Generation of a proteomic platform of high processivity for the identification and characterization of recombinant VHHs against CD105

Milagros Quintana Caceda

UNIVERSIDAD PERUANA CAYETANO HEREDIA

Université Grenoble Alpes

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITE GRENOBLE ALPES

préparée dans le cadre d'une cotutelle entre l'*Université Grenoble Alpes* et l´*Universidad Peruana Cayetano Heredia*

Spécialité : **Chimie-Biologie**
Arrêté ministériel : le 25 mai 2016

Présentée par

Milagros QUINTANA CACEDA

Thèse dirigée par **Yoann ROUPIOZ, DR, CEA-CNRS UGA** et **José R. ESPINOZA, LID, UPCH**

préparée au sein **de l'équipe Chimie pour la Reconnaissance et l'Etude des Assemblages Biologiques (CREAB) du laboratoire Systèmes Moléculaires et NanoMatériaux pour l'Énergie et la Santé (SYMMES, UMR 5819 UGA-CEA-CNRS) au CEA-Grenoble**
et **de l'équipe Nanobodies du Unidad de Biotecnología Molecular (UBM) au Laboratorios de Investigación y Desarrollo (LID), UPCH**

dans **l'École Doctorale de Chimie et Sciences du Vivant** et **la Escuela Doctoral Franco Peruana en Ciencias de la Vida**

Génération d'une plateforme protéomique de traitement haut débit pour l'identification et la caractérisation des VHH recombinants contre CD105

Thèse soutenue publiquement le **31 mars 2022**,
devant le jury composé de :

Mme Patricia HERRERA
Professeure à l'Universidad Peruana Cayetano Heredia, Lima-Pérou, Présidente du jury
Mme Virginie FAURE
Maîtresse de conférences à l'Université Grenoble Alpes, Grenoble, Examinatrice
M. Alain ROUSSEL
Responsable de « Interactions Hote- Pathogène », LISM, Marseille, Rapporteur
Mme Martha VALDIVIA
Professeure à l'Universidad Nacional Mayor de San Marcos, Lima-Pérou, Rapporteur
M. Yoann ROUPIOZ
Directeur de Recherche au CNRS, Grenoble, Directeur de thèse
M. José R ESPINOZA
Professeur à l'Universidad Peruana Cayetano Heredia, Lima-Pérou, Directeur de thèse

Abstract

Generation of a proteomic platform of high processivity for the identification and characterization of recombinant VHHs against CD105

Screening of a VHH cDNA library raised from alpacas immunised with a lysate of T24 cells (human urine bladder cancer cell line over-expressing CD105), with soluble CD105 by phage display and phage-ELISA resulted in 19 VHHs that bind to human CD105 protein, an endothelial co-receptor of TFG-β involved in the regulation of angiogenesis and tumour development. The selected anti-CD105 VHH cDNAs were amplified, sequenced and analysed with ExPASY, GeneDoc and BLAST tools. The nineteen sequences showed the typical structural features of a VHH and all of them were different. The VHH cDNAs were subcloned from the phagemid vector pHEN2 into pET22b(+) vector for expression and purification of recombinant VHH proteins by affinity chromatography on Ni-NTA agarose columns. Thirteen recombinant VHH proteins maintained the ability to bind CD105 as determined by ELISA. These anti-CD105 VHHs were microarrayed on biochips to evaluate their specificity and affinity to CD105 membrane-bound on cells by Surface Plasmon Resonance imaging (SPRi). SPRi data showed that the recombinant VHHs bind SC cells that overexpress CD105 on their surface but not or little THP-1 cells that do not, proving their specificity. The average of SPRi reflectivity variation among the 13 VHHs at 2 hours assays was 4.59 with SC cells and 2.69 with the THP-1 cells, cell density average was 578 cells/area with SC cells and 367 cells/area with THP-1 cells and the average of the approximation to the apparent affinity constant was 1.1019 with SC cells and 1.0626 with THP-1 cells. The VHH were grouped according to their specificity and affinity approximation, six had high specificity and affinity approximation (3 of them even higher than the positive control), five were specific but had low affinity approximation and two had low specificity and affinity approximation. The most efficient (six top-ranked) anti-CD105 VHHs, that bind with high affinity and specificity to CD105 in solution (sCD105) and on cells (membrane-bound CD105), are proposed as CD105 detection tools for research both in studies that seek a better understanding of the CD105 function *in vitro* and *in vivo* and in application as nanoprobes for the diagnosis and therapeutics of diseases in which angiogenesis plays an important role.

Keywords: VHH, CD105, SPRi, biochip, nano-probes

Résumé

Génération d'une plateforme protéomique de traitement haut débit pour l'identification et la caractérisation des VHH recombinants contre CD105

Le criblage d'une bibliothèque d'ADNc de VHH provenant d'alpagas immunisés avec un lysat de cellules T24 (lignée cellulaire de cancer de la vessie humaine surexprimant CD105), avec CD105 soluble par l'exposition sur phage (phage display) et phage-ELISA a permis de trouver 19 VHHs qui se lient à la protéine CD105 humaine, un co-récepteur endothélial de TFG-β impliqué dans la régulation de l'angiogenèse et du développement tumoral. Les ADNc des VHH anti-CD105 sélectionnés ont été amplifiés, séquencés et analysés avec les outils ExPASY, GeneDoc et BLAST. Les dix-neuf séquences présentaient les caractéristiques structurelles typiques d'un VHH et toutes étaient différentes. Les ADNc de VHH ont été sous-clonés du vecteur phagemid pHEN2 dans le vecteur pET22b(+) pour l'expression et la purification des protéines VHH recombinantes par chromatographie d'affinité sur des colonnes d'agarose Ni-NTA. Treize protéines VHH recombinantes ont conservé la capacité de lier CD105 comme déterminé par ELISA. Ces VHH anti-CD105 ont été immobilisés sur des biopuces afin d'évaluer leur spécificité et leur affinité avec le CD105 lié à la membrane des cellules par l´imagerie de la résonance des plasmons de surface (SPRi). Les données SPRi ont montré que les VHHs se lient aux cellules SC qui surexpriment CD105 à leur surface mais pas ou peu aux cellules THP-1 qui ne le font pas, prouvant ainsi leur spécificité. La moyenne de la variation de la réflectivité SPRi parmi les 13 VHHs à 2 heures d'essais était de 4,59 avec les cellules SC et de 2,69 avec les cellules THP-1, la moyenne de la densité cellulaire était de 578 cellules/superficie de référence avec les cellules SC et de 367 cellules/superficie de référence avec les cellules THP-1 et la moyenne de l'approximation de la constante d'affinité apparente était de 110,19 x 10^{-2} avec les cellules SC et de 106,26 x 10^{-2} avec les cellules THP-1. Les VHH ont été répartis en groupes en fonction de leur spécificité et de leur approximation d'affinité. Six avaient une spécificité et une approximation d'affinité élevées (trois encore plus que le contrôle positif), cinq étaient spécifiques mais avaient une approximation d'affinité faible et deux avaient une spécificité et une approximation d'affinité faibles. Les VHH anti-CD105, les plus efficaces qui se lient avec une affinité et une spécificité élevées au CD105 en solution (sCD105) et sur les cellules (CD105 lié à la membrane), ont été proposés comme outils de détection du CD105 pour la recherche à la fois dans les études qui cherchent à mieux comprendre la fonction du CD105 *in vitro* et *in vivo* et dans l'application comme nanosondes pour le diagnostic et la thérapeutique des maladies dans lesquelles l'angiogenèse joue un rôle important.

Mots clés: VHH, CD105, SPRi, biopuce, nano-sondes

Acknowledgements

The present work was funded by the project "Generación de una plataforma proteómica de alta procesividad para la identificación de nanocuerpos recombinantes de alpaca generados contra antígenos de relevancia biomédica" granted to José R. Espinoza by the National Council of Science and Technology (CONCYTEC) of Peru, Contract "Convenio 111-2015-FONDECYT". In addition, I received a doctoral scholarship from the Escuela Doctoral Franco Peruana en Ciencias de la Vida, contrato de subvención N° 086-2017 - FONDECYT for the development of my thesis project

Many thanks to my thesis directors, José Espinoza and Yoann Roupioz for their unconditional support and for sharing their knowledge and professional experience for the realization of this research work, and above all, for the trust placed in me and their time… merci beaucoup du fond du cœur.

Thanks to all my more than colleagues, friends, from UBM, LID, UPCH, especially to Eduardo Gushiken for providing the VHH cDNA library obtained during his master's research project. To Teresa, Pilar, Agueda, Dalia, Diego for their invaluable help in performing and improving the experiments regardless of the time. To Dra. Paty whose advice based on her great professional experience provided an invaluable contribution to this research work.

Thanks to the CREAB team, from SyMMES of CEA, Grenoble, for allowing me to be part of the group during my stay in France, the weekly meetings and productive discussions enriched my professional formation. To Arnaud, Yanxia, Aurélie, Elodie, Pierre, Martial, for their constant support and sincere interest in the successful completion of my thesis project. To Raph and Loic for helping me in learning the experimental procedures and the good handling of the equipment and to José, for always providing what was necessary for the realization of my work and especially for being my translator in some occasions… thank you for your patience. To Christine and Didier for welcoming me at Biomade, where I was able to carry out fundamental experiments for the continuity of my work. To Annette, Abdu, Ricardo, Lucile, Daniel, Jonathan, Larry, Marie, Maud, Élise, Charlotte, Eric, Marine, Sophie, Erenildo, Yosr, for sharing their time and expertise to always achieve the best results.

Thanks to Jorge Morales for his help with data processing and MATLAB software.

Thanks to my family, whose unconditional support was fundamental for the development and completion of this work. To my parents Teobaldo and Petronila, my sisters Hilda, Gloria, Maria, Alicia, Ofelia and Rocio and my brother Teobaldo, my nephews and nieces, especially Milagros and Noelia for their help with the English writing. To my brothers-in-law, aunts, uncles and cousins. To each one of them, thank you very much for your understanding and love.

Thanks to all the people who, in one way or another, contributed, manually .or animically, to the realisation of my thesis project.

Muchas gracias a todos.

Content

List of figures

Abbreviations

aa: amino acid
Abs: Antibodies
Ags: Antigens
ALK: Activin receptor-Like Kinase
AML: Acute Myeloid Leukaemia
AVM: ArterioVenous Malformations
ATCC: American Type Culture Collection
BLAST: Basic Local Alignment Search Tool
BMP: Bone Morphogenetic Protein
bp: base pair
BSA: Bovine Serum Albumin
CAIX: Carbonic Anhydrase IX
cAMP: cyclic Adenosine MonoPhosphate
CCD: Charge-Coupled Device
CD: Cluster of Differentiation
CD105: Cluster of Differentiation 105
cDNA: complementary DNA
CDR or HV: Complementary determining region or Hypervariable region
CEA: CarcinoEmbryonic Antigen
CH: Constant domain of Heavy chain
CL: Constant domain of Light chain
CL: Cell Lysate
CN: Negative Control
COVID-19: Coronavirus disease 19
COX2: Cyclooxygenase 2
CP: Positive Control
CTLA-4: Cytotoxic T-Lymphocyte Antigen 4
CXCR7: CXC chemokine receptor 7
DENV: Dengue Virus
DMSO: DiMethyl SulfOxide
dNTPs: Deoxyribonucleotide triphosphates
E: protein Eluates
ECM: Extracellular Matrix
EDTA: EthyleneDiamineTetraacetic Acid
EIS: Electrochemical Impedance Spectroscopy
EMBL-EBI: European Molecular Biology Laboratory-European Bioinformatics Institute
ENG: CD105 gene
Eng-iKOe: Endothelial-specific endoglin knockout

ELISA: Enzyme-Linked ImmunoSorbent Assays
eNOS: enzyme Nitric Oxide Synthase
ExPASy: Expert Protein Analysis System
Fab: Fragment antigen binding
FBS: Fetal Bovine Serum
Fc: Fragment Crystallizable
FGF: Fibroblast growth factor
FR: Framework Region
FT: Flow through
GRAVY: Grand Average of Hydropathy
GST: Glutathione S-transferase
GvHD: Graft-versus-Host Disease
H: Heavy chain of antibodies
HCAb: Heavy Chain Antibody
HEK293: Human Embryonic Kidney cells
HEPES: (4-(2-HydroxyEthyl)-1-PiperazineEthaneSulfonic acid
HER2: Human Epidermal growth factor Receptor-2
HHT: Hereditary haemorrhagic telangiectasia
HIF1: Hypoxia Inducible Factor 1
HRE: Hypoxia Response Elements
HRP: HorseRadish Peroxidase
HUVEC: Human Umbilical Vein Endothelial Cells
IBPs: Ice-Binding Proteins
ICAM-1: InterCellular Adhesion Molecule 1
IC: Induced Clone
ID1: Inhibitor of DNA binding 1
Ig: Immunoglobulin
IL: Interleukin
IMAC: Immobilised Metal Affinity Chromatography
IMDM: Iscove's Modified Dulbecco's Medium
IPTG: IsoPropyl- β-D-1-ThioGalactopyranoside
kDa: kiloDalton
KLF6: Kruppel Like Factor 6
L: Light chain of antibodies
L-CD105: Long CD105
LED: Light-Emitting Diode
LSPR: Localised Surface Plasmon Resonance
M: molecular Marker
MBP: Maltose Binding Protein
MCS: Multiple Cloning Site

MERS-CoV: Middle East Respiratory Syndrome CoronaVirus
MHC-II: Major Histocompatibility Complex class II
MMP-14 or MT1-MMP: Matrix MetalloProteinase 14 or Membrane Type I-MMP
MOPS: 3-(N-MOrpholino) PropaneSulfonic acid
MP-SPR: MultiParametric Surface Plasmon Resonance
mRNA: messenger RNA
MSC: Mesenchymal Stem Cells
MUC1: Mucin 1
MUSCLE: Multiple Sequence Comparison by Log-Expectation
MWCO: Molecular Weight CutOff
NCBI: National Center for Biotechnology Information
NF-κB: Nuclear Factor kappa B
NHL: Non-Hodgkin's lymphoma
NHS: N-HydroxySuccinimide
NIC: Non Induced Clone
Ni-IDA: Nickel-IminoDiacetic Acid
Ni-NTA: Nickel-NitriloTriacetic Acid
NS1: Non-structural protein 1
PAI-1: Plasminogen Activator Inhibitor type 1
PBMCs: Peripheral Blood Mononuclear Cells
PBS: Phosphate Buffer Saline
PCR: Polymerase Chain Reaction
PDB: Protein Data Bank
PD-L1: Programmed Death-ligand 1
PDZ: Post synaptic density protein (PSD95), Drosophila disc large tumor suppressor (Dlg1), and Zonula occludens-1 protein (Zo-1)
PE: Protein Eluate from Pellet
PEEK: PolyEther Ether Ketone
PKC: Protein kinase C
PSMA: Prostate Specific Membrane Antigen
PWFT: Pellet Wash Flow through
RBD: Receptor Binding Domain
RBS: Ribosome Binding Site
RGD: Arginine-Glycine-Aspartic acid
ROS: Reactive Oxygen Species
ROI: Region Of Interest
RPM: Revolutions Per Minute
RPMI: Roswell Park Memorial Institute
RT°: Room Temperature
RT-PCR: Reverse Transcription PCR

RU: Resonance Unit
SARS-CoV-2: Severe Acute Respiratory Syndrome CoronaVirus 2
S-CD105: Short CD105
sCD105: soluble CD105
scFv: single-chain variable fragment
SDS-PAGE: Sodium Dodecyl Sulfate - PolyAcrylamide Gel Electrophoresis
SE: Protein Eluate from Supernatant
sFlt-1 or sVEGFR-1: Soluble Fms-like tyrosine kinase -1 or soluble form of VEGF receptor
SH3: SRC Homology 3
SIB: Swiss Institute of Bioinformatics
SMAD: MAD (Mothers Against Decantaplegic) and SMA (Small body size)
SPR: Surface Plasmon Resonance
SPRi: Surface Plasmon Resonance imaging
SPRi RV: SPRi Reflectivity Variation
SUMO: Small Ubiquitin-like MOdifier
SWFT: Supernatant Wash Flow through TBE: Tris-Borate-EDTA
Tm: Temperature melting
TGF-β: Transforming Growth Factor β
TGFβR-I: TGF-β Receptor 1
TMB: 3, 3´, 5, 5´-Tetrametilbencidina
TNFα: Tumor Necrosis Factor α
TPA: 12-O-tetradecanoylphorbol-13-acetate
VCAM1: Vascular Cells Adhesion Molecule 1
VEGF: Vascular Endothelial Growth Factor
VH: Variable domain of Heavy chain
VHH: Variable domain of HCAbs
VL: Variable domains of Light chain
VSMC: Vascular Smooth Muscle Cells
W: Wash.
ZP: Zona Pellucida

Chapter I: Generalities

1.1 Antibodies (Ab)

The manufacture of proteins engineered to bind to their target with high affinity and specificity has been inspired from natural antibodies.

The antibodies or immunoglobulins (Ig) are a vast family of glycoproteins (gamma globulin type) of approximately150 kDa, produced by the immune system, that bind with high affinity and specificity antigens (Ag), substances that are generally exogenous and recognized as dangerous or harmful to the organism. Antibodies are synthesised by the lymphocytes B and constitutes a main component of the humoral response of the adaptative immune system. Antibodies bind to the antigen and guide the response of the effector cells of the immune system for elimination of such antigen or antigen-bearing organisms. Immunoglobulins are found soluble in blood. lymph and other body fluids and attached on B-lymphocytes membrane, where they act as membrane receptors, showing a a structure that is identical to the soluble form[1-3].

1.1.1 Structure

The typical antibody structure is composed of two basic dimeric units denominated as heavy chain (H) and light chain (L) that are linked by disulphide bonds[4]. The heavy chain, 55 kDa per unit, is bigger in size than the light chain, 25 kDa per unit, and together form a Y-shaped structure, the monomer of Ig[5-7] (Figure 1.1). Five different isotypes are found in mammals that are clasiffied by the form of the heavy chain they possess, these are: IgG, IgM, IgA, IgE and IgD[8]. Each isotype performs different functions to direct the adequate immune response to a specific type of antigen[3]. The five different types of heavy chain are encoded by the genes γ, μ, α, ε and δ, while the light chain protein are encoded by λ and κ genes, that are present in every Ig isotype[3,5,8].

1.1.1.1 Structural domains

The heavy chain and the light chain possess structural constant domains (CH and CL) and variable domains (VH and VL). The heavy chains of Ig G, Ig A and Ig D present 3 constant domains (CH1, CH2, CH3) in tandem, while Ig M and Ig E present 4 (CH1, CH2, CH3, CH4)[4]. As for the light chains, only 1 constant domain (CL) is present in all isotypes. The heavy chains as well as the light chains present 1 variable domain in each chain (VH and VL). These variable domains are located at opposite sides of the Y shaped structure. Furthermore, there is a region of the heavy chain that gives flexibility to the structure and maintains the union between variable domains of the chains, the hinge region. Its high proline content confers it that high flexibility that allows its movement and rotation as two arms and makes it capable of easily combining with the molecules of a cellular membrane. The antibodies present the characteristic “immunoglobulin fold”, where two β sheets parallel to the longitudinal axis of the domain generate a “sandwich” shape, staying together by interactions between cysteines well conserved over the course of evolution, as well as other charged amino acids[9]. One of the sheets is made up of 4 antiparallel β-strands and the other of 5. Between consecutive β-strands there are loops of variable length[3].

The region of binding the antigens, located at the ends of the Y-shaped structure, is called Fragment antigen binding (Fab) and consists of a constant domain and the variable domain in the light and heavy chains[10]. In this way, the paratope or antigen binding site, consists of variable domains of the heavy and light chains (VH and VL) at the N terminus of the antibody monomer.The base of the Y-shaped structure modulates the activity of the immune cell assuring that the antibody generates an appropriate immune response to the given antigen. This region was called Fragment Crystallizable (Fc) and is composed of two or three constant domains of both heavy chains, depending on the type of antibody[11,12].

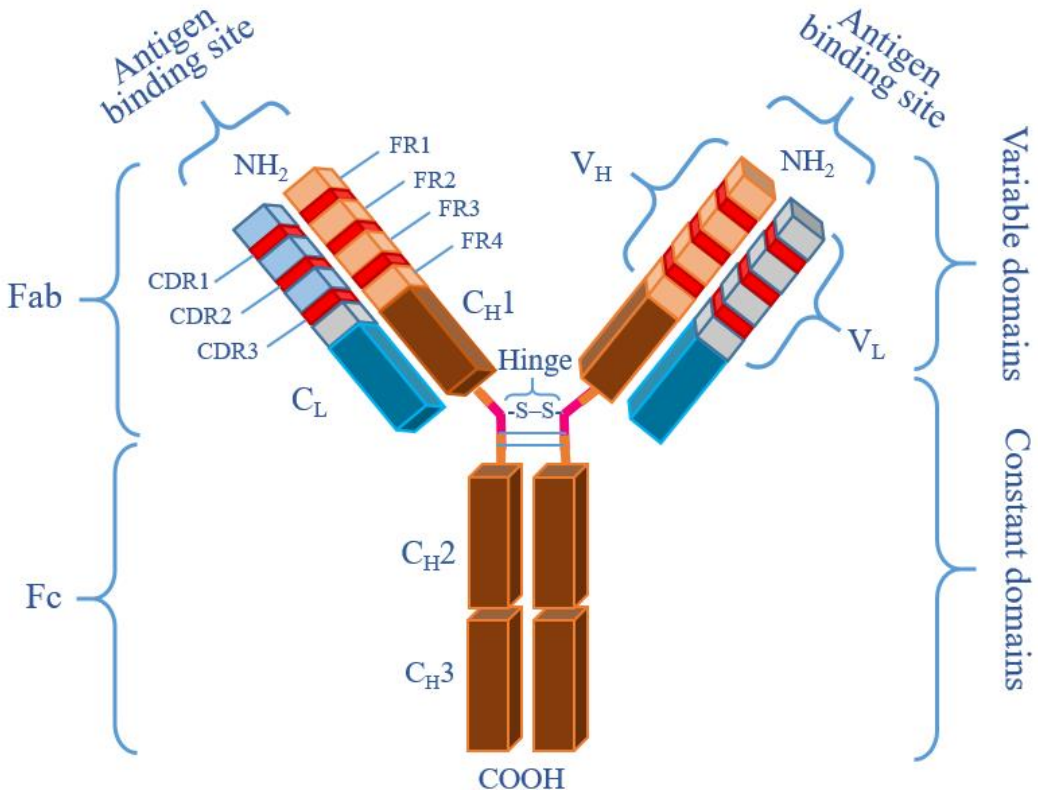


Figure 1.1– Conventional structure of an antibody (Ig monomer). In brown and blue, the constant domains of the heavy chain and light chain are shown respectively. In salmon/red and light blue/red, the variable region of the heavy and light chains. The conserved regions or FR (Framework Region) are signaled in light blue (V_L) and in salmon (V_H), the sequences of hypervariability, called CDR (*Complementary determining regions*) which are sequences that form *loops* of contact with the antigen are shown in red. In pink, the hinge region is signaled. The fragment of antigen binding (FAB) and the crystallizable fraction (FC) are indicated. (Adapted from Awwad & Angkawinitwong[13]).

1.1.1.2 Variable Domains

The variable domains present a high variability in their amino acid sequences to recognise different antigens. In humans up to 10^9 different antibody sequences are produced in a certain moment[14]. During B-lymphocyte maturation, a complex recombination process develops, called V(D)J recombination, which generates the variable domain of the heavy chain by the combination of 3 genes: V, D and J, and the light chain by the combination of V and J[15,16]. Furthermore, a process of random nucleotide insertion occurs that alters the sequence size and reading frame.

The differences of the variable domains in the distinct antibodies are located in three loops known as hypervariable regions (HV-1, HV-2 and HV-3) or complementarity

determining regions (CDR1, CDR2 and CDR3). The CDRs are maintained separated by regions of conserved frame, the framework regions (FR1, FR2, FR3 and FR4) (Figure 1.2). Usually, the CDR3 of the heavy chain is the major contributor to antigen-antibody binding specificity as it possesses greater variability in length and higher mobility than CDR1 and CDR2. The conformation of the paratope is related to the structure of the epitope, and its surface may well be a cavity, a groove or a flat surface that tends to bind to other small molecules, lineal epitopes or structural epitopes respectively[17,18].

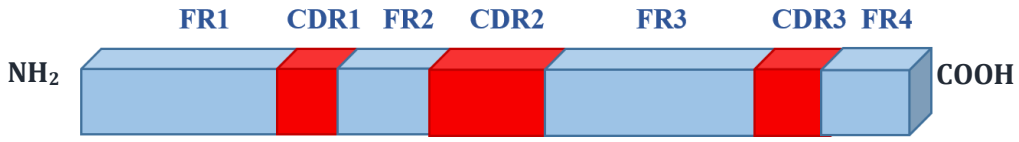


Figure 1.2- Scheme of the distribution of framework (FR) and hypervariable (CDR) regions in the variable domains of an antibody. Four FR regions (FR1, FR2, FR3 and FR4) interspersed by three CDR regions (CDR1, CDR2 and CDR3) are shown. The FRs in light blue and the CDRs in red (Adapted from Paul[19]).

1.1.2 Function

Antibodies are produced by clonal lines of B-lymphocytes that respond specifically to a unique antigen and, as part of the humoral immune system, are circulating in the blood stream. They can contribute to the immunity in three ways by: i) neutralisation, binding to the cells to avoid the entry of harmful pathogen; ii) opsonization, coating the pathogen to stimulate its elimination by the macrophages and other cells; iii) lysis, triggering the direct destruction of the pathogen via stimulation of other immune responses such as the complement pathway[20].

1.2 Heavy chain antibodies (HCAbs)

In addition to conventional antibodies, camelids and sharks produce unusual antibodies that lack light chains and a constant domain of the heavy chain (CH1), for which they are smaller (95 kDa approximately) than any other antibodies. These functional antibodies that only have heavy chains are called Heavy Chain Antibodies (HCAbs). They have been reported for the first time by a group of researchers directed by Hamers-Casterman of the Université Libre de Bruxelles, when investigating the immunodefense of dromedaries and water buffalos against parasites[21]. In the study of affinity of the antibodies present in camel serum, two different IgG fragments were obtained; one of 170 kDa and another of 100 kDa. In the first, after its reduction, the heavy (fragment of 50 kDa) and light (fragments of 30 kDa) chains were recognised while, in the second only the heavy chains (fragments of 43kDa and 46kDa) were recognised, for which it was deduced that these latter isotopes lacked a light chain. Furthermore, the most surprising attribute of these "incomplete antibodies" was their maintained affinity for the antigen. Therefore, it was assumed that these antibodies present changes that replace or compensate for the lack of VL and CH1 in a way that it does not affect its recognition of the antigen.

1.2.1 Structure

HCAbs present a similar structure to the heavy chains of the antibodies, with constant and variable domains, differing mainly in the absence of the first constant domain CH1. They are of the IgG type. The variable domain of HCAbs, VHH (with another H to differentiate it

from the variable domain of the conventional antibody heavy chain) is the most important aspect of these heavy chain antibodies (Figure 1.3).

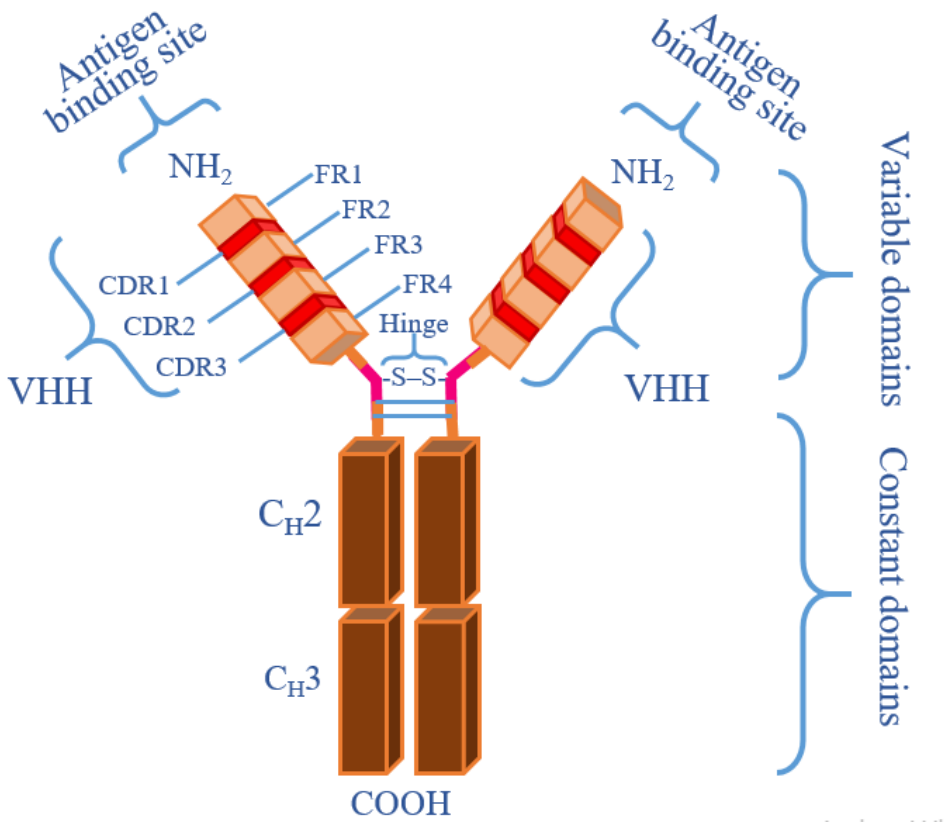


Figure 1.3- Structure of a heavy chain antibody (HCAb). In brown, the constant domains are shown and in salmon (FRs)/red (CDRs) the variable domain (VHH: Variable domain of the HCAb Heavy chain, with an extra H to differentiate it from the variable domain of the conventional antibody heavy chain). They are of the IgG type and are present in the serum of species of the family Camelidae, lacking a light chain and a CH1 region. (Adapted from Harmsen & De Haard[22]).

1.2.2 Variable domain of HCAbs (VHH)

The VHH domain is a polypeptide of about 110 to 140 amino acids (12 to 15 kDa) at the N-terminus where binding of the HCAbs to the antigen occurs. It can fold independently from the rest of the protein and maintain the capacity to bind with affinity and specificity to its antigen[22,23]. In other words, the VHH domain is the portion of HCAb that can interact with the antigen and behave as an independent monomer of the rest of the HCAb. Since it can be isolated without significant loss of its antigen recognition properties, independent VHH can be produced recombinantly[23,24]. Recombinant VHHs are also known as single domain antibody fragments, nanoantibodies or simply "nanobodies" as they were initially called.

1.2.2.1 Structure

The VHH has a secundary structure composed of 9 β-strands connected by loops, as the VH domains of antibodies with some specific changes in key positions that allow the appropriate folding and that favours the stability of the protein in the absence of the light chain and the region CH1[25]. They have 3 hypervariable loops, the CDR (CDR1, CDR2 and CDR3) with a high proportion of amino acids variability in one position for being involved

in the antigen recognition and 4 framework regions (FR1, FR2, FR3 and FR4) with sequences of amino acids more stable than the CDRs because they form the core structure of the VHH. By possessing a β-sheet structure, it serves as a scaffold to maintain the CDRs in contact position with the antigen (Figure 1.4).

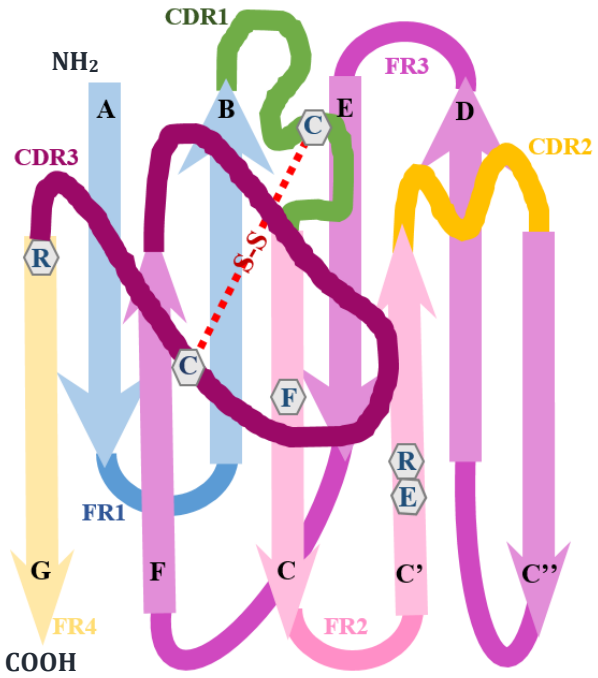


Figure 1.4- Scheme of the variable domain (VHH) of the HCabs. Composed of 9 β-strands: A, B, C, C', C", D, E, F, G and 8 loops. The two β sheets are organized in four antiparallel β-strands (A, B, D and E) and five antiparallel β-strands (C, C', C", F and G) that are connected by a disulfuric bridge between the cysteines (C) present in the CDR1 and the CDR3 (in red). The CDR1, CDR2 and CDR3 connect to the β strands B-C, C'-C" and F-G respectively. CDR3 possesses more variability, length and mobility among the CDR for which it is considered to be the major contributor to the binding specificity between the antigen and the antibody. The region FR1 is shown in light blue, CDR1 in green, FR2 in pink, CDR2 in yellow, FR3 in lilac, CDR3 in maroon and FR4 in melon. The phenylalanine (F) at position 42 can form the *Stretched twisted turn* structure. The W103R substitution, as well as G44E and L45R decrease the hydrophobicity and allow the stability of the VHH. (Adapted from Muyldermans[26]).

To maintain the stability of the VH domain structure, the conventional antibodies present a hydrophobic FR2 very conserved that interacts with the VL domain. The FR2 of the VHH presents amino acids substitutions as V37F, F37Y, G44E, L45R, W47G, V47F, G49Q, L50A or I51V[22,23,27-29]. These substitutions of the amino acids involved in the formation of the hydrophobic interface with the VL domain of conventional antibodies, change the hydrophobicity of the region FR2 making it hydrophilic. The hydrolific FR2 of the VHH allows their exposition to the medium, in the absence of the light chain, and avoids the interaction with other hydrophobic regions that could cause the loss of the correct folding.

Other substitutions related to the VHH stability are: L11S, P14A and A83P. These have been reported as mainly present in camels, while in other camelids they are not maintained in all the sequences[30]. Furthermore, W103R substitution decreases the hydrophobicity in the VHH and is determinant in its nature of single domain. In the VH domain, this hydrophobic residue of W is highly preserved as it interacts directly with the VL domain, (inexistent in the VHH domain), which means that this substitution is crucial

for stability of the VHH and is distinctive for them. In addition to FR2, the VH domain presents a slightly hydrophobic region formed by amino acids that interact with the CH1 domain. However, the VHH due to not having contact with a CH1 domain, substitutes some of the amino acids of this hydrophobic region, by hydrophilic residues[23,31] as in the case of the L11S substitution.

The differences between the VH and the VHH are also evidenced in the size and variability of the CDRs[31]. Unlike the VH, the N-terminus of the VHH CDR1 is more variable, which extends from position 27 to 35 while in VH it goes from position 31 to 35 very conservatively[30,32]. Furthermore, CDR1 is even more variable than CDR2, unlike the VH where CDR2 possesses more variability than CDR1[27].

The CDR3 of the VHH is more variable and larger than that of VH. The VHH possess a CDR3 of 16 to 18 amino acids[27,28] while that of the VH possess on average 12 amino acids[32,33]. The CDR3 is often stabilised by a disulfide bond with an additional cysteine at CDR1 or the FR2, generating a loop of the folded CDR3 through the old interface of FR2[34]. Furthermore, if at the 42 position of the FR2 a phenylalanine is present it can form a structure called "*Stretched twisted turn*"[35], in which the base of the loop twists and turns it toward the β-strand B C' of the FR2. This structure seems to give stability, thus avoiding contact of the hydrophobic phenylalanine residue with water molecules[26]. However, if it is an amino acid with an uncharged polar R-group, such as the tyrosine, present at position 42, the CDR3 is unfolded from the structure[36]. The majority of the VHHs have CDR3 loops similar in length to VH[37].

The substitution of amino acids in the VH domains of the conventional antibodies at key positions to resemble the properties and characteristics of the VHH was called camelisation[38] and improved thermal stability and reduced hydrophobicity in human antibody VH as well as in rabbit antibody VH (rVH)[39]. While the process of amino acids substitution in VHH to resemble the human VH, the "humanization" of the VHH, causes them loose stability, solubility and affinity for the antigen, unless a protein with VL domain sequence is added, after which they regain their stability and affinity[40,41].

1.2.2.2 Properties of the VHH proteins

The VHHs, the domains of antigen recognition of HCAbs, do not interact with any other domain and possess various unique properties that define them as powerful and promising research tools, distinct from the antibodies, offering a much more practical starting point[37,42]. Firstly, they are much smaller than antibodies, almost 10 times less. VHH is a single domain polypeptide of 15 kDa approximately; hence, they have the capacity to enter sites inaccessible for antibodies, an attractive attribute for using in medicine. Secondly, high specificity and affinity can be evaluated using a broad spectrum of approaches that vary from phage display to Surface Plasmon Resonance imaging (SPRi). Thirdly, they are structurally and biochemically stable, as well as highly soluble with low cross immunoreactivity and high homology to the human VH3 gene family. And most importantly, they can be cloned, genetically or chemically modified, and recombinantly produced in various cells with relative ease and low production cost at

large scale[37,42,43]. The bacteria expression systems allow the production of purified VHHs in quantities of milligrams per liter of culture, offering an unlimited supply system[24]. In that regard, VHH libraries are produced by recombinant DNA technology, that is, direct amplification of lymphocytes mRNA and cDNA amplification with VHH-specific primers, and then selected by phage display[44]. Thus, a large panel of peptides or proteins can be quickly and easily evaluated by applying the VHHs[45].

With all these attributes, the VHHs have an increasing use in both basic and medical research, as well as, in biotechnological applications[37,42,46]. In fact, VHHs are biotechnological tools that have been the subject of many studies. VHHs are produced *in vitro,* in their majority, for diagnosis and therapeutical purposes. Furthermore, the VHHs can be used to track in real time and manipulate the location and activity of the target proteins in eukaryote cells[47-49].

A longer CDR3, with convex shape of the antigen-binding site and small size, allows them to access epitopes that may be occult and not exposed to the conventional antibodies. Indeed, unlike conventional antibodies that usually detect linear or flat epitopes, many VHHs bind to concave and discontinuous epitopes that are only formed on the folded protein[50]. This property makes them valuable tools as probes to assess the conformational states of the target proteins and correlate the structural dynamic of the protein *in vitro* with the behavior of the proteins in the cells. Another notable property of VHH is their capacity to refold efficiently with complete restoration of their antigen binding properties after the thermal denaturalization. This property allows studying the VHH folding process[51].

1.2.2.3 Stability of recombinant VHHs

VHHs are reported to be stable soluble monomers since maintain their antigen binding properties and functionality in PBS (phosphate buffer saline) at 4° C for months and at -20° C for longer periods, even years. Moreover, they are able to maintain their antigen recognition and binding capacity after incubation for one week in PBS at 37° C[52]. They appear stable to mild chemical stress, since is reported that maintain their binding capacity in the range of 0.25 - 4 M urea[51]. Denaturation of the VHHs occurred at 6 M urea or at 2 -3 M guanidine chloride. In addition, the VHHs can withstand pressure above 400 Megapascals (MPa) without loss of functionality[53].

The VHHs are reported to show remarkable stability under temperature stress, retaining functionality even at 90° C[51]. However, in general, the VHH melting temperature or Tm (temperature at which half of the proteins in a system are denatured and half are folded[54]) is approximately 60° C[53-55]. Although, conventional antibodies and derivatives such as scFv (single-chain fragment variable) and FAB have similar Tm[56-59], the VHHs show greater thermostability and renaturation capacity than them. This is attributed to the fact that antibody denaturation exposes hydrophobic regions that mediate the interaction between heavy and light chains and these induce protein aggregation and precipitation, whereas the VHH do not possess these hydrophobic regions (as they lack a light chain) rather, these regions are hydrophilic[53]. Furthermore, the disulphide bridges present in

the VHH are suggested to stabilise the tertiary structure of the protein, making them more compact, so, the substitution of certain amino acids by cysteines were proposed to induce the formation of disulphide bridges that would increase protein stability[60-62].

On the other hand, cysteine substitutions at positions 54 and 78 appear to protect the VHHs from the proteolytic action of pepsin and chymotrypsin, possibly by blocking access to the cleavage sites due to a more compact structure. This property suggests the potential of VHHs for therapy by oral ingestion[63].

1.2.3 Antibodies vs nanobodies

The antibodies are the main biomolecules widely used as recognition agents to detect biomarkers and as analytical tools in clinical settings[43,64,65]. Antibodies are used in immunodiagnosis as binding agents of relevant antigens in a variety of platforms, including enzyme-linked immunosorbent assays (ELISA), immunoblots (Western Blot), immunocytochemistry and immunoprecipitation assays[66,67].

The binding properties of conventional antibodies are defined by the six complementarity determining regions (CDRs) situated inside the variable domains of heavy chains (VH) and light chains (VL). The large size, the composition of multichain and the requirements for the formation of essential disulphide bonds, complicate the production of the recombinant IgG. Smaller antibody fragments have been developed, Fab which has a molecular mass of approximately 50 kDa and consists of the light chain and a truncated heavy chain, that includes the two variable domains and scFv (25 kDa), the smallest practical derivative of a conventional antibody, which was produced by connecting VH and VL with an artificial peptide linker (Figure 1.5). Since the association between the variable domains of the light and heavy chains depends on hydrophobic interactions, the production and applications of the isolated individual variable domains of conventional antibodies are hampered by the poor solubility and aggregation of the protein. Although, it is possible, in principle, to modify the amino acid sequence of VH or VL to eliminate the hydrophobic regions on the protein surface[41], such work is labour intensive and can be misguided by the unintended effects of such mutations[24].

As the expression of the variable domain of the conventional antibodies is complex for its tendency to the formation of protein aggregates because of the hydrophobicity of the regions that mediate the VH-VL interaction[69], the discovery of the VHHs brought many advantages over antibodies, FAB or scFv. This is, an easier, more stable and economic production of the functional recombinant protein in a prokaryotic system such as *E. coli*. Furthermore, due to their relatively small size, it was supposed they could diffuse further through the membrane, which proposes them as potential intracellular therapeutic agents. Other advantages are the stability under unfavorable conditions, low immunogenicity, ease of sequence manipulation, humanization capacity, among others.

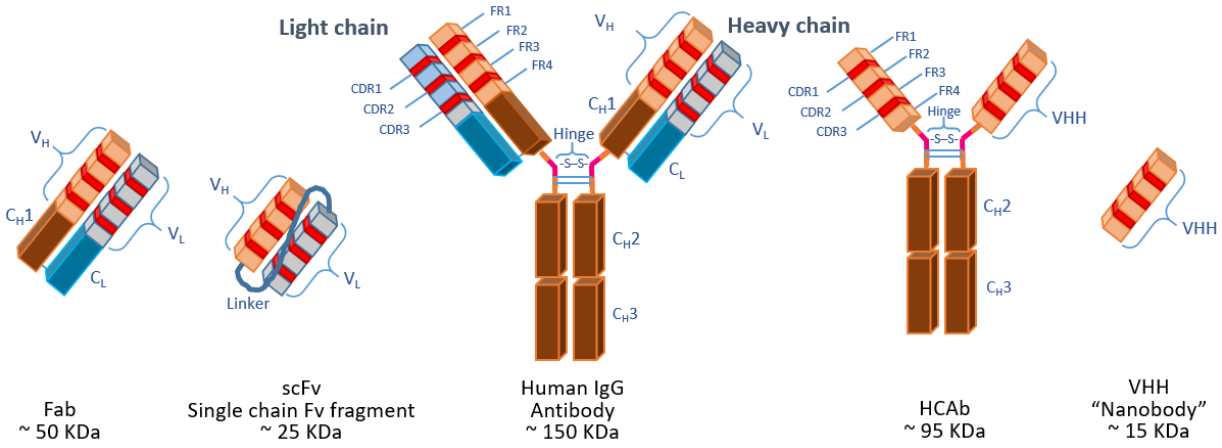


Figure 1.5- Comparison between antibodies type IgG, HCab and nanobodies (VHH). The structure of an IgG-type antibody is composed of heavy and light chains with constant (CH1, CH2, CH3 and CL) and variable (VH and VL) domains. On the left, the structures FAB (*Fragment antigen binding*) composed of the constant (CH1 and CL) and variable (VH and VL) domains, and scFv (*single chain Fragment variable*) composed only of the variable domains (VH and VL). The structure of the HCAb is simply composed of the heavy chains with constant (CH2 and CH3) and variable (VHH) domains. On the right, the independent and functional variable domain of the HCab, the VHH (initially referred to commercially as nanobody. (Adapted from Boulenouar *et al*[68]).

1.3 CD105

CD105, also known as Endoglin, END, ENG, FLJ41744, HHT1, ORW or ORW1[70] is a type I membrane glycoprotein located on the cell surface. It is a type III accessory receptor for transforming growth factor β (TGF-β) that binds TGF-β1 and TGF-β3 isoforms, and other members of the TGF- β family, in human endothelial cells. It can form complexes with type I and II receptors (TGFβR-I and -II), and acts as a modulator of TGF-β interactions with its signalling receptors[71-73]. In 1993, at the 5th International Workshop on Human Leucocyte Differentiation Antigens (HLDA) in Boston, this protein was assigned as the Cluster of Differentiation 105 (CD105), due to it increasing its expression in the differentiation from monocyte to macrophage[74].

1.3.1 Structure

CD105 is a transmembrane homodimer complex of 180 kDa with subunits linked by disulfide bridges. The primary structure of each monomer of 658 amino acid residues encompasses a signal peptide of 25 amino acid residues, a great extracellular domain of 561 amino acid residues, a transmembrane region of 25 amino acid residues and a cytoplasmic domain of 47 amino acid residues rich in serine (S) and threonine (T)[75-79] (Figure 1.6).

