universität wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Evaluation of the anti-GFP tandem nano body LaG-16-G4S-LaG2 for its applicability in protein purification"

verfasst von / submitted by

Georgi Nikolov, BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2018 / Vienna 2018

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	A 066 830
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Masterstudium Molekulare Mikrobiologie, Mikrobielle Ökologie und Immunbiologie/ Master's degree programme Molecular Microbiology, Microbial Ecology and Immunbiology
Betreut von / Supervisor:	Prof. Dr. Thomas Marlovits, Privatdoz.

1. INTRODUCTION

1.1. Protein purification systems- advances and obstacles

The development of techniques and methods for separating and purifying biological macromolecules has been an important prerequisite for many advancements in the fields of biochemistry and biotechnology over the past five decades. In the field of chromatography, the development of new porous resin supports, new crosslinked beaded agaroses and new bonded porous silicas has enabled rapid growth in high resolution techniques (high performance liquid chromatography, HPLC; fast protein liquid chromatography, FPLC). An expanded bed adsorption enables rapid isolation of target proteins directly from cell lysate or cell cultures.

A wide variety of chromatographic column packing materials, such as gel-filtration media, ion exchangers, reversed phase packings, hydrophobic interaction adsorbents and affinity chromatography adsorbents are commercially available today in various media diameters in order to satisfy the different requirements of efficiency, capacity and cost. [1]

In most cases, interest is focused on one particular biological activity and function of an enzyme. Therefore, traditional animal or microbial sources have been replaced by genetically engineered microorganisms or cultured eukaryotic cells. Protein products of eukaryotic origin, cloned and expressed in bacteria such as *Escherichia coli*, may either be located in the cytoplasm or secreted through the cell membrane. In the latter case, they are either collected in the periplasmic space or they are truly extracellular, secreted into the culture medium. At this stage, a considerable degree of purification has already been achieved by choosing a secreting strain. The recombinant protein can also be equipped with an affinity tag such as a His-tag or a fusion protein such as Protein A, glutathione-S-transferase, or maltose-binding protein in order to facilitate purification. [1]

However, many difficulties remain in finding optimum conditions for protein extraction and sample pretreatment, as well as in choosing suitable methods for monitoring protein concentration and biological activity. A suitable solution for improving the efficiency of purification is the technique presented in the following paragraphs- utilizing the LaG16-G4S-LaG2 tandem nanobody.

1.2. Nanobodies – comparison to conventional IgG, structural features, properties and applications

Immunoglobulin G (IgG) is the major serum isotype of antibodies and consists of four subtypes (IgG1,2,3 and 4) based on antigenic differences in the H chains and on the number and location of disulfide bonds. IgG1 makes up most (65%) of the total IgG It is the only isotype able to cross the placental barrier to provide immunity to the fetus. [2] It plays a number of other roles such as bacterial opsonization, complement fixation, bacterial killing enhancement and neutralization of bacterial toxins and viruses.

The general structure of IgG antibodies consists of two identical heavy (H)-chain and two identical light (L)-chain polypeptides. It is conserved and widely spread in mammals. (**Figure 1.1**)[3]

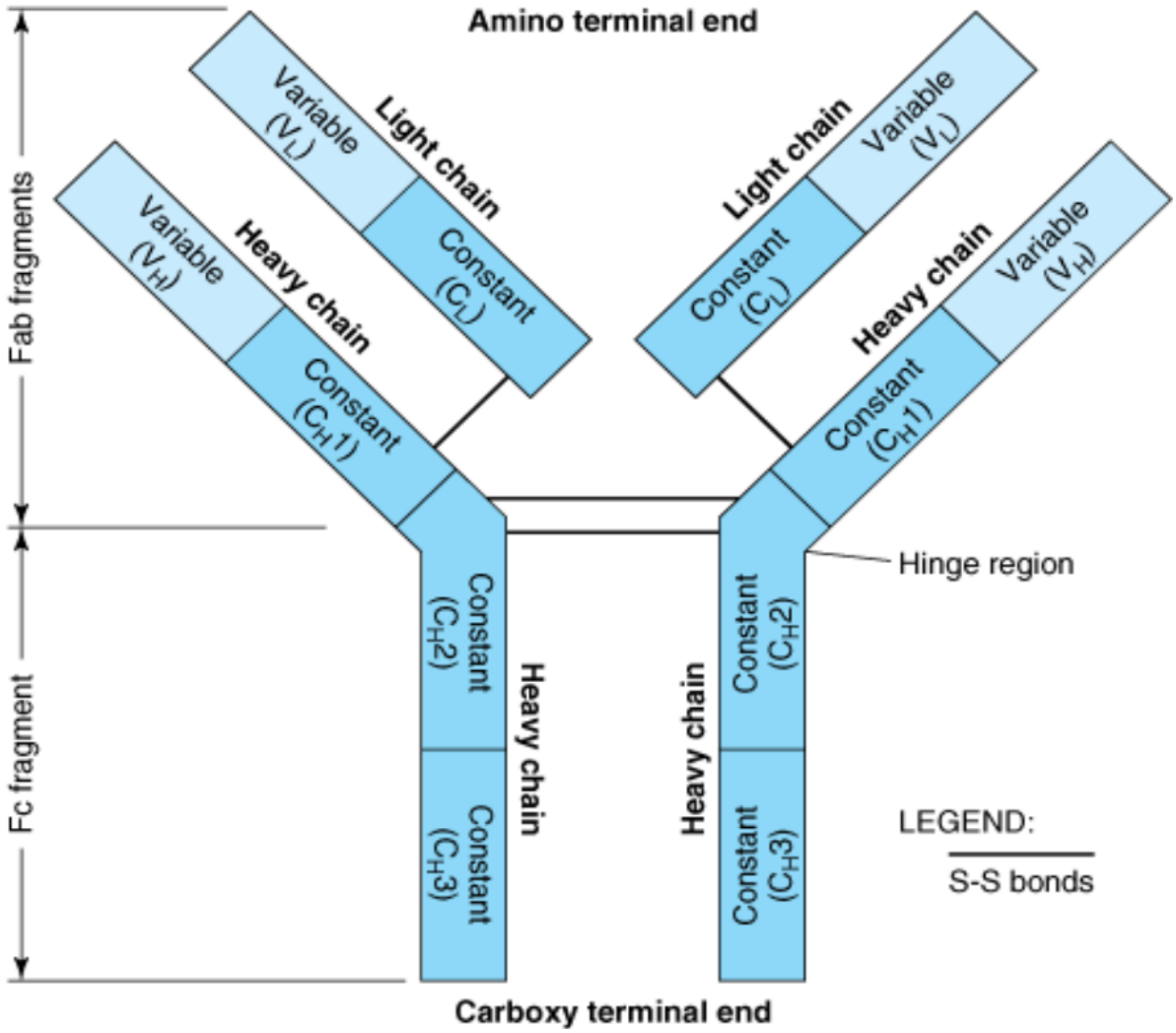

Figure 1.1: Schematic representation of naturally occurring antibody IgG in sera of camelids.

The Y-shaped IgG molecule consists of two light chains and two heavy chains. Each light chain contains a variable region and a constant region. Each heavy chain consists of a variable region and a constant region that is divided into three domains: C_H1, C_H2 and C_H3. The C_H2 domain contains the complement-binding site and the C_H3 domain is the site of attachment of IgG to receptors on neutrophils and macrophages. The antigen binding site is formed by the variable regions of both the light and heavy chains.
Picture adapted from Levinson W: Review of Medical Microbiology and Immunology, 10th Edition (2010)

The L-chain of these antibodies is comprised of two domains, while the H chain- consists of four. The N-terminal domains of the H and L chains vary between different antibodies (designated as VH and VL). These two domains constitute the variable fragment (Fv) that recognizes antigens. The remaining H and L sequences are much more conserved and are designated as CH and CL, respectively. The last two CH regions are important for recruitment and priming of immune cells (e.g. macrophages, natural killer cells) or for effector functions (e.g. complement activation). [4]

However, one exception to these structural rules can be found in the sera of Camelidae. In addition to producing the conventional heterotetrameric antibodies, they are also able to generate special IgG antibodies, known as heavy-chain antibodies (HCAbs). These antibodies are devoid of the L-chain polypeptide and lack the first constant domain (CH1) (**Figure 1.1**). Such characteristics as well as a smaller size (~90kDa rather than 150kDa for conventional antibodies) and their compact architecture allow an improved adaptation for accessing hidden targets. [5]

Nanobodies (Nbs) are the recombinant single variable domains derived from the camelid HCAbs. Unlike the monoclonal antibodies, they are easily produced in various bacterial expression systems. In general, they are cloned after a secretion signal in order to be produced in the periplasm. Their subsequent purification is also straightforward due to the fact that Nbs are usually cloned in frame with a His6 purification tag. Thus, after periplasmic extraction, the nanobodies may be easily purified using metal affinity chromatography. Nbs also have an excellent shelf life, allowing storage for months at 4°C and even longer at -20°C while maintaining full antigen-binding capacity.

Nanobodies possess several qualities which make them ideal binder molecules: they are a renewable source, allow for economic production, possess a small size, are stable and prove soluble in aqueous solutions. Furthermore, they allow reversible refolding and are specific with a high affinity for only one target antigen surface patch. These unique properties allow multiple types of applications for Nbs: as a research tool, in diagnostics and therapeutics. An example for this applicability is the study by Ries et al. (2012) which utilized Nbs with high

binding affinity to GFP as a delivery mechanism to transport bright organic fluorophores to GFP-tagged proteins for use in single-molecular nanoscopy.[6] In this example, the small size and high affinity to GFP allowed the combination of genetic tagging with the high photon yield of organic dyes and minimal linkage error. Another study by Koide et al. (2009) showed how Nbs could be used to assist the crystallization process and structural determination of flexible or aggregating proteins. Again, it was due to the small size, robust structure, high yield and targeting clefts on the surface of the antigen that allowed Nbs to tackle these crystallization challenges. [7]

The same qualities will allow the Nbs to have special advantages as therapeutic agents, as well.[8] In some cases, the Nbs will be able to access hidden and essential epitopes on pathogenic agents, such as viruses, bacteria or parasites. In addition, the added value of the Nbs is their capacity to discriminate the cognate target from closely related variants, allowing a high specificity unreachable by most small organic antagonists.[5] Such a case was investigated by a study by Koch-Nolte et al. (2007), which attempted to block the function of a specific leukocyte ecto-enzyme (ART2.2) *in vivo*. These characteristics make nanobodies a useful class of molecules for research and various medical diagnostic and therapeutic applications.

So far, nanobody-based affinity capture has been used to study protein-protein interactions[9], analyze *in vivo* DNA-protein interactions at the genomic scale after chromatin immunoprecipitations, track protein conformations, trace bacterial infections and purify recombinant proteins. Nanobodies have also been employed for scavenging toxins from plasma[10] and for depletion of abundant serum proteins, a step required for proteomic analysis of blood samples[11]. They also exhibit a good track record as chaperones to crystallize challenging proteins.[12–15]Indeed, these robust, single-domain antibody fragments often lock proteins in a particular conformation.[16] The ability of the nanobodies to stabilize intrinsically flexible regions or shield aggregating surfaces from contact with solvents are key features to allow the crystallization process.[7]

Biosensor development has also benefited from nanobody application due to its need for small, robust, highly specific affinity probes, where size and directional immobilization matter. Nanobodies are particularly suited for controlled and reproducible coupling because site-specific functional groups are easily introduced, allowing covalent and oriented binding with minimal loss of specificity and affinity.[17] Their small size provides high binding capacity

surfaces resulting in higher sensitivity and their high stability and refolding competence allows stringent washing and regeneration conditions.

This would be of particular interest for the study of membrane protein complexes. These are a type of proteins which are usually present at low levels in biological membranes and rarely are the major peptide constituent of the membrane. Some cases have managed to easily isolate large amounts of pure protein, for example, bacteriorhodopsin (bR)- the only protein present in the purple membrane of *Halobacteria salinaria*.[18] However, most membrane proteins cannot be readily obtained in sufficient amounts from their environments and thus alternative methods, such as overexpression, are needed. A major problem with overexpressing them in *E. coli* and other systems, is the aggregation of protein in the cytoplasm.[19] Therefore, high yields of functional and stable protein are rarely obtained.[18] Thus, this is a field which could greatly benefit from a novel method of purifying complex targets.

These structural qualities and practical applications make nanobodies a prime candidate for many future studies and experiments which would require a small, highly flexible and specific binding molecule.

1.3. The LaG-16-G4S-LaG-2 tandem nanobody - production, qualities and application

Owing to the combination of their unique qualities, nanobodies are a highly demanded and indispensable research, diagnostic and therapeutic tool. However, rapid and robust techniques for isolating repertoires of nanobodies have proven elusive, leading to demand highly exceeding supply. In a study by Fridy et al. (2014), a highly-optimized pipeline was invented to allow the rapid production of high-affinity nanobodies against selected proteins (**Figure 1.2.1).** Their approach was based on high-throughput DNA sequencing of a marrow lymphocyte VHH cDNA library from an immunized llama combined with mass spectrometric (MS) identification of high-affinity VHH regions derived from the serum of the same animal.

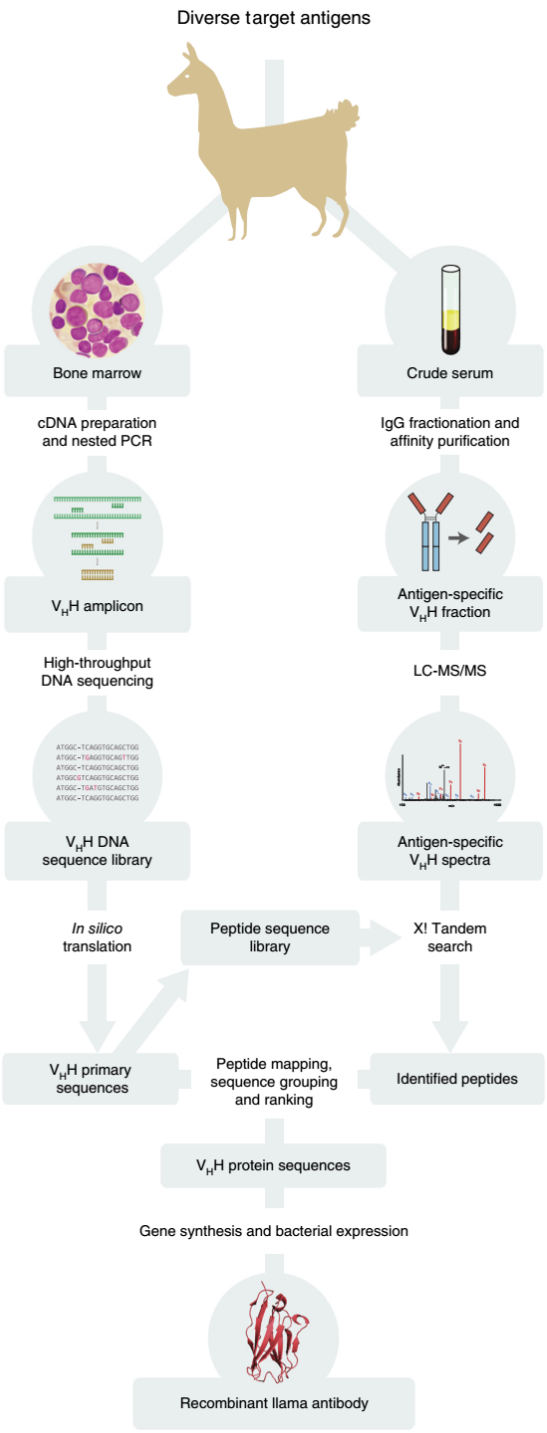

Figure 1.2.1: Overview of nanobody identification and production pipeline.
Example nanobody structure was obtained from PDB.
Picture adapted from Fridy et. al, (2014)

In order to generate nanobodies of maximal utility, the study chose GFP and mCherry tags as first target antigens due to their central role in biological studies. Upon immunizing individual

llama with the antigens and confirmation of an immune response, serum was collected and affinity-purified in order to obtain VHH-containing fractions.[20] These were then further purified and analyzed by liquid chromatography-MS and tandem MS. (**Figure 1.2.2**) A bioinformatic pipeline was later developed in order to identify the highest-probability matches from a pool of related VHH sequences, thus overcoming challenges from the lack of a dedicated database for VHH cDNA and the difficult distinguishing of specific VHH sequences via MS-based protein identification. This automated ranking pipeline, alongside manual inspection, allowed for the identification of 44 high-probability hits against GFP and their classification as LaG 1-44.[21-24] (**Figure 1.2.3**)

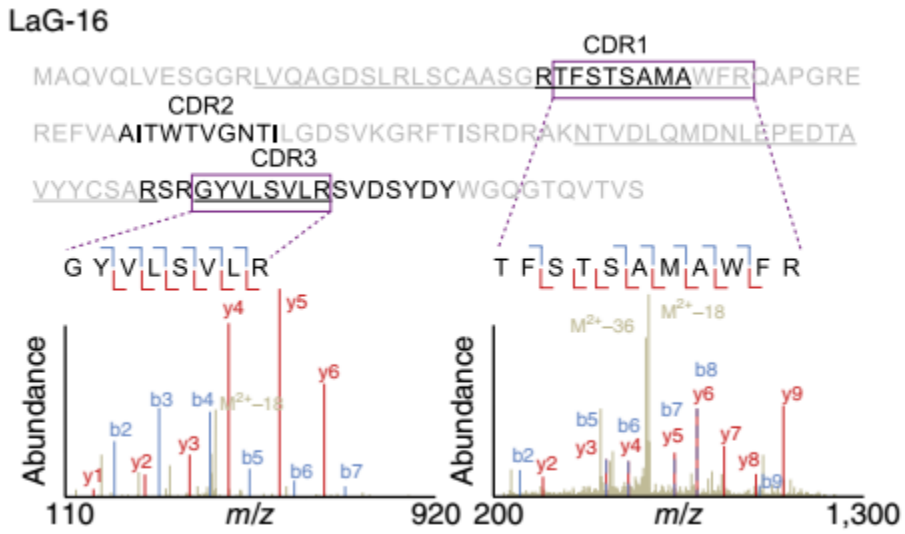

Figure 1.2.2: Characterization of the LaG16 nanobody via tandem MS.
Tandem mass spectra of identified peptides (boxed), mapped to the CDR regions of the LaG16 sequence. The MS-covered regions are underlined. Dashed lines indicate overlapping peaks.
Picture adapted from Fridy et. al, (2014)

