Journal of Immunological Methods 350 (2009) 54–62

Contents lists available at ScienceDirect

Journal of Immunological Methods

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

Research paper

# A single-step procedure of recombinant library construction for the selection of efficiently produced llama VH binders directed against cancer markers

Damjana Kastelic [a,b], Snježana Frković-Grazio [c], Daniel Baty [d,1], Gilles Truan [a], Radovan Komel [b], Denis Pompon [a,*]

[a] CNRS, Centre de Génétique Moléculaire (CGM) FRE3144, Unité associée à l'Université Pierre et Marie Curie, Gif-sur-Yvette, France
[b] Medical Center for Molecular Biology, Institute of Biochemistry, Faculty of Medicine, University of Ljubljana, Slovenia
[c] Department of Pathology, Institute of Oncology, Ljubljana, Slovenia
[d] CNRS, Laboratoire d'Ingénierie des Systèmes Macromoléculaires, UPR9027, GDR2352, Marseille, France

ARTICLE INFO

*Article history:*
Received 13 March 2009
Received in revised form 28 August 2009
Accepted 31 August 2009
Available online 8 September 2009

*Keywords:*
Heavy chain antibody
Variable domain
Phage display
Llama VHH
Llama VH
Cancer biomarkers

ABSTRACT

Heavy chain antibodies are naturally occurring in *camelidae* (camels and llamas). Their variable domain (VHH) can be efficiently produced as a recombinant protein in *E. coli* with a large range of applications in the fields of diagnostics and immunotherapy. Standard cloning approach involves resolution of VHH from the heavy chain variable domain of conventional antibodies (VH) by a nested PCR amplification followed by a phage display based selection. Present work illustrates that in contrast to usual finding, specific, good affinity and efficiently expressed VH domain of conventional antibodies can be selected from the co-amplification products of VH and VHH cDNAs. Sequence analysis illustrated that following the two first rounds of selection against cancer markers, similar number of VH and VHH binders were observed. However, after a third round, the more specific binders directed against p53, VEGF, BCL-2 proteins surprisingly contain only VH specific hallmarks. Characterisation of the specificity, affinity and productivity of selected VH binders is described. Because llama VHs show higher sequence and structural homology with the human VH III group than llama VHHs (Vu et al., 1997), they constitute very interesting agents in therapeutic applications, especially in human immunotherapy and cancer treatment.



## 1. Introduction

Heavy chain antibodies (HcAb) naturally occur in *camelidae* (camels and llamas). In their serum conventional heterodimer antibodies and homodimer HcAb coexist (Hamers-Casterman et al., 1993) in relative amounts depending on species (van der Linden et al., 2000). In HcAb, the light chain and first constant domain (CH1) are absent and their variable domains, referred to as VHH (Arbabi Ghahroudi et al., 1997), constitute the smallest (about 120 amino acids) naturally occurring fragment capable of binding an antigen. VHH can be obtained from *E. coli* (Olichon and Surrey, 2007), *S. cerevisiae* (Frenken et al., 2000), *Lactobacillus* (Hultberg et al., 2007), *A. awamori* (Joosten et al., 2005), *P. pastoris* (Rahbarizadeh et al., 2006) and tobacco plant (Rajabi-Memari et al., 2006) with high yields. They are generally stable (Perez et al., 2001), soluble and bind specifically and with a high affinity to their targets (Desmyter et al., 2002; Hmila et al., 2008). They are able to recognise haptens (Frenken et al., 2000; Yau et al., 2003; Alvarez-Rueda et al., 2007) and cryptic epitopes, such as enzyme active site (Lauwereys et al., 1998), which are not accessible to conventional (larger) antibodies. VHHs have

*Abbreviations:* CDR, complementarity determining region; ELISA, enzyme-linked immunosorbent assay; FR, framework region; GST, glutathione S-transferase; IgG, immunoglobulin G; $K_D$, equilibrium dissociation constant; MPBST, milk phosphate-buffered saline-Tween 20; PCR, polymerase chain reaction; Rmax, maximum response; RT, room temperature; RU, resonance unit; SDS, sodium dodecyl sulfate; SPR, surface plasmon resonance; VH, heavy chain variable domain of the conventional antibody; VHH, variable domain of a heavy chain antibody; VL, light chain variable domain of the conventional antibody.

* Corresponding author. Centre National de la Recherche Scientifique (CNRS), Centre de Génétique Moléculaire (CGM), 91190 Gif-sur-Yvette, France. Tel.: +33 1 69 82 36 80; fax: +33 1 69 82 36 82.
*E-mail address:* pompon@cgm.cnrs-gif.fr (D. Pompon).
[1] Present address: Institut National de la Santé et de la Recherche Médicale, Laboratoire Stress Cellulaire, U624, Marseille, France.

0022-1759/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.jim.2009.08.016

been successfully used as immunotherapeutics (Harmsen et al., 2006) and to develop cancer therapeutic approaches (Cortez-Retamozo et al., 2004). They constitute promising agents in new generation of therapeutic antibodies (Behar et al., 2008). They can be used in muscular dystrophy disease treatment (Chartier et al., 2009).

VHHs are clearly distinguishable from the VHs of the conventional antibodies, by the presence of several amino acid signatures, located in the solvent-exposed surface which is normally covered by the variable domain of the light chain (VL). Using the Kabat numbering (Kabat et al., 1991), VHHs differ from VHs by the V37Y/F, G44E/Q, L45R, W47F/G/S/L substitutions on conserved VH residues belonging to the second framework (FR2) (Vu et al., 1997; Harmsen et al., 2000). The importance of these residues for the solubility was highlighted by Davies and Riechmann (1994). Expression of the VH domain from human antibodies in absence of light chain was found frequently inefficient (Worn and Pluckthun, 1998) because of aggregation, poor-solubility and stickiness. This led to the usual finding that the llama VHs are similarly not suitable for high efficiency expression. This finding was supported by Davies and Riechmann (1996) who reported that substitution of the four amino acids in human VH aligning with the hallmarks of camelid VHH allowed to improve the solubility of the expressed domain and reduce aggregation. Until now, soluble llama conventional VHs were isolated by Tanha et al (2002), but only a single structure of llama VH was resolved by NMR (PBD code-1IEH) (Vranken et al., 2002).

Based on these previous findings, standard methods for llama VHH library construction were designed to avoid VH contamination in VHH libraries. Following mRNA isolation and CH2 gene specific reverse transcription, usually two successive PCRs are performed. The first is used to discriminate VH from VHH based on amplicon sizes which differ by ~300 bp in length. Shorter amplicons (620 and 690 bp) encoding VHHs are subsequently re-amplified through a second nested PCR with primers annealing at the codons of FR1 and FR4, because all variable domains of camelid heavy chain antibodies belong to a single family (family III). Afterwards VHH fragments are cloned and displayed for selection. Such a nested PCR approach could be mutagenic and disadvantageous for molecular diversity, but is expected to eliminate VH derived contributors which could lead to sticky proteins through exposed hydrophobic amino acids on their surface lacking VL domain.