The signal peptide that constitutes the first amino acid residues at the amino terminal (N-terminus), directs the protein from the endoplasmic reticulum and Golgi apparatus towards the cell membrane, where it performs its function. Once in the membrane, it is degraded to form the mature and functional protein[79].

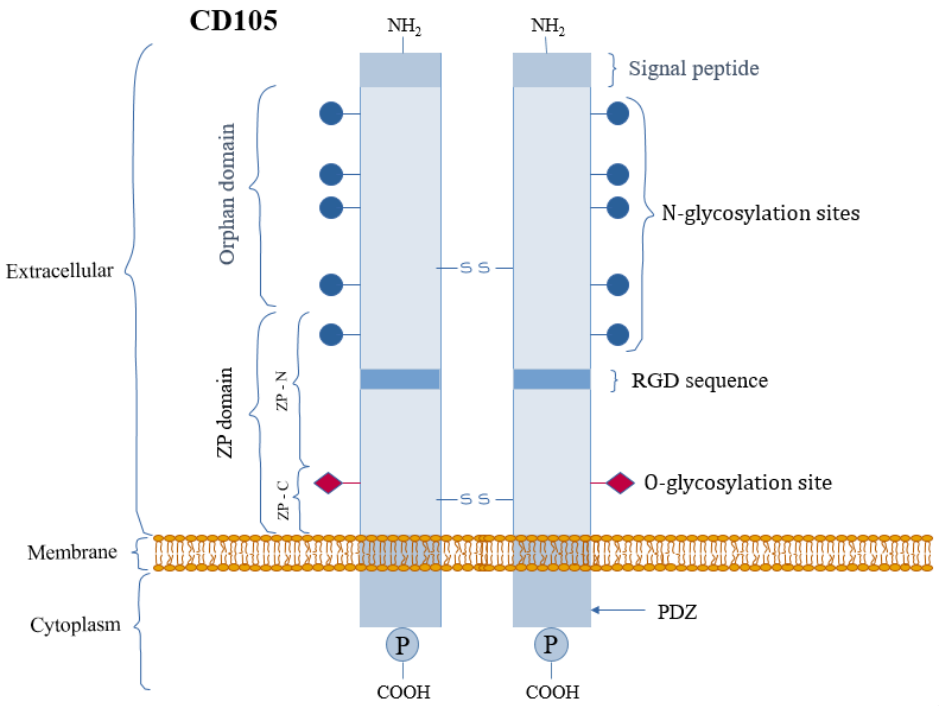


Figure 1.6- Scheme of the structure of CD105, homodimer membrane glycoprotein. From N terminus to C terminus are shown: the signal peptide, the extracellular domain consisting of the orphan domain and the Zona Pellucida (ZP) domain, the transmembrane domain and the cytoplasmic domain phosphorylated at S-T residues. The extracellular domain contains N-glycosylation (blue circles) and O-glycosylation (red diamond) sites and RGD sequence while the cytoplasmic domain contains PDZ binding motifs. The homodimers are linked by disulfide bonds between the C350 cysteine, located between the ZP and orphan domains, and the C582 cysteine located very close to the membrane. (Adapted from Lopez-Novoa & Bernabéu[70]).

The three-dimensional structure of the extracellular domain of CD105 is organised as a dome made of antiparallel monomers that locks a cavity at an extreme[70] and it was also determined the presence of two well-defined domains[77]. A domain at the N-terminus that comprises the amino acid residues E26-I359, the so-called orphan domain for not showing significative homology to other protein families and domains in the *GenBank*. This domain is responsible for binding the ligand BMP9 (bone morphogenetic protein 9), one of the members of the superfamily of TGF-β, reason for which it was also considered as a ligand binding domain[80,81]. The other domain, located close to the membrane and consisting of 260 amino acid residues with 8 highly conserved cysteine residues, common among the members of the extracellular protein family of the Zona Pellucida (ZP), is the ZP domain[77,82,83]. The ZP domain spans the amino acid residues Q360-G586 that is divided in two subdomains: a region that comprises the amino acid residues Q360-S457, the ZP-N subdomain; and a region that comprises the amino acid residues P458-G586, the ZP-C subdomain.

In the primary structure of the extracellular domain, the identification of consensus motifs suggest the presence of five N-glycosylation sites and one O-glycosylation site, located close to the transmembrane region, that can bind glycans and glycosaminoglucuronans to the extracellular domain. It also contains the tripeptide Arginine-Glycine-Aspartic acid (RGD) at position R399-G400-D401. The RGD motif is known as a cellular recognition site for numerous adhesive proteins present in the

extracellular matrix such as the integrin family (αvβ3)[84-88]. At least, C350, residue located in the linker region between the ZP and orphan domain (amino acid residues 338 to 362), and C582 residue, located adjacent to transmembrane region, are involved in the CD105 dimerization by disulfide bonds[70,76]. In addition, the ZP domain is implicated in the interaction of CD105 with TGF-β, TGFβR-II and ALK5 receptors[89,90]. This structural model of CD105 made it possible to define regions important for the functions of this protein[91].

The cytoplasmic domain is constitutively phosphorylated at S and T amino acid residues by serin and threonine kinases, including the type I and II TGF-β receptors[92,93]. Furthermore, it contains a Serine-Serine-Methionine-Alanine (S-S-M-A) consensus binding motif for the PDZ (Post synaptic density protein (PSD95), Drosophila disc large tumor suppressor (Dlg1), and Zonula occludens-1 protein (Zo-1)) domain[94] at the C-terminus that mediates the interaction of CD105 with several PDZ domain-containing proteins and the phosphorylation of CD105 at distal T residues[89,95].

Other kinases, besides the TGF-β receptors, can phosphorylate CD105 in this domain, such as the Protein Kinase C (PKC) which phosphorylate at S621, and the Src kinase (name derived from sarcoma) which phosphorylate at Y612 and Y614[79]. The phosphorylation of CD105 influences the subcellular localisation and cell migration, possibly by modulating the CD105 interactions with actin adhesive proteins such as zyxin and ZRP-1, thus, modifying the cytoskeleton organization and the adhesion properties of CD105-expressing cells. The conserved distal extreme of the cytoplasmic domain of CD105 interacts, aside from zyxin and ZRP-1, also with β-arrestin 2 which regulates CD105 internalization through the endocytic vesicles and Tctex2b. Therefore, the cytosolic interactions mediate functions, such as F-actine dynamics, the focal adhesion composition, protein transport through endocytic vesicles[70], and cellular processes such as cell division and migration involved in angiogenesis[96-99].

CD105 as a member of the ZP protein family shares a high sequence homology in the transmembrane and cytoplasmic domains with betaglycan, a type III TGF-β receptor. In fact, the cytoplasmic domains of these proteins are the most highly conserved region among homologue members in different mammalian species. The C-terminus cytoplasmic domains of human CD105 and betaglycan contain PDZ-binding motifs (S-S-M-A and S-S-T-A, respectively). Thus, in the TGF-β receptor system classification, both CD105 and betaglycan are considered as type III receptors, a class of non-signaling receptors, whose function is not well understood as serine/threonine kinases type I and II receptors that are involved in signal transduction via SMADs[72].

1.3.2 Isoforms

The human CD105 protein-encoding gene is located in the cytogenic band at 9q34.11 of chromosome 9[100]. The CD105 gene (*Eng*) contains 15 exons: the 2-8 exons encode the orphan domain, the 9-11 exons encode the ZP-N domain, the 12-14 exons encode the ZP-C domain and the 15 exon encode the cytoplasmic domain[79]. The expression of CD105 is under control of a proximal promoter that lacks TATA and CAAT boxes but contains GC-

rich regions that join the Sp1 transcription factor, which participates in the basal transcription. The transcript described is 3.4 kbp[70].

Two isoforms of CD105 are generated by alternative splicing, the predominantly expressed long isoform, L-CD105 (658 aa), and the short isoform, S-CD105 (625 aa), both have been detected in human and mice tissues[101]. S-CD105 is generated by a retention of the intron located between the exons 13 and 14, in an alternative splicing process that involves the factor ASF/SF2 with the minor spliceosome. The retention of this intron introduces a premature stop codon, which gives rise to a shorter protein[76,102,103]. In humans, S-CD105 and L-CD105 proteins differ in their cytoplasmic tails sequences that contain 14 and 47 amino acids respectively, with a tract of only 7 specific residues exclusive to S-CD105. As the cytoplasmic domain of S-CD105 is shorter, it does not contain neither the PDZ domain of interaction with other proteins nor other potential phosphorylation sites as S621, for which its phosphorylation capacity is less than L-CD105[79,92]. Since L-CD105 is the predominantly expressed isoform, most functional studies refer to this isoform.

In addition to the membrane-bonds CD105, the proteolytic action of the matrix metalloproteinase 14 (MMP-14) also known as membrane type I-MMP (MT1-MMP) over full-length CD105, give rise to a 65 kDa soluble form of CD105 that is released to the extracellular, the soluble CD105 (sCD105) (Figure 1.7)[70,104,105].

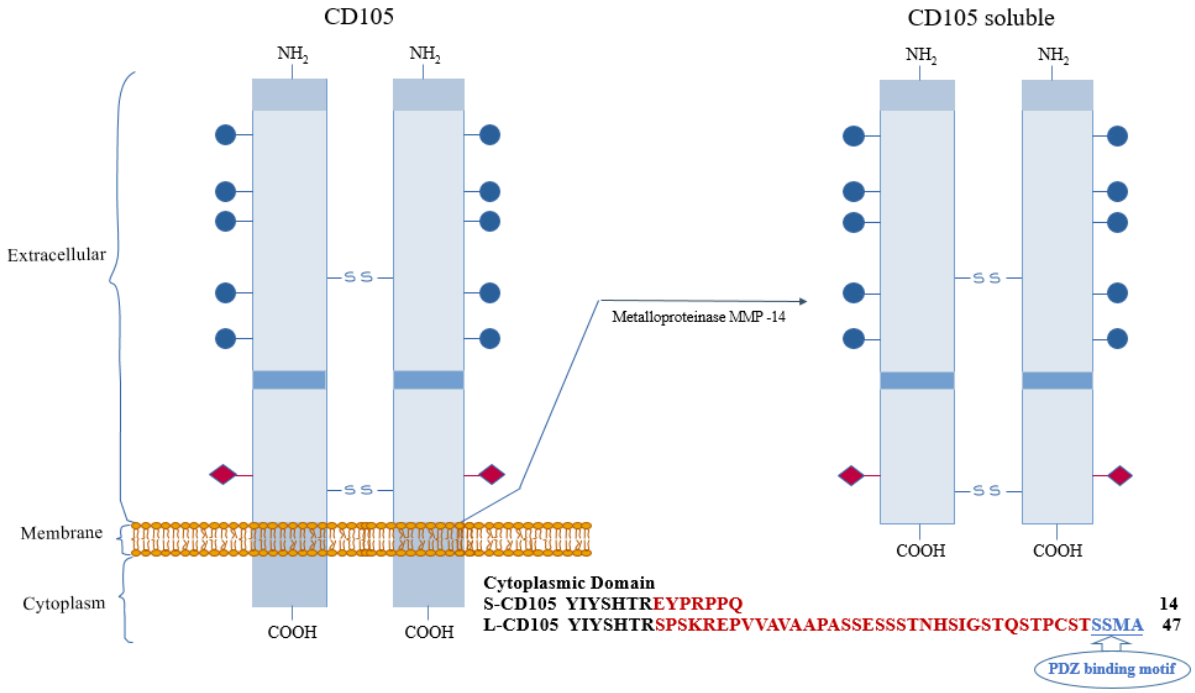


Figure 1.7-.Types of CD105. a) Isoforms of membrane-bound CD105, the long form (L-CD105) differs from the short form (S-CD105) in the length and amino acid sequence of the cytoplasmic tail that contain 14 and 47 amino acids respectively. The specific residues exclusive to S-CD105 (only 7) and to L-CD105 are shown in red, S-CD105 not contain neither the PDZ domain nor other potential phosphorylation sites as S621, b) CD105 in free form or soluble CD105 (sCD105). The 65 kDa sCD105 is released into the extracellular by the proteolytic action of the matrix metalloproteinase 14 (MMP-14 or MT1-MMP) at position 586 of full-length membrane-bound CD105. The extracellular domain of CD105 and sCD105 are similar in size. (Adapted from Lopez-Novoa & Bernabéu[70]).

Even with four potential cleavage sites (G-L) at positions 299, 387, 453, 586 of the extracellular domain of CD105, MMP14 cleaves at position 586 (the closest to the membrane), making the size of CD105 soluble equivalent to the size of the extracellular domain of membrane CD105[106]. sCD105 can be found free or bound to other factors of TGF- β, such as TGF-β1, BMP9 and BMP10. At crucial moments, these factors can be sequestered by sCD105, preventing binding to their respective receptors, which inactivates angiogenic pathways and increases the production of enzyme nitric oxide synthase (eNOS)[78,80,106-109].

1.3.3 Expression

Towards the end of the 80's, CD105 was identified as a marker antigen for human endothelial cells by employing a monoclonal antibody denominated 44G4 produced against a B pre-cells leukaemic cell line[110,111]. Afterwards, it was shown that CD105 is expressed on hematopoietic cells such as B cells precursors[112], proerythroblasts[113], myeloid lineage cells, macrophages[114,115] and bone marrow stromal cells[116-119]. In addition, CD105 is expressed also in the syncytiotrophoblasts of the terminal placenta[120,121], in fibroblasts[110,111,122], in cartilage chondrocytes[123], in kidney mesangial cells[124,125], in liver stellate cells[126], in epidermal keratinocytes of hair follicles and in sweat glands[127]. CD105 is expressed at lower levels in vascular smooth muscle cells (VSMC)[128] and pericytes[129,130]. It can also be expressed in caveolae[131]. Caveolin-1 (Cav-1), the major protein of caveolae, colocalizes with CD105, and interacts and cooperates with TGF-β signalization in endothelial cells[131,132].

High level of CD105 is expressed in the vascular endothelium in the early stage (at fourth week of gestation) of development of the embryonic vascular system and endocardium[120,133,134] and continuous in arteries, veins and capillaries of the adult[135]. It is also expressed at high levels in cells involved in the maintenance of the structure and function of the blood vessels such as perivascular stroma fibroblasts[136] and the smooth muscle cells that surround the vessels[88,120], which are associated to angiogenesis.

The expression of CD105 increases in response to certain physiological or pathological processes, such as: the transition from monocyte to macrophage[114], the development of preeclampsia[121], in tumoral and metastatic processes as in ovary cancer[137], melanoma[73,127], and prostate cancer[138,139]. In atherosclerosis and wound repair in smooth muscle cells[129,140], blood cell-mediated vascular repair[141], myogenic differentiation in embryo development[142], erythropoiesis in hematopoietic progeny cells of bone marrow[143-145], ischaemia-reperfusion in kidney and heart[146], etc.

The presence of CD105 in almost all vascular cell types and its high expression in related processes points out the important role that CD105 plays in vascular physiology[79].

1.3.4 Function

CD105 is an auxiliary receptor for the TGF-β protein family that play a key role in different physiological processes such as development, cell proliferation, extracellular matrix synthesis, angiogenesis or immune responses and their de-regulation can result in

tumour development. Although TGF-β is a potent inhibitor of cell proliferation, as well as, an inducer of both apoptosis and extracellular matrix protein synthesis, CD105 expression can counteract these effects in various cells types. The important role of CD105 in the regulation of TGF-β-dependent vascular remodelling and angiogenesis has been demonstrated[70,72,79,88].

CD105 binds with high affinity various members of the TGF-β family, including TGF-β1 and TGF-β3 (but not TGF-β2), activin-A, BMP-2, BMP-7 and BMP-9. Among these, BMP-9 can bind to CD105 even in absence of type I and II signaling receptors. In turn, CD105 can also bind to type I and II receptors in absence of the ligand[89]. CD105 forms a protein complex with type I (ALK1 or ALK5) and type II TGF-β receptors and the ligand (TGF-β)[147] (Figure 1.8).

Increased CD105 expression on endothelial cells is related to: autoimmune diseases[148], epidermal lesions[149], endothelial damage[140,150], active angiogenesis processes[151], embryogenesis[144] and tumoral processes [152,153].

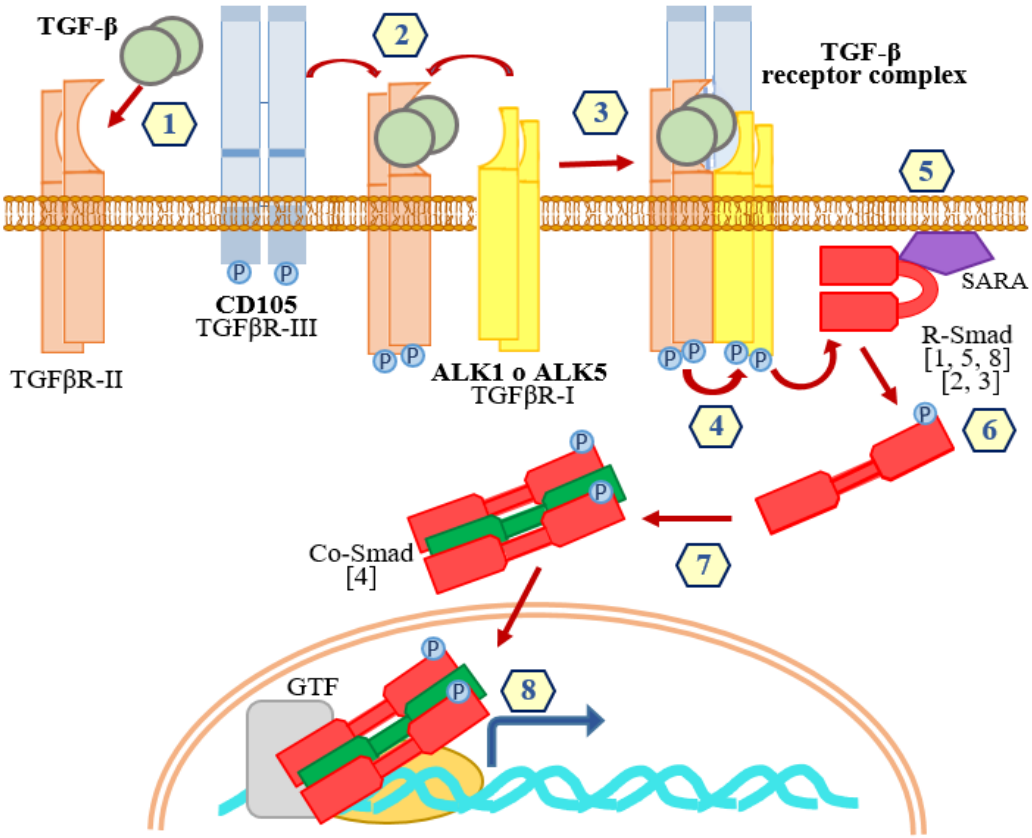


Figure 1.8- Schematic diagram of the CD105 participation in the TGF-β signalling from the membrane to the nucleus in the endothelial cell. The TGFβ receptor II (TGFβR-II) dimer in the endothelial membrane captures circulating TGF-β in blood (1) and autophosphorylates at S-T residues (2). The ALK5 or ALK1 dimer (TGFβR-I) and CD105 (TGFβR-III) are recruited to form the TGF-β receptor complex. (3). TGFβR-II phosphorylates S-T residues of the GS domain of TGFβR-I to activate the complex (4). R-Smads (receptor-dependent Smads) are attracted close to the membrane by anchoring proteins such as SARA (5) and are phosphorylated at the C-terminus, changing their conformation (6). This creates a binding site for Co-Smad (cooperating Smad) that can associate in the cytoplasm with R-Smads and enter the nucleus as a complex or can associate in the nucleus after translocation of activated R-Smads (7). The complex interacts with target genes (ID1 o PAI-1, among others), general transcription factors (GTFs) and specific coactivators or co-repressors, modulating gene expression in response to the cellular environment (8). (Adapted from Garrido[147]).

1.3.4.1 Equilibrium modulator between the TGF-β signalling (ALK1 and ALK5)

CD105 modulates TGF-β binding and signalling by association with ALK1 and ALK5 mediating endothelial activation/resolution states during angiogenesis[88,122,154,155]. These type I receptors activate the signaling pathways through Smad1, Smad5 and Smad8 (ALK1) or Smad2 and Smad3 (ALK5) to regulate, among others, the proangiogenic inhibitor of DNA binding 1 (ID1) or the target genes such as plasminogen activator inhibitor type 1 (PAI-1), respectively.

The ID1 protein inhibits the cell differentiation and activates the action of metalloproteases that degrade the extracellular matrix (ECM), favouring vessel ramification and growth[156,157]. While PAI-1 inhibits the proliferation and migration of endothelial cells and promotes stabilization of new vessels favouring the basal membrane formation and recruitment of pericytes and muscles cells[90,156,158,159] and thus favours the resolution and quiescence. Thus, the activation of the pathway through ALK1 induces endothelial proliferation and migration, whereas the activation of the pathway through ALK5 inhibits cell proliferation and migration[79] (Figure 1.9).

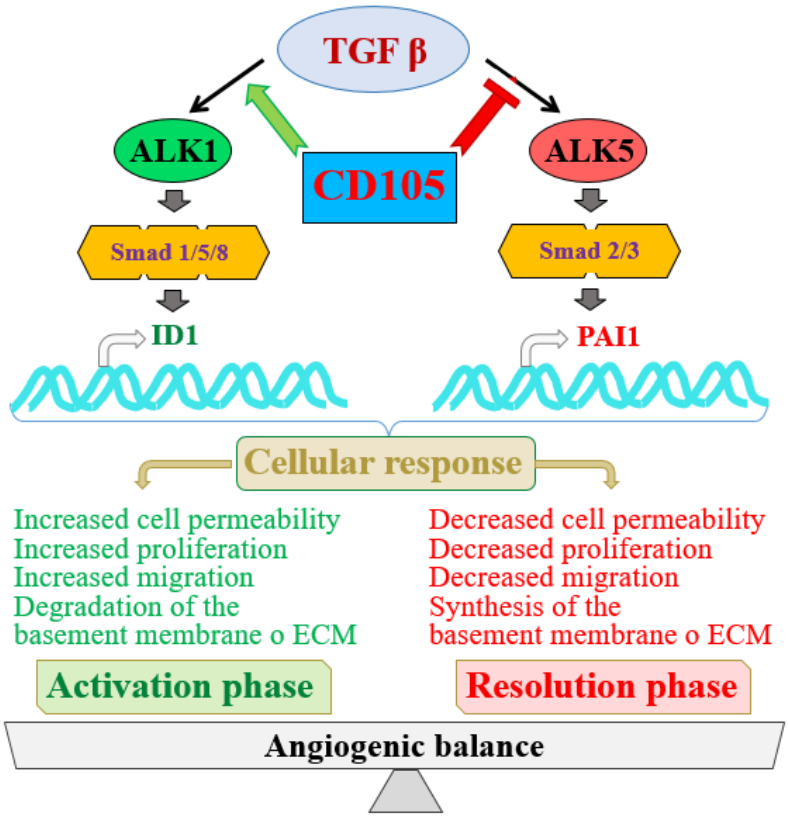


Figure 1.9- Schematic diagram of the balance of CD105-mediated angiogenesis in endothelial cells. When CD105 modulates TGF-β binding to the ALK1 receptor, Smads1/5/8 phosphorylation occurs and ID1 expression is induced, which stimulates angiogenesis activation processes (in green). Whereas if CD105 modulates the binding of TGF-β to the ALK5 receptor, Smads 2/3 are phosphorylated and PAI-1 expression is induced (among others), inhibiting activation processes, which favours the endothelial resolution-quiescence phase (in red). The increase in CD105 favours the ALK1 pathway (green arrow) and disfavours the ALK5 pathway (red truncated arrow). The effect on angiogenesis will depend on the balance between both signalling (ALK1 or ALK5) according to the requirements of the environment at a given cellular moment. (Adapted from Gallardo-Vara[79]).

The balance between ALK1 and ALK5 signalling in the endothelial and vascular smooth muscle cells plays a crucial role in the regulation of vascular remodelling and in the maturation and activation of angiogenesis[108], although the exact molecular mechanisms are still being studied. In the ALK1/ALK5 balance, CD105 is essential to promote TGF-β1/ALK1 and BMP-9/ALK1 signalling in endothelial cells. In addition, CD105 inhibits TGF-β/ALK5/Smad3-mediated cellular responses and enhances ALK5/Smad2-mediated responses. CD105 also enhances BMP-7 signalling through Smad1/Smad5 pathway in myoblasts[70]. In absence of CD105, the ALK5 pathway predominantes and thus maintains the quiescent state of the endothelium. Otherwise, in the presence of high levels of CD105, the ALK1 pathway is stimulated, and the ALK5 pathway is indirectly inhibited, promoting the activation state of the angiogenesis[160]. Hence, CD105 is a critical modulator of the balance between ALK1 and ALK5 signalling[79].

1.3.4.2 Angiogenesis

CD105 expression increases during angiogenesis, playing a fundamental role in this process[152]. Lack of CD105 is enough for the blood vessels to fail to develop and when CD105 expression is altered, the vessels develop little and poorly, and the angiogenesis is slow. CD105 knock-out mice (mice with CD105 inactive by homozygous inactivation of the gene) die at a very early stage of their fetal development (mid-gestation) by defective angiogenesis and severe cardiovascular abnormalities. Whereas CD105+/- heterozygotes mice were viable, although develop THH1-like vascular abnormalities[161-163]. On the other hand, its excess is characteristic in tumours, where many vessels are formed but of poor quality. The amount of CD105 is determinant for the formation of vessels in normal quantity and quality[164,165].

Experiments with transgenic mice Eng-iKO^{e} (Endothelial-specific endoglin knockout) with the CD105 gene knockout) showed that the mice had a loss of identity to specific markers in retinal arteries and veins after the formation of arteriovenous malformations (AVM), and also that AVMs are developed when an angiogenic stimulus is combined with CD105 depletion[166]. Thus, CD105 is not only related to the regulation of angiogenesis and vascular remodeling but also to an appropriate response of endothelial cells proliferation and possibly to the specification of arteries and veins[79].

In addition, CD105 modulates the release of the angiogenic factor VEGF (vascular endothelial growth factor)[167,168]. And it also regulates the expression and activity of the endothelial enzyme nitric oxide synthase (eNOS), which is involved in both angiogenesis and vascular tone[169,170]. CD105 has also been related to hypertension because causes vessel dilation due to the effect on the nitric oxide production[171].

1.3.4.3 Interaction with other factors that promote signalization

CD105 in presence of type I and II receptors binds TGF-β1, TGF-β3, activin, BMP2 and BMP7 and modulates TGF-β signalling (as described previously) by interactions with these receptors. However, CD105 can interact directly with BMP9 and BMP10[78-81]. The effects that mediate BMP9/BMP10/ALK1 on endothelial cells have yet to be determined as pro- and anti- angiogenic roles have been proposed[172-176]. It is suggested that BMP9

and BMP10, at the range 0,5-15 ng/mL in blood, may act via ALK1 as suppressors of cell migration and proliferation, thus, maintaining the quiescent state of the endothelium[177,178].

1.3.5 Diseases and CD105

1.3.5.1 Tumors and Cancer

TGF-β has a paradoxical role not only in the endothelial cell, where it can exert opposite effects depending on whether it interacts with ALK1 or ALK5, but also in cancer. It is a tumour suppressor in the pre-malignant phase of carcinogenesis, as it inhibits cell growth and induces apoptosis or differentiation, but in cancer cells, which have lost their growth inhibition, TGF-β modulates cell invasion, angiogenesis, immune regulation or interactions between tumour cells and their microenvironment, making them more malignant[70].

CD105 is currently considered a modulator of the TGF-β response with important functions in cancer. As it is highly expressed on the vascular endothelium associated to tumors, it is important for the prognosis in selected neoplasia, and is a potential candidate as a main vascular target for antiangiogenic cancer therapy[151]. CD105 expression on tumour cells seems to play an role in cancer progression, influencing cell proliferation, motility, invasiveness and tumorigenicity[179,180]. Its role in tumour cells is distinct from its role in tumour vasculature. Inside cancerous cells, CD105 would act as a tumour suppressor[80,138]. However, in endothelial cells, it shows a proangiogenic role by favouring tumour growth in solid vascularized tumours[70,181]. CD105 gene promotor contains HRE (hypoxia response elements) sequences that can bind to HIF1 (hypoxia inducible factor 1), increasing CD105 expression in response to hypoxia, a typical condition of solid tumors[88,182] and hence favors tumor growth by the stimulation of angiogenesis.

Regarding the soluble form of CD105, an association between the increment of sCD105 in blood and the severity of certain tumoral disorders such as metastatic breast cancer, colorectal cancer, lung cancer, acute myeloid leukemia and prostate cancer has been reported[180-184]. The increased level of sCD105 may be explained by the activation of the angiogenesis and the inflammation generated in these pathological conditions that induce expression and activity of metalloproteases (responsible of facilitating migration processes) that consequently cleave membrane-bound CD105 releasing sCD105 into circulation[79].

sCD105 can be used either as a prognosis marker in oesophagus[185], stomach[186] and ovary tumors[187], and as a diagnosis marker in prostate cancer, where high levels of sCD105 in urine correspond to the most advanced stages of tumour progression[188,189]. However, in general, sCD105 is considered as a negative prognostic marker in certain types of cancer. In a study, breast cancer patients undergoing chemotherapy decreased sCD105 levels as a response to treatment, hence, it not being a recommended prognostic marker for this disease in patients with treatment[181].

1.3.5.2 Hereditary haemorrhagic telangiectasia type 1 (HHT1)

The Hereditary haemorrhagic telangiectasia (HHT), also known as Rendu-Osler-Weber syndrome, is an autosomal dominant multisystemic dysplastic vascular disorder. This syndrome is characterized by genetic alteration of the blood vessels that leads to the formation of abnormal vascular structures that produce: telangiectasias (small dilated blood vessels), that may appear close to surface of the oral mucosa, nasal septum, gastrointestinal tract or finger skin; AVMs, in organs like lung, brain and liver or direct connections that predispose to arteriovenous communication, epistaxis (recurrent and spontaneous bleeding) bleeding or haemorrhages and anemia [190-193]. Thus, for the diagnosis of HHT, based on the Curaçao criteria, epistaxis, telangiectasias, AVMs in internal organs and family history are considered.

HHT is genetically heterogeneous, as different genes are involved in its pathogenesis, which may present mutations that determine different forms or types of a same disease. The mutations in the CD105 gene (*ENG*), located on the long arm of human chromosome 9 (9q33- q34.1) are responsible for the hereditary haemorrhagic telangiectasia type 1 (HHT1)[194-197]. Mutations in the ACVRL1/ALK1 gene, located on the long arm of chromosome 12 (12q11-q14) originate the HHT type 2 (HHT2)[198,199]. And mutations in chromosomes 5 and 7, whose genes have not yet been identified, are related with the origin of the HHT type 3 (HHT3) and the HHT type 4 (HHT4) respectively[200-202]. In addition, mutations in the MADH4/SMAD4 gene give rise to HHT-associated familial juvenile polyposis[203] and mutations in BMP9/GDF2 produce a syndrome similar to HHT or HHT5[204]. The mutations in the *ENG* or ACVRL1/ALK1 genes, both belonging to the TGF-β/BMP9 signalling are the most frequent in patients con HHT (over 90%)[205]. Currently, there are more than 500 mutations in the CD105 gene that affect the integrity of the protein domains and have been reported and described in the public database of the University of Utah, USA (http://www.arup.utah.edu/database/ENG/ENG_welcome. php). Most of these mutations are deletions, insertions and splicing defects, which would produce a truncated protein, that, if expressed, it would be retained intracellularly, leading to haploinsufficiency of CD105, the key protagonist in HHT1[206,207].

The loss of CD105 expression in HHT1, produces a disorder in the endothelium, even when endothelial cells proliferation and migration are activated, due to its indirect influence over determined proteins or the expression of one or other genes,. This leads to vascular weakness resulting in the disappearance of most of the capillary net, which increases blood flow and pressure, and produces a widening of the connected vessels, formation of AVMs and subsequent characteristic bleeding[79,166,208,209]. Despite knowledge of the symptoms and genes involved in the disease, there is no definitive cure for HHT, with only palliative measures and drug trials available. There are two strategies, intervention at "ground zero" to minimise by iron (Fe) and blood transfusions to counteract anaemia. Among the latter, there are three different strategies, favouring coagulation with antifibrinolytic agents, increasing ENG and ALK1 transcription with specific oestrogen receptor modulators, antioxidants, or immunosuppressants and impairing the abnormal angiogenic process with antibodies or blocking drugs.

1.3.5.3 Antagonism of CD105: Senescence, Hypertension and Preeclampsia

Exist three proven types of CD105[210], L-CD105, S-CD105 (with opposite effects) and sCD105 (Figure 1.7). The long form, L-CD105 has the effects generally attributed to CD105 and the short form, S-CD105 the opposite

As the cells age, there is an in S-CD105 increase would cause a reduced dilatation capacity of blood vessels due to decreased eNOS expression and increased COX2 (cyclooxygenase 2). Elevation of S-CD105 would contribute with the vascular pathology of ageing. In transgenic murine models that overexpressed S-CD105 was observed alteration in the vascular homeostasis by imbalance in nitric oxide produicing hypertension[102]. In endothelial cells, S-CD105 inhibits the ALK1 signaling pathway, which leads to inhibition of angiogenesis due to the non-expression of ID1. Instead, S-CD105 has a preferential interaction with ALK5, which in turn decreases cell proliferation and increases the synthesis of the extracellular matrix by induction of PAI-1 expression. In L6E9 myoblasts, S-CD105 uses the similar signiling patway leading to the typical fibrosis during cellular aging[102,211]. S-CD105 expression is increased in endothelial cell senescence[102] and in response to oxidative stress and ageing in macrophage[212], and mouse liver and lung cells[101].

CD105 soluble (sCD105) seems to antagonize with CD105 membrane-bound (L-CD105) since it shows that inhibits angiogenesis *in vitro*[106,107] and it can induce vascular permeability and hypertension *in vivo* [104,107,213]. Utilizing transgenic mouse models overexpressing sCD105, it has proven that the increment of sCD105 does not influence the expression of markers such as ICAM-1 (intercellular adhesion molecule 1) and VCAM1 (vascular cells adhesion molecule 1) and the eNOS in aorta, pointing out to other factors that may cause endothelial disorder in addition to sCD105[214-216]. Furthermore, certain studies have shown that the increment of sCD105 can cause a local inflammatory response, and increment of NFκB and IL6, inflammation-related factors[217]. The increment of sCD105 level in blood seems associated to vascular pathologies such as preeclampsia, hypertension, systemic sclerosis, dilated cardiomyopathy, acute myocardial infarction and atherosclerosis[171,213,218-220]. Also, sCD105 is associated with inflammatory and autoimmune pathologies such as multiple sclerosis, psoriasis, and rheumatoid arthritis[149,221].

Preeclampsia is a serious complication during pregnancy that is characterized clinically by hypertension (systolic/diastolic pression >140/90 mmHg) and proteinuria (≥300 mg of protein per day) after 20 weeks of gestation[222,223] and cellularly by increased serum levels of sFlt1 (soluble form of the receptor of VEGF)[224] and sCD105[225]. High circulating levels of sCD105 are considered as a marker for the diagnosis of preeclampsia in the early stages of this disease[79]. The anti-angiogenic capacity of sCD105 is related to the severity of this disease. sCD105 induces vasoconstriction and damage to the endothelium and inhibits vasodilation suggesting a key role in preeclampsia, including hypertension[222,226,227].

On the other hand, studies have associated the increment of sCD105 with hypoxia, where the high levels of reactive oxygen species (ROS) produced under these conditions promote the formation of oxysterols from cholesterol. These oxysterols stimulate the expression of the metalloprotease MMP14 at the membrane level, which results in the proteolytic cleavage of CD105 and the release of sCD105 to the extracellular medium, increasing its level in blood[104].

To understand the role of sCD105 in the physiopathology of preeclampsia and other cardiovascular diseases such as hypertension, where sCD105 levels are increased, a study was conducted to identify possible mediators of sCD105 activity. It was proven that sCD105 induces BMP4 expression *in vitro* and *in vivo*, and that sCD105-induced blood pressure elevation was abolished in the presence of Noggin BMP4 inhibitor in sCD105-overexpressing mice, suggesting that BMP4 is a downstream mediator of sCD105. Furthermore, serum levels of sCD105 and BMP4 are positively correlated in pregnant women with or without preeclampsia[171]. This has opened up the picture for directing studies to elucidate the identity of the factors involved in sCD105-mediated processes as well as the mechanism of its functions. Research to reveal new ways to comprehend the mechanisms involved in CD105-associated pathologies are in constant development. The growing attractiveness of CD105 is obvious, so, generating tools to enable its investigation and study is critical.

1.4 Background

The CD105, dimeric glycoprotein of 180 kDa, binds the TGF-β1 and TGF-β3 and participates in the regulation of the angiogenesis[78]. It is an important marker of endothelial cells and also mesenchymal stem cells (MSC)[228,229] that are pluripotent cells capable of differentiation in osteocytes, in chondrocytes, adipocytes, myoblasts among others[230]. CD105 plays also an important role in the fibrogenesis as evidenced by overexpression of this protein in biopsies from patients with renal and hepatic fibrosis[231,232]. sCD105, the soluble form produced by cleavage of the extracellular region, is found in the systemic circulation in different conditions of endothelium damage or inflammation[107] and its presence is associated with pre-eclampsia, hypercholesteremia and some types of cancer[151,227,233]. However, as most of the studies about CD105 have been directed towards membrane isoforms, that is, its role in angiogenesis or senescence, little is known about the exact regulation and role that this form of CD105 can perform outside the membrane. The need to clarify the mechanisms involved has led in recent years to new studies that have advanced the knowledge of CD105 traits in order to understand their functional role and establish a basis for the treatment or prevention of syndromes and diseases involved. Thus, to study in more detail the role of CD105 *in vivo*, in the development and in different angiogenic processes, murine animal models for the HHT1 have been generated. These studies have proven that CD105 is essential for early embryonic development, as homozygous mutation provoking a failure of vitelline sac vessels formation and failure of the embryonate progress, causing premature embryonic death. Currently, there are two main murine models: the heterozygotes for the CD105 gene (*ENG*) and the KO (Knock out) models. In *ENG* heterozygotes model, the penetrance

and phenotype of *ENG* is not complete; meanwhile, the KO models are inducible for *ENG*, both systemically and in certain cell lines, allowing an comprehensive study of the effect of *ENG* gene and CD105 protein loss[134,161,162,234-236].

In the case of sCD105, it is known to be incremented in several cardiovascular, inflammatory[237] and tumoral pathologies; and mainly, it has an anti-angiogenic function. In fact, sCD105 is generally and paradoxically considered a negative prognostic factor in some types of cancer, at the same time fulfilling an anti-angiogenic function in preeclampsia and in other cardiovascular diseases. To understand this contradiction, further studies are needed to evaluate the effect of sCD105 at vascular and tumoral level to elucidate its mechanism of action and find possible future therapies.

In this regard, the cellular and molecular biology of the vascular endothelium group of the Centre for Biological Research (CIB) of the Spanish National Research Council (CSIC) directed by Dr. Bernabeu, as well as, the group of researchers from the Department of Physiology and Pharmacology of the University of Salamanca in Spain under the direction of Dr. Lopez-Novoa have dedicated many years to the study of the structure, functions and mechanisms involving CD105. Both researchers, in 2012, described in detail CD105 «Endoglin» in the Atlas of Genetics and Cytogenetics in Oncology and Hematology, Card: ENGID40452ch9q34. Gallardo-Vara at CSIC, looked to deepen the research in the regulation of the sCD105 release during endothelial damage and its function over the endothelial cells and the angiogenesis processes and the vascular remodeling[79].

Since some of the medicines used for the treatment of the HHT have anti-angiogenic properties, the question was raised whether sCD105 could also have a beneficial action in the prevention and treatment of the arteriovenous malformations. It was concluded that high concentrations of sCD105 inhibit the endothelial migration and the angiogenesis by decreasing the expression of certain pro-angiogenic proteins released into the environment such as VEGF and FGF (Fibroblast Growth Factor). It was also determined that the transcription factor KLF6 can be considered as a “master regulator” during the process of the endothelial damage by activating genes implicated in the angiogenesis and vascular remodeling such as CD105, ALK1, IL-6 and MMP14 genes[79]. This research group, in search of a definitive cure for this disease, has described a full review[205] of the main strategies and the clinical trials developed with drugs to alleviate the HHT.

Another study evaluated the functional effect of high plasmatic levels of sCD105 on the physiopathology of the preeclampsia employing a pregnant mice model, in which the sCD105 levels in maternal blood during pregnancy replicated the human preeclampsia conditions. It was shown that the pregnant wild type mice carrying transgenic fetuses with expression of human sCD105 (fWT(hsEng+)) present elevated plasmatic levels of sCD105 with a temporary profile similar to that of human preeclampsia. Furthermore, the fWT(hsEng+) mice also present placental alterations comparable to those caused by the poor remodeling of the spiral arteries characteristic of preeclampsia. *In vitro* and *ex vivo* experiments, performed on a human trophoblasts cell line and human placenta explants, showed that sCD105 interferes with the trophoblasts invasion and the associated pseudo

vasculogenesis, a process in which the cytotrophoblasts change from an epithelial to an endothelial phenotype, being both events related to the spiral arteries remodeling. The findings provide a useful animal model for future research about preeclampsia and reveal a much more relevant role of sCD105 in preeclampsia than initially proposed[238].

On the other hand, the CD105 deficiency has shown leads to a reduced and defective angiogenesis, but little is known about the effect of its augmented expression, which is characteristic of various types of cancer. Angiogenesis is essential for tumors growth, so it is assumed that elevated levels of pro-angiogenic molecules, such as CD105, are associated with increased tumor growth that leads to a poor cancer prognosis. However, in a study it was shown that overexpression of CD105 does not stimulate the vascularization in various *in vitro* and *in vivo* models but on the contrary, maintains the endothelial cells in an active phenotype resulting in alteration of the proper endothelial stabilization and recruitment of mural cells. In a context of continuously increased angiogenesis, such as in tumors, the overexpression of CD105 results in altered vases with an incomplete mural coverage allowing blood extravasation. Furthermore, these alterations allow the intravasation of tumoral cells, the subsequent development of metastasis and, therefore, a worse prognostic of cancer[179].

The relevance that CD105 possesses in medicine continues to grow and not only due to its biological importance in fundamental endothelial cells processes[91], but also for its antigenic capacity, which enables its use as a biomarker to generate biotechnological tools that allow its detection for diagnosis, study and treatment of involved diseases. In the generation of tools, that is, designed proteins that bind to their targets with high affinity and specificity, the natural antibodies have been the inspiration. The six-hypervariable CDRs regions in the VH and VL domains determine their properties and binding capacity to the antigen, the target.

The potential of CD105 as a therapeutic target for human cancer has be established. In both *in vitro* and *in vivo* studies, it has been shown that immunotoxin- conjugated anti-CD105 monoclonal antibodies induce long-lasting suppressions of tumour growth and metastasis in the oncological animal models[151,239]. In a study, an anti-CD105 antibody was recognized as a good candidate for the generation of vascular or tumour targeting agents, as it showed a strong binding to human CD105-expressing endothelial cells (HUVEC), but not to a CD105- negative cell line (HEK293)[240]. In another study that explored the possibilities of using an anti-CD105 as an angiogenesis inhibitor, it was shown to inhibit the proliferation of human endothelial cells and the formation of capillary-like structures, offering high specificity against tumors with reduced side effects, and a lower risk of provoking drug resistance compared to conventional therapy[241].

The large size, multi-stranded composition and its complex folding make it difficult for functional recombinant antibodies to be generated. Smaller antibodies fragments, Fab and scFv, have been developed, the poor protein solubility and aggregation complicate the production and application of such variable domains in isolation even though modifications can be made to eliminate the hydrophobicity. The VHH have many

attributes that make advantageous and versatile research tools[37,42]. The generation of the VHHs is not complex and well stablished[242], which is why more and more research is relying and support on them. Even in the most recent research aimed at combating the COVID-19 pandemic. A study determined that the VHHs represent promising tools to prevent SARS-CoV-2 mortality when vaccines are compromised due to the continuous evolution of the receptor-binding domain (RBD) of the virus; furthermore, their use is a potential alternative to prevent the viral escape as they can recognise epitopes that are often inaccessible to conventional antibodies. The isolated anti-RBD VHHs were highly neutralising and according to researchers, rival the most potent antibodies against SARS-CoV-2 that have been made to date. They are capable of overcoming the SARS-CoV-2 mutations through two distinct mechanisms: a increased avidity for the ACE2 binding domain and the recognition of conserved epitopes that are usually largely inaccessible to human antibodies[243].

1.5 PhD objectives and cooperative context

At the Molecular Biology Unity of the Universidad Peruana Cayetano Heredia, the Nanobodies group directed by Dr. Espinoza develops research aimed towards the generation of recombinant VHH proteins from alpacas immunised with antigens of biomedical interest, such as Fas2[244] from *Fasciola hepatica* and human CD105. To obtain the VHH, libraries of phages carrying VHH cDNA have been performed through the phage display technique. Thus, Gushiken in 2016 obtained a library of 2.7 x 10^7 phages displaying anti-CD105 VHHs[245]. For this purpose, an alpaca was immunised with a lysate of cells overexpressing CD105, the T24 cell line from a urine bladder tumour, producing a polyclonal immune response to the T24 lysate and CD105. From RNA extracted from the animal blood, a cDNA library of 1.2 x 10^8 VHH sequences was first obtained, and then, by phage display, the anti-CD105 VHH carrying phages library. Two anti-CD105 VHHs were selected by phage-ELISA and the three-dimensional structures of the proteins were predicted from their sequences.

The need for further assays that allow an exhaustive screening of the cDNA library, in a way that more VHH sequences can be captured for their expression as anti-CD105 VHH proteins that broaden the repertoire of proteins with potential use in biosensors, was the motivation for the first part of this study.