Clone ID	MW (Da)	K_d (nM)
LaG-2	15,919	19[a], 16
LaG-3	15,329	25
LaG-6	15,700	310
LaG-9	16,062	3.5
LaG-10	15,748	97
LaG-12	16,090	56
LaG-14	16,002	1.9
LaG-16	16,306	0.7
LaG-17	15,823	50
LaG-19	15,528	24.6[a]
LaG-21	15,452	7
LaG-24	14,763	41
LaG-26	16,221	2.6
LaG-27	15,565	9.5
LaG-29	15,449	110
LaG-30	16,159	0.5
LaG-35	16,010	23.5[a]
LaG-37	16,329	24
LaG-41	15,471	0.9
LaG-42	15,490	600
LaG-43	16,167	11
LaG-5	15,589	14,200[a]
LaG-8	15,953	20,000[a]
LaG-11	16,221	22,900[a]
LaG-18	16,459	3,800[a]
LaG-16–G_4S–LaG-2	30,791	0.036
LaG-16–3×Flag–LaG-2	32,972	0.268
LaG-41–G_4S–LaG-2	29,956	0.150

Figure 1.2.3: Characteristics of LaG monomers and dimers.
LaG-16-G4S-LaG-2 tandem nanobody has been highlighted in yellow, alongside its MW (30kDa).
Figure adapted from Fridy et. al, (2014)

Epitopes on the GFP, which are recognized by the 12 highest affinity LaGs, were identified via chemical shift perturbation. This method allows the mapping of binding sites on a protein by following changes in its characteristic “fingerprinting” spectrum that result from adding an unlabeled ligand to a N15- labeled protein sample. The study then identified three distinct groups of nanobodies (Group 1-3) based on the binding epitopes they recognize. Of particular interest are the Group 1 (LaG16) and Group 3 (LaG2) nanobodies, respectively. Their binding epitopes on GFP[25,26] are located on opposite sides of the GFP molecule, thus avoiding almost all overlapping. (**Figure 1.2.4**)

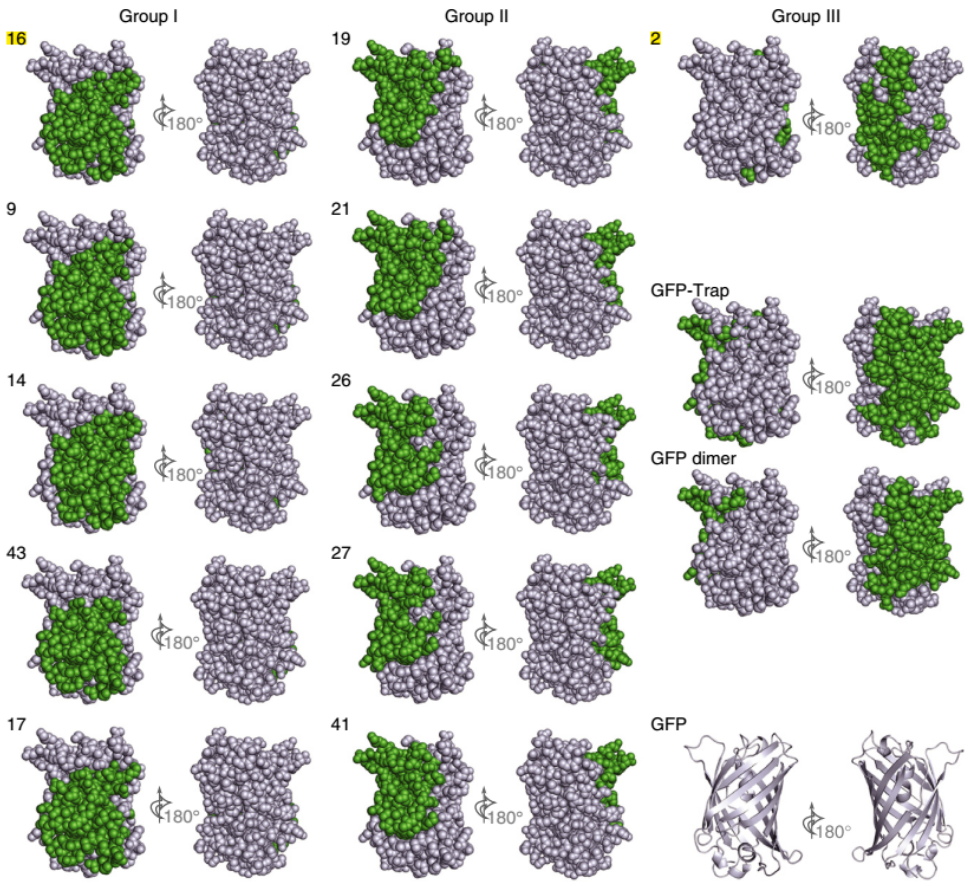

Figure 1.2.4: Mapping nanobody binding epitopes on GFP by NMR.
For each nanobody, two opposite sides of the GFP are shown (via a 180°C rotation on a vertical axis), with the binding site of the respective nanobody colored in green. The Lag16 and LaG2 nanobodies are highlighted in yellow.
Picture adapted from Fridy et. al, (2014)

It is the combination of these qualities that make the tandem LaG-16-G4S-LaG-2 construct a nanobody of interest. Due to its high specificity and the lack of GFP-epitope overlaps, it can be utilized as a mechanism to reach and interact with difficult targets.

1.4. FimGt/DsF - a stable protein-ligand complex

In general, purifying protein complexes has been made easily doable by the development of a number of affinity tags such as the cmyc-, FLAG-, GFP-, MPB- and His- tags.[27,28] These systems, however, suffer from dynamic binding equilibria and measurable dissociation rate constants which enable the competitive elution of bound target proteins when an excess of suitable free

ligand is present.[29-31] This, of course, hampers the isolation and studying of low-abundance protein complexes. In a study by Giese et.al. (2012), a novel affinity purification system derived from type 1 pili of *E. coli* was introduced. These pili are composed of four structural protein subunits termed FimH, FimG, FimF and FimA. In an assembled pilus, these subunits interact non-covalently by a mechanism called donor strand complementation. The mechanism is based on the completion of the imperfect, immunoglobulin-like fold of one subunit by an N-terminal extension, termed donor strand, of the successive subunit.[32-36] What is of particular interest is that the complex formed between FimGt, an N-terminally truncated variant of FimG lacking its own donor strand, and a peptide corresponding to the donor strand of the neighboring subunit FimF (DsF) was found to be an extremely stable protein-ligand complex. (**Figure 1.3.1**)

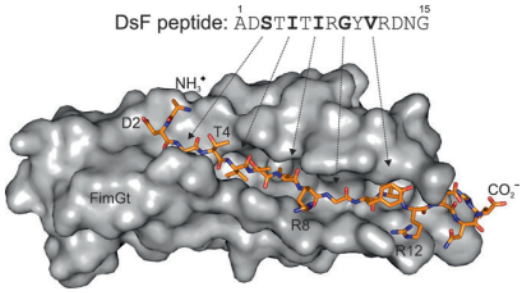

Figure 1.3.1: Crystal structure of the FimGt/DsF complex.
FimGt is shown as a gray surface and the DsF peptide as a stick representation. Residues that point towards FimGt are in bold and indicated with arrows.
Picture adapted from Giese et.al, (2012)

The FimGt/DsF system was then utilized as an affinity purification tool for extracting protein complexes directly from cell extracts. The donor strand of FimF can be used as the affinity tag (termed DsF tag) and FimGt as the binding partner. This was demonstrated in the same study by purifying DsF-tagged *E. coli* tryptophan synthase, a heterotetrameric complex of low cellular abundance.[37] (**Figure 1.3.2**)

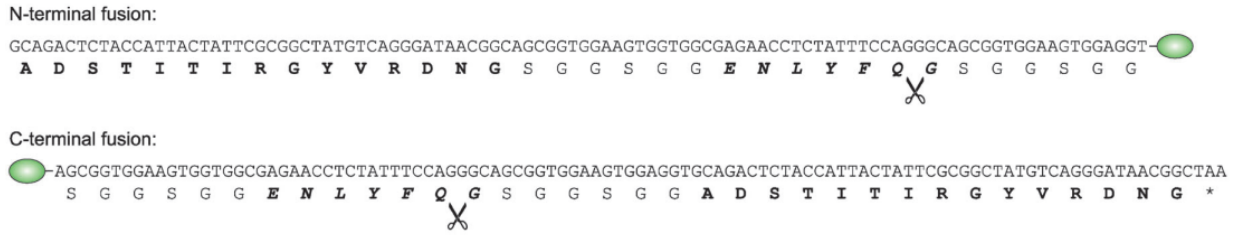

Figure 1.3.2: Nucleotide and amino acid sequences of the DsF tag fused to TrpA.
The tag can be fused to either the N or C terminus of a target protein. The DSF sequence is highlighted in bold and the TEV protease recognition site is shown in bold and italics.
Picture adapted from Giese et.al, (2012)

This new system proved to be superior to other commonly used purification methods, particularly in terms of isolating low-abundance protein complexes such as the before mentioned tryptophan synthase. As pure FimGt can be easily obtained recombinantly in high yields from *E. coli*, it renders the production of FimG-t-sepharose efficient and straightforward. (**Figure 1.3.3**) Thus, the proposed method is suitable for isolating pure and highly active multi-subunit protein complexes for structure determination or enzymatic assays. It also begs the question whether the FimGt approach can be combined with a second purification system, in order to further improve on its qualities and efficiency.

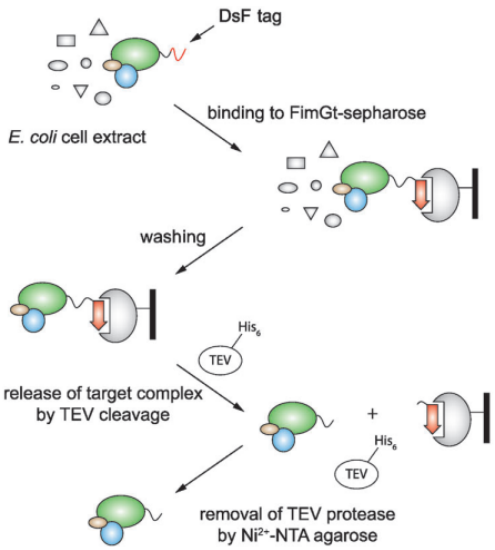

Figure 1.3.3: General workflow for the application of the FimGt/DsF system in affinity chromatography.
Picture adapted from Giese et.al, (2012)

1.5. Aim of the project

The LaG-16-G4S-LaG-2 (L16L2) is a novel recombinant nanobody that has only recently shown great potential as a mechanism for binding GFP with a high specificity and efficiency, demonstrating a high affinity for the protein. This opens up many possibilities for the isolation of various low abundance proteins, which have not yet been accessible for structural studies, using a method that would be faster and cheaper than existing methods. Furthermore, the binding strength between the FimGt truncated domain and the DsF peptide of type 1 pili of *E. coli* is outstanding. By combining both methods, a whole new way on how we perform protein isolation and purification could be opened up.

The aim of my project was to analyze the strength of the L16L2-GFP bond by cloning the GFP protein onto model proteins of the Type 3 secretion system of *Salmonella enterica* and attempting to isolate them via affinity chromatography. As a side question, FimGt would be cloned at three specific positions of the L16L2 nanobody while the DsF peptide would be cloned at two specific positions on the GFP molecule, thus allowing to monitor whether this double affinity construct would successfully increase the binding strength of the model protein capture system.

2. MATERIAL AND METHODS

1. Cloning genes of interest in different expression vectors

For the cloning of several genes (*FimGt, DSF, eGFP*) into the according vectors conventional cloning with restriction enzymes (RE) as well as Gibson assembly was used for large constructs. To amplify the genes of interest, *Escherichia* genomic DNA (purified per the Bacterial Genomic Miniprep kit protocol from Sigma) was used as templates in polymerase chain reaction (PCR). Primers were designed according the experimental need as described in the following cloning strategies. All the genes, vectors and primers used for the cloning experiment can be found in the Cloning List (**Section 7.2**) and therefore only the principle strategy is described here.

Conventional cloning using restriction enzymes

The conventional cloning strategy is shown in Figure 2.1 and performed as follows: RE are proteins cutting specific recognition sites. In this study, only enzymes leaving a 3´-or 5´-overhang (sticky ends) were used. To be able to clone a DNA sequence of interest into a multiple cloning site (MCS) of a vector, insert and vector had to be treated with two restriction enzymes creating compatible ends. For this purpose, primers had to be designed already containing homology region to the insert sequence and the restriction sites. Using PCR, restriction sites were added to both ends of a dsDNA, which was then digested by the RE, ligated with the vector and transformed to *E. coli* Top10 competent cells.

Insert generation

For the PCR, a 50 µl set up was prepared containing ddH_2O, 1 x Buffer HI-Fi with 2 mM Mg^{2+} (BIOLINE), Deoxynucleotides (dNTPs), DMSO, forward and reverse primers with the restriction enzyme sequence included and VELOCITY DNA polymerase 1 U (BIOLINE):

1 x Master Mix		
Double distilled water (ddH_2O)	33	µl
5 x Buffer Hifi [10 mM Mg^{2+}]	10	µl
DMSO	1.5	µl
dNTPs	1.0	µl
genomic DNA [50 ng/µl]	2.0	µl
Primer forward [10 µM]	1.0	µl
Primer reverse [10 µM]	1.0	µl
VELOCITY DNA Polymerase [2 U/µl]	0.5	µl
total	**50**	**µl**

The Velocity DNA polymerase that was used provided fast amplification of DNA, allowing to synthesize up to 66 base pairs (bp) per second, as per the manufacturer's instructions. The time of extension and the annealing temperature of the primers were chosen accordingly to the insert size and the primer melting temperatures. The reaction was performed in an Eppendorf Mastercycler with the steps shortly sketched as follows:

step	note	temp [°C]	t [sec]	cycle
1	initial denaturation	98	120	
2	denaturation	98	30	
3	primer annealing	/	30	x 35
4	extension	72	/	
5	final extension	72	150	
6	cooling	4	∞	

The PCR product was loaded on a 1 % Agarose gel and the band corresponding to the theoretical molecular weight was cut out and purified using the Wizard® Gel and PCR Clean-Up System protocol (Promega). The concentration was measured using a NanoDrop spectrophotometer (peQLab).
For the vector an overnight culture of *E. coli (*strain DH5α) containing the sequence of interest was grown and Plasmid Miniprep was performed.

Double digest
The double digest was performed in a 50 µl set-up containing 1 µg DNA of either the insert amplified via PCR or the vector, fast digest restriction enzymes (Thermo Scientific) and 1 x Fast Digest Buffer (Thermo Scientific). The digestion mix was incubated for 1 h at 37 °C. Additionally single cutting reactions of the vector were set up as negative control and investigated on a 1 % Agarose gel. The reactions were purified using the Wizard® PCR Clean-Up System protocol (Promega) to remove the small cut fragments.

Ligation, transformation and sequencing
The next step of the cloning protocol was the ligation of the insert into the linearized vector, ensuring a 3 x molar excess of the insert. This step is catalyzed using the T4 DNA Ligase (Thermo Scientific). Depending on the ratio between vector and insert DNA the desired amount of DNA was calculated as follows:

$$\frac{Vector\ bp}{Insert\ bp} = x \quad -\rightarrow \quad \frac{ng\ vector}{x} \times 3 = ng\ insert$$

1 µl T4 DNA Ligase (5 U/µl) and 1 µl 10 x T4 DNA Ligase buffer were added to a 10 µl reaction set-up. The mixture was incubated for 2 h at room temperature and immediately afterwards transformed into chemically competent cells *(E. coli*, strain DH5α) as described below (Section 2.2). The next day colonies were picked, grown for plasmid Minipreps (2.3) and sent for sequencing to check the consistency of the new construct. The plasmids containing the correct insert in the desired position were stored at -20 °C for future experiments.

Gibson assembly for larger constructs

The Gibson assembly method was used in this study for cloning of large constructs of up to 9 kbp. The reaction was performed under isothermal conditions utilizing three enzymatic activities: a 5' exonuclease which generates long single stranded overhangs, a polymerase for filling the gap of the annealed regions, and a DNA ligase which is needed for the nick filling of the gaps. The whole set-up is sketched in Figure 2.2, adapted from the New England Biolabs website.

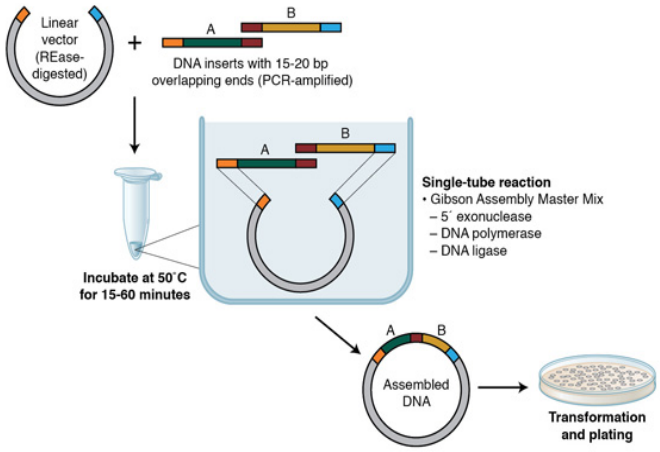

Figure 2.2: Sketch for Gibson assembly
Via PCR, DNA inserts are generated containing overlapping homology regions to the linearized vector of interest. In an isothermal reaction, the insert DNA and the vector are assembled and can be transformed quickly after
(Adapted from www.neb.com)

Insert preparation

For the PCR of the insert the same set up was used as for conventional cloning described earlier. The added primers were designed with 40 bp homology region to the insert of interest and 20 bp homology regions to the corresponding vector. Due to the high melting temperature of the primers a two-step amplification program was used. A short first step with a low primer annealing temperature was used to amplify few templates from the plasmid and a long second step with a high primer annealing temperature (almost at the extension temperature) was

used to amplify based on the templates generated in the low temperature step. The extension time differed according to the length of insert used (**Section 7.2**). Steps are shortly sketched as followed:

step	note	temp [°C]	t [sec]	cycles
	initial denaturation	98	120	
1	denaturation	98	30	x 5
	primer annealing	55	30	
	extension	72	25	
	final extension	72	150	
2	denaturation	98	30	x 35
	primer annealing	69	30	
	extension	72	25	
	final extension	72	150	
	cooling	4	∞	

The PCR product was loaded on a 1 % agarose gel and the corresponding band was cut out and purified using the Wizard® Gel and PCR Clean-Up System protocol (Promega). The concentration was measured using a NanoDrop spectrophotometer (peQLab).

Vector preparation

The corresponding vector had to be linearized. This was performed utilizing PCR by designing the primers at the precise position where the insert should assemble or via RE digestion as described under conventional cloning. When the PCR product was used, a DpnI digest for 1 h at 37 °C was performed in the PCR buffer to remove the template vector. Afterwards, the samples were loaded on a 1 % Agarose gel and purified using the Wizard® kit. In case RE were

used, the overhangs were filled up with Klenow fragment (Thermo Scientific). Therefore, the digested vector was purified, 1 µl dNTPs and 0.5 µl Klenow were added to a 40 µl reaction mix and incubated for 30 min at 37 °C. Afterwards the samples were loaded on a 1 % Agarose gel and purified using the Wizard® kit.