In contrast to described VHH advantages, llama VHs show a higher similarity with the human VH III family than VHH (Vu et al., 1997). Because of this better similarity, they are expected to be less immunogenic in humans and could represent a significant advantage for immunotherapy and other therapeutic applications. Here we describe a single step PCR method to construct a llama VH–VHH mixed library and an adapted selection procedure. This work illustrates on several examples (cancer protein markers p53, BCL-2 and VEGF), that VH good affinity binders can be easily selected and efficiently produced in *E. coli*. Analysis of the sequences at different stages of the selection process is presented as well as the specificity, affinity and expression of selected binders. Results are discussed on the basis of predicted structural features.

## 2. Materials and methods

### 2.1. Llama antibody library construction and characterisation

Adult female llama (*Lama glama*) was immunised by a single sub-coetaneous injection of human protein extract followed by three booster injections every 3 weeks. Protein extract was prepared from post-operative samples of six patients suffering from different types and stages of gastric cancer. Native proteins were extracted using liquid nitrogen homogenization and standard biochemical techniques (Ha et al., 2002). Finally, extracted proteins (1.5 mg) were dissolved in PBS buffer containing 10 mM EDTA, 0.5% sodium cholate and 0.1% NP-40 detergent and mixed with an equal volume of Freund's incomplete adjuvant. The blood of immunised animals was collected and lymphocytes were separated by Ficoll-Histopaque-1077 (Sigma-Aldrich) discontinuous gradient centrifugation followed by washing with PBS, and were afterwards stored at −70 °C until further use.

Total RNA was isolated from $10^7$ leucocytes by acid guanidinium thiocyanate-phenol chloroform extraction (using RNable; Eurobio, Courtaboeuf, France), verified by electrophoresis and stored at −80 °C in DEPC treated water.

Isolated total RNA was used as a template for cDNA synthesis (Super Script II™, Reverse Transcriptase, Invitrogen) using CH2 domain gene-specific primer 5′-GGTAC-GTGCTGTTGAACTGTTCC-3′, annealing at the CH2 exon of the heavy chains of all llama immunoglobulins. PCR was performed on cDNA using DyNAZYME EXT DNA Polymerase (Finnzymes) with an equimolar mixture of degenerated PAGE purified primers (Sigma Aldrich) annealing at the FR1 and FR4 of conventional (IgG1) and heavy chain antibodies (IgG2 and IgG3):

(I.) 5′-CCTATAGGCCCAGCCGGCCATGGCCGAGGTG-CAGCTGGTGSAG-3′

(II.) 5′-CCACGATTCTGCGGCCGCTGAGGAGACRGT-GACCTGGGTCC-3′

containing BglI and NotI restriction enzyme sites (underlined). PCR protocol consisted of an initial denaturation step at 94 °C for 2 min followed by 37 cycles of 94 °C for 1 min, 60 °C for 1 min and 72 °C for 100 s, and a final step extension at 72 °C for 10 min. The resulting unique ~450 bp PCR fragment was purified from 1.5% highly pure agarose gel and digested with BglI and NotI (New England Biolabs), re-purified on gel and ligated (T4 DNA Ligase, New England Biolabs) into SfiI and NotI digested pHEN2 phagemid (between the PelB leader sequence and the M13 viral gene III). Vector comprises the β LacIq (LacZ) promoter, a pelB leader sequence, C-terminal myc-tag (EQKLI-SEEDLN) and $(His)_6$ tag to facilitate purification, amber stop codon and bacterial phage M13-derived g3 gene. After electrotransformation into TG1 *E. coli* (Stratagene, USA) the cells were plated on selective plates. Library size was calculated by plating serial dilution aliquots on 2YT/ampicillin (100 μg/ml) containing 2% glucose (2YTAG) agar plates and incubated overnight at 30 °C. Colonies were scraped from the plates with liquid 2YTAG and library was stored at −80 °C in the presence of 20% glycerol. Fifty clones were sequenced (GATC company) for diversity analysis using 5′-M13 reverse and 3′-M13 forward universal primers. All *in situ* PCR reactions were performed using Taq DNA polymerase (New England Biolabs).

### 2.2. Phage production and panning

Phages were rescued as described by (Clackson and Lowman, 2007). Briefly, cells grown in 2YTAG at 37 °C were inoculated in log-phase with M13KO7 helper phages (New England Biolabs) and incubated at 37 °C for 30 min without shaking, before the medium was replaced by 2YT containing 100 μg/ml ampicillin and 25 μg/ml kanamycin. Cells were grown overnight at 28 °C under rapid shaking. The supernatant was PEG precipitated (20% polyethylene glycol 6000 in 2.5 M NaCl in water) on ice for 1 h. Phages were spun-down and pellet was resuspended in 1 ml PBS and centrifuged again for 2 min at 14000 *g*. Phages were titrated and stored at 4 °C. They were used within a week.

The library was panned against human recombinant commercially available proteins (all were expressed as GST-tagged fusion proteins): p53 (1–393), IL-8 (hBA-72), TGFα (hBA-50), VEGF (BA-h165), BCL-2 (D21) (all Tebu-Bio). For each antigen 8 wells of polystyrene 96-well plates (Maxisorp NUNC, Denmark) were coated overnight at 4 °C with 100 μl of 10 μg/ml target protein in PBS for the first round, and 2 μg/ml for subsequent panning. Wells were washed three times with 200 μl PBS and three times with PBS supplemented by 0.1% Tween (PBST), and blocked with 200 μl 2% BSA (w/v) in PBS for 2 h at RT. $10^{10}$ phages were diluted in blocking solution and incubated with the immobilised antigens for 2 h at RT. Wells were washed 20 times with PBS and 10 times with PBST. Bound phages were eluted with freshly prepared 100 mM triethylamine (100 μl/well) for 8 min at room temperature, transferred to tube and immediately neutralised with a half volume of 1 M TRIS–HCl pH 7.4. Subsequently, log-phase TG1 bacteria in 2YT medium were inoculated with neutralised phages and incubated for 30 min at 37 °C. After spinning down, the pellet was resuspended in 500 μl fresh 2YTAG medium, plated on 2YTAG agar plates and incubated overnight at 30 °C. Additionally after each panning, phages were titrated. After overnight incubation, the selection plates were scraped and bacteria were resuspended in 2YTAG containing 20% glycerol and stocked at −80 °C until further use.

#### 2.2.1. Expression of soluble antibodies and screening by ELISA

After second and third round of panning, clones were screened for enrichment in 96-well plates (Nunc 163320). 100 μl 2YTAG per well was inoculated with a randomly chosen single colony from the plates of phage-infected TG1 spread after selection. After overnight growing at 30 °C the plates were duplicated in 150 μl 2YTAG per well. Plates were grown at 37 °C until log-phase before being inoculated with M13KO7 helper phage. Plates were incubated for 30 min at 37 °C and afterwards spun down. Pellets were resuspended in 150 μl 2YT with 100 μg/ml ampicillin and 25 μg/ml kanamycin. Plates were grown overnight at 28 °C. After incubation plates were spun down and 50 μl supernatant per well was diluted with 50 μl 2% MPBST and used for phage ELISA screening. Maxisorp plates with 1 μg/ml antigen, blocked with 2% (w/v) skim milk (2% MPBS) were incubated for 2 h with diluted phage solution. After extensive washing, phages were detected using HRP/Anti-M13 monoclonal antibody conjugate (GE Healthcare) in 2% MPBST. Tetramethylbenzidine (Sigma) was used as a substrate. After second selection round, 5 clones were sequenced for each antigen.

