UNIVERSITÀ DI PARMA

UNIVERSITÀ DEGLI STUDI DI PARMA

DOTTORATO DI RICERCA IN
'BIOTECNOLOGIE E BIOSCIENZE'

CICLO XXXVI

Engineered vectors for single-domain antibody selection from immune and synthetic libraries

Coordinatore:
Chiar.mo Prof. Marco Ventura

Tutore:
Chiar.mo Prof. Angelo Bolchi

Dottorando: Ylenia Ciummo

Anni Accademici 2020/2021-2022/2023

UNIVERSITÀ DI PARMA

UNIVERSITÀ DEGLI STUDI DI PARMA

Ph.D. program in
Biotechnology and Life Sciences

XXXVI cycle

Engineered vectors for single-domain antibody selection from immune and synthetic libraries

Coordinator:
Prof. Marco Ventura

Tutor:
Prof. Angelo Bolchi

PhD student: Ylenia Ciummo

Academic Years 2020/2021-2022/2023

'[…] *To foster science, to enjoy its many material and cultural benefits, I believe our society must sustain whole scientific communities,* […] *not try to identify a small cadre of individuals who are somehow specially destinated to make breakthrough discoveries or innovations-or to be awarded Nobel prizes'*

George P. Smith, Nobel Lecture, 2018

Index

Abstract

The revolutionary and serendipitous discovery of heavy chain antibodies in the blood of camelids in 1993 aroused great interest. This discovery has marked a breakthrough in the field of antibodies because the variable domain of this class of antibodies, known as VHH or nanobody or single-domain antibodies, has unique properties. This variable domain is very small (~15 kDa) compared to conventional antibodies (~150 kDa), it is soluble due to the absence of exposed hydrophobic amino acid residues, and it can be expressed individually in bacteria as a recombinant protein while retaining target affinity comparable to that of the whole antibody. These distinctive properties have made single-domain antibodies a versatile tool with a wide range of applications, from basic research to diagnostics and therapy.

During my PhD, I focused on engineering vectors suitable for displaying nanobodies on phages and bacteria. I systematically tested the functionality of these vectors by generating libraries encompassing both synthetic and immune nanobodies.

To get this research project started, I chose an efficient cloning system that could guarantee the creation of highly complex libraries: Gateway® technology.

Commercially available vectors for ligand selection by phage display do not support the use of this cloning method and, as they are mostly phagemid-based, require long times to perform repeated selection steps. I strategically modified an existing phage vector, already adapted for phage display technology, to take advantages of Gateway cloning. This cloning system enabled the construction of a phage library of synthetic humanized nanobodies, using a synthetic sequence encoding a nanobody with a humanized scaffold framework (hCFW) as a template. Despite achieving a sufficiently complex library and enriching it with nanobodies targeting the thioredoxin protein from the bacterium *Pyrococcus furiosus* (*Pf*Trx), subsequent attempts to isolate individual clones specifically directed against the target proved challenging. This limitation may be due to the low binding affinity of the nanobodies in the initial library, which probably requires the inclusion of maturation steps.

The destination genetic cassette (DEST) proposed by Invitrogen for modifying vectors used with the Gateway cloning system is unsuitable for classical phagemid vectors used in phage display. Counter-selection of the destination cassette, which contains the gene encoding the toxin CcdB, proves ineffective in the bacterial strains used for phage display beacause they also carry the gene for the antitoxin (CcdA) on the F plasmid, which can neutralise the toxin. To overcome this issue, I designed and implemented the modification of a phagemid vector by mutagenizing the classic destination cassette. This modification allowed overexpression of the ccdB gene upon induction with the IPTG. Under toxin overexpression conditions, the resulting vector facilitated the selection of recombinant clones, even in normally resistant bacterial strains. To validate the functionality of this new vector, I constructed an immune nanobody library from a llama vaccinated with *Pf*Trx and isolated a specific ligand for this target protein.

Finally, I investigated the potential of the bacterial display system by evaluating three different *Escherichia coli* surface proteins for their suitability in this selection process. I

prepared three expression vectors capable of producing and transporting the three proteins to the bacterial outer membrane. These proteins were fused to a passenger protein, specifically a domain of the FbaB protein from *Streptococcus pyogenes* (SpyCatcher). This setup allowed for the evaluation of the effectiveness of the three autotransporters in delivering a foreign protein to the outer membrane of *E. coli* using the SpyTag-SpyCatcher binding system. After identifying the most effective autotransporter, I transformed the corresponding expression vector into a destination vector by incorporating our modified cassette that overexpressing the CcdB toxin to ensure compatibility with the Gateway technology. Using this Gateway cloning approach, I demonstrated the ability of the vector to display a specific nanobody on the bacterial surface. The next step is to use this vector to generate a nanobody library and perform selection through bacterial display methodology.

Introduction

Introduction

Over the decades, the interest in antibodies in the scientific community has grown exponentially. Their intriguing features make them attractive as ideal molecules for therapeutic and diagnostic applications due to the high affinity and specificity for their targets[1].

The introduction of hybridoma technology in 1975 by Kölher and Milstein[2] was revolutionary. Since then, it has been possible to isolate monoclonal antibodies (mAb) efficiently. This breakthrough has significantly contributed to addressing research, therapeutic, and diagnostic challenges. Since the approval of the first therapeutic mAb muromonab-CD3 (Orthoclone OKT3) by the Food and Drug Administration (FDA) in 1986, the field of therapeutic antibody development and its market have undergone an important evolution[3]. Currently, nearly 1200 antibody therapeutics are involved in clinical studies and nearly 75 are in regulatory review or approved, as evidenced by the annual report of commercial mAb therapeutics[4].

However, the problems related to the hybridoma method and advancements in molecular biology and immunology[5], along with ethical concerns related to immunization[6–8], have led to the development of recombinant antibodies and antibody fragments. In particular, the introduction of antibody fragments, such as Fab and scFvs, combined with the power of phage display technology in the late 1980s, opened new opportunities to isolate engineering antibodies for therapeutic purposes[9,10].

1 Heavy chain antibodies

In the wake of recombinant antibody applications, in 1989 a serendipitous discovery by Raymond Hamers-Casterman, Serge Muyldermans and colleagues of the Vrije Universiteit Brussel (VUB) was revolutionary.

In a short communication, they demonstrated the presence of special IgG antibodies in the serum of camelids, in addition to the conventional heterotetrameric IgG. They observed that two classes of IgG were completely devoid of the light chain (L), as no separation between the heavy and light chains was observed after treatment with reducing agents, and also lacked the first constant domain (CH1) (**Figure 1**). They also demonstrated that these new classes could generate a broad repertoire of antibodies in the event of an immune response from the animal[11].

They called these antibody variants heavy chain antibodies (hcAbs) and these findings opened a new era in antibody engineering.

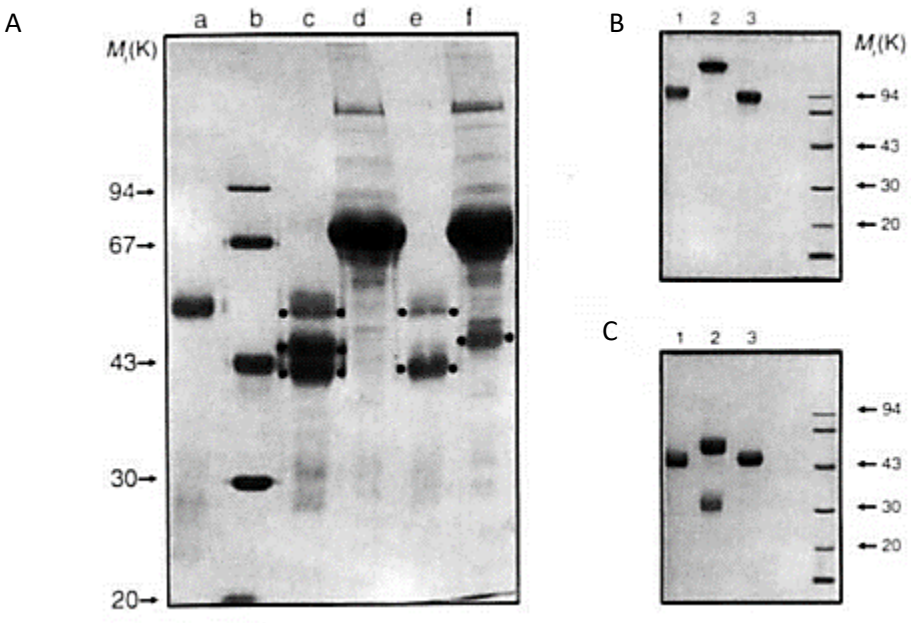


Figure 1 Purification of camel IgG subclasses. A. Adsorption of *C. dromedarius* serum on protein A shows the presence of three IgG subclasses: 50, 46 and 43 kDa (lane c). The three components are absent in the non-adsorbed fraction (lane d). Adsorption on protein G shows the 50 kDa and 43 kDa components (lane e) while the 46 kDa component is present in the non-adsorbed fraction (lane f). Size marker (lane b). B. Analysis of IgG fractions containing three subclasses 43 kDa (lane 1), 46 kDa (lane 2), and 50 kDa (lane 3), on SDS-PAGE in the absence of dithiothreitol (DTT). C. Analysis of the same subclasses on SDS-PAGE in the presence of dithiothreitol (DTT). Figure adapted from reference [11].

1.1 Characterization of heavy chain antibodies

Heavy chain antibodies are present in the biological family of *Camelidae,* in particular in Old World camelids including *Camelus dromedarius* and *Camelus bactrianus* and in New World camelids including *Lama glama*, *Lama pacos*, *Lama guanicoe*, and *Lama vicugna.* The percentage of the hcAb and the conventional IgG in the serum of camelids is variable. A higher ratio, 50-80%, was observed in camels, while a lower value was noted in South American species, ranging from 10-25%[12]. Such a significant value underscores the importance of hcAb in the immune protection of camelids.

These antibody variants belong to the immunoglobulin γ antibody class, specifically to the IgG2 and IgG3 isotypes based on their differential affinities to protein A and protein G, respectively[13].

The origin of functional hcAb in camelids can be explained by analysing the IgH gene locus. The structural organization of this locus consist of multiple variable, diversity, joining and constant genes. In IGHG genes encoding IgG heavy chain antibodies, a nucleotide G-to-A point mutation disrupts the consensus splicing site (GT) at the 5' end of the intron between the CH1-hinge exons[14]. This mutation results in the elimination of the CH1 domain by splicing, explaining the absence of light chains.

Normally heavy and light chains of antibodies are expressed in the endoplasmic reticulum (ER). The nascent heavy chain is retained in the ER by binding BiP protein

(heavy chain binding protein), which associates with the first constant domain CH1 until a light chain replaces BiP, resulting in mature antibody secretion. BiP plays a central role in ER, and actually is involved in the translocation, secretion, quality control and degradation of secreted proteins[15]. The absence of CH1 domain in hcAbs causes a failure in BiP binding and results in their secretion without the light chain.

As mentioned above, the H locus harbours dedicated germ line genes, from which the variable domains of hcAb originate, called IGHVH, and interspersed with IGHV germ line genes of conventional antibodies[16]. The variation sequence of variable domain is introduced during B-cell lymphopoiesis in which both genes (IGHVH and IGHV) are involved in V-D-J. In this process, one IGHVH/IGHV is rearranged with one IGHD gene and one IGHJ gene for the assembling of VHH or VH domain[17,18] (**Figure 2**).

D gene and J gene pool are common for both variable domains, while IGHVH genes and IGHV genes differ in their sequence. Actually, IGHVH genes contain certain sequences (e.g. palindromes) that are unstable and may have led to faster gene evolution and to a rapid expansion of the IGHVH repertoire[19].

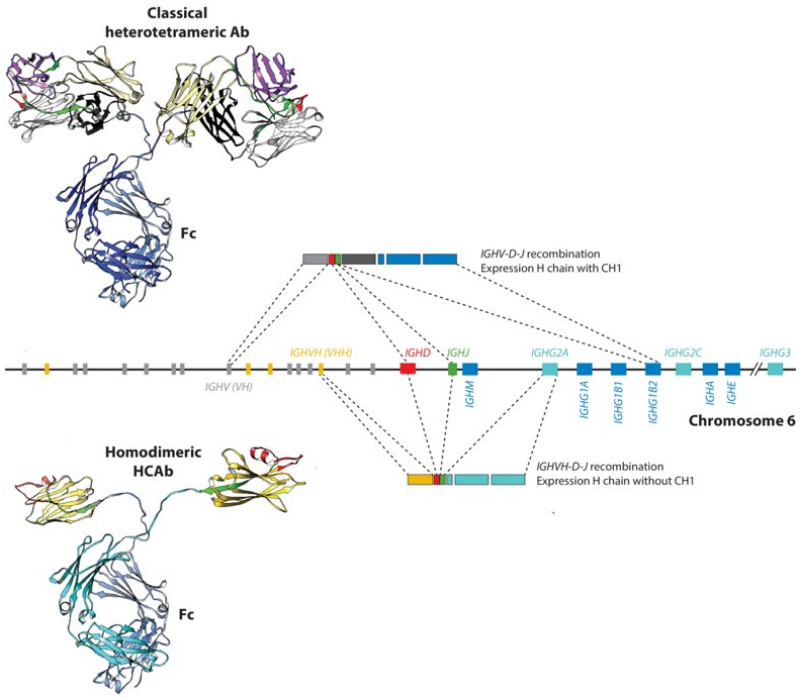


Figure 2 Schematic representation of IgH locus genes of Chromosome 6 in Bactrian camels: IGHV and IGHVH genes upstream of IGHD and IGHJ genes; downstream genes encoding constant heavy chain domains of various immunoglobulins. H-chain expression in the classical heterotetrameric antibody after recombination of IGHV D-J genes in B cells (upper) and in homodimeric hcAb (bottom) when an IGHVH gene is involved in the recombination. Figure adapted from reference [34].

The *Camelidae* family is not the only one with heavy chain antibodies in immune repertoire. In 1995, a new antigen receptor (NAR) was discovered in the nurse shark (*Ginglymostoma cirratum*)[20]. Structural analysis showed that NARs have a similar structure to hcAbs: two heavy chains covalently associated and the absence of light chains. Two heavy chains consist of six domains, five of which are constant (CNAR1-5)

and one is variable (VNAR). Moreover, since this molecule shares several functional features with immunoglobulins, it was named IgNAR[21]. It has been demonstrated that following immunization an IgNAR response is induced with a similar kinetic as in IgM response[22]. The question arises: why do two such different species produce similar antibodies? Phylogenetic analysis has shown that hcAbs in camelids emerged and diverged from conventional antibodies approximately 25 million years ago after *Tylopoda* split from other mammals and before the camel and llama speciation[23]. The origin of shark IgNAR is poorly understood, but IgNAR is found in all elasmobranchs and thus emerged at least 220 million years ago probably from IgW isotype[24,25]. The analysis clearly shows the evolutionary distance between the two species and the presence of heavy chain antibodies is due to convergent evolution on the basis of certain requirements: the first constant domain (CH1) should be absent or modified, a repertoire of single V domain must be generated and the selected V domain should be soluble[24].

Finally, heavy chain antibodies have been identified in human and mouse sera, but in this case their presence is associated with pathological disorders[26].

1.2 Structural features of heavy chain antibodies

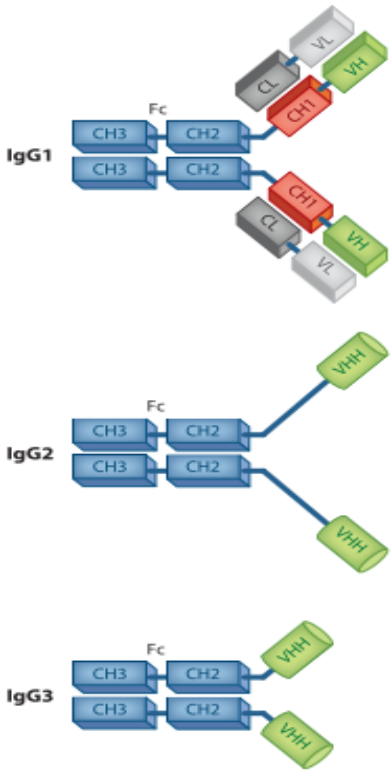


Figure 3. Schematic representation of conventional antibody (IgG1) and two types of homodimeric heavy chain antibodies (IgG2 and IgG3) subclasses of in sera of camelids. Figure adapted from reference [17].

IgG antibodies have a conserved mammalian heterotetrameric architecture composed of two identical heavy (H) and two identical light (L) polypeptide chains and stabilized by disulphide bonds. Papain digestion allows the identification of two Fab fragments (antigen-binding fragments), each of which binds the antigen and contains the variable (VL) and constant (CL) domains of the L chain and the first two variable (VH) and constant (CH1) domains of the H chain, and a single Fc fragment (fragment crystallizable) containing the constant domains (CH2 and CH3) of the H chain[27].

HcAbs exhibit the same Fc region of the conventional IgG antibodies: two constant domains CH3 and CH2 conserved in sequence and structure. Each domain folds into a constant structure of a three-strand-four strands β sheet joined by an intrachain disulfide bond. This region is important for effector functions. The absence of the CH1 domain in hcAbs results in a reduced molecular mass compared to conventional antibodies (90 kDa rather than 150 kDa) and the variable domain is connected to constant domains by a hinge region. The hinge region of IgG2 is longer than that of IgG3. It contains Pro-Gln repeats that probably replace the CH1 domain[17] (**Figure 3**). The loss of the light chains results in the variable region

of the antibodies having a single domain called VHH, which retains full antigen-binding capacity (**Figure 4**).

The folded VHH domain comprises nine β-strand (A-B-C-C'-C''-D-E-F-G) organized in a 4-strand/5-strand β-sheet connected by loops and by a conserved disulphide bond between Cys23 and Cys94. Specifically, the sequence of β-sheet is conserved and divided into four conserved structural framework regions (FR) surrounding three hypervariable domains, the loops, called complementary-determining regions (CDR). CDR regions are interested in the interaction with the antigen.

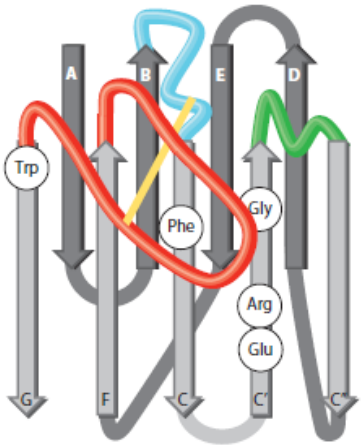


Figure 4. Schematic structural organization of a VHH domain: A-B-C-C'-C''-D-E-F-G β-strands represent framework regions (FRs) in grey; loops indicate CDR 1-2-3 in blue, green and red respectively. Figure adapted from reference [17].

As described, the structural organization of VHH is similar to VH of conventional antibodies with notable differences in framework 2 (FR2) and in CDRs. In VH FR2 conserved hydrophobic amino acids (Val47, Gly49, Leu50, Trp52) form a hydrophobic surface to facilitate the interaction with the VL (variable domain of light chain); instead in VHH, these residues are substituted by hydrophilic or smaller amino acids (Phe42, Glu49, Arg50, and Gly52) to render the domain more soluble in the absence of VL partner[28]. Another difference concerns the three CDR loops. CDR1 and CDR3 have longer loops than in conventional antibodies to compensate for the lack of VL domain and the three corresponding CDRs[29]. This aspect provides a large antigen-interacting surface of 600-800 A^2 comparable to the classical interaction surface VH-VL (600-900 A^2). The enlarged CDR1 derives from the germ line genes while the longer CDR3 is due to the selection during B cells maturation, in particular, the use of more than one D gene segment during V-D-J recombination[17]. An extended loop implies greater flexibility, and this is entropically counterproductive for antigen binding. To stabilize the CDR3 loop, many camel VHH contain an extra pair of cysteine that forms an interloop disulfide bond[30]. However, the presence of the extra disulfide bond may vary from species to species: in dromedary is present between CDR1 and CDR3 and in rare cases between CDR3 and FR2; in llama it is less frequent because their CDR3 are shorter even if there has been noticed a disulfide bond between CDR2 and CDR3[31].

The analysis of the V gene in alpaca identified three subgroups IGHV1, IGHV2, and IGHV3 based on their degree of homology with human IgHV clans I, II, III. Up to a few years ago, all VHH genes were reported to be homologous to human IGHV3 genes[16]. Subsequently, was demonstrated the presence of IGHV1 and IGHV4 genes in alpaca and llama[32].

2 From serendipitous discovery to single-domain antibodies

Since the first publication on the natural hcAbs in camelids in 1993, considerable efforts have been directed toward understanding the various aspects of hcAbs, spanning their biology, evolution, immunogenetics, and potential applications in research and medicine.

The single variable antigen-binding domain (VHH) of hcAbs has captured the attention of many researchers and opened a new era of antibody engineering. Among the interesting features described below, one has aroused the greatest curiosity, namely the possibility of expressing only the single variable domain in a recombinant manner. One aspect not to be underestimated is the ability to maintain the same binding affinity as a classical antibody, despite being only one-tenth of its size.

To address challenges associated with the use of mAbs, the field of single-domain antibodies has grown exponentially. It is not a coincidence that Muyldermans and colleagues, in 2001, after a decade since the discovery of hcAbs, founded a spin-off of the Vlaams Interuniversitair Instituut voor Biotechnologie (VIB), named Ablynx, specialized in the development and production of Nanobodies[33], a registered trademark of the company.

Since 2018, Ablynx is part of Sanofi, boasting over than 500 granted patents regarding the discovery, generation, optimization, formatting, manufacture, administration, formulation and clinical use of Nanobodies.

2.1 Single domain antibodies: pros and cons

The sdAbs have represented a great promise due to the unique characteristics of their structure.

Firstly, the small size (approximately 15 kDa) and an extended CDR3 contribute to forming a singular paratope with a convex surface capable of binding the concave epitopes of antigens often inaccessible for classical antibodies[34]. An example of this is their association with enzyme catalytic sites, modulating catalytic activity or stabilizing conformation[17,35]. Despite their small size and the presence of only three CDRs, sdAbs exhibit antigen specificity and affinity comparable to mAbs, typically in the nanomolar or picomolar equilibrium dissociation constant range[17]. Additionally, their small size allows better tissue penetration, particularly in tumor tissues where conventional antibodies may struggle[36], and facilitates crossing the blood-brain barrier[37].

SdAbs also possess interesting biochemical characteristics. They are robust molecules and resist under stringent conditions, chemical and thermal denaturation. For instance, sdAbs denature in 2-3 M guanidinium solution and at 60-80°C[38] and show the capacity to renaturate subsequently. This thermal stability is due to the extended CDR3 and can be enhanced by the addition of a disulphide bond and specific amino acid mutations in the N-terminal sequence[39,40]. In addition, sdAbs have a good shelf life: they can be stored for months at 4°C and -20°C, and for several weeks at 37°C, while retaining close to 100% antigen binding activity; they can tolerate exposure to pH changes remaining stable in the presence of proteases[41].

Another advantage of sdAbs concerns the simplicity of production and purification. Typically, they are expressed in simple bacterial systems, such as *Escherichia coli*, in using a secretion signal that enables expression in the periplasm of the bacterium, where the oxidizing environment favours the correct folding of the protein and disulfide bond formation. Usually, yields of several milligrams per liter can be obtained from a single flask of culture. While cytoplasmatic expression is reported, the reducing environment of the bacterial cytoplasm could prevent the formation of disulphide bonds and thus proper sdAb folding[42,43]. Inclusion bodies may also be encountered, necessitating an in vitro refolding protocol, although not all refolded protein is functional in antigen binding. SdAbs can also be produced by eukaryotic expression systems such as yeast (i.e. *Saccharomyces cerevisiae*, *Pichia pastoris*)[44,45], mammalian cell lines[46] and plants[47].

A notable property that makes nanobodies candidates for drug development is their low immunogenicity. This aspect is due to several intrinsic features: their small size, which decreases the number of immunogenic epitopes, rapid blood clearance and high sequence identity with human IGHV3 family comparable to humanized murine VH domains[48]. However, given the camelid origin of sdAbs, the immunogenicity risk of these proteins is a potential obstacle. To mitigate the risk of adverse reaction of immune system the humanization process is to evaluate[49]. Nevertheless, a recent study of two non-humanised monomeric sdAbs for human applications demonstrated their low immunogenic risk profile, encouraging the development of sdAbs as potential drugs[50].

Therefore, their modular and monomeric nature makes them easy to engineer into a multi-domain format to increase affinity or avidity for antigens and use in the treatment of disease[51,52].

Despite the many advantages, there are also limitations to consider. The single-domain nature of about 110-130 amino acids makes each residue more significant, requiring careful study during engineering to preserve stability; For example, complete humanization often compromises the antigen binding affinity, sdAb stability and expression yield. Other disadvantages include their short serum half-life and the complete lack of effector function, which affect the pharmacokinetic of the therapeutic antibodies. Various strategies can be applied to overcome these limitations: polyethylene glycol (PEG)-ylation, fusion to human serum albumin or an anti-serum albumin to increase the retention in the serum and conjugation to Fc region to enhance effector function[53].

2.2 Single domain antibodies generation

Antigen-specific sdAbs can be retrieved by three different sources: immune, naïve and synthetic libraries.

The widely used approach to obtain target-specific, high-affinity sdAbs involves the construction of an immune library. The first step is to immunise an animal of the *Camelidae* family with the antigen of interest for a period of about 2 to 3 months. Follow the collection of an aliquot of anti-coagulated camelid blood (50-100 ml) and the extraction of mRNA that is converted by reverse transcription in cDNA. This cDNA is used to amplify VHH repertoire, which is then ligated into any vector of interest and

transformed into an appropriate expression system. A good immune sdAbs library should contain 10^6-10^7 individual transformants[54]. Immune libraries facilitate the isolation of high-affinity antibodies because the animal's immune response allows affinity maturation in vivo. The quality of the antigen used for immunization is crucial; soluble, properly folded recombinant protein is preferred, although, in some cases, DNA immunization may be a valid option[55,56]. If the antigen is difficult to obtain, transgenic mice capable of producing antigen-specific hcAbs can be used[57].

In cases where immunization is not feasible, for instance, if the antigen is toxic, pathogenic, harmful to the animal or environment, or low immunogenic, solutions are offered by naive or synthetic libraries.

The preparation of naïve library starts from a large pool of blood more than 1 L from several non-immunized animals. The lack of somatic maturation *in vivo* due to immunization makes sdAb libraries nonspecific, necessitating large-sized libraries (10^9-10^{10} individual clones) in order to retrieve high-affinity binders. Often, in vitro maturation techniques, such as DNA shuffling, error-prone PCR, and randomized primer, are required[58,59].

Synthetic libraries are constructed completely *in vitro* and represent an attractive alternative to circumvent animal use. A distinction must be made between synthetic and semi-synthetic libraries. The main difference lies in the design of their construction: in a semi-synthetic library, a characterized sdAb is used as starting point and the CDR3 is randomized according to a specific design; in a synthetic library, a stable and well-expressed scaffold is selected and preferably one of which the crystal structure is available[60]. In general, the two main factors to consider when designing libraries are the framework and the CDRs. Framework selection is focused on stability; actually, a reported universal or sdAb-derived framework or a framework generated from consensus sequence derived from the natural repertoire can be used for the library construction. CDRs design is focused on variability considering the natural diversity of natural repertoire or manipulating the sequence partially or completely by randomization[60,61]. Also, in this case large libraries of different clones are a necessary to compensate for the lack of somatic maturation[54].

Although various design strategies for synthetic libraries are reported in literature, the most detailed study on the construction and validation of synthetic sdAb library is the one conducted of Moutel and co-workers in 2016. Starting from a highly stable and robust sdAb scaffold, they introduced mutations in the framework regions substituting seven amino acids to make them more similar to the human VH3; then, they introduced variability in CDRs. The length of CDR1 and CDR2 was kept constant, using amino acids that reflect natural diversity at each position. For CDR3, instead, all positions were randomized and to cover natural variability of camelid sdAb, variable lengths (9, 12, 15 and 18 amino acids) were used. A large library size was obtained ($3x10^9$ individual recombinant clones) and high-affinity binders, with a constant dissociation from 50 pM to 10 nM, were isolated[62].

