org/10.1371/JOURNAL.PONE.0067915

Rosano, G. L., & Ceccarelli, E. A. (2014). Recombinant protein expression in Escherichia coli: Advances and challenges. *Frontiers in Microbiology*, *5*(APR), 79503. https://doi.org/10.3389/FMICB.2014.00172/BIBTEX

Schimek, C., Egger, E., Tauer, C., Striedner, G., Brocard, C., Cserjan-Puschmann, M., & Hahn, R. (2020). Extraction of recombinant periplasmic proteins under industrially relevant process conditions: Selectivity and yield strongly depend on protein titer and methodology. *Biotechnology Progress*, *36*(5), e2999. https://doi.org/10.1002/BTPR.2999

Sequeira, A. F., Turchetto, J., Saez, N. J., Peysson, F., Ramond, L., Duhoo, Y., Blémont, M., Fernandes, V. O., Gama, L. T., Ferreira, L. M. A., Guerreiro, C. I. P. I., Gilles, N., Darbon, H., Fontes, C. M. G. A., & Vincentelli, R. (2017). Gene design, fusion technology and TEV cleavage conditions influence the purification of oxidized disulphide-rich venom peptides in Escherichia coli. *Microbial Cell Factories*, *16*(1), 1–16. https://doi.org/10.1186/S12934-016-0618-0/FIGURES/7

Singhvi, P., Saneja, A., Srichandan, S., & Panda, A. K. (2020). Bacterial Inclusion Bodies: A Treasure Trove of Bioactive Proteins. *Trends in Biotechnology*, *38*(5), 474–486. https://doi.org/10.1016/J.TIBTECH.2019.12.011

Stability evaluation of anti-Coronavirus VHH antibody using circular dichroism spectrometer | JASCO Global. (n.d.). Retrieved June 19, 2023, from https://www.jasco-global.com/solutions/stability-evaluation-of-anti-coronavirus-vhh-antibody-using-circular-dichroism-spectrometer/

Studier, F. W. (2005). Protein production by auto-induction in high density shaking cultures. *Protein Expression and Purification*, *41*(1), 207–234. https://doi.org/10.1016/J.PEP.2005.01.016

Wang, W., Yuan, J., & Jiang, C. (2021). Applications of nanobodies in plant science and biotechnology. *Plant Molecular Biology*, *105*(1–2), 43. https://doi.org/10.1007/S11103-020-01082-Z

Watanabe, H., Itagaki, F., Shimizu, Y., Iikuni, S., & Ono, M. (2019). Synthesis and evaluation of a radioiodinated BODIPY derivative as a thiol-labeling agent. *Journal of Labelled Compounds and Radiopharmaceuticals*, *62*(13), 885–891. https://doi.org/10.1002/JLCR.3809

7. ANNEXES

7.1. Protein sequence of the used fusion constructs.

Table 2. Amino acid sequence of the used constructs
Enterokinase cleavage site, 6xHis tag, VHH, TEV cleavage site, DsbC sequence, **Dsbc Signal Peptide**, DsbA sequence, DsbA Signal Peptide, **Thioredoxin sequence**

