Nanobody selection from a highly variable gene library using mRNA/cDNA display

Inaugural-Dissertation

To obtain the academic degree

Doctor rerum naturalium (Dr. rer. nat.)

Submitted to the Department of Biology, Chemistry, Pharmacy

Of Freie Universität Berlin

By

Raphael Möhrle

2023

The presented work was performed between July 2019 and June 2023 under supervision of Dr. phil. nat. Magdalena Schacherl at the Institute of Medical Physics and Biophysics of Charité – Universitätsmedizin Berlin.

1st reviewer: Dr. Magdalena Schacherl

2nd reviewer: Prof. Dr. Markus Wahl

Date of defense: 19.09.2023

Declaration of independence

Herewith I certify that I have prepared and written my thesis independently and that I have not used any sources and aids other than those indicated by me. This dissertation has not yet been presented to any other examination authority in the same or a similar form and has not yet been published.

Berlin, 29.06.2023

_____

Raphael Möhrle

Acknowledgements

First and foremost, I would like to thank my supervisor Dr. Magdalena Schacherl for giving me the opportunity to work on this exciting but challenging project, and for supporting me with my experiments and this thesis.

I want to thank Prof. Dr. Christian Spahn for hosting our group in his institute, and all my IMPB colleagues for a nice and welcoming atmosphere in the lab. Especially, I want to thank our TAs Helena Seibel, Birgit Schroeer, and Christine Gotthold. Helena supported me in cloning and protein expressions, Birgit Schroeer helped me with RNA and yeast handling, did not get exhausted from my endless questions, and was always open for some sourdough talk, and Christine Gotthold prepared the stable TREx293 cell line.

I want to thank Dr. Yollete Guillén Schlippe for sharing her expertise in mRNA display with me, and Dr. Andrea Schmidt for her support in crystallization.

Prof. Dr. Markus Wahl agreed to review my thesis, which I appreciate greatly.

During my project, many students joined our small group and all of them, Kasia, Rosa, Aurora, Sophie, and Lara, made working in the lab more enjoyable and interesting.

My parents Sabine and Manfred and my brother Dominik have always supported me during my time at university and during my PhD thesis, have always kept faith in me, and always welcomed me back home with a cold Schenkelbier, and I would not have been able to even start this project without their help.

I would not have been able to finish this thesis without the constant support of my fiancée Marie, who kept me well-fed and in the best possible mood especially during the final and most challenging weeks of this project.

Table of Contents

Abstract

Besides their function in vertebrate immune systems, antibodies are useful tools in immunohistochemistry, structural biology, diagnostics, and as therapeutics. With a molecular weight of 15 kDa, single-domain antibodies or Nanobodies are the smallest functional antibody units. Nanobodies are derived from heavy chain antibodies, which can be found in camelid species like dromedaries, camels, llamas, or alpacas. Compared to canonical antibodies, Nanobodies provide several advantages like cheap and fast production in bacterial hosts, enhanced stability, low-immunogenicity, and high tissue penetration and are promising molecules for research and as therapeutics.

Generation of new Nanobodies is traditionally performed by immunization of camelids with the target antigen. In the past, several Nanobody selection methods have been established that avoid animal immunization entirely. Additionally, these animal-free selection methods, called phage display, yeast display, or mRNA/cDNA display, provide benefits like the ability to generate Nanobodies against non-immunogenic or toxic targets. Starting point of every animal-free selection is a Nanobody gene library. The design of these Nanobody gene libraries greatly influences the success of a selection. Larger, more variable gene libraries increase the chance of selecting high-affinity binders, while smaller, more conservative library designs usually deliver more well-behaved proteins.

For this study, we designed a highly variable Nanobody gene library that uses only few invariant positions, based on multiple sequence alignments of published Nanobody sequences. This Nanobody gene library provides a huge sequence space, which enables selection of Nanobodies for a variety of target antigens. The mRNA/cDNA display selection method is performed entirely *ex vivo* and allows the use of very large gene libraries. Therefore, this selection method is the right fit for our highly variable Nanobody gene library and was chosen for this project.

The proteins retinal guanylate cyclase 1 (retGC1) and its main regulator guanylate cyclase activating protein 1 (GCAP1) are localized in the outer segments of photoreceptor cells and play important roles in the visual phototransduction cycle. At low intracellular Ca^{2+} concentrations, GCAP1 strongly activates retGC1, which then synthesizes cGMP. This is essential to restore the photoreceptor dark state. Multiple mutations in the retGC1 and GCAP1 genes are associated with rare retinopathies like Leber's congenital amaurosis or autosomal dominant cone-rod dystrophy. RetGC1- and GCAP1-binding Nanobodies could provide useful tools to gain a better understanding of the function of these two proteins and the mechanisms that influence retinopathies. To date, no retGC1 purification protocol has been published. In this study, we provide a solubilization and purification protocol for retGC1 after expression in a stable TREx293 cell line.

To improve the understanding of retGC1 regulation by GCAP1, we attempted structural characterization of different ion-bound GCAP1 variants by X-ray crystallography. Initial diffracting crystals could be obtained for GCAP1 wild type and the disease-related M26R variant.

Starting in late 2019, SARS-CoV2 caused the worldwide COVID19 pandemic. In early 2020, we joined the research on the corona virus by choosing the SARS-CoV2 main protease (M^{pro}) as a target antigen for Nanobody selection by mRNA/cDNA display. M^{pro} cleavage of viral polyproteins is essential for the formation of new virus particles and virus activities have been reduced by M^{pro} peptide inhibitors.

With the help of a commercial Nanobody gene library with a more traditional, less variable design, we were able to establish an mRNA/cDNA display Nanobody selection procedure that allows performing single selection cycles within four days. Optimization of the selection procedure and the implementation of several purification steps to maintain a high library quality enabled the Nanobody selection from highly variable gene libraries of up to 10^{13} Nb sequences.

Zusammenfassung

Neben ihrer Funktion im Immunsystem von Wirbeltieren sind Antikörper nützliche Werkzeuge in Immunhistochemie, Strukturbiologie, Diagnose und als Therapeutika. Mit einem Molekulargewicht von 15 kDa sind Einzeldomänantikörper, auch Nanobodies genannt, die kleinste funktionelle Antikörpereinheit. Nanobodies sind von Schwere-Ketten-Antikörpern abgeleitet, die in Cameliden wie Dromedaren, Kamelen, Lamas oder Alpakas vorkommen. Nanobodies bieten einige Vorteile gegenüber konventionellen Antikörpern wie die einfache, preisgünstige Produktion in Bakterien, hohe Stabilität, niedrige Immunogenität und Gewebedurchdringung, weshalb Nanobodies vielversprechende Moleküle für Forschung und als Therapeutika darstellen.

Neue Nanobodies werden gewöhnlich durch Immunisierung von Cameliden mit Zielantigenen generiert. Um diesen Immunisierungsschritt zu umgehen, wurden in der Vergangenheit einige tierfreie Methoden zur Selektionierung von Nanobodies etabliert. Zusätzlich bieten diese tierfreien Methoden, Hefe-, Phagen, oder mRNA/cDNA-Display, die Möglichkeit, Nanobodies gegen wenig immunogene oder toxische Zielantigene zu generieren. Ausgangsmaterial für jede tierfreie Nanobodyselektion ist eine Nanobody Genbibliothek. Das Design dieser Bibliotheken beeinflusst den Erfolg einer Selektion maßgeblich. Während größere, variablere Genbibliotheken größere Chancen bieten, hochaffine Binder zu generieren, liefern kleinere, weniger variable Bibliotheken gut gefaltete Proteine.

Für dieses Projekt wurde eine hochvariable Nanobody-Genbibliothek konzeptioniert, die nur wenige invariante Positionen beinhaltet, die aufgrund von Sequenzvergleichen publizierter Nanobodysequenzen festgelegt wurden. Diese Genbibliothek bietet einen großen Sequenzraum, was die Nanobodyselektionierung gegen eine Vielzahl von Zielantigenen ermöglicht. Die mRNA/cDNA-Display Selektionsmethode findet vollständig *ex vivo* statt, was das Benutzen sehr großer Genbibliotheken zulässt, weshalb diese Methode für dieses Projekt ausgewählt wurde.

Die Proteine retinale Guanylatzyklase 1 (retGC1) und ihr Regulator, das Guanylatzyklase aktivierende Protein 1 (GCAP1), sind in den äußeren Segmenten von Photorezeptorzellen lokalisiert und spielen eine wichtige Rolle im visuellen Phototransduktionszyklus. Bei niedrigen intrazellulären Ca^{2+}-Konzentrationen stimuliert GCAP1 die Aktivität von retGC1. Aktivierte retGC1 synthetisiert dann cGMP, was für die Wiederherstellung des Dunkelzustandes in Photorezeptorzellen essenziell ist. Einige Mutationen in den Genen für retGC1 und GCAP1 kommen in Verbindung mit seltenen Retinopathien wie der Leberschen kongenitalen Amaurose 1 oder der autosomal-dominanten Zapfen-Stäbchen-Dystrophie vor. Nanobodies, die retGC1 oder GCAP1 binden, könnten helfen, ein besseres Verständnis für die beiden Proteine und deren Einfluss in Retinopathien zu gewinnen. Deshalb wurden diese beiden Proteine als Zielantigene für die Selektionierung von Nanobodies ausgewählt.

Bisher wurde noch kein Protokoll für die Aufreinigung von retGC1 publiziert. In dieser Arbeit präsentieren wir ein Protokoll für die Solubilisierung und Aufreinigung der retGC1, nachdem diese in einer stabilen TREx293-Zelllinie exprimiert wurde.

Um einen besseren Einblick in die Regulation von retGC1 durch GCAP1 zu erlangen, wurde versucht, GCAP1 in verschiedenen ionengebundenen Varianten mittels der Röntgenkristallographie strukturell zu charakterisieren. Erste streuende Kristalle konnten für das GCAP1 Wildtypprotein und die M26R-Variante produziert werden.

Seit Ende 2019 verbreitete sich das SARS-CoV2 weltweit und verursachte die COVID19 Pandemie. Im Frühjahr 2020 haben wir im Zuge der Pandemieforschung die SARS-CoV2 Main Protease (M^{pro}) als weiteres Zielantigen ausgewählt. Die proteolytische Spaltung von Coronaviruspolyproteinen durch M^{pro} ist essenziell für die Produktion neuer Viruspartikel und Virusaktivität kann durch M^{pro}-Peptidinhibitoren reduziert werden.

Mithilfe einer kommerziellen Nanobody-Genbibliothek mit traditionellem, weniger variablen Design konnten wir ein mRNA/cDNA-Display Nanobodyselektionsprotokoll erarbeiten, das die Durchführung eines Selektionszyklus innerhalb von vier Tagen ermöglicht. Optimierung der Methode und die Implementierung mehrerer Aufreinigungsschritte, die zur Aufrechterhaltung einer hohen Qualität der Genbibliothek dienen, ermöglichen Nanobodyselektionen von hochvariablen Genbibliotheken mit bis zu 10^{13} Nanobodysequenzen.

1. Introduction

1.1 Immunoglobulins are powerful research tools

1.1.1   Canonical Abs and their functional fragments

Recognition and elimination of foreign substances is essential for the health of any organism. In vertebrates, the task of specific detection of foreign molecules, or antigens, is performed by proteins called immunoglobulins. These immunoglobulins are also known as antibodies (Ab) (Ehrlich, 1900) and are tasked with the function to detect foreign molecules with high affinity and specificity and mark them for degradation by cells of the adaptive immune system (Hazzard et al., 1968; Talmage, 1959).
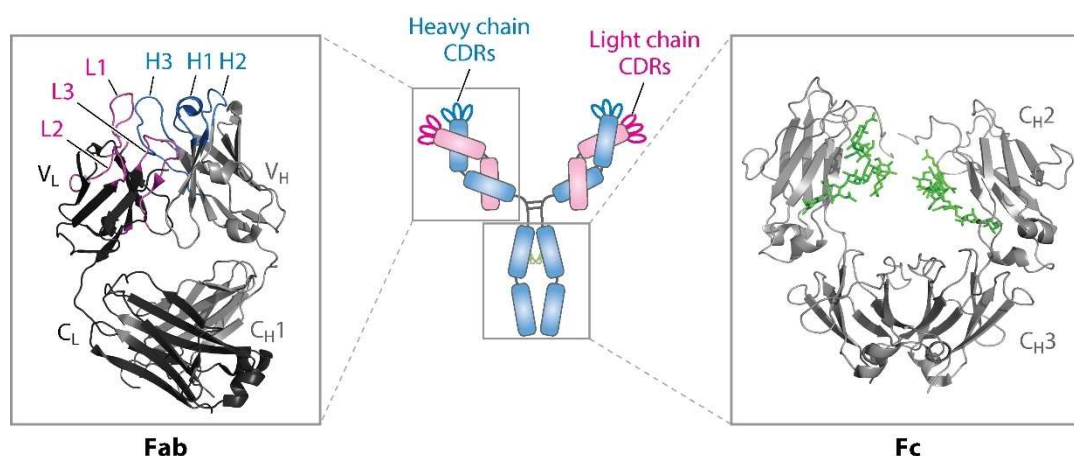
Figure 1.1 Domain organization of a canonical Ab. An immunoglobulin G (IgG) (middle) is comprised of two heavy chains (blue) and two light chains (pink) and the antigen-recognizing complementarity-determining regions CDRs visualized as loops. A cartoon representation of an antigen-binding fragment (left) highlights the CDR-loops of the heavy chain (H1, H2, H3 in blue) and of the light chain (L1, L2, L3 in pink) (left) and of the crystallizable fragment (Fc), where post-translational modifications are highlighted in green (right). The image was adapted from Tiller & Tessier, 2015.

Throughout all vertebrate species, most Abs show similar domain architecture. One canonical Ab consists of four protein chains, two heavy and two light chains, which are stabilized by disulfide bonds. Each heavy chain has a molecular weight (MW) of 50 kDa and each light chain of 25 kDa, which adds up to a MW of 150 kDa for the entire Ab. Abs show different regions of globular fold, arranged in a characteristic Y-shape (Figure 1.1, middle; Davies & Chacko, 1993). The region, where both heavy chains are connected is called crystallizable fragment (Fc) (Figure 1.1, right), which is well sequence-conserved and is essential for the formation of the characteristic Y-shape. The arms of the Y contain one heavy and one light chain and are called fragment antigen-binding (Fab) (Figure 1.1, left). On the top ends of a Fab, the complementarity-determining regions (CDR) are located. Each Fab contains six CDR loops, three on the variable domain of the heavy and three on the variable domain of the light

chain. These CDRs are crucial for antigen binding. The Ab surface, which interacts with its antigen is called paratope, while the interaction surface on the antigen site is called epitope. For their traditional function as part of the adaptive immune system, Abs are usually secreted and contain a multitude of post-translational modifications, such as glycosylation (Figure 1.1, right (green)).

The ability of Abs to detect their antigens with high specificity and affinity can be utilized in various ways. Therapeutic Abs can be used to treat a variety of diseases like cancer (Pillarisetti et al., 2020), migraine (Markham, 2018), or Alzheimer's disease (McDade et al., 2022). Besides the pharmaceutical use, Abs can be used to detect their antigens in methods like ELISA (Engvall & Perlmann, 1971), Western blot (Renart et al., 1979), or in affinity chromatography of proteins, e.g. over the DYKDDDDK- or Flag-tag (Munro & Pelham, 1984). If coupled with a fluorophore, Abs can also be used to visualize certain structures within a cell or tissue in immunohistochemistry. By removing the Fc region, 50 kDa Fab fragments are generated. Even smaller fragments can be generated from only the CDR-containing domains (uppermost rectangles in Figure 1.1) on a single 30 kDa protein chain, called single-chain variable fragment (scFv). Canonical Abs can be recombinantly produced in mammalian cell culture like Chinese hamster ovarian cells (CHO cells). Because scFvs contain fewer posttranslational modifications, they can be synthesized in bacteria like *Escherichia coli*, even with relatively high yield (Martineau et al., 1998).

In nature, novel Abs are generated by B lymphocytes, or B cells. Different B cells produce different Abs and only B cells with target-binding abilities Abs are maintained by the animal's immune system. Sequence variability of the many Abs is provided by different V (variable), D (diversity) and J (joining) exon cassettes, which are rearranged into the Abs' variable fragments (Brack et al., 1978; Max et al., 1979). Like this, a library of Abs presented by B cells can be generated. After being presented to the antigen, only target-interacting cells are maintained and propagated, while B cells that did not bind an antigen undergo apoptosis (Hasbold & Klaus, 1990). The Abs, displayed by the maintained B lymphocytes are affinity matured by somatic mutations in the following B cell generations (Valbuena et al., 1978), and Abs with even higher affinities can be generated.

These steps to generate novel Abs, namely sequence variability, selection by application of selective pressure, and affinity maturation can also be simulated in animal-free selection methods (McCafferty et al., 1990; Thompson et al., 1996). These methods enable the selection of Abs or Ab fragments against targets that would not yield high-affinity binders by animal immunization. Targets with low immunogenicity would probably not trigger an immune response in the animal and no high-affinity Abs would be generated. Animal-free selection methods also enable the generation of Abs against toxic target antigens (Indrawattana et al., 2010; Prado et al., 2016). Different Ab DNA libraries can be designed as starting material for the selections, enabling the use of human immunoglobulin scaffolds. When applied to patients, these Abs show reduced immunogenicity (Silacci et al., 2005). Animal-free

Ab selection methods (Chapter 1.2) are crucial for the generation of Abs or Ab fragments against such targets.

1.1.2   Heavy chain Abs and their fragments

In the late 1980s, Abs with a different domain architecture were discovered in dromedary serum (Hamers-Casterman et al., 1993). This type of Ab consists only of two heavy chains and was therefore named heavy chain Ab (hcAb). Those heavy chain Abs were later found in several camelid species like llamas, alpacas, and camels. Additionally, hcAbs have been discovered in cartilaginous fish (Greenberg et al., 1995). Those hcAbs are of slightly different architecture and are thought to have developed independently and are called IgNAR (immunoglobulin new antigen receptor).

The smallest antigen-binding immunoglobulin unit can be constructed from the variable fragment of hcAbs and is called $V_H H$ (**V**ariable domain of a **H**eavy chain of **H**eavy chain Ab), single domain Ab (sdAb) or Nanobody (Nb). With a MW of roughly 15 kDa, Nbs are only a tenth of the size of a canonical Ab and are even smaller than the Ab fragments Fab or scFv.
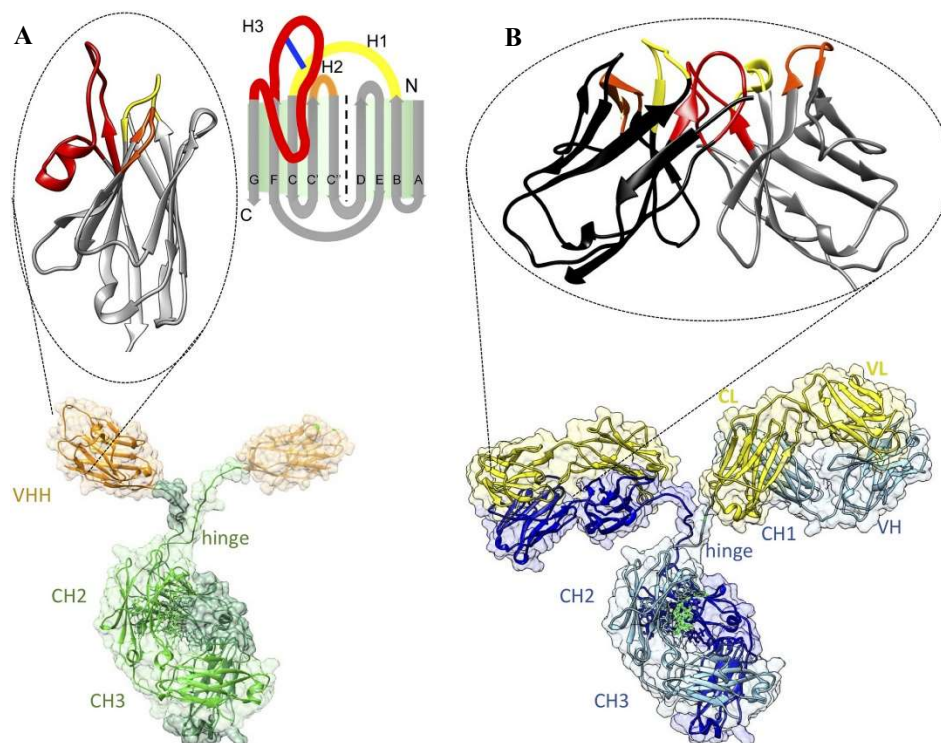
Figure 1.2 Structural comparison of a heavy chain Ab: and a canonical Ab. A: Heavy chain antibodies contain a constant region with the domains CH1 and CH2. CH2 is connected to the variable domain (VHH) via a hinge region. Focus on a VHH highlights the CDR loops in yellow (CDR1), orange (CDR2) and red (CDR3). Protein topology of a VHH (insert) shows that a VHH contains nine β-strands that are connected by the CDR loops (H1, H2, H3) and five other loops. The β-strands are labeled alphabetically from N- to C-terminus. B: A canonical Ab in composed of two heavy and two light chains, which are organized in constant domains 1, 2, and 3 on the heavy chains (CH1, CH2, CH3), one constant domain per heavy (CH1) and per light (CL1) chain, and a variable domain per heavy (VH) and light (VL) chain. VL and VH are essential for target recognition and are highlighted with colored CDR complementary to the VHH. The image was adapted from Muyldermans, 2021).

While the constant regions of hcAbs are quite similar to those of canonical Abs (Figure 1.2), the variable domains show some differences. A canonical Ab presents a paratope that spans over the variable domains of both heavy and light chains, a hcAb presents a single variable domain (VHH), which are present in a slightly elongated, rugby ball-shaped structure (Figure 1.2).

Nbs are organized in four framework regions (FR), which influence the protein fold, and three complementarity-determining regions (CDR). The CDR loops provide the majority of interactions with the antigen and are crucial for antigen detection.

The interaction between a Nb and its antigen is usually dominated by CDR3. Different CDR3 lengths and sequences can influence the binding geometry. The interaction geometry between Nb and antigen can be classified into three binding modes (Figure 1.3). Especially with short CDR3s, like a GFP-binder with a 6 amino acid CDR3 (Kirchhofer et al., 2009), the Nb usually presents a concave surface. Medium sized CDR3s like the 12 amino acid CDR3 of a Nb that recognizes the β_2-adrenoceptor (Rasmussen et al., 2011) can protrude into cavities presented by the antigen. Even longer CDR3s can fold over, which leads to a convex binding surface. With this binding mode, Nbs can recognize clefts on the antigen surface, like a lysozyme binding hcAb with a 16 amino acid CDR3 (De Genst et al., 2006).
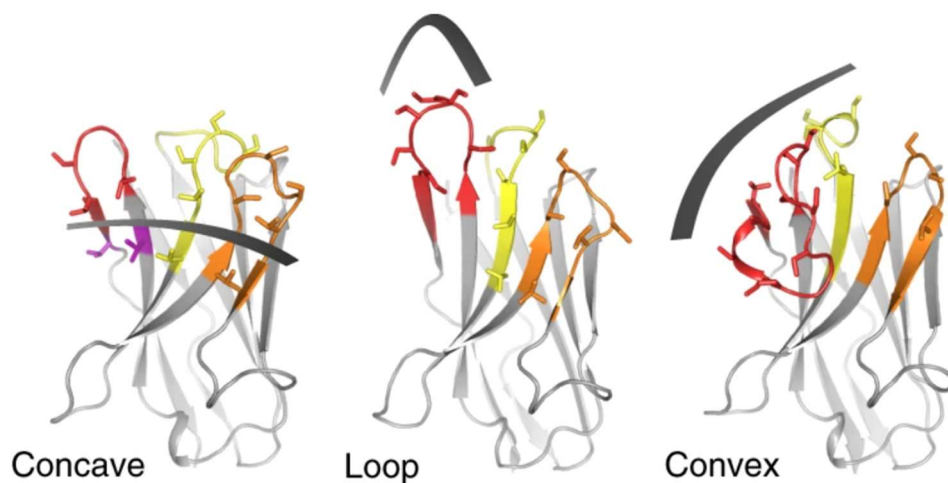
Figure 1.3 Graphical representation of different binding modes of Nbs. Positioning of the antigen surface (grey line) is shown on cartoon representations of different Nbs with highlighted CDR1 (yellow), CDR2 (orange) and CDR3 (red). The described binding modes were named concave, loop, and convex following of the Nb paratope shape. The image was adapted from Zimmermann et al., 2018.

The small size and few posttranslational modifications of Nbs provide a multitude of advantages compared to canonical Abs or their Fab and scFv fragments. While canonical Abs need to be synthesized in mammalian cell culture to generate properly folded and modified immunoglobulins, Nbs can be synthesized in bacterial hosts like *E. coli* with high yield (Arbabi Ghahroudi et al., 1997). This reduces production costs greatly and requires less laborious and less complicated expression and purification procedures. To provide Nbs with an essential disulfide bond, they can be synthesized in the cytoplasm of *E. coli* strains like Origami2 (Novagen), where mutations in the enzymes thioredoxin reductase and

glutathione reductase enable cytosolic disulfide bond formation (Stewart et al., 1998). Alternatively, strains like wk6 (Salema et al., 2013) can be used for Nb synthesis and co-translational translocation to the periplasm (Baneyx, 1999). This provides the Nb proteins with an oxidizing milieu, which enables disulfide bond formation.

Many characterized Nbs with longer CDR3s usually bind clefts or cavities of their antigens (De Genst et al., 2006), which often leads to very high affinities. In some cases, Nbs reach affinities even in the sub-picomolar range (Schoof et al., 2020). The cavity recognition allows some Nbs to bind directly to the catalytic center of enzymes and inhibit their activities (Lauwereys et al., 1998). While Nbs excel in cleft or cavity recognition, there are only a few cases known, where they can recognize linear epitopes, the anti-ALFA (Götzke et al., 2019) or anti-Spot-tag (Braun et al., 2016) Nbs being examples of rare exceptions. The cavity or cleft recognition of most Nbs usually leads to the recognition of different epitopes than by canonical Abs, and Nbs and Abs that recognize the same antigen usually do not compete.

Because of their stable, well-conserved fold, Nbs usually exhibit beneficial properties like enhanced physical and chemical stability. Some Nbs have been found that show denaturing temperatures (T_m) exceeding 95 °C (Ma et al., 2021) or the ability to withstand high concentrations of denaturing chemicals like guanidinium hydrochloride (Dumoulin et al., 2002). Additionally, many Nbs show the ability to re-fold after denaturation (Pérez et al., 2001).

Because of their high similarity to canonical Ab V_H, Nbs often exhibit low immunogenicity in humans, which enables their use as therapeutics (Kijanka et al., 2015). By mutating single amino acids or by grafting Nb CDRs onto a human V_H scaffold, immunogenicity can be reduced even further (Vincke et al., 2009). Nbs have been shown to demonstrate high tissue penetration and some possess the ability to clear the blood-brain barrier (Muruganandam et al., 2002). The ability to bind their antigens with high affinity and specificity, combined with the mentioned beneficial properties make Nbs powerful pharmaceutical tools. For example, Nbs specifically targeting the epidermal growth factor receptor (EGFR), which is upregulated in tumor cells, have been shown to inhibit cancer growth (Roovers et al., 2011). The first EMA- and FDA-approved Nb therapeutic, Caplacizumab, specifically targets the so-called von Willebrand factor and is used to treat acquired thrombotic thrombocytopenic purpura (Scully et al., 2019).

Camelids being the only mammals naturally producing heavy chain Abs means that the generation of novel hcAbs or Nbs requires immunization of rather large animals like llamas, alpacas, camels, or dromedaries. While novel canonical Abs can be generated by immunization of small mammals like mice or rabbits, generation of novel Nbs can be quite laborious and expensive, without even considering the ethical aspects of potentially harming or stressing these animals by frequently introducing foreign proteins to their bodies. In the past decades, several animal-free methods have been developed that

enable the generation of novel Ab fragments or Nbs (Hanes & Plückthun, 1997; Roberts & Szostak, 1997). These methods will be discussed in chapter 1.2.

1.2 *In vitro* selection methods – directed protein evolution

Directed protein evolution is mainly influenced by three factors – sequence variability, selective pressure, and maturation. During antibody selection in vertebrates, sequence variability is provided by a multitude of V, D, and J exon cassettes, which are rearranged to generate novel Ab genes (Tonegawa, 1983). Selective pressure is introduced by displaying Abs on the surface of B lymphocytes (Nussenzweig et al., 1987) and only propagating active, antigen-binding B-cells. Over generations of B cells, Abs are (affinity) matured (Papavasiliou et al., 1997).

To obtain highly affine Abs from animal-free selection methods, these three principles must by applied. Sequence variability is provided by a (Nb) gene library containing many different (Nb) sequences. These gene libraries can have different properties or sizes depending on the used *in vitro* selection method. Gene libraries can be constructed from a pool of Ab genes harvested from animals before (naïve library) or after immunization (immune library), or by an *in silico* design (synthetic library) (Muyldermans, 2021), which is usually based on a known Ab/Nb scaffold and sequence variability in the CDRs (McMahon et al., 2018). More detailed insights into the construction of different Nb libraries is provided in section 1.3. To apply selective pressure on the Abs/Nbs in the libraries, they are translated into proteins and used for selection based on their affinity to the target antigen, which is usually immobilized on a solid surface. This process is called bio-panning (Parmley & Smith, 1988). By using washing procedures of varying stringency, the selective pressure can be adjusted. *In vitro* selections are usually performed in several cycles. PCRs can introduce mutations in every single cycle, which resembles affinity maturation.

All *in vitro* selection methods share one common principle, which is called genotype-phenotype coupling. After affinity-based selections, the protein sequences must be retrieved. However, it is impossible to determine protein sequence out of a library of different proteins. To solve this problem, the phenotypic information of the Ab i.e., the Ab protein, must be coupled to its genetic information i.e., Ab DNA or RNA sequence (Nemoto et al., 1997).

1.2.1   Cell-mediated display methods

There are several strategies to provide this genotype-phenotype coupling. The most prominent and commonly used one is phage-display. Here, the Ab/Nb gene is integrated into the gen of a viral coat protein (Devlin et al., 1990; Smith, 1985). In every selection cycle, a bacterial host (*Escherichia coli*) is infected with these phages. *E. coli* cells then produce virus particles that display the Ab/Nb protein on

their surface. The displayed proteins can then be used for affinity-based selection. In the end, the phage DNA is isolated, and the Ab genes can be sequenced.

A similar approach is used for cell-surface display methods (yeast display or bacterial surface display) (Boder & Wittrup, 1997). Here, the cells are not infected with virus, but are transformed with plasmids containing Ab/Nb genes. In yeast display, the Abs or Nbs are fused to the C-terminus of the Aga2p mating adhesion receptor (Boder & Wittrup, 1997), which is located on the surface of the yeast cell wall. Before selection, expression of the Nb is introduced and the Nb is displayed on the cell surface (Boder & Wittrup, 1997). These methods avoid the infection step to amplify the Nb-containing phages but require a transformation step of the used cells before the first selection cycle. Transformation or infection steps limit the feasible DNA library size to roughly 10^9 sequences for the methods of phage and cell surface display (Muyldermans, 2021).

1.2.2 Cell-free display methods

Cell-free display methods like ribosome display or mRNA display use a different approach. Here, the Ab/Nb protein chain is not linked to a cell or a virus particle, but directly to its own mRNA. The most common cell-free selection methods are ribosome display and mRNA display. In ribosome display, the Nb/Ab mRNA lacks a Stop-codon but contains a 3'-stem loop. The stem-loop stalls the ribosome, and the lack of a Stop-codon prevents proper translation termination (Hanes & Plückthun, 1997). This leads to the formation of a Nb protein-mRNA-ribosome hybrid molecule.
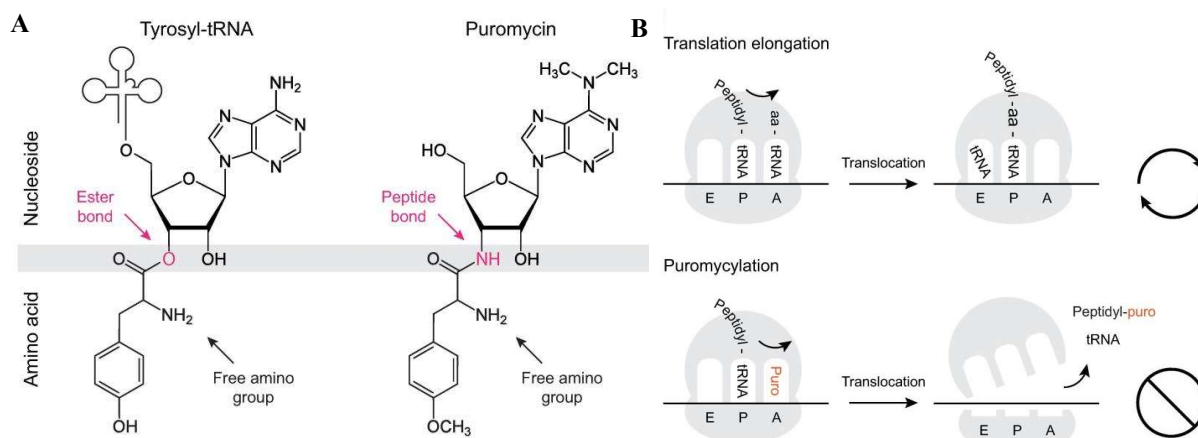
Figure 1.4 Puromycylation of nascent protein chains in the ribosome. A: Comparison of the molecule structures of a Tyrosyl-tRNA (left) and the antibiotic puromycin (right), with nucleoside parts (top) and amino acid parts (bottom). At the position of the ester bond in the tyrosyl-tRNA, puromycin has a peptide bond. B: Schematic representation of translation elongation (top) and Puromycylation of the nascent protein chain (bottom). Puromycin can enter the ribosomal A site. Then, the nascent protein chain (Peptidyl-) is covalently linked to Puromycin. Because bond cleavage and translocation are prevented, further translation elongation is prevented and a puromycylated polypeptide is released from the ribosome. The image was adapted from Semenkov et al., 1992.

In mRNA display, the Nb mRNA is modified before translation. The antibiotic puromycin is introduced to the 3'-end of the mRNA. Puromycin is structurally similar to the 3'-adenosine of an aminoacylated

Tyrosyl-tRNA (Figure 1.4 A; (Roberts et al., 1959)). During translation, puromycin can enter the ribosomal A site, and the ribosome translocates the nascent protein chain onto the puromycin. Puromycin contains a peptide bond instead of an ester bond in the aminoacylated tRNA. Because the ribosome is not able to cleave this peptide bond, the nascent protein chain stays attached to puromycin (Figure 1.4 B; Semenkov et al., 1992). Since puromycin has been covalently linked to the mRNA before translation, the Nb phenotype (nascent protein chain) is covalently coupled with the Nb genotype (mRNA). In contrast to ribosome display, the protein-mRNA hybrid molecule can be released from the ribosome (R. W. Roberts & Szostak, 1997).

In mRNA display, there are different approaches to modify the mRNA with puromycin. In an enzymatic approach, a DNA splint is used to mediate generation of a single nucleic acid strand containing the Nb mRNA and the puromycin molecule (Roberts & Szostak, 1997; Figure 1.5 A). Usually, an oligo or poly(dA) sequence and a spacer in front of the puromycin are added as well. The oligo(dA) sequence enables oligo(dT) affinity purification, and the spacer facilitates the puromycin to enter the ribosome. Because of the enzymatic reaction, this strategy requires careful purification steps before translation. To avoid the enzymatic reaction, a 3'-puromycin labeled and 5'-psoralen labeled oligonucleotide can be used. Psoralen can intercalate in nucleic acid base pairs and, after UV-irradiation at 365 nm, form a covalent bond between the strands (Kurz et al., 2000; Figure 1.5 B). This reaction generates a single-stranded nucleic acid, but with a backbone imperfection at the position where psoralen was inserted. Because this reaction lacks the use of an enzyme, in theory the nucleic acids can simply be precipitated and reconstituted in a buffer of choice, resulting in higher yields of puromycin-labeled mRNA, even if the efficiency of the UV-crosslinking is lower than that of the splint ligation.
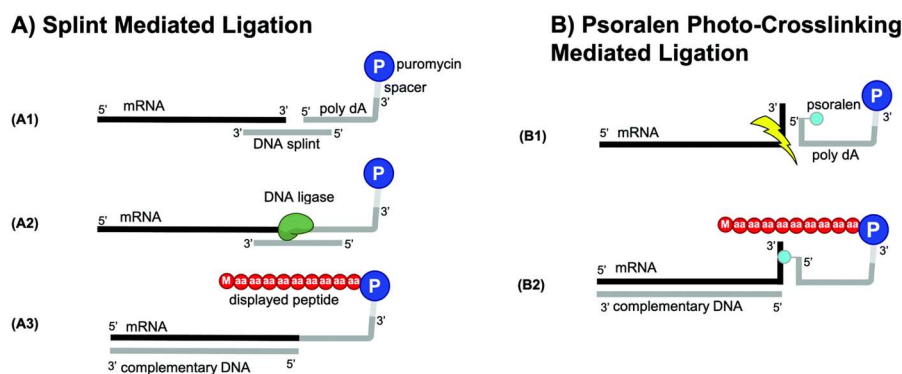
Figure 1.5 Schematic representations of different methods to incorporate puromycin in mRNA. A: A DNA splint, which base-pairs the 3'-end of the mRNA and the 5'-end of the puromycin-containing (P in blue circle) oligonucleotide, is used to connect mRNA and oligo on a single strand after ligation. After translation and subsequent reverse transcription, an mRNA/cDNA-protein hybrid is generated (The nascent protein chain is depicted as red aa circles). B: The 5'-end of the Puromycin oligo base-pairs the 3'-end of the mRNA. A psoralen molecule (small, turquoise circle) located at the 5'-end of the oligo covalently connects mRNA and oligo after UV-irradiation. The image was adapted from Kamalinia et al., 2021.

For our Nb selection method by mRNA/cDNA display, we chose an approach that uses a 5'-psoralen and 3'-puromycin labeled, linear oligonucleotide, which was covalently linked to the Nb mRNA by UV-crosslink, similar to the strategy in Figure 1.5 B. This method combines a non-enzymatic reaction with a relatively simple oligo design and a quickly established crosslink reaction.

1.2.3 Comparison of different animal-free selection methods

One main difference between the mentioned display methods is the fusion partner of the Nanobody polypeptide. Ranging from relatively small polymeric molecules, like mRNA, to large cells, like those of *Saccharomyces cerevisiae*, the size-difference of these molecules/cells obviously influences antigen-recognition and the selection greatly (Newton et al., 2020; Figure 1.6).
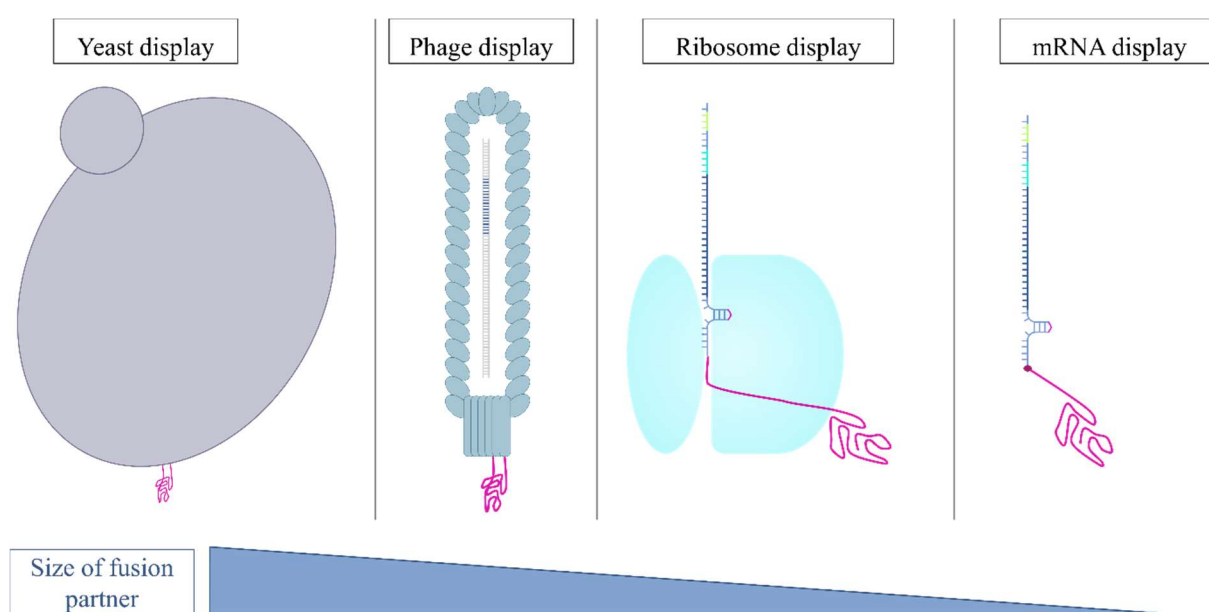
Figure 1.6 Size-comparison of different display methods. Schematic representation of yeast surface display, phage display, ribosome display, and mRNA display of a Nb protein (pink), not drawn to scale. Nbs or other proteins or peptides can be displayed on large cells, like yeast, smaller structures like bacteriophages, or on a single mRNA molecule via ribosome display or directly in mRNA display.

In vitro selection methods like ribosome display or mRNA display avoid any transformation or infection step and are entirely performed cell-free. This increases the maximum library size from ~10^9 different Nb sequences to ~10^{14} (Roberts & Szostak, 1997). While ribosome and mRNA display are restricted only by the reaction volumes and the efficiency of protein-mRNA hybrid formation, phage and cell surface display are limited by transformation/infection efficiencies and by the number of clones that can be picked (Newton et al., 2020).

Another advantage of cell-free selection methods is the possibility to incorporate unnatural amino acids, which further increases the sequence space. Unnatural amino acids can be introduced by substituting the corresponding tRNAs or tRNA synthetases during translation (Li et al., 2002). Additionally, modifications like macrocyclization can be used in cell-free selections of peptides to enhance serum

stability and membrane permeability (Passioura et al., 2014). Cell-free selection methods cannot only be used to select target-binding proteins or peptides. Depending on the experimental design, covalent bond forming or removing enzymes can be selected (Seelig & Szostak, 2007).

The choice of fusion partner can influence the selection procedure greatly (Newton et al., 2020). In the case of phage display, some Nbs located on the phage surface might alter the phages' infectivity (Kamalinia et al., 2021). In cell-free selection methods, the mRNA sequences and secondary structures can lead to unspecific binding. This problem can be mitigated by a reverse transcription step before translation, generating a mRNA/cDNA-protein hybrid molecule without any single-stranded mRNA (Kurz et al., 2001). In all selection methods, unspecific binders must be removed in a counter selection, where the displayed Nb library is incubated with the "empty" solid phase without the target antigen (Benhar et al., 2000; Smith & Petrenko, 1997). This step is performed to reduce the number of false-positive selection results. Cell-based methods include a protein quality control step in every selection cycle, because misfolded or truncated proteins are degraded and not displayed on the cell/phage surface (Frydman, 2003; Brodsky & McCracken, 1999). Cell-free methods lack this protein quality control step and are prone to overamplify unspecific binders (Polz & Cavanaugh, 1998), which makes a good counter selection indispensable.

In cell-free selection methods, protein quality can be enhanced by using disulfide bond enabling reagents like the disulfide bond enhancer (New England Biolabs) during translation. This provides the formation of a disulfide bond, which is essential for Nb stability. However, Nb genes encoding misfolded proteins are still translated and greatly increase the risk of false-positive selection results if the misfolded protein binds unspecifically to the target antigen or the solid phase used for the panning step (Moreno-Gonzalez & Soto, 2011). Additionally, sequences containing premature Stop-codons, are not eliminated. These can be introduced by PCR errors (Cha & Thilly, 1993) or already be present in the initial library. Therefore, strict washing procedures are essential for cell-free selection methods.

Another point that needs consideration is library bias (Ja et al., 2005). Most synthetic Nb DNA libraries are constructed from a single Nb fold (constant FRs) and show variability in the CDRs. The used Nb scaffold might not translate to well-behaved proteins with all different lengths or geometries of CDRs in the library (Zimmermann et al., 2018). More restrictive library designs lead to higher degrees of library bias. However, less restrictive designs with variable FRs might contain many misfolded proteins, which decreases the likelihood to select well-behaved proteins. Additionally, different Nb sequences might differ in PCR-efficiency, expression levels or transformation/infection efficiency (Takahashi et al., 2003).

All display methods have certain advantages and disadvantages, and there is not one method that clearly beats the others. Because of the large library size of our novel highly variable Nanobody DNA library (chapter 1.3.2), which was designed for this study, a cell-free selection method was required. The mRNA

display method has the additional benefits of a stable, covalent genotype-phenotype coupling and the smallest fusion partner (Figure 1.6). Considering all these advantages, we opted for the method of mRNA display to select Nbs from our novel library design against the mentioned antigens retGC1, myrGCAP1, and SARS-CoV2-M^{pro}.

1.2.4   Detailed look into our Nb selection strategy by mRNA/cDNA display

For our Nb selections against retGC1, myrGCAP1 and SARS-CoV2-M^{pro}, we based our mRNA/cDNA display procedure on a protocol that was shown to generate well-behaved, antigen-binding Nbs (Doshi et al., 2014). Here, Nbs could be generated even for membrane proteins, such as our target retGC1. Because of our decision not to use splint ligation to modify the Nb mRNA library with puromycin, but to use a psoralen-mediated UV-crosslink approach, we combined the mentioned method with a psoralen-mediated UV-crosslinking procedure (Seelig, 2011).

In the following, I will give a detailed insight into every single step of the Nb selection procedure by mRNA/cDNA display.

The first step to enter the first selection cycle is a PCR, which amplifies and modifies the initial Nb DNA library. The initial library can have the form of linear DNA fragments or plasmids containing Nb genes. During this first PCR, or several PCRs, the library is prepared for subsequent reactions (Figure 1.7). A T7 promoter, enabling transcription by T7 RNA polymerase, as well as a ribosome binding site for translation, and a Flag-affinity tag are added to the Nb gene 5'-end, and the annealing site for the puromycin containing oligonucleotide is added to the 3'-end.

After PCRs, a Nb mRNA library is generated by T7 RNA polymerase-mediated *in vitro* transcription. Because the T7 promoter serves as recognition site for the RNA polymerase, it is lost after transcription, and single-stranded mRNA molecules are generated with a ribosome binding site at the 5'-end (Figure 1.7).

The Nb mRNA library is then modified with Puromycin by psoralen-mediated UV-crosslink. An oligonucleotide containing a 5'-psoralen and a 3'-puromycin label binds the Nb mRNA library by base-pairing. Psoralen intercalates in the base pairs and by irradiation at a wavelength of 365 nm forms a covalent bond between the mRNA and the oligonucleotide (Figure 1.7). This step generates a Nb mRNA library with a 3'-puromycin label, which prepares the library for *in vitro* translation.

The actual display of Nb proteins on their own mRNA happens during *in vitro* translation of Puromycin-labeled Nb mRNA using the PURExpress *invitro* protein synthesis kit (New England Biolabs; Shimizu et al., 2001). By mimicking the structure of a tyrosyl-tRNA, Puromycin can enter the A site of a ribosome and the nascent protein chain is translocated onto the puromycin molecule. Since the ribosome cannot cleave the peptide bond in the puromycin, a covalent hybrid molecule consisting of the Nb protein, and the corresponding mRNA is generated (Figure 1.7).

The mRNA moiety of the hybrid molecule might cause some problems for selection and the subsequent sequence regeneration. Single-stranded mRNA is prone to RNase-degradation, which would lead to truncated or absent Nb sequences in the sequence regeneration step. Additionally, single-stranded RNA is likely to form secondary structures and to bind nucleic acids or proteins with its unpaired bases. Binding to the antigen, or the solid phase during selection would lead to false-positive results. To enhance stability and to eliminate unpaired bases, a cDNA strand is synthesized by reverse transcriptase (Figure 1.7). This step generates Nb mRNA/cDNA-protein hybrid molecules, which give the selection method its name of mRNA/cDNA display.
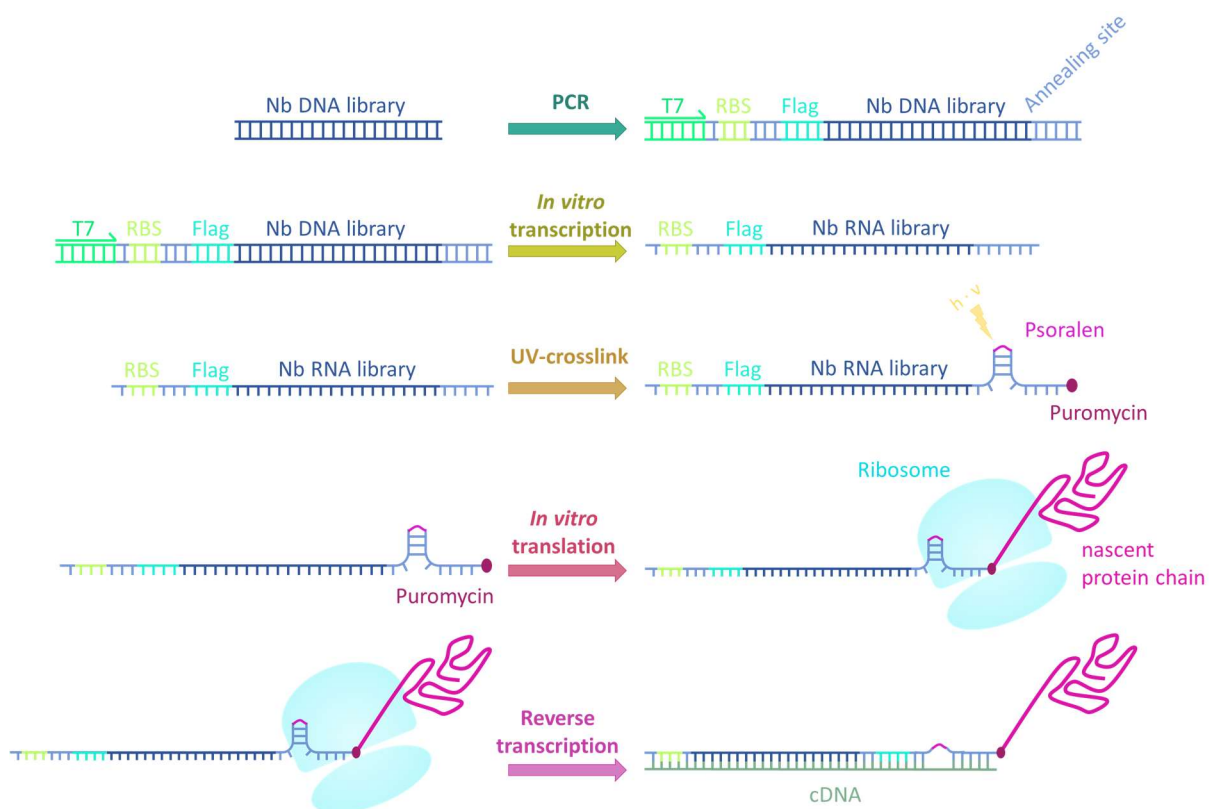
Figure 1.7 Overview of the single reaction steps to generate a Nb mRNA/cDNA-protein hybrid library. The initial Nb DNA library is amplified by PCR, and a T7 promoter, ribosome binding site (RBS), a Flag-affinity-tag, and a 3'-annealing site are added to the Nb genes. During *in vitro* transcription, a Nb mRNA library is generated. Then, a puromycin-containing oligo is annealed to the Nb mRNA 3'-end and covalently linked by psoralen-mediated UV-crosslink. During *in vitro* translation, Nb mRNA and nascent protein chain are covalently linked. The mRNA is then reverse transcribed, generating a Nb mRNA/cDNA-protein hybrid library.

The Nb mRNA/cDNA-protein hybrid molecules can then be used for the affinity-based selection against the target antigens. As a first selection step, unspecific solid phase-binders are removed in a counter selection, where the Nb mRNA/cDNA-protein library is incubated with the Strep-tactin magnetic beads, which serve as solid phase. By removing unspecific solid-phase binders, the number of false-positives is greatly reduced.

After the solid-phase binders are removed, the library is finally ready to generate complexes with the target antigen, which happens in solution. The complexes of Nb hybrids and Strep-tagged target antigen are bound to the Strep-tactin magnetic beads (Figure 1.8). Several stringent wash steps remove all unspecific binders. After washing, the target antigen in specifically eluted by biotin, which has a higher affinity to the Strep-tactin than the Strep-tagged antigen. The elution fraction then, in theory, contains only complexes of target antigen molecules with specifically binding Nb hybrids.
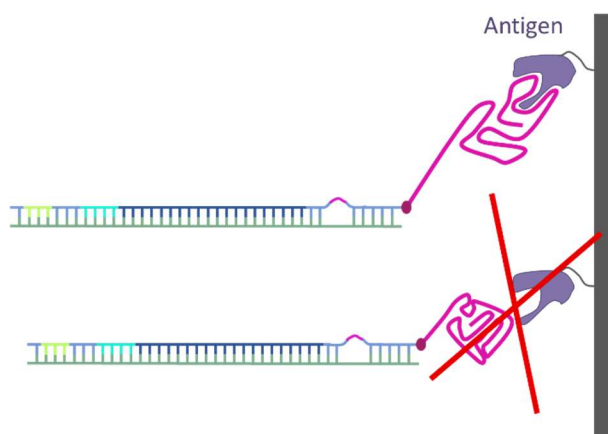
Figure 1.8 Bio-panning step in Nb selection. During the Nb selection step against the target antigen, only high-affinity binders can recognize the antigen. Nbs that do not bind the antigen, or only with low affinity, are removed by stringent wash steps.

To enter a next selection cycle, the Nb cDNA sequences are regenerated by PCR. After several selection cycles (Figure 1.9), selected Nb sequences are regenerated and prepared for subcloning into an expression vector. Subcloned Nb DNA sequences are analyzed by Sanger sequencing. Promising Nb sequences will then be selected for overexpression in *E. coli* host cells and protein purification. Nb proteins that can be expressed and purified in usable amounts can then be used to confirm complex formation and kinetic analysis of those complexes by fiber optics surface plasmon resonance spectroscopy (FO-SPR). FO-SPR experiments are performed using the White FOx instrument (FOx Biosystems).
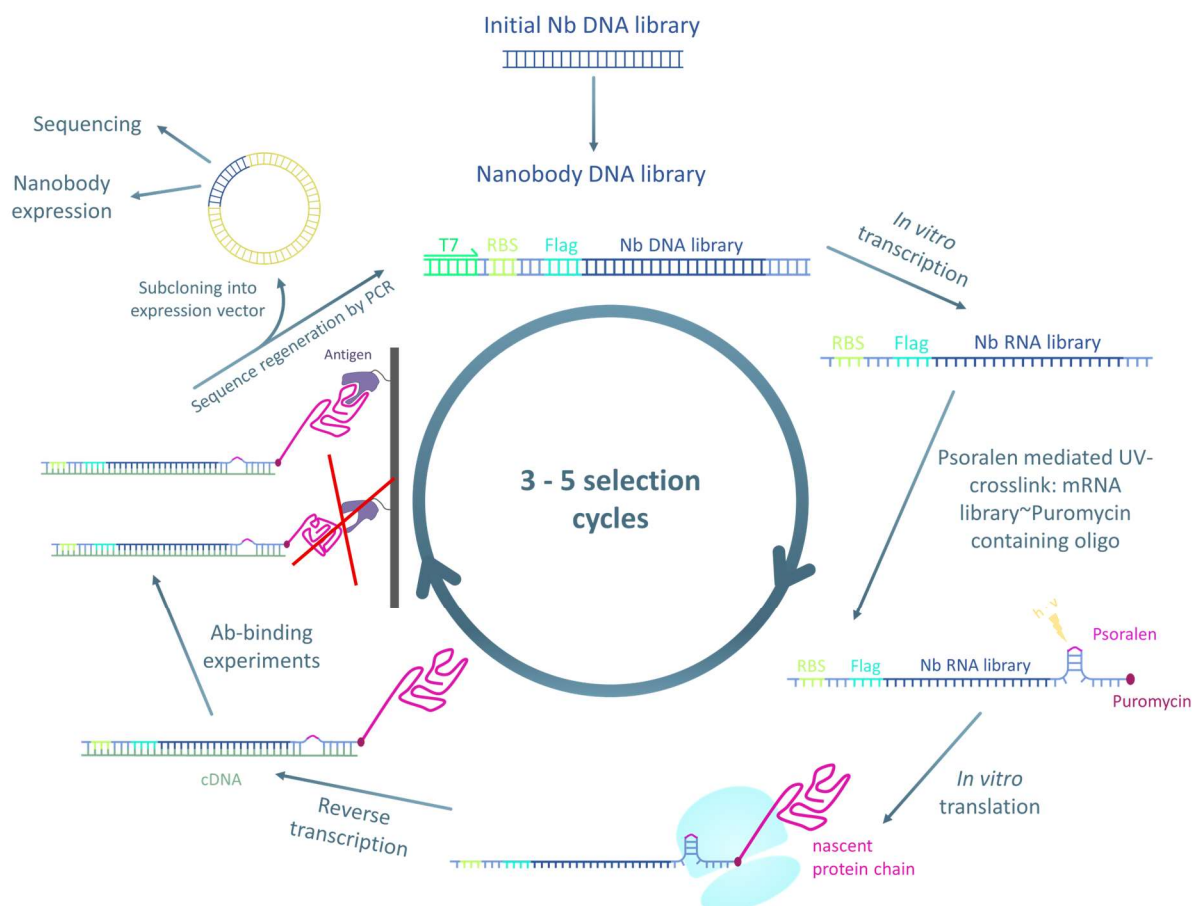
Figure 1.9 Overview image of the established *in vitro* mRNA/cDNA display Nb selection procedure. The single reaction steps are arranged clockwise. The input material, a Nb DNA library (top) is first amplified and modified by PCR, then transcribed into a Nb mRNA library, and fused to a Puromycin containing oligonucleotide by Psoralen-mediated UV-crosslink. Puromycin-containing mRNA molecules are then *in vitro* translated, generating Nb mRNA-protein hybrids (bottom). The Nb mRNA is reverse transcribed. This step generates Nb mRNA/cDNA-protein hybrid molecules, which can then be used for affinity-based selection against a target antigen (left). Selected Nb DNA sequences can then be regenerated/amplified by PCR. After three to five selection cycles, the selected Nb genes are subcloned into an expression vector. The genes are analyzed by Sanger sequencing and selected Nbs can be expressed in *E. coli*.

1.3 Design of Nb DNA libraries for directed protein evolution

1.3.1   Three different types of Nb libraries

A crucial task for the successful animal-free selection of high-affinity Nbs is the choice and design of the initial Nb gene library (Muyldermans, 2021). In general, there are three types of Nb libraries. Immune libraries can be obtained from immunized animals, possess a relatively small library size of roughly 10^6, and already contain many different antigen-binding Abs with different affinities to the target antigen (Saerens et al., 2004). These libraries can then be used for further affinity maturation in animal-free selection methods like phage, yeast, or mRNA display (Ryckaert et al., 2010; Sheikholeslami et al., 2010). The *in vitro* methods ribosome display and mRNA display could then be

used to expand the sequence space by introducing unnatural amino acids (Hartman et al., 2007). Size of immune libraries is limited by the amount of blood that can be harvested from the immunized animals. Other disadvantages are the inability to generate an immune library against non-immunogenic target antigens and the need of large quantities of target antigen to immunize the animals (Muyldermans, 2021). However, after successful immunization, immune libraries contain already lots of antigen-binders, which results in a high success rate (Saerens et al., 2004).

Animal-derived Nb libraries obtained without the immunization step are called naïve libraries. One naïve library can be used for different antigens since it is not affinity-matured for one specific target (Yan et al., 2015). Naïve libraries enable the use of non-immunogenic or toxic antigens and can present larger library sizes. However, blood from several animals is required to obtain large library sizes. A naïve Nb gene library with a library size of roughly 10^{11} requires the harvest of peripheral blood lymphocytes from 24 camels (Yan et al., 2015). Animal-derived libraries provide the advantage of containing only native, well-behaved protein sequences that underwent quality control in the animals' B lymphocytes.

Synthetic Nb libraries provide a different, entirely animal-free approach. Here, usually a single, known Nb is used as library scaffold, and single amino acid positions are randomized (McMahon et al., 2018). Variability in traditional synthetic Nb library designs is restricted to the CDRs. The theoretical protein sequence space increases with a factor of a^n with each randomized position n that allows a different amino acids. So, if all 20 amino acids were allowed in twelve positions, a protein sequence space of 20^{12} or $4.1 \cdot 10^{15}$ is achieved, which already exceeds the feasible library sizes of all mentioned display methods (Muyldermans, 2021; Seelig, 2011). This means that the entire sequence space of highly variable libraries cannot be covered in actual selections. Because synthetic Nb libraries do not require an immunization step, they also allow selection of Nbs against non-immunogenic or toxic target antigens. Synthetic Nb libraries provide a nearly unlimited theoretical sequence space, especially if the variability is extended to the FRs. Therefore, a single synthetic library can be applied for many antigens (Contreras et al., 2023).

The major drawback of these Nb libraries lies in its synthetic nature. Because the Nb sequences do not undergo any quality control steps, many of them might not translate to well-behaved proteins (Rosenblum & Cooperman, 2014). This can partially be circumvented in selections by phage or yeast display, which provide at least the bacterial or yeast protein quality control (Baneyx & Mujacic, 2004). The use of high-variability synthetic libraries is especially challenging in ribosome and mRNA display and requires the application of large libraries to increase the chance of selecting well-behaved Nbs (Kamalinia et al., 2021).

1.3.2   Our novel Nb gene library design for use in *in vitro* selection by mRNA/cDNA display

By using a single Nb scaffold, traditional synthetic Nb gene libraries introduce a high degree of library bias by being limited to a single Nb fold (Zimmermann et al., 2020). Additionally, not all randomly generated CDRs might be compatible with this certain fold, because different lengths and geometries of CDRs might require different Nb backbones.

For this study, we designed a novel, highly variable Nb gene library, called c-lib, which uses a different approach. By aligning several known Nb protein sequences, some well-conserved amino acids could be determined in the FRs. These amino acids are thought to be essential for well-behaved Nb folds and are invariant in our library design. For instance, two cysteines in FR1 and FR3, respectively, form a necessary disulfide bond and are invariant positions. Other, less-conserved positions were randomized, introducing sequence variability to the FRs. CDRs are designed to be not only sequence-randomized but also show length-polymorphisms. CDR1 can contain either 4, 6, 8, or 10 amino acids, CDR2 can have lengths of 5 to 8, and CDR3 can have lengths between 3 and 23 amino acids (Figure 1.10).

Figure 1.10 Schematic representation of our novel Nb gene library design. The framework regions (FR) contain invariant (grey) and partially or entirely random (pink) positions. The complementarity-determining regions (CDR) are entirely sequence-randomized and show length polymorphisms (pink arrows). A disulfide bond between two invariant cysteine residues (Cys) is formed in every Nb of the library design.

Longer sequence stretches of completely randomized sequences, like the CDRs, provide a high chance of introducing Stop-Codons. To mitigate this risk, the CDRs are designed with NNK codon usage, instead of NNN, with N allowing any nucleotide and K only guanine or thymine. This eliminates the UAA and UGA, retaining only the UAG Stop-codon. Using NNK triplets reduces the Stop-codon probability from 4.7 % to 3.1 % per amino acid position. A KGRFT-motif in FR3 that is well conserved among Nbs was designed to be invariant to facilitate DNA library synthesis by GeneArt (Thermo Fisher) (Figure 1.11).
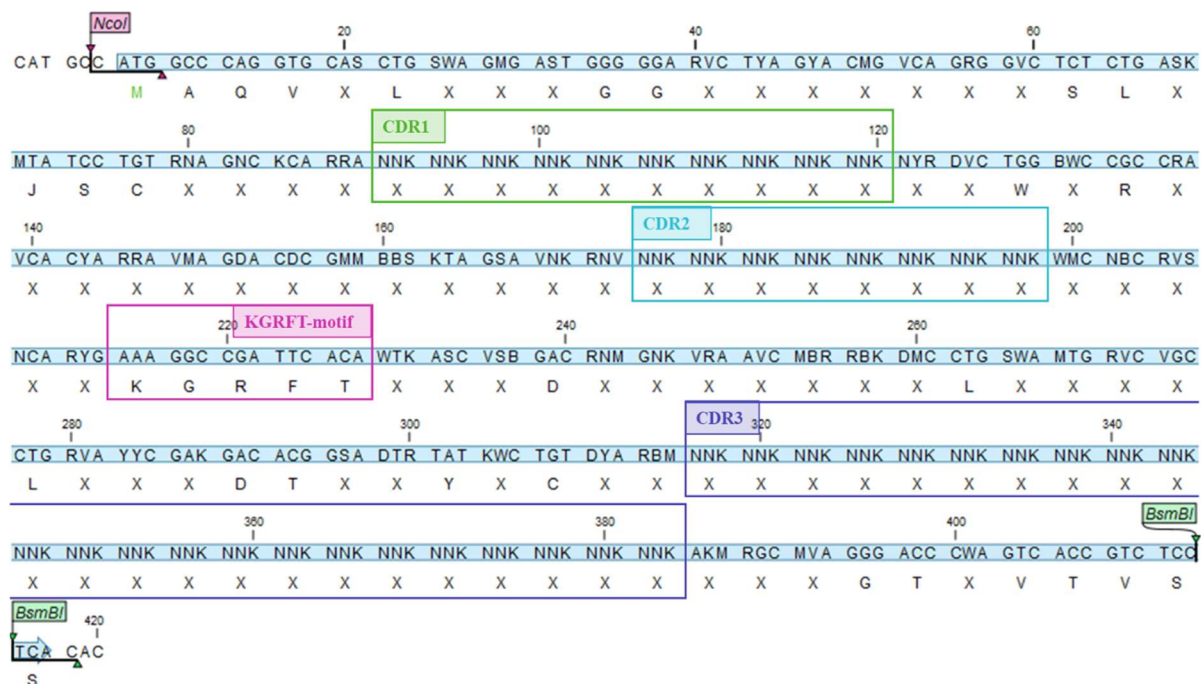
Figure 1.11 Theoretical library design. The theoretical variable Nb DNA sequence is shown in IUPAC nucleic acid code (blue box) with the corresponding protein sequence in one letter amino acid code below. Variable positions are labeled with X. The entirely sequence-randomized CDRs are highlighted in green (CDR1), cyan (CDR2) and purple (CDR3). An invariant KGRFT-motif is localized in FR3 (pink).

1.3.3 The yeast Nb library provides a traditional Nb gene library design

In contrast to our novel library design of the c-lib, the yeast Nb gene (y-lib) (McMahon et al., 2018), which is commercially distributed by Kerafast, uses the traditional approach with a single Nb scaffold and some variation in the CDRs. The heavy chain Ab, which serves as a scaffold for the Nb library design is derived from *Llama glama* genes *IGHV1S1–S5* (McMahon et al., 2018). The y-lib uses highly variable positions with a certain distribution (Figure 1.12, grey box), especially in CDR1 and CDR3. These highly variable regions are flanked by positions with lower variability. To allow different CDR3 geometries, and therefore different binding modes, the highly variable region in CDR3 is designed in different lengths of seven, eleven, and 15 amino acids (Figure 1.12). The amino acid distribution for the highly variable positions and the restraints on the less-variable positions has been obtained from comparing 93 known Nb sequences (McMahon et al., 2018).
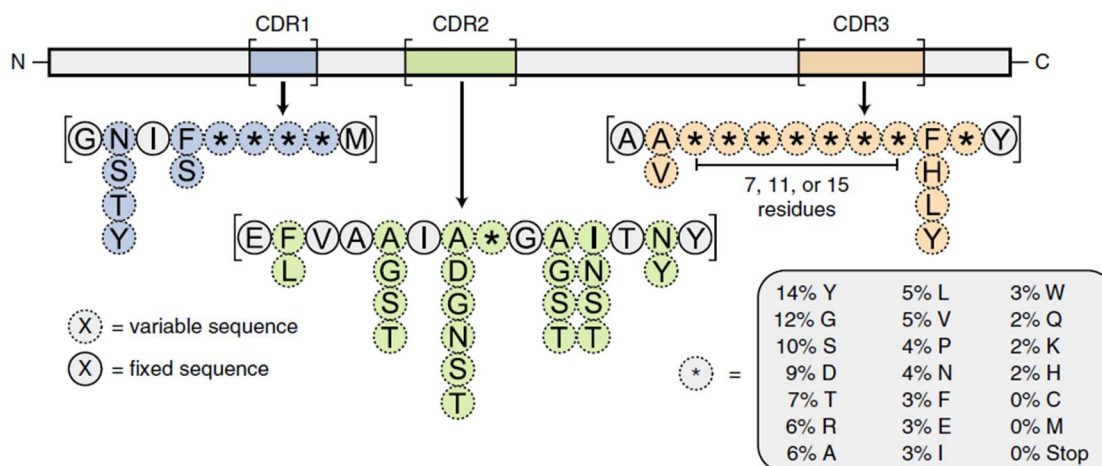
Figure 1.12 Schematic design of the y-lib Nb gene library. The library uses a single llama-derived scaffold (grey) and partially variable complementarity-determining regions CDR1 (blue), CDR2 (green), and CDR3 (orange). Stacks of one letter coded amino acids mark variable positions. Positions marked with an asterisk are more variable and use 18 amino acids with the shown distribution (grey box; amino acids in one letter code). CDR3 is present in different lengths, with seven, eleven, or 15 highly variable positions. The image was adapted from McMahon et al., 2018.