Multiple display techniques have been developed to retrieve the best antibody from large libraries. The oldest and most robust selection method is phage display. It was first described by George Smith[63] in 1985 and later adapted for scFv by Sir Gregory Winter[64].

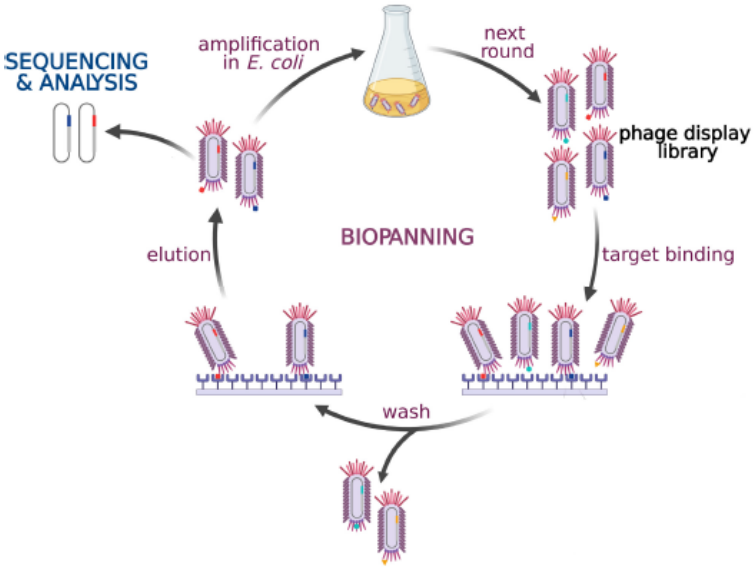


Figure 5. Schematic representation of biopanning. Figure adapted from reference [78].

This method involves the fusion of proteins and peptides to coat proteins on the phage surface. The power of phage display technology also arises from its capability to connect the displayed molecule (phenotype) with DNA sequence that encodes it (genotype). Studies on this molecular technique led to the two scientists being awarded the Nobel Prize in 2018: 'for the phage display of peptides and antibodies'[65]. The library is typically cloned into a phagemid vector in frame with the sequence encode a coat phage protein, gIIIp or gVIIIp, and transformed in an *E. coli* strain. In order to obtain phages that display sdAbs on the surface, an infection with M13 helper phage is necessary. After that, phage particles are used for biopanning, where they are incubated in a microtiter plate with the immobilized antigen of interest. Only those virions carrying a target-specific sdAb are retained and recovered for further pannings, while unbound phages are washed out. After multiple rounds of panning, usually 3-5 depending on the nature of the library, the single specific binders are analyzed by an enzyme-linked immunosorbent assay (ELISA) and the high-affinity sdAb candidate is sequenced and used for further analysis[66] (**Figure 5**). This is the common general strategy for phage display-based sdAbs or antibodies discovery, but it is possible to use different and optimised advanced strategies to improve antibody selection[67].

Alternative screening methods have been developed that allow the exposure of antibodies on the cell surface. These include yeast and bacterial display.

Yeast display enables the selection of libraries by fluorescence-activated cell sorting (FACS) providing high-throughput screening of large surface-displayed libraries and the biophysical characterization of individual displayed antibodies without the need for their subcloning[68,69]. Furthermore, since the expression system of yeasts is similar to that of mammals, the correct folding of proteins is guaranteed[70].

Bacterial display system consists of three main factors: the host cells to bind fusion proteins; the carriers (i.e. outer membrane proteins) that contain a signal peptide to transfer the passenger from intracellular location to cell surface, and the passenger (the target foreign protein displayed on the cell surface)[71]. The most common bacterium employed in this system is *E. coli* due to the availability of genetic tools and the high transformation efficiency, although other bacteria, especially some Gram-positive species, have been used[72,73]. A large family of proteins secreted by Gram-negative bacteria and involved in the selection of sdAbs library on *E. coli* surface is represented

by autotransporters (ATs)[74]. Large cell surface allows easy selection and sorting of displayed libraries. In particular, the selection of high affinity binders can be performed by FACS or MACS, even though in bacterial display is preferred the use of MACS[75–77].
Display platforms described above, known as cell-based display, have certain limitations due to the molecular cloning and transformation procedures. This implies a decrease in library complexity and consequently a restriction of library diversity. For these reasons, cell-free display systems that do not require cloning procedures have been developed, including ribosome display and CIS display[78].
Ribosome display shares the common concept to the other display system: the association of the phenotype and genotype with the generation of non-covalent ternary complex (mRNA-ribosome-polypeptide). DNA libraries are initially transcribed into mRNA lacking the stop codon, ensuring that the synthesized peptide and its encoding mRNA are not released from the ribosome. The ribosomal complexes are used for affinity selection and subsequently mRNAs are dissociated from the complexes, recovered and converted into cDNA for a next cycle of enrichment or analyzed[79].
CIS display is a library selection system that connects an expressed library to its own DNA sequence and it is based on the bacterial initiator protein RepA, which is able to bind to the DNA from which it has expressed. A DNA template that encodes the peptide library is fused to a DNA fragment encoding RepA. After *in vitro* transcription and translation, RepA locates its own binding site on the template from which it was transcribed and the expressed library is fused to its coding through RepA binding and can be easily sequenced[80,81].

2.3 Applications of single domain antibodies

Over the past 30 years, the field of sdAb applications has grown considerably as its favourable properties facilitate its production. Today sdAbs are employed in research, diagnostic and therapy.

2.3.1 Single domain antibodies as research tools

One of the interesting sdAb application is their key role in the structural characterization of their antigen. They are considered crystallization chaperones thanks to rigid folding[82]. In particular, sdAb is able to reduce the antigen mobility and to fix the conformational variants of the target facilitating the crystallization of the complex. Currently, more than one hundred cocrystal of sdAb-antigen have been solved, including important example involving sdAb against G-protein-coupled receptors (GPCRs). The first structural determination of the active β_2 adrenoreceptor was made by identification of nanobody 80[83,84] (**Figure 6**). After this discovery, several sdAbs have been developed and used as active-state chaperones for a variety of G-coupled receptors[85–88].

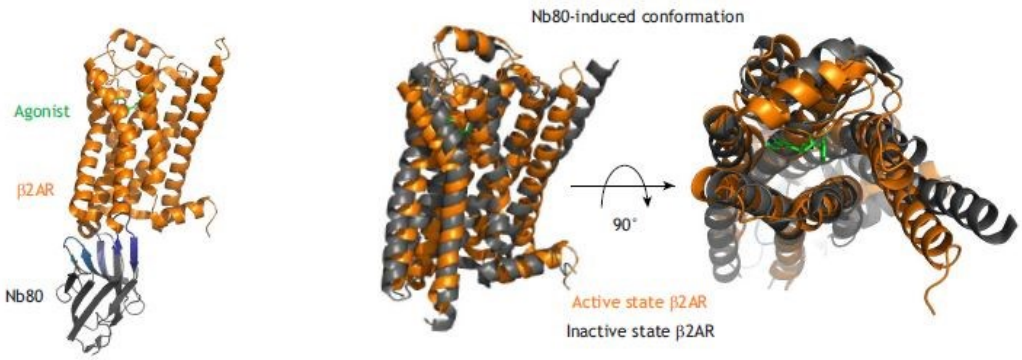


Figure 6. Example of a nanobody as a conformational chaperone: Nb80 stabilises the active state of the β_2 adrenoreceptor (β_2AR) (left). Comparison of the active state induced by Nb80 (orange) and the inactive state of (β_2AR) (gray) (right). Figure adapted from reference [84].

Recently, cryogenic electron microscopy (cryo-EM) has emerged as an attractive method in protein structure determination[89]. Because structure determination is difficult for small proteins (< 100 kDa) in cryo-EM analysis, sdAbs are used as a fiducial marker[90]. Several examples are reported in literature in which sdAbs can be engineered to assist structural analysis of difficult proteins. Regarding GPCRs for example, a sdAb (Nb6) has been used to obtain cryo-EM structures of inactive-state GPCRs[91]. Another approach involves the use of sdAbs as fiducial markers in more complex scaffolds called Legobodies[92].

Due to their property to access to various cell compartments, sdAbs can be applied as intrabodies allowing the visualization and/or modulation of endogenous targets but also quantification of changes in the level endogenous proteins in living cells[93,94]. One challenge in working with intrabodies is their delivery into living cells. There are two options: transferring sdAbs-encoding gene into an expression vector or introducing them as purified protein. To avoid genetic manipulation of cells, delivery of intrabodies as proteins is preferred. A widely used method to facilitate intrabodies cellular uptake is coupling them to cell-penetrating peptides (CPPs)[95].

One of the applications of intrabodies is their fusion to a fluorescent protein, allowing the imaging of antigens in living cells. In particular, the intrabody can be genetically coupled to GFP o RFP obtaining fluorescent probes known as chromobodies (Cbs)[96]. The work of Rothbauer is the first example of fusion of a sdAb to the monomeric red fluorescent protein (mRFP), demonstrating the capacity of chromobodies to recognize and trace antigens in different cellular compartments [96]. After that, many target specific chromobodies has been reported in the literature. These include actin-Cb used as cytoskeletal imaging probes for the visualization of cytoskeletal remodeling during the different developmental stages in zebrafish embryos[97] and the monitoring neuronal actin plasticity in mouse[98], and β-catenin Cb to monitoring the induction of Wnt/ β-catenin signal pathway[99,100].

SdAbs can be used as guides in targeted protein degradation. The fusion of an anti-GFP sdAb to the F-box domain of the ubiquitination system causes the targeted degradation of GFP-tagged proteins[101]. This system applies to various cells and organisms, such as mammalian cells[102], plants[103], and zebrafish[104]. Similarly, joining an anti-GFP sdAb to the

DNA binding domain and a second anti-GFP sdAb to the activation domain of the transcription factors might lead to the activation of the reporter gene only in the presence of the antigen (GFP) expression in the cell[105].

2.3.2 Single domain antibodies as diagnostic tools

Specificity, robustness and low production costs make sdAbs suitable for *in vitro* and *in vivo* diagnostic applications.
sdAbs variants can be adapted for their use in *in vitro* assays: the fusion with tags (i.e. myc tag, His6 tag) or enzyme (i.e. alkaline phosphatase, horseradish peroxidase, nanoluciferase) allows their use for target recognition in ELISA, western blot and flow cytometry experiments[106,107]. sdAbs-based sandwich ELISA is widely used to detect biomarkers proteins, viruses and pathogens, while competitive inhibition ELISA is preferred to monitor the presence of pesticides on food.
Excellent results were obtained for sdAbs applications in *in vivo* imaging. Actually, the small size, ease of modification, rapid tissue penetration and rapid renal clearance make them the best tracers for *in vivo* diagnostic methods[108]. For example, if a labelled sdAb is intravenously injected, it rapidly extravasate and diffuse in tissue to associate with its target without sticking to non-targeted cells. Moreover, the excess is cleared by the kidneys, allowing the decrease of background signal[109,110]. For these reasons, sdAbs seem to be interesting tools for SPECT (single-photon emission computerized tomography) or PET/CT (positron emission tomography/computerized tomography) imaging. Many studies demonstrate their application as radioactive tracers in PET/SPECT. A successful sdAb probe for PET imaging is directed against HER2, a common breast cancer biomarker, and conjugated to Ga-68[111]. This probe completed phase 1 clinical trial and now is undergoing a phase 2 trial[112,113]. The field of application has extended to cardiovascular disease. A sdAb probe labelled with Tc-99m for SPECT imaging was developed against vascular cell adhesion molecule-1 (VCAM-1), an atherosclerotic lesions biomarker. This study makes sdAb as potential radiotracers for cardiovascular applications[114].
The sdAbs properties mentioned above and in addition the short half-life *in vivo* opened the opportunity to obtain sdAbs for brain imaging. Actually, despite the blood brain barrier (BBB) poses a challenge for the delivery of protein with a molecular weight greater than 400 Da in brain[115], some examples of sdAb probes for *in vivo* are reported[116,117].

2.3.3 Single domain antibodies as therapeutics

The beneficial properties of sdAbs, comparable to those of mAbs, make them attractive agents for the development of novel therapeutic applications. The interesting in using these antibodies as potential therapeutics started with the foundation of Ablynx spin-off. Since then, despite some limiting properties such as short serum half-lives and rapid renal clearance, which are undesirable in therapy, many strategies have been developed to increase the persistence of sdAbs in the blood: conjugation to Fc domain of conventional antibodies, polyethylene glycol (PEG)-ylation, coupling to abundant serum

protein (i.e. albumin)[53]. This has led to identification of several sdAbs, many of which have entered clinical trials. In 2018, the EMA's approval of Caplacizumab (by Ablynx-Sanofi) for the treatment of thrombotic thrombocytopenic purpura (TTP) was the first demonstration of the power of sdAbs as drugs. In 2019, it was approved by the FDA and entered the market[118–120]. To date, two additional antibodies have been approved for clinical use: Ozoralizumab, a trivalent anti-tumour necrosis factor alpha sdAb developed by Taisho Pharmaceuticals (under license from Ablynx) for the treatment of rheumatoid arthritis, was approved in Japan in September 2022[121]; Ciltacabtagene, the first sdAb-based CAR-T for the treatment of multiple myeloma, was approved by FDA in February 2022[122].

As evidenced by the approved sdAbs, today sdAbs are characterised against many diseases.

In cancer therapy, sdAbs are used as drugs in several approaches thanks to their ability to penetrate tumour tissues. One promising avenue concerns the engineering of chimeric antigen receptor (CAR) T-cells expressing specific sdAbs against tumour antigens[123]. After the transfer into patients, modified T-cells are able to bind tumour cells and the sdAb is linked to the intracellular signal transduction domains by a hinge spacer allowing the activation of immune response[124]. In addition to Ciltacabtagene mentioned above, other studies have been carried out to develop Nb-expressing CAR T cells that target other tumour biomarkers, such as EGFR[125] and HER2[126].

Furthermore, sdAbs can be used as antagonists: they control cells proliferation targeting overexpressed receptors on tumor cells and suppressing signal transduction. Examples are sdAbs against HER2, EGFR and HGF, that showed an inhibition of solid tumor growth *in vitro* and *in vivo*[127–129]. SdAbs can also inhibit metastasis targeting receptors involved in angiogenesis[130,131] or can be used to bind immune checkpoints that are up-regulated by tumor cells[132].

The loss of the fc region should not necessarily be seen as a disadvantage, as it can provide insights into the study of new functions, such as the development of immunotoxins. Actually, it is reported that the fusion of a sdAb against VEGFR2 to *Pseudomonas* exotoxin A is able to kill cancer cells[133].

SdAbs are involved also in radionuclide therapy whereby a radiolabaled sdAb can accumulate at the tumor site inducing the destruction of cancer cells[134]. An interesting example is anti-HER2 sdAb linked to I-131 for the treatment of breast and gastric cancer that is now undergoing phase 2 clinical trials[135]. An alternative approach is photodynamic therapy (PDT), a minimally invasive and non-toxic treatment, that induces cell death by activating a sdAb-conjugated photosensitizer by light exposure[136].

Another important area of application for sdAbs is infectious diseases caused by viruses, bacteria and parasites.

Regarding viruses, the recent outbreak of the SARS-coV-2 has shown versatility of sdAbs. Actually, many groups have focused on the development of sdAbs against the receptor binding domain (RBD) of the spike protein of the virus, the ideal target in order to obtain therapeutic agents. In a few years, efforts to find candidates to block or neutralize the virus entry led to a new and constant publications. The first neutralizing sdAbs against SARS-CoV-2 initially targeted MERS-CoV and SARS-CoV-1 but exhibited cross-reactivity

and neutralizing efficacy against SARS-CoV-2[137] . These were followed by sdAbs isolated from synthetic or llama immune libraries[138–141]. Furthermore, it was demonstrated that neutralizing sdAbs are able to recognize several independent epitopes of spike protein highlighting their potential even in the case of mutate virus forms[142,143].

Against influenza virus sdAbs were developed; Cardoso et al. reported an example on the isolation of a bivalent anti-H5N1 NA sdAb that showed an *in vitro* antiviral potency higher than monovalent sdAb[144].

For the treatment of respiratory syncytial virus, ALX-0171 sdAb showed interesting results in animal models but failed to produce significant results in phase 2b clinical trials[145].

Promising sdAbs can be obtained also for the development of therapeutics against bacteria and bacteria-produced toxins. Usually, sdAbs to combat bacteria are obtained against bacterial surface proteins to block bacterial infection in the host cell. Based on this principle, three high affinity sdAbs against the internalin B (InIB) protein of *Listeria monocytogenes* contribute to the *in vitro* neutralization of bacterium infection[146]. Another interesting example is bispecific sdAb for Shiga Toxin-producing *E. coli* (STEC) infections: two sdAbs against the Shiga toxin-2 B subunit (Stx2B) and the C-terminus of Intimin (IntC280) were genetically fused and the bispecific format demonstrated the ability to neutralise Stx2B or IntC280 binding to the respective host receptors[147].

It is possible to isolate sdAbs against parasites: sdAbs against variant surface glycoprotein (VSG) of *Trypanosome brucei* can induce parasites killing[148]; two high affinity sdAbs specific for Pfs230, a surface protein expressed during the sexual stage development of *Plasmodium falciparum*, are able to significantly block malaria parasite transmission and decrease the formation of exflagellation centers[149].

A final area of sdAbs application is the treatment of snake and scorpion envenoming. Actually, various formats have been reported[150–152]. In this case it is very important the isolation of high affinity toxin-neutralizing sdAbs, a rapid extravasation intravenously of sdAb and a good tissue penetration.

3 Bibliography

1. Wu H, An LL. Tailoring kinetics of antibodies using focused combinatorial libraries. *Methods in molecular biology.* 2003; 207:213-233. doi:10.1385/1-59259-334-8:213
2. Köhler G, Milstein C. Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature.* 1975; 256:495-497. doi:10.1038/256495a0
3. Leavy O. Therapeutic antibodies: past, present and future. *Nature Reviews Immunology.* 2010;10(5):297. doi:10.1038/nri2763
4. Kaplon H, Crescioli S, Chenoweth A, Visweswaraiah J, Reichert JM. Antibodies to watch in 2023. *MAbs*. 2023;15(1). doi:10.1080/19420862.2022.2153410
5. Chames P, Van Regenmortel M, Weiss E, Baty D. Therapeutic antibodies: successes, limitations and hopes for the future. *Br J Pharmacol*. 2009;157(2):220-233. doi:10.1111/J.1476-5381.2009.00190.X
6. Bradbury ARM, Dübel S, Knappik A, Plückthun A. Animal- versus in vitro-derived antibodies: avoiding the extremes. *MAbs*. 2021;13(1). doi:10.1080/19420862.2021.1950265
7. Gray A, Bradbury ARM, Knappik A, Plückthun A, Borrebaeck CAK, Dübel S. Animal-free alternatives and the antibody iceberg. *Nature Biotechnology*. 2020;38(11):1234-1239. doi:10.1038/s41587-020-0687-9
8. Laustsen AH, Greiff V, Karatt-Vellatt A, Muyldermans S, Jenkins TP. Animal Immunization, in Vitro Display Technologies, and Machine Learning for Antibody Discovery. *Trends Biotechnol*. 2021;39(12):1263-1273. doi:10.1016/J.TIBTECH.2021.03.003
9. McCafferty J, Griffiths AD, Winter G, Chiswell DJ. Phage antibodies: filamentous phage displaying antibody variable domains. *Nature*. 1990;348(6301):552-554. doi:10.1038/348552a0
10. Hoogenboom HR. Selecting and screening recombinant antibody libraries. *Nature Biotechnology*. 2005;23(9):1105-1116. doi:10.1038/nbt1126
11. Hamers-Casterman C, Atarhouch T, Muyldermans S, et al. Naturally occurring antibodies devoid of light chains. *Nature.* 1993; 363:446-448. doi:10.1038/363446a0
12. Blanc MR, Anouassi A, Abed MA, et al. A one-step exclusion-binding procedure for the purification of functional heavy-chain and mammalian-type γ-globulins from camelid sera. *Biotechnol Appl Biochem*. 2009;54(4):207-212. doi:10.1042/ba20090208
13. Daley LP, Gagliardo LF, Duffy MS, Smith MC, Appleton JA. Application of monoclonal antibodies in functional and comparative investigations of heavy-chain immunoglobulins in new world camelids. *Clin Diagn Lab Immunol*. 2005;12(3):380-386. doi:10.1128/CDLI.12.3.380-386.2005
14. Woolven BP, Frenken LGJ, Van Der Logt P, Nicholls PJ. The structure of the llama heavy chain constant genes reveals a mechanism for heavy-chain antibody formation. *Immunogenetics*. 1999;50(1-2):98-101. doi:10.1007/S002510050694

15. Dudek J, Benedix J, Cappel S, et al. Functions and pathologies of BiP and its interaction partners. *Cell Mol Life Sci*. 2009;66(9):1556-1569. doi:10.1007/S00018-009-8745-Y
16. Achour I, Cavelier P, Tichit M, Bouchier C, Lafaye P, Rougeon F. Tetrameric and Homodimeric Camelid IgGs Originate from the Same IgH Locus. *The Journal of Immunology*. 2008;181(3):2001-2009. doi:10.4049/JIMMUNOL.181.3.2001
17. Muyldermans S. Nanobodies: Natural single-domain antibodies. *Annu Rev Biochem*. 2013; 82:775-797. doi:10.1146/annurev-biochem-063011-092449
18. Nguyen VK, Muyldermans S, Hamers R. The specific variable domain of camel heavy-chain antibodies is encoded in the germline. *J Mol Biol*. 1998;275(3):413-418. doi:10.1006/JMBI.1997.1477
19. Nguyen V, Su C, Muyldermans S, Van Der Loo W. Heavy-chain antibodies in Camelidae; a case of evolutionary innovation. *Immunogenetics*. 2002;54(1):39-47. doi:10.1007/S00251-002-0433-0
20. Greenberg AS, Avila D, Hughes M, Hughes A, McKinney EC, Flajnik MF. A new antigen receptor gene family that undergoes rearrangement and extensive somatic diversification in sharks. *Nature.* 1995; 374:168-173. doi:10.1038/374168a0
21. Roux KH, Greenberg AS, Greene L, et al. Structural analysis of the nurse shark (new) antigen receptor (NAR): Molecular convergence of NAR and unusual mammalian immunoglobulins. *PNAS*. 1998;95(20):11804-11809. doi:10.1073/pnas.95.20.11804
22. Dooley H, Flajnik MF. Shark immunity bites back: affinity maturation and memory response in the nurse shark, Ginglymostoma cirratum. *Eur J Immunol*. 2005;35(3):936-945. doi:10.1002/EJI.200425760
23. Conrath KE, Wernery U, Muyldermans S, Nguyen VK. Emergence and evolution of functional heavy-chain antibodies in Camelidae. *Dev Comp Immunol*. 2003;27(2):87-103. doi:10.1016/S0145-305X(02)00071-X
24. Flajnik MF, Deschacht N, Muyldermans S. A Case Of Convergence: Why Did a Simple Alternative to Canonical Antibodies Arise in Sharks and Camels? *PLoS Biol*. 2011;9(8): e1001120. doi:10.1371/JOURNAL.PBIO.1001120
25. Flajnik MF, Kasahara M. Origin and evolution of the adaptive immune system: genetic events and selective pressures. *Nature Reviews Genetics.* 2010; 11:47-59. doi:10.1038/nrg2703
26. Cogné M, Preud'homme JL, Guglielmi P. Immunoglobulin gene alterations in human heavy chain diseases. *Res Immunol*. 1989;140(5-6):487-502. doi:10.1016/0923-2494(89)90115-6
27. Padlan EA. Anatomy of the antibody molecule. *Mol Immunol*. 1994;31(3):169-217. doi:10.1016/0161-5890(94)90001-9
28. Harmsen MM, Ruuls RC, Nijman IJ, Niewold TA, Frenken LGJ, De Geus B. Llama heavy-chain V regions consist of at least four distinct subfamilies revealing novel sequence features. *Mol Immunol*. 2000;37(10):579-590. doi:10.1016/S0161-5890(00)00081-X

29. Muyldermans S, Atarhouch T, Saldanha J, Barbosa JARG, Hamers R. Sequence and structure of VH domain from naturally occurring camel heavy chain immunoglobulins lacking light chains. *Protein Engineering, Design and Selection*. 1994;7(9):1129-1135. doi:10.1093/PROTEIN/7.9.1129
30. Govaert J, Pellis M, Deschacht N, et al. Dual beneficial effect of interloop disulfide bond for single domain antibody fragments. *Journal of Biological Chemistry*. 2012;287(3):1970-1979. doi:10.1074/jbc.M111.242818
31. Harmsen MM, Ruuls RC, Nijman IJ, Niewold TA, Frenken LGJ, De Geus B. Llama heavy-chain V regions consist of at least four distinct subfamilies revealing novel sequence features. *Mol Immunol*. 2000;37(10):579-590. doi:10.1016/S0161-5890(00)00081-X
32. Klarenbeek A, El Mazouari K, Desmyter A, et al. Camelid Ig V genes reveal significant human homology not seen in therapeutic target genes, providing for a powerful therapeutic antibody platform. *MAbs*. 2015;7(4):693-706. doi:10.1080/19420862.2015.1046648
33. Wolfson W. Ablynx Makes Nanobodies from Llama Bodies. *Chem Biol*. 2006;13(12):1243-1244. doi: 10.1016/j.chembiol.2006.12.003
34. Muyldermans S. Applications of Nanobodies. *Annu Rev Anim Biosci*. 2021; 9:401-421. doi:10.1146/ANNUREV-ANIMAL-021419-083831
35. Desmyter A, Transue TR, Ghahroudi MA, et al. Crystal structure of a camel single-domain VH antibody fragment in complex with lysozyme. *Nature Structural Biology 1996 3:9*. 1996;3(9):803-811. doi:10.1038/nsb0996-803
36. Yang EY, Shah K. Nanobodies: Next Generation of Cancer Diagnostics and Therapeutics. *Front Oncol*. 2020; 10:539832. doi:10.3389/FONC.2020.01182
37. Ruiz-López E, Schuhmacher AJ. Transportation of Single-Domain Antibodies through the Blood–Brain Barrier. *Biomolecules*. 2021;11(8):1131. doi:10.3390/BIOM11081131
38. Dumoulin M, Conrath K, Meirhaeghe A Van, et al. Single-domain antibody fragments with high conformational stability. *Protein Science*. 2002;11(3):500-515. doi:10.1110/PS.34602
39. Kunz P, Zinner K, Mücke N, Bartoschik T, Muyldermans S, Hoheisel JD. The structural basis of nanobody unfolding reversibility and thermoresistance. *Scientific Reports.* 2018;8(1):1-10. doi:10.1038/s41598-018-26338-z
40. Kunz P, Flock T, Soler N, et al. Exploiting sequence and stability information for directing nanobody stability engineering. *Biochimica et Biophysica Acta (BBA) - General Subjects*. 2017;1861(9):2196-2205. doi: 10.1016/J.BBAGEN.2017.06.014
41. Hussack G, Hirama T, Ding W, Mackenzie R, Tanha J. Engineered Single-Domain Antibodies with High Protease Resistance and Thermal Stability. *PLoS One*. 2011;6(11):28218. doi: 10.1371/journal.pone.0028218
42. de Marco A. Recombinant expression of nanobodies and nanobody-derived immunoreagents. *Protein Expr Purif*. 2020; 172:105645. doi:10.1016/J.PEP.2020.105645
43. Billen B, Vincke C, Hansen R, et al. Cytoplasmic versus periplasmic expression of site-specifically and bioorthogonally functionalized nanobodies using expressed

protein ligation. *Protein Expr Purif*. 2017; 133:25-34. doi: 10.1016/J.PEP.2017.02.009