Isothermal Gibson Assembly

For the reaction mix 1 µl of linear vector (10 - 30 ng) were used and the amount of insert was calculated as followed, to achieve a 3 x molar excess:

1) $$pmols = \frac{(weight\ in\ ng) \times 1000}{base\ pairs \times 650\ daltons}$$

2) $$amount\ insert\ [\mu l] = \frac{pmol\ (vector)}{pmol\ (insert)} \ -\rightarrow when\ 1\ \mu l\ vector\ is\ used \ .$$

The vector was mixed with the insert according to the calculations. 3 x the volume of the Gibson assembly stock solution (in-house from the Molecular Biology Service) were added and the solution was incubated for 1 h at 50 °C. Afterwards the whole set-up was transformed into chemically competent *E. coli* cells as described right after in 2.2.

2. Transformation of Plasmid Deoxyribonucleic acid (DNA) in Escherichia coli

Plasmid amplification was performed in chemical competent *E. coli* (strain DH5α) prepared from the in house Molecular Biology service. The competent cells were thawed from -80 °C slowly on ice before usage and later transformed with the following conditions: 50-100 ng of plasmid DNA was added to 150 µl competent cells and the sample was incubated on ice for 30 min. Afterwards heat shock was performed for 30 sec. at 42 °C and immediately afterwards cells were put on ice again for 5 min. For recovery and to express the antibiotic resistance from the plasmid DNA, 950 µl of pre-warmed SOC medium were added and cells were grown for at least 30 min (for Ampicillin resistance) or 60 min (for Chloramphenicol) at 37°C with 850 rpm. The media was then spread on two separate plates, one only with 100ul and the second with the remaining 850ul, and incubated overnight in a 37°C incubator.

3. Plasmid Minipreps and Sequencing

After transformation, single colonies were picked from the plate and grown in 5 ml LB with selective antibiotics for at least 6 h at 37 °C. The bacteria were harvested with 4,600 x g for 15 minutes in a Heraeus Multifuge 3 S-R tabletop centrifuge. Plasmid isolation was performed

as per instructions from the QIAprep Spin Miniprep Kit protocol (Qiagen). The concentration of the plasmid was measured using a NanoDrop spectrophotometer (peQLab).

Sequencing of the plasmid DNA was performed in-house by the Molecular Biology Service via Sanger sequencing. Around 150 ng plasmid DNA were needed per reaction and sequenced with 1 ul of 5 µM primer of interest.

4. SDS- Polyacrylamide-Gel electrophoresis and Coomassie Brilliant Blue staining

The SDS-Page of varying percentage were prepared by mixing 4 % and 10-12-15-20 % (depending on experimental needs) polyacrylamide solutions (see buffer and reagents) in separate beakers. The separation gel was mixed first and inserted into a gel casting station via a syringe, followed by the same procedure for the stacking gel several minutes later. A comb was inserted and gels were left to polymerize at room temperature for about 15 mins, then used immediately or stored at 4 °C in a moisture chamber.
Laemmli buffer (see buffer and reagents) was added to the samples according their optical density as followed:

$$Volume\,Laemmli\,buffer = \frac{Optical\,Density\,600\,nm}{0.02\,\frac{units}{\mu l}}.$$

The samples were boiled for 5 min at 95 °C with added Laemmli buffer, in order to denature proteins. Next, the samples were loaded onto a prepared fixed percentage gel with a commercial protein standard (PAGE ruler; Thermo Scientific). SDS-PAGE was conducted at 50 mA per gel for about 30 min in 1 x SDS Running Buffer.
To visualize the bands on the SDS-PAGE Coomassie Brilliant Blue staining was performed. Therefore, the gel was placed in a small vessel and washed with ddH_2O twice for 2 min in the microwave. PageBlue™ Protein Staining solution (Thermo Scientific) was poured onto the gel and incubated for 20 min with shaking. The gel was then destained in ddH_2O until bands became visible.

5. Western Blotting

Polyacrylamide gels were run according to the SDS-PAGE protocol, rinsed with ddH_2O and incubated in 1x Protein Transfer Buffer (PTB) (see buffer and reagents) for 10 min. Pre-cut Polyvinylidene fluoride (PVDF) membranes were incubated 3 min in methanol and then in PTB for 5 minutes. The semi-dry transfer was carried out by building up a sandwich (from top to

bottom) containing 3 soaked filter papers, the gradient gel, the PVDF membrane and again 3 soaked Whatman filter papers. The transfer was set up for 45 min with a constant voltage of 13 V. After transfer, the membrane was blocked for 20 minutes with 5% bovine milk and PBS-T to avoid unspecific antibody binding. The primary antibody was diluted as per its tested dilution (see buffer and reagents) in 5% bovine milk and the membrane was incubated in it overnight at 4 °C. The next day the membrane was washed three times with PBS-T for 20 min followed by incubation with a commercially available secondary antibody (goat anti-rabbit HRP, 1:10000, Sigma) for 1 hour at room temperature. Afterwards the membrane was washed again three times for 20 min with PBS-T. The western blotting substrates (ECL Plus Solution A and B, Pierce) were mixed as per the manufacturer's instructions and applied to the membrane for 3 minutes. The Western blots were developed either by exposing a photosensitive film on top or by using a Molecular Imager™ (Bio-Rad).

2.6 LaG-16-G4S-LaG-2 recombinant nanobody extraction and purification

For the isolation of the LaG-16-G4S-LaG-2 recombinant nanobody, a protocol developed by the Fridy lab from The Rockefeller University was used and adjusted to some extent. All the used buffers can be found under supporting material (buffer and reagents), therefore only shortened names are used to describe the used buffers.

Procedure:

50 ml overnight pre-culture of the *E. coli* strain, carrying a plasmid (pET21-pelB-VHH) with the *lac* operon, was inoculated from a glycerol stock and grown overnight in LB supplemented with required antibiotics at 37°C, 220 rpm. The next day, the culture was diluted 1:100 in 1L LB and grown at 37°C, 220 rpm until a required OD600 of 0.6 was reached. The cultures were then transferred to 18°C, induced with 0.1mM IPTG and grown further for 18-24 hours. 1L cultures were harvested on the next day by centrifugation for 10 minutes at 5000 rpm. The resulting pellet was then re-suspended in 10 ml TES buffer per 1L of bacterial culture along with 20 ml TES diluted 1:4 with ddH2O and the additional 50x stock solution of cOmplete Mini, EDTA-free protease inhibitor cocktail tablets (Roche, 1:100). The pellet was then incubated 45 minutes at 4°C on a shaker. This facilitated the lysis of the cells and the release of the recombinant nanobody in the resulting supernatant. To separate the supernatant from the cell material, two centrifugation steps were performed in immediate succession: 6000 g for 10 minutes, followed by 48000 g for 15 minutes. 0.15 M NaCl was then added to the collected supernatant and the SN was loaded onto a 5ml HisTrap HP column (GE Healthcare Life Sciences). The flow-through was collected and the column washed two times with 5 CV

(column volumes) of Wash Buffer 1 and 2. The recombinant nanobody was then eluted with 4 CV Elution Buffer and could be further purified via dialysis into PBS (via 3.5k MWCO tubing) or via a Superdex 75 gel filtration column (GE Healthcare Life Sciences) equilibrated with TNB storage buffer. The resulting purified nanobody was frozen in liquid nitrogen and stored at -80°C for a prolonged period.

2.7 eGFP construct extraction and purification

In order to isolate the various eGFP constructs, an in-house protocol was used with small modifications. A quick overview of the strategy can be seen on Figure 2.3.

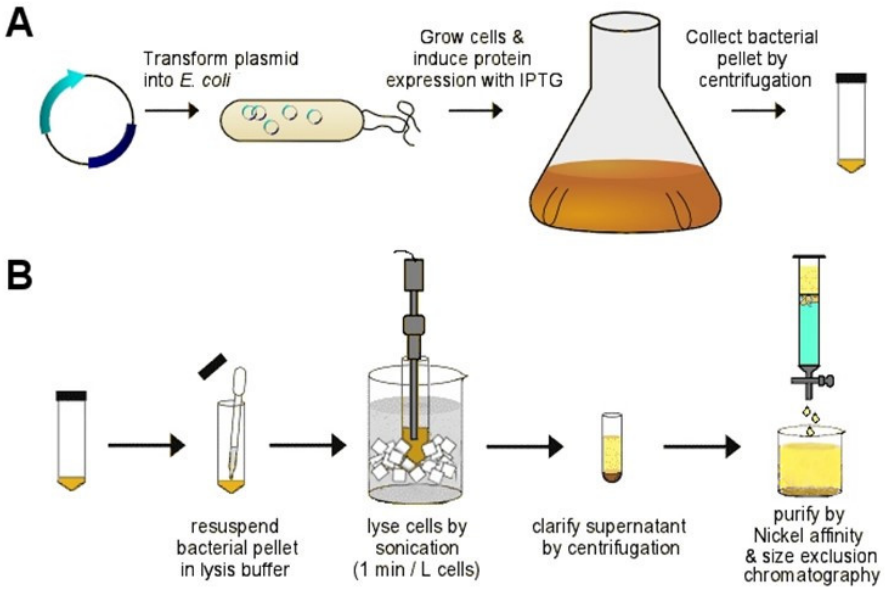

Figure 2.3: A schematic of the purification approach.

The sample is prepared by pelleting a cell culture, expressing the protein of interest. The pellet is then re-suspended and lysed via sonication, the supernatant is separated via centrifugation and loaded onto a HisTrap HP column for further purification.

(Adapted from www.neb.com)

As previously stated, all used buffers can be found under supporting materials (buffers and reagents), therefore only shortened names will be used to describe them.

Procedure:

50 ml overnight culture of the *E. coli* strain, carrying the constructed plasmid of interest was inoculated from a glycerol stock, supplemented with the required antibiotics and grown

overnight at 37°C, 220 rpm. The next day the culture was diluted 1:100 in 1L LB and grown at 37°C, 220 rpm until an OD600 of 0.6-0.8 was reached. The cultures were then induced with 1mM IPTG and grown overnight, with samples being taken pre- and post-induction for SDS-PAGE to check expression levels. On the next day, the cell pellet was collected by centrifugation at 4600 rpm for 15 minutes and either frozen at -20°C or directly proceeded to extraction. The 2-step-extraction protocol was adapted from an in-house protocol. The pellet was re-suspended in 20 ml buffer B1 with cOmplete Mini, EDTA-free protease inhibitor cocktail tablets (Roche) and then proceeded to sonication for 1 min at 20% amplitude and constant pulse (Fisher Scientific). The sample was then centrifuged at 15000 g for 30 minutes to separate the supernatant from the pellet. The supernatant was then collected in a separate tube and the remaining pellet was again re-suspended in 20 ml of buffer B1 with cOmplete Mini, EDTA-free protease inhibitor cocktail tablets (Roche). The second step of the extraction then proceeded identically to the first one. The supernatants of both steps were mixed and filtered through a 0.45 um syringe-filter. It then proceeded to His Trap purification identical to the recombinant nanobody, except for the buffers being used: B1 and B1+500 mM Imidazole. As in some cases, the His tag needed to be removed post-purification, it was cleaved off via adding GST-TEV enzyme to the unified elution fractions. The samples were incubating at 4°C overnight. The protein construct was purified further over a Superdex 75 or Superdex 200 gel filtration column (GE Healthcare Life Sciences), depending on the size of the target protein. For long term storage, the protein containing fractions were pooled, frozen in liquid nitrogen and stored at -80°C.

2.8 Maltose binding protein-eGFP extraction and purification

To isolate the maltose binding protein-eGFP construct, a modified version of the eGFP purification protocol was used. As previously stated, all used buffers can be found under supporting materials (buffers and reagents), therefore only shortened names will be used to describe them.

Procedure:

The procedure was performed identically to the eGFP extraction and purification protocol. The samples were subjected to a MBP Trap purification step which was similar to a regular His Trap purification, the difference being in the column (1ml MBP Trap HP, GE Healthcare Life Sciences) and the buffers (MBP Binding buffer, MBP Elution buffer) used. The MBP-eGFP construct was further purified over a Superdex 200 filtration column (GE Healthcare Life

Sciences). For long term storage, the protein fraction was frozen in liquid nitrogen and stored at -80°C.

2.9 Modified sandwich ELISA for quantifying of protein-protein interactions

In order to perform a simple quantification of the interactions between the proteins of interests (LaG16 and MBP-eGFP), a modified version of the sandwich ELISA protocol developed by Abcam was used. As previously stated, all used buffers can be found under supporting materials (buffers and reagents), therefore only shortened names will be used to describe them.

Procedure:

A 96-well PVC microtiter plate was coated with the "capture" protein- LaG16 at a concentration of ~5 ug/ml in TNB storage buffer. The plate was then covered with an adhesive plastic and incubated overnight at 4°C. The next day, the coating solution was removed and the plate was washed 3 times with 200 ul PBS per well. The solution and washes were removed via flicking the plate over a sink and then patting it on a paper towel. The remaining protein binding sites in the coated wells were blocked by adding 200 ul of TNB Blocking buffer per well. The plate was again covered with an adhesive plastic and incubated for 2 hours at room temperature. The blocking solution was then washed away with PBS and 100 ul of the second protein of interest (eGFP/MBP-eGFP) were added in both an equimolar ratio and serial dilutions to different wells. The plate was then incubated for 90 min at 37°C after which the samples were removed and the wells washed twice with PBS. The fluorescence of each well was then measured on a TECAN microplate reader, allowing the detection of any captured eGFP/MBP-eGFP complexes.

3. LISTS

3.1 Common solutions, Buffers, Reagents etc.

PBS pH 7.5	
NaCl	137.0 mM
KCl	2.7 mM
Na_2HPO_4	8.2 mM
KH_2HPO_4	1.5 mM
Adjust pH to 7.5 with 6.4 %	

Solutions provided by the in-house Media Kitchen

1 x TE	
Tris pH 8.0	10 mM
EDTA	1 mM

LB-Medium	
Bacto Tryptone	10.0 g
Yeast Extract	5.0 g
NaCl	5.0 g
Fill up with Mono Q water to 1	

0.5 M EDTA pH 8.0 (adjusted with NaOH)

1 M Tris pH 7.5 (adjusted with HCl)

1 M Tris pH 8.0 (adjusted with HCl)

1 M $MgCl_2$

5 M NaCl

Super Optimal Broth (SOB)-	
Bacto Tryptone	20.0 g
Yeast Extract	5.0 g
NaCl	8.5 mM

Self-made solutions

Chloramphenicol (CM) [25 mg/ml]

25 mg/ml stock solution in EtOH, used 1:1000

Ampicillin (Amp) [100 mg/ml]

100 mg/ml stock solution in Mono Q, used 1:1000

Tetracycline (Tet) [15 mg/ml]

15 mg/ml stock solution in EtOH, used 1:1000

Blocking Solution for Immunoblotting

3 % Bovin Serum Albumin in PBST

PBS-T

PBS provided by the Media Kitchen with 0.1 % Tween

Gels (4-20 %) for SDS-PAGE

	4 % gel	20 % gel
ddH_2O	30.0 ml	3.0 ml
4 x Resolving Gel buffer (1.5 M Tris-HCl pH 8.8)	12.5 ml	12.5 ml
10 % SDS	0.5 ml	0.5 ml
TEMED Tetramethylethylendiamin (PlusOne)	50 µl	20 µl
Acrylamide (Protogel 37.5:1 Acrylam.:Bis-Acrylam.	6.5 ml	33.5 ml

5x Laemmli sample buffer
300 mM Tris-Cl pH 6.8
10% SDS
50% glycerol
100 mM DTT
0.05% bromophenol blue

Wash Buffer 1 (nanobody isolation)	
Hepes (pH 8.0)	20 mM
NaCl	0.9 M

1 x Protein Transfer Buffer	
Glycine	2.9 g
Tris base	5.8 g
SDS	0.37 g
ddH_2O	800 ml
Fill up with Methanol to 1 L	

Wash Buffer 2 (nanobody isolation)	
Hepes (pH 8.0)	20 mM
NaCl	0.15 M
Imidazole	10 mM

Elution buffer (nanobody isolation)	
Hepes (pH 8.0)	20 mM
NaCl	0.15 M
Imidazole	250 mM

Binding buffer (nanobody isolation)	
Hepes (pH 8.0)	20 mM
NaCl	0.15 M

Storage buffer TNB (nanobody storage)	
Hepes (pH 8.0)	20 mM
NaCl	150 mM
Tween-20	0.05%
EDTA	0.25 mM

10 x SDS Running buffer	
Glycine	288.4 g
Tris base	60.6 g
SDS	20.0 g
Fill up with Mono Q water to 1	

4. Results

4.1 Large-scale expression and purification of LaG-16-G4S-LaG-2 tandem nanobody

In order to study and implement the L16-L2 tandem nanobody in further experiments, a highly pure sample needed to be obtained. As already shown in (Fridy et al., 2014), the pelB fused nanobody could be expressed in Arctic Express (DE3) and BL21 *E. coli* expression vectors. Here, the protein was expressed in a heterologous manner using IPTG induction, as seen in **figure 4.1.1**, and exhibited a relative size of 30 kDa - a value which agrees with data previously shown by the Rout lab. [4]

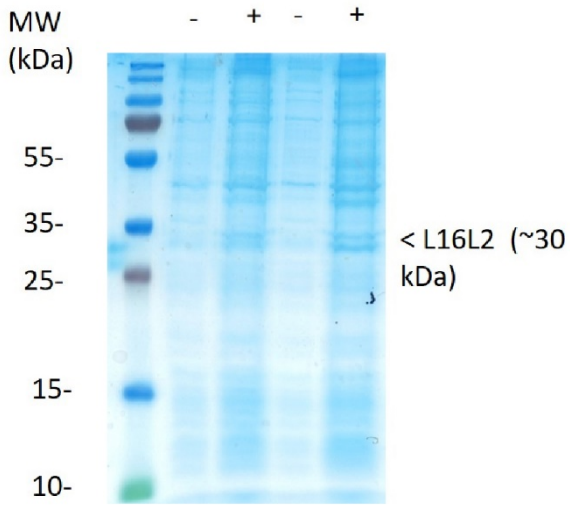

Figure 4.1.1 Heterologous expression of recombinant nanobody and expected size of protein.

To test whether the tandem nanobody could be purified, a NiNTA purification system was employed. The tandem nanobody was expressed in a BL21 vector at 18°C overnight. Samples were taken prior to and after induction with IPTG, indicated by – and + respectively.