Finally, after the third selection round, positive clones were tested for detection of soluble antibody fragments. Antibody expression was induced in 96-well plates with 1 mM IPTG. After protein production, supernatant was used as previously on maxisorp Nunc 96-well plates with immobilised antigen. Positive clones were detected using mouse Anti-6his (Sigma) and anti mouse donkey polyclonal-HRP (Tebu-Bio). After the third enrichment cycle positive clones were tested for cross reaction using non-related proteins. Clones fulfilling all selection criteria were sequenced. We attempted to determine the cDNA origin of the two isolated binders (VH anti-p53 and VH anti-BCL-2) with PCR using CDR3 specific oligonucleotide primers 5′-CTATAGTCGTATGTACTACTTTGG-3′ (for VH anti-p53) and 5′-CCTATTTGGTACTTGGAAGCGG-3′ (for VH anti-BCL-2) and CH2 specific primer.

### 2.3. Selection and purification of recombinant antibodies

Three positive clones with unique sequences were produced in larger scale (6 L). Briefly, TG1 *E. coli* was grown overnight with shaking at 37 °C. Next day a 1/100 dilution was made in 6 L 2YTAG and cells grown with shaking at 37 °C until OD = 0.9. After spinning-down, cells were resuspended in 1.5 L 2YT/ampicillin containing 1 mM IPTG. Cells were grown at 28 °C for 3 h and periplasmic proteins were extracted according to a previous protocol (Alvarez-Rueda et al., 2007). Briefly, after 1-h incubation on ice cells were spun down and resuspended in 1/10th of volume of ice-cold PBS, 1 M NaCl, 1 mM EDTA. Cells were again incubated on ice for 30 min. After spinning-down, cells were removed and the supernatant was transferred to a fresh tube and centrifuged again at 14 000 × *g* for 10 min to remove debris. Periplasmic protein fraction was precipitated in 60% saturation with ammonium sulphate, resuspended in 50 mM phosphate buffer pH 7.0 with 300 mM NaCl and dialysed overnight at 4 °C against the same buffer. Antibodies were purified on Talon Metal Affinity Resin (Clontech) and eluted with buffer containing 150 mM imidazole and dialysed again against phosphate buffer. The purity of the proteins was checked by SDS-PAGE. Concentration of VHs after purification to homogeneity was determined from the UV absorption at 280 nm, and theoretical extinction coefficient was calculated from their amino acid content. Possible aggregation of isolated VH domains was verified using size exclusion chromatography on a Superdex 75 column (internal diameter of 1 cm, length 32 cm) (Pharmacia) eluted with 50 mM phosphate buffer pH 7.0 with 300 mM NaCl at a flow rate of 0.5 mL/min. For comparison, a control VHH domain (containing four VHH specific hallmarks, CDR3 length of 8 residues, size of 12 kDa) was analysed under the same conditions. The column was calibrated using the monomeric (12 kDa) and the corresponding disulfide bond cross-linked cytochrome *b5* dimer as standards.

### 2.4. Surface plasmon resonance (SPR) analysis

Single domain antibodies were analysed for antigen binding by real-time surface plasmon resonance (SPR) using BIACORE™ 2000 BIOSENSOR and BIACORE X100. Corresponding antigens p53 and VEGF (Tebu-Bio) in 10 mM sodium acetate buffer pH 4.5 were covalently immobilized on

CM5 sensor chip (Biacore AB, Uppsala, Sweden) using EDC/NHS activation according to the manufacturer's instructions. A control flow cell surface was prepared with the same chemical procedure without antigens. All analysis were performed at 25 °C, at a flow rate of 50 µL/min and using HBS-EP (0.15 M HEPES, 150 mM NaCl, 3 mM EDTA and 0.05% Tween) as running buffer.

Injections of VH anti-p53 and VH anti-BCL-2 were performed at 7 different concentrations from 1 mg/ml to 0.0156 mg/ml, in HBS-EP during 5 min. Between each run flow cell surface was regenerated with 10 mM glycine–HCl pH 1.5 and 0.1% SDS. Sensorgrams were corrected by subtracting the signal from the reference flow cell and were analyzed using the BIA evaluation version 3.2 software (Biacore AB) and home made softwares. For VH anti-VEGF the steady-state response was rapidly reached experimentally. For VH anti-p53 binding, the SPR signal was fitted by a bi-exponential law and the initial slope and end-point steady state responses calculated from the fits. Ratios between the initial rates for VHs binding and the steady state responses were plotted against the VH concentration ($C_{VH}$) to obtain an estimate of the $k_{on} \times C_{VH} + k_{off}$. Rate constant for the major slow phase of the VH anti-p53 dissociation was evaluated by linear fitting of sensorgrams for times ranging from 1 to 10 min after the start of washing. The fast phase of signal change immediately following the washing start was thus not taken in account in the calculation.

### 2.5. Structural models

The modeling used as template the only available structure (PDB entry 1IEH) for a llama VH domain (Vranken et al., 2002). The alignment between all three selected VHs (p53, VEGF and BCL-2) and 1IEH was generated with the STRAP software (Gille and Frommel, 2001). Modeling was performed with the Modeller 9v5 software (Eswar et al., 2007). Comparative protein structure modeling was performed in two independent rounds. A first round was designed to build the core of the protein, generating 5 models with mild optimization, particularly of the CDR3 loop. A second round of optimization was entirely devoted to loop modeling. Briefly, the loop model method of Modeller was invoked on the CDR3 loop (residues 95–103) of the five previously generated models. From each starting model, 10 models were constructed and ranked *via* their calculated energies (molpdf and DOPE scores). VH anti-BCL-2 was used as a presentation model.

## 3. Results

### 3.1. Library construction and characterisation

In conventional method of recombinant heavy chain antibody library construction, separation of VH and VHH domains is employed and most often two successive PCR are used. Oligonucleotide primers used in first PCR anneal at the CH2 region and at the framework 1 of all isotypes and amplify both DNA fragments of conventional and non-conventional antibodies. Because of the presence of CH1 domain in conventional antibodies (IgG1) and different hinge size in non-conventional antibodies, three bands are observed. The 900 bp band represents conventional VH domain including the CH1 region (subtype IgG1), while two bands (620 and 690 bp) represent VHH where the CH1 region is absent and hinge length is different (subtypes IgG2 and IgG3). The smaller fragments are gel-purified and a second PCR was performed to amplify only the VHH genes without CH2 or hinge derived sequences.