Using one Nb scaffold and restraining amino acid distribution according to known Nbs enhances the likelihood to select well-behaved Nbs (McMahon et al., 2018). On the other hand, these restraints introduce a high level of library bias, which decreases the chance of selecting high-affinity binders (Rodi et al., 2002).

Another factor that needs consideration is the fact that the y-lib was designed for the use in yeast display (McMahon et al., 2018). In theory, every synthetic Nb gene library could be utilized in different display methods. However, by using the library in yeast display, Nb proteins undergo quality control during translation and translocation to the cell wall (Vashist & Ng, 2004). This eliminates misbehaved Nb proteins or sequences that potentially contain premature Stop-codons. Additionally, with its small size of roughly $5 \cdot 10^8$, the y-lib provides only a limited sequence space, which further reduces the likelihood of selecting high-affinity target binders. The library has been shown to generate Nbs that specifically bind their target antigens (human serum albumin and β_2 adrenergic receptor) with moderate affinities of up to 44 nM (McMahon et al., 2018). But this success does not necessarily translate to different antigens and display methods.

In the following, we will attempt to translate the success of the y-lib to selections by mRNA/cDNA display against our selected target antigens. Additionally, the smaller library size and published success will help us to establish and validate our single reaction steps for our Nb selection procedure.

1.4 The vertebrate visual signal transduction cycle and the roles of retinal guanylate cyclase 1 and guanylate cyclase activating protein 1

Vertebrate retinas are composed of different cell types, with two kinds of photoreceptor cells, so called rods and cones. In these cells, incoming photons are converted into neurological signals, which then can be processed by neuronal cells and evaluated in the brain. Rods and cones are elongated cells that are partitioned into an outer and an inner segment (Detwiler & Laurens, 1921) and connected by a thin cilium that mainly regulates protein transport (Yang & Li, 2005; Figure 1.13).
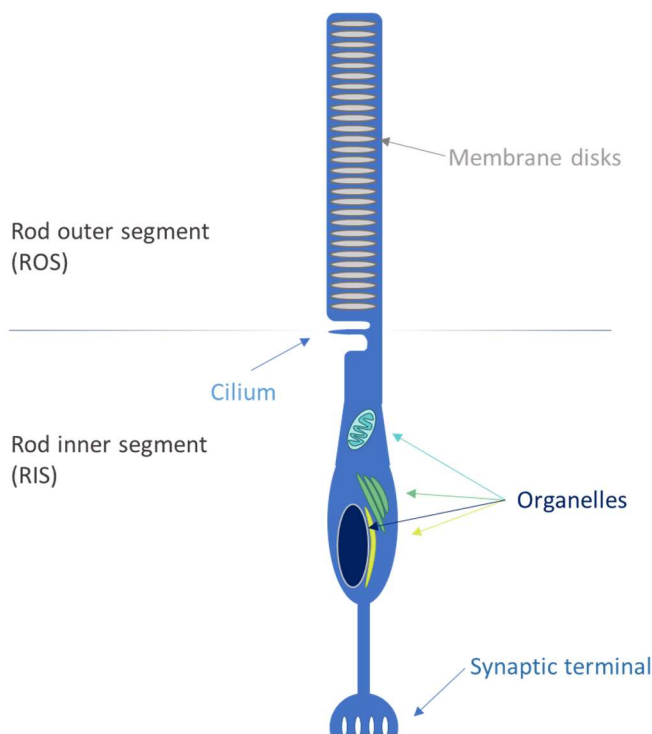
Figure 1.13 Schematic architecture of a retinal rod cell. A rod cell is partitioned in an inner (RIS) and outer segment (ROS), which are connected by a thin cilium. The inner segment contains the nucleus, mitochondria, and all other organelles, as well as a synaptic terminal to communicate with adjacent cells. The outer segment contains membrane discs, where incoming photons trigger a signal cascade that leads to a change in intracellular Ca^{2+}-concentrations.

The inner segment contains the nucleus and all the other organelles (Young & Droz, 1968). Since the cell organelles are located in the inner segment, all proteins are synthesized here, and outer segment-proteins need to be transported across the connecting cilium. Additionally, the inner segment contains synaptic terminals, where the photoreceptor cell can communicate with adjacent neuronal cells by releasing neurotransmitters (Marc et al., 1990).

The outer segment is responsible for the actual reception of photons and the signal translation into an electrophysiological signal. Retinal outer segments contain many membrane discs (Nickell et al., 2007) that accommodate transmembrane proteins like the GPCR (G-protein coupled receptor) rhodopsin. Rhodopsin consists of a 38 kDa 7-transmembrane-domain protein, called opsin, and the chromophore 11-*cis*-retinal and is the first responder to an incoming photon in rod cells. Cone cells use similar but slightly different opsin variants. The energy of an incoming photon is absorbed and converted into an isomerization of the 11-*cis*-retinal into the all-*trans*-retinal (Bernstein et al., 1987), which leads to a conformational change in the opsin (Farrens et al., 1996). One activated rhodopsin can stimulate up to

800 G-proteins called transducin (Bruckert et al., 1992), which exchanges its bound GDP for GTP (Fung et al., 1981). This leads to a separation of the GTP-bound α-subunit from the β-γ complex. Two activated (GTP-bound) α-subunits can then activate one phosphodiesterase 6 (PDE6) (Qureshi et al., 2018), which is a membrane-bound enzyme, as well. PDE6 hydrolyzes cyclic GMP (cGMP) (Yee & Liebman, 1978), which in turn regulates cyclic nucleotide gated (CNG) ion channels. The presence of cGMP opens these channels, which allows cations like K^+, Na^+, or Ca^{2+} to enter the cell. After cGMP hydrolysis, the channels close and the concentration of free Ca^{2+}-ions, $[Ca^{2+}]$, decreases, because they are constantly removed from the cell by an ion pump (Yau & Nakatani, 1985) (Figure 1.14). This Ca-depletion leads to a reduction of cations in the cell and hyperpolarizes the photoreceptor cell (Hagins et al., 1970). This electrophysiological signal is then translated into a reduced release of neurotransmitters, which is then processed as the signal of an illuminated cell.
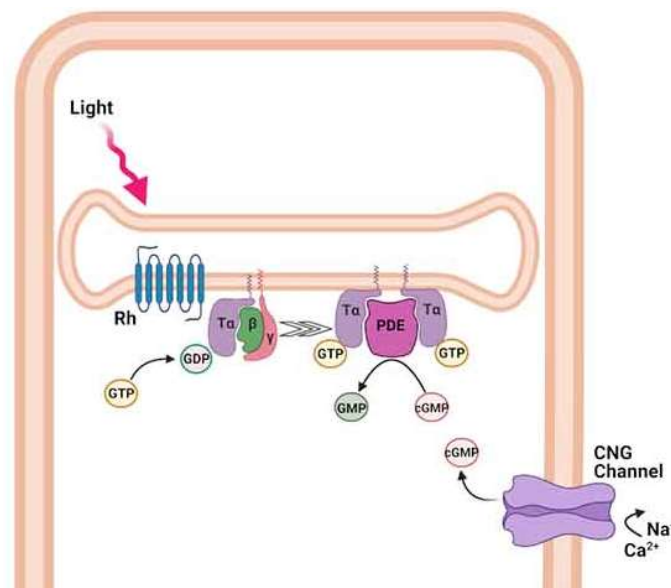
Figure 1.14 Photoreceptor light-state signal cascade. An incoming photon triggers a chain reaction in retinal rod outer segments. A photon (Light) is absorbed by the GPCR rhodopsin (Rh), where the isomerization of 11-*cis*-retinal to all-*trans*-retinal leads to a conformational change activating the rhodopsin. Activated rhodopsin can stimulate several hundred G-proteins called transducin (T), which exchanges GDP for GTP. Transducin's GTP-bound α-subunit separates from the β-γ-complex. Two GTP-bound α-subunits can induce phosphodiesterase 6 (PDE) activity, which hydrolyzes cGMP to GMP. At depleted cGMP-concentrations, cyclic nucleotide-gated (CNG) ion channels close. Because Ca^{2+} is constantly pumped out of the cytoplasm, $[Ca^{2+}]$ sinks rapidly, which reduces the cytosolic cation concentration and hyperpolarizes the rod cell. The image was adapted from Genovese et al., 2021.

After illumination, the dark state is restored by inactivating the mentioned proteins and converting them back to their initial state, and by restoring cGMP levels, which open the cGMP-gated ion channels and re-establish dark-state Ca^{2+} concentrations. Rhodopsin is phosphorylated by a rhodopsin kinase and then bound by the inactivating protein arrestin (Gross & Burns, 2010). Before the next excitation, all-*trans*-retinal is exchanged and rhodopsin is de-phosphorylated. Transducin is self-inactivated by its own GTPase activity. After GTP is hydrolyzed to GDP, the transducin α-β-γ complex is reorganized, which

stops PDE activation. cGMP levels are restored by the enzyme retinal guanylate cylase 1 (retGC1), which is tightly regulated by the guanylate cyclase activating protein 1 (GCAP1).

During an initial dark state, GCAP1 binds three Ca^{2+}-ions, which are present at a high concentration. Calcium-bound GCAP1 inhibits retGC1, while leaving a small basal activity, maintaining high cGMP levels in the outer segments, and keeping the cGMP-gated ion channels open. During illumination, PDE6 hydrolyses cGMP with a higher rate than the basal retGC1 activity, hence cGMP levels drop. As mentioned, cGMP-gated ion channels close and $[Ca^{2+}]$ is decreased. Once $[Ca^{2+}]$ reaches its minimum, GCAP1 binds the more abundant Mg^{2+}-ions. This leads to a conformational change in GCAP1 structure, which triggers a strong activation of retGC1. The combination of retGC1 synthesizing cGMP with a high rate and the loss of PDE activation by transducin after GTP hydrolysis restores the dark-state cGMP levels. Then, cGMP-gated ion channels can open again and reestablish dark-state $[Ca^{2+}]$ (Figure 1.15).
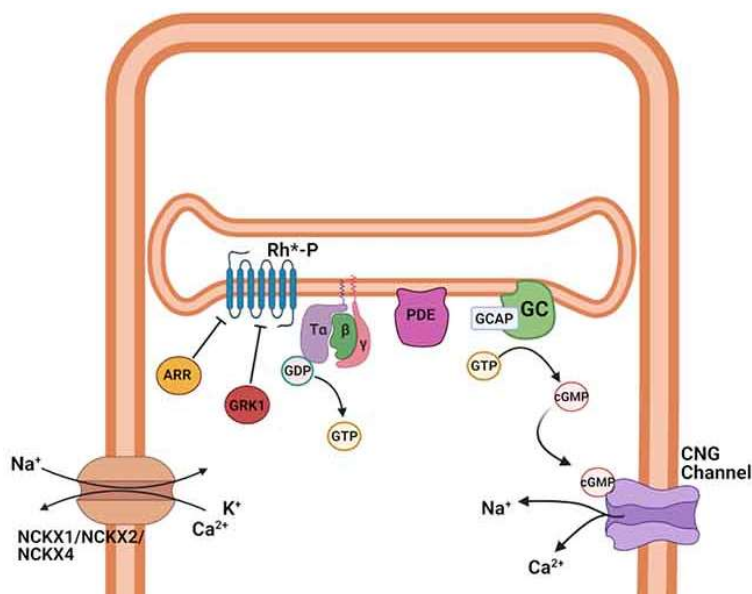
Figure 1.15 Restoration of the photoreceptor dark state. Active rhodopsin (Rh*) is phosphorylated (Rh*-P) by a rhodopsin kinase or G-protein coupled receptor kinase 1 (GRK1), which recruits arrestin (ARR), resulting in rhodopsin inactivation. Transducin (T) self-inactivates by its GTPase activity. After GTP hydrolysis to GDP, the transducin α-β-γ-complex reassembles and stops phosphodiesterase (PDE) activation. At low $[Ca^{2+}]$, guanylate cyclase activating protein 1 (GCAP) strongly stimulates guanylate cyclase (GC). Active GC replenishes the cGMP levels, opens the cyclic nucleotide gated (CNG) ion channels, and dark state $[Ca^{2+}]$ can be restored. The image was adapted from Genovese et al., 2021.

With their crucial roles in the restoration of the dark state, retGC1 and GCAP1 are essential proteins for the sense of vision and the maintenance of a healthy retina. Mutations in the genes encoding for retGC1 (GUCY2D) and GCAP1 (GUCA1A) can lead to severe pathologies (Dizhoor et al., 2016; Sokal et al., 1998), which makes them interesting target antigens for Nb selections.

1.4.1 Retinal guanlylate cyclase 1 (retGC1)

As mentioned above, the main role of the retinal guanylate cyclase 1 (retGC1) is cGMP synthesis, which opens CNG-channels and restores dark state Ca^{2+}-levels.

Catalytically active retGC1 is present in homodimeric form composed of two 120 kDa monomers. RetGC1 is composed of an N-terminal, disc lumen-facing domain, a single transmembrane helix, and intracellular domains, essential for catalytic function, GTP-binding, and interaction with GCAPs. Interestingly, retGC1 is associated with so called membrane rafts (Nair et al., 2002), membrane structures of different lipid composition, which show high detergent resistance (Fiedler et al., 1993). This makes retGC1 very difficult to purify, since solubilization in detergent is usually inefficient. To date, no successful purification of retGC1 has been shown. While experiments regarding catalysis and GCAP binding could be performed with crude retGC1-containing cell lysates, protein structure determination by X-ray crystallography or cryogenic electron microscopy (cryo EM) or the selection of anti-retGC1 Ab fragments requires pure retGC1 (Kitiratschky et al., 2009; Yu et al., 1999).

Analysis of domain organization and interactions of retGC1 are based on crosslinking mass spectrometry (XL-MS) (Rehkamp et al., 2021). Domain architecture of retGC1 is mainly based on homologies to other guanylate cyclases (Figure 1.16). The N-terminal, disc lumen facing, portion of retGC1 contains an unstructured leader sequence (LS) and a lumen-facing domain (LD) of unknown function (Dizhoor et al., 1994). As mentioned, each monomer traverses the disc membrane with a single α-helix (TmD) (Shyjan et al., 1992). The juxtamembrane domain (JmD) is essential for interaction with GCAP1 (Lange et al., 1999) and is simply named by its membrane-adjacent location. The kinase homology domain (KHD) shows high sequence-similarity with known kinases (Shyjan et al., 1992). Based on this homology it can be concluded that the KHD is essential for nucleotide binding, in this case GTP. KHD has also been shown to interact with GCAP1 (Peshenko et al., 2010) and GCAP2 (Laura & Hurley, 1998), a secondary regulator of retGC1. The dimerization domain (DD) is an α-helical domain, which has been shown to be essential for dimerization and catalytic function of retGC1 (Zägel et al., 2013). Cyclization of GTP to cGMP is catalyzed in the core catalytic domain (CCD), which is highly conserved among guanylate cyclases (R. B. Yang et al., 1995). GCAP2 has been shown to bind the interface between the CCD and a C-terminal extension (Pettelkau et al., 2012); Figure 1.16).
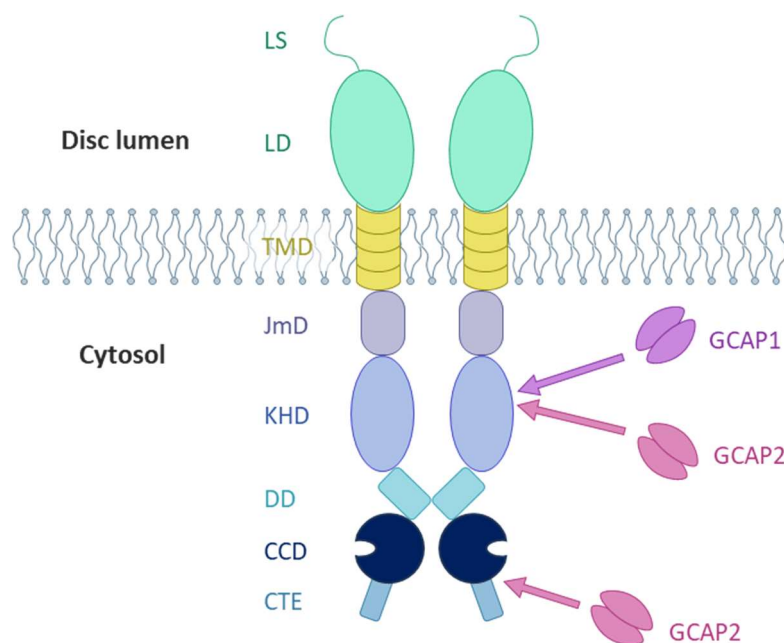
Figure 1.16 Domain organization of a retinal guanylate cyclase 1 (retGC1) dimer. An N-terminal leader sequence (LS) and a lumen domain (LD) are facing into the lumen of rod outer segment membrane discs. A transmembrane domain (TmD) traverses the disc membrane with a single α-helix per monomer. The cytosolic portion contains a juxtamembrane domain (JmD), kinase homology domain (KHD), dimerization domain (DD), core catalytic domain (CCD) and a C-terminal extension (CTE). The regulating proteins GCAP1 and GCAP1 have been shown to interact with the KHD. Additionally, interactions between the CTE and GCAP2 have been detected.

Mutations in the GUCY2D gene coding for retGC1 are associated with several retinopathies. The R838S mutation has been found in patients diagnosed with cone-rod dystrophy 6 (CORD6) (Wilkie et al., 2000). RetGC1 R838S shows a decreased basal activity in dark-adapted photoreceptors (Dizhoor et al., 2016). The R838S mutation has been shown to alter the affinity of retGC1 to GCAP1, and retGC1 R838S preferably binds the activating, Mg^{2+}-bound GCAP1. This leads to a constitutive cGMP synthesis (Peshenko et al., 2004). After illumination, constitutively active retGC1 counteracts PDE6 cGMP hydrolysis, which leads to an inefficient closure of CNG channels, incomplete Ca^{2+}-depletion and lack of hyperpolarization. So, retGC1 R838S containing rods are not able to successfully emit the signal of an illuminated photoreceptor. Additionally, the R838S mutation decreases retGC1's affinity to the retinal degeneration 3 (RD3) protein (Dizhoor et al., 2016), which mediates retGC1 transport from rod inner to outer segments (Molday et al., 2013).

Leber's congenital amaurosis 1 (LCA1) (Leber, 1869) is often related to the retGC1 P858S variant (Dharmaraj et al., 2000). RetGC1 shows decreased basal activity and barely any GCAP1-mediated stimulation (Tucker et al., 2004). RetGC1 P858S is inept to restore dark state cGMP levels and CNG channels remain closed. Because the dark state cannot be restored, retGC1 P858S containing rods are trapped in the illuminated state. Altered affinity to RD3 could also be detected for this variant (Zulliger et al., 2015).

In LCA1, CORD6 and other retinopathies, affected photoreceptors are degraded over time. This irreversible process is called retinal degeneration (Dharmaraj et al., 2000; Wilkie et al., 2000). Residues R838 and R858 are both localized between the DD, and the CCD (Duda et al., 2016). To date, the exact molecular mechanisms of these pathological mutations are not entirely understood.

Since Nbs could facilitate protein purification, stabilize certain conformations, or inhibit enzymes. So, high-affinity binders would be valuable tools to further characterize retGC1 structure and function. Enhanced knowledge about the disease-related retGC1 variants would lead to a better understanding of the mentioned retinopathies and could aid the development of therapeutics.

1.4.2 Guanylate cyclase activating protein 1 (GCAP1)

The main retGC1 regulator is a protein called guanylate cyclase activating protein 1 (GCAP1), which controls retGC1 activity depending on the intracellular $[Ca^{2+}]$ (Gorczyca et al., 1994; Koch & Stryer, 1988). GCAP1 has a MW of roughly 24 kDa and possesses four EF hand motifs, three of which can bind bivalent cations (Palczewski et al., 1994; Figure 1.17). The N-terminally located EF-1 does not bind any ions but is essential for interaction with retGC1 (Otto-Bruc et al., 1997). Depending on the ion bound, GCAP1 can either inhibit retGC1, or strongly stimulate retGC1 activity. Ca^{2+}-bound GCAP1 (GCAP1-Ca), which is present during dark state, inhibits retGC1, while still allowing a basal activity. In illuminated photoreceptors, $[Ca^{2+}]$ drops drastically from 500 nM to 50 nM and GCAP1, while $[Mg^{2+}]$ remains at roughly 1 mM during all stages of the visual cycle (Chen et al., 2003; Gray-Keller & Detwiler, 1994). This triggers a conformational change in GCAP1 and Mg^{2+}-bound GCAP1 (GCAP1-Mg) strongly stimulates cGMP synthesis (Sokal et al., 2001).

GCAP1-Ca shows an altered migration pattern in SDS-PAGE (Sokal et al., 2005). In contrast to the ion-free GCAP1 species that migrates at 24 kDa, the GCAP1-Ca migrates at roughly 19 kDa. The altered migration behavior is probably caused by some structural elements that are stable even in SDS-containing buffers (Sokal et al., 2005).

GCAP1 has been found to be N-acylated, mostly with the fatty acid myristic acid (Otto-Bruc et al., 1997; Palczewski et al., 1994). Myristoylation has been shown to be necessary for GCAP1 stability and proper retGC1 regulation (Stephen et al., 2007). Myristoylation is not thought to be a membrane anchor or provide any interactions with retGC1, since it is buried within the GCAP1 structure and not solvent-exposed in GCAP1-Ca or GCAP1-Mg (Lim et al., 2016; Stephen et al., 2007). The N-myristoylated form of GCAP1 has been named myrGCAP1. When GCAP1 is overexpressed in *E. coli*, the D6S variant enables *in vivo* myristoylation by a yeast N-myristoyl transferase (Hwang & Koch, 2002). Because the N-myristoylation is essential for native GCAP1 folding, the myrGCAP1 D6S variant is called wild type GCAP1.

In contrast to retGC1, GCAP1 is easily overexpressed in bacteria and can be purified in large quantities. Protein structures of different states of wild type GCAP1 were determined by solution NMR (Lim et al., 2016) and X-ray crystallography (Stephen et al., 2007), and ion-dependent conformational changes in the EF-hands can be visualized (Lim et al., 2016).
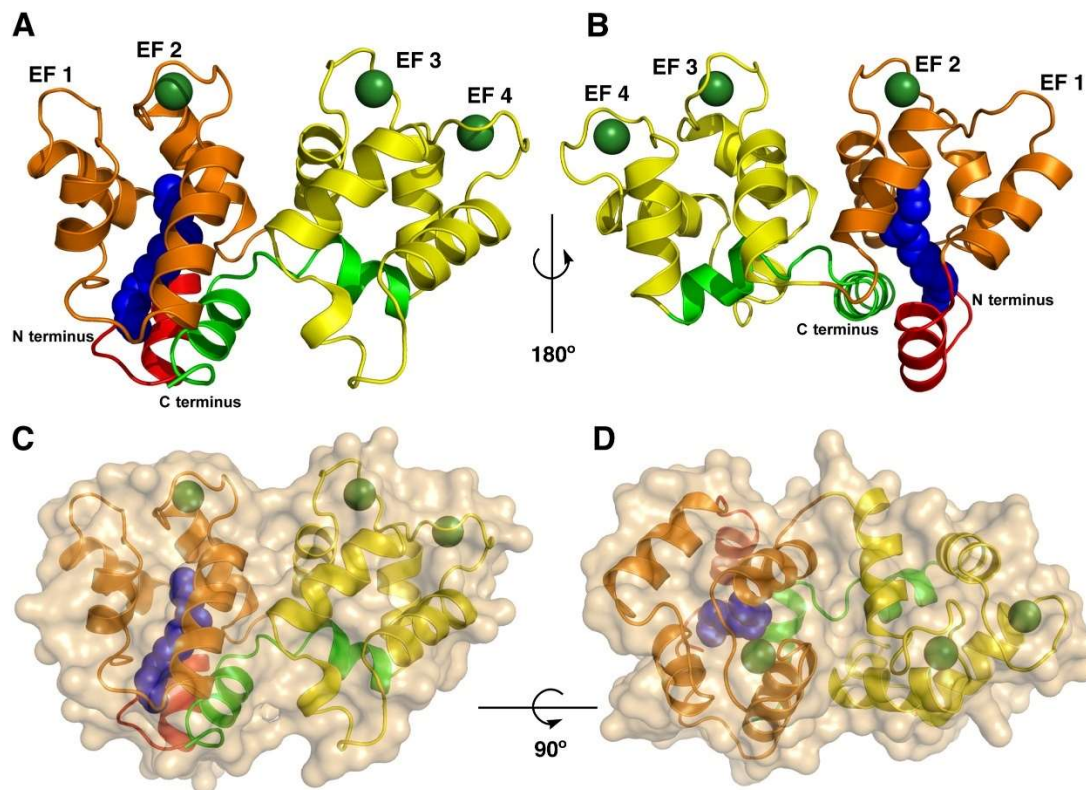
Figure 1.17 Protein structure of a myristoylated guanylate cyclase activating protein 1 (myrGCAP1). Cartoon representation of a myrGCAP1 protein structure as front and back view (A and B), with Ca^{2+}-ions (green spheres) bound by EF-hands 2 – 4. A myristoyl group (blue spheres) is buried within the protein. The N-terminus of myrGCAP1 is shown in red. Color transitions over orange and yellow to the green C-terminus. C and D show the myrGCAP1 cartoon image within its surface representation in two different views (C and D). The image was adapted from Stephen et al., 2007.

Several mutations in the GUCA1A gene, which encodes for GCAP1, have been found in patients diagnosed with different retinopathies. The Y99C mutant is associated with an autosomal dominant cone dystrophy (adCORD) (Payne et al., 1998). Y99 is localized adjacent to EF-hand 3 results. The mutation to cysteine results in a decreased affinity for Ca^{2+}, which leads to a constitutive activation of retGC1 (Sokal et al., 1998). Similar to Y99C, many disease-related mutants have been shown to influence ion sensitivity and exhibit deviated retGC1 regulation.

One exception is the M26R mutation, which has been shown to eliminate retGC1 activation, entirely (Peshenko et al., 2014). M26R is positioned in EF-1 and the mutation to arginine influences GCAP1's interaction with retGC1 and has been shown to interfere with GCAP2-mediated activation of retGC1 (Peshenko et al., 2015).

Selecting Nbs against GCAP1 wild type and M26R could mediate crystallization and aid structural analysis of different GCAP1 variants. Additional purposes for anti-GCAP1 Nanobodies could be the stabilization of retGC1-GCAP1 complexes. Structural analysis of these complexes could provide new insights in retGC1 regulation by GCAP1-Ca and GCAP1-Mg, which would greatly improve the understanding of the visual phototransduction cycle and the mechanisms that influence the mentioned retinopathies.

1.4.3   Treatment of rare retinopathies

The rare inherited retinal dystrophies usually have an early onset and can already be diagnosed in infants in electroretinograms (Brecelj & Stirn-Kranjc, 1999; Perrault et al., 1996). To date, there are no FDA- or EMA-approved treatments. Potential approaches to treat these rare retinopathies have been established in mouse or macaque models (Boye et al., 2010; McCullough et al., 2019) with several approaches in clinical trial (Chiu et al., 2021). In theory, Nbs could also be used to treat these retinopathies. A Nanobody that specifically binds and inhibits retGC1 or GCAP1 mutants in heterozygous patients might restore native retGC1 function and regulation by GCAP1 and might decelerate retinal degeneration.

Therapeutical Nbs for the treatment of retinopathies would have to be administered to the retina and would need to function in the cytoplasm of photoreceptors. Even though Nanobodies lack most posttranslational modifications, they still usually require one stabilizing disulfide bond (Dingus et al., 2022). Nbs and other Ab fragments that are stable within the reducing, cytoplasmic environment, have been described as so called intrabodies (Lobato & Rabbitts, 2003; Zhang et al., 2021). While the development of intracellular Nbs is possible, this provides additional challenges. To function inside the photoreceptor cytoplasm, selected Nanobodies must first enter the cell through endocytosis, escape the endosome (D'Astolfo et al., 2015; Schumacher et al., 2018) and then be stable in the cytoplasmic conditions (Dingus et al., 2022).

1.5 SARS-CoV2-main protease M^{pro}

After being detected in Chinese patients in late 2019 (Zhou et al., 2020), the SARS-CoV2 (severe acute respiratory syndrome causing coronavirus 2) rapidly spread across the entire world in 2020, causing the global COVID19 (coronavirus disease 2019) pandemic, which resulted in millions of fatalities (Simonsen & Viboud, 2021). Immediately, large efforts have been put into the development of vaccines and treatments to mitigate the spread of the virus (Polack et al., 2020; M. Wang et al., 2020).

The genome of SARS-CoV2 consists of roughly 30 000 nucleotides and is organized in two overlapping open reading frames (ORF) encoding for all non-structural proteins (nsp), and several smaller ORFs for the structural proteins Spike, Envelope, Membrane, and Nucleocapsid (Figure 1.18; Kim et al., 2020).

The two ORFs containing the nsp genes are separated by a frameshift and both produce long polyproteins containing the single nsps (Kelly et al., 2020). For a virus particle to form, these nsps must be cleaved into individual proteins by two proteases (Khailany et al., 2020), papain-like protease (PL^pro) (nsp3), and 3-chymotrypsin-like, or main protease (M^pro) (nsp5). Both proteases autocatalytically cleave themselves from the polyprotein 1a before processing all other nsps from both polyproteins (Cannalire et al., 2022).
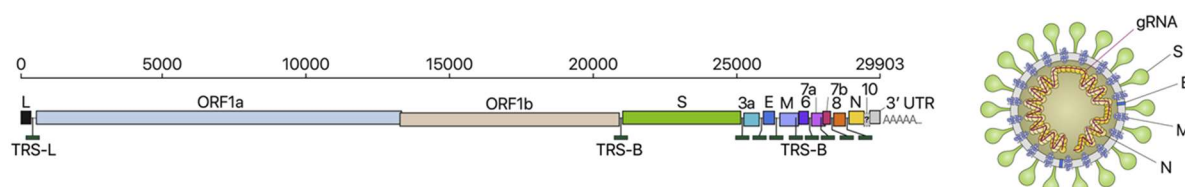
Figure 1.18 Genotype and phenotype representations of the SARS-CoV2. The genome of SARS-CoV2 contains two overlapping ORFs, which encode all non-structural proteins, and the structural proteins, which are organized in separate ORFs. Transcription regulating sequences (TRS) are labeled below the ORFs. A schematic virus particle with the structural proteins Spike (S), Envelope (E), Membrane (M), and Nucleocapsid (N) is depicted on the right. The image was adapted from Kim et al., 2020.

Since both proteases are essential for the formation of virus particles, they provide popular therapeutic targets and have been shown to be targeted by small molecule or peptide inhibitors (Owen et al., 2021; Shen et al., 2022; Shin et al., 2020). The proteases could also serve as targets for the generation of high-affinity Nbs (Sun et al., 2022). Potential Nanobody inhibitors might impede the formation of new virus particles (Armstrong et al., 2021). PL^pro is, even after proteolytic release, part of a larger, 215 kDa multi-domain enzyme with different catalytic functions (Klemm et al., 2020). Nbs could either be generated against the entire nsp3, or against PL^pro alone. Selecting Nbs against the entire nsp3 would decrease the likelihood of selecting Nbs specific for PL^pro, since the large protein probably provides many different epitopes. Binding and inhibiting functions of Nbs selected against PL^pro alone would require reproduction against the entire nsp3.

M^pro is a souble protein with a molecular mass of 34 kDa and is active as a homodimer (Goyal & Goyal, 2020; Jin et al., 2020). It is structurally (Figure 1.19) and biochemically well characterized. M^pro is highly conserved among different coronaviruses and can be expressed in *E. coli* (Goyal & Goyal, 2020; Jin et al., 2020). M^pro is a popular target for selection of peptide or Nb inhibitors (Qiao et al., 2021; Sun et al., 2022), which can be used to develop potent therapeutics. Because M^pro can be easily produced in *E. coli* and is not part of a larger muti-domain protein complex, we selected M^pro as our target for Nb selection.
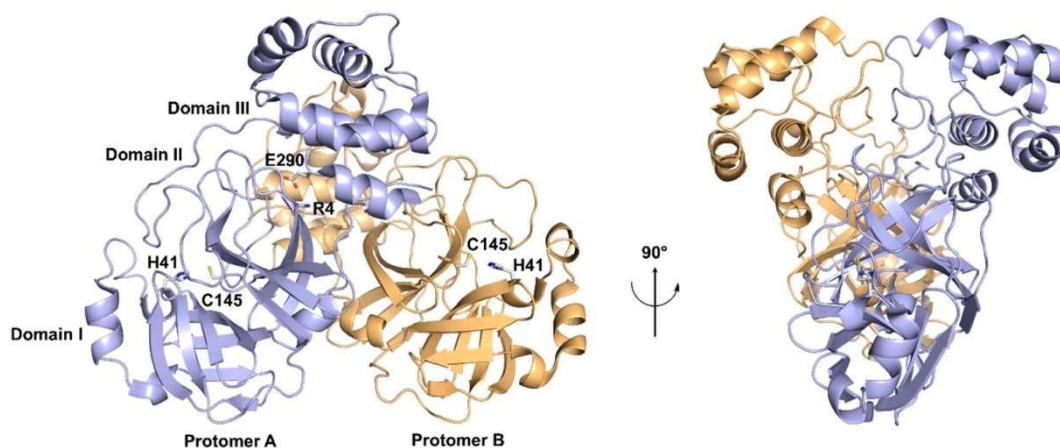
Figure 1.19 Protein structure of SARS-CoV2 main protease (M^pro^). Cartoon representation of the protein structure of a M^pro^ dimer with two protomers (blue and orange). Amino acids arginine 4 and glutamate 290 are essential for dimerization. Histidine 41 and cysteine 145 are located in the catalytic pocket. M^pro^ structure is presented in a front (left) and a side view (right). The image was adapted from Cannalire et al., 2022.

1.6 Aims of this study

1.6.1 Establishing mRNA/cDNA display for the selection of high-affinity Nbs from a novel high-variability DNA library

The main goal for this study is to establish a protocol for Nb selections by mRNA/cDNA display. Single reaction steps from published mRNA/cDNA display protocols will be adapted for selections of Nbs from the highly-variable c-lib Nb gene library, which was designed for this study. The elaborated protocol allows selection against a multitude of target antigens. Two proteins essential for retinal phototransduction, retGC1 and GCAP1, as well as the SARS-CoV2 main protease were chosen as target antigens. Selected Nbs could serve as tools for protein purification, structural and biochemical characterization, or as therapeutics.

1.6.2 Structural analysis of retGC1 and GCAP1 wild type and mutants and their roles in rare retinopathies

Several retGC1 and GCAP1 mutations are associated with retinopathies like LCA1 or adCORD. How these mutations alter photoreceptor signal transduction is not entirely characterized. To provide a deeper understanding of these diseases, retGC1 and GCAP1 wild type and mutants must be structurally and biochemically characterized. Especially for retGC1, little structural information is known, mainly because it has never been shown to be purified from a natural source or after heterologous expression. In the following, retGC1 will be expressed in different cell culture models, and a purification procedure will be established. If substantial amounts of retGC1 can be produced with high purity, Nbs will be selected against retGC1 by our established mRNA/cDNA display selection approach. Selected Nbs would then provide tools for retGC1 purification, and for structural analysis by cryo-EM.

Nbs will also be selected against different ion-bound forms of myristoylated GCAP1. Selected Nbs could then be used as analytical tools or to develop therapeutics. Because protein structures are not known for all ion-bound forms, myrGCAP1 wild type and the disease-related M26R mutant will be analyzed by X-ray crystallography. Obtained protein structures could provide a better understanding of GCAP1's role in phototransduction and in rare retinopathies.

2. Materials and Methods

2.1 Molecular cloning methods

Sequence handling and *in silico* cloning was performed using CLC Main Workbench 23.0.3 (Qiagen).

2.1.1   DNA amplification by PCR

Molecular cloning was performed either by restriction ligation cloning or Gibson Assembly. All DNA fragments were amplified by Q5 DNA Polymerase (New England Biolabs). PCR reaction mixes were prepared according to Table 2.1, and PCRs were performed according to following program (Table 2.2). The primers used for PCRs in this study are listed in the appendix (Table 6.5).

Table 2.1 Standard PCR reaction mix.

Component	Volume	Final concentration
Water	(23 – x) µl	
5x Q5 Polymerase buffer	10 µl	1x
5x Q5 GC-rich enhancer	10 µl	1x
dNTP mix (10 mM per dNTP)	1 µl	200 µM per dNTP
10 µM Primer 1	2.5 µl	500 nM
10 µM Primer 2	2.5 µl	500 nM
DNA template	x µl	
2 U/µl Q5 DNA Polymerase	0.5 µl	0.02 U/µl
Σ	50 µl	

Table 2.2 Standard PCR program.

Step	Temperature [°C]	Time	Cycles
Initial denaturation	**95**	3 min	
Denaturation	**95**	15 s	
Annealing	**Annealing temp (see table x)**	20 s	30 PCR cycles
Elongation	**72**	30 s/kb	
Final elongation	**72**	3 min	
End	**10**	Hold	

PCR products were purified by spin-column purification (GeneJET PCR purification kit, Thermo Fisher) according to manufacturer's instructions.

2.1.2 Restriction ligation cloning

For restriction ligation cloning, a linearized vector and a DNA insert must have complementary single-stranded overhangs, which can anneal at low temperatures and are then fused by a DNA ligase.

Vectors were linearized by restriction digest with either one or two restriction endonucleases. The inserts were prepared by PCR and subsequent restriction digest with the same restriction endonucleases or ones that generate equal sticky ends. Restriction digests were set up as follows, with N as number of restriction enzymes (Table 2.3).

Table 2.3 Standard reaction mix for a restriction digest of DNA with N restriction endonucleases.

Component	Volume	Final concentration
Water	$(36 - 3 \cdot N - x)$ µl	
1 µg Template DNA	x µl	25 ng/µl
10 x Digestion buffer	4 µl	1 x
20 U/ µl Restriction endonucleases	3 µl	1.5 U/µl
Σ	40 µl	

Digested products were separated from the excised fragments by agarose gel electrophoresis and purified from the gel by melting the desired gel fragment and spin-column purification (GeneJET Gel Extraction Kit, Thermo Fisher) according to manufacturer's instructions.

For the subsequent ligation, insert and vector are used with a 4:1 molar ratio, meaning the insert is used with a 3-fold molar excess (Table 2.4). T4 ligase buffer is stored as 5 µl aliquots to decrease the number of freeze-thaw cycles. After thawing, the buffer was checked for (ATP-) precipitate, which needs to be dissolved completely.

Table 2.4 Standard reaction mix of a ligation of two DNA fragments, one linearized vector and a digested insert fragment, by a T4 DNA ligase.

Component	Volume
Water	$(8 - y - x)$ µl
50 ng linearized vector	y µl
3-fold molar excess of insert	x µl
10x T4 ligase buffer	1 µl
400 U/µl T4 DNA ligase	1 µl
Σ	10 µl

The reaction mix was incubated at 16 °C overnight. The next day, the ligation mix was either transferred to -20 °C for storage or directly used to transform *E. coli* TOP10 cells.

2.1.3 Cloning by Gibson Assembly

For Gibson Assembly cloning, vector and insert(s) need to share equal overlaps of roughly 20 base pairs on both ends, which can be introduced by PCR.

Either the vector was linearized by restriction digest (Table 2.3) and both overlaps were introduced to the insert by PCR, or the vector was linearized by PCR and the overlaps were introduced to both, vector and insert by PCR (Table 2.1).

For the cloning reaction, 2x Gibson Assembly Master Mix (BioCat) was mixed with 50 ng of the linearized vector and a 3-fold molar excess of insert (Table 2.5).

Table 2.5 Standard reaction mix of a Gibson Assembly reaction.

Component	Volume
Water	(5 - y - x) µl
50 ng linearized vector	y µl
3-fold molar excess of insert	x µl
2x Gibson Assembly Master Mix	5 µl
Σ	10 µl

The reaction mix was incubated at 50 °C for one hour and either transferred to -20 °C for storage or directly used to transform *E. coli* TOP10 cells.

2.1.4 Site-directed mutagenesis

The gene of the GCAP1 M26R variant was generated by site directed mutagenesis. The single-point mutation was introduced by PCR amplification of the entire pStrep1(amp)-GCAP1 D6S plasmid using the primers GCAP1 M26R fw and GCAP1 M26R rev (Appendix Table 6.5).

PCRs were performed according to Table 2.1. After PCR, adenine-N6-methylated template DNA is removed by DpnI digest (Table 2.3). DpnI is a restriction endonuclease that cleaves methylated GATC sites, specifically.

This yields only plasmids that carry the mutated gene of interest, which are then used to transform *E. coli* TOP10 cells according to following protocol.

2.1.5 Transformation of chemically competent *E. coli*

Chemically competent *E. coli* were transformed by 42°C heat-shock. Therefore, previously prepared chemically competent cells were thawed (50 µl aliquots) and added to 10 µl of ligation reaction, Gibson Assembly mix, or 50 ng of an isolated plasmid, and incubated on ice for 30 min. Then, the 30 s, 42 °C heat-shock was applied, followed by another incubation on ice for 5'. As a next step, 900 µl of SOC-

medium (37 °C) were added and the transformation mix was incubated at 37 °C for 1 h while shaking at 200 rpm. In the end, the transformation mix was plated onto an LB-agar plate supplemented with proper antibiotics. The plate was incubated at 37 °C overnight.

2.2 Preparation of chemically competent *E. coli*

To enhance transformation efficiency, it is necessary to remove as much detergent as possible from used glassware.

As a first step, seed stocks were prepared. Therefore, *E. coli* cells were streaked onto an SOB-agar plate and diluted across the plate in order to obtain single colonies. Colonies were grown at 23 °C overnight.

The next day, overnight cultures of 2 ml SOB were inoculated with single colonies from the plate and grown at 23 °C overnight.

On day three, glycerol was added to a concentration of 15 %, 1 ml aliquots were prepared in Nunc™ cryovials (VWR), frozen in a dry ice/ethanol bath and stored at -80 °C.

250 ml of SOB medium were inoculated with one 1 ml seed stock vial and grown at 16 °C until an OD_{600} of 0.3 was reached, which took roughly 18 hours. Cells were harvested by centrifugation at 4 000 rcf, 4°C for 15 min. The supernatant was discarded, the cell pellet was re-suspended in 80 ml of ice cold CCMB80 buffer (Appendix Table 6.2), incubated on ice for 20 min and centrifuged again. After centrifugation, the cell pellet was re-suspended in 10 ml of ice cold CCMB80 (Appendix Table 6.2) and the OD_{600} was measured to estimate cell concentration. The cell suspension was diluted to an OD_{600} of 5 and incubated on ice for another 20 min. After incubation, 50 µl aliquots were prepared, snap frozen in liquid nitrogen and stored at -80 °C.

2.3 Heterologous expression and purification of target proteins

2.3.1   Expression and purification of myristoylated GCAP1 wildtype and mutants

All GCAP1 constructs contain a D6S mutation, which enables *in vivo* myristoylation by a yeast N-myristoyl transferase. To synthesize myristoylated GCAP1 proteins, *E. coli* was co-transformed with the pNMT plasmid (Carlson & Hurley, 2012) carrying the gene for the N-myristoyl transferase and a pStrep1-GCAP1 plasmid containing the gene for GCAP1 with a Strep-tag.

The pStrep1-GCAP1 vector was generated by restriction ligation cloning of the GCAP1 gene from a pET22b-GCAP1 plasmid. This is an expression vector containing the GCAP1 gene with a C-terminal hexa-histidine-tag (His-tag). Because the His-tag is incompatible with some downstream processes, a GCAP1 construct with a C-terminal Strep-tag was produced. The pET22b-GCAP1 and pStrep1 vectors

were digested with XbaI and XhoI restriction endonucleases at 37 °C overnight and then ligated by T4 DNA ligase at 16 °C overnight.

The pStrep1 and pNMT plasmids both carry the gene for kanamycin (kan) resistance and are therefore incompatible. A pStrep1/amp-GCAP1 (wildtype) were generated by restriction ligation cloning utilizing the DraII restriction sites in the pStrep1-GCAP1 and pET22b vectors and subsequent T4 DNA ligase, as mentioned.

E. coli BL21(DE3) codon plus RIL cells were co-transformed with the pStrep1/amp and pNMT-vectors by 90 s, 42 °C heat-shock and plated onto an LB-agar plate supplemented with 100 mg/l ampicillin (selection marker for the pStrep1/amp plasmid), 50 mg/l kan (selection marker for the pNMT plasmid), 35 mg/l chloramphenicol (chl) (selection marker for the BL21 codon plus RIL cells), and 1 % glucose.

Transformed cells were grown on a LB agar plate supplemented with 100 mg/l ampicillin, 50 mg/l kan, 35 mg/l chl, and 1 % glucose at 37 °C overnight. The next day, 100 ml of 2 YT medium supplemented with ampicillin, kan, chl and glucose at the mentioned concentrations were inoculated with single colonies from the agar plate and grown overnight at 37 °C, while shaking at 200 rpm.

On day three, 4 – 6 l of 2 YT supplemented with ampicillin, kan, chl, and 0.1 % glucose were inoculated with the overnight culture to an OD_{600} of 0.1 and grown at 37 °C, 200 rpm. Once an OD_{600} of 0.5 was reached, myristic acid was added to a concentration of 100 mg/l and protein expression was induced by adding IPTG (isopropyl-β-D-thiogalactopyranoside) to a concentration of 1 mM. Proteins were expressed at 18 °C for 16 to 18 hours.

After protein expression, cells were harvested by centrifugation at 4 000 rcf, 4 °C for 15 min, then washed in 100 ml ice cold HBS and centrifuged again. Pellets were snap-frozen in liquid nitrogen and stored at -80 °C until further use.

Cell pellets were re-suspended in 15 ml of re-suspension buffer per gram of cell pellet. Cells were lysed by sonication (Bandelin Sonopuls HD2200) and the lysate was cleared by centrifugation at 20 000 rcf, 4 °C, for 20 min and then by ultracentrifugation at 100 000 rcf, 4 °C, for 1 h. The cleared lysate was loaded onto a Strep-Tactin (IBA) affinity chromatography column. After the column was washed, proteins were eluted by applying elution buffer, which contains 6 mM Desthiobiotin.

To exchange buffers and to separate purified proteins by their hydrodynamic radius, size-exclusion chromatography (SEC) was performed using an ÄKTA FPLC (Cytiva) and a Superdex® 75 10/300 column.

2.3.2   Heterologous retGC1 expression in *sf21* insect cell culture

retGC1 was expressed heterologously in *sf21* insect cells and in a TREx293 cell line, which stably integrated the gene coding for retGC1.

Before infection, a pFL-vector containing the retGC1 gene (GUCY2D) was generated by restriction ligation. The pFL-vector and PCR amplified inserts were digested by the restriction endonucleases *EcoRI* and *XbaI* and were ligated by a T4 DNA ligase.

E. coli EMBacY cells were transformed with the MultiBac fusion plasmid pFL-retGC1-CStrep by 42 °C heat-shock to generate the first generation of Bacmids. Successful transformation was assessed by blue-white screening. Bacmids were generated by inoculating 2 ml of LB medium supplemented with kan (50 µg/ml), tetracyclin (tet) (10 µg/ml), gentamycin (10 µg/ml) and growing the culture to saturation. Cells were harvested by centrifugation and lysed by alkaline lysis. Bacmids were precipitated using isopropanol.

Precipitated and solubilized Bacmids were used to transfect *sf21* insect cells, growing on a 6-well plate, using a X-tremeGENE HP (Roche) transfection reagent. After incubation for 60 h, the supernatant was collected and used for virus amplification.

To generate virus suitable for infection of an expression culture, $12.5 \cdot 10^6$ cells in 25 ml were infected with 3 ml of prepared Bacmid. The virus propagation culture was grown as a suspension culture at 28 °C. 60 h after proliferation arrest, cells were pelleted, and the supernatant was collected and used to infect a protein expression culture.

An expression culture with a cell concentration of $0.5 \cdot 10^6$ was infected with the previously generated virus at a ratio of 1:70 (virus:culture volume). Cell concentration was monitored every day, and after 48 h after proliferation arrest, cells were harvested by centrifugation at 300 rcf, 4 °C, for 20 min, washed in PBS, and centrifuged again. Cell pellets were shock frozen in liquid nitrogen and stored at -80 °C.

2.3.3 Establishing a stable retGC1-harboring TREx 293 cell line

The TREx 293 cell lines, which stably integrated the GUCY2D gene coding for the human retGC1, were generated in our lab by Christine Gotthold and Magdalena Schacherl according to the Flp-In™ System (Thermo Fisher).

2.3.4 retGC1 expression in TREx 293 cells

TREx 293 cells were seeded in cell culture flasks at a density of $5 \cdot 10^4$ cells/cm^2 and grown until 80 % confluence. Once the desired confluence was reached, retGC1 expression was induced by supplementing tet to a concentration of 0.1 µg/ml. Proteins were expressed for 48 h.

Cells were harvested by rinsing with ice-cold PBS and subsequent centrifugation at 300 rcf for 20 min. The cell pellet was washed in ice-cold PBS, centrifuged again, snap-frozen in liquid nitrogen and stored at -80 °C.

2.3.5   Purification of human retGC1 from *sf21* or TREx 293

Cell pellets containing retGC1 were re-suspended in 4 ml of re-suspension buffer per gram of cell pellet. Cells were lysed by sonication using a Sonopuls HD2200 sonicator (Bandelin) and the lysate was cleared by centrifugation at 20 000 rcf, 4 °C, for 20 min. Cell membranes were pelleted by ultracentrifugation at 200 000 rcf, 4 °C, for 1 h. Membranes were solubilized in a buffer containing detergent, SMA (styrene/maleic acid alternating co-polymer), or DIBMA (Diisobutylene/maleic acid alternating co-polymer) at 4 °C for 1 h. Un-solubilized membranes and proteins were pelleted by ultracentrifugation at 200 000 rcf, 4 °C, for 1 h. After ultracentrifugation, the supernatant was loaded onto a Strep-Tactin (IBA) affinity chromatography column. After the column was washed, proteins were eluted by applying elution buffer, which contains 6 mM Desthiobiotin.

To exchange buffers and to separate purified proteins by their hydrodynamic radius, size-exclusion chromatography (SEC) was performed using an ÄKTA FPLC (Cytiva) and a Superose 6 10/300 SEC column (Cytiva).

2.3.6   Heterologous expression and purification of SARS-CoV2 main protease Mpro

SARS-CoV2 M^{pro} was expressed with an N-terminal GST-tag and a C-terminal Strep-Tag in *E. coli* BL21(DE3) cells.

Therefore, chemically competent *E. coli* BL21(DE3) were transformed with a pGEX4-M^{pro}-Strep vector by 42 °C heat-shock and plated onto an LB-agar plate supplemented with 100 mg/l ampicillin and 1 % glucose and incubated at 37 °C overnight.

On day two, a 100 ml pre-culture was prepared by inoculating 100 ml LB supplemented with ampicillin and 1 % glucose with a single colony from the agar plate and grown overnight at 37 °C, 200 rpm.

Then, the pre-culture was expanded to a 4 l expression culture, by inoculating 4 l of LB medium supplemented with ampicillin and 0.1 mM glucose to an OD_{600} of 0.1. The expression culture was grown at 37 °C, 200 rpm, until an OD_{600} of 0.6 was reached. Then, M^{pro}-expression was induced by the addition of IPTG to a concentration of 0.2 mM IPTG. M^{pro} was expressed at 16 °C for 16 hours.

On the final day, cells were harvested by centrifugation at 3500 rcf, 4 °C, for 15 min. The cell pellet was re-suspended in 100 ml of ice-cold HBS. Cells were pelleted again, snap frozen in liquid nitrogen and stored at -80 °C.

Cell pellets containing SARS-CoV2 M^{pro} were re-suspended in 15 ml of re-suspension buffer per gram of cell pellet. Cells were lysed by sonication (Bandelin Sonopuls HD2200) and the lysate was cleared by centrifugation at 20 000 rcf, 4 °C, for 20 min and then by ultracentrifugation at 100 000 rcf, 4 °C, for 1 h. The cleared lysate was loaded onto a Strep-Tactin (IBA) affinity chromatography column. After the

column was washed, proteins were eluted by applying elution buffer, which contains 6 mM Desthiobiotin.

To exchange buffers and to separate purified proteins by their hydrodynamic radius, size-exclusion chromatography (SEC) was performed using an ÄKTA FPLC (Cytiva) and a Superdex 200 10/300 column (Cytiva).

2.3.7 Heterologous expression and purification of Nbs

Nbs were heterologously expressed in *E. coli* wk6 cells from a pHen6-vector, which contains an N-terminal PelB-sequence. This allows the Nbs to be transported into the periplasm, co-translationally. The pHen6-Nb expression vector was generated by restriction ligation or Gibson Assembly. For restriction ligation, inserts were produced by PCR and subsequent restriction digest and ligated into a digested pHen6-vector via NcoI and XhoI restriction sites. For Gibson Assembly, the same digested vector was used to avoid introducing mutations into the vector by PCR. Overhangs necessary for Gibson Assembly were introduced by PCR. Gibson Assembly was performed as mentioned.

Selected Nbs carrying a frameshift in the first framework regions were cloned into a pHen6-Nb vector by Gibson Assembly using primers. This generated a pHen6 vector, which contains a full-length and in-frame Nb gene.

Chemically competent *E. coli* wk6 were transformed with a pHen6-vector containing a Nb gene by 42 °C heat-shock and plated onto LB-agar plates containing 100 mg/l ampicillin, 1 % glucose and 1 mM $MgCl_2$ and grown at 37 °C overnight.

Pre-cultures for Nb expression were prepared in TB-medium, supplemented with ampicillin, 1 % glucose and mM $MgCl_2$, and grown at 37 °C, 200 rpm, overnight.

4 l expression cultures were set up in TB medium containing ampicillin, 0.1 % glucose, and 1 mM $MgCl_2$ with an OD_{600} of 0.1 and grown to an OD_{600} of 0.7. Protein expression was induced by 1 mM IPTG. Temperature was reduced to 28 °C and Nbs were expressed for 16 to 18 hours.

Cells were harvested by centrifugation, washed in TBS and, after another centrifugation, snap-frozen in liquid nitrogen and stored at -80 °C.

Cell pellets were re-suspended in TES buffer (Table 6.2) to extract the periplasm. 15 ml of buffer were used for a pellet from 1 l of expression culture with an OD_{600} of 20. The cell suspension was rotated at 4 °C for 1 h. Then, the 30 ml of TES/4 buffer were added followed by another rotating step at 4 °C of 1 h. Periplasm and spheroplasts were separated by centrifugation at 7500 rcf, 4 °C for 30 min. The supernatant (periplasmic fraction) was loaded onto an affinity chromatography column containing 1 ml of Ni-NTA Superflow resin (Qiagen), which was pre-equilibrated in Nb affinity chromatography buffer 1 (Table 6.2). After the column was washed with 20 ml of Nb affinity chromatography buffer 1

(Table 6.2) and 40 ml of Nb affinity chromatography buffer 2 (Table 6.2), Nbs were eluted in 4 ml Nb affinity Elution buffer (Table 6.2) by gravity flow and 1 ml of Tris-HCl pH 7.5 were added, immediately. The elution fraction was concentrated to 2 ml using a FFF filter unit with a MWCO (molecular weight cut-off) of 3 kDa. Then, SEC was performed using an ÄKTA FPLC (Cytiva) and a Superdex 75 pg 10/300 SEC column. Purified Nbs were stored at 4 °C until further use.

2.4 Preparation of Nb libraries DNA for selections by mRNA/cDNA display

Our own designed, high-variability Nb DNA library was synthesized by the GeneArt gene synthesis service (Thermo Fisher). The Nb library was delivered in the form of purified, lyophilized DNA and was, after reconstitution in water, ready to use for PCR.

2.4.1   Preparation of the y-lib Nb gene library from *Saccharomyces cerevisiae*

In contrast to our own combinatorial Nb DNA library, which was synthesized upon our design and delivered in from of pure DNA, the commercial y-lib (McMahon et al., 2018) was delivered in form of $2 \cdot 10^{10}$ transformed yeast cells, which equals 40 library sizes.

2.4.1.1 Expansion of the yeast culture

In order to use this library for PCR, those plasmids that carry the Nb genes had to be purified from the yeast cells in high quantity and purity. A protocol therefore had first to be established. Common plasmid preparation protocols from yeast usually yield either small amounts of plasmid DNA or lots of contaminations of genomic DNA.

As a first step, the yeast cells were expanded in order to obtain aliquots of at least $5 \cdot 10^{10}$ cells. Therefore, the cell pellet was thawed and transferred to 1 l of $Y_{Glc}4.5_{-Trp}$ medium in a 2 l baffled flask, which was previously sterilized by baking at 180 °C. Because the pYDS-Nb vector, which carries the different Nb genes, contains the Trp operon, a Tryptophan-lacking growth medium was used to apply selective pressure. The culture was incubated overnight at 28 °C, while shaking at 200 rpm.

The next day, the culture was expanded to 4 l in four baffled flasks and incubated another day at 28°C, 200 rpm.

On day three, OD_{600} (optical density at a wavelength of 600 nm) was measured to estimate cell concentration of the culture. 1 OD_{600} unit correlates to roughly $1.5 \cdot 10^7$ cells/ml. Cells were harvested by centrifugation at 3 500 rcf, 4 °C for 15 min. Supernatant was decanted and the cell pellet was re-suspended in the residual liquid. Volume was measured and DMSO was added to a concentration of 10 % to achieve the conditions of the freezing medium. A total cell number of $6 \cdot 10^{10}$ was harvested. One aliquot of $5 \cdot 10^{10}$ cells and three of $5 \cdot 10^9$ cells were prepared.

2.4.1.2 Specific purification of Nb containing plasmids

One aliquot of transformed yeast cells was thawed and re-suspended in 30 ml of buffer P1 (Qiagen Plasmid Maxi kit). To remove the cell wall from the yeast cells and obtain spheroblast, 150 mg zymolyase (Biozol) were solubilized in 1 ml of buffer P1 and added to cell suspension, resulting in a zymolyase concentration of 5 mg/ml. Cell walls were digested at 37°C for 2 h, while shaking at 150 rpm. After 2 h, a white precipitate formed, which was separated from the supernatant by centrifugation at 2 000 rcf, 4°C for 10 min. Plasmid DNA was successfully purified from both supernatant and pellet, if treated separately. However, a lot of genomic DNA contamination was obtained when combining both fractions for further treatment.

Cells were lysed by adding 30 ml of Lysis buffer (Qiagen Plasmid Maxi kit). After 5 min, 30 ml of Neutralization buffer (Qiagen Plasmid Maxi kit) were mixed in to stop the lysis. A white precipitate formed containing cell debris and majority of the genomic DNA, as usual in plasmid preparations. The precipitation was removed by centrifugation at 20 000 rcf, 4°C for 10 min. This step was repeated until the supernatant was entirely clear. The supernatant was mixed with 30 ml of Binding buffer (Qiagen Plasmid Maxi kit) and loaded onto a spin-column (Qiagen Plasmid Maxi kit) via a vacuum manifold and applying a weak 300 mbar vacuum. After loading, the column was washed twice with 500 µl Wash buffer (Qiagen Plasmid Maxi kit). The columns were dried by centrifugation (in a 1.5 ml Eppendorf tube) at 15 000 rcf for 2 min. To elute the DNA from the columns, 50 µl of water were applied directly onto the filters, incubated for 1' and then centrifuged at 13 000 rcf for 1 min. This procedure yielded up to 15 µg of relatively pure plasmid DNA.

2.5 Selection of Nbs by mRNA/cDNA display

2.5.1   Library amplification and sequence augmentation by PCR

As a first step, the Nb-DNA libraries were amplified by PCR according to Table 2.1 and Table 2.2. During PCR, affinity tags, Shine-Dalgarno sequence, T7 promoter, and an annealing site for a Puromycin-containing oligonucleotide were added to the Nb-DNA libraries.

PCR products were purified by spin-column purification (GeneJET PCR purification kit, Thermo Fisher) according to manufacturer's protocol. By preparing eight 50 µl PCR reactions, the rection volume was increased to 400 µl. This enabled the use of a library size of 10^{13} Nb sequences.

To ensure that only DNA fragments of the proper sizes were used in mRNA/cDNA display, the entire PCR product was used in agarose gel electrophoresis, where the fragment of the desired sizes were extracted from the agarose gel using a GeneJET gel extraction kit (Thermo Fisher) according to manufacturer's instruction.

2.5.2   Nb mRNA generation by *in vitro* transcription

After PCR, the Nb-DNA libraries were transcribed into mRNA, *in vitro,* using a T7 RiboMAX™ Express Large Scale RNA Production System (Promega) (Table 2.6).

Table 2.6 Standard *in vitro* transcription reaction mix for mRNA synthesis by a T7 RNA Polymerase.

Component	Volume
RNase-free water	(16 – x) µl
2x T7 express buffer	20 µl
500 ng DNA	x µl
10 xT7 enzyme mix	4 µl
Σ	40 µl

The transcription mix was incubated at 37°C for 2-4 h, then 4 µl of Turbo DNase (2 U/µl) (ThermoFisher) were added to the transcription reaction to remove template DNA at 37°C for 10 min.

mRNA was purified by spin-column purification (Monarch® RNA Cleanup Kit (50 µg), New England BioLabs) according to the manufacturer's protocol. This protocol yielded up to 150 µg mRNA. Therefore, three spin-columns were used. The mRNA was eluted from the columns by applying 100 µl of RNase-free water (70°C) for one minute.

2.5.3   Fusion of mRNA to a Puromycin-containing oligonucleotide by Psoralen-mediated UV-crosslink

After *in vitro* transcription, the Nb-mRNA was linked to a Puromycin-containing oligonucleotide (Figure 2.1) by Psoralen-mediated UV crosslink (Table 2.7). The oligonucleotide with the sequence depicted in Figure 2.1 was synthesized by Keck Oligo Synthesis facility.

Table 2.7 Reaction mix for the Psoralen-mediated UV-crosslink of Nb mRNA and the Pso-Pur oligonucleotide.

Component	Volume
RNase-free water	(116 – x) µl
10 x XL buffer	15 µl
60 µg mRNA	x µl
3.5x molar excess of 100 µM Pur oligonucleotide	19 µl (12.7 µM final concentration)
Σ	150 µl

The crosslink reaction mix was incubated at 70°C for 3 minutes and then cooled down to 25°C with a cooling rate of 3°C/s to anneal the Psoralen and Puromycin-containing oligonucleotide to the mRNA.

Psoralen-mU-mA-mG-mC-mC-mG-mG-mU-mG-A-A-A-A-A-A-A-A-A-A-A-A-A-A-A-Sp$_{18}$-A-C-C-Puromycin

Figure 2.1 Puromycin-containing oligonucleotide. Representation of the oligonucleotide sequence of the Psoralen- and Puromycin-containing oligonucleotide used in the UV crosslink reaction with mU, mA, mG, and mC as methylated ribonucleotides, A and C as unmethylated ribonucleotides, and Sp$_{18}$ as a linker containing 18 ethylene glycol units.

Then, the reaction mix was transferred to a Costar 96-well, round-bottom, plate (75 µl per well). The oligonucleotide was cross-linked to the mRNA by excitation at 365 nm for 20 min. For the excitation, a 5 W UV-lamp was used with a distance of 2.5 cm to the top of the plate. The plate was placed in ice water to prevent the reaction from heating up.

2.5.4 Purification of cross-linked mRNA by gel extraction and Oligo(dT) affinity

After UV-crosslink, the RNA was purified by spin-column purification (Monarch® RNA Cleanup Kit (50 µg), New England BioLabs) according to the manufacturer's protocol.

To remove the excess of Psoralen-Puromycin oligonucleotide from the cross-linked mRNA (XL-mRNA) and any byproducts, the XL-mRNA was purified from a denaturing 7 M urea, 9.5% polyacrylamide gel. After electrophoresis, a gel fragment at to the desired fragment size was excised from the gel and nucleic acids were purified from the gel by electro elution and subsequent spin-column purification.

Because mRNA and XL-RNA show similar electrophoretic behavior, they were co-purified. In a later step, mRNA and XL-mRNA were separated by an oligo(dT)-purification step. This is possible because the Psoralen-Puromycin-oligonucleotide carries a 15x-adenine tail.

For the electro elution, the gel fragment was divided into small pieces and transferred into a dialysis tube, which was pre-equilibrated with TBE buffer and closed on one end. 800 µl TBE buffer were pipetted on top, air bubbles were removed, and the dialysis tube was closed on the remaining open end. The dialysis tube was put into an agarose submarine electrophoresis chamber filled with TBE buffer. By applying a voltage of 120 V for 1 h, nucleic acids were eluted from the gel fragments.

XL-mRNA and mRNA were purified by spin-column purification (Monarch® RNA Cleanup Kit (50 µg), New England BioLabs) according to the manufacturer's protocol. This protocol was expected to yield up to 150 µg mRNA. Therefore, three spin-columns were used. The mRNA was eluted from the columns by applying 50 µl of RNase-free water (70°C) for one minute.

For the oligo(dT)-purification of one 60 µg crosslink reaction, 20 µl 50% slurry were used. The settled resin was washed four times in 500 µl oligo(dT)-binding buffer. The supernatant over the settled beads should then be clear. The electro-elution fraction was then mixed with 450 µl of oligo(dT)-binding

buffer and bound to the beads for 20 minutes while rotating. After binding, the slurry was transferred to Costar® Spin-X® Centrifuge Tube Filters (0.22 µm Pore CA Membrane). The spin-columns were cleared by centrifugation at 800 rcf for 1 min. The beads were retained in the column and were washed three times with oligo(dT)-wash buffer. In the end, XL-mRNA and XL-mRNA-protein hybrids were eluted with 30 µl oligo(dT)-elution buffer, which was pre-heated to 70°C.

2.5.5   Production of Nb protein-mRNA hybrid molecules by *in vitro* translation

For *in vitro* synthesis of proteins and mRNA-protein hybrid molecules, the PURE method (Protein synthesis Using Recombinant Elements) was chosen, using the PURExpress® protein synthesis kit (New England BioLabs) (Table 2.8).

Table 2.8 Standard *in vitro* translation reaction mix. For Fluorotect-labeling of the nascent protein chain, the reaction mix was adjusted (green).

Component	Volume
RNase-free water	$(14.6 - x)$ µl $(14.1 - x)$ µl
Solution A	30 µl
Solution B	22.5 µl
40 U/µl RNase inhibitor	1.9 µl
25x Disulfide bond enhancer 1	3 µl
25x Disulfide bond enhancer 2	3 µl
XL-mRNA	x µl (up to 250 µM)
(fluorotect-Lys)	0.5 µl
Σ	75 µl

Proteins were synthesized at 37°C for 2 h. After translation, the reaction mix was incubated at room temperature for 15 minutes. To enhance hybrid formation, KCl and $MgCl_2$ were added to final concentrations of 375 mM and 37.5 mM, respectively, followed by another incubation at room temperature for 15 min.

In the next step, mRNA-protein hybrids (and untranslated XL-mRNA) were purified over POROS™ oligo(dT) affinity resin (Thermo Fisher) to remove un-displayed proteins. Because the Psoralen-Puromycin-oligonucleotide carries a 15x-adenine tail, the cross-linked mRNA can bind the resin.

For the oligo(dT)-affinity-purification of one 75 µl translation reaction, 10 µl 50% slurry were washed four times in 500 µl oligo(dT)-binding buffer. The supernatant over the settled beads should then be clear. The translation reaction was then mixed with nine times its volume of oligo(dT)-binding buffer and bound to the beads for 20 minutes while rotating. After binding, the slurry was transferred to Costar® Spin-X® Centrifuge Tube Filters (0.22 µm Pore CA Membrane). The spin-columns were cleared by centrifugation at 800 rcf for 1 min. The beads were retained in the column and were washed

three times with oligo(dT)-wash buffer. In the end, XL-mRNA and XL-mRNA-protein hybrids were eluted with 30 µl oligo(dT)-elution buffer, which was pre-heated to 42°C.

2.5.6 Generation of mRNA/cDNA-protein hybrid molecules by reverse transcription

To enhance mRNA stability, the Nb-RNA library was reverse transcribed by a SuperScriptIV reverse transcriptase (Thermo Fisher) (Table 2.9). This step yields the Nb mRNA-cDNA-protein hybrid molecule, which can be used for selections.

Table 2.9 RNA-dNTP-primer mix to use in a reverse transcription reaction with a SuperscriptIV reverse transcriptase.