The presence of a 6x His tag allowed for capturing the nanobody on a Ni Sepharose High Performance HisTrap column and eluting it via an imidazole gradient. All steps of the protein extraction process were collected and analyzed on an SDS-PAGE gel in order to detect the presence of the tandem nanobody (**figure 4.1.2**). As expected, the protein exhibited a relative size of 30 kDa, as previously mentioned.

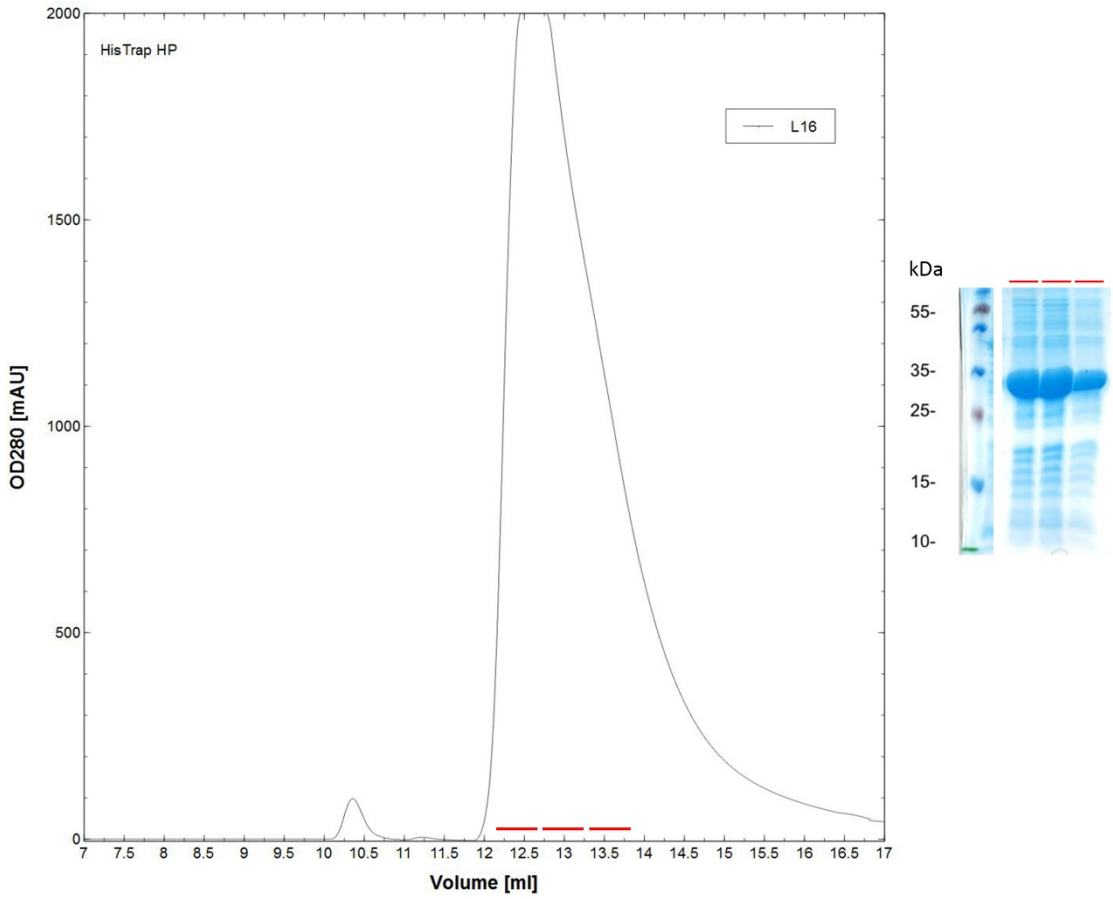

Figure 4.1.2 SDS-Page analysis of L16-L2 purification profile over a His-Trap column.

The SDS-PAGE analysis of the tandem nanobody purification shows its' strong presence in the eluted fractions which correspond to the large peak on the right. The small waste peak in the beginning and the silent region in between both peaks suggest that little-to-no amount of protein was lost prior to elution.

Having eluted the tandem nanobody via HisTrap, the sample was further purified via size exclusion chromatography on a Superdex75 column (**Figure 4.1.3**). The molecular weight of the purified protein eluted within the peak fractions at 13 ml. The size could be calculated based on a standard curve generated from a standard run and amounted to a relative mass of around 30 kDa. This finding reaffirms the data shown by Fridy et al. (2014).

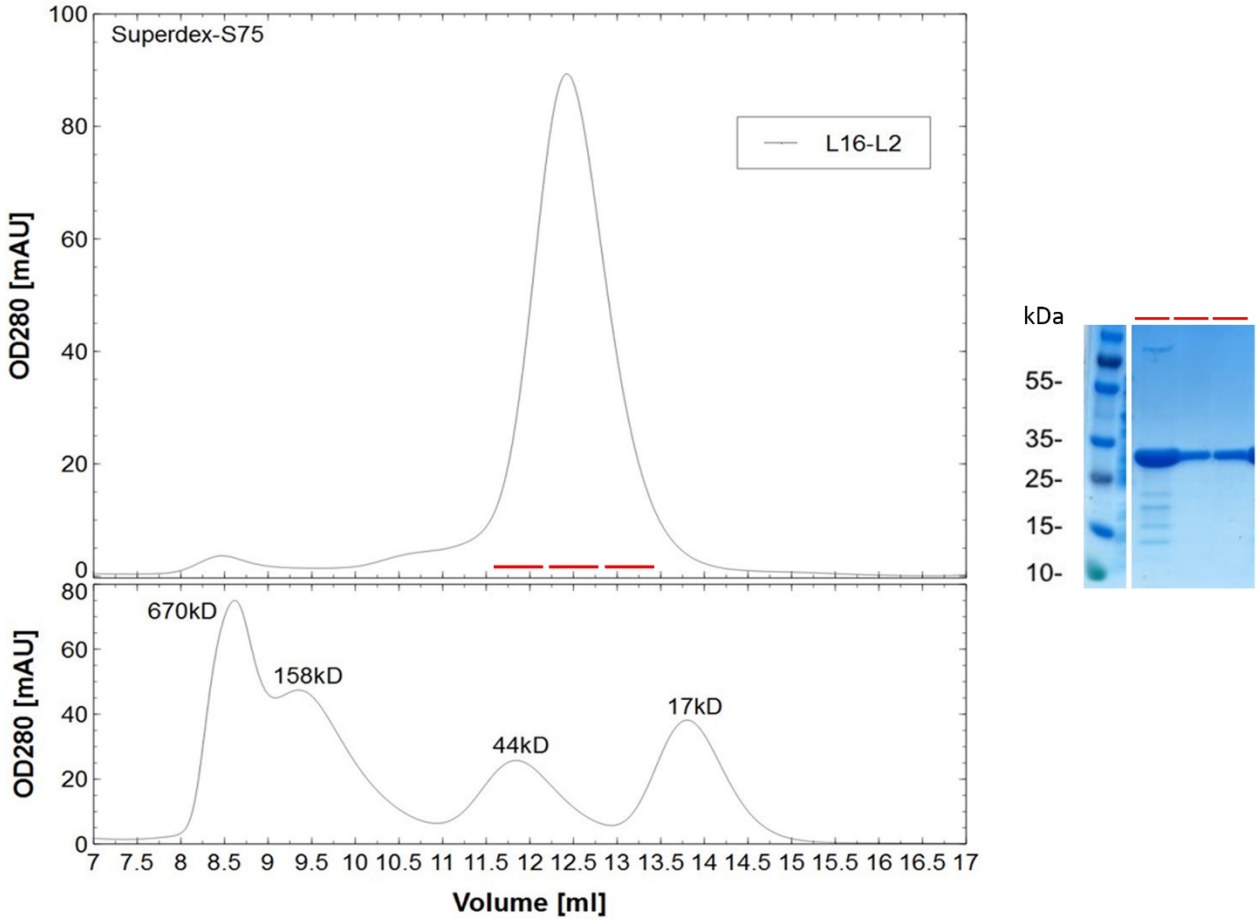

Figure 4.1.3 Purified L16-L2 tandem nanobody Superdex 75 (SD75) profile and SDS-PAGE analysis.

Size exclusion chromatography (Superdex 75) profile and SDS-PAGE analysis of peak fractions. L16-L2 elutes at 13 ml; the molecular mass was calculated via a standard curve to 39 kDa (Supplementary figure C1). The standard run SD75 profile is shown below the actual sample run; the gel image shows the approximate mass and relates to the peak via the black lines.

4.2 Stability of the L16-L2 tandem nanobody under varying temperature conditions.

The L16-L2 tandem nanobody was tested under several temperature conditions in order to discern how best to store it. As is generally known, optimally, antibodies should be divided into small aliquots and frozen at -20°C or -80°C for long-term storage, only being thawed upon need.[38] Once thawed, they can normally be kept and at 4°C and are viable for up to 2 weeks. They, however, lose efficiency with multiple freeze-thaw cycles.[5]

Upon testing several storage conditions (temperature and time) for the tandem nanobody, as seen in figure 4.2.1, it was determined that it is stable under all conditions tested, with little

variation present. The variation could be attributed to sample loading. Overall the protein is stable under all conditions, however higher molecular species tend to accumulate at room temperature indicated by arrows on the graph. Therefore, storage at -20°C for longer periods of time and at 4°C for shorter periods was chosen.

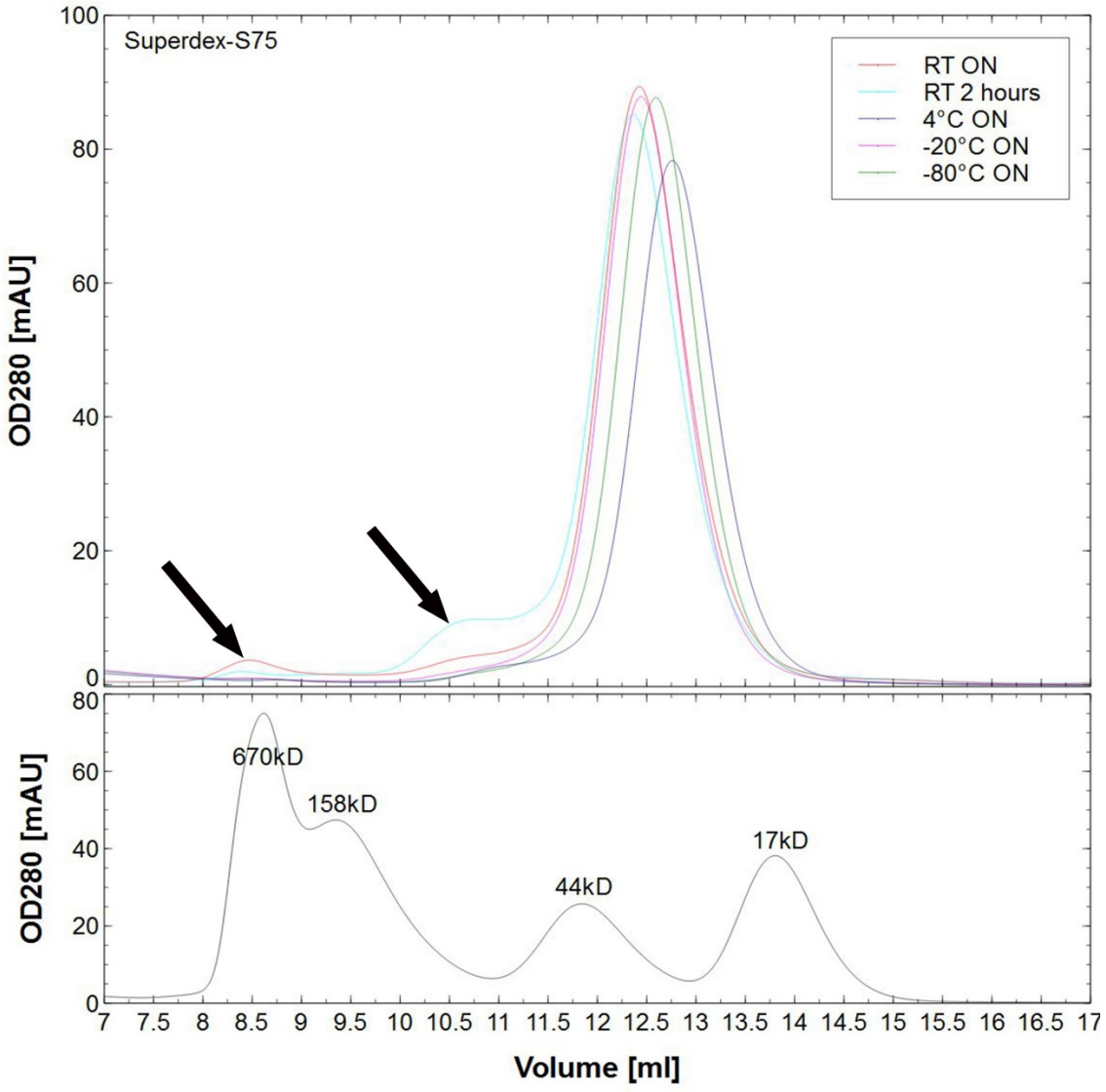

Figure 4.2.1: Elution profile of L16-L2 nanobody under various temperature and time conditions from Superdex 75 sample runs.

Several aliquots were stored at room temperature both overnight and for 2 hours and at 4°C, -20°C and -80°C overnight in order to test the durability of the protein. The nanobody exhibited similar stability under all tested temperature conditions. Arrows indicate accumulation of higher molecular species.

4.3 Expression and purification of L16L2-FimGt construct variants.

The FimGt peptide was cloned in three major sites of the L16L2-encoding plasmid, prior to the N-terminus of the L16 monobody (Variant 1), between the L16 and L2 monobodies (Variant 2) and after the C-terminus of the L2 monobody (Variant 3). All three variants were then expressed in a heterologous manner, identical to the original L16L2 nanobody. With to the addition of FimGt, the new constructs were approximated to have a rough molecular weight of 50 kDa, as indicated on **figure 4.3.1**.

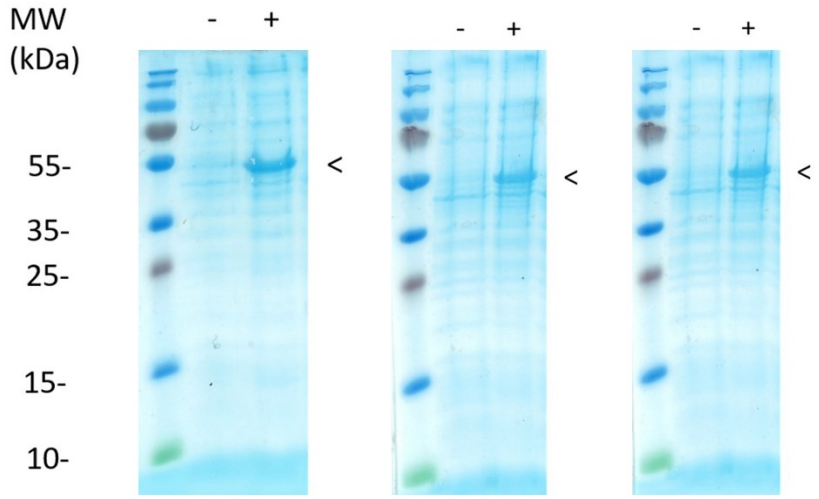

Figure 4.3.1 Heterologous expression (non-induced and induced) of FimGt+L16L2 construct variants.

The three constructs were expressed in a *E. coli* BL21 vector at 18°C overnight and have an expected size of ~50kDa. Samples were taken before and after induction with IPTG.

As the 6x His-tag was left intact, the constructs were proceeded to purification on a Ni Sepharose HP column. The purification, however, was not as successful as with the native L16L2 nanobody. The proteins eluted in minimal quantities and did not produce strong peaks on the chromatography profile or bands on an SDS-PAGE gel analysis, as shown in **figure 4.3.2**. Nevertheless, the collected protein fractions were loaded on a size exclusion column (SD75) using the TNB storage buffer.

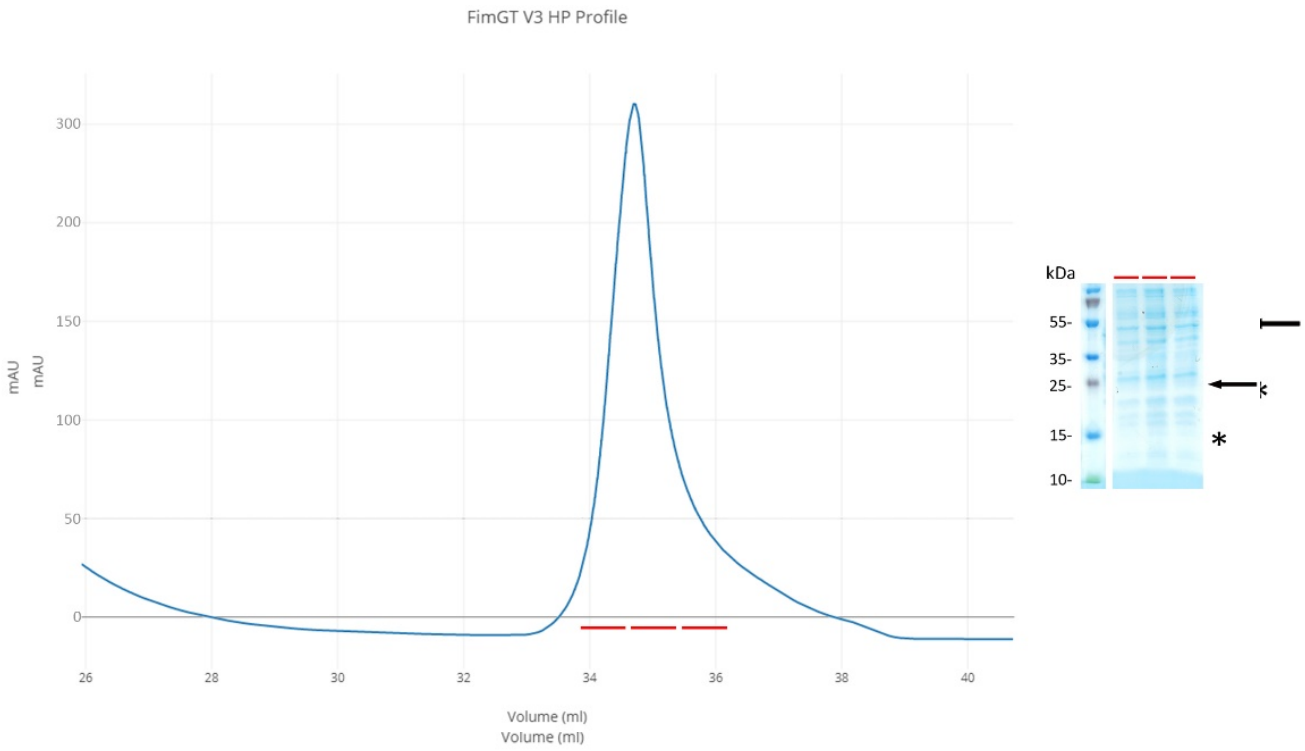

Figure 4.3.2. SDS-PAGE analysis (right) of FimGt constructs V1,2 and 3 and HisTrap profiles (left).