44. Wang Y, Li X, Chen X, Nielsen J, Petranovic D, Siewers V. Expression of antibody fragments in Saccharomyces cerevisiae strains evolved for enhanced protein secretion. *Microb Cell Fact*. 2021;20(1):1-17. doi:10.1186/S12934-021-01624-0/FIGURES/6
45. Chen Q, Zhou Y, Yu J, et al. An efficient constitutive expression system for Anti-CEACAM5 nanobody production in the yeast Pichia pastoris. *Protein Expr Purif*. 2019; 155:43-47. doi: 10.1016/J.PEP.2018.11.001
46. Zhao Y, Wang Y, Su W, Li S. Construction of Synthetic Nanobody Library in Mammalian Cells by dsDNA-Based Strategies. *ChemBioChem*. 2021;22(20):2957-2965. doi:10.1002/CBIC.202100286
47. Vanmarsenille C, Elseviers J, Yvanoff C, et al. In planta expression of nanobody-based designer chicken antibodies targeting Campylobacter. *PLoS One*. 2018;13(9): e0204222. doi: 10.1371/JOURNAL.PONE.0204222
48. Klarenbeek A, El Mazouari K, Desmyter A, et al. Camelid Ig V genes reveal significant human homology not seen in therapeutic target genes, providing for a powerful therapeutic antibody platform. *MAbs*. 2015;7(4):693-706. doi:10.1080/19420862.2015.1046648
49. Vincke C, Loris R, Saerens D, Martinez-Rodriguez S, Muyldermans S, Conrath K. General Strategy to Humanize a Camelid Single-domain Antibody and Identification of a Universal Humanized Nanobody Scaffold. *Journal of Biological Chemistry*. 2009;284(5):3273-3284. doi:10.1074/JBC.M806889200
50. Ackaert C, Smiejkowska N, Xavier C, et al. Immunogenicity Risk Profile of Nanobodies. *Front Immunol*. 2021;12. doi:10.3389/fimmu.2021.632687
51. Arbabi-Ghahroudi M. Camelid single-domain antibodies: Historical perspective and future outlook. *Front Immunol*. 2017;8(NOV):315863. doi:10.3389/FIMMU.2017.01589
52. Wang J, Kang G, Yuan H, Cao X, Huang H, de Marco A. Research Progress and Applications of Multivalent, Multispecific and Modified Nanobodies for Disease Treatment. *Front Immunol*. 2022;12. doi:10.3389/fimmu.2021.838082
53. Steeland S, Vandenbroucke RE, Libert C. Nanobodies as therapeutics: big opportunities for small antibodies. *Drug Discov Today*. 2016;21(7):1076-1113. doi:10.1016/J.DRUDIS.2016.04.003
54. Muyldermans S. A guide to: generation and design of nanobodies. *FEBS Journal*. 2021;288(7):2084-2102. doi:10.1111/febs.15515
55. Eden T, Menzel S, Wesolowski J, et al. A cDNA immunization strategy to generate nanobodies against membrane proteins in native conformation. *Front Immunol*. 2018;8(JAN):316032. doi:10.3389/FIMMU.2017.01989
56. Peyrassol X, Laeremans T, Gouwy M, et al. Development by Genetic Immunization of Monovalent Antibodies (Nanobodies) Behaving as Antagonists of the Human ChemR23 Receptor. *The Journal of Immunology*. 2016;196(6):2893-2901. doi:10.4049/JIMMUNOL.1500888

57. Janssens R, Dekker S, Hendriks RW, et al. Generation of heavy-chain-only antibodies in mice. *Proc Natl Acad Sci U S A*. 2006;103(41):15130-15135. doi:10.1073/PNAS.0601108103
58. Olichon A, De Marco A. Preparation of a naïve library of camelid single domain antibodies. *Methods in Molecular Biology*. 2012; 911:65-78. doi:10.1007/978-1-61779-968-6_5
59. Liu W, Song H, Chen Q, et al. Recent advances in the selection and identification of antigen-specific nanobodies. *Mol Immunol*. 2018; 96:37-47. doi: 10.1016/J.MOLIMM.2018.02.012
60. Valdés-Tresanco MS, Molina-Zapata A, Pose AG, Moreno E. Structural Insights into the Design of Synthetic Nanobody Libraries. *Molecules*. 2022;27(7):2198. doi:10.3390/MOLECULES27072198
61. Saerens D, Pellis M, Loris R, et al. Identification of a Universal VHH Framework to Graft Non-canonical Antigen-binding Loops of Camel Single-domain Antibodies. *J Mol Biol*. 2005;352(3):597-607. doi:10.1016/J.JMB.2005.07.038
62. Moutel S, Bery N, Bernard V, et al. NaLi-H1: A universal synthetic library of humanized nanobodies providing highly functional antibodies and intrabodies. *Elife*. 2016;5(JULY). doi:10.7554/ELIFE.16228
63. Smith GP. Filamentous Fusion Phage: Novel Expression Vectors That Display Cloned Antigens on the Virion Surface. *Science*. 1985;228(4705):1315-1317. doi:10.1126/SCIENCE.4001944
64. McCafferty J, Griffiths AD, Winter G, Chiswell DJ. Phage antibodies: filamentous phage displaying antibody variable domains. *Nature*. 1990;348(6301):552-554. doi:10.1038/348552a0
65. Smith GP. Phage Display: Simple Evolution in a Petri Dish (Nobel Lecture). *Angewandte Chemie International Edition*. 2019;58(41):14428-14437. doi:10.1002/ANIE.201908308
66. Pardon E, Laeremans T, Triest S, et al. A general protocol for the generation of Nanobodies for structural biology. *Nat Protoc*. 2014;9(3):674-693. doi:10.1038/nprot.2014.039
67. Ledsgaard L, Ljungars A, Rimbault C, et al. Advances in antibody phage display technology. *Drug Discov Today*. 2022;27(8):2151-2169. doi:10.1016/J.DRUDIS.2022.05.002
68. McMahon C, Baier AS, Pascolutti R, et al. Yeast surface display platform for rapid discovery of conformationally selective nanobodies. *Nature Structural & Molecular Biology*. 2018;25(3):289-296. doi:10.1038/s41594-018-0028-6
69. Rothbauer U. Speed up to find the right ones: rapid discovery of functional nanobodies. *Nature Structural & Molecular Biology*. 2018;25(3):199-201. doi:10.1038/s41594-018-0038-4
70. Mei M, Li J, Wang S, et al. Prompting Fab Yeast Surface Display Efficiency by ER Retention and Molecular Chaperon Co-expression. *Front Bioeng Biotechnol*. 2019; 7:492629. doi:10.3389/FBIOE.2019.00362

71. Han L, Zhao Y, Cui S, Liang B. Redesigning of Microbial Cell Surface and Its Application to Whole-Cell Biocatalysis and Biosensors. *Applied Biochemistry and Biotechnology*. 2017;185(2):396-418. doi:10.1007/S12010-017-2662-6
72. Lee SY, Choi JH, Xu Z. Microbial cell-surface display. *Trends Biotechnol*. 2003;21(1):45-52. doi:10.1016/S0167-7799(02)00006-9
73. Salema V, Fernández LÁ. Escherichia coli surface display for the selection of nanobodies. *Microb Biotechnol*. 2017;10(6):1468-1484. doi:10.1111/1751-7915.12819
74. Nicolay T, Vanderleyden J, Spaepen S. Autotransporter-based cell surface display in Gram-negative bacteria. *Crit Rev Microbiol*. 2015;41(1):109-123. doi: 10.3109/1040841X.2013.804032
75. Salema V, Marín E, Martínez-Arteaga R, et al. 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*. 2013;8(9). doi: 10.1371/journal.pone.0075126
76. Salema V, Mañas C, Cerdán L, et al. High affinity nanobodies against human epidermal growth factor receptor selected on cells by E. coli display. *MAbs*. 2016;8(7):1286-1301. doi:10.1080/19420862.2016.1216742
77. Salema V, López-Guajardo A, Gutierrez C, Mencía M, Fernández LÁ. Characterization of nanobodies binding human fibrinogen selected by E. coli display. *J Biotechnol*. 2016; 234:58-65. doi: 10.1016/J.JBIOTEC.2016.07.025
78. Jaroszewicz W, Morcinek-Orłowska J, Pierzynowska K, Gaffke L, Węgrzyn G. Phage display and other peptide display technologies. *FEMS Microbiol Rev*. 2022;46(2). doi:10.1093/femsre/fuab052
79. Li R, Kang G, Hu M, Huang H. Ribosome Display: A Potent Display Technology used for Selecting and Evolving Specific Binders with Desired Properties. *Mol Biotechnol*. 2019;61(1):60-71. doi:10.1007/S12033-018-0133-0
80. Odegrip R, Coomber D, Eldridge B, et al. CIS display: In vitro selection of peptides from libraries of protein-DNA complexes. *Proc Natl Acad Sci U S A*. 2004;101(9):2806-2810. doi:10.1073/PNAS.0400219101
81. Patel S, Mathonet P, Jaulent AM, Ullman CG. Selection of a high-affinity WW domain against the extracellular region of VEGF receptor isoform-2 from a combinatorial library using CIS display. *Protein Engineering, Design and Selection*. 2013;26(4):307-315. doi:10.1093/PROTEIN/GZT003
82. Uchański T, Pardon E, Steyaert J. Nanobodies to study protein conformational states. *Curr Opin Struct Biol*. 2020; 60:117-123. doi: 10.1016/J.SBI.2020.01.003
83. Rasmussen SGF, Choi HJ, Fung JJ, et al. Structure of a nanobody-stabilized active state of the β2 adrenoceptor. *Nature*. 2011;469(7329):175-180. doi:10.1038/nature09648
84. Frecot DI, Froehlich T, Rothbauer U. 30 years of nanobodies – an ongoing success story of small binders in biological research. *J Cell Sci*. 2023;136(21). doi:10.1242/JCS.261395

85. Manglik A, Kobilka BK, Steyaert J. Nanobodies to Study G Protein–Coupled Receptor Structure and Function. *Annual Review ofPharmacology and Toxicology*. 2017; 57:19-37. doi:10.1146/ANNUREV-PHARMTOX-010716-104710
86. Kruse AC, Ring AM, Manglik A, et al. Activation and allosteric modulation of a muscarinic acetylcholine receptor. *Nature*. 2013;504(7478):101-106. doi:10.1038/nature12735
87. Burg JS, Ingram JR, Venkatakrishnan AJ, et al. Structural basis for chemokine recognition and activation of a viral G protein-coupled receptor. *Science.* 2015;347(6226):1113-1117. doi:10.1126/SCIENCE.AAA5026
88. Huang W, Manglik A, Venkatakrishnan AJ, et al. Structural insights into µ-opioid receptor activation. *Nature*. 2015;524(7565):315-321. doi:10.1038/nature14886
89. Murata K, Wolf M. Cryo-electron microscopy for structural analysis of dynamic biological macromolecules. *Biochimica et Biophysica Acta (BBA) - General Subjects*. 2018;1862(2):324-334. doi:10.1016/J.BBAGEN.2017.07.020
90. Wentinck K, Gogou C, Meijer DH. Putting on molecular weight: Enabling cryo-EM structure determination of sub-100-kDa proteins. *Curr Res Struct Biol*. 2022; 4:332-337. doi: 10.1016/J.CRSTBI.2022.09.005
91. Robertson MJ, Papasergi-Scott MM, He F, et al. Structure determination of inactive-state GPCRs with a universal nanobody. *Nature Structural & Molecular Biology*. 2022;29(12):1188-1195. doi:10.1038/s41594-022-00859-8
92. Wu X, Rapoport TA. Cryo-EM structure determination of small proteins by nanobody-binding scaffolds (Legobodies). *Proc Natl Acad Sci U S A*. 2021;118(41): e2115001118. doi:10.1073/PNAS.2115001118
93. Keller BM, Maier J, Secker KA, et al. Chromobodies to Quantify Changes of Endogenous Protein Concentration in Living Cells. *Molecular & Cellular Proteomics*. 2018;17(12):2518-2533. doi:10.1074/MCP.TIR118.000914
94. Traenkle B, Segan S, Fagbadebo FO, Kaiser PD, Rothbauer U. A novel epitope tagging system to visualize and monitor antigens in live cells with chromobodies. *Scientific Reports*. 2020;10(1):1-13. doi:10.1038/s41598-020-71091-x
95. Frankel AD, Pabo CO. Cellular uptake of the tat protein from human immunodeficiency virus. *Cell*. 1988;55(6):1189-1193. doi:10.1016/0092-8674(88)90263-2
96. Rothbauer U, Zolghadr K, Tillib S, et al. Targeting and tracing antigens in live cells with fluorescent nanobodies. *Nat Methods*. 2006;3(11):887-889. doi:10.1038/nmeth953
97. Panza P, Maier J, Schmees C, Rothbauer U, Söllner C. Live imaging of endogenous protein dynamics in zebrafish using chromobodies. *Development (Cambridge)*. 2015;142(10):1879-1884. doi:10.1242/DEV.118943
98. Wegner W, Ilgen P, Gregor C, et al. In vivo mouse and live cell STED microscopy of neuronal actin plasticity using far-red emitting fluorescent proteins. *Scientific Reports*. 2017;7(1):1-10. doi:10.1038/s41598-017-11827-4
99. Dietrich J, Sommersdorf C, Gohlke S, et al. Okadaic acid activates Wnt/β-catenin-signaling in human HepaRG cells. *Arch Toxicol*. 2019;93(7):1927-1939. doi:10.1007/s00204-019-02489-4

100. Traenkle B, Emele F, Anton R, et al. Monitoring Interactions and Dynamics of Endogenous Beta-catenin With Intracellular Nanobodies in Living Cells. *Molecular & Cellular Proteomics*. 2015;14(3):707-723. doi:10.1074/MCP.M114.044016
101. Caussinus E, Kanca O, Affolter M. Fluorescent fusion protein knockout mediated by anti-GFP nanobody. *Nature Structural & Molecular Biology*. 2011;19(1):117-121. doi:10.1038/nsmb.2180
102. Shin YJ, Park SK, Jung YJ, et al. Nanobody-targeted E3-ubiquitin ligase complex degrades nuclear proteins. *Scientific Reports*. 2015;5(1):1-11. doi:10.1038/srep14269
103. Baudisch B, Pfort I, Sorge E, Conrad U. Nanobody-directed specific degradation of proteins by the 26s-proteasome in plants. *Front Plant Sci*. 2018; 9:312434. doi:10.3389/FPLS.2018.00130
104. Yamaguchi N, Colak-Champollion T, Knaut H. ZGrad is a nanobody-based degron system that inactivates proteins in zebrafish. *Elife*. 2019;8. doi:10.7554/ELIFE.43125
105. Tang JCY, Szikra T, Kozorovitskiy Y, et al. A Nanobody-Based System Using Fluorescent Proteins as Scaffolds for Cell-Specific Gene Manipulation. *Cell*. 2013;154(4):928-939. doi: 10.1016/J.CELL.2013.07.021
106. Bruce VJ, McNaughton BR. Evaluation of Nanobody Conjugates and Protein Fusions as Bioanalytical Reagents. *Anal Chem*. 2017;89(7):3819-3823. doi: 10.1021/acs.analchem.7b00470
107. Ren W, Li Z, Xu Y, et al. One-Step Ultrasensitive Bioluminescent Enzyme Immunoassay Based on Nanobody/Nanoluciferase Fusion for Detection of Aflatoxin B 1 in Cereal. *J Agric Food Chem*. 2019;67(18):5221-5229. doi: 10.1021/ACS.JAFC.9B00688
108. Salvador JP, Vilaplana L, Marco MP. Nanobody: outstanding features for diagnostic and therapeutic applications. *Anal Bioanal Chem*. 2019;411(9):1703-1713. doi:10.1007/S00216-019-01633-4
109. de Vos J, Devoogdt N, Lahoutte T, Muyldermans S, Muyldermans S. Camelid single-domain antibody-fragment engineering for (pre)clinical in vivo molecular imaging applications: adjusting the bullet to its target. *Expert Opin Biol Ther*. 2013;13(8):1149-1160. doi:10.1517/14712598.2013.800478
110. Debie P, Lafont C, Defrise M, et al. Size and affinity kinetics of nanobodies influence targeting and penetration of solid tumours. *Journal of Controlled Release*. 2020; 317:34-42. doi: 10.1016/J.JCONREL.2019.11.014
111. Xavier C, Vaneycken I, D'Huyvetter M, et al. Synthesis, Preclinical Validation, Dosimetry, and Toxicity of 68Ga-NOTA-Anti-HER2 Nanobodies for iPET Imaging of HER2 Receptor Expression in Cancer. *Journal of Nuclear Medicine*. 2013;54(5):776-784. doi:10.2967/JNUMED.112.111021
112. Keyaerts M, Xavier C, Heemskerk J, et al. Phase I Study of 68 Ga-HER2-Nanobody for PET/CT Assessment of HER2 Expression in Breast Carcinoma. *J Nucl Med*. 2016; 57:27-33. doi:10.2967/jnumed.115.162024

113. Keyaerts M, Xavier C, Everaert H, et al. Phase II trial of HER2-PET/CT using 68Ga-anti-HER2 VHH1 for characterization of HER2 presence in brain metastases of breast cancer patients. *Annals of Oncology*. 2019;30: iii25-iii26. doi:10.1093/annonc/mdz095.081
114. Broisat A, Hernot S, Toczek J, et al. Nanobodies targeting mouse/human VCAM1 for the nuclear imaging of atherosclerotic lesions. *Circ Res*. 2012;110(7):927-937. doi:10.1161/CIRCRESAHA.112.265140
115. Ruiz-López E, Schuhmacher AJ. Transportation of Single-Domain Antibodies through the Blood–Brain Barrier. *Biomolecules*. 2021;11(8). doi:10.3390/BIOM11081131
116. Li T, Vandesquille M, Koukouli F, et al. Camelid single-domain antibodies: A versatile tool for in vivo imaging of extracellular and intracellular brain targets. *Journal of Controlled Release*. 2016; 243:1-10. doi:10.1016/J.JCONREL.2016.09.019
117. Pothin E, Lesuisse D, Lafaye P. Brain Delivery of Single-Domain Antibodies: A Focus on VHH and VNAR. *Pharmaceutics 2020*. 2020;12(10):937. doi:10.3390/PHARMACEUTICS12100937
118. Peyvandi F, Scully M, Kremer Hovinga JA, et al. Caplacizumab for Acquired Thrombotic Thrombocytopenic Purpura. *New England Journal of Medicine*. 2016;374(6):511-522. doi:10.1056/nejmoa1505533
119. Scully M, Cataland SR, Peyvandi F, et al. Caplacizumab Treatment for Acquired Thrombotic Thrombocytopenic Purpura. *New England Journal of Medicine*. 2019;380(4):335-346. doi:10.1056/nejmoa1806311
120. Morrison C. Nanobody approval gives domain antibodies. *Nature reviews drug discovery.* 2019; 18:485-487.
121. Keam SJ. Ozoralizumab: First Approval. *Drugs*. 2023;83(1):87-92. doi:10.1007/S40265-022-01821-0
122. Mullard A. FDA approves second BCMA-targeted CAR-T cell therapy. *Nat Rev Drug Discov*. 2022;21(4):249. doi: 10.1038/D41573-022-00048-8
123. Safarzadeh Kozani P, Naseri A, Mirarefin SMJ, et al. Nanobody-based CAR-T cells for cancer immunotherapy. *Biomarker Research*. 2022;10(1):1-18. doi:10.1186/S40364-022-00371-7
124. Benmebarek MR, Karches CH, Cadilha BL, Lesch S, Endres S, Kobold S. Killing Mechanisms of Chimeric Antigen Receptor (CAR) T Cells. *International Journal of Molecular Sciences*. 2019;20(6):1283. doi:10.3390/IJMS20061283
125. Albert S, Arndt C, Feldmann A, et al. A novel nanobody-based target module for retargeting of T lymphocytes to EGFR-expressing cancer cells via the modular UniCAR platform. *Oncoimmunology*. 2017;6(4). doi:10.1080/2162402X.2017.1287246
126. Jamnani FR, Rahbarizadeh F, Shokrgozar MA, et al. T cells expressing VHH-directed oligoclonal chimeric HER2 antigen receptors: Towards tumor-directed oligoclonal T cell therapy. *Biochimica et Biophysica Acta (BBA) - General Subjects*. 2014;1840(1):378-386. doi:10.1016/J.BBAGEN.2013.09.029

127. Vosjan MJWD, Vercammen J, Kolkman JA, Stigter-Van Walsum M, Revets H, Van Dongen GAMS. Nanobodies targeting the hepatocyte growth factor: Potential new drugs for molecular cancer therapy. *Mol Cancer Ther*. 2012;11(4):1017-1025. doi:10.1158/1535-7163.MCT-11-0891/83956
128. Roovers RC, Vosjan MJWD, Laeremans T, et al. A biparatopic anti-EGFR nanobody efficiently inhibits solid tumour growth. *Int J Cancer*. 2011;129(8):2013-2024. doi:10.1002/IJC.26145
129. Omidfar K, Amjad Zanjani FS, Hagh AG, Azizi MD, Rasouli SJ, Kashanian S. Efficient growth inhibition of EGFR over-expressing tumor cells by an anti-EGFR nanobody. *Mol Biol Rep*. 2013;40(12):6737-6745. doi:10.1007/S11033-013-2790-1
130. Behdani M, Zeinali S, Karimipour M, et al. Expression, purification, and characterization of a diabody against the most important angiogenesis cell receptor: Vascular endothelial growth factor receptor 2. *Adv Biomed Res*. 2012;1(1):34. doi:10.4103/2277-9175.100126
131. Maussang D, Mujić-Delić A, Descamps FJ, et al. Llama-derived Single Variable Domains (Nanobodies) Directed against Chemokine Receptor CXCR7 Reduce Head and Neck Cancer Cell Growth in Vivo. *Journal of Biological Chemistry*. 2013;288(41):29562-29572. doi:10.1074/JBC.M113.498436
132. Zhang F, Wei H, Wang X, et al. Structural basis of a novel PD-L1 nanobody for immune checkpoint blockade. *Cell Discovery.* 2017;3(1):1-12. doi:10.1038/celldisc.2017.4
133. Behdani M, Zeinali S, Karimipour M, et al. Development of VEGFR2-specific Nanobody Pseudomonas exotoxin A conjugated to provide efficient inhibition of tumor cell growth. *N Biotechnol*. 2013;30(2):205-209. doi:10.1016/J.NBT.2012.09.002
134. D'Huyvetter M, Xavier C, Caveliers V, Lahoutte T, Muyldermans S, Devoogdt N. Radiolabeled nanobodies as theranostic tools in targeted radionuclide therapy of cancer. *Expert Opin Drug Deliv*. 2014;11(12):1939-1954. doi:10.1517/17425247.2014.941803
135. D'Huyvetter M, De Vos J, Caveliers V, et al. Phase I Trial of 131I-GMIB-Anti-HER2-VHH1, a New Promising Candidate for HER2-Targeted Radionuclide Therapy in Breast Cancer Patients. *Journal of Nuclear Medicine*. 2021;62(8):1097-1105. doi:10.2967/JNUMED.120.255679
136. Heukers R, van Bergen en Henegouwen PMP, Oliveira S. Nanobody–photosensitizer conjugates for targeted photodynamic therapy. *Nanomedicine*. 2014;10(7):1441-1451. doi: 10.1016/J.NANO.2013.12.007
137. Wrapp D, De Vlieger D, Corbett KS, et al. Structural Basis for Potent Neutralization of Betacoronaviruses by Single-Domain Camelid Antibodies. *Cell*. 2020;181(5):1004-1015.e15. doi: 10.1016/j.cell.2020.04.031
138. Custódio TF, Das H, Sheward DJ, et al. Selection, biophysical and structural analysis of synthetic nanobodies that effectively neutralize SARS-CoV-2. *Nature Communications*. 2020;11(1):1-11. doi:10.1038/s41467-020-19204-y

139. Hanke L, Vidakovics Perez L, Sheward DJ, et al. An alpaca nanobody neutralizes SARS-CoV-2 by blocking receptor interaction. *Nature Communications.* 2020;11(1):1-9. doi:10.1038/s41467-020-18174-5
140. Ma H, Zeng W, Meng X, et al. Potent Neutralization of SARS-CoV-2 by Hetero-Bivalent Alpaca Nanobodies Targeting the Spike Receptor-Binding Domain. *J Virol*. 2021;95(10). doi:10.1128/JVI.02438-20/ASSET/CD6E41F9-C0FF-47C3-9E81-670449F79817
141. Huo J, Le Bas A, Ruza RR, et al. Neutralizing nanobodies bind SARS-CoV-2 spike RBD and block interaction with ACE2. *Nature Structural & Molecular Biology*. 2020;27(9):846-854. doi:10.1038/s41594-020-0469-6
142. Maeda R, Fujita J, Konishi Y, et al. A panel of nanobodies recognizing conserved hidden clefts of all SARS-CoV-2 spike variants including Omicron. *Communications Biology*. 2022;5(1):1-16. doi:10.1038/s42003-022-03630-3
143. Xu J, Xu K, Jung S, et al. Nanobodies from camelid mice and llamas neutralize SARS-CoV-2 variants. *Nature*. 2021;595(7866):278-282. doi:10.1038/s41586-021-03676-z
144. Cardoso FM, Ibañez LI, Van den Hoecke S, et al. Single-Domain Antibodies Targeting Neuraminidase Protect against an H5N1 Influenza Virus Challenge. *J Virol*. 2014;88(15):8278-8296. doi:10.1128/JVI.03178-13
145. Cunningham S, Piedra PA, Martinon-Torres F, et al. Nebulised ALX-0171 for respiratory syncytial virus lower respiratory tract infection in hospitalised children: a double-blind, randomised, placebo-controlled, phase 2b trial. *Lancet Respir Med*. 2021;9(1):21-32. doi:10.1016/S2213-2600(20)30320-9
146. King MT, Huh I, Shenai A, Brooks TM, Brooks CL. Structural basis of VHH-mediated neutralization of the food-borne pathogen Listeria monocytogenes. *Journal of Biological Chemistry*. 2018;293(35):13626-13635. doi:10.1074/jbc.RA118.003888
147. Lu Z, Liu Z, Li X, et al. Nanobody-Based Bispecific Neutralizer for Shiga Toxin-Producing E. coli. *ACS Infect Dis*. 2022;8(2):321-329. doi:10.1021/ACSINFECDIS.1C00456
148. Hempelmann A, Hartleb L, Van Straaten M, et al. Nanobody-mediated macromolecular crowding induces membrane fission and remodeling in the African trypanosome. *Cell Rep*. 2021;37. doi: 10.1016/j.celrep.2021.109923
149. Dietrich MH, Gabriela M, Reaksudsan K, et al. Nanobodies against Pfs230 block Plasmodium falciparum transmission. *Biochemical Journal*. 2022;479(24):2529-2546. doi:10.1042/BCJ20220554
150. Richard G, Meyers AJ, McLean MD, Arbabi-Ghahroudi M, MacKenzie R, Hall JC. In Vivo Neutralization of α-Cobratoxin with High-Affinity Llama Single-Domain Antibodies (VHHs) and a VHH-Fc Antibody. *PLoS One*. 2013;8(7): e69495. doi: 10.1371/JOURNAL.PONE.0069495
151. Bailon Calderon H, Yaniro Coronel VO, Cáceres Rey OA, et al. Development of Nanobodies Against Hemorrhagic and Myotoxic Components of Bothrops atrox Snake Venom. *Front Immunol*. 2020; 11:655-655. doi:10.3389/FIMMU.2020.00655

152. Hmila I, Saerens D, Abderrazek R Ben, et al. A bispecific nanobody to provide full protection against lethal scorpion envenoming. *The FASEB Journal*. 2010;24(9):3479-3489. doi:10.1096/FJ.09-148213

Aim of the project

Aim of the project

Since their discovery, single-domain antibodies, also known as nanobodies, have attracted considerable interest in various areas of biological and medical research, and now represent a relevant class of biomolecules[1].
The isolation of nanobodies with specific properties requires the construction of either immune or synthetic libraries followed by an accurate selection process. Two critical considerations in this procedure are the size of the library and the selection system employed. The larger the library and the more robust the selection method, the easier it will be to isolate target-specific nanobodies[2].
My doctoral project is related to this aspect and aims to develop library construction vectors that allow the most efficient cloning procedures for the generation of highly complex libraries. In addition, the vectors developed must provide a more efficient and faster selection mechanism compared to traditional vectors. This is crucial to reduce the time and cost of the entire process required to isolate high-affinity nanobodies for the target of interest.