Fig. 1 illustrates the comparative flow chart for the two steps (classical) and single step (this work) approaches. Following llama immunisation and reverse transcription using a CH2 gene-specific primer, a single PCR was performed using a pair of primers annealing at the FR1 (primer I) and at FR4 (primer II) of all three llama isotypes (IgG1, IgG2 and IgG3) thus leading to co-amplification of VH and VHH (Fig. 1, left part). In this single step PCR approach, the designed primer pair allows amplification of a unique and well-defined band of 400–500 bp. Using the two-step protocol, the same starting material gave rise to two poorly sized bands corresponding to VH and VHH amplification products (Fig. 1, right part). Nested PCR amplification of the smaller size band led to selective amplification of VHH sequences. Yield of the second PCR appeared to be low and poorly reproducible, suggesting a high risk of diversity lost. This was confirmed by sequencing of randomly picked clones following cloning which evidenced occurrences of repeated sequences. This prompted us to focus on amplification products resulting from the single step procedure: following digestion of the primer encoded restriction sites, ligation of the PCR product into a phagemid vector and transformation into TG1 *E. coli*, a recombinant library of $8 \times 10^5$ independent clones was easily obtained. About 90% of clones within the library contained an inserted fragment of the proper size as determined by *in situ* PCR on individual colonies.

#### 3.1.1. Sequences

Prior to any selection, 50 randomly chosen clones were sequenced and were found to belong to mammalian III family as evidenced by specific amino acid hallmarks at positions 9, 18, 67 and 82. All clones contain Gly9, all except one contain Leu18, all except two contain Phe67 and all contain Met82. Sequencing revealed that 58% of clones contain the full set (V/I37, G44, L45 and W47) of VH specific hallmarks when 42% of clones contain the full VHH signatures at the same positions (Y/F37, E/Q44, R45, L/F47). This is consistent with observations of (van der Linden et al., 2000). Other clones contain two or three of amino acids specific for each group (Table 1) at the same position. Tryptophan at position 103 (Kabat numbering) is characteristic residue of conventional antibodies, forming the hydrophobic interface with VL domain. It cannot be excluded, as previously hypothesized, that such clones could result from PCR-mediated cross-over. However, detailed examination of sequences (not shown) did not bring any evidence supporting the formation of chimerical structures in the library. Clones were classified into VH or VHH groups based on the observation of at least two specific amino acids at signature positions. In addition, all VHH clones were found to contain the highly conserved Arg at position 45 (Muyldermans et al., 1994). The length of the CDR3 was found to range between 3 and 19 residues with an average length of 12. Among VHH and VH groups the most frequent length of CDR3 was 10 and 13 residues, respectively.

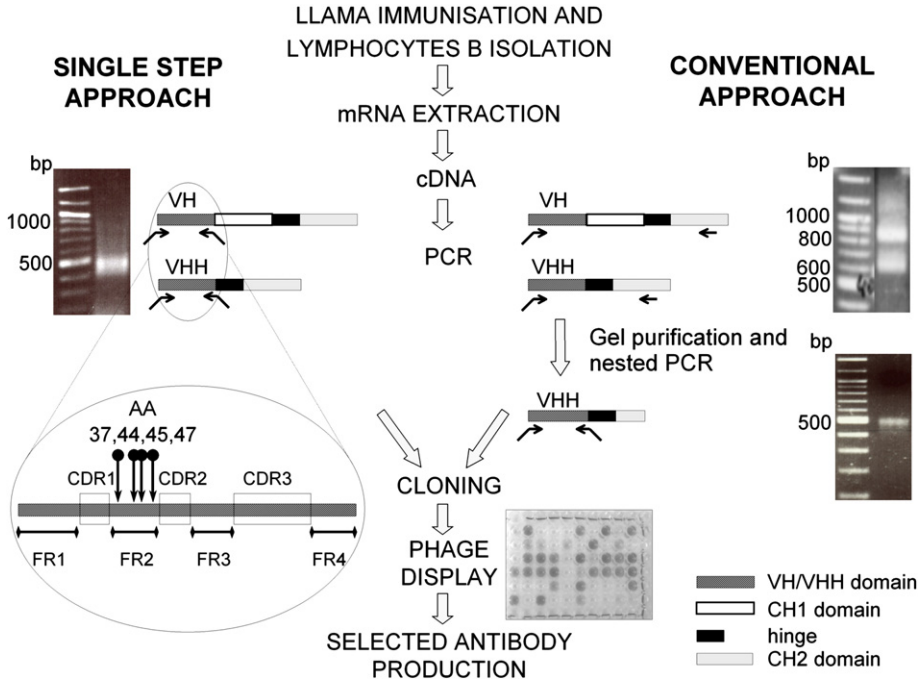


**Fig. 1.** Schematic representation of two alternate approaches for recombinant llama single domain library construction. On the right side the conventional method is presented, employing two successive PCRs in order to separate VH and VHH. Only VHH are cloned and used for phage display selection. On the left side a single PCR approach is used to amplify both VH and VHH at once. Selection is performed using directly the mixed library of single heavy chain domains and positive clones (represented with full circles in the 96-well plate) are used for antibody production. Large circle on the left represents magnification of VH and VHH corresponding amino-acid sequences. Three CDR loops are boxed. Differential amino acid hallmarks between VH and VHH are represented with arrows. Oligonucleotide primers, annealing at different parts of antibody genes, are represented with arrows under the different domains.

### *3.2. Phage selection*

The mixed VH and VHH antibody repertoire was displayed on phages following infection with a helper phage. In order to isolate specific binders for proteins: p53, VEGF, BCL-2, TGFα and IL-8, three rounds of display selection on solid-phase coated antigen were performed. After the second and third selection round, periplasmic extracts were tested by ELISA assay and positive clones were sequenced. Sequencing of 5 positive clones for each antigen, after the second selection round, revealed unique sequence for each targeted antigen. Those sequences are either of VH or VHH types with relatively short CDR3 ranging from 6 to 8 residues.

After the third enrichment round, 34 extracts (18%) out of 188 selected clones against protein p53 were shown to be specific binders. For the selection targeting protein BCL-2, 7 clones (7.5%) out of 94 were found to be specific. Similarly 13 clones (14%) out of 94 were positive protein VEGF and 28 clones (15%) out of 188 for protein TGFα. Positive clones giving clear signal with protein p53 were tested for cross reactions with blocking agents and other studied proteins. Approximately half clones (12 out of 25) exhibited specific binding with p53 and no cross reaction with other proteins. The same result was obtained for the proteins BCL-2 and VEGF. In contrast, selected binders directed against proteins IL-8 and TGFα were found to exhibit significant cross-reaction and were finally discarded.

Clones with unique sequences obtained after the third round against antigens p53, VEGF and BCL-2, were processed for expression of their encoded binder protein. For all positive clones, sequencing revealed a unique sequence for each antigen. All selected binders were found to include only specific VH hallmarks (V37, G44, L45, W47) and contained surprisingly long CDR3 of 14 and 13 amino acids for VH directed against p53 (Fig. 2). Attempts to check the presence of isolated sequences in the unamplified cDNA pool using CDR3-specific PCR primers failed to amplify any VH or VHH sequence type. This is consistent with previous observations that such strongly asymmetric PCR amplification is extremely difficult and noisy (Tanha et al., 2002).