HisTrap chromatography of FimGt constructs V1,2 and 3. SDS-PAGE analysis confirmed that the protein constructs (**arrow**) were either dissociated (**asterisk**) during the purification process or were generally being expressed in lower quantities than the original L16L2 tandem nanobody.

The resulting profiles from each of the three FimGt construct purifications are shown in **figure 4.3.3**. In each of the three cases the proteins were calculated to have a rough molecular weight of ~150 kDa. This far exceeded the expected size of 50 kDa, indicating that the proteins were not purified successfully. This, alongside the unusual HisTrap profiles from **figure 4.3.2,** led us to believe that the protein may have dissociated or somehow degraded and aggregated during the purification process. After multiple attempts to stabilize the FImGt-L16-L2 constructs, focus shifted onto utilization of the original L16-L2 tandem nanobody.

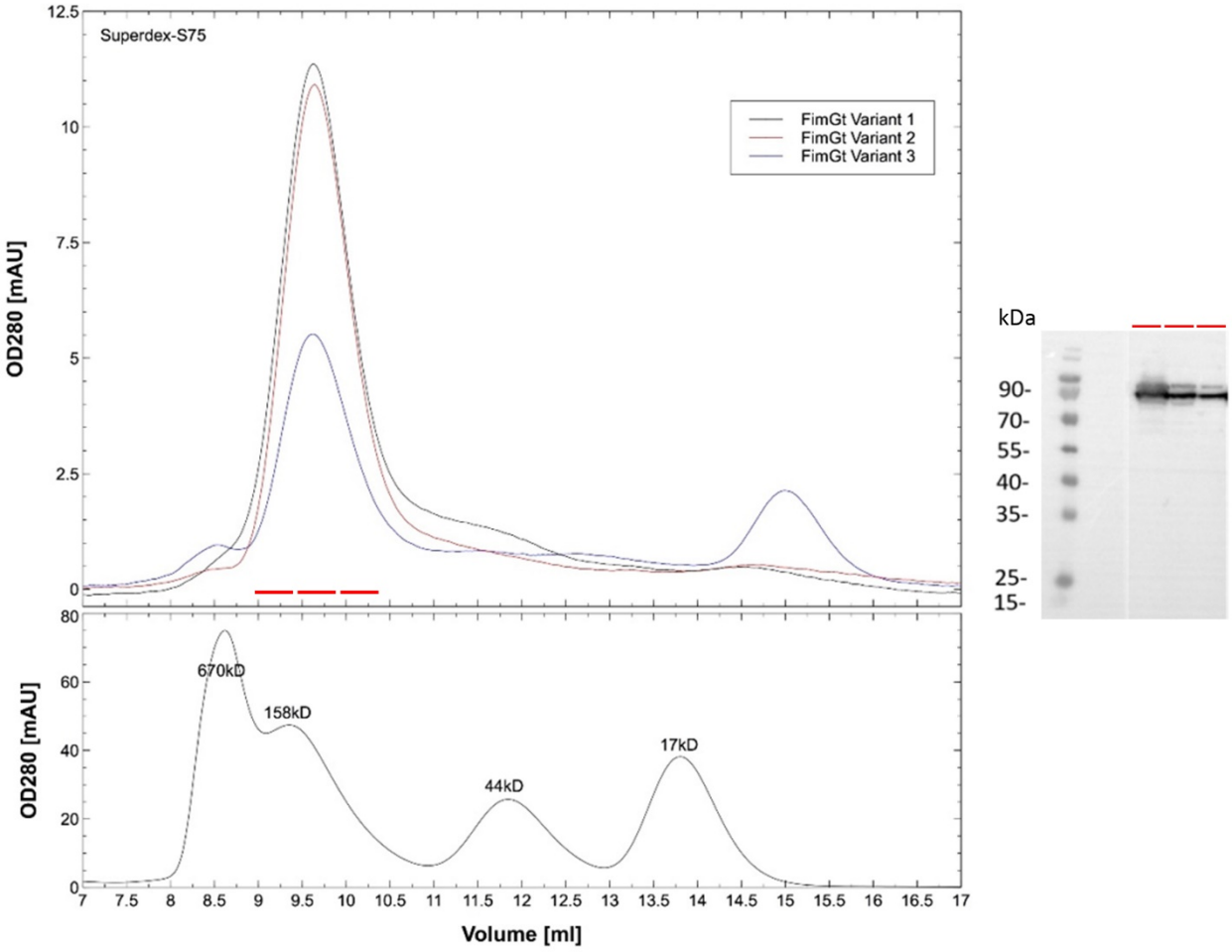

Figure 4.3.3 Elution profile and Western blot of FimGt+L16L2 Variants 1-3 from Superdex 75 column.

The size exclusion chromatography (SD75) of the three FimGt constructs, alongside a Western blot with anti-His antibody, confirmed that the attempts were unsuccessful. As seen above, the strong peaks could be calculated to ~150 kDa, a value much higher than the expected molecular weight. This could be attributed to possible contaminants in the sample in addition to the constructs being unstable and possibly dissociating or aggregating.

4.4 Expression and purification of Maltose binding protein -eGFP construct.

In order to test the binding of the L16-L2 tandem nanobody to any GFP tagged protein, a Maltose binding protein (MBP)-eGFP construct was used. The protein was provided by Prof. Thomas Marlovits and was stored in a *E. coli* BL21 expression strain. Its' rough molecular weight was calculated to approximately 70 kDa, as can be seen in **figure 4.4.1**.

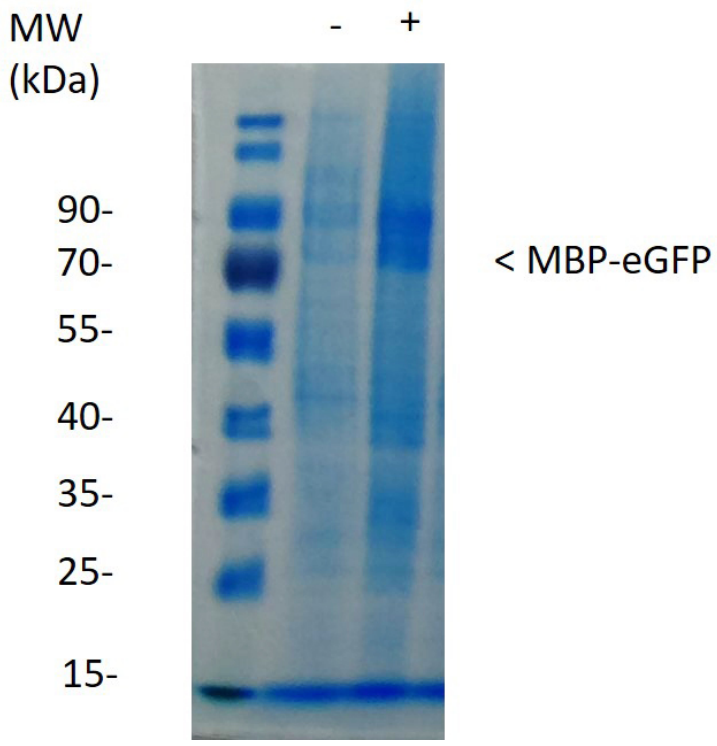

Figure 4.4.1 SDS-PAGE analysis of MBP-eGFP protein construct.

The construct was readily available in-house in a *E. coli* BL21 expression strain. The protein was expressed overnight at 16°C. Samples for the gel were taken before and after induction with IPTG.

Having expressed the protein successfully, it was then purified via an MBP-Trap HP column (GE Healthcare Life Sciences) designed for separating recombinant proteins, which have been tagged with MBP. The column was washed with MBP Binding buffer and the protein was eluted using MBP Elution buffer. The resulting chromatography profile and eluted fractions can be seen in **figure 4.4.2**

.

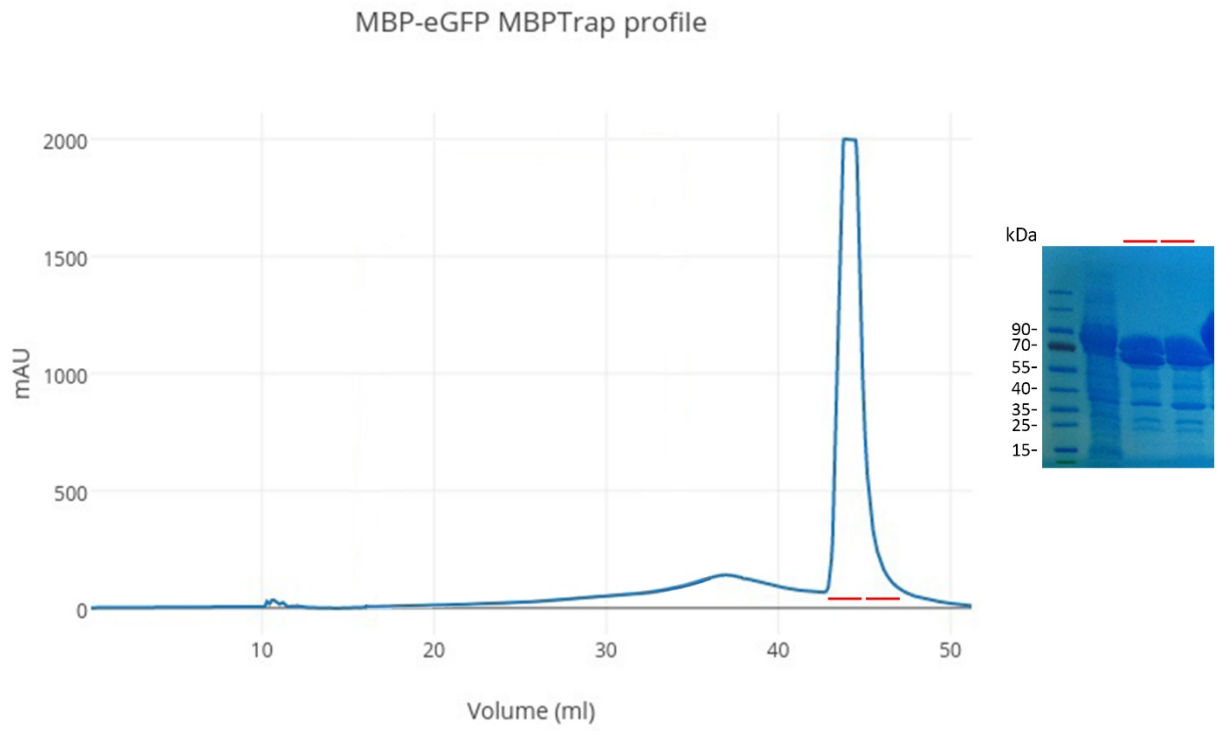

Figure 4.4.2. MBP-eGFP MBP-Trap chromatography profile and SDS-PAGE profile of MBP-eGFP construct.

MBP-Trap chromatography of the MBP-eGFP recombinant protein and the strong bands on the SDS-PAGE analysis confirmed the successful elution.

4.5 Co-elution of the tandem nanobody L16L2 and target protein MBP-eGFP.

Both proteins were incubated in an equimolar ratio for 1 hour at room temperature. The resulting complex was loaded onto a Superdex 200 column and eluted using Wash Buffer. Control samples for each of the proteins were also eluted using identical conditions as for the complex. Profiles of all three elutions can be seen in figure 4.5.1.

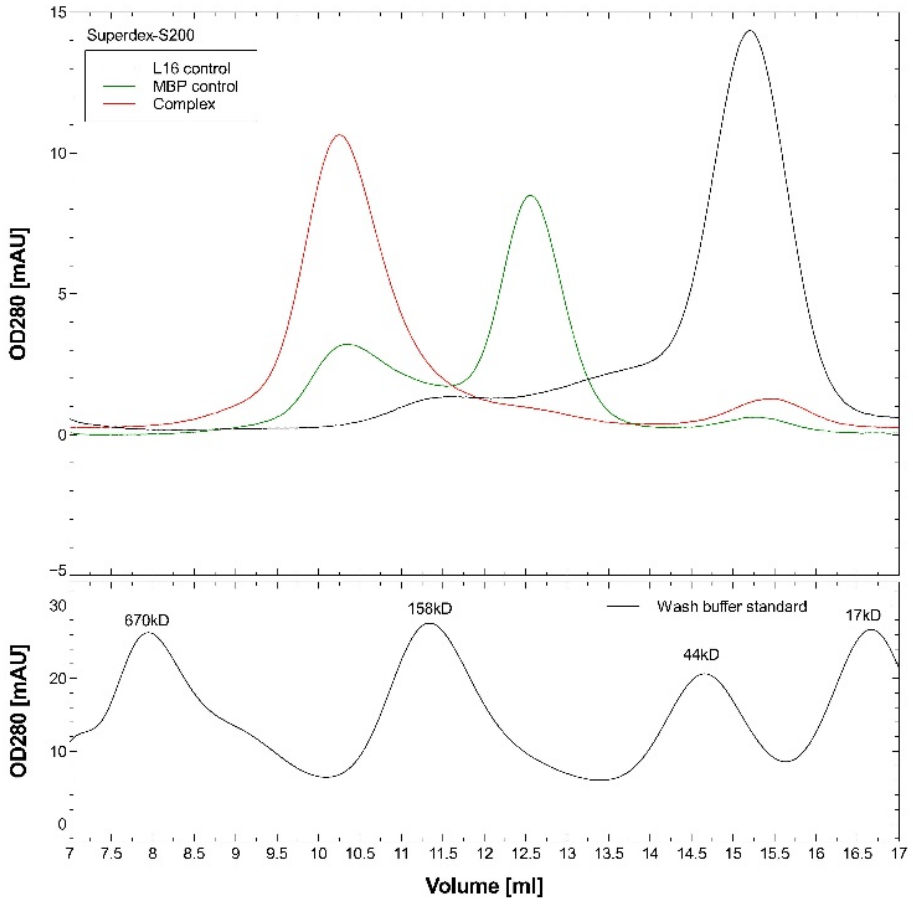

Figure 4.5.1. Co-elution of L16L2 tandem nanobody and MBP-eGFP construct.

Both control samples elute at different sizes (34 kDa for L16-L2 and 100 kDa for MBP-eGFP) while the complex eluted at a size higher than either of the two initial proteins (~ 300 kDa). The calculation was done via a standard curve (Supplementary figure C2).

In order to determine whether the proteins had indeed bound together, co-detection over Western blot would be needed. Fractions of each of the three separate runs were collected and loaded on a SDS gel for Western blot analysis. Two gels were ran using an anti-GFP and anti-His antibody for detecting any GFP and His-tags in the samples, respectively. The result, which is shown in figure 4.5.2, confirmed that L16L2 and MBP-eGFP were present in the complex SE fractions, thus unambiguously demonstrating that the tandem nanobody had successfully bound to the GFP construct and could be used for further pull-out and co-elution experiments.

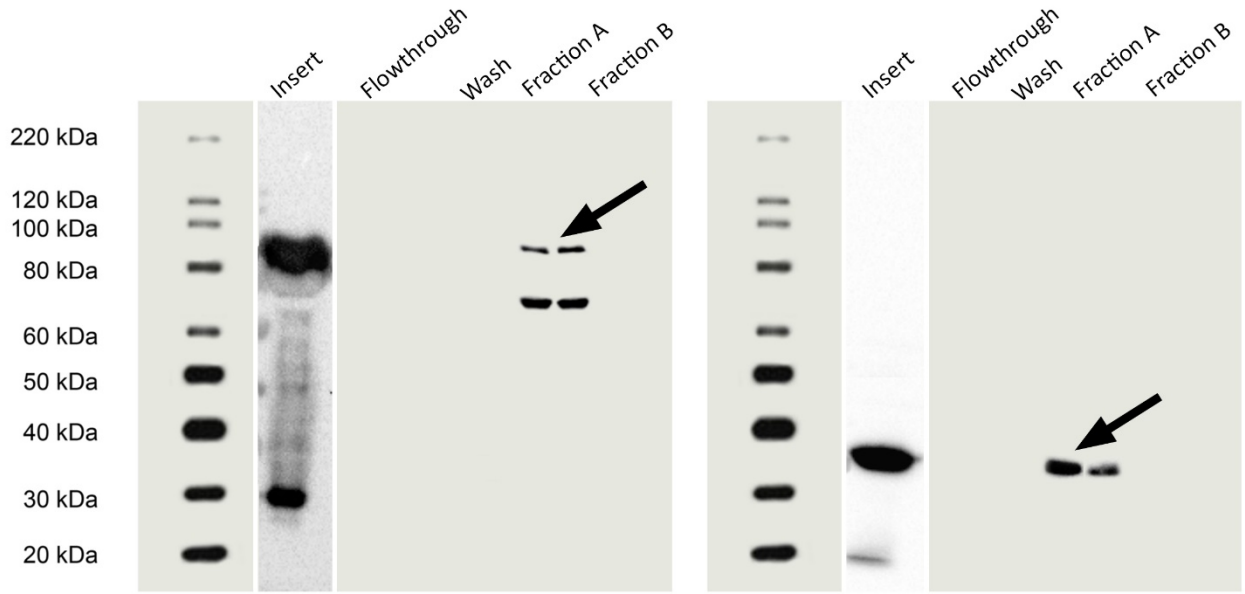

Figure 4.5.2. Western blot analysis of L16L2-MBP-eGFP complex fractions.

Anti-His (top) and anti-GFP (bottom) antibodies confirmed the presence of each of the two proteins (with their respective sizes) in the complex fractions.

5. DISCUSSION

The experimental work during my master thesis involved the biochemical investigation of the L16L2-GFP protein isolation system, evaluating its performance as means to purify low abundance complexes and improving upon it. Furthermore, first experiments concerning the combination of the L16L2-GFP-system with the FimGt-DsF system were performed.

Generally, membrane proteins are a challenging area of the proteasome and remain one of the most under studied. Currently developed synthetic systems for studying them consist of a solubilizing component and must satisfy the hydrophobic nature of the transmembrane segments while bringing loop regions into contact with an aqueous phase. [19] They still meet a number of difficulties during sample preparation and are quite time-consuming, which is why the L16L2-GFP system could prove an interesting solution to these problems.

The L16L2-GFP system allowed for the purification of a model recombinant protein – MBP attached to GFP[39] – via affinity chromatography.