In particular, the engineering has been applied to three different types of vectors using three different approaches:

- For the preparation of nanobodies intended for therapeutic or diagnostic use in humans, humanisation is a crucial step to mitigate immune rejection issues in patients. Starting from a functional humanized consensus framework scaffold (hCFW)[3,4], I decided to build a humanised synthetic nanobodies library. The goal is to yield nanobodies that are potentially suitable for human administration. To enhance the likelihood of isolating high-quality ligands, I opted for the efficient and versatile Gateway technology[5] as the cloning method and phage display as the selection method due to its robustness and widespread use. A novel phage vector, designed to speed up the phage display selection process, was employed as the cloning vector (Chapter 1).

- While preparing phage or phagemid vector libraries, a drawback of the Gateway cloning technology is that it lacks the ability to directly select recombinants in bacterial strains employed in the phage display selection system. Such strains are resistant to the CcdB toxin encoded by the destination cassette characteristic in this cloning system[5]. To solve this problem, I decided to modify the classical Gateway cassette (namely DEST) within a phagemid vector. This modification facilitates the overexpression of the toxin, inducing deleterious levels of toxicity even in typically resistant bacterial cells. As a proof of concept, I evaluated the performance of this novel phagemid vector in constructing a nanobody library from an immunised llama and subjected it to selection through phage display (Chapter 2).

- The bacterial display screening technology demands the exposure of potential ligands on the bacterial cell surface in a properly folded state for screening purposes.

In pursuit of this goal, I assess the display functionality of three distinct *E. coli* autotransporters using SpyTag-SpyCatcher technology[6]. Following the identification of the most efficient autotransporter, I proceeded to modify the corresponding expression vector, rendering it compatible with Gateway cloning. In this modification, I introduced the same cassette facilitating the overexpression of the CcdB toxin as elucidated in Chapter 2. To determine the adaptability of these modifications for the construction of a nanobody library, I transplanted a coding sequence (CDS) encoding a particular nanobody into this vector. The subsequent evaluation focused on its successful exposure to the bacterial surface, constituting the key assessment in Chapter 3.

Bibliography

1. Muyldermans S. Applications of Nanobodies. *Annu Rev Anim Biosci*. 2021; 9: 401-421. doi:10.1146/ANNUREV-ANIMAL-021419-083831
2. Muyldermans S. A guide to: generation and design of nanobodies. *FEBS Journal*. 2021;288(7):2084-2102. doi:10.1111/febs.15515
3. Ferrari D, Garrapa V, Locatelli M, Bolchi A. A Novel Nanobody Scaffold Optimized for Bacterial Expression and Suitable for the Construction of Ribosome Display Libraries. *Mol Biotechnol*. 2020;62(1):43-55. doi:10.1007/S12033-019-00224-Z
4. Garrapa V. *A Novel Scaffold for Nanobody Selection and Expression Suitable for Diagnostic and Therapeutic Applications*. 2019. Accessed January 9, 2024. https://www.repository.unipr.it/handle/1889/3722
5. Life Technologies Carlsbad CA. Gateway ® Technology with Clonase ® II A universal technology to clone DNA sequences for functional analysis and expression in multiple systems. *User guide*.
6. Zakeri B, Fierer JO, Celik E, et al. Peptide tag forming a rapid covalent bond to a protein, through engineering a bacterial adhesin. *PNAS*. 2012;109(12): E690-E697. doi:10.1073/PNAS.1115485109

Chapter 1

Chapter 1: Construction of a humanised synthetic nanobody library in a novel phage-display vector

1 Introduction

Antibodies and their derivatives have emerged as the most important classes of therapeutics in recent decades. Monoclonal antibodies have conquered the world market, but the development of antibody fragments such as scFv and Fab using recombinant DNA technology has also contributed to the expansion of this market.
In 1993, the serendipitous discovery of heavy chain antibodies (hcAbs) in members of the *Camelidae* family[1] opened up new possibilities in the area of therapeutic antibodies. HcAbs can bind their target with a single variable heavy domain called VHH or nanobody (Nb) or single domain antibody (sdAb), and the possibility of expressing it in recombinant form, while retaining the binding capacity of the whole immunoglobulin, led to the development of a new antibody fragment. Several important advantages such as the small size, high solubility, and high specificity make nanobodies an interesting class of biomolecules for different applications. Another important feature is low immunogenicity. They also share a high sequence identity with the human VH III family, making them attractive for therapeutic applications[2]. However, given their nature, they might elicit an immune response in case of administration in humans. The potential antigenicity is not a problem for research and diagnostic applications but could become one for *in vivo* use[3]. So, for therapeutic applications, humanization must be considered. The humanisation process consists of replacing certain amino acid residues of VHHs with those commonly found in the corresponding positions of the human VH domain, while retaining four amino acids characteristic of the FR2 region (mostly Phe42, Glu49, Arg50, Gly52) that are important for their structure and stability[4]. However, this process remains challenging: humanisation could affect anyway the stability of the nanobody and even their target affinity, as the structural residues could be involved in binding to the target or their modification could lead to a different exposure of the CDRs[5].
In recent years, the design of synthetic libraries has emerged as an alternative to the use of animals, allowing the isolation of ligands against toxic or poorly immunogenic targets and offering advantages in terms of cost and time[6]. Therefore, the construction of VHH synthetic libraries with CDRs, scrambled in sequence, inserted in a validated and already humanised scaffold makes it possible to avoid the efforts of the humanisation process and, after the selection of target-specific one, obtain nanobodies ready for use in therapeutic applications[7].
Among the various selection strategies, phage display remains the most common and widely used method for the discovery of new antibodies[8]. M13, which belongs to the class of filamentous bacteriophages (Ff), is the most widely used bacteriophage in phage display. These bacteriophages infect *E. coli* host cells via a specific interaction between the F-pilus of the host cell and the phage coat protein pIII. In the phage display system, the coat protein pIII is used as antibody fusion partner because can accommodate large

proteins without affecting the infectious function of the phage[9]. So, the antibody library (in this case nanobody library) is expressed on the surface of M13 phage fused in-frame to the phage pIII protein.

An important aspect is the choice of phage display formats: phage vector or phagemid vector respectively type 3 and type 3+3 according to Smith's classification[10]. Typically, for the selection of antibody libraries, it is preferable to use the phagemid vector: the engineered pIII protein encoding gene is located in the phagemid vector and requires the use of the phage helper, which provides *in trans* all the structural components for the production of complete phage particles. This format has the advantage of producing a single copy of the surface peptide-pIII protein (monovalent display), which facilitates the selection of stronger related ligands, but also increases the panning selection time by up to 4 days[11,12].

In this chapter, we present the construction of a humanized synthetic nanobody library using a functional humanized consensus framework scaffold (hCFW)[13,14].

To accelerate phage display steps, we employed an engineered phage vector M13-flash[14], which maintains rapid panning steps while acquiring the monovalent characteristics of a phagemid vector (type 33[10]). The ultimate aim is to achieve a library complexity of 10^9 individual clones, so we choose an efficient cloning method, Gateway® technology, developed by Invitrogen, which enables the creation of large libraries. Once a library of the desired complexity has been obtained, it will be possible to isolate nanobodies that are potentially ready for human applications.

2 Materials and Methods

2.1 Synthesis of humanized synthetic nanobody library

Humanized synthetic nanobody library was generated by assembly PCR a humanized synthetic nanobody used as a template, whose sequence is shown in **Table 1**. PAGE-purified oligonucleotides (Microsynth, Balgach, Switzerland), listed in **Table 2**, and Phusion DNA polymerase (Thermo Fisher Scientific) were used; Individual PCR products were purified from agarose gel using a Macherey-Nagel kit (Düren, Germany).

Name	Sequence
Humanized synthetic nanobody sequence	GAGGTACAGCTGCTTGAATCGGGTGGTGGCCTGGTTCAACCAGGCGGCA GTTTACGCCTGTCATGTGCAGCGTCAGGTTTCATTTTTTCGGGTTATGC TATTGGTTGGTTCCGTCAAGCACCAGGGAAAGGCCTGGAGTGGGTGTCG GCCATTACTTGGGGTGGTGGTAGTACCTACTATGCGGATTCCGTGAAAG GACGCTTCACCATTTCTCGGGACAATTCGAAGAACACCCTGTATTTGCA GATGAACTCCCTGCGTGCGGAAGATACGGCTGTCTACTATTGCGCTGCG GATCCTAGTTCCCATGGTCTTCACCCTTATCGTGCTGGTGATTTTGGTT ATTGGGGCCAGGGTACTCTGGTCACCGTAAGC

Table 1 Nucleotide sequence of synthetic humanised nanobody used as template for PCR assembly. Sequences are shown in standard 5'-3' orientation.

Primer	Sequence
attB1_hNblib_FW	GGGGACAAGTTTGTACAAAAAAGCAGGCTCCGAGGTACAGCTGCTTG
attB2_hNblib_RE	GGGGACCACTTTGTACAAGAAAGCTGGGTAGCTTACGGTGACCAGAG
FR1-RE	ACCTGACGCTGCACATGAC
hFR2-RE	AATGGCCGACACCCACTCCAGGCCTTTCCCTGGTGCTTGACG
FR3-RE	CGCAGCGCAATAGTAGACAG
hCDR1-FW	GTCATGTGCAGCGTCAGGTTYCAYTTTTTCGRDTWATRCTATKGSTTGGKWCC GTCAAGCACCAGGGAAAG
hCDR2-FW	GGAGTGGGTGTCGGCCATTAVTWSGRGTGRTGRTAVKACCTACTATGCGGATT CCGTGAAAGGA
hCDR3-FW	CTGTCTACTATTGCGCTGCGSVTMSTVGTNHYNHYVGTNHYNHYYHTWMTVNT BCTVNTRVKTWTRRTTATTGGGGCCAGGGTACTC

Table 2. Oligonucleotides used for assembling PCR steps. Sequences are shown in a standard 5'-3' orientation. Standard one-letter nomenclature is used for specific nucleotide subsets, where: M represents A, or C; Y represents C or T; R represents A, or G; D represents A, G or T; W represents A, or T; K represents G or T; S represents C or G; V represents A, C or G; H represents A, C or T; B represents C, G or T; and N represents any of the four nucleotides.

PCR1: the oligonucleotides attB1_hNblib_FW and FR1-RE were used to amplify the first portion of the library clone containing the *att*B region and FR1. The following PCR conditions were used: initial denaturation 98°C, 1 s; 25 cycles of 10 s at 98°C, 30 s at 65°C, 30 s at 72°C; final extension 72°C 7 min; 4°C ∞.

PCR2: the oligonucleotides hCDR1_FW and hFR2_RE were fused in the presence of dNTPs, PCR Buffer and Phusion DNA Polymerase in a final volume of 50 µl. The conditions were the same as for PCR1 except for 20 cycles and 66°C annealing temperature. After 20 cycles, the purified PCR1 product (100 ng), dNTPs and DNA polymerase were added, and another 20 cycles were performed. After that, the external oligonucleotides attB1_hNblib_FW and hFR2-RE, dNTPs and DNA polymerase were added and 28 cycles of amplification were performed to generate a final product containing the FR1, CDR1 and FR2 regions.
PCR3: the oligonucleotides hCDR2-FW and FR3-RE primers were used to amplify the initial template (10 ng) (conditions as PCR1 except for 28 cycles and 64°C annealing temperature) and to generate an amplicon containing CDR2 and FR3.
PCR4: the oligonucleotides hCDR3-FW and attB2_hNblib_RE were used to amplify the initial template (10 ng), under the same conditions as for PCR3. The amplicon containing CDR3 and FR4 was obtained.
PCR5: 40 ng and 25 ng (equimolar amounts) of the purified PCR2 and PCR3 products respectively, were fused (conditions as PCR3 except for 20 cycles) in the presence of dNTPs, PCR buffer and Phusion DNA Polymerase in a final volume of 50 µl. The fused product was amplified after the addition of the external oligonucleotides attB1_hNblib_FW and FR3-RE, dNTPs and Phusion DNA Polymerase (conditions as PCR1 except 28 cycles and of 70°C annealing temperature). A final product containing the FR1, CDR1, FR2, CDR2 and FR3 regions was obtained.
PCR6: 35 ng and 65 ng (equimolar amounts) of the purified PCR4 and PCR5 products were fused by 25 cycles of PCR amplification (conditions as PCR5 except 59°C annealing temperature and 1 min extension time) in the presence of dNTPs, Phusion buffer and Phusion DNA Polymerase in a final volume of 50 µl. Then the external oligonucleotides attB1_hNblib_FW and attB2_hNblib_RE, dNTPs and DNA polymerase were added and 28 cycles of amplification were performed to obtain the final fusion product corresponding to the complete humanized nanobody library. PCR6 was performed 10 times to have a representative library.

2.2 Cloning procedures

All cloning procedures used for the library construction refer to Gateway® technology user's guide[15]. This cloning method is based on two site-specific recombinations mediated by bacteriophage lambda proteins and allows high cloning efficiency.
The two recombination reactions are the BP reaction, in which the library is transferred into a donor vector (pDONR222, Invitrogen, Waltham, MA, USA) and the LR reaction into a compatible destination vector (**Figure 1**). This cloning method uses both positive and negative selection in the selection of recombinant clones and in vector propagation: different selection markers carried by vectors allow the selection of recombinant plasmid, the ccdB gene present in donor and destination vectors eliminate non-recombinant clones. The ccdB gene is the key element of the system: it is located on F' episome of *E. coli* and encoding the toxic protein CcdB, a DNA gyrase inhibitor. This protein is part of a toxin-antitoxin system. The antitoxin CcdA protect the cells by the toxic effects of CcdB. However, cells lacking the F' episome suffer from CcdB toxicity[16,17].

The M13-flash destination (M13-flash DEST) vector, previously engineered and tested by colleagues[14], was used in the LR reaction to obtain the library suitable for phage display selection.

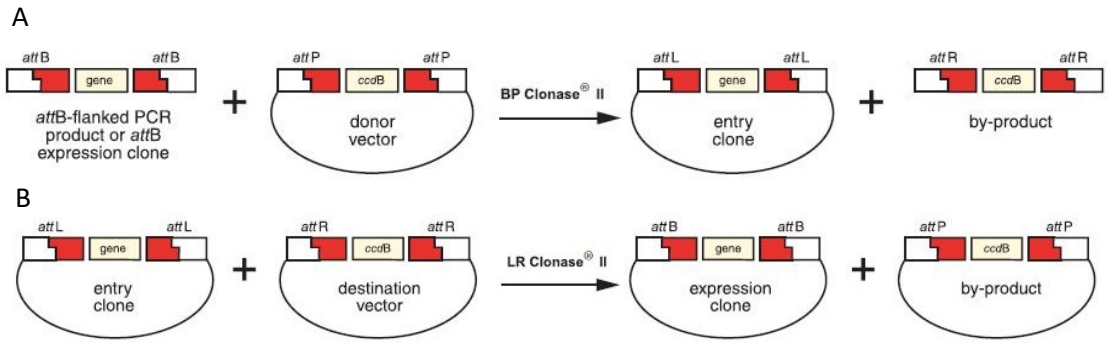


Figure 1[15]**.** Schematic representation of Gateway® technology. A: BP reaction, B: LR reaction

2.3 Phage display procedures

2.3.1 Biopanning

All recombinant proteins used in the phage display were kindly provided by the collaborators.

50 µl of an overnight stationary culture of TG1 *E. coli* strain cells were inoculated into 5 ml of 2xYT medium and, after reaching an A_{600} value of 0.5, stored at 4°C until use. Panning steps were performed in 96-well polystyrene microtitre plates according to the following protocol: a single well was coated overnight at 4°C with 2 µg/ml casein-glutathione dissolved in 50 mM carbonate buffer. Blocking was performed by incubation for 1 hour at room temperature with 2% skimmed milk (Sigma-Aldrich) in phosphate-buffered saline (PBS, pH 7.4), followed by incubation for 1 hour at room temperature with 500 nM GST-*Pf*Trx in PBS. The GST protein, which has a high affinity for glutathione, allows the target to be exposed to avoid any conformational deformation that might occur with direct adhesion. An aliquot of 10^{11} phages in PBS+ 10% skimmed milk was incubated for 2 hours at 700 rpm with shaking. After each incubation step, the wells were washed with PBS containing 0.3% (v/v) Tween20. Target-bound phages were eluted by the addition of 0.2 M glycine-HCl buffer (pH 2.2) and incubation at 700 rpm with shaking for 15 minutes. Then, phages solution was neutralized with 1 M TrisHCl (pH 9) and 50 µl was added to 3 ml of TG1 A_{600} 0.5 cells and infection was performed for 30 min at 37°C. To infected TG1 cells, 7 ml of 2xYT media containing 1 mM IPTG were added and cells were incubated o/n at 30°C. The next day cells were centrifuged (3200 rpm, 15 min), the supernatant containing the phages was precipitated with a solution of 20 % PEG6000 - 2.5 M NaCl and used for the next panning step.

2.3.2 Polyclonal phage-ELISA

The phage ELISA was performed according to the following protocol: 96-well polystyrene microtitre plates were coated overnight at 4°C with 500 nM of the purified

PfTrx and GST-GFP proteins in PBS, provided by colleagues. After blocking for 1 hour at room temperature with 2% skimmed milk (Sigma-Aldrich) in PBS (pH 7.4), a 10^{10} dilution in PBS of eluted phages from each round of panning was added and incubated for 1 hour at room temperature. After that, the plate was incubated with an anti-M13 (pVIII)-HRP (GE Healthcare, Little Chalfont, USA) diluted 1:5000 in PBS for 1 hour at room temperature. The signal was detected by adding 2,2'-azino-di- [3-ethylbenzthiazoline sulfonate (ABTS) substrate (KPL, Gaithersburg, MD, USA) and the absorbance was measured at 415 nm with a microplate reader (iMark, Biorad). Wells were washed three times with PBS containing 0.3% (v/v) Tween20 after each incubation step.

2.4 Western blot procedures

Approximately 10^{10} purified phages were loaded into each lane of 11% SDS-PAGE gel and transferred onto PVDF membrane (Bio-Rad). Then, the membrane was blocked with TBS (20 mM Tris HCl, 150 mM NaCl, pH 7.4) + 5% skim milk for 2 hours at room temperature. Monoclonal mouse anti-g3p antibody (MoBiTec, Goettingen, Germany) was added diluted 1:1000 in TBS+ 5% skim milk and incubated overnight at 4°C. The next day goat anti-mouse IgG antibody conjugated with StarBright Blue 700 (Bio-Rad,) diluted 1:15000 in TBS + 5% BSA were added. After 1 hour of incubation, the membrane was analysed with ChemiDoc Imaging System (BioRad).

3 Results

3.1 Development and validation of the cloning procedure

To achieve a synthetic gene library in the M13 phage vector with considerable complexity, it is crucial to employ an efficient cloning procedure. After testing various methods, the one that, in our opinion, yielded the highest efficiency was the cloning method known as Gateway® technology, developed by Invitrogen (data not shown,[14]). This technology necessitates the conversion of the vector of interest, where the gene library is to be inserted, into a format called DEST (from destination vector), containing a gene coding for the CcdB toxin (a bacterial DNA gyrase inhibitor) at the cloning site. Through site-specific recombination, the genes to be cloned are inserted into the vector by replacing the gene coding for the toxin, enabling after bacterial transformation the selection of almost only recombinant clones. Our target phage vector, M13 flash, was then transformed into the DEST vector (M13 flash-DEST) (**Figure 2**).

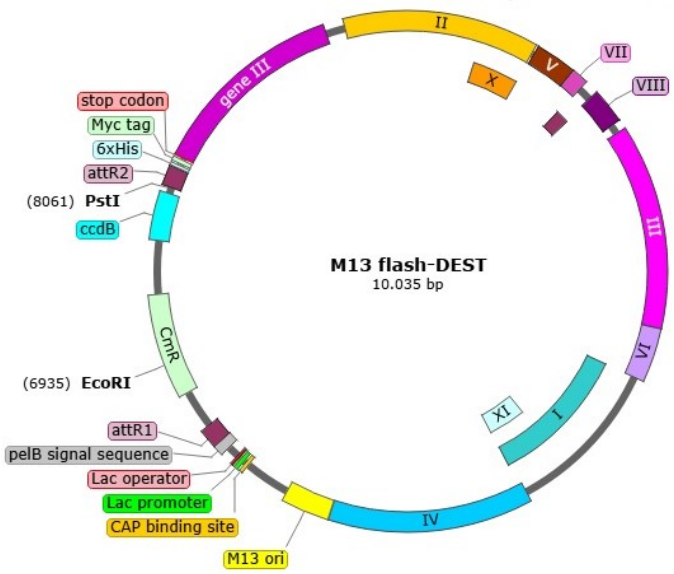


Figure 2. Representation of M13 flash- DEST vector modifications. SnapGene® software was used to design the vector figures (from Dotmatics; available at snapgene.com).

Unfortunately, the counter-selection of non-recombinants, fundamental to the Gateway technology, cannot be implemented for M13 flash-DEST. The bacteria (e.g., TG1 or XL1B strains) typically used for the propagation of this phage vector must contain the F-plasmid, which, on the one hand, encodes the external receptor for phage infection but, on the other hand, contains the gene coding for the antitoxin CcdA, making these strains insensitive to the toxin. The use of the OmniMAX strain (Invitrogen, Waltham, MA, USA), a strain containing the F-plasmid in which the antitoxin gene has been deleted, did not allow the desired selection. Despite the strain being sensitive to the toxin it allowed the replication of non-recombinant phage, as evidenced by the formation of plaques (data not shown). However, requiring this technology for library preparation, a variation in the procedure was introduced: before the bacterial transformation of the recombination products, digestion with restriction enzymes containing corresponding unique cutting sites in the DEST cassette was performed. By taking advantage of the higher transformation efficiency of circular molecules compared to linear ones, the linearization of only the non-recombinant constructs allowed their presence to be minimized among the clones obtained after the entire cloning process (data not shown).

3.2 Construction of a humanised synthetic nanobody library

To construct the coding library for humanized synthetic nanobodies, the desired protein template was initially designed using a constant protein scaffold. This scaffold consisted of the Framework (FW) regions of a universal template nanobody[13], suitably mutagenized to humanize it and make it as similar as possible to the FW of human IgG VH III (hCFW)[14]. The three hypervariable CDR regions were obtained by partially

randomizing amino acids that occur frequently (frequency higher than 10%) in natural nanobody sequences (**Figure 3**).

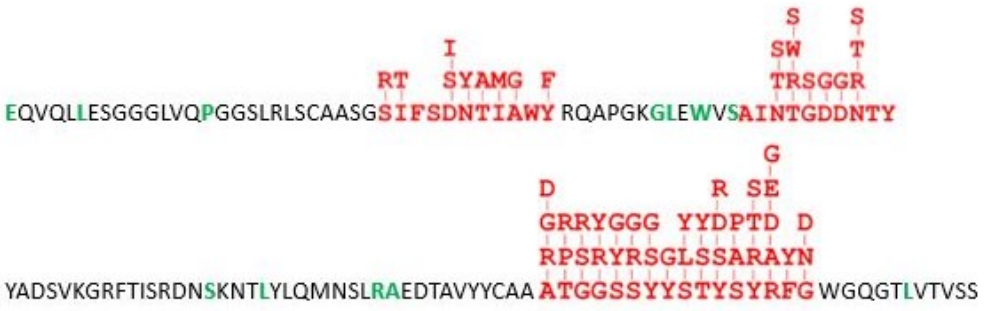


Figure 3. Humanised nanobody consensus sequence with framework regions in black and CDRs in red. In the CDRs, all amino acids occurring with a frequency higher than 10% are shown. Humanising amino acid residues are highlighted in green. The humanization strategy followed is as reported in reference [14].

The coding genes for the hNb library were generated through PCR assembly using eight oligonucleotides and a coding sequencefor a specific humanized Nb (synthesized ad hoc) as a template. The oligonucleotides were optimized for *E. coli* codons, and those corresponding to the three CDR regions had partially degenerate nucleotide sequences (**Figure 4**, see 'Materials and methods' for details). The terminally paired oligonucleotides allowed the *att*B1 and *att*B2 recombination regions required for the Gateway cloning procedure to be inserted into the final amplicon.

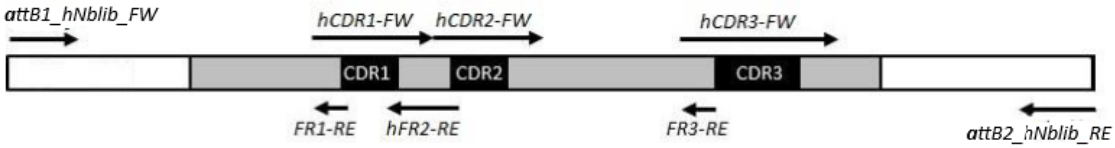


Figure 4. Schematic representation of the construction of the humanised artificial library by PCR assembly. FR regions in grey, CDR regions in black, and primers used are indicated by arrows.

The resulting amplicon was initially transferred into the pDONR222 vector (Invitrogen, Waltham, MA, USA) via BP reaction, maintaining a 1:1 vector: insert molar ratio. The first bacterial transformation was performed using the Top10 strain of *E. coli*, sensitive to the toxic effects of the ccdB cassette, enabling the selection of recombinant clones according to the classical strategy of the Gateway cloning system and the preparation of the hNb DONR library. A small aliquot of the transformed bacteria was plated to estimate the final complexity of the hNb DONR library, resulting in $3.6x10^8$ individual clones. One hundred randomly selected clones were sequenced to verify the correctness of the constructs and to define the percentage of non-recombinant clones. In the sample of picked clones, there were no non-recombinants, and approximately 20% of clones

had incorrect inserts, presumably due to primer errors and/or incorrect assembly of the intermediate amplicons in the PCR assembly reaction.
The hNb DONR library thus produced was harvested, and after extraction, the corresponding plasmid DNA was used in an LR reaction together with the engineered phage vector M13 flash-DEST (1:1 molar ratio). This second recombination reaction allowed for the transfer of the genes coding for humanized Nb into a destination vector usable in the phage-display selection technique. The product of the LR reaction was transformed into the *E. coli* TG1 strain and, for the reasons mentioned above, it was first subjected to a complete digestion reaction with the restriction enzymes EcoRI and PstI, having unique sites in the DEST gateway cassette. The resulting hNb DEST phage library was titrated and found to have a complexity of $3.2x10^8$individual clones. The absence of non-recombinant clones was once again confirmed by sequencing one hundred randomly selected clones, demonstrating also that over 75% of clones turned out to be correct. The correct amino acid sequences of nanobodies encoded by clones are given in the Appendix 1.
Although the desired final complexity of 10^9 clones was not achieved, the order of 10^8 is still within the acceptable range for a synthetic Nb library, by employing a phage display methodology for selection.

3.3 Preliminary screening of humanised synthetic nanobody library

Phages derived from the hNb M13 library should expose nanobodies on their surface. To confirm this exposure and to isolate specific hNb, four rounds of phage display panning were performed against a specific target, namely *Pyrococcus furiosus* thioredoxin (*Pf*Trx) protein (see 'Materials and Methods'). For each panning step, the target concentration and number of washes were kept constant.
At the end of the procedure, a polyclonal phage-ELISA test was conducted to verify the progressive enrichment of the library with phages exposing *Pf*Trx binders. The phages collected after every panning step were analysed against *Pf*Trx protein and a non-correlate protein (GFP protein), used as a negative control. As shown in **Figure 5**, there was an increase in the ELISA signal against *Pf*Trx from the first to the third enriched library (L1-L3 M13hNb library), compared to the starting library (L0 M13hNb library). The ELISA signal against the negative control indicates the unspecific signal contribution. The decrease in the *Pf*Trx signal after the fourth panning (L4) indicates complete enrichment of the library.