Component	Volume
RNase-free water	$(22 - x)$ µl
Template RNA	x µl
dNTP mix (200 µM per dNTP)	2 µl
2 µM RT Primer	2 µl
Σ	26 µl

The RNA-dNTP-Primer mix was prepared, incubated at 42 °C for 5 min and then on ice for 1 min. In the meantime, the cDNA synthesis containing reaction buffer DTT RNase inhibitor and the reverse transcriptase was prepared (Table 2.10). In the final Nb selection against SARS-CoV2 M^{pro} was removed to improve the stability of the disulfide bond containing Nbs. The DTT volume was substituted with RNase-free water.

Table 2.10 cDNA-synthesis mix for reverse transcription reactions.

Component	Volume
SuperScriptIV buffer	8 µl
DTT	2 µl
RNase OUT	2 µl
SuperScriptIV reverse transcriptase	2 µl
Σ	14 µl

RNA-dNTP-primer-mix and cDNA-synthesis mix were combined and incubated at 42 °C for 1 h and terminated by putting the reaction mix on ice.

2.5.7 Selection of full-length Nb proteins by purification over Ni-IDA

To ensure that only full-length Nb proteins were used in selection experiments, synthesized proteins were selected for a C-terminal His_6-tag.

The reverse transcription product was mixed with an equal volume of His-binding buffer (Appendix Table 6.2) and bound to pre-equilibrated PureCube Ni-IDA MagBeads (Cube Biotech) for 1 h. Then, the beads were washed three times with His-wash buffer (Appendix Table 6.2). After washing, the His-

tag containing mRNA-cDNA-protein hybrid molecules were eluted by applying His-elution buffer (Appendix Table 6.2) for 20 min.

2.5.8 *In vitro* selection of Nbs against their target antigen

All steps during selection were conducted in RNase-free protein low-bind tubes. All buffers used during counter selection and Nb selection are listed in the appendix (Table 6.2)

To reduce the number of bead-binding Nb sequences before the actual selection against the target protein, a counter-selection against empty MagStrep "type3" (iba lifesciences) Strep-Tactin magnetic beads was performed, after anti-His selection or directly after reverse transcription. Therefore, the reverse transcription reaction or the His-elution fraction was mixed with an equal volume of Selection Binding buffer and bound to 1 µl of pre-equilibrated Strep-Tactin beads (20 µl of 5% slurry) and incubated for 1 h at room temperature while rotating.

As a next step, the counter-selection supernatant was mixed with 0.15 nmol of target protein in 50 µl Selection Binding buffer to form Nb-target complexes in solution. Complex formation was performed at room temperature for 1 h while rotating.

During complex-formation, 0.5 µl Strep-Tactin beads (10 µl slurry) were pre-equilibrated in 200 µl Selection Binding buffer supplemented with BSA and DNA to concentrations of 2 µg/ml and 10 µg/ml, respectfully. After 1 h, the Strep-tagged target proteins were bound to the pre-equilibrated Strep-Tactin magnetic beads at room temperature for 4 min, while rotating. Binding time was optimized to reduce the number of bead-binders while ensuring efficient target binding to the beads. After 5 min, beads and supernatant were separated by a magnetic separator. The supernatant, which contains all the non-target binding Nbs, as well as some complexes that did not bind to the Strep-Tactin within the 4 min binding period, was transferred to a fresh tube for analysis. Beads were washed once with Selection binding buffer plus BSA, then three times with Selection wash 1 buffer, and in the end twice with Selection wash 2 buffer. Strep-tagged target proteins were eluted from the beads by applying 30 µl of Selection elution buffer and rotating at room temperature for 15 min.

Nb sequences in the elution fractions can then be amplified via PCR for either another selection cycle or sub-cloning into an expression vector.

2.6 Preparation of myrGCAP1 protein crystals for X-ray crystallography

In an attempt to determine the protein structure of myrGCAP1 wild type and M26R as apo protein, in Mg-bound and Ca-bound form, protein crystals were grown.

After myrGCAP1 wild type and M26R were SEC purified in three different sets of buffers (supplemented with EDTA/EGTA, $MgCl_2$, or $CaCl_2$), crystals screens were set up for these three conditions. Protein concentrations between 5 and 10 mg/ml were used for the initial screens Index

(Hampton Research), JCSG++ (Jena Bioscience), SaltRx (Hampton Research), and PEG/ion (Hamton Research). Crystals were grown by sitting drop (300 nl droplets; protein:precipitant ratio 1:1) vapor diffusion at 18 °C for two to three weeks.

For one condition (myrGCAP1-Ca, JCSG++ H2), optimization screens were setup. 1.5 µl protein (10 mg/ml) was mixed with 1.5 µl precipitant solution and crystals were grown again by sitting drop vapor diffusion at 18 °C for 3 weeks. A 24-well grid screen was set up around the original condition (100 mM BIS-TRIS pH 5.5, 1% PEG 3350, 1 M ammonium sulfate) with four different pH-values (5.0 – 6.0) and six ammonium sulfate concentrations (750 mM – 1.375 M). A PEG 3350 concentration of 1 % was used in all conditions.

2.7 Analytical methods

2.7.1 Detection of nucleic acids and/or proteins

2.7.1.1 In gel detection of DNA

DNA was detected in agarose or polyacrylamide gels by SYBR safe (Thermo Fisher) staining. During gel preparation, SYBR safe was added to the gel mix at a concentration of 1:10 000.

2.7.1.2 In gel detection of RNA

RNA was detected in agarose or polyacrylamide gel by SYBR safe staining, as mentioned, or by Toluidine blue staining. After electrophoresis, 30 ml of Toluidine blue staining solution were added to the gel and incubated for 15 min, while shaking. The gel was de-stained by shaking in demineralized water.

2.7.1.3 In gel detection of proteins

For visualization in SDS-PAGE, proteins were stained unspecifically with Quick Coomassie Stain (Protein Ark). Gels were shaken in 30 ml of staining solution for 30 min. Excess staining solution in the gel was removed by shaking the gel in de-mineralized water for at least 2 h.

Proteins, which were synthesized in in vitro translation reactions using the PURExpress® in vitro protein synthesis kit (New England Biolabs), were detected by in gel fluorescence of FluoroTect™ GreenLys in vitro Translation Labeling System (Promega). Per 100 µl translation reaction, 1 µl of FluoroTect™ was supplemented. The fluorescently labeled lysine could be detected using a Typhoon™ laser-scanner platform (Cytiva). The fluorescent dye, BODIPY-FL, was excited at 488 nm and emission was detected using a 520 nm ± 20 nm bandpass filter.

2.7.1.4 Specific protein detection by Western blot

Semi-dry Western blot was the method of choice to detect proteins specifically. After SDS-PAGE, proteins were transferred to a PVDF (polyvinylidene fluoride) (Merck Millipore) membrane. 3 mm Gel

Blotting Papers (Whatman), the PVDF membrane and the polyacrylamide gel were soaked in buffer, and a blotting sandwich was assembled, with the anode on the bottom, then two blotting papers, then the PVDF membrane, the gel, two blotting papers, and the cathode on top. Proteins were transferred with a current of 200 mA. After transfer, the membrane was blocked with 10 ml of Immobilon® Block-CH (Millipore) by shaking for 1 h. Membranes were washed three times with TBST for 5 min each. Then, the primary Ab was applied in TBST at the desired concentration (1:500 – 1:5 000) for 1 h, while shaking.

If the primary Ab was already conjugated to horseradish peroxidase (HRP), the membrane was immediately used for protein detection by enhanced chemiluminescence (ECL). If not, the membrane was washed again, as mentioned, and then the secondary Ab was applied in TBST at a desired concentration (1:5 000 – 1:10 000) for 1 h.

For ECL detection, the membrane was washed once with TBST and twice with TBS. Then, the ECL solution was applied for 10 s and chemiluminescence was detected by an AI600 chemiluminescent imager (GE Healthcare).

2.7.1.5 Specific detection of Puromycin-labeled mRNA by immuno-Northern blot

mRNA, which was UV-crosslinked to a Puromycin-containing oligonucleotide, was specifically detected in an anti-Puromycin immuno-Northern blot. At first, RNA was separated according to its size in denaturing urea-PAGE using 7 M urea, 9.5 % polyacrylamide gels in TBE buffer. After electrophoresis, RNA was transferred to a positively charged Nylon membrane (Supplier). Therefore, a blotting sandwich was prepared with the anode on the bottom, then two Gel Blotting Papers (Whatman) soaked in 0.5x TBE (TBE buffer diluted 1:1 in filtered water), then the Nylon membrane and the polyacrylamide gel, which were both soaked in 0.5x TBE, again two soaked Gel Blotting Papers and the cathode on top. RNA was transferred at a constant current of 250 mA for 20 min. After transfer, RNA was crosslinked to the membrane by UV-irradiation at 256 nm with a distance of 5 cm for 5 min. After crosslink, the membrane was blocked in TBST + 5 % milk powder for 1 h and washed three times with 10 ml TBST for 5 min, each. Then the anti-Puromycin Ab was applied, diluted 1:2500 for 1 h. After incubation with the anti-Puromycin Ab, the membrane was washed again three times in TBST, and the secondary anti-mouse-HRP conjugated Ab 1:5000 for 1 h. For ECL detection, the membrane was washed once with TBST and twice with TBS. Then, the ECL solution was applied for 10 s and chemiluminescence was detected by AI600.

2.7.2   **Differential light scanning fluorimetry**

Protein denaturation temperatures were determined by differential scanning fluorimetry (DSF) using a Prometheus NT.48 (Nanotemper). Protein autofluorescence was measured at 330 and 350 nm while

temperature was increased from 20 °C to 95 °C with a rate of 2 °C/min. The inflection point of the ratio of 330 nm/350 nm was used to measure the denaturation temperature of a protein.

2.7.3 Fiber-optic surface plasmon resonance spectroscopy

Protein-protein interactions were qualitatively analyzed by fiber-optic surface plasmon resonance (FO-SPR) spectroscopy using a White FOx biosensor instrument (FOx Biosystems). Two different SPR sensor surfaces were used, a carboxyl- (COOH) and an NTA-surface. After immobilization of Streptavidin on the EDC/NHS-activated (1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide/N-hydroxysuccinimide) COOH-surface, biotinylated Nbs were bound to the Streptavidin. The NTA-sensor surface was first loaded with Ni^{2+}, before Nbs were captured via their C-terminal His-tag. After Nb capture, association of the target antigen was measured.

Standard protocols for SPR spectroscopy measurements are listed in Tables 2.11 and 2.12. Standard measurements with important steps are visualized in Figures 2.2 and 2.3.

Table 2.11 Standard program for an SPR measurement on COOH-sensors.

Condition	Shaking speed [min⁻¹]	Time [s]	Comment
Buffer (MES pH 6)	1000	600	Hydration of the sensors
EDC/NHS	1000	300	Activation of the COOH surface
Streptavidin capture buffer (MES pH 6)	1000	300	Buffer equilibration
Streptavidin	1000	180	Streptavidin immobilization
100 mM ethanolamine	1000	150	Quenching of the surface
Capture buffer (HBS)	1000	450	Wash step
Biotinylated Nb	1000	180	Capture of biotinylated Nb
Interaction buffer (HBS + 1 mM MgCl₂)	1000	150	Removing of un-bound proteins; buffer equilibration
Antigen	1000	180	Antigen association
Interaction buffer (HBS + 1 mM MgCl₂)	1000	180	Antigen dissociation
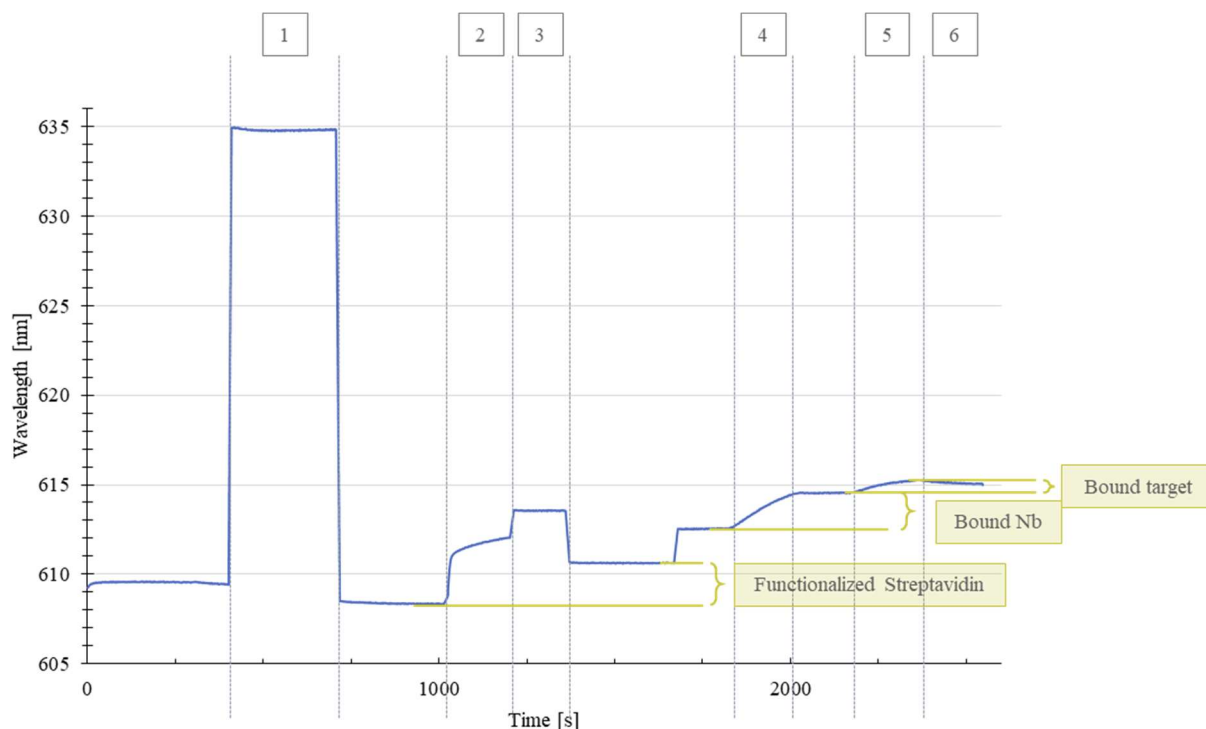
Figure 2.2 Standard SPR-measurement with Streptavidin functionalization of COOH-sensors. An SPR-measurement on a COOH-sensor surface requires following steps. **1:** Activation of the COOH-surface by EDC and NHS; **2:** Immobilization of Streptavidin; **3:** Surface quenching by ethanolamine; **4:** Capture of biotinylated Nbs; **5:** Association of the target antigen; **6:** Dissociation of the target antigen. Changes in surface layer thickness is measured by a change in refractive index. The y-axis value of a datapoint describes the maximum of surface plasmon absorption at a certain wavelength. Higher refractive indices correlate with maxima at higher wavelengths.

Table 2.12 Standard program for an SPR measurement on NTA-sensors.

Condition	Shaking speed [min^{-1}]	Time [s]	Comment
Buffer (HBS)	1000	600	Hydration of the sensors
100 mM NiSO$_4$	1000	300	Ni2+ loading
Capture buffer (HBS)	1000	300	Buffer equilibration
Nb-His	1000	900	Capture His-tagged Nb
Interaction buffer (HBS + 1 mM MgCl$_2$)	1000	300	Removing of un-bound proteins; buffer equilibration
2nd protein/interaction partner	1000	600	Protein-protein interaction; association phase
Interaction buffer (HBS + 1 mM MgCl$_2$)	1000	300	Dissociation phase
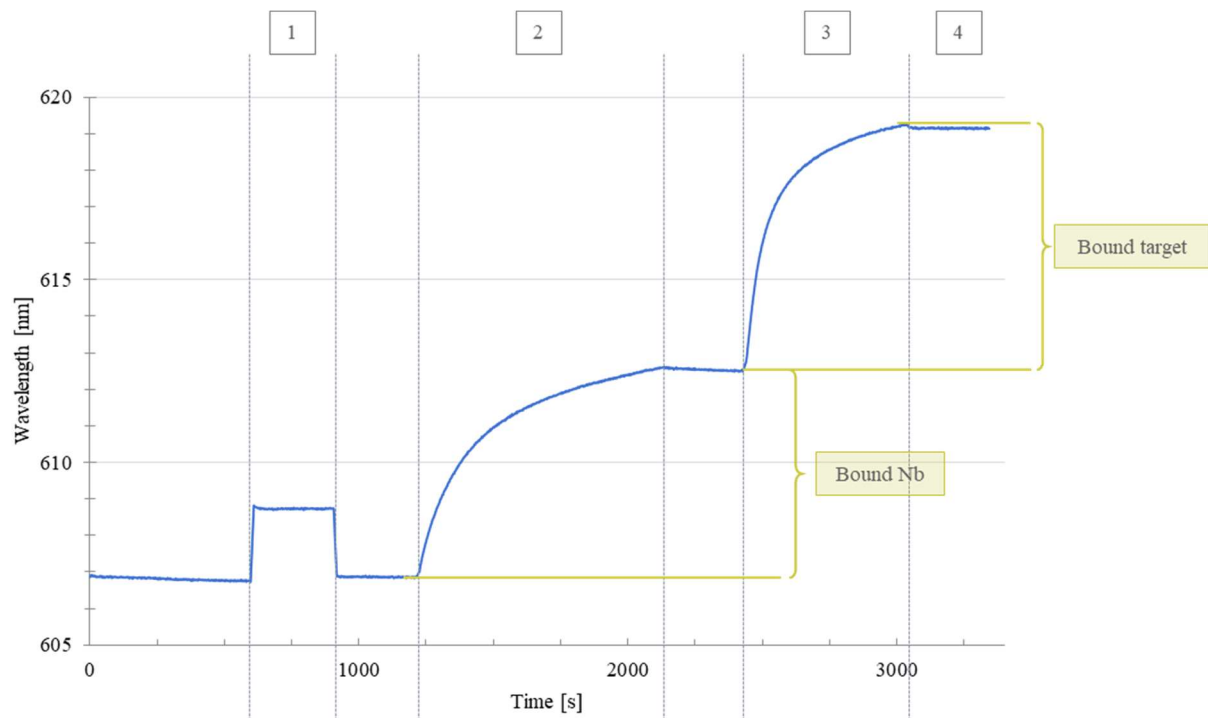
Figure 2.3 Standard SPR measurement with NTA sensors. SPR-measurements on NTA-sensor surfaces require following steps. **1:** Ni^{2+}-loading of NTA-surface; **2:** Capture of His-tagged Nbs; **3:** Association of the target antigen; **4:** Dissociation of the target antigen. Changes in surface layer thickness is measured by a change in refractive index. The y-axis value of a datapoint describes the maximum of surface plasmon absorption at a certain wavelength. Higher refractive indices correlate with maxima at higher wavelengths.

3. Results

3.1 Preparation of target antigens

3.1.1   Retinal guanylyl cyclase 1 (retGC1)

To date, no successful retGC1 purification has been published. For biochemical analysis and initial structural experiments by XL-MS, crude cell lysates were sufficient. However, for structural analysis by cryo-EM or for Nb selections by mRNA/cDNA display, a protocol for retGC1 expression and purification must established. In the following, retGC1 was transiently expressed in *sf21* insect cell culture, and a TREx293 cell line was generated, which stably integrated the GUCY2D gene encoding retGC1. After retGC1 expression in *sf21* and TREx293, a solubilization and purification protocol was established.

3.1.1.1 Expression of retGC1 in *sf21* insect cell culture

For transient retGC1 expression in *sf21*, a pFL-retGC1-Strep plasmid was prepared by restriction ligation over EcoRI and XbaI restriction sites. The C-terminal Strep-tag could later be used for detection and affinity chromatography.

The constructed pFL-retGC1-Strep plasmid was used to prepare the virus generation V_0 in *E. coli* EMbacY cells. The V_0 virus could then be used to infect *sf21* to amplify the virus, generate later virus generations and to induce retGC1 expression in *sf21*. After transformation of *E. coli* EMbacY cells, positive cloning of the Bacmid was assessed by blue-white screening. Positive (white) clones were used to inoculate 2 ml overnight cultures.

Bacmids were purified from the *E. coli* cultures by alkaline cell lysis and subsequent isopropanol precipitation. Afterwards, $1.5 \cdot 10^6$ *sf21* cells were transfected with the bacmids isolated from overnight culture. Because the generated bacmids also contain the gene encoding YFP, positively transfected/infected cells could be assessed by YFP-fluorescence. After 60 hours at 27 °C, some cells started to show YFP fluorescence indicating successful transfection (Figure 3.1).
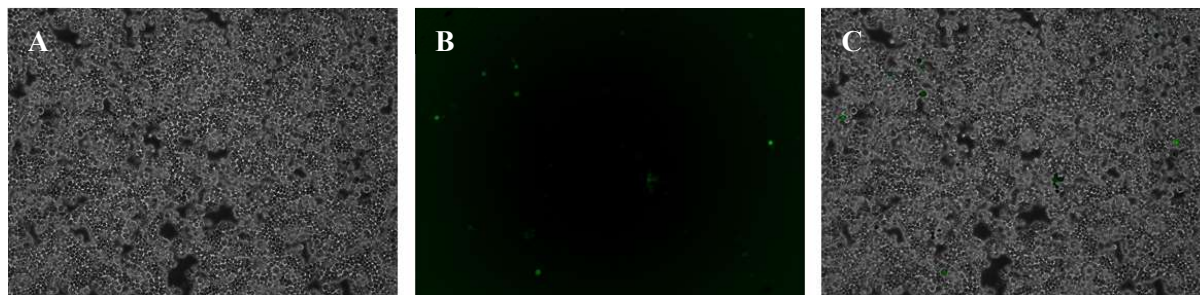
Figure 3.1 retGC1 V_0 generation. Microscopy images of *sf21* insect cells of *sf21* cells transfected with pFL-retGC1-Strep bacmid were taken 60 h after transfection. A: Light microscopic image showed an adherent *sf21* culture; B: Successfully infected cells emit YFP-fluorescence. C: Superimposition of both images.

The supernatant contains the V_1 virus generation, was harvested, and used to infect more cells to propagate the virus. The virus was amplified one more time and the resulting V_2 virus was used to infect large *sf21* cultures and induce retGC1 expression. Later virus generations showed increased potency, which is visible as an increase in infected, YFP-fluorescence emitting, cells (Figure 3.2).
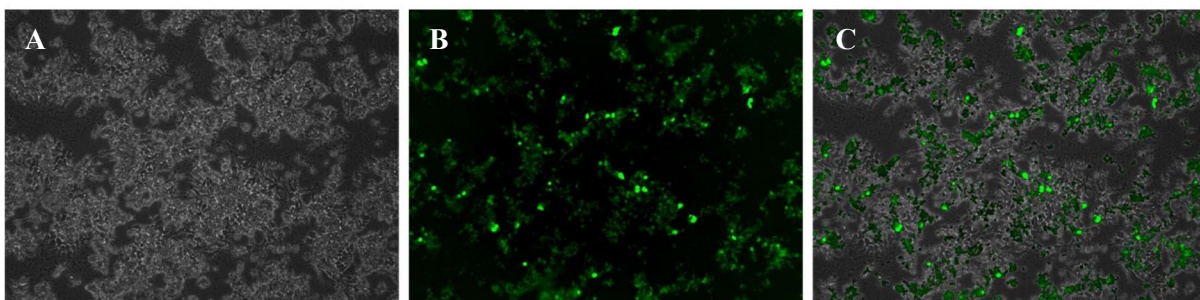
Figure 3.2 retGC1 V_2 generation. Microscopy images of *sf21* insect cells of *sf21* cells transfected with pFL-retGC1-Strep bacmid were taken 60 h after transfection. A: Light microscopic image showed an adherent *sf21* culture; B: Successfully infected cells emit YFP-fluorescence. C: Superimposition of both images.

Cells from two different expression cultures were harvested 24 h, 48 h, 72, and 96 h after proliferation arrest. Harvested cells were lysed (Total lysate) and centrifuged (Soluble fraction). Both fractions were analyzed by anti-Strep-tag Western blot (Figure 3.3). In culture 2, retGC expression appeared 24 h after proliferation arrest, in culture 1 after 48 h. This is indicated by the 100-110 kDa species in the anti-Strep-tag Western blot, which roughly correlates with the MW of retGC1 monomers of 120 kDa. The presence of two different species at 100 and 110 kDa might indicate different phosphorylation states or differently processed retGC1.
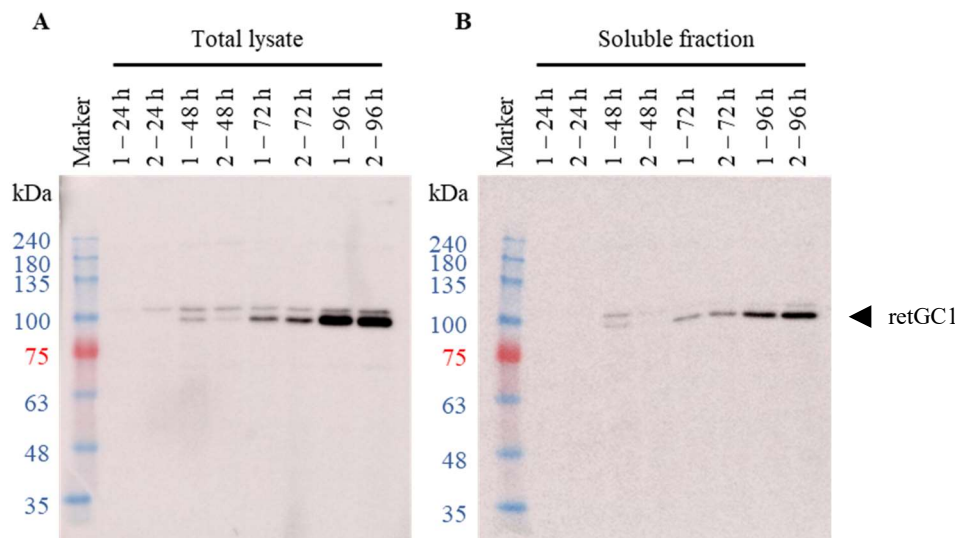
Figure 3.3 Anti-Strep-tag Western blot analysis of retGC1 expression in *sf21*. A: RetGC1 shows first expression 24 h after proliferation arrest in culture 2 and 48 h after proliferation arrest in culture 1. Expression levels increased until 96 h. RetGC1 was present in two different species at 100 and 110 kDa. B: A smaller but detectable fraction of retGC1 was present in the soluble protein fractions.

Expression levels of retGC1 increased over time in both cultures. A certain fraction of retGC1 was also present as soluble proteins, as indicated by the 100 kDa species in the soluble fraction (Figure 3.3 B).

After *sf21* cells successfully expressed soluble retGC1 proteins, the next step was to find a suitable detergent, SMA or DIBMA to solubilize retGC1 out of the *sf21* cell membranes. After detergents like DDM (n-Dodecyl-β-D-maltoside) were not able to solubilize retGC1 in previous experiments, the denaturing detergent lauroylsarcosine (Sarcosyl), as well as the non-denaturing detergents LDAO (N,N-Dimethyl-n-dodecylamine N-oxide), CYMAL-5 (5-Cyclohexylpentyl-β-D-maltosid), and Triton-X100 were used to solubilize retGC1. Additionally, Sokalan CP9 (DIBMA) and XIRAN (SMA) were used in different concentrations.

The denaturing detergent Sarcosyl was the only detergent to be able to solubilize at least a detectable amount of retGC1 (Figure 3.4 A). All other detergent lacked detectable signal in the solubilized fraction. Sokalan CP9 solubilization solubilized retGC1 dimers, which are shown at higher MWs (~300 kDa). Higher concentrations of Sokalan CP9 showed more successful solubilization, with 5 % Sokalan CP9 yielding the strongest signal (Figure 3.4 B). The high-MW signal might correlate with retGC1 dimers in Sokalan CP9 discs. XIRAN solubilization did not yield any detectable signal. (Figure 3.4 B).
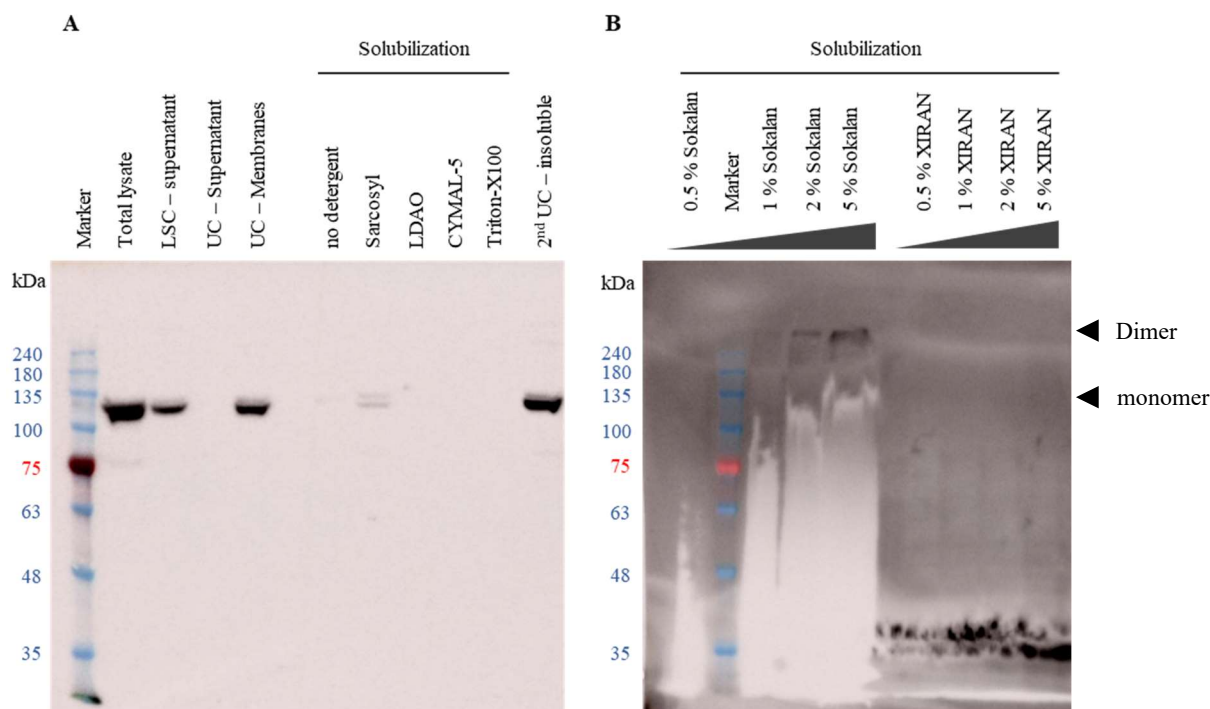
Figure 3.4 Solubilization of retGC1 from *sf21* membranes. A: *Sf21* membranes were isolated after retGC1 expression (UC – membranes), and retGC1 was solubilized with the detergents Sarcosyl, LDAO, CYMAL-5, and Triton-X100. B: RetGC1 was solubilized in different concentrations of the Sokalan CP9 and XIRAN. All samples were analyzed by anti-Strep-tag Western blot.

After Solubilization in 5 % Sokalan, retGC1 was purified by Strep-tactin affinity chromatography and analytic SEC on a Superose 6 PC 3.2/30 column. RetGC1 eluted in a single peak at 1.2 ml (Figure 3.5), even earlier than a 669 kDa protein in the supplier's column calibration. This means that retGC1 probabaly formed larger aggregates, which eluted very early. RetGC1 dimers have a molecular mass of roughly 240 kDa. Even when accounting for the added mass of Sokalan and the close lipid environment of retGC1, the expected molecular mass of the solubilized DIBMALPs (DIBMA lipid particle) was not expected to exceed 300 – 350 kDa. Because solubilized retGC1 dimers were expected to elute much later, fraction collection was started after 1.5 ml and the peak fractions were lost.
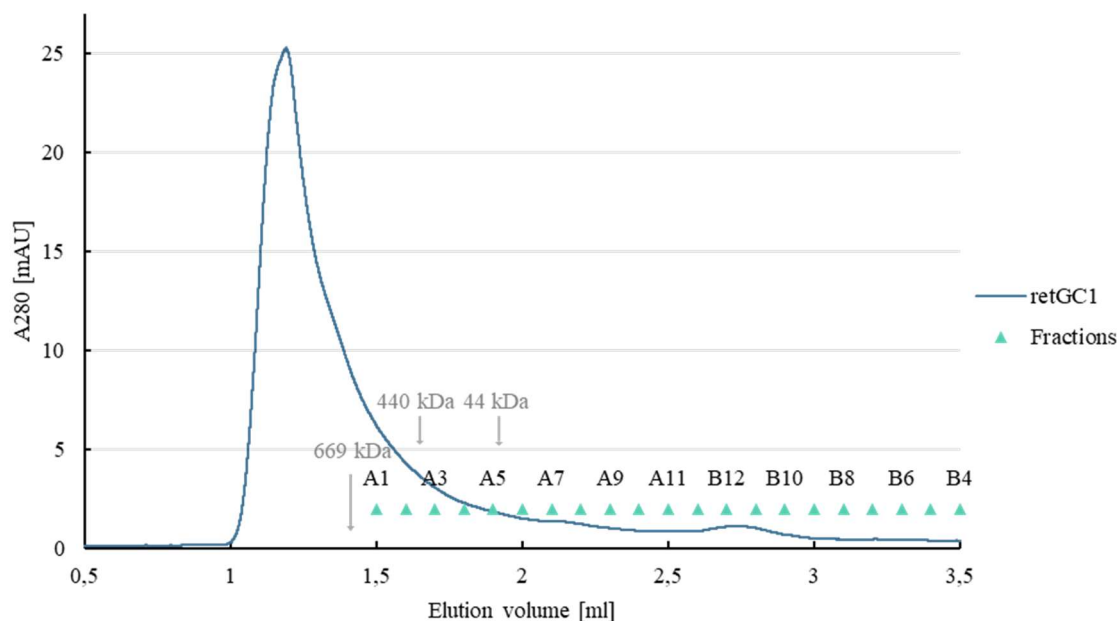
Figure 3.5 SEC-purification of retGC1. retGC1 was transiently expressed in *sf21* insect cells, solubilized with 5 % Sokalan CP9, purified by Strep-tactin affinity chromatography and SEC on a Superose 6 PC 3.2/30 column. RetGC1 eluted in an early peak at 1.2 ml, which is earlier than the calibration with a 669 kDa protein, and probably represents retGC1 aggregates. 100 µl fractions were collected starting from 1.5 ml (green triangles). MWs from the supplier's column calibration are shown as grey arrows.

3.1.1.2 Expression of retGC1 in a TREx293 stable cell line

After retGC1 could not be purified from *sf21* membranes, a TREx293 cell line, which stably integrated the *GUCY2D* gene encoding for retGC1, was established by Christine Gotthold and Magdalena Schacherl. This cell line allows induction or retGC1 expression by adding 100 ng/ml tet to the growth medium for the adherent TREx293 cell culture. RetGC1 was expressed with an N-terminal Strep-tag and a C-terminal eGFP. So, retGC1 expression could be detected either by an anti-Strep-tag Western blot or by SDS-PAGE and in-gel eGFP fluorescence.

To analyze time-dependent expression levels, cells were harvested 48 h and 72 h after induction of retGC1 expression. RetGC1 was expressed to saturation after 48 h, as indicated by the 140 kDa species in the anti-Strep-tag Western blot that did not increase noticeably after 72 h (Figure 3.6 A). The dominant 140 kDa species correlates nicely with the theoretical MW of the Strep-retGC1-eGFP construct of roughly 150 kDa. Only a small fraction of retGC1 seemed to be present as soluble proteins, as the weaker signal in the soluble fraction shows (Figure 3.6 A).
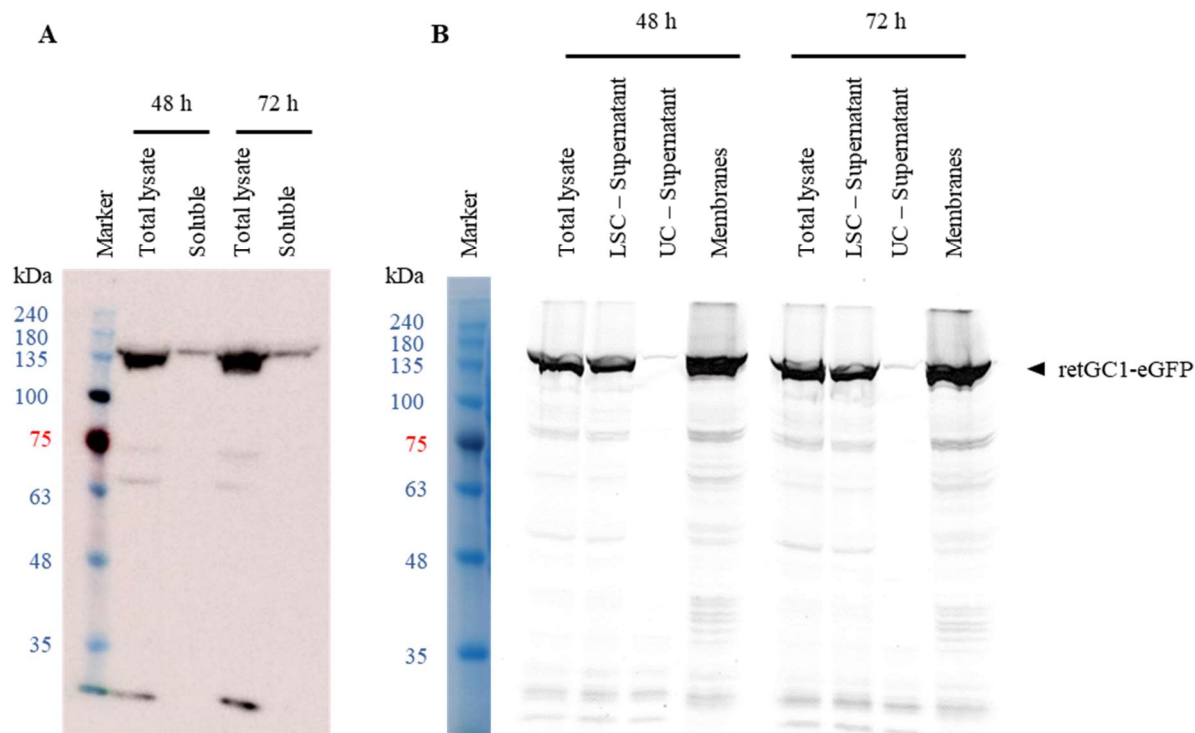
Figure 3.6 Overexpression of retGC1 in TREx293 cells with subsequent membrane isolation. A: Anti-Strep-tag Western blot analysis of Strep-retGC1-eGFP expressions in TREx293 48 h and 72 h after induction of protein expression. Total lysate and soluble protein fractions were analyzed. B: In-gel eGFP fluorescence of Strep-retGC1-eGFP expression and membrane isolation and comparison to a pre-stained protein marker. Strep-retGC1-eGFP was expressed for 48 h and 72 h.

After the expression test, TREx293 membranes were isolated and analyzed for the presence of retGC1 by in-gel eGFP fluorescence (Figure 3.6 B). RetGC1 again showed similar expression levels after 48 h and 72 h. Interestingly, most of retGC1-eGFP remained soluble after low-speed centrifugation (Figure 3.6 B; LSC-supernatant). The N-terminal localization of the Strep-tag and the C-terminal localization of the eGFP suggest that the excess of insoluble retGC1 in the anti-Strep-tag Western blot contains a truncated or denatured C-terminus.

After all misfolded proteins were removed by low-speed centrifugation (LSC), the membranes were pelleted by ultracentrifugation (UC). After UC, the majority of retGC1 is present in the membrane fraction and only little retGC1 remains in the supernatant (Figure 3.6 B).

Because the mammalian TREx293 cells have different cell membrane compositions than the *sf21* insect cells, Sokalan-mediated retGC1-solubilization had to be confirmed with TREx293 membranes (Figure 3.7). RetGC1-containing TREx293 membranes were solubilized in Sokalan CP9 concentrations of 1 %, 2 %, and 5 %. Similar to the previous solubilization from *sf21* membranes, larger amounts of retGC1 were solubilized from TREx293 membranes at higher Sokalan CP9 concentrations. Signal intensity, as well as shape of the 140 kDa species were enhanced (Figure 3.7). In addition to the 140 kDa species, a secondary, ~30 kDa species was detected, which correlates well with the MW of eGFP (27 kDa).

Interestingly, Sokalan CP9-mediated solubilization from TREx293 membranes yielded a 140 kDa in SDS-PAGE, compared to the higher MW species in the solubilization from *sf21* membranes. A reason might be the different membrane compositions. Other possibilities could be that *sf21*-retGC1 formed higher-MW aggregates, or that the C-terminal eGFP impedes the formation of retGC1 homodimers.
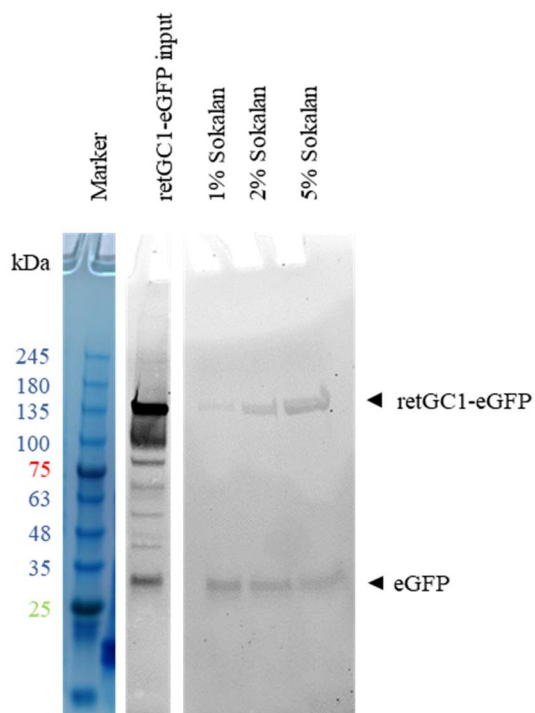
Figure 3.7 retGC1 solubilization in Sokalan CP 9. Strep-retGC1-eGFP was solubilized from isolated TREx293 membranes (retGC1-eGFP input) using 1 %, 2 %, and 5 % of Sokalan CP9. Solubilization levels were analyzed by eGFP fluorescence in SDS-PAGE. A pre-stained protein marker was used as a protein size standard.

Solubilized retGC1 was purified by affinity chromatography over the N-terminal Strep-tag and subsequently by SEC using a Superose 6 10/300 column (Cytiva). The size-exclusion chromatogram (Figure 3.8 A) showed an increase in A_{280} after roughly 14 ml, which plateaued after 15.9 ml with small spikes at 17.2 ml (Figure 3.8 A). Pure retGC1 was present in the early SEC fraction after the increase (fractions 16 to 19), as indicated by the pure 140 kDa species in the SDS-PAGE (Figure 3.8 B).

Free eGFP was collected in fractions 21 to 25, indicated by the ~30 kDa species in the SDS-PAGE (Figure 3.8 B). The peaks from retGC1 and GFP were probably not separated enough, which explains the plateau.
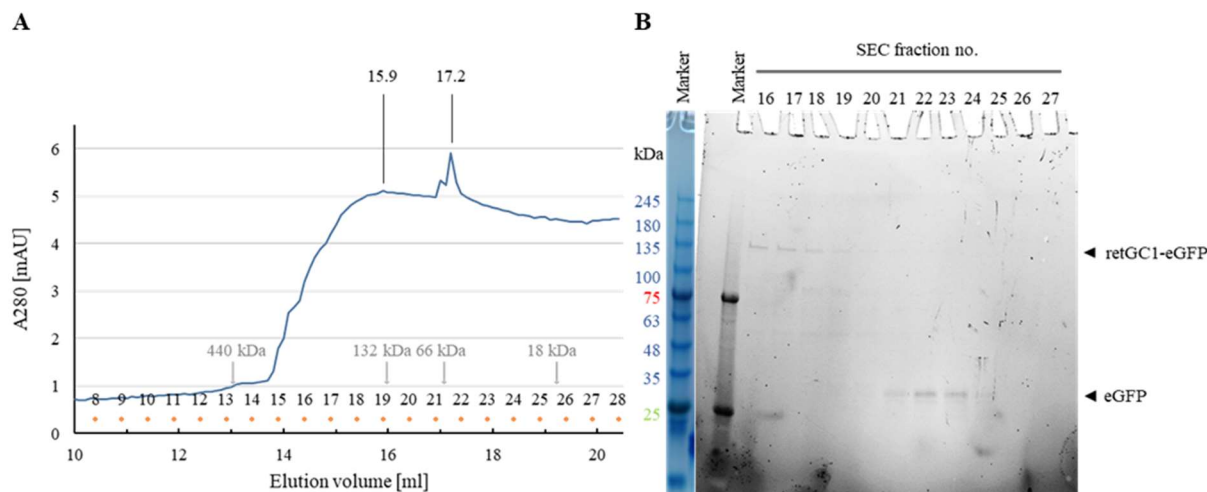
Figure 3.8 Size-exclusion chromatography of retGC1. A: Size-exclusion chromatogram of a Strep-retGC1-eGFP SEC-purification over a Superose 6 10/300 column (Cytiva). Proteins eluted after 14 ml, reached a plateau after 15.9 ml and showed two sharp peaks at 17.2 ml B: SDS-PAGE analysis of the collected SEC fractions. Proteins were visualized by eGFP-fluorescence and compared to a pre-stained protein marker. Molecular weights from column calibration are indicated by grey arrows.

The purified amounts of retGC1 were insufficient for either structural analysis or Nanobody selection by mRNA/cDNA display, and culture volumes must be scaled up largely to generate sufficient amounts of retGC1.

Nevertheless, retGC1 has been shown to be solubilized from *sf21* and TREx293 membranes in a DIBMALP, with subsequent purification by affinity and size-exclusion chromatography for the first time.

3.1.2 Expression and purification of *in vivo* myristoylated guanylyl cyclase activating protein 1 (GCAP1)

For Nb selections by mRNA/cDNA display, as well as for protein crystallography, human myrGCAP1 was overexpressed in *E. coli* BL21(DE3) Codon Plus RIL (C+ RIL) and purified by affinity and size-exclusion chromatography.

A C-terminally Strep-tagged GCAP1-construct was prepared for purification Strep-Tactin-affinity chromatography and for panning during Nb selection on Strep-Tactin magnetic beads. GCAP1 was cloned into a pStrep1 vector by restriction ligation over NdeI and XhoI restriction sites.

Because N-terminal myristoylation is essential for properly folded and active GCAP1, the posttranslational modification was introduced on the GCAP1 D6S variant by a yeast N-myristoyl transferase (NMT), which was co-expressed. NMT is expressed under T7 promoter control, which can be induced by the addition of IPTG to the growth medium. Because the pNMT vector contains the gene for kan resistance, the GCAP1 expression vector must contain a different antibiotic resistance to maintain selective pressure for both plasmids. Unfortunately, pStrep1-GCAP1 also contains the kan

resistance gene. For GCAP1 expression and *in vivo* myristoylation, the resistance gene in pSrep1-GCAP1 was replaced by the gene for ampicillin resistance. The resistance gene exchange was performed by restriction ligation over two DraIII restriction sites in the pStrep1-GCAP1 vector and in a PET22b-GCAP1 vector. One of the ligation products was a pStrep1(amp)-GCAP1 vector containing the ampicillin resistance gene.

For overexpression of myrGCAP1-Strep, BL21(DE3) C+ RIL were co-transformed with pStrep(amp)-GCAP1 D6S and pNMT, and protein expression was induced by IPTG addition. MyrGCAP1 was produced in large quantities, with the majority of myrGCAP1 being soluble, as indicated by anti-Strep-tag Western blot (Figure 3.9). The different species in the harvest fractions represent myrGCAP1 dimers, which can be stable in SDS-containing buffers (48 kDa), myrGCAP1 (25 kDa), and myrGCAP1-Ca (19 kDa), which retains some secondary structure elements in SDS-containing buffers and shows altered electrophoretic behavior (Viviano et al., 2016).
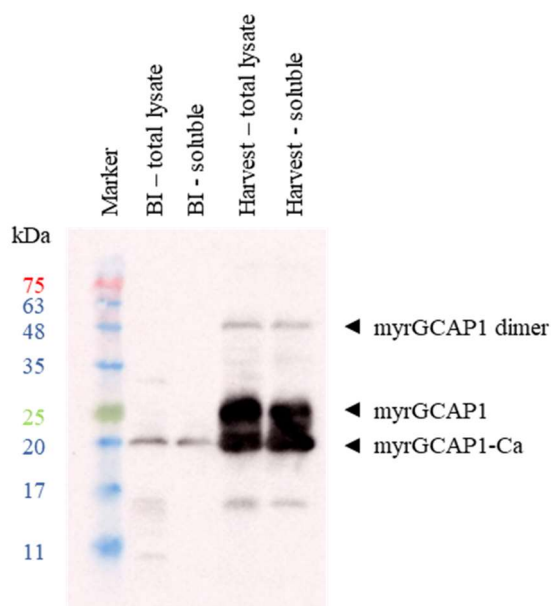
Figure 3.9 Overexpression of myrGCAP1. MyrGCAP1 wild type was overexpressed in *E. coli* BL21(DE3) C+ RIL. Expression levels and solubility were analyzed by anti-Strep-tag Western blot.

After overexpression, large quantities of myrGCAP1 were purified by Strep-tactin-affinity chromatography and subsequent SEC using an ÄKTA FPLC and a Superdex 75 pg 10/300 column (Cytiva). During affinity chromatography, Ca^{2+}- and Mg^{2+}-ions were removed by using buffers supplemented with 1 mM EDTA and EGTA.

To generate myrGCAP1-Ca, a buffer supplemented with 1 mM $CaCl_2$ was used for SEC. For myrGCAP1-Mg, the SEC buffer was supplemented with 1 mM $MgCl_2$. In the myrGCAP1-apo purification EDTA and EGTA concentrations were lowered to 50 µM.

All three variants showed three distinct peaks in the SEC. The two earlier peaks coincide for all three variants at 8.2 and 9.8, respectively. MyrGCAP1-Ca showed a third peak at 10.5 ml, -apo and -Mg at 11.2 ml (Figure 3.10). The altered chromatographic behavior is probably caused by a Ca-dependent conformational change. The latest peak contains myrGCAP1 monomers, which confirmed by SDS-

PAGE, as the fractions of the latest peak are completely devoid of the SDS-stable dimer (Figure 3.11). Because GCAP1 has a very high affinity for Ca^{2+}, small amounts of Ca^{2+} in SDS-PAGE sample or running buffer suffice to generate myrGCAP1-Ca, which migrates at 19 kDa. Only samples containing high concentrations of EDTA and EGTA (Figure 3.11 A; Strep-Wash, Strep-Elution and Concentrated Elution) are devoid of the 19 kDa, myrGCAP1-Ca, species.
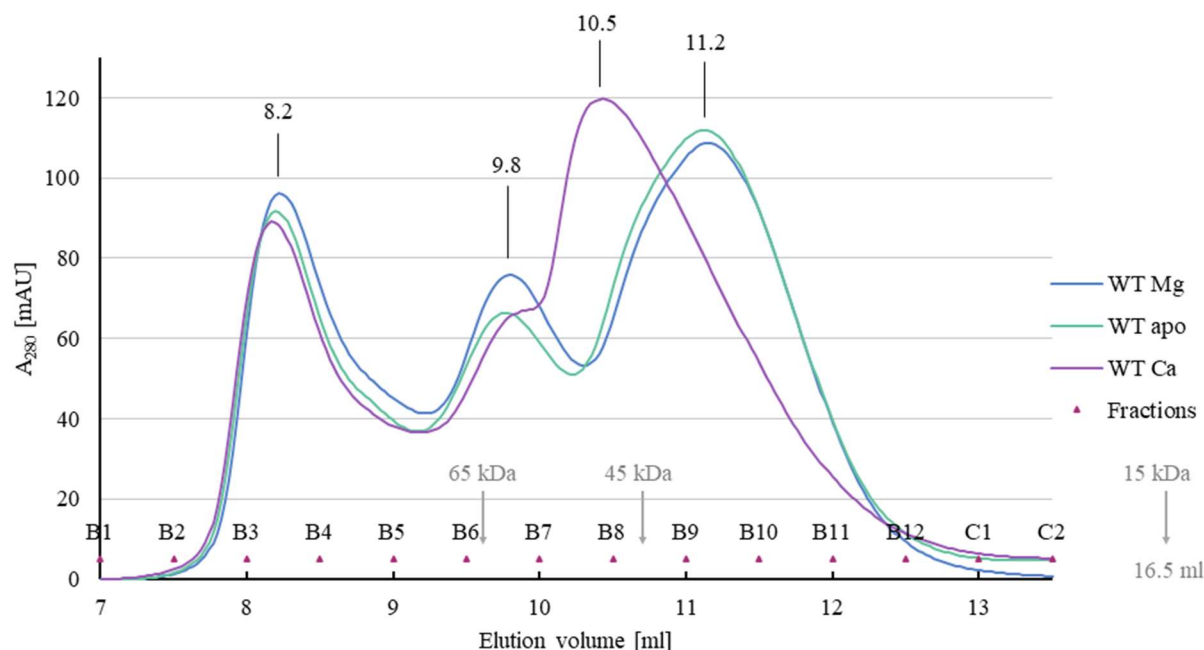
Figure 3.10 SEC-purifications of myrGCAP1-apo, -Ca, and -Mg. SEC-purification of myrGCAP1-variants over a Superdex 75 pg 10/300 column. MyrGCAP1-apo (green), myrGCAP1-Mg (blue), and GCAP1-Ca (purple) eluted in three peaks. The two earlier peaks coincided at 8.2 and 9.8 ml, respectively. MyrGCAP1-Ca showed a third peak after 10.5 ml, -apo and -Mg after 11.2 ml. Collected fractions are depicted above the x-axis (purple trianlges). Molecular weights from column calibration are indicated by grey arrows.
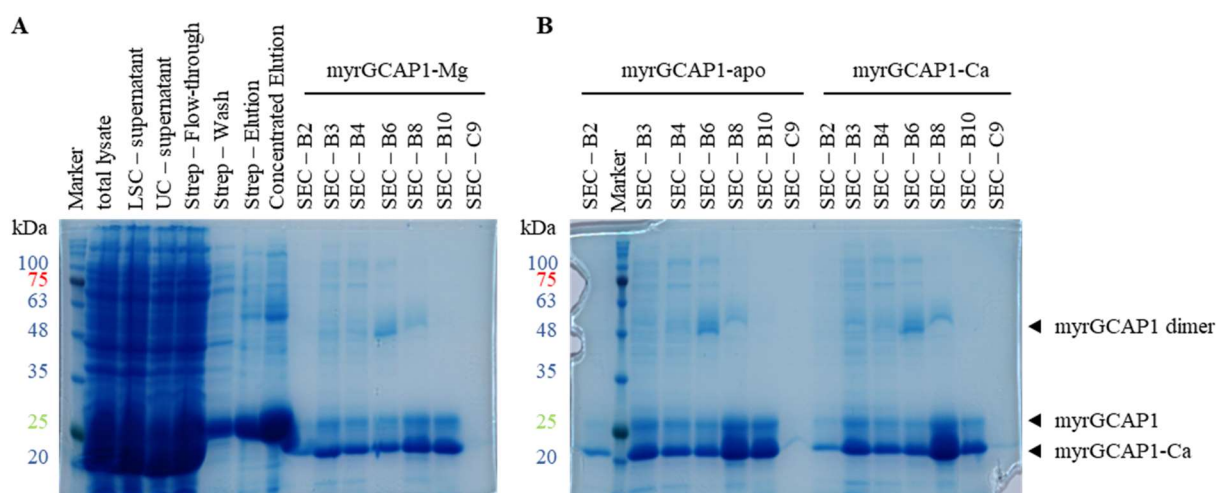
Figure 3.11 myrGCAP1-Mg, -apo, and -Ca purification. A: Fractions from myrGCAP1 affinity chromatography and SEC-purification of myrGCAP1-Mg and B: SEC-fractions from myrGCAP1-apo and -Ca were analyzed by Coomassie-stained SDS-PAGE.

Besides an altered migration behavior in SEC and SDS-PAGE, myrGCAP1-Ca also showed increased thermostability, compared to myrGCAP1-apo and -Mg. While myrGCAP1-apo and -Mg were denatured at temperatures of 53.3 °C or 55.5 °C, respectively, myrGCAP1-Ca could not be denatured in nanoDSF measurements with maximum temperatures of 95 °C (Figure 3.12).
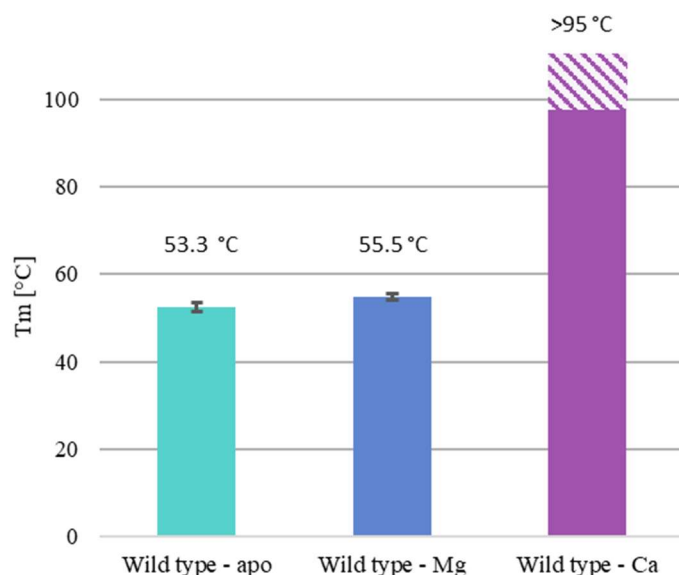
Figure 3.12 Thermal stability of wild type myrGCAP1-apo, -Mg, and -Ca. Melting temperatures (T_m) of myrGCAP1-apo (green), Mg-bound (blue), and Ca-bound (purple) were determined from four replicate nanoDSF measurements with the standard deviation as error values. MyrGCAP1-apo denatured at 53.3 °C, myrGCAP1-Mg at 55.5 °C. MyrGCAP1-Ca was not denatured until the limit of 95 °C (dark purple) and therefore has an unknown T_m >95 °C (striped).

3.1.3 Heterologous expression and purification of SARS-CoV2-main protease (M^pro)

During the COVID-19 pandemic, research on coronaviruses became more relevant and we decided to include SARS-CoV2-main protease (M^pro) as one of our target antigens for Nb selection. M^pro is essential for the synthesis of new virus particles by cleaving the coronavirus polyproteins into the single non-structural proteins. Before being able to select Nbs against SARS-CoV2-M^pro, we were required to produce large quantities of M^pro with high purity.

An M^pro gene, which was codon-optimized for *E. coli* was synthesized (GenScript) and cloned into a pGEX4T1 expression vector by restriction ligation over BamHI and XhoI restriction and T4 DNA ligation. The M^pro gene was synthesized with flanking BamHI and XhoI restriction sites. Therefore, no PCR amplification was required. The generated construct contained a GST-M^pro-Strep tag fusion protein with an M^pro-specific cleavage site between the GST-tag and M^pro. This enabled M^pro-cleavage from the fusion protein, similar to the cleavage of M^pro from the SARS-CoV2 polyprotein.
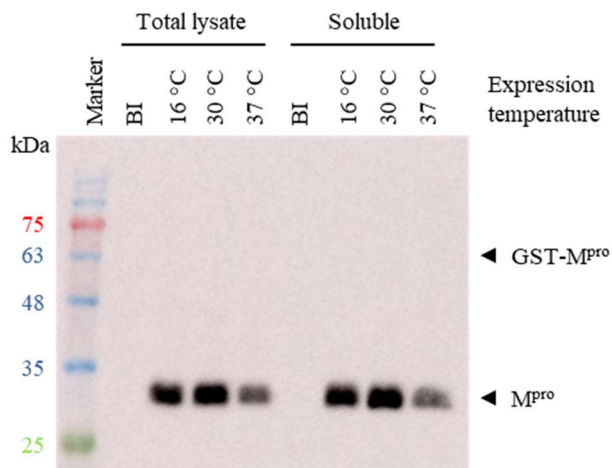
Figure 3.13 Test expression of SARS-CoV2-M^pro. GST-M^pro-Strep was overexpressed in *E. coli* BL21(DE3) at 16 °C for 16 h, 30 °C for 5 h, and 37 °C for 3 h. After cell harvest, total lysates and soluble protein fractions were analyzed by anti-Strep-tag Western blot. Due to its autocatalytic activity, M^pro cleaved itself off GST.

At first, the right conditions for M^pro overexpression were determined. Protein expression at 16 °C for 16 h and at 30 °C for 5 h yielded comparable amounts of soluble M^pro, indicated by ~33 kDa species in anti-Strep-tag Western blot. Only expression at 37 °C for 3 h had a worse yield (Figure 3.13). Catalytic activity of M^pro was confirmed by cleavage of M^pro from the GST-tag, which is shown by the ~33 kDa species in the Western blot, which correlate well with the theoretical MW of M^pro of 34 kDa. GST-M^pro fusion proteins have a theoretical MW of 62 kDa.
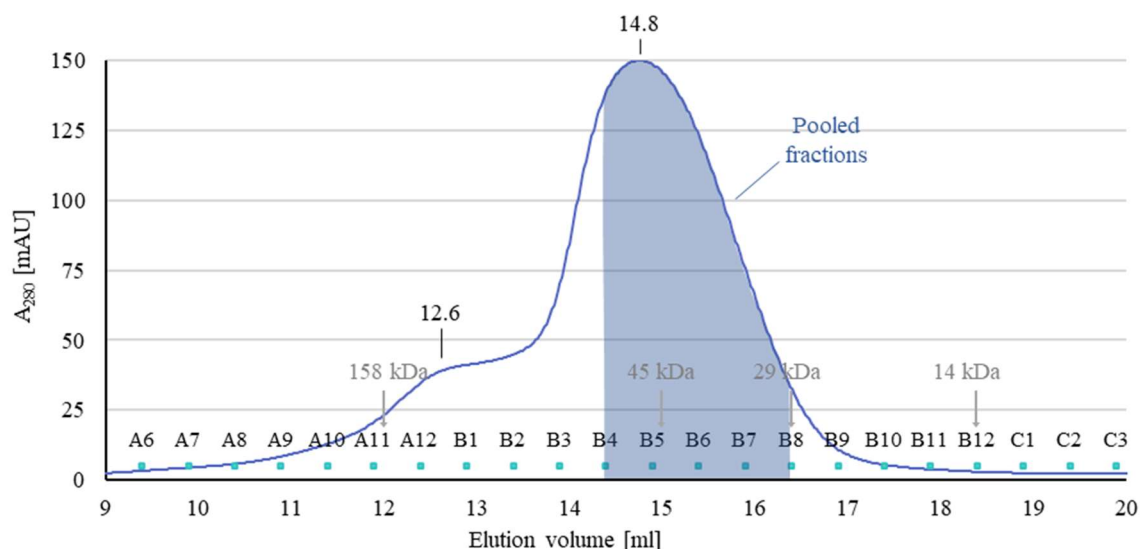
Figure 3.14 Purification of SARS-CoV2 M^pro via SEC. M^pro-Strep was SEC-purified on a Superdex 200 10/300 column. M^pro-Strep eluted with a peak at 14.8 ml and a smaller shoulder at 12.6 ml. Fractions B4 to B7 were collected and pooled. Molecular weights from the manufacturer's column calibration are indicated by grey arrows.

Overexpressed M^pro-Strep was purified by Strep-tactin-affinity chromatography und subsequent SEC using an ÄKTA FPLC and a Superdex 200 10/300 column (Cytiva). The SEC resulted in a main peak at 14.8 ml and a small shoulder at 12.6 ml (Figure 3.14). M^pro can form dimers with low affinity (Johansen-Leete et al., 2022). So, the 12.6 ml shoulder might contain M^pro-homodimers, which are not stable in SDS-PAGE, and migrate as 34 kDa species in fraction A12 (Figure 3.15). The main peak

fractions B5 and B7 contain large amounts of M^{pro}, which was confirmed by SDS-PAGE (Figure 3.15) The SEC fractions B4 to B7 were pooled, which yielded 7.6 mg of M^{pro} with high purity.
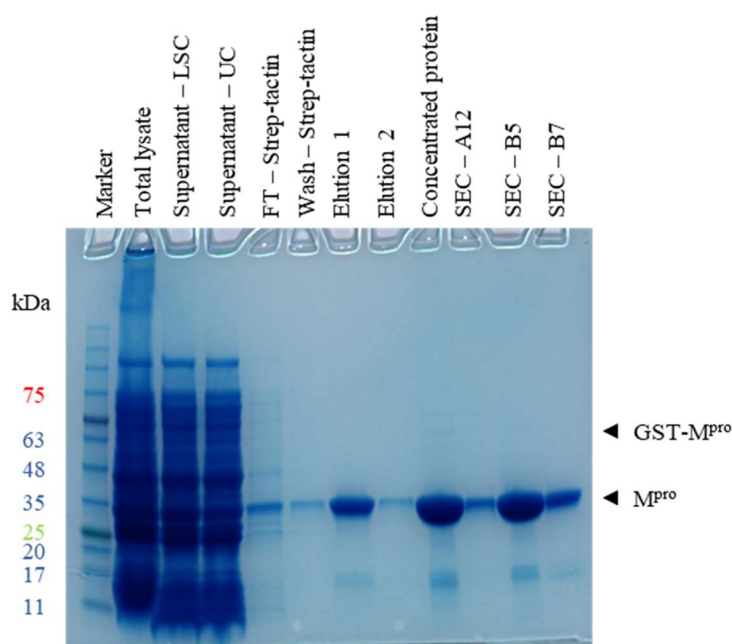
Figure 3.15 Purification of SARS-CoV2 M^{pro}. GST-M^{pro}-Strep was expressed in *E. coli* BL21(DE3). M^{pro} autocatalytically cleaved itself off GST (35 kDa). M^{pro}-Strep was successfully purified by Strep-tactin affinity chromatography and subsequent SEC. The collected fractions B5 and B7 contain large quantities of M^{pro}.

3.2 Preparation of the yeast display Nanobody library from *Saccharomyces cerevisiae*

The yeast display Nanobody library (y-lib) (McMahon et al., 2018) was, as the name suggests, optimized for the use in yeast display. The commercially available y-lib was delivered in the form of yeast cells transformed with a pYDS plasmid containing the Nb genes. Before we could use the y-lib for mRNA/cDNA display, the pYDS-Nb plasmids had to be isolated from the transformed yeast cells. Existing protocols for plasmid usually have a low yield and poor quality. Therefore, a protocol needed to be established, which allows plasmid isolation from yeast with high yield and purity.

At first, the yeast culture was extended and during cell harvest, aliquots with cell numbers of $5 \cdot 10^9$ to $5 \cdot 10^{10}$ were prepared. Then, plasmids were purified using different variations of the mentioned protocol (chapter 2.4.1.2).

In the first plasmid preparation (Figure 3.16 A-1), a high cell number was used. $5 \cdot 10^{10}$ cells were re-suspended in 30 ml of buffer P1 (Qiagen Plasmid Maxi Kit). Zymolyase (Zymo Research) was then added to digest the yeast cell wall. In this protocol variant, precipitate and supernatant were processed together. This had the effect that mostly genomic was purified, which is indicated by high MW species in agarose gel electrophoresis in the undigested sample and a characteristic smear after restriction digest (Figure 3.16 A-1). The yeast 2µ plasmid, which is present in the majority of yeast strains, was co-purified and migrated at roughly 5 000 bp in the agarose gel electrophoresis (Figure 3.16 A-1).

In the second plasmid isolation attempt (Figure 3.16 A-2), the cell number was reduced to $5 \cdot 10^9$ cells, while maintaining the volume of 30 ml re-suspension buffer. After zymolyase digest, precipitate and supernatant were still processed together. Reducing the cell number resulted in an overall reduced DNA yield. Genomic DNA was almost removed entirely and traces of the pYDS-Nb plasmid were purified, as indicated by the 8 kb species in agarose gel electrophoresis (Figure 3.16 A-2). With this protocol, the 2µ plasmid was purified with larger quantities than the desired pYDS-Nb plasmid (Figure 3.16 A-2).

In the third plasmid preparation (Figure 3.16 A-3), the lower cell number was maintained. After zymolyase digest, precipitate and supernatant were separated by centrifugation. Then, the precipitate was re-suspended in 30 ml buffer P1 and the cell lysis and spin-column purification were performed as mentioned. Here, detectable amounts of pYDS-Nb were purified, but genomic DNA was the dominant species (Figure 3.16 A-3).

In the fourth and final plasmid preparation (Figure 3.16 A-4), again the lower cell number was used and after zymolyase digest, the precipitation was removed by centrifugation. Then, the 30 ml of supernatant were used for alkaline cell lysis and subsequent spin-column purification of plasmid DNA. This protocol yielded ~15 µg of desired pYDS-Nb plasmid with a smaller amount of genomic DNA contamination and barely detectable amounts of 2µ plasmid (Figure 3.16 A-4).