Having cloned the FimGt insert onto three distinct places of the L16L2-encoding plasmid, purification and isolation of the newly formed construct was attempted. This was, however, met with multiple setbacks due to the degradation of the complex during affinity chromatography (**Figure 4.3.2**) and generally low levels of expression in the designated bacterial vector. (**Figure 4.3.1**) The plasmid was then also ordered in order to see whether there was a problem with the in-house cloning methods, however the issues persisted with the new plasmid as well.

Each of the purification attempts of these protein complexes were analyzed both via SDS-PAGE gel and size exclusion chromatography (**Figure 4.3.3**). All three experiments showed a much heavier protein complex than was expected, indicating possible aggregation of the proteins. After multiple attempts to stabilize the complex, focus was shifted on the original L16L2-GFP system.

The stability of the L16L2 tandem nanobody was tested under varying temperature conditions in order to determine how to proceed with further experiments and to see how best to store it for prolonged periods of time. The sample was divided into aliquots, which were stored under various conditions and later analyzed via size exclusion chromatography (**Figure 4.2.1**) The original tandem nanobody construct proved to be stable under all conditions tested, thus it was decided to store the protein at -80°C.

The binding between the proteins was tested via an adapted ELISA protocol from Abcam. The two binding partners were incubated in an ELISA plate in varying molar ratios. The concentration of the L16L2 tandem nanobody was kept constant. Due to its similarities to capture antibodies used in regular ELISA tests, it was utilized in a similar manner in the experiments at hand. The binding was then analyzed by monitoring the fluorescence of each well on a Tecan microplate reader with the results showing a strong binding efficiency when both proteins were at an equimolar ratio. (**Figure 4.5**)

After determining the proper molar ratio between the “capture” and target protein, the tandem nanobody was tested on a model protein: MBP bound to eGFP. Having incubated the two proteins for an appropriate amount of time, the proteins were loaded on size exclusion chromatography and analyzed further via SDS-PAGE and Western blot procedure (**Figure**

4.6.1). It was observed that both proteins formed a complex which was confirmed via antibodies against His6 tag and GFP (**Figure 4.6.2**).

In this thesis, I present the results of using the L16L2-GFP capture system to purify a model protein -MBP. Although the premise of the original aim - combining the strong binding affinities of the L16L2-GFP and the FimGt system - was very compelling, the respective constructs proved difficult to handle in many situations. Ultimately, the experiments were unsuccessful, namely capturing and purifying structural proteins (InvC, OrgB, SpaO, InvG) or sub-complexes of the Type 3 secretion system needle complex.

Future experiments utilizing the constructs should carefully determine optimal conditions for cloning and expressing the GFP attached to the target proteins. It should also be noted that the nanobodies are to be handled with extensive care as they can easily lose their function and dissociate if mishandled.

I would also like to point out that if successful, this capture system would be greatly beneficial as it can replace the time and resource-consuming conventional purification methods that are currently in use for purifying the Type 3 secretion system needle complex or comparable low abundance (membrane) complexes.

There are still many questions left open to be addressed in future studies:

- How can the L16L2-GFP capture system be easily applied for purifying a specific low abundance protein?

- What are the reasons behind the sudden dissociation or aggregation of the FimGt variants?

- In which expression vector should the GFP fusion-protein constructs be optimally cloned in order to achieve high expression and solubility?

- How can the capture system be further optimized and stabilized under specific buffer conditions required for the stability of respective target complexes?

Many of these questions can be answered by applying the capture system to complexes studied in the Marlovits lab e.g. of Type 1 secretion system. This can prove especially useful as it would show the behavior of the system when applied to various different targets.

I hope that these experiments help to understand the possibilities of applying the LaG-16-G4S-LaG-2 tandem nanobody to purify complex of low abundance and bypass the difficulties that are currently met in the respective studies. While it is a potent capture system, it is still too early in its developmental cycle and has yet to be proven as a firm replacement of more conventional methods. Although not all initial goals were met with success, this study should hopefully shed a light on the potential of this system.

6. ABSTRACT

Nanobodies are an unconventional type of antibodies. Also known as heavy-chain antibodies, they are found in the sera of *Camelidae* and are devoid of the L-chain polypeptide, while also lacking the first constant domain - CH1. These properties, as well as their relatively smaller size allow them to be well adapted for accessing binding sites of difficult targets. Nanobodies can be readily produced in bacterial expression systems and subsequently purified via immobilized metal affinity chromatography via respective affinity tags. Futhermore, nanobodies possess a number of qualities: they are a renewable source, allow their economic production, are small in size, stable and are soluble in aqueous solutions, show the ability to refold, are specific and possess high affinity for only one target patch. These properties allow researchers to apply nanobodies in many different fields such as life science research, diagnostics and therapeutics.

A prime example of a versatile nanobody is the tandem nanobody LaG-16-G4S-LaG-2, developed by the Rout lab, as this group has shown its high specificity and affinity for GFP. The strong affinity of the tandem nanobody for this marker molecule would allow its utilization in the purification of low-abundance proteins or protein complexes. The Marlovits lab is interested in these types of targets for structural studies. Consequenctly, the main research focus of the Marlovits group – the Type 3 secretion system needle complex from *S. enterica* – represents an excellent model case to test and optimize the new means of affinity purification.

Here I present experiments aimed at rendering the LaG-16-G4S-LaG-2-GFP protein capture system and its subsequent combination with the FimGt-DsF binding complex accessible for future structural studies. Furthermore, I show how the system can be improved upon in the future. I am certain that the experiments at hand will help to illuminate the crucial steps of the novel protein purification approach and can lead to a better understanding of the possible applications of nanobodies as protein capturing agents.

7. ZUSAMMENFASSUNG

Bei „nanobodies“ handelt es sich um eine spezielle Art von Antikörpern. Sie bestehen ausschließlich aus einem bestimmten Subkomplex üblicher Antikörper, der sogenannten „heavy chain“, und sind im Blutserum von *Camelidae* zu finden. Charakteristisch für nanobodies ist das Fehlen des L-Ketten Polypeptids und der ersten konstanten Domäne, CH1. Aufgrund dessen sind sie besonders klein und besitzen die Fähigkeit sich an Substrate anzupassen, welche schwer zugängliche Bindestellen besitzen. Nanobodies können in bakteriellen Expressionssystemen produziert und anschließend mit Immobilisierter-Metallionen-Affinitätschromatographie (IMAC) über einen an die Nanobodies gelinkten Affinitätstag aufgereinigt werden. Nanobodies besitzen folgende Qualitäten: sie sind regenerierbar, kommerziell produzierbar, besitzen eine geringe Größe, sind stabil und wasserlöslich, besitzen die Fähigkeit zur Rückfaltung nach Denaturierung und binden spezifisch und mit hoher Affinität nur ein gewünschtes Substrat. Aufgrund dieser Eigenschaften werden nanobodies in so unterschiedlichen Bereichen wie der biotechnologischen Forschung, der Diagnostik und in der Therapie verwendet.

Einer der bekanntesten nanobodies ist der vom Rout Labor etablierte tandem nanobody LaG-16-G4S-LaG-2, welcher eine hohe Spezifität und Affinität für GFP besitzt. Der nanobody ermöglicht folglich die Aufreinigung von Proteinen oder Proteinkomplexen die nur in sehr geringen Mengen in den untersuchten Zellen vorkommen und mit dem Fluoreszentprotein getagt wurden. Die Marlovits Gruppe ist besonders am Typ 3 Sekretionssystem von *S. enterica* interessiert, im Speziellen an strukturellen Analysen. Es handelt sich dabei um ein elegantes Modell-System, welches erlaubt, die neuen Möglichkeiten zur Proteinaufreinigung mit Hilfe von nanobodies auszutesten und zu optimieren.

In dieser Masterarbeit werden die Experimente zur Etablierung des LaG-16-G4S-LaG-2-GFP „protein capture system“ in der Kombination mit dem FimGt-DsF Bindungskomplex in der

Proteinreinigung präsentiert. Zudem wird besprochen, wie das System in Zukunft optimiert werden kann. Es wird auf essentielle Schritte im Zuge der neuartigen Proteinaufreinigung eingegangen, welche zu einem besseren Verständnis der möglichen Anwendung von nanobodies als „protein capture agents“ führen.

8. SUPPLEMENTARY

8.1. Plasmid maps

Here are listed the plasmids including the cloned genes for expression of the individual components in this study.

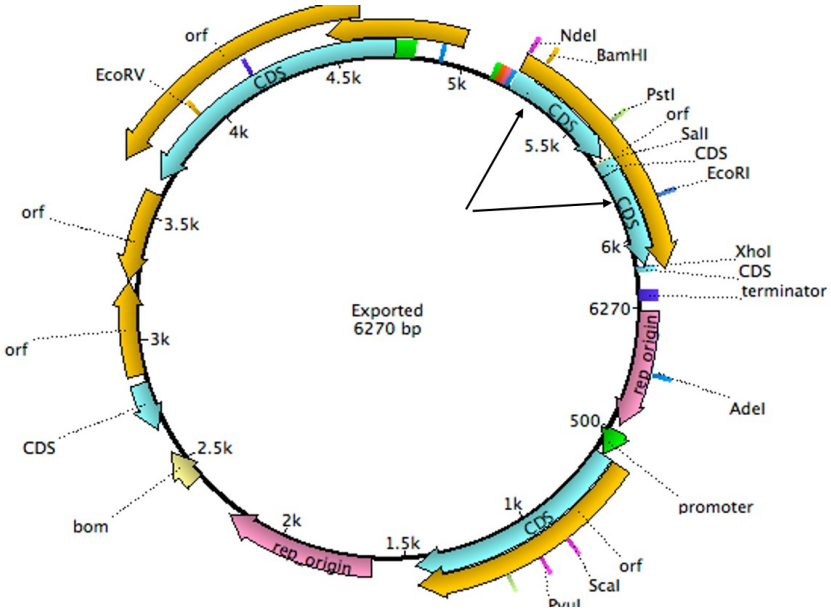

Figure P.1 Plasmid map of the LaG-16-G4S-LaG-2 tandem nanobody.

The coding sequence of the L16L2 tandem nanobody has been highlighted with arrows. (Plasmid was supplied by the Rout lab)

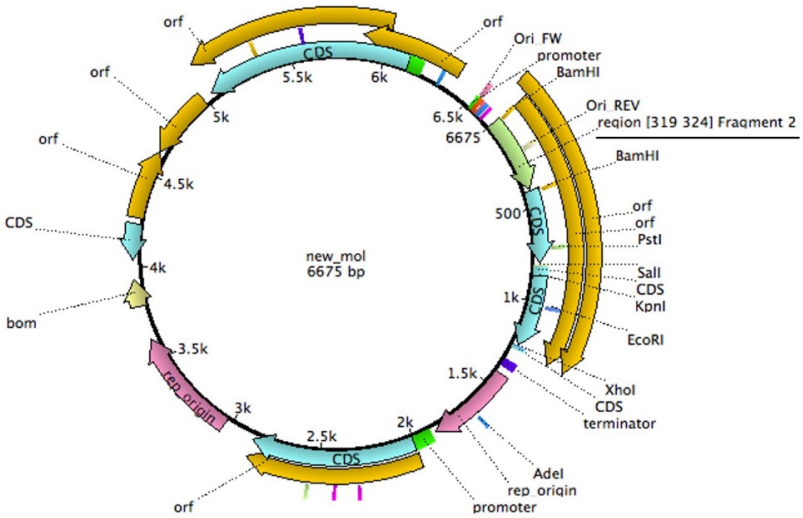

Figure P.2 Plasmid map of the LaG-16-G4S-LaG-2 tandem nanobody with the FimGt insert at an N-terminal position to the L16 monobody sequence.

The coding sequence for the FimGt insert has been underlined.

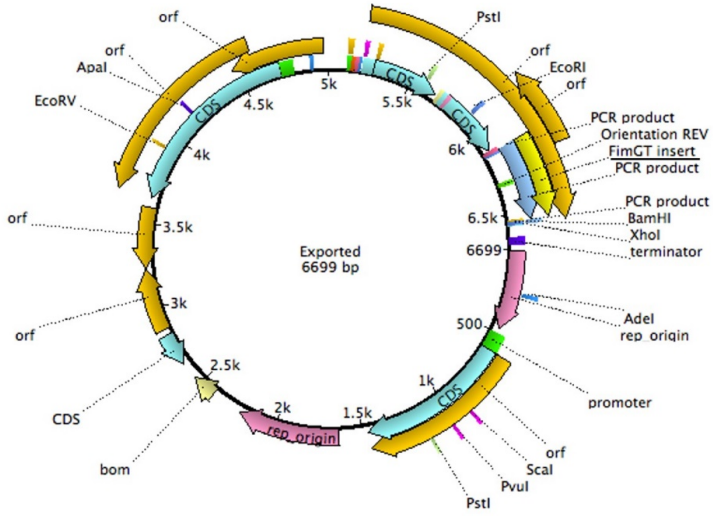

Figure P.3 Plasmid map of the LaG-16-G4S-LaG-2 tandem nanobody with the FimGt insert at a C-terminal position to the L16 monobody sequence.

The coding sequence for the FimGt insert has been underlined.

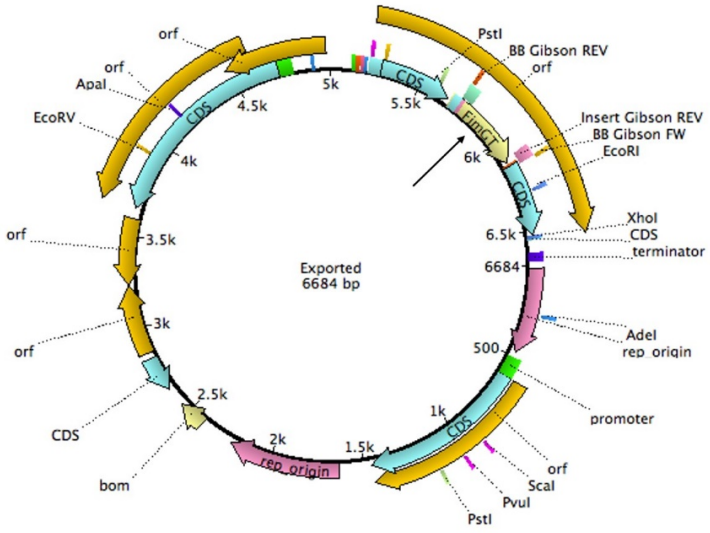

Figure P.4 Plasmid map of the LaG-16-G4S-LaG-2 tandem nanobody with the FimGt insert at an N-terminal position to the L2 monobody sequence.

The coding sequence for the FimGt insert has been indicated by an arrow.

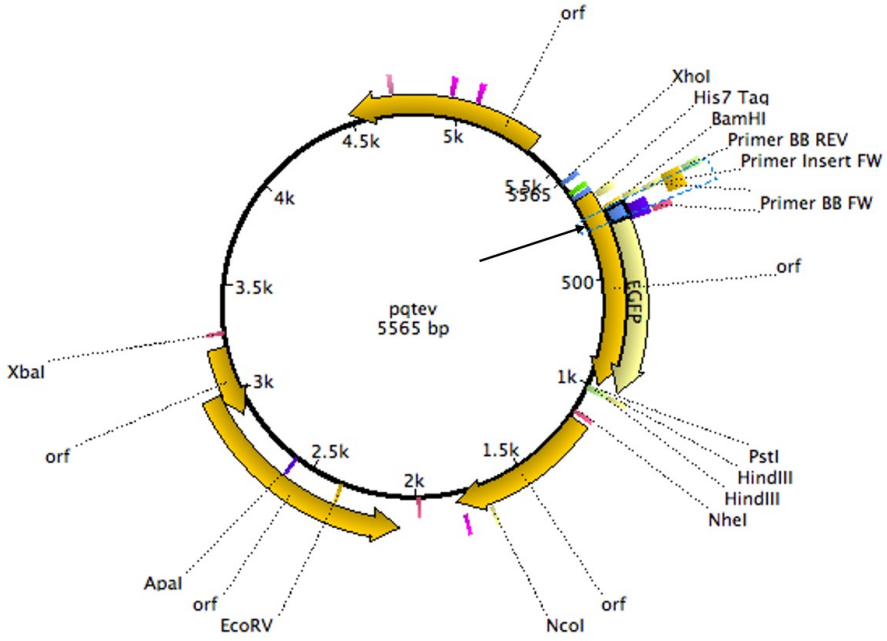

Figure P.5 Plasmid map of eGFP with the DsF insert at an N-terminal position to the protein sequence.

The coding sequence for the DsF insert has been indicated by an arrow.

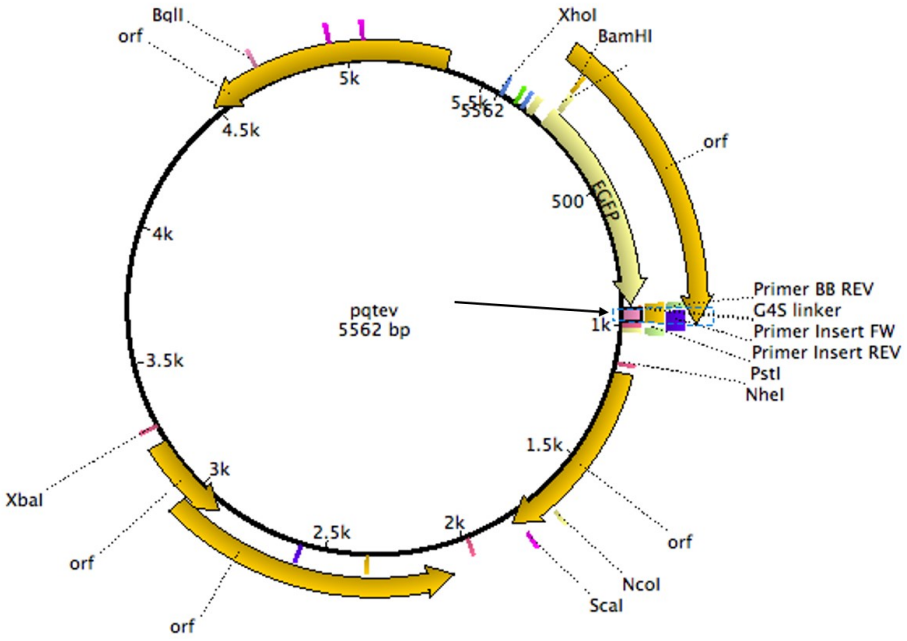

Figure P.6 Plasmid map of eGFP with the DsF insert at a C-terminal position to the protein sequence.

The coding sequence for the DsF insert has been indicated by an arrow.