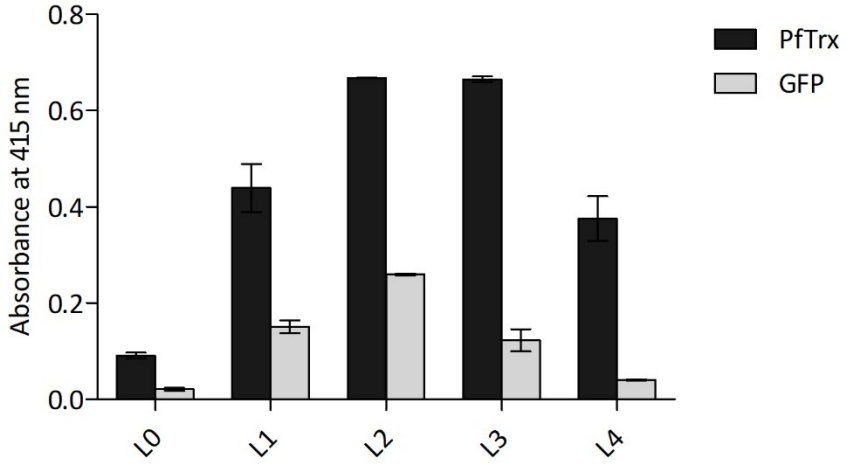


Figure 5. Polyclonal Phage ELISA on phages deriving from the four panning steps (L1-L4) against *Pf*Trx. GFP protein as a negative control. The absorbance values are represented as the mean ± SD of duplicate measurements.

Ninety-six randomly picked clones from the third panning were individually analyzed by monoclonal phage ELISA, but unfortunately, none of the clones showed binding specificity to the *Pf*Trx protein (data not shown).

3.4 Analysis of the valence of M13-flash

The new M13 phage vector used for the preparation of the hNb library, based on its genetic features, should ideally function as a monovalent vector. In other words, it should not expose more than one copy of hNb-pIII on the surface of the phage particle, which normally contains 5 copies of the pIII phage protein.

After failing to isolate individual clones with a specific affinity toward the target, the Nb-pIII:pIII ratio was analysed by western blot. Proteins extracted from an aliquot of phages (approximately 10^{10}) from the hNb library before the panning steps (L0 M13hNb library) and from the enriched library after the third panning of the phage display (L3 M13hNb library) were analysed using a monoclonal antibody against the pIII protein of M13. Proteins extracted from the non-recombinant M13 phage were used as controls. As observed from the hybridization signal (**Figure 6**), the pIII protein was the predominant form on the phage surface in all analysed samples. The signal corresponding to the hNb-pIII fusion protein was not detected in any of the M13 hNb libraries analysed.

Although the L3 M13hNb library exhibited an obvious ability to bind the target (*Pf*Trx), as evidenced by the polyclonal phage-ELISA (**Figure 5**), estimating the Nb-pIII:pIII ratio for this library was challenging and was certainly much lower than the theoretical 1:5.

The low level of exposure of Nb on the phage surface could be attributed to various factors. One cause might be a lower expression by the phage of the hNb-pIII fusion protein compared to that encoded by the endogenous phage pIII gene, potentially due to its less efficient expression signals. Another possible reason could be a low suppression efficiency by the propagating bacterial strain (e.g., TG1). Due to the

presence of the supE suppressor in the genome of this *E. coli* strain, the TAG stop codon interposed between the gene coding for Nb and the pIII protein gene must be suppressed to allow expression of the fusion protein. Inefficient suppression would result in a reduction in the presence of the fusion protein on the phage surface.
Even the addition of the IPTG inducer during phage amplification, meant to activate the lac promoter controlling Nb-pIII expression, did not show an appreciable increase in fusion protein expression on the phage surface (data not shown).

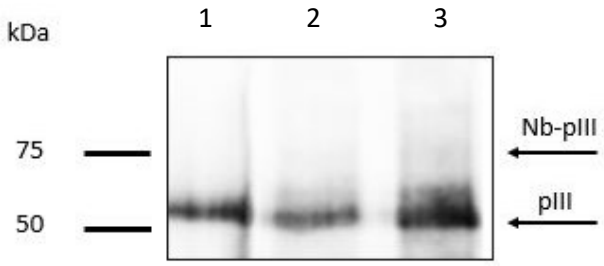


Figure 6. Western blot performed on ~10^{10} phage particles with monoclonal mouse anti-g3p as primary antibody and anti-mouse StarBright Blue 700 as secondary antibody. Lane 1: M13 flash (control), lane 2: L0 M13hNb library, lane 3: L3 M13hNb. The arrows indicate the position of the expected bands for the phage proteins pIII and Nb-pIII.

4 Discussion

The use of antibodies for therapeutic applications in humans requires the humanization of the sequence to try to eliminate the risk of adverse reactions by the immune system. The humanization process is challenging because it can often adversely affect the stability and binding affinity of the antibody[5]. In recent years, synthetic libraries have emerged as a viable alternative to the use of animals. Furthermore, the construction of already humanized synthetic libraries from a stable framework makes it possible to avoid the subsequent humanization process on any specific antibodies isolated after an in vitro selection procedure[7,18].
Here, we present the construction of a humanized synthetic nanobody library selectable by phage display to isolate nanobodies potentially ready for use in human therapeutic applications.
Using a validated and functional human consensus framework (hCFW)[13,14] scaffold as a template, a library of humanized Nb was constructed by assembly PCR (**Figure 4**). The oligonucleotides corresponding to the three CDR regions had a partially degenerated sequence to introduce some degree of hypervariability.
Synthetic libraries are not target-specific and, to ensure maximum variability and increase the probability of obtaining good ligands, they must be large[6]. Our goal was to obtain a library complexity of 10^9 individual clones, which can only be achieved with

efficient cloning strategies. As a cloning method, we chose Gateway® technology, a highly efficient method for transferring DNA into multiple target vector systems. To clone a gene of interest into a preferred expression vector, it is essential that this is first modified appropriately[19]. Gateway® cloning system is based on two site-specific recombinations mediated by bacteriophage λ-proteins: the BP and LR reactions. The high cloning efficiency is also due to the dual selection system ensured by the presence of a selection marker (positive selection) and the ccdB cassette (negative selection). Given the toxic effects of the gene cassette encoding the corresponding CcdB toxin, the choice of the *E. coli* strain to be transformed is important[20].

The library of amplicons coding for hNb, suitably modified at the 5' and 3' ends, was transferred into the pDONR222 vector through a BP reaction and, subsequently, the inserts of the obtained library (hNb-DONR library) were transferred into the final destination vector through an LR reaction. To make this second library selectable by phage display technology, the LR reaction was performed in a new phage vector: M13-flash converted into a destination vector (M13-flash DEST). The final complexity of the resulting library was $3.2x10^8$ individual clones, an order at the lower limit of the accepted range.

Since the DEST cassette proposed by Invitrogen does not facilitate the selection of recombinant phages, a restriction enzyme digestion step was incorporated into the Gateway cloning procedure prior to the LR reaction to eliminate the background of non-recombinant phages. While this additional step allows for the desired library to be obtained, it also extends the time required for the entire process. Direct cloning into a phage vector would necessitate the modification of the DEST cassette, such as replacing the gene encoding the ccdB toxin with one encoding a toxin capable of killing the bacterium before amplification of the infecting phage can occur.

However, we initiated a preliminary selection of this humanized nanobody library by phage display. The M13 flash phage vector speeds up the selection steps because, unlike phagemid vectors, it does not require the use of the helper phage and is designed to maintain the monovalent property of the phagemid vector[14]. Four rounds of panning against the *Pf*Trx protein were performed in less than a week and, as shown by polyclonal phage ELISA analysis, the library was enriched against the target (**Figure 5**).

However, the analysis by phage monoclonal ELISA of 96 clones randomly selected from the library resulting from the third panning did not reveal any clones with significant binding affinity against the *Pf*Trx target protein (data not shown).

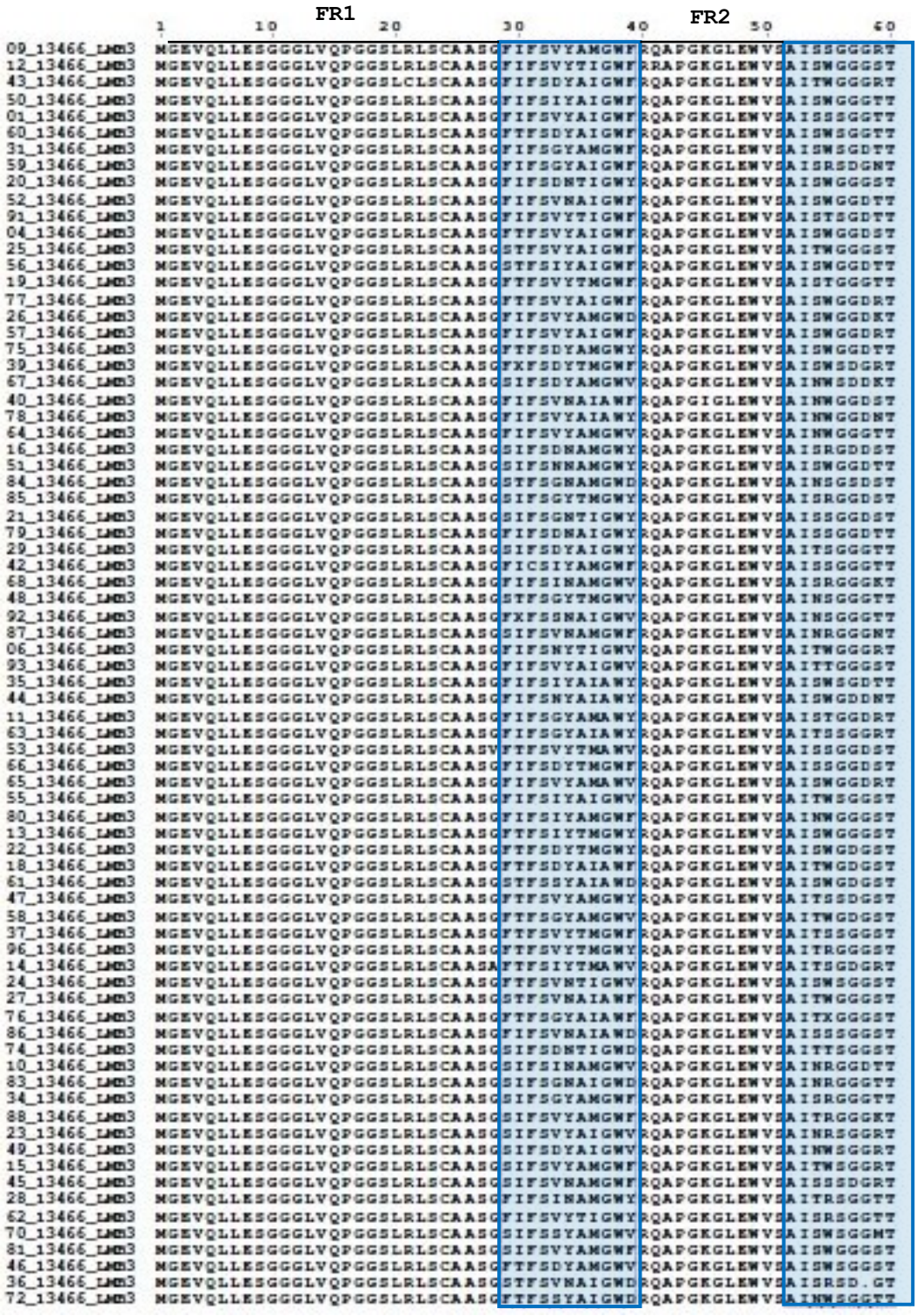
This unsatisfactory result could be related to several factors. Despite the high complexity of the library, there might not be any hNb with sufficiently high affinity for the target. The presence of a weakly positive signal in polyclonal phage ELISA obtained with the enriched library after the third panning could result from the presence of many low-affinity binders within it, which cannot be detected individually by the monoclonal phage ELISA. Isolating a good binder would probably require performing an in-vitro clone maturation procedure on the enriched hNb library, to compensate for the absence of the natural antibody maturation phenomenon that occurs in vivo and allows the amplification of the most closely related antibodies.

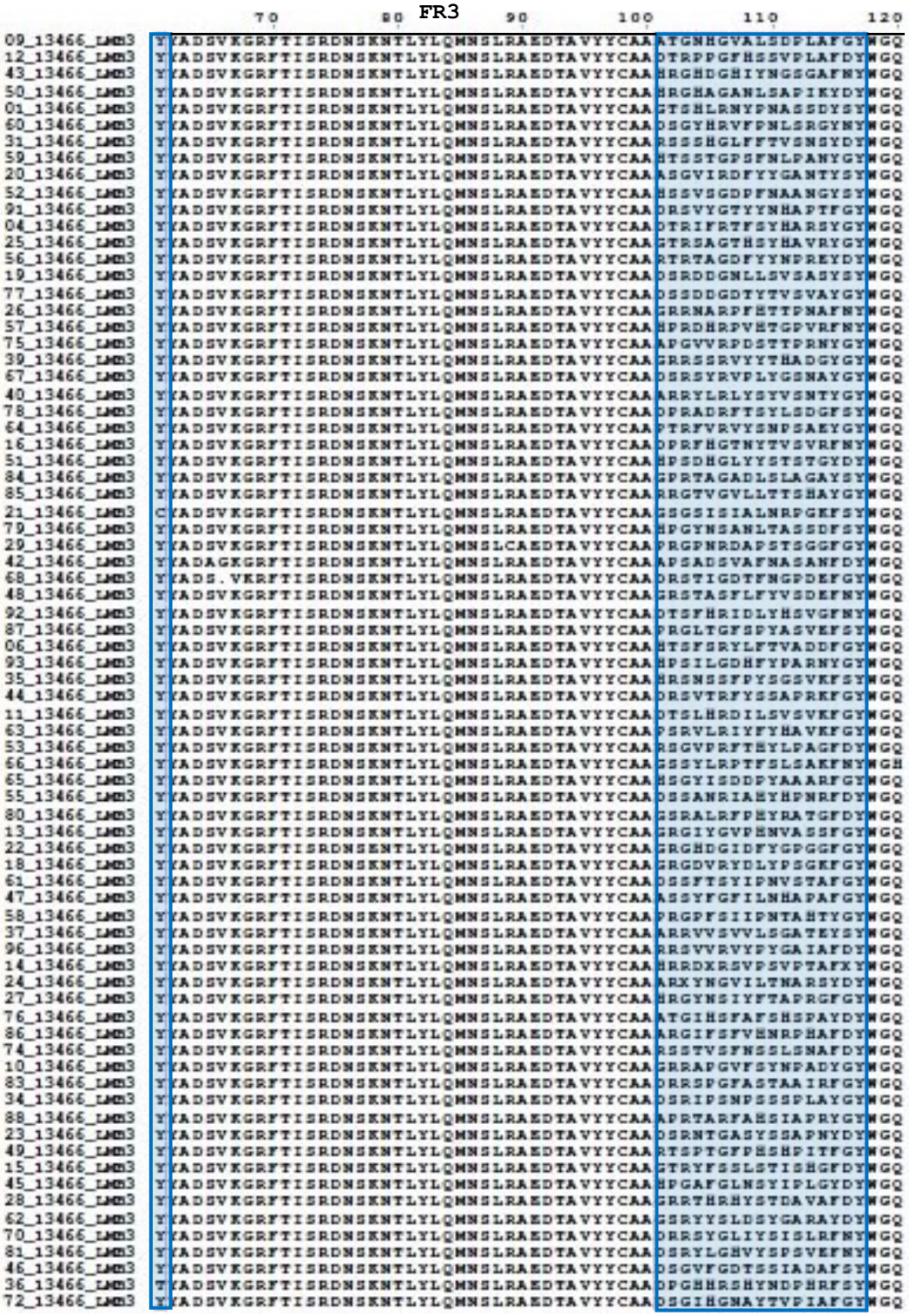
Furthermore, western blot analysis (**Figure 6**) failed to detect the Nb-pIII:pIII ratio on the phage surface, which theoretically should be 1:5[21]. The Nb-pIII:pIII, which is undoubtedly less than 1:5, means that in any panning step, far fewer phage clones are subjected to selection than are actually used, significantly reducing the chances of selecting an efficient binder.

To increase the copy number of Nb-pIII on the phage surface, a polyvalent display could be considered by fusing the Nb library directly with the endogenous pIII from the M13 genome. This solution would require the preparation of a different destination vector, with the possible disadvantage, in this case, of selecting Nb not only based on affinity but also on high avidity for the target.

Appendix 1

Alignment of the amino acid sequences of nanobodies encoded by selected clones from of the humanised Nb library. Framework regions are depicted in black, while the CDR regions (CDR1, CDR2, CDR3) are highlighted in blue.




FR3
70
80
90
100
110
120

	FR4
09_13466_LMm3	GTLVTVSS
12_13466_LMm3	GTLVTVSS
43_13466_LMm3	GTLVTVSS
50_13466_LMm3	GTLVTVSS
01_13466_LMm3	GTLVTVSS
60_13466_LMm3	GTLVTVSS
31_13466_LMm3	GTLVTVSS
59_13466_LMm3	GTLVTVSS
20_13466_LMm3	GTLVTVSS
52_13466_LMm3	GTLVTVSS
91_13466_LMm3	GTLVTVSS
04_13466_LMm3	GTLVTVSS
25_13466_LMm3	GTLVTVSS
56_13466_LMm3	GTLVTVSS
19_13466_LMm3	GTLVTVSS
77_13466_LMm3	GTLVTVSS
26_13466_LMm3	GTLVTVSS
57_13466_LMm3	GTLVTVSS
75_13466_LMm3	GTLVTVSS
39_13466_LMm3	GTLVTVSS
67_13466_LMm3	GTLVTVSS
40_13466_LMm3	GTLVTVSS
78_13466_LMm3	GTLVTVSS
64_13466_LMm3	GTLVTVSS
16_13466_LMm3	GTLVTVSS
51_13466_LMm3	GTLVTVSS
84_13466_LMm3	GTLVTVSS
85_13466_LMm3	GTLVTVSS
21_13466_LMm3	GTLVTVSS
79_13466_LMm3	GTLVTVSS
29_13466_LMm3	GTLVTVSS
42_13466_LMm3	GTLVTVSS
68_13466_LMm3	GTLVTVSS
48_13466_LMm3	GTLVTVSS
92_13466_LMm3	GTLVTVSS
87_13466_LMm3	GTLVTVSS
06_13466_LMm3	GTLVTVSS
93_13466_LMm3	GTLVTVSS
35_13466_LMm3	GTLVTVSS
44_13466_LMm3	GTLVTVSS
11_13466_LMm3	GTLVTVSS
63_13466_LMm3	GTLVTVSS
53_13466_LMm3	GTLVTVSS
66_13466_LMm3	GTLVTVSS
65_13466_LMm3	GTLVTVSS
55_13466_LMm3	GTLVTVSS
80_13466_LMm3	GTLVTVSS
13_13466_LMm3	GTLVTVSS
22_13466_LMm3	GTLVTVSS
18_13466_LMm3	GTLVTVSS
61_13466_LMm3	GTLVTVSS
47_13466_LMm3	GTLVTVSS
58_13466_LMm3	GTLVTVSS
37_13466_LMm3	GTLVTVSS
96_13466_LMm3	GTLVTVSS
14_13466_LMm3	GTLVTVSS
24_13466_LMm3	GTLVTVSS
27_13466_LMm3	GTLVTVSS
76_13466_LMm3	GTLVTVSS
86_13466_LMm3	GTLVTVSS
74_13466_LMm3	GTLVTVSS
10_13466_LMm3	GTLVTVSS
83_13466_LMm3	GTLVTVSS
34_13466_LMm3	GTLVTVSS
88_13466_LMm3	GTLVTVSS
23_13466_LMm3	GTLVTVSS
49_13466_LMm3	GTLVTVSS
15_13466_LMm3	GTLVTVSS
45_13466_LMm3	GTLVTVSS
28_13466_LMm3	GTLVTVSS
62_13466_LMm3	GTLVTVSS
70_13466_LMm3	GTLVTVSS
81_13466_LMm3	GTLVTVSS
46_13466_LMm3	GTLVTVSS
36_13466_LMm3	GTLVTVSS
72_13466_LMm3	GTLVTVSS

5 Bibliography

1. Hamers-Casterman C, Atarhouch T, Muyldermans S, et al. Naturally occurring antibodies devoid of light chains. *Nature*. 1993; 363:446-448. doi:10.1038/363446a0
2. Muyldermans S. Nanobodies: Natural single-domain antibodies. *Annu Rev Biochem*. 2013; 82:775-797. doi:10.1146/annurev-biochem-063011-092449
3. Ackaert C, Smiejkowska N, Xavier C, et al. Immunogenicity Risk Profile of Nanobodies. *Front Immunol*. 2021;12. doi:10.3389/fimmu.2021.632687
4. Vincke C, Loris R, Saerens D, Martinez-Rodriguez S, Muyldermans S, Conrath K. General Strategy to Humanize a Camelid Single-domain Antibody and Identification of a Universal Humanized Nanobody Scaffold. *Journal of Biological Chemistry*. 2009;284(5):3273-3284. doi:10.1074/JBC.M806889200
5. Soler MA, Medagli B, Wang J, et al. Effect of Humanizing Mutations on the Stability of the Llama Single-Domain Variable Region. 2021; 11:163. doi:10.3390/biom11020163
6. Valdés-Tresanco MS, Molina-Zapata A, Pose AG, Moreno E. Structural Insights into the Design of Synthetic Nanobody Libraries. *Molecules 2022, Vol 27, Page 2198*. 2022;27(7):2198. doi:10.3390/MOLECULES27072198
7. Moutel S, Bery N, Bernard V, et al. NaLi-H1: A universal synthetic library of humanized nanobodies providing highly functional antibodies and intrabodies. *Elife*. 2016;5(JULY). doi:10.7554/ELIFE.16228
8. Zhang Y. Evolution of phage display libraries for therapeutic antibody discovery. *MAbs*. 2023;15(1). doi:10.1080/19420862.2023.2213793
9. Kehoe JW, Kay BK. Filamentous phage display in the new millennium. *Chem Rev*. 2005;105(11):4056-4072. doi:10.1021/CR000261R
10. Smith GP, Petrenko VA. Phage display. *Chem Rev*. 1997;97(2):391-410. doi:10.1021/CR960065D
11. Lee CMY, Iorno N, Sierro F, Christ D. Selection of human antibody fragments by phage display. *Nature Protocols*. 2007;2(11):3001-3008. doi:10.1038/nprot.2007.448
12. Jaroszewicz W, Morcinek-Orłowska J, Pierzynowska K, Gaffke L, Węgrzyn G. Phage display and other peptide display technologies. *FEMS Microbiol Rev*. 2022;46(2). doi:10.1093/femsre/fuab052
13. Ferrari D, Garrapa V, Locatelli M, Bolchi A. A Novel Nanobody Scaffold Optimized for Bacterial Expression and Suitable for the Construction of Ribosome Display Libraries. *Mol Biotechnol*. 2020;62(1):43-55. doi:10.1007/S12033-019-00224-Z
14. Garrapa V. *A Novel Scaffold for Nanobody Selection and Expression Suitable for Diagnostic and Therapeutic Applications*. 2019. Accessed January 9, 2024. https://www.repository.unipr.it/handle/1889/3722
15. Life Technologies Carlsbad CA. Gateway ® Technology with Clonase ® II A universal technology to clone DNA sequences for functional analysis and expression in multiple systems. *User guide*.

16. Bahassi EM, Salmon MA, van Melderen L, Bernard P, Couturier M. F plasmid CcdB killer protein ccdB gene mutants coding for non-cytotoxic proteins which retain their regulatory functions. *Mol Microbiol*. 1995;15(6):1031-1037. doi:10.1111/J.1365-2958.1995.TB02278.X
17. Bernard P, Gabarit P, Bahassi EM, Couturier M. Positive-selection vectors using the F plasmid ccdB killer gene. *Gene*. 1994;148(1):71-74. doi:10.1016/0378-1119(94)90235-6
18. Zimmermann I, Egloff P, Hutter CAJ, et al. Synthetic single domain antibodies for the conformational trapping of membrane proteins. *Elife*. 2018;7. doi:10.7554/ELIFE.34317
19. Hartley JL, Temple GF, Brasch MA. DNA Cloning Using In Vitro Site-Specific Recombination. *Genome Res*. 2000;10(11):1788-1795. doi:10.1101/GR.143000
20. Katzen F. Gateway® recombinational cloning: a biological operating system. *Expert Opin Drug Discov*. 2007;2(4):571-589. doi:10.1517/17460441.2.4.571
21. Kramer RA, Cox F, van der Horst M, et al. A novel helper phage that improves phage display selection efficiency by preventing the amplification of phages without recombinant protein. *Nucleic Acids Res*. 2003;31(11): e59-e59. doi:10.1093/NAR/GNG058

Chapter 2

Chapter 2: A novel destination cassette for extending Gateway cloning to phage display vectors

1 Introduction

Currently, antibodies and their derivatives play crucial roles in research, diagnostics, and therapy. Notably, the exploration of antibody fragments like Fab, Nb, and scFv has generated significant interest within the scientific community. While constructing antibody libraries or fragments enables the isolation of unique clones, selecting molecules with desired properties necessitates the use of an in vitro selection method.
Among the various library selection methods, phage display stands out as the most robust and widely employed. Since its discovery in 1985 by George Smith, initially for displaying peptides[1] and the publication of the first antibody fragment displayed on phage surface in 1990[2], phage display technology has been highly successful in discovering numerous antibodies for both research and therapeutic applications[3].
Filamentous bacteriophages (Ff) that include f1, fd and M13 infect *E. coli* cells that carry F-plasmid are involved in this process, but the most popular phage used is M13.
The system is based on the display of proteins or peptides on the surface of bacteriophages by genetic fusion of the coding sequence of the foreign protein with that of a phage coat protein. One of the key coat proteins for the display of foreign proteins on the phage surface is pIII, responsible for phage infectivity[4].
Phage and phagemid vectors are utilized as gene cloning vectors for this purpose. However, phagemids are more commonly used than phages for several reasons. These include the smaller size of a phage genome, which can accommodate larger foreign DNA fragments. Additionally, phagemids are easier to manipulate in vitro and can be more readily transferred into a bacterial cell. Regarding their morphology, phagemid vectors contain two origins of replication: one is a classical origin of an *E. coli* plasmid and the other the origin of M13 phage. They also include an antibiotic-resistance gene, a gene encoding a phage coat protein (typically gene III) with a secretion signal peptide, the expression of which is controlled by an inducible bacterial promoter (commonly the lac promoter) and a multiple cloning site for the insertion of foreign DNA. They do not contain any of the other genes required for a complete and functional phage and are therefore unable to assemble phage particles on their own, requiring the support of a helper phage that provides *in trans* all the structural proteins required for phage packaging[5]. These structural features result in a monovalent display, where only one copy of the protein/peptide is produced per phage particle. This monovalent display allows discrimination between low and high-affinity binders[6].
Generally, an important aspect not to be underestimated when constructing libraries of antibodies or antibody fragments, whether of natural origin (immune or naïve library) or artificially generated (synthetic library), is the size we aim to achieve. Indeed, libraries must be sufficiently complex to isolate high-affinity ligands with a high probability.

This requires the use of simple and efficient cloning methods. Among the various systems proposed as alternatives to conventional digestion-ligation, the Gateway cloning method has emerged as one of the most robust and efficient[7].

In Gateway technology (Invitrogen), the cloning of the DNA of interest into a target vector involves two cloning steps, known as BP and LR reactions. The two steps involve two site-specific recombination reactions to transfer the DNA of interest first into an intermediate vector, called the donor vector, and then into a desired destination vector, provided it has been modified by insertion of the ccdB cassette[8].