By applying the fourth protocol, 15 µg of plasmid DNA were isolated from $5 \cdot 10^9$ cells, which equaled a culture volume of 330 ml. So, the plasmid yield of this isolation procedure was roughly 45 µg per liter of yeast culture.
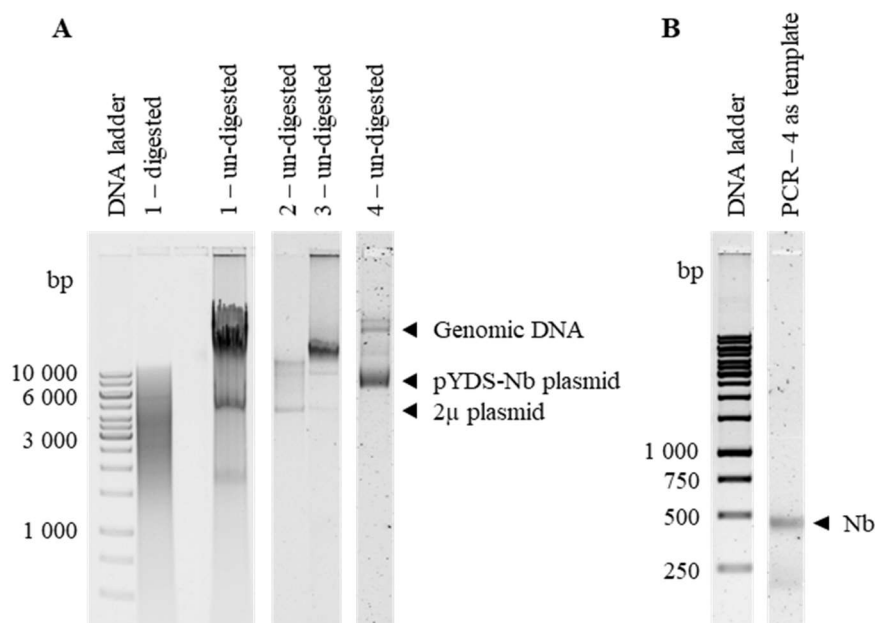
Figure 3.16 Isolation of pYDS-Nb plasmids from yeast. A: Agarose gel electrophoresis of isolated DNA from four different procedures. Plasmid isolation protocol 1 produced mostly genomic DNA, which generates a smeared signal after restriction digest. In protocol 2, genomic DNA contamination was reduced, and the yeast 2μ plasmid was isolated. Protocol 3 purified mostly genomic DNA. With protocol 4, large quantities of pYDS-Nb were isolated and little genomic DNA. B: Nb genes were PCR amplified from pYDS-Nb plasmid isolated by protocol 4.

To assess if the quality of the purified plasmid DNA was sufficient for the use in mRNA/cDNA display, Nb genes were amplified by PCR (Figure 3.16 B). Nb genes were amplified specifically, as indicated by the 400 – 450 bp fragment, which correlates nicely with the size of a Nb gene (Figure 3.16 B). The amplified Nb genes are of high purity, as no secondary products were visible (Figure 3.16 B). To avoid overamplification of the DNA library, Nb genes were amplified with a reduced number of PCR cycles, which explains the relatively weak signal.

The mentioned protocol of plasmid DNA purification from yeast combines a high yield (45 μg per liter of culture) with a purity and quality sufficient for downstream PCRs, which is unprecedented.

3.3 Analysis of the Nanobody DNA libraries

Because the two Nanobody gene libraries, c-lib and y-lib, contain randomized Nb genes, both libraries were screened for well-behaved Nbs that overexpressed in *E. coli* and were translocated to the periplasm. Additionally, Nb genes were Sanger sequenced to gain first insights into library composition and differences in variability between the two libraries.

3.3.1 Analysis of our highly variably Nanobody gene library

Our own combinatorial Nanobody gene library, c-lib, is composed of synthetic Nb genes, and the fraction of well-behaved Nbs is not known. Therefore, several Nb genes were subcloned into a pHen6-

Nb-His vector and used to induce Nb expression in *E. coli* wk6. The expression cultures were harvested, and the *E. coli* periplasm was isolated. Nbs in the periplasmic extracts were analyzed by anti-His-tag dot/Western blot (Figure 3.17)
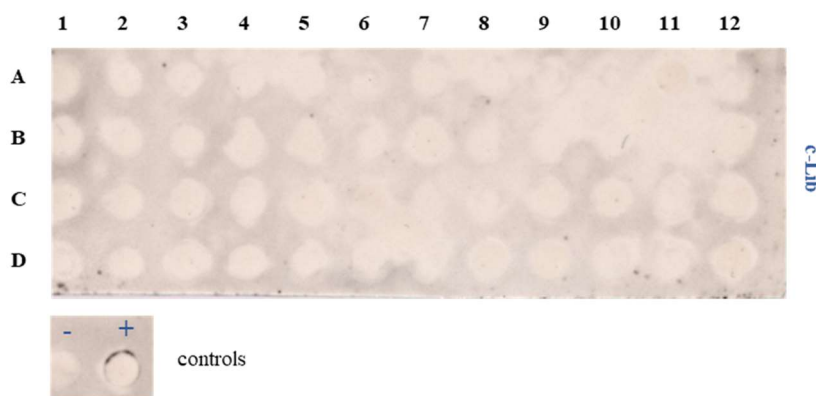
Figure 3.17 Nb expression tests of Nbs from the y-lib. 48 Nbs of the c-lib (bottom) were expressed in *E. coli* wk6 and the periplasmic fractions were analyzed by Anti-His-tag dot/Western blot. An uninduced culture served as negative control and an Nb that is known to be well-behaved as positive control, which generated a small signal around the edge.

Unfortunately, no Nbs could be detected. However, the faint signal of the positive control suggests that some Nbs might simply remain undetected.

Additionally, Nb sequences were analyzed by Sanger sequencing (LGC genomics). In total, 58 Nb genes were analyzed. Of those 58 sequences only 17 showed full-length Nb sequences. Twelve sequences contained truncated Nb sequences (deletion within the Nb gene). The majority of those, nine sequences, showed a deletion of the entire CDR3 with a characteristic CEER amino acid sequence at the end of FR3, which we encountered often in Nb sequencings. 27 sequences contained premature Stop-Codons, and two sequences contained a frameshift, which led to an out-of-frame C-terminus (Table 3.1).

Table 3.1 Distribution of different sequencing results of the c-lib.

Sequencing result	Amount	Percentage
Full-length Nb	17	29 %
Truncated Nb	12	21 %
Premature Stop-codon	27	47 %
Frameshift	2	3 %
Σ	58	100 %

The c-lib was designed with a high level of variability. All CDRs are entirely-sequence randomized and also show length polymorphisms. In this library, the variability is, however, not restricted to the CDRs. All FRs contain several variable positions. The most conserved regions are the termini and a KGRFT-motif in FR2 (Figure 3.18), which was essential for successful library construction. The residues C23 and C101 can form a stabilizing disulfide bond and are designed to be conserved, as well.

However, Nb cLib-B07 has a mutation at this position, which might lead to a misfolded Nb protein. Throughout the analyzed sequences there are several deviations from the library design and not all conserved residues appear to be fixed in the sequenced Nbs (Figure 3.18).
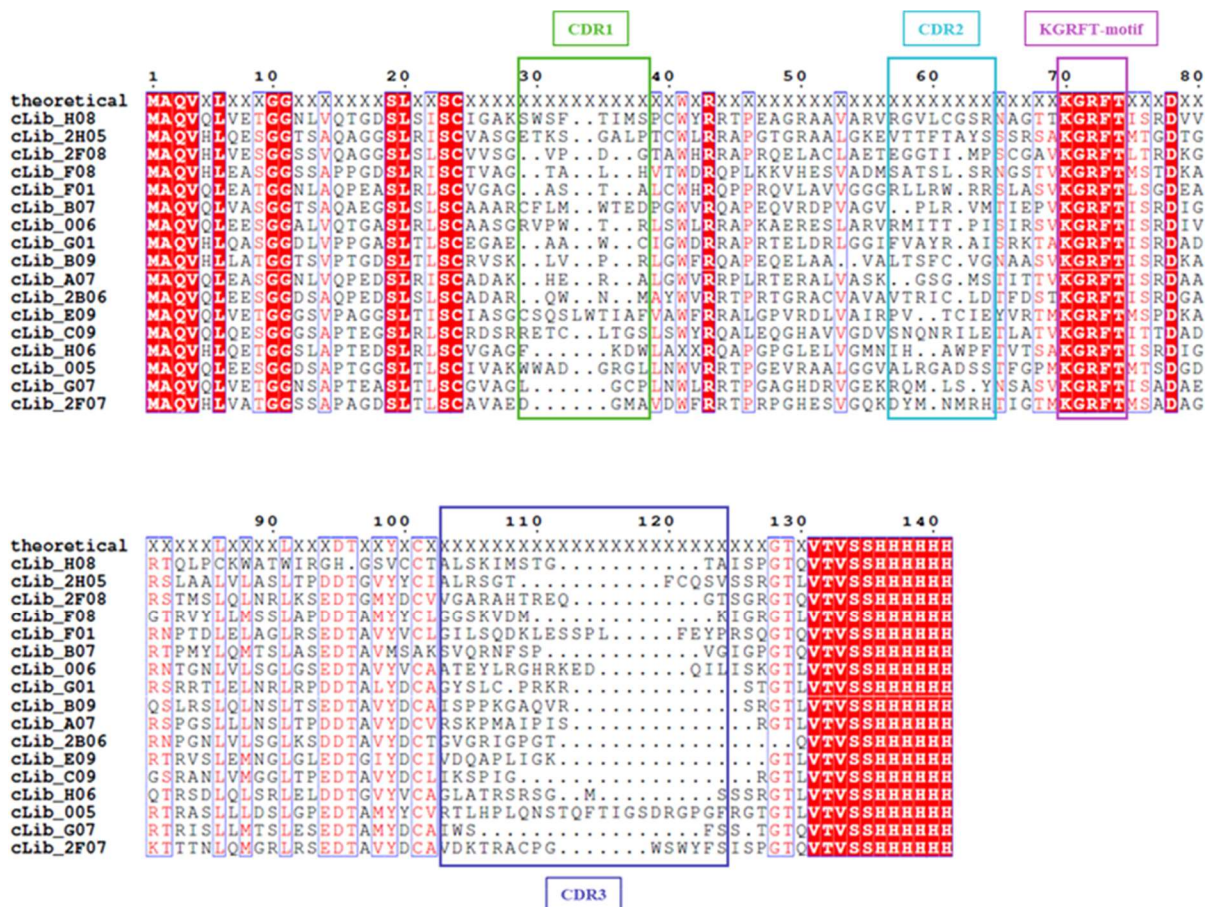
Figure 3.18 Sequence analysis of the c-lib. Multiple protein sequence alignment were generated with 17 subcloned and Sanger sequenced Nb genes from our own Nb library. Identical amino acid positions are marked in white letters and red background, position similarity in red letters and no similarity in black letters. Nb CDR1 (green), CDR2 (cyan), and CDR3 (blue) are highlighted in colored boxes. The amino acid sequence of the theoretical Nb library design is depicted as the top sequence. Multiple sequence alignment was performed by Clustal Omega multiple sequence alignment tool (Madeira et al., 2022) and visualized using the Esprit server (Robert & Gouet, 2014).

3.3.2 Analysis of the yeast Nanobody library

Similar to the c-lib, the y-lib was screened for well-behaved Nbs that overexpressed in *E. coli* and were translocated to the periplasm. Nb genes were subcloned into a pHen6-Nb-His expression vector, which was used to transform *E. coli* wk6, and Nb were expressed. After Nb expression, *E. coli* periplasm was extracted, and insoluble proteins were removed by centrifugation.
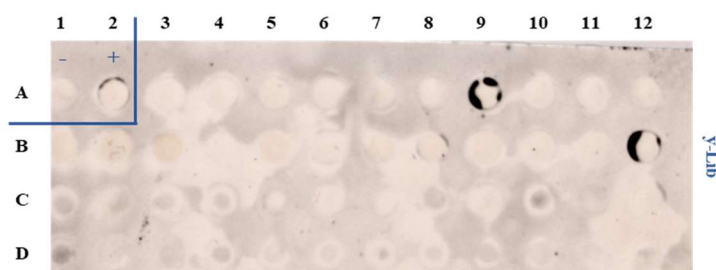
Figure 3.19 Nb expression tests of Nbs from the y-lib. 46 Nbs from the y-lib were expressed in *E. coli* wk6 and the periplasmic fractions were analyzed by Anti-His-tag dot/Western blot. An uninduced culture served as negative control and an Nb that is known to be well-behaved as positive control. Nbs A9, B8, and B12 could be detected.

Only three soluble Nbs could be detected in an anti-His-tag dot/Western blot of periplasmic extracts. Nbs A9, B8, and B12 yielded signal with varying intensities (Figure 3.19). These Nbs were validated by SDS-PAGE and subsequent anti-His-tag Western blot (Figure 3.20). The Western blot confirmed expression of Nbs A9, B8, and B12 with MWs of roughly 14 kDa. Nb B12 showed the highest expression levels (Figure 3.20).
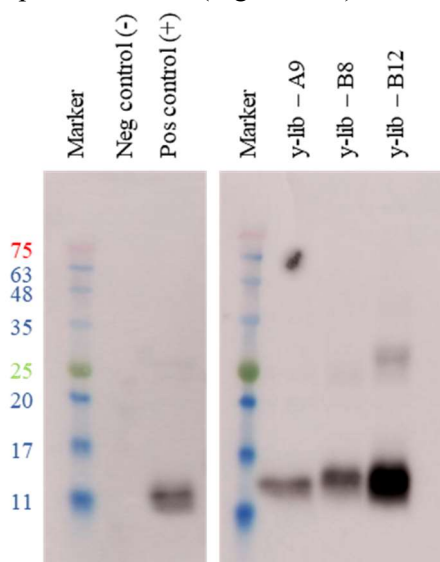
Figure 3.20 Expression of Nbs A9, B8, and B12. Expression and periplasmic localization of the y-lib Nbs A9, B8, and B12 was analyzed by SDS-PAGE and anti-His-tag Western blot of periplasmic fractions of Nb expression cultures.

After subcloning of the Nb genes from the y-lib, 40 Nb genes were Sanger sequenced (LGC genomics). Of those 40 sequences, 13 yielded full-length Nb genes. The remaining 27 contained truncated or frameshifted sequences. To compare the 13 full-length Nb sequences, protein sequence alignments were generated by the Clustal Omega Multiple sequence alignment tool (Madeira et al., 2022) and visualized using the Espript server (Robert & Gouet, 2014).

Of the 13 full-length Nb genes, Nbs ylib_033, yLib_038, and yLib_008 showed some deviations from the Nb scaffold, which was designed to be entirely conserved outside of the CDRs. By design, all CDR1 and CDR2 sequences show restricted variability with fixed lengths. CDR3 shows high levels of sequence-, and length-variability (Figure 3.21).
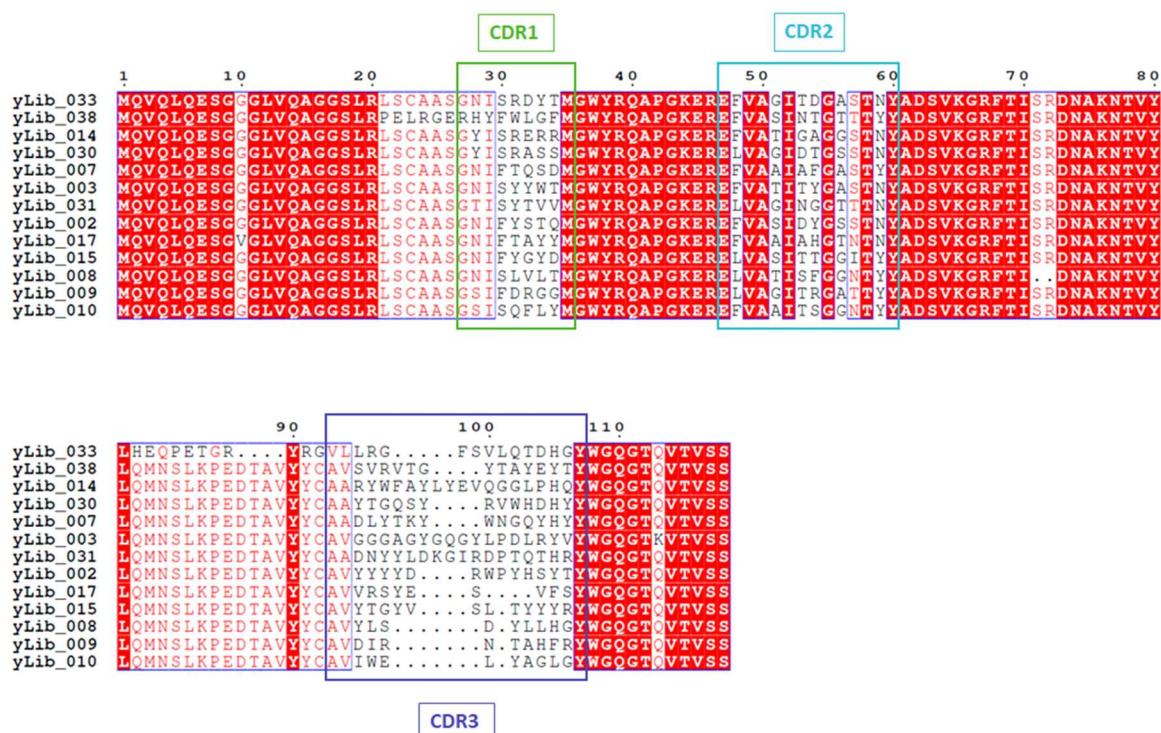
Figure 3.21 Sequence analysis of 13 Nbs from the y-lib. Identical amino acid positions marked in white letters and red background, position similarity in red letters and no similarity in black letters. Nb CDR1 (green), CDR2 (cyan), and CDR3 (blue) are highlighted in colored boxes. Multiple sequence alignment was performed by Clustal Omega multiple sequence alignment tool (Madeira et al., 2022) and visualized using the Esprit server (Robert & Gouet, 2014).

The analysis of Nb expression, and periplasmic translocation was greatly influenced by differences in expression levels, cell concentrations, efficiency of periplasmic extraction, protein transfer onto the PVDF-membrane, and protein detection. Additionally, the analyzed sample sizes in both experiments were quite low compared to the actual Nb library sizes. So, these experiments give some insights into the Nb libraries but do not describe the quality of the analyzed libraries adequately.

Sanger sequencing of a few Nb genes already revealed the main differences between the c-lib and y-lib. The y-lib uses a single Nb scaffold with invariant FRs and even some invariant positions in the CDRs. The c-lib uses a much higher degree of variability, with invariant positions only at the termini and in the KGRFT-motif in FR3. Analysis of Nb expression and periplasmic translocation indicated that the y-lib contains at least a few well-behaved proteins, while no well-behaved Nbs were detected in the c-lib. However, these results seem to be hardly representative, and should not be overemphasized.

3.4 Establishing of an mRNA/cDNA Nanobody selection protocol

Before the first Nanobody selection cycles were performed, protocols for the single reaction steps needed to be established. As mentioned, modification of the mRNA with puromycin by UV-crosslink was performed according to Seelig, 2011, while all other reactions like PCR, *in vitro* transcription,

translation, and reverse transcription, were set up similar to Doshi et al., 2014. The counter selection, bio-panning, and washing steps were newly established.

Even though protocols already existed for some steps, establishing an mRNA/cDNA selection procedure requires meticulous optimization for every library and every target antigen. The first steps in each selection cycle were two PCRs. Depending on the composition of the template DNA library, the outcome of a PCR can vary greatly.
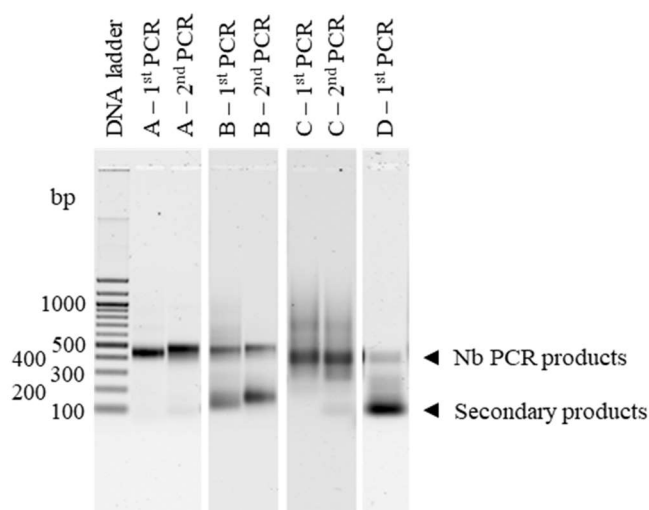
Figure 3.22 Different examples of entry PCRs into Nb selection cycles. PCR products were analyzed by agarose gel electrophoresis. A: Optimal PCRs with one distinct product of ~460 bp (1st PCR) and ~490 bp (2nd PCR), B: PCRs with the desired Nb fragments and a secondary low-MW product, C: PCRs without a visible ~30 bp shift in the 2nd PCR, and D: very dominant secondary product and almost complete loss of desired product.

Single fragments of roughly 460 bp and 490 bp were expected in PCR1 and PCR2 in each selection cycle (Figure 3.22 A). Because both Nb libraries contain some length variability the PCR products were stretched out over roughly 30 bp. In some PCRs, low-MW secondary products appeared (Figure 3.22 B), or a smeared signal was visible (Figure 3.22 C). If low-MW secondary products appeared, they were usually maintained through the entire selection process and into the next selection cycles, which often led to a gradual loss of the main PCR product (Figure 3.22 D).

In vitro transcriptions were expected to yield Nb mRNA of 450-480 nt, which shifted to 480-510 nt after being crosslinked to the puromycin oligonucleotide (Figure 3.23). The low-MW secondary product from PCRs was usually maintained after *in vitro* transcription and UV-crosslink, as seen in the transcription and UV-crosslink of selection cycle 3 (Figure 3.23).

In the transcription and crosslink products of selection cycle 3 (Figure 3.23), low-MW secondary products were present with sizes of 150 nt and 180 nt, respectively. The presence of the secondary transcription product suggests an intact 5'-end of the PCR product, since the T7 promoter, which is essential for transcription, is located at the very 5'-end of the PCR products. The secondary transcription product was also crosslinked to the puromycin oligo, indicated by the 180 nt species in the XL-RNA. Because the annealing site is located at the 3'-end of the mRNA, this suggests an intact 3'-end of the mRNA. This means that the low-MW secondary products have intact 5'- and 3'-ends but contain a ~300 nt deletion within the Nb gene.
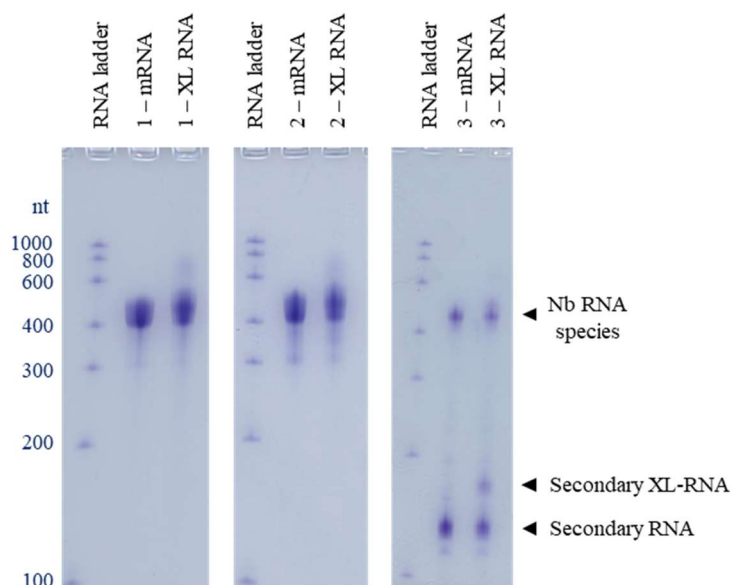
Figure 3.23 *In* vitro transcription and XL-RNA. Nb mRNA and XL-RNA were analyzed by toluidine blue stained urea PAGE. Transcription and UV-crosslink reactions of selection cycle 1 and 2 yielded large quantities of mRNA (~400-500 nt) and XL-RNA (430-530 nt), respectively. In selection cycle 3, a low-MW secondary RNA was more dominant (150 nt), which was crosslinked to the puromycin oligo and is shown as ~180 nt species.

After *in vitro* transcription and UV-crosslink, the puromycin-labeled Nb mRNA was *in vitro* translated using the PURExpress kit (New England Biolabs). This step generated Nb mRNA-protein hybrid molecules. To visualize the Nbs, nascent protein chains were labeled with the FluoroTect Green Lys *in vitro* Translation Labeling System (Promega) during translation. This allows visualization of synthesized proteins in SDS-PAGE by in-gel fluorescence. Nb mRNA-protein hybrid molecules have been shown to migrate at roughly 180 kDa in SDS-PAGE (Doshi et al., 2014).

Because of the low-efficiency of mRNA-protein hybrid formation and weak fluorescence, the Nb mRNA-protein hybrids were only visible after oligo(dT)-affinity purification, which removed background fluorescence and concentrated the Nb mRNA-protein hybrids. Because only the protein portion of the hybrid molecules was fluorescently labeled, the mRNA-protein hybrid molecules were digested by RNase A. A loss of fluorescent (protein) signal at 180 kDa after RNase digest confirmed the 180 kDa species were actually mRNA-protein hybrid molecules.

Nb mRNA-protein hybrid molecules were successfully synthesized in all three selection cycles, indicated by the fluorescent signal at 180 kDa, which was lost after RNase digest (Figure 3.24). The translation of third selection cycle 3 produced a secondary product, which migrated at 75 kDa (Figure 3.24). Formation of this secondary product were probably caused by the secondary product from UV-crosslink, which migrated at 180 nt in the urea-PAGE.
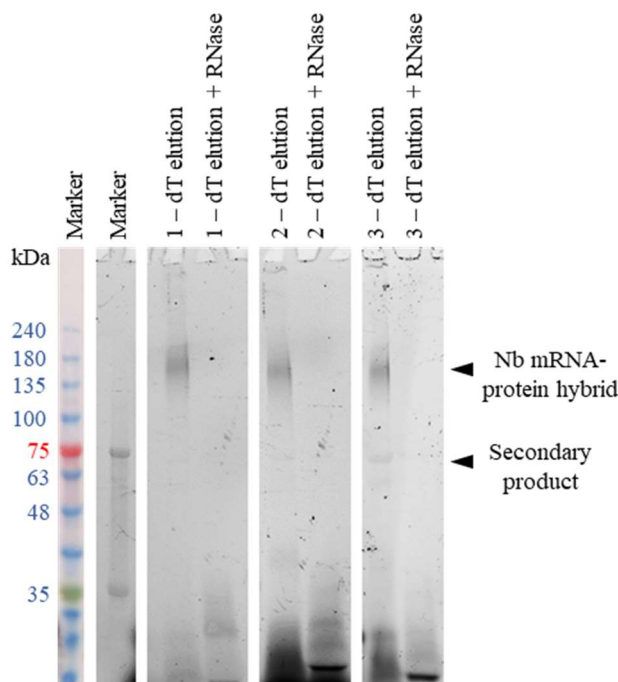
Figure 3.24 Analysis of Nb protein-mRNA hybrids by in-gel fluorescence. Fluorescent protein-labeling was visualized in SDS-PAGE and compared to a pre-stained protein marker (left). Nb mRNA-protein hybrids were synthesized in every selection cycle (1, 2, and 3), and were visible as 180 kDa in the oligo(dT) elution fraction. Removal of the 180 kDa species after RNase digest indicates RNA presence. In selection cycle 3, a secondary product was synthesized, which migrated at 75 kDa.

The synthesized and purified mRNA-protein hybrid molecules were then used for *in vitro* reverse transcription, counter selection against Strep-Tactin magnetic beads, complex formation with myrGCAP1-Strep and panning of the complexes on Strep-Tactin magnetic beads. After several washing steps, complexes of target antigen and Nb mRNA/cDNA-protein hybrids were eluted from the beads. Then, the Nb DNA sequences were regenerated by PCR, which started the next selection cycle.

After all single selection steps were established, a three-cycle Nb selection against myrGCAP1-Mg from the c-lib was performed. The selected Nb sequences were subcloned into a pHen6-vector. 38 of the subcloned Nb genes were sequenced.

Of the 38 analyzed sequences, 16 correlated with the full-length Nb design at least partially (Figure 3.25). Sequences Nb005 and 009 were entirely identical but showed a twelve amino acid deletion in FR1. Interestingly, Nb006 had a 23 amino acid CDR3, which was the maximum CDR3 length in the library design, but showed a CDR1 that was shorter than the minimal CDR1 length of four amino acids. Additionally, Nb006 lacked the essential cysteine in FR3.

Sequences Nb001, 004, and 008 followed the library design perfectly with CDR1s between four and eight amino acids, CDR2s of six or seven amino acids and CDR3s between eleven and 20 amino acids. All CDR lengths agree with the c-lib library design. Nb013 showed some deviations from the library design and lacked a CDR3. The remaining nine sequences all showed the mentioned CEER motif in FR3 and lacked CDR3.

The most promising Nbs, Nb001, 004, and 008 were expressed in *E. coli* wk6. All three Nbs were expressed, as indicated by the 11 – 17 kDa species in the total lysate fractions in the anti-His-tag Western blot but were not translocated to the periplasm (Figure 3.26).
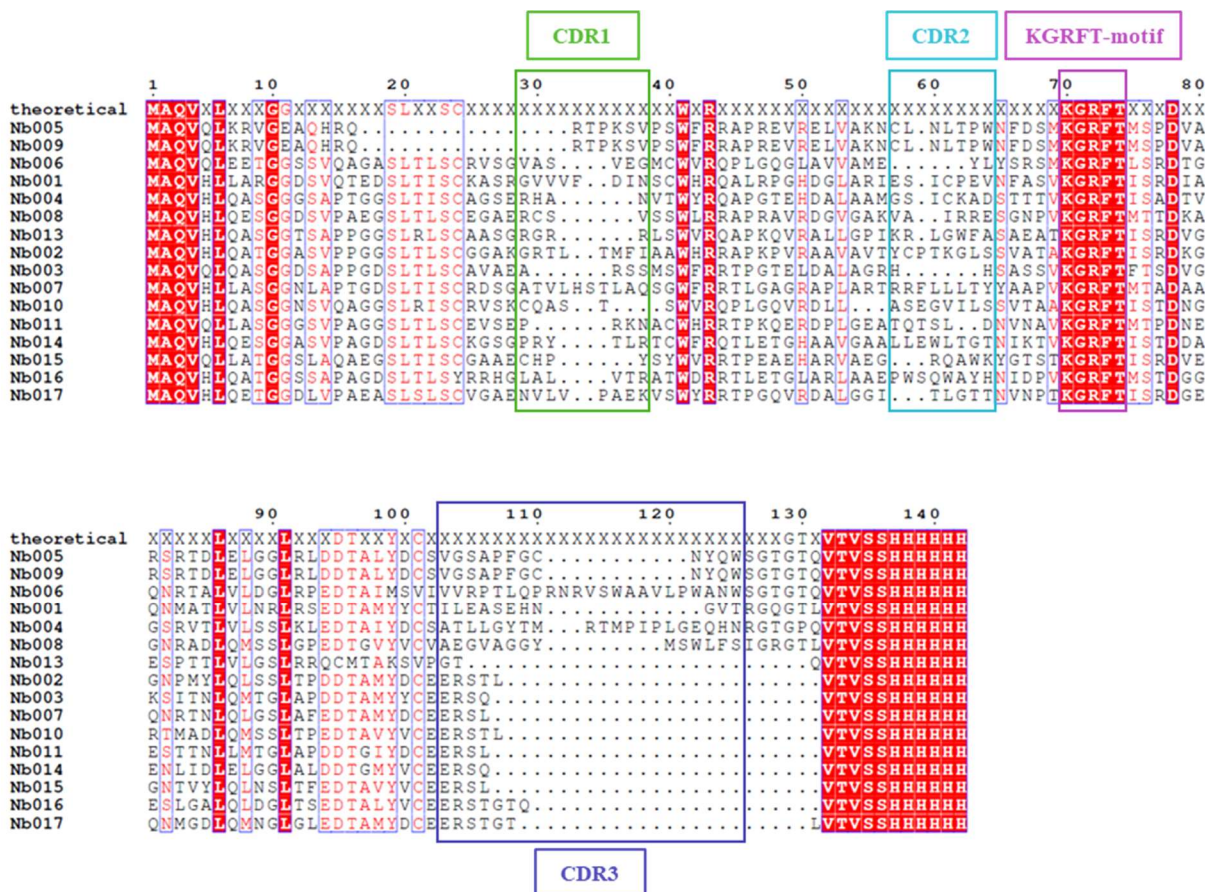
Figure 3.25 Multiple protein sequence alignment of Nb genes obtained from anti-myrGCAP1 selection from our own library. CDRs are marked in green (CDR1), cyan (CDR2), and blue (CDR3) boxes. The theoretical Nb sequence according to our library design is displayed as the top sequence. Multiple sequence alignment was generated by Clustal Omega multiple sequence alignment tool (Madeira et al., 2022) and visualized using the Espript server (Robert & Gouet, 2014).
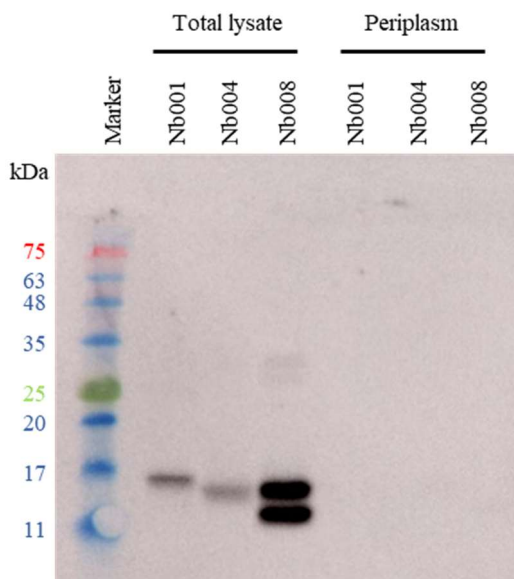

Figure 3.26 Test expressions of three anti-myrGCAP1-Mg Nbs. Nbs 001, 004, and 008 were expressed in *E. coli* wk6. Overall protein expression (Total lysate) and translocation to the periplasm (Periplasm) were analyzed by anti-His-tag Western blot.

3.5 *In vitro* selection of anti-myrGCAP1 Nbs from the y-lib

After all reaction steps of an Nb selection cycle were established and no myrGCAP1-Mg binder were selected from the c-lib, Nbs were selected against myrGCAP1-Mg from the yeast display Nb library (McMahon et al., 2018), y-lib.

To enter the first selection cycle, Nb genes were amplified by PCR from the isolated pYDS-Nb plasmids. In this PCR, Nb genes were successfully amplified, but with low efficiency, as indicated by the faint ~450 bp species (Figure 3.27). In the 2nd PCR, Nb genes were further amplified and modified for the following reactions. A T7 promoter, ribosome binding site, and a Flag-tag were added to the 5'-end, an annealing site for the puromycin oligonucleotide to the 5'-end. The added sequences increase the size of the Nb sequences to ~480 bp, which was confirmed by agarose gel electrophoresis. Additionally, the 2nd PCR of selection cycle 1 produced several high-MW products (Figure 3.27).

The PCRs of the selection cycle 2 generated almost pure Nb genes, with a smear in the 2nd PCR. In selection cycle 3, the Nb PCR products are still present. However, low-MW secondary products were more dominant (Figure 3.27). The secondary products were probably caused by overamplification of the DNA library.
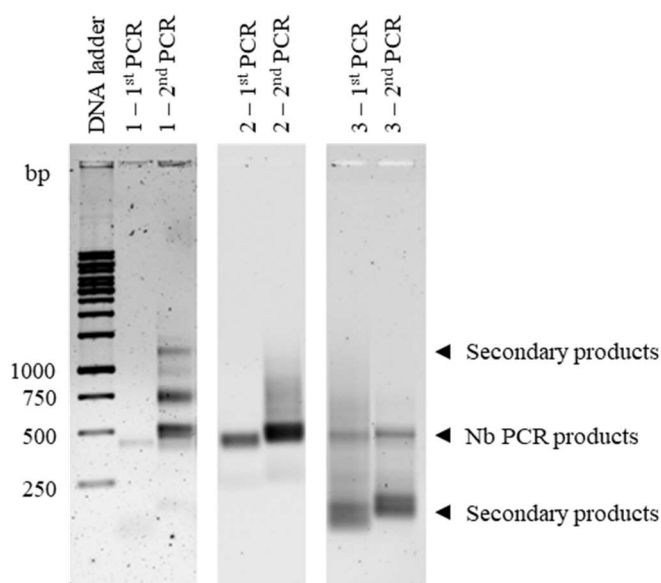
Figure 3.27 Entry PCRs into a 3-cycle Nb selection. Nb genes from the y-lib were amplified from isolated pYDS-Nb plasmids and generated ~480 bp (1st PCR) and ~510 bp (2nd PCR) species. The 1st PCR in selection cycle 1 generated several higher-MW secondary products. Both PCRs in cycle 3 produced low-MW secondary products, which were more dominant than the desired Nb PCR products.

The *in vitro* transcriptions and UV-crosslinks generated the desired Nb mRNA and XL-RNA molecules throughout three selection cycles, as shown in urea-PAGE (Figure 3.28). Even though the low-MW secondary products were barely present after PCR, and the Nb sequences were purified by gel extraction, the low-MW appeared after transcription and UV-crosslink of the second selection cycle. As expected, the low-MW species became more dominant in the third selection cycle (Figure 3.28).
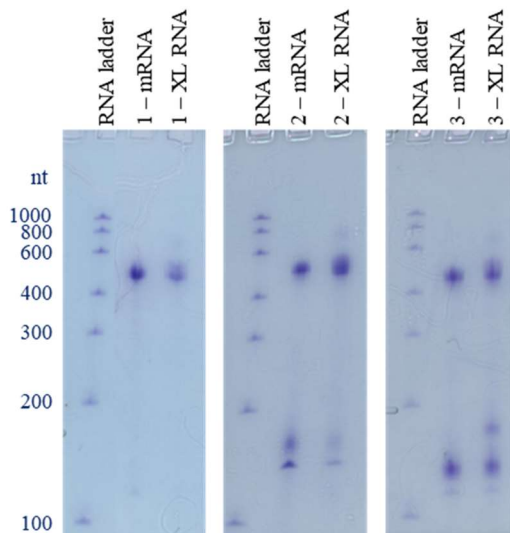
Figure 3.28 mRNA and XL-RNA synthesis. Y-lib Nb mRNA and crosslinked mRNA (XL-RNA) from the selection cycles 1, 2, and 3 against myrmyrGCAP1 were synthesized and analyzed by toluidine blue-stained urea-PAGE. In selection cycle 1, pure mRNA and XL-RNA were synthesized. In selection cycles 2 and 3, secondary products were present after transcription and UV-crosslink.

After UV-crosslink, the XL-RNA was used in *in vitro* translations to generate Nb mRNA-protein hybrid molecules. Successfully mRNA-displayed Nbs were then purified by oligo(dT)-affinity purification, and the purification fractions were analyzed by in-gel protein-fluorescence (Figure 3.29).
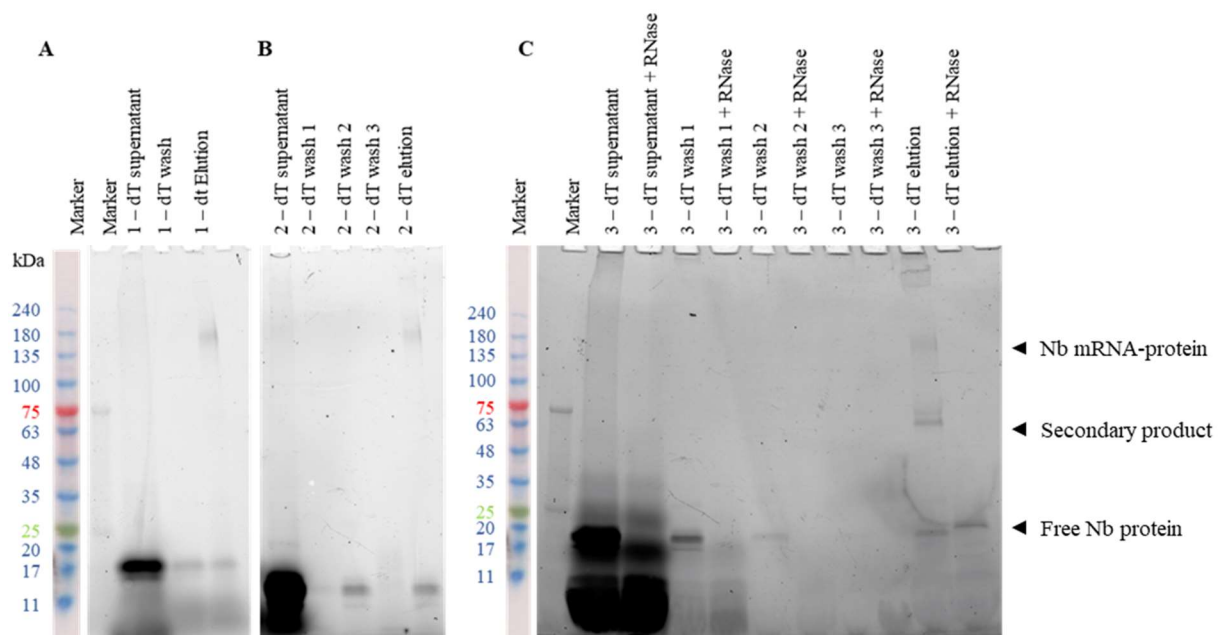
Figure 3.29 *In vitro* translation and oligo(dT) purification of Nb mRNA-protein hybrids. In-gel fluorescence of *in vitro* translated Nb mRNA/cDNA hybrid molecules on SDS-PAGE. A: In the first Nb selection cycle from the y-lib Nb mRNA-protein hybrids were synthesized and visible after oligo(dT)-purification as 180 kDa species in the elution fraction. B: In the second selection cycle, mRNA-protein hybrids could be visualized after oligo(dT)-purification, as well. C: All fractions of the oligo(dT)-purification were analyzed. Nb mRNA-protein hybrids were again only visible in the oligo(dT)-elution fraction and were removed by RNase A digest. Free Nb proteins could be visualized as ~20 kDa species.

Nb mRNA-protein hybrid molecules were synthesized in each selection cycle but were only visible after olgio(dT)-purification. In the selection cycle 3, a secondary mRNA-protein hybrid species was produced, which migrated at 75 kDa in SDS-PAGE and was also lost after RNase A digest (Figure 3.29

C). Presence of protein fluorescence combined with the loss after RNase A digest suggest that this secondary species is composed of mRNA and protein.

After oligo(dT)-purification, a cDNA strand was synthesized on the Nb mRNA by *in vitro* reverse transcription, which generated mRNA/cDNA-protein hybrid molecules. They were then used for counter selection against Strep-Tactin magnetic beads and Nb selection against myrGCAP1-Mg. After the first two selections, Nb genes were regenerated by PCR and used for the next selection cycle. After the third cycle, Nb genes were amplified and modified for subcloning by PCR (Figure 3.30). Nb genes were subcloned into a pHen6 vector by restriction ligation over NcoI and XhoI/SalI restriction sites – XhoI and SalI generate compatible overhangs after restriction digest.

The PCR generated Nb genes flanked with NcoI and SalI restriction sites, as shown in agarose gel electrophoresis as 450 bp species (Figure 3.30). A low-MW secondary product appeared in this PCR, as well. By extracting the NcoI-Nb-SalI fragments from the gel, the secondary product was removed.
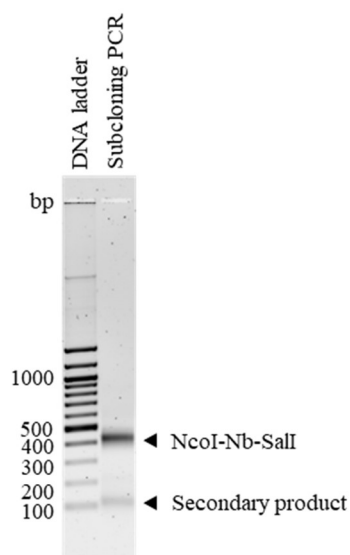
Figure 3.30 Analysis of PCR product for subcloning. After three cycles of anti-myrGCAP1 selection, Nb genes were amplified by PCR. During PCR, NcoI and SalI restriction sites were introduced. The PCR generated a dominant product, migrating at 450 bp in agarose gel electrophoresis, which correlates well to the size of the Nb genes. A secondary product appeared at ~100 bp.

After PCR and restriction digest, pHen6-Nb-Avi-His plasmids were cloned by restriction ligation and used to transform *E. coli* TOP10. 95 clones were sequenced by Direct Colony Sequencing (Eurofins). Only two of the analyzed Nb sequences (D10 and F7) matched a full-length Nb sequence with all FRs, CDRs and the C-terminal Avi-His tag in-frame (Figure 3.32, two top sequences).

Those two Nbs were expressed in *E. coli* wk6 (Figure 3.31 A) and purified by Ni-NTA-affinity chromatography and SEC over a Superdex 75 pg 10/300 column (Cytiva) (Figure 3.31 B). Nb F7 showed one major peak at 13.76 ml in the SEC chromatogram, while the chromatogram of Nb D10 had two peaks, a smaller one at 11.35 ml and a taller one at 12.78 ml (Figure 3.31 B).
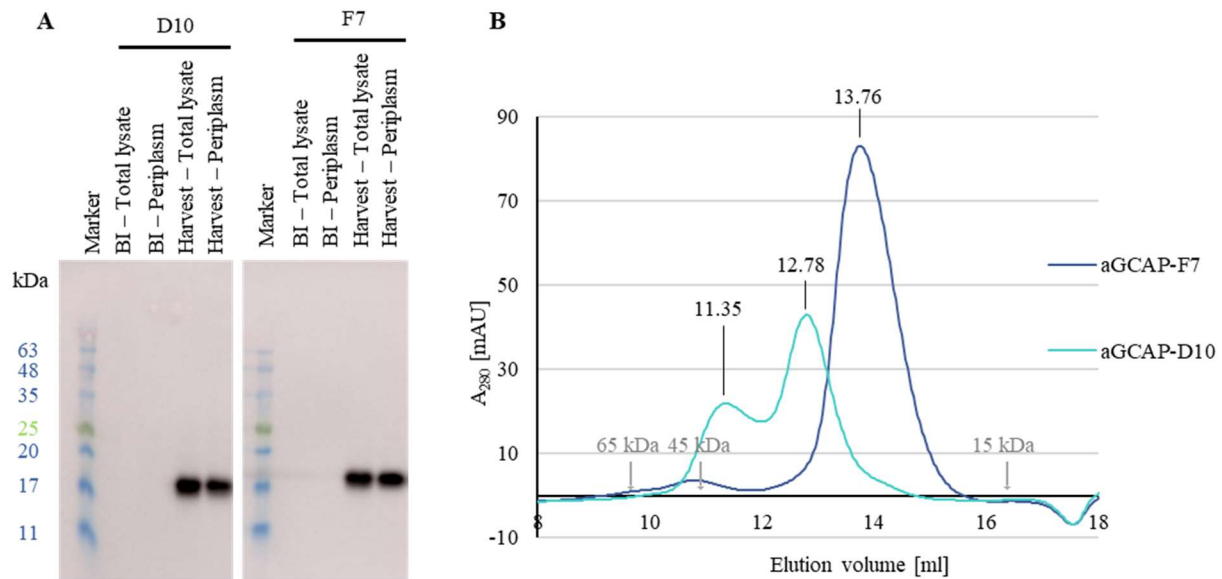
Figure 3.31 Expression and purification analysis of Nbs aGCAP1-Mg-F7 and D10. A: Expression of Nbs D10 and F7 in *E. coli* wk6 produced well-behaved Nbs, which were translocated to the periplasm, as indicated by the ~17 kDa species in the Harvest – Periplasm fractions in the anti-His-tag Western blot. B: After affinity chromatography, both Nbs were purified by SEC on a Superdex 75 pg 10/300 column. Nb D10 eluted with two peaks after 11.35 ml and 12.78 ml. Nb F7 eluted with a single peak at 13.76 ml.

Nine of the sequenced Nb genes contained an N-terminus and FR1 according to library design and an intact sequence from CDR1 to the His-tag, after a frameshift. Because *E. coli* ribosomes have been shown to possess frameshifting abilities (Flower & McHenry, 1990), we hypothesized that the *E. coli* ribosomes in the PURE system frameshifted and synthesized full-length Nb proteins. Therefore, frameshift-containing Nb DNA sequences were corrected for the frameshift by Gibson Assembly cloning. Because FR1 is entirely conserved in the y-lib, primers were designed to add the selected Nb sequences (CDR1 to His-tag) to a pHen6-Avi-His-vector containing the successfully cloned Nb FR1.

Gibson assembled pHen6-Nb-Avi-His sequences were used for transformation of *E. coli* TOP10, subsequent plasmid isolation and Sanger sequencing. Rescue of the selected Nb sequences yielded in-frame Nb genes from N-terminal PelB translocation sequence to the C-terminal Avi-His-tag and was confirmed by Sanger sequencing (Figure 3.32).

Figure 3.32 Comparison of rescued Nb sequences. After anti-myrGCAP1 selection from the y-lib, eleven frameshift-containing Nbs were rescued by Gibson Assembly cloning, generating full-length Nb sequences. Multiple sequence alignment was performed by Clustal Omega multiple sequence alignment tool (Madeira et al., 2022) and visualized using the Esprit server (Robert & Gouet, 2014).

Noticeably, the Nbs aGCAP-B3 and B10 were completely identical. As expected, variability of the Nb sequences is limited to CDR3 and certain positions in CDRs 1 and 2 (Figure 3.32). FR3 (between CDR2 and CDR3) showed some unexpected variability, which was either introduced during library construction or as PCR errors during selection. CDR3s of aGCAP-B3/B10 and -A9 are missing one amino acid from the shortest theoretical Nb design. All other CDRs agree with the theoretical library design.

After the Nb sequences were successfully rescued, they were expressed and *in vivo* biotinylated in *E. coli* wk6 cells (co-transformation with BirA plasmid containing biotin ligase BirA), purified, and used to analyze complex formation with myrGCAP1 by SPR spectroscopy.
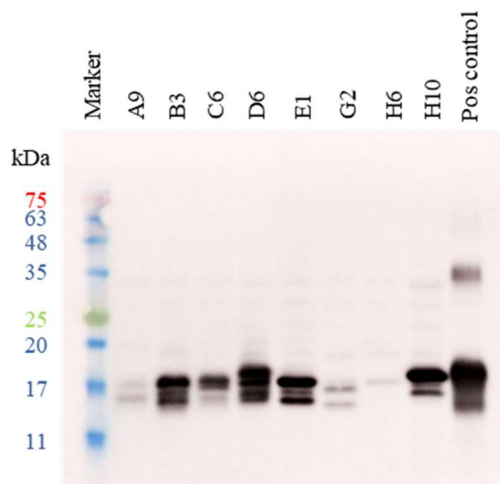
Figure 3.33 Test expression of rescued Nb sequences. Periplasmic fractions from Nb expression tests in *E. coli* wk6 were analyzed. Rescued anti-myrGCAP1 Nbs were *in vivo* biotinylated on their C-terminal Avi-tag. Nb presence and biotinylation was assessed by anti-biotin Western blot.

Nbs B3, C6, D6, E1 and H10 were overexpressed, biotinylated, and translocated to the *E. coli* periplasm, as indicated in the anti-biotin Western blot of periplasmic fractions of Nb expression cultures (Figure 3.33). No substantial amounts of Nbs A9, G2, and H6 were expressed and translocated to the periplasm.

Nbs B3, C6, D6, E1 and H10 were then purified by affinity chromatography over Ni-NTA resin and afterwards by SEC over a Superdex 75 pg 10/300 column (Cytiva).
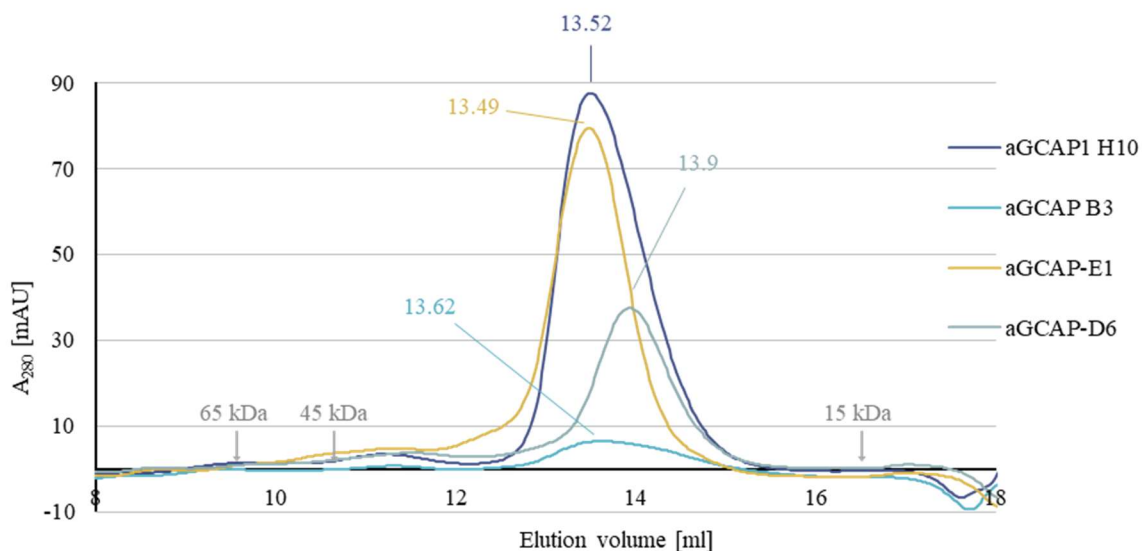
Figure 3.34 Size-exclusion chromatography purification of rescued anti-GCAP1 Nbs. After expression in *E. coli* wk6, anti-GCAP1 Nbs H10 (blue), B3 (teal), E1 (gold), and D6 (grey) were purified by affinity chromatography and SEC on a Superdex 75 pg 10/300 column. All Nbs eluted with single peaks at 13.49 (E1), 13.52 (H10), 13.62 (B3) and 13.9 (D6). Nb B3 was purified with barely detectable amounts.

Substantial amounts of Nbs H10, E1, and D6 were purified and showed distinctive peaks in the chromatograms at 13.52 ml (H10), 13.49 ml (E1), and 13.9 ml (D6). Nb B3 produced only a small peak at 13.62 ml (Figure 3.34).

3.5.1   Analysis of myrGCAP1 and anti-myrGCAP1 Nb complexes by SPR spectroscopy

Anti-myrGCAP1-Mg Nbs F7, D6, D10, E1, and H10 were purified from *E. coli* wk6 periplasm. Complex formation of these Nbs with myrGCAP1-Mg was analyzed by FO-SPR.

At first, Nbs F7 and D10 were analyzed. The Nbs were bound to SPR-sensors functionalized with Streptavidin over their biotinylated Avi-tag, and complexes were formed with C-terminally His-tagged myrGCAP1-Mg. Nbs D6, E1, and H10 (and F7 to validate the previous result) were bound to SPR-sensors functionalized with Ni-NTA, which were not available for the previous experiment. Here, C-terminally Strep-tagged myrGCAP1-Mg was used for complex formation.

The SPR measurement detects changes in the refractive index around the sensor, which are measured as wavelength shifts (Surface plasmon absorption maxima are located at higher wavelengths in media with higher refractive indices). These changes in the refractive index are either generated by a change in

buffer – buffer changes generate sudden wavelength shifts – or by a change in the thickness of the layer bound to the sensor, which is visible as rate dependent growth curves.
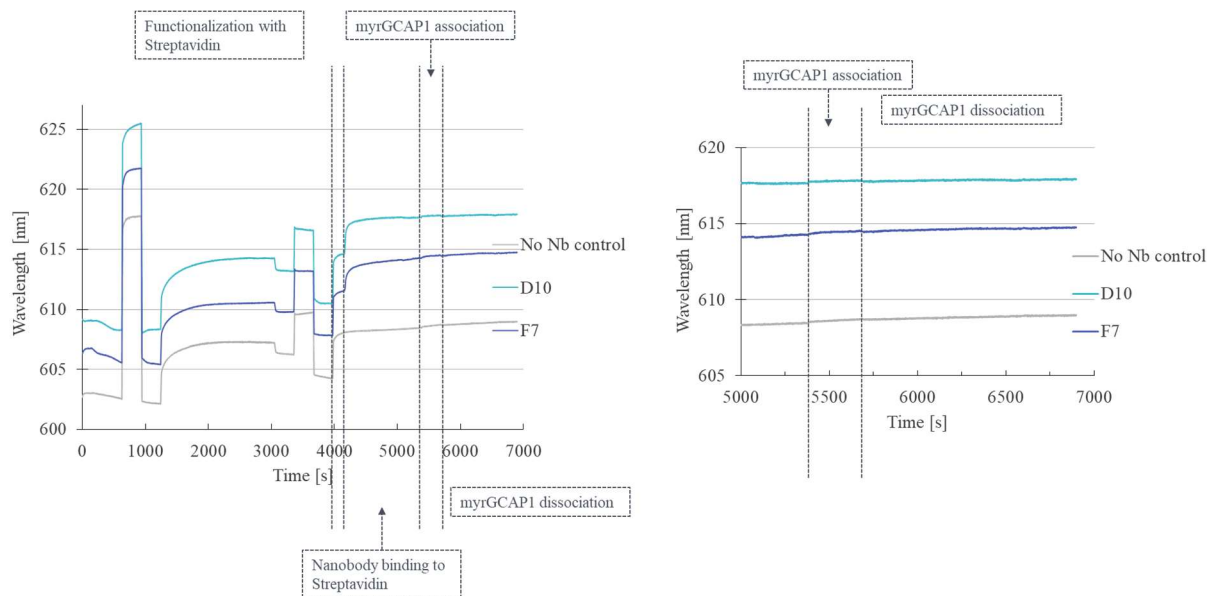
Figure 3.35 Analysis of Nb-myrGCAP1 complex formation. Complex formation of myrGCAP1 and selected Nbs anti-GCAP1-D10 (teal) and -F7 (blue) were analyzed by FO-SPR and compared to a no-Nb negative control (left). Biotinylated Nbs were bound to Streptavidin-functionalized SPR-sensors. The myrGCAP1 association and dissociation steps were enlarged for enhanced visibility (right). No complexes were formed, indicated by the lack of wavelength increase in the myrGCAP1 association step after 5400 s.

After Streptavidin was successfully immobilized on the COOH-sensor surface (Time $t < 4000$ s), the biotinylated Nbs were bound to Streptavidin (4000 s $< t <$ 5100 s), resulting in an upwards shift of ~3 nm (Figure 3.35, blue and teal). In the 'no-Nb control' (Figure 3.35, grey), no wavelength shift was generated in the same timeframe.

If Nbs D10 and F7 were myrGCAP1-binders, another wavelength shift was expected in the myrGCAP1-association step (Figure 3.35; 5400 s $< t <$ 5700 s). No wavelength shift was generated, meaning that the selected Nbs do not have substantial binding affinities to myrGCAP1-Mg.

To analyze complex formation of myrGCAP1-Mg with the 'rescued' Nbs D6, E1, and H10, SPR-sensors with an NTA-surface were available. So, after the NTA molecules were loaded with Ni^{2+}, the mentioned Nbs – and F7 to validate the previous results (Figure 3.35, blue) – were immobilized on the Ni-NTA sensor surface via a C-terminal His-tag (Figure 3.36; $t < 2100$ s). After 2440 s, an upwards-shift was expected for myr-GCAP1-binding Nbs, which could not be detected (Figure 3.36; 2440 s $< t <$ 2800 s). This indicates that the 'rescued' Nbs do not have substantial binding affinities to myrGCAP1-Mg, as well.
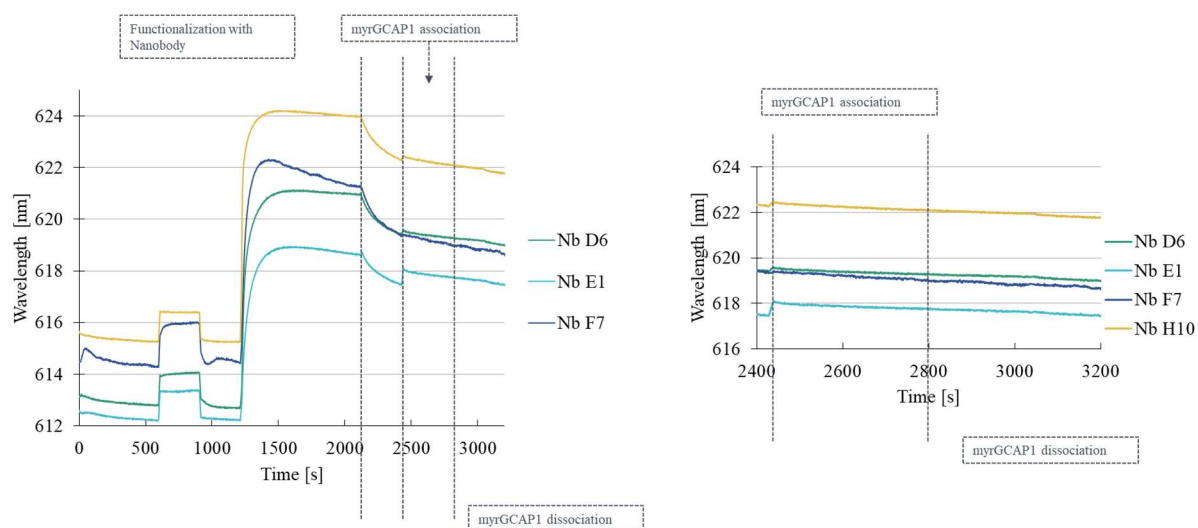
Figure 3.36 Analysis of Nb-myrGCAP1 complex formation. Complex formation was assessed by SPR. Anti-GCAP1 Nbs D6 (green), E1 (teal), F7 (blue), and H10 (yellow) were bound to the Ni-NTA sensor surface followed by a complex formation step (myrGCAP1 association) (left). The myrGCAP1 association and dissociation steps were enhanced for betted visibility(right). No complexes were formed, indicated by the lack of wavelength increase in the myrGCAP1 association step after 2440 s.

'Rescuing' the mentioned Nbs did not result in any Nbs that showed substantial affinities for myrGCAP1-Mg, and the hypothesis that the ribosomes in the PURE system introduced frameshifts during Nb translation was disregarded. The more likely scenario is that the ribosomes stalled, when encountering a premature Stop-codon, which allowed the puromycin to enter the A site, and the prematurely stopped nascent protein chain was displayed on the mRNA. This process then generated short peptides, which bound either myrGCAP1-Mg or the Strep-tactin magnetic beads unspecifically.

So far, no target-binding Nbs could be selected with the established mRNA/cDNA display selection procedure. To remove non-specific binders and to enrich the actual target-binding Nbs, the single steps of the selection protocol were optimized before starting another selection cycle.

3.6 Optimization of single steps of the Nb selection procedure

After several attempts to select Nbs by mRNA/cDNA display, no well-behaved, target-binding Nbs could be selected. Therefore, we decided to optimize the single steps of the selection procedure before performing the next Nb selections.

For the PCRs, two parameters were altered. First, primer concentration was reduced from 1 µM to 500 nM (Figure 3.37 A). This led to a reduction of the low-MW secondary product migrating at ~150 bp but did not improve the smeared signal. Secondly, the PCR was cycle-optimized in order to prevent overamplification of the library DNA. In cycle-optimization PCRs, samples were taken after eight, ten, twelve, etc. PCR cycles and analyzed by agarose gel electrophoresis.
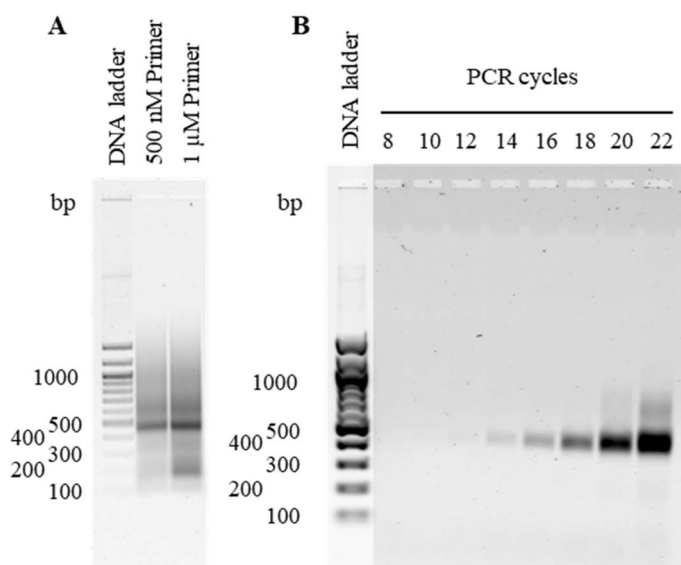
Figure 3.37 Analysis of optimization entry PCRs. A: Nb genes were amplified by PCR with primer concentrations of 1 µM and 500 nM. Both PCRs amplifies the Nb genes, as indicated by the ~450 bp species, but produced smeared signals. The smear was slightly weaker in the PCR with a primer concentration of 500 nM. B: The optimal cycle number of the PCR was determined (right). Samples were taken after every second PCR cycle from 8 to 22 cycles. All PCRs were analyzed by agarose gel electrophoresis. Detectable amounts of Nb were synthesized after 12 PCR cycles. After 20 cycles, secondary PCR products appeared.

In a PCR from a library DNA template, every sequence is doubled in each PCR cycle, as long as the primers are present in abundance, which generates perfect DNA double strands. In each PCR cycle, the number of consumed primers equals the number of template DNA strands. Once the number of amplified Nb genes gets close to the number of remaining primer molecules, not every Nb gene can be doubled. In the denaturing step of a PCR cycle, every DNA double strand is denatured. Because of the limiting primer amounts, some DNA single strands remain unbound by a primer. When the temperature is reduced in the annealing step, unbound single strands can anneal because of their sequence similarity. But, because the DNA library is composed of many similar but unequal Nb genes, two different Nb sequences can anneal – with several mismatches. This rehybridization in overamplified PCRs is known to interfere with PCR performance (Mathieu-Daude, 1996). Especially in PCRs that use DNA libraries as template, overamplification must be avoided.

The right cycle number is dependent on several factors like template concentration, primer concentration, and primer annealing efficiency. Because template concentration and library composition can vary after each selection, the optimal cycle number has to be determined for every single PCR.

In this example of a cycle-optimization PCR (Figure 3.37 B), the desired ~500 bp product started to appear after 12 PCR cycles. After 20 PCR cycles, the signal smeared to higher MWs, and after 22 cycles, a distinct ~700 bp fragment appeared. The optimal cycle number with maximum amplification and without secondary products in this PCR was 18.

Because optimal purity and quantity of the crosslinked mRNA are essential for the synthesis of mRNA-protein hybrid molecules during *in vitro* translation reactions, the efficiency of the UV-crosslink was improved. The most important parameters that could be optimized in this reaction were UV-irradiation time and the ratio of Puromycin-containing oligonucleotide to Nb mRNA. In previous crosslinks, a ratio of 4.5:1 (3.5-fold molar excess of Puromycin-containing oligonucleotide) and an irradiation time of

20 min were used according to Seelig, 2011. Because the high excess of the ribosome-binding antibiotic Puromycin might be detrimental to the subsequent *in vitro* translation, we assessed if the crosslink-efficiency is maintained, or even increased, at lower excesses of Puromycin-containing oligonucleotide.
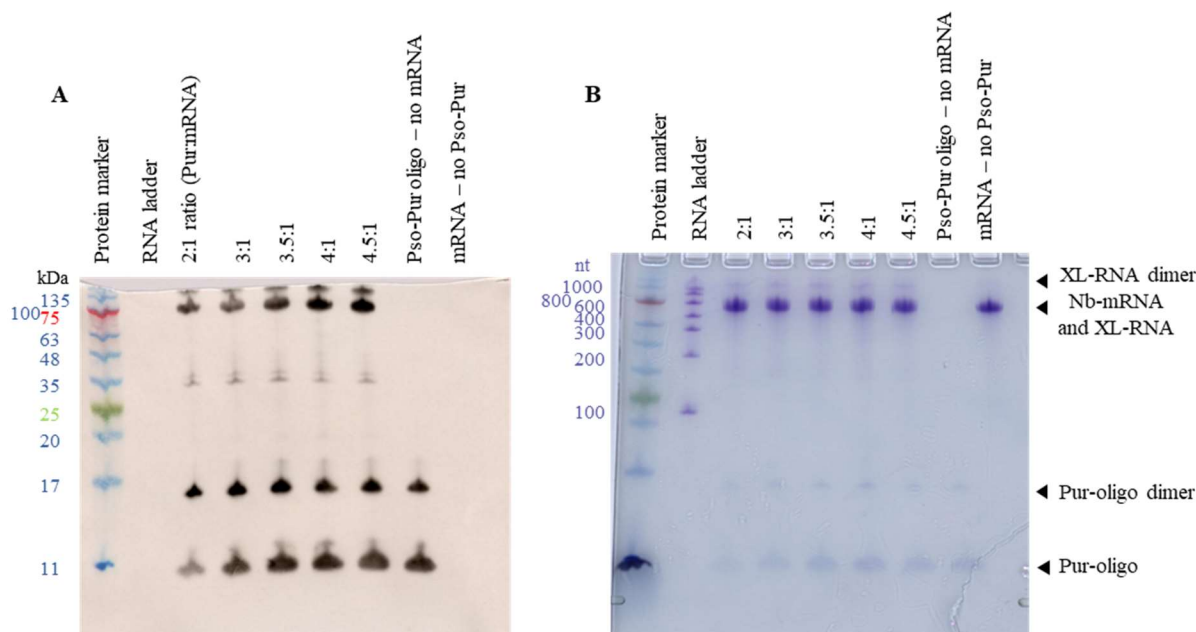
Figure 3.38 Optimization of Puromycin-containing oligo to Nb mRNA in UV-crosslinks. Puromycin- (Pur) and Psoralen- (Pso) containing oligos were used at 2:1 to 4.5:1 (oligo:mRNA) ratios. Crosslinked mRNA was analyzed by anti-Puromycin Northern blot (left) and toluidine blue-stained urea-PAGE (right).