8.2. Nucleotide sequence and genomic location of genes used in this study

- Tandem nanobody in plasmid (reverse complement)

tgccggccacgatgcgtccggcgtagaggatcgagatctcgatcccgcgaaattaatacgactcactataggggaattgt
gagcggataacaattcccctctagaaataattttgtttaactttaagaaggagatatacataTGAAATACTTATTGCCTA
CGGCAGCCGCTGGATTGTTATTACTCGCGGCCCAGCCGGCCATGGCTGGATCCATGGCCCAGGTTCAGTTGGTGGAATCT
GGCGGACGTCTGGTTCAGGCTGGTGATAGTTTACGCCTGTCCTGTGCTGCAAGCGGTCGTACCTTCTCTACTAGCGCGAT
GGCGTGGTTTCGTCAAGCGCCGGGTCGGGAACGCGAGTTTGTGGCAGCGATTACGTGGACCGTGGGTAACACGATTCTGG
GCGATTCGGTGAAAGGCCGCTTTACCATCAGCCGTGATCGTGCCAAGAATACCGTGGATCTCCAGATGGACAATCTGGAA
CCCGAAGATACGGCGGTCTATTACTGCAGTGCCCGTTCACGTGGGTATGTGCTGTCAGTTTTGCGCAGCGTTGATTCCTA
CGATTATTGGGGTCAGGGCACCCAAGTCACCGTATCAGGCCCTGGGTCGACAGGCGGAGGTGGATCTGGCGGAGGTGGAT
CGGGCGGAGGTGGATCAGGTACCATGGCTCAAGTCCAACTGGTGGAGAGTGGAGGTGGCTTAGTCCAGGCTGGCGGCAGC
CTTCGCCTGTCTTGCGCAGCATCTGGTCGCACGTTTAGCAACTATGCCATGGGCTGGTTTCGCCAGGCACCGGGCAAAGA
ACGGGAATTCGTTGCCGCGATCTCATGGACAGGGGTAAGCACCTATTATGCGGACTCCGTGAAAGGGCGCTTCACCATTT
CCCGCGATAACGATAAGAATACCGTATATGTGCAGATGAACAGTCTGATTCCGGAAGATACGGCCATCTACTATTGTGCA
GCCGTTCGTGCCCGCTCGTTCTCGGATACTTATTCGCGCGTTAATGAGTACGACTATTGGGGTCAAGGAACTCAGGTAAC
GGTCCTCGAGCACCACCACCACCACCACTGAGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCA
CCGCTGAGCAAtaactagcataaccccttggggcctctaaacgggtcttgaggggtt

- FimG shortened (from structure)

GCCAAACCGTGTACGGTTTCCACCACCAATGCCACGGTTGATCTCGGCGATCTTTATTCTTTCAGTCTTATGTCTGCCGG
GGCGGCATCGGCCTGGCATGATGTTGCGCTTGAGTTGACTAATTGTCCGGTGGGAACGTCGAGGGTCACTGCCAGCTTCA
GCGGGGCAGCCGACAGTACCGGATATTATAAAAACCAGGGGACCGCGCAAAACATCCAGTTAGAGCTACAGGATGACAGT
GGCAACACATTGAATACTGGCGCAACCAAAACAGTTCAGGTGGATGATTCCTCACAATCAGCGCACTTCCCGTTACAGGT
CAGAGCATTGACAGTAAATGGCGGAGCCACTCAGGGAACCATTCAGGCAGTGATTAGCATCACCTATACCTACAGC

- FimGt Variant 1

tgccggccacgatgcgtccggcgtagaggatcgagatctcgatcccgcgaaattaatacgactcactatagggaattgt
gagcggataacaattcccctctagaaataattttgtttaactttaagaaggagatatacataTGAAATACTTATTGCCTA
CGGCAGCCGCTGGATTGTTATTACTCGCGGCCCAGCCGGCCATGGCTGGATCCATGGCCAAACCGTGTACGGTTTCCACC
ACCAATGCCACGGTTGATCTCGGCGATCTTTATTCTTTCAGTCTTATGTCTGCCGGGGCGGCATCGGCCTGGCATGATGT
TGCGCTTGAGTTGACTAATTGTCCGGTGGGAACGTCGAGGGTCACTGCCAGCTTCAGCGGGGCAGCCGACAGTACCGGAT
ATTATAAAAACCAGGGGACCGCGCAAAACATCCAGTTAGAGCTACAGGATGACAGTGGCAACACATTGAATACTGGCGCA
ACCAAAACAGTTCAGGTGGATGATTCCTCACAATCAGCGCACTTCCCGTTACAGGTCAGAGCATTGACAGTAAATGGCGG
AGCCACTCAGGGAACCATTCAGGCAGTGATTAGCATCACCTATACCTACAGCGGTACCGGCGGAGGTGGATCTGGCGGAG
GTGGATCTGGTACCGCCCAGGTTCAGTTGGTGGAATCTGGCGGACGTCTGGTTCAGGCTGGTGATAGTTTACGCCTGTCC
TGTGCTGCAAGCGGTCGTACCTTCTCTACTAGCGCGATGGCGTGGTTTCGTCAAGCGCCGGGTCGGGAACGCGAGTTTGT
GGCAGCGATTACGTGGACCGTGGGTAACACGATTCTGGGCGATTCGGTGAAAGGCCGCTTTACCATCAGCCGTGATCGTG
CCAAGAATACCGTGGATCTCCAGATGGACAATCTGGAACCCGAAGATACGGCGGTCTATTACTGCAGTGCCCGTTCACGT
GGGTATGTGCTGTCAGTTTTGCGCAGCGTTGATTCCTACGATTATTGGGGTCAGGGCACCCAAGTCACCGTATCAGGCCC
TGGGTCGACAGGCGGAGGTGGATCTGGCGGAGGTGGATCGGGCGGAGGTGGATCAGGTACCATGGCTCAAGTCCAACTGG
TGGAGAGTGGAGGTGGCTTAGTCCAGGCTGGCGGCAGCCTTCGCCTGTCTTGCGCAGCATCTGGTCGCACGTTTAGCAAC
TATGCCATGGGCTGGTTTCGCCAGGCACCGGGCAAAGAACGGGAATTCGTTGCCGCGATCTCATGGACAGGGGTAAGCAC
CTATTATGCGGACTCCGTGAAAGGGCGCTTCACCATTTCCCGCGATAACGATAAGAATACCGTATATGTGCAGATGAACA
GTCTGATTCCGGAAGATACGGCCATCTACTATTGTGCAGCCGTTCGTGCCCGCTCGTTCTCGGATACTTATTCGCGCGTT
AATGAGTACGACTATTGGGGTCAAGGAACTCAGGTAACGGTCCTCGAGCACCACCACCACCACCACTGAGATCCGGCTGC
TAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAAtaactagcataaccccttggggcctctaaac
gggtcttgaggggtt

Forward primer- CGGCCCAGCCGGCCATGGCTGGATCCATGGCCAAACCGTGTACGGTTTCCAC
Reverse primer-
CAACTGAACCTGGGCGGTACCAGATCCACCTCCGCCAGATCCACCTCCGCCGGTACCGCTGTAGGTATAGGTGATGC

- FimGt Variant 2

tgccggccacgatgcgtccggcgtagaggatcgagatctcgatcccgcgaaattaatacgactcactatagggaattgt
gagcggataacaattcccctctagaaataattttgtttaactttaagaaggagatatacataTGAAATACTTATTGCCTA
CGGCAGCCGCTGGATTGTTATTACTCGCGGCCCAGCCGGCCATGGCTGGATCCATGGCCCAGGTTCAGTTGGTGGAATCT
GGCGGACGTCTGGTTCAGGCTGGTGATAGTTTACGCCTGTCCTGTGCTGCAAGCGGTCGTACCTTCTCTACTAGCGCGAT
GGCGTGGTTTCGTCAAGCGCCGGGTCGGGAACGCGAGTTTGTGGCAGCGATTACGTGGACCGTGGGTAACACGATTCTGG
GCGATTCGGTGAAAGGCCGCTTTACCATCAGCCGTGATCGTGCCAAGAATACCGTGGATCTCCAGATGGACAATCTGGAA
CCCGAAGATACGGCGGTCTATTACTGCAGTGCCCGTTCACGTGGGTATGTGCTGTCAGTTTTGCGCAGCGTTGATTCCTA
CGATTATTGGGGTCAGGGCACCCAAGTCACCGTATCAGGCCCTGGGTCGACAGGCGGAGGTGGATCTGGCGGAGGTGGAT
CGGGCGGAGGTGGATCAGGTACCATGGCTCAAGTCCAACTGGTGGAGAGTGGAGGTGGCTTAGTCCAGGCTGGCGGCAGC
CTTCGCCTGTCTTGCGCAGCATCTGGTCGCACGTTTAGCAACTATGCCATGGGCTGGTTTCGCCAGGCACCGGGCAAAGA
ACGGGAATTCGTTGCCGCGATCTCATGGACAGGGGTAAGCACCTATTATGCGGACTCCGTGAAAGGGCGCTTCACCATTT
CCCGCGATAACGATAAGAATACCGTATATGTGCAGATGAACAGTCTGATTCCGGAAGATACGGCCATCTACTATTGTGCA
GCCGTTCGTGCCCGCTCGTTCTCGGATACTTATTCGCGCGTTAATGAGTACGACTATTGGGGTCAAGGAACTCAGGTAAC
GGTC
GGTACCGGCGGAGGTGGATCTGGCGGAGGTGGATCTGGTACCGCCAAACCGTGTACGGTTTCCACCACCAATGCCACGGT
TGATCTCGGCGATCTTTATTCTTTCAGTCTTATGTCTGCCGGGGCGGCATCGGCCTGGCATGATGTTGCGCTTGAGTTGA
CTAATTGTCCGGTGGGAACGTCGAGGGTCACTGCCAGCTTCAGCGGGGCAGCCGACAGTACCGGATATTATAAAAACCAG
GGGACCGCGCAAAACATCCAGTTAGAGCTACAGGATGACAGTGGCAACACATTGAATACTGGCGCAACCAAAACAGTTCA
GGTGGATGATTCCTCACAATCAGCGCACTTCCCGTTACAGGTCAGAGCATTGACAGTAAATGGCGGAGCCACTCAGGGAA
CCATTCAGGCAGTGATTAGCATCACCTATACCTACAGCCTCGAGCACCACCACCACCACCACTGAGATCCGGCTGCTAAC
AAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAAtaactagcataaccccttggggcctctaaacgggt
cttgaggggtt

Forward-
GGTAACGGTCGGTACCGGCGGAGGTGGATCTGGCGGAGGTGGATCTGGTACCGCCAAACCGTGTACGGTTTCC

Reverse- GTGGTGGTGGTGGTGGTGCTCGAGGCTGTAGGTATAGGTGATGC

•FimGt Variant 3

tgccggccacgatgcgtccggcgtagaggatcgagatctcgatcccgcgaaattaatacgactcactataggggaattgt
gagcggataacaattcccctctagaaataattttgtttaactttaagaaggagatatacataTGAAATACTTATTGCCTA
CGGCAGCCGCTGGATTGTTATTACTCGCGGCCCAGCCGGCCATGGCTGGATCCATGGGTACCGGCGGAGGTGGATCTGGC
GGAGGTGGATCTGGTACCGCCCAGGTTCAGTTGGTGGAATCTGGCGGACGTCTGGTTCAGGCTGGTGATAGTTTACGCCT
GTCCTGTGCTGCAAGCGGTCGTACCTTCTCTACTAGCGCGATGGCGTGGTTTCGTCAAGCGCCGGGTCGGGAACGCGAGT
TTGTGGCAGCGATTACGTGGACCGTGGGTAACACGATTCTGGGCGATTCGGTGAAAGGCCGCTTTACCATCAGCCGTGAT
CGTGCCAAGAATACCGTGGATCTCCAGATGGACAATCTGGAACCCGAAGATACGGCGGTCTATTACTGCAGTGCCCGTTC
ACGTGGGTATGTGCTGTCAGTTTTGCGCAGCGTTGATTCCTACGATTATTGGGGTCAGGGCACCCAAGTCACCGTATCAG
GCCCTGGGTCGACAGGCGGAGGTGGATCTGGCGGAGGTGGATCTGCCAAACCGTGTACGGTTTCCACCACCAATGCCACG
GTTGATCTCGGCGATCTTTATTCTTTCAGTCTTATGTCTGCCGGGGCGGCATCGGCCTGGCATGATGTTGCGCTTGAGTT
GACTAATTGTCCGGTGGGAACGTCGAGGGTCACTGCCAGCTTCAGCGGGGCAGCCGACAGTACCGGATATTATAAAAACC
AGGGGACCGCGCAAAACATCCAGTTAGAGCTACAGGATGACAGTGGCAACACATTGAATACTGGCGCAACCAAAACAGTT
CAGGTGGATGATTCCTCACAATCAGCGCACTTCCCGTTACAGGTCAGAGCATTGACAGTAAATGGCGGAGCCACTCAGGG
AACCATTCAGGCAGTGATTAGCATCACCTATACCTACAGCGGCGGAGGTGGATCTGGCGGAGGTGGATCTGGTACCATGG
CTCAAGTCCAACTGGTGGAGAGTGGAGGTGGCTTAGTCCAGGCTGGCGGCAGCCTTCGCCTGTCTTGCGCAGCATCTGGT
CGCACGTTTAGCAACTATGCCATGGGCTGGTTTCGCCAGGCACCGGGCAAAGAACGGGAATTCGTTGCCGCGATCTCATG
GACAGGGGTAAGCACCTATTATGCGGACTCCGTGAAAGGGCGCTTCACCATTTCCCGCGATAACGATAAGAATACCGTAT
ATGTGCAGATGAACAGTCTGATTCCGGAAGATACGGCCATCTACTATTGTGCAGCCGTTCGTGCCCGCTCGTTCTCGGAT
ACTTATTCGCGCGTTAATGAGTACGACTATTGGGGTCAAGGAACTCAGGTAACGGTCCTCGAGCACCACCACCACCACCA
CTGAGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAAtaactagcataacccc
ttggggcctctaaacgggtcttgaggggtt

Forward primer-
CACCCAAGTCACCGTATCAGGCCCTGGGTCGACAGGCGGAGGTGGATCTGGCGGAGGTGGATCTGCCAAACCGTGTACGG
TTTCCAC

Reverse primer-
CCAGTTGGACTTGAGCCATGGTACCAGATCCACCTCCGCCAGATCCACCTCCGCCGCTGTAGGTATAGGTGATGCTAATC
ACTG

• eGFP sequence

CTCGAGAAATCATAAAAAATTTATTTGCTTTGTGAGCGGATAACAATTATAATAGATTCAATTGTGAGCGGATAACAATT
TCACACAGAATTCATTAAAGAGGAGAAATTAACTATGAAACATCACCATCACCATCACCATAGCGATTACGACATCCCCA
CTACTGAGAATCTTTATTTTCAGGGATCCGGGTCGACTGTTGATAGAGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTG
GTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCAC
CTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGA
CCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGC
TACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACAC
CCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACT
ACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATC
GAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAA
CCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGA
CCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAGGCCGCTCGACCTGCAGCCAAGCTTAATTAGCTGAG
CTTGGACTCCTGTTGATAGATCCAGTAATGACCTCAGAACTCCATCTGGATTTGTTCAGAACGCTCGGTTGCCGCCGGGC
GTTTTTTATTGGTGAGAATCCAAGCTAGCTTGGCGAGATTTTCAGGAGCTAAGGAAGCTAAAATGGAGAAAAAAATCACT
GGATATACCACCGTTGATATATCCCAATGGCATCGTAAAGAACATTTTGAGGCATTTCAGTCAGTTGCTCAATGTACCTA
TAACCAGACCGTTCAGCTGGATATTACGGCCTTTTTAAAGACCGTAAAGAAAAATAAGCACAAGTTTTATCCGGCCTTTA
TTCACATTCTTGCCCGCCTGATGAATGCTCATCCGGAATTTCGTATGGCAATGAAAGACGGTGAGCTGGTGATATGGGAT

AGTGTTCACCCTTGTTACACCGTTTTCCATGAGCAAACTGAAACGTTTTCATCGCTCTGGAGTGAATACCACGACGATTT
CCGGCAGTTTCTACACATATATTCGCAAGATGTGGCGTGTTACGGTGAAAACCTGGCCTATTTCCCTAAAGGGTTTATTG
AGAATATGTTTTTCGTCTCAGCCAATCCCTGGGTGAGTTTCACCAGTTTTGATTTAAACGTGGCCAATATGGACAACTTC
TTCGCCCCCGTTTTCACCATGGGCAAATATTATACGCAAGGCGACAAGGTGCTGATGCCGCTGGCGATTCAGGTTCATCA
TGCCGTTTGTGATGGCTTCCATGTCGGCAGAATGCTTAATGAATTACAACAGTACTGCGATGAGTGGCAGGGCGGGGCGT
AATTTTTTTAAGGCAGTTATTGGTGCCCTTAAACGCCTGGGGTAATGACTCTCTAGCTTGAGGCATCAAATAAAACGAAA
GGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCCTGAGTAGGACAAATCCGCCCT
CTAGATTACGTGCAGTCGATGATAAGCTGTCAAACATGAGAATTGTGCCTAATGAGTGAGCTAACTTACATTAATTGCGT
TGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGC
GGTTTGCGTATTGGGCGCCAGGGTGGTTTTTCTTTTCACCAGTGAGACGGGCAACAGCTGATTGCCCTTCACCGCCTGGC
CCTGAGAGAGTTGCAGCAAGCGGTCCACGCTGGTTTGCCCCAGCAGGCGAAAATCCTGTTTGATGGTGGTTAACGGCGGG
ATATAACATGAGCTGTCTTCGGTATCGTCGTATCCCACTACCGAGATATCCGCACCAACGCGCAGCCCGGACTCGGTAAT
GGCGCGCATTGCGCCCAGCGCCATCTGATCGTTGGCAACCAGCATCGCAGTGGGAACGATGCCCTCATTCAGCATTTGCA
TGGTTTGTTGAAAACCGGACATGGCACTCCAGTCGCCTTCCCGTTCCGCTATCGGCTGAATTTGATTGCGAGTGAGATAT
TTATGCCAGCCAGCCAGACGCAGACGCGCCGAGACAGAACTTAATGGGCCCGCTAACAGCGCGATTTGCTGGTGACCCAA
TGCGACCAGATGCTCCACGCCCAGTCGCGTACCGTCTTCATGGGAGAAAATAATACTGTTGATGGGTGTCTGGTCAGAGA
CATCAAGAAATAACGCCGGAACATTAGTGCAGGCAGCTTCCACAGCAATGGCATCCTGGTCATCCAGCGGATAGTTAATG
ATCAGCCCACTGACGCGTTGCGCGAGAAGATTGTGCACCGCCGCTTTACAGGCTTCGACGCCGCTTCGTTCTACCATCGA
CACCACCACGCTGGCACCCAGTTGATCGGCGCGAGATTTAATCGCCGCGACAATTTGCGACGGCGCGTGCAGGGCCAGAC
TGGAGGTGGCAACGCCAATCAGCAACGACTGTTTGCCCGCCAGTTGTTGTGCCACGCGGTTGGGAATGTAATTCAGCTCC
GCCATCGCCGCTTCCACTTTTTCCCGCGTTTTCGCAGAAACGTGGCTGGCCTGGTTCACCACGCGGGAAACGGTCTGATA
AGAGACACCGGCATACTCTGCGACATCGTATAACGTTACTGGTTTCACATTCACCACCCTGAATTGACTCTCTTCCGGGC
GCTATCATGCCATACCGCGAAAGGTTTTGCACCATTCGATGGTGTCGGAATTTCGGGCAGCGTTGGGTCCTGGCCACGGG
TGCGCATGATCTAGAGCTGCCTCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTC
ACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCGC
AGCCATGACCCAGTCACGTAGCGATAGCGGAGTGTATACTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGT
GCACCACATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGCGCTCTTCCGCTTCCTCGCTCA
CTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAA
TCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGC
GTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGAC
TATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTG
TCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCG
CTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCA
ACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCT
ACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGT
TACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGC
AGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAA
AACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTT
TAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGA
TCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGC
CCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGC
CGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTT
CGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCA
TTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCC
TCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCA
TGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGT
TGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTC
TTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTT
CAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCG
ACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGG
ATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTCT
AAGAAACCATTATTATCATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCCTTTCGTCTTCAC

- DSF peptide sequence

Gcggatagcacgattactatccgcggctatgtcagggataacggc

- eGFP-DsF Variant 1

GAGCGGATAACAATTTCACACAGAATTCATTAAAGAGGAGAAATTAACTATGAAACATCACCATCACCATCACCATAGCG
ATTACGACATCCCCACTACTGAGAATCTTTATTTTCAGGGATCGGAgcggatagcacgattactatccgcggctatgtca
gggataacggcGGCGGAGGTGGATCTGGCGGAGGTGGATCTGGAGGACGGGTCGACTGTTGATAGAGTGAGCAAGGGCGA
GGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTA

Insert REV primer-
CTATCAACAGTCGACCCGTCCTCCAGATCCACCTCCGCCAGATCCACCTCCGCCgccgttatccctgacatagccgcgga
tagtaatcgtgctatccgcTCCGATCCCTGAAAATAAAGA

Insert FW primer-
TCTTTATTTTCAGGGATCGGAgcggatagcacgattactatccgcggctatgtcagggataacggcGGCGGAGGTGGATC
TGGCGGAGGTGGATCTGGAGGACGGGTCGACTGTTGATAG

- eGFP-DsF Variant 2

CACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGGGCGGAGGTGGATC
TGGCGGAGGTGGATCTGGAGGAgcggatagcacgattactatccgcggctatgtcagggataacggcTGAGGCCGCTCGA
CCTGCAGCCAAGCTTAATTAGCTGAGCTTGGACTCCTGTTGATAGATCCAGTAAT

Insert FW primer-
ATGGACGAGCTGTACAAGGGCGGAGGTGGATCTGGCGGAGGTGGATCTGGAGGAgcggatagcacgattactatccgcgg
ctatgtcagggataacggcTGAGGCCGCTCGACCTGCAGC

Insert REV primer-
GCTGCAGGTCGAGCGGCCTCAgccgttatccctgacatagccgcggatagtaatcgtgctatccgcTCCTCCAGATCCAC
CTCCGCCAGATCCACCTCCGCCCTTGTACAGCTCGTCCAT

8.3. Western Blots

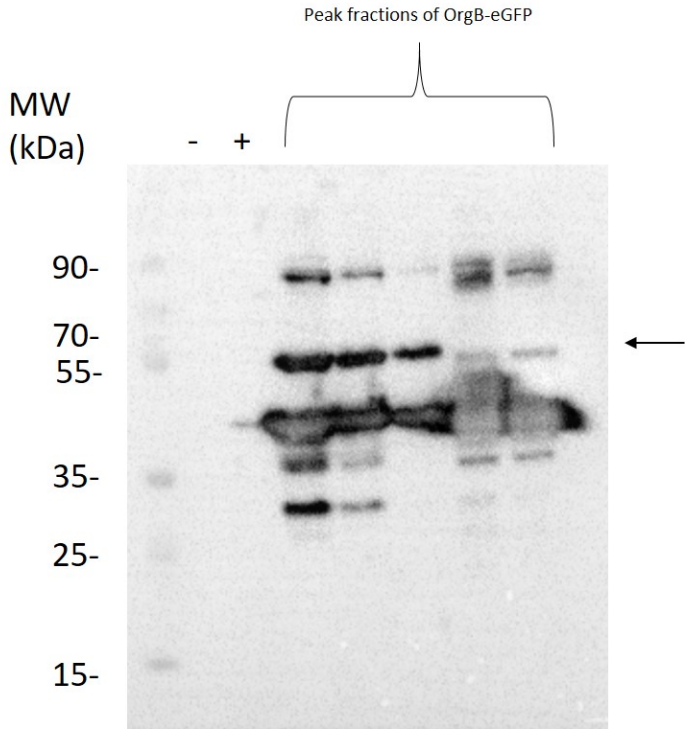

Figure S.1: Western blotting raw data of OrgB-eGFP complex pull down using the L16L2 tandem nanobody. (anti-GFP)

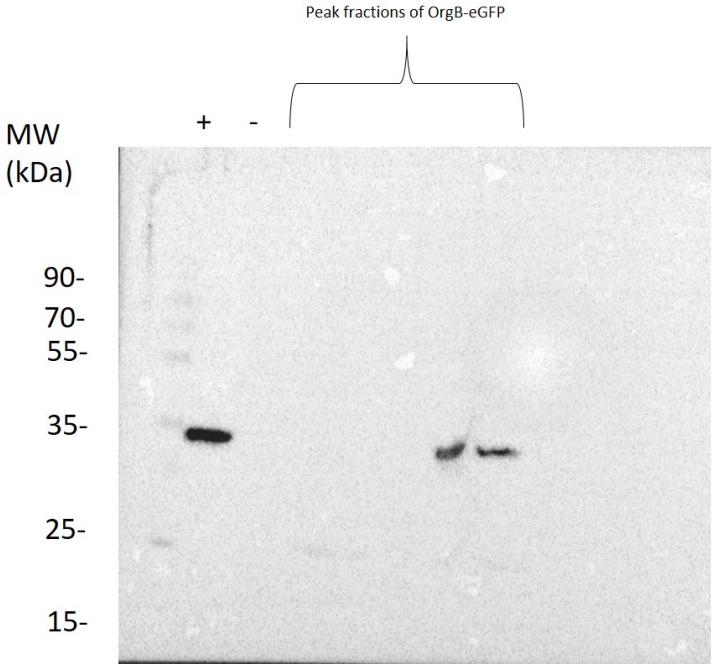

Figure S.2: Western blotting raw data of OrgB-eGFP complex pull down using the L16L2 tandem nanobody. (anti-His)

8.4. Standard curves for Superdex 75 and Superdex 200

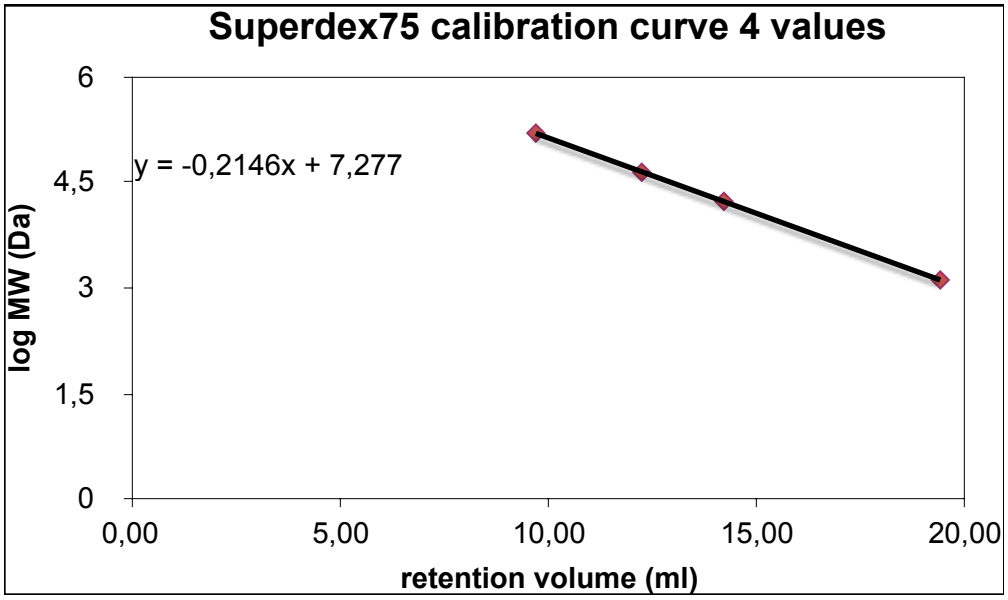

Figure C1: Standard curve for calculating protein mass from a Superdex 75 column run.

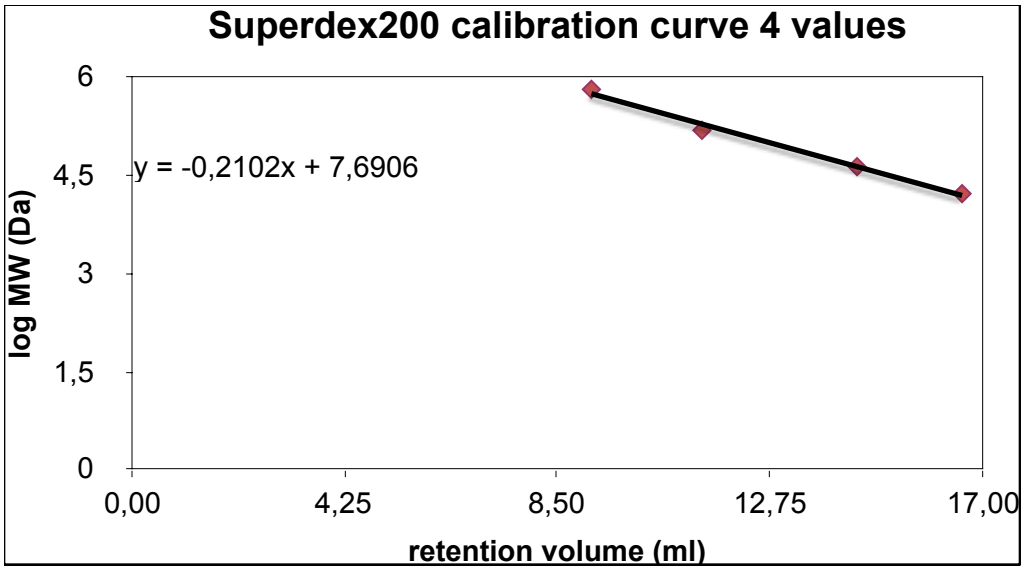

Figure C2: Standard curve for calculating protein mass from a Superdex 200 column.

8. LITERATURE

1. S., E., S., T., O, U. & MD, Ozge E. Onur, MD, Arzu Denizbasi, MD, Haldun Akoglu, MD, Cigdem Ozpolat, MD, Ebru Akoglu, M. *Protein purification - Principles, High Resolutio Methods, and Applications*. (2011). doi:10.1073/pnas.0703993104

2. Urassa DP(1), Nystrom L, Carlstedt A, Msamanga GI, L. G. Pubmed_Result. *Interaction between disease concepts and the organisation of health care* **5,** 134–7 (2009).

3. Levinson, W. E. *Review of Medical Immunology*. (2010).

4. Fridy, P. C. *et al*. A robust pipeline for rapid production of versatile nanobody repertoires. *Nat Methods* **11,** 1253–1260 (2014).

5. Muyldermans, S. Nanobodies: Natural Single-Domain Antibodies. 1–23 (2013). doi: 10.1146/annurev-biochem-063011-092449

6. Ries, J., Kaplan, C., Platonova, E., Eghlidi, H. & Ewers, H. A simple, versatile method for GFP-based super-resolution microscopy via nanobodies. *Nat. Methods* **9,** 582–584 (2012).

7. Mcmaster, M. L., Kristinsson, S. Y., Turesson, I., Bjorkholm, M. & Landgren, O. Engineering of recombinant crystallization chaperones. *Clin. Lymphoma* **9,** 19–22 (2010).

8. Nanobodies and their potential applications.

9. Rothbauer, U. *et al*. A Versatile Nanotrap for Biochemical and Functional Studies with Fluorescent Fusion Proteins. *Mol. Cell. Proteomics* **7,** 282–289 (2007).

10. Adams, H. *et al*. Specific immuno capturing of the staphylococcal superantigen toxic-shock syndrome toxin-1 in plasma. *Biotechnol. Bioeng.* **104,** 143–151 (2009).

11. Klooster, R. *et al*. Improved anti-IgG and HSA affinity ligands: Clinical application of VHH antibody technology. *J. Immunol. Methods* **324,** 1–12 (2007).

12. Klooster, R. *et al*. Selection and characterization of KDEL-specific VHH antibody fragments and their application in the study of ER resident protein expression. *J.*

Immunol. Methods **342,** 1–12 (2009).

13. Baranova, E. *et al*. SbsB structure and lattice reconstruction unveil Ca2+ triggered S-layer assembly. *Nature* **487,** 119–122 (2012).

14. Tereshko, V. *et al*. Toward chaperone-assisted crystallography: protein engineering enhancement of crystal packing and X-ray phasing capabilities of a camelid single-domain antibody (VHH) scaffold. *Protein Sci.* **17,** 1175–87 (2008).

15. Loris, R. *et al*. Crystal structure of the intrinsically flexible addiction antidote MazE. *J. Biol. Chem.* **278,** 28252–28257 (2003).

16. Rasmussen, S. G. F. *et al*. Tructure of a Nanobody-Stabilized Active State of the Beta2 Adrenoceptor. *Nature* **469,** 175–80 (2011).

17. Sukhanova, A. *et al*. Oriented conjugates of single-domain antibodies and quantum dots: Toward a new generation of ultrasmall diagnostic nanoprobes. *Nanomedicine Nanotechnology, Biol. Med.* **8,** 516–525 (2012).

18. Drew, D., Froderberg, L., Baars, L. & de Gier, J. W. Assembly and overexpression of membrane proteins in Escherichia coli. *Biochim.Biophys.Acta.* **1610,** 3–10 (2003).

19. Seddon, A. M., Curnow, P. & Booth, P. J. Membrane proteins, lipids and detergents: Not just a soap opera. *Biochimica et Biophysica Acta - Biomembranes* (2004). doi:10.1016/j.bbamem.2004.04.011

20. Hamers-Casterman, C. *et al*. Naturally occurring antibodies devoid of light chains. *Nature* **363,** 446–8 (1993).

21. Alvarez-Rueda, N. *et al*. Generation of llama single-domain antibodies against methotrexate, a prototypical hapten. *Mol. Immunol.* **44,** 1680–1690 (2007).

22. Conrath, K. E., Lauwereys, M., Galleni, M. & Fre, J. ^{N_L} -Lactamase Inhibitors Derived from Single-Domain Antibody Fragments Elicited in the Camelidae. **45,** 2807–2812 (2001).

23. Rothbauer, U. *et al*. Targeting and tracing antigens in live cells with fluorescent nanobodies. *Nat Meth* **3,** 887–889 (2006).

24. Arbabi Ghahroudi, M., Desmyter, A., Wyns, L., Hamers, R. & Muyldermans, S. Selection and identification of single domain antibody fragments from camel heavy-chain antibodies. *FEBS Lett*. **414,** 521–526 (1997).

25. Battistutta, R., Negro, A. & Zanotti, G. Crystal structure and refolding properties of the mutant F99S/M153T/V163A of the green fluorescent protein. *Proteins Struct. Funct. Genet*. **41,** 429–437 (2000).

26. Kirchhofer, A. *et al*. Modulation of protein properties in living cells using nanobodies. *Nat Struct Mol Biol* **17,** 133–138 (2010).

27. Arifuzzaman, M. *et al*. Large-scale identification of protein – protein interaction of Escherichia coli K-12. *Genome Res.* **16,** 686–691 (2006).

28. Gloeckner, C. J., Boldt, K., Schumacher, A., Roepman, R. & Ueffing, M. A novel tandem affinity purification strategy for the efficient isolation and characterisation of native protein complexes. *Proteomics* **7,** 4228–4234 (2007).

29. Nieba, L. *et al*. BIACORE Analysis of Histidine-Tagged Proteins Using a Chelating NTA Sensor Chip. *Anal. Biochem.* **252,** 217–228 (1997).

30. Atanasov, A. G. *et al*. Direct protein-protein interaction of 11ß-hydroxysteroid dehydrogenase type 1 and hexose-6-phosphate dehydrogenase in the endoplasmic reticulum lumen. *Biochim. Biophys. Acta - Mol. Cell Res.* **1783,** 1536–1543 (2008).

31. Einhauer, A. & Jungbauer, A. Affinity of the monoclonal antibody M1 directed against the FLAG peptide. *J. Chromatogr. A* **921,** 25–30 (2001).

32. Le Trong, I. *et al*. Structural Basis for Mechanical Force Regulation of the Adhesin FimH via Finger Trap-like?? Sheet Twisting. *Cell* **141,** 645–655 (2010).

33. Puorger, C. *et al*. Infinite Kinetic Stability against Dissociation of Supramolecular Protein Complexes through Donor Strand Complementation. *Structure* **16,** 631–642 (2008).

34. Puorger, C., Vetsch, M., Wider, G. & Glockshuber, R. Structure, Folding and Stability of FimA, the Main Structural Subunit of Type 1 Pili from Uropathogenic Escherichia coli Strains. *J. Mol. Biol.* **412,** 520–535 (2011).

35. Choudhury, D. *et al.* X-ray Structure of the FimC-FimH Chaperone-Adhesin Complex from Uropathogenic Escherichia coli *Science (80-.).* **285,** 1061 LP-1066 (1999).

36. Gossert, A. D. *et al.* NMR Structure of the Escherichia coli Type 1 Pilus Subunit FimF and Its Interactions with Other Pilus Subunits. *J. Mol. Biol.* **375,** 752–763 (2008).

37. Giese, C., Zosel, F., Puorger, C. & Glockshuber, R. The most stable protein-ligand complex: Applications for one-step affinity purification and identification of protein assemblies. *Angew. Chemie - Int. Ed.* **51,** 4474–4478 (2012).

38. Lutz, H. U. *Naturally Occuring Antibodies (NAbs).* (2012).

39. Feilmeier, B. J., Iseminger, G., Schroeder, D., Webber, H. & Phillips, G. J. Green fluorescent protein functions as a reporter for protein localization in Escherichia coli. *J. Bacteriol.* **182,** 4068–76 (2000).