**Table 1**
VH and VHH specific amino-acid residues at positions 37, 44, 45, 47, 103 (Kabat numbering) and percentage of each type in a recombinant library before selection.

| Type | Percentage | 37 | 44 | 45 | 47 | 103 |
|---|---|---|---|---|---|---|
| VH 58% | 40% | V/I | G | L | W | W |
| | 6% | V | G | P | W | W |
| | 2% | V | G | S | W | W |
| | 2% | V | E | P | W | W |
| | 2% | V | K | L | S | W |
| | 2% | V | G | L | W | A |
| | 2% | I | G | L | E | R |
| | 2% | V | G | L | W | R |
| VHH 42% | 20% | Y/F | Q/E | R | L/F | W |
| | 6% | F | D | R | F/L | W |
| | 4% | Y | L | R | L | W |
| | 4% | F/Y | Q | R | L | G |
| | 4% | Y | Q | R | L | L |
| | 2% | F | P | R | F | W |
| | 2% | F | E | R | A | A |

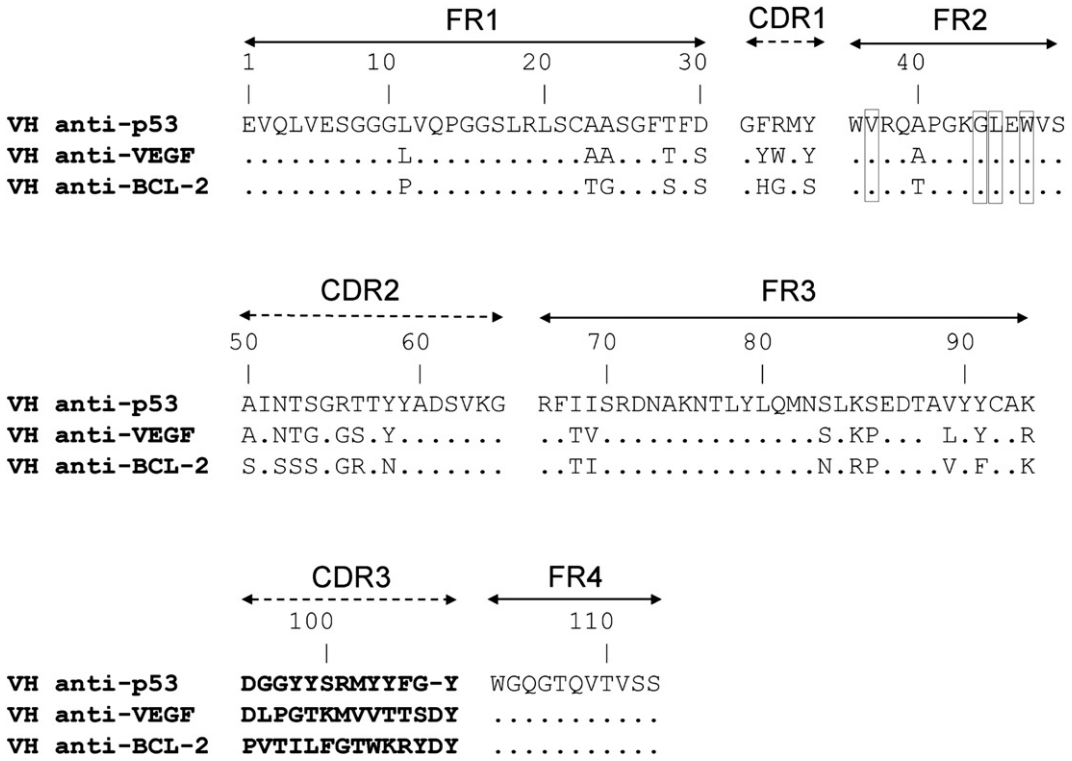


**Fig. 2.** Sequence alignment of the three selected VH binders against proteins p53 (GenBank accession no. FJ810218), BCL-2 (GenBank accession no. FJ810219), and VEGF (GenBank accession no. FJ810220), respectively. VH specific hallmarks V37, G44, L45, W47 are marked boxes. Unchanged positions are represented in dots. CDR3 loop is marked in bold letters. Kabat numbering system is used.

### 3.3. Production of soluble VHs

The three binders for proteins p53, VEGF and BCL-2 (Fig. 2), all containing specific VH hallmarks, were produced in large scale in *E. coli* TG1 strain. The antibodies, which carry a 6His tag to facilitate purification, were found to be transported into the periplasm (data not shown). Following cell disruption and precipitation of the periplasmic protein fraction with ammonium sulphate, metal ion affinity purification yielded highly pure, single band with molecular mass of approximately 17 kDa based on Coommasie stained SDS polyacrylamide gel (Fig. 3). VH protein yields was in the range of 3 mg/L of bacterial culture (2.70 mg/L for VH anti-p53, 1.69 mg/L for VH anti-BCL-2 and 4.55 mg/L for VH anti-VEGF).

Gel filtration on Superdex-75 column evidenced the monomeric nature of isolated VH domains. The monomer peak was found to exhibit an apparent MW of about 12 kDa (theoretical MW of VH anti-p53 is 16.2 kDa). For comparison a VHH protein containing all specific hallmarks was analysed under the same conditions. Similarly, a single peak was observed at approximately 10 kDa (theoretical value is 15.4 kDa). Small shoulders present in both VH and VHH elution profiles at larger MW that the main peak likely result from minor impurities also visualized on SDS page gels. Retention times for these shoulders do not match the ones for potential dimers. Similar elution profiles were observed for VH anti-BCL-2 and VH anti-VEGF.

### 3.4. Affinity characterisation

Specificity and affinity of isolated VH single domain antibodies against proteins p53, BCL-2 and VEGF was determined. ELISA, performed with immobilised antigen and VH-antibodies concentration ranging from 3 μM to 1 nM, confirmed estimated affinity constants, determined by Biacore surface plasmon resonance analysis. Isolated VH binders were found to be specific for their corresponding antigens as no cross-reaction was evidenced with other proteins or with control surface. Affinity constants ($K_D$) were estimated to be in the 10–25 nM range for the VH anti-p53, but larger (30–100 nM range) for the VH anti-VEGF (by Biacore) and VH anti-BCL-2 (by ELISA). Noticeably, the VH anti-p53 features a non-classical binding profile (Fig. 4) and a relatively low association constant (about 5000 $M^{-1}s^{-1}$) based on analysis of initial binding rates. This might suggest that some slow conformational reorganisation is occurring during the binding to the antigen. Dissociation kinetic was found to be biphasic with a major slow phase (about 0.47 $h^{-1}$) allowing globally a good affinity in the nM range. A 1:1 stoichiometry was observed between the isolated VH anti-p53 and VH anti-VEGF and their corresponding antigens.

### 3.5. Modeling

Modeling was performed for all three characterised VH using the unique available llama VH structure as template.