UV-crosslink experiments were conducted with ratios of 2:1 to 4.5:1 (oligo:mRNA) and were analyzed by anti-Puromycin Northern blot (Figure 3.38 A) and toluidine blue-stained urea-PAGE (Figure 3.38 B). To be able to assess fragment sizes, a Riboruler RNA marker (Thermo Fisher) and a pre-stained protein marker (Applichem) were used in the toluidine blue-stained gel and in the Northern blot. The desired fragment has a size of ~500 nt and migrates at 75-100 kDa compared to the pre-stained protein marker. Additional Puromycin-positive fragments were visible at ~10 and ~16 kDa (different states of the oligonucleotide), 35 kDa (probably caused by a secondary transcription products), and at ~150 kDa, which might be a dimer of the XL-RNA. Ratios of 4:1 and 4.5:1 yielded the strongest signal of the desired fragment. Based on this, the high excess of Puromycin-containing oligonucleotide was maintained.

After the optimal ratio of Puromycin-containing oligonucleotide to Nb mRNA was determined, UV-irradiation time was optimized. According to (Seelig, 2011), an irradiation time of 20 min was used in past experiments. But different annealing conditions of the fragments, a different UV-lamp, and a different irradiation setup might alter the optimal irradiation time. Reducing irradiation time would prevent potential radiation damage of the mRNA. Annealed mRNA and puromycin oligo were irradiated

at 365 nm for 1 min, 2 min, 4 min, 8 min, 16 min, and 24 min. A non-irradiated sample served as negative control.

Nb XL-RNA was already generated after 1 min of irradiation time, as indicated by the ~75 kDa species in Northern blot (Figure 3.39 A) and toluidine blue-stained urea-PAGE (Figure 3.39 B), which was not present in the unirradiated sample. The amount of XL-RNA slightly increased with longer irradiation times, with a maximum after 24 min. Exceeding irradiation times of 20 min would increase the risk of radiation damage, and an irradiation time between 15 and 20 min was assessed to be a good compromise between optimal crosslinking-efficiency and minimal radiation damage.

Secondary products of higher and lower MW were generated to the same extend in all samples, and purification of the desired XL-RNA was necessary.
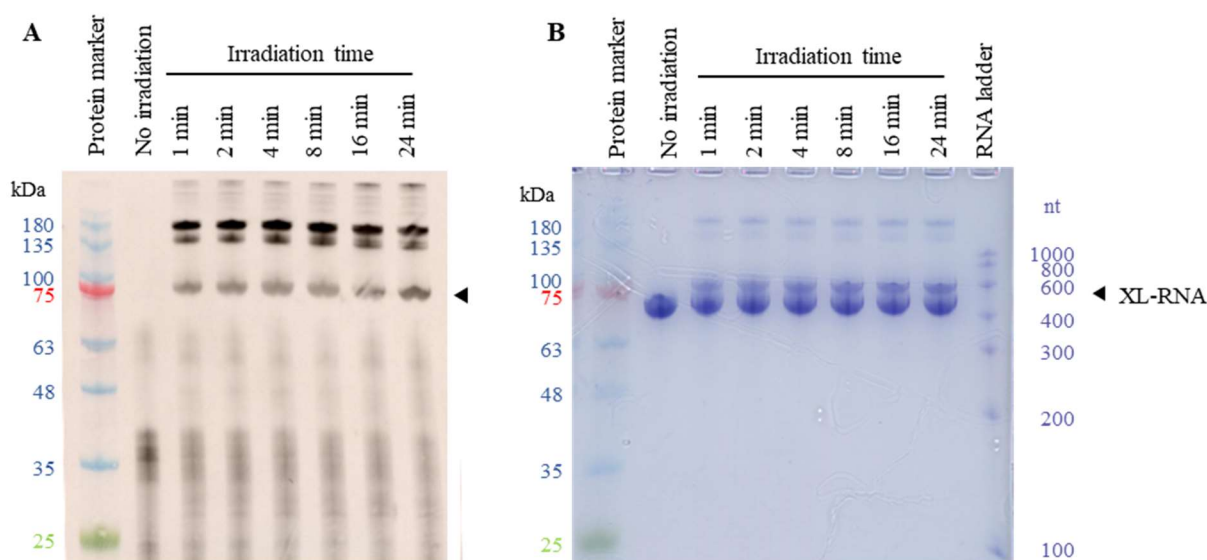
Figure 3.39 Optimization of UV-crosslinking times. Nb mRNA and the puromycin-containing oligonucleotide were crosslinked with irradiation times between 1 and 24 minutes. A non-irradiated sample (No irradiation) served as negative control. Crosslinked samples were analyzed by anti-Puromycin Northern blot (left) and toluidine blue-stained urea-PAGE (right).

After optimization of the crosslinking efficiency, the purity of the XL-RNA was increased by several purification steps. First, unwanted species of XL-RNA and the large excess of Puromycin-containing oligonucleotide were removed by excising RNA of the desired size (Figure 3.39) from an unstained urea-polyacrylamide gel – the desired fragment was chosen based on co-migration with the 75 kDa pre-stained protein marker species. RNA was then released from the gel pieces by electro elution into a dialysis bag filled with electrophoresis buffer.
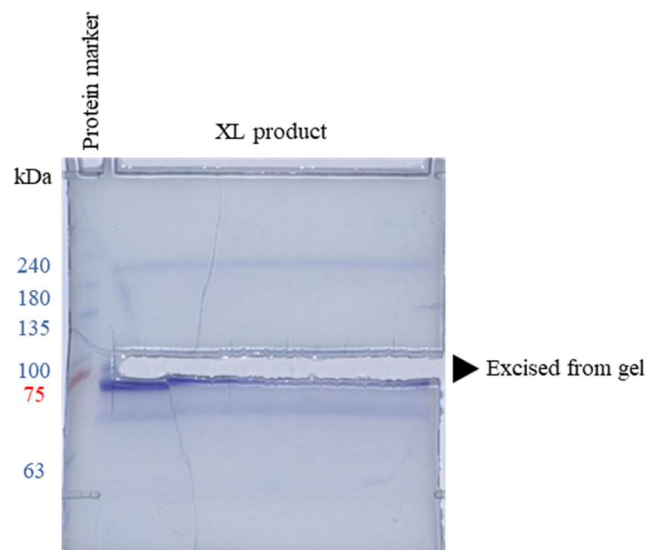
Figure 3.40 Excision of XL-RNA from urea-PAGE. After UV-crosslink (XL), the XL-RNA is separated by size in urea-PAGE. The desired fragment that migrates at roughly 75 to 100 kDa compared to the pre-stained protein marker was excised from the gel. After extraction of the XL-RNA, the RNA was stained with toluidine blue.

Because un-crosslinked Nb mRNA and XL-RNA partially overlap in size, XL-RNA could not be separated from mRNA by gel extraction. To eliminate all un-crosslinked Nb mRNA molecules, the XL-RNA was further purified by oligo(dT)-affinity (Figure 3.41), utilizing the 15-times A-tail in the Puromycin-containing oligonucleotide. Un-crosslinked mRNA stayed in the unbound oligo(dT) flow-through fraction. Un-crosslinked mRNA is visible in the toluidine blue-stained urea-PAGE (Figure 3.41 A), but undetected in the anti-Puromycin Northern blot (Figure 3.41 B). So, un-crosslinked mRNA and was not maintained through this purification step, pure XL-RNA was present in the oligo(dT) elution fraction. XL-RNA is visible in both, toluidine blue-stained urea-PAGE and anti-Puromycin Northern blot (Figure 3.41).
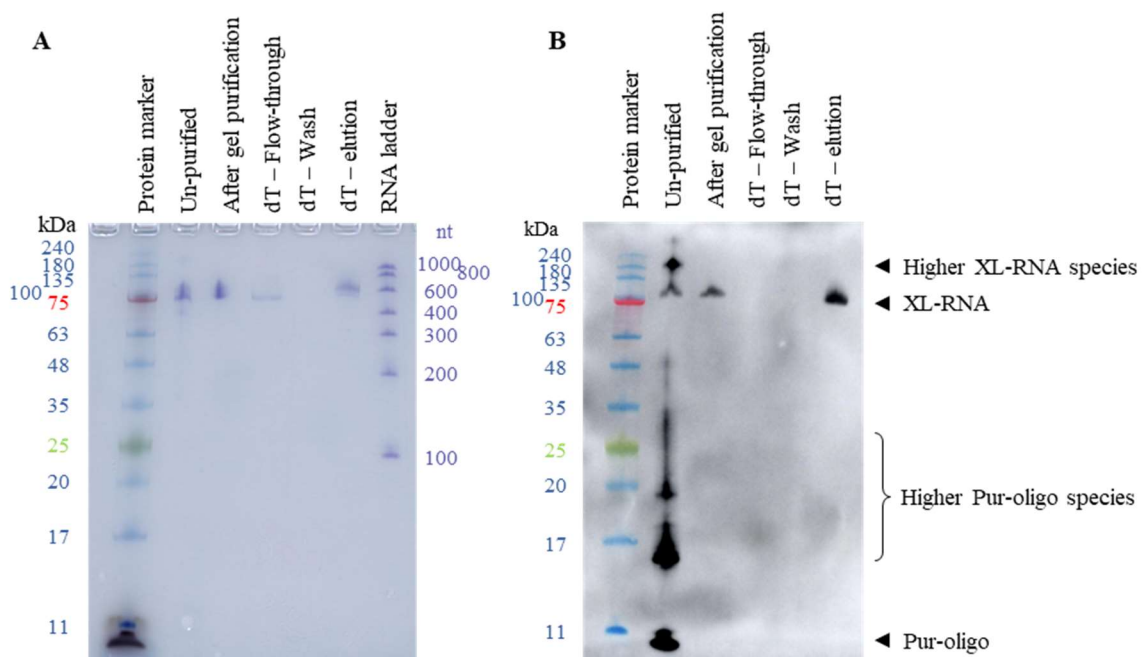
Figure 3.41 Analysis of UV-crosslink and purification of Nb XL-RNA. XL mRNA was excised from a urea-polyacrylamide gel and oligo(dT)-purified. Unpurified XL-RNA and fractions from gel extraction and oligo(dT) purification were analyzed by A: toluidine blue-stained urea-PAGE and B: Anti-puromycin Northern blot.

Using purified XL-RNA was then used in *in vitro* translations to generate Nb mRNA-protein hybrid molecules, which were purified over oligo(dT) resin and then used in reverse transcription.

In the reverse transcription step, temperature was lowered to reduce the heat-stress imposed on the Nb proteins. Nbs in general show a broad range of denaturing temperatures. Some Nbs denature at temperatures as low as 38 °C, but mostly in the range of 48 °C to 78 °C (Kunz et al., 2017). Temperature during the reverse transcription reaction was decreased from 55 °C to 42 °C. Additionally, DTT was removed from the reaction mix to avoid reduction of Nb-stabilizing disulfide bonds. To test, whether the reverse transcription still functions under these conditions, a reverse transcription reaction was performed with a Cy5-labeled reverse transcription primer.

In both reverse transcription reactions, similar amounts of Nb mRNA/cDNA-protein hybrids were synthesized, as shown by the 180 kDa species in SDS-PAGE (Figure 3.42). This means that the temperature reduction and removal of DTT did not decrease the efficiency of the reaction.

A secondary product, which migrated at 75 kDa, was present in both reactions but has a slightly higher intensity in the 42 °C reaction (Figure 3.42). Its size correlates nicely with the MW of the SuperScript IV reverse transcriptase (74 kDa). Due to insufficient denaturing before SDS-PAGE, Cy5-labeled primers were not completely removed from the enzyme. To entirely denature the reverse transcriptase and remove the primer molecules, the 95 °C denaturing step needs to be further increased from the 15 min used for the analyzed samples.
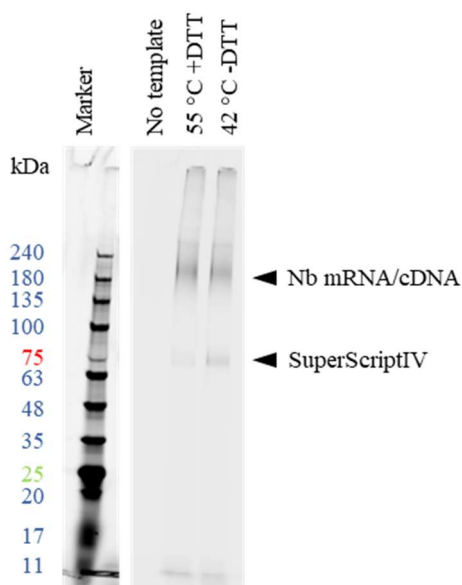
Figure 3.42 Temperature-optimization of *in vitro* reverse transcription. Synthesized cDNA was visualized by in-gel (SDS-PAGE) Cy5-fluorescence, introduced by a Cy5-labeled primer. Equal amounts of Nb cDNA were synthesized at 55 °C than at 42 °C (no DTT). Insufficient denaturation of the SuperScript IV reverse transcriptase before SDS-PAGE led to Cy5-primer remaining bound to SuperScript IV, which produced the 75 kDa species in SDS-PAGE.

The generated mRNA/cDNA-protein hybrid molecules were then used for counter selection against Strep-tactin magnetic beads. During this step, the mRNA/cDNA-protein library was incubated with the beads for 1 hour at room temperature, while rotating. Bead binders remained on the beads, and the supernatant was used for complexes formation with the target antigen. Library and target were mixed in Selection bind buffer (Appendix Table 6.3) and incubated at room temperature, while rotating for 1 hour.

The next step was to bind the target antigen and bound the mRNA/cDNA-protein hybrids to Strep-tactin magnetic beads.

Extended binding times increase the risk of acquiring false-positive results from unspecific interactions, but insufficient binding times lead to a large fraction of unbound target molecules. Therefore, binding time of M^{pro}-Strep to the Strep-tactin magnetic beads was optimized. To determine the optimal binding time, 5 µg (150 pmol) of M^{pro}-Strep were incubated with 1 µl of Strep-tactin magnetic beads (20 µl slurry) for 10 s, 1 min, 5 min, 20 min, while rotating. After a washing step in 500 µl Selection bind buffer, M^{pro}-Strep was eluted by incubating the beads with 30 µl of BXT buffer (IBA Lifesciences) for 15 min.
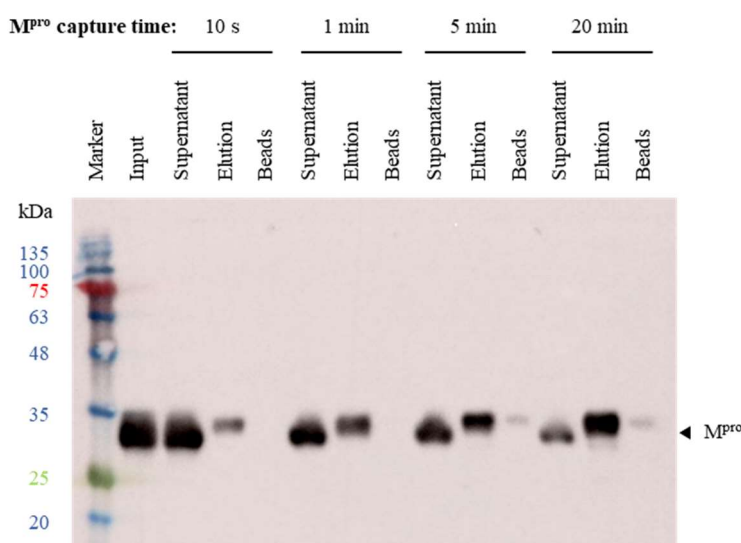
Figure 3.43 Optimization of M^{pro}-Strep binding time to Strep-tactin magnetic beads. 5 µg of M^{pro}-Strep were captured on Strep-Tactin magnetic beads for 10 s, 1 min, 5 min, and 20 min, before washing and elution. Unbound M^{pro} stays in the Supernatant fraction, successfully bound and eluted M^{pro} is found in the Elution fraction and the Beads fraction contains M^{pro} that has not been eluted from the beads. Samples were analyzed by anti-Strep-tag Western Blot.

After 10 s binding, the majority of M^{pro}-Strep remained unbound, as is shown by the strong 34 kDa species in the supernatant fraction (Figure 3.43). The ratio of bound to unbound M^{pro} (Elution to Supernatant) shifts towards the bound fraction with higher binding times. After 5 min equal amounts of M^{pro}-Strep were bound and unbound. After 20 min, the majority of M^{pro}-Strep was bound to the Strep-tactin beads. After 5 min and 20 min, the elution was not complete and a small portion of M^{pro}-Strep remains on the beads (Figure 3.43). As a compromise between optimal M^{pro}-Strep binding and reduction of unspecific interactions, 5 min was used as binding time in future experiments.

3.7 *In vitro* selection of Nbs against SARS-CoV2 main protease (M^{pro})

After the mRNA/cDNA display Nb selection protocol was optimized, Nbs were selected against SARS-CoV2-M^{pro}. Because M^{pro} possesses some favorable properties over GCAP1, such as the lack of any different ion-bound states or posttranslational modifications, and the easy assessment of M^{pro} activity by cleavage from the GST-M^{pro} fusion protein, M^{pro} was chosen as model target for the first Nb selections with the optimized mRNA/cDNA display selection procedure.

3.7.1 Selection from the y-lib library

To select Nbs from the y-lib, at first the Nb genes had to amplified from the isolated pYDS-Nb plasmids by cycle-optimized PCR.
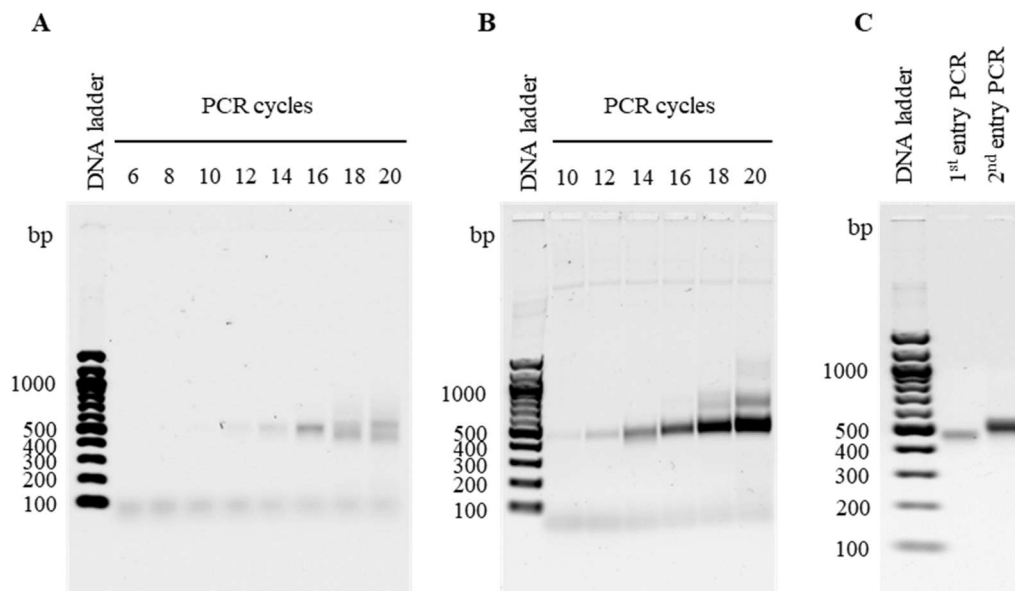
Figure 3.44 Cycle-optimization and preparative entry PCRs of the first anti-M^{pro} selection from the y-lib. A: Cycle-optimization of the first entry PCR. Nb fragments (~500 bp) appeared after 12 PCR cycles, secondary products after 18. B: In the second entry PCR, Nb fragments appeared after 10 PCR cycle, secondary products after 18. C: Products of preparative PCRs after cycle-optimization generated Nb fragments that migrated at ~480 bp (1st PCR) and ~520 bp (2nd PCR). All PCR products were analyzed by agarose gel electrophoresis.

Cycle-optimized PCRs finally amplified Nb sequences without any secondary products. The initial PCR (Figure 3.44 A) did not yield large amounts of the desired product, but after 16 PCR cycles, a pure ~480 bp fragment was generated. Secondary products started to appear after 18 cycles. Primer dimers were visible in all samples, as indicated by a ~80 bp species. The second entry PCR (Figure 3.44 B) generated larger amounts of desired product that started to appear after 10 PCR cycles. Here, byproducts started to appear after 16 PCR cycles. Preparative PCRs were conducted with 16 PCR cycles (1st PCR) and 14 PCR cycles (2nd PCR), respectively. Preparative PCR products were agarose gel purified and their quality was analyzed by agarose gel electrophoresis (Figure 3.44 C). Both preparative PCRs produced pure Nb fragments, and the 2nd PCR was shown to increase fragment size, as expected.

The product of the 2nd PCR was then used in *in vitro* transcription to produce Nb mRNA, which was then crosslinked to the Puromycin-containing oligonucleotide. Starting from the second Nb selection cycle, primers were chosen in a way that one PCR was sufficient to amplify selected Nb and add all sequence components necessary for downstream reactions. Those PCRs were sequence-optimized and agarose gel purified, as well.

After *in vitro* transcription and UV-crosslink, the XL-RNA was purified by PAGE-purification, electro elution, and subsequent purification over oligo(dT) resin. Purified XL-RNA was analyzed by urea-PAGE and subsequent anti-Puromycin Northern blot (Figure 3.45). The mentioned purification steps yielded pure XL-RNA without any low-MW impurities originating in different states of un-crosslinked Puromycin-containing oligonucleotide. Removing the large excess of Puromycin was expected to improve ribosome performance during *in vitro* translation.
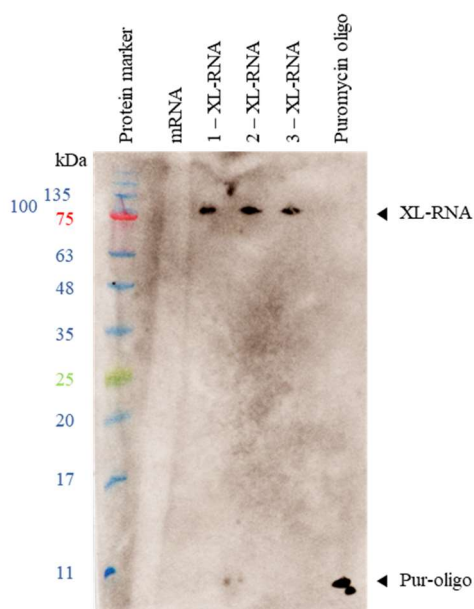
Figure 3.45 XL-RNA from all three Nb selection cycles against M^{pro}-Strep from the y-lib. All three UV-crosslink reactions during anti-M^{pro} selection from the y-lib produced similar amounts of pure Nb XL-RNA, as indicated by the 75 kDa species in the anti-puromycin Northern blot.

Unfortunately, the fluorescently labeled Nb proteins could not be detected in all three Nb selection cycles, probably due to long storage times of the Fluorotect Green Lys reagent, which was replaced after this selection. Therefore, Nb mRNA-protein hybrid molecules could not be visualized during this anti-M^{pro} selection. Because *in vitro* translation, oligo(dT)-purification, and reverse transcription procedures already have been established, the following steps were conducted without visualization.

After three selection cycles, Nb genes were retrieved by PCR. During this PCR, NcoI- and SalI-restriction sites were added to the termini of the Nb genes. 14 PCR cycles delivered the most optimal PCR product (Figure 3.46). After a preparative PCR, the product was digested with the restriction endonucleases NcoI and SalI and agarose gel-purified. Then, the digested Nb genes were ligated into a previously prepared pHen6 expression vector, which was then used to transform *E. coli* TOP10.
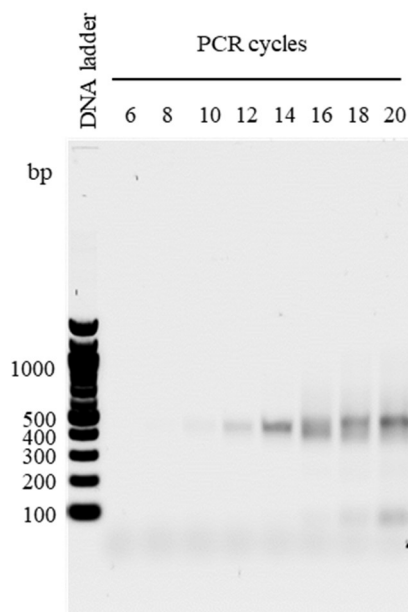
Figure 3.46 Cycle-optimization of Nb subcloning PCR. NcoiI and SalI restriction sites were added to the Nb genes by PCR. After 10 PCR cycles, NcoI-Nb-SalI fragments were detected in agarose gel electrophoresis. After 16 PCR cycles, secondary products were generated.

Single colonies of transformed *E. coli* TOP10 were used to inoculate sequencing cultures. 95 clones were Sanger sequenced directly from re-suspended colonies by the direct colony sequencing (Eurofins). Unfortunately, none of the analyzed sequences contained a full-length Nb gene. Besides some low-quality sequencing results, most sequences contained premature Stop-codons. Apparently, only truncated Nbs were selected via unspecific interactions.

3.7.2 Selection from our designed combinatorial Nb library

After the Nb selection from the y-lib yielded only unspecific binders, we concluded that this library was probably not suited for our purposes. Because of the small library size of $5 \cdot 10^9$ sequences, the likelihood of selecting actual target-binding Nbs is quite low. Additionally, using such a small library in mRNA/cDNA-display does not take advantage of the main benefit of this method, which is the ability to use large DNA libraries. Therefore, the c-lib was used for anti-M^{pro} selection with the optimized mRNA/cDNA display procedure.

Previous results revealed that the c-lib contains many Nb sequences with deletions or premature Stop-codons (chapter 3.3.1). As seen in the selections from the y-lib, these truncated sequences can be carried throughout because the selection process of unspecific interactions and generate false-positive results. To deplete unspecific binders and to further enrich for target-binding Nbs, the selection cycle number was increased to five.

As in previous selection cycles, Nb genes were amplified and prepared for downstream processes by PCR. Both PCRs were cycle-optimized (Figure 3.47 A and B) and, after the product of PCR2 was gel-purified, both products were compared by agarose gel electrophoresis (Figure 3.47 C). Both products showed single fragments, which were stretched out over the expected ~40 bp due to length-polymorphisms in the CDRs. A ~30 bp upward-shift was visible after PCR2.
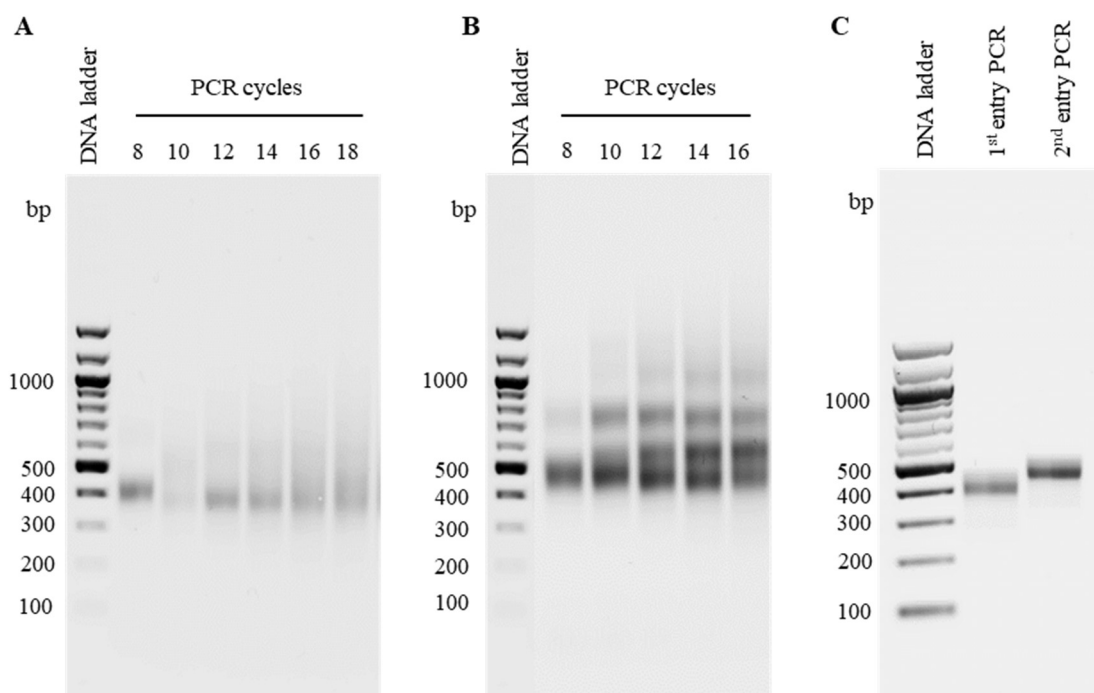
Figure 3.47 Entry PCRs into the first Nb selection cycle against Mpro from our Nb DNA library.
A: Cycle-optimization of the first entry PCR into Nb selections from the c-lib. Nb fragments (~520 bp)
appeared after 8 PCR cycles, secondary products after 12. B: In the second entry PCR, Nb fragments
appeared after 8 PCR cycle, as well, secondary products after 10. C: Products of preparative PCRs after
cycle-optimization generated Nb fragments that migrated at ~420 bp (1st PCR) and ~460 bp (2nd PCR).
All PCR products were analyzed by agarose gel electrophoresis.

After PCR2, Nb mRNA was synthesized by T7 RNA polymerase-mediated *in vitro* transcription. Nb
mRNA was then crosslinked to the Puromycin-containing oligonucleotide.

Nb mRNA was synthesized without any secondary products, as indicated by the 400-500 nt species in
urea-PAGE (Figure 3.48 A). The UV-crosslink produced several secondary products, as shown in the
anti-puromycin Northern blot (Figure 3.48 B). Nb XL-RNA was purified by gel extraction and
oligo(dT)-affinity purification.

Nb mRNA-protein hybrid molecules were synthesized by *in vitro* translation and oligo(dT)-purified.
This yielded pure Nb mRNA-protein hybrids, as the 180 kDa species in SDS-PAGE showed (Figure
3.48 C).
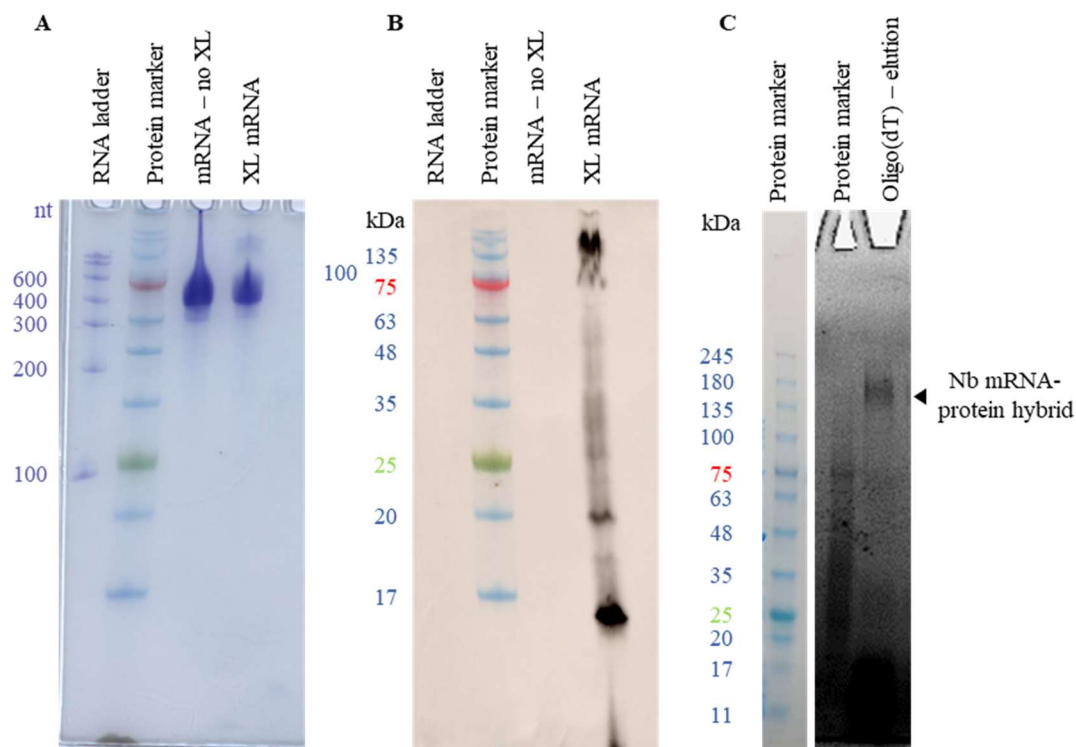
Figure 3.48 Synthesis of Nb mRNA, XL-RNA, and mRNA-protein hybrids. A: Toluidine blue stained urea PAGE and B: anti-Puromycin immuno-Northern blot of Nb mRNA and UV-crosslinked Nb mRNA compared to an RNA ladder and a pre-stained protein marker. The *in vitro* transcription generated large amounts of Nb mRNA, as indicated by the species that migrated between 400 nt and 600 nt in the toluidine blue urea-PAGE. After UV-crosslink, several puromycin-containing products were generated, as indicated in the Norhtern blot. Nb XL-RNA was purified before translation. C: Nb mRNA-protein hybrids were synthesized labeled by FluoroTect Green Lys. After oligo(dT) purification, Nb protein-fluorescence was visualized in SDS-PAGE.

After *in vitro* translation and subsequent oligo(dT)-purification, a cDNA strand was synthesized by a SuperScriptIV reverse transcriptase. Because of the absence of a matching Cy5-labeled primer for this library during the first selection cycle and because of the success in all previous reverse transcription reactions, the reaction was not visualized.

During this Nb selection, we introduced selective pressure for full-length Nbs by implementing a Ni-NTA-affinity purification of C-terminally His-tagged Nbs. After this purification step, bead binders were removed during counter selection against Strep-Tactin magnetic beads. Then, Nb-M^{pro} complexes were formed, which were then panned on Strep-tactin magnetic beads. After stringent wash steps with Selection wash buffers 1 and 2 (Appendix Table 6.3), the remaining complexes were eluted by Selection elution buffer (Appendix Table 6.3). The elution fraction was then used in a PCR to enter the second selection cycle.

The choice of suitable primers allowed amplification and modification of Nb genes by a single PCR. PCR products from all five selection cycles were compared in a single agarose gel (Figure 3.49). A high purity was maintained through all five selection cycles and there were no secondary products visible in

any of the PCR products (Figure 3.49). Because of the length polymorphisms in all CDRs, the fragments were, as expected, stretched out over roughly 50 bp. Only the PCR product of the fifth selection cycle seems to be slightly condensed.
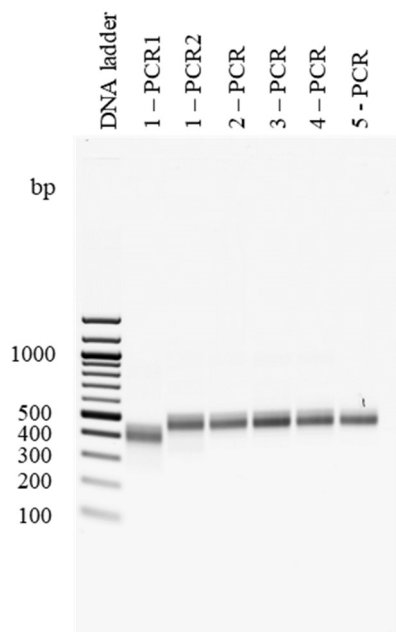
Figure 3.49 Preparative PCRs from the 5-cycle anti-M^{pro} selection from the c-lib. PCR products of five cycles of the Nb selection against SARS-CoV2-M^{pro} by the optimized mRNA/cDNA selection protocol were analyzed by agarose gel electrophoresis. PCR1 of selection cycle 1 produced ~400 bp fragments, all other PCRs ~450 bp fragments, which was confirmed by agarose gel electrophoresis.

PCR products were transcribed into Nb mRNA, which was then used for UV-crosslink. After UV-irradiation, crosslinked mRNA was purified by gel extraction (Figure 3.50) and oligo(dT)-affinity purification. This two-step XL-RNA purification procedure was applied in all five anti-M^{pro} selection cycles and yielded between 4.5 and 8 µg of high-quality and -purity XL-RNA.

Since the puromycin oligonucleotide contains the oligo(A)-sequence, the anti-Puromycin Northern blot would not deliver any additional information after oligo(dT)-purification. Therefore, this detection step was abandoned in the following selection cycles.
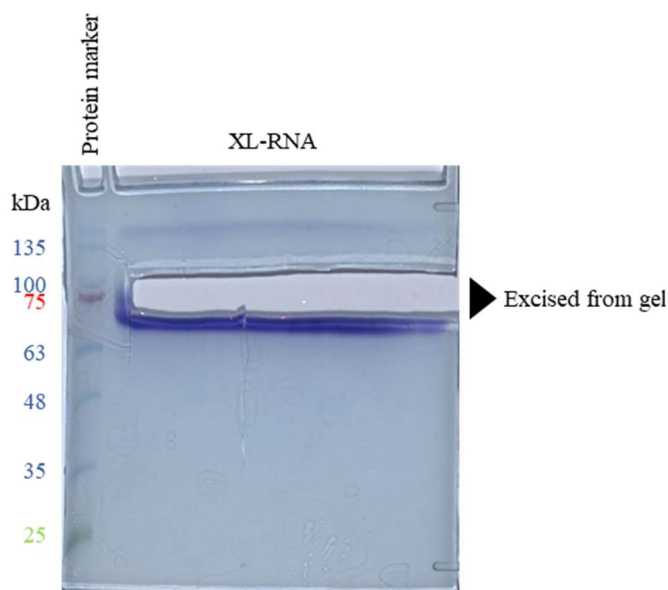
Figure 3.50 Excision of Nb XL-RNA from urea-PAGE. After UV-crosslink, the Nb XL-RNA was removed from urea-PAGE to remove the excess of puromycin-containing oligonucleotide and higher orders of XL-RNA.

Purified XL-RNA was used in *in vitro* translation with concentrations between 150 and 250 nM to generate Nb mRNA-protein hybrid molecules, which were fluorescently labeled by Fluorotect Green Lys (Promega). After oligo(dT)-affinity purification, Nb mRNA-protein hybrid molecules were visualized by in-gel fluorescence in SDS-PAGE. A distinct fragment that migrated at the expected ~180 kDa was visible in SDS-PAGE (Figure 3.51 A).

During reverse transcription, a mRNA/cDNA-protein hybrid molecules were marked by a Cy5-labeled primer. Nb mRNA/cDNA-protein hybrids were successfully synthesized, as indicated by the 180 kDa species in (Cy5-) SDS-PAGE (Figure 3.51 B). The secondary species at 75 kDa correlates with the MW of the SuperScriptIV reverse transcriptase. Because of insufficient denaturation, bound primer molecules were visualized. After reverse transcription, Nb mRNA/cDNA-protein hybrids were purified over their C-terminal His-tag by Ni-NTA affinity. This step was implemented to enrich full-length Nbs. The majority of hybrid molecules did not carry a C-terminal His-tag, as suggested by the equal signal intensities of the reverse transcription product and the unbound fraction of the Ni-NTA purification (Ni-NTA – Supernatant) (Figure 3.51 B).
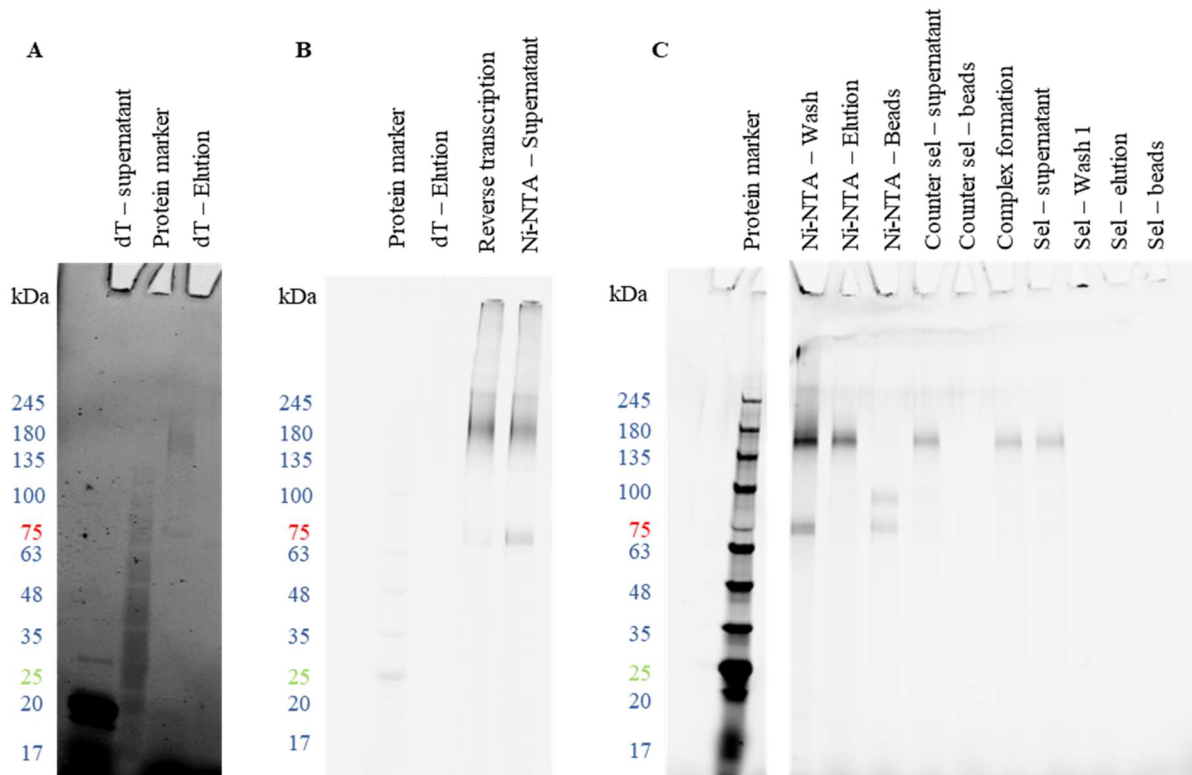
Figure 3.51 Synthesis of Nb mRNA-protein and mRNA/cDNA-protein hybrid molecules. A: Nb mRNA-protein hybrid molecules were synthesized and labeled proteins were visualized after oligo(dT) purification. B: The in vitro reverse transcription generated mRNA/cDNA-protein hybrids. Majority of the hybrids did not bind to the Ni-NTA magnetic beads, as indicated by the retained 180 kDa species in the Ni-NTA – Supernatant fraction. C: The Nb mRNA/cDNA-protein library was partially purified over Ni-NTA magnetic beads, as indicated by the 180 kDa species in the Ni-NTA – Elution fraction. Purified Nbs contain an in-frame His-tag. Detectable amounts of the Nb mRNA/cDNA-protein hybrids were maintained retained through counter selection and complex formation with Mpro-Strep. Nb cDNA was labeled with Cy5 over a labeled primer. Nb mRNA/cDNA-protein hybrid molecules were visualized by Cy5-fluorescence in SDS-PAGE (B and C).

A small fraction of Nb hybrid molecules contained a C-terminal His-tag and were purified, which could be visualized by enhancing the contrast in (Cy5) SDS-PAGE (Ni-NTA – Elution) (Figure 3.51 C).

The Ni-NTA elution fraction was then used for counter selection. Counter selection supernatant for complex formation and panning. During counter selection, only a small fraction did bind to the Strep-tactin magnetic beads, indicated by the faint 180 kDa signal in the 'Counter selection – beads' fraction (Figure 3.51 C). The majority of Nbs did not form complexes M^{pro}, as indicated by the equal intensity of the 180 kDa species in the 'Complex' and 'Sel – Supernatant' fractions (Figure 3.51 C). The 'Wash', 'elution', and 'beads' fractions of the selection did contain no substantial Nb amounts. Nevertheless, Nb genes could be amplified by PCR after each selection cycle, as indicated by the 400-500 bp in agarose gel images of the PCRs (Figure 3.49; Figure 3.52).
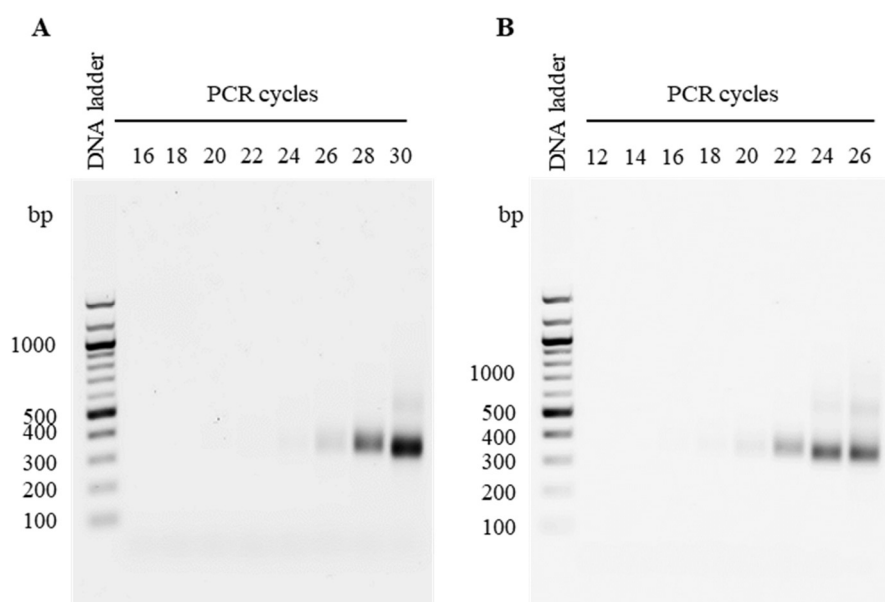
Figure 3.52 PCR-optimizations of subcloning PCRs. For subcloning, NcoI and SalI restriction sites were added to the selected Nb genes. Nb genes were subcloned after 3 and 5 selection cycles. A: Subcloning PCR after 3 selection cycles produced detectable amounts of NcoI-Nb-SalI fragments after 24 PCR cycles, and secondary products after 30. B: Subcloning PCR after 5 selection cycles generated detectable amounts of the desired fragment after 20 PCR cycles, secondary products after 24. PCR products were analyzed by agarose gel electrophoresis.

After three Nb selection cycles against SARS-CoV2-M^{pro}, NcoI and SalI restriction sites were adding by PCR, which produced detectable amounts of desired ~400 bp product after 24 PCR cycles (Figure 3.52 A). After 30 cycles, a secondary product at ~600 bp started to appear. The optimal PCR cycle number that produced a maximum amount of primary product without generating any secondary product was 28 PCR cycles.

In a subcloning PCR after five Nb selection cycles, the desired product started to appear after 20 PCR cycles (Figure 3.52 B). Secondary products were generated after 24 PCR cycles, meaning that the optimal cycle number for this PCR was 22. Earlier appearance of the PCR product after five selection cycles than after three suggests an enrichment of the DNA library. However, for a more precise determination of the enrichment, quantitative PCRs would be required.

After cycle-optimization, preparative PCRs with the optimal cycle number were performed. The desired product was digested by the restriction endonucleases NcoI and SalI, and the digested fragments were purified by agarose gel extraction. The digested Nb genes were ligated into a previously prepared pHen6-Avi-His vector. After transforming *E. coli* TOP10 with the generated plasmids, 95 colonies were used to inoculate sequencing colonies, which were sequenced, as mentioned.

After three selection cycles, the majority of the 95 clones contained premature Stop-codons. Only seven sequences were generated that were in frame and without Stop-Codon from the N-terminal PelB sequence to the C-terminal Avi-His tag. Of those seven Nb genes, five contained the characteristic

CEER-motif and deletion of CDR3 (Figure 3.53). Only Nbs aM^pro-010 and -023 contained a proper CDR3. Apparently, aM^pro-010 contained a deletion in CDR1 and an L20E mutation. Nb aM^pro-023 entirely agreed with the theoretical library design with CDR lengths of four amino acids (CDR1), six amino acids (CDR2) and five amino acids (CDR3) (Figure 3.53, bottom sequence).
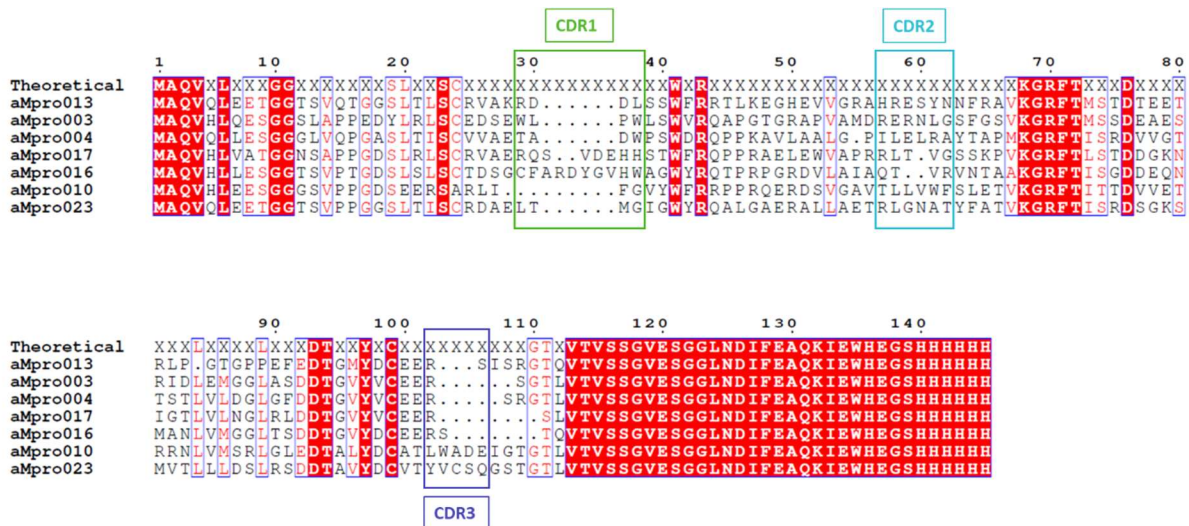
Figure 3.53 Multiple protein sequence alignment of mRNA/cDNA-display-selected Nb genes after three selection cycles compared to the theoretical Nb design. The depicted Nb genes were selected after three selections cycles after SARS-CoV2-M^pro. Multiple sequence alignment was performed by Clustal Omega multiple sequence alignment tool (Madeira et al., 2022) and visualized using the Esprit server (Robert & Gouet, 2014).

Nbs aM^pro-010 and -023 were expressed in *E. coli* wk6. Both Nbs were expressed, but were not translocated to the periplasm, as indicated by the 15 – 25 kDa species in the total lysate fraction and the absence in the Periplams fraction (Figure 3.54). This would have been essential for properly folded and disulfide bond containing Nbs. This means that no M^pro-binding Nbs could have been selected after three mRNA/cDNA selection cycles. The Nbs were probably not entirely processed and still contained the PelB translocation signal peptide, as indicated by the higher-MW species at 20 – 25 kDa (Figure 3.54).
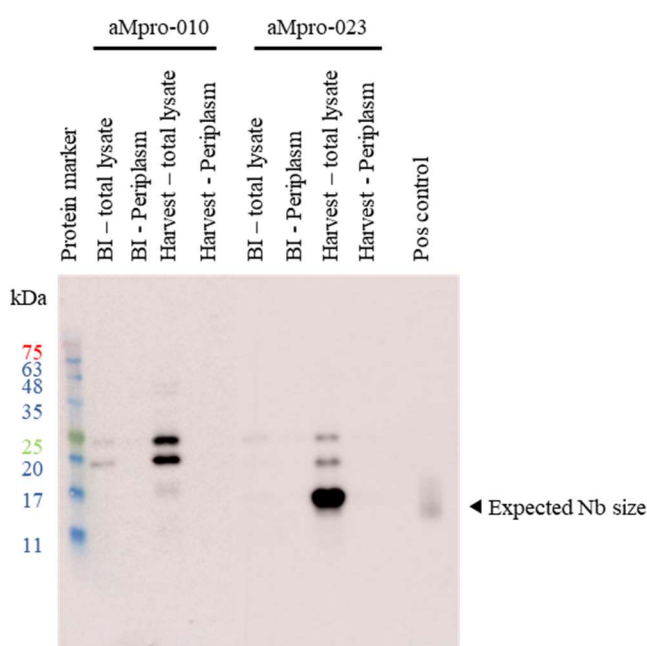
Figure 3.54 Anti-M^pro Nb-010 and -023 test expressions. His-tagged Nbs aM^pro-010 and -023 expression was analyzed by anti-His Western blot. Samples were taken before IPTG-mediated induction of protein expression (BI) and at cell harvest after protein expression. Nb expression was detected (Harvest – total lysate), but no Nbs were translocated to the periplasm (Harvest - Periplasm).

95 subcloned Nb genes were Sanger sequenced after five selection cycles, as well. Only 70 sequencing results showed sufficient quality. Of those, 27 sequences had premature Stop-Codons. 33 sequences were Nb genes with CDR3-deletions and with a characteristic CEER-motif. Eight full-length, in-frame, Nb genes were obtained (Figure 3.55).
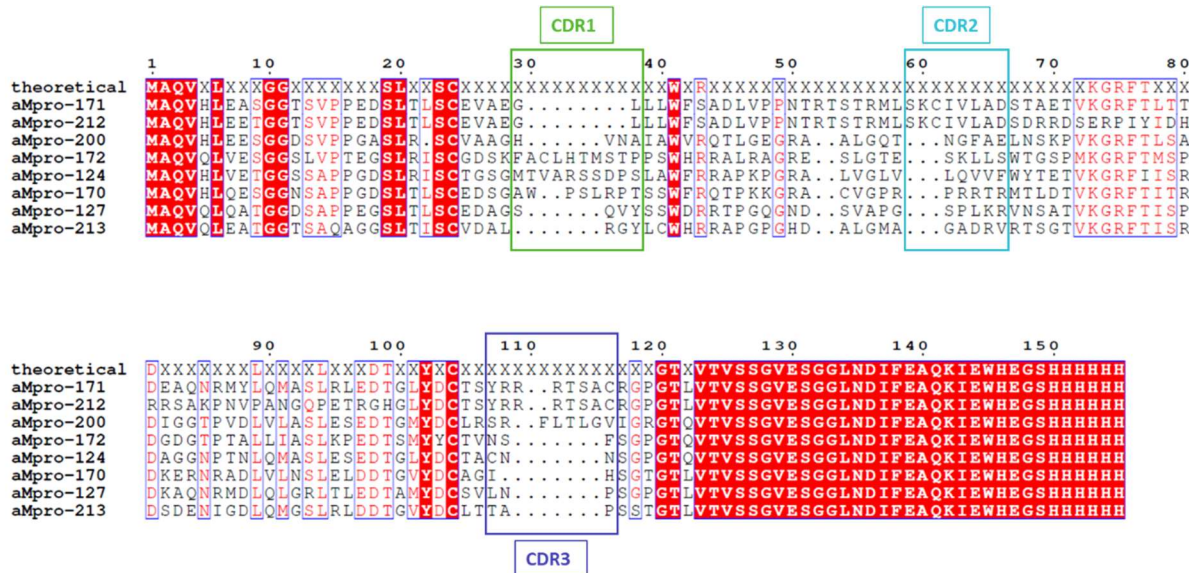
Figure 3.55 Multiple protein sequence alignment of selected Nb genes. The depicted Nb sequences were selected after five mRNA/cDNA-display selection cycles against SARS-CoV2-Mpro with CDR1 (green), CDR2 (cyan), and CDR3 (blue) highlighted. Multiple sequence alignment was performed by Clustal Omega multiple sequence alignment tool (Madeira et al., 2022) and visualized using the Esprit server (Robert & Gouet, 2014).

Noticeably, Nbs aM^pro-171 and 212 showed completely identical CDRs. However, Nb 212 deviated from the library design between positions 69 and 99. This was probably caused by two frameshifts at

these positions. Most of the Nbs showed very short CDR3, with Nbs 124, 127, 170, 172, and 213 having CDRs even shorter than the minimal four amino acids from the library design. Interestingly, Nbs 172, 124, and 170 showed longer CDR1 in combination with very short CDR3 (Figure 3.55).

The eight selected Nbs were attempted to by synthesized in *E. coli* wk6. All Nbs showed some expression, but none were successfully processed and transported into the periplasm. For most Nbs, a ladder pattern was detected with the most dominant signal not matching the expected Nb size of ~15 kDa (Figure 3.56; Figure 3.57). Again, the Nbs were probably not properly unprocessed, as indicated by the different species between ~18 and 25 kDa.
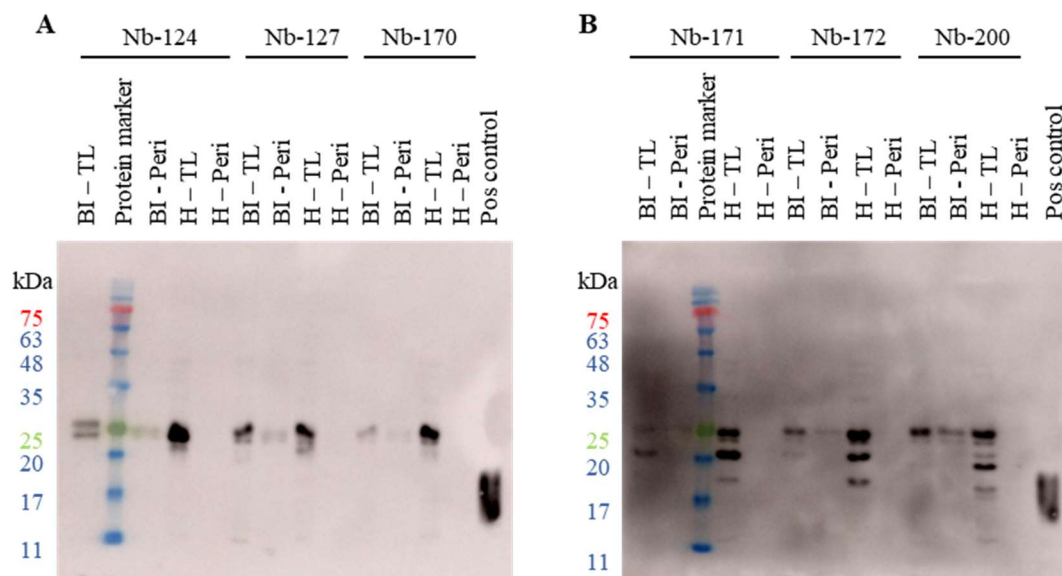
Figure 3.56 Test expressions of selected aM^{pro}-Nbs. Eight Nbs were selected against M^{pro}. These Nbs were expressed in *E. coli* wk6. Samples were again taken before IPTG-mediated induction of protein expression (BI) and at cell harvest after protein expression. Total cell lysates and periplasmic fractions were analyzed by anti-His Western blot. None of the Nbs were expressed and translocated to the periplasm (H – Peri fractions).
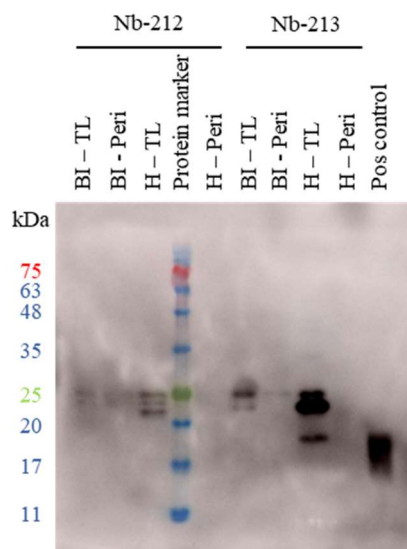
Figure 3.57 Test expressions of anti-M^{pro} Nbs 212 and 213. Selected Nbs were expressed in *E. coli* wk6. Expression levels and translocation to the periplasm were analyzed by anti-His-tag Western blot. Both Nbs were expressed but were not translocated to the periplasm, which is indicated by the absence of a detectable 15 – 20 kDa species in the periplasmic fraction after cell harvest (H – Peri).

After *E. coli* wk6 was not able to synthesize any well-behaved anti-M^{pro} Nbs, we tried to express anti-M^{pro}-200 in different *E. coli* expression strains. Because the synthetic Nb genes from the DNA library were not codon-optimized for *E. coli*, BL21(DE3) Codon Plus RIL cells were used, which carry additional tRNAs for the amino acids arginine (R), isoleucine (I), and leucine (L) and remove codon bias. To analyze if case RNA stability was a limiting factor, BL21(DE3) Star were used, which have an RNase E deletion and show increased RNA stability. *E. coli* Origami2 carry mutations in the glutathione reductase and thioredoxin reductase, which enables cytoplasmic disulfide bond formation. *E. coli* Arctic Express were designed to express proteins at low temperatures of 10 – 13 °C and contain the cold-activated chaperonins Cpn10 and Cpn60 from the psychrophilic bacterium *Oleispira antarctica*.
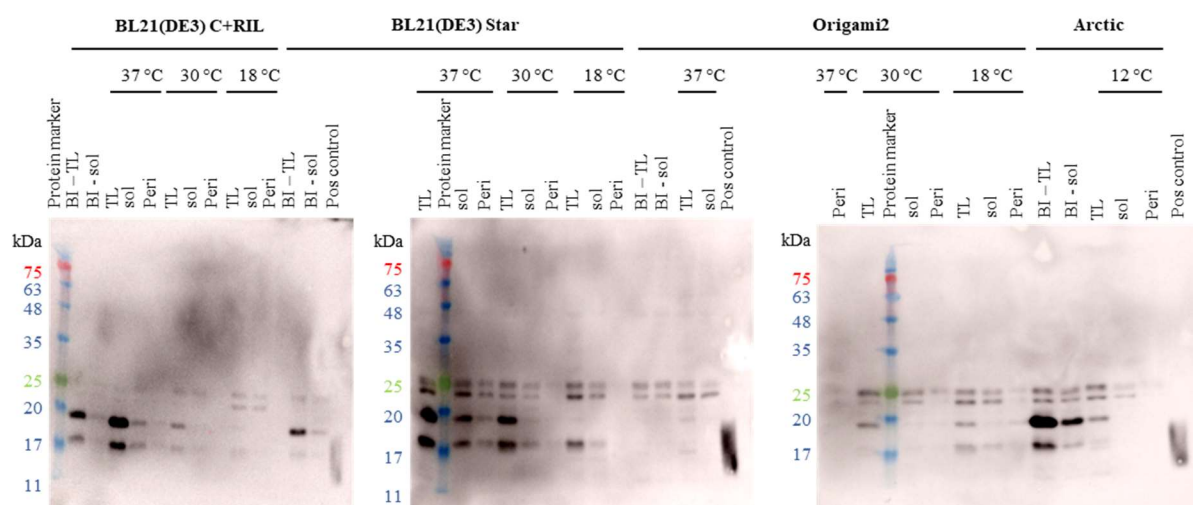
Figure 3.58 Nb test expressions in different *E. coli* strains. Anti-M^{pro} Nb 200 was expressed in *E. coli* BL21(DE3) Codon Plus RIL (C+RIL), BL21(DE3) Star, Origami2, and Arctic Express. Samples taken before protein expression was induced (BI) and directly at cell harvest after protein expression (all other samples) were analyzed. Cells were lysed by sonication and the total and cleared (soluble) lysate fractions were analyzed. Additionally, the periplasmic fraction was used. Expression levels and periplasmic translocation was assessed by anti-His Western blot. None of the cells were able to synthesize well-behaved Nbs.
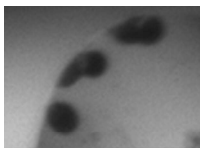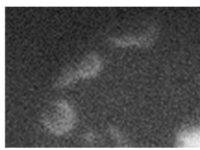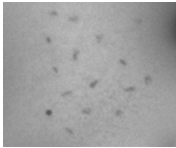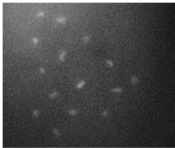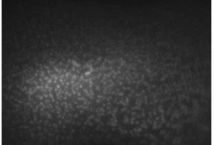
Unfortunately, none of the cells were able to synthesize well-behaved M^{pro} Nb-200 with the necessary quality or quantity (Figure 3.58). BL21(DE3) Codon Plus RIL, Star, and Arctic express show large amounts of expression before induction, which should be avoided for the sake of protein quality. All samples show either no soluble Nb-200, at all, or a ladder pattern, similar to expression in *E. coli* wk6. So, none of the used cells were able to synthesize substantial amounts of aM^{pro} Nb.

Unfortunately, the selected anti-M^{pro} Nbs could not be expressed as well-behaved proteins with substantial amounts in different *E. coli* host strains.

3.8 Crystallization of myrGCAP1

Commercially available crystallization screens (Jena Bioscience; Hampton Research) were used to screen for initial myrGCAP1 crystals. Crystals were grown by sitting drop, vapor diffusion at 18 °C with protein concentrations of 5 – 10 mg/ml in droplet sizes of 500 nl (250 nl protein solution + 250 nl precipitant solution). Initial crystals with varying size and regularity appeared after two to three weeks.
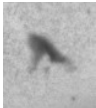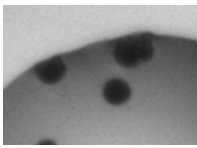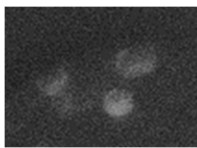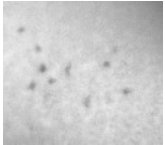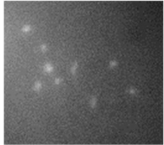
Table 3.2 Initial crystals of myrGCAP1 wild type (GCAP1-apo, GCAP1-Ca, and GCAP1-Mg).

Condition no. (GCAP1 form)	Condition composition	Crystal image	UV image	Shape
1 (Ca^{2+})	2 % isopropanol; 100 mM MES pH 6.5; 200 mM ZnOAc			Disk-like
2 (apo)	25 % PEG 3350; 200 mM AmSO$_4$; 100 mM BIS-TRIS pH 5.5			Three intergrown crystals
3 (Mg^{2+})	25 % PEG 3350; 200 mM AmSO$_4$; 100 mM HEPES pH 7.5			Small; irregular
4 (Ca^{2+})	100 mM BIS-TRIS pH 5.5; 1 % PEG 3350; 1 M AmSO$_4$			Uniform; Cubic; very small

Crystals grew for all three myrGCAP1 wild type variants. GCAP1-Ca produced some disc-shaped crystals, which appeared to be more gelatinous than crystalline (Table 3.2, condition #1). Additionally, myrGCAP1-Ca produced many tiny, cube-shaped crystals that were too small for further use (Table 3.2, condition #4). The apo protein formed one bigger crystal that was comprised of three intergrown crystals (Table 3.2, condition #2). MyrGCAP1-Mg produced several small, non-uniform crystals (Table 3.2, condition #3). All crystals appeared to be protein crystals, as indicated by UV-fluorescence. Crystals from conditions 1 – 3 were frozen and analyzed for X-ray diffraction. Because crystals from condition #4 were too small for further use, optimization screens were set up around the mentioned condition.
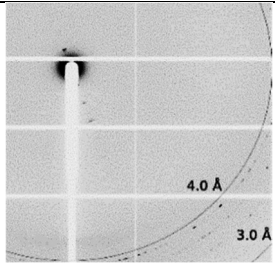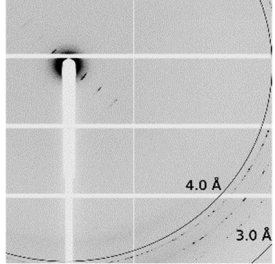
Table 3.3 Initial crystals of myrGCAP1 M26R mutant (GCAP1-apo, GCAP1-Ca, and GCAP1-Mg).

Condition no. (GCAP1 form)	Condition composition	Crystal image	UV image	Shape
5 (Ca^{2+})	2 M AmSO$_4$; 100 mM BIS-TRIS pH 5.5			Irregular
6 (apo)	2% isopropanol 100 mM MES pH 6.5 200 mM Zinc acetate			Disk-like
7 (Mg^{2+})	25 % PEG 3350; 200 mM AmSO$_4$; 100 mM HEPES pH 7.5			Small; irregular

The myrGCAP1 M26R mutant produced protein crystals of irregular shapes in conditions #5 (myrGCAP1-Ca) and #7 (myrGCAP1-Mg). The apo protein generated some disk-like protein crystals in condition #6 (Table 3.3). Those crystals were frozen and analyzed for X-ray diffraction, as well.