The use and application of protein microarrays based on Surface Plasmon Resonance imaging (SPRi) in research has increased strongly in a few years due to its advantages, such as being fast and real time, having high specificity and sensitivity, requiring no labelling, using small volumes and being compatible with many types of biological materials[246-249]. And specially, for this study, because it is a suitable biosensing technology for testing of biomolecular interaction assays such as antigen-antibody interaction (in the case, antigen-VHH) and for analytes detection[250,251] using a microarray format on a biochip[252]. Microarrays of antibodies have been generated to determine the profile of proteins present in fluid and tissue samples of varied origin. The printed antibodies act as specific probes to the detection and capture of biomarker proteins, such as those

associated to many pathologies like cancer, psoriatic arthritis, infectious agents, signaling, pathogens, viruses, etc., in microliter scale reactions[253-258]. The construction of such microarrays presents critical points such as the stoichiometry of the proteins, microarray design, composition of the detection signal and data processing, factors that must be adjusted to the objectives of the assays and the nature of the sample[259-261]. Thus, more research is directed so.

Goode in 2015 conducted his work using nanobodies (VHHs) to make impedimetric immunosensors[262]. His research looked to contribute in the development of biosensors that could be regenerated and reutilized repeatedly to diminish the cost of testing and improve their adoption in diagnostic and field applications. It also sought to explore the role of the bioreceptor, an antibody mimetic receptor, in the generation of signals by electrochemical impedance spectroscopy (EIS), given that the recent evolution towards the use of these receptors (VHHs) may have deep implications for development of biosensors. The study determined that the regeneration and therefore the reutilization of impedimetric immunosensors was possible. In addition, it highlighted the importance of the spacing and the physical limitations at nanoscale that can influence the signal generation in immunosensors, so the bioreceptors were re-designed to include peptide spacer arms with terminal cysteines to allow both the oriented conjugation over the transducer surface and the precise positioning above it.

The development of biochips with protein microarrays for the cellular analysis through SPRi biosensors considers other relevant aspects. The Chemistry for Recognition and Studies of Biological Architectures (CREAB) group of the Molecular Systems and Nanomaterials for Energy and Health (SyMMES) research unit of CEA-Grenoble, Université Grenoble Alpes in France, directed by Dr. Roupioz, has dedicated many researches to the development of SPRi biochips and biosensors as new tools for the analysis, diagnosis and investigation of relevant cellular processes and cell detection.

These include research by Suraniti *et al*, which demonstrated the use of SPRi for the detection of cell binding to a functionalized gold surface, using B or T lymphocyte populations on an array of electrochemically grafted antibodies. This study showed that SPRi detects only the cells expressing cell markers bound by specific antibody arrays on a biochip, which is particularly convenient for the real time study of label-free cells that physically interact on a surface (interacting within a layer of 100-200nm above the surface). Regarding the functionalization of the active surfaces towards biological samples, it was demonstrated that the coupling of pyrrole to biomolecules and subsequent electropolymerizing on a gold surface is a versatile approach to biochips design. Another notable aspect was the design of the protein microarrays for the analysis of the cellular samples where thin gold films electrochemically arrayed with pyrrole-coupled antibodies can be coupled easily to SPRi detection. These conductive films placed on glass prisms enabled the SPRi bioanalysis[263].

Another research by Milgram *et al*, proposed a new approach to the study of cellular secretion of molecules, which was based on a biochip that allowed real-time SPRi

monitoring of antibodies produced by cultured B cells *in vitro*. For this purpose, a biochip microarrayed with antigenic proteins (hen egg lysozyme, HEL) and control proteins was incubated with hybridomas of B cells secreting HEL-specific antibodies (Ig). Protein secretion was monitored and detected in a shorter time than with an ELISA and the independence of the SPR response with respect to the density of the immobilised probes on the surface was noted. This biosensor proposed a method capable of assessing, in real time and without labelling, secretion activities of non-adherent live cells and also, because of its flexibility, suitable for multi-target detection, as several probes can be evaluated on a single biochip[264].

Alvarado in 2018 evaluated diverse protein functionalisation strategies for the generation of innovative SPR biochips[265]. He provided novel strategies including the development of an indirect photo-functionalization method that allowed the generation of protein microarrays in fully aqueous conditions, preserving the functionality of the grafted proteins. It was also designed and evaluated a microstructured SPRi biochip for real-time monitoring of the cell secretions. Finally, an evaluation of many functionalization strategies to the design of a nanostructured fibre bundle SPRi biochip was performed. Among the approaches, the generation of photoreactive self-assembled monolayers was the most suitable to this system. This nanostructured biochip may help for the development of *in vivo* biosensor systems.

The versatility of cell detection by SPRi opens several fields of application and allows the design of new miniaturised systems for detection analysis or kinetic studies of adhesion between cells and surfaces. For other hand, the growing attractiveness of CD105 is obvious, so generating tools to enable its research and study is critical. The development of a SRPi biosensor for the detection of the antigen-antibody (VHH) interaction between the receptor previously deposited on the gold surface (anti-CD105 VHH) and the complementary analyte (CD105), would allow to evaluate in real time and without the need for labelling, the recognition capacity of each anti-CD105 VHH to its antigen expressed on cells and not only in a free form. This could determine the application of these VHHs as tools to screen the ability of the anti-CD105 VHH candidates arrayed on a SPRi biochip for detecting and capturing CD105 on cells that express it, that is, as anti-CD105 nanoprobes. The generation of a SPRi biosensor for CD105 that will contribute to the establishment of basis for future biomedical research to aid in the detection, diagnosis or treatment of diseases or syndromes where CD105 plays an important role or for studies that seek to better understand the mechanisms and the way in which CD105 participates, was the motivation for the second part of this study.

1.5.1 Goal

To develop a proteomics platform “biochip” that allows the rapid real-time assessment and characterisation of alpaca VHHs specific against CD105 on cells by SPRi for the selection of the most efficient VHHs in detection and capture of CD105 and the application of such VHHs as nanoprobes in diagnostic and therapeutic research studies.

1.5.1.1 Specific goals

- Screening of a VHH cDNA library from alpacas immunised with a lysate of human bladder cancer cells (T24 cell line overexpressing CD105) for the selection of CD105-specific VHHs.
- Sequencing of anti-CD105 VHH cDNAs for the sequences analysis and *in silico* characterisation of recombinant VHH proteins.
- Subcloning of the anti-CD105 VHH sequences into the pET22b(+) vector (Qiagen) for expression and purification of the recombinant proteins by affinity chromatography on Ni-NTA Agarose matrix.
- Functionalisation of the selected recombinant anti-CD105 VHH proteins for their arraying on a biochip that allows the design and generation of the proteomics platform (protein microarray) for rapid, real-time detection of CD105.
- Recognition and characterisation of the antigen-antibody interactions of the anti-CD105 VHHs on the biochip with the antigen on cells that express it using SPRi to determine those with higher affinity and specificity that allow their application in research studies, diagnosis or treatment of diseases that involve CD105 expression.

1.5.2 Justification and Importance

The generation of VHH proteins that bind with high affinity and specificity to their target, overcoming many of the weaknesses of conventional antibodies, represent novel tools for biomedical or biotechnological applications at nanoscale. The membrane protein CD105, is a marker of human endothelial cells that is also expressed in other cells such as haematopoietic cells or bone marrow stromal cells and is a potential therapeutic target for human cancer. The design of a platform of highly specific anti-CD105 VHH proteins (VHH microarray biochip) as a recognition and capture instrument in an SPRi-based biosensor would provide a basis for future studies that require selective, real-time detection of CD105 in a medium, not only in free form (sCD105) but also on cells expressing it (membrane CD105). Studies that lead to a better knowledge of the biological role of CD105 in its various forms, to elucidate the mechanisms in which it is involved or its interaction with other proteins and to understand its influence on the regulation of angiogenesis, tumour formation or in diseases such as renal and hepatic fibrosis, pre-eclampsia and hypertension. Also, studies to detect CD105 for the diagnosis and treatment of the diseases involved or to prove its capacity as a therapeutic target for cancer. Last but not least, the highly specific anti-CD105 VHHs obtained could be evaluated for their application in tissue engineering, as a bridge in the attachment of osteoblast progenitor cells (where CD105 is an important marker[119]) to a scaffold, in specifically and without generating an immune response that allows them to partially or fully differentiate into tissue with growth factors.

References

1) Sanabria, V. & Landa, A. (2007) Anticuerpos: sus propiedades, aplicaciones y Perspectivas. Médicas UIS 20, 15-30

2) Litman, G. W., Rast, J. P., Shamblott, M. J., Haire, R. N., Hulst, M., Roess, W., . . . Amemiya, C. T. (1993). Phylogenetic diversification of immunoglobulin genes and the antibody repertoire. *Mol Biol Evol 10(1),* 60-72.

3) Coico, R. & Sunshine, G. (2013) Immunobiology - A Short Course. *Journal of Chemical Information and Modeling 53,* 1689-1699.

4) Woof, J. & Burton, D. (2004). Human antibody-Fc receptor interactions illuminated by crystal structures. *Nat Rev Immunol 4(2),* 89-99.

5) Padlan, E. A. (1994) Anatomy of the antibody molecule. *Mol Immunol 31(3),* 169–217.

6) Stanfield, R. L., Fieser, T. M., Lerner, R. A. & Wilson IA. (1990) Crystal structures of an antibody to a peptide and its complex with peptide antigen at 2.8 A. *Science 248(4956),* 712–719.

7) Ramos-Bello, D. & Llorente, L. (2009) Cincuentenario del descubrimiento de la estructura química de los anticuerpos. *Reumatol Clin 5(6),* 280–284.

8) Market, E. & Papavasiliou, N. (2003) V(D)J Recombination and the Evolution of the Adaptive Immune System. *PLoS Biology1(1),* e16.

9) Barclay, A. (2003). Membrane proteins with immunoglobulin-like domains--a master superfamily of interaction molecules. *Semin Immunol 15(4),* 215-223.

10) Putnam, F. W., Liu, Y. S. & Low, T. L. (1979) Primary structure of a human IgA1 immunoglobulin. IV. Streptococcal IgA1 protease, digestion, Fab and Fc fragments, and the complete amino acid sequence of the alpha 1 heavy chain. *J Biol Chem 254(8),* 2865-2874.

11) Huber, R. (1980) Spatial structure of immunoglobulin molecules. *Klin Wochenschr 58(22),* 1217-1231.

12) Porter, R. R. (1958) Separation and isolation of fractions of rabbit gamma-globulin containing the antibody and antigenic combining sites. *Nature 182(4636),* 670–671.

13) Awwad, S. & Angkawinitwong, U. (2018) Overview of Antibody Drug Delivery. *Pharmaceutics 10(3),* 83.

14) Knappik, A., Ge, L., Honegger, A., Pack, P., Fischer, M., Wellnhofer, G., . . . Virnekäs, B. (2000) Fully synthetic human combinatorial antibody libraries (HuCAL) based on modular consensus frameworks and CDRs randomized with trinucleotides. *J Mol Biol 296(1),* 57–86.

15) Lewis, S. M. & Wu, G. E. (1997) The Origins of V(D)J Recombination Minireview. *Cell 8(2),* 159–62.

16) Parham, P. (2014) The Immune System. New York: Garland Science (4th ed)

17) Braden, B. & Poljack, R. (1995) Structural features of the reactions between antibodies and protein antigens. *FASEB 9(1),* 9–16.

18) Sundberg, E. J. & Mariuzza, R. A. (2002) Molecular recognition in antibody-antigen complexes. *Adv Protein Chem 61,* 119–160.

19) Paul, W. E. (2013) Fundamental Immunology. Lippincott Williams & Wilkins (7th ed).

20) Ravetch, J. & Bolland, S. (2001). IgG Fc receptors. Annu Rev Immunol 19, 275-290.

21) Hamers-Casterman C., Atarhouch T, Muyldermans S, Robinson G, Hamers C, Bajyana Songa, E . . . Hammers, R (1993) Naturally occurring antibodies devoid of light chains. Nature 363(6428), 446–448.

22) Harmsen, M. M. & De Haard, H. J. (2007) Properties, production, and applications of camelid single-domain antibody fragments. *Applied Microbiology and Biotechnology 77(1)*, 13-22.

23) Muyldermans, S., Atarhouch, T., Saldanha, J., Barbosa, J. & Hamers, R. (1994) Sequence and structure of VH domain from naturally occurring camel heavy chain immunoglobulins lacking light chains. *Protein Eng 7(9),* 1129–1135.

24) Dmitriev, O. Y., Lutsenko, S. & Muylderman, S. (2016) Nanobodies as Probes for Protein Dynamics *in Vitro* and in Cells. *J Biological Chemistry 291(8),* 3767–3775

25) Holliger, P. & Hudson, P. J. (2005) Engineered antibody fragments and the rise of single domains. *Nat Biotechnol 23(9),* 1126–1136.

26) Muyldermans, S. (2013) Nanobodies: Natural Single-Domain Antibodies. *Annu Rev Biochem 82(1),* 775–797.

27) Vu, K. B., Ghahroudi, M. A., Wyns, L. & Muyldermans, S. (1997) Comparison of llama V(H) sequences from conventional and heavy chain antibodies. *Mol Immunol 34(16–17),* 1121–1131.

28) Harmsen, M. M., Ruuls, R. C., Nijman, I. J., Niewold, T., Frenken, L. G. J., De Geus, B. (2000) Llama heavy-chain V regions consist of at least four distinct subfamilies revealing novel sequence features. *Mol Immunol 37(10),* 579–590.

29) Maass, D. R., Sepulveda, J., Pernthaner, A. & Shoemaker, C. B. (2007) Alpaca (Lama pacos) as a convenient source of recombinant camelid heavy chain antibodies (VHHs). *J Immunol Methods 324(1–2),* 13–25.

30) Nguyen, V. K., Hamers, R., Wyns, L. & Muyldermans, S. (2000) Camel heavy-chain antibodies: diverse germline V(H)H and specific mechanisms enlarge the antigen-binding repertoire. *EMBO J 19(5),* 921–930.

31) Muyldermans, S. (2001) Single domain camel antibodies: Current status. *Rev Mol Biotechnol 74(4),* 277–302.

32) Kabat, E., Te-Wu, T., Foeller, C., Perry, H. & Goterman, K. (1991) Sequences of proteins of immunological interest. New York: Bethesda (5th ed).

33) Wu, T. T., Johnson, G. & Kabat, E. A. (1993) Length distribution of CDRH3 in antibodies. *Proteins Struct Funct Genet 16(1),* 1–7.

34) Desmyter, A., Transue, T. R., Ghahroudi, M. A., Thi, M. H., Poortmans, F., Hamers, R., . . . Wyns, L. (1996) Crystal structure of a camel single-domain VH antibody fragment in complex with lysozyme. *Nature Structural Biology 3(9),* 803-811.

35) Sircar, A., Sanni, K. A., Shi, J. & Gray, J. J. (2011) Analysis and modeling of the variable region of camelid single-domain antibodies. *J Immunol 186(11),* 6357–6367.

36) Kirchhofer, A., Helma, J., Schmidthals, K., Frauer, C., Cui, S., Karcher, A., . . . Rothbauer, U. (2010) Modulation of protein properties in living cells using nanobodies. *Nat Struct Mol Biol 17(1),* 133–138.

37) Wesolowski, J., Alzogaray, V., Reyelt, J., Unger, K., Juárez, K., Cauerhff, A.,. . . Koch-Nolte, F. (2009) Single domain antibodies: promising experimental and therapeutic tools in infection and immunity. *Med Microbiol Immunol 198(3),* 157-174.

38) Davies, J. & Riechmann, L. (1994) "Camelising" human antibody fragments: NMR studies on VH domains. *FEBS Lett 339(3),* 285–290.

39) Shinozaki, N., Hashimoto, R., Noda, M., & Uchiyama, S. (2018). Physicochemical improvement of rabbit derived single-domain antibodies by substitutions with amino acids conserved in camelid antibodies. *Journal of bioscience and bioengineering 125(6),* 654–661.

40) Conrath, K., Vincke, C., Stijlemans, B., Schymkowitz, J., Decanniere, K., Wyns, L., . . . Loris, R. (2005) Antigen binding and solubility effects upon the veneering of a camel VHH in framework-2 to mimic a VH. *J Mol Biol 350(1),* 112–125.

41) Barthelemy, P. A., Raab, H., Appleton, B. A., Bond, C. J., Wu, P., Wiesmann, C. & Sidhu, S. S. (2008) Comprehensive analysis of the factors contributing to the stability and solubility of autonomous human VH domains. *J Biol Chem.283(6),* 3639–3654.

42) Helma, J., Cardoso, M. C., Muyldermans, S. & Leonhardt, H. (2015) Nanobodies and recombinant binders in cell biology. *J Cell Biol 209(5),* 633–644.

43) Leow, C. H., Fischer, K., Leow, C. Y., Cheng, Q., Chuah, C. & McCarthy, J. (2017) Single Domain Antibodies as New Biomarker Detectors. *Diagnostics 7,* 52.

44) Derda, R., Tang, S. K., Cory Li, S., Ng, S., Matochko, W. & Jafari, M. R. (2011) Diversity of phage-displayed libraries of peptides during panning and amplification. *Molecules 16,* 1776-1803.

45) Chaikuad, A., Keates, T., Vincke, C., Kaufholz, M., Zenn, M., Zimmermann, B., . . . Müller, S. (2014) Structure of cyclin G-associated kinase (GAK) trapped in different conformations using nanobodies. *Biochem J 459*, 59–69.

46) Yang, E. Y. & Shah, K. (2020) Nanobodies: Next Generation of Cancer Diagnostics and Therapeutics. *Frontiers in Oncology 10,* 1182.

47) Broisat, A., Hernot, S., Toczek, J., De Vos, J., Riou, L. M., Martin, S., . . . Devoogdt, N. (2012) Nanobodies targeting mouse/human VCAM1 for the nuclear imaging of atherosclerotic lesions. *Circ Res 110,* 927–937

48) Caussinus, E., Kanca, O. & Affolter, M. (2012) Fluorescent fusion protein knockout mediated by anti-GFP nanobody. *Nat Struct Mol Biol 19,* 117–121

49) Rothbauer, U., Zolghadr, K., Tillib, S., Nowak, D., Schermelleh, L., Gahl, A., . . . Leonhardt, H. (2006) Targeting and tracing antigens in live cells with fluorescent nanobodies. *Nat Methods 3,* 887–889.

50) Rudolph, M. J., Vance, D. J., Cheung, J., Franklin, M. C., Burshteyn, F., Cassidy, M. S., . . . Mantis, N. J. (2014) Crystal structures of ricin toxin's enzymatic subunit (RTA) in complex with neutralizing and non-neutralizing single chain antibodies. *J Mol Biol 426,* 3057–3068.

51) van der Linden, R. H., Frenken, L. G., de Geus, B., Harmsen, M. M., Ruuls, R. C., Stok, W., . . . Verrips, C.T. (1999) Comparison of physical chemical properties of llama VHH antibody fragments and mouse monoclonal antibodies. *Biochim Biophys Acta 1431,* 37–46.

52) 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(3),* 521–526.

53) Dumoulin, M., Conrath, K., Van Meirhaeghe, A., Meersman, F., Heremans, K., Frenken, L. G. J., . . . Matagne, A. (2002) Single-domain antibody fragments with high conformational stability. *Protein Sci 11(3),* 500–515.

54) Rees, D. C. & Robertson, A. D. (2001) Some thermodynamic implications for the thermostability of proteins. *Protein Sci 10(6),* 1187–1194.

55) Pérez, J. M. J., 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(1),* 74–83.

56) Yasui, H., Ito, W. & Kurosawa Y. (1994) Effects of substitutions of amino acids on the thermal stability of the Fv fragments of antibodies. *FEBS Lett 353(2),* 143–146.

57) Shimba, N., Torigoe, H., Takahashi, H., Masuda, K., Shimada, I., Arata, Y. & Sarai, A. (1995) Comparative thermodynamic analyses of the Fv, Fab* and Fab and Fab fragments of anti-dansyl mouse monoclonal antibody. *FEBS Lett 360(3),* 247–250.

58) Young, N. M., MacKenzie, C. R., Narang, S. A., Oomen, R. P. & Baenziger, J. E. (1995) Thermal stabilization of a single-chain Fv antibody fragment by introduction of a disulphide bond. *FEBS Lett 377(2),* 135–139.

59) Welfle, K., Misselwitz, R., Hausdorf, G., Höhne, W. & Welfle, H. (1999) Conformation, pH inducedconformational changes, and thermal unfolding of anti-p24 (HIV-1) monoclonal antibody CB4-1 and its Fab and Fc fragments. *Biochim Biophys Acta Protein Struct Mol Enzymol 1431(1),* 120–131.

60) Saerens, D., Conrath, K., Govaert, J. & Muyldermans, S. (2008) Disulfide Bond Introduction for General Stabilization of Immunoglobulin Heavy-Chain Variable Domains. *J Mol Biol 377(2),* 478–488.

61) Hagihara, Y., Mine, S. & Uegaki, K. (2007) Stabilization of an immunoglobulin fold domain by an engineered disulfide bond at the buried hydrophobic region. *J Biol Chem 282(50),* 36489–36495.

62) Zabetakis, D., Olson, M. A., Anderson, G. P., Legler, P. M. & Goldman, E. R. (2014) Evaluation of disulfide bond position to enhance the thermal stability of a highly stable single domain antibody. *PLoS One 9(12),* 1–14.

63) Hussack, G., Hirama, T., Ding, W., MacKenzie, R. & Tanha, J. (2011) Engineered single domain antibodies with high protease resistance and thermal stability. *PLoS One 6(11),* e28218 .

64) Jovčevska, I. & Muyldermans, S. (2020) The Therapeutic Potential of Nanobodies. *BioDrugs 34,* 11–26.

65) Bannas, P., Hambach, J. & Koch-Nolte, F. (2017) Nanobodies and Nanobody-Based Human Heavy Chain Antibodies As Antitumor Therapeutics. *Front Immunol 8,* 1603.

66) Demerdash, Z. A., Diab, T. M., Aly, I. R., Mohamed, S. H., Mahmoud, F. S., Zoheiry, M. K., . . . El-Bassiouny, A. E. (2011) Diagnostic efficacy of monoclonal antibody based sandwich enzyme linked immunosorbent assay (ELISA) for detection of *Fasciola gigantica* excretory/secretory antigens in both serum and stool. *Parasites Vectors 4,* 176.

67) Fujiwara, K., Yoshizaki, Y., Shin, M., Miyazaki, T., Saita, T. & Nagata, S. (2012) Immunocytochemistry for vancomycin using a monoclonal antibody that reveals accumulation of the drug in rat kidney and liver. *Antimicrob Agents Chemother 56,* 5883–5891.

68) Boulenouar, H., Amar, Y., Bouchoutrouch, N., Faouzi, M. E. A., Cherrah, Y. & Sefrioui, H. (2020). Nanobodies and their medical applications. *Genet Mol Res 19(1),* GMR18452.

69) Ward, E. S., Güssow, D., Griffiths, A. D., Jones P. T. & Winter, G. (1989) Binding activities of a repertoire of single immunoglobulin variable domains secreted from *Escherichia coli. Nature 341(6242)*, 544–546.

70) Lopez-Novoa, J. M. & Bernabéu, C. (2012) ENG (endoglin). *Atlas Genet Cytogenet Oncol Haematol* http://atlasgeneticsoncology.org/Genes/ENGID40452ch9q34.Html. Revised July 2020

71) Santibañez, J. F., Quintanilla, M. & Bernabéu, C. (2011) TGF-β/TGF-β receptor system and its role in physiological and pathological conditions*. Clin Sci (Lond) 121(6),* 233-251

72) Jerkic, M., Rivas, J. V., Carrón, R., Sevilla, M. A., Rodríguez-Barbero, A., Bernabéu, C., . . . López Novoa, J. M. (2002) Endoglina, un componente del complejo de receptores de TGF-β, es un regulador de la estructura y función vascular. *Nefrología XX11,* 2.

73) Altomonte, M., Montagner, R., Fonsatti, E., Colizzi, F., Cattarossi, I., Brasoveanu, L. I., . . . Maio, M. (1996) Expression and structural features of endoglin (CD105), a transforming growth factor beta1 and beta3 binding protein, in human melanoma. *Br J Cancer 74(10),* 1586-1591.

74) Letarte, M., Greaves, A. & Vera S. (1995) CD105 (endoglin) cluster report. In Schlossman, S.F. *et al* editors. Leukocyte typing V: white cell differentiation antigens. (pp1756–1759). Oxford, UK: Oxford University Press.

75) Gougos, A. & Letarte, M. (1990) Primary structure of endoglin, an RGD-containing glycoprotein of human endothelial cells. *J Biol Chem 265,* 8361-8364.

76) Bellón, T., Corbi, A., Lastres, P., Cales, C., Cebrian, M., Vera, S., . . . Bernabéu, C. (1993) Identification and expression of two forms of the transforming growth factor beta-binding protein endoglin with distinct cytoplasmic regions. *Eur J Immunol 23,* 2340-2345.

77) Llorca, O., Trujillo, A., Blanco, F. J. & Bernabéu, C. (2007) Structural model of human endoglin, a transmembrane receptor responsible for hereditary hemorrhagic telangiectasia. *J Mol Biol 365(3)*, 694-705.

78) Gregory, A. L., Xu, G., Sotov, V. & Letarte, M. (2014) Review: the enigmatic role of endoglin in the placenta. *Placenta 35(Suppl),* 93-99.

79) Gallardo-Vara, E. M. (2018) *Endoglina soluble: mecanismo de generación y función en células endoteliales y su efecto en el remodelado vascular*. Universidad Complutense de Madrid, Facultad de Ciencias Quimicas, Departamento de Bioquímica y Biología Molecular: Thesis to obtain the degree of Doctor en Ciencias Biológicas.

80) Castonguay, R., Werner, E. D., Matthews, R. G., Presman, E., Mulivor, A. W., Solban, N., . . . Grinberg, A. V. (2011) Soluble endoglin specifically binds bone morphogenetic proteins 9 and 10 via its orphan domain, inhibits blood vessel formation, and suppresses tumor growth. *J Biol Chem 286(34),* 30034-30046.

81) Alt, A., Miguel-Romero, L., Donderis, J., Aristorena, M., Blanco, F. J., Round, A., . . . Marina, A. (2012) Structural and functional insights into endoglin ligand recognition and binding. *PLoS One 7(2),* e29948

82) Bork, P. & Sander, C. (1992) A large domain common to sperm receptors (ZP2 and ZP3) and TGF-beta type III receptor. *FEBS letters 300,* 237-240.

83) Jovine, L., Darie, C. C., Litscher, E. S. & Wassarman, P. M. (2005) Zona pellucida domain proteins. *Annu Rev Biochem 74,* 83-114.

84) Guerrero-Esteo, M., Lastres, P., Letamendia, A., Pérez-Álvarez, M. J., Langa, C., López, L. A., . . . Bernabéu, C. (1999) Endoglin overexpression modulates cellular morphology, migration, and adhesion of mouse fibroblasts. *Eur J Cell Biol 78,* 614-623.

85) Guo, B., Rooney, P., Slevin, M., Li, C., Parameshwar, S., Liu, D., . . . Kumar, S. (2004) Overexpression of CD105 in rat myoblasts: Role of CD105 in cell attachment, spreading and survival. *Int J Oncol 25(2),* 285-291.

86) Rossi, E., Sanz-Rodriguez, F., Eleno, N., Düwell, A., Blanco, F. J., Langa, C., . . . Bernabéu, C. (2013) Endothelial endoglin is involved in inflammation: role in leukocyte adhesion and transmigration. *Blood 121(2),* 403-415.

87) Rossi, E., Smadja, D. M., Boscolo, E., Langa, C., Arevalo, M. A., Pericacho, M., . . . Bernabéu, C. (2016) Endoglin regulates mural cell adhesion in the circulatory system. *Cell Mol Life Sci 73(8),* 1715-1739.

88) López-Novoa, J. M. & Bernabéu, C. (2010) The physiological role of endoglin in the cardiovascular system. *Am J Physiol Heart Circ Physiol 299(4),* 959-974.

89) Guerrero-Esteo, M., Sanchez-Elsner, T., Letamendia, A. & Bernabéu, C. (2002) Extracellular and cytoplasmic domains of endoglin interact with the transforming growth factor-beta receptors I and II. *J Biol Chem 277(32),* 29197–29209.

90) Blanco, F. J., Santibanez, J. F., Guerrero-Esteo, M., Langa, C., Vary, C. P. & Bernabéu, C. (2005) Interaction and functional interplay between endoglin and ALK-1, two components of the endothelial transforming growth factor-beta receptor complex. *J Cell Physiol 204(2),* 574-584.

91) García-Pozo, L., Miquilena-Colina, M. E., Lozano-Rodríguez, T. & García-Monzón, C. (2008) Endoglina: estructura, funciones biológicas y papel en la fibrogénesis. *Rev Esp Enferm Dig (Madrid) 100(6)*, 355-360.

92) Lastres, P., Martín-Perez, J., Langa, C. & Bernabéu, C. (1994) Phosphorylation of the human-transforming growth-factor-beta-binding protein endoglin. *Biochem J 301(3)*, 765-768.

93) Pan, C.C., Kumar, S., Shah, N., Hoyt, D. G., Hawinkels, L. J., Mythreye, K., Lee, N. Y. (2014) Src-mediated posttranslational regulation of endoglin stability and function is critical for angiogenesis. *J Biol. Chem 289(37),* 25486-25496.

94) Kennedy, M. B. (1995) Origin of PDZ (DHR, GLGF) domains. *Trends Biochem Sci* 20, 350

95) Koleva, R. I., Conley, B. A., Romero, D., Riley, K. S., Marto, J. A., Lux, A. & Vary, C. P. (2006) Endoglin structure and function: Determinants of endoglin phosphorylation by transforming growth factor-beta receptors. *J Biol Chem 281(35),* 25110-25123.

96) Conley, B. A., Koleva, R., Smith, J. D., Kacer, D., Zhang, D., Bernabéu, C. & Vary, C. P. (2004) Endoglin controls cell migration and composition of focal adhesions: function of the cytosolic domain. *J Biol Chem 279(26),* 27440-27449.

97) Sanz-Rodriguez, F., Guerrero-Esteo, M., Botella, L., Banville, D., Vary, C. & Bernabéu, C. (2004) Endoglin Regulates Cytoskeletal Organization through Binding to ZRP-1, a Member of the Lim Family of Proteins. *J Biol Chem 279(31),* 32858-32868.

98) Lee, N. Y. & Blobe, G. C. (2007) The interaction of endoglin with beta-arrestin2 regulates transforming growth factor-beta-mediated ERK activation and migration in endothelial cells. *J Biol Chem 282(29),* 21507-21517.

99) Lee, N. Y., Golzio, C., Gatza, C. E., Sharma, A., Katsanis, N. & Blobe, G. C. (2012) Endoglin regulates PI3-kinase/Akt trafficking and signaling to alter endothelial capillary stability during angiogenesis. *Mol Biol Cell 23(13),* 2412-2423.

100) Fernández-Ruiz, E., St Jacques, S., Bellón, T., Letarte, M. & Bernabéu, C. (1993) Assignment of the human endoglin gene (END) to chromosome 9q34-qter. *Cytogenet Cell Genet 64,* 204-207.

101) Pérez-Gómez, E., Eleno, N., López-Novoa, J. M., Ramirez, J. R., Velasco, B., Letarte, M., . . . Quintanilla, M. (2005). Characterization of murine S-endoglin isoform and its effects on tumor development. *Oncogene. 24(27),* 4450-4461.

102) Blanco, F. J., Grande, M. T., Langa, C., Oujo, B., Velasco, S., Rodriguez-Barbero, A., . . . Bernabéu, C. (2008) S-endoglin expression is induced in senescent endothelial cells and contributes to vascular pathology. *Circ Res 103(12),* 1383-1392.

103) Blanco, F. J. & Bernabéu, C. (2011) Alternative splicing factor or splicing factor-2 plays a key role in intron retention of the endoglin gene during endothelial senescence. *Aging Cell 10(5),* 896-907.

104) Valbuena-Díez, A. C., Blanco, F. J., Oujo, B., Langa, C., Gonzalez-Nuñez, M., Llano, E., . . . Bernabéu, C. (2012) Oxysterol-induced soluble endoglin release and its involvement in hypertension. *Circulation 126(22),* 2612-2624.

105) Gallardo-Vara, E., Blanco, F. J., Roque, M., Friedman, S. L., Suzuki, T., Botella, L. M. & Bernabéu, C. (2016) Transcription factor KLF6 upregulates expression of metalloprotease MMP14 and subsequent release of soluble endoglin during vascular injury. *Angiogenesis 19,* 155–171.

106) Hawinkels, L. J., Kuiper, P., Wiercinska, E., Verspaget, H. W., Liu, Z., Pardali, E., . . . ten Dijke, P. (2010) Matrix metalloproteinase-14 (MT1-MMP)-mediated endoglin shedding inhibits tumor angiogenesis. *Cancer Res 70(10),* 4141-4150.

107) Venkatesha, S., Toporsian, M., Lam, C., Hanai, J., Mammoto, T., Kim, Y. M., . . . Karumanchi, S. A. (2006). Soluble endoglin contributes to the pathogenesis of preeclampsia. *Nat Med 12(6),* 642-649.

108) ten Dijke, P. & Arthur, H. M. (2007) Extracellular control of TGFbeta signalling in vascular development and disease. *Nat Rev Mol Cell Biol 8(11),* 857-869.

109) Rathouska, J., Jezkova, K., Nemeckova, I. & Nachtigal, P. (2015) Soluble endoglin, hypercholesterolemia and endothelial dysfunction. *Atherosclerosis 243(2),* 383-388.

110) Gougos, A. & Letarte, M. (1988a) Identification of a human endothelial cell antigen with monoclonal antibody 44G4 produced against a pre-B leukemic cell line. *J Immunol 141,* 1925-1933.

111) Gougos, A. & Letarte, M. (1988b). Biochemical characterization of the 44G4 antigen from the HOON pre-B leukemic cell line. *J Immunol. 141,* 1934-1940.

112) Zhang, H., Shaw, A. R., Mak, A. & Letarte, M. (1996) Endoglin is a component of the transforming growth factor (TGF) beta-receptor complex of human pre-B leukemic cells*. J Immunol 156,* 564–573.

113) Bürhing, H. J., Müller, C. A., Letarte, M., Gougos, A., Saalmüller, A., von Agthoven, A. J. & Busch, F. W. (1991) Endoglin is expressed on a subpopulation of inmature erythroid cells of normal human bone marrow. *Leukemia 5,* 841-847

114) Lastres, P., Bellón, T., Cabañas, C., Sánchez-Madrid, F., Acevedo, A., Gougos, A., . . . Bernabéu, C. (1992) Regulated expression on human macrophages of endoglin, an Arg-Gly-Asp-containing surface antigen. *Eur J Immunol 22,* 393-397.

115) O'Connell, P. J., McKenzie, A., Fisicaro, N., Rockman, S. P., Pearse, M. J., d'Apice, A. J. (1992) Endoglin: a 180 kD endothelial cell and macrophage restricted differentiation molecule. *Clin Exp Immunol 90,* 154–159.

116) Robledo, M. M., Hidalgo, A., Lastres, P., Arroyo, A. G., Bernabéu, C., Sánchez-Madrid, F. & Teixidó, J. (1996) Characterization of TGF-beta 1-binding proteins in human bone marrow stromal cells. *Br J Haematol 93(3),* 507-514.

117) St Jacques, S., Cymerman, U., Pece, N. & Letarte, M. (1994a) Molecular characterization and *in situ* localization of murine endoglin reveal that it is a transforming growth factor-β binding protein of endothelial and stromal cells. *Endocrinology 134,* 2645-2657.

118) Rokhlin, O. W., Cohen, M. B., Kubagawa, H., Letarte, M. & Cooper, M. D. (1995) Differential expression of endoglin on fetal and adult hematopoietic cells in human bone marrow*. J Immunol 154,* 4456–4465.

119) Saito, M. T., Salmon, C. R., Amorim, B.R., Ambrosano, M. B., Casati, M. Z., Sallum, E. A., . . . Silvério, K.G. (2014) Characterization of highly osteoblast/cementoblast cell clones from a CD105-enriched periodontal ligament progenitor cell population. *J Periodontol 85(6),* 205–211.

120) Gougos, A., St Jacques, S., Greaves, A., O'Connell, P.J., d'Apice, A.J.F., Bürhing, H.J., . . . Letarte, M. (1992) Identification of distinct epitopes of endoglin, an RGD-containing glycoprotein of endothelial cells, leukemic cells, and syncytiotrophoblasts. *Int Immunol 4(1),* 83-92.

121) St-Jacques, S., Forte, M., Lye, S. J. & Letarte, M. (1994b) Localization of endoglin, a transforming growth factor-beta binding protein, and of CD44 and integrins in placenta during the first trimester of pregnancy. *Biol Reprod 51,* 405–413 b.

122) Bernabéu, C., Conley, B. A. & Vary, C. P. (2007) Novel biochemical pathways of endoglin in vascular cell physiology. *J Cell Biochem 102(6),* 1375-1388.

123) Parker, W. L., Goldring, M. B. & Philip, A. (2003) Endoglin is expressed on human chondrocytes and forms a heteromeric complex with betaglycan in a ligand and type II TGFbeta receptor independent manner. *J Bone Min. Res 18,* 289–302.

124) Rodriguez-Barbero, A., Obreo, J., Eleno, N., Rodriguez-Pena, A., Duwel, A., Jerkic, M., . . . Lopez-Novoa, J. M. (2001) Endoglin expression in human and rat mesangial cells and its upregulation by TGF-beta1. *Biochem Biophys Res Commun 282,* 142–147.

125) Díez-Marqués, L., Ortega-Velázquez, R., Langa, C., Rodríguez-Barbero, A., López-Novoa, J. M., Lamas, S. & Bernabéu, C. (2002) Expression of endoglin in human

mesangial cells: Modulation of extracellular matrix synthesis. *Biochem Biophys Acta 158,* 36-44.

126) Meurer, S. K., Tihaa, L., Lahme, B., Gressner, A. M. & Weiskirchen, R. (2005) Identification of endoglin in rat hepatic stellate cells: new insights into transforming growth factor beta-receptor signaling. *J Biol Chem 280,* 3078–3087.

127) Quintanilla, M., Ramirez, J. R., Pérez-Gómez, E., Romero, D., Velasco, B., Letarte, M., . . . Bernabéu, C. (2003) Expression of the TGF-beta coreceptor endoglin in epidermal keratinocytes and its dual role in multistage mouse skin carcinogenesis. *Oncogene 22(38),* 5976-5985

128) Adam, P. J., Clesham, G. J. & Weissberg, P. L. (1998) Expression of endoglin mRNA and protein in human vascular smooth muscle cells. *Biochem Biophys Res Commun 247(1),* 33–37

129) Conley, B. A., Smith, J. D., Guerrero-Esteo, M., Bernabéu, C. & Vary, C. P. (2000) Endoglin, a TGF-beta receptor-associated protein, is expressed by smooth muscle cells in human atherosclerotic plaques. *Atherosclerosis 153,* 323–335.

130) Rivera, L. B. & Brekken, R. A. (2011) SPARC promotes pericyte recruitment via inhibition of endoglin dependent TGF-beta1 activity. *J Cell Biol 193,* 1305–1319.

131) Toporsian, M., Gros, R., Kabir, M. G., Vera, S., Govindaraju, K., Eidelman, D. H., . . . Letarte, M. (2005) A role for endoglin in coupling eNOS activity and regulating vascular tone revealed in hereditary hemorrhagic telangiectasia. *Circ Res 96,* 684–692.

132) Santibanez, J.F., Blanco, F. J., Garrido-Martin, E. M., Sanz-Rodriguez, F. del Pozo, M. A. & Bernabéu, C. (2008) Caveolin-1 interacts and cooperates with the transforming growth factor-b type I receptor ALK1 in endothelial caveolae. *Cardiovascular Research 77,* 791–799

133) Qu, R., Silver, M. M. & Letarte, M. (1998) Distribution of endoglin in early human development reveals high levels on endocardial cushion tissue mesenchyme during valve formation. *Cell Tissue Res 292,* 333–343.

134) Arthur, H. M., Ure, J., Smith, A. J., Renforth, G., Wilson, D. I., Torsney, E., . . . Diamond, A. G. (2000) Endoglin, an ancillary TGF beta receptor, is required for extraembryonic angiogenesis and plays a key role in heart development. *Dev Biol 217(1),* 42-53.

135) Mahmoud, M., Borthwick, G.M., Hislop, A. A. & Arthur, H. M. (2009) Endoglin and activin receptor-like-kinase 1 are co-expressed in the distal vessels of the lung: implications for two familial vascular dysplasias, HHT and PAH. *Lab Invest 89(1),* 15-25.

136) Matsubara, S., Bourdeau, A., terBrugge, K.G., Wallace, C. & Letarte, M. (2000) Analysis of endoglin expression in normal brain tissue and in cerebral arteriovenous malformations. *Stroke 31(11),* 2653-2660.

137) Henriksen, R., Gobl, A., Wilander, E., Oberg, K., Miyazono, K. & Funa, K. (1995) Expression and prognostic significance of TGF-b isotypes, latent TGF-b1 binding protein, TGF-b type I and typeII receptors, and endoglin in normal ovary and ovarian neoplasms. *Lab Invest 73(2),* 213-220

138) Liu, Y., Jovanovic, B., Pins, M., Lee, C. & Bergan, R. C. (2002) Over expression of endoglin in human prostate cancer suppresses cell detachment, migration and invasion. *Oncogene 21(54),* 8272-8281.

139) Craft, C. S., Romero, D., Vary, C. P. & Bergan, R. C. (2007) Endoglin inhibits prostate cancer motility via activation of the ALK2-Smad1 pathway. *Oncogene 26(51),* 7240-7250.

140) Botella, L. M., Sánchez-Elsner, T., Sanz-Rodriguez, F., Kojima, S., Shimada, J., Guerrero-Esteo, M., . . . Bernabéu, C. (2002) Transcriptional activation of endoglin and transforming growth factor-beta signaling components by cooperative interaction between Sp1 and KLF6: their potential role in the response to vascular injury. *Blood 100(12),* 4001-4010.

141) van Laake, L. W., van den Driesche, S., Post, S., Feijen, A., Jansen, M. A., Driessens, M. H., . . . Mummery, C. L. (2006) Endoglin has a crucial role in blood cell-mediated vascular repair. *Circulation 114(21),* 2288-2297.

142) Mancini, M. L., Verdi, J. M., Conley, B. A., Nicola, T., Spicer, D. B., Oxburgh, L. H. & Vary, C. P. (2007) Endoglin is required for myogenic differentiation potential of neural crest stem cells. *Dev Biol. 308(2),* 520-533

143) Chen, C. Z., Li, M., de Graaf, D., Monti, S., Go¨ttgens, B., Sanchez, M. J., . . . Lodish, H. F. (2002) Identification of endoglin as a functional marker that defines long-term repopulating hematopoietic stem cells. *Proc Natl Acad Sci USA 99(24),* 15468-15473.

144) Perlingeiro, R. C. (2007) Endoglin is required for hemangioblast and early hematopoietic development. *Development 134(16),* 3041-3048.

145) Nasrallah, R., Knezevic, K., Thai, T., Thomas, S. R., Göttgens, B., Lacaud, G., . . . Pimanda, J. E. (2015) Endoglin potentiates nitric oxide synthesis to enhance definitive hematopoiesis. *Biol Open 4(7),* 819-8 29

146) Docherty, N. G., López-Novoa, J. M., Arevalo, M., Düwel, A., Rodriguez-Peña, A., Pérez-Barriocanal, F., Bernabéu, C. & Eleno, N. (2006) Endoglin regulates renal ischaemia-reperfusion injury. *Nephrol Dial Transplant 21(8),* 2106-2119.

147) Garrido, E. (2011) *Estudio de la regulación transcripcional de ALK1, receptor tipo I del TGF-β, en la célula endotelial*. Universidad Complutense de Madrid, Facultad de Biología, Departamento de Bioquímica y Biología Molecular: Thesis to obtain the degree of Doctor en Ciencias Biológicas.

148) Marazuela, M., Sánchez-Madrid, F., Acevedo, A., Larrañaga, E., de Landázuri, M. O. (1995) Expression of vascular adhesion molecules on human endothelial in autoimmune thyroid disorders. *Clin Exp Immunol 102(2),* 328-334.

149) van de Kerkhof, P. C., Rulo, H. F., van Pelt, J. P., van Vlijmen-Willems, I. M., De Jong, E. M. (1998) Expression of endoglin in the transition between psoriatic uninvolved and involved skin. *Acta Derm Venereol 78(1),* 19-21.

150) Ma, X., Labinaz, M., Goldstein, J., Miller, H., Keon, W. J., Letarte, M. & O'Brien, E. (2000) Endoglin is overexpressed after arterial injury and is required for transforming growth factor-beta-induced inhibition of smooth muscle cell migration. *Arterioscler Thromb Vasc Biol 20(12*), 2546-2552.

151) Fonsatti, E., Altomonte, M., Nicotra, M. R., Natali, P. G., Maio, M. (2003) Endoglin (CD105): a powerful therapeutic target on tumor-associated angiogenetic blood vessels. *Oncogene 22(42),* 6557-6563.

152) Duff, S. E., Li, C., Garland, J. M., Kumar, S. (2003) CD105 is important for angiogenesis: Evidence and potential applications. *FASEB J 17,* 984-992.

153) Bauman, T. M., Huang, W., Lee, M. H. & Abel, E. J. (2016) Neovascularity as a prognostic marker in renal cell carcinoma. *Hum Pathol 57,* 98-105.

154) Cheifetz, S., Bellón, T., Calés, C., Vera, S., Bernabéu, C., Massagué, J. & Letarte, M. (1992) Endoglin is a component of the transforming growth factor-beta receptor system in human endothelial cells. *J Biol Chem 267(27)* 19027-19030.

155) Barbara, N. P., Wrana, J. L. & Letarte, M. (1999) Endoglin is an accessory protein that interacts with the signaling receptor complex of multiple members of the transforming growth factor-beta superfamily. *J Biol Chem. 274(2),* 584–594

156) Goumans, M. J., Valdimarsdottir, G., Itoh, S., Rosendahl, A., Sideras, P. & ten Dijke, P. (2002) Balancing the activation state of the endothelium via two distinct TGF-beta type I receptors. *EMBO J 21(7),* 1743-1753.