Construct	Protein sequence of the fusion protein
His-VHH	MAHHHHHHVGTGSNDDDDKSPDPNWELAEVQLQESGGGLVQAGGSLRLSCVASGRTSSISYMAWFRQAPGKEREFVAVITWSGGTTYYADSVKGRFTISRDNAKNTVYLQMNSLKAEDTAVYYCAASRKYYSGSYLPESLFDYWGQGTQVTVSS
Trx-His-VHH	**MSDKIIHLTDDSFDTDVLKADGAILVDFWAEWCGPCKMIAPILDEIADEYQGKLTVAKLNIDQNPGTAPKYG IRGIPTLLLFKNGEVAATKVGALSKGQLKEFLDANLA**GSGSGHMHHHHHHGNGGSASENLYFQAEVQLQESGGGLVQAGGSLRLSCVASGRTSSISYMAWFRQAPGKEREFVAVITWSGGTTYYADSVKGRFTISRDNAKNTVYLQMNSLKAEDTAVYYCAASRKYYSGSYLPESLFDYWGQGTQVTVSS
DsbC-His-VHH	MKDDAAIQQTLAKMGIKSSDIQPAPVAGMKTVLTNSGVLYITDDGKHIIQGPMYDVSGTAPVNVTNKMLLKQLNALEKEMIVYKAPQEKHVITVFTDITCGYCHKLHEQMADYNALGITVRYLAFPRQGLDSDAEKEMKAIWCAKDKNKAFDDVMAGKSVAPASCDVDIADHYALGVQLGVSGTPAVVLSNGTLVPGYQPPKEMKEFLDEHQKMTSGKGSTSGSGHHHHHHGNGGSASENLYFQAEVQLQESGGGLVQAGGSLRLSCVASGRTSSISYMAWFRQAPGKEREFVAVITWSGGTTYYADSVKGRFTISRDNAKNTVYLQMNSLKAEDTAVYYCAASRKYYSGSYLPESLFDYWGQGTQVTVSS
PS-DsbC-VHH	**MKKGFMLFTLLAAFSGFAQA**DDAAIQQTLAKMGIKSSDIQPAPVAGMKTVLTNSGVLYITDDGKHIIQGPMYDVSGTAPVNVTNKMLLKQLNALEKEMIVYKAPQEKHVITVFTDITCGYCHKLHEQMADYNALGITVRYLAFPRQGLDSDAEKEMKAIWCAKDKNKAFDDVMAGKSVAPASCDVDIADHYALGVQLGVSGTPAVVLSNGTLVPGYQPPKEMKEFLDEHQKMTSGKGSTSGSGHHHHHHGNGGSASENLYFQAEVQLQESGGGLVQAGGSLRLSCVASGRTSSISYMAWFRQAPGKEREFVAVITWSGGTTYYADSVKGRFTISRDNAKNTVYLQMNSLKAEDTAVYYCAASRKYYSGSYLPESLFDYWGQGTQVTVSS
PS-DsbA-His-VHH	**MKKIWLALAGLVLAFSASA**AQYEDGKQYTTLEKPVAGAPQVLEFFSFFCPHCYQFEEVLHISDNVKKKLPEGVKMTKYHVNFMGGDLGKDLTQAWAVAMALGVEDKVTVPLFEGVQKTQTIRSASDIRDVFINAGIKGEEYDAAWNSFVVKSLVAQQEKAAADVQLRGVPAMFVNGKYQLNPQGMDTSNMDVFVQQYADTVKYLSEKKGSTSGSGHHHHHHGNGGSASENLYFQAEVQLQESGGGLVQAGGSLRLSCVASGRTSSISYMAWFRQAPGKEREFVAVITWSGGTTYYADSVKGRFTISRDNAKNTVYLQMNSLKAEDTAVYYCAASRKYYSGSYLPESLFDYWGQGTQVTVSSS
DsbA-His-VHH	MKAQYEDGKQYTTLEKPVAGAPQVLEFFSFFCPHCYQFEEVLHISDNVKKKLPEGVKMTKYHVNFMGGDLGKDLTQAWAVAMALGVEDKVTVPLFEGVQKTQTIRSASDIRDVFINAGIKGEEYDAAWNSFVVKSLVAQQEKAAADVQLRGVPAMFVNGKYQLNPQGMDTSNMDVFVQQYADTVKYLSEKKGSTSGSGHHHHHHGNGGSASENLYFQAEVQLQESGGGLVQAGGSLRLSCVASGRTSSISYMAWFRQAPGKEREFVAVITWSGGTTYYADSVKGRFTISRDNAKNTVYLQMNSLKAEDTAVYYCAASRKYYSGSYLPESLFDYWGQGTQVTVSS

7.2. Monitoring of TEV cleavage reaction by LC-MS

The cleavage reaction depicted in Figure 15 was carried out at a temperature of 25 ˚C. TEV protease was added to the reaction mixture at a ratio of 1:20 to the mass of the starting material (Trx-His-VHH). After 7.5 hours, the major peak representing the starting material (Trx-His-VHH) is observed at approximately 7.2 minutes on the chromatogram (indicated with an arrow). Moreover, two additional peaks begin to emerge, which correspond to the reaction products, Trx-His and VHH. After 24 h, 2 peaks that represent the reaction products have become much more visible which indicates increase in their concentration and signifies the efficiency of the reaction. The peak at approximately 2.8 minutes corresponds to TEV protein (indicated with an arrow).

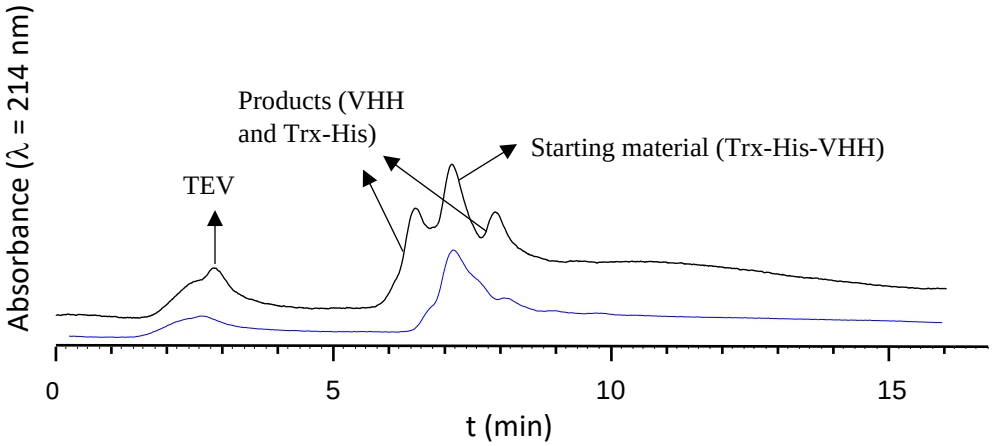

Figure 15. Overlapped LC-MS chromatograms obtained during the monitoring of reaction progress of cleavage by TEV protease. The blue curve corresponds to a sample taken after 7,5 h after the beginning of the reaction, while the black curve represents a sample taken 24 h after the beginning of the reaction.