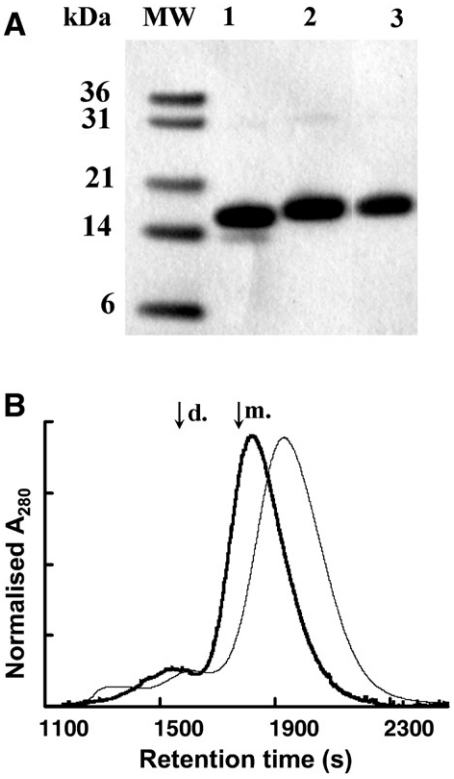


**Fig. 3.** (A) 4–12% polyacrylamide SDS electrophoresis gel of purified VHs following ammonium sulphate precipitation and metal affinity column. A single band of approximately 17 kDa is observed for each purified VH domain (lane 1: VH anti-p53; lane 2: VH anti-BCL-2; lane 3: VH anti-VEGF). (B) Size exclusion chromatography of VH anti-p53 (bold lane) and VHH protein (thin lane). Arrows indicate the elution time of b5 calibration protein (monomer and dimer).

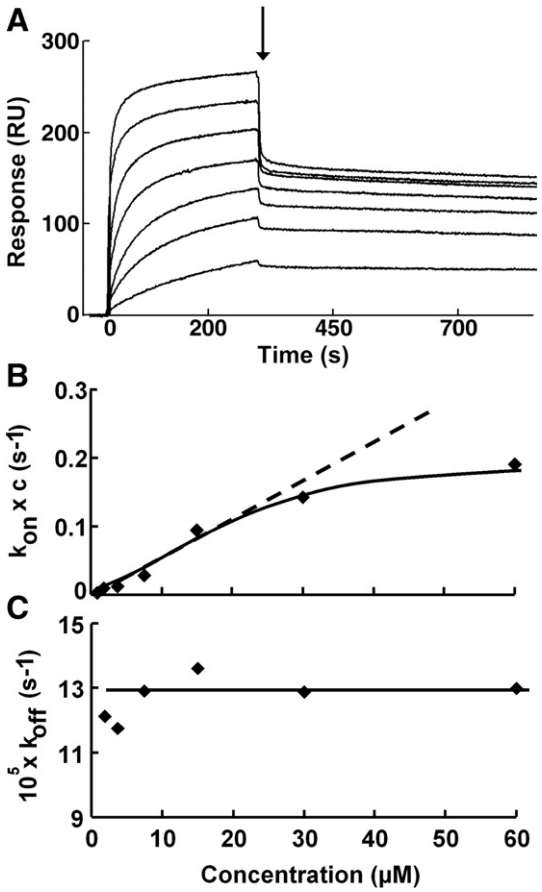


**Fig. 4.** (A) Sensorgrams for the binding of VH anti-p53 (60, 30, 15, 7.5, 3.75, 1.8 and 0.93 μM) on immobilised p53 antigen (1300 RU). Start of washing (300 s after injection) is marked with an arrow. (B) Initial rates for the association of the VH as function of VH concentrations. Rates were calculated from the biexponential fitting of SPR traces. (C) Apparent first order rate constant for the slow phase of VH dissociation (see Materials and methods section) as function of the VH concentration used during the loading phase.

Fifty structures containing different CDR3 loop conformations were calculated for each VH and the ten lowest energy models were retained in each case for further analysis. Fig. 5 was generated using one of the lowest energies VH anti BCL-2 calculated model. In this calculated model CDR3 loop has clearly the ability to fully cover the exposed hydrophobic patch, represented by V37, G44, L45, W47 resulting into a soluble VH, efficient expression and lack of aggregation.

## 4. Discussion

In llama serum approximately 40% of homodimer heavy chain antibodies are present, the remaining representing conventional four chain heterodimer antibodies. Since the discovery that VHH are able to specifically recognize antigens in the absence of the VL counterpart, attention was focused on the production and exploitation of their variable domain, which is small, easy to express and efficiently produced in *E. coli*. Little attention has been paid on llama VH domains, although their greater sequence similarity with human homologs represents a potential advantage and up to date, recombinant libraries were systematically constructed as VHH only libraries (Arbabi Ghahroudi et al., 1997, van Koningsbruggen et al., 2003).

This prompted us to reinvestigate the potentiality of llama VH from HcAb and to analyse how a VH/VHH mixed llama library evolves during a phage display based selection process. For this purpose, we designed a simple and efficient approach involving a single PCR step to preserve sequence diversity. Using subsequent phage-display selection, we succeeded to obtain specific binders against 3 cancer markers (p53, VEGF, BCL-2). In contrast, selection directed against TGFα and IL-8 failed to isolate specific binders due to cross-reactivity and probably also to the small size (8 kDa) of one of the antigen.

Selected antibodies against p53, VEGF and BCL-2 were found specific and efficiently produced in *E. coli.* Surprisingly, they all contain a VH characteristic signature, when the initial library and subpopulations following the first and second round of selection exhibited equilibrated population of VH and VHH. This suggests that the final round, which included not only selection for binders, but also screening for selectivity was discriminating, favouring VH solutions. It can be also pointed out that the finally selected VH clones are "pure VH" featuring all specific VH signatures when initial or

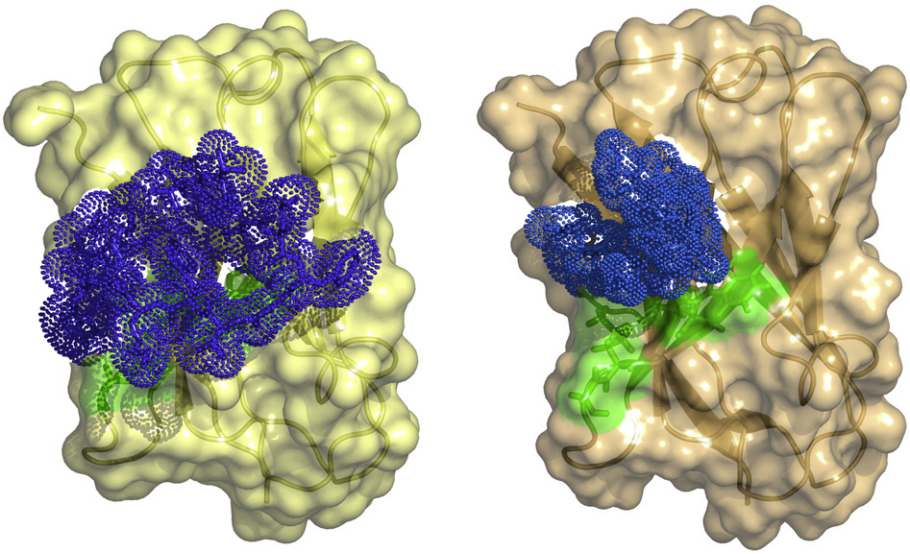

**Fig. 5.** Structure and modeling of the VH single chain antibodies. The surface of VH hallmarks (V37, G44, L45, W47) is coloured in green, CDR3 is represented with blue dots. Left, the selected VH anti BCL-2 was modeled using the 1IEH pdb entry as initial framework, sequence substitution and two rounds of energy minimalization using MODELLER. The structure was randomly picked among the structures of lower energy. Model shows that hydrophobic patch can be covered by the long CDR3. Right, for comparison the NMR-solved structure of the llama VH protein with pdb accession number 1IEH is shown.