157) ten Dijke, P., Korchynskyi, O., Valdimarsdottir, G. & Goumans, M. J. (2003) Controlling cell fate by bone morphogenetic protein receptors. *Mol Cell Endocrinol 211(1-2),* 105-113.

158) Boehm, J. R., Kutz, S. M., Sage, E. H., Staiano-Coico, L. & Higgins, P. J. (1999) Growth state-dependent regulation of plasminogen activator inhibitor type-1 gene expression during epithelial cell stimulation by serum and transforming growth factor-beta1. *J Cell Physiol 181(1),* 96-106

159) Goumans, M. J., Lebrin, F. & Valdimarsdottir, G. (2003) Controlling the angiogenic switch: a balance between two distinct TGF-b receptor signaling pathways. *Trends Cardiovasc Med 13(7),* 301-307.

160) Lebrin, F., Goumans, M. J., Jonker, L., Carvalho, R. L., Valdimarsdottir, G., Thorikay, M., . . . ten Dijke, P. (2004) Endoglin promotes endothelial cell proliferation and TGF-beta/ALK1 signal transduction. *EMBO J 23(20),* 4018-4028.

161) Bourdeau, A., Dumont, D.J. & Letarte, M. (1999) A murine model of hereditary hemorrhagic telangiectasia*. J Clin Invest 104,* 1343-1351

162) Li, D. Y, Sorensen, L. K., Brooke, B. S., Urness, L. D., Davis, E. C., Taylor, D. G., . . . Wendel, D. P. (1999) Defective angiogenesis in mice lacking endoglin. *Science 284(5419),* 1534-1537.

163) Sorensen, L.K., Brooks, B.S., Li, D.Y. & Umess, L.D. (2003) Loss of distinct arterial and venous boundaries in mice lacking endoglin, a vascular-specific TGFbeta coreceptor. *Dev Biol 261,* 235-250.

164) Núñez-Gómez, E. & Lopez-Novoa, J.M. (2015) Papel de endoglina en los eventos fisiológicos involucrados en la revascularización postisquémica. *Angiología 67,* 206-215

165) Núñez-Gómez, E., Pericacho, M., Ollauri-Ibáñez, C., Bernabéu, C. & López-Novoa, J. M. (2017) The role of endoglin in post-ischemic revascularization. *Angiogenesis 20(1),* 1-24.

166) Mahmoud, M., Allinson, K. R., Zhai, Z., Oakenfull, R., Ghandi, P., Adams, R. H., . . . Arthur, H. M. (2010) Pathogenesis of arteriovenous malformations in the absence of endoglin. *Circ Res 106(8),* 1425-1433.

167) Jerkic, M., Rodríguez-Barbero, A., Prieto, M., Toporsian, M., Pericacho, M., Rivas-Elena, J. V., . . . López-Novoa, J. M. (2006) Reduced angiogenic responses in adult Endoglin heterozygous mice. *Cardiovasc Res 69(4),* 845-854

168) Park, S., Dimaio T. A., Liu, W., Wang, S., Sorenson, C. M. & Sheibani, N. (2013) Endoglin regulates the activation and quiescence of endothelium by participating in canonical and non-canonical TGF-β signaling pathways. *J Cell Sci 126(6),* 1392-1405.

169) Santibañez, J. F., Letamendia, A., Perez-Barriocanal, F., Silvestri, C., Saura, M., Vary, C. P., . . . Bernabéu, C. (2007) Endoglin increases eNOS expression by modulating Smad2 protein levels and Smad2-dependent TGF-beta signaling. *J Cell Physiol 210,* 456–468.

170) Zucco, L., Zhang, Q., Kuliszewski, M. A., Kandic, I., Faughnan, M. E., Stewart, D. J. & Kutryk, M. J. (2014) Circulating angiogenic cell dysfunction in patients with hereditary hemorrhagic telangiectasia. *PLoS One 9(2),* e89927.

171) Gallardo-Vara, E., Gamella-Pozuelo, L., é-Roque, L., Bartha, J. L., Garcia-Palmero, I., Casal, J. I. & Bernabéu, C. (2020) Potential Role of Circulating Endoglin in Hypertension via the Upregulated Expression of BMP4. *Cells 9(4),* 988.

172) David L, Mallet C, Mazerbourg S, Feige JJ, Bailly S. (2007) Identification of BMP9 and BMP10 as functional activators of the orphan activin receptor-like kinase 1 (ALK1) in endothelial cells. *Blood 109(5),* 1953-1961

173) Suzuki, Y., Ohga, N., Morishita, Y., Hida, K., Miyazono, K. & Watabe, T. (2010) BMP-9 induces proliferation of multiple types of endothelial cells *in vitro* and *in vivo*. *J Cell Sci 123(10),* 1684-1692.

174) Park, J. H., Choi, M. R., Park, K. S., Kim, S. H., Jung, K. H. & Chai, Y. G. (2012) The characterization of gene expression during mouse neural stem cell differentiation in vitro. *Neurosci Lett 506(1),* 50-54.

175) van Meeteren, L. A. & ten Dijke, P. (2012) Regulation of endothelial cell plasticity by TGF-β. *Cell Tissue Res 347(1),* 177-186.

176) Nolan-Stevaux, O., Zhong, W., Culp, S., Shaffer, K., Hoover, J., Wickramasinghe, D. & Ruefli-Brasse, A. (2012) Endoglin requirement for BMP9 signaling in endothelial cells reveals new mechanism of action for selective anti-endoglin antibodies. *PLoS One 7(12),* e50920.

177) David, L., Feige, J. J. & Bailly, S. (2009) Emerging role of bone morphogenetic proteins in angiogenesis. *Cytokine Growth Factor Rev 20(3),* 203-212.

178) Atri, D., Larrivée, B., Eichmann, A. & Simons, M. (2014) Endothelial signaling and the molecular basis of arteriovenous malformation. *Cell Mol Life Sci 71,* 867–883.

179) Ollauri-Ibáñez, C., Núñez-Gómez, E., Egido-Turrión, C., Silva-Sousa, L., Díaz-Rodríguez, E. Rodríguez-Barbero, A., . . . Pericacho, M. (2020) Continuous endoglin (CD105) overexpression disrupts angiogenesis and facilitates tumor cell metastasis. *Angiogenesis 23,* 231–247.

180) Pérez-Gómez, E., Villa-Morales, M., Santos, J., Fernández-Piqueras, J., Gamallo, C., Dotor, J., . . . Quintanilla, M. (2007) A role for endoglin as a suppressor of malignancy during mouse skin carcinogenesis. *Cancer Res 67(21),* 10268-10277.

181) Takahashi, N., Kawanishi-Tabata, R., Haba, A., Tabata, M., Haruta, Y., Tsai, H., Seon, B. K. (2001) Association of serum endoglin with metastasis in patients with colorectal, breast, and other solid tumors, and suppressive effect of chemotherapy on the serum endoglin. *Clin Cancer Res 7(3),* 524-532.

182) Sánchez-Elsner, T., Botella, L. M., Velasco, B., Langa, C. & Bernabéu, C. (2002) Endoglin expression is regulated by transcriptional cooperation between the

hypoxia and transforming growth factor-beta pathways. *J Biol Chem 277(46),* 43799-43808.

183) Li, C., Guo, B., Wilson, P. B., Stewart, A., Byrne, G., Bundred, N. & Kumar, S. (2000) Plasma levels of soluble CD105 correlate with metastasis in patients with breast cancer. *Int J Cancer 89,* 122–126.

184) Calabrò, L., Fonsatti, E., Bellomo, G., Alonci, A., Colizzi, F., Sigalotti, L., . . . Maio, M. (2003) Differential levels of soluble endoglin (CD105) in myeloid malignancies. *J Cell Physiol 194(2),* 171-175

185) Bellone, G., Solerio, D., Chiusa, L., Brondino, G., Carbone, A., Prati, A., . . . Dei Poli, M. (2007) Transforming growth factor-beta binding receptor endoglin (CD105) expression in esophageal cancer and in adjacent nontumorous esophagus as prognostic predictor of recurrence. *Ann Surg Oncol 14(11),* 3232-3242.

186) Mysliwiec, P., Pawlak, K., Bandurski, R. & Kedra, B. (2009) Soluble angiogenesis markers in gastric tumor patients. *Folia Histochem Cytobiol 47(1),* 81-86.

187) Odegaard, E., Davidson, B., Engh, V., Onsrud, M. & Staff, A. C. (2008) Assessment of endoglin and calprotectin as potential biomarkers in ovarian carcinoma and borderline tumors of the ovary. *Am J Obstet Gynecol 199(5),* 533.

188) Fujita, K., Ewing, C. M., Chan, D. Y., Mangold, L. A., Partin, A. W., Isaacs, W. B. & Pavlovich, C. P. (2009) Endoglin (CD105) as a urinary and serum marker of prostate cancer. *Int J Cancer 124(3),* 664-669.

189) Karam, J. A., Svatek, R. S., Karakiewicz, P. I., Gallina, A., Roehrborn, C. G., Slawin, K. M. & Shariat, S. F. (2008) Use of preoperative plasma endoglin for prediction of lymph node metastasis in patients with clinically localized prostate cancer. *Clin Cancer Res 14(5),* 1418-1422.

190) Di Cosola, M., Cazzolla, A. P., Scivetti, M., Testa, N. F., Lo Muzio, L., Favia, G., . . . Bascones, A. (2005) Síndrome de Rendu-Osler-Weber o Telangiectasia Hemorrágica Hereditaria (HHT) Descripción de dos casos y revisión de la literatura. *Av Odontoestomatol 21(6),* 297-303.

191) Shovlin, C. L., Hughes, J. M., Tuddenham, E. G., Temperley, I., Perembelon, Y. F., Scott, J., . . . Seidman, J. G. (1994) A gene for hereditary haemorrhagic telangiectasia maps to chromosome 9q3. *Nat Genet 6,* 205-209.

192) McDonald, J., Damjanovich, K., Millson, A., Wooderchak, W., Chibuk, J. M., Stevenson, D. A., . . . BayrakToydemir, P. (2011) Molecular diagnosis in hereditary hemorrhagic telangiectasia: findings in a series tested simultaneously by sequencing and deletion/duplication analysis. *Clin Genet 79(4),* 335-344.

193) McDonald, J., Wooderchak-Donahue, W., VanSant Webb, C., Whitehead, K., Stevenson, D. A. & BayrakToydemir, P. (2015) Hereditary hemorrhagic telangiectasia: genetics and molecular diagnostics in a new era. *Front Genet 6,* 1.

194) Fernández-Ruiz, E., St Jacques, S., Bellón, T., Letarte, M., Bernabéu, C. (1993) Assignment of the human endoglin gene (END) to chromosome 9q34-qter. *Cytogenet Cell Genet 64,* 204-207.

195) McDonald, M. T., Papenberg, K. A., Ghosh, S., Glatfelter, A. A., Biesecker, B. B., Helmbold, E. A., . . . Marchuk, D. A. (1994) A disease locus for hereditary haemorrhagic telangiectasia maps to chromosome 9q33-34. *Nat Genet 6,* 197-204.

196) Shovlin, C. L., Hughes, J. M., Scott, J., Seidman, C. E., . . . Seidman, J. G. (1997) Characterization of endoglin and identification of novel mutations in hereditary hemorrhagic telangiectasia. *Am J Hum Genet 61(1),* 68- 79.

197) McAllister, K. A., Grogg, K. M., Johnson, D. W., Gallione, C. J., Baldwin, M. A., Jackson, C. E., . . . Marchuk, D. A. (1994) Endoglin, a TGF-β binding protein of endothelial cells, is the gene for hereditary haemorrhagic telangiectasia type 1. *Nature Genet 8,* 345-351.

198) Johnson, D. W., Berg, J. N., Gallione, C. J., McAllister, K. A., Warner, J. P., Helmbold, E. A., . . . Marchuk, D. A. (1995) A second locus for hereditary hemorrhagic telangiectasia maps to chromosome 12. *Genome Res 5(1),* 21-28.

199) Johnson, D. W., Berg, J. N., Baldwin, M. A., Gallione, C. J., Marondel, I., Yoon, S. J., Marchuk, D. A. (1996) Mutations in the activin receptor-like kinase 1 gene in hereditary haemorrhagic telangiectasia type 2. *Nat Genet 13(2),* 189-195.

200) Cole, S. G., Begbie, M. E., Wallace, G. M. & Shovlin, C. L. (2005) A new locus for hereditary haemorrhagic telangiectasia (HHT3) maps to chromosome 5. *J Med Genet 42(7),* 577-582.

201) Bayrak-Toydemir, P., McDonald, J., Akarsu, N., Toydemir, R. M., Calderon, F., Tuncali, T., . . . Mao, R. (2006a) A fourth locus for hereditary hemorrhagic telangiectasia maps to chromosome 7. *Am J Med Genet A 140(20),* 2155-2162.

202) Bayrak-Toydemir, P., McDonald, J., Markewitz, B., Lewin, S., Miller, F., Chou, L. S., . . . Mao, R. (2006b) Genotype-phenotype correlation in hereditary hemorrhagic telangiectasia: mutations and manifestations. *Am J Med Genet A 140(5),* 463-470.

203) Gallione, C. J., Repetto, G. M., Legius, E., Rustgi, A. K., Schelley, S. L., Tejpar, S., . . . Marchuk, D. A. (2004) A combined syndrome of juvenile polyposis and hereditary haemorrhagic telangiectasia associated with mutations in MADH4 (SMAD4). *Lancet 363(9412),* 852-859.

204) Wooderchak-Donahue, W. L., McDonald, J., O'Fallon, B., Upton, P. D., Li, W., Roman, B. L. & Bayrak-Toydemir, P. (2013) BMP9 mutations cause a vascular-anomaly syndrome with phenotypic overlap with hereditary hemorrhagic telangiectasia. *Am J Hum Genet 93(3),* 530-537.

205) Albiñana, V., Cuesta, A. M., de Rojas-P, I., Gallardo-Vara, E., Recio-Poveda, L., Bernabéu, C. & Botella, L. M. (2020) Review of Pharmacological Strategies with Repurposed Drugs for Hereditary Hemorrhagic Telangiectasia Related Bleeding. *J Clin Med 9,* 1766.

206) Cymerman, U., Vera, S., Pece-Barbara, N., Bourdeau, A., White, R.I. Jr., Dunn, J. & Letarte, M. (2000) Identification of hereditary hemorrhagic telangiectasia type 1 in newborns by protein expression and mutation analysis of endoglin. *Pediatr Res 47(1),* 24-35.

207) Paquet, M. E., Pece-Barbara, N., Vera, S., Cymerman, U., Karabegovic, A., Shovlin, C. & Letarte, M. (2001) Analysis of several endoglin mutants reveals no endogenous mature or secreted protein capable of interfering with normal endoglin function. *Hum Mol Genet 10(13),* 1347-1357

208) Lebrin, F., Srun, S., Raymond, K., Martin, S., van den Brink, S., Freitas, C., . . . Mummery, C. L. (2010) Thalidomide stimulates vessel maturation and reduces epistaxis in individuals with hereditary hemorrhagic telangiectasia. *Nat Med 16(4),* 420-428.

209) Tual-Chalot, S., Oh, S. P., Arthur, H. M. (2015) Mouse models of hereditary hemorrhagic telangiectasia: recent advances and future challenges. *Front Genet 6,* 25.

210) Núñez-Gómez, E. (2016) *Las isoformas de endoglina en la regulación de la angiogénesis.* Universidad de Salamanca, Departamento de Fisiología y Farmacología: Thesis to obtain the degree of Doctor with mention International Doctorate

211) Velasco, S., Alvarez-Muñoz, P., Pericacho, M., Dijke, P. T., Bernabéu, C., López-Novoa, J. M. & Rodríguez-Barbero, A. (2008) L- and S-endoglin differentially modulate TGFbeta1 signaling mediated by ALK1 and ALK5 in L6E9 myoblasts. *J Cell Sci 121(6),* 913-919.

212) Aristorena, M., Blanco, F. J., de Las Casas-Engel, M., Ojeda-Fernandez, L., Gallardo-Vara, E., Corbi, A., . . . Bernabéu C. (2014) Expression of endoglin isoforms in the myeloid lineage and their role during aging and macrophage polarization. *J Cell Sci 127(12),* 2723-2735.

213) Blázquez-Medela, A. M., García-Ortiz, L., Gómez-Marcos, M. A., Recio-Rodríguez, J. I., Sánchez-Rodríguez, A., López-Novoa, J. M. & Martínez-Salgado, C. (2010) Increased plasma soluble endoglin levels as an indicator of cardiovascular alterations in hypertensive and diabetic patients. *BMC Med 8,* 86.

214) Nemeckova, I., Serwadczak, A., Oujo, B., Jezkova, K., Rathouska, J., Fikrova, P., . . . Nachtigal, P. (2015) High soluble endoglin levels do not induce endothelial dysfunction in mouse aorta. *PLoS One 10(3),* e0119665.

215) Jezkova, K., Rathouska, J., Nemeckova, I., Fikrova, P., Dolezelova, E., Varejckova, M., . . . Nachtigal, P. (2016) High Levels of Soluble Endoglin Induce a Proinflammatory and Oxidative-Stress Phenotype Associated with Preserved NO-Dependent Vasodilatation in Aortas from Mice Fed a High-Fat Diet. *J Vasc Res 53(3-4),* 149-162.

216) Rathouska, J., Fikrova, P., Mrkvicova, A., Blazickova, K., Varejckova, M., Dolezelova, E., . . . Nachtigal, P. (2017) High soluble endoglin levels do not induce changes in structural parameters of mouse heart. *Heart Vessels 32,* 1013–1024.

217) Varejckova, M., Gallardo-Vara, E., Vicen, M., Vitverova, B., Fikrova, P., Dolezelova, E., . . . Nachtigal, P. (2017) Soluble endoglin modulates the pro-inflammatory mediators NF-κB and IL-6 in cultured human endothelial cells. *Life Sci 175,* 52-60.

218) Cruz-González, I., Pabón, P., Rodríguez-Barbero, A., Martín-Moreiras, J., Pericacho, M., Sánchez, P. L., . . . López-Novoa, J. M. (2008) Identification of serum endoglin as a novel prognostic marker after acute myocardial infarction. *J Cell Mol Med 12(3),* 955-961.

219) Nachtigal, P., Zemankova, Vecerova, L., Rathouska, J. & Strasky, Z. (2012) The role of endoglin in atherosclerosis. *Atherosclerosis 224(1),* 4-11.

220) Fujimoto, M., Hasegawa, M., Hamaguchi, Y., Komura, K., Matsushita, T., Yanaba, K., . . . Sato, S. (2006) A clue for telangiectasis in systemic sclerosis: elevated serum soluble endoglin levels in patients with the limited cutaneous form of the disease. *Dermatology 213(2),* 88-92.

221) Karampoor, S., Zahednasab, H., Ramagopalan, S., Mehrpour, M. & Keyvani, H. (2016) Angiogenic factors are associated with multiple sclerosis. *J Neuroimmunol 301*, 88-93.

222) Henderson, J. T., Thompson, J. H., Burda, B. U., Cantor, A., Beil, T. & Whitlock, E. P. (2017) Screening for Preeclampsia: A Systematic Evidence Review for the U.S. Preventive Services Task Force [Internet]. Rockville (MD): Agency for Healthcare Research and Quality (US). *Evidence Synthesis 148.*

223) Akaishi, R., Yamada, T., Morikawa, M., Nishida, R. & Minakami, H. (2014) Clinical features of isolated gestational proteinuria progressing to pre-eclampsia: retrospective observational study. *BMJ Open 4(4),* e004870.

224) Ahmad, S. & Ahmed, A. (2004) Elevated placental soluble vascular endothelial growth factor receptor-1 inhibits angiogenesis in preeclampsia. *Circ Res 95(9),* 884-891.

225) Staff, A. C., Braekke, K., Johnsen, G. M., Karumanchi, S. A., & Harsem, N. K. (2007) Circulating concentrations of soluble endoglin (CD105) in fetal and maternal serum and in amniotic fluid in preeclampsia. *Am J Obstet Gynecol 197(2), 176*.

226) Hagmann, H., Thadhani, R., Benzing, T., Karumanchi, S. A. & Stepan, H. (2012) The promise of angiogenic markers for the early diagnosis and prediction of preeclampsia. *Clin Chem 58(5),* 837-845.

227) Oujo, B., Perez-Barriocanal, F., Bernabéu, C. & Lopez-Novoa, J. M. (2013) Membrane and soluble forms of endoglin in preeclampsia. *Curr Mol Med 13(8),* 1345-57.

228) Calloni, R., Viegas, G. S., Türck, P., Bonatto, D. & Henriques, J. A. (2014) Mesenchymal stromal cells from unconventional model organisms. *Cytotherapy 16(1),* 3-16.

229) Calloni, R., Cordero, E. A., Henriques, J. A., Bonatto, D. (2013) Reviewing and updating the major molecular markers for stem cells. Stem Cells Dev 22(9), 1455–1476.

230) Bunnell, B. A., Flaat, M., Gagliardi, C., Patel, B. & Ripoll, C. (2008) Adipose-derived Stem Cells: Isolation, Expansion and Differentiation. *Methods 45(2),* 115–120.

231) Roy-Chaudhury, P., Simpson, J. G. & Power, D. A. (1997) Endoglin, a transforming growth factor-beta-binding protein, is upregulated in chronic progressive renal disease. *Exp Nephrol 5,* 55-60.

232) García-Monzón, C., Sánchez-Madrid, F., García-Buey, L., García-Arroyo, A., García-Sánchez, A. & Moreno-Otero, R. (1995) Vascular adhesion molecule expression in viral chronic hepatitis: Evidence of neoangiogenesis in portal tracts. *Gastroenterology 108,* 231-241.

233) Blann, A. D., Wang, J. M., Wilson, P. B. & Kumar, S. (1996) Serum levels of the TGF-beta receptor are increased in atherosclerosis. *Atherosclerosis 120(1–2),* 221–226.

234) Torsney, E., Charlton, R., Diamond, A. G., Burn, J., Soames, J. V. & Arthur HM. (2003). Mouse model for hereditary hemorrhagic telangiectasia has a generalized vascular abnormality. *Circulation 107(12),* 1653-1657.

235) Allinson, K. R., Carvalho, R. L., van den Brink, S., Mummery, C. L. & Arthur, H. M. (2007). Generation of a floxed allele of the mouse Endoglin gene. *Genesis 45(6),* 391-395.

236) Ojeda-Fernández, L., Recio-Poveda, L., Aristorena, M., Lastres, P., Blanco, F. J., Sanz-Rodríguez, F., . . . Botella, L. M. (2016). Mice Lacking Endoglin in Macrophages Show an Impaired Immune Response. *PLoS Genet 12(3),* e1005935

237) Kapur, N. K., Morine, K. J. & Letarte, M. (2013) Endoglin: a critical mediator of cardiovascular health. *Vasc Health Risk Manag 9,* 195-206.

238) Pérez-Roque, L., Núñez-Gómez, E., Rodríguez-Barbero, A., Bernabéu, C., López-Novoa, J. M. & Pericacho, M. (2021) Pregnancy-Induced High Plasma Levels of Soluble Endoglin in Mice Lead to Preeclampsia Symptoms and Placental Abnormalities. *Int J Mol Sci 22,* 165.

239) Maio, M., Altomonte, M. & Fonsatti, E. (2006). It is the prime time for endoglin (CD105) in the clinical setting?. *Cardiovasc Res 69,* 781-783.

240) Ahmadvand, D., Rassaee, M. J., Rahbarizadeh, F. & Mohammadi, M. (2008) Production and Characterization of a High-Affinity Nanobody against Human Endoglin. *Hybridoma 27(5),* 353-360.

241) Ahmadvand, D., Rassaee, M., Rahbarizadeh, F., Sheikholislami, F. & Kontermann, R. (2009) Cell selection and characterization of a novel human endothelial cell specific nanobody. *Molecular Immunology 46(8-9),* 1814-1823.

242) Pardon, E., Laeremans, T., Triest, S., Rasmussen, S. G., Wohlkönig, A., Ruf, A., . . . Steyaert, J. (2014) A general protocol for the generation of Nanobodies for structural biology. *Nat Protoc 9,* 674–693.

243) Xu, J., Xu, K., Jung, S., Conte, A., Lieberman, J., Muecksch, F., . . . Casella, R. (2021) Nanobodies from camelid mice and llamas neutralize SARS-CoV-2 variants. *Nature 595,* 278-282.

244) Barreto, T., Alfonso, Y., Lafaye, P., Perez, A., Herrera-Velit, P. & Espinoza, J. R. (2018) Single-chain antibodies from alpaca for the detection of *Fasciola hepatica* antigens. *Revista Peruana de Medicina Experimental y Salud Publica 35(4),* 573-580.

245) Gushiken, E. (2016) *Generación de anticuerpos de dominio único específicos para CD105 humano.* Universidad Peruana Cayetano Heredia, Perú. Thesis to obtain the degree of Magister en Ciencias, mention Biochemistry and Molecular Biology.

246) Hill, R. T. (2015) Plasmonic biosensors. *Wiley Interdiscip Rev Nanomed Nanobiotechnol 7,* 152–168.

247) Cherif, B., Roget, A., Villiers, C. L., Calemczuk, R., Leroy, V., Marche, P. N., . . . Villiers, M. B. (2006) Clinically related protein-peptide interactions monitored in real time on novel peptide chips by surface plasmon resonance imaging. *Clin Chem 52,* 255–262.

248) Rich, R. L. & Myszka, D. G. (2011) Survey of the 2009 commercial optical biosensor literature. *J Mol Recognit 24,* 892–914.

249) Nguyen, H. H., Park, J., Kang, S. & Kim, M. (2015) Surface plasmon resonance: a versatile technique for biosensor applications. *Sensors 15,* 10481–10510.

250) Kodoyianni, V. (2011) Label-free analysis of biomolecular interactions using SPR imaging. *Biotechniques 50,* 32–40.

251) Abadian, P. N., Kellev, C. P. & Goluch, E. D. (2014) Cellular analysis and detection using Surface plasmon resonance techniques. *Anal Chem 86,* 2799–2812.

252) Mallevre, F., Temp,er, V., Mathey, R., Leroy, L., Roupioz, Y., Fernandes, T. F., . . . Livache, T. (2016) Real-time toxicity testing of silver nanoparticles to *Salmonella enteritidis* using surface plasmon resonance imaging: A proof of concept. *NanoImpact 1,* 55–59.

253) Li, S., Song, G., Bai, Y., Song, N., Zhao, J., Jian Liu, J. & Hu, C (2021) Applications of Protein Microarrays in Biomarker Discovery for Autoimmune Diseases. *Front. Immunol 12*, 645632.

254) Ruano-Gallego, D., García-Villadangos, M., Moreno-Paz, M., Gómez-Elvira, J., Postigo, M., Simón-Sacristán, M., . . . Parro, V. (2021) A multiplex antigen microarray for simultaneous IgG and IgM detection against SARS-CoV-2 reveals higher sero prevalence than reported. *Biotechnology 14*, 1228–1236.

255) Bacarese-Hamilton, T., Mezzasoma, L., Ingham, C., Ardizzoni, A., Rossi, R., Bistoni, F. & Crisanti, A. (2002) Detection of allergen specific IgE on microarray by use of signal amplification techniques. *Clin Chem 48,* 1367-1370.

256) Cleary, M. D., Singh, U., Blader, I. J., Brewer, J. L. & Boothroyd J. C. (2002) *Toxoplasma gondii* asexual development: identification of developmentally regulated genes and distinct patterns of gene expression. *Eukaryot Cell 1,* 329-340.

257) De Avalos, S. V., Blader, I. J., Fisher, M., Boothroyd, J. C. & Burleigh, B. A. (2002) Immediate/early response to *Trypanosoma cruzi* infection involves minimal modulation of host cell transcription. *J Biol Chem 277,* 639-644.

258) Mezzasoma, L., Bacarese-Hamilton, T., Di Cristina, M., Rossi, R., Bistoni, F. & Crisanti, A. (2002) Antigen microarray for serodiagnosis of infectious diseases. *Clin Chem 48*, 121-130.

259) Arenkov, P., Kukhtin, A., Gemmell, A., Voloshchuk, S., Chupeeva, V. & Mirzabekov, A. (2000) Protein microchips: use for immunoassay and enzymatic reactions. *Anal Biochem 278,* 123–131.

260) MacBeath, G. & Schreiber, S. L. (2000) Printing proteins as microarrays for high-throughput function determination. *Science 289(1),* 60-63.

261) Uttamchandani, M., Chen, G. Y., Lesaicherre, M. L. & Yao, S. Q. (2004) Site peptide immobilization strategies for the rapid detection of kinase activity on microarrays. *Methods Mol Biol 264,* 191–204.

262) Goode, J. (2015) *Development of biosensors using novel bioreceptors; Investigation and optimisation of fundamental parameters at the nanoscale.* University of Leeds, England, United Kingdom: Thesis to obtain the degree of Doctor of Philosophy

263) Suraniti, E., Sollier, E., Calemczuk, R., Livache, T., Marche, P. N., Villiers, M-B & Roupioz, Y. (2007) Real-time detection of lymphocytes binding on an antibody chip using SPR Imaging. *Lab Chip 7,* 1206–1208.

264) Milgram, S., Cortes, S., Villiers, M. B., Marche, P., Buhot, A., Livache, T. & Roupioz, Y. (2011). On chip real time monitoring of B-cells hybridoma secretion of immunoglobulin. *Biosensors & bioelectronics* 26(5), 2728–2732.

265) Alvarado, R. (2018) *Stratégies de fonctionnalisation pour le développement de biopuces innovantes.* Communaute Université Grenoble Alpes, France: Thesis to obtain the degree of Doctor, Spécialité Chimie-Biologie.

Chapter II: Selection of the anti-CD105 VHH

Abstract

A VHH cDNA library was synthesised from RNA extracted from peripheral blood mononuclear cells (PBMCs) of alpacas immunised with a lysate of human urine bladder cancer cells (T24 cell line). The library was screened with recombinant sCD105 by phage display resulting in 376 recombinant monoclonal phages that were evaluated by Phage-ELISA, 19 VHH-phages were positive to CD105, 9 phages showed a strong signal, 10 phages showed a signal just above the cut-off to be considered as positive, OD_{450nm} values three times higher than control. This variation in phage-ELISA values suggested that the screening resulted in different anti-CD105 VHH clones.The 19 VHH cDNAs were amplified and sequenced with specific primers. The DNA sequences were edited and translated with the ExPASY and GeneDoc tools and submitted to BLAST to determine their identity and uniqueness. The 19 sequences were found to be different VHH sequences, displaying the characteristic structural features such as four frameworks regions (FRs), three complementary-determining regions (CDRs), substitutions in FR2 which increase the hydrophilicity of the region, a CD3 of 16-18 amino acids and a conserved secondary structure consisting of 9 β-sheets and 8 loops as modelled by SwissModel. The VHH sequences were grouped according to the greater or lesser similarity between them.

2.1 Scientific context

2.1.1 Phage display

Phage display is a screening technique that relies on the interaction of a protein with high affinity and specificity to a target molecule and since its creation has been applied in a variety of areas, as fully and successfully described in several reviews[1-21]. In the field of biomedicine phage display was extensively used in the discovery of drugs and new antibiotics[22-27]; the identification of receptors ligands capable of crossing the blood-brain barrier[28,29], organ-or tissue-specific peptides[30,31], peptides that recognise cerebrovascular changes[32], protein mimicking peptides, such as, ice-binding proteins (IBPs)[33], the protein hormone erythropoietin[34,35], enzyme inhibitors[36,37]; for diagnosis and therapy against pathogens and toxins[5,38,39]; in cancer theranostics[40] and in cancer probes and therapies[41]; in the study of protein-protein interactions, DNA-protein or cDNA expression[13,42,43]; in neurobiology[13]; for the design of human antibodies[44,45] and vaccines[46]; in optimising the specificity and mapping of antibody epitopes[47-50]; in mapping human vasculature[51]; in the search for ligands that bind to target proteins by covalent bond formation[52] or peptide ligands that bind to small molecules such as, prostaglandin E2[53], 15-ketocholestane[54], taxol[55] and even explosive dinitrotoluene derivatives[56]; it has also been used in agriculture[57]; in nanostructured electronics[10], and more applications in other fields.

The phage display technique involves the production of a set of phages that express different peptides, proteins or antibodies on their surface. From this pool, a process called bio-panning selects phages that bind to the desired target. In order to perform phage display (exposure or presentation), the cDNA sequences of the peptides, proteins or

antibodies are cloned into the phage genome to be expressed as a fusion protein with the coat proteins of the phage[58,59]. Thus, phage libraries can be constructed, evaluated and screened by binding to the desired target. The screened ligands that bind the target can be isolated for identification by DNA sequencing[60,61] and then analysed to gain a better understanding of the ligand-target interaction and to obtain information for their application[21].

The main advantages of this technique include that is cost-effective, fast and bio-panning rounds can be repeated as many times as more rigorous selection is required. Also, its ability to produce a high number of diverse exogenous peptides or proteins that display on the phage surface by rapid standard molecular biology techniques rather than using genetically modified proteins or individual peptide variants.

Multiple proteins, antibodies and peptides of different sizes have been displayed by this technique, such as: the 60 kDa alkaline phosphatase[62], the 11 kDa cytochrome b562[63], the 6.5 kDa SH3 (Src homology 3) domain[64], the 7 kDa mustard trypsin inhibitor[65], and more recent ones include B-cell epitopes against coronavirus disease 19 (COVID-19)[49], antibodies against various epitopes of the receptor binding domain (RBD) of SARS-CoV-2[50,66], monoclonal antibodies specific for MERS-CoV nucleoprotein[67], serotype-specific monoclonal antibodies for dengue virus (DENV) non-structural protein 1 (NS1)[68]; among several others.

2.1.2 Production of recombinant VHH proteins

The production of VHHs against antigens is a straightforward procedure that has been established[69,70]. In general, the first step is to generate a VHH cDNA library from mRNA of PBMCs from the serum of the animal immunised with the antigen of interest, once the circulating immunoglobulin in the serum has been raised against the antigen. Then, the library is screened for the VHHs that specifically bind the antigen by phage display and phage-ELISA. Finally, the selected VHH are cloned and expressed in a vector/host expression system and the recombinant VHH are tested to determine the affinity, specificity and other characteristics related to the binding to the antigen.

2.1.2.1 Generation of VHH cDNA libraries

Alpacas (and other camelids) are immunised with a preparation of the antigenic protein with adjuvants. The immunisation is monitored by ELISA, when a high titre of circulating immunoglobulins against the antigenic protein is obtained, PBMCs, which are mainly the antibody-expressing lymphocytes, are isolated from blood samples. The mRNA is extracted from the PBMCs and cDNAs are synthesized by reverse transcription[71,72]. Through PCR using primers specific for the VHH domain of the HCAbs, VHH cDNA are amplified and subsequently cloned into a specific vector according to the antibody selection method[73], such as, a phagemid vector[74]. The VHH cDNA library is ready for screening (Figure 2.1)

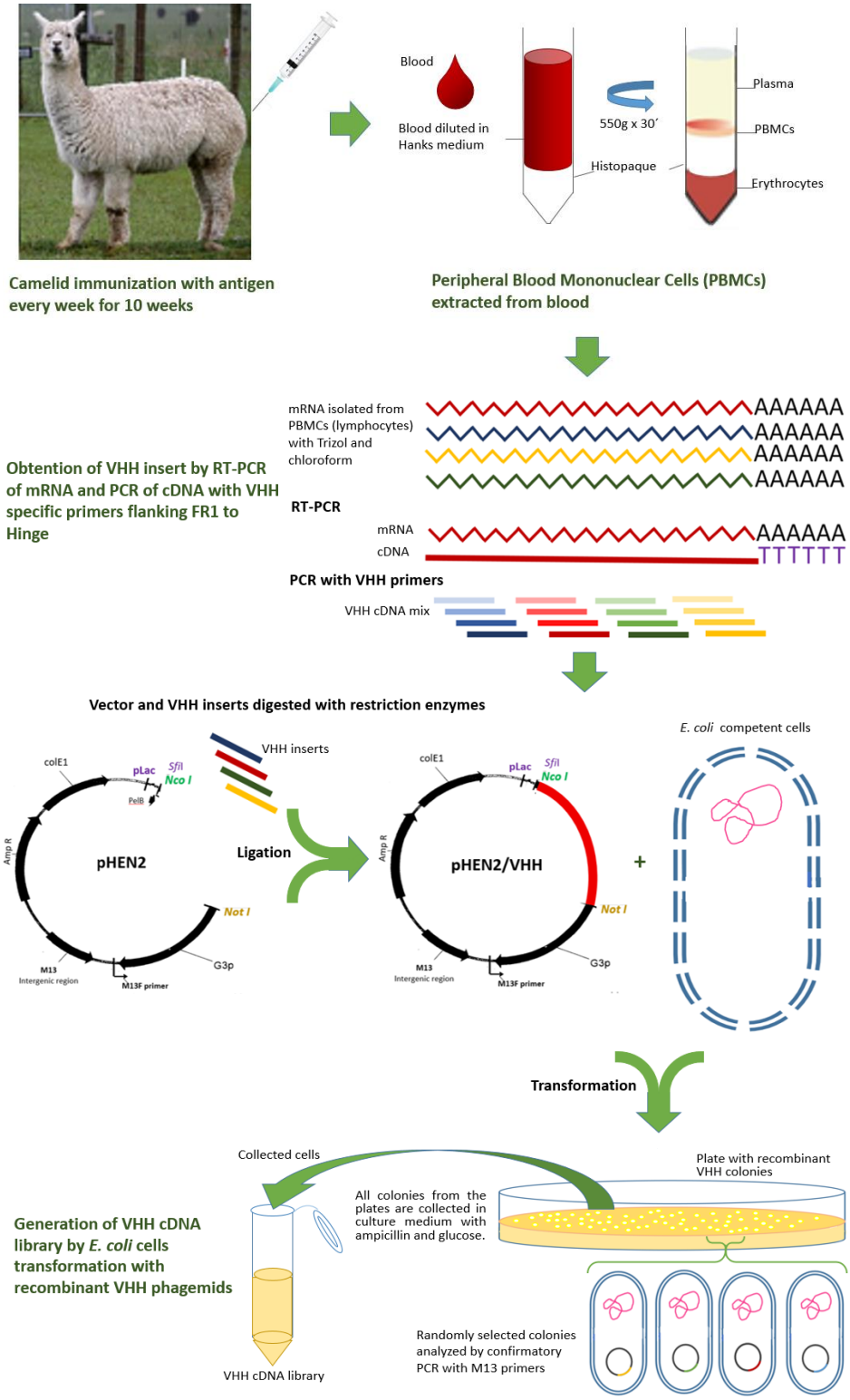


Figure 2.1- Generation of VHH cDNA libraries. Camelids are immunised with the antigen of interest. After the immune response develops, peripheral blood mononuclear cells (PBMCs) are extracted to obtain lymphocytes. mRNA isolated from lymphocytes is used to obtain VHH cDNAs by reverse transcription PCR (RT-PCR) and PCR with primers specific for the VHH region. Both the VHH cDNAs PCR products (inserts) and the phage vector are cleaved with the same digestion enzymes and subsequently join them by ligation. *Escherichia coli* (*E. coli*) cells are transformed with the recombinant phage vector (vector with VHH cDNA insert) and analysed by verification PCR. Recombinant colonies are collected in culture broth supplemented with ampicillin and glucose. The VHH cDNA library is ready for screening by phage display. (Adapted from Pardon *et al*[69] and Dmitriev *et al*[70])

2.1.2.2 Selection of specific VHH

The widely used technique for the selection of VHH is phage display[11,16,20,75]. Through this technique, the VHH cDNA sequences are cloned into a phagemid vector carrying the gene encoding one of the five phage coat proteins (pIII), so that, when the phages are assembled, VHH protein is expressed as a fusion to its coat proteins[73,74]. The phages are exposed to the antigen of interest in order to pick those expressing VHH that bind to the antigen. The VHH-displayed phages library is screened by selection rounds of binding and washing the phages to the immobilised antigen, rounds of bio-panning[69,73]. After three or four rounds of bio-panning, the *E. coli* cells with the mixture of phages that bind to the chosen antigen (polyclonal phages) are plated on agar plates. Individual colonies are picked and expanded for the production of monoclonal VHH-displayed phages (Figure 2.2a). The monoclonal phages are tested by Phage-ELISA to confirm antigen binding[76].

In Phage-ELISA, the phage groups are incubated on an antigen-coated plate and then, after washing, revealed with an enzyme-conjugated anti-phage antibody directed against the major capsid protein and a chromogenic substrate specific for the chosen enzyme. If these groups bind to the desired antigen (as compared to phage control libraries), the enzyme produces a colour change. Each selection is sequenced and the obtained coding DNA sequences can be analysed for identity and uniqueness[69,74,76]. The different sequences identified as VHH are selected for expression and subsequent affinity testing of the recombinant VHH protein (Figure 2.2b).

2.1.2.3 Affinity testing of selected VHH

Selected VHH being simple, soluble and exempt from the need for post-translational modification or assembly (as required by Fab or scFV) can be expressed in both prokaryotic and eukaryotic vector/host systems. In *E. coli*, the VHHs are attached to a polyhistidine tag (polyHis-tag) to allow purification by affinity chromatography in a nickel column and also contain a secretion signal sequence to target them to the periplasm to facilitate both purification and disulfide bond formation[70]. Large amounts of VHH proteins can be obtained per litre of bacterial culture and requirements are easily standardised[77]. PolyHis-tagged VHH can also be purified by affinity chromatography using protein A[78]. Purified recombinant proteins can be analysed to provide a more detailed understanding of the antigen-antibody interaction and to assess their affinity and specificity.

2.1.3 Applications of recombinant VHH

The VHHs are small, soluble, stable antibodies that can be produced in a cost-effective process. The VHHs have the capacity to enter tissues and a short half-life in blood circulation. They are amenable of humanization by genetic engineering, which would decrease immunogenicity without losing affinity for their antigen[79-84]. VHHs can be linked to peptide tags, toxins, Fc domains and other VHHs and conjugated at specific sites to drugs, radionuclides, photosensitisers or nanoparticles[85]. These characteristics make VHH suited for developing applications in the diagnosis and therapy of diseases, and as research tools[72,79,85-104].

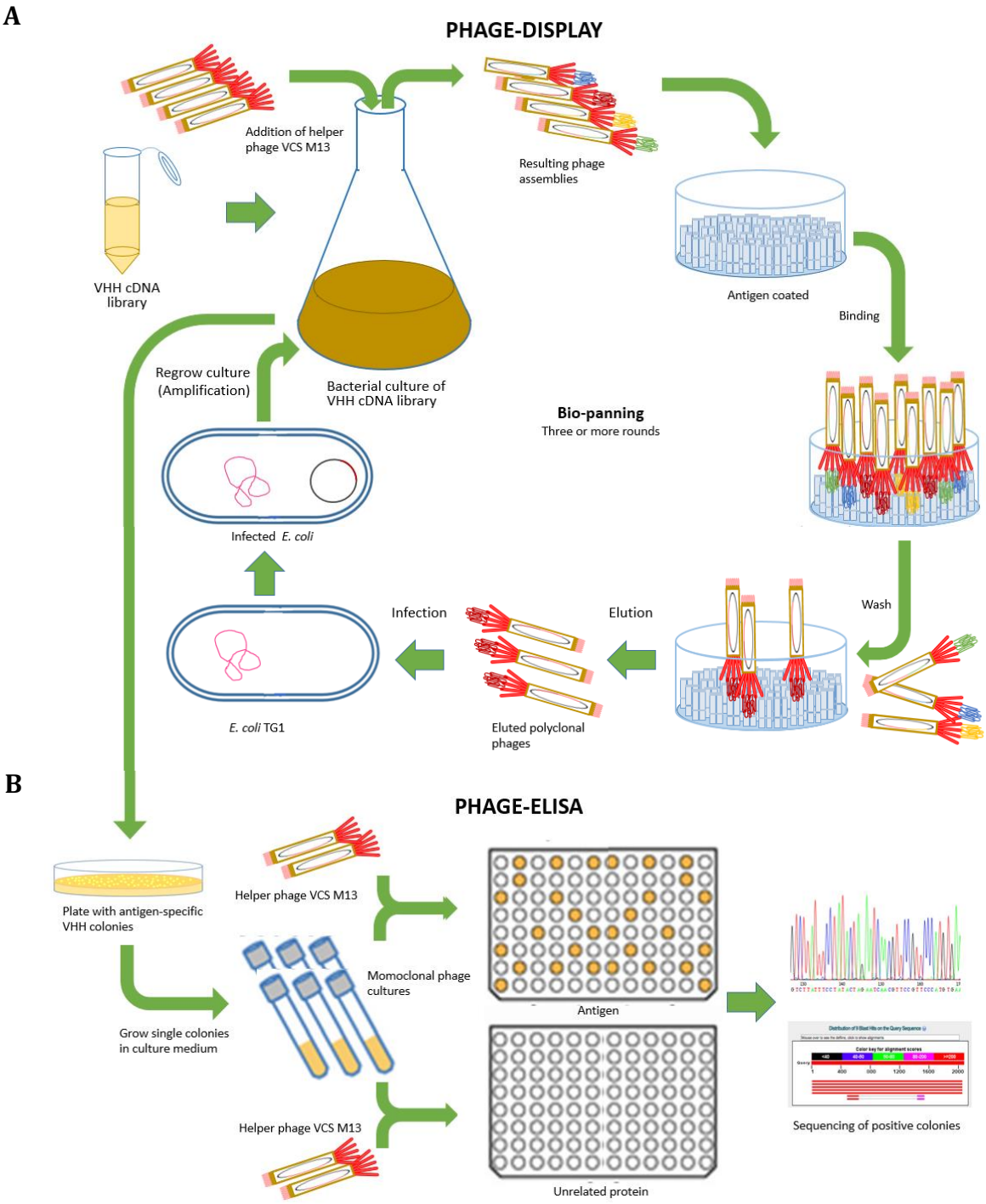


Figure 2.2- Selection of antigen-specific VHH by phage display and phage-ELISA. a. Phage display of filamentous bacteriophage M13 using a phagemid vector carrying the gene encoding one of the five phage coat proteins (pIII). Phage libraries are obtained by transforming *E. coli* cells with phagemids containing the various variants of VHH cDNAs and rescuing the phages with VCS M13 helper phages. Thus, phages displaying the specific binding VHH against the immobilized antigen (polyclonal phages) can be selected and isolated by several rounds of biopanning. These include binding, washing, elution, infection and amplification. b. Phage-ELISA. The culture of *E. coli* infected with the eluted phages displaying the antigen-specific binding VHH is grown on plates to obtain individual colonies to allow individualisation of the phages (monoclonal phages) which are subsequently screened by ELISA and sequenced prior to expression and purification of their proteins. (Adapted from Herng Leow *et al*[4]).