The purpose of the ccdB cassette is to counter-select recombinants by allowing the expression of the CcdB toxin in bacteria. The bacterial toxin CcdB (Controller of Cell Death protein B) is part of a toxin-antitoxin system involved in maintaining the F plasmid in *E. coli*[9]. Specifically, the ccd operon of plasmid F encodes two proteins: CcdB, the toxin that targets the GyrA subunit of DNA gyrase, an essential type II topoisomerase of E. coli; CcdA, the antitoxin that interacts with CcdB and neutralises its activity[10]. Based on this, the selection mechanism is effective when a toxin-sensitive bacterial strain, i.e. one that does not express the CcdA antitoxin, is used to transform the products of the BP and LR reactions.

The bacterial strains used in phage display contain the F plasmid required for phage infection and subsequent propagation. However, this renders them insensitive to the toxic effects of CcdB due to the presence of the ccdA antitoxin gene carried by the F plasmid. Consequently, a phagemid carrying a ccdB cassette can only be used to isolate recombinants if it is transformed into a sensitive strain lacking the F plasmid. If recombinant phagemids are to be used in phage displays, it is necessary to transfer all selected recombinant clones into a strain containing the F plasmid[11].

To avoid these time-consuming steps, here we present the modification of ccdB cassette in the target phagemid to allow direct selection of the recombinant in any strain carrying the F plasmid after transformation with the LR reaction. Then, we demonstrate the functionality of the new cassette in phagemid recombinant selection. As a proof of concept, an immune nanobody library was generated in this new vector using VHH sequences derived from a llama immunised with a specific target protein.

2 Materials and Methods

2.1 Molecular biology reagents and bacterial strains

Standard protocols were used for all basic recombinant DNA procedures[12]. Enzymes were from Takara Bio Europe (Saint-Germain-en Laye, France) and New England Biolabs (Ipswich, MA, USA). Molecular cloning was carried out in Top10 (Invitrogen, Waltham, MA, USA) and TG1 (LGC Biosearch Technologies, Hoddesdon, UK) *E. coli* strains, and recombinant protein expression was carried out in BL21 (DE3) codon plus *E. coli* strains (Stratagene; La Jolla, CA, USA), using pIT2 (http://www.geneservice.co.uk/products/proteomic/datasheets/tomlinsonIJ.pdf) and pET28 (Merck Millipore, Darmstadt, Germany) plasmids.

2.2 Modification of pIT2 phagemid vector into a destination vector

The Gateway attP1-2 cassette nucleotide sequence (destination cassette) was amplified by PCR from the pDEST32 vector by introducing sequences overlapping the cloning site of the pIT2 vector at the 5' and 3' ends. PCR was performed using 1 unit of Phusion DNA Polymerase (New England Biolabs, Ipswich, MA, USA) in the presence of attR1-pDEST32-FW and attR2-pDEST32RE oligonucleotides (4 nM each) (see **Table 1**), plus dNTPs (0.8 mM), 1x PCR buffer, and 10 ng of the pDEST32 vector in a final volume of 50 µl and the following PCR condition: initial denaturation at 98°C 3 min; 25 cycles of 30 s at 98°C, 30 s at 55°C, 90 s at 72°C; final extension at 72°C for 10 minutes. The pIT2 phagemid was digested with EcoRI and NotI restriction enzymes using a standard protocol[12]. The ligation of the Gateway attP1-2 cassette nucleotide into the pIT2 phagemid was performed with InFusion HD cloning kit according to the user's manual instructions[13]. This cloning step generated a destination pIT2 vector suitable for Gateway® technology (Invitrogen, Waltham, MA, USA): pIT2 DEST. Because the CcdB, expressed by destination cassette, is a toxic protein, this vector was constructed and propagated in TG1 *E. coli* strain.

Starting from 50 ng of the pIT2 DEST vector template, we conducted site-specific mutagenesis through circular PCR using Dest-pLac-ccdB2-FW and Dest-pLac-ccdB2-RE oligonucleotides (0.5 µM each) (refer to **Table 1**), dNTPs (200 µM) and 1 unit of Phusion Hot Start II High Fidelity DNA polymerase (Thermo Scientific, Waltham, MA, USA), and the following PCR conditions: an initial denaturation at 98°C for 3 min, 25 cycles of 30 s at 98°C, 30 s at 62°C, and 3 min at 72°C, and a final extension at 72°C for 5 min. The primers were strategically designed to incorporate the desired modifications and included a 20-base complementary region at the 5' end to facilitate self-annealing. Following purification from agarose gel utilizing a Macherey-Nagel kit (Düren, Germany), the resulting amplicon underwent self-ligation using the T5 Exonuclease DNA Assembly (TEDA) cloning procedure[14]. This cloning step yielded a novel pIT2 phagemid vector: pIT2-ccdB.

Primer	Sequence
attR1-pDEST32-FW	cccagccggCCATGGcgaatcaaACAAGTTTGTACAAA
attR2-pDEST32RE	tgatgatgtGCGGCCGCccatAaaACCACTTTGTACAAG
Dest-pLac-ccdB2-FW	catcAaggagacagtcataATGCAGTTTAAGGTTTACACCTATA
Dest-pLac-ccdB2-RE	Ttatgactgtctcctgatgctgccaacttagcgg

Table 1. Oligonucleotides sequence used in PCRs. Sequences are shown in a standard 5'-3' orientation.

2.3 llama immunization protocol, B cell isolation, total RNA extraction and cDNA synthesis (preclinics GmbH, Potsdam, Germany).

The target antigen chosen for immunization was thioredoxin protein from *P. furiosus* bacteria (*Pf*Trx), which was expressed in *E. coli* and purified by affinity chromatography as described in[15]. For generating an immune library, a male llama (Lama glama) was immunized subcutaneously with 300 µg *Pf*TRX in 1 mL physiological saline mixed with an equal volume of Freund's adjuvant (InvivoGen SAS, Toulouse, France). Three weeks after priming with complete Freund's adjuvant, the animal received three boost immunizations using incomplete Freund's adjuvant in a two-week interval. To monitor the development of immune responses, serum samples were gained directly before prime immunization and on days 32 and 42 of the immunization project. Four days after the fourth injection (day 53), a final blood sample of 200 mL was gained by venipuncture using a sterile vacuum bottle (megro GmbH & Co. KG, Wesel, Germany) prepared with 2 mL heparin sodium solution (B. Braun Melsungen AG, Melsungen, Germany), along with a final serum sample. Subsequently, peripheral blood mononuclear cells (PBMCs) were isolated by density gradient centrifugation on lymphocyte separation medium (LSM, Corning Inc., Corning, NY, USA) with a density of 1.077 to 1.080 g/mL at 20°C. The isolated cells from the interphase were collected and washed in DPBS (Corning). Finally, the cells were counted and lysed in RA1 buffer (Macherey-Nagel, Düren, Germany) with 1 % β-mercaptoethanol and stored frozen until subsequent RNA and cDNA preparation. RNA was isolated from the lysed PBMCs using NucleoSpin RNA Midi kit for RNA purification (Macherey-Nagel, Düren, Germany) according to the manufacturer's instructions. For the preparation of cDNA, RevertAid H Minus First Strand cDNA Synthesis Kit (K1632, Thermo Fisher Scientific Inc.) was used according to the manufacturer's instructions. cDNA reactions were performed using both oligo(dT)18 and random hexamer primers, and the resulting cDNA reactions were pooled afterwards.

2.4 Amplification of VHH CDS and construction of the immune library in the novel vector

To amplify all the coding sequences (CDS) of immunoglobulin heavy chains (VHs and VHHs), ten PCR reactions were performed using the cDNA, from Llama immunised, as a template. Every single PCR reaction was conducted using CALL001 and CALL002 (Eurofins Genomics) oligonucleotides (0.4 µM each), listed in Table 2, along with 1 unit

of Phusion DNA polymerase and dNTPs (200 µM). The PCR conditions included an initial denaturation at 98°C for 3 min; followed by 25 cycles of 1 min at 98°C, 1 min at 57°C, and 90 s at 72°C; and a final extension at 72°C for 7 min. The resulting amplicons, predicted to be 700 bp and 1000 bp in size, underwent separation through agarose gel electrophoresis. The smallest amplicon corresponding to the VHH CDS was purified from the agarose gel using a gel purification kit (Macherey-Nagel Düren, Germany). Subsequently, it was reamplified using the primers GW-attB1-Llamalib-FW and GW-attB2-Llamalib-RE (Eurofins Genomics) (refer to **Table 2**). This new amplification step provided the VHH sequences with attB sites at the 5' and 3' ends necessary for the first Gateway BP recombination. Following the manufacturer's instructions[8], the *att*B-VHH sequences were cloned into the pDONR222 vector (Invitrogen, Waltham, MA, USA) using the BP clonase reaction. The entry library was obtained by transforming the Top10 *E. coli* strain with the resulting constructs. The collected entry library was then amplified and subjected to plasmid purification. An aliquot of this plasmid preparation was used to perform the LR clonase reaction using the novel pIT2-ccdB phagemid vector as the destination vector. After the transformation of the TG1 *E. coli* strain, the expression library was selected in the presence of 1 mM IPTG.

Primer	Sequence
CALL001	GTCCTGGCTGCTCTTCTACAAGG
CALL002	GGTACGTGCTGTTGAACTGTTCC
GW-attB1-Llamalib-FW	GGGGACAAGTTTGTACAAAAAAGCAGGCTatGTgCAGCTGCAGGAGTCTGG
GW-attB2-Llamalib-RE	CCgctACCACTTTGTACAAGAAAGCTGGGTtGgagACGGTGACCtgggt

Table 2. Oligonucleotides sequence used in PCRs. Sequences are shown in a standard 5'-3' orientation

2.5 Phage display selection

The phage display procedure reported by Pardon et al[16] with some modifications was performed with some modifications.

2.5.1 Amplification of immune library phages

Before each round of panning, 10^7 of VHH expression-library bacterial cells were amplified in 2x TY medium supplemented with 100 µg/ml carbenicillin and 2% (wt/vol) glucose. The phage rescue was performed by infecting 10 ml of log-phase grown cells with an excess of 10-fold of the helper phage KM13, in 2x TY medium with 100 µg/ml carbenicillin and 25 µg/ml kanamycin. Before each panning step, the rescued phage library was purified with a 20 % PEG6000 - 2.5 M NaCl solution and titrated.

2.5.2 Panning procedures

The panning procedures were executed in 96-well polystyrene microtitre plates. Initially, 50 µl of a 2 µg/ml casein-glutathione solution, prepared as previously outlined [17] and dissolved in 50 mM carbonate buffer, was introduced into the wells and left to incubate at 4°C overnight. Following three washes with PBS containing 0.3% (v/v) Tween20, the

wells underwent blocking with 2% skim milk in PBS buffer for 1 hour at room temperature. After three additional washes, 500 nM of the target protein *Pf*Trx, fused to the GST carrier protein and provided by collaborators, was applied. Utilizing the GST protein, known for its high affinity to glutathione, facilitated the exposure of the target, preventing any conformational deformations that might occur with direct adhesion. Following a 1-hour incubation at room temperature and five subsequent washes, an aliquot of 10^{11} phages from the library was introduced to the well containing the target. The plate underwent a 2-hour incubation on a shaking platform (700 rpm). After washes to eliminate unadsorbed phages, 100 µL of 0.2 M glycine-HCl buffer (pH 2.2) was added to each selection well. The plate was then incubated for 15 minutes at room temperature, shaking at 700 rpm, to elute phages bound to the target. The eluted phages were neutralized with 1 M TrisHCl (pH 9), and 50 µl were employed to infect 3 ml of exponentially growing TG1 cells. An enriched library amplification was carried out, as described previously for the initial VHH expression library. Throughout the three panning rounds, the binding stringency of the selected phages was heightened by reducing the amount of target bound to the selection well (500, 100, and 50 nM, respectively) and increasing the number of unbound phage washes (3, 5, and 10 washes, respectively).

2.5.3 Polyclonal and monoclonal phage ELISA

Polyclonal phage ELISA was performed employing the casein-glutathione/GST-*Pf*Trx system[17] to immobilize the target onto 96-well polystyrene microtitre plates. A 10^{10}-fold dilution in PBS of eluted phages from each panning round was applied and left to incubate for 1 hour at room temperature. Subsequently, the plate was incubated with an anti-M13 (pVIII)-HRP (GE Healthcare, Little Chalfont, USA) solution, diluted 1:5000 in PBS, for an additional hour at room temperature. The signal was identified by introducing 2,2'-azino-di- [3-ethylbenzthiazoline sulfonate (ABTS) substrate (KPL, Gaithersburg, MD, USA), and the absorbance was gauged at 415 nm using a microplate reader (iMark, Biorad). Following each incubation step, wells were subjected to three washes with PBS containing 0.3% (v/v) Tween20.
For monoclonal phage ELISA, performed after the conversion of individual clones into phages using KM13 helper phage[18], the same casein-glutathione/GST-*Pf*Trx system as in the polyclonal phage ELISA was employed.

2.6 Expression and purification of anti-*Pf*Trx nanobody

Standard procedures were used for the expression of anti-*Pf*Trx nanobody. Specifically, induction was performed by overnight cultures (LB, 50 mg/l of kanamycin and 34 mg/l of chloramphenicol) in autoinducing LB medium at 20°C overnight. After cell harvesting, bacterial cell lysis was performed by sonication by sonication (Sonicator 3000, Misonix) in 50 ml of Tris-buffered saline (25 mM Tris pH 8.0, 0.3 M NaCl, freshly supplemented with protease inhibitors). The bacterial lysate was centrifuged (10000 rpm, 30 min at 4°C) and the resulting soluble fraction was used in affinity chromatography on His Trap FF column (GE Healthcare) thanks to the presence of 6xHis-tag at the C-terminus of the

protein. Bacterial lysates and metal-affinity chromatography fractions were analysed by SDS-PAGE on 15% polyacrylamide gels.

2.7 Peptide ELISA

Peptide ELISA was conducted using the casein-glutathione/GST-*Pf*Trx system to immobilize the target on 96-well polystyrene microtitre plates. Serial 1:2 dilutions of the purified anti-*Pf*Trx nanobody were introduced into the wells. The bound nanobody was then identified using an anti-His tag monoclonal antibody-HRP conjugated (proteintech®), which was appropriately diluted at 1:5000 in PBS buffer. For the visualization of the signal, 2,2'-azino-di- [3-ethylbenzthiazoline sulfonate (ABTS) substrate (KPL, Gaithersburg, MD, USA) was added, and the absorbance was measured at 415 nm using a microplate reader (iMark, Biorad).

3 Results

3.1 Mutagenesis of the pIT2 phagemid vector

In order to enable the direct use of E. coli F' strains in the recombinant selection process following an LR reaction, as envisioned by the Gateway cloning, in the phagemid of interest, we chose to modulate the expression of the CcdB toxin. This regulation should allow its overexpression upon induction, reaching intracellular levels that cannot be neutralised by basal levels of CcdA antitoxin produced by the corresponding gene present on the F' episome. The phagemid pIT2, previously modified into a destination vector compatible with Gateway cloning (pIT2-DEST) (see 'Materials and Methods'), underwent mutagenesis. Specifically, the destination cassette was mutagenized by inverse-PCR (see Materials and Methods for details), as described below. The recombination sites (attR1 and attR2) required for cloning were retained, the gene conferring chloramphenicol resistance was deleted, and the gene coding for CcdB was transferred under the Lac promoter, already present in the phagemid. Furthermore, the insertion of a second ribosome binding site allows the formation of a bicistronic mRNA, the second sequence of which codes for the toxin. Consequently, the expression of the CcdB toxin can be repressed or induced by the addition of glucose or IPTG to the culture medium, respectively. The resulting phagemid vector, named pIT2-ccdB, and verified by sequencing contained the modified region of interest, as shown in **Figure 1**.

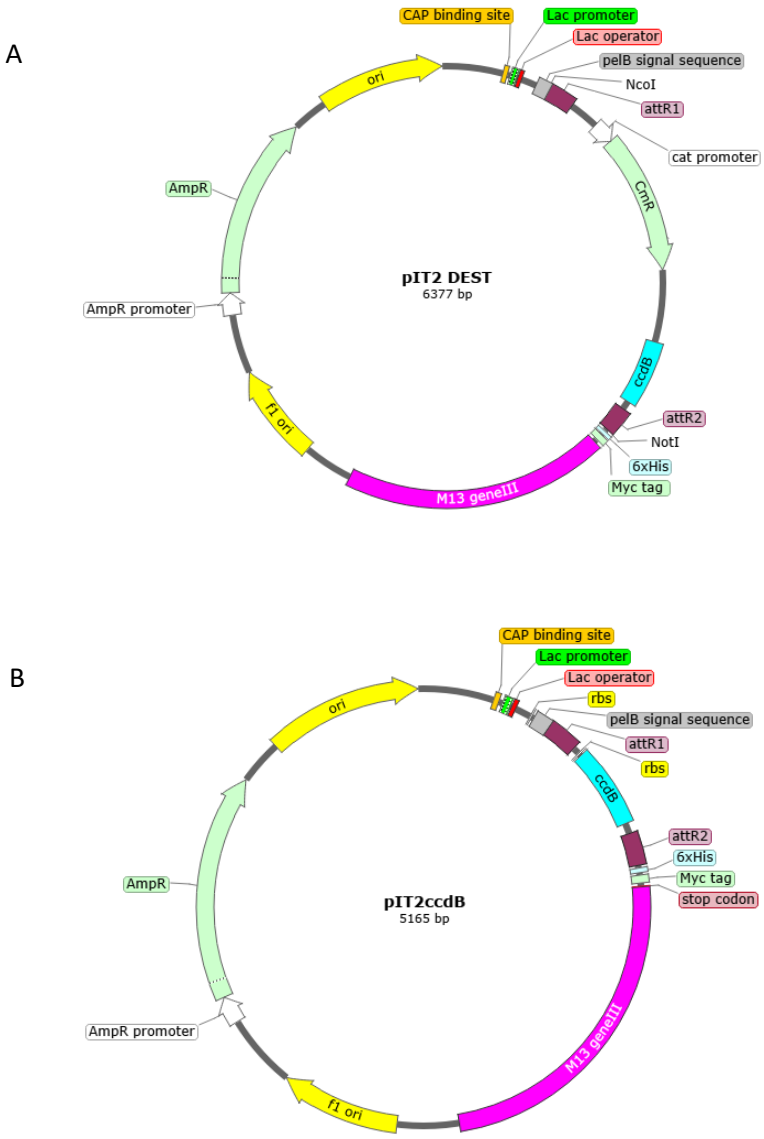


Figure 1. Representation of pIT2 phagemid vector modifications. A: pIT2 destination vector, B: pIT2-ccdB. SnapGene® software was used to design the vector figures (from Dotmatics; available at snapgene.com).

The ability of the pIT2-ccdB phagemid vector to kill host bacteria containing the F' episome, even in the presence of basal levels of CcdA antitoxin, was assessed. TG1 strain was transformed with the pIT2-ccdB vector and then plated on a selective medium supplemented with antibiotics in the presence of either 2% glucose or 1 mM IPTG. By counting the number of colonies grown under conditions suppression and induction of toxin expression, we observed a 10,000-fold reduction in the number of colonies due to toxin toxicity. The same behavior was also observed in other normally toxin-resistant bacterial strains (data not shown). This modification allows the propagation of the pIT2ccdB vector under normal growth conditions in the presence of glucose. After the second Gateway cloning reaction (LR reaction), however, recombinant selection can be

achieved by the addition of IPTG in any bacterial strain, even those containing the F plasmid.

3.2 Construction of an immune nanobody library in the novel pIT2-ccdB phagemid vector

To evaluate the efficacy of the novel destination vector, we utilized it to construct an immune library of Llama Nanobodies (Nbs) employing Gateway technology. The target antigen chosen for immunization was thioredoxin protein from *P. furiosus* bacteria (*Pf*Trx), which was expressed in *E. coli* and purified by affinity chromatography as described in[15]. The immunization protocol, outlined in the Materials and Methods section, was followed to administer the protein to a llama. Following immunization, B cells were isolated from a 200-ml blood sample from the immunized animal. Subsequently, mRNA extraction was performed, and the extracted mRNA was converted into cDNA using the methods described in the Materials and Methods. The entire procedure, from animal immunization to cDNA production, was conducted by preclinics GmbH (Potsdam, Germany).

Using cDNA from immunized Llama as a template, sequences encoding the variable domains of all immunoglobulin heavy chains (VHs and VHHs) were amplified using the oligonucleotides CALL001 and CALL002. Following gel electrophoresis and purification, the VHH-specific amplicon, slightly smaller at 700 bp compared to the VHH counterpart (1000 bp), underwent a second round of PCR. Specific oligonucleotides (refer to 'Materials and Methods') were employed in this step to introduce the *att*B1 and *att*B2 regions at the 5' and 3' ends, facilitating the BP reaction— the initial Gateway cloning reaction—into the intermediate vector pDONR222. Upon transformation into the *E. coli* Top10 strain, the resulting library, with a complexity of 5.8x 10^6 individual clones, underwent screening through sequencing a representative subset of clones. The entry library was then utilized in the LR reaction— the subsequent Gateway cloning reaction— using the novel phagemid vector pIT2-ccdB. This reaction was subsequently introduced into the *E. coli* TG1 strain, and recombinant selection was achieved by adding the inducer IPTG. The final library complexity in this case was 3.2x10^6 individual clones. A sequencing analysis of 96 randomly chosen clones validated their recombinant nature. The correct amino acid sequences of nanobodies encoded by clones are given in the Appendix 2.

3.3 Enrichment of the nanobody library by phage display

Phages derived from the phage library prepared using the helper phage KM13, as described in 'Materials and Methods', were subjected to three rounds of panning. In each round, the target was immobilised using the casein-glutathione/GST-*Pf*Trx system. The three panning were preceded by a counter-panning step using only casein-glutathione. Phages eluted after each panning round were used to infect *E. coli* TG1 cells, resulting in the isolation of three enriched phagemid libraries. After phage conversion of the original immune-library and the three enriched libraries, the phages were analysed by a polyclonal phage-ELISA. Given the expectation that the libraries were

enriched with phages displaying anti-*Pf*Trx nanobodies on their surface, the target antigen was employed as the coating agent in the well. The target antigen was immobilized using the same casein-glutathione/GST-*Pf*Trx system. Casein-glutathione alone served as a negative control.
As shown in **Figure 2**, a clear signal indicating binding to the target was detected after the third panning cycle. Notably, this signal was statistically significant compared to the signals obtained from the previous enriched libraries.

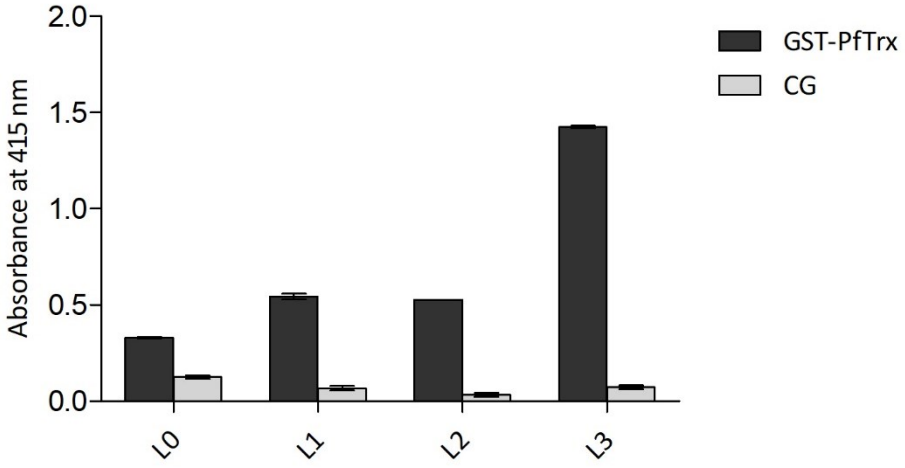


Figure 2. Polyclonal phage ELISA on phages derived from three panning rounds against *Pf*Trx. Casein-glutathione (CG) as a negative control. The absorbance values are represented as the mean ± SD of duplicate measurements.

Ninety-two phage clones were randomly selected from the third enriched library. After conversion to phage using the helper phage, they were analysed by monoclonal phage ELISA. Similar to the previous polyclonal phage ELISA, the casein glutathione/GST-*Pf*Trx system was used. Among the 92 clones, only three exhibited a signal significantly higher than the background. Gene sequencing showed that these three clones had identical sequences. The amino acid sequence common to all the clones is shown in **Figure 3**.

Figure 3. Amino acids sequence of clone 57 isolated from the third library. CDR 1, 2, 3 regions are highlighted in red.

3.4 Expression, purification and binding analysis of anti-*Pf*Trx nanobody

As the three clones shared an identical sequence, clone 57 was selected for subsequent expression and purification analysis.

In order to enhance the expression yield of clone 57, it was transferred into a pET vector exploiting again the versatility offered by Gateway technology. Briefly, the gene sequence of clone 57, including the *att*B sequences, was amplified by PCR using specific oligonucleotides (see **Table 2**). The resulting amplicon was cloned into the entry vector pBluescript KS-DONR derived, previously modified (unpublished), via the BP reaction. The LR reaction was then used to transfer the insert from this intermediate clone into the pET28 vector, already modified into a destination vector (pET28 DEST-unpublished). After confirming its sequence through sequencing analysis, a resultant recombinant clone was introduced into the *E. coli* Bl21 (DE3) codon plus strain for subsequent expression. Upon inducing the bacteria with IPTG and lysing the cells, the protein was isolated from the soluble fraction using affinity chromatography, resulting in the production of 2 mg of pure protein per liter of bacterial culture.

To confirm the ability of the purified nanobody 57 to bind the target, a peptide ELISA was performed using the casein-glutathione/GST-*Pf*Trx coating system and analysing different dilutions of the nanobody. As shown in **Figure 4**, the nanobody 57 is able to detect the target, although the resulting titration curve did not allow the accurate assessment of the corresponding Kd.

Further analysis such as surface plasmon resonance (SPR) or Isothermal Titration Calorimetry (ITC) will be required to accurately estimate the binding affinity of this nanobody against the target.

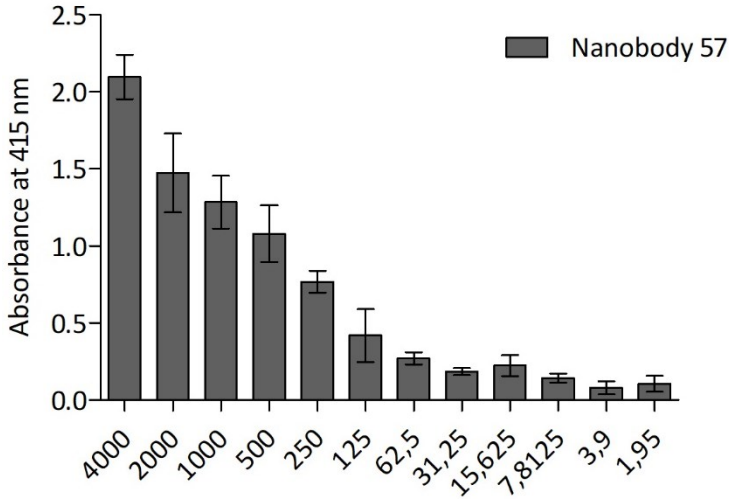


Figure 4. Peptide ELISA testing of the nanobody 57 against *Pf*Trx protein. Bars represent absorbance measured at 415 nm; abscissa indicates the serial-tested dilution of nanobody 57 (nM). The absorbance values are represented as the mean ± SD of triplicate measurements.