partially selected libraries contained a large amount of clones with partial VH or VHH signatures. Different groups previously reported selection of antibodies that lack VHH specific hallmarks, featuring high affinity, specificity and productivity (Conrath et al., 2001; Harmsen and De Haard, 2007; Lafaye P. personal communication), those domains contain hydrophilic residues at position 103 (Kabat numbering), instead of highly preserved hydrophobic tryptophan present in conventional antibodies and they were named VHH-like conventional VHs. This finding was also supported by the report of Tanha et al. (2002), who demonstrated that high affinity VH, originating from germline of conventional antibodies can be selected. In our case, attempt to detect mature anti-p53 immunoglobulin in the serum of cancer tissue immunized llama was unsuccessful, consistently with a potential naïve nature of isolated VH. Even if we failed to experimentally demonstrate that our isolated VH clones were present as such within the unamplified cDNA pool, sequence alignment and clustering with all available *Lama pacos* germline sequences clearly evidenced that isolated sequences belong to the conventional VH and not to the VHH gene cluster.

Molecular features allowing suitable affinity, selectivity and productivity of single domain llama antibodies were examined. Since CDR3 is the most implicated loop in antigen recognition (Desmyter et al., 2001, De Genst et al., 2006), we hypothesize that CDR3 length might also play a critical role for the solubility of VH domains. In the initial library, the CDR3 lengths were very dispersed with a minimum of three residues. After the first and second selection round the average length reached 6–8 residues and finally 13–14 residues after the final selection. Previous studies showed that VHH domains contain a significantly longer CDR3 loop than VHs (Vu et al., 1997), possibly to compensate affinity loss for targets due to the lack of VL counterpart. Interestingly, all finally selected VH binders contain unusually long CDR3 instead of dispersed length (average of 10 residues) normally found in VH (Vu et al., 1997). Modeling of the selected VH proteins using the sole available VH structural coordinate as molecular template highly suggests that the long CDR3 loop could cover the hydrophobic patch from the skeleton and expose to solvent a fairly hydrophilic surface (CDR3).Thus, long CDR3 (above 12 residues) could shield the hydrophobic area, located in FR2 and favour isolation of VH domains, meeting requirements for biotechnology. This provides a potential way for building optimized VH library by selecting longer PCR products before cloning and selection. However, VH were *in vivo* matured in the presence of VL and afterwards cloned and expressed in the absence of it. In such a situation, the CDR3 loop might adopt a different conformation than in the native four-chain antibody. Observation of fairly unusual binding kinetic traces might be related to the need of additional CDR3 conformational change during the antigen binding step. However, the high affinity of isolated VH anti-p53, indicates that such induced fit could led to a high stability complex with long dissociation time. This does not exclude that the original combination of VH and VL would result in even higher affinity. Until now, therapeutic *camelidae* single domain antibody research has been focusing on VHH or camelised human VH. Present work clearly illustrates on several examples that llama VHs directed against disease markers can also be efficiently isolated and expressed. In addition, the suggestion that the length of the CDR3 loop can be a critical factor for the isolation of VH-only llama antibodies with optimal biotechnological properties seems to be important for applications.

## Competing interests statement

None of the authors has any potential financial conflict of interest related to this manuscript.

## Acknowledgments

This work was financially supported by a grant of Ad-Futura Slovene Human Resources (Development and Scholarship Foundation) and by grants of EGIDE, PICS 3294 (French–Slovene program) and J1-9740-0381 (Slovenian Research Agency). We thank Martine Chartier for mRNA isolation. Thanks to Nicolas Soler, Aude Laisne and Jacqueline Loeper for critical reading of the manuscript. We are grateful to Patrick England from Institute Pasteur for providing SPR analysis in the frame of INSERM "Atelier de formation No. 183".

## References

Alvarez-Rueda, N., Behar, G., Ferre, V., Pugniere, M., Roquet, F., Gastinel, L., Jacquot, C., Aubry, J., Baty, D., Barbet, J., Birkle, S., 2007. Generation of llama single-domain antibodies against methotrexate, a prototypical hapten. Mol. Immunol. 44, 1680.

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

Behar, G., Siberil, S., Groulet, A., Chames, P., Pugniere, M., Boix, C., Sautes-Fridman, C., Teillaud, J.L., Baty, D., 2008. Isolation and characterization of anti-FcgammaRIII (CD16) llama single-domain antibodies that activate natural killer cells. Protein Eng. Des. Sel. 21, 1.

Chartier, A., Raz, V., Sterrenburg, E., Verrips, C.T., van der Maarel, S.M., Simonelig, M., 2009. Prevention of oculopharyngeal muscular dystrophy by muscular expression of llama single-chain intrabodies in vivo. Hum. Mol. Genet. 18, 1849.

Clackson, T., Lowman, H., 2007. Phage Display a Practical Approach. Oxford University Press.

Conrath, K.E., Lauwereys, M., Galleni, M., Matagne, A., Frere, J.M., Kinne, J., Wyns, L., Muyldermans, S., 2001. Beta-lactamase inhibitors derived from single-domain antibody fragments elicited in the camelidae. Antimicrob. Agents Chemother. 45, 2807.

Cortez-Retamozo, V., Backmann, N., Senter, P.D., Wernery, U., De Baetselier, P., Muyldermans, S., Revets, H., 2004. Efficient cancer therapy with a nanobody-based conjugate. Cancer Res. 64, 2853.

Davies, J., Riechmann, L., 1994. 'Camelising' human antibody fragments: NMR studies on VH domains. FEBS Lett. 339, 285.

Davies, J., Riechmann, L., 1996. Single antibody domains as small recognition units: design and in vitro antigen selection of camelized, human VH domains with improved protein stability. Protein Eng. 9, 531.

De Genst, E., Silence, K., Decanniere, K., Conrath, K., Loris, R., Kinne, J., Muyldermans, S., Wyny, L., 2006. Molecular basis for the preferential cleft recognition by dromedary heavy-chain antibodies. Proc. Natl. Acad. Sci. U. S. A. 103, 4586.

Desmyter, A., Decanniere, K., Muyldermans, S., Wyns, L., 2001. Antigen specificity and high affinity binding provided by one single loop of a camel single-domain antibody. J. Biol. Chem. 276, 26285.