The VHHs against different biomedically important targets have been generated[105], such as: N14, C24, C9 (anti-PSMA, prostate specific membrane antigen)[106] and C3 (anti-PSMA)[107] for prostate cancer; cAb-CEA5 (anti-CEA, carcinoembryonic antigen)[108] and JJB-B2 (anti-CEA)[109] for breast, ovarian, colorectal and pancreatic cancer. RR-B7, RR-D40 (anti-MUC1, mucin 1)[110,111], C3, E2 (anti-PD-L1, Programmed Death-ligand 1)[112] and K2 (anti-PD-L1)[113] for advanced solid tumours; Anti-TNF-VHH (anti-TNFα, tumor necrosis factor α)[114] for solid tumours. 11A4 (anti-HER2, human epidermal growth factor receptor-2)[115] and K24 (anti-CAIX, carbonic anhydrase IX)[116] for breast cancer; VHH6 (anti-CD7)[117,118] for leukaemia; HuNb1-IgG4 (anti-CD47)[119] for colorectal and ovarian cancer, acute myeloid leukaemia (AML) and non-Hodgkin's lymphoma (NHL). Nb_7, Nb_21, Nb_22 (anti-CD33)[120] for AML; H11 (anti-CTLA-4, Cytotoxic T-Lymphocyte Antigen 4)[121] for different types of cancer; MU738, JK2, WF211, MU523, and MU1067 (anti-CD38)[122] for multiple myeloma. NB5 (anti-CXCR7, CXC chemokine receptor 7)[123] for head and neck cancer; 127D1, 163E3 (anti-CXCR2)[124] for immunological patterns and solid tumours; V36, 81, 51, and B10 (anti-CD11b)[125] for immunological patterns. VHH4 (anti-MHC-II, major histocompatibility complex class II)[126] for graft-versus-host disease (GvHD); ALX-0081 (anti-von Willebrand factor)[127] for anti-thrombotic treatment; SRB-37, SRB-85 (anti-CD19)[128] for haematological diseases.

In contrast to VHHs, the structural properties of conventional antibodies, as mainly their size, are a limitation for their application in cancer therapy, where antibody-drug conjugates need to reach the tumour site to deliver the cytotoxic effect[105,129,130]. In addition, antigen recognition domains that are non-covalently bound by a hydrophobic interface hinder specificity. Therefore, the VHHs appear to be suited to overcome the limitations of conventional antibodies and their derivatives without losing efficacy and specificity[85,105,131].

In addition to cancer therapy, the VHHs are being developed to treat infectious diseases such as those caused by *E. coli*[132], the fungus *Malassezia furfur*[133], influenza A subtype H5N1[134], *Clostridium difficile*[135], and most recently, SARS-CoV-2[136]. The VHHs conjugated to cytotoxic drugs are also being developed as a signalling or checkpoint blockade therapy[105] or in VHH-mediated redirection of various types of transport and delivery systems (liposomes, micelles, nanoparticles, etc) for different types of cargoes (drugs, toxins, interference RNA, etc) to established sites[105,137-139].

The VHHs are also being applied to help crystallization of hard proteins and to resolve their structures, as membrane receptors, transporters, proteins of bacteria secretion systems, etc[140-146]. To identify preferred epitopes, the best co-crystals between VHH and the target or to reveal unknown target confomers[147], to map the location of protein domains by cryoelectronic microscopy[148], to disrupt domain interactions when the epitope is at the interface of interacting domains[149], etc.

2.2 Procedures

2.2.1 Obtaining the anti-CD105 VHH cDNA library

The VHH cDNA library was previously done by Gushiken in 2016[71] as follows: An alpaca was immunised with a lysate of T24, a cell line derived from a urine bladder cancer that over-express CD105 protein. The cDNA was synthesised by reverse transcription of mRNA extracted from alpaca lymphocytes. RT-PCR products were PCR amplified with VHH-specific primers (VHH-*Sfi* and VHH-R1*Not*)[150]. The VHH cDNAs were cloned into the phage pHEN2 obtaining a cDNA library of 1.2 x 10^8 VHH sequences. Phage display was used to capture the phages carrying anti-CD105 VHH. For this purpose, three rounds of bio-panning selection were performed using 1 µg, 0.8 µg and 0.6 µg of immobilised antigen on microtitre plates. Each phage-infected culture was placed on plates with 2xYT medium and the colonies obtained were recovered together as a whole to form the phage library carrying anti-CD105 VHH cDNA sequences.

2.2.2 Screening of anti-CD105 VHH cDNA library

20 µl of the final anti-CD105 VHH cDNA library were inoculated in 1 ml of 2xYT medium supplemented with 100 µg.ml^{-1} ampicilline and 2% Glucose (2xYT supp) and incubated at 37° C with gentle agitation overnight. 200 µl of culture were dispersed over 2xYT supp plates and incubated at 37° C overnight. Ninety-four single colonies were picked and placed in 100 µl of 2XYT supp medium in a 96-well flat-bottom culture plate. The plate was sealed with a porous film to allow air entry and incubated at 30° C with gentle agitation overnight. 2 µl of the culture from each well of the plate was inoculated in 248 µl of fresh 2XYT supp medium in a deep concave bottom 96-well plate. It was incubated at 37° C and 250 rpm for 2.5 hours to reach an OD_{600nm} between 0.5 and 0.6.

The initial plate (mother plate) was kept as stock at -70° C. Any plate obtained from the mother plate was referred to as a replicate plate. To each replicate plate culture, 3 x 10^9 of VCS-M13 helper phage dissolved in 50 µl of 2xYT supp medium were added to have a multiplicity of infection of 20. After resting for 30 minutes at 37° C, the infection period, the plate was centrifuged at 1000 g for 10 minutes and the precipitate from each well was re-suspended in 150 µl of 2xYT supp medium with Kanamycin 100 µg/ml. The plate was incubated at 30° C overnight at 125 rpm and centrifuged at 1000 g for 10 minutes. The supernatant, with the phages, were used for PHAGE-ELISA.

2.2.3 PHAGE-ELISA

Inside each well of the Maxisorb plate, 50 ng of the antigen (CD105 protein) dissolved in PBS 1X pH 7.4 were fixed and in another plate (control plate) 50 ng of an unrelated protein (recombinant Fas2) were fixed too. After fixation, each well of both plates were washed 5 times with 300 µl of Phage-ELISA wash solution. 100 µl of Phage-ELISA blocking solution were added to each well and the plates were left for 1 hour at 37° C. After that time, the plates were washed 5 times. 50 µl of concentrated Phage-ELISA blocking solution (4% milk) and 50 µl of a different phage (from each supernatant obtained in the previous process) were added to each well. The phages were inoculated once on both the CD105

plate and the unrelated protein control plate. The plates were incubated at 37° C for one hour and washed 5 times. The retained phages were treated with 200 µl of the anti-P8 secondary antibody diluted 1:2000 in 2% milk, incubated at 37° C for one hour and then washed 5 times. 100 µl of TMB substrate solution were added to each well, and incubated for 10 minutes in the dark. The reaction was stopped with 50 µl of 2N sulfuric acid. Absorbances were obtained by an ELISA plate reader at OD_{450nm} Positive reactions, meaning, with phage carrying anti-CD105 VHH, were considered those that had a 3-fold higher absorbance with respect to the control[151,152], the unrelated protein, recombinant Fas2. The positive control was the culture of pHEN-EngVHH17 recombinant clone[71]. Phages from this culture were used for the determination of the minimum amount of antigen needed for a good signal in the Phage-ELISA, prior to library screening, 25, 50, 100, 200 and 300 ng of CD105 antigen were tested.

2.2.4 Extraction of recombinant pHEN2-VHH DNA

The Wizard plus SV Minipreps DNA Purification System Kit (PROMEGA) was used and the manufacturer's alkaline lysis protocol described below was followed.

Each selected recombinant colony, meaning, carrying a different anti-CD105 VHH sequence, were inoculated in 5 ml LB medium supplemented with 100 $ug.ml^{-1}$ Ampicilline (LB supp) and incubated at 37° C overnight on shaking at 250 rpm. Cells were collected by centrifugation at 12000 rpm for 5 minutes at 4° C and re-suspended in 250 µl of Cell Resuspension Solution using vortexing. 250 µl of Cell Lysis Solution were added and gently mixed by inversion 5 times (without vortexing) and incubated until the cell suspension cleared (approximately 1 to 5 minutes). 10 µl of Alkaline Protease Solution were added, mixed by inversion 5 times, and allowed to stand for 5 minutes. 350 µl of Neutralization Solution were added, immediately mixed by inversion four times to disperse the entire solution and centrifuged at 13000 rpm for 10 minutes. The supernatant, approximately 850 µl, were transferred to a minicolumn over a microtube and centrifuged at 13000 rpm for 1 minute. 750µl of Column Wash Solution, previously diluted with 95% ethanol, were added to the minicolumn and centrifuged at 13000 rpm for 1 minute. The minicolumn was washed again with 250µl of Column Wash Solution and centrifuged at 13000 rpm for 2 minutes. The minicolumn was transferred to a new sterile microtube and the DNA was eluted by adding 100µl of nuclease-free ultrapure water and centrifuged at 13000 rpm for 1 minute. The DNA obtained was visualized by 1.5% agarose gel electrophoresis in 1X TBE buffer and stained with ethidium bromide. Electrophoresis was performed at 80 V for 90 minutes in the presence of molecular-weight size marker GeneRuler100 bp Plus DNA ladder (ThermoScientific). Finally, DNA was quantified by Nanodrop at 260 nm and stored at -20° C until use.

2.2.5 PCR amplification of anti-CD105 VHH at pHEN2 with AHis and M13R primers for sequencing

The PCR amplification of the VHHs proceeded in 50 µl 1X PCR reaction buffer (Invitrogen) containing 10 ng DNA, 1mM $MgCl_2$, 0.5 µM of each primer (AHIS and M13), 200 µM of each dNTP (dATP, dCTP, dGTP, dTTP) and 0.2 units of Platinum Taq DNA (Invitrogen). An initial denaturation step at 94° C for 2 minutes, followed by 30 amplification cycles of 94° C for 45 seconds, 64° C for 90 seconds and 72° C for 120 seconds, the final extension 10 minutes at 72° C and a final cooling step at 4° C were used. The reactions were performed in triplicate, at the end of the PCR they were pooled. The amplification products were resolved by 1.2% agarose gel electrophoresis in 1X TBE buffer at 90 V for 70 minutes and visualized by staining with ethidium bromide. Molecular-weight size marker GeneRuler100 bp Plus DNA ladder (ThermoScientific) were used to estimate the size of the amplification products.

2.2.6 Purification of PCR products for sequencing

Purification of PCR products was performed with QIAquick Gel Extraction Kit (Qiagen) following manufacturer procedure[153]: PCR product reaction was diluted 1:6 v/v in PB buffer and loaded into a minicolumn that was previously placed in a 2 ml microtube. It was centrifuged for 1 minute at 13000 rpm and the eluted buffer was discarded. The minicolumn was washed by adding 750 ul of Buffer PE and centrifuging for 1 minute. After discarding the eluted buffer, the minicolumn was centrifuged for a further 1 minute. The minicolumn was placed in a new sterile 1.5 ml microtube and the DNA was eluted by adding 50 µl of Buffer EB or nuclease-free ultrapure water and centrifuging at 13000 rpm for 1 minute. The DNA concentration was quantified by absorbance at 260 nm in Nanodrop.

2.2.7 VHH sequence analysis

The anti-CD105 VHH sequences were obtained through the sequencing service of the Macrogen company (https://dna.macrogen.com/) using the AHis and M13R primers and 20 µl of the PCR product purified at 50 ng/µl. Sequences were edited using Mega 7 software and the tools Reverse Complement from Bioinformatics.org. (https://www.bioinformatics.org/sms/rev_comp.html), and LALIGN from ExPaSY (Expert Protein Analysis System) server of the Swiss Institute of Bioinformatics (SIB), (https://www.expasy.org). The nucleotide sequences of each VHH were aligned with the MUSCLE (Multiple Sequence Comparison by Log-Expectation) programme of the European Molecular Biology Laboratory-European Bioinformatics Institute (EMBL-EBI) (https://www.ebi.ac.uk/Tools/msa/muscle/), analysed and translated for characterisation using the tool Translate from ExPaSY server.

To obtain the three-dimensional structure of the VHH, homology modeling was performed by the SwissModel Expasy server[154]. The identity of each sequence was determined by homology with other camelid VHH sequences reported in the databases through the sequence alignment software BLAST (Basic Local Alignment Search Tool[155])

from the National Center for Biotechnology Information (NCBI), (https://blast.ncbi.nlm.nih.gov).

The tool ProtParam from ExPaSY server calculated the physico-chemical characteristics of each VHH. These were: molecular weight (MW), isoelectric point (pI), amino acid composition, total number of negatively charged amino acid residues (D+E), total number of positively charged amino acid residues (R+K), atomic composition, total number of atoms, molar extinction coefficient[156], half-life, instability index[157], aliphatic index[158] and Grand Average of Hydropathy (GRAVY)[159]. Modelling the three-dimensional structure of the VHH was performed by the SwissModel Expasy server[154].

2.3 Results and Discussion

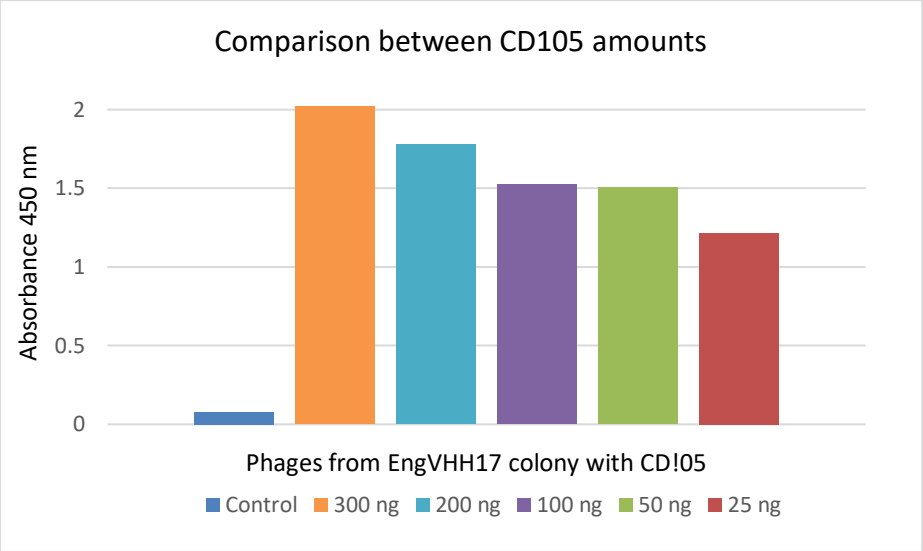
The VHH region of HCAbs is of interest because it is the smallest independent portion of the antibody with the capacity to bind the antigen with high affinity and specificity. To obtain VHHs that bind to human CD105 membrane protein, VHH cDNA library was screened with the recombinant sCD105.

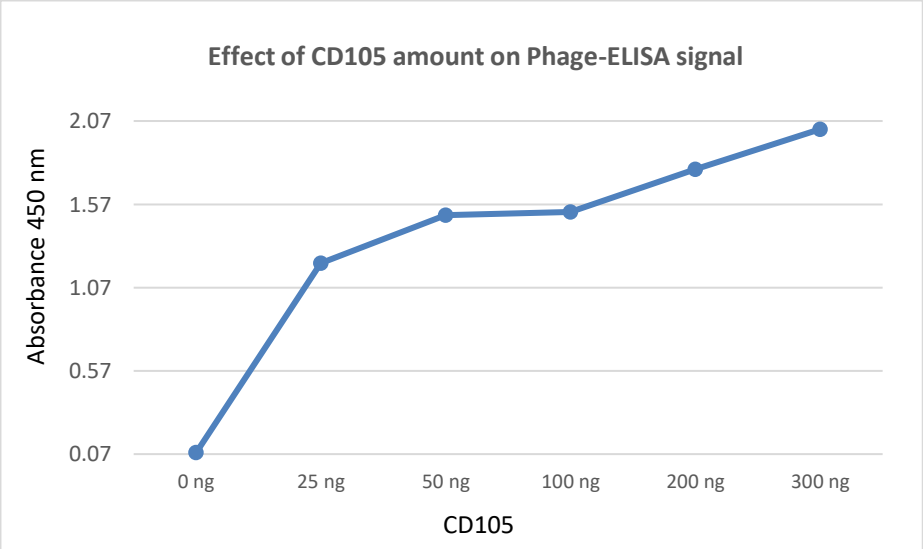
For this purpose, *E. coli* culture infected with polyclonal phages, from the third round of bio-panning of the cDNA library (library titre of 2.7 x 10^7 phage/ml) with immobilised sCD105, was plated. 376 colonies carrying monoclonal phages were randomly picked and screened by phage-ELISA in four 96-well plates with the CD105 antigen, and in four plates with the unrelated recombinant protein Fas2 as negative control to discard those reacting with the blocking agent or with the well, that is, positive for both CD105 and the negative control. Colonies whose absorbance value for CD105 was at least 3 times higher than the value for the negative control were considered positive[151,152].

In order to optimize VHH library screening with CD105, a titration of EngVHH17, a VHH with proven specificity to CD105 previously obtained[71], was performed with 5 amounts of antigen ranging from 25 ng to 300 ng (Figure 2.3), the latter amount being used initially to ensure a positive response and it was not limiting for the reaction. The antigen, the membrane protein CD105, is expensive, so determining the minimum amount of antigen needed to obtain a good signal and reliable positive ELISA reaction was of utmost. It was determined that 50 ng of antigen provided a signal as well as 100 ng (almost 20 times more than control) and better than 300 ng, whose signal exceeds the optimal absorbance range. This is because in the high absorbance zone, deviations occur as a consequence of the light that manages to pass is almost zero and is confused with the background noise or as a product of the species that form the sample interacting and deviating from linearity[160].

Nineteen colonies were considered as positive, meaning, nineteen colonies carrying VHH-dsplayed phages with OD_{450nm} higher than the cut-off value (3 times stronger than the control) were considered as colonies carrying anti-CD105 VHH-displayed phages (Figure 2.4A and B).

Phage-ELISA of EngVHH17 with different amounts of CD105





PHAGE ELISA	EngVHH17 phages	
CD105	Abs 450 nm	St. Desv.
Control	0.078	0.0006
25 ng	1.215	0.0015
50 ng	1.505	0.0040
100 ng	1.524	0.0060
200 ng	1.779	0.0025
300 ng	2.021	0.0085

Figure 2.3- Determination of antigen concentration for phage-ELISA.

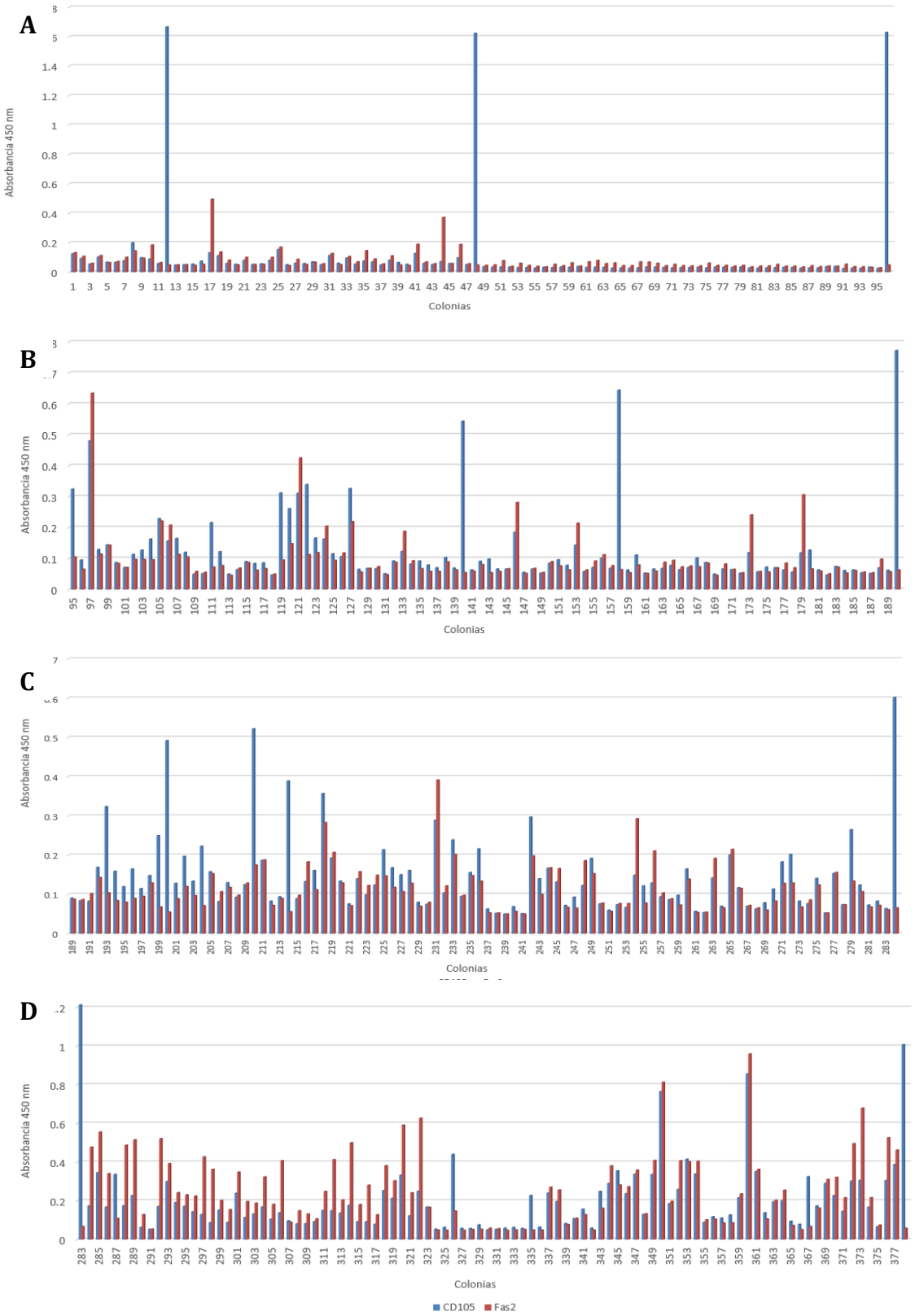


Figure 2.4A- Screning of recombinant phage library for selection of CD105-specific phages by phage-ELISA. Phages that reacted with CD105 with an absorbance 3-fold higher than the negative control, Fas2 were considered positive. a) Phages from colonies 12 and 48 were positive, b) Phages from colonies 95, 111, 119, 122, 140 and 158 were positive, c) Phages from colonies 193, 199, 200, 204, 210 and 214 were positive and d) Phages from colonies 283, 287, 326, 335 and 367 were positive. The phages from colony EngVHH17 (Control (+)) were located in the last well. The control without phages (Control (-)) was placed in the penultimate well.

a

PLACA I CON CD105

	1	2	3	4	5	6	7	8	9	10	11	12
A	0.125	0.098	0.133	0.154	0.096	0.128	0.037	0.031	0.029	0.032	0.028	0.038
B	0.093	0.09	0.113	0.051	0.055	0.06	0.034	0.035	0.028	0.034	0.03	0.039
C	0.056	0.059	0.06	0.061	0.076	0.052	0.036	0.034	0.031	0.03	0.032	0.026
D	0.104	1.666	0.054	0.059	0.07	0.073	0.035	0.039	0.036	0.035	0.033	0.029
E	0.068	0.049	0.08	0.07	0.05	0.059	0.032	0.033	0.035	0.035	0.033	0.027
F	0.067	0.051	0.053	0.051	0.082	0.098	0.032	0.035	0.032	0.032	0.027	0.035
G	0.078	0.055	0.057	0.114	0.066	0.053	0.028	0.034	0.03	0.035	0.03	0.041
H	0.2	0.075	0.081	0.061	0.054	1.622	0.034	0.029	0.031	0.029	0.029	1.629

PLACA I CON FAS2

	1	2	3	4	5	6	7	8	9	10	11	12
A	0.133	0.094	0.496	0.17	0.107	0.19	0.047	0.054	0.045	0.043	0.041	0.042
B	0.108	0.185	0.137	0.045	0.071	0.07	0.05	0.044	0.045	0.043	0.043	0.041
C	0.061	0.065	0.082	0.088	0.146	0.058	0.079	0.064	0.07	0.062	0.053	0.055
D	0.113	0.049	0.05	0.053	0.09	0.372	0.04	0.041	0.069	0.046	0.041	0.038
E	0.065	0.05	0.101	0.068	0.057	0.059	0.061	0.071	0.061	0.047	0.041	0.037
F	0.073	0.052	0.053	0.058	0.111	0.189	0.046	0.08	0.044	0.042	0.036	0.033
G	0.102	0.046	0.053	0.127	0.049	0.058	0.04	0.059	0.054	0.047	0.044	0.041
H	0.145	0.053	0.101	0.052	0.046	0.049	0.036	0.063	0.045	0.037	0.038	0.049

b

PLACA II CON CD105

	1	2	3	4	5	6	7	8	9	10	11	12
A	0.324	0.127	0.216	0.312	0.326	0.092	0.098	0.096	0.063	0.102	0.072	0.074
B	0.095	0.163	0.122	0.261	0.065	0.079	0.066	0.078	0.111	0.087	0.07	0.061
C	0.48	0.229	0.05	0.311	0.068	0.07	0.066	0.143	0.053	0.05	0.063	0.063
D	0.129	0.157	0.064	0.339	0.067	0.103	0.185	0.058	0.066	0.066	0.056	0.054
E	0.144	0.165	0.09	0.166	0.051	0.069	0.055	0.071	0.068	0.064	0.118	0.052
F	0.087	0.12	0.084	0.163	0.092	0.544	0.066	0.101	0.078	0.053	0.127	0.07
G	0.071	0.05	0.086	0.115	0.123	0.063	0.053	0.068	0.064	0.119	0.063	0.062
H	0.113	0.051	0.047	0.106	0.082	0.091	0.085	0.644	0.071	0.057	0.047	0.772

PLACA II CON FAS2

	1	2	3	4	5	6	7	8	9	10	11	12
A	0.105	0.097	0.072	0.095	0.219	0.067	0.054	0.076	0.054	0.073	0.056	0.072
B	0.065	0.096	0.077	0.148	0.056	0.058	0.058	0.063	0.079	0.085	0.07	0.053
C	0.634	0.221	0.046	0.425	0.068	0.058	0.067	0.214	0.052	0.046	0.085	0.061
D	0.114	0.208	0.069	0.112	0.074	0.089	0.281	0.063	0.059	0.082	0.07	0.056
E	0.143	0.113	0.087	0.119	0.048	0.063	0.052	0.092	0.088	0.065	0.306	0.054
F	0.084	0.105	0.062	0.205	0.088	0.054	0.068	0.112	0.094	0.054	0.066	0.098
G	0.071	0.058	0.067	0.094	0.188	0.059	0.055	0.077	0.073	0.241	0.059	0.057
H	0.097	0.055	0.05	0.118	0.093	0.08	0.09	0.064	0.076	0.058	0.051	0.063

c

PLACA III CON CD105

	1	2	3	4	5	6	7	8	9	10	11	12
A	0.091	0.115	0.158	0.094	0.076	0.08	0.063	0.132	0.067	0.057	0.079	0.154
B	0.084	0.148	0.082	0.389	0.14	0.075	0.052	0.072	0.149	0.054	0.114	0.074
C	0.083	0.25	0.13	0.089	0.099	0.289	0.05	0.093	0.122	0.142	0.183	0.265
D	0.169	0.492	0.093	0.133	0.124	0.104	0.069	0.123	0.129	0.07	0.202	0.124
E	0.324	0.128	0.125	0.161	0.214	0.239	0.051	0.192	0.094	0.201	0.083	0.073
F	0.159	0.197	0.522	0.357	0.168	0.095	0.297	0.076	0.087	0.117	0.077	0.083
G	0.12	0.134	0.187	0.193	0.15	0.156	0.14	0.06	0.098	0.07	0.141	0.064
H	0.165	0.223	0.083	0.134	0.161	0.216	0.167	0.074	0.165	0.063	0.053	0.692

PLACA III CON FAS2

	1	2	3	4	5	6	7	8	9	10	11	12
A	0.088	0.095	0.153	0.09	0.071	0.07	0.053	0.166	0.077	0.054	0.06	0.156
B	0.087	0.129	0.107	0.056	0.158	0.08	0.053	0.067	0.293	0.055	0.083	0.074
C	0.102	0.068	0.118	0.098	0.123	0.392	0.05	0.065	0.078	0.192	0.128	0.134
D	0.143	0.055	0.098	0.183	0.149	0.122	0.057	0.186	0.211	0.066	0.129	0.107
E	0.104	0.089	0.129	0.112	0.147	0.202	0.05	0.153	0.104	0.215	0.068	0.068
F	0.084	0.12	0.175	0.283	0.118	0.098	0.198	0.078	0.089	0.115	0.086	0.072
G	0.08	0.097	0.188	0.207	0.107	0.148	0.1	0.057	0.073	0.072	0.124	0.061
H	0.09	0.071	0.072	0.129	0.128	0.134	0.168	0.077	0.139	0.066	0.053	0.066

d

PLACA IV CON CD105

	1	2	3	4	5	6	7	8	9	10	11	12
A	1.21	0.05	0.15	0.1	0.09	0.17	0.05	0.08	0.34	0.09	0.2	0.15
B	0.17	0.17	0.09	0.08	0.09	0.05	0.06	0.11	0.13	0.12	0.2	0.3
C	0.35	0.3	0.24	0.08	0.08	0.06	0.06	0.16	0.34	0.11	0.1	0.31
D	0.17	0.19	0.11	0.09	0.25	0.44	0.06	0.06	0.77	0.13	0.08	0.17
E	0.34	0.17	0.13	0.15	0.21	0.06	0.23	0.25	0.19	0.22	0.33	0.07
F	0.18	0.14	0.17	0.15	0.33	0.06	0.06	0.29	0.26	0.86	0.17	0.31
G	0.23	0.13	0.11	0.14	0.12	0.08	0.24	0.36	0.42	0.35	0.29	0.39
H	0.06	0.09	0.14	0.18	0.25	0.05	0.2	0.24	0.34	0.14	0.23	1.01

PLACA IV CON FAS2

	1	2	3	4	5	6	7	8	9	10	11	12
A	0.07	0.05	0.2	0.09	0.18	0.17	0.06	0.08	0.36	0.1	0.2	0.22
B	0.48	0.52	0.16	0.15	0.28	0.05	0.05	0.11	0.13	0.11	0.26	0.5
C	0.56	0.39	0.35	0.13	0.13	0.05	0.05	0.13	0.41	0.09	0.07	0.68
D	0.34	0.24	0.2	0.11	0.38	0.15	0.05	0.05	0.81	0.09	0.05	0.22
E	0.11	0.23	0.19	0.25	0.3	0.05	0.05	0.16	0.2	0.24	0.07	0.08
F	0.49	0.23	0.32	0.41	0.59	0.05	0.05	0.38	0.41	0.96	0.16	0.53
G	0.52	0.43	0.18	0.21	0.24	0.05	0.27	0.28	0.4	0.36	0.31	0.46
H	0.13	0.36	0.41	0.5	0.63	0.06	0.26	0.27	0.41	0.11	0.32	0.06

e

Colony-well		CD105 Abs_{450nm}	St. Desv	Control Abs_{450nm}	St. Desv	CD105/control ratio
12	D2-I	1.666	0.0032	0.049	0.002	34
48	H6-I	1.622	0.0027	0.049	0.002	33
95	A1-II	0.324	0.0038	0.105	0.003	3
111	A3-II	0.216	0.0024	0.072	0.003	3
119	A4-II	0.312	0.0021	0.095	0.004	3
122	D4-II	0.339	0.0033	0.112	0.003	3
140	F6-II	0.544	0.0023	0.054	0.003	10
158	H8-II	0.644	0.0022	0.064	0.003	10
193	E1-III	0.324	0.0018	0.104	0.004	3
199	C2-III	0.25	0.0035	0.068	0.004	4
200	D2-III	0.492	0.0031	0.055	0.002	9
204	H2-III	0.223	0.0029	0.071	0.003	3
210	F3-III	0.522	0.0021	0.175	0.003	3
214	B4-III	0.389	0.0037	0.056	0.004	7
283	A1-IV	1.21	0.0029	0.07	0.003	17
287	E1-IV	0.34	0.0031	0.11	0.003	3
326	D6-IV	0.44	0.0019	0.15	0.002	3
335	E7-IV	0.23	0.0041	0.05	0.003	5
367	E11-IV	0.33	0.0034	0.07	0.003	5

Figure 2.4B- Absorbance readings at 450 nm on phage-ELISA plates. The first with CD105 and the second with Fas2 (unrelated protein control). In yellow, the positives that showed a signal for CD105 more than 3-fold higher than the control. In purple, the positive control EngVHH17. a) Colonies from 1 to 94, b) Colonies from 95 to 188, c) Colonies from 189 to 282, d) Colonies from 283 to 376 and e) 19 selected colonies and their CD105/control ratio.

The phage-ELISA, on the first plate resulted in 2 positive colonies: n°12 (named D2-I because of its position on plate 1) and n° 48 (H6-I) with a signal almost 35 times stronger than the control. In addition, on the second, third and fourth plates, 17 positive colonies were found. 7 had an OD_{450nm} 17 times stronger than the control in the case of n° 283 (A1-IV), almost 10 times in n° 140 (F6-II), n° 158 (H8-II) and n° 200 (D2-III), and almost 6 times in n° 214 (B4-III), n° 335 (E7-IV) and n° 367 (E11-IV). While the other 10, n° 95 (A1-II), n° 111 (A3-II), n° 119 (A4-II), n° 122 (D4-II), n° 193 (E1-III), n° 199 (C2-III), n° 204 (H2-III), n° 210 (F3-III), n° 287 (E1-IV) and n° 326 (D6-IV), had the minimum to be considered positive. The variation in OD_{450nm} values was suggestive not only the diversity of anti-CD105 VHH-displayed phages fished out from the phage display library screening but also the different degrees of affinity of the VHH to the antigen.

The cDNA sequences of the 19 anti-CD105 VHH were obtained by sequencing performed by Macrogen using the AHis and M13R primers flanking the amber codon and the ribosome-binding site (RBS) respectively in the pHEN2 vector. (Figure 2.5 and 2.6). The sequences were designated as: Seq VHH1 (D2-I), Seq VHH2 (H6-I), Seq VHH3 (F6-II), Seq VHH4 (H8-II), Seq VHH5 (A1-II), Seq VHH6 (D4-II), Seq VHH7 (A4-II), Seq VHH8 (A3-II), Seq VHH9 (D2-III), Seq VHH10 (B4-III), Seq VHH11(H2-III), Seq VHH12 (C2-III), Seq VHH13 (E1-III), Seq VHH14 (F3-III), Seq VHH15 (A1-IV), Seq VHH16 (E11-IV), Seq VHH17 (E1-IV), Seq VHH18(D6-IV) y Seq VHH19 (E7-IV).

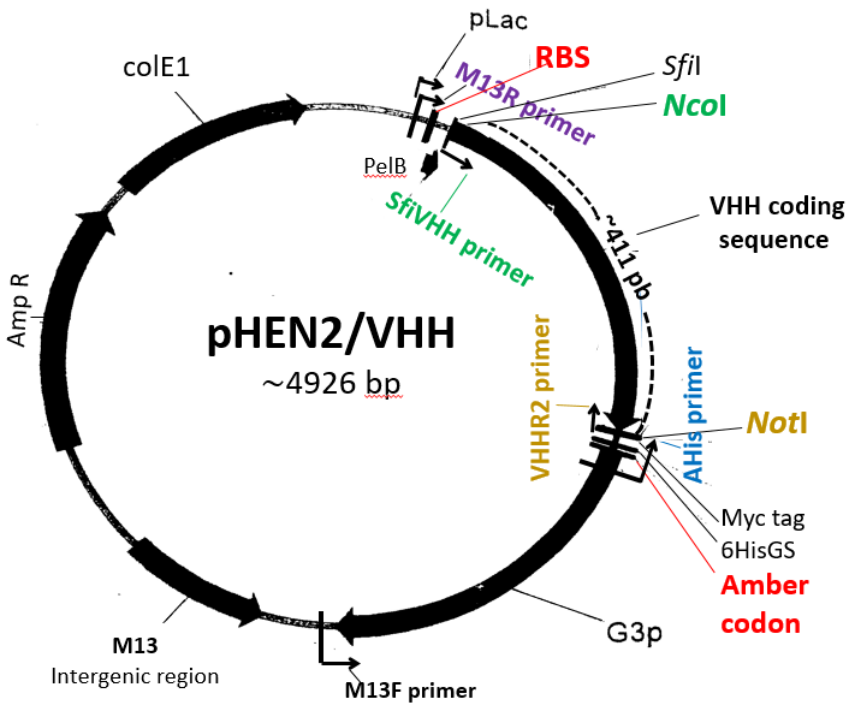


Figure 2.5- Scheme of pHEN2 phagemid vector showing VHH sequence located between the *Not*I (in brown) and *Nco*I (in green) enzyme cleavage sites. The location of the primers AHis (in blue) and M13R (in purple) and SfiVHH (in green) and VHHR2 (in brown) are also shown. RBS and amber codon in red

The average 630 bp anti-CD105 VHH sequences were edited with Mega software and aligned with the MUSCLE program in order to identify AHis and M13 primer sequences of pHEN2 vector and the VHH-specific primers (SfiVHH and ALLVHHR2) bearing restriction sites for the subcloning in the expression vector. The VHH protein coding sequences averaged 410 bp and 443 bp with the 6X His tag. (Figure 2.6). The VHH protein sequences had approximately 136 amino acids without a 6X His tag and 147 amino acids with a 6X His tag. All sequences were different, even though there were sequences that differed only in a single amino acid, such as 2 and 15 (F107Y), 4 and 18 (R30T), 12 and 18 (A50G), and in 2 amino acids such as 4 and 12 (R30T, G50A) (Figure 2.7a in brown on light green background).

The 6X His tag of the VHH proteins was added considering the sequence of the corresponding site in the expression vector, that is, after the *NotI* restriction site (Figure 2.6 and 2.7a) since recombinant VHH proteins would be produced from this system and not from pHEN 2 even though it also has it. In pHEN2, the 6X His tag is found after 17 amino acids AAAKQKLISEEDLNGAA whereas in pET22 it is after 5 amino acids AAALE, this smaller increment means a minimal change in the protein structure and higher stability, so it was preferred not to consider this sequence from pHEN2 for recombinant VHH.

The consensus sequence of the 19 aligned VHH protein sequences (Figure 2.7b) showed homology to the conserved domains of immunoglobulins, specifically to the variable region of the heavy chain (VH) (Figure 2.7c).

Sequences with a single amino acid difference could configure virtually identical VHH proteins depending on the location of the amino acid change in the sequence. However, with the single amino acid change in CDR3 (as in the case of VHH2 and VHH15) and in CDR1 (as in the case of VHH4 and VHH18) these would be distinct VHH proteins that may or may not have a different behaviour in their interaction with the antigen. Unlike if located in an FR region, where the change would not necessarily involve a different behaviour in the interaction.

The residue variations at positions 37, 44, 45 and 47 of FR2 in reference to VH is a characteristic of VHH[161] and they were present in all the VHH sequences (again verifying that they were indeed VHH sequences). As previously reported, these variations decrease hydrophobicity of the VHH region that in the corresponding VH region stabilises the interaction to VL. VHH lack a light chain, so, these residues increases the hydrophilicity, the solubility and prevents protein aggregation with other proteins or with other regions of the same protein.