Table 3.4 Diffraction patterns obtained from a myrGCAP1 wild type (apo) and M26R (GCAP1-Ca) crystals.

Protein	Condition no. and composition	Cryo-protectant	Diffraction image
Wild type apo	#2: 25 % PEG 3350; 200 mM AmSO$_4$; 100 mM BIS-TRIS pH 5.5	20 % glycerol	
M26R Ca^{2+}	#6: 2% isopropanol 100 mM MES pH 6.5 200 mM Zinc acetate	40 % ethylene glycol	

Only two analyzed crystals showed X-ray diffraction, wild type myrGCAP1-apo crystal obtained in crystallization condition #2 and a M26R myrGCAP1-Ca crystal formed in crystallization condition #6 (Table 3.4). However, both crystals diffracted only to insufficient resolutions of ~10 Å, and the crystals require further optimization. Hopefully, well-diffracting crystals can be obtained from the set-up optimization screens, which will be analyzed in the future.

4. Discussion

The main goals of this study were the production of the retinal proteins retGC1 and GCAP1, as well as the SARS-CoV2 main protease (M^{pro}), and to establish an mRNA/cDNA display-based Nb selection protocol to obtain high-affinity binders against these protein targets.

RetGC1 was expressed in *sf21* and a stable TREx293 cell line, solubilized using the DIBMA Sokalan CP9. Purification from the TREx293 membranes by Strep-tactin-affinity chromatography and SEC yielded small amounts of pure retGC1. Wild type myrGCAP1 and the disease-associated M26R mutant were purified with large quantities of the apo, Ca^{2+}-bound, and Mg^{2+}-bound forms. For structural analysis, initial crystals were grown that showed weak diffraction up to roughly 10 Å. SARS-CoV2-M^{pro} could be purified in large quantities. Purified M^{pro} showed catalytic activity as it cleaved itself off an N-terminal GST-tag.

A Nanobody selection procedure by mRNA/cDNA display was developed that allows completion of a single selection cycle in approximately four days. Within each selection cycle, several reaction and purification steps were introduced that improve Nb library purity before selection, namely PCR cycle-optimization, purification of XL-RNA by gel extraction and oligo(dT)-purification and purification of full-length Nbs over a C-terminal His-tag.

4.1 Production of target antigens

4.1.1   Retinal guanylate cyclase 1

RetGC1 is a membrane-embedded protein that natively resides in disc membranes of ROS (Toibana et al., 1982), and is associated to detergent-resistant membrane environments (Nair et al., 2002), so called membrane rafts, which makes it difficult to solubilize and purify.

For biochemical characterization and generation of initial structural insights, retGC1 has been studied from crude cell lysates or membrane preparations (Peshenko & Dizhoor, 2004; Rehkamp et al., 2021). Because Nanobody selection and structural analysis by cryo-EM require purified proteins, we attempted to express retGC1 in different host organisms and purify it by affinity chromatography and SEC. RetGC1 was transiently expressed in *sf21* insect cells and in a stable TREx293 cell line, which was allows inducible retGC1 expression. After solubilization of retGC1 in non-protein-denaturing detergents was unsuccessful, we attempted to solubilize retGC1 using the DIBMA Sokalan CP9, which generated increasing retGC1 solubilization rates at increased Sokalan CP9 concentrations (1 – 5 %). In contrast to solubilization with detergents, which strip off and replace the surrounding lipids, DIBMAs extract membrane proteins together with their native membrane environment in DIBMA lipid particles, or DIBMALPs (Oluwole et al., 2017), similar to SMAs (Jamshad et al., 2011). Apparently, solubilizing

retGC1 alongside its membrane environment was beneficial for retGC1 solubility and long-term stability.

After successful solubilization, retGC1 was purified by Strep-tactin-affinity chromatography and subsequent SEC. RetGC1 was purified only as aggregated proteins from *sf21*, as the early elution in the SEC suggested. Even though, *sf21* cells are generally a useful host organism for the expression of membrane proteins (Damhof et al., 1994; Salem et al., 2018), they apparently do not provide the right membrane environment for retGC1. While mammalian membrane rafts are enriched in cholesterol (Elliott et al., 2003; Martin et al., 2005), insect cell membranes usually have lower cholesterol concentrations (Marheineke et al., 1998), which might be detrimental for proper retGC1 folding.

The mammalian TREx293 cells provided a more suitable membrane environment for retGC1 expression, and we were able to solubilize and purify retGC1 from TREx293 membranes. Affinity chromatography and SEC of TREx293-expressed retGC1 yielded small amounts of pure retGC1, but protein concentration was very low and could not be determined by spectrophotometry. For a single Nb selection with our established mRNA/cDNA display selection procedure, 150 nmol of target antigen would be required, which equals 18.6 µg of retGC1.

This means that retGC1 amounts must be increased, either by improving of the solubilization, or by increasing the culture volume. Solubilization of retGC1 could be improved by using different SMA or DIBMA derivates (Overduin et al., 2021; Swainsbury et al., 2017). A scale-up of the adherent TREx293 cell culture would be challenging, as incubator capacities might be limiting. Alternatively, a TREx293 suspension culture (Xiao et al., 2013) might be more suitable to produce substantial amounts of retGC1.

4.1.2 Guanylate cyclase activating protein 1

In contrast to the membrane-embedded retGC1, GCAP1 is a relatively small (24 kDa), soluble and well-behaved protein. Its posttranslational modification, N-myristoylation, can be introduced on the GCAP1 D6S mutant by co-expression of a yeast N-myristoyltransferase in the bacterial host *E. coli* (Hwang & Koch, 2002). Therefore, we were able to produce myristoylated GCAP1 wild type and the mutant M26R mutant in the apo, Mg^{2+}, and Ca^{2+}-bound form for Nb selection and for crystallization experiments.

We were able to grow first crystals of myrGCAP1 wild type and the disease-associated M26R variant. Crystals of the apo wild type protein and the Ca^{2+}-bound M26R variant showed some weak diffraction, which indicates that the crystals were intergrown and must be further optimized. Optimization experiments were already set up with the goal to yield single-lattice, well-diffracting crystals suitable for structure determination of myrGCAP1 wild type and M26R. Determining the protein structures of the different ion-bound versions of myrGCAP wild type and M26R would help in gaining a better understanding of retGC1 regulation by GCAP1 and the mechanism that influence rare retinopathies.

4.1.3 SARS-CoV2 M^{pro}

During the pandemic, research on SARS-CoV2 became a main focus of the Charité – Universitätsmedizin Berlin. We joined the pandemic related research by attempting to select Nbs against SARS-CoV2 M^{pro}. As Nb target, M^{pro} provides some advantages compared to GCAP1 and was chosen as a model target to establish and optimize our mRNA/cDNA display Nb selection procedure. M^{pro} is a small, soluble protein, which is, in contrast to GCAP1, not present in any different ion-bound forms. Especially, GCAP1's high affinity for Ca^{2+} (Dell'Orco et al., 2010) makes it difficult to maintain GCAP1-Mg or -apo, since very small Calcium amounts suffice to generate GCAP1-Ca, as seen in SDS-PAGE of purified GCAP1-Mg or -apo (Sokal et al., 2005); Figure 3.11). Additionally, proteolytic M^{pro}-activity can be easily assessed by split fluorescent protein assays, where proteolytic cleavage enables reconstitution and fluorescence of a split fluorescent protein (Callahan et al., 2010), or qualitatively by cleavage of M^{pro} from a polyprotein with an easy readout in SDS-PAGE. To analyze GCAP1-activity, retGC1 cyclase assays would be necessary (Wilkie et al., 2001), which would be more laborious and difficult.

We successfully established M^{pro} expression in *E. coli* BL21(DE3) and purification by Strep-tactin-affinity chromatography and SEC in mg quantities. The used M^{pro} construct contained an N-terminal GST-tag and a C-terminal Strep-tag, and an M^{pro}-specific cleavage site between the GST-tag and M^{pro}. Proteolytic activity of M^{pro} was indicated by the cleavage of M^{pro} from the GST-M^{pro} polyprotein (Figure 3.15; (Lee et al., 2020), similar to M^{pro} cleavage from a SARS-CoV2 polyprotein (Cannalire et al., 2022).

4.2 Enabling the use of the y-lib in mRNA/cDNA display

In addition to the highly variable c-lib, which was specially designed for this study, we used the commercially available y-lib (McMahon et al., 2018). The y-lib uses a fixed Nb scaffold with variable positions only in the CDRs and has a library size of $5 \cdot 10^8$ Nb DNA sequences. Before the y-lib could be used in mRNA/cDNA display, the Nanobody containing pYDS plasmids needed to be isolated from transformed yeast cells.

Published protocols for plasmid isolation from yeast have maximum yields of roughly 2.75 µg per liter of culture but usually show insufficient qualities and require plasmid amplification in a bacterial host after isolation from yeast (Holm et al., 1986; Singh & Anthony Weil, 2002). To isolate a single library size, $5 \cdot 10^8$ of pYDS plasmid, 2.2 ng would be sufficient. Because, by chance, multiple copies of some plasmids might be isolated, some others would be lost. In order to represent the entire gene library, higher amounts of plasmid DNA should be isolated. Published protocols would allow the plasmid isolation of roughly 1000 library sizes, but because of the low quality in some protocols, we wanted to improve the plasmid isolation procedure from yeast cells.

The plasmid isolation procedure established in this work included the separation of spheroplasts from the zymolyase digested cell wall prior to alkaline lysis and spin-column purification. This step increased the yield of isolated plasmid DNA from 2.75 μg per liter of culture (Singh et. al., 2002) to 45 μg per liter of culture. The quality of the isolated pYDS-Nb plasmid was analyzed by agarose gel electrophoresis and specific PCR-amplification of Nb genes, which yielded a single fragment that matched the size of Nb genes. The enhanced quality of the isolated plasmid along with the high yield made the y-lib accessible for the use in mRNA/cDNA display.

4.3 Development of a fast Nb selection protocol by mRNA/cDNA display

New Nbs are traditionally selected by immunization of camelids like llamas or alpacas with the target antigens. In order to avoid the immunization step, several animal-free selection methods have been developed like phage, yeast, or mRNA/cDNA display. These animal-free methods enable Nb selections even against toxic target antigens or targets with low-immunogenicity. In contrast to the cell-mediated methods of phage and yeast display, the *in vitro* selection method mRNA/cDNA display provides the ability to select Nbs from large Nb gene libraries, which was essential for Nb selections from the highly variable Nb gene library that was designed for this study. Therefore, we opted for mRNA/cDNA display as our Nb selection method of choice.

We developed an mRNA/cDNA display Nb selection protocol that enables Nb selection from large, highly variable Nanobody libraries. The maximum library size that can be used in our mRNA/cDNA display procedure is 10^{13} DNA sequences, which is the limit for many published mRNA display protocols, as well (Anzai et al., 2019; Norman et al., 2021; Xu et al., 2002).

Our protocol enables the execution of one selection cycle in roughly four days (Table 4.1). Faster protocols have been published, where times per selection cycle could be reduced to two days (Mochizuki et al., 2011), or even to a few hours (Ishizawa et al., 2013). However, these protocols usually save time by combining multiple reactions in single steps and thus omit important purification steps. To maintain integrity of the Nb gene library and to avoid the formation of secondary products after amplification steps, the established purification steps have been essential. So, our developed protocol provides a good compromise between high library integrity and a fast selection procedure.

In our established mRNA/cDNA display selection procedure, three mRNA/cDNA display selection cycles can be conducted in less than three weeks, The selection cycles are then followed by subcloning of the genes of selected binders into expression vectors for heterologous expression in *E. coli* and Sanger sequencing, which might take another seven to ten days. Accordingly, target-binding Nbs could be selected within one month, which is much faster than immunization of llamas or alpacas, where serum can usually be harvested after roughly two months (Maass et al., 2007).

Table 4.1 Overview of single steps of one Nanobody selection cycle by mRNA/cDNA display.

Day	Reaction step	Approx. Duration per step
Day 1	Optimization PCRs	3 h
	Preparative PCRs + purification	4 h
Day 2	*In vitro* transcription + purification	5 h
	UV-crosslink	2 h
Day 3	Urea-PAGE and excision of XL-RNA	3 h
	Electro elution from gel	1.5 h
	Oligo(dT) purification	1 h
Day 4	*In vitro* translation	2.5 h
	Oligo(dT)-affinity purification	1 h
	Ni-NTA-affinity purification	1.5 h
	Counter selection	1 h
	Complex formation in solution	1 h
	Pull-down + washing + elution	1 h

One critical step in our selection procedure is the two-step purification of the XL-RNA by gel extraction and oligo(dT)-affinity purification, which enables isolation by size and presence of the puromycin-containing oligonucleotide. While other mRNA/cDNA display selection procedures only purify the modified mRNA by spin-column purification (Doshi et al., 2014), or simply precipitate the XL-RNA without removing the excess of puromycin or any unwanted species (Seelig, 2011), our protocol ensures purity of the Nb gene library throughout the selection process.

4.3.1 Suitability of the y-lib for mRNA/cDNA display based Nb selections

The y-lib is an inexpensive, easily accessible initial library for Nanobody selections and served as a valuable tool to establish and optimize most of the steps of our newly established selection protocol. Therefore, we isolated the y-lib from yeast cells and then used it for Nanobody selections against the targets myrGCAP1 and SARS-CoV2 M^{pro}.

Several Nb sequences were retrieved after three rounds of selection against myrGCAP1-Mg. Two full-length, in-frame Nbs were obtained, overexpressed, purified, and probed for their affinity for myrGCAP1-Mg. Additionally, nine Nb genes showed frameshifts in the FR1 region, but had intact, in-frame, downstream sequences from CDR1 to the C-terminus. Because of the known frameshifting ability of *E. coli* ribosomes (Flower & McHenry, 1990; Larsen et al., 1994), we suspected that the *E. coli* ribosomes in the PURE system might have generated full-length, in-frame, Nb proteins from the frameshift containing Nb mRNAs. Before heterologous overexpression, the Nb genes were 'rescued' by Gibson Assembly, meaning the Nb sequence downstream of the frameshift was cloned into an expression vector, in frame with a Nb FR1.

The rescued Nbs were overexpressed, purified, and analyzed for affinity for myrGCAP1-Mg, as well. Neither the selected nor the rescued Nbs showed substantial affinity for myrGCAP1-Mg. Apparently, the synthesis of full-length (in-frame) Nb proteins by ribosomal frame shift is unlikely, and we suspect that these truncated Nb sequences were falsely selected based on unspecific interactions (Wang & Liu, 2011).

In Nb selections from the y-lib against M^pro, only Nanobody sequences were selected that contained either frameshifts or premature Stop-codons. One reason for the lack of high-affinity Nbs in selections from the y-lib against myrGCAP1 and M^pro might be the absence of specific binders against the target proteins in the rather small initial library ($5 \cdot 10^8$ Nb sequences), as larger libraries greatly increase the likelihood of selecting high-affinity binders (Kamalinia et al., 2021). Additionally, using such a small gene library ($5 \cdot 10^8$ Nb sequences) in mRNA/cDNA display does not take advantage of the main benefit of *in vitro* selections, which is the employment of large gene libraries (Newton et al., 2020).

Another reason might be the quality of the library and its use outside of yeast display. The y-lib is optimized to be used in yeast display (McMahon et al., 2018) and might require protein quality control mechanisms acting in the yeast cells after protein synthesis and before translocation to the cell wall (Boder et al., 2012). To enrich for properly folding Nbs in the first place, even before plasmid isolation, the y-lib should first be amplified and pre-selected for displayed Nbs. This step would deplete frameshifted and prematurely stopped Nb sequences and enrich well-behaved Nbs before the first selection cycle.

Nevertheless, the y-lib has been a useful tool to establish most of the single steps of the selection procedure. To increase the chances of selecting binders from this Nanobody library, the yeast display selection method should be employed. Alternatively, a preselection for well-behaved, full-length, Nbs in combination with Nb selections by mRNA/cDNA display would greatly increase the likelihood of selecting high-affinity binders from the y-lib.

4.3.2 Nanobody selection using the highly variable c-lib

In nature, new Ab sequences are generated in B lymphocytes by recombination of several exons (Brack et al., 1978). Our design of the synthetic c-lib was based on multiple sequence alignments of published Nbs, which revealed high sequence variability and length polymorphisms for the CDRs and certain variable positions in the more conserved FRs (Figure 1.10). The synthesis of this highly variable gene library by GeneArt (Thermo Fisher) proved to be quite challenging. After personal communication with GeneArt, we implemented an invariant sequence in FR2 in the form of a 'KGRFT'-motif to facilitate library construction. GeneArt's quality control revealed a viability of 62 % for the c-lib (Appendix Figure 6.9), meaning that 30 of 48 analyzed DNA sequences agreed with the theoretical library design. However, this DNA-level analysis did not account for any premature Stop-codons. Therefore, we subcloned the c-lib into a vector, Sanger sequenced 58 Nb genes, and translated the DNA into amino

acid sequence, *in silico* (Figure 3.18). This revealed that only 29 % of the analyzed sequences corresponded to full-length Nb proteins, and 47 % of the sequences contained premature Stop-codons (Table 3.1). To gain a deeper insight into the composition and quality of the c-lib, a very large number of sequences would have to be analyzed, which would require next-gen sequencing methods like nanopore (Lood et al., 2020) or illumina sequencing (Choi et al., 2023).

In order to enrich the full-length Nb sequences and to deplete shortened or prematurely stopped Nb sequences, we implemented frequent purification steps like PCR cycle-optimization, purification of XL-RNA, and selection for in-frame Nbs over the C-terminal His-tag. Additionally, a counterselection step was implemented to deplete unspecific bead-binders.

Our optimized mRNA/cDNA display selection procedure enabled the use of highly variable gene libraries like the c-lib that was designed for this study. Nanobody selections against the SARS-CoV2 M^{pro} yielded several Nb sequences after three and after five selection cycles. After expression, none of the selected Nbs were successfully translocated to the *E. coli* periplasm, which is necessary for the formation of a stabilizing disulfide bond (Dingus et al., 2022). Alternatively, a selected Nb was expressed in *E. coli* Origami2, which enables cytosolic disulfide bond formation and can enable cytoplasmic Nb expression (D. Li et al., 2019). This did not yield any substantial amount of soluble Nb (Figure 3.58).

In the c-lib library design, CDR sequences are entirely sequence-randomized and use NNK-codons (N: any nucleotide; K: guanine or thymine). This codon choice enables the incorporation of all 20 amino acids while reducing the Stop-codon rate from 4.7 % – all three Stop-codons would be possible in NNN-codon usage – to 3.1 % per position. However, the resulting Nb sequences are entirely synthetic and do not match the *E. coli* codon usage, which can influence protein expression greatly (Grosjean & Fiers, 1982; Robinson et al., 1984). The use of an *E. coli* strain with added tRNAs like BL21(DE3) C+ RIL (Kleber-Janke & Becker, 2000) removes the codon bias and may enable expression of genes with unfavorable codon usage. Nb expression in *E. coli* BL21(DE3) C+RIL did not produce any soluble Nbs in the *E. coli* periplasm (Figure 3.58).

Additionally, a selected Nb was attempted to be expressed in the RNase E-deficient *E. coli* BL21(DE3) Star strain to increase mRNA stability (Mädje et al., 2012), and in *E. coli* Arctic express, which aids protein expression and folding by cold-activated chaperonins (Pacheco et al., 2012). Nb expression in these two stains did not yield any substantial amounts of soluble and periplasmically translocated Nbs (Figure 3.58).

Most likely, only poorly expressing Nbs were selected because of lack of selective pressure for well-behaved, full-length Nbs (Kamalinia et al., 2021), which would be provided in cell-mediated selection methods like phage or yeast display (Boder et al., 2012). By implementing a purification step over a C-

terminal His-tag, we preselected for full-length, in-frame Nb sequences. However, additional selective pressure must be provided to enrich for stably folding proteins.

To further optimize our developed Nb selection procedure and to introduce selective pressure for well-behaved proteins, we will combine mRNA/cDNA display with phage display. At first, the large, highly variable Nb gene library will be used in three to five selection cycles, where the binders are initially enriched, and the library size is reduced. The reduced library will then be used in phage display for another two to three selection cycles (Figure 4.1), in order to enrich for stably folding Nb proteins under the selective pressure of the *E. coli* protein quality control (Gottesman et al., 1997) and the translocation to the periplasm prior to the display on the phage surface.
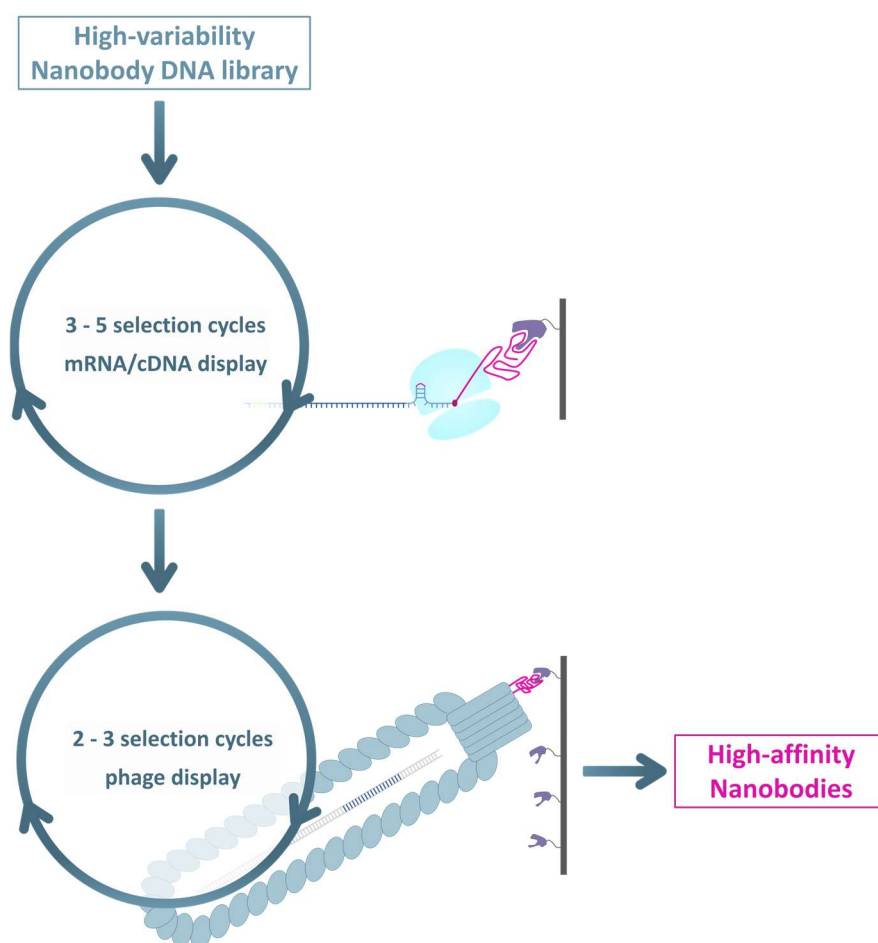
Figure 4.1 Proposed workflow for future Nanobody selections from a high-variability DNA library. The initial high-variability Nanobody DNA library is first used in three to five mRNA/cDNA display selection cycles and subsequently in two to three phage display selection cycles. This procedure should yield high-affinity Nanobodies.

To enable the c-lib for the use in phage display, a maximum library size of $10^8 – 10^9$ must be reached (Derda et al., 2011). An enrichment rate of 10^2 per selection cycle can be hypothesized (Hanes & Plückthun, 1997), meaning that after each selection cycle, the library size is reduced by a factor of 10^2. Consequently, a library size of 10^7 would be obtained after three cycles (Figure 4.2), which would be

feasible for the use in phage display (Zahra et al., 1999). Taken that the enrichment factor in mRNA/cDNA display selection would, in theory, be as low as 10 per selection cycle, a library size of 10^8 would be reached only after five selection cycles. To exactly determine the library size after selections and to calculate the actual enrichment factor, a quantitative PCR step could be implemented (Moore et al., 2003).
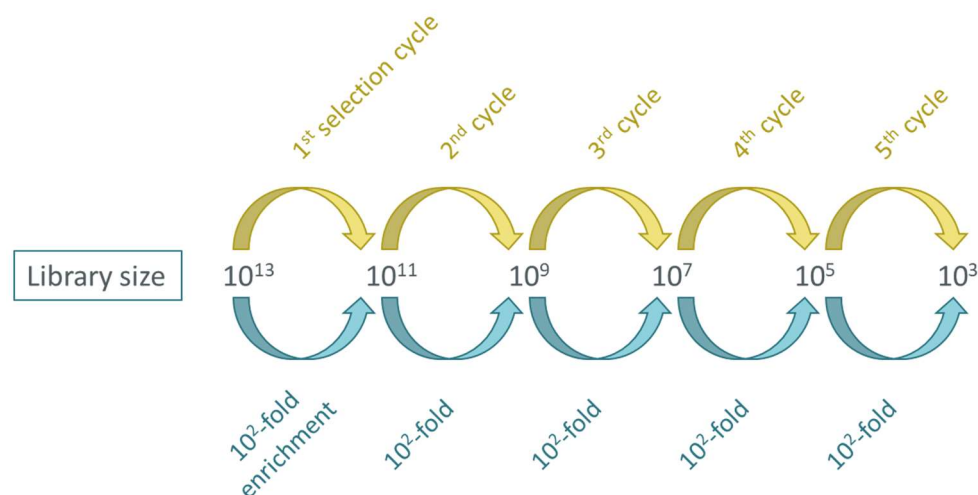
Figure 4.2 Visualization of library size through five selection cycles. An estimated enrichment factor of 10^2 per selection cycle (Hanes & Plückthun, 1997) was used to calculate the remaining library size after one to five selection cycles from an initial library size of 10^{13}.

In vitro selection by mRNA/cDNA display is a very challenging method comprising many individual steps, as shown in this work. Overamplification of the gene library can cause problems like library bias by preferential amplification of certain sequences (Polz & Cavanaugh, 1998) or the synthesis of unwanted secondary products (Cha & Thilly, 1993). Additionally, a lack of selective pressure can lead to the unspecific selection of shortened sequences (Kamalinia et al., 2021).

In this study, we developed a Nb selection procedure by mRNA/cDNA that avoids overamplification of the Nb gene library by cycle-optimization of every single PCR. Additionally, a high library quality was maintained by including several purification steps. Nb DNA was purified by agarose gel extraction prior to *in vitro* transcription, XL-RNA was purified by extraction from urea-polyacrylamide gels followed by oligo(dT)-affinity-purification, and after *in vitro* translation, Nb mRNA-protein hybrid molecules were again oligo(dT)-affinity-purified. To preselect for in-frame Nb proteins, a purification over a C-terminal His-tag was implemented, and unspecific binders were depleted by a counterselection step. Despite this stringent protocol, no high-affinity Nbs could be selected. Therefore, a combined approach of mRNA/cDNA display and phage display (Figure 4.1) will be established in the future, which integrates advantages from both selection methods, namely the use of large gene libraries and the application of selective pressure to enrich for stably folding, high-affinity binders that will serve for structural studies and protein purifications of the target proteins, and potentially for the development of therapeutics against rare retinopathies or COVID19.

5. Conclusion and Outlook

In this work, I was able to produce first human retGC1 protein from a recombinant host with purity suitable for structural studies and Nb selections. More efforts must be made to obtain sufficient protein quantities, e.g., by further improving membrane solubilization or by scaling up the culture volume, optimally by introducing a TRex293 suspension culture.

I obtained first diffracting, but possibly intergrown crystals of myrGCAP1 that need to be further optimized in order to determine the 3D structure of wild type myrGCAP1 and the disease-related variant M26R. This will enable us to elucidate the mechanisms of retGC1 regulation by GCAP1 and the disease mechanisms of variants like GCAP1 M26R.

Finally, I was successful in establishing an mRNA/cDNA display-based Nb selection method, which yielded first full-length Nbs against myrGCAP1-Mg and SARS-CoV2 M^{pro}. Since the nanobodies exhibited poor biophysical behavior, we will introduce selective pressure for good folding properties via phage display in the future.

6. Appendix

6.1 Abbreviations

Abbreviation	Meaning
A	adenine
Ab	antibody
adCORD	autosomal dominant cone-rod dystrophy
Ala, A	alanine
amp	ampicillin
Arg, R	arginine
Asn, N	asparagine
Asp, D	aspartic acid, aspartate
BamHI	*Bacillus amyloliquefaciens* H restriction enzyme I
BIS-TRIS	2-[Bis(2-hydroxyethyl)amino]-2-(hydroxymethyl)propane-1,3-diol
bp	base pair
BSA	bovine serum albumin
C	cytosine
CCD	core catalytic domain
cDNA	complementary DNA
CDR	complementarity-determinong region
cGMP	cyclic guanosine monophosphate
chl	chloramphenicol
c-lib	our combinatorial Nanobody gene library
CNG	cyclic nucleotide-gated
CORD6	Cone-rod dystrophy 6
cryo EM	cryo-electron microscopy
CTE	C-terminal extension
CYMAL-5	5-Cyclohexylpentyl-β-D-maltosid
Cys, C	cysteine

DD	dimerization domain
DDM	n-Dodecyl-β-D-maltoside
DIBMA	Diisobutylene/maleic acid alternating co-polymer
DMSO	dimethylsulfoxide
DNA	Deoxyribonucleic acid
DraIII	*Deinococcus radiophilus* restriction enzyme III
DSF	differential scanning fluorimetry
DTT	dithiotreitol
E. coli	*Escherichia coli*
EcoRI	*Escherichia coli* strain R restriction enzyme I
EDC	1-Ethyl-3-(3-dimethylaminopropyl)-carbodiimide
EDTA	ethylenediaminetetraacetic acid
eGFP	enhanced green fluorescent protein
EGTA	ethylene glycol-bis(β-aminoethyl ether)-N,N,N′,N′-tetraacetic acid
ELISA	enzyme-linked immunosorbent assay
Fab	fragment antigen-binding
Fc	fragment crystallizable
FO-SPR	fiber optics surface plasmon resonance
FPLC	fast protein lipid chromatography
FR	framework region
G	guanine
GCAP1	guanylate cyclase activating protein 1
GDP	guanosine diphosphate
GFP	green fluorescent protein
Gln, Q	glutamine
Glu, E	glutamic acid, glutamate
Gly, G	glycine
GPCR	G-protein coupled receptor
GTP	Guanosine triphosphate

HBS	HEPES buffered saline
hcAb	heavy chain antibody
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
His, H	histidine
Ig	immunoglobulin
IgNAR	immunoglobulin new antigen receptor
Ile, I	isoleucine
IPTG	isopropyl β-d-1-thiogalactopyranoside
JmD	juxtamembrane domain
kan	kanamycin
kDa	kilodalton
LB	lysogeny broth
LCA1	Leber's congenital amaurosis
LD	lumen-facing domain
LDAO	N,N-Dimethyl-n-dodecylamine N-oxide
Leu, L	leucine
LS	leader sequence
LSC	low-speed centrifucation
Lys, K	lysine
MES	2-(N-morpholino)ethanesulfonic acid
Met, M	methionine
Mpro	main protease
mRNA	messenger RNA
MS	mass spectrometry
MWCO	molecular weight cut-off
myr	myristoylated
Nb	Nanobody
NcoI	*Nocardia corallina* restriction enzyme I
NdeI	*Neisseria denitrificans* restriction enzyme I

NHS	N-hydroxysuccinimide
NMR	nuclear magnetic resonance
nsp	non-structural protein
nt	nucleotide
NTA	nitrilotriacetic acid
OD	optical density
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PDE6	phosphodiesterase 6
PEG	polyethylene glycol
pH	negative decadic logarithm of the Oxonium-ion concentration
Phe, F	phenylalanine
Plpro	Papain-like protease
Pro, P	proline
PURE	protein synthesis using recombinant elements
PVDF	polyvinylidene fluoride
RD3	Retinal degeneration 3
retGC1	retinal guanylate cyclase 1
RNA	ribonucleic acid
SalI	*Streptomyces albus* restriction enzyme I
Sarcosyl	lauroylsarcosine
SARS-CoV2	severe acute respiratory syndrome - coronavirus 2
scFv	single-chain variable fragment
SDS	sodium dodecyl sufate
SEC	size-exclusion chromatography
Ser, S	serine
sf21	spodoptera frugiperda
SMA	styrene/maleic acid alternating co-polymer

SOB	super optimal broth
SOC	super optimal broth with catabolic repression
T	thymine
TB	terrific broth
TBS	Tri-buffered saline
tet	tetracycline
Thr, T	threonine
Tm	denaturing temperature of proteins
TMD	transmembrane domain
Tris	tris(hydroxymethyl)aminomethane
tRNA	transfer RNA
Trp, W	tryptophan
Tyr, Y	tyrosin
U	uracil
UC	ultracentrifugation
UV	ultraviolet
Val, V	valine
VHH	variable domain of a heavy chain of a heavy chain antibody
XbaI	*Xanthomonas badrii* restriction enzyme I
XhoI	*Xanthomonas holcicola* restriction enzyme I
XL-MS	crosslinked mass spectrometry
XL-RNA	crosslinked RNA
YFP	yellow fluorescent protein

6.2 Materials

Table 6.1 *E. coli* strains used in this work.

E. coli strain	Genotype	Reference
BL21(DE3)	*E. coli* str. B F⁻*omp*T *hsd*S$_B$ (r$_B$⁻, m$_B$⁻) *gal dcm* (DE3)	(Studier & Moffatt, 1986)
BL21(DE3) C+ RIL	*E. coli* str. B F⁻*omp*T *hsd*S$_B$ (r$_B$⁻, m$_B$⁻) *gal dcm* (DE3) (Cam^R)	(Rosano & Ceccarelli, 2009)
wk6	*E. coli* Δ (*lac-pro*), *galE, strA, nalr;* F' *lac1q Z* Δ M15*, pro+*	(Zell & Fritz, 1987)
Origami2	*E. coli* str. K-12 *Δ(ara-leu)7697 ΔlacX74 ΔphoA Pvu*II *phoR araD139 ahpC galE galK rpsL* F'[*lac⁺ lacI�q pro*] (DE3) *gor522*::Tn*10 trxB* (Str^R, Tet^R)	Novagen
Arctic express	E. coli B F– ompT hsdS(rB – mB –) dcm+ Tetr gal endA Hte [cpn10 cpn60 Gentr]	Agilent technologies
BL21(DE3) Star	*E. coli* str. B F⁻*omp*T *hsd*S$_B$ (r$_B$⁻, m$_B$⁻) *gal dcm rne*131 (DE3)	Invitrogen
TOP10	*E.* coli str. K-12 F⁻*mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) φ80*lacZ*ΔM15 Δ*lac*X74 *rec*A1 *ara*D139 Δ(*ara-leu*)7697 *gal*U *gal*K λ⁻*rps*L(Str^R) *end*A1 *nup*G	(Casadaban & Cohen, 1980)

Table 6.2 List of buffers.

Buffer	Composition
TBS	50 mM Tris-HCl pH 7.5 150 mM NaCl
TBST	50 mM Tris-HCl pH 7.5 150 mM NaCl 0.05 % Tween-20
HBS	25 mM HEPES-NaOH pH 7.5 150 mM NaCl
CCMB80	10 mM KOAc pH 6.4 80 mM CaCl$_2$ 20 mM MnCl$_2$ 10 mM MgCl$_2$ 10 % glycerol
Cell lysis buffer	20 mM HEPES-NaOH pH 7.5 300 mM NaCl
Strep affinity chromatography Wash	20 mM HEPES-NaOH pH 7.5 300 mM NaCl
MyrGCAP1 Wash	20 mM HEPES-NaOH pH 7.5 300 mM NaCl 1 mM EDTA 1 mM EGTA

Strep affinity chromatography Elution	20 mM HEPES-NaOH pH 7.5 150 mM NaCl 6 mM Desthiobiotin
MyrGCAP1 Elution	20 mM HEPES-NaOH pH 7.5 150 mM NaCl 6 mM Desthiobiotin 1 mM EDTA 1 mM EGTA
retGC1 SEC buffer	20 mM HEPES-NaOH pH 7.5 100 mM NaCl 50 mM KCl
MyrGCAP1-apo SEC	20 mM HEPES-NaOH pH 7.5 150 mM NaCl 50 μM mM EDTA 50 μM EGTA
MyrGCAP1-Ca SEC	20 mM HEPES-NaOH pH 7.5 150 mM NaCl 1 mM $CaCl_2$
MyrGCAP1-Mg SEC	20 mM HEPES-NaOH pH 7.5 150 mM NaCl 1 mM $MgCl_2$
TES	200 mM Tris pH 8 500 μM EDTA 500 mM sucrose
Nb affinity chromatography buffer 1	50 mM Na_2HPO_4 pH 7 1 M NaCl
Nb affinity chromatography buffer 1	50 mM NaH_2PO_4 pH 6 1 M NaCl
Nb affinity Elution	50 mM NaOAc pH 4.66 1 M NaCl
Nb SEC	20 mM HEPES pH 8 150 mM NaCl
E-blot (Western blot)	25 mM Tris 200 mM Glycine 20 % EtOH
TB medium buffer	0.17 M KH_2PO_4 0.73 M K_2HPO_4

Table 6.3 List of buffers for Nb selection.

Buffer	Composition
His-binding	20 mM Tris pH 8 300 mM NaCl 0.05 % Tween-20
His-wash	20 mM Tris pH 8 150 mM NaCl

	0.05 % Tween-20
His-elution	20 mM Tris pH 8 150 mM NaCl 150 mM imidazole 0.05 % Tween-20
Counter selection	20 mM Tris pH 8 300 mM NaCl 0.05 % Tween-20
Selection bind	20 mM Tris pH 8 300 mM NaCl 0.05 % Tween-20
Selection wash 1	20 mM Tris pH 8 300 mM NaCl 0.05 % Tween-20 2 µg/ml RNase-free BSA
Selection wash 2	20 mM Tris pH 8 150 mM NaCl 0.05 % Tween-20
Selection elution	100 mM Tris pH 8 1 mM EDTA 50 mM biotin

Table 6.4 List of media.

Buffer	Composition
LB	10 g/l Bacto tryptone 5 g/l yeast extract 10 g/l NaCl
TB	12 g/l Bacto tryptone 14 g/l yeast extract 0.4 % (v/v) glycerol 100 ml TB medium buffer (Table 6.2)
SOB	20 g/l Bacto tryptone 5 g/l yeast extract 10 mM NaCl 2.5 mM KCl 20 mM $MgSO_4$
2 YT	16 g/l Bacto tryptone 10 g/l yeast extract 5 g/l NaCl
$Y_{Glc}4.5_{-Trp}$	3.8 g/l -Trp drop-out media supplement 6.7 g/l Yeast Nitrogen Base 10.4 g/l sodium citrate 7.4 g citric acid monohydrate 20 g/l glucose Adjust pH to 4.5

Table 6.5 Primer list.

Primer name	Sequence	Purpose
EcoRI-retGC1-fw	ACG AAT TCA TGA CCG CCT GCG	Cloning of retGC1 into pFL-Cstrep vector
XbaI-retGC1-rev	ACT CTA GAA GAG AAC TGG CCC GG	Cloning of retGC1 into pFL-Cstrep vector
NdeI-GCAP1	GGA ATT CCA TAT GGG TAA CAT CAT GAG CGG CAA G	Cloning of GCAP into pStrep1 or PET22b
XhoI-GCAP1	CCG CTC GAG GCC GTC AGC TTC AGC	Cloning of GCAP into pStrep1 or pET22b
GCAP1 M26R fw	GAA ATT CAG GAG CGA GTG CCC GAG	Mutagenesis of GCAP1 M26R
GCAP1 M26R rev	CGG TCC TGA ATT TCT TGT ACC AC	Mutagenesis of GCAP1 M26R
Nb rescue fw	CTG CGC GGC GAG CGG C	Rescue of frameshifted Nb sequences
Nb rescue rev	GGT CAC CTG GGT GCC CTG GC	Rescue of frameshifted Nb sequences
Nb rescue bb fw	GCC AGG GCA CCC AGG TGA CC	Rescue of frameshifted Nb sequences – plasmid amplification
Nb rescue bb rev	GCC GCT CGC CGC GCA G	Rescue of frameshifted Nb sequences – plasmid amplification
PCR1 c-lib fw	CCA TGG ACT ACA AAG ATG ACG ATG ACA AAT CCT CAG GCT CTG GTA TGG CCC AGG TGC	c-Lib 1st entry PCR
PCR1 c-lib rev	GAA CCA CCG TGA TGG TGA TGG TGG TGA CCA GAG CCT GAG GAG ACG GTG AC	c-Lib 1st entry PCR
PCR2 c-lib fw	TTA ATA CGA CTC ACT ATA GGG TTT AAC TTT AAG AAG GAG GTG ATA TAC CAT GGA CTA CAA AGA TGA C	c-Lib 2nd entry PCR
PCR2 c-lib rev	GAA CCA CCG TGA TGG TGA TGG TGG TGA CCA GAG CCT GAG GAG ACG GTG AC	c-Lib 2nd entry PCR
PCR1 y-lib fw	CCA TGG ACT ACA AAG ATG ACG ATG ACA AAT CCT CAG CTG ACG CAC AGG TG	y-Lib 1st entry PCR
PCR1 y-lib rev	TTA TAG CCG TGG CCA CCA GAA CCA CCG CT GCT CAC GGT CAC CTG G	y-Lib 1st entry PCR
PCR2 y-lib fw	TTA ATA CGA CTC ACT ATA GGG TTT AAC TTT AAG AAG GTG ATA TAC CAT GGA CTA CAA AGA TGA C	y-Lib 2nd entry PCR
PCR2 y-lib rev	TTA TAG CCG TGG CCA CCA GAA CCA CCG CT GCT CAC GGT CAC CTG G	y-Lib 2nd entry PCR
RT-Nb	TTT TTT TTT TTT TTT GCC ACC AGA ACC ACC TTT GTC ATC	Reverse transcription - both libraries
RT-Nb-Chis	TTT TTT TTT TTT TTT GCC ACC AGA ACC ACC GTG ATG	Reverse transcription - both libraries with C-terminal His-tag
Cy5_RT-Nb	Cy5-TTT TTT TTT TTT TTT GCC ACC AGA ACC ACC TTT GTC ATC	Cy5-labeled Reverse transcription - both libraries
Cy_5RT-N-CHis	Cy5-TTT TTT TTT TTT TTT GCC ACC AGA ACC ACC GTG ATG	Cy5-labeled Reverse transcription - both libraries with C-terminal His-tag
NcoI-c-lib fw	GAC TTG GCC ATG GCC CAG GTG C	Nb subcloning - c-lib
SalI-c-lib rev	GTC GTG ACG TCG ACT CCT GAG GAG ACG GTG AC	Nb subcloning - c-lib
NcoI-y-lib fw	GAT GGC CAT GGC CCA GGT GCA GCT GCA GGA AAG C	Nb subcloning - y-lib
SalI-y-lib rev	GTC GTG ACG TCG ACG CTG CTC ACG GTC ACC TGG	Nb subcloning - y-lib
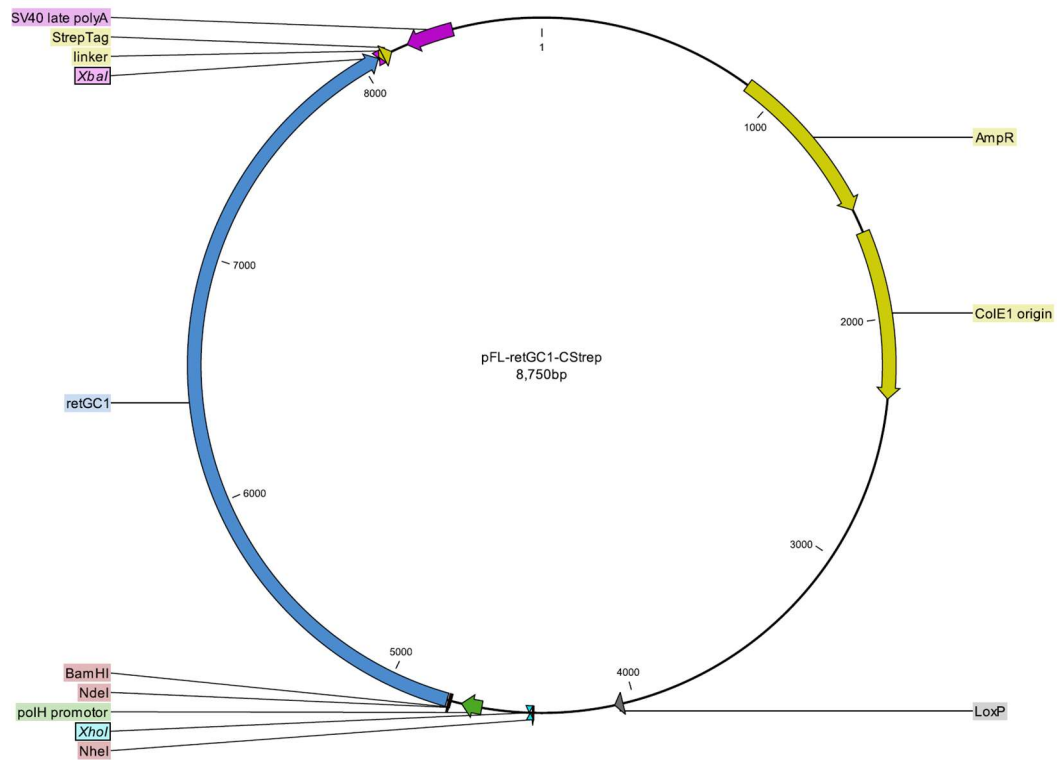
Figure 6.1 Plasmid map of pFL-retGC1-Strep with the gene for retGC1 highlighted in blue.
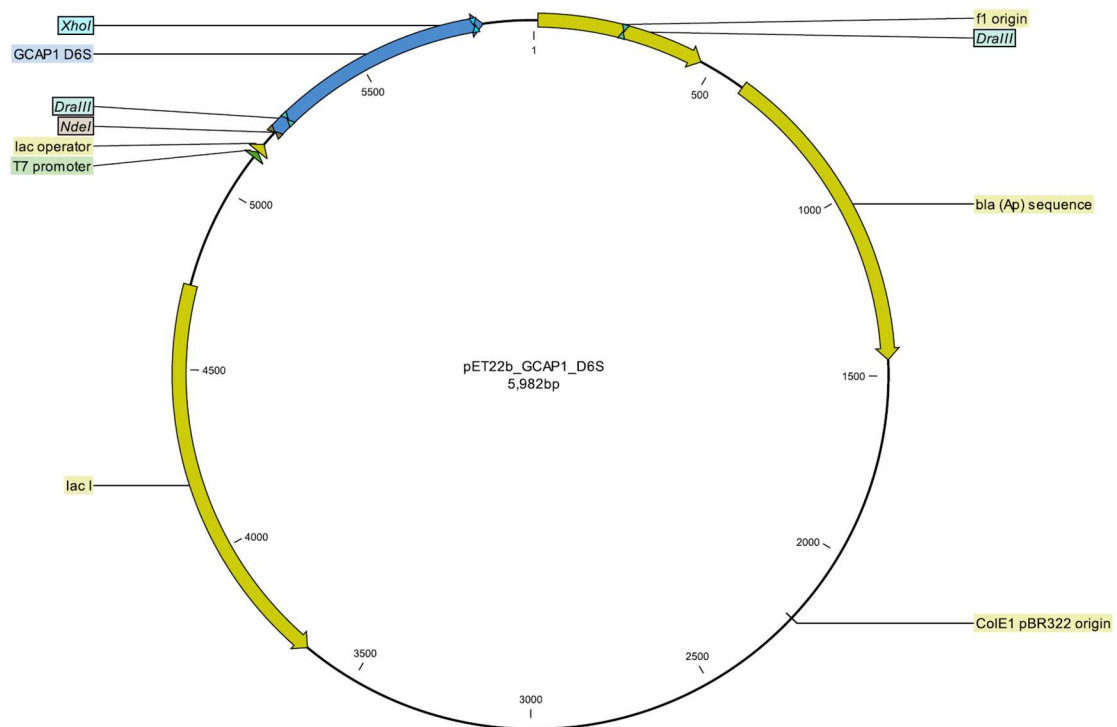
Figure 6.2 Plasmid map of pET22b-GCAP1 D6S with the gene for GCAP1 highlighted in blue.
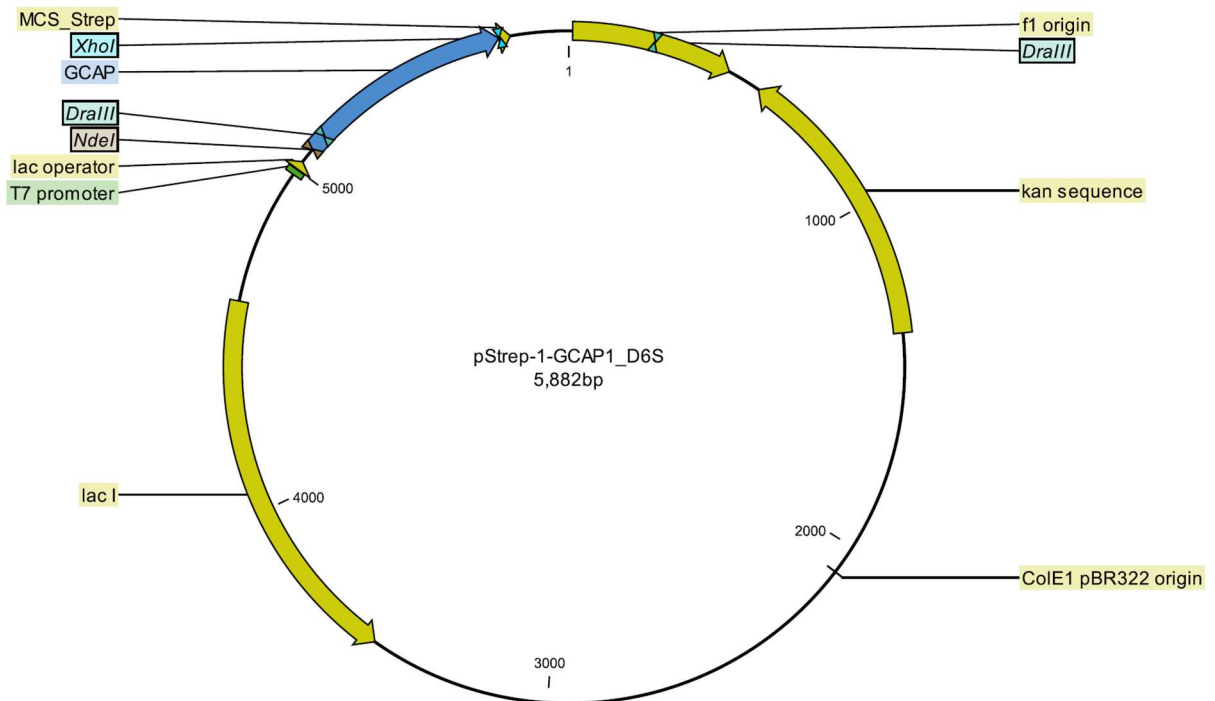
Figure 6.3 Plasmid map of pStrep1-GCAP1 D6S with the gene for GCAP1 highlighted in blue.
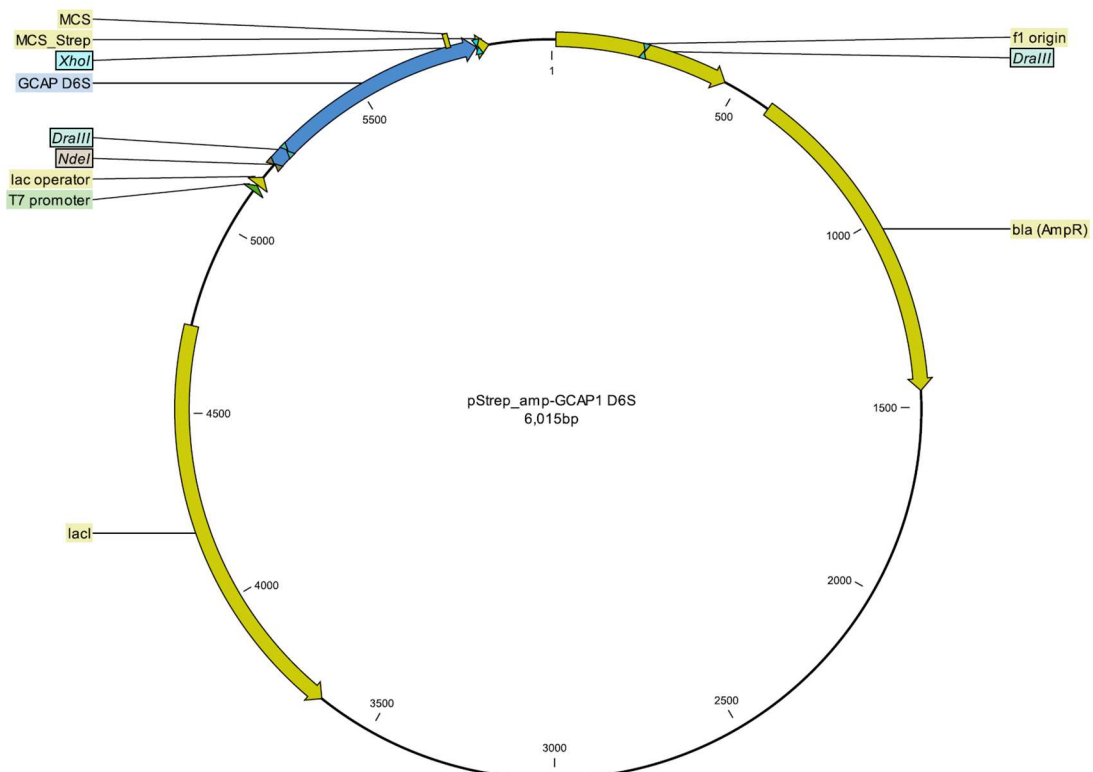
Figure 6.4 Plasmid map of pStrep(amp)-GCAP1 D6S with the gene for GCAP1 highlighted in blue.
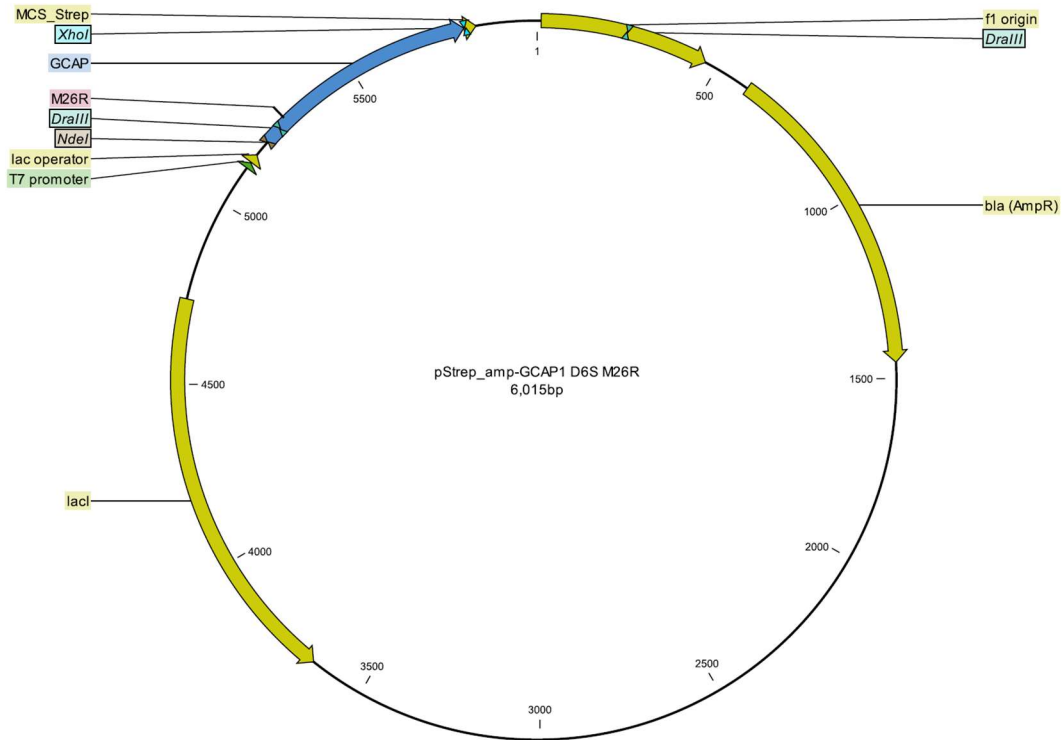
Figure 6.5 Plasmid map of pStrep(amp)-GCAP1 D6S M26R with the gene for GCAP1 highlighted in blue.
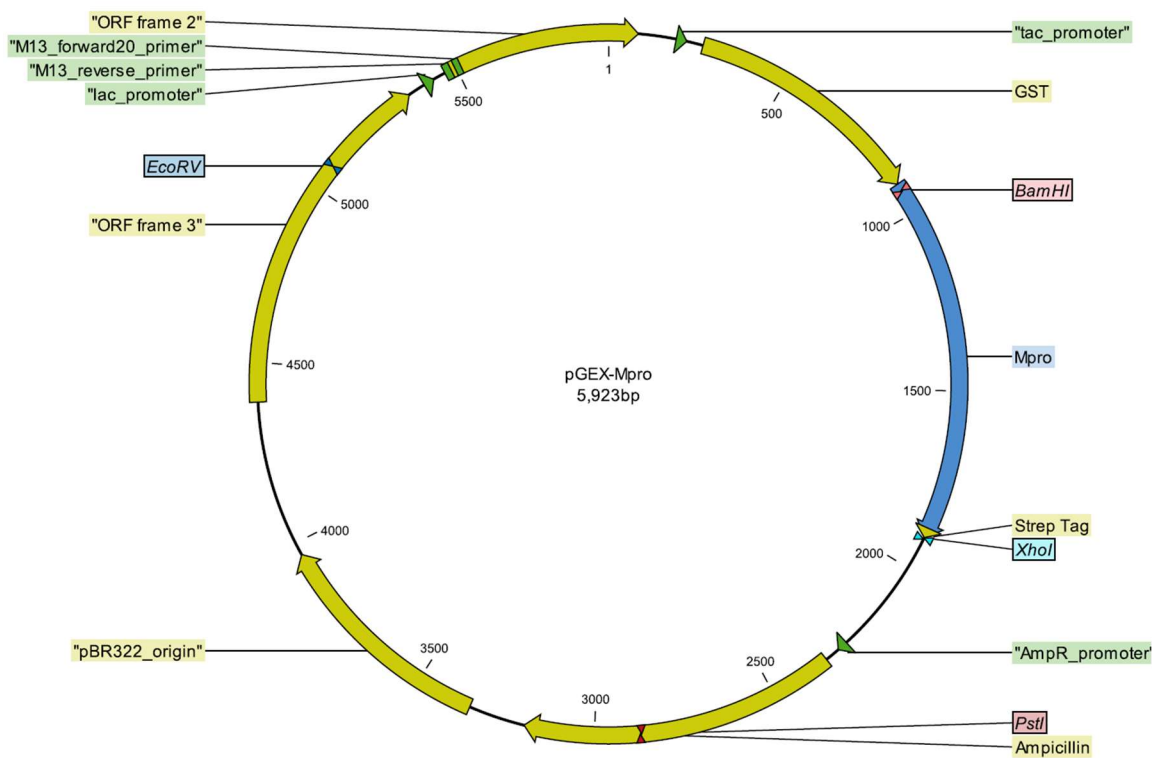
Figure 6.6 Plasmid map of pGEX-Mpro with the M^{pro}-gene marked in blue.
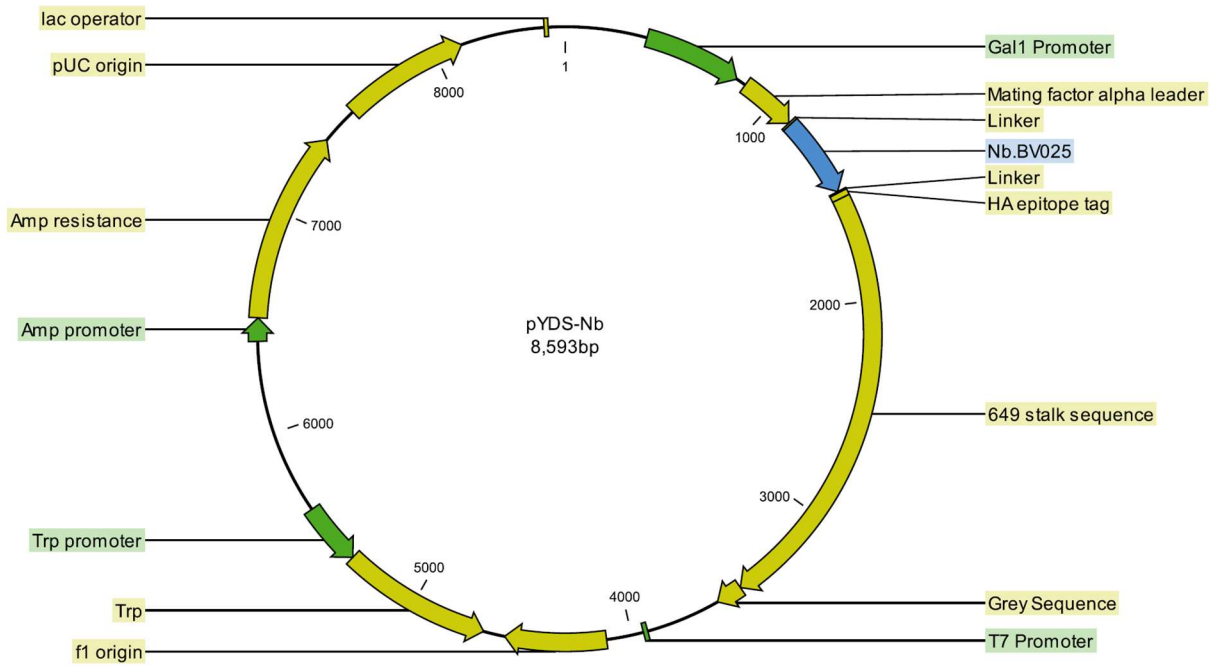
Figure 6.7 pYDS-Nb plasmid map with a Nb gene (Nb.BV025) highlighted in blue.
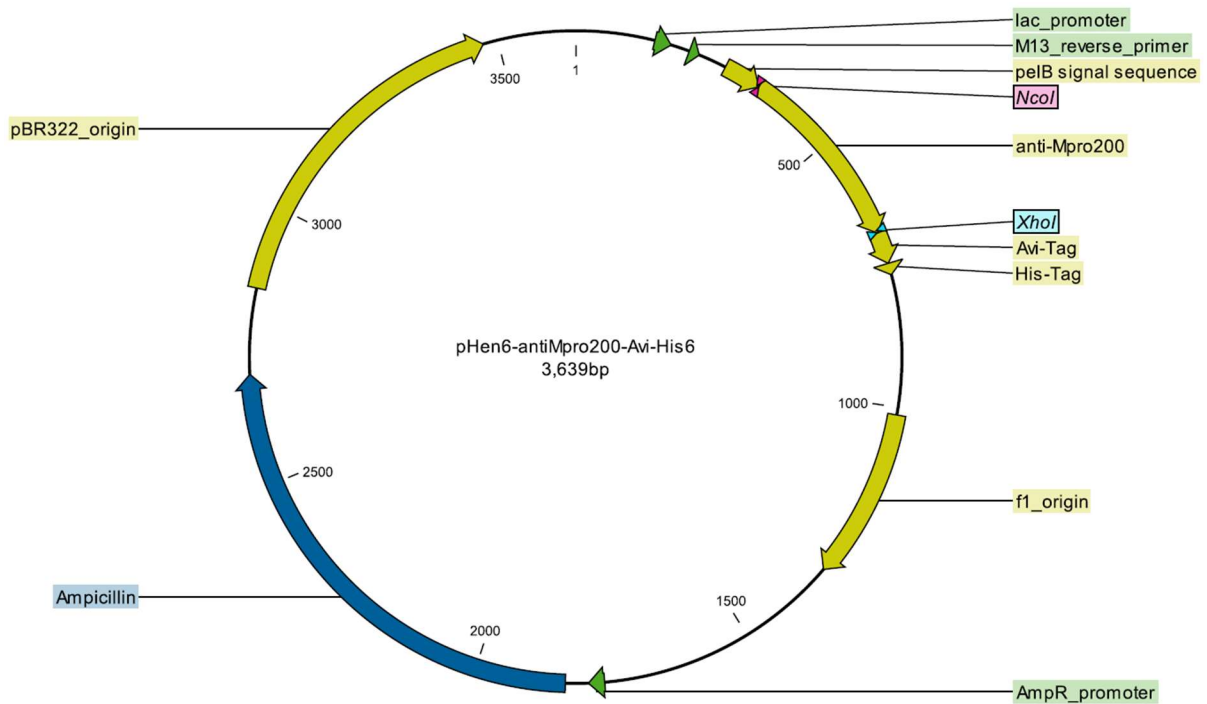
Figure 6.8 Plasmid map of pHen6-Nb-Avi His with a Nb gene (anti-M^{pro}-200) marked in blue.

ProtParam analysis of retGC1-Strep

```
        10         20         30         40         50         60
MTACARRAGG LPDPGLCGPA WWAPSLPRLP RALPRLPLLL LLLLLQPPAL SAVFTVGVLG

        70         80         90        100        110        120
PWACDPIFSR ARPDLAARLA AARLNRDPGL AGGPRFEVAL LPEPCRTPGS LGAVSSALAR

       130        140        150        160        170        180
VSGLVGPVNP AACRPAELLA EEAGIALVPW GCPWTQAEGT TAPAVTPAAD ALYALLRAFG

       190        200        210        220        230        240
WARVALVTAP QDLWVEAGRS LSTALRARGL PVASVTSMEP LDLSGAREAL RKVRDGPRVT

       250        260        270        280        290        300
AVIMVMHSVL LGGEEQRYLL EAAEELGLTD GSLVFLPFDT IHYALSPGPE ALAALANSSQ

       310        320        330        340        350        360
LRRAHDAVLT LTRHCPSEGS VLDSLRRAQE RRELPSDLNL QQVSPLFGTI YDAVFLLARG

       370        380        390        400        410        420
VAEARAAAGG RWVSGAAVAR HIRDAQVPGF CGDLGGDEEP PFVLLDTDAA GDRLFATYML

       430        440        450        460        470        480
DPARGSFLSA GTRMHFPRGG SAPGPDPSCW FDPNNICGGG LEPGLVFLGF LLVVGMGLAG

       490        500        510        520        530        540
AFLAHYVRHR LLHMQMVSGP NKIILTVDDI TFLHPHGGTS RKVAQGSRSS LGARSMSDIR

       550        560        570        580        590        600
SGPSQHLDSP NIGVYEGDRV WLKKFPGDQH IAIRPATKTA FSKLQELRHE NVALYLGLFL

       610        620        630        640        650        660
ARGAEGPAAL WEGNLAVVSE HCTRGSLQDL LAQREIKLDW MFKSSLLLDL IKGIRYLHHR

       670        680        690        700        710        720
GVAHGRLKSR NCIVDGRFVL KITDHGHGRL LEAQKVLPEP PRAEDQLWTA PELLRDPALE

       730        740        750        760        770        780
RRGTLAGDVF SLAIIMQEVV CRSAPYAMLE LTPEEVVQRV RSPPPLCRPL VSMDQAPVEC

       790        800        810        820        830        840
ILLMKQCWAE QPELRPSMDH TFDLFKNINK GRKTNIIDSM LRMLEQYSSN LEDLIRERTE

       850        860        870        880        890        900
ELELEKQKTD RLLTQMLPPS VAEALKTGTP VEPEYFEQVT LYFSDIVGFT TISAMSEPIE

       910        920        930        940        950        960
VVDLLNDLYT LFDAIIGSHD VYKVETIGDA YMVASGLPQR NGQRHAAEIA NMSLDILSAV

       970        980        990       1000       1010       1020
GTFRMRHMPE VPVRIRIGLH SGPCVAGVVG LTMPRYCLFG DTVNTASRME STGLPYRIHV

      1030       1040       1050       1060       1070       1080
NLSTVGILRA LDSGYQVELR GRTELKGKGA EDTFWLVGRR GFNKPIPKPP DLQPGSSNHG

      1090       1100       1110       1120
ISLQEIPPER RRKLEKARPG QFSSRDLQSS VDGSWSHPQF EK
```

Number of amino acids: 1122

Molecular weight: 122230.83

Theoretical pI: 6.99

Amino acid composition: [CSV format]

Ala (A)	112	10.0%
Arg (R)	91	8.1%
Asn (N)	22	2.0%
Asp (D)	56	5.0%
Cys (C)	18	1.6%
Gln (Q)	35	3.1%
Glu (E)	65	5.8%
Gly (G)	97	8.6%
His (H)	28	2.5%
Ile (I)	41	3.7%
Leu (L)	149	13.3%
Lys (K)	28	2.5%
Met (M)	26	2.3%
Phe (F)	36	3.2%
Pro (P)	86	7.7%
Ser (S)	73	6.5%
Thr (T)	49	4.4%
Trp (W)	16	1.4%
Tyr (Y)	19	1.7%
Val (V)	75	6.7%
Pyl (O)	0	0.0%
Sec (U)	0	0.0%
(B)	0	0.0%
(Z)	0	0.0%
(X)	0	0.0%

Total number of negatively charged residues (Asp + Glu): 121
Total number of positively charged residues (Arg + Lys): 119

Atomic composition:

Carbon	C	5439
Hydrogen	H	8678
Nitrogen	N	1552
Oxygen	O	1563
Sulfur	S	44

Formula: $C_{5439}H_{8678}N_{1552}O_{1563}S_{44}$
Total number of atoms: 17276

Extinction coefficients:

Extinction coefficients are in units of M^{-1} cm^{-1}, at 280 nm measured in water.