Desmyter, A., Spinelli, S., Payan, F., Lauwereys, M., Wyns, L., Muyldermans, S., Cambillau, C., 2002. Three camelid VHH domains in complex with porcine pancreatic alpha-amylase. Inhibition and versatility of binding topology. J. Biol. Chem. 277, 23645.

Eswar, N., Webb, B., Marti-Renom, M.A., Madhusudhan, M.S., Eramian, D., Shen, M.Y., Pieper, U., Sali, A., 2007. Comparative protein structure modeling using MODELLER. Curr. Protoc. Protein Sci. Chapter 2, Unit 2.9.

Frenken, L.G., van der Linden, R.H., Hermans, P.W., Bos, J.W., Ruuls, R.C., de Geus, B., Verrips, C.T., 2000. Isolation of antigen specific llama VHH antibody fragments and their high level secretion by *Saccharomyces cerevisiae*. J. Biotechnol. 78, 11.

Gille, C., Frommel, C., 2001. STRAP: editor for STRuctural Alignments of Proteins. Bioinformatics 17, 377.

Ha, G.H., Lee, S.U., Kang, D.G., Ha, N.Y., Kim, S.H., Kim, J., Bae, J.M., Kim, J.W., Lee, C.W., 2002. Proteome analysis of human stomach tissue: separation of soluble proteins by two-dimensional polyacrylamide gel electrophoresis and identification by mass spectrometry. Electrophoresis 23, 2513.

Hamers-Casterman, C., Atarhouch, T., Muyldermans, S., Robinson, G., Hamers, C., Songa, E.B., Bendahman, N., Hamers, R., 1993. Naturally occurring antibodies devoid of light chains. Nature 363, 446.

Harmsen, M.M., De Haard, H.J., 2007. Properties, production, and applications of camelid single-domain antibody fragments. Appl. Microbiol. Biotechnol. 77, 13.

Harmsen, M.M., Ruuls, R.C., Nijman, I.J., Niewold, T.A., Frenken, L.G., de Geus, B., 2000. Llama heavy-chain V regions consist of at least four distinct subfamilies revealing novel sequence features. Mol. Immunol. 37, 579.

Harmsen, M.M., van Solt, C.B., van Zijderveld-van Bemmel, A.M., Niewold, T.A., van Zijderveld, F.G., 2006. Selection and optimization of proteolytically stable llama single-domain antibody fragments for oral immunotherapy. Appl. Microbiol. Biotechnol. 72, 544.

Hmila, I., Abdallah, R.B., Saerens, D., Benlasfar, Z., Conrath, K., Ayeb, M.E., Muyldermans, S., Bouhaouala-Zahar, B., 2008. VHH, bivalent domains and chimeric Heavy chain-only antibodies with high neutralizing efficacy for scorpion toxin AahI′. Mol. Immunol. 45, 3847.

Hultberg, A., Tremblay, D.M., de Haard, H., Verrips, T., Moineau, S., Hammarstrom, L., Marcotte, H., 2007. Lactobacillli expressing llama VHH fragments neutralise *Lactococcus* phages. BMC Biotechnol. 7, 58.

Joosten, V., Gouka, R.J., van den Hondel, C.A., Verrips, C.T., Lokman, B.C., 2005. Expression and production of llama variable heavy-chain antibody fragments (V(HH)s) by *Aspergillus awamori*. Appl. Microbiol. Biotechnol. 66, 384.

Kabat, E.A., Wu, T.T., Perry, H.M., Gottesman, K.S., Foeller, C., 1991. Sequences of Proteins of Immunological Interest. Diane Books Publishing Company.

Lauwereys, M., Arbabi Ghahroudi, M., Desmyter, A., Kinne, J., Holzer, W., De Genst, E., Wyns, L., Muyldermans, S., 1998. Potent enzyme inhibitors derived from dromedary heavy-chain antibodies. Embo J. 17, 3512.

Muyldermans, S., Atarhouch, T., Saldanha, J., Barbosa, J.A., Hamers, R., 1994. Sequence and structure of VH domain from naturally occurring camel heavy chain immunoglobulins lacking light chains. Protein Eng. 7, 1129.

Olichon, A., Surrey, T., 2007. Selection of genetically encoded fluorescent single domain antibodies engineered for efficient expression in *Escherichia coli*. J. Biol. Chem. 282, 36314.

Perez, J.M., Renisio, J.G., Prompers, J.J., van Platerink, C.J., Cambillau, C., Darbon, H., Frenken, L.G., 2001. Thermal unfolding of a llama antibody fragment: a two-state reversible process. Biochemistry 40, 74.

Rahbarizadeh, F., Rasaee, M.J., Forouzandeh, M., Allameh, A.A., 2006. Over expression of anti-MUC1 single-domain antibody fragments in the yeast *Pichia pastoris*. Mol. Immunol. 43, 426.

Rajabi-Memari, H., Jalali-Javaran, M., Rasaee, M.J., Rahbarizadeh, F., Forouzandeh-Moghadam, M., Esmaili, A., 2006. Expression and characterization of a recombinant single-domain monoclonal antibody against MUC1 mucin in tobacco plants. Hybridoma (Larchmt) 25, 209.

Tanha, J., Dubuc, G., Hirama, T., Narang, S.A., MacKenzie, C.R., 2002. Selection by phage display of llama conventional V(H) fragments with heavy chain antibody V(H)H properties. J. Immunol. Methods 263, 97.

van der Linden, R., de Geus, B., Stok, W., Bos, W., van Wassenaar, D., Verrips, T., Frenken, L., 2000. Induction of immune responses and molecular cloning of the heavy chain antibody repertoire of *Lama glama*. J. Immunol. Methods 240, 185.

van Koningsbruggen, S., de Haard, H., de Kievit, P., Dirks, R.W., van Remoortere, A., Groot, A.J., van Engelen, B.G., den Dunnen, J.T., Verrips, C.T., Frants, R.R., van der Maarel, S.M., 2003. Llama-derived phage display antibodies in the dissection of the human disease oculopharyngeal muscular dystrophy. J. Immunol. Methods 279, 149.

Vranken, W., Tolkatchev, D., Xu, P., Tanha, J., Chen, Z., Narang, S., Ni, F., 2002. Solution structure of a llama single-domain antibody with hydrophobic residues typical of the VH/VL interface. Biochemistry 41, 8570.

Vu, K.B., Ghahroudi, M.A., Wyns, L., Muyldermans, S., 1997. Comparison of llama VH sequences from conventional and heavy chain antibodies. Mol. Immunol. 34, 1121.

Worn, A., Pluckthun, A., 1998. Mutual stabilization of VL and VH in single-chain antibody fragments, investigated with mutants engineered for stability. Biochemistry 37, 13120.

Yau, K.Y., Groves, M.A., Li, S., Sheedy, C., Lee, H., Tanha, J., MacKenzie, C.R., Jermutus, L., Hall, J.C., 2003. Selection of hapten-specific single-domain antibodies from a non-immunized llama ribosome display library. J. Immunol. Methods 281, 161.