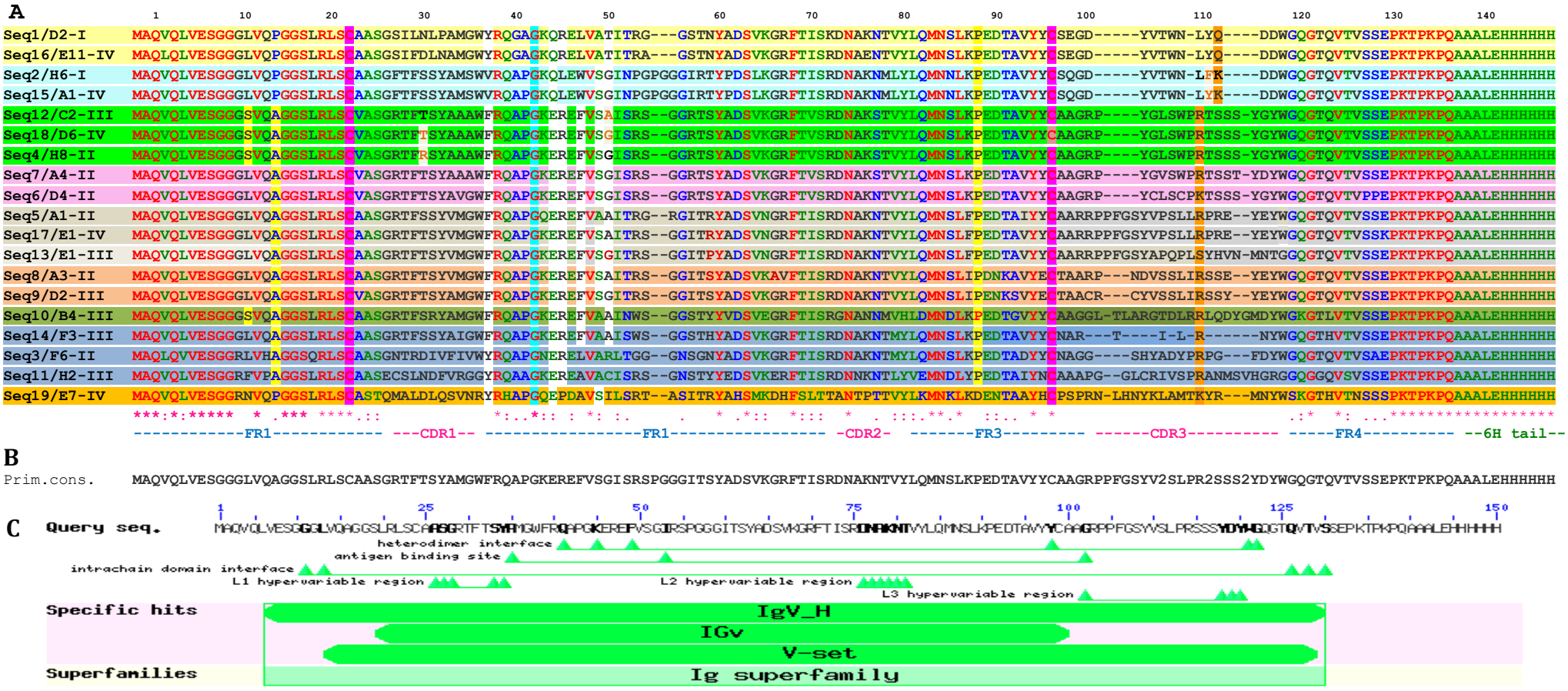
CLUSTAL multiple sequence alignment by MUSCLE (3.8)

```
                 M13R primer
Seq19/E7-IV      CAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCAAATTCTATTTCAAGGAGACAGTCATAATGAAATACCTATTGCCTACGGCAGCCGCTGGATTGTTATTACTCGCGGCCC
Seq11/H2-III     CAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCAAATTCTATTTCAAGGAGACAGTCATAATGAAATACCTATTGCCTACGGCAGCCGCTGGATTGTTATTACTCGCGGCCC
Seq3/F6-II       CAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCAAATTCTATTTCAAGGAGACAGTCATAATGAAATACCTATTGCCTACGGCAGCCGCTGGATTGTTATTACTCGCGGCCC
Seq2/H6-I        CAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCAAATTCTATTTCAAGGAGACAGTCATAATGAAATACCTATTGCCTACGGCAGCCGCTGGATTGTTATTACTCGCGGCCC
Seq15/A1-IV      CAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCAAATTCTATTTCAAGGAGACAGTCATAATGAAATACCTATTGCCTACGGCAGCCGCTGGATTGTTATTACTCGCGGCCC
Seq1/D2-I        CAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCAAATTCTATTTCAAGGAGACAGTCATAATGAAATACCTATTGCCTACGGCAGCCGCTGGATTGTTATTACTCGCGGCCC
Seq16/E11-IV     CAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCAAATTCTATTTCAAGGAGACAGTCATAATGAAATACCTATTGCCTACGGCAGCCGCTGGATTGTTATTACTCGCGGCCC
Seq10/B4-III     CAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCAAATTCTATTTCAAGGAGACAGTCATAATGAAATACCTATTGCCTACGGCAGCCGCTGGATTGTTATTACTCGCGGCCC
Seq5/A1-II       CAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCAAATTCTATTTCAAGGAGACAGTCATAATGAAATACCTATTGCCTACGGCAGCCGCTGGATTGTTATTACTCGCGGCCC
Seq13/E1-III     CAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCAAATTCTATTTCAAGGAGACAGTCATAATGAAATACCTATTGCCTACGGCAGCCGCTGGATTGTTATTACTCGCGGCCC
Seq17/E1-IV      CAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCAAATTCTATTTCAAGGAGACAGTCATAATGAAATACCTATTGCCTACGGCAGCCGCTGGATTGTTATTACTCGCGGCCC
Seq14/F3-III     CAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCAAATTCTATTTCAAGGAGACAGTCATAATGAAATACCTATTGCCTACGGCAGCCGCTGGATTGTTATTACTCGCGGCCC
Seq6/D4-II       CAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCAAATTCTATTTCAAGGAGACAGTCATAATGAAATACCTATTGCCTACGGCAGCCGCTGGATTGTTATTACTCGCGGCCC
Seq7/A4-II       CAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCAAATTCTATTTCAAGGAGACAGTCATAATGAAATACCTATTGCCTACGGCAGCCGCTGGATTGTTATTACTCGCGGCCC
Seq4/H8-II       CAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCAAATTCTATTTCAAGGAGACAGTCATAATGAAATACCTATTGCCTACGGCAGCCGCTGGATTGTTATTACTCGCGGCCC
Seq12/C2-III     CAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCAAATTCTATTTCAAGGAGACAGTCATAATGAAATACCTATTGCCTACGGCAGCCGCTGGATTGTTATTACTCGCGGCCC
Seq18/D6-IV      CAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCAAATTCTATTTCAAGGAGACAGTCATAATGAAATACCTATTGCCTACGGCAGCCGCTGGATTGTTATTACTCGCGGCCC
Seq8/A3-II       CAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCAAATTCTATTTCAAGGAGACAGTCATAATGAAATACCTATTGCCTACGGCAGCCGCTGGATTGTTATTACTCGCGGCCC
Seq9/D2-III      CAGGAAACAGCTATGACCATGATTACGCCAAGCTTGCATGCAAATTCTATTTCAAGGAGACAGTCATAATGAAATACCTATTGCCTACGGCAGCCGCTGGATTGTTATTACTCGCGGCCC
                 ******************************************************************************************************************

                 VHHSfi(NcoI)primer
Seq19/E7-IV      AGCCGGCCATGGCTCAGGTGCAGCTCGTGGAGTCAGGGGGAAGGAATGTACAACCTGGGGGTTCACTCAGACTCTCCTGCGCATCCACTCAAATGGCTTTAGACCTTCAATCCGTGAACC
Seq11/H2-III     AGCCGGCCATGGCTCAGGTGCAGCTCGTGGAGTCCGGGGGAAGATTTGTGCCAGCCGGGGGCTCCCTTAGACTCTCCTGTGCAGCCTCTGAGTGCAGTTTGAACGATTTTGTCCGAGGTG
Seq3/F6-II       AGCCGGCCATGGCGCAGTTGCAGGTCGTGGAGTCCGGGGGGAGGTTGGTGCACGCTGGGGGTTCTCAGAGACTCTCCTGTGCAGCCTCTGGAAATACCAGAGATATCGTTTTCATAGTCT
Seq2/H6-I        AGCCGGCCATGGCTCAGGTGCAGCTCGTGGAGTCAGGGGGAGGCTTGGTGCAACCTGGGGGATCCCTGAGACTCTCCTGTGCAGCCTCTGGATTCACCTTCAGTAGCTATGCTATGAGTT
Seq15/A1-IV      AGCCGGCCATGGCTCAGGTGCAGCTCGTGGAGTCAGGGGGAGGCTTGGTGCAACCTGGGGGATCCCTGAGACTCTCCTGTGCAGCCTCTGGATTCACCTTCAGTAGCTATGCTATGAGTT
Seq1/D2-I        AGCCGGCCATGGCTCAGGTGCAGCTCGTGGAGTCCGGGGGAGGCTTGGTGCAGCCTGGGGGGTCTCTGAGACTCTCCTGTGCAGCCTCTGGAAGCATCCTCAATTTGCCTGCTATGGGCT
Seq16/E11-IV     AGCCGGCCATGGCTCAGTTGCAGCTCGTGGAGTCCGGGGAGGCTTGGTGCAGCCTGGGGGGTCTCTGAGACTCTCCTGTGCAGCCTCTGGAAGCATCTTCGATTTAAATGCTATGGGCT
Seq10/B4-III     AGCCGGCCATGGCGCAGGTGCAGCTCGTGGAGTCAGGGGGAGGATCGGTGCAGGCTGGGGGCTCTCTGAGACTCTCCTGTGCAGCCTCTGGACGCACCTTCAGTAGGTATGCCATGGGCT
Seq5/A1-II       AGCCGGCCATGGCGCAGGTGCAGCTCGTGGAGTCCGGGGGAGGATTGGTGCAGGCTGGGGGCTCTCTGAGACTCTCCTGTGCAGCCTCTGGACGCACGTTCAGTAGCTATGTAATGGGCT
Seq13/E1-III     AGCCGGCCATGGCGCAGGTGCAGCTCGTGGAGTCCGGGGGAGGATTGGTGCAGGCTGGGGGCTCTCTGAGACTCTCCTGTGCAGCCTCTGGACGCACCTTCACTAGCTATGTCATGGGCT
Seq17/E1-IV      AGCCGGCCATGGCGCAGGTGCAGCTCGTGGAGTCCGGGGGAGGATTGGTGCAGGCTGGGGGCTCTCTGAGACTCTCCTGTGCAGCCTCTGGACGCACCTTCACTAGTTATGTCATGGGCT
Seq14/F3-III     AGCCGGCCATGGCGCAGGTGCAGCTCGTGGAGTCAGGGGGAGGATTGGTGCAGGCTGGGGGCTCTCTGAGACTCTCCTGTGCAGCCTCTGGACGCACCTTCAGTAGCTATGCCATAGGCT
Seq6/D4-II       AGCCGGCCATGGCGCAGGTGCAGCTCGTGGAGTCCGGGGAGGATTGGTGCAGGCTGGGGGCTCTCTGAGACTCTCCTGTGTAGCCTCTGGACGCACCTTCACTAGTTATGCCGTGGGCT
Seq7/A4-II       AGCCGGCCATGGCGCAGGTGCAGCTCGTGGAGTCAGGGGAGGATTGGTGCAGGCTGGGGGCTCTCTGAGACTCTCCTGTGTAGCCTCTGGACGCACCTTCACTAGTTATGCCGCGGCCT
Seq4/H8-II       AGCCGGCCATGGCGCAGGTGCAGCTCGTGGAGTCCGGGGGAGGATCGGTGCAGGCTGGGGGCTCTCTGAGACTCTCCTGTGTAGCCTCTGGACGCACCTTCAGAAGTTATGCCGCGGCCT
Seq12/C2-III     AGCCGGCCATGGCGCAGGTGCAGCTCGTGGAGTCAGGGGGAGGATCGGTGCAGGCTGGGGGCTCTCTGAGACTCTCCTGTGTAGCCTCTGGACGCACCTTCACAAGTTATGCCGCGGCCT
Seq18/D6-IV      AGCCGGCCATGGCGCAGGTGCAGCTCGTGGAGTCAGGGGAGGATCGGTGCAGGCTGGGGGCTCTCTGAGACTCTCCTGTGTAGCCTCTGGACGCACCTTCACAAGTTATGCCGCGGCCT
Seq8/A3-II       AGCCGGCCATGGCGCAGGTGCAGCTCGTGGAGTCCGGGGGAGGATTGGTGCAGGCTGGGGGCTCTCTGAGACTCTCCTGTGCAGCCTCTGGACGCACCTTCACTAGTTATGTCATGGGCT
Seq9/D2-III      AGCCGGCCATGGCGCAGGTGCAGCTCGTGGAGTCCGGGGGAGGATTGGTGCAGGCTGGGGGCTCTCTGAGACTCTCCTGTGTAGCCTCTGGACGCACCTTCACTAGTTATGCCATGGGCT
                 ************* *** ***** ********** *****  *    ** *   * **** * ** *  *********** * * ** **

Seq19/E7-IV      GGTACCGCCACGCCCCAGGGCAGGAACCTGACGCAGTCTCAATCCTTA------GCAGAACTGCTTCGATTACACGATATGCCCACTCCATGAAAGACCACTTCTCCCTCACCACAGCCA
Seq11/H2-III     GGTACCGTCAGGCCGCAGGGAAGGAACGCGAGGCAGTCGCATGTATTA------GTAGGAGTGGTAATAGTACATACTACGAGGACTCCGTGAAGGAACGATTCACCATCTCCAGAGACA
Seq3/F6-II       GGTACCGCCAGGCTCCAGGGAACGAGCGTGAGTTGGTCGCACGGCTTA------CGGGAGGTGGTAATAGCGGAAACTACGCAGACTCCGTGAAGGGCCGATTCACCATCTCTAGAGACA
Seq2/H6-I        GGGTCCGCCAGGCTCCAGGAAAGCAGCTCGAGTGGGTCTCAGGTATTAACCCTGGTCCTGGTGGTGGTATACGAACGTATCCAGACTCCCTGAAGGGCCGATTCACCATCTCCAGAGACA
Seq15/A1-IV      GGGTCCGCCAGGCTCCAGGAAAGCAGCTCGAGTGGGTCTCAGGTATTAACCCTGGTCCTGGTGGTGGTATACGAACGTATCCAGACTCCCTGAAGGGCCGATTCACCATCTCCAGAGACA
Seq1/D2-I        GGTACCGCCAGGGCGCAGGGAAGCAGCGCGAGTTGGTTGCAACTATTA------CTC---GTGGTGGTAGTACAAACTATGCAGACTCCGTGAAGGGCCGATTCACCATCTCCAAAGACA
Seq16/E11-IV     GGTACCGCCAGGGCGCAGGGAAGCAGCGCGAGTTGGTTGCAACCATTA------CTC---GTGCTGGTAGCACAAACTATGCAGACTCCGTGAAGGGCCGATTCACCATCTCCAGAGACA
Seq10/B4-III     GGTTCCGCCAGGCTCCAGGGAAGGAGCGTGAGTTTGTAGCAGCTATTA------ACTGGAGTGGTGGTAGCACATACTATGTAGACTCCGTCGAGGGTCGATTCACCATCTCCAGAGGTA
Seq5/A1-II       GGTTCCGCCAGGCTCCAGGGCAGGAGCGTGAGTTTGTAGCAGCTATTA------CCAGGGGTCGTGGTATTACACGATATGCAGACTCCGTGAATGGCCGATTCACCATCTCCAGAGACA
Seq13/E1-III     GGTTCCGCCAGGCTCCAGGGAAGGAGCGTGAGTTTGTATCAGGTATTA------CCAGGAGTGGTGGTATTACACCATATGCAGACTCCGTGAATGGCCGATTCACCATCTCCAGAGACA
Seq17/E1-IV      GGTTCCGCCAGGCTCCAGGGAAGGAGCGTGAGTTTGTATCAGCTATTA------CCAGGAGTGGTGGTATTACACGATATGCAGACTCCGTGAATGGCCGATTCACCATCTCCAGAGACA
Seq14/F3-III     GGTTCCGTCAGGCTCCAGGGAAGGAGCGTGAGTTTGTAGCAGCTATTA------GCTGGAGTGGTGGTAGCACACACTATGCAGACTCCGTGAAGGGCCGATTCACCATCTCCAGAGACA
Seq6/D4-II       GGTTCCGCCAGGCTCCAGGGAAGGAGCGTGAGTTTGTATCAGGTATTA------GTAGGAGTGGTGGTAGAACATCCTATGCTGACTCCGTGAAGGGCCGATTCACCGTCTCCAGAGACA
Seq7/A4-II       GGTTCCGCCAGGCTCCAGGGAAGGAGCGTGAATTTGTATCAGGTATTA------GTAGGAGTGGTGGTAGAACATCCTATGCTGACTCCGTGAAGGGCCGATTCACCGTCTCCAGAGACA
Seq4/H8-II       GGTTCCGCCAGGCTCCAGGGAAGGAGCGTGAATTTGTATCAGGTATTA------GTAGGAGTGGTGGTAGAACATCCTATGCTGACTCCGTGAAGGGCCGATTCACCGTCTCCAGAGACA
Seq12/C2-III     GGTTCCGCCAGGCTCCAGGGAAGGAGCGTGAATTTGTATCAGCTATTA------GTAGGAGTGGTGGTAGAACATCCTATGCTGACTCCGTGAAGGGCCGATTCACCGTCTCCAGAGACA
Seq18/D6-IV      GGTTCCGCCAGGCTCCAGGGAAGGAGCGTGAATTTGTATCAGGTATTA------GTAGGAGTGGTGGTAGAACATCCTATGCTGACTCCGTGAAGGGCCGATTCACCGTCTCCAGAGACA
Seq8/A3-II       GGTTCCGCCAGGCTCCAGGGAAGGAGCGTGAGTTTGTATCAGCTATTA------CCAGGAGTGGTGGTATAACATCCTATGCTGACTCCGTGAAGGCCGTATTCACCATCTCCAGAGACA
Seq9/D2-III      GGTTCCGCCAGGCTCCAGGGAAGGAGCGTGAGTTTGTATCAGGTATTA------CTAGGAGTGGTGGTATAACATCCTATGCTGACTCCGTGAAGGGCCGATTCACCATCTCCAGAGACA
                 **  *** ** *   ****  *  * *  **    **  **    ***            *  *   *    *   **    ***** *  * *     *** ** ** * * **  *
```

Figure 2.6- DNA sequences of the 19 anti-CD105 VHH obtained by phage display and phage-ELISA. The coding sequences of the VHH proteins are shown highlighted.



Protein clasification				
Immunoglobulin heavy chain variable domain-containing protein (domian architecture ID 10141748)				
immunoglobulin heavy chain variable domain-containing protein similar to immunoglobulin heavy chains				
List of domain hits				
(+) Name	**Accession**	**Description**	**Interval**	**E-value**
(+) IgV-H	cd04981	Immunoglobulin (Ig) heavy chain (H), variable (V) domain ...	6-130	5.11E-61
(+) Igv	smart00406	Immunoglobulin V-Type ...	19-100	1.65E-31
(+) V-set	pfam07686	Immunoglobulin V-set domain; This domain is found in antibodies as well as neural protein PO ...	13-129	1.94E-15
Blast search parameters				
Data Source		Live blast search RID=SXGCCCER01R		
User Options		Database: CDSEARCH/cdd. Low complexity filter:no, Composition Based Adjustment: yes, E-value threshold: 0.01, Maximum number of hits:500		
References: Marchier-Bauer et al[162-164] and Marchier-Bauer& Bryant[165].				

Figure 2.7- Amino acid sequences of the 19 anti-CD105 VHH proteins. a) Alignment of the sequences, b) Consensus sequence, c) Conserved domains of the immunoglobins on VHH consense sequence of 19 VHH proteins.

The V37F, V37Y, G44E, L45R and W47G substitutions at FR2[166,167] were present in most of the sequences, being the substitution by F (phenylalanine) at position 37 the one that was more present (12 sequences,) than the substitution by Y (tyrosine). Regarding positions 44 and 45, the substitutions for E (glutamic acid) and R (arginine) were present in almost all the sequences, except in the sequences VHH 1, 16, 2 and 15 where Q (glutamine) was present at position 44 and in the sequences VHH 2, 15 and 19, which had L (leucine) or P (proline) at position 45. The substitution by G (glycine) at position 47 was not present, rather it was F (Figure 2.7a, on white background).Thus, some substitutions such as V47F, G49Q or L50A have also been reported in FR2[150,166-168] that would increase the hydrophilicity of this region. At position 49, S (serine), and not Q, was the uncharged polar amino acid residue that was mostly present (12 sequences). Whereas at position 50, G was the amino acid that was predominant (8 sequences) over A (alanine) (6 sequences) and threonine (T) (2 sequences) (Figure 2.7a). The implications of these changes in the folding and stability of the recombinant proteins will become evident in the assays measuring their functionality in the interaction with their antigen.

Other substitutions such as L11S, P14A and A83P, which are related to protein stability, were also evidenced but only in some sequences and not all substitutions as they occur mainly in camel VHH and not much in llama or alpaca VHH[167,169]. L11S was present in only 4 sequences, while P14A was present in the majority (14 sequences). Regarding A83P, M (methionine) was present and not P, P was present rather at position 88 in all sequences (figure 2.7a on yellow background). Also, the W103R substitution characteristic of VHH and contributing to hydrophilicity was present in most of the sequences but at position 108, where in addition to R, other polar amino acids were present, such as K (lysine), Q or S (Figure 2.7a in orange).

Position 42 in FR2 presents G (Figure 2.7a in red on light blue background), an amino acid with an uncharged polar R group and not F (phenylalanine), so that CDR3 would not form the so-called "Stretched twisted turn" structure[170], but rather it would remain extended and unbound to the rest of the protein structure[171]. On the other hand, the cysteines involved in the stabilising disulphide bridge of the more flexible VHH CDRs are located at position 22 of FR1 and 96 of FR3 of the sequences (Figure 2.7a in red on fuchsia background).

Other differences evidenced between the VH and VHH in the sequences were the size and variability of the CDRs. CDR1 was more variable than CDR2 as reported for VHH[167] and CDR1 and CDR3 were more extensive[161,166-168], with CDR1 being composed of approximately 9 amino acid residues and CD3 of approximately 18 amino acid residues. (Figure 2.7a).

The sequences were submitted to the NCBI BLAST service to determine their identity. All matched camelid antibody sequences reported in the database, mostly with identity percentages between 73 and 85% (Figure 2.8); being the sequence AHY24762.1 immunoglobulin heavy chain variable region, partial [Lama glama], with which there

were more matches, including VHH 3 with 68.7% identity, VHH 6 with 75.0%, VHH 7 with 77.2% and VHH 14 with 85.7%.

	Description	Max Score	Total Score	Query cover	Expect value	Per. Ident	Accession
Seq1/D2-I 141 aa	Immunoglobulin heavy chain variable region [Lama glama]	213	213	92%	6e-69	82.31%	AHY24765.1
Seq2/H6-I 144 aa	Immunoglobulin heavy chain variable region [Vicugna pacos]	207	207	90%	3e-66	78.95%	ALD14657.1
Seq3/F6-II 145 aa	Immunoglobulin heavy chain variable region [Lama glama]	182	182	92%	1e-56	68.66%	AHY24762.1
Seq4/H8-II 147 aa	Immunoglobulin heavy chain variable region [Lama glama]	203	203	92%	7e-65	76.47%	AHY24763.1
Seq5/A1-II 150 aa	Immunoglobulin variable heavy region IIY-E1 RTA-E1 [Vicugna pacos]	204	204	91%	3e-65	73.38%	AHA34193.1
Seq6/D4-II 147 aa	Immunoglobulin heavy chain variable region [Lama glama]	205	205	92%	2e-65	75.00%	AHY24762.1
Seq7/A4-II 147 aa	Immunoglobulin heavy chain variable region [Lama glama]	207	207	92%	1e-66	77.21%	AHY24762.1
Seq8/A3-II 147 aa	Immunoglobulin variable heavy region IIY-E3 RTA-E3 [Vicugna pacos]	199	199	91%	5e-63	75.74%	AHA34194.1
Seq9/D2-III 147 aa	Immunoglobulin heavy chain variable region [Vicugna pacos]	201	201	91%	9e-64	77.04%	ALD14660.1
Seq10/B4-III 151 aa	immunoglobulin variable heavy region IIY-E1 RTA-E1 [Vicugna pacos]	202	202	91%	3e-64	73.19%	AHA34193.1
Seq11/H2-III 150 aa	Immunoglobulin heavy chain variable region [Vicugna pacos]	156	156	90%	3e-46	62.50%	ACS73867.1
Seq12/C2-III 147 aa	Chain C, Nanobody Mu1053 [Lama glama]	203	203	91%	2e-64	76.12%	5F1K_C
Seq13/E1-III 151 aa	Immunoglobulin variable heavy region IIY-E3 RTA-E3 [Vicugna pacos]	197	197	91%	5e-62	74.64%	AHA34194.1
Seq14/F3-III 137 aa	Immunoglobulin heavy chain variable region [Lama glama]	229	229	91%	3e-75	85.71%	AHY24762.1
Seq15/A1-IV 144 aa	Immunoglobulin heavy chain variable region [Vicugna pacos]	205	205	90%	1e-65	78.20%	ALD14657.1
Seq16/E11-IV 141 aa	Immunoglobulin heavy chain variable region [Lama glama]	213	213	92%	1e-68	81.54%	AHY24765.1
Seq17/E1-IV 150 aa	Immunoglobulin variable heavy region IIY-E3 RTA-E3 [Vicugna pacos]	209	209	91%	8e-67	78.99%	AHA34194.1
Seq18/D6-IV 147 aa	Immunoglobulin heavy chain variable region [Lama glama]	201	201	92%	5e-64	75.74%	AHY24763.1
Seq19/E7-IV 148 aa	Immunoglobulin heavy chain variable region [Lama glama]	130	130	92%	7e-36	58.91%	AHY24765.1

Figure 2.8- Results of the BLAST analysis on the 19 VHH sequences.

Other sequences with more than two matches were AHA34194.1 with the sequences VHH8 (75.74%), VHH13 (74.64%) and VHH17 (78.99); and AHY24765.1 with VHH1 (82.31%), VHH16 (81.54%) and VHH19 (58.91%). The latter sequence was the most distinct of the 19 VHH and had the lowest percentage of identity with the sequences reported in the databases. The sequence AHA34193.1, ALD14657.1 and AHY24763.1 had two matches, the first with VHH5 (73.38%) and VHH10 (73.19%), the second with VHH2 (78.95%) and VHH15 (78.20%) and the third with VHH4 (76.47%) and VHH18 (75.74%) (Figure 2.9). This is explained since these last 2 pairs of sequences varied by only one amino acid. It should be noted that the sequence with the highest score for each VHH was considered, however, there were other sequences with high similarity as well. In addition, there were sequences such as AHY24762.1 (immunoglobulin heavy chain variable region, partial [Lama glama]), which corresponded to all the VHHs but with different percentages of identity. In general, all of the so-called highest scoring sequences were present in the BLAST analysis of each VHH. The variety in the repertoire of these sequences identified as VHH not only increased the diversity of anti-CD105 VHH to express but also held the promise of gaining more information on the antigen-antibody, CD105-VHH interaction, providing better knowledge for application in CD105 detection assays by SPRi.

Max, Score Sequence	Match	VHH Sequence
AHY24762.1	4	3, 6, 7, 14
AHA34194.1	3	8, 13, 17
AHY24765.1	3	1, 16,19
AHY24763.1	2	4, 18
ALD14657.1	2	2, 15
AHA34193.1	2	5, 10
ALD14660.1	1	9
ACS73867.1	1	11
5F1K_C	1	12

Figure 2.9 Sequences from the database with the highest similarity scores to the VHH sequences and distribution of matches.

The characteristics of each anti-CD105 VHH protein without and with 6X His-tag (the recombinant proteins that would be expressed from the pET22b(+) expression vector), were predetermined with the ProtParam program of the Expasy server[172] (Figure 2.10 a and 2.10 b). The expressed recombinant proteins would be composed on average of 147 amino acid residues with a molecular weight of 15 kDa and an isoelectric point of 8.5. The extinction coefficient would be 25121 $M^{-1}.cm^{-1}$ at 280 nm in water. This coefficient indicates how much light a protein absorbs at a given wavelength, so its estimation will be very useful to determine by spectrophotometry, the concentration of VHH proteins when they are purified.

Seq VHH	Name	Number of amino acids	Molecular weight	Theoretical pI	Atomic composition	Total number of atoms	Extinction coefficients (£) Abs 0.1% (=1 g/l)	Estimated half-life	Instability index	Aliphatic index	Grand average of hydropathy (GRAVY)
1	D2-I	130	14042.69	7.71	Carbon C 612 Hydrogen H 961 Nitrogen N 171 Oxygen O 198 Sulfur S 5	1947	27055 1.927 (1) 26930 1.918 (2)	30 h (mammalian reticulocytes, *in vitro*) >20 h (yeast, *in vivo*) >10 h (E. coli, *in vivo*)	33.36 Stable	68.23	-0.501
2	H6-I	133	14501.3	8.58	Carbon C 642 Hydrogen H 987 Nitrogen N 175 Oxygen O 197 Sulfur S 6	2007	31065 2.142 (1) 30940 2.134(2)	30 h (mammalian reticulocytes, *in vitro*) >20 h (yeast, *in vivo*) >10 h (*E. coli, in vivo*)	36.04 Stable	63.01	-0.426
3	F6-II	134	14596.2	8.58	Carbon C 633 Hydrogen H 976 Nitrogen N 188 Oxygen O 201 Sulfur S 5	2003	23045 1.579 (1) 22920 1.570 (2)	30 h (mammalian reticulocytes, *in vitro*) >20 h (yeast, *in vivo*) >10 h (*E. coli, in vivo*	32.10 Stable	57.54	-0.666
4	H8-II	136	14609.3	9.78	Carbon C 639 Hydrogen H 991 Nitrogen N 185 Oxygen O 201 Sulfur S 4	2020	28545 1.954 (1) 28420 1.945 (2)	30 h (mammalian reticulocytes, *in vitro*) >20 h (yeast, *in vivo*) >10 h (*E. coli, in vivo*	36.86 Stable	53.09	-0.499
5	A1-II	139	15369.3	9.64	Carbon C 679 Hydrogen H 1052 Nitrogen N 196 Oxygen O 203 Sulfur S 5	2135	23045 1.499 (1) 22920 1.491(2)	30 h (mammalian reticulocytes, *in vitro*) >20 h (yeast, *in vivo*) >10 h (*E. coli, in vivo*	44.74 Unstable	63.17	-0.432
6	D4-II	136	14617.4	9.44	Carbon C 640 Hydrogen H 999 Nitrogen N 183 Oxygen O 198 Sulfur S 6	2026	23170 1.585 (1) 22920 1.568(2)	30 h (mammalian reticulocytes, *in vitro*) >20 h (yeast, *in vivo*) >10 h (*E. coli, in vivo*)	43.58 Unstable	55.88	-0.460
7	A4-II	136	14679.4	9.65	Carbon C 643 Hydrogen H 997 Nitrogen N 185 Oxygen O 202 Sulfur S 4	2031	28545 1.945(1) 28420 1.936(2)	30 h (mammalian reticulocytes, *in vitro*) >20 h (yeast, *in vivo*) >10 h (*E. coli, in vivo*	37.81 Stable	54.49	-0.500
8	A3-II	136	14695.5	8.97	Carbon C 641 Hydrogen H 1013 Nitrogen N 181 Oxygen O 205 Sulfur S 5	2045	20065 1.365(1) 19940 1.357(2)	30 h (mammalian reticulocytes, *in vitro*) >20 h (yeast, *in vivo*) >10 h (*E. coli, in vivo*	45.79 Unstable	68.82	-0.331
9	D2-III	136	14831.7	9.3	Carbon C 647 Hydrogen H 1017 Nitrogen N 183 Oxygen O 203 Sulfur S 7	2057	23170 1.562(1) 22920 1.545(2)	30 h (mammalian reticulocytes, *in vitro*) >20 h (yeast, *in vivo*) >10 h (*E. coli, in vivo*)	48.86 Unstable	64.49	-0.340
10	B4-III	140	15155.9	8.6	Carbon C 660 Hydrogen H 1025 Nitrogen N 191 Oxygen O 206 Sulfur S 7	2089	27055 1.785(1) 26930 1.777(2)	30 h (mammalian reticulocytes, *in vitro*) >20 h (yeast, *in vivo*) >10 h (*E. coli, in vivo*	35.34 Stable	61.71	-0.419
11	H2-III	136	14869.6	8.3	Carbon C 630 Hydrogen H 1007 Nitrogen N 193 Oxygen O 208 Sulfur S 8	2046	9190 0.618(1) 8940 0.601(2)	30 h (mammalian reticulocytes, *in vitro*) >20 h (yeast, *in vivo*) >10 h (*E. coli, in vivo*	50.20 Unstable	63.17	-0.503
12	C2-III	136	14609.3	9.78	Carbon C 639 Hydrogen H 991 Nitrogen N 185 Oxygen O 201 Sulfur S 4	2020	28545 1.954 (1) 28420 1.945 (2)	30 h (mammalian reticulocytes, *in vitro*) >20 h (yeast, *in vivo*) >10 h (*E. coli, in vivo*	36.86 Stable	53.09	-0.499
13	E1-III	140	15063.9	9.47	Carbon C 660 Hydrogen H 1025 Nitrogen N 189 Oxygen O 204 Sulfur S 6	2084	16055 1.066(1) 15930 1.057 (2)	30 h (mammalian reticulocytes, *in vitro*) >20 h (yeast, *in vivo*) >10 h (*E. coli, in vivo*	47.18 Unstable	58.50	-0.397
14	F3-III	126	13742.4	9.59	Carbon C 603 Hydrogen H 937 Nitrogen N 175 Oxygen O 184 Sulfur S 5	1904	25565 1.860(1) 25440 1.851(2)	30 h (mammalian reticulocytes, *in vitro*) >20 h (yeast, *in vivo*) >10 h (*E. coli, in vivo*	45.64 Unstable	61.98	-0.439
15	A1-IV	133	14517.3	8.56	Carbon C 642 Hydrogen H 987 Nitrogen N 175 Oxygen O 198 Sulfur S 6	2008	32555 2.242(1) 32430 2.234(2)	30 h (mammalian reticulocytes, *in vitro*) >20 h (yeast, *in vivo*) >10 h (*E. coli, in vivo*	37.17 Stable	63.01	-0.507
16	E11-IV	130	14151.7	5.01	Carbon C 615 Hydrogen H 956 Nitrogen N 172 Oxygen O 202 Sulfur S 5	1950	27055 1.912(1) 26930 1.903(2)	30 h (mammalian reticulocytes, *in vitro*) >20 h (yeast, *in vivo*) >10 h (*E. coli, in vivo*	34.13 Stable	66.77	-0.511
17	E1-IV	139	15315.3	9.82	Carbon C 678 Hydrogen H 1054 Nitrogen N 194 Oxygen O 202 Sulfur S 5	2133	23045 1.505(1) 22920 1.497(2)	30 h (mammalian reticulocytes, *in vitro*) >20 h (yeast, *in vivo*) >10 h (*E. coli, in vivo*	42.77 Unstable	61.73	-0.432
18	D6-IV	136	14595.3	9.78	Carbon C 638 Hydrogen H 985 Nitrogen N 188 Oxygen O 201 Sulfur S 4	2017	28545 1.956(1) 28420 1.947(2)	30 h (mammalian reticulocytes, *in vitro*) >20 h (yeast, *in vivo*) >10 h (*E. coli, in vivo*	36.23 Stable	52.35	-0.515
19	E7-IV	137	15268.3	9.91	Carbon C 659 Hydrogen H 1053 Nitrogen N 199 Oxygen O 203 Sulfur S 8	2122	16055 1.052 (1) 15930 1.043(2)	30 h (mammalian reticulocytes, *in vitro*) >20 h (yeast, *in vivo*) >10 h (*E. coli, in vivo*	41.97 Unstable	59.85	-0.726

Figure 2.10A- Theoretical physico-chemical characteristics of VHH proteins (without 6X His-tag) determined *in silico*. £ Extinction coefficients are in units of M^{-1} cm^{-1}, at 280 nm measured in water. (1) Assuming all pairs of Cys residues form cystines, (2) Assuming all Cys residues are reduce

Amino acid composition	1		2		3		4		5		6		7		8		9		10		11		12		13		14		15		16		17		18		19	
	D2-I		H6-I		F6-II		H8-II		A1-II		D4-II		A4-II		A3-II		D2-III		B4-III		H2-III		C2-III		E1-III		F3-III		A1-IV		E11-IV		E1-IV		D6-IV		E7-IV	
	amount	porcentage	amount	porcentage	amount	porcentage	amount	porcentage	amount	porcentage	amount	porcentage	amount	porcentage	amount	porcentage	amount	porcentage	amount	porcentage	amount	porcentage	amount	porcentage	amount	porcentage	amount	porcentage	amount	porcentage	amount	porcentage	amount	porcentage	amount	porcentage	amount	porcentage
Ala (A)	9	6.9%	7	5.3%	12	9.0%	13	9.6%	12	8.6%	10	7.4%	12	8.8%	12	8.8%	9	6.6%	12	8.6%	13	9.4%	13	9.6%	11	7.9%	12	9.5%	7	5.3%	10	7.7%	11	7.9%	12	8.8%	11	8.0%
Arg (R)	5	3.8%	5	3.8%	9	6.7%	10	7.4%	13	9.4%	9	6.6%	10	7.4%	8	5.9%	9	6.6%	10	7.1%	11	7.9%	10	7.4%	9	6.4%	8	6.3%	5	3.8%	6	4.6%	12	8.6%	10	7.4%	8	5.8%
Asn (N)	6	4.6%	6	4.5%	8	6.0%	2	1.5%	4	2.9%	3	2.2%	2	1.5%	5	3.7%	4	2.9%	5	3.6%	8	5.8%	2	1.5%	6	4.3%	5	4.0%	6	4.5%	6	4.6%	4	2.9%	2	1.4%	9	6.6%
Asp (D)	6	4.6%	6	4.5%	7	5.2%	3	2.2%	3	2.2%	3	2.2%	4	2.9%	4	2.9%	2	1.5%	7	5.0%	5	3.6%	3	2.2%	3	2.1%	3	2.4%	6	4.5%	7	5.4%	3	2.2%	3	2.2%	4	2.9%
Cys (C)	2	1.5%	2	1.5%	2	1.5%	2	1.5%	2	1.4%	4	2.9%	2	1.5%	2	1.5%	4	2.9%	2	1.4%	5	3.6%	2	1.5%	2	1.4%	2	1.6%	2	1.5%	2	1.5%	2	1.4%	2	1.5%	2	1.5%
Gln (Q)	10	7.7%	10	7.5%	8	6.0%	8	5.9%	9	6.5%	8	5.9%	8	5.9%	8	5.9%	8	5.9%	6	4.3%	6	4.3%	8	5.9%	9	6.4%	7	5.6%	10	7.5%	10	7.7%	8	5.8%	8	5.9%	7	5.1%
Glu (E)	5	3.8%	4	3.0%	5	3.7%	5	3.7%	7	5.0%	5	3.7%	5	3.7%	7	5.1%	7	5.1%	6	4.3%	9	6.5%	5	3.7%	5	3.6%	5	4.0%	4	3.0%	6	4.6%	6	4.3%	5	3.7%	3	2.2%
Gly (G)	15	11.5%	16	12.0%	16	11.9%	15	11.0%	14	10.1%	16	11.8%	15	11.0%	12	8.8%	14	10.3%	19	13.6%	15	10.8%	15	11.0%	16	11.4%	13	10.3%	16	12.0%	14	10.8%	14	10.1%	16	11.8%	6	4.4%
His (H)	0	0.0%	0	0.0%	2	1.5%	0	0.0%	0	0.0%	0	0.0%	0	0.0%	0	0.0%	0	0.0%	1	0.7%	1	0.7%	0	0.0%	1	0.7%	2	1.6%	0	0.0%	0	0.0%	0	0.0%	0	0.0%	6	4..4%
Ile (I)	3	2.3%	3	2.3%	3	2.2%	1	0.7%	4	2.9%	1	0.7%	1	0.7%	5	3.7%	5	3.7%	2	1.4%	4	2.9%	1	0.7%	3	2.1%	4	3.2%	3	2.3%	3	2.3%	3	2.2%	1	0.7%	2	1.5%
Leu (L)	10	7.7%	10	7.5%	7	5.2%	6	4.4%	8	5.8%	7	5.1%	6	4.4%	7	5.1%	7	5.1%	10	7.1%	7	5.0%	6	4.4%	7	5.0%	7	5.6%	10	7.5%	10	7.7%	8	5.8%	6	4.4%	11	8.0%
Lys (K)	7	5.4%	7	5.3%	5	3.7%	6	4.4%	3	2.2%	7	5.1%	6	4.4%	6	4.4%	6	4.4%	5	3.6%	5	3.6%	6	4.4%	4	2.9%	6	4.8%	7	5.3%	5	3.8%	5	3.6%	6	4.4%	9	6.7%
Met (M)	3	2.3%	4	3.0%	3	2.2%	2	1.5%	3	2.2%	2	1.5%	2	1.5%	3	2.2%	3	2.2%	5	3.6%	3	2.2%	2	1.5%	4	2.9%	3	2.4%	4	3.0%	3	2.3%	3	2.2%	2	1.5%	6	4.4%
Phe (F)	1	0.8%	4	3.0%	3	2.2%	4	2.9%	6	4.3%	4	2.9%	4	2.9%	4	2.9%	4	2.9%	4	2.9%	3	2.2%	4	2.9%	6	4.3%	4	3.2%	3	2.3%	2	1.5%	6	4.3%	4	2.9%	1	0.7%
Pro (P)	6	4.6%	9	6.8%	7	5.2%	7	5.1%	9	6.5%	9	6.6%	7	5.1%	6	4.4%	5	3.7%	5	3.6%	7	5.0%	7	5.1%	10	7.1%	5	4.0%	9	6.8%	5	3.8%	9	6.5%	7	5.1%	9	6.6%
Ser (S)	12	9.2%	13	9.8%	10	7.5%	20	14.7%	13	9.4%	16	11.8%	18	13.2%	17	12.5%	18	13.2%	12	8.6%	15	10.8%	20	14.7%	14	10.0%	14	11.1%	13	9.8%	12	9.2%	14	10.1%	20	14.7%	15	10.9%
Thr (T)	10	7.7%	8	6.0%	8	6.0%	10	7.4%	9	6.5%	10	7.4%	11	8.1%	10	7.4%	10	7.4%	9	6.4%	5	3.6%	10	7.4%	11	7.9%	9	7.1%	8	6.0%	10	7.7%	10	7.2%	10	7.4%	13	9.5%
Trp (W)	3	2.3%	4	3.0%	2	1.5%	3	2.2%	2	1.4%	2	1.5%	3	2.2%	2	1.5%	2	1.5%	3	2.1%	0	0.0%	3	2.2%	1	0.7%	3	2.4%	4	3.0%	3	2.3%	2	1.4%	3	2.2%	1	0.7%
Tyr (Y)	7	5.4%	6	4.5%	8	6.0%	8	5.9%	8	5.8%	8	5.9%	8	5.9%	6	4.4%	8	5.9%	7	5.0%	6	4.3%	8	5.9%	7	5.0%	6	4.8%	7	5.3%	7	5.4%	8	5.6%	8	5.9%	7	5.1%
Val (V)	10	7.7%	9	6.8%	9	6.7%	11	8.1%	10	7.2%	12	8.8%	12	8.8%	12	8.8%	11	8.1%	10	7.1%	11	7.9%	11	8.1%	11	7.9%	8	6.3%	9	6.8%	9	6.9%	11	7.9%	11	8.1%	7	5.1%

Total number of negatively charged residues (Asp + Glu)	11	10	12	9	10	8	9	11	9	13	14	9	8	8	10	13	9	8	7
Total number of positively charged residues (Arg + Lys)	12	12	14	16	16	15	16	14	15	15	16	16	13	14	12	11	17	16	17

Figure 2.10A- Theoretical physico-chemical characteristics of VHH proteins (without 6X His-tag) determined *in silico* (continuation)

Seq VHH	Name	Number of amino acids	Molecular weight	Theoretical pI	Atomic composition	Total number of atoms	Extinction coefficients (£) Abs 0.1% (=1 g/l)	Estimated half-life	Instability index	Aliphatic index	Grand average of hydropathy (GRAVY)
1	D2-I	141	15321.05	7.01	Carbon C 668 Hydrogen H 1036 Nitrogen N 194 Oxygen O 211 Sulfur S 5	2114	27055 1.766 (1) 26930 1.758 (2)	30 h (mammalian reticulocytes, *in vitro*) >20 h (yeast, *in vivo*) >10 h (*E. coli, in vivo*	31.00 Stable	67.80	-0.557
2	H6-I	144	15779.67	7.83	Carbon C 698 Hydrogen H 1062 Nitrogen N 198 Oxygen O 210 Sulfur S 6	2174	31065 1.969 (1) 30940 1.961 (2)	30 h (mammalian reticulocytes, *in vitro*) >20 h (yeast, *in vivo*) >10 h (*E. coli, in vivo*	33.53 Stable	62.99	-0.533
3	F6-II	145	15874.51	7.86	Carbon C 689 Hydrogen H 1051 Nitrogen N 211 Oxygen O 214 Sulfur S 5	2170	23045 1.452 (1) 22920 1.444 (2)	30 h (mammalian reticulocytes, *in vitro*) >20 h (yeast, *in vivo*) >10 h (*E. coli, in vivo*	29.91 Stable	57.93	-0.708
4	H8-II	147	15928.67	9.77	Carbon C 696 Hydrogen H 1069 Nitrogen N 211 Oxygen O 213 Sulfur S 4	2193	28545 1.792(1) 28420 1.784 (2)	30 h (mammalian reticulocytes, *in vitro*) >20 h (yeast, *in vivo*) >10 h (*E. coli, in vivo*	36.74 Stable	53.13	-0.594
5	A1-II	150	16647.67	9.46	Carbon C 735 Hydrogen H 1127 Nitrogen N 219 Oxygen O 216 Sulfur S 5	2302	23045 1.384 (1) 22920 1.377 (2)	30 h (mammalian reticulocytes, in vitro) >20 h (yeast, in vivo) >10 h (E. coli, in vivo)	39.69 Stable	63.13	-0.491
6	D4-II	147	15895.8	9.41	Carbon C 696 Hydrogen H 1074 Nitrogen N 206 Oxygen O 211 Sulfur S 6	2193	23170 1.458 (1) 22920 1.442(2)	30 h (mammalian reticulocytes, *in vitro*) >20 h (yeast, *in vivo*) >10 h (*E. coli, in vivo*	38.55 Stable	56.39	-0.518
7	A4-II	147	15957.71	9.51	Carbon C 699 Hydrogen H 1072 Nitrogen N 208 Oxygen O 215 Sulfur S 4	2198	28545 1.789(1) 28420 1.781(2)	30 h (mammalian reticulocytes, *in vitro*) >20 h (yeast, *in vivo*) >10 h (*E. coli, in vivo*	35.21 Stable	55.1	-0.554
8	A3-II	147	15973.84	8.6	Carbon C 697 Hydrogen H 1088 Nitrogen N 204 Oxygen O 218 Sulfur S 5	2212	20065 1.256(1) 19940 1.248(2)	30 h (mammalian reticulocytes, *in vitro*) >20 h (yeast, *in vivo*) >10 h (*E. coli, in vivo*)	39.60 Stable	68.37	-0.398
9	D2-III	147	16110.07	9.14	Carbon C 703 Hydrogen H 1092 Nitrogen N 206 Oxygen O 216 Sulfur S 7	2224	23170 1.438(1) 22920 1.423(2)	30 h (mammalian reticulocytes, *in vitro*) >20 h (yeast, *in vivo*) >10 h (*E. coli, in vivo*	42.44 Instable	64.35	-0.406
10	B4-III	151	16434.33	7.85	Carbon C 716 Hydrogen H 1100 Nitrogen N 214 Oxygen O 216 Sulfur S 7	2256	27055 1.646(1) 26930 1.639(2)	30 h (mammalian reticulocytes, *in vitro*) >20 h (yeast, *in vivo*) >10 h (*E. coli, in vivo*)	33.00 Stable	62.72	-0.478
11	H2-III	150	16147.93	7.71	Carbon C 686 Hydrogen H 1082 Nitrogen N 216 Oxygen O 221 Sulfur S 8	2213	9190 0.569(1) 8940 0.554(2)	30 h (mammalian reticulocytes, *in vitro*) >20 h (yeast, *in vivo*) >10 h (*E. coli, in vivo*	43.75 Instable	63.13	-0.556
12	C2-III	147	15887.62	9.65	Carbon C 695 Hydrogen H 1066 Nitrogen N 208 Oxygen O 214 Sulfur S 4	2187	28545 1.797 (1) 28420 1.789 (2)	30 h (mammalian reticulocytes, *in vitro*) >20 h (yeast, *in vivo*) >10 h (*E. coli, in vivo*	34.33 Stable	53.81	-0.553
13	E1-III	151	16342.26	9.26	Carbon C 716 Hydrogen H 1100 Nitrogen N 212 Oxygen O 217 Sulfur S 6	2251	16055 0.982(1) 15930 0.975 (2)	30 h (mammalian reticulocytes, *in vitro*) >20 h (yeast, *in vivo*) >10 h (*E. coli, in vivo*	43.83 Instable	58.81	-0.458
14	F3-III	137	15020.79	9.43	Carbon C 612 Hydrogen H 952 Nitrogen N 178 Oxygen O 187 Sulfur S 5	2071	25565 1.702(1) 25440 1.694(2)	30 h (mammalian reticulocytes, *in vitro*) >20 h (yeast, *in vivo*) >10 h (*E. coli, in vivo*	39.22 Stable	62.04	-0.502
15	A1-IV	144	15795.67	7.82	Carbon C 698 Hydrogen H 1062 Nitrogen N 198 Oxygen O 211 Sulfur S 6	2175	32555 2.061(1) 32430 2.053(2)	30 h (mammalian reticulocytes, *in vitro*) >20 h (yeast, *in vivo*) >10 h (*E. coli, in vivo*	34.57 Stable	62.99	-0.562
16	E11-IV	141	15430.04	6.03	Carbon C 671 Hydrogen H 1031 Nitrogen N 195 Oxygen O 215 Sulfur S 5	2117	27055 1.753(1) 26930 1.745(2)	30 h (mammalian reticulocytes, *in vitro*) >20 h (yeast, *in vivo*) >10 h (*E. coli, in vivo*	31.72 Stable	66.45	-0.567
17	E1-IV	150	16593.66	9.69	Carbon C 734 Hydrogen H 1129 Nitrogen N 217 Oxygen O 215 Sulfur S 5	2300	23045 1.389(1) 22920 1.381(2)	30 h (mammalian reticulocytes, *in vitro*) >20 h (yeast, *in vivo*) >10 h (*E. coli, in vivo*	39.86 Stable	61.8	-0.49
18	D6-IV	147	15873.59	9.65	Carbon C 694 Hydrogen H 1064 Nitrogen N 208 Oxygen O 214 Sulfur S 4	2184	28545 1.798(1) 28420 1.790(2)	30 h (mammalian reticulocytes, *in vitro*) >20 h (yeast, *in vivo*) >10 h (*E. coli, in vivo*	33.76 Stable	53.13	-0.568
19	E7-IV	148	16546.7	9.80	Carbon C 715 Hydrogen H 1128 Nitrogen N 222 Oxygen O 216 Sulfur S 8	2289	16055 0.970 (1) 15930 0.963(2)	30 h (mammalian reticulocytes, *in vitro*) >20 h (yeast, *in vivo*) >10 h (*E. coli, in vivo*	37.42 Stable	39.08	-0.764