Ext. coefficient   117435
Abs 0.1% (=1 g/l)   0.961, assuming all pairs of Cys residues form cystines

Ext. coefficient   116310
Abs 0.1% (=1 g/l)   0.952, assuming all Cys residues are reduced

Estimated half-life:

The N-terminal of the sequence considered is M (Met).

The estimated half-life is: 30 hours (mammalian reticulocytes, in vitro).
                           >20 hours (yeast, in vivo).
                           >10 hours (Escherichia coli, in vivo).

Instability index:

The instability index (II) is computed to be 51.98
This classifies the protein as unstable.

Aliphatic index: 95.41

Grand average of hydropathicity (GRAVY): -0.069

The ProtParam tool is part of the ExPASy proteomics server (Gasteiger et al., 2005).

ProtParam analysis of GCAP1 D6S

```
        10         20         30         40         50         60
MGNIMSGKSV EELSSTECHQ WYKKFMTECP SGQLTLYEFR QFFGLKNLSP WASQYVEQMF

        70         80         90        100        110        120
ETFDFNKDGY IDFMEYVAAL SLVLKGKVEQ KLRWYFKLYD VDGNGCIDRD ELLTIIRAIR

       130        140        150        160        170        180
AINPCSDSTM TAEEFTDTVF SKIDVNGDGE LSLEEFMEGV QKDQMLLDTL TRSLDLTRIV

       190        200        210
RRLQNGEQDE EGASGRETEA AEADGLEWSH PQFEK
```

Number of amino acids: 215

Molecular weight: 24764.77

Theoretical pI: 4.48

Amino acid composition: [CSV format]

```
Ala (A)   10     4.7%
Arg (R)   10     4.7%
Asn (N)    7     3.3%
Asp (D)   16     7.4%
Cys (C)    4     1.9%
Gln (Q)   11     5.1%
Glu (E)   24    11.2%
Gly (G)   15     7.0%
His (H)    2     0.9%
Ile (I)    9     4.2%
Leu (L)   21     9.8%
Lys (K)   12     5.6%
Met (M)    8     3.7%
Phe (F)   13     6.0%
Pro (P)    4     1.9%
Ser (S)   15     7.0%
Thr (T)   13     6.0%
Trp (W)    4     1.9%
Tyr (Y)    7     3.3%
Val (V)   10     4.7%
Pyl (O)    0     0.0%
Sec (U)    0     0.0%

 (B)       0     0.0%
 (Z)       0     0.0%
 (X)       0     0.0%
```

Total number of negatively charged residues (Asp + Glu): 40
Total number of positively charged residues (Arg + Lys): 22

Atomic composition:

```
Carbon      C      1094
Hydrogen    H      1679
Nitrogen    N       283
Oxygen      O       349
Sulfur      S        12
```

Formula: $C_{1094}H_{1679}N_{283}O_{349}S_{12}$
Total number of atoms: 3417

Extinction coefficients:

Extinction coefficients are in units of $M^{-1} cm^{-1}$, at 280 nm measured in water.

Ext. coefficient   32680
Abs 0.1% (=1 g/l)   1.320, assuming all pairs of Cys residues form cystines

Ext. coefficient   32430
Abs 0.1% (=1 g/l)   1.310, assuming all Cys residues are reduced

Estimated half-life:

The N-terminal of the sequence considered is M (Met).

The estimated half-life is: 30 hours (mammalian reticulocytes, in vitro).
                            >20 hours (yeast, in vivo).
                            >10 hours (Escherichia coli, in vivo).

Instability index:

The instability index (II) is computed to be 37.04
This classifies the protein as stable.

Aliphatic index: 72.56

Grand average of hydropathicity (GRAVY): -0.491

ProtParam analysis of GCAP1 D6S M26R

```
        10         20         30         40         50         60
MGNIMSGKSV EELSSTECHQ WYKKFRTECP SGQLTLYEFR QFFGLKNLSP WASQYVEQMF

        70         80         90        100        110        120
ETFDFNKDGY IDFMEYVAAL SLVLKGKVEQ KLRWYFKLYD VDGNGCIDRD ELLTIIRAIR

       130        140        150        160        170        180
AINPCSDSTM TAEEFTDTVF SKIDVNGDGE LSLEEFMEGV QKDQMLLDTL TRSLDLTRIV

       190        200        210
RRLQNGEQDE EGASGRETEA AEADGLEWSH PQFEK
```

Number of amino acids: 215

Molecular weight: 24789.77

Theoretical pI: 4.53

Amino acid composition: [CSV format]

Ala (A)	10	4.7%
Arg (R)	11	5.1%
Asn (N)	7	3.3%
Asp (D)	16	7.4%
Cys (C)	4	1.9%
Gln (Q)	11	5.1%
Glu (E)	24	11.2%
Gly (G)	15	7.0%
His (H)	2	0.9%
Ile (I)	9	4.2%
Leu (L)	21	9.8%
Lys (K)	12	5.6%
Met (M)	7	3.3%
Phe (F)	13	6.0%
Pro (P)	4	1.9%
Ser (S)	15	7.0%
Thr (T)	13	6.0%
Trp (W)	4	1.9%
Tyr (Y)	7	3.3%
Val (V)	10	4.7%
Pyl (O)	0	0.0%
Sec (U)	0	0.0%
(B)	0	0.0%
(Z)	0	0.0%
(X)	0	0.0%

Total number of negatively charged residues (Asp + Glu): 40
Total number of positively charged residues (Arg + Lys): 23

Atomic composition:

Carbon	C	1095
Hydrogen	H	1682
Nitrogen	N	286
Oxygen	O	349
Sulfur	S	11

Formula: $C_{1095}H_{1682}N_{286}O_{349}S_{11}$
Total number of atoms: 3423

Extinction coefficients:

Extinction coefficients are in units of $M^{-1} cm^{-1}$, at 280 nm measured in water.

Ext. coefficient   32680
Abs 0.1% (=1 g/l)   1.318, assuming all pairs of Cys residues form cystines

Ext. coefficient   32430
Abs 0.1% (=1 g/l)   1.308, assuming all Cys residues are reduced

Estimated half-life:

The N-terminal of the sequence considered is M (Met).

The estimated half-life is: 30 hours (mammalian reticulocytes, in vitro).
                         >20 hours (yeast, in vivo).
                         >10 hours (Escherichia coli, in vivo).

Instability index:

The instability index (II) is computed to be 37.17
This classifies the protein as stable.

Aliphatic index: 72.56

Grand average of hydropathicity (GRAVY): -0.520

```
        10         20         30         40         50         60
MSPILGYWKI KGLVQPTRLL LEYLEEKYEE HLYERDEGDK WRNKKFELGL EFPNLPYYID

        70         80         90        100        110        120
GDVKLTQSMA IIRYIADKHN MLGGCPKERA EISMLEGAVL DIRYGVSRIA YSKDFETLKV

       130        140        150        160        170        180
DFLSKLPEML KMFEDRLCHK TYLNGDHVTH PDFMLYDALD VVLYMDPMCL DAFPKLVCFK

       190        200        210        220        230        240
KRIEAIPQID KYLKSSKYIA WPLQGWQATF GGGDHPPKSD LVPRGSSAVL QSGFRKMAFP

       250        260        270        280        290        300
SGKVEGCMVQ VTCGTTTLNG LWLDDVVYCP RHVICTSEDM LNPNYEDLLI RKSNHNFLVQ

       310        320        330        340        350        360
AGNVQLRVIG HSMQNCVLKL KVDTANPKTP KYKFVRIQPG QTFSVLACYN GSPSGVYQCA

       370        380        390        400        410        420
MRPNFTIKGS FLNGSCGSVG FNIDYDCVSF CYMHHMELPT GVHAGTDLEG NFYGPFVDRQ

       430        440        450        460        470        480
TAQAAGTDTT ITVNVLAWLY AAVINGDRWF LNRFTTTLND FNLVAMKYNY EPLTQDHVDI

       490        500        510        520        530        540
LGPLSAQTGI AVLDMCASLK ELLQNGMNGR TILGSALLED EFTPFDVVRQ CSGVTFQGPA
```

Number of amino acids: 549

Molecular weight: 61940.25

Theoretical pI: 6.08

Amino acid composition: [CSV format]

Ala (A)	29	5.3%
Arg (R)	21	3.8%
Asn (N)	25	4.6%
Asp (D)	35	6.4%
Cys (C)	16	2.9%
Gln (Q)	21	3.8%
Glu (E)	26	4.7%
Gly (G)	42	7.7%
His (H)	14	2.6%
Ile (I)	24	4.4%
Leu (L)	58	10.6%
Lys (K)	33	6.0%
Met (M)	19	3.5%
Phe (F)	27	4.9%
Pro (P)	29	5.3%
Ser (S)	29	5.3%
Thr (T)	30	5.5%
Trp (W)	8	1.5%
Tyr (Y)	25	4.6%
Val (V)	38	6.9%
Pyl (O)	0	0.0%
Sec (U)	0	0.0%
(B)	0	0.0%
(Z)	0	0.0%
(X)	0	0.0%

Total number of negatively charged residues (Asp + Glu): 61
Total number of positively charged residues (Arg + Lys): 54

Atomic composition:

Carbon	C	2787
Hydrogen	H	4295
Nitrogen	N	727
Oxygen	O	802
Sulfur	S	35

Formula: $C_{2787}H_{4295}N_{727}O_{802}S_{35}$
Total number of atoms: 8646

Extinction coefficients:

Extinction coefficients are in units of $M^{-1} cm^{-1}$, at 280 nm measured in water.

Ext. coefficient   82250
Abs 0.1% (=1 g/l)   1.328, assuming all pairs of Cys residues form cystines

Ext. coefficient   81250
Abs 0.1% (=1 g/l)   1.312, assuming all Cys residues are reduced

Estimated half-life:

The N-terminal of the sequence considered is M (Met).

The estimated half-life is: 30 hours (mammalian reticulocytes, in vitro).
                           >20 hours (yeast, in vivo).
                           >10 hours (Escherichia coli, in vivo).

Instability index:

The instability index (II) is computed to be 32.82
This classifies the protein as stable.

Aliphatic index: 83.61

Grand average of hydropathicity (GRAVY): -0.178

ProtParam analysis of M^{pro}-Strep

```
         10         20         30         40         50         60
SDLVPRGSSA VLQSGFRKMA FPSGKVEGCM VQVTCGTTTL NGLWLDDVVY CPRHVICTSE

         70         80         90        100        110        120
DMLNPNYEDL LIRKSNHNFL VQAGNVQLRV IGHSMQNCVL KLKVDTANPK TPKYKFVRIQ

        130        140        150        160        170        180
PGQTFSVLAC YNGSPSGVYQ CAMRPNFTIK GSFLNGSCGS VGFNIDYDCV SFCYMHHMEL

        190        200        210        220        230        240
PTGVHAGTDL EGNFYGPFVD RQTAQAAGTD TTITVNVLAW LYAAVINGDR WFLNRFTTTL

        250        260        270        280        290        300
NDFNLVAMKY NYEPLTQDHV DILGPLSAQT GIAVLDMCAS LKELLQNGMN GRTILGSALL

        310        320        330
EDEFTPFDVV RQCSGVTFQG PASWSHPQFE K
```

Number of amino acids: 331

Molecular weight: 36459.59

Theoretical pI: 6.05

Amino acid composition: [CSV format]

Ala (A)	19	5.7%
Arg (R)	12	3.6%
Asn (N)	21	6.3%
Asp (D)	18	5.4%
Cys (C)	12	3.6%
Gln (Q)	16	4.8%
Glu (E)	10	3.0%
Gly (G)	28	8.5%
His (H)	8	2.4%
Ile (I)	11	3.3%
Leu (L)	31	9.4%
Lys (K)	12	3.6%
Met (M)	10	3.0%
Phe (F)	18	5.4%
Pro (P)	16	4.8%
Ser (S)	21	6.3%
Thr (T)	24	7.3%
Trp (W)	4	1.2%
Tyr (Y)	11	3.3%
Val (V)	29	8.8%
Pyl (O)	0	0.0%
Sec (U)	0	0.0%
(B)	0	0.0%
(Z)	0	0.0%
(X)	0	0.0%

Total number of negatively charged residues (Asp + Glu): 28
Total number of positively charged residues (Arg + Lys): 24

Atomic composition:

Carbon	C	1618
Hydrogen	H	2498
Nitrogen	N	436
Oxygen	O	481
Sulfur	S	22

Formula: $C_{1618}H_{2498}N_{436}O_{481}S_{22}$
Total number of atoms: 5055

Extinction coefficients:

Extinction coefficients are in units of M^{-1} cm^{-1}, at 280 nm measured in water.

Ext. coefficient   39140
Abs 0.1% (=1 g/l)   1.074, assuming all pairs of Cys residues form cystines

Ext. coefficient   38390
Abs 0.1% (=1 g/l)   1.053, assuming all Cys residues are reduced

Estimated half-life:

The N-terminal of the sequence considered is S (Ser).

The estimated half-life is: 1.9 hours (mammalian reticulocytes, in vitro).
                           >20 hours (yeast, in vivo).
                           >10 hours (Escherichia coli, in vivo).

Instability index:

The instability index (II) is computed to be 30.16
This classifies the protein as stable.

Aliphatic index: 80.63

Grand average of hydropathicity (GRAVY): -0.059

ProtParam analysis of Nb anti-M^{pro}-Avi-His

```
        10         20         30         40         50         60
MAQVHLEESG GDSVPPGASL RSCVAAGHVN AIAWVRQTLG EGRAALGQTN GFAELNSKPV

        70         80         90        100        110        120
KGRFTLSADI GGTPVDLVLA SLESEDTGMY DCLRSRFLTL GVIGRGTQVT VSSGVESGGL

       130        140
NDIFEAQKIE WHEGSHHHHH H
```

Number of amino acids: 141

Molecular weight: 14894.56

Theoretical pI: 5.78

Amino acid composition: [CSV format]

```
Ala (A)  12    8.5%
Arg (R)   7    5.0%
Asn (N)   4    2.8%
Asp (D)   6    4.3%
Cys (C)   2    1.4%
Gln (Q)   5    3.5%
Glu (E)  10    7.1%
Gly (G)  19   13.5%
His (H)   9    6.4%
Ile (I)   5    3.5%
Leu (L)  13    9.2%
Lys (K)   3    2.1%
Met (M)   2    1.4%
Phe (F)   4    2.8%
Pro (P)   4    2.8%
Ser (S)  13    9.2%
Thr (T)   8    5.7%
Trp (W)   2    1.4%
Tyr (Y)   1    0.7%
Val (V)  12    8.5%
Pyl (O)   0    0.0%
Sec (U)   0    0.0%

(B)       0    0.0%
(Z)       0    0.0%
(X)       0    0.0%
```

Total number of negatively charged residues (Asp + Glu): 16
Total number of positively charged residues (Arg + Lys): 10

Atomic composition:

```
Carbon     C     645
Hydrogen   H    1014
Nitrogen   N     194
Oxygen     O     205
Sulfur     S       4
```

Formula: $C_{645}H_{1014}N_{194}O_{205}S_4$
Total number of atoms: 2062

Extinction coefficients:

Extinction coefficients are in units of $M^{-1} cm^{-1}$, at 280 nm measured in water.

Ext. coefficient   12615
Abs 0.1% (=1 g/l)   0.847, assuming all pairs of Cys residues form cystines

Ext. coefficient   12490

Abs 0.1% (=1 g/l)   0.839, assuming all Cys residues are reduced

Estimated half-life:

The N-terminal of the sequence considered is M (Met).

The estimated half-life is: 30 hours (mammalian reticulocytes, in vitro).
                            >20 hours (yeast, in vivo).
                            >10 hours (Escherichia coli, in vivo).

Instability index:

The instability index (II) is computed to be 39.53
This classifies the protein as stable.

Aliphatic index: 82.98

Grand average of hydropathicity (GRAVY): -0.204

ThermoFisher
S C I E N T I F I C

Library Analysis of Amplified Library via Sanger Sequencing:

colonies from transformation plates were picked and sequence analysed.

48 readable sequence data were analysed

18 of the 48 clones contained erroneous sequences with deletions (17x), insertions (1x) within the reading frame.

Resulting library viability: 62%   (30 out of 48 total clones)

Sequenced Clones of Peer Group:

```
                        10        20        30        40        50        60        70        80
               ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
16ACVQRC_195350 CATGCCATGGCCCAGGTGCASCTGSWAGMGASTGGGGGARVCTYAGYACMGVCAGRGGVCTCTCTGASKMTATCCTGTRN 80
1953507_P1.A03  CATGCCATGGCCCAGGTGCACCTGCAAGCGAGTGGGGGAGCCTTAGCACCGGCAGGGGCCTCTCTGACGATATCCTGTGT 80
1953507_P1.A07  CATGCCATGGCCCAGGTGCAGCTGGAAGCGAGTGGGGGAAACTTAGTACAGCCAGAGGACTCTCTGAGGATATCCTGTGC 80
1953507_P1.A08  CATGCCATGGCCCAGGTGCACCTGCTAGCGACTGGGGGAACCTCAGCACCGGCAGAGGCCTCTCTGACGATATCCTGTAA 80
1953507_P1.B07  CATGCCATGGCCCAGGTGCAGCTGGTAGCGAGTGGGGGAACCTCAGCACAGGCAGAGGCTCTCTGAGTCTATCCTGTGC 80
1953507_P1.B08  CATGCCATGGCCCAGGTGCACCTGCAAGAGACTGGGGGAGACTCAGCACAGACAGAGGCCTCTCTGACGCTATCCTGTGT 80
1953507_P1.B09  CATGCCATGGCCCAGGTGCACCTGCTAGCGACTGGGGGAACCTCAGTACCGACAGGGGACTCTCTGACGCTATCCTGTAG 80
1953507_P1.C01  CATGCCATGGCCCAGGTGACCTGCAAATGGACCGCCTGGAATCCGATGACACGGCAGTATATGACTGTATAAGAACTTTG 80
1953507_P1.C02  CATGCCATGGCCCAGGTGCAGCTGCAAGAGACTGGGGGAAACTCAGCACCGGCAGGGGCCTCTCTGACTCTATCCTGTAC 80
1953507_P1.C04  CATGCCATGGCCCAGGTGCAGCTGCAAGCGACTGGGGGAGGCTTAGCACCGCCAGAGGGCTCTCTGAGGCTATCCTGTGA 80
1953507_P1.C05  CNTGCCATGGCCCNGGTGCAGCTGGTAGCGAGTGGGGGAAGCTCAGTACAGCCNGAGGCTCTCTGAGTCTATCCTGTGC 80
1953507_P1.C06  CATGCCATGGCCCAGGTGCAGCTGGAAGCGAGTGGGGGAGCTTAGCACCGGCAGAGGACTCTCTGACTCTATCCTGTAG 80
1953507_P1.C07  CATGCCATGGCCCAGGTGCAGCTGGTAGCGACTGGGGGAAGCTTAGCACCGCCAGGGGCTCTCTGAGTATATCCTGTGA 80
1953507_P1.C09  CATGCCATGGCCCAGGTGCAGCTGCAAGAGAGTGGGGGAGGGCTCAGCACCGACAGAGGGCTCTCTGAGGCTATCCTGTAG 80
1953507_P1.D04  CATGCCATGGCCCAGGTGCAGCTGCAAGCGACTGGGGGAGCCTCAGTACCGGCAGGGGACTCTCTGAGTCTATCCTGTGG 80
1953507_P1.D06  CATGCCATGGCCCAGGTGCAGCTGGTAGAGAGTGGGGGAAGCTTAGCACAGACAGGGGGCCTCTCTGACGATA-CCTGTGG 79
1953507_P1.D07  CATGCCATGGCCCAGGTGCAGCTGCTAGCGAGTGGGGGAACCTCAGTACAGGCAGGGGCCTCTCTGACGATATCCTGTGT 80
1953507_P1.D08  CATGCCATGGCCCAGGTGCACTTGCAAGAGACTGGGGGAAACTCAGTACCGGCAGGGGGCTCTCTGAGGCTATCCTGTAT 80
1953507_P1.E06  CATGCCATGGCCCAGGTGCAGCTGCTAGCGACTGGGGGAAGCTCAGCACAGGCAGGGGCTCTCTGACGATATCCTGTGT 80
1953507_P1.E08  CATGCCATGGCCCAGGTGCACCTGCNANAGACTGGGGGAACCTCAGTACAGGCAGGGGACTCTCTGACGATATCCTGTAC 80
1953507_P1.E09  CATGCCATGGCCCAGGTGCAGCTGGTAGAGACTGGGGGAGGCTCAGTACCGGCAGGGGGCTCTCTGACGATATCCTGTAT 80
1953507_P1.F01  CATGCCATGGCCCAGGTGCAGCTGGAAGCGACTGGGGGAAACTTAGCACCAGCAGGCTCTCTGAGGCGTATCCTGTGT 80
1953507_P1.F02  CATGCCATGGCCCAGGTGCAGCTGGAAGCGAGTGGGGGAGCTTAGCACAGGCAGGGGACTCTCTGAGGCTATCCTGTGT 80
1953507_P1.F04  CATGCCATGGCCCAGGTGCAGCTGCAAGCGACTGGGGGAGACTCAGTACAGGCAGGGGCCTCTCTGACTATATCCTGTGG 80
1953507_P1.F08  CATGCCATGGCCCAGGTGCAGCTGCAAGAGAGTGGGGGAAGCTCAGCACCGCCAGGGGGCTCTCTGACGATATCCTGTAC 80
1953507_P1.G01  CATGCCATGGCCCAGGTGCACCTGCAAGCGAGTGGGGGAGACTTAGTACCGCCAGGGGCCTCTCTGACGCTATCCTGTGA 80
1953507_P1.G02  CATGCCATGGCCCAGGTGCACCTGCTAGAGAGTGGGGGAAACTCAGTACCGCCAGAGGGCTCTCTGACGCTATCCTGTAG 80
1953507_P1.G04  CATGCCATGGCCCAGGTGCAGCTGGTAGAGAGTGGGGGAGCCTCAGTACAGGCAGAGGCCTCTCTGACGATATCCTGTGC 80
1953507_P1.G06  CATGCCATGGCCCAGGTGCAGCTGGTAGCGAGTGGGGGAACCTTAGTACCGGCAGGGGACTCTCTGAGTCTATCCTGTGA 80
1953507_P1.G07  CATGCCATGGCCCAGGTGCAGCTGGTAGAGACTGGGGGAAACTCAGCACCGACAGAGGCCTCTCTGACGCTATCCTGTGG 80
1953507_P1.G09  CATGCCATGGCCCAGGTGCAGCTGCTAGCGACTGGGGGAGACTCAGCACAGCCAGAGGGCTCTCTGAGTCTATCCTGTGC 80
1953507_P1.H01  CATGCCATGGCCCAGGTGCACCTGCTAGAGAGTGGGGGAGCCTCAGCACAGCCAGAGGCACTCTCTGACTATATCCTGTGT 80
1953507_P1.H02  CATGCCATGGCCCAGGTGCAGCTGCCTAGAGAGTGGGGGAACCTTAGTACCGGCAGGGGACTCTCTGACTCTATCCTGTAC 80
1953507_P1.H04  CATGCCATGGCCCAGGTGCACCTGCAAGCGACTGGGGGAAACTCAGTACCGCCAGAGGCCTCTCTGAGTCTATCCTGTAT 80
1953507_P1.H06  CATGCCATGGCCCAGGTGCACCTGCAAGAGACTGGGGGAAACTCAGTACCGACAGGGGGACTCTCTGAGGCTATCCTGTGT 80
1953507_P1.H08  CATGCCATGGCCCAGGTGCAGCTGGTAGAGACTGGGGGAACTTAGTACAGACAGGGCAGTCTCTGAGTATATCCTGTAT 80
1953507_P2.A08  CATGCCATGGCCCAGGTGCAGCTGCAAGCGACTGSAGGGGAAACTTAGCACCGCCAGAGGCCTCTCTGAGGATATCCTGTGT 80
1953507_P2.B06  CATGCCATGGCCCAGGTGCAGCTGGAAGAGAGTGGGGGAGACTCAGCACCGACAGACAGAGGACTCTCTGAGTCTATCCTGTGC 80
1953507_P2.B09  CATGCCATGGCCCAGGTGCAGCTGCTAGAGAGTGGGGGAGACTTAGCACAGGCAGGGGACTCTCTGAGGATATCCTGTGG 80
1953507_P2.C02  CATGCCATGGCCCAGGTGCAGCTGGTAGCGAGTGGGGGAGACTCAGCACCGGCAGGGGCCTCTCTGACTATATCCTGTGA 80
1953507_P2.C07  CATGCCATGGCCCAGGTGCAGCTGGTAGCGAGTGGGGGAACCTCAGCACCGGCAGGGGCTC...-.............. 62
1953507_P2.E05  CATGCCATGGCCCAGGTGCAGCTGGTAGCGACTGGGGGAAGCTCAGCACAGACAGGGGCCTCTCTGACTCTATCCTGTAC 80
1953507_P2.E06  CATGCCATGGCCCAGGTGCACCTTCAAGAGAGTGGGGGAAGCTCAGTACCGGCAGGGGACTCXCTGACTATATCCTGTGT 80
1953507_P2.E08  CATGCCATGGCCCAGGTGCAGCTGCAAGCGACTGGGGGAAGCTTAGCACATCCAGGGGCTCTCTGACTCTATCCTGTAA 80
1953507_P2.F04  CATGCCATGGCCCAGGTGCACCTGCTAGAGACTGGGGGAGACTCAGCACAGACAGAGGGCTCTCTGACGCTATCCTGTAT 80
1953507_P2.F07  CATGCCATGGCCCAGGTGCACCTGGTAGCGACTGGGGGAGCTCAGCACCGGCAGGGGACTCTCTGACGCTATCCTGTGC 80
1953507_P2.F08  CATGCCATGGCCCAGGTGCACCTGGTAGAGAGTGGGGGAAGCTCAGTACAGGCAGGGGGCTCTCTGAGTCTATCCTGTGT 80
1953507_P2.G04  CATGCCATGGCCCAGGTGCACCTGGAAGCGACTGGGGGAAACTCAGTACAGACAGAGGCTCTCTGACGCTATCCTGTGA 80
1953507_P2.H05  CATGCCATGGCCCAGGTGCACCTGCAAGAGAGTGGGGGAACCTCAGCACAGGCAGGGGCTCTCTGAGGATATCCTGTGT 80
```

```
              90        100       110       120       130       140       150       160
         ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
16ACVQRC_195350  AGNCKCARRANNKNNKNNKNNKNNKNNKNNKNNKNNKNNYRDVCTGGBWCCGCRAVCACYARRAVMAGDACDCGMMB  160
1953507_P1.A03   AGGCGCAGGATAGTGGAGTGGT..................ATGTGCTGGGACCGGCCGAACAC.....-......AGGAC  129
1953507_P1.A07   AGACGCAAAACATGAGCGTGCT.......CTGGGCTGGGTCCGCCGACCACTAAGAACAGAACGCGCCC  142
1953507_P1.A08   AGCCTCAAGAAGGATGGCTCGGGCGTTTGATATT......GTGACCTGGTACCGCCAAGCACCAGGAGCAGTACTCGCAG  154
1953507_P1.B07   AGCCGCAAGATGTTTTTTGATGTGGACTGAGGAT......CCAGGCTGGGTCCGCCAAGCACCCAAACCAGAACACGAAG  142
1953507_P1.B08   AGACTCAGAAGGTCCTCGTGTT..........CTATGCTGGTTCCGCCAAACACCAAACCAGAACACGAAG  142
1953507_P1.B09   AGTCTCAAAATTGGTGCCTCGT..........TTGGGCTGGTTCCGCCAAGCACCAGAACAAGAACTCGCCG  142
1953507_P1.C01   ACTACGGTTGCGTTGATGATGTAGCATACTGAGCGGTATCGGCTG....................................  125
1953507_P1.C02   AGTCTCAGGAGTGATGTATTAT..................ACAACCTGGGACCGCCNNNNCCAGGAGAAGGACACGCAC  142
1953507_P1.C04   AGACTCAGGAGTGATGACTTAG..........TTAAACTGGCTCCGCCGACCACCAGGACCGTACACGACG  142
1953507_P1.C05   AGGCGCAAGATCTTCGTCGTCGTTGTGTCTTANG......GTGGACTGGGTCCGCCAAGCACCCAAAGAAGAACACGCCC  154
1953507_P1.C06   AGCCTCAGAAGTTCTTAGGGTGCATTTGGGTCTT......GCAAACTGGTTCCGCCGACACTAGGACACAGAAC-CGCAC  153
1953507_P1.C07   AGTCTCAAGATTTAGTTCTTCT..........TTGTGCTGGGTCCGCCGAGCACTAGGACAAGCACGCGAAT  142
1953507_P1.C09   ACACTCAAGACGGGAGACTTGTTTGACGGGGTCT......TTGAGCTGGTACCGCCAAGCACTAGAACAAGGACACCCAG  154
1953507_P1.D04   AGCCGCAAAACATGTGCTGCGTCTGGGG..........TTGAGCTGGTTCCGCCAAACACCAAGACAAGTACGCGCAC  148
1953507_P1.D06   AGGCGCAAAAGAGGTGAGGGGGCATGGT..........ACGGACTGGGACCGCCAAGCACTAGGAAAAGGACACGAAG  147
1953507_P1.D07   ACACGCAAAACAGTGTGTGTTTCATTTGCATATT......ACGGGCTGGGTCCGCCGAACACTAAAAC.-GGACTCGAAC  152
1953507_P1.D08   AGGCGCAGAACCGCATTTG.....CTGTACTGGTACCGCCGACGACCAAGAAAAGACACGCAAC  139
1953507_P1.E06   AGTCGCAAAACCTTCTCCTGCTTATAATTGTCGG......TCAGGCTGGTTCCGCCGACCACCAGGACAGTACTCGCCG  154
1953507_P1.E08   AGCCGCAAAACGTTGGTTGGGGATGAGGTTTCG......ACGTACTCGTTCCGCCGACACCAAGACAANANATCGACT  154
1953507_P1.E09   AGCCTCAGGATGTAGTCAGTCGCTTTGGACGATTGCTTTTGTGGCTGGTTCCGCCGAGCACTAGGACCAGTACGCGACC  160
1953507_P1.F01   AGGCGCAGGAGCTTCGACGGCT..........CTATGCTGGCACCGCCAACCACCAAGACACTAGTCGCAG  142
1953507_P1.F02   AGGCGCAAGATCTATTCGTGTT..........CCATACTGGGTCCGCCAAGCACCAGAACCAGAACACGCAG  142
1953507_P1.F04   AGCCTCAAGACGGGGTCCGATGACGGGT......ACATGCTGGCTCCGCCAAACACCAAGAGAAGGACGCGAAG  148
1953507_P1.F08   AGTCGCAGGAACGGCGCTGCAT..........GTAACCTGGGACCGCCACCACTAAAAAAAGTACACGAAT  142
1953507_P1.G01   AGGCGCAGAAGCTGCTTGGTGT..........ATAGGATGGGACCGCCGAGCACCAAGAACAGAACTCGACC  142
1953507_P1.G02   AGCCGCAGAATAGGGGACTAAG..........TCGGACTGGGACCGCCGACCACCAGGAGACAGTACTCGCAT  142
1953507_P1.G04   AGTCTCAAAAACGAGGCGTAAG..........TCGTCCTGGCTCCGCCGAACACTAGAACAGTACGCGCAG  142
1953507_P1.G06   AGTCTCAAGAGTTCCTAATGTT..........CTATCCTGGGACCGCCAACCACCAGGAGCAGTCGAAC  142
1953507_P1.G07   AGTCGCAGGACTGGGTTGTCCT..........TTAAACTGGCTCCGCCGACCACCAGGAGCAGGACACGACC  142
1953507_P1.G09   AGTCTCAGGAGATCTTTTGTCGTGTACT......CCGACCTGGTACCGCCGACCACTAAAAAAAGTACACGCCT  148
1953507_P1.H01   AGCCTCAGGATGGAGTTTTCATGCTACGTTTACT......TTATGCTGGCACCGCCGAGCACTAAGAGAAGAACTCGCCG  154
1953507_P1.H02   AGTCCAAGAGTTTGTCGTGTTGGGACG......ACGAGCTGGCTCCGCCGAACACCAAAACAGGACGCGACG  148
1953507_P1.H04   AGTCGCAAGACAGTAGGAGTCGATGTGT......GTAGGCTGGTTCCGCCGACCACCAGAAGCACTAGTCGCCT  148
1953507_P1.H06   AGGCGCAGGATTTAAGGATTGG..........TTGGCCNNNNTCCGCCAAGCACCAGGACCAGGACTCGAAC  142
1953507_P1.H08   AGGCGCAAAATCTTGGAGTTTTACGATTATGTCG......CCATGCTGGTACCGCCAACACCAGAAGGACGACGCGCCG  154
1953507_P2.A08   AGACGCAGAATATCATGTTAATGTTCCG......ACAGGCTGGCTCCGCCAACCACTAGGACCCAGGACGCGACC  148
1953507_P2.B06   AGACGCAGAGCAGTGGAATATG..........GCGTACTGGGTCCGCCGAACACCAAGAACAGGACGCGCCT  142
1953507_P2.B09   AGGCTCAGGACCGATTCCGCCTGTTCCGGATTGGGGTTTGCTGGACTGNTTCNNCCAAACACCAAGAACAGAACGCGCGT  160
1953507_P2.C02   AGACGCAGGATTGCTGGCTACGCTGTTTCATGCG......ATGGCCTGGTACCGCCGACCAC-AGGAGCAGGACTCGACC  153
1953507_P2.C07   ....................................CTGGCTCCGCCAACACCAGAACAAGTACG.C...  93
1953507_P2.E05   AGTCGCAAAAGAGGAGTTTGGGACGTTGCGTGAT......GCGGGCTGGGTCCGCCGACCACTAAGACCACTACCACGAAC  154
1953507_P2.E06   AGGCGCAAAGATGCAGTGGAAGAGAGG..........GTGAGCTGGGTCCGCCAAGCACCAGAATACGCGAAC  147
1953507_P2.E08   AGTCGCAAAATGTAGTGGGCATTTGCTTGGGTAGCGGAAGTTATGCTGGCACCGCCAACCACTAGGAGCAGAACACGCAG  160
1953507_P2.F04   AGTCGCAAAACCGGTGCAGTGT..........TTGTGCTGGTACCGCCAAACACTAGGAGCAGGACACGAAC  142
1953507_P2.F07   AGTCGCAGAAGATGGTATGGCT..........GTGGACTGGTTCCGCCGAACACCAAGACCAGGACACGAAT  142
1953507_P2.F08   AGTCTCAGGAGTTCCGGATGGT......ACAGCCTGGCACCGCCGAGCACCAAGCAAGAACTCGCCT  142
1953507_P2.G04   AGGCGCAAGACATCATTATTCGAAGGGT......CTGGGCTGGGTCCGCCAACCACTAGGACAGTACACGACT  148
1953507_P2.H05   AGCCTCAGGAGAGACTAAGAGTGGTGCTTTGCCG......ACATGCTGGCTCCGCCGAGCACCAGGAACAGGACGCGCCG  154


              170       180       190       200       210       220       230       240
         ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
16ACVQRC_195350  BSKTAGSAVNKRNVNNKNNKNNKNNKNNKNNKNNKWMCNBCRVSNCARYGRAAGRCCGATKCRYAWTKASCVSBGAC  240
1953507_P1.A03   ATATACTGGTGGCTTGGACG..................TACTTCGCGCCAATGAAAGGCCGATTTACAATTAGCCGGGAC  191
1953507_P1.A07   TGGTAGCAAGTAAGGGGTCGGGGATGAGT.........ACCATCACGACAGTGAAAGGCCGATTTACAATTAGCCGTGAC  213
1953507_P1.A08   GGGTAGCAGCGGCAACTGTGAGTTTGGCTCCTTCTTGTTCCGCGCCCAGCGAAAGGCCGATTTACAATGACCCGGAC  234
1953507_P1.B07   CCGTAGCAGGTGTCCCGTTGCGTGTGATG......ACCATCCGAGCCAGTGAAAGGCCGATTTACAATTAGCCGTGAC  225
1953507_P1.B08   CCTTAGCACTTGCGGTGCATTGTTCTCGT.........AACATCGGCACAGTGAAAGGCCGATTTACAATTAGCGCGAAC  213
1953507_P1.B09   CGGTAGCACTGACAAGTTTTTGTGTTGGG......AACGCCGCCTCAGTGAAAGGCCGATTTACAATTAGCCGTGAC  213
1953507_P1.C01   ................................................................................  125
1953507_P1.C02   CGGTAGCACGGGACCTGTGTGGGATTGCGAATTCG...TCCGTCGGGACAACGAAAGGCCGATTTACAATTAGCGGGAC  219
1953507_P1.C04   CGGTAGGAGGTGACTCGCATGGGTTTACGGAG......TACTTCGCGTCAGTGAAAGGCCGATTCACAATTAGCAGTGAC  216
1953507_P1.C05   GCGTAGCAACGAGCTCGTATTGGTTTATTCGTTCG...TCCCTCACCGCAGCGAAAGGCCGATTTACAATTAGCCCGCAC  231
1953507_P1.C06   TGGTAGGAGATATAGTTCCTATGCATATG......TCCCTCAAGGCAACGAAAGGCCGATTTACAATTAGCCGTGAC  224
1953507_P1.C07   GCGTAGGAGAGGTGATGGGGAAGGGTGAGAGT......TACATCGACGCAATGAAAGGCCGATTTACAATTAGCCGTGAC  216
1953507_P1.C09   TGGTAGGAGATGTGTCTAATCAGAATAGGATTTTGGAGACCCTCGCCACAGTGAAAGGCCGATTTACAATTACCCACGGAC  234
1953507_P1.D04   CCTTAGCAAAGACGTCGAAGCGTTTGTTGACG......TACCCCAACTCAGCGAAAGGCCGATTCACAATGACCCTGG..  220
1953507_P1.D06   TCGTAGGACTNGACGGGTCTTTCGTGGGGGGCTTGGTGTCCGTCGGCGAAGAGTGAAAGGCCGATTTACAATTAGCAGGGAC  227
1953507_P1.D07   TGTTAGGAGATTGAATGGATGGGGTATACTATT...GCGTCCGGCGGGTCAGCGAAAGGCCGATTCACATTTAGCCGTGAC  229
1953507_P1.D08   CGTTAGGACGGAAAGGCGTGTCTAAGCTG..........AACCTCGACTCAGTGAAAGGCCGATTCACATTTACCCGTGAC  209
1953507_P1.E06   CGTTAGGAGGGGTGCGGAGGGGTCAGGAGCAGGTTCTGTACTCCGGGGCAGTGAAAGGCCGATTTACAATTAGCCGTGAC  234
1953507_P1.E08   GGGGAATTAGAACGTTGCCGCCTTATTGTACG......TACAGCGCCACAACGAAAGGCCGATTTACAATGACCAGTGAC  228
1953507_P1.E09   TGGTAGCAATTAGAACCTGTTACTTGTATTGAG......TACGTCAGGACAATGAAAGGCCGATTCACAATGAGCCCGGAC  234
1953507_P1.F01   TCGTAGGAGGTGGACGGCTGCTTCGGTGGCGGCGG...TCCCTCGCGTCAGTGAAAGGCCGATTCACATTGAGCGGCGAC  219
1953507_P1.F02   TCGTAGCAAAGAAAGGGACTAAGGTTAGGGTG......ACCAGCGGGTCAATGAAAGGCCGATTTACAATTAGCCGTGAC  216
1953507_P1.F04   TCGTAGGAGCTAAATGTTTGCGTGAGTCGTCG......TCCACCGAGGCAACGAAAGGCCGATTCATAATGACCCGTGAC  222
1953507_P1.F08   CCGTAGCAGATATGAGTGCTACGTCTCTTTCGCGG...AACGGCAGCACAGTGAAAGGCCGATTCACAATGAGCACGGAC  219
1953507_P1.G01   GCTTAGGAGAGGTATCTTTGTGGCGTATCGGGCGATT...TCCCGCAAGACAGCGAAAGGCCGATTTACAATTAGCCGTGAC  219
1953507_P1.G02   TCTTAGCAAGTAACCGGGGTGTGCTTCTG.........AACGCCGGCGCAATGAAAGGCCGATTCACAATGAGCCGTGAC  213
1953507_P1.G04   CCTTAGCAGACCTGCCTGTCTTCTTACGTGTATGCCT...TCCGTCAGCACAGTGAAAGGCCGATTCACATTGAGCA-CGAC  218
1953507_P1.G06   GCGTAGCAATTGAGTATACTCAGCGTTCTACG......ACCTGCAGCCCAATGAAAGGCCGATTTACAATTAGCCGTGAC  216
1953507_P1.G07   GCGTAGGAGAGAAGCGTCAGATGTTGTCTTAT......AACTCCGACCCAGTGAAAGGCCGATTTACAATTAGCGGCCGAC  216
1953507_P1.G09   GCTTAGGAGGGAAGTCTAGGGGGTTGTTT......TACGTCACCCAGTGAAAGGCCGATTTACAATTAGCCGTGAC  219
1953507_P1.H01   TGGTAGCAACTGAACCGGTACGTCTGGCT.........AACCTCGGGTCAACGAAAGGCCGATTCACAATTACCACGGAC  225
1953507_P1.H02   GCGTAGCACCTGCAGGTTCGGCAGGGAT..........TCCTTCGGGGCAGTGAAAGGCCGATTTACAATTAGCGGCGAC  218
1953507_P1.H04   TCGTAGGAGGGGGGGTGTACGTCTTGGTCGTCT......TCCAGCACCGCAGTGAAAGGCCGATTCACAATGACCACCGAC  222
1953507_P1.H06   TGGTAGGAATGAACATTCATGCGTGGCCCTTTT......ACCGTCACGTCAGCGAAAGGCCGATTTACAATTAGCCGTGAC  216
1953507_P1.H08   CCGTAGCACGGGTGCGTGGTGGTTCTTTGTGGTTCGCGTAACGCCGGCAACGAAAGGCCGATTTACAATTAGCCGGGAC  234
1953507_P2.A08   TGTTAGGAGGGGATACCTTAGACTATGTCGAAT......AACGTCAGCGCAGTGAAAGGCCGATTCACATTTAGCCCTGAC  222
1953507_P2.B06   GCGTAGCAGTTGCGGTTACGCGTATTTGTTTGGAT...ACCTTCGACTCAACGAAAGGCCGATTTACAATTAGCCGGGAC  219
1953507_P2.B09   GGNTANGAACGGTGCCTGTGTATCCGGAGAGTTCT...ACCCCCACGACAGCGAAAGGCCGATTCACAATTANCGNNNNAC  237
1953507_P2.C02   CGGTAGGAAGTGAACTTGCTCTGTGTCGTACGGTGAATTACGGCGCAGTGAAAGGCCGATTTACAATGACCACTGAC  233
1953507_P2.C07   -....GCAGGCGTAGCACGTGACCATTATTCGGGGGAGTACCTCGACGCAACGAAAGGCCGATTTACAATTAGCCGGGAC  168
1953507_P2.E05   GGGTAGCACCTGGCAAGTTTGCGAATTTT.........ACCGTCACCCAGCGAAAGGCCGATTTACAATTAGCCCCGAC  225
1953507_P2.E06   CCGTAGCAGGGGTGAGGTGTAGGGGGACTCAT......ACCCGCCGTCAACGAAAGGCCGATTTACAATTAGCCGTGAC  221
1953507_P2.E08   TGTTAGCAGATAAATGTCTGTGGTATTGTATTACG...TCCGGCAAGGCAACGAAAGGCCGATTTACAATTAGCCGGGAC  237
1953507_P2.F04   TGGTAGGAACTAGGCGGTTGCGTGGGGTTGGTACTGGGAACGCGGTCAACGAAAGGCCGATTCATTACCCGTGAC  222
1953507_P2.F07   CGGTAGCACAGAAAGGATTATATGAATATGCGTCAT...ACCATCGGGACAATGAAAGGCCGATTCACAATGAGCGCTGAC  219
1953507_P2.F08   GCTTAGGAGACGGAGGGGGGGGTACGATTATGCCT...TCCTGCGGGGCAGTGAAAGGCCGATTCACATTGACCAGGGAC  219
1953507_P2.G04   GGTTAGGAATTGGACTTCGGGTGTGTTGTTTCGGTTGACCG-CGACCCAACGAAAGGCCGATTCACAATTACCCCGGAC  227
1953507_P2.H05   CGTTAGGAAAGGAGGTTACTACGTTTACTGCTTATAGTTCCAGCAGGTCAGCGAAAGGCCGATTCACAATGACCGGGGAC  234
```

```
                    250       260       270       280       290       300       310       320
                ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
16ACVQRC_195350 RNMGNKVRAAVCMBRRBKDMCCTGSWAMTGRVCVGCCTGRVAYYCGAKGACACGGSADTRTATKWCTGTDYARBMNNKNN 320
1953507_P1.A03  ATCGTTGAAACCATGAGTACCCTGGAACTGAACAGCCTGGAACCCGAGGACACGGCAGTATATGACTGTACAAGCATGGT 271
1953507_P1.A07  GCCGCGAGAAGCCCGGGTTCCCTGCTACTGAACAGCCTGACACCCGATGACACGGCAGTATATGACTGTGTAAGAAGTAA 293
1953507_P1.A08  GAAGGGCGAACCATGGTTTACCTGCTAATGAGCAGCCTGACACTCGAGGACACGGCAGTGTATGACTGTACAATATCGGG 314
1953507_P1.B07  ATAGGGCGAACCCCAATGTACCTGCAAATGACCAGCCTGGCATCCGAGGACACGGCAGTTATGTCTGCGAAGAGCGGTTCA 305
1953507_P1.B08  ATCGTTGGAAACAGAGTGAACCTGGTACTGAACGGCCTGGCACCCGAGGACACGGCAT...................... 271
1953507_P1.B09  AAAGCGCAAAGCCTAAGGTCCCTGCAACTGAACAGCCTGACATCCGAGGACACGGCAGTGTATGACTGTGCAATCTCTCC 293
1953507_P1.C01  .............................................................................. 125
1953507_P1.C02  ACCGAGGGAAGCAGGAGTGACCTGCTAATGACCGCCTGACACTCGAGGACACGGCAGTATATTACTGCGAAG........ 292
1953507_P1.C04  GAACAGGAAAGCAGAGTGAACCTGGTACTGACCAGCCTGAGACTCGATGACACGGCAGTGTATGACTGCGAAGAGCGTTC 296
1953507_P1.C05  AGCGCGAGAAGCCGAGCGGACCTGCAAATGACACCGCCTGGAACCCGAGGACACGGCATTGTATTACTGCGAAGAGC.... 307
1953507_P1.C06  GCAGATCAAAACACAGTGAACCTGCTACTGCCCAGCCTGACATCCGAGGACACGGGAATGTATTTCTGCGAAGAG...... 299
1953507_P1.C07  AACGGGAAAAGCCTGGC-TACCCTGCTACTGACCAGCCTGAGATCCGATGACACGGCAGTA-ATGACTGCGAAGAGCGTTC 294
1953507_P1.C09  GCAGATGGAAGCCGGGCCAACCTGTAATGGGCGGCCTGCACACCCGAGGACACGGCAGTATATGACTGTTTAATCAAGAG 314
1953507_P1.D04  ........-.....................TACTGGGCCTGGGATCCGAGGACACGGGGATTGTATTACTGCGAAGAGCGTTC 272
1953507_P1.D06  GACGGTCAAACCACAATGGACCTGGTAATGACCCGCCTGAGATTCGATGACACGGCAGTGGTCTGTGCAATCCGTGG 307
1953507_P1.D07  ACCGATGGAA-CCGGACGTCCCTGCAACTGAACAGCCTGGGATTCCGATGACACGGCAGTATATGACTGTACAGTCGTGTT 308
1953507_P1.D08  AGAGGGCGAAACCTAGTGTCCCTGGTAATGAACCGCCTGGAATTCGATGACACGGCAGTGTATGACTGTATAAGAACTTT 289
1953507_P1.E06  ATAGGGTAAAAGCCTAATGAACCTGCAACTGACCCGCCTGAGACTCGAGGACACGGCAATATATGTCTGCGAAG....... 307
1953507_P1.E08  GGAGTGGGAAAGCCCAACTGCCCTGGAAATGGACAGCCTGGGACTCGATGACACGGCAATGTATGTCTGCGAA........ 300
1953507_P1.E09  AAAGCTAGAACCAGGGTGTCCCTGGAAATGAAGGGCCTGGACTCGAGGACACGGGAATATATGACTGTATAGTCGATCA 314
1953507_P1.F01  GAAGCGAGAAACCCGACGGACCTGGAACTCGGCCGGCCTGAGATCCGAGGACACGGCAGTATATGTCTGTTAGGCATTTT 299
1953507_P1.F02  AGAGGGCGAAACCTAGTGTCCCTGGTAATGAACCGCCTGGAATTCGATGACACGGCAGTATATGACTGTATAAGAACTTT 296
1953507_P1.F04  AGCGTTCAAAACACAGTGACCCTGGAACTGACCCGCCTGGAACTCGAGGACACGGCATTATATGTCTGCGAAGAG..... 297
1953507_P1.F08  AAAGCTGGAACCCGAGTTTAACCTGCTAATGAGCAGCCTGGCACCCGATGACACGGCAATGTATTACTGTTAGGAGGGAG 299
1953507_P1.G01  GCCGATCGAAGCAGAAGGACCCTGGAACTGAACCGCCTGAGACCCGATGACACGGCATTGTATGACTGTGCTCTATTC 299
1953507_P1.G02  GGCGCTGAAACCACGAGCGAACCTGCAACTGGGCAGCCTGGGATCCGAGGACACGGGAATGTATGTCTGTGCAACCTCGTT 293
1953507_P1.G04  ACCGCGCGAAACCCAACTGACCTGC-AATGAGCAGCCTGCAAACCCGAGGACACGGCAGTATATTACTGCGAAGAGCGTTC 297
1953507_P1.G06  ATCGTGCGAACCCGAGCTTCCCTGGAAATGACCGGCCTGCACACCCGAGGACACGGCATTATATGTCTGCGAAGAGCGTTC 296
1953507_P1.G07  GCCGAGAGAACCAGAATTTCCCTGCTAATGACCAGCCTGGAATCCGAGGACACGGCAATGTATGACTGTGCAATCTGGTC 296
1953507_P1.G09  GCCGGGAAAAGCAGGAGTGCCCCTGCAACTGAACGGCCTGCACATTCGAGGACACGGCAGTATGTATTTAGTATTGGG 299
1953507_P1.H01  ATCCGCGCAACCACGGTTTCCCTGGTACTGGGCGCCGCCTGGAATCCGAGGACACGGGAATGTATGACTGC.......... 294
1953507_P1.H02  GTACCGAAAACCATGACGTACCTGGTAATGAACCGCCTGAGATCGATGACACGGCAGTATTTTCTGCGAAGAG..... 293
1953507_P1.H04  GTAGCTCGAAGCCTGGTGTACCTGCTACTGACCCGTGCTACATTCGAGGACACGGGAATGTATGTCTGCGAAGA...... 296
1953507_P1.H06  ATCGGGCAAACCAGGAGTGACCTGCAACTGAGCCGCCTGGAATCGATGACACGGGAGTGTATGTCTGTGCAGGACTTGC 296
1953507_P1.H08  GTAGTGCGAA-CCCAACTACCCTGCAAATGGGCGGACCTGG-ATCCGAGGACACGGCAGTGTATG-CTGTACAGCCTTGTC 311
1953507_P2.A08  ACCGAGCAAAACAGGGGCGAACCTGGAACTGGCCAGCCTGGCATTCGATGACACGGCATTGTATGACTGTGCAGCCGTTC 302
1953507_P2.B06  GGCGCTCGAAACCCAGGTAACCTGGTACTGAGCGGCCTGAAATCGATGACACGGCAGTATATGACTGTACAGCGTGGG 299
1953507_P2.B09  NGANANGAAAACCCAAGCGAACTGGAAATGGNCCGCCTGGCACTCNATGACACGNGAATTATATGTCTGTGTANNATTGTT 317
1953507_P2.C02  GACCATCAAAGCACAACCGGCCCTGGTTCTGACCGGCCTGACATCCGAGGACACGGGAATGTATGACTGCGAAGAG...... 308
1953507_P2.C07  GGCGGTAGAAGCACAAGTAACCTGCAACTGAGCCGGCCTGGCATCCGAGGACACGGCATTGTATGACTGTTCAATCGAGCT 248
1953507_P2.E05  ATCGCGAGAAGCCCGGCCGAACCTGGAACTGACCCGCCTGAA-TCCGAGGACACGGCA-TTTATTTCTGTGCAGTACGGTG 303
1953507_P2.E06  GCAGGTGGAACCCGGGCTGCCCTGCAACTGAGCAGCCTGGAACTCGAGGACACGGCAGTATATGACTGTACAGCAGAGCT 301
1953507_P2.E08  ATAGCATAGAACCCGAATCGTACCTGCAAATGGCCGGCCTGGCATCCGAGGACACGGCAATGTATGTCTGTACAATATTTCG 317
1953507_P2.F04  AGAGGGCGAAACCTAGTGTCCCTGGTAATGAACCGCCTGGAATCCGATGACACGGCAGTATATGACTGTATAAGAACTTT 302
1953507_P2.F07  GCAGGTAAAACCACGACCAACCTGCAAATGGGCCGCCTGAGATCCGAGGACACGGCAGTATATGACTGTGCAGTAGATAA 299
1953507_P2.F08  AAAGGGAGAAGCACGATGTCCCTGCAACTGAACCGCCTGAAATCCGAGGACACGGGAATGTATGACTGTGTAGTAGGTGC 299
1953507_P2.G04  AAAGCGGGAAACCGGGCTGAACCTGCAAATGAGCAGCCTGGGACACGGCAGGTGTATGTCTGTAGCCGAGGG 307
1953507_P2.H05  ACAGGGAGAAGCCTGGCGGCCCTGGTACTGGCCAGCCTGACACCCGATGACACGGGAGTGTATTACTGTATAGCCTTGCG 314
```

```
                    330       340       350       360       370       380       390       400
                ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
16ACVQRC_195350 KNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKAKMRGCMVAGGGACCC 400
1953507_P1.A03  GGTCAGTGTTAGAGTACGGCTAATGCGTTGCTTTAGGCG.....................AGCAGCAGAGGGACCC 326
1953507_P1.A07  GCCGATGGCGATTCCT...............................................ATCAGCAGAGGGACCC 325
1953507_P1.A08  GTTGGTTGGTATGGTGAGGTATCCTTGGAAGGTTTAGATTTTATGACG...............ATAGGCACAGGGACCC 379
1953507_P1.B07  GCGGAATTTTTCTCCTGTGGGT.........................................ATCGGGCCAGGGACCC 343
1953507_P1.B08  ..........................................................TATATGACTGCGAAGAGCGTTCCC 295
1953507_P1.B09  TCCGAAGGGTGCTCAGGTG...........................................AGAAGCCAGGGACCC 328
1953507_P1.C01  ...........................................................AGAGGCAGAGGGACCC 141
1953507_P1.C02  ..........................................................-...AGCGTTCCCC 302
1953507_P1.C04  TCAGTGG.......................................................ATCAGCACAGGGACCC 319
1953507_P1.C05  ..........................................................GTTCGCAGAGGGACCC 323
1953507_P1.C06  ..........................................................-...CGTTCGACCC 309
1953507_P1.C07  GCAT..........................................................ATCGGGCCAGGGACCC 314
1953507_P1.C09  TCCG.........................................................ATAGGCCGAGGGACCC 334
1953507_P1.D04  CTGTGTAGGACATACGGCTAGGCGGCCG..................................AGCGGCAGAGGGACCC 316
1953507_P1.D06  TAAGGGGGGTTTGTTTCGG...........................................ATCAGCAGAGGGACCC 342
1953507_P1.D07  TATGCTGTACTATGATGACGATGGTGGATAATTCT...........................AGCAGCCAAGGGACCC 359
1953507_P1.D08  GACTACGGTTGCGTTGATGATGTAGCATACTGAGCGGTATCGGCTG................AGAGGCAGAGGGACCC 351
1953507_P1.E06  ...........................................................-..AGCGTTCACCC 319
1953507_P1.E08  ...........................................................-.GAGCGTTCACCC 312
1953507_P1.E09  GGCTCCGTTG.....................................................ATAGGCAAGGGACCC 340
1953507_P1.F01  GTCGCAGGATAAGCTTGAGTCTTCTCCTTTGTTTGAGTATCCT..................AGAAGCCAGGGACCC 358
1953507_P1.F02  GACTACGGTTGCGTTGATGATGTAGCATACTGAGCGGTATCGGCTG................AGAGGCAGAGGGACCC 358
1953507_P1.F04  ..........................................................CGTTCCAGGGACCC 313
1953507_P1.F08  TAAGGGTTGATATGAAG.............................................ATCGGCAGAGGGACCC 331
1953507_P1.G01  TCTTTGTCCGCGGAAG..............................................AGAAGCACAGGGACCC 331
1953507_P1.G02  GCTTCGTTCGGTGTCTAATTGT.........................................AGCGGCCAGGGACCC 331
1953507_P1.G04  GTACG.........................................................AGCGGCAGAGGGACCC 318
1953507_P1.G06  GGTTATGCAT.....................................................AGAGGCCCAGGGACCC 322
1953507_P1.G07  GTTT...........................................................AGCAGCACAGGGACCC 316
1953507_P1.G09  CGG............................................................ATCAGCCAGGGACCC 318
1953507_P1.H01  .........................................................-...GAAGAGCGTT 304
1953507_P1.H02  ..........................................................CGTTCCCGAGGGACCC 309
1953507_P1.H04  ..........................................................-..GCGTTCACCC 307
1953507_P1.H06  TACTCGTAGTAGGAGTGGGATGTCG.....................................AGCAGCGAGGGACCC 337
1953507_P1.H08  GAAGATTATGAGTACTGGTACTGCT.....................................ATAAGCCCAGGGACCC 352
1953507_P2.A08  GTGGAGGTTGCGGCGCTTGCGATTGGTTTGTATTAGGATTCGTATTTGCGT...........ATAGGCGAGGGACCC 369
1953507_P2.B06  TCGT..........................................................ATCGGGCCAGGGACCC 319
1953507_P2.B09  GCGTAATTGGNATATTGACAGCAGCATGATANTAGGNNNANGATGTTTACG...........ATCGGNCCAGGNACCC 382
1953507_P2.C02  ..........................................................CGTTCCCAAGGGACCC 324
1953507_P2.C07  GCATATTATTCGGACGTCT...........................................ATAGGCCCAGGGACCC 283
1953507_P2.E05  GAGGACGTTTGCTTTGATGCCGTTGGAT..................................AGAAGCCAAGGGACCC 347
1953507_P2.E06  TATGTTGGTTAGTGTGACTGGTATGAATAATGCG............................AGAGGCACAGGGACCC 351
1953507_P2.E08  GTCTGATTTG....................................................AGAAGCCAGGGACCC 343
1953507_P2.F04  GACTACGGTTGCGTTGATGATGTAGCATACTGAGCGGTATCGGCTG................AGAGGCAGAGGGACCC 364
1953507_P2.F07  GACGAGGGCGTGTCCTGGTTGGTCGTGGTATTTTTCG.........................ATCAGCCCAGGGACCC 352
1953507_P2.F08  TAGGGCGCATACTCGTGAGCAGGGGACT..................................AGCGGCCGAGGGACCC 343
1953507_P2.G04  TGTTGCGGGGGGGTTATATGTCGTGGCTGTTTTCG...........................ATAGGCAGAGGGACCC 357
1953507_P2.H05  GAGTGGTACGTTTTGTCAGTCGGTT.....................................AGCAGCAGAGGGACCC 355
```

```
                                     410        420
                                 ....|....|....|....|.
       16ACVQRC_195350           WAGTCACCGTCTCCTCACACC 421
       1953507_P1.A03            AAGTCACCGTCTCCTCACACC 347
       1953507_P1.A07            TAGTCACCGTCTCCTCACACC 346
       1953507_P1.A08            TAGTCACCGTCTCCTCACACC 400
       1953507_P1.B07            AAGTCACCGTCTCCTCACACC 364
       1953507_P1.B08            TAGTCACCGTCTCCTCACACC 316
       1953507_P1.B09            TAGTCACCGTCTCCTCACACC 349
       1953507_P1.C01            TAGTCACCGTCTCCTCACACC 162
       1953507_P1.C02            AAGTCACCGTCTCCTCACACC 323
       1953507_P1.C04            TAGTCACCGTCTCCTCACACC 340
       1953507_P1.C05            TAGTCACCGTCTCCTCACACC 344
       1953507_P1.C06            TAGTCACCGTCTCCTCACACC 330
       1953507_P1.C07            AAGTCACCGTCTCCTCACACC 335
       1953507_P1.C09            TAGTCACCGTCTCCTCACACC 355
       1953507_P1.D04            AAGTCACCGTCTCCTCACACC 337
       1953507_P1.D06            AAGTCACCGTCTCCTCACACC 363
       1953507_P1.D07            AAGTCACCGTCTCCTCACACC 380
       1953507_P1.D08            TAGTCACCGTCTCCTCACACC 372
       1953507_P1.E06            TAGTCACCGTCTCCTCACACC 340
       1953507_P1.E08            AAGTCACCGTCTCCTCACACC 333
       1953507_P1.E09            TAGTCACCGTCTCCTCACACC 361
       1953507_P1.F01            AAGTCACCGTCTCCTCACACC 379
       1953507_P1.F02            TAGTCACCGTCTCCTCACACC 379
       1953507_P1.F04            TAGTCACCGTCTCCTCACACC 334
       1953507_P1.F08            TAGTCACCGTCTCCTCACACC 352
       1953507_P1.G01            TAGTCACCGTCTCCTCACACC 352
       1953507_P1.G02            AAGTCACCGTCTCCTCACACC 352
       1953507_P1.G04            AAGTCACCGTCTCCTCACACC 339
       1953507_P1.G06            AAGTCACCGTCTCCTCACACC 343
       1953507_P1.G07            AAGTCACCGTCTCCTCACACC 337
       1953507_P1.G09            TAGTCACCGTCTCCTCACACC 339
       1953507_P1.H01            CAGTCACCGTCTCCTCACACC 325
       1953507_P1.H02            AAGTCACCGTCTCCTCACACC 330
       1953507_P1.H04            TAGTCACCGTCTCCTCACACC 328
       1953507_P1.H06            TAGTCACCGTCTCCTCACACC 358
       1953507_P1.H08            AAGTCACCGTCTCCTCACACC 373
       1953507_P2.A08            TAGTCACCGTCTCCTCACACC 390
       1953507_P2.B06            AAGTCACCGTCTCCTCACACC 340
       1953507_P2.B09            AAGTCNNCGTCTCCTCACACC 403
       1953507_P2.C02            TAGTCACCGTCTCCTCACACC 345
       1953507_P2.C07            AAGTCACCGTCTCCTCACACC 304
       1953507_P2.E05            TAGTCACCGTCTCCTCACACC 368
       1953507_P2.E06            TAGTCACCGTCTCCTCACACC 372
       1953507_P2.E08            TAGTCACCGTCTCCTCACACC 364
       1953507_P2.F04            TAGTCACCGTCTCCTCACACC 385
       1953507_P2.F07            AAGTCACCGTCTCCTCACACC 373
       1953507_P2.F08            AAGTCACCGTCTCCTCACACC 364
       1953507_P2.G04            TAGTCACCGTCTCCTCACACC 378
       1953507_P2.H05            TAGTCACCGTCTCCTCACACC 376
```

Figure 6.9 Library documentation: After synthesis of the c-lib by GeneArt (Thermo Fisher), 48 total clones were sequenced, 30 of which yielded full-length sequences.