4 Discussion

A characteristic aspect of all types of immune libraries, whether they consist of natural ligands, e.g. animal-derived antibodies, or antibody fragments, is complexity. It is crucial to obtain sufficiently complex libraries to ensure the isolation of high-affinity binders.
The versatility of Gateway cloning technology allows for more efficient library generation than classical restriction and ligation methods[19].
Regarding Gateway technology, the conventional approach of transforming phage display vectors, including phagemid or M13 phage, with the classic DEST cassette, as described in the reference protocol[8], does not allow direct selection of recombinants within the bacterial cells used for phage display. The bacterial cells used in this selection method are resistant to the CcdB toxin at the levels produced by non-recombinant clones because the expression of the native CcdA antitoxin blocks its toxicity.
To overcome this limitation, we modified the classical cassette to allow induction of toxin expression at toxic levels even in normally resistant cells. In the new phagemid vector, instead of exploit the endogenous constitutive promoter of the ccdB gene present in the DEST cassette, we used the inducible lac promoter already present in the vector. In non-recombinant phagemid vectors, the addition of IPTG therefore results in significant expression of the CcdB toxin, whereas recombinant vectors do not; instead, they express the ligand fused to the pIII protein of the phage. Amplification of the non-recombinant vector for use in the Gateway reaction can be performed by adding glucose to the culture medium to repress toxin expression. Such amplification can also be achieved in sensitive cells such as *E. coli* Top10 and BL21 strains.
By comparing the number of colonies in toxin-insensitive strains carrying phagemid pIT2-ccdB under conditions of repression and induction of toxin expression, we

observed a counter-selection efficiency comparable to that of the original system (DEST) in CcdB toxin-sensitive cells. We also observed efficient counter-selection of recombinants when the vector was used to construct a nanobody immune-library. Sequencing analysis of 96 clones obtained after transfer of the library into the new pIT2-ccdB vector (LR reaction) revealed the absence of non-recombinants and confirmed the correct phase-insertion of coding sequences from the donor immune-library. Furthermore, the isolation of a target-specific Nb from the immuno-library, following the phage selection procedure, showed that the new vector retained all the properties necessary for its use of this selection method.

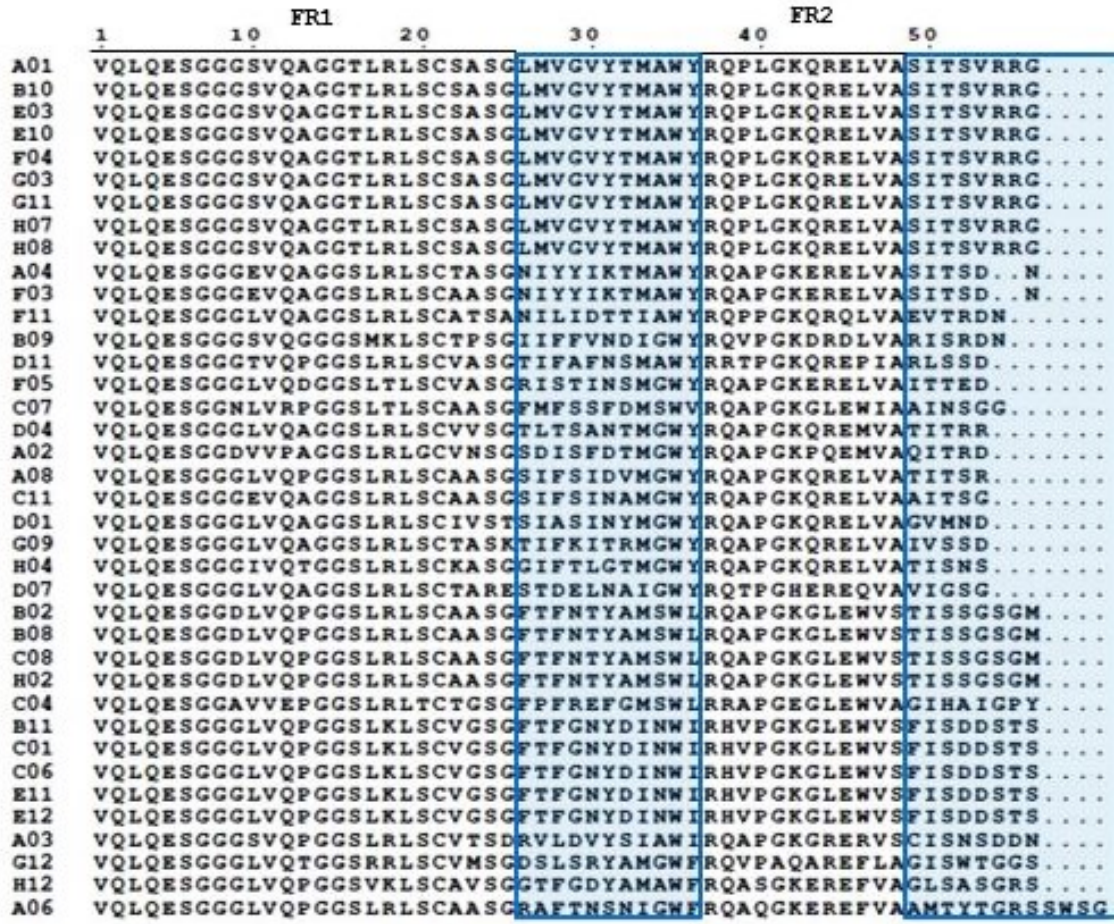
As an alternative to modifying the destination cassette, one could have focused on modifying the bacterial strains used for phage display. These strains must retain the F' episome to allow infection by the helper phage but should be devoid of the gene encoding the CcdA antitoxin, which is also located on the F plasmid. However, using homologous recombination procedures to knock out the ccdA gene in the bacterium is more complex. In addition, it cannot be excluded that such a gene modification could potentially alter the bacterial transformability and/or infectivity by the phage, especially in strains that are normally efficient in the phage display procedure.

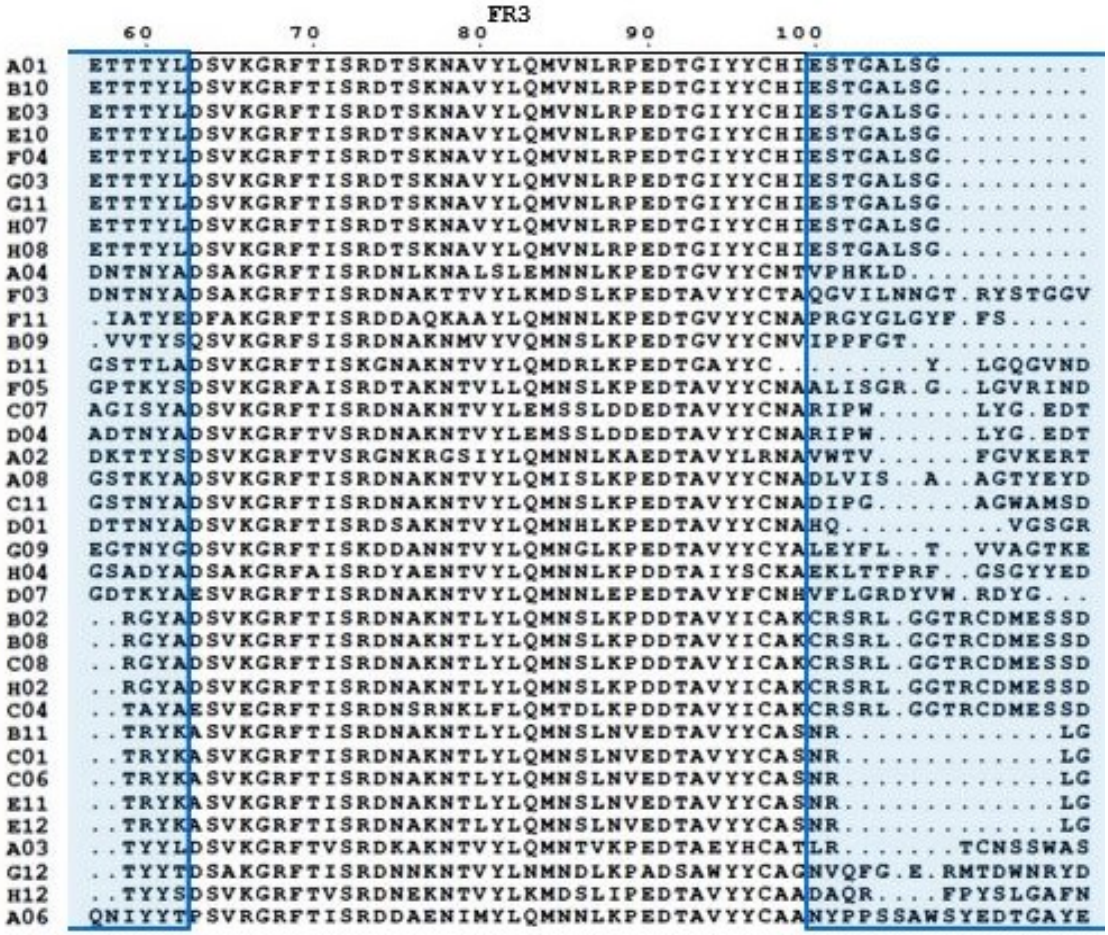
The ability to utilize Gateway cloning technology in the production of phage display libraries offers advantages in both generating highly complex synthetic libraries, which can be challenging for conventional cloning systems, and streamlining the preparation of immune libraries from immunized animals. Even with small amounts of starting material (e.g. cDNA from RNA extracted from the animal's B lymphocytes) and a minimum number of cloning reactions and cellular transformations, sufficiently complex libraries can be obtained (10^6 or more is generally recommended). Furthermore, the presence of an intermediate library in the donor plasmid facilitates the transfer of ligand encoding sequences into other plasmids suitable for alternative selection systems such as bacterial display, yeast display and ribosome display.

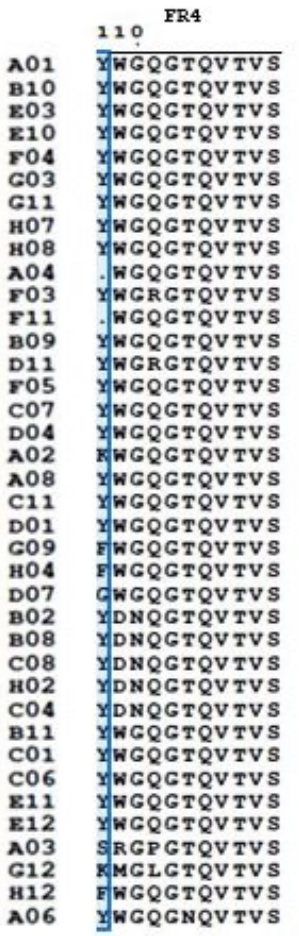
In conclusion, we believe that the novel target cassette used in this study to modify a phagemid vector can be applied to other target vectors where the classical DEST cassette is not readily applicable, such as phage M13 type 3 or 33 vectors.

Appendix 2

Alignment of the amino acid sequences of nanobodies encoded by selected clones from the immune Nb library.
Framework regions are depicted in black, while the CDR regions (CDR1, CDR2, CDR3) are highlighted in blue.




FR3
60
70
80
90
100
A01 ETTTYLDSVKGRFTISRDTSKNAVYLQMVNLRPEDTGIYYCHIESTGALSG.....
B10 ETTTYLDSVKGRFTISRDTSKNAVYLQMVNLRPEDTGIYYCHIESTGALSG.....
E03 ETTTYLDSVKGRFTISRDTSKNAVYLQMVNLRPEDTGIYYCHIESTGALSG.....
E10 ETTTYLDSVKGRFTISRDTSKNAVYLQMVNLRPEDTGIYYCHIESTGALSG.....
F04 ETTTYLDSVKGRFTISRDTSKNAVYLQMVNLRPEDTGIYYCHIESTGALSG.....
G03 ETTTYLDSVKGRFTISRDTSKNAVYLQMVNLRPEDTGIYYCHIESTGALSG.....
G11 ETTTYLDSVKGRFTISRDTSKNAVYLQMVNLRPEDTGIYYCHIESTGALSG.....
H07 ETTTYLDSVKGRFTISRDTSKNAVYLQMVNLRPEDTGIYYCHIESTGALSG.....
H08 ETTTYLDSVKGRFTISRDTSKNAVYLQMVNLRPEDTGIYYCHIESTGALSG.....
A04 DNTNYADSAKGRFTISRDNLKNALSLEMNNLKPEDTGVYYCNTVPHKLD.....
F03 DNTNYADSAKGRFTISRDNAKTTVYLKMDSLKPEDTAVYYCTAQGVILNNGT.RYSTGGV
F11 .IATYEDFAKGRFTISRDDAQKAAYLQMNNLKPEDTGVYYCNAPRGYGLGYF.FS.....
B09 .VVTYSQSVKGRFSISRDNAKNMVYVQMNSLKPEDTGVYYCNVIPPFGT.....
D11 GSTTLADSVKGRFTISKGNAKNTVYLQMDRLKPEDTGAYYC.....Y..LGQGVND
F05 GPTKYSDSVKGRFAISRDTAKNTVLLQMNSLKPEDTAVYYCNAALISGR.G..LGVRIND
C07 AGISYADSVKGRFTISRDNAKNTVYLEMSSLDDEDTAVYYCNARIPW.....LYG.EDT
D04 ADTNYADSVKGRFTVSRDNAKNTVYLEMSSLDDEDTAVYYCNARIPW.....LYG.EDT
A02 DKTTYSDSVKGRFTVSRGNKRGSIYLQMNNLKAEDTAVYLRNAVWTV.....FGVKERT
A08 GSTKYADSVKGRFTISRDNAKNTVYLQMISLKPEDTAVYYCNADLVIS..A..AGTYEYD
C11 GSTNYADSVKGRFTISRDNAKNTVYLQMNSLKPEDTAVYYCNADIPG.....AGWAMSD
D01 DTTNYADSVKGRFTISRDSAKNTVYLQMNHLKPEDTAVYYCNAHQ.....VGSGR
G09 EGTNYGDSVKGRFTISKDDANNTVYLQMNGLKPEDTAVYYCYALEYFL..T..VVAGTKE
H04 GSADYADSAKGRFAISRDYAENTVYLQMNNLKPDDTAIYSCKAEKLTTPRF..GSGYYED
D07 GDTKYAESVRGRFTISRDNAKNTVYLQMNNLEPEDTAVYFCNHVFLGRDYVW.RDYG...
B02 ..RGYADSVKGRFTISRDNAKNTLYLQMNSLKPDDTAVYICAKCRSRL.GGTRCDMESSD
B08 ..RGYADSVKGRFTISRDNAKNTLYLQMNSLKPDDTAVYICAKCRSRL.GGTRCDMESSD
C08 ..RGYADSVKGRFTISRDNAKNTLYLQMNSLKPDDTAVYICAKCRSRL.GGTRCDMESSD
H02 ..RGYADSVKGRFTISRDNAKNTLYLQMNSLKPDDTAVYICAKCRSRL.GGTRCDMESSD
C04 ..TAYAESVEGRFTISRDNSRNKLFLQMTDLKPDDTAVYICAKCRSRL.GGTRCDMESSD
B11 ..TRYKASVKGRFTISRDNAKNTLYLQMNSLNVEDTAVYYCASNR.....LG
C01 ..TRYKASVKGRFTISRDNAKNTLYLQMNSLNVEDTAVYYCASNR.....LG
C06 ..TRYKASVKGRFTISRDNAKNTLYLQMNSLNVEDTAVYYCASNR.....LG
E11 ..TRYKASVKGRFTISRDNAKNTLYLQMNSLNVEDTAVYYCASNR.....LG
E12 ..TRYKASVKGRFTISRDNAKNTLYLQMNSLNVEDTAVYYCASNR.....LG
A03 ..TYYLDSVKGRFTVSRDKAKNTVYLQMNTVKPEDTAEYHCATLR.....TCNSSWAS
G12 ..TYYTDSAKGRFTISRDNNKNTVYLNMNDLKPADSAWYYCAGNVQFG.E.RMTDWNRYD
H12 ..TYYSDSVKGRFTVSRDNEKNTVYLKMDSLIPEDTAVYYCAADAQR....FPYSLGAFN
A06 QNIYYTPSVRGRFTISRDDAENIMYLQMNNLKPEDTAVYYCAANYPPSSAWSYEDTGAYE


FR4
110
A01 YWGQGTQVTVS
B10 YWGQGTQVTVS
E03 YWGQGTQVTVS
E10 YWGQGTQVTVS
F04 YWGQGTQVTVS
G03 YWGQGTQVTVS
G11 YWGQGTQVTVS
H07 YWGQGTQVTVS
H08 YWGQGTQVTVS
A04 .WGQGTQVTVS
F03 YWGRGTQVTVS
F11 .WGQGTQVTVS
B09 YWGQGTQVTVS
D11 YWGRGTQVTVS
F05 YWGQGTQVTVS
C07 YWGQGTQVTVS
D04 YWGQGTQVTVS
A02 FWGQGTQVTVS
A08 YWGQGTQVTVS
C11 YWGQGTQVTVS
D01 YWGQGTQVTVS
G09 FWGQGTQVTVS
H04 FWGQGTQVTVS
D07 GWGQGTQVTVS
B02 YDNQGTQVTVS
B08 YDNQGTQVTVS
C08 YDNQGTQVTVS
H02 YDNQGTQVTVS
C04 YDNQGTQVTVS
B11 YWGQGTQVTVS
C01 YWGQGTQVTVS
C06 YWGQGTQVTVS
E11 YWGQGTQVTVS
E12 YWGQGTQVTVS
A03 SRGPGTQVTVS
G12 KMGLGTQVTVS
H12 FWGQGTQVTVS
A06 YWGQGNQVTVS

5 Bibliography

1. Smith GP. Filamentous Fusion Phage: Novel Expression Vectors That Display Cloned Antigens on the Virion Surface. *Science (1979)*. 1985;228(4705):1315-1317. doi:10.1126/SCIENCE.4001944
2. McCafferty J, Griffiths AD, Winter G, Chiswell DJ. Phage antibodies: filamentous phage displaying antibody variable domains. *Nature 1990 348:6301*. 1990;348(6301):552-554. doi:10.1038/348552a0
3. Zhang Y. Evolution of phage display libraries for therapeutic antibody discovery. *MAbs*. 2023;15(1). doi:10.1080/19420862.2023.2213793
4. Kehoe JW, Kay BK. Filamentous phage display in the new millennium. *Chem Rev*. 2005;105(11):4056-4072. doi:10.1021/CR000261R
5. Qi H, Lu H, Qiu HJ, Petrenko V, Liu A. Phagemid Vectors for Phage Display: Properties, Characteristics and Construction. *J Mol Biol*. 2012;417(3):129-143. doi:10.1016/J.JMB.2012.01.038
6. Jaroszewicz W, Morcinek-Orłowska J, Pierzynowska K, Gaffke L, Węgrzyn G. Phage display and other peptide display technologies. *FEMS Microbiol Rev*. 2022;46(2). doi:10.1093/femsre/fuab052
7. Hartley JL, Temple GF, Brasch MA. DNA Cloning Using In Vitro Site-Specific Recombination. *Genome Res*. 2000;10(11):1788-1795. doi:10.1101/GR.143000
8. Life Technologies Carlsbad CA. Gateway ® Technology with Clonase ® II A universal technology to clone DNA sequences for functional analysis and expression in multiple systems. *User guide*.
9. Bernard P, Couturier M. Cell killing by the F plasmid CcdB protein involves poisoning of DNA-topoisomerase II complexes. *J Mol Biol*. 1992;226(3):735-745. doi:10.1016/0022-2836(92)90629-X
10. Afif H, Allali N, Couturier M, Van Melderen L. The ratio between CcdA and CcdB modulates the transcriptional repression of the ccd poison–antidote system. *Mol Microbiol*. 2001;41(1):73-82. doi:10.1046/J.1365-2958.2001.02492.X
11. Lehtonen SI, Taskinen B, Ojala E, et al. Efficient preparation of shuffled DNA libraries through recombination (Gateway) cloning. *Protein Engineering, Design and Selection*. 2015;28(1):23-28. doi:10.1093/protein/gzu050
12. Sambrook J, W Russell D. Molecular Cloning: A Laboratory Manual. *Cold Spring Harb Lab Press Cold Spring Harb NY*. Published online 2001.
13. Takara Bio USA Inc. In-Fusion® HD Cloning Kit. *User manual*. 1(102518):1-15.
14. Xia Y, Li K, Li J, Wang T, Gu L, Xun L. T5 exonuclease-dependent assembly offers a low-cost method for efficient cloning and site-directed mutagenesis. *Nucleic Acids Res*. 2019;47(3). doi:10.1093/nar/gky1169
15. Canali E, Bolchi A, Spagnoli G, et al. A high-performance thioredoxin-based scaffold for peptide immunogen construction: proof-of-concept testing with a human papillomavirus epitope. *Scientific Reports 2014 4:1*. 2014;4(1):1-11. doi:10.1038/srep04729
16. Pardon E, Laeremans T, Triest S, et al. A general protocol for the generation of Nanobodies for structural biology. *Nat Protoc*. 2014;9(3):674-693. doi:10.1038/nprot.2014.039

17. Sehr P, Zumbach K, Pawlita M. A generic capture ELISA for recombinant proteins fused to glutathione S-transferase: validation for HPV serology. *J Immunol Methods*. 2001;253(1-2):153-162. doi:10.1016/S0022-1759(01)00376-3
18. Kim H, Ho M. Isolation of Antibodies to Heparan Sulfate on Glypicans by Phage Display. *Curr Protoc Protein Sci*. 2018;94(1):e66. doi:10.1002/CPPS.66
19. Katzen F. Gateway® recombinational cloning: a biological operating system. *Expert Opin Drug Discov*. 2007;2(4):571-589. doi:10.1517/17460441.2.4.571

Chapter 3

Chapter 3: Engineering of autotransporter proteins to enhance their application in bacterial display system

1 Introduction

The isolation of antibodies or antibody fragments with optimal performance in their intended applications, whether for scientific, diagnostic or therapeutic, can be achieved by constructing large libraries and choosing an appropriate selection system.

Although phage display has become the most robust and widely used method for library selection and has revolutionised the field of new antibody development, alternative display systems have been developed to ensure the generation of affinity ligands by screening large libraries.

A viable alternative is the display of recombinant proteins and peptides on the bacterial surface, a system that exploits the natural surface proteins of bacteria. The basic principle involves exposing the protein of interest (termed the passenger) by genetically fusing it to anchoring motifs of intrinsic cell surface proteins (referred to as carrier proteins)[1].

Depending on their compatibility with the protein of interest, different bacteria are used as hosts in this system. Although some Gram-positive bacteria have been employed[2,3], *Escherichia coli* is the most widely used Gram-negative bacteria. Its rapid growth rate, ease of genetic manipulation and high transformation efficiency are advantages that make *E. coli* suitable for the construction of large libraries (> 10^{11})[4]. The surface display system in *E. coli* poses challenges given its Gram-negative characteristics, where the cell envelope comprises an inner membrane (IM) and an outer membrane (OM) separated by the periplasm. Carrier proteins need to offer strong anchoring to the OM to guarantee the appropriate exposure of the passenger[5]. In addition, the characteristics of the passenger, such as its size, folding, and the potential formation of disulphide bridges, can influence.

Several carrier protein platforms, including outer membrane proteins (OMPs), lipoproteins, ice nucleation proteins, and autotransporter proteins, have been developed to address this challenge [5].

The first example of display of protein on the *E. coli* cell surface combined with flow-cytometric sorting was reported by Francisco et al in 1993. A scFv fragment was fused to the C-terminus of Lpp-OmpA, a chimeric protein derived from *E. coli* lipoprotein and outer membrane protein A[6]. However, a limitation of the OMP-based display system is that it can tolerate only small proteins without losing stability.

This limitation can be overcome by employing autotransporter (AT) proteins, constituting a large family secreted by Gram-negative bacteria, and encompassing numerous virulence factors. They are classified as type Va due to the type V secretion pathway, which mediates the transition from the inner to the outer membrane[7]. AT proteins comprise, within a single polypeptide, all the information required to traverse

the two membranes, with three discernible functional domains: an N-terminus signal sequence guiding the exportation of the protein across IM; a passenger domain that is secreted; a β-domain (referred to as translocator domain) anchored to the OM, facilitating the secretion of the passenger domain through the OM[8]. Although the mechanism of secretion remains unclear, ATs in the unfolded state are translocated to the periplasm thanks to the signal sequence driving the Sec-dependent translocation system. Subsequently, the insertion of the β-domain in the outer membrane as well as the secretion of the passenger domain are mediated by the β-barrel assembly machine (BAM)[9]. Given that the passenger domain seems dispensable for the secretion and translocation process, substituting it with a protein of interest, potentially a large one, renders ATs appealing for surface recombinant protein exposure.

This study presents the modification of three autotransporters already used for the exposure of foreign proteins: (i) the adhesin autotransporter EhaA[10]; (ii) the Intimin[11], which, despite its similarity to ATs, has an opposite topological organisation in the OM, with the passenger domain located at the C-terminus[12,13]; and (iii) the antigen 43 (Ag43) autotransporter, which has the characteristic, compared to the other ATs, of having high copy number on the cell surface[14,15].

As a proof of concept, to test the display functionality of the three ATs, the passenger domains were replaced by the SpyCatcher protein. This protein is part of the SpyTag-SpyCatcher system[16]. SpyTag and SpyCatcher are two polypeptides belonging to the same collagen adhesion domain (CnaB2) of the fibronectin-binding protein (FbaB) of *Streptococcus pyogenes* (Spy). The CnaB2 domain spontaneously forms an isopeptide bond between Lys31 and Asp117 during protein assembly and so do its two polypeptides SpyTag and SpyCatcher, even though they are physically separated. Furthermore, the cloning of SpyTag and SpyCatcher in frames at the N/C termination of two fusion partners leads to their rapid directional ligation in vitro[16].

The idea was to expose one of the polypeptides (namely SpyCatcher) on the bacterial surface by fusion with ATs and to evaluate its correct and efficient exposure by reaction with a reporter protein fused to the other polypeptide (namely SpyTag), which was introduced externally to the modified bacteria.

Since the ultimate goal is to construct a nanobody library suitable for bacterial display system, the expression vector associated with the best-performing autotransporter among the three tested was subsequently modified into a destination vector compatible with the Gateway cloning[17]. The vector's capacity to present an anti-GFP nanobody on the surface of E. coli was then assessed.

2 Materials and Methods

2.1 Bacterial strain, plasmid, growth and induction conditions

The *E. coli* strain used in this chapter was BL21 star (DE3) (Stratagene; La Jolla, CA, USA). All modifications of the pHEA (a gift from Luis Ángel Fernández (Addgene plasmid # 168297; http://n2t.net/addgene:168297; RRID: Addgene_168297)), pNeae2 (a gift from Luis Ángel Fernández (Addgene plasmid # 168300; http://n2t.net/addgene:168300; RRID: Addgene_168300)), pAg43 plasmids by cloning the SpyCatcher002 encoding sequence and/or the ccdB cassette were performed by Genscript (New Jersey, United States).
Plasmid-carrying bacteria were grown at 30°C in LB liquid medium containing of chloramphenicol (25 µg/ml) and 2% glucose. For induction, the same medium (containing LB and chloramphenicol 25 µg/ml) without glucose was inoculated with a 1:100 dilution of the overnight culture. After reaching an OD_{600} of 0.5, culture was induced with 0.5 mM IPTG and incubated for 3 hours at 30°C.

2.2 Western blot and SpyCatcher-SpyTag reactions

Cell protein extracts were prepared by harvesting 2 ml of bacteria after induction, resuspended in 100 µl of 10 mM Tris HCl pH 8.0, mixed with sample buffer 4X and boiled for 15 minutes. After, the samples were sonicated for 5 seconds (Sonicator 3000, Misonix), centrifugated for 5 minutes 14000 g and loaded onto 10% SDS-PAGE gel. The gel was transferred onto a PVDF membrane (Bio-Rad), and then the membrane was blocked with TBS (20 mM Tris HCl, 150 mM NaCl, pH 7.4) + 5% BSA for 2 hours at room temperature. Monoclonal mouse anti-myc antibody (Invitrogen, Waltham, MA, USA) was added diluted 1:5000 in TBS+ 5% BSA and incubated overnight at 4°C. The next day goat anti-mouse IgG antibody conjugated with StarBright Blue 700 (Bio-Rad,) diluted 1:15000 in TBS + 5% BSA were added. After 1 hour of incubation, the membrane was analysed with ChemiDoc Imaging System (BioRad).
For SpyCatcher-SpyTag reactions, 2 ml of bacterial cells after induction were harvested and resuspended in 100 µl of 10 mM Tris HCl pH 8.0. An excess of purified GFP-SpyTag002 protein was added to the cells and the reaction was incubated for 20 minutes at room temperature. After the reaction, the fusion was analysed by western blot as described above.