Figure 2.10B- Theoretical physico-chemical characteristics of VHH proteins (with 6X His-tag) determined *in silico*. £ Extinction coefficients are in units of M^{-1} cm^{-1}, at 280 nm measured in water. (1) Assuming all pairs of Cys residues form cystines, (2) Assuming all Cys residues are reduce

Amino acid composition	1		2		3		4		5		6		7		8		9		10		11		12		13		14		15		16		17		18		19	
	D2-I		H6-I		F6-II		H8-II		A1-II		D4-II		A4-II		A3-II		D2-III		B4-III		H2-III		C2-III		E1-III		F3-III		A1-IV		E11-IV		E1-IV		D6-IV		E7-IV	
	amount	porcentage	amount	porcentage	amount	porcentage	amount	porcentage	amount	porcentage	amount	porcentage	amount	porcentage	amount	porcentage	amount	porcentage	amount	amount	porcentage	amount	porcentage	amount	porcentage	amount	porcentage	amount	porcentage	amount	porcentage	amount	porcentage	amount	porcentage	amount	porcentage
Ala (A)	12	8.5%	10	6.9%	15	10.30%	16	10.90%	15	10.00%	13	8.80%	15	10.20%	15	10.20%	12	8.20%	15	9.90%	16	10.70%	16	10.90%	14	9.30%	15	10.90%	10	6.90%	13	9.20%	14	9.30%	15	10.20%	14	9.50%
Arg (R)	5	3.5%	5	3.5%	9	6.20%	10	6.80%	13	8.70%	9	6.10%	10	6.80%	8	5.40%	9	6.10%	10	6.60%	11	7.30%	10	6.80%	9	6.00%	8	5.80%	5	3.50%	6	4.30%	12	8.00%	10	6.80%	8	5.40%
Asn (N)	6	4.3%	6	4.2%	8	5.50%	2	1.40%	4	2.70%	3	2.00%	2	1.40%	5	3.40%	4	2.70%	5	3.30%	8	5.30%	2	1.40%	6	4.00%	5	3.60%	6	4.20%	6	4.30%	4	2.70%	2	1.40%	9	6.10%
Asp (D)	6	4.3%	6	4.2%	7	4.80%	3	2.00%	3	2.00%	3	2.00%	4	2.70%	4	2.70%	2	1.40%	7	4.60%	5	3.30%	3	2.00%	3	2.00%	3	2.20%	6	4.20%	7	5.00%	3	2.00%	3	2.00%	4	2.70%
Cys (C)	2	1.4%	2	1.4%	2	1.40%	2	1.40%	2	1.30%	4	2.70%	2	1.40%	2	1.40%	4	2.70%	2	1.30%	5	3.30%	2	1.40%	2	1.30%	2	1.50%	2	1.40%	2	1.40%	2	1.30%	2	1.40%	2	1.40%
Gln (Q)	10	7.1%	10	6.9%	8	5.50%	8	5.40%	9	6.00%	8	5.40%	8	5.40%	8	5.40%	8	5.40%	6	4.00%	6	4.00%	8	5.40%	9	6.00%	7	5.10%	10	6.90%	10	7.10%	8	5.30%	8	5.40%	7	4.70%
Glu (E)	6	4.3%	5	3.5%	6	4.10%	6	4.10%	8	5.30%	6	4.10%	6	4.10%	8	5.40%	8	5.40%	7	4.60%	10	6.70%	6	4.10%	6	4.00%	6	4.40%	5	3.50%	7	5.00%	7	4.70%	6	4.10%	4	2.70%
Gly (G)	15	10.6%	16	11.1%	16	11.00%	15	10.20%	14	9.30%	16	10.90%	15	10.20%	12	8.20%	14	9.50%	19	12.60%	15	10.00%	15	10.20%	16	10.60%	13	9.50%	16	11.10%	14	9.90%	14	9.30%	16	10.90%	6	4.10%
His (H)	6	4.3%	6	4.2%	8	5.50%	6	4.10%	6	4.00%	6	4.10%	6	4.10%	6	4.10%	6	4.10%	7	4.60%	7	4.70%	6	4.10%	7	4.60%	8	5.80%	6	4.20%	6	4.30%	6	4.00%	6	4.10%	12	8.10%
Ile (I)	3	2.1%	3	2.1%	3	2.10%	1	0.70%	4	2.70%	1	0.70%	1	0.70%	5	3.40%	5	3.40%	2	1.30%	4	2.70%	1	0.70%	3	2.00%	4	2.90%	3	2.10%	3	2.10%	3	2.00%	1	0.70%	2	1.40%
Leu (L)	11	7.8%	11	7.6%	8	5.50%	7	4.80%	9	6.00%	8	5.40%	7	4.80%	8	5.40%	8	5.40%	11	7.30%	8	5.30%	7	4.80%	8	5.30%	8	5.80%	11	7.60%	11	7.80%	9	6.00%	7	4.80%	12	8.10%
Lys (K)	7	5.0%	7	4.9%	5	3.40%	6	4.10%	3	2.00%	7	4.8%	6	4.10%	6	4.10%	6	4.10%	5	3.30%	5	3.30%	6	4.10%	4	2.60%	6	4.40%	7	4.90%	5	3.50%	5	3.30%	6	4.10%	9	6.10%
Met (M)	3	2.1%	4	2.8%	3	2.10%	2	1.40%	3	2.00%	2	1.40%	2	1.40%	3	2.00%	3	2.00%	5	3.30%	3	2.00%	2	1.40%	4	2.60%	3	2.20%	4	2.80%	3	2.10%	3	2.00%	2	1.40%	6	4.10%
Phe (F)	1	0.7%	4	2.8%	3	2.10%	4	2.70%	6	4.00%	4	2.70%	4	2.70%	4	2.70%	4	2.70%	4	2.60%	3	2.00%	4	2.70%	6	4.00%	4	2.90%	3	2.10%	2	1.40%	6	4.00%	4	2.70%	1	0.70%
Pro (P)	6	4.3%	9	6.2%	7	4.80%	7	4.80%	9	6.00%	9	6.1%	7	4.80%	6	4.10%	5	3.40%	5	3.30%	7	4.70%	7	4.80%	10	6.60%	5	3.60%	9	6.20%	5	3.50%	9	6.00%	7	4.80%	9	6.10%
Ser (S)	12	8.5%	13	9.0%	10	6.90%	20	13.60%	13	8.70%	16	10.90%	18	12.20%	17	11.60%	18	12.20%	12	7.90%	15	10.00%	20	13.60%	14	9.30%	14	10.20%	13	9.00%	12	8.50%	14	9.30%	20	13.60%	15	10.10%
Thr (T)	10	7.1%	8	5.6%	8	5.50%	10	6.80%	9	6.00%	10	6.8%	11	7.50%	10	6.80%	10	6.80%	9	6.00%	5	3.30%	10	6.80%	11	7.30%	9	6.60%	8	5.60%	10	7.10%	10	6.70%	10	6.80%	13	8.80%
Trp (W)	3	2.1%	4	2.8%	2	1.40%	3	2.00%	2	1.30%	2	1.40%	3	2.00%	2	1.40%	2	1.40%	3	2.00%	0	0.00%	3	2.00%	1	0.70%	3	2.20%	4	2.80%	3	2.10%	2	1.30%	3	2.00%	1	0.70%
Tyr (Y)	7	5.0%	6	4.2%	8	5.50%	8	5.40%	8	5.30%	8	5.40%	8	5.40%	6	4.10%	8	5.40%	7	4.60%	6	4.00%	8	5.40%	7	4.60%	6	4.40%	7	4.90%	7	5.00%	8	5.30%	8	5.40%	7	4.70%
Val (V)	10	7.1%	9	6.2%	9	6.20%	11	7.50%	10	6.70%	12	8.20%	12	8.20%	12	8.20%	11	7.50%	10	6.60%	11	7.30%	11	7.50%	11	7.30%	8	5.80%	9	6.20%	9	6.40%	11	7.30%	11	7.50%	7	4.70%

Total number of negatively charged residues (Asp + Glu)	12	11	13	9	11	9	10	12	10	14	15	9	9	9	11	14	10	9	8	
Total number of positively charged residues (Arg + Lys)	12	12	14	16	16	16	16	14	15	15	16	16	13	14	12	11	17	16	17	

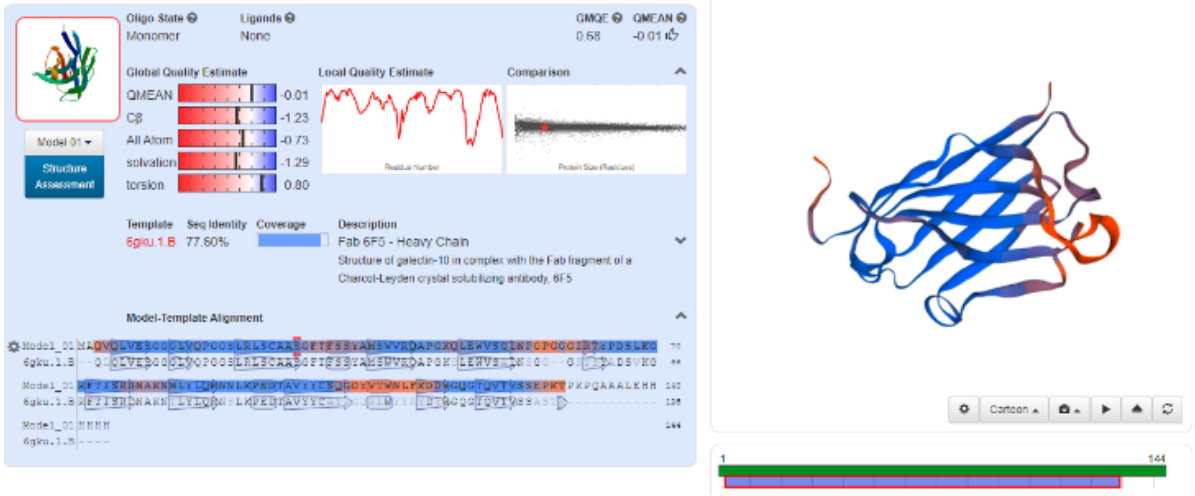
Figure 2.10B- Theoretical physico-chemical characteristics of VHH proteins (with 6X His-tag) determined *in silico*. (Continuation)

Recombinant VHHs are reported as stable, soluble antibody fragments with high affinity and specificity for their antigen. Thus, they maintain their ability to recognise and bind to their antigen after being stored for months in PBS at 4° C or longer if stored at -20° C. They are even able to maintain between 80% and 100% of this capacity when incubated at 37° C in PBS for one week[173]. According to the instability index, the recombinant VHH produced in pET22b (that is, with the 6X His tag), would be stable in solution after purification, with the exception of VHH 9, 11 and 13. The 6X His tag appears to contribute to the stability as the instability index is higher in recombinant VHHs without the tag. Thus, VHH 5, 6, 8, 9, 11, 13, 14, 17 and 19 show instability indices with values above the minimum, meaning, values higher than 40 (Figure 2.10 a). This prediction reinforced the choice of the pET22b(+) vector for the expression of the recombinant VHH due to the stability that this system would provide to the purified proteins. The half-life of VHHs in *E. coli in vivo* was adequate (more than 10 hours), as being an important fact to control both the induction and bacterial growth timing in order to optimise the conditions of VHH protein expression. The hydropathy index (GRAVY) classifies proteins as hydrophilic with an average value of -0.54 and thermostable with an average aliphatic index of 60. The thermostability of proteins is related to their ability to maintain the correct folding when the temperature of the medium containing them changes. For VHH, it has been reported that the melting temperature (Tm) is approximately 60° C and that thermal denaturation is reversible, with partial loss of binding capacity due in part to incorrect refolding of the long loops (CDR), which are responsible for antigen recognition[174,175]. However, VHH capable of fully maintaining their antigen binding capacity at 90° C have also been reported[81]. In general, higher thermostability and renaturation capacity of VHH may be due to lack of hydrophobic regions that mediate the interaction of heavy and light chains that conventional antibodies possess[174,176]. These characteristics reinforce that the sequences obtained correspond to potentially functional VHH coding sequences.

The structure of the VHH was modeled using SwissModel program, which perform modelling based on proteins reported in the database. The structural model of the 19 VHH proteins showed the composition of 9 β-sheets linked by 8 loops, a characteristic structure of the variable domains (VHH) of the HCAbs and the variable domains of heavy chain (VH) of the antibodies (Figure 2.11). The structural modelling was based on the VHH Domain BS-1 (4tvs.1) for VHH 4, 5, 6, 12, 13, 17 and nanobody MU1053 (5f1k.1) for VHH 3, 7, 8, 10, 18. Other base proteins were B39 VHH (4nc2.1.B) for VHH 1 and 11 and nanobody MU551 (5f1o.1.B) for VHH 16 and 19. All the main template proteins were VHH, thereby, also proving the identity of the 19 sequences obtained as VHH coding sequences (Figure 2.12). The analysis of the 19 VHH sequences showed the diversity and the greater or lesser similarity between them, allowing their distribution in groups. Thus, VHH 2 and 15, as well as VHH 4, 18 and 12 were classified in the groups with the highest similarity between them (with only one different amino acid), followed by VHH 1 and 16. VHH 6 and 7, VHH 5, 17 and 13, and VHH 8 and 9 also showed high similarity between them. VHH 3, 11 and 14 showed similarity among them, even though they showed the least similarity with the rest of VHH, while VHH 10 and VHH 19 were the least similar among all VHH.

SeqVHH-1/D2-I

SeqVHH-2/H6-I



SeqVHH-3/F6-II

Figure 2.11- Images of the three-dimensional structures of the VHHs modeled by Swiis Model. The images of the 3D structures of the VHH 4 to VHH19 sequence are shown in annexes (A.8).

Template Accession	Protein Description	VHH Sequence
4tvs.1	VHH Domain BS-1, LAP1(aa356-583), H.sapiens, bound to VHH-BS1	4, 5, 6, 12, 13, 17
5f1k.1	Nanobody MU1053, human CD38 in complex with nanobody MU1053	3, 7, 8, 10, 18
4nc2.1.B	B39 VHH, Crystal structure of TcdB-B1 bound to B39 VHH	1.11
5f1o.1.B	Nanobody MU551 , human CD38 in complex with nanobody MU551	16,19
6gku.1.B	Fab 6F5 Heavy chain, Fab fragment of a Charcot-Leyden crystal solubilizing antibody, 6F5	2
4nbx.1.B	A20.1 VHH, Crystal Structure of Clostridium difficile Toxin A fragment TcdA-A1 Bound to A20.1	9
5lhn.1.B	Camelid-Derived Antibody Fragment Nb7, catalytic domain of murine urokinase in complex with the allosteric inhibitory nanobody Nb7	14
6cr1.1.B	Heavy chain of adalimumab EFab (VH IgE CH2), Homo sapiens, immune systen	15

Figure 2.12- Template sequences for the modeling of the three-dimensional structure of VHH proteins according to Swiss Model.

2.4 Conclusions

The screening of a VHH cDNA library from an alpaca immunised with a lysate of CD105-overexpressing cells (T24) resulted in 376 VHH displayed phages that were subject to selection by anti-CD105 phage-ELISA. Nineteen anti-CD105-VHH displayed phages were captured, nine with strong signal (up to 35-fold higher than the control) and ten with a signal with the minimum ratio (3-fold). The nine anti-CD105-VHH displayed phages with strong signal were: n° 12 (D2-I), n° 48 (H6-I), n° 140 (F6-II), n° 158 (H8-II), n° 200 (D2-III), n° 214 (B4-III), n° 283 (A1-IV), n° 335 (E7-IV) and n° 367 (E11-IV). The ten anti-CD105-VHH displayed phages with the minimum ratio were: n° 95 (A1-II), n° 111 (A3-II), n° 119 (A4-II), n° 122 (D4-II), n° 193 (E1-III), n° 199 (C2-III), n° 204 (H2-III), n° 210 (F3-III), n° 287 (E1-IV) and n° 326 (D6-IV). The cDNA of the 19 anti-CD105-VHH were sequenced for analysis and identification. The BLAST analysis determined that the sequences matched camelid antibody sequences (mainly llamas and vicuñas) in the majority with percentages of identity between 73 and 85% and corroborated that all the sequences encoded different proteins, even if they differed in only one amino acid. The theoretical physico-chemical characteristics of each VHH protein predicted *in silico* were established to gain further insight into them and thus be able to improve the production of recombinant VHH proteins. SwissModel modelling of the three-dimensional structure of the protein sequences based on VHH proteins reported in the Protein Data Bank (PDB), in addition to the BLAST analysis of the sequences and the *in silico* characterisation of the predicted proteins, allowed the identification of the 19 different sequences as cDNAs encoding VHH proteins. The similarity between sequences allowing them to be segregated into groups. Two groups in which the VHH sequences differ by only one amino acid, a group consisting of VHH2 and VHH15 and the other of VHH4, VHH12 and VHH18. Others group were VHH 1 and 16; VHH 6 and VHH7; VHH 5, VHH 17 and VHH13; VHH 8 and VHH9 and VHH 3, VHH11 and VHH14. The latter group showed similarity among themselves and the least similarity with the rest of the VHH sequences. VHH 10 and VHH 19 were the least similar among all the VHH sequences.

References

1) Zhao, F., Shi, R., Liu, R., Tian, Y. & Yan, Z. (2021) Application of phage-display developed antibody and antigen substitutes in immunoassays for small molecule contaminants analysis: A mini-review. *Food Chemistry 339,* 128084.

2) Sioud, M. (2019) Phage Display Libraries: From Binders to Targeted Drug Delivery and Human Therapeutics. *Molecular Biotechnology 61,* 286–303.

3) Ledsgaard, L., Kilstrup, M., Karatt-Vellatt, A., McCafferty, J. & Laustsen, A. H. (2018) Basics of Antibody Phage Display Technology. *Toxins 10*, 236.

4) Herng Leow, C., Fischer, K., Yee Leow, C., Cheng, Q., Chuah, C. & McCarthy, J. (2017) Single Domain Antibodies as New Biomarker Detectors. *Diagnostics 7,* 52.

5) Kuhn, P., Fuhner, V., Unkauf, T. Marcal, G., Garcia Moreira, M. G. S., Frenzel, A., . . . Hust, M. (2016) Recombinant antibodies for diagnostics and therapy against pathogens and toxins generated by phage display. *Proteomics Clin Appl 10,* 922–948.

6) Roncolato, E. C., Campos, L. B., Pessenda, G., Costa e Silva, L., Furtado, G. P. & Barbosa, J. E. (2015) Phage display as a novel promising antivenom therapy: A review. *Toxins 93,* 79-84.

7) Chan, C. E. Z., Lim, A. P. C., MacAry, P. A. & Hanson, B. J. (2014) The role of phage display in therapeutic antibody discovery. *International Immunology 26(12)*, 649–657.

8) Sundell, G. N. & Ivarsson, Y. (2014) Interaction Analysis through proteomic phage display. *BioMed Research International 2014*, 176172,

9) Clementi, N.; Mancini, N.; Solforosi, L.; Castelli, M.; Clementi, M. & Burioni, R. (2012) Phage Display-based Strategies for Cloning and Optimization of Monoclonal Antibodies Directed against Human Pathogens. *Int. J Mol Sci13,* 8273-8292.

10) Rakonjac, J., Bennett, N. J., Spagnuolo, J., Gagic, D. & Russel, M. (2011). Filamentous bacteriophage: biology, phage display and nanotechnology applications. *Curr Issues Mol Biol 13,* 51–76.

11) Pande, J., Szewczyk, M. M. & Grover, A. K. (2010) Phage display: concept, innovations, applications and future. *Biotechnol Adv 28,* 849–858.

12) Bratkovic, T. (2010). Progress in phage display: evolution of the technique and its application. *Cell Mol Life Sci 67(5),* 749–767.

13) Bradbury, A. R. (2010) The use of phage display in neurobiology. *Curr Protoc Neurosci Chapter 5,* Unit 5.12.

14) Petrenko, V. A. & Vodyano, V. J. (2003) Phage display for detection of biological threat agents. *J Microbiol Methods 53,* 253–262.

15) Konthur, Z. & Crameri, R. (2003) High-throughput applications of phage display in proteomic analyses. *Targets 2(6),* 261–270.

16) Azzazy, H. M. & Highsmith, W. E. Jr. (2002) Phage display technology: clinical applications and recent innovations. *Clin Biochem 35(6),* 425–445.

17) Stricker, N. & Li, M. (2001) *Phage Display Technologies*. John Wiley & Sons, Ltd

18) Schmitz, U., Versmold, A., Kaufmann, P. & Frank, H. G. (2000) Phage Display: A Molecular Tool for the Generation of Antibodies— A Review. *Placenta 21,* S106-S112.

19) Hoogenboom, H. R., P de Bruïne, A., Hufton, S. E., Hoet, R. M., Arends, J. W. & Roovers, R. C. (1998) Antibody phage display technology and its applications. *Immunotechnology 4(1),* 1-20.

20) Smith, G. P. & Petrenko, V. A. (1997). Phage display. *Chem Rev 97(2),* 391–410.

21) Katz, B. A. (1997) Structural and mechanistic determinants of affinity and specificity of ligands discovered or engineered by phage display. *Annu Rev Biophys Biomol Struct 26,* 27–45

22) Molek, P., Strukelj, B. & Bratkovic, T. (2011) Peptide phage display as a tool for drug discovery: targeting membrane receptors. *Molecules 16(1),* 857–887.

23) Garnock-Jones, K. P. (2010) Ecallantide: in acute hereditary angioedema. *Drugs 70(11),* 1423–1431.

24) Sergeeva, A., Kolonin, M. G., Molldrem, J. J., Pasqualini, R. & Arap, W. (2006) Display technologies: application for the discovery of drug and gene delivery agents. *Adv Drug Deliv Rev 58,* 1622–1654.

25) Lehmann, A.. (2008) Ecallantide (DX-88), a plasma kallikrein inhibitor for the treatment of hereditary angioedema and the prevention of blood loss in on-pump cardiothoracic surgery. *Expert Opin Biol Ther 8(8),* 1187–1199.

26) Macdougall, I. C. (2008) Novel erythropoiesis-stimulating agents: a new era in anemia management. *Clin J Am Soc Nephrol 3(1*), 200–207.

27) Christensen, D. J., Gottlin, E. B., Benson R. E. & Hamilton, P. T. (2001) Phage display for target-based antibacterial drug discovery. *Drug Discov Today 6(14),* 721–727.

28) van Rooy, I., Cakir-Tascioglu, S., Couraud, P-O., Romero, I. A., Weksler, B., Storm, G., . . . Mastrobattista, E. (2010) Identification of peptide ligands for targeting to the bloodbrain-barrier. *Pharm. Res. 27,* 673–682

29) Wu, L-P., Ahmadvand, D. Su, J., Hall, A., Tan, X. Farhangrazi, Z. S. & Moghimi, S. M (2019) Crossing the blood-brain-barrier with nanoligand drug carriers self-assembled from a phage display peptide. *Nature Communications 10,* 4635

30) Pasqualini, R. & Ruoslahti, E. (1996) Organ targeting in vivo using phage display peptide libraries. *Nature. 380(6572),* 364-366.

31) Cureton, N., Korotkova, I., Baker, B., Greenwood, S., Wareing, M., Kotamraju, V. R., Teesalu, T., . . . Harris, L. K. (2017) Selective targeting of a novel vasodilator to the uterine vasculature to treat impaired uteroplacental perfusion in pregnancy. *Theranostics 7,* 3715–3713

32) Mann, A. P. Scodeller, P., Hussain, S., Braun, G. B., Mölder, T., Toome, K., . . . Ruoslahti, E. (2017) Identification of a peptide recognizing cerebrovascular changes in mouse models of Alzheimer's disease. *Nat Commun 8,* 1403.

33) Stevens, C. A., Bachtiger, F., Kong, X-D., Abriata, L. A., Sosso, G. C., Gibson, M. I. & Klok, H-A. (2021) A minimalistic cyclic ice-binding peptide from phage display. *Nat Commun 12,* 2675.

34) Wrighton, N. C., Farrell, F. X., Chang, R., Kashyap, A. K., Barbone, F. P., Mulcahy, L.S., . . . Dower, W. J. (1996) Small peptides as potent mimetics of the protein hormone erythropoietin. *Science 273,* 458–464.

35) Livnah, O., Stura, E. A., Johnson, D. L., Middleton, S. A., Mulcahy, L. S., Wrighton, N. C., . . . Wilson, I. A.(1996) Functional mimicry of a protein hormone by a peptide agonist: The EPO receptor complex at 2.8 A. *Science 273,* 464–471,

36) Hyde-DeRuyscher, R., Paige, L. A., Christensen, D. J., Hyde-DeRuyscher, N., Lim, A., Fredericks, Z. L., . . . Hamilton, P. T. (2000) Detection of small-molecule enzyme inhibitors with peptides isolated from phage-displayed combinatorial peptide libraries. *Chem Biol 7,* 17–25

37) Kay, B. K., & Hamilton, P. T. (2001) Identification of enzyme inhibitors from phage-displayed combinatorial peptide libraries. *Comb Chem High Throughput Screen 4*, 535–543.

38) Hairul Bahara, N. H., Tye, G. J., Choong, Y. S., Ong, E. B., Ismail, A. & Lim, T.S. (2013). Phage display antibodies for diagnostic applications. *Biologicals 41(4),* 209–216.

39) Chan, C. E., Zhao, B. Z., Cazenave-Gassiot, A., Pang, S. W., Bendt, A. K., Wenk, M. R., . . . Hanson, B. J. (2013) Novel phage display-derived mycolic acid-specific antibodies with potential for tuberculosis diagnosis. *J Lipid Res 54(10),* 2924-2932.

40) Kwak, M. H., Yi, G., Yang, S. M., Choe, Y., Choi, S., Lee, H., . . . Park, J. M. (2020) A Dodecapeptide Selected by Phage Display as a Potential Theranostic Probe for Colon Cancers. *Translational Oncology 13,* 100798.

41) Brissette, R., Prendergast, J. K. & Goldstein, N. I. (2006) Identification of cancer targets and therapeutics using phage display. *Current Opinion in Drug Discovery & Development, 9(3),* 363-369.

42) Hertveldt, K., Belien, T. & Volckaert, G. (2009) General M13 phage display: M13 phage display in identification and characterization of protein-protein interactions. *Methods Mol Biol 502,* 321–339

43) Sidhu, S. S, & Koide, S. (2007) Phage display for engineering and analyzing protein interaction interfaces. *Curr Opin Struct Biol 17,* 481–487

44) Lonberg, N. (2008a) Fully human antibodies from transgenic mouse and phage display platforms. *Curr Opin Immunol 20,* 450–459.

45) Lonberg, N. (2008b) Human monoclonal antibodies from transgenic mice. *Handb Exp Pharmacol 181,* 69–97.

46) De Berardinis, P. & Haigwood N. L. (2004) New recombinant vaccines based on the use of prokaryotic antigen-display systems. *Expert Rev Vaccines 3*, 673–679.

47) Hogan, S., Rookey, K. & Ladner, R. (2005) URSA: ultra rapid selection of antibodies from an antibody phage display library. *Biotechniques 38,* 536–538.

48) Rowley, M. J., O'Connor, K. & Wijeyewickrema, L. (2004) Phage display for epitope determination: a paradigm for identifying receptor-ligand interactions. *Biotechnol Annu Rev 10,* 151–188

49) Guo, J. Y., Liu, I. J., Lin, H. T., Wang, M-J., Chang, Y-L., Lin, S-C., . . . Wu, H. C. (2021). Identification of COVID-19 B-cell epitopes with phage-displayed peptide library. *J Biomed Sci 28,* 43

50) Noy-Porat, T., Makdasi, E., Alcalay, R., Mechaly, A., Levy, Y., Bercovich-Kinori, A., . . . Rosenfeld, R. (2020) A panel of human neutralizing mAbs targeting SARS-CoV-2 spike at multiple epitopes. *Nat Commun 11,* 4303.

51) Arap, W., Kolonin, M. G., Trepel, M., Lahdenranta, J., Cardó-Vila, M., Giordano, R. J., . . . Pasqualini, R. (2002). Steps toward mapping the human vasculature by phage display. *Nat. Med. 8,* 121–127

52) Chen, S., Lovell, S., Lee, S., Fellner, M., Mace, P. D. & Bogyo, M. (2021) Identification of highly selective covalent inhibitors by phage display. *Nature Biotechnology 39,* 490–498.

53) Yan, D., Han, W., Bai, Q., Zhao, X., Han, X., Du, B. & Zhu, X. (2011) Prostaglandin E(2) binding peptide screened by phage displaying: a new therapeutic strategy in rheumatoid arthritis. *Lipids in Health and Disease 10,* 75.

54) Islam, M. O., Lim, Y. T., Chan, C. E., Cazenave-Gassiot, A. Croxford, J. L., Wenk, M. R., . . . Hanson, B. J. (2012) Generation and characterization of a novel recombinant antibody against 15-ketocholestane isolated by phage-display. *Int J Mol Sci 13(4),* 4937–4948.

55) Rodi, D. J., Janes, R. W., Sanganee, H. J., Holton, R. A., Wallace, B. A. & Makowski, L. (1999) Screening of a library of phage-displayed peptides identifies human bcl-2 as a taxol-binding protein. *J Mol Biol 285(1),* 197–203.

56) Jang, H. J., Na, J. H., Jin, B. S., Lee, W. K., Lee, W. H., Jung, H. J., . . . Yu, Y. G. (2010) Identification of dinitrotoluene selective peptides by phage display cloning. *Bull Korean Chem Soc 31(12)*, 3703-3706.

57) Kushwaha, R., Payne, C. M. & Downie, A. B. (2013) Uses of phage display in agriculture: a review of food-related protein-protein interactions discovered by biopanning over diverse baits. *Comput Math Methods Med 2013,* 653759

58) Sidhu, S. S. (2000) Phage display in pharmaceutical biotechnology. *Curr Opin Biotechnol 11(6),* 610–616

59) Kay, B. K., Winter, J., & McCafferty, J. (1996). *Phage display of peptides and proteins: a laboratory manual.* Elsevier Academic Press.

60) Noren, K. A. & Noren, C. J. (2001) Construction of high-complexity combinatorial phage display peptide libraries. *Methods 23(2),* 169–178.

61) Willats, W.G. (2002).Phage display: practicalities and prospects. *Plant Mol Biol 50,* 837–854.

62) McCafferty, J., Jackson, R. H. & Chiswell, D. J. (1991) Phage-enzymes: expression and affinity chromatography of functional alkaline phosphatase on the surface of bacteriophage. *Protein Eng 4(8),* 955–961.

63) Ku, J. & Shultz, P.G. (1995) Alternate protein frameworks for molecular recognition. *Proc Natl Acad Sci USA 92(14),* 6552–6556.

64) Hiipakka, M., Poikonen, K. & Saksela, K. (1999) SH3 domains with high affinity and engineered ligand specificities targeted to HIV-1 Nef. *J Mol Biol 293(5)*, 1097–1106.

65) Volpicella, M., Ceci, L. R., Gallerani, R., Jongsma, M. A. & Beekwilder, J. (2001) Functional expression on bacteriophage of the mustard trypsin inhibitor MTI-2. *Biochem Biophys Res Commun 280(3),* 813–817.

66) Pan, Y., Du, J., Liu, J., Hai Wu, H., Gui, F., Zhang, N., . . . Yang, X. (2021).Screening of potent neutralizing antibodies against SARS-CoV-2 using convalescent patients-derived phage-display libraries. *Cell Discov 7,* 57.

67) Lim, C. C., Woo, P. C. Y. & Lim, T. S. (2019) Development of a Phage Display Panning Strategy Utilizing Crude Antigens: Isolation of MERS-CoV Nucleoprotein human antibodies. *Scientific Reports 9,* 6088.

68) Lebani, K., Martina L. Jones M. L., Watterson, D., Ranzoni, A.,Traves, R. J., . . . Mahler, S. M. (2017). Isolation of serotype-specific antibodies against dengue virus non-structural protein 1 using phage display and application in a multiplexed serotyping assay. *PloS one 12,* e0180669

69) Pardon, E., Laeremans, T., Triest, S., Rasmussen, S.G., Wohlkönig, A., Ruf, A., . . . Steyaert, J. (2014) A general protocol for the generation of Nanobodies for structural biology. *Nat Protoc 9,* 674–693.

70) Dmitriev, O. Y., Lutsenko, S. & Muylderman, S. (2016) Nanobodies as Probes for Protein Dynamics *in Vitro* and in Cells. *J Biological Chemistry 291(8),* 3767–3775.

71) Gushiken, E. (2016) *Generación de anticuerpos de dominio único específicos para CD105 humano.* Universidad Peruana Cayetano Heredia, Perú. Thesis to obtain the degree of Magister en Ciencias, mention Biochemistry and Molecular Biology.

72) Miyazaki, N., Kiyose, N., Akazawa, Y., Takashima, M., Hagihara, Y., Inoue, N., . . . Ito, Y. (2015) Isolation and characterization of antigen-specific alpaca (Lama pacos) VHH antibodies by biopanning followed by high throughput. *J Biochem Adv 158(3),* 205–215.

73) Enever, C., Coulstock, E., Swider, M. P. & Hamilton, B. (2014) Single-Domain Antibodies: An Overview. In: Dübel, S. & Reichert, J.M., *Handbook of Therapeutic Antibodies.* Wiley-VCH Verlag GmbH & Co. KGaA.

74) Hust, M., Frenzel, A. Tomszak, F., Kügler, J. & Dübel, S. (2014) Antibody Phage Display. In: Dübel, S. & Reichert, J. M., *Handbook of Therapeutic Antibodies.* Wiley-VCH Verlag GmbH & Co. KGaA.

75) Smith, G. P. (1985) Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science 228(4705),* 1315–1317.

76) Hammers, C. M. & Stanley, J. R. (2014) Antibody Phage Display: Technique and Applications. *J Invest Dermatol 134(2),* e17.

77) Muyldermans, S. (2013) Nanobodies: Natural Single-Domain Antibodies. *Annu Rev Biochem 82(1),* 775–797.

78) Evazalipour, M., Tehrani, B. S., Abolhassani, M., Morovvati, H. & Omidfar, K. (2012) Camel heavy chain antibodies against prostate-specific membrane antigen. *Hybridoma (Larchmt) 31(6),* 424–429.

79) Kijanka, M., Dorresteijn, B., Oliveira, S. & van Bergen en Henegouwen, P. M. (2015) Nanobody-based cancer therapy of solid tumors. *Nanomedicine (Lond) 10(1),* 161–174.

80) Tanha, J., Xu, P., Chen, Z., Ni, F., Kaplan, H., Narang, S. A. & MacKenzie, C. R. (2001) Optimal design features of camelized human single-domain antibody libraries. *J Biol Chem 276(27),* 24774–24780.

81) van Der Linden, R. H., Frenken, L. G., de Geus, B., Harmsen, M. M., Ruuls, R. C., Stok, W., . . . Verrips, C. T. (1999) Comparison of physical chemical properties of llama VHH antibody fragments and mouse monoclonal antibodies. *Biochim Biophys Acta 1431(1),* 37–46.

82) Fang, T., Lu, X., Berger, D., Gmeiner, C., Cho, J., Schalek, R., . . . Lichtman, J. (2018) Nanobody immunostaining for correlated light and electron microscopy with preservation of ultrastructure. *Nat Methods 15(12),* 1029–1032.

83) Tijink, B. M., Laeremans, T., Budde, M., Stigter-van Walsum, M., Dreier, T., de Haard, H. J., . . . van Dongen, G. A. (2008) Improved tumor targeting of anti-epidermal growth factor receptor Nanobodies through albumin binding: taking advantage of modular Nanobody technology. *Mol Cancer Ther 7(8),* 2288–2297.

84) Jank, L., Pinto-Espinoza, C., Duan, Y., Koch-Nolte, F., Magnus, T. & Rissiek, B. (2019) Current approaches and future perspectives for nanobodies in stroke diagnostic and therapy. *Antibodies 8(1),* 5.

85) Bannas, P., Hambach, J. & Koch-Nolte, F. (2017) Nanobodies and nanobodybased human heavy chain antibodies as antitumor therapeutics. *Front Immunol 8,* 1603.

86) Muyldermans S. (2021). Applications of Nanobodies. *Annual review animal biosciences 9,* 401–421.

87) Hashem Boroojerdi, M., Rahbarizadeh, F., Safarzadeh Kozani, P., Kamali, E. & Safarzadeh Kozani, P. (2020) Strategies for having a more effective and less toxic CAR T-cell therapy for acute lymphoblastic leukemia. *Med Oncol. 37(11),* 100.

88) Yang, E. Y. & Shah, K. (2020) Nanobodies: Next Generation of Cancer Diagnostics and Therapeutics. *Frontiers in Oncology 10,* 1182.

89) Jovcevska, I. & Muyldermans, S. (2020) The therapeutic potential of nanobodies. *BioDrugs 34(1),* 11–26.

90) Beghein, E. & Gettemans, J. (2017) Nanobody technology: a versatile toolkit for microscopic imaging, protein-protein interaction analysis, and protein function exploration. *Front Immunol 8,* 771.

91) Dekempeneer, Y., Keyaerts, M., Krasniqi, A., Puttemans, J., Muyldermans, S., Lahoutte, T., . . . Devoogdt, N. (2016) Targeted alpha therapy using short-lived alpha-particles and the promise of nanobodies as targeting vehicle. *Expert Opin Biol Ther 16(8),* 1035–1047.

92) Van Audenhove, I. & Gettemans, J. (2016) Nanobodies as versatile tools to understand, diagnose, visualize and treat cancer. *E Bio Medicine 8,* 40–48.

93) Steeland, S., Vandenbroucke, R. E. & Libert, C. (2016) Nanobodies as therapeutics: big opportunities for small antibodies. *Drug Discov Today 21(7),* 1076–1113.

94) Romao, E., Morales-Yanez, F., Hu, Y., Crauwels, M., De Pauw, P., Hassanzadeh, G. G., . . . Muyldermans, S. (2016). Identification of Useful Nanobodies by Phage Display of Immune Single Domain Libraries Derived from Camelid Heavy Chain Antibodies. *Current pharmaceutical design* 22(43), 6500–6518.

95) D'Huyvetter, M., Xavier, C., Caveliers, V., Lahoutte, T., Muyldermans, S. & Devoogdt, N. (2014) Radiolabeled nanobodies as theranostic tools in targeted radionuclide therapy of cancer. *Expert Opin Drug Deliv 11(12),* 1939–1954.

96) De Meyer, T., Muyldermans, S. & Depicker, A. (2014) Nanobody-based products as research and diagnostic tools. *Trends Biotechnol 32(5),* 263–270.

97) Chakravarty, R., Goel, S. & Cai, W. (2014) Nanobody: the "magic bullet" for molecular imaging?. *Theranostics 4(4),* 386–398.

98) Oliveira, S., Heukers, R., Sornkom, J., Kok, R. J. & van Bergen En Henegouwen, M. (2013) Targeting tumors with nanobodies for cancer imaging and therapy. *J Control Release 172(3),* 607–617.

99) Hassanzadeh-Ghassabeh, G., Devoogdt, N., De Pauw, P., Vincke, C., & Muyldermans, S. (2013). Nanobodies and their potential applications. *Nanomedicine (London, England) 8(6),* 1013–1026.

100) Unciti-Broceta, J. D., Del Castillo, T., Soriano, M., Magez, S., & Garcia-Salcedo, J. A. (2013). Novel therapy based on camelid nanobodies. *Therapeutic delivery 4(10),* 1321–1336.

101) Vincke, C., & Muyldermans, S. (2012). Introduction to heavy chain antibodies and derived Nanobodies. *Methods in molecular biology 911*, 15–26.

102) Van Bockstaele, F., Holz, J. B. & Revets, H. (2009) The development of nanobodies for therapeutic applications. *Curr Opin Investig Drugs 10(11),* 1212–1224.

103) De Genst, E., Saerens, D., Muyldermans, S., & Conrath, K. (2006) Antibody repertoire development in camelids. *Developmental and comparative immunology 30(1-2),* 187–198.

104) Zarebski, L. M., Urrutia, M., & Goldbaum, F. A. (2005). Llama single domain antibodies as a tool for molecular mimicry. *J mol biol 349(4),* 814–824.

105) Safarzadeh Kozani, P., Safarzadeh Kozani, P. & Rahbarizadeh, F. (2021a) The Potential Applicability of Single-Domain Antibodies (VHH): From Checkpoint Blockade to Infectious Disease Therapy. *Trends in Med Sci 1(2),* e114888.

106) Saerens, D., Kinne, J., Bosmans, E., Wernery, U., Muyldermans, S. & Conrath, K. (2004) Single domain antibodies derived from dromedary lymph node and peripheral blood lymphocytes sensing conformational variants of prostate-specific antigen. *J Biol Chem 279(50),* 51965–51972.

107) Zare, H., Rajabibazl, M., Rasooli, I., Ebrahimizadeh, W., Bakherad, H., Ardakani, L. S. & Gargari, S. L. M. (2014) Production of nanobodies against prostate-specific membrane antigen (PSMA) recognizing LnCaP cells. *Int J Biol Markers 29(2),* e169–179.

108) 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(8),* 2853–2857.

109) Kaliberov, S. A., Kaliberova, L. N., Buggio, M., Tremblay, J. M., Shoemaker, C. B. & Curiel, D. T. (2014) Adenoviral targeting using genetically incorporated camelid single variable domains. *Lab Invest 94(8),* 893–905.

110) Rahbarizadeh, F., Rasaee, M. J., Forouzandeh Moghadam, M., Allameh, A. A. & Sadroddiny, E. (2004) Production of novel recombinant single-domain antibodies against tandem repeat region of MUC1 mucin. *Hybrid Hybridomics 23(3),* 151–159.

111) Rahbarizadeh, F., Rasaee, M. J., Forouzandeh, M., Allameh, A., Sarrami, R., Nasiry, H. & Sadeghizadeh, M. (2005) The production and characterization of novel heavychain antibodies against the tandem repeat region of MUC1 mucin. *Immunol Invest 34(4),* 431–452.

112) Broos, K., Keyaerts, M., Lecocq, Q., Renmans, D., Nguyen, T., Escors, D., . . . Devoogdt, N. (2017) Non-invasive assessment of murine PD-L1 levels in syngeneic tumor models by nuclear imaging with nanobody tracers. *Oncotarget 8(26),* 41932–41946.

113) Broos, K., Lecocq, Q., Xavier, C., Bridoux, J., Nguyen, T. T., Corthals, J., . . . Breckpot, K. (2019) Evaluating a single domain antibody targeting human PD-L1 as a nuclear imaging and therapeutic agent. *Cancers 11(6),* 872.

114) Ji, X., Peng, Z., Li, X., Yan, Z., Yang, Y., Qiao, Z.. & Liu, Y. (2017) Neutralization of TNFα in tumor with a novel nanobody potentiates paclitaxel-therapy and inhibits metastasis in breast cancer. *Cancer Lett 386,* 24–34.

115) Kijanka, M., Warnders, F. J., El Khattabi, M., Lub-de Hooge, M., van Dam, G. M., Ntziachristos, V., . . . van Bergen En Henegouwen, P. M. P. (2013) Rapid optical imaging of human breast tumour xenografts using anti-HER2 VHHs site-directly conjugated to IRDye 800CW for image-guided surgery. *Eur J Nucl Med Mol Imaging 40(11),* 1718–1729.

116) Araste, F., Ebrahimizadeh, W., Rasooli, I., Rajabibazl, M. & Mousavi Gargari, S. L. (2014) A novel VHH nanobody against the active site (the CA domain) of tumor-associated, carbonic anhydrase isoform IX and its usefulness for cancer diagnosis. *Biotechnol Lett 36(1),* 21–28.

117) Tang, J., Li, J., Zhu, X., Yu, Y., Chen, D., Yuan, L., . . . Yang, L. (2016) Novel CD7 specific nanobody-based immunotoxins potently enhanced apoptosis of CD7-positive malignant cells. *Oncotarget 7(23),* 34070–34083.

118) You, F., Wang, Y., Jiang, L., Zhu,X., Chen, D., Yuan, L., . . . Yang, L. (2019) A novel CD7 chimeric antigen receptor-modified NK-92MI cell line targeting T-cell acute lymphoblastic leukemia. *Am J Cancer Res 9(1),* 64–78.

119) Ma, L., Zhu, M., Gai, J., Li, G., Chang, Q., Qiao, P., . . . Wan, Y. (2020) Preclinical development of a novel CD47 nanobody with less toxicity and enhanced anti-cancer therapeutic potential. *J Nanobiotechnology 18(1),* 12.

120) Romao, E., Krasniqi, A., Maes, L., Vandenbrande, C., Sterckx, Y. G., Stijlemans, B., . . . Muyldermans, S. (2020) Identification of nanobodies against the acute myeloid leukemia marker CD33. *Int J Mol Sci 21(1)*, 310.

121) Ingram, J. R., Blomberg, O. S., Rashidian, M., Ali, L., Garforth, S., Fedorov, E., . . . Dougan, M. (2018) Anti-CTLA-4 therapy requires an Fc domain for efficacy. *Proc Natl Acad Sci USA 115(15),* 3912–3917.

122) Fumey, W., Koenigsdorf, J., Kunick, V., Menzel, S., Schutze, K., Unger, M., . . . Koch-Nolte, F. (2017) Nanobodies effectively modulate the enzymatic activity of CD38 and allow specific imaging of CD38(+) tumors in mouse models *in vivo*. *Sci Rep 7(1),* 14289.

123) Maussang, D., Mujic-Delic, A., Descamps, F. J., Stortelers, C., Vanlandschoot, P., Stigter-van Walsum, M., . . . Smit, M. J. (2013) Llama-derived single variable domains (nanobodies) directed against chemokine receptor CXCR7 reduce head and neck cancer cell growth *in vivo. J Biol Chem 288(41),* 29562–29572.

124) Bradley, M. E., Dombrecht, B., Manini, J., Willis, J., Vlerick, D., De Taeye, S., . . . Cromie, K. D. (2015) Potent and efficacious inhibition of CXCR2 signaling by biparatopic nanobodies combining two distinct modes of action. *Mol Pharmacol 87(2),* 251–262.

125) Rossotti, M., Tabares, S., Alfaya, L., Leizagoyen, C., Moron, G. & Gonzalez-Sapienza G. (2015) Streamlined method for parallel identification of single domain antibodies to membrane receptors on whole cells. *Biochim Biophys Acta 1850(7),* 1397–1404.