7. References

Anzai, H., Terai, T., Jayathilake, C., Suzuki, T., & Nemoto, N. (2019). A novel immuno-PCR method using cDNA display. *Analytical Biochemistry*, *578*, 1–6. https://doi.org/10.1016/j.ab.2019.04.017

Arbabi Ghahroudi, M., Desmyter, A., Wyns, L., Hamers, R., & Muyldermans, S. (1997). Selection and identification of single domain antibody fragments from camel heavy-chain antibodies. *FEBS Letters*, *414*(3), 521–526. https://doi.org/10.1016/S0014-5793(97)01062-4

Armstrong, L. A., Lange, S. M., Cesare, V. D., Matthews, S. P., Nirujogi, R. S., Cole, I., Hope, A., Cunningham, F., Toth, R., Mukherjee, R., Bojkova, D., Gruber, F., Gray, D., Wyatt, P. G., Cinatl, J., Dikic, I., Davies, P., & Kulathu, Y. (2021). Biochemical characterization of protease activity of Nsp3 from SARS-CoV-2 and its inhibition by nanobodies. *PLOS ONE*, *16*(7), e0253364. https://doi.org/10.1371/JOURNAL.PONE.0253364

Baneyx, F. (1999). Recombinant protein expression in Escherichia coli. *Current Opinion in Biotechnology*, *10*(5), 411–421. https://doi.org/10.1016/S0958-1669(99)00003-8

Baneyx, F., & Mujacic, M. (2004). Recombinant protein folding and misfolding in Escherichia coli. *Nature Biotechnology 2004 22:11*, *22*(11), 1399–1408. https://doi.org/10.1038/nbt1029

Benhar, I., Azriel, R., Nahary, L., Shaky, S., Berdichevsky, Y., Tamarkin, A., & Wels, W. (2000). Highly efficient selection of phage antibodies mediated by display of antigen as Lpp-OmpA′ fusions on live bacteria 1 1Edited by J. Wells. *Journal of Molecular Biology*, *301*(4), 893–904. https://doi.org/10.1006/jmbi.2000.4021

Bernstein, P. S., Law, W. C., & Rando, R. R. (1987). Isomerization of all-trans-retinoids to 11-cis-retinoids in vitro. *Proceedings of the National Academy of Sciences*, *84*(7), 1849–1853. https://doi.org/10.1073/pnas.84.7.1849

Boder, E. T., Raeeszadeh-Sarmazdeh, M., & Price, J. V. (2012). Engineering antibodies by yeast display. *Archives of Biochemistry and Biophysics*, *526*(2), 99–106. https://doi.org/10.1016/j.abb.2012.03.009

Boder, E. T., & Wittrup, K. D. (1997). Yeast surface display for screening combinatorial polypeptide libraries. *Nature Biotechnology*, *15*(6), 553–557. https://doi.org/10.1038/nbt0697-553

Boye, S. E., Boye, S. L., Pang, J., Ryals, R., Everhart, D., Umino, Y., Neeley, A. W., Besharse, J., Barlow, R., & Hauswirth, W. W. (2010). Functional and Behavioral Restoration of Vision by Gene Therapy in the Guanylate Cyclase-1 (GC1) Knockout Mouse. *PLoS ONE*, *5*(6), e11306. https://doi.org/10.1371/journal.pone.0011306

Brack, C., Hirama, M., Lenhard-Schuller, R., & Tonegawa, S. (1978). A complete immunoglobulin gene is created by somatic recombination. *Cell*, *15*(1), 1–14. https://doi.org/10.1016/0092-8674(78)90078-8

Braun, M. B., Traenkle, B., Koch, P. A., Emele, F., Weiss, F., Poetz, O., Stehle, T., & Rothbauer, U. (2016). Peptides in headlock – a novel high-affinity and versatile peptide-binding nanobody for proteomics and microscopy. *Scientific Reports 2016 6:1*, *6*(1), 1–10. https://doi.org/10.1038/srep19211

Brecelj, J., & Stirn-Kranjc, B. (1999). ERG and VEP follow-up study in children with Leber's congenital amaurosis. *Eye*, *13*(1), 47–54. https://doi.org/10.1038/eye.1999.10

Brodsky, J. L., & McCracken, A. A. (1999). ER protein quality control and proteasome-mediated protein degradation. *Seminars in Cell & Developmental Biology*, *10*(5), 507–513. https://doi.org/10.1006/SCDB.1999.0321

Bruckert, F., Chabre, M., & Vuong, T. M. (1992). Kinetic analysis of the activation of transducin by photoexcited rhodopsin. Influence of the lateral diffusion of transducin and competition of guanosine diphosphate and guanosine triphosphate for the nucleotide site. *Biophysical Journal*, *63*(3), 616–629. https://doi.org/10.1016/S0006-3495(92)81650-8

Callahan, B. P., Stanger, M. J., & Belfort, M. (2010). Protease Activation of Split Green Fluorescent Protein. *ChemBioChem*, *11*(16), 2259–2263. https://doi.org/10.1002/cbic.201000453

Cannalire, R., Cerchia, C., Beccari, A. R., Di Leva, F. S., & Summa, V. (2022). Targeting SARS-CoV-2 Proteases and Polymerase for COVID-19 Treatment: State of the Art and Future Opportunities. *Journal of Medicinal Chemistry*, *65*(4), 2716–2746. https://doi.org/10.1021/acs.jmedchem.0c01140

Carlson, L.-A., & Hurley, J. H. (2012). In vitro reconstitution of the ordered assembly of the endosomal sorting complex required for transport at membrane-bound HIV-1 Gag clusters. *Proceedings of the National Academy of Sciences*, *109*(42), 16928–16933. https://doi.org/10.1073/pnas.1211759109

Casadaban, M. J., & Cohen, S. N. (1980). Analysis of gene control signals by DNA fusion and cloning in Escherichia coli. *Journal of Molecular Biology*, *138*(2), 179–207. https://doi.org/10.1016/0022-2836(80)90283-1

Cha, R. S., & Thilly, W. G. (1993). Specificity, efficiency, and fidelity of PCR. *Genome Research*, *3*(3), S18–S29. https://doi.org/10.1101/gr.3.3.S18

Chen, C., Nakatani, K., & Koutalos, Y. (2003). Free Magnesium Concentration in Salamander Photoreceptor Outer Segments. *The Journal of Physiology*, *553*(1), 125–135. https://doi.org/10.1113/jphysiol.2003.053280

Chiu, W., Lin, T.-Y., Chang, Y.-C., Isahwan-Ahmad Mulyadi Lai, H., Lin, S.-C., Ma, C., Yarmishyn, A. A., Lin, S.-C., Chang, K.-J., Chou, Y.-B., Hsu, C.-C., Lin, T.-C., Chen, S.-J., Chien, Y., Yang, Y.-P., & Hwang, D.-K. (2021). An Update on Gene Therapy for Inherited Retinal Dystrophy: Experience in Leber Congenital Amaurosis Clinical Trials. *International Journal of Molecular Sciences*, *22*(9), 4534. https://doi.org/10.3390/ijms22094534

Choi, H. L., Yang, H. R., Shin, H. G., Hwang, K., Kim, J. W., Lee, J. H., Ryu, T., Jung, Y., & Lee, S. (2023). Generation and Next-Generation Sequencing-Based Characterization of a Large Human Combinatorial Antibody Library. *International Journal of Molecular Sciences*, *24*(6), 6011. https://doi.org/10.3390/ijms24066011

Contreras, M. A., Serrano-Rivero, Y., González-Pose, A., Salazar-Uribe, J., Rubio-Carrasquilla, M., Soares-Alves, M., Parra, N. C., Camacho-Casanova, F., Sánchez-Ramos, O., & Moreno, E. (2023). Design and Construction of a Synthetic Nanobody Library: Testing Its Potential with a Single Selection Round Strategy. *Molecules*, *28*(9), 3708. https://doi.org/10.3390/molecules28093708

Damhof, R. A., Feijlbrief, M., Welling-Wester, S., & Welling, G. W. (1994). Purification of the integral membrane glycoproteins D of Herpes simplex virus types 1 and 2, produced in the recombinant baculovirus expression system, by ion-exchange high-performance liquid chromatography. *Journal of Chromatography A*, *676*(1), 43–49. https://doi.org/10.1016/0021-9673(94)80454-0

D'Astolfo, D. S., Pagliero, R. J., Pras, A., Karthaus, W. R., Clevers, H., Prasad, V., Lebbink, R. J., Rehmann, H., & Geijsen, N. (2015). Efficient Intracellular Delivery of Native Proteins. *Cell*, *161*(3), 674–690. https://doi.org/10.1016/j.cell.2015.03.028

Davies, D. R., & Chacko, S. (1993). Antibody Structure. *Accounts of Chemical Research*, *26*(8), 421–427. https://doi.org/10.1021/AR00032A005

De Genst, E., Silence, K., Decanniere, K., Conrath, K., Loris, R., Kinne, J., Muyldermans, S., & Wyns, L. (2006). Molecular basis for the preferential cleft recognition by dromedary heavy-chain antibodies. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(12), 4586–4591. https://doi.org/10.1073/PNAS.0505379103/SUPPL_FILE/05379TABLE_4.DOC

Dell'Orco, D., Behnen, P., Linse, S., & Koch, K.-W. (2010). Calcium binding, structural stability and guanylate cyclase activation in GCAP1 variants associated with human cone dystrophy. *Cellular and Molecular Life Sciences*, *67*(6), 973–984. https://doi.org/10.1007/s00018-009-0243-8

Derda, R., Tang, S., Li, S. C., Ng, S., Matochko, W., & Jafari, M. (2011). Diversity of Phage-Displayed Libraries of Peptides during Panning and Amplification. *Molecules*, *16*(2), 1776–1803. https://doi.org/10.3390/molecules16021776

Detwiler, S. R., & Laurens, H. (1921). Studies on the retina. Histogenesis of the visual cells in amblystoma. *The Journal of Comparative Neurology*, *33*(5), 493–508. https://doi.org/10.1002/cne.900330503

Devlin, J. J., Panganiban, L. C., & Devlin, P. E. (1990). Random peptide libraries: a source of specific protein binding molecules. *Science (New York, N.Y.)*, *249*(4967), 404–406. https://doi.org/10.1126/science.2143033

Dharmaraj, S., Silva, E., Pina, A. L., Li, Y. Y., Yang, J.-M., Carter, R. C., Loyer, M., El-Hilali, H., Traboulsi, E., Sundin, O., Zhu, D., Koenekoop, R. K., & Maumenee, I. H. (2000). Mutational analysis and clinical correlation in Leber congenital amaurosis. *Ophthalmic Genetics*, *21*(3), 135–150. https://doi.org/10.1076/1381-6810(200009)2131-ZFT135

Dingus, J. G., Tang, J. C., Amamoto, R., Wallick, G. K., & Cepko, C. L. (2022). A general approach for stabilizing nanobodies for intracellular expression. *ELife*, *11*. https://doi.org/10.7554/eLife.68253

Dizhoor, A. M., Lowe, D. G., Olshevskaya, E. V., Laura, R. P., & Hurley, J. B. (1994). The human photoreceptor membrane guanylyl cyclase, RetGC, is present in outer segments and is regulated by calcium and a soluble activator. *Neuron*, *12*(6), 1345–1352. https://doi.org/10.1016/0896-6273(94)90449-9

Dizhoor, A. M., Olshevskaya, E. V., & Peshenko, I. V. (2016). The R838S Mutation in Retinal Guanylyl Cyclase 1 (RetGC1) Alters Calcium Sensitivity of cGMP Synthesis in the Retina and Causes Blindness in Transgenic Mice. *Journal of Biological Chemistry*, *291*(47), 24504–24516. https://doi.org/10.1074/jbc.M116.755553

Doshi, R., Chen, B. R., Vibat, C. R. T., Huang, N., Lee, C. W., & Chang, G. (2014). In vitro nanobody discovery for integral membrane protein targets. *Scientific Reports*. https://doi.org/10.1038/srep06760

Duda, T., Pertzev, A., Makino, C. L., & Sharma, R. K. (2016). Bicarbonate and Ca2+ Sensing Modulators Activate Photoreceptor ROS-GC1 Synergistically. *Frontiers in Molecular Neuroscience*, *9*. https://doi.org/10.3389/fnmol.2016.00005

Dumoulin, M., Conrath, K., Meirhaeghe, A. Van, Meersman, F., Heremans, K., Frenken, L. G. J., Muyldermans, S., Wyns, L., & Matagne, A. (2002). Single-domain antibody fragments with high conformational stability. *Protein Science*, *11*(3), 500–515. https://doi.org/10.1110/PS.34602

Ehrlich, P. (1900). Croonian lecture.—On immunity with special reference to cell life. *Proceedings of the Royal Society of London*, *66*(424–433), 424–448. https://doi.org/10.1098/rspl.1899.0121

Elliott, M. H., Fliesler, S. J., & Ghalayini, A. J. (2003). Cholesterol-Dependent Association of Caveolin-1 with the Transducin α Subunit in Bovine Photoreceptor Rod Outer Segments: Disruption by Cyclodextrin and Guanosine 5'- *O* -(3-Thiotriphosphate). *Biochemistry*, *42*(26), 7892–7903. https://doi.org/10.1021/bi027162n

Engvall, E., & Perlmann, P. (1971). Enzyme-linked immunosorbent assay (ELISA) quantitative assay of immunoglobulin G. *Immunochemistry*, *8*(9), 871–874. https://doi.org/10.1016/0019-2791(71)90454-X

Farrens, D. L., Altenbach, C., Yang, K., Hubbell, W. L., & Khorana, H. G. (1996). Requirement of Rigid-Body Motion of Transmembrane Helices for Light Activation of Rhodopsin. *Science*, *274*(5288), 768–770. https://doi.org/10.1126/science.274.5288.768

Fiedler, K., Kobayashi, T., Kurzchalia, T. V, & Simons, K. (1993). Glycosphingolipid-Enriched, Detergent-Insoluble Complexes in Protein Sorting in Epithelial Cells1. *Biochemistry*, *32*, 6365–6373. https://pubs.acs.org/sharingguidelines

Flower, A. M., & McHenry, C. S. (1990). The gamma subunit of DNA polymerase III holoenzyme of Escherichia coli is produced by ribosomal frameshifting. *Proceedings of the National Academy of Sciences*, *87*(10), 3713–3717. https://doi.org/10.1073/pnas.87.10.3713

Frydman, J. (2003). Folding of Newly Translated Proteins In Vivo: The Role of Molecular Chaperones. *Https://Doi.Org/10.1146/Annurev.Biochem.70.1.603*, *70*, 603–648. https://doi.org/10.1146/ANNUREV.BIOCHEM.70.1.603

Fung, B. K. K., Hurley, J. B., & Stryer, L. (1981). Flow of information in the light-triggered cyclic nucleotide cascade of vision. *Proceedings of the National Academy of Sciences of the United States of America*, *78*(1 II), 152–156. https://doi.org/10.1073/PNAS.78.1.152

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., & Bairoch, A. (2005). Protein Identification and Analysis Tools on the ExPASy Server. In *The Proteomics Protocols Handbook* (pp. 571–607). Humana Press. https://doi.org/10.1385/1-59259-890-0:571

Genovese, F., Reisert, J., & Kefalov, V. J. (2021). Sensory Transduction in Photoreceptors and Olfactory Sensory Neurons: Common Features and Distinct Characteristics. *Frontiers in Cellular Neuroscience*, *15*, 414. https://doi.org/10.3389/FNCEL.2021.761416/BIBTEX

Gorczyca, W. A., Gray-Keller, M. P., Detwiler, P. B., & Palczewski, K. (1994). Purification and physiological evaluation of a guanylate cyclase activating protein from retinal rods. *Proceedings of the National Academy of Sciences*, *91*(9), 4014–4018. https://doi.org/10.1073/pnas.91.9.4014

Gottesman, S., Wickner, S., & Maurizi, M. R. (1997). Protein quality control: triage by chaperones and proteases. *Genes & Development*, *11*(7), 815–823. https://doi.org/10.1101/gad.11.7.815

Götzke, H., Kilisch, M., Martínez-Carranza, M., Sograte-Idrissi, S., Rajavel, A., Schlichthaerle, T., Engels, N., Jungmann, R., Stenmark, P., Opazo, F., & Frey, S. (2019). The ALFA-tag is a highly versatile tool for nanobody-based bioscience applications. *Nature Communications 2019 10:1*, *10*(1), 1–12. https://doi.org/10.1038/s41467-019-12301-7

Goyal, B., & Goyal, D. (2020). Targeting the Dimerization of the Main Protease of Coronaviruses: A Potential Broad-Spectrum Therapeutic Strategy. *ACS Combinatorial Science*, *22*(6), 297–305. https://doi.org/10.1021/acscombsci.0c00058

Gray-Keller, M. P., & Detwiler, P. B. (1994). The calcium feedback signal in the phototransduction cascade of vertebrate rods. *Neuron*, *13*(4), 849–861. https://doi.org/10.1016/0896-6273(94)90251-8

Greenberg, A. S., Avila, D., Hughes, M., Hughes, A., McKinney, E. C., & Flajnik, M. F. (1995). A new antigen receptor gene family that undergoes rearrangement and extensive somatic diversification in sharks. *Nature*, *374*(6518), 168–173. https://doi.org/10.1038/374168a0

Grosjean, H., & Fiers, W. (1982). Preferential codon usage in prokaryotic genes: the optimal codon-anticodon interaction energy and the selective codon usage in efficiently expressed genes. *Gene*, *18*(3), 199–209. https://doi.org/10.1016/0378-1119(82)90157-3

Gross, O. P., & Burns, M. E. (2010). Control of Rhodopsin's Active Lifetime by Arrestin-1 Expression in Mammalian Rods. *The Journal of Neuroscience*, *30*(9), 3450–3457. https://doi.org/10.1523/JNEUROSCI.5391-09.2010

Hagins, W. A., Penn, R. D., & Yoshikami, S. (1970). Dark Current and Photocurrent in Retinal Rods. *Biophysical Journal*, *10*(5), 380–412. https://doi.org/10.1016/S0006-3495(70)86308-1

Hamers-Casterman, C., Atarhouch, T., Muyldermans, S., Robinson, G., Hammers, C., Songa, E. B., Bendahman, N., & Hammers, R. (1993). Naturally occurring antibodies devoid of light chains. *Nature*, *363*(6428), 446–448. https://doi.org/10.1038/363446a0

Hanes, J., & Plückthun, A. (1997). *In vitro* selection and evolution of functional proteins by using ribosome display. *Proceedings of the National Academy of Sciences*, *94*(10), 4937–4942. https://doi.org/10.1073/pnas.94.10.4937

Hartman, M. C. T., Josephson, K., Lin, C. W., & Szostak, J. W. (2007). An Expanded Set of Amino Acid Analogs for the Ribosomal Translation of Unnatural Peptides. *PLOS ONE*, *2*(10), e972. https://doi.org/10.1371/JOURNAL.PONE.0000972

Hasbold, J., & Klaus, G. G. B. (1990). Anti-immunoglobulin antibodies induce apoptosis in immature B cell lymphomas. *European Journal of Immunology*, *20*(8), 1685–1690. https://doi.org/10.1002/eji.1830200810

Hazzard, W. R., Crockford, P. M., Buchanan, K. D., Vance, J. E., Chen, R., & Williams, R. H. (1968). A Double Antibody Immunoassay for Glucagon. *Diabetes*, *17*(4), 179–186. https://doi.org/10.2337/diab.17.4.179

Holm, C., Meeks-Wagner, D. W., Fangman, W. L., & Botstein, D. (1986). A rapid, efficient method for isolating DNA from yeast. *Gene*, *42*(2), 169–173. https://doi.org/10.1016/0378-1119(86)90293-3

Hwang, J.-Y., & Koch, K.-W. (2002). Calcium- and Myristoyl-Dependent Properties of Guanylate Cyclase-Activating Protein-1 and Protein-2. *Biochemistry*, *41*(43), 13021–13028. https://doi.org/10.1021/bi026618y

Indrawattana, N., Sookrung, N., Kulkeaw, K., Seesuay, W., Kongngoen, T., Chongsa-nguan, M., Tungtrongchitr, A., & Chaicumpa, W. (2010). Human monoclonal ScFv that inhibits cellular entry and metalloprotease activity of tetanus neurotoxin. *Asian Pacific Journal of Allergy and Immunology*, *28*(1), 85–93.

Ishizawa, T., Kawakami, T., Reid, P. C., & Murakami, H. (2013). TRAP Display: A High-Speed Selection Method for the Generation of Functional Polypeptides. *Journal of the American Chemical Society*, *135*(14), 5433–5440. https://doi.org/10.1021/ja312579u

Ja, W. W., Olsen, B. N., & Roberts, R. W. (2005). Epitope mapping using mRNA display and a unidirectional nested deletion library. *Protein Engineering, Design and Selection*, *18*(7), 309–319. https://doi.org/10.1093/PROTEIN/GZI038

Jamshad, M., Lin, Y.-P., Knowles, T. J., Parslow, R. A., Harris, C., Wheatley, M., Poyner, D. R., Bill, R. M., Thomas, O. R. T., Overduin, M., & Dafforn, T. R. (2011). Surfactant-free purification of membrane proteins with intact native membrane environment. *Biochemical Society Transactions*, *39*(3), 813–818. https://doi.org/10.1042/BST0390813

Jin, Z., Du, X., Xu, Y., Deng, Y., Liu, M., Zhao, Y., Zhang, B., Li, X., Zhang, L., Peng, C., Duan, Y., Yu, J., Wang, L., Yang, K., Liu, F., Jiang, R., Yang, X., You, T., Liu, X., … Yang, H. (2020). Structure of Mpro from SARS-CoV-2 and discovery of its inhibitors. *Nature 2020 582:7811*, *582*(7811), 289–293. https://doi.org/10.1038/s41586-020-2223-y

Johansen-Leete, J., Ullrich, S., Fry, S. E., Frkic, R., Bedding, M. J., Aggarwal, A., Ashhurst, A. S., Ekanayake, K. B., Mahawaththa, M. C., Sasi, V. M., Luedtke, S., Ford, D. J., O'Donoghue, A. J., Passioura, T., Larance, M., Otting, G., Turville, S., Jackson, C. J., Nitsche, C., & Payne, R. J. (2022). Antiviral cyclic peptides targeting the main protease of SARS-CoV-2. *Chemical Science*, *13*(13), 3826–3836. https://doi.org/10.1039/D1SC06750H

Kamalinia, G., Grindel, B. J., Takahashi, T. T., Millward, S. W., & Roberts, R. W. (2021). Directing evolution of novel ligands by mRNA display. *Chemical Society Reviews*, *50*(16), 9055–9103. https://doi.org/10.1039/D1CS00160D

Kelly, J. A., Olson, A. N., Neupane, K., Munshi, S., Emeterio, J. S., Pollack, L., Woodside, M. T., & Dinman, J. D. (2020). Structural and functional conservation of the programmed −1 ribosomal frameshift signal of SARS coronavirus 2 (SARS-CoV-2). *Journal of Biological Chemistry*, *295*(31), 10741–10748. https://doi.org/10.1074/jbc.AC120.013449

Khailany, R. A., Safdar, M., & Ozaslan, M. (2020). Genomic characterization of a novel SARS-CoV-2. *Gene Reports*, *19*, 100682. https://doi.org/10.1016/j.genrep.2020.100682

Kijanka, M., Dorresteijn, B., Oliveira, S., & Van Bergen En Henegouwen, P. M. P. (2015). Nanobody-based cancer therapy of solid tumors. *Nanomedicine*, *10*(1), 161–174. https://doi.org/10.2217/NNM.14.178/ASSET/IMAGES/LARGE/FIGURE3.JPEG

Kim, D., Lee, J. Y., Yang, J. S., Kim, J. W., Kim, V. N., & Chang, H. (2020). The Architecture of SARS-CoV-2 Transcriptome. *Cell*, *181*(4), 914-921.e10. https://doi.org/10.1016/J.CELL.2020.04.011

Kirchhofer, A., Helma, J., Schmidthals, K., Frauer, C., Cui, S., Karcher, A., Pellis, M., Muyldermans, S., Casas-Delucchi, C. S., Cardoso, M. C., Leonhardt, H., Hopfner, K. P., & Rothbauer, U. (2009). Modulation of protein properties in living cells using nanobodies. *Nature Structural & Molecular Biology 2009 17:1*, *17*(1), 133–138. https://doi.org/10.1038/nsmb.1727

Kitiratschky, V. B. D., Behnen, P., Kellner, U., Heckenlively, J. R., Zrenner, E., JÃ¤gle, H., Kohl, S., Wissinger, B., & Koch, K.-W. (2009). Mutations in the *GUCA1A* gene involved in hereditary cone dystrophies impair calcium-mediated regulation of guanylate cyclase. *Human Mutation*, *30*(8), E782–E796. https://doi.org/10.1002/humu.21055

Kleber-Janke, T., & Becker, W.-M. (2000). Use of Modified BL21(DE3) Escherichia coli Cells for High-Level Expression of Recombinant Peanut Allergens Affected by Poor Codon Usage. *Protein Expression and Purification*, *19*(3), 419–424. https://doi.org/10.1006/prep.2000.1265

Klemm, T., Ebert, G., Calleja, D. J., Allison, C. C., Richardson, L. W., Bernardini, J. P., Lu, B. G., Kuchel, N. W., Grohmann, C., Shibata, Y., Gan, Z. Y., Cooney, J. P., Doerflinger, M., Au, A. E., Blackmore, T. R., Noort, G. J. van der H. van, Geurink, P. P., Ovaa, H., Newman, J., … Komander, D. (2020). Mechanism and inhibition of the papain-like protease, PLpro, of SARS-CoV-2. *The EMBO Journal*, *39*(18), e106275. https://doi.org/10.15252/EMBJ.2020106275

Koch, K.-W., & Stryer, L. (1988). Highly cooperative feedback control of retinal rod guanylate cyclase by calcium ions. *Nature*, *334*(6177), 64–66. https://doi.org/10.1038/334064a0

Kunz, P., Flock, T., Soler, N., Zaiss, M., Vincke, C., Sterckx, Y., Kastelic, D., Muyldermans, S., & Hoheisel, J. D. (2017). Exploiting sequence and stability information for directing nanobody stability engineering. *Biochimica et Biophysica Acta (BBA) - General Subjects*, *1861*(9), 2196–2205. https://doi.org/10.1016/J.BBAGEN.2017.06.014

Kurz, M., Gu, K., Al-Gawari, A., & Lohse, P. A. (2001). cDNA–Protein Fusions: Covalent Protein–Gene Conjugates for the In Vitro Selection of Peptides and Proteins. *ChemBioChem*, *2*(9), 666–672. https://doi.org/https://doi.org/10.1002/1439-7633(20010903)2:9<666::AID-CBIC666>3.0.CO;2-#

Kurz, M., Gu, K., & Lohse, P. A. (2000). Psoralen photo-crosslinked mRNA–puromycin conjugates: a novel template for the rapid and facile preparation of mRNA–protein fusions. *Nucleic Acids Research*, *28*(18), e83–e83. https://doi.org/10.1093/NAR/28.18.E83

Lange, C., Duda, T., Beyermann, M., Sharma, R. K., & Koch, K.-W. (1999). Regions in vertebrate photoreceptor guanylyl cyclase ROS-GC1 involved in Ca $^{2+}$ -dependent regulation by guanylyl cyclase-activating protein GCAP-1. *FEBS Letters*, *460*(1), 27–31. https://doi.org/10.1016/S0014-5793(99)01312-5

Larsen, B., Wills, N. M., Gesteland, R. F., & Atkins, J. F. (1994). rRNA-mRNA base pairing stimulates a programmed -1 ribosomal frameshift. *Journal of Bacteriology*, *176*(22), 6842–6851. https://doi.org/10.1128/jb.176.22.6842-6851.1994

Laura, R. P., & Hurley, J. B. (1998). The Kinase Homology Domain of Retinal Guanylyl Cyclases 1 and 2 Specifies the Affinity and Cooperativity of Interaction with Guanylyl Cyclase Activating Protein-2. *Biochemistry*, *37*(32), 11264–11271. https://doi.org/10.1021/bi9809674

Lauwereys, M., Ghahroudi, M. A., Desmyter, A., Kinne, J., Hölzer, W., De Genst, E., Wyns, L., & Muyldermans, S. (1998). Potent enzyme inhibitors derived from dromedary heavy-chain antibodies. *The EMBO Journal*, *17*(13), 3512–3520. https://doi.org/10.1093/EMBOJ/17.13.3512

Leber, Th. (1869). Ueber Retinitis pigmentosa und angeborene Amaurose. *Albrecht von Graefes Archiv Für Ophthalmologie*, *15*(3), 1–25. https://doi.org/10.1007/BF02721213

Lee, J., Worrall, L. J., Vuckovic, M., Rosell, F. I., Gentile, F., Ton, A.-T., Caveney, N. A., Ban, F., Cherkasov, A., Paetzel, M., & Strynadka, N. C. J. (2020). Crystallographic structure of wild-type SARS-CoV-2 main protease acyl-enzyme intermediate with physiological C-terminal autoprocessing site. *Nature Communications*, *11*(1), 5877. https://doi.org/10.1038/s41467-020-19662-4

Li, D., Ji, F., Huang, C., & Jia, L. (2019). High Expression Achievement of Active and Robust Anti-β2 microglobulin Nanobodies via E.coli Hosts Selection. *Molecules*, *24*(16), 2860. https://doi.org/10.3390/molecules24162860

Li, S., Millward, S., & Roberts, R. (2002). In Vitro Selection of mRNA Display Libraries Containing an Unnatural Amino Acid. *Journal of the American Chemical Society*, *124*(34), 9972–9973. https://doi.org/10.1021/ja026789q

Lim, S., Peshenko, I. V., Olshevskaya, E. V., Dizhoor, A. M., & Ames, J. B. (2016). Structure of Guanylyl Cyclase Activator Protein 1 (GCAP1) Mutant V77E in a Ca2+-free/Mg2+-bound Activator State. *Journal of Biological Chemistry*, *291*(9), 4429–4441. https://doi.org/10.1074/jbc.M115.696161

Lobato, M. N., & Rabbitts, T. H. (2003). Intracellular antibodies and challenges facing their use as therapeutic agents. *Trends in Molecular Medicine*, *9*(9), 390–396. https://doi.org/10.1016/S1471-4914(03)00163-1

Lood, C., Gerstmans, H., Briers, Y., Noort, V. van, & Lavigne, R. (2020). Quality control and statistical evaluation of combinatorial DNA libraries using nanopore sequencing. *BioTechniques*, *69*(5), 379–383. https://doi.org/10.2144/btn-2020-0060

Ma, H., Zeng, W., Meng, X., Huang, X., Yang, Y., Zhao, D., Zhou, P., Wang, X., Zhao, C., Sun, Y., Wang, P., Ou, H., Hu, X., Xiang, Y., & Jin, T. (2021). Potent Neutralization of SARS-CoV-2 by Hetero-Bivalent Alpaca Nanobodies Targeting the Spike Receptor-Binding Domain. *Journal of Virology*, *95*(10). https://doi.org/10.1128/JVI.02438-20/ASSET/CD6E41F9-C0FF-47C3-9E81-670449F79817/ASSETS/IMAGES/LARGE/JVI.02438-20-F010.JPG

Maass, D. R., Sepulveda, J., Pernthaner, A., & Shoemaker, C. B. (2007). Alpaca (Lama pacos) as a convenient source of recombinant camelid heavy chain antibodies (VHHs). *Journal of Immunological Methods*, *324*(1–2), 13–25. https://doi.org/10.1016/j.jim.2007.04.008

Madeira, F., Pearce, M., Tivey, A. R. N., Basutkar, P., Lee, J., Edbali, O., Madhusoodanan, N., Kolesnikov, A., & Lopez, R. (2022). Search and sequence analysis tools services from EMBL-EBI in 2022. *Nucleic Acids Research*, *50*(W1), W276–W279. https://doi.org/10.1093/NAR/GKAC240

Mädje, K., Schmölzer, K., Nidetzky, B., & Kratzer, R. (2012). Host cell and expression engineering for development of an E. coli ketoreductase catalyst: Enhancement of formate dehydrogenase activity for regeneration of NADH. *Microbial Cell Factories*, *11*(1), 7. https://doi.org/10.1186/1475-2859-11-7

Marc, R., Liu, W., Kalloniatis, M., Raiguel, S., & van Haesendonck, E. (1990). Patterns of glutamate immunoreactivity in the goldfish retina. *The Journal of Neuroscience*, *10*(12), 4006–4034. https://doi.org/10.1523/JNEUROSCI.10-12-04006.1990

Marheineke, K., Grünewald, S., Christie, W., & Reiländer, H. (1998). Lipid composition of *Spodoptera frugiperda* (Sf9) and *Trichoplusia ni* (Tn) insect cells used for baculovirus infection. *FEBS Letters*, *441*(1), 49–52. https://doi.org/10.1016/S0014-5793(98)01523-3

Markham, A. (2018). Erenumab: First Global Approval. *Drugs*, *78*(11), 1157–1161. https://doi.org/10.1007/S40265-018-0944-0/METRICS

Martin, R. E., Elliott, M. H., Brush, R. S., & Anderson, R. E. (2005). Detailed Characterization of the Lipid Composition of Detergent-Resistant Membranes from Photoreceptor Rod Outer Segment Membranes. *Investigative Opthalmology & Visual Science*, *46*(4), 1147. https://doi.org/10.1167/iovs.04-1207

Martineau, P., Jones, P., & Winter, G. (1998). Expression of an antibody fragment at high levels in the bacterial cytoplasm. *Journal of Molecular Biology*, *280*(1), 117–127. https://doi.org/10.1006/JMBI.1998.1840

Mathieu-Daude, F. (1996). DNA rehybridization during PCR: the 'Cot effect' and its consequences. *Nucleic Acids Research*, *24*(11), 2080–2086. https://doi.org/10.1093/nar/24.11.2080

Max, E. E., Seidman, J. G., & Leder, P. (1979). Sequences of five potential recombination sites encoded close to an immunoglobulin kappa constant region gene. *Proceedings of the National Academy of Sciences*, *76*(7), 3450–3454. https://doi.org/10.1073/pnas.76.7.3450

McCafferty, J., Griffiths, A. D., Winter, G., & Chiswell, D. J. (1990). Phage antibodies: filamentous phage displaying antibody variable domains. *Nature*, *348*(6301), 552–554. https://doi.org/10.1038/348552a0

McCullough, K. T., Boye, S. L., Fajardo, D., Calabro, K., Peterson, J. J., Strang, C. E., Chakraborty, D., Gloskowski, S., Haskett, S., Samuelsson, S., Jiang, H., Witherspoon, C. D., Gamlin, P. D., Maeder, M. L., & Boye, S. E. (2019). Somatic Gene Editing of *GUCY2D* by AAV-CRISPR/Cas9 Alters Retinal Structure and Function in Mouse and Macaque. *Human Gene Therapy*, *30*(5), 571–589. https://doi.org/10.1089/hum.2018.193

McDade, E., Cummings, J. L., Dhadda, S., Swanson, C. J., Reyderman, L., Kanekiyo, M., Koyama, A., Irizarry, M., Kramer, L. D., & Bateman, R. J. (2022). Lecanemab in patients with early Alzheimer's disease: detailed results on biomarker, cognitive, and clinical effects from the randomized and open-label extension of the phase 2 proof-of-concept study. *Alzheimer's Research and Therapy*, *14*(1), 1–17. https://doi.org/10.1186/S13195-022-01124-2/FIGURES/9

McMahon, C., Baier, A. S., Pascolutti, R., Wegrecki, M., Zheng, S., Ong, J. X., Erlandson, S. C., Hilger, D., Rasmussen, S. G. F., Ring, A. M., Manglik, A., & Kruse, A. C. (2018). Yeast surface display platform for rapid discovery of conformationally selective nanobodies. *Nature Structural and Molecular Biology*, *25*(3). https://doi.org/10.1038/s41594-018-0028-6

Mochizuki, Y., Biyani, M., Tsuji-Ueno, S., Suzuki, M., Nishigaki, K., Husimi, Y., & Nemoto, N. (2011). One-Pot Preparation of mRNA/cDNA Display by a Novel and Versatile Puromycin-Linker DNA. *ACS Combinatorial Science*, *13*(5), 478–485. https://doi.org/10.1021/co2000295

Molday, L. L., Djajadi, H., Yan, P., Szczygiel, L., Boye, S. L., Chiodo, V. A., Gregory-Evans, K., Sarunic, M. V., Hauswirth, W. W., & Molday, R. S. (2013). RD3 gene delivery restores guanylate cyclase localization and rescues photoreceptors in the Rd3 mouse model of Leber congenital amaurosis 12. *Human Molecular Genetics*, *22*(19), 3894–3905. https://doi.org/10.1093/hmg/ddt244

Moore, J. B., Blanchard, R. K., & Cousins, R. J. (2003). Dietary zinc modulates gene expression in murine thymus: Results from a comprehensive differential display screening. *Proceedings of the National Academy of Sciences*, *100*(7), 3883–3888. https://doi.org/10.1073/pnas.0330670100

Moreno-Gonzalez, I., & Soto, C. (2011). Misfolded protein aggregates: Mechanisms, structures and potential for disease transmission. *Seminars in Cell & Developmental Biology*, *22*(5), 482–487. https://doi.org/10.1016/J.SEMCDB.2011.04.002

Munro, S., & Pelham, H. R. (1984). Use of peptide tagging to detect proteins expressed from cloned genes: deletion mapping functional domains of Drosophila hsp 70. *The EMBO Journal*, *3*(13), 3087–3093. https://doi.org/10.1002/j.1460-2075.1984.tb02263.x

Muruganandam, A., Tanha, J., Narang, S., & Stanimirovic, D. (2002). Selection of phage-displayed llama single-domain antibodies that transmigrate across human blood-brain barrier endothelium. *The FASEB Journal*, *16*(2), 1–22. https://doi.org/10.1096/FJ.01-0343FJE

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

Nair, K. S., Balasubramanian, N., & Slepak, V. Z. (2002). Signal-Dependent Translocation of Transducin, RGS9-1-Gβ5L Complex, and Arrestin to Detergent-Resistant Membrane Rafts in Photoreceptors. *Current Biology*, *12*(5), 421–425. https://doi.org/10.1016/S0960-9822(02)00691-7

Nemoto, N., Miyamoto-Sato, E., Husimi, Y., & Yanagawa, H. (1997). In vitro virus: Bonding of mRNA bearing puromycin at the 3′-terminal end to the C-terminal end of its encoded protein on the ribosome in vitro. *FEBS Letters*, *414*(2), 405–408. https://doi.org/10.1016/S0014-5793(97)01026-0

Newton, M. S., Cabezas-Perusse, Y., Tong, C. L., & Seelig, B. (2020). *In Vitro* Selection of Peptides and Proteins—Advantages of mRNA Display. *ACS Synthetic Biology*, *9*(2), 181–190. https://doi.org/10.1021/acssynbio.9b00419

Nickell, S., Park, P. S.-H., Baumeister, W., & Palczewski, K. (2007). Three-dimensional architecture of murine rod outer segments determined by cryoelectron tomography. *Journal of Cell Biology*, *177*(5), 917–925. https://doi.org/10.1083/jcb.200612010

Norman, A., Franck, C., Christie, M., Hawkins, P. M. E., Patel, K., Ashhurst, A. S., Aggarwal, A., Low, J. K. K., Siddiquee, R., Ashley, C. L., Steain, M., Triccas, J. A., Turville, S., Mackay, J. P., Passioura, T., & Payne, R. J. (2021). Discovery of Cyclic Peptide Ligands to the SARS-CoV-2 Spike Protein Using mRNA Display. *ACS Central Science*, *7*(6), 1001–1008. https://doi.org/10.1021/acscentsci.0c01708

Nussenzweig, M. C., Shaw, A. C., Sinn, E., Danner, D. B., Holmes, K. L., Morse, H. C., & Leder, P. (1987). Allelic Exclusion in Transgenic Mice That Express the Membrane form of Immunoglobulin μ. *Science*, *236*(4803), 816–819. https://doi.org/10.1126/SCIENCE.3107126

Oluwole, A. O., Danielczak, B., Meister, A., Babalola, J. O., Vargas, C., & Keller, S. (2017). Solubilization of Membrane Proteins into Functional Lipid-Bilayer Nanodiscs Using a Diisobutylene/Maleic Acid Copolymer. *Angewandte Chemie International Edition*, *56*(7), 1919–1924. https://doi.org/10.1002/anie.201610778

Otto-Bruc, A., Buczyłko, J., Surgucheva, I., Subbaraya, I., Rudnicka-Nawrot, M., Crabb, J. W., Arendt, A., Hargrave, P. A., Baehr, W., & Palczewski, K. (1997). Functional Reconstitution of Photoreceptor Guanylate Cyclase with Native and Mutant Forms of Guanylate Cyclase-Activating Protein 1. *Biochemistry*, *36*(14), 4295–4302. https://doi.org/10.1021/bi963000d

Overduin, M., Trieber, C., Prosser, R. S., Picard, L.-P., & Sheff, J. G. (2021). Structures and Dynamics of Native-State Transmembrane Protein Targets and Bound Lipids. *Membranes*, *11*(6), 451. https://doi.org/10.3390/membranes11060451

Owen, D. R., Allerton, C. M. N., Anderson, A. S., Aschenbrenner, L., Avery, M., Berritt, S., Boras, B., Cardin, R. D., Carlo, A., Coffman, K. J., Dantonio, A., Di, L., Eng, H., Ferre, R., Gajiwala, K. S., Gibson, S. A., Greasley, S. E., Hurst, B. L., Kadar, E. P., … Zhu, Y. (2021). An oral SARS-CoV-2 M^pro inhibitor clinical candidate for the treatment of COVID-19. *Science*, *374*(6575), 1586–1593. https://doi.org/10.1126/science.abl4784

Pacheco, B., Crombet, L., Loppnau, P., & Cossar, D. (2012). A screening strategy for heterologous protein expression in Escherichia coli with the highest return of investment. *Protein Expression and Purification*, *81*(1), 33–41. https://doi.org/10.1016/j.pep.2011.08.030

Palczewski, K., Subbaraya, I., Gorczyca, W. A., Helekar, B. S., Ruiz, C. C., Ohguro, H., Huang, J., Zhao, X., Crabb, J. W., Johnson, R. S., Walsh, K. A., Gray-Keller, M. P., Detwiler, P. B., & Baehr, W. (1994). Molecular cloning and characterization of retinal photoreceptor guanylyl cyclase-activating protein. *Neuron*, *13*(2), 395–404. https://doi.org/10.1016/0896-6273(94)90355-7

Papavasiliou, F., Casellas, R., Suh, H., Qin, X.-F., Besmer, E., Pelanda, R., Nemazee, D., Rajewsky, K., & Nussenzweig, M. C. (1997). V(D)J Recombination in Mature B Cells: A Mechanism for Altering Antibody Responses. *Science*, *278*(5336), 298–301. https://doi.org/10.1126/science.278.5336.298

Parmley, S. F., & Smith, G. P. (1988). Antibody-selectable filamentous fd phage vectors: affinity purification of target genes. *Gene*, *73*(2), 305–318. https://doi.org/10.1016/0378-1119(88)90495-7

Passioura, T., Katoh, T., Goto, Y., & Suga, H. (2014). Selection-Based Discovery of Druglike Macrocyclic Peptides. *Annual Review of Biochemistry*, *83*(1), 727–752. https://doi.org/10.1146/annurev-biochem-060713-035456

Payne, A. M., Downes, S. M., Bessant, D. A. R., Taylor, R., Holder, G. E., Warren, M. J., Bird, A. C., & Bhattacharya, S. S. (1998). A Mutation in Guanylate Cyclase Activator 1A (GUCA1A) in an Autosomal Dominant Cone Dystrophy Pedigree Mapping to a New Locus on Chromosome 6p21.1. *Human Molecular Genetics*, *7*(2), 273–277. https://doi.org/10.1093/HMG/7.2.273

Pérez, J. M. J., Renisio, J. G., Prompers, J. J., Van Platerink, C. J., Cambillau, C., Darbon, H., & Frenken, L. G. J. (2001). Thermal unfolding of a llama antibody fragment: A two-state reversible process. *Biochemistry*, *40*(1), 74–83. https://doi.org/10.1021/BI0009082/SUPPL_FILE/BI0009082_S.PDF

Perrault, I., Rozet, J. M., Calvas, P., Gerber, S., Camuzat, A., Dollfus, H., Châtelin, S., Souied, E., Ghazi, I., Leowski, C., Bonnemaison, M., Paslier, D. Le, Frézal, J., Dufier, J.-L., Pittler, S., Munnich, A., & Kaplan, J. (1996). Retinal–specific guanylate cyclase gene mutations in Leber's congenital amaurosis. *Nature Genetics*, *14*(4), 461–464. https://doi.org/10.1038/ng1296-461

Peshenko, I. V., & Dizhoor, A. M. (2004). Guanylyl Cyclase-activating Proteins (GCAPs) Are Ca2+/Mg2+ Sensors. *Journal of Biological Chemistry*, *279*(17), 16903–16906. https://doi.org/10.1074/jbc.c400065200

Peshenko, I. V., Moiseyev, G. P., Olshevskaya, E. V., & Dizhoor, A. M. (2004). Factors that Determine Ca $^{2+}$ Sensitivity of Photoreceptor Guanylyl Cyclase. Kinetic Analysis of the Interaction between the Ca $^{2+}$ -Bound and the Ca $^{2+}$ -Free Guanylyl Cyclase Activating Proteins (GCAPs) and Recombinant Photoreceptor Guanylyl Cyclase 1 (RetGC-1). *Biochemistry*, *43*(43), 13796–13804. https://doi.org/10.1021/bi048943m

Peshenko, I. V., Olshevskaya, E. V., & Dizhoor, A. M. (2015). Evaluating the role of retinal membrane guanylyl cyclase 1 (RetGC1) domains in binding guanylyl cyclase-activating proteins (GCAPs). *Journal of Biological Chemistry*, *290*(11), 6913–6924. https://doi.org/10.1074/jbc.M114.629642

Peshenko, I. V., Olshevskaya, E. V., Lim, S., Ames, J. B., & Dizhoor, A. M. (2014). Identification of target binding site in photoreceptor guanylyl cyclase-activating protein 1 (GCAP1). *Journal of Biological Chemistry*, *289*(14), 10140–10154. https://doi.org/10.1074/jbc.M113.540716

Peshenko, I. V., Olshevskaya, E. V., Yao, S., Ezzeldin, H. H., Pittler, S. J., & Dizhoor, A. M. (2010). Activation of Retinal Guanylyl Cyclase RetGC1 by GCAP1: Stoichiometry of Binding and Effect of New LCA-Related Mutations. *Biochemistry*, *49*(4), 709–717. https://doi.org/10.1021/bi901495y

Pettelkau, J., Schröder, T., Ihling, C. H., Olausson, B. E. S., Kölbel, K., Lange, C., & Sinz, A. (2012). Structural Insights into Retinal Guanylylcyclase–GCAP-2 Interaction Determined by Cross-Linking and Mass Spectrometry. *Biochemistry*, *51*(24), 4932–4949. https://doi.org/10.1021/bi300064v

Pillarisetti, K., Powers, G., Luistro, L., Babich, A., Baldwin, E., Li, Y., Zhang, X., Mendonça, M., Majewski, N., Nanjunda, R., Chin, D., Packman, K., Elsayed, Y., Attar, R., & Gaudet, F. (2020). Teclistamab is an active T cell–redirecting bispecific antibody against B-cell maturation antigen for multiple myeloma. *Blood Advances*, *4*(18), 4538–4549. https://doi.org/10.1182/bloodadvances.2020002393

Polack, F. P., Thomas, S. J., Kitchin, N., Absalon, J., Gurtman, A., Lockhart, S., Perez, J. L., Pérez Marc, G., Moreira, E. D., Zerbini, C., Bailey, R., Swanson, K. A., Roychoudhury, S., Koury, K., Li, P., Kalina, W. V., Cooper, D., Frenck, R. W., Hammitt, L. L., … Gruber, W. C. (2020). Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine. *New England Journal of Medicine*, *383*(27), 2603–2615. https://doi.org/10.1056/NEJMoa2034577

Polz, M. F., & Cavanaugh, C. M. (1998). Bias in Template-to-Product Ratios in Multitemplate PCR. *Applied and Environmental Microbiology*, *64*(10), 3724–3730. https://doi.org/10.1128/AEM.64.10.3724-3730.1998

Prado, N. D. R., Pereira, S. S., Da Silva, M. P., Morais, M. S. S., Kayano, A. M., Moreira-Dill, L. S., Luiz, M. B., Zanchi, F. B., Fuly, A. L., Huacca, M. E. F., Fernandes, C. F., Calderon, L. A., Zuliani, J. P., Da Silva, L. H. P., Soares, A. M., Stabeli, R. G., & Fernandes, C. F. C. (2016). Inhibition of the Myotoxicity Induced by Bothrops jararacussu Venom and Isolated Phospholipases A2 by Specific Camelid Single-Domain Antibody Fragments. *PLOS ONE*, *11*(3), e0151363. https://doi.org/10.1371/JOURNAL.PONE.0151363

Qiao, J., Li, Y.-S., Zeng, R., Liu, F.-L., Luo, R.-H., Huang, C., Wang, Y.-F., Zhang, J., Quan, B., Shen, C., Mao, X., Liu, X., Sun, W., Yang, W., Ni, X., Wang, K., Xu, L., Duan, Z.-L., Zou, Q.-C., … Yang, S. (2021). SARS-CoV-2 M ^{pro} inhibitors with antiviral activity in a transgenic mouse model. *Science*, *371*(6536), 1374–1378. https://doi.org/10.1126/science.abf1611

Qureshi, B. M., Behrmann, E., Schöneberg, J., Loerke, J., Bürger, J., Mielke, T., Giesebrecht, J., Noé, F., Lamb, T. D., Hofmann, K. P., Spahn, C. M. T., & Heck, M. (2018). It takes two transducins to activate the cGMP-phosphodiesterase 6 in retinal rods. *Open Biology*, *8*(8). https://doi.org/10.1098/RSOB.180075

Rasmussen, S. G. F., Choi, H. J., Fung, J. J., Pardon, E., Casarosa, P., Chae, P. S., Devree, B. T., Rosenbaum, D. M., Thian, F. S., Kobilka, T. S., Schnapp, A., Konetzki, I., Sunahara, R. K., Gellman, S. H., Pautsch, A., Steyaert, J., Weis, W. I., & Kobilka, B. K. (2011). Structure of a nanobody-stabilized active state of the β2 adrenoceptor. *Nature 2011 469:7329*, *469*(7329), 175–180. https://doi.org/10.1038/nature09648

Rehkamp, A., Tänzler, D., Tüting, C., Kastritis, P. L., Iacobucci, C., Ihling, C. H., Kipping, M., Koch, K. W., & Sinz, A. (2021). First 3D-Structural Data of Full-Length Guanylyl Cyclase 1 in Rod-Outer-Segment Preparations of Bovine Retina by Cross-Linking/Mass Spectrometry. *Journal of Molecular Biology*, *433*(10), 166947. https://doi.org/10.1016/J.JMB.2021.166947

Renart, J., Reiser, J., & Stark, G. R. (1979). Transfer of proteins from gels to diazobenzyloxymethyl-paper and detection with antisera: a method for studying antibody specificity and antigen structure. *Proceedings of the National Academy of Sciences*, *76*(7), 3116–3120. https://doi.org/10.1073/pnas.76.7.3116

Robert, X., & Gouet, P. (2014). Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Research*, *42*(W1), W320–W324. https://doi.org/10.1093/NAR/GKU316

Roberts, 15, Abelson, P. H., Cowie, D. B., Bolton, E. T., Britten, R. J., & Institu, C. (1959). INHIBITION BY PUROMYCIN OF AMINO ACID INCORPORATION INTO PROTEIN. *Proceedings of the National Academy of Sciences of the United States of America*, *45*(12), 1721. https://doi.org/10.1073/PNAS.45.12.1721

Roberts, R. W., & Szostak, J. W. (1997). RNA-peptide fusions for the *in vitro* selection of peptides and proteins. *Proceedings of the National Academy of Sciences*, *94*(23), 12297–12302. https://doi.org/10.1073/pnas.94.23.12297

Robinson, M., Lilley, R., Little, S., Emtage, J. S., Yarranton, G., Stephens, P., Millican, A., Eaton, M., & Humphreys, G. (1984). Codon usage can affect efficiency of translation of genes in *Escherichia coli*. *Nucleic Acids Research*, *12*(17), 6663–6671. https://doi.org/10.1093/nar/12.17.6663

Rodi, D. J., Soares, A. S., & Makowski, L. (2002). Quantitative Assessment of Peptide Sequence Diversity in M13 Combinatorial Peptide Phage Display Libraries. *Journal of Molecular Biology*, *322*(5), 1039–1052. https://doi.org/10.1016/S0022-2836(02)00844-6

Roovers, R. C., Vosjan, M. J. W. D., Laeremans, T., El Khoulati, R., De Bruin, R. C. G., Ferguson, K. M., Verkleij, A. J., Van Dongen, G. A. M. S., & Van Bergen En Henegouwen, P. M. P. (2011). A biparatopic anti-EGFR nanobody efficiently inhibits solid tumour growth. *International Journal of Cancer*, *129*(8), 2013–2024. https://doi.org/10.1002/IJC.26145

Rosano, G. L., & Ceccarelli, E. A. (2009). Rare codon content affects the solubility of recombinant proteins in a codon bias-adjusted Escherichia coli strain. *Microbial Cell Factories*, *8*(1), 41. https://doi.org/10.1186/1475-2859-8-41

Rosenblum, G., & Cooperman, B. S. (2014). Engine out of the chassis: Cell-free protein synthesis and its uses. *FEBS Letters*, *588*(2), 261–268. https://doi.org/10.1016/J.FEBSLET.2013.10.016

Ryckaert, S., Pardon, E., Steyaert, J., & Callewaert, N. (2010). Isolation of antigen-binding camelid heavy chain antibody fragments (nanobodies) from an immune library displayed on the surface of Pichia pastoris. *Journal of Biotechnology*, *145*(2), 93–98. https://doi.org/10.1016/j.jbiotec.2009.10.010

Saerens, D., Kinne, J., Bosmans, E., Wernery, U., Muyldermans, S., & Conrath, K. (2004). Single Domain Antibodies Derived from Dromedary Lymph Node and Peripheral Blood Lymphocytes Sensing Conformational Variants of Prostate-specific Antigen. *Journal of Biological Chemistry*, *279*(50), 51965–51972. https://doi.org/10.1074/jbc.M409292200

Salem, T. Z., Zhang, F., Sahly, N., & Thiem, S. (2018). Effect of Temporal Expression of Integral Membrane Proteins by Baculovirus Expression Vector System. *Molecular Biotechnology*, *60*(8), 576–584. https://doi.org/10.1007/s12033-018-0099-y

Salema, V., Marín, E., Martínez-Arteaga, R., Ruano-Gallego, D., Fraile, S., Margolles, Y., Teira, X., Gutierrez, C., Bodelón, G., & Fernández, L. Á. (2013). Selection of Single Domain Antibodies from Immune Libraries Displayed on the Surface of E. coli Cells with Two β-Domains of Opposite Topologies. *PLoS ONE*, *8*(9). https://doi.org/10.1371/JOURNAL.PONE.0075126

Schoof, M., Faust, B., Saunders, R. A., Sangwan, S., Rezelj, V., Hoppe, N., Boone, M., Billesbølle, C. B., Puchades, C., Azumaya, C. M., Kratochvil, H. T., Zimanyi, M., Deshpande, I., Liang, J., Dickinson, S., Nguyen, H. C., Chio, C. M., Merz, G. E., Thompson, M. C., … Manglik, A. (2020). An ultrapotent synthetic

nanobody neutralizes SARS-CoV-2 by stabilizing inactive Spike. *Science (New York, N.y.)*, *370*(6523), 1473. https://doi.org/10.1126/SCIENCE.ABE3255

Schumacher, D., Helma, J., Schneider, A. F. L., Leonhardt, H., & Hackenberger, C. P. R. (2018). Nanobodies: Chemical Functionalization Strategies and Intracellular Applications. *Angewandte Chemie International Edition*, *57*(9), 2314–2333. https://doi.org/10.1002/anie.201708459

Scully, M., Cataland, S. R., Peyvandi, F., Coppo, P., Knöbl, P., Kremer Hovinga, J. A., Metjian, A., de la Rubia, J., Pavenski, K., Callewaert, F., Biswas, D., De Winter, H., & Zeldin, R. K. (2019). Caplacizumab Treatment for Acquired Thrombotic Thrombocytopenic Purpura. *New England Journal of Medicine*, *380*(4), 335–346. https://doi.org/10.1056/NEJMOA1806311/SUPPL_FILE/NEJMOA1806311_DISCLOSURES.PDF

Seelig, B. (2011). MRNA display for the selection and evolution of enzymes from in vitro-translated protein libraries. *Nature Protocols*. https://doi.org/10.1038/nprot.2011.312

Seelig, B., & Szostak, J. W. (2007). Selection and evolution of enzymes from a partially randomized non-catalytic scaffold. *Nature 2007 448:7155*, *448*(7155), 828–831. https://doi.org/10.1038/nature06032

Semenkov, Y., Shapkina, T., Makhno, V., & Kirillov, S. (1992). Puromycin reaction for the A site-bound peptidyl-tRNA. *FEBS Letters*, *296*(2), 207–210. https://doi.org/10.1016/0014-5793(92)80380-Y

Sheikholeslami, F., Rasaee, M. J., Shokrgozar, M. A., Dizaji, M. M., Rahbarizadeh, F., & Ahmadvande, D. (2010). Isolation of a Novel Nanobody Against HER-2/neu Using Phage Displays Technology. *Laboratory Medicine*, *41*(2), 69–76. https://doi.org/10.1309/LM0WXKM0R0DVUZWF

Shen, Z., Ratia, K., Cooper, L., Kong, D., Lee, H., Kwon, Y., Li, Y., Alqarni, S., Huang, F., Dubrovskyi, O., Rong, L., Thatcher, G. R. J., & Xiong, R. (2022). Design of SARS-CoV-2 PLpro Inhibitors for COVID-19 Antiviral Therapy Leveraging Binding Cooperativity. *Journal of Medicinal Chemistry*, *65*(4), 2940–2955. https://doi.org/10.1021/acs.jmedchem.1c01307

Shimizu, Y., Inoue, A., Tomari, Y., Suzuki, T., Yokogawa, T., Nishikawa, K., & Ueda, T. (2001). Cell-free translation reconstituted with purified components. *Nature Biotechnology*, *19*(8), 751–755. https://doi.org/10.1038/90802

Shin, D., Mukherjee, R., Grewe, D., Bojkova, D., Baek, K., Bhattacharya, A., Schulz, L., Widera, M., Mehdipour, A. R., Tascher, G., Geurink, P. P., Wilhelm, A., van der Heden van Noort, G. J., Ovaa, H., Müller, S., Knobeloch, K.-P., Rajalingam, K., Schulman, B. A., Cinatl, J., … Dikic, I. (2020). Papain-like protease regulates SARS-CoV-2 viral spread and innate immunity. *Nature*, *587*(7835), 657–662. https://doi.org/10.1038/s41586-020-2601-5

Shyjan, A. W., de Sauvage, F. J., Gillett, N. A., Goeddel, D. V., & Lowe, D. G. (1992). Molecular cloning of a retina-specific membrane guanylyl cyclase. *Neuron*, *9*(4), 727–737. https://doi.org/10.1016/0896-6273(92)90035-C

Silacci, M., Brack, S., Schirru, G., Mårlind, J., Ettorre, A., Merlo, A., Viti, F., & Neri, D. (2005). Design, construction, and characterization of a large synthetic human antibody phage display library. *PROTEOMICS*, *5*(9), 2340–2350. https://doi.org/10.1002/pmic.200401273

Simonsen, L., & Viboud, C. (2021). A comprehensive look at the covid-19 pandemic death toll. *ELife*, *10*. https://doi.org/10.7554/ELIFE.71974

Singh, M. V., & Anthony Weil, P. (2002). A method for plasmid purification directly from yeast. *Analytical Biochemistry*, *307*(1), 13–17. https://doi.org/10.1016/S0003-2697(02)00018-0

Smith, G. P. (1985). Filamentous Fusion Phage: Novel Expression Vectors That Display Cloned Antigens on the Virion Surface. *Science*, *228*(4705), 1315–1317. https://doi.org/10.1126/science.4001944

Smith, G. P., & Petrenko, V. A. (1997). Phage Display. *Chemical Reviews*, *97*(2), 391–410. https://doi.org/10.1021/cr960065d

Sokal, I., Dupps, W. J., Grassi, M. A., Brown, J., Affatigato, L. M., Roychowdhury, N., Yang, L., Filipek, S., Palczewski, K., Stone, E. M., & Baehr, W. (2005). A Novel GCAP1 Missense Mutation (L151F) in a Large Family with Autosomal Dominant Cone-Rod Dystrophy (adCORD). *Investigative Opthalmology & Visual Science*, *46*(4), 1124. https://doi.org/10.1167/iovs.04-1431

Sokal, I., Li, N., Klug, C. S., Filipek, Sb., Hubbell, W. L., Baehr, W., & Palczewski, K. (2001). Calcium-sensitive Regions of GCAP1 as Observed by Chemical Modifications, Fluorescence, and EPR Spectroscopies. *Journal of Biological Chemistry*, *276*(46), 43361–43373. https://doi.org/10.1074/jbc.M103614200

Sokal, I., Li, N., Surgucheva, I., Warren, M. J., Payne, A. M., Bhattacharya, S. S., Baehr, W., & Palczewski, K. (1998). GCAP1(Y99C) Mutant Is Constitutively Active in Autosomal Dominant Cone Dystrophy. *Molecular Cell*, *2*(1), 129–133. https://doi.org/10.1016/S1097-2765(00)80121-5

Stephen, R., Bereta, G., Golczak, M., Palczewski, K., & Sousa, M. C. (2007). Stabilizing Function for Myristoyl Group Revealed by the Crystal Structure of a Neuronal Calcium Sensor, Guanylate Cyclase-Activating Protein 1. *Structure*. https://doi.org/10.1016/j.str.2007.09.013

Stewart, E. J., Åslund, F., & Beckwith, J. (1998). Disulfide bond formation in the Escherichia coli cytoplasm: an in vivo role reversal for the thioredoxins. *The EMBO Journal*, *17*(19), 5543–5550. https://doi.org/10.1093/EMBOJ/17.19.5543

Studier, F. W., & Moffatt, B. A. (1986). Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *Journal of Molecular Biology*, *189*(1), 113–130. https://doi.org/10.1016/0022-2836(86)90385-2

Sun, Z., Wang, L., Li, X., Fan, C., Xu, J., Shi, Z., Qiao, H., Lan, Z., Zhang, X., Li, L., Zhou, X., & Geng, Y. (2022). An extended conformation of SARS-CoV-2 main protease reveals allosteric targets. *Proceedings of the National Academy of Sciences of the United States of America*, *119*(15), e2120913119. https://doi.org/10.1073/PNAS.2120913119/SUPPL_FILE/PNAS.2120913119.SAPP.PDF

Swainsbury, D. J. K., Scheidelaar, S., Foster, N., van Grondelle, R., Killian, J. A., & Jones, M. R. (2017). The effectiveness of styrene-maleic acid (SMA) copolymers for solubilisation of integral membrane proteins from SMA-accessible and SMA-resistant membranes. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, *1859*(10), 2133–2143. https://doi.org/10.1016/j.bbamem.2017.07.011

Takahashi, T. T., Austin, R. J., & Roberts, R. W. (2003). mRNA display: ligand discovery, interaction analysis and beyond. *Trends in Biochemical Sciences*, *28*(3), 159–165. https://doi.org/10.1016/S0968-0004(03)00036-7

Talmage, D. W. (1959). Immunological Specificity. *Science*, *129*(3364), 1643–1648. https://doi.org/10.1126/science.129.3364.1643

Thompson, J., Pope, T., Tung, J.-S., Chan, C., Hollis, G., Mark, G., & Johnson, K. S. (1996). Affinity Maturation of a High-affinity Human Monoclonal Antibody Against the Third Hypervariable Loop of Human Immunodeficiency Virus: Use of Phage Display to Improve Affinity and Broaden Strain Reactivity. *Journal of Molecular Biology*, *256*(1), 77–88. https://doi.org/10.1006/jmbi.1996.0069

Tiller, K. E., & Tessier, P. M. (2015). Advances in Antibody Design. *Annual Review of Biomedical Engineering*, *17*(1), 191–216. https://doi.org/10.1146/annurev-bioeng-071114-040733

Toibana, M., Tsukahara, I., & Ogawa, K. (1982). Cytochemical demonstration of guanylate cyclase activity in retinal photoreceptors with special reference to changes under light and dark adaptation. *ACTA HISTOCHEMICA ET CYTOCHEMICA*, *15*(1/2), 5–20. https://doi.org/10.1267/ahc.15.5

Tonegawa, S. (1983). Somatic generation of antibody diversity. *Nature*, *302*(5909), 575–581. https://doi.org/10.1038/302575a0

Tucker, C. L., Ramamurthy, V., Pina, A.-L., Loyer, M., Dharmaraj, S., Li, Y., Maumenee, I. H., Hurley, J. B., & Koenekoop, R. K. (2004). Functional analyses of mutant recessive GUCY2D alleles identified in Leber congenital amaurosis patients: protein domain comparisons and dominant negative effects. *Molecular Vision*, *10*, 297–303.

Valbuena, O., Marcu, K. B., Weigert, M., & Perry, R. P. (1978). Multiplicity of germline genes specifying a group of related mouse κ chains with implications for the generation of immunoglobulin diversity. *Nature*, *276*(5690), 780–784. https://doi.org/10.1038/276780a0

Vashist, S., & Ng, D. T. W. (2004). Misfolded proteins are sorted by a sequential checkpoint mechanism of ER quality control. *Journal of Cell Biology*, *165*(1), 41–52. https://doi.org/10.1083/jcb.200309132

Vincke, C., Loris, R., Saerens, D., Martinez-Rodriguez, S., Muyldermans, S., & Conrath, K. (2009). General strategy to humanize a camelid single-domain antibody and identification of a universal humanized nanobody scaffold. *Journal of Biological Chemistry*, *284*(5), 3273–3284. https://doi.org/10.1074/jbc.M806889200

Viviano, J., Krishnan, A., Wu, H., & Venkataraman, V. (2016). Electrophoretic mobility shift in native gels indicates calcium-dependent structural changes of neuronal calcium sensor proteins. *Analytical Biochemistry*, *494*, 93–100. https://doi.org/10.1016/J.AB.2015.11.005

Wang, H., & Liu, R. (2011). Advantages of mRNA display selections over other selection techniques for investigation of protein–protein interactions. *Expert Review of Proteomics*, *8*(3), 335–346. https://doi.org/10.1586/epr.11.15

Wang, M., Cao, R., Zhang, L., Yang, X., Liu, J., Xu, M., Shi, Z., Hu, Z., Zhong, W., & Xiao, G. (2020). Remdesivir and chloroquine effectively inhibit the recently emerged novel coronavirus (2019-nCoV) in vitro. *Cell Research*, *30*(3), 269–271. https://doi.org/10.1038/s41422-020-0282-0

Wilkie, S. E., Li, Y., Deery, E. C., Newbold, R. J., Garibaldi, D., Bateman, J. B., Zhang, H., Lin, W., Zack, D. J., Bhattacharya, S. S., Warren, M. J., Hunt, D. M., & Zhang, K. (2001). Identification and Functional Consequences of a New Mutation (E155G) in the Gene for GCAP1 That Causes Autosomal Dominant Cone Dystrophy. *The American Journal of Human Genetics*, *69*(3), 471–480. https://doi.org/10.1086/323265

Wilkie, S. E., Newbold, R. J., Deery, E., Walker, C. E., Stinton, I., Ramamurthy, V., Hurley, J. B., Bhattacharya, S. S., Warren, M. J., & Hunt, M. J. (2000). Functional characterization of missense mutations at codon 838 in retinal guanylate cyclase correlates with disease severity in patients with autosomal dominant cone-rod dystrophy. *Human Molecular Genetics*, *9*(20), 3065–3073. https://doi.org/10.1093/hmg/9.20.3065

Xiao, S., White, J. F., Betenbaugh, M. J., Grisshammer, R., & Shiloach, J. (2013). Transient and Stable Expression of the Neurotensin Receptor NTS1: A Comparison of the Baculovirus-Insect Cell and the T-REx-293 Expression Systems. *PLoS ONE*, *8*(5), e63679. https://doi.org/10.1371/journal.pone.0063679

Xu, L., Aha, P., Gu, K., Kuimelis, R. G., Kurz, M., Lam, T., Lim, A. C., Liu, H., Lohse, P. A., Sun, L., Weng, S., Wagner, R. W., & Lipovsek, D. (2002). Directed Evolution of High-Affinity Antibody Mimics Using mRNA Display. *Chemistry & Biology*, *9*(8), 933–942. https://doi.org/10.1016/S1074-5521(02)00187-4

Yan, J., Wang, P., Zhu, M., Li, G., Romão, E., Xiong, S., & Wan, Y. (2015). Characterization and applications of Nanobodies against human procalcitonin selected from a novel naïve Nanobody phage display library. *Journal of Nanobiotechnology*, *13*(1), 33. https://doi.org/10.1186/s12951-015-0091-7

Yang, J., & Li, T. (2005). The ciliary rootlet interacts with kinesin light chains and may provide a scaffold for kinesin-1 vesicular cargos. *Experimental Cell Research*, *309*(2), 379–389. https://doi.org/10.1016/J.YEXCR.2005.05.026

Yang, R. B., Foster, D. C., Garbers, D. L., & Fülle, H. J. (1995). Two membrane forms of guanylyl cyclase found in the eye. *Proceedings of the National Academy of Sciences*, *92*(2), 602–606. https://doi.org/10.1073/PNAS.92.2.602

Yau, K.-W., & Nakatani, K. (1985). Light-induced reduction of cytoplasmic free calcium in retinal rod outer segment. *Nature*, *313*(6003), 579–582. https://doi.org/10.1038/313579a0

Yee, R., & Liebman, P. A. (1978). Light-activated phosphodiesterase of the rod outer segment. Kinetics and parameters of activation and deactivation. *Journal of Biological Chemistry*, *253*(24), 8902–8909. https://doi.org/10.1016/S0021-9258(17)34263-1

Young, R. W., & Droz, B. (1968). THE RENEWAL OF PROTEIN IN RETINAL RODS AND CONES. *Journal of Cell Biology*, *39*(1), 169–184. https://doi.org/10.1083/JCB.39.1.169

Yu, H., Olshevskaya, E., Duda, T., Seno, K., Hayashi, F., Sharma, R. K., Dizhoor, A. M., & Yamazaki, A. (1999). Activation of Retinal Guanylyl Cyclase-1 by Ca2+-binding Proteins Involves Its Dimerization. *Journal of Biological Chemistry*, *274*(22), 15547–15555. https://doi.org/10.1074/jbc.274.22.15547

Zägel, P., Dell'Orco, D., & Koch, K. W. (2013). The dimerization domain in outer segment guanylate cyclase is a Ca 2+-sensitive control switch module. *Biochemistry*, *52*(30), 5065–5074. https://doi.org/10.1021/BI400288P/SUPPL_FILE/BI400288P_SI_001.PDF

Zahra, D. G., Vancov, T., Dunn, J. M., Hawkins, N. J., & Ward, R. L. (1999). Selectable in-vivo recombination to increase antibody library size — an improved phage display vector system. *Gene*, *227*(1), 49–54. https://doi.org/10.1016/S0378-1119(98)00593-9

Zell, R., & Fritz, H. J. (1987). DNA mismatch-repair in Escherichia coli counteracting the hydrolytic deamination of 5-methyl-cytosine residues. *The EMBO Journal*, *6*(6), 1809–1815. https://doi.org/10.1002/j.1460-2075.1987.tb02435.x

Zhang, W., Shan, H., Jiang, K., Huang, W., & Li, S. (2021). A novel intracellular nanobody against HPV16 E6 oncoprotein. *Clinical Immunology*, *225*, 108684. https://doi.org/10.1016/j.clim.2021.108684

Zhou, P., Yang, X.-L., Wang, X.-G., Hu, B., Zhang, L., Zhang, W., Si, H.-R., Zhu, Y., Li, B., Huang, C.-L., Chen, H.-D., Chen, J., Luo, Y., Guo, H., Jiang, R.-D., Liu, M.-Q., Chen, Y., Shen, X.-R., Wang, X., … Shi, Z.-L. (2020). A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*, *579*(7798), 270–273. https://doi.org/10.1038/s41586-020-2012-7

Zimmermann, I., Egloff, P., Hutter, C. A. J., Arnold, F. M., Stohler, P., Bocquet, N., Hug, M. N., Huber, S., Siegrist, M., Hetemann, L., Gera, J., Gmür, S., Spies, P., Gygax, D., Geertsma, E. R., Dawson, R. J. P., & Seeger, M. A. (2018). Synthetic single domain antibodies for the conformational trapping of membrane proteins. *ELife*, *7*. https://doi.org/10.7554/ELIFE.34317

Zimmermann, I., Egloff, P., Hutter, C. A. J., Kuhn, B. T., Bräuer, P., Newstead, S., Dawson, R. J. P., Geertsma, E. R., & Seeger, M. A. (2020). Generation of synthetic nanobodies against delicate proteins. *Nature Protocols 2020 15:5*, *15*(5), 1707–1741. https://doi.org/10.1038/s41596-020-0304-x

Zulliger, R., Naash, M. I., Rajala, R. V. S., Molday, R. S., & Azadi, S. (2015). Impaired Association of Retinal Degeneration-3 with Guanylate Cyclase-1 and Guanylate Cyclase-activating Protein-1 Leads to Leber Congenital Amaurosis-1. *Journal of Biological Chemistry*, *290*(6), 3488–3499. https://doi.org/10.1074/jbc.M114.616656