2.3 Confocal microscopy analysis (in collaboration with Prof. Massimiliano Bianchi, University of Parma)

Induced bacteria (1 ml) were harvested by centrifugation (5000 rpm, 5 minutes) and resuspended in 100 µl of LB medium. An excess of purified GFP-SpyTag002 protein was added to the cells and the reaction was incubated for 20 minutes at room temperature. After that, the cells that reacted with the protein were collected by centrifugation, followed by a washing step to remove excess GFP-SpyTag002 protein, and then the pellet was resuspended in 100 µl of 10 mM Tris HCl, 150 mM NaCl pH 8.0 and 5 µl were

used for microscopy analysis with STELLARIS 5 confocal microscope (Leica Microsystems).

3 Results

3.1 Comparison of EhaA, Intimin and Ag43 β-domains for display of SpyCatcher002 on *E. coli* surface

To test the ability of the three autotransporters (namely EhaA, Intimin and Ag43 β-domains) to display foreign proteins on *E. coli* OM, a gene fusion was designed with the Spycatcher002 protein as it easily detectable by SpyTag/SpyCatcher technology. The passenger domain of each autotransporter was replaced by SpyCatcher002 protein (**Figure 1**). The modifications were designed as follows. The coding sequence (CDS) for SpyCatcher002 was cloned into the expression vector pHEA[10] coding for EhaA such that it was in frame with the signal sequence in the N-terminal region and the C-terminal fragment (β-domain) of EhaA. Similarly, in the pNeae2[11] vector expressing Intimin, the CDS for SpyCatcher002 was cloned in frame with the N-terminus fragment of Intimin comprising the signal sequence, the β domain and a passenger fragment. Finally, in the expression vector pAG43 coding for the AT Ag43, the CDS for SpyCatcher002 was inserted in frame with the signal sequence at the N-terminus and with the β-domain at its C-terminus.

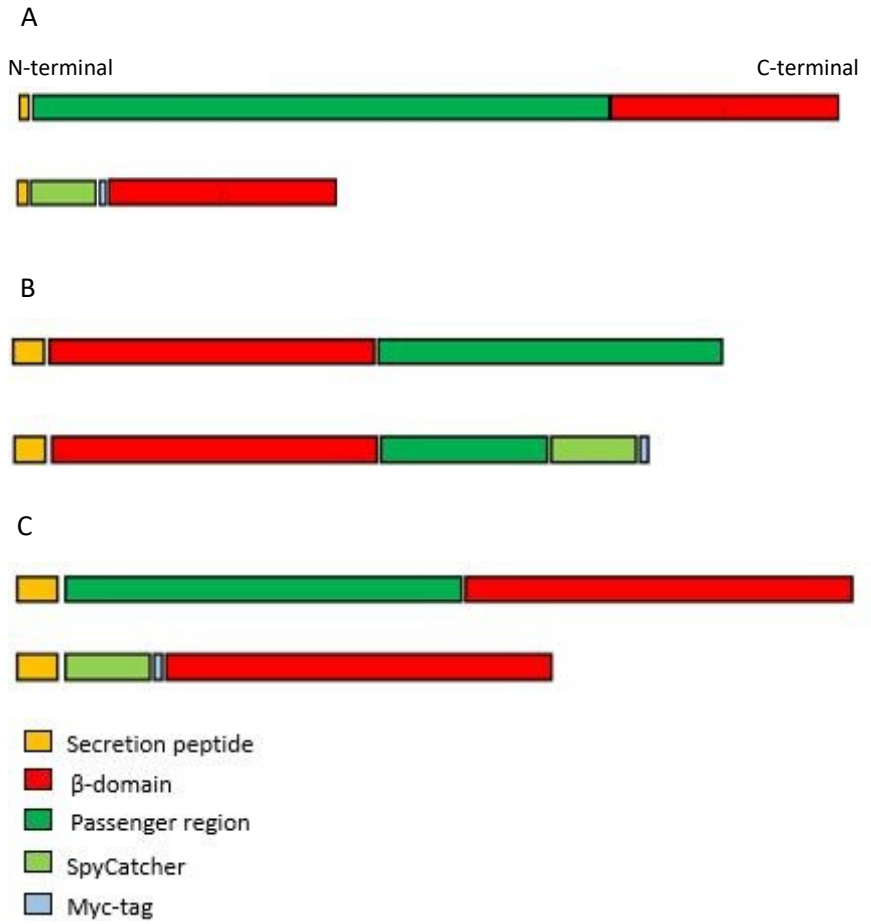


Figure 1. Schematic representation of the modifications designed on the three autotransporters. A: EhaA autotransporter; B: Intimin; C: Ag43 autotransporter. In each panel, the schematic structure of the original autotransporter is shown in the upper figure and the corresponding modification in the lower figure.

The expression of the three modified constructs in *E. coli* Bl21 star cells, strain deficient in outer membrane protease OmpT, was analysed by western blot after induction with 0.5 mM IPTG at 30°C for 3h. In addition, to verify that the cells expressed the SpyCatcher002 protein on the surface, western blot analysis was also performed after addition of the purified GFP-SpyTag002 fusion reporter protein externally (**Figure 2**). The monoclonal anti-Myc tag antibody (Invitrogen, Waltham, MA, USA) was used as a probe for western blot analysis, having inserted a Myc Tag into all protein constructs in addition to the SpyCatcher002 protein.

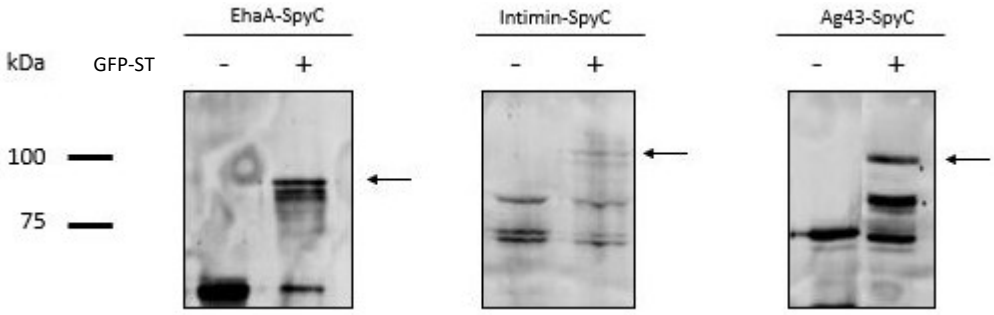


Figure 2. Western blot of whole-cell protein extracts from *E. coli* Bl21 star. *E. coli* cells were incubated with (+) or without (-) GFP-SpyTag protein. Arrows indicate the expected position of the reaction products of the three autotransporter-Spycatcher002 with GFP-SpyTag002. The molecular weight (in kDa) is shown on the left.

As shown in **Figure 2**, the three modified vectors allowed the expression on the surface of *E. coli* of the autotransporters fused with the SpyCather002 protein, as evidenced by the fact that the addition of the GFP-SpyTag002 protein to the outside of the intact bacteria resulted in the formation of a product of the SpyTag-SpyCatcher reaction corresponding to a molecular weight of the expected size.

Analysis by confocal microscopy (in collaboration with Prof. Massimiliano Bianchi, University of Parma) confirmed cell surface exposure of AT for all three modified cell types.

Indeed, the addition of the GFP-SpyTag002 fusion protein to the intact cells expressing the autotransporter-Spycatcher002 proteins revealed the presence of a fluorescent halo decorating the cell surface (**Figure 3**).

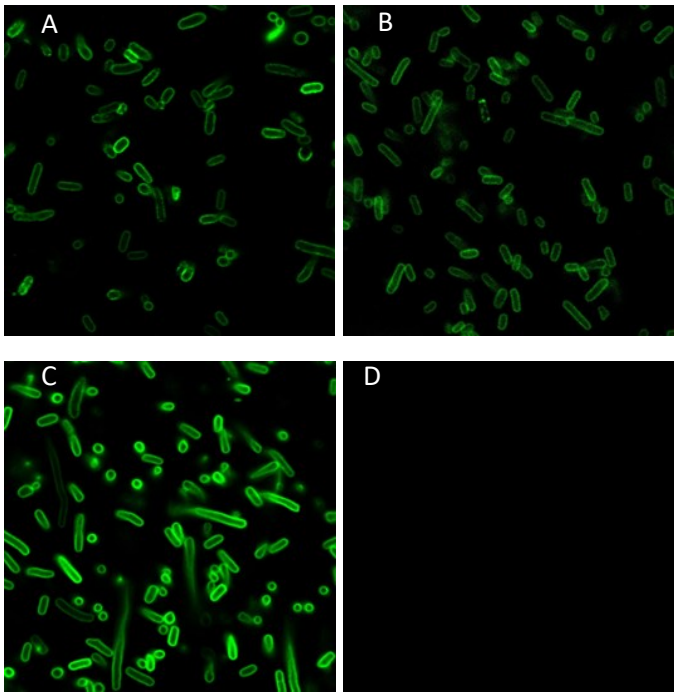


Figure 3. Confocal microscopy images of autotransporters-SpyCatcher displayed on *E. coli* cells after the reaction with GFP-SpyTag protein. A. EhaA-SpyC+GFPSpyTag; B. Intimin-SpyC+GFPSpyTag; C. Ag43-SpyC+GFPSpyTag; D. Negative control: non-recombinant Bl21 star cells incubated with GFPSpyTag protein.

However, as can be seen from the confocal microscopy images, the bacteria expressing the Ag43-SpyCatcher002 autotransporter reserve a greater fluorescent signal on the surface (**Figure 3C**), probably due to the fact that this AT is one of the most abundant proteins on the outer membrane of *E. coli*[14].
On the basis of these results, the Ag43 autotransporter was chosen for further testing in order to establish the possibility of its application in bacterial display technology.

3.2 Modification of pAg43 vector into a destination format ad its validation in Gateway cloning procedure

Based on the promising outcomes observed with Ag43-SpyCatcher, a modification of the corresponding pAg43 vector was designed for its application in the construction of nanobody bacterial-display libraries by means of the efficient Gateway cloning technology.
This modification was accomplished by inserting the new Gateway ccdB cassette (refer to Chapter 2 for details) into the pAg43 vector between the SfiI and BglII restriction sites. The resulting modified vector was labeled as pAg43-ccdB.
The efficiency of the pAg43-ccdB vector in the recombinant selection process specific to the Gateway system was then assessed. The *E. coli* Bl21 star strain, inherently susceptible to the detrimental effects of the ccdB cassette, was transformed with the pAg43-ccdB vector and subsequently plated on a selective medium containing antibiotic with the addition of 2% glucose or 1 mM IPTG.
As expected, by counting the number of colonies grown under both repression and induction conditions revealed a significant 10,000-fold reduction in colony numbers upon CcdB toxin expression. This observation underscores the efficacy of the ccdB cassette in this bacterial display vector.
The destination vector pAg43-ccdB was employed in the second Gateway cloning reaction (LR reaction) to assess its cloning efficiency through site-specific recombination and the accuracy of inserting a CDS encoding a nanobody in-frame with the CDS for the Ag43 autotransporter.
Specifically, the LR reaction was performed between pBluescript KS-DONR carrying an anti-GFP nanobody (unpublished) and the novel pAg43-ccdB vector at a 1:1 molar ratio. The reaction was then transformed into the Bl21 star strain, and recombinant selection was conducted by adding the inducer IPTG. Sequencing analysis of some clones grown after this selection confirmed their recombinant nature within the pAG43 plasmid and the accurate fusion of the CDSs.

3.3 Analysis of the ability of the new vector to expose the anti-GFP nanobody on the surface of *E. coli*

To ensure that the anti-GFP nanobody was correctly exposed on the surface of the bacteria with the appropriate molecular conformation, its capability to bind to the corresponding target, GFP, was tested.
First, a preliminary binding assay was conducted, where some recombinant clones were induced under the same conditions as outlined in section 3.1 (for details, refer to 'Materials and Methods') and then treated by adding an excess of GFP-SpyTag protein externally to the cells. Cells expressing SpyCatcher002 protein fused with the Ag43 autotransporter served as a positive control. After a 30-minute incubation at room temperature and removal of the excess GFP-SpyTag, the fluorescence of bacterial pellet was observed using the Chemidoc Imaging System (Biorad) (**Figure 4**).

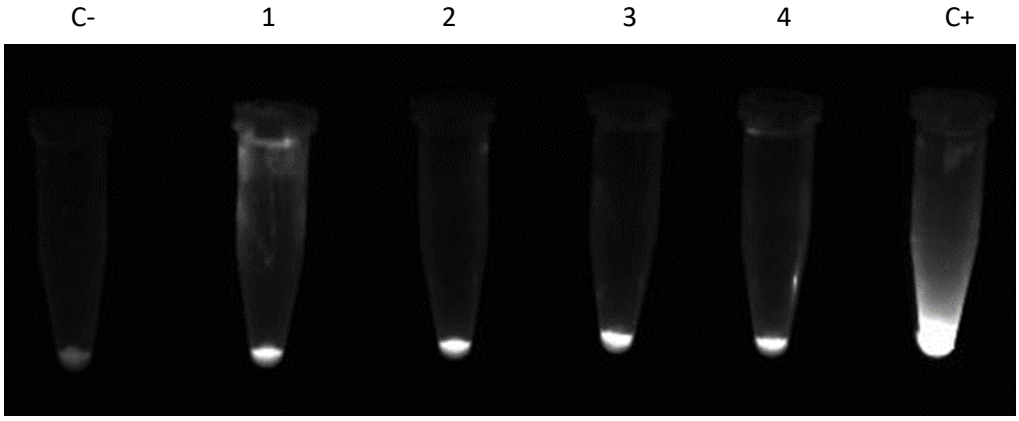


Figure 4. Reaction of induced Bl21 star cells carrying the Ag43 -anti-GFP Nb with the GFP-SpyTag protein (samples 1-4). C-: negative control, non-recombinant Bl21 star cells with GFP-SpyTag protein added; C+: positive control, reaction of Bl21 star induced harbouring Ag43-SpyC with GFP-SpyTag.

As shown in **Figure 4**, all analyzed clones expressing Ag43-antiGFP Nb on the surface demonstrated the ability to bind the target.
The capability of the Ag43 autotransporter to surface-expose the functional anti-GFP nanobody was further evaluated through confocal microscopy analysis. Following induction, cells were treated with an excess of GFP-Spytag protein, as detailed in the Materials and Methods section. As depicted in **Figure 5**, the fluorescence resulting from bound GFP is localized to the cell surface, similar to that observed with Ag43-spaycatcher-expressing bacteria.
The pAg43-ccdB vector is now ready for application in the assembly of antibody libraries for screening through bacterial display technology.

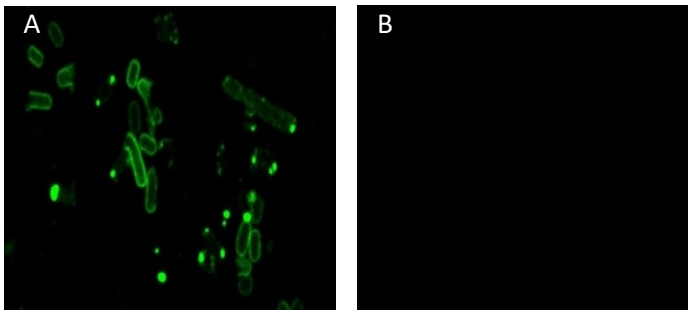


Figure 5. Confocal microscopy image of Ag43-antiGFP Nb displayed on *E. coli* cells after the reaction with GFP-SpyTag protein (A). B. Negative control: non-recombinant Bl21 star cells incubated with GFP-SpyTag protein.

4 Discussion

The isolation of nanobodies, or more generally of antibodies, with high affinity for the target, can be achieved by constructing immune, naive or synthetic libraries. Large libraries offer the possibility of isolating good binders, but the choice of selection system is important.

Over the years, in addition to the extensively employed phage display method, various selection systems have emerged. Among these alternatives, one notable approach is the direct presentation of antibodies on the surface of bacterial cells, primarily utilizing *E. coli*[4].

However, in this display system, the selection of an appropriate anchoring motif to the outer bacterial membrane is crucial. It needs to maintain the stability of the cell envelope, facilitate the translocation of the proteins to be displayed across the two membranes (IM and OM), a characteristic feature of Gram-negative bacteria, and ensure that the exposed proteins assume the correct molecular conformation.

A family of bacterial proteins that meets these criteria is the autotransporter (AT) family. The singular polypeptide constituting these proteins contains all the structural elements required for translocation across the inner membrane and self-anchoring to the outer membrane. The native passenger domain of ATs can be readily substituted genetically with a protein of interest, rendering these proteins appealing for protein engineering applications[8].

With the aim of the construction of nanobody libraries suitable for bacterial display system, we focused our attention on autotransporter proteins.

For our study, we opted for the EhaA autotransporter, the enterohaemorrhagic *E. coli* intimin (EHEC), and the antigen 43 (Ag43) autotransporter. We selected the first two due to their proven ability to exhibit nanobody libraries on the surface of *E. coli* cells[11,18]; and the third was chosen because it ranks as the most abundant outer membrane protein in *E. coli*[14].

Initially, we assessed the capacity of the three autotransporters to expose a foreign protein on the bacterial surface. The passenger domains were replaced by the

Spycatcher002 protein, easily detectable as part of the SpyTag-SpyCatcher reaction system[16].

Western blot analysis demonstrated that all three autotransporters could express the fully folded and functional SpyCatcher002 protein. However, microscopic analysis revealed that Ag43 exhibited the highest levels of the correct protein on the surface.

The Western blot analysis of the bacterial surface proteins, conducted before and after the addition of a fluorescent protein containing the SpyTag interactor, showed that all three autotransporters were capable of expressing the fully folded and functional SpyCatcher002 protein. However, fluorescence microscopy analysis indicated that Ag43 exhibited the highest levels of correctly folded protein on the surface. This difference in behaviour cannot be attributed exclusively to transcription efficiency, as all three transport systems share the same promoter and induction system. The variance is likely due to differences in their translocation mechanisms across the bacterial double membrane. Previous studies have estimated the number of passenger proteins translocated to the surface using the three autotransporters, EhaA, EHEC, and Ag43, resulting in 6000, 8000 [11], and 50,000[15] molecules per bacterial cell, respectively These values agree with those observed by microscopic analysis.

The Ag43 autotransporter was then selected for more in-depth analysis. To construct a nanobody library, an efficient cloning system is important, and Gateway cloning technology represents a valid strategy[19].Therefore, the pAg43 vector was adapted for utilization in this cloning strategy. Specifically, it was modified into a destination vector through the insertion of the new ccdB cassette (see Chapter 2 for details). The functionality of the resulting pAg43-ccdB vector was then assessed in the susceptible Bl21 star strain. As expected, a substantial reduction in colony growth was observed under toxin-induced conditions. As a proof of concept, the second cloning reaction (LR reaction) was performed by transferring an anti-GFP nanobody to evaluate the vector's suitability in the Gateway cloning of antibody molecules. Analysis of the obtained recombinants confirmed the accurate transfer of the nanobody coding sequence from the corresponding donor vector. Recombinant analysis confirmed the correct insertion of the coding sequence from the donor vector. Again, the capacity of Ag43 to surface-expose the anti-GFP nanobody by binding the GFP-SpyTag target protein, externally added to the bacteria, was established, and confocal microscopy confirmed the surface localization. Having validated the capability of Ag43 to present nanobodies on the surface of *E. coli*, the next step involves testing the vector's efficacy in creating a nanobody library and its subsequent selection, for instance, through cell sorting, to isolate a specific nanobody for a target of interest.

5 Bibliography

1. Lee SY, Choi JH, Xu Z. Microbial cell-surface display. *Trends Biotechnol*. 2003;21(1):45-52. doi:10.1016/S0167-7799(02)00006-9
2. Kronqvist N, Löfblom J, Jonsson A, Wernérus H, Ståhl S. A novel affinity protein selection system based on staphylococcal cell surface display and flow cytometry. *Protein Engineering, Design and Selection*. 2008;21(4):247-255. doi:10.1093/PROTEIN/GZM090
3. Fleetwood F, Devoogdt N, Pellis M, et al. Surface display of a single-domain antibody library on Gram-positive bacteria. *Cellular and Molecular Life Sciences*. 2013;70(6):1081-1093. doi:10.1007/S00018-012-1179-Y
4. Daugherty PS. Protein engineering with bacterial display. *Curr Opin Struct Biol*. 2007;17(4):474-480. doi: 10.1016/J.SBI.2007.07.004
5. van Bloois E, Winter RT, Kolmar H, Fraaije MW. Decorating microbes: Surface display of proteins on Escherichia coli. *Trends Biotechnol*. 2011;29(2):79-86. doi: 10.1016/j.tibtech.2010.11.003
6. Francisco JA, Campbell R, Iverson BL, Georgiou G. Production and fluorescence-activated cell sorting of Escherichia coli expressing a functional antibody fragment on the external surface. *Proceedings of the National Academy of Sciences*. 1993;90(22):10444-10448. doi:10.1073/PNAS.90.22.10444
7. Leo JC, Grin I, Linke D. Type V secretion: mechanism(s) of autotransport through the bacterial outer membrane. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 2012;367(1592):1088-1101. doi:10.1098/RSTB.2011.0208
8. Wells TJ, Tree JJ, Ulett GC, Schembri MA. Autotransporter proteins: novel targets at the bacterial cell surface. *FEMS Microbiol Lett*. 2007;274(2):163-172. doi:10.1111/J.1574-6968.2007. 00833.X
9. Van Ulsen P, Zinner KM, Jong WSP, Luirink J. On display: autotransporter secretion and application. *FEMS Microbiol Lett*. 2018;365(18). doi:10.1093/FEMSLE/FNY165
10. Marín E, Bodelón G, Fernández LÁ. Comparative analysis of the biochemical and functional properties of C-terminal domains of autotransporters. *J Bacteriol*. 2010;192(21):5588-5602. doi:10.1128/JB.00432-10
11. Salema V, Marín E, Martínez-Arteaga R, et al. 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*. 2013;8(9): e75126. doi: 10.1371/JOURNAL.PONE.0075126
12. Fairman JW, Dautin N, Wojtowicz D, et al. Crystal structures of the outer membrane domain of intimin and invasin from enterohemorrhagic E. coli and enteropathogenic Y. pseudotuberculosis. *Structure*. 2012;20(7):1233-1243. doi: 10.1016/j.str.2012.04.011
13. Luo Y, Frey EA, Pfuetzner RA, et al. Crystal structure of enteropathogenic Escherichia coli intimin–receptor complex. *Nature.* 2000;405(6790):1073-1077. doi:10.1038/35016618
14. Kjærgaard K, Hasman H, Schembri MA, Klemm P. Antigen 43-mediated autotransporter display, a versatile bacterial cell surface presentation system. *J Bacteriol*. 2002;184(15):4197-4204. doi:10.1128/JB.184.15.4197-4204.2002

15. Van Der Woude MW, Henderson IR. Regulation and Function of Ag43 (Flu). *Annu Rev Microbiol.* 2008; 62:153-169. doi: 10.1146/ANNUREV.MICRO.62.081307.162938
16. Zakeri B, Fierer JO, Celik E, et al. Peptide tag forming a rapid covalent bond to a protein, through engineering a bacterial adhesin. *PNAS*. 2012;109(12): E690-E697. doi:10.1073/PNAS.1115485109
17. Life Technologies Carlsbad CA. Gateway ® Technology with Clonase ® II A universal technology to clone DNA sequences for functional analysis and expression in multiple systems. *User guide*.
18. Salema V, Fernández LÁ. Escherichia coli surface display for the selection of nanobodies. *Microb Biotechnol*. 2017;10(6):1468-1484. doi:10.1111/1751-7915.12819
19. Hartley JL, Temple GF, Brasch MA. DNA Cloning Using In Vitro Site-Specific Recombination. *Genome Res*. 2000;10(11):1788-1795. doi:10.1101/GR.143000

Conclusions and future perspectives

Conclusions and future perspectives

Interest in antibodies, particularly their fragments, is continually increasing. Currently, recombinant antibodies are pivotal in the fields of basic research, diagnostics, and therapy. In particular, single-domain antibodies (or nanobodies) have held a significant role since their initial discovery. My doctoral research has focused on the engineering of vectors that can be used in the construction and selection of nanobody libraries.

The isolation of nanobodies with high affinity for the target necessitates the creation of libraries with substantial complexity, particularly in the case of synthetic libraries constructed entirely in vitro. To meet this requirement, it is essential to use an efficient cloning system.

Aiming to build a library of synthetic humanized nanobodies, I opted for the Gateway® technology, developed by Invitrogen, as the chosen cloning method.

I constructed a library of sequences coding for humanised nanobodies by assembling PCR. As a template, I utilized a synthetic sequence encoding a nanobody with a humanized framework (hCFW), and I employed eight oligonucleotides with partially degenerated sequences at the level of the three CDRs to introduce the characteristic hypervariability.

As a selection method, I chose the robust and most widely used phage display system.

I utilized a new phage vector, appropriately modified for Gateway cloning, to clone the synthetic library, which underwent four rounds of panning to enrich it with nanobodies against the *Pf*Trx protein, used as a target. Despite the achieved enrichment, I was unable to isolate individual clones directed against the target, likely due to the low binding affinity of the artificial nanobodies in the initial library. A potential solution to this outcome could involve conducting *in vitro* affinity maturation on the enriched library to compensate for the absence of antibody maturation that typically occurs *in vivo.* Moreover, since the new phage vector is designed for monovalent display, enabling the selection of highly affinitive ligands, the idea of using a polyvalent display vector could be considered. While this would allow the selection of lower affinity ligands, it comes with the disadvantage of favouring ligands based on their avidity.

An inherent limitation of Gateway cloning is the inefficiency of counter-selection of the destination cassette in bacterial strains utilized for phage display. The bacteria commonly employed for phage vector propagation carry the essential F plasmid for phage infection, but this plasmid also bears the genes encoding the CcdB toxin and the CcdA antitoxin. This renders the strains insensitive to the CcdB toxin.

To overcome this limitation, I modified the classical destination cassette of the phagemid pIT2 vector by PCR mutagenesis. This involved positioning the ccdB gene under the control of the Lac promoter already inherent in the phagemid. In the resulting vector, named pIT2ccdB, the expression of the CcdB toxin is regulated by the addition of a repressor (glucose) or an inducer (IPTG). Under conditions of toxin induction, I observed efficient recombinant selection, even in resistant strains like TG1, comparable to the effectiveness of the original system (DEST) in CcdB toxin-sensitive cells.

Using the new vector, I assembled an immune library of nanobodies from llama cDNA post-immunization with the recombinant *Pf*Trx protein. After three rounds of panning, I successfully identified specific nanobodies targeting the *Pf*Trx, affirming the capability of the new vector to operate within the phage display selection system. Moreover, the novel ccdB cassette, regulated by the Lac promoter, holds promise for application in vectors where the original DEST cloning cassette is impractical, such as the M13 type 3 or 33 phage vectors.

Furthermore, I assessed the efficacy of three autotransporters, EhaA, Intimin, and antigen 43, as surface proteins on the outer membrane of *E. coli* to present ligands, aiming to establish an effective bacterial display system. As a validation, I utilized these autotransporters to present the SpyCatcher protein on the bacterial surface and confirmed its correct exposure through a reaction with a reporter protein fused to the SpyTag peptide (GFP-SpyTag).

I selected the antigen-43 autotransporter as the most efficient and incorporated it into a corresponding destination vector by integrating the new ccdB cassette, ensuring compatibility with Gateway technology. I evaluated the performance of this vector in Gateway cloning and its capability to present an anti-GFP nanobody on the surface of the bacterial cells.

The subsequent phase of this research will involve assessing the effectiveness of the vector in constructing nanobody libraries for selection using bacterial display technology.

In conclusion, I expect that the vectors I have developed will prove valuable in streamlining the creation of highly complex nanobody libraries, thereby expediting and improving the selection of specific ligands.