126) Van Elssen, C., Rashidian, M., Vrbanac, V., Wucherpfennig, K. W., Habre, Z. E., Sticht, J., . . . Ploegh, H. L. (2017) Noninvasive imaging of human immune responses in a human xenograft model of graft-versus-host disease. *J Nucl Med 58(6),* 1003–1008.

127) Bartunek, J., Barbato, E., Heyndrickx, G., Vanderheyden, M., Wijns, W. & Holz, J. B. (2013) Novel antiplatelet agents: ALX-0081, a nanobody directed towards von Willebrand factor. *J Cardiovasc Transl Res 6(3),* 355–363.

128) Banihashemi, S. R., Hosseini, A. Z., Rahbarizadeh, F. & Ahmadvand, D. (2018) Development of specific nanobodies (VHH) for CD19 immuno-targeting of human B-lymphocytes. *Iran J Basic Med Sci 21(5*), 455–464.

129) Safarzadeh Kozani, P., Safarzadeh Kozani, P. & Rahbarizadeh, F. (2021b) Novel antigens of CAR T cell therapy: New roads; old destination. *Transl Oncol 14(7),* 101079.

130) Oldham, R. K. & Dillman, R. O. (2008) Monoclonal antibodies in cancer therapy:25 years of progress. *J Clin Oncol 26(11),* 1774–1777.

131) Bannas, P., Lenz, A., Kunick, V., Fumey, W., Rissiek, B., Schmid, J., . . . Koch-Nolte, F. (2015) Validation of nanobody and antibody based in vivo tumor xenograft NIRF-imaging experiments in mice using ex vivo flow cytometry and microscopy. *J Vis Exp 98,* e52462.

132) 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(3),* 544–551.

133) Dolk, E., van der Vaart, M., Lutje Hulsik, D., Vriend, G., de Haard, H., Spinelli, S., . . . Verrips, T. (2005) Isolation of llama antibody fragments for prevention of dandruff by phage display in shampoo. *Appl Environ Microbiol 71(1),* 442–450.

134) Ibañez, L. I., De Filette, M., Hultberg, A., Verrips, T., Temperton, N., Weiss, R. A., . . . Saelens, X. (2011) Nanobodies with *in vitro* neutralizing activity protect mice against H5N1 influenza virus infection. *J Infect Dis 203(8),* 1063–1072.

135) Hussack, G., Arbabi-Ghahroudi, M., van Faassen, H., Songer, J. G., Ng, K. K., MacKenzie, R. & Tanha, J. (2011) Neutralization of Clostridium difficile toxin A with single-domain antibodies targeting the cell receptor binding domain. *J Biol Chem 286(11),* 8961–8976.

136) Wrapp, D., De Vlieger, D., Corbett, K. S., Torres, G. M., Wang, N., Van Breedam, W., . . . McLellan, J. S. (2020) Structural basis for potent neutralization of betacoronaviruses by single-domain camelid antibodies. *Cell 181(5),* 1004–1015.

137) Safarzadeh Kozani, P., Safarzadeh Kozani, P. & Rahbarizadeh, F. (2021c) Aptamer-assisted Delivery of Nucleotides with Tumor-Suppressing Properties for Targeted Cancer Therapies. *Trends in Med Sci In Press,* e114909.

138) He, F., Wen, N., Xiao, D., Yan, J., Xiong, H., Cai, S., . . .Liu, Y. (2020) Aptamer-based targeted drug delivery systems: Current potential and challenges. *Curr Med Chem 27(13),* 2189–2219.

139) Cui, Y., Cui, P., Chen, B., Li, S. & Guan, H. (2017) Monoclonal antibodies: Formulations of marketed products and recent advances in novel delivery system. *Drug Dev Ind Pharm 43(4),* 519–530.

140) Banner, D. W., Gsell, B., Benz, J., Bertschinger, J., Burger, D., Brack, S., . . . Ruf, A. (2013) Mapping the conformational space accessible to BACE2 using surface mutants and cocrystals with Fab fragments, Fynomers and Xaperones. *Acta Crystallogr D Biol Crystallogr 69,* 1124–1137.

141) Baranova, E., Fronzes, R., Garcia-Pino, A., Van Gerven, N., Papapostolou, D., Péhau-Arnaudet, G., . . . Remaut, H. (2012) SbsB structure and lattice reconstruction unveil Ca2_ triggered S-layer assembly. *Nature 487,* 119–122.

142) Ehrnstorfer, I. A., Geertsma, E. R., Pardon, E., Steyaert, J. & Dutzler, R. (2014) Crystal structure of a SLC11 (NRAMP) transporter reveals the basis for transition-metal ion transport. *Nat Struct Mol Biol 21,* 990–996.

143) Hassaine, G., Deluz, C., Grasso, L., Wyss, R., Tol, M. B., Hovius, R., . . . Nury, H. (2014) X-ray structure of the mouse serotonin 5-HT3 receptor. *Nature 512,* 276–281.

144) Korotkov, K. V., Pardon, E., Steyaert, J. & Hol, W. G. (2009) Crystal structure of the N-terminal domain of the secretin GspD from ETEC determined with the assistance of a nanobody. *Structure 17,* 255–265

145) Park, Y. J., Pardon, E., Wu, M., Steyaert, J. & Hol, W. G. (2012) Crystal structure of a heterodimer of editosome interaction proteins in complex with two copies of a cross-reacting nanobody. *Nucleic Acids Res 40,* 1828–1840.

146) Pathare, G. R., Nagy, I., S´ledz´, P., Anderson, D. J., Zhou, H. J., Pardon, E., . . . Baumeister, W. (2014) Crystal structure of the proteasomal deubiquitylation module Rpn8-Rpn11. *Proc Natl Acad Sci USA 111,* 2984–2989.

147) Chaikuad, A., Keates, T., Vincke, C., Kaufholz, M., Zenn, M., Zimmermann, B., . . . Müller, S. (2014) Structure of cyclin G-associated kinase (GAK) trapped in different conformations using nanobodies. *Biochem J 459*, 59–69

148) Schotte, L., Thys, B., Strauss, M., Filman, D. J., Rombaut, B. & Hogle, J. M. (2015) Characterization of poliovirus neutralization escape mutants of single-domain antibody fragments (VHHs). *Antimicrob Agents Chemother 59*, 4695–4706.

149) Koromyslova, A. D. & Hansman, G. S. (2015) Nanobody binding to a conserved epitope promotes norovirus particle disassembly. *J Virol 89,* 2718–2730.

150) Maass, D. R., Sepulveda, J., Pernthaner, A. & Shoemaker, C. B. (2007) Alpaca (Lama pacos) as a convenient source of recombinant camelid heavy chain antibodies (VHHs). *J Immunol Methods 324(1–2),* 13–25.

151) Cabezas, S., Rojas, G., Pavon, A., Alvarez, M., Pupo, M., Guillen, G. & Guzmán, M. (2008) Selection of phage-displayed human antibody fragments on Dengue virus particles captured by a monoclonal antibody: Application to the four serotypes. *J Virol Methods 147(2),* 235–243.

152) Lamdan, H., Ayala, M., Rojas, G., Munoz, Y., Morera, Y., Guirola, O., . . . Gavilondo, J (2011) Isolation of a novel neutralizing antibody fragment against human vascular endothelial growth factor from a phage-displayed human antibody repertoire using an epitope disturbing strategy. *J Biotechnol 151(2),* 166–174.

153) Qiagen (2018) *Quick-Start protocol. QIAquick® Gel Extraction Kit QIAquick® PCR & Gel Cleanup Kit.* Obtained from https://www.qiagen.com/us/knowledge-and-support/knowledge-hub/search/resources/?categories=RESOURCES&page=0& filters=%7B%7D& query=Quick-Start%20protocol.%20QIAquick. (Revised July 2018)

154) Biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G., Schmidt, T., . . . Schwede, T. (2014). SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic acids research 42*, W252–W258.

155) Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research 25(17),* 3389–3402

156) Gill, S. C., & von Hippel, P. H. (1989). Calculation of protein extinction coefficients from amino acid sequence data. *Analytical biochemistry 182*(2), 319–326.

157) Guruprasad, K., Reddy, B. V., & Pandit, M. W. (1990). Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting in vivo stability of a protein from its primary sequence. *Protein engineering 4*(2), 155–161.

158) Ikai, A. (1980). Thermostability and aliphatic index of globular proteins. *Journal of biochemistry 88*(6), 1895–1898.

159) Kyte, J., & Doolittle, R. F. (1982). A simple method for displaying the hydropathic character of a protein. *Journal of molecular biology 157*(1), 105–132.

160) Villegas, W., Acereto, P. & Vargas, M. (2006). Análisis Ultravioleta-visible. La teoría y la práctica en el ejercicio profesional. Departamento Editorial Universidad Autónoma de Yucatán, Mérida, México.

161) Harmsen, M. M. & De Haard, H. J. (2007) Properties, production, and applications of camelid single-domain antibody fragments. *Appl Microbiol Biotechnol 77(1),* 13–22.

162) Marchler-Bauer, A., Bo, Y., Han, L., He, J., Lanczycki, C. J., Lu, S., . . . Bryant, S. H. (2017) CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic acids research 45*(D1), D200–D203.

163) Marchler-Bauer, A., Derbyshire, M. K., Gonzales, N. R., Lu, S., Chitsaz, F., Geer, L. Y., . . . Bryant, S. H. (2015) CDD: NCBI's conserved domain database. *Nucleic acids research 43*, D222–D226.

164) Marchler-Bauer, A., Lu, S., Anderson, J. B., Chitsaz, F., Derbyshire, M. K., DeWeese-Scott, C., . . . Bryant, S. H. (2011) CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic acids research 39*, D225–D229.

165) Marchler-Bauer, A. & Bryant, S. H. (2004) CD-Search: protein domain annotations on the fly. *Nucleic acids research 32*, W327–W331.

166) Muyldermans, S., Atarhouch, T., Saldanha, J., Barbosa, J. & Hamers, R. (1994) Sequence and structure of VH domain from naturally occurring camel heavy chain immunoglobulins lacking light chains. *Protein Eng 7(9),* 1129–1135.

167) Vu, K. B., Ghahroudi, M. A., Wyns, L. & Muyldermans, S. (1997) Comparison of llama V(H) sequences from conventional and heavy chain antibodies. *Mol Immunol 34(16–17*), 1121–1131.

168) Harmsen, M. M., Ruuls, R. C., Nijman, I. J., Niewold, T., Frenken, L. G. J., De Geus, B. (2000) Llama heavy-chain V regions consist of at least four distinct subfamilies revealing novel sequence features. *Mol Immunol 37(10),* 579–590.

169) Nguyen, V. K., Hamers, R., Wyns, L. & Muyldermans, S. (2000) Camel heavy-chain antibodies: diverse germline V(H)H and specific mechanisms enlarge the antigen-binding repertoire. *EMBO J 19(5),* 921–930.

170) Sircar, A., Sanni, K. A., Shi, J. & Gray, J. J. (2011) Analysis and modeling of the variable región of camelid single-domain antibodies. *J Immunol 186(11),* 6357–6367.

171) Kirchhofer, A., Helma, J., Schmidthals, K., Frauer, C., Cui, S., Karcher, A., . . . Rothbauer, U. (2010) Modulation of protein properties in living cells using nanobodies. *Nat Struct Mol Biol 17(1*), 133–138.

172) Gasteiger E., Hoogland C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., Bairoch A. (2005) Protein Identification and Analysis Tools on the ExPASy Server. In: Walker, J. M., *The Proteomics Protocols Handbook* (pp 571-607). Humana Press.

173) 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(3),* 521–526.

174) Dumoulin, M., Conrath, K., Van Meirhaeghe, A., Meersman, F., Heremans, K., Frenken, L. G. J., . . . Matagne, A. (2002) Single-domain antibody fragments with high conformational stability. *Protein Sci 11(3),* 500–515.

175) Pérez, J. M. J., 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(1),* 74–83.

176) Yasui, H., Ito, W. & Kurosawa Y. (1994) Effects of substitutions of amino acids on the thermal stability of the Fv fragments of antibodies. *FEBS Lett 353(2),* 143–146.

Chapter III: Subcloning, expression and purification of the anti-CD105 VHH

Abstract

The nineteen different anti-CD105 VHH cDNAs cloned in the phagemid vector pHEN2 were subcloned into the plasmid vector pET-22b(+) (Novagen) for expression as recombinant proteins in *Escherichia coli* strain BL21. The VHH insertion in the correct reading frame into the vector was checked by digestion with the restriction enzymes and by PCR amplification using two pairs of primers specific for the vector and the VHH sequence. The expression conditions of the recombinant VHH proteins were standardised by varying the induction time, inducer concentration and antibiotic concentration. The VHH expression by induction with 1mM IPTG was visualised in 15% SDS-PAGE. Thirteen VHHs were positive to CD105 by ELISA from a total of nineteen recombinant VHH proteins analysed. They were purified by affinity chromatography in Ni-NTA agarose columns from the supernatant of the bacterial cultures obtaining approximately 5 mg of each protein per 100 ml culture and they were shown as a single band in 15% SDS-PAGE gels stained with Coomassie blue. The pure recombinant VHH proteins were used for functionalisation and application on microarrays (as biochips) to determine the binding affinity and specificity by Surface Plasmon Resonance imaging (SPRi).

3.1 Scientific context

3.1.1 Purified proteins

Obtaining purified proteins for application in various experimental studies, such as structure, function, *in vitro* biological assays, recognition or blocking tools, drugs, diagnostics, therapy, etc. has always been of great importance and interest to researchers, and nowadays, it has become a necessity. Purification involves a separation of the protein from its source due to differences in their physical properties[1]. Traditionally proteins are obtained from tissues or biological samples where they are found in abundance, however, with technological advances, it is now common to obtain them by expression in model organisms such as bacteria: *Escherichia coli* (*E. coli*)[2-4], *Bacillus subtilis*, among others[5,6]; yeasts: *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*[7], or cell cultures of insects or mammals[8-10].

3.1.2 Recombinant proteins in *E. coli*

E. coli is widely used as a host for the expression of recombinant non-glycosylated proteins for various purposes in research and industry, such as diagnostics, therapeutics, vaccines, etc. The production of recombinant proteins in *E. coli* culture is facilitated by using chemically defined media and relatively low-cost procedures for purification of recombinant proteins. There is a variety of systems of expression in *E. coli* due to the very complete genetic and physiological characterisation of this organism[11].

There are many advantages offered by this microorganism as a host, so the knowledge and experience acquired for its management over the course of multiple studies is extensive. Other advantages are the genetic manipulation, growth faster than yeast, insect and mammal cell lines, low cost nutritional requirements, cultivation does not require complex or sophisticated equipment, a wide variety of stable expression vectors and the approval by regulatory entities for its use as a host for the production of biopharmaceuticals and proteins for therapeutic use[2,4,12-15].

The disadvantages of using *E. coli* as a host for producing recombinant proteins include proteases produced by the microorganism that can destroy recombinant proteins, and the secreted toxins restrict its use in food. Protein secretion into the culture medium is not very efficient and has limited capacity for the correct folding of proteins with numerous disulphide bridges. Additionally, the major disadvantage accounts for lacking the capacity to perform common post-translational modifications in proteins of eukaryotic origin (such as glycosylation), which are often necessary for the stability, biological activity and functionality of the manufactured recombinant protein. Even though, these disadvantages, numerous complex proteins are produced in *E. coli* with the biological activity required to be used[3,4,16-18].

In order to overcome these disadvantages, genetic modifications have been made in *E. coli* strains. Such as: the IS186 insertion in the promoter of the protease-deleting lon gene[19] and a deletion in the outer membrane protease ompT gene[20] to decrease proteolytic degradation of recombinant proteins; or modifications in thioredoxin reductase (trxB), or glutathione reductase (gor) to promote proper folding of disulphide bridges in the cytoplasm[21]. Strains capable of post-transcriptional modifications such as phosphorylation by carrying the tyrosine kinase gene have also been developed[4,22]. To improve protein production performance in *E. coli*, protein-production strains, that is, *E. coli* strains with improved protein-producing characteristics, have also been isolated[23-25].

3.1.3 Expression systems in *E. coli*

The parameters usually considered for the expression of recombinant proteins in *E. coli* are: the stability of the expression vector and the transcript (mRNA), the transcriptional and translational efficiency, the half-life of the protein in the proteolytic environment and the folding. It is also necessary to consider factors that can influence expression such as: the number of copies of the expression vector, the characteristics of the gene to be expressed, the promoter used, the strain used, the composition of the culture medium, the fermentation requirements and the culture strategy[4,6,17,26,27].

Expression systems that favour secretion of the protein towards the culture supernatant are frequently applied on a large scale because no cell disruption is required to obtain it[28,29], even if the product is obtained diluted. There are also expression vectors designed to direct the export of the recombinant protein into the periplasmic space. With these, the

isolation process is facilitated but purification is complicated by the large amount of contaminating cellular debris and carries a risk of loss of activity by proteases[30]. Other systems involve the production of the protein within the cytoplasm, so they do not require solubilisation or renaturation for their correct folding, however, a long and complex process is needed to obtain purified proteins[3,26]. And, there are other expression systems that involve the production of insoluble recombinant proteins in the form of cytoplasmic inclusion bodies[16,30].

These expression systems have a high production yield but require renaturation and refolding of the proteins, which is a very costly process. For the expression of recombinant proteins, taking the advantages provided by the periplasm over the cytosol, plasmid vectors have been developed where recombinant genes can be cloned fused to secretion signals that guide the transport of the recombinant protein to this compartment and the outside[4]. In addition, a set of vectors have been designed to improve the expression of proteins with solubility problems through parallel cloning of the same PCR product into 12 different expression vectors that assess protein expression under different promoter strengths, fusion tags, as well as solubility-enhancing proteins[31].

On the other hand, in the last decades of the 20th century, different strains of *E. coli* have been evaluated for their characteristics in recombinant protein production and line B emerged as the best. BL21(DE3), a derived strain, has now become the host of choice for recombinant protein production. Therefore, the genome of the B line strains have been sequenced to understand the molecular basis of phenotypes useful for recombinant protein synthesis[2,32,33]. Often, secretion of recombinant proteins yields better results than production in the cytoplasm[34]. *E. coli* line B secretes more proteins than other laboratory strains, due in part to the presence of a second T2S secretion system[2,35].

3.1.4 Expression Vectors

Expression vectors for *E. coli* are usually plasmids that typically contain: an origin of replication that determines the number of copies of the vector, a promoter that regulates the transcription of the gene encoding the protein of interest, a multiple cloning site (MCS) where the restriction enzyme recognition sequences are located, a selection marker that can be an antibiotic resistance gene (such as ampicillin, kanamycin and tetracycline) and a transcription termination gene. At the laboratory level, ampicillin is the most commonly used marker for the selection of recombinants, but at the industrial level, it is not always used due to its probable allergenic effect in the case of the production of therapeutics for human consumption. As an alternative to the use of antibiotics, mutant strains of *E. coli* have been developed in which an essential gene has been deleted and this deleted essential gene is fused with the lac operator[10,36], or combined vectors have been designed to increase their stability inside the cell[37] or the use of triclosan as a selection agent[38].

The lac promoter, whose expression is regulated by isopropyl-b-D-thiogalactopyranoside (IPTG), is often found as part of the expression vector. When higher production is required, such as at industrial level, the tryptophan promoter (trp operon) is

preferentially used in the vector, as it can be induced in culture media lacking this amino acid.

3.1.4.1 pET system (Novagen)

This system was developed for the cloning and expression of recombinant proteins in *E. coli*. The sequence of interest is cloned into pET plasmids under the control of transcriptional and translational signals (optionally) from the bacteriophage T7, therefore, to induce expression, the host cell must have a source of T7 RNA polymerase. For this, the host can be infected with λCE6, a phage carrying the T7 RNA polymerase gene under the control of the λ pL and pI promoters, or an expression host can be transformed with the plasmid containing a chromosomal copy of the T7 RNA polymerase gene under lacUV5 control. In the latter case, expression is induced by the addition of IPTG to the bacterial culture.

T7 RNA polymerase is so selective and active that, when fully induced, almost all of the cell resources are converted for gene expression, such that the protein of interest can comprise more than 50% of the total cellular protein a few hours after induction. However, it is possible to attenuate expression levels by reducing the inducer concentration. Another advantage of the system is the transcriptional silence of the cloned sequence in the uninduced state. It is recommended that the sequences of interest are initially cloned into a host without T7 RNA polymerase to eliminate instability of the recombinant plasmid due to the production of proteins potentially toxic to the host cell. Although in some cases, such as with harmless proteins, it is possible to clone directly into expression hosts[39]. The pET expression systems consist of the host strains BL21, BL21(DE3) and BL21(DE3)pLysS whose genome has been sequenced (GenBank entry CP001509.3)[40] and plasmid vectors.

The pET plasmid vectors were originally constructed by Studier and colleagues[41-43] and renovated by Novagen to improve the cloning, detection and purification of the proteins of interest. There are two categories of vectors: transcription vectors, designed for the expression of genes with their own prokaryotic ribosome binding site and the AUG start codon, and translation vectors, which contain the ribosome-binding site of the major capsid protein of phage T7 and are used for the expression of genes without their own ribosome-binding site. The pET-22b(+) vector has the ampicillin resistance gene, the T7 lac promoter, a signal sequence that directs the synthesised protein to the periplasm and the polyHis tag (6X His tag) at the carboxyl terminus for protein purification[39] (Figure 3.1).

The applications of expressed proteins are very varied, whether analytical quantities of the protein are needed (for activity studies, mutant characterisation, ligand interaction screening or antigen preparation) or large quantities (for structural studies, use as a reagent or affinity matrix preparation). Although only a combination of vector, host strain and culture conditions may work best for large-scale purification[14].

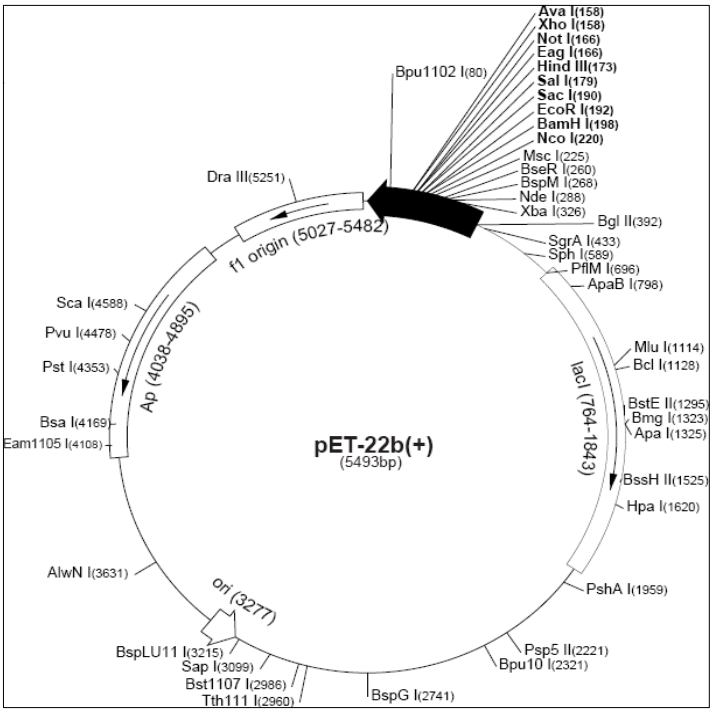


Figure 3.1- Plasmid expression vector pET-22b(+)(Novagen)[39].

3.1.5 Protein Purification

The development of expression processes has required the parallel development of purification processes that seek to obtain the maximum amount of functional protein with the fewest contaminants and the minimum possible steps. The most important consideration for a protein purification scheme is the application that will be given to that protein[1]. Therefore, it is necessary to choose a system that allows obtaining the protein in sufficient quantity and quality for the experimental assays. It is also necessary to know the characteristics, folding and behaviour of the protein to avoid possible processes that affect its correct structure, and therefore its function, such as degradation, defolding, aggregation, etc. during purification or storage.

The conditions of the protein solvent medium at each purification step are essential to maintain the stability and function of the protein. Thus, it is important to select the most appropriate buffers for the requirements of each step. The most frequently used biological buffers are phosphate buffer saline (PBS) (pH 5.8-8.0), MOPS (pH 6.5-7.9), HEPES (pH 6.8-8.2) and Tris (pH 7.5-9.0)[44], as they work well at physiological pH. However, it will always depend on the nature of the protein and its application. In addition, if necessary, compounds that play a role related to the purity and stability of the protein can be added, such as protease inhibitors (prevent degradation), reducing agents (protect against oxidative damage), metal chelators (inactivate metalloproteases), osmolytes (stabilise the protein and increase solubility) or ionic stabilisers (increase solubility)[1]. As a general recommendation, the protein should be kept cold during the purification process at 4° C, as this temperature ensures structural integrity of the proteins and reduces the risk of degradation.

3.1.6 Recombinant protein purification strategies in *E. coli*

There are two general ways in which protein purification can be divided according to its expression form: soluble and insoluble protein purification. The strategy used for purification will depend on this. On the other hand, recombinant proteins expressed in *E. coli* can be secreted into the extracellular medium or remain intracellular, either in the periplasm or the cytoplasm. Synthesised and secreted recombinant proteins are always soluble. Recombinant proteins produced at high levels are often located in the cytoplasm forming inclusion bodies (insoluble aggregates of misfolded proteins, which lack biological activity and can be observed by phase microscopy as dark intracellular particles between 0.2– 0.6 mm in size[45]), although they can also be found in soluble form. In the periplasm, proteins tend to remain in soluble form, as this is a more favourable environment for folding and disulphide bridge formation. The extracellular secreted proteins are generally more stable and smaller[4].

The denatured recombinant protein is the main component of the inclusion bodies and initial separation of the inclusion bodies by centrifugation is an effective purification step. The native recombinant molecule is solubilised using chaotropic agents such as urea or guanidinium chloride[46] through a process known as extraction. After recovery of the protein of interest, it is necessary to restore the tertiary structure of the protein through a renaturation step, which consists of decreasing the concentration of this agent by dialysis, direct dilution or using a chromatographic column[47,48]. Renaturation can be a simple process for small proteins that do not contain disulphide bridges, but it can also be a complex and low-recovery process, so that, the optimal renaturation conditions are protein-dependent and can become a tedious and undesirable process. Despite this, there has been increasing advocacy for the production of recombinant proteins in inclusion bodies as their formation allows a high yield and purity of the product and the production of denatured toxic proteins, being in some cases the only alternative for the production of a protein[49,50]. In that sense, a 39 amino acid signal sequence of the *E. coli* TorA protein (ssTorA) has been reported to promote the formation of inclusion bodies even in highly soluble proteins such as thioredoxin and maltose binding protein (MBP) and could help in particular protein production strategies[51].

In the case of purification of soluble proteins that remain intracellularly, it is necessary to rupture the cells that contain them. Methods such as freeze/thaw, ultrasound, buffering with salts and lysozyme, high-pressure homogenisation or permeabilisation with organic solvents are used. The chosen method will depend on the nature, composition, folding and stability of the protein. Nucleic acids, lipids, polysaccharides and other proteins, as well as other small molecules are the contaminants present in protein extracts. Most can be separated from proteins by molecular size-based methods such as dialysis, ultrafiltration or gel filtration[4].

For the purification of soluble proteins secreted into the medium, methods are mainly based on the use of fusion partners and affinity tags[52,53]. This consists of the addition of a small fusion peptide to recombinant proteins (usually at their C-terminal end) that allows

improving solubility and increasing soluble protein yield in the case of fusion partners and, in addition, protein purification by affinity chromatography in the case of affinity tags. As some affinity tags can be mentioned the GST-tag (glutathione S-transferase)[54], MBP-tag, SUMO-tag (members of the small ubiquitin like modifier (SUMO) family of proteins), S-Tag[55], T7-Tag, HB-tag (heparin)[56,57] and poly His-tag, which is 4 or more consecutive Histidine residues (such as 6X His-tag)[58-60], the latter being one of the most widely employed affinity tags. Combinations that improve the yield of purified proteins have been reported, such as NT-11, sequence of the first 11 amino acid residues of a duplicated carbonic anhydrase from *Dunaliella* with His-tag, which improved the solubility of all recombinant proteins tested[61]. Or MBP-Pyr, a truncated maltotriose-binding protein from *Pyrococcus furiosus* with a modified His-tag composed of intercalated histidine and glutamate residues (HE-tag) that achieved higher expression and solubility of HE-MBP(Pyr) fusion proteins than H-MBP(Pyr)[62].

The insertion of a protease cleavage site between the fusion peptide and the protein of interest allows removal of the tag if required. Affinity purification allows obtaining high levels of purity using a single chromatographic step[63], that is, this method may be the only step required to achieve an adequate degree of purity, depending on the intended application of the protein.

3.1.7 Affinity chromatography for protein purification

In general, affinity chromatography is a robust purification technique used in the first steps of the purification scheme. It is based on the specific and reversible binding of a protein to a matrix-bound ligand. The ligand can bind directly to the protein of interest, such as, for example, cAMP-resin binds to the cAMP-binding PKA RIα and RIIβ proteins[64] or to a molecular tag that is linked to the protein. The stationary phase consists of an inert matrix (typically agarose or cross-linked polyacrylamide) covalently bound to a ligand that binds specifically to the protein.

Purification can be selective, when a specific ligand is used for a protein or for a molecular tag covalently attached to the protein, or non-selective, when the ligand binds to a group of proteins with similar binding capacity, such as for example, lectin binds to SARS-CoV-2 glycoprotein S expressed to produce a vaccine for COVID 19[65]. In both, proteins are poured onto a column under conditions that promote binding between the protein (or tag) and its ligand. The bound protein is washed under conditions that break non-specific interactions between contaminating proteins and the matrix, but not specific interactions between the protein of interest and the ligand. The bound protein is then eluted using a buffer containing a protein competing molecule or under conditions that disrupt all protein/protein interactions. The protein-competing molecule binds to the ligand, displacing the protein of interest. Methods to elute proteins from the stationary phase by breaking all protein/protein interactions include a change in pH or ionic strength of the buffer. These methods can affect the stability of the protein and it is suggested that the eluted protein be neutralised immediately or diluted to minimise damage. For some

variants of affinity chromatography, particular elution conditions have been described to maximise the amount of functional protein recovered[1,66].

Ritchie compiled a list of the main protein purification suppliers and methods cited in publications that were reviewed by Labome (https://www.labome.com). Affinity and molecular exclusion methods are the most cited in the literature. GE Healthcare is the largest supplier for all methods, Qiagen is the most significant supplier of HIS-tag-based protein purification methods and Sigma Aldrich of those based on FLAG-tags[1].

3.1.7.1 The QIAexpress Ni-NTA system (Qiagen)

This protein purification system is based on the high selectivity of the patented Ni-NTA (Nickel-Nitrilotriacetic Acid) resin for His-tagged proteins. It has the advantage to allow the purification of almost any His-tagged protein, produced from any expression system and under native or denaturing conditions. Moreover, it does not depend on the three-dimensional structure of the protein so that purification can be done in a single step, either from dilute solutions or from crude lysates. NTA, which has four chelation sites for nickel ions (Ni^{+2}), binds nickel more strongly than metal chelation purification systems that have only three sites available for interaction with metal ions (Figure 3.2a). The additional chelation site prevents Ni^{+2} leaching and results in higher binding capacity and higher purity protein preparations[60].

The purification of proteins with poly His-tags by Immobilised Metal Affinity Chromatography (IMAC) consists of 3 steps: a) binding, b) washing and c) elution.

a) Binding

The imidazole is part of the histidine structure (Figure 3.2b) and can bind to the Ni^{+2} immobilised by the NTA groups in the matrix. Thus, recombinant proteins with one or more His-tags at the N-terminus or C-terminus can bind to the Ni-NTA groups on the matrix (Figure 3.2c) with a very high affinity independent of folding. It is only necessary that the proteins contain more than two available Histidine residues to interact with Ni^{+2}, although the fewer the number of accessible Histidine residues, the weaker the binding strength. Therefore, proteins without His-tags but with very exposed Histidine residues nearby on the surface can also bind, however, this interaction is much weaker than binding with 6xHis-tags and can be separated by washing under stringent conditions.

In general, low imidazole concentrations prevent non-specific and low-affinity binding of contaminating proteins to the Ni-NTA matrix, leading to higher purity in fewer steps. For purification under native conditions, binding can be optimised for each protein by adjusting the imidazole concentration and/or the buffer pH. Usually up to 20 mM imidazole are used for most proteins but if this affects the binding of the His-tag recombinant protein it is reduced to 1-5 mM. In addition, the Ni-NTA matrix is equilibrated with a buffer containing 10-20 mM imidazole prior to binding to "shield" it and decrease or eliminate non-specific binding of weakly interacting contaminating proteins.

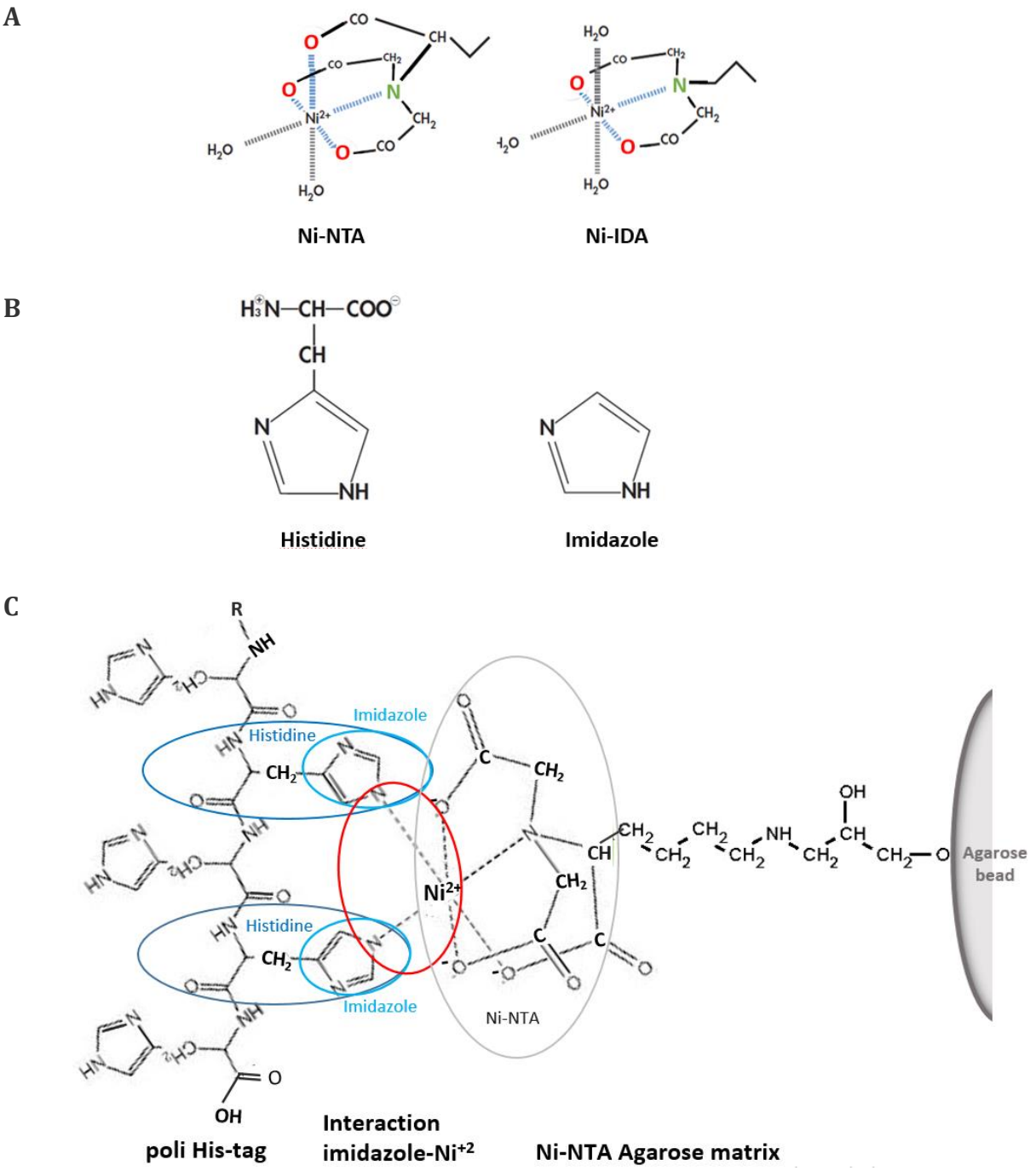


Figure 3.2- Structure and interactions of Ni-NTA with poly His-tag recombinant proteins. a) Interactions of Ni^{+2} with two different metal chelating matrices, b) Chemical structure of amino acid Histidine and Imidazole group, c) Interaction between the Ni-NTA. Agarose with neighbouring Imidazole residues in the poly His-tag. Adapted from Qiagen[60] and Magdeldin & Moser[67]).

Binding of the purified proteins to the Ni-NTA matrix can be carried out by batch or column procedures. The first, batch procedure involves binding the protein to the Ni-NTA matrix in solution and then packing the protein-matrix complex into a column for washing and elution steps. If 6xHis-tag recombinant proteins are secreted into the medium, the protein concentration is low or the 6xHis-tag of the proteins is not fully accessible for binding, it is advisable to use this strategy to favour binding. In column procedure, the Ni-NTA matrix is first introduced into the column and equilibrated with buffer. Then, the

solution with the proteins is slowly applied to the column. The washing and elution steps are identical to those of the batch procedure.

b) Washing

Wash buffers may include reagents that allow efficient removal of non-specific binding contaminants. The optimum pH and imidazole concentration conditions for wash buffers vary slightly for each protein and must be determined empirically. Contaminant proteins that interact with Ni-NTA groups of the matrix can be washed under stringent conditions that are achieved by addition of imidazole at a concentration of 10-50 mM or by lowering the pH to 6.3. Normally in bacterial expression systems, recombinant proteins are expressed in large amounts, and copurified contaminant proteins in very small amounts, so it is not necessary to use very stringent washing conditions. In eukaryotic expression systems, the relative abundance of contaminating proteins is higher, and even higher in non-denaturing procedures, so it is necessary to increase the stringency of washing. Gradual decrease of the wash buffer pH or progressive increase of the imidazole concentration are recommended. As the binding of contaminating proteins is higher under native conditions than under denaturing conditions, it is recommended to use imidazole in low concentrations in the binding and washing buffers.

c) Elution

Purified proteins can be eluted by addition of high concentrations of imidazole as a competitor, by lowering the pH or by addition of ion chelating agents. If the imidazole concentration is increased to 100-300 mM, 6xHis-tag proteins will dissociate because they can no longer compete for binding sites on the Ni-NTA matrix. In the same way, the histidine residues of the 6xHis-tag have a pKa of about 6.0 so if the pH is reduced (pH 4.5-5.3) they will be protonated and the 6xHis-tag protein will no longer be able to bind Ni^{+2} and it will then dissociate from the Ni-NTA matrix. Elution conditions are reproducible, but must be determined for each 6xHis-tag protein. Monomers usually elute at about pH 5.9, while aggregates and proteins containing more than one 6xHis-tag elute at about pH 4.5. EDTA chelates Ni^{+2} and removes it from the NTA groups, causing the 6xHis-tag protein elute as a protein-metal complex. Although all elution methods (imidazole, pH and EDTA) are equally effective, imidazole is the mildest and is recommended under native conditions, when a reduction in pH may damage the protein, or when the presence of metal ions in the eluate may have an adverse effect on the purified protein.

3.2 Procedures

3.2.1 PCR amplification of VHH sequences at pHEN2 for subcloning

For a 50 ul reaction, 10 ng of DNA, 1X PCR reaction buffer (Invitrogen), 2 mM $MgCl_2$, 0.5 μM of each primer VHHSfi(*NcoI*) and ALLVHHR2-*Not*), 200 μM of each dNTP and 0.2 units of PlatinumTaq DNA (Invitrogen) were used. The conditions were: an initial denaturation step 94° C for 2 minutes, followed by 30 amplification cycles of 94° C for 30 seconds, 58° C for 30 seconds and 72° C for 30 seconds, extension 10 minutes at 72° C and a final cooling step at 4° C. Products were visualised by 1.2% agarose gel electrophoresis in 1X

TBE buffer at 80V for 60 minutes in the presence of a molecular-weight size marker and stained with ethidium bromide. Product purification was performed with QIAquick Gel Extraction Kit according to the protocol proposed by the manufacturer (Qiagen[68,69]) and described in section 3.2.3. The quantification was performed by Nanodrop at 260 nm.

3.2.2 Double digestion with *Not*I and *Nco*I

Digestion of each amplified VHH used 1 ug of DNA and digestion of pET-22b(+) (Novagen) expression vector used 1.5 ug of DNA. Reactions were 20 ul and used: 1X enzyme digestion buffer (NEB), 1 Unit of *NotI* (NEB) and 1 Unit of *NcoI* (NEB). Each sample was incubated at 37° C overnight, then incubated at 65° C for 15 minutes for inactivation. Products were visualised by 1.5 % agarose gel electrophoresis in 1X TBE buffer at 80 V for 120 minutes, in the presence of a molecular-weight size marker, and stained with ethidium bromide.

3.2.3 Recovery and purification of DNA fragments from agarose gel

DNA fragments of the corresponding size (VHH 450bp and pET 5438bp) were cut from the gel and placed in a new microtube to be recovered with the QIAquick Gel Extraction Kit according to the protocol proposed by the manufacturer(Qiagen[68,69]) and detailed below. Each recovered gel fragment with the DNA of interest was weighed and 3 volumes of QG buffer was added for each mg of gel (1 mg ~1 ul) and incubated at 50° C for 10 minutes or until completely dissolved. One volume of isopropanol (initial gel volume) was added to the sample and mixed. The minicolumn was placed in a 2 ml microtube, the sample was added and was centrifuged for 1 minute at 15142 g. The minicolumn was removed from the microtube, the decanted liquid was discarded and the minicolumn was returned to the tube. 0.5 ml of QG buffer was added to the minicolumn, it was centrifuged for 1 minute and the decanted liquid was discarded. For washing, 0.75 ml of Buffer PE was added and centrifuged for 1 minute. The minicolumn was transferred to a new sterile 1.5 ml microtube and DNA was eluted by adding 50µl of EB Buffer (10 mM Tris-Cl, pH 8.5) or nuclease-free ultrapure water and centrifuging at 13000 rpm for 1 minute at room temperature (RT°). The products were quantified by Nanodrop at 260 nm.

3.2.4 Ligation

The cut VHH DNA fragments were referred to as inserts as they will be inserted into the vector after ligation. The insert/vector ratio was 5:1 according to the following formula:

[(ng vector x Kb insert)/Kb vector]x molar ratio (insert/vector) = ng insert

The reactions were 10 ul and used 50 ng of pET-22b(+) vector, 20.68 ng of insert, 1X T4 DNA ligase buffer (NEB) and 1 Unit of recombinant T4 bacteriophage DNA ligase (NEB). It was used as a positive control of ligation to λ/*Hind*III and as a negative control (background control) to a reaction where no insert was added (self-ligated plasmid reaction). Each reaction was incubated at 16° C overnight and then inactivated at 65° C for 15 minutes. The samples were immediately used for bacterial transformation.

3.2.5 Preparation of competent cells

E. coli strain BL21 bacteria were grown on an LB plate at 37° C overnight. One colony was picked and inoculated in 4 ml of LB medium and then, it was incubated at 37° C overnight on shaking (250 rpm). The culture was added to 96 ml of fresh LB medium and incubated at 37° C on shaking (250 rpm) until the culture reached an OD_{600nm} between 0.85 and 0.98 (approximately 2 hours 15 minutes). The bacterial culture was distributed into two sterile, pre-frozen 50 ml Corning tubes for centrifugation at 4000 g for 15 minutes at 4° C. After removal of the medium, the bacteria were re-suspended in 20 ml of ice-cold solution of 0.1 M $CaCl_2$ in a single tube and left on ice for 30 minutes. It was then centrifuged at 4000 g for 15 minutes at 4° C. The supernatant was discarded, the bacteria were re-suspended in 2 ml of ice-cold solution of 0.1M $CaCl_2$ and 15% glycerol and distributed in aliquots of 100 µl into 1.5 ml microtubes for their cryopreservation at -70° C or for immediate use for bacterial transformation. Each tube contained approximately 9 x 10^8 bacteria according to tube 3 of the MacFarland scale. Competent bacteria were evaluated for their transformation efficiency using plasmid pUC19.

3.2.6 Transformation of *E. coli* by the $CaCl_2$ method

5 µl of ligation product or recombinant plasmid were added to a tube containing 100 µl of competent cells. Each tube with the samples were mixed gently and incubated on ice for 30 minutes. Cells were heat shocked by placing them at 42° C for 45 seconds and immediately returning them to ice for 2 minutes. 900 µl of SOC medium were added on the cells and incubated at 37° C on shaking at 250 rpm for 45 minutes. Bacteria were collected by centrifugation at 12902 g for 15 seconds and resuspended in 100 µl of LB medium. Bacteria were resuspended and spread onto LB supplemented with 150 $ug.ml^{-1}$ ampicilline (LB supp) plates and incubated at 37° C overnight. Bacterial dilutions of 1/10 and 1/100 were also used.

3.2.7 Selection of recombinant colonies

Recombinant colonies, that is, those containing the plasmid with the insert, were selected using Ampicillin as a selective agent. With sterile toothpicks, the selected colonies were streaked on LB supp plates divided into sectors. These plates were referred to as master plates and each sector was labelled with the corresponding numbering identifying each unique recombinant colony streaked in each sector. The plates were incubated at 37° C overnight. Each toothpick with the remnant of the colony was introduced into the PCR tube containing the corresponding reaction. Each PCR tube was carefully labelled with the colony number that was analyzed. PCR-positive recombinant colonies were grown in LB supp and cryopreserved with 15% glycerol to store them at -70° C.

3.2.8 PCR for analysis of recombinant colonies

The presence of the insert in the selected recombinant colonies was verified by PCR amplification with T7F and T7R primers flanking the cloning site of the pET expression vector. For 10 µl PCR reactions, 0.5 µM of each primer, 1X PCR reaction buffer, 3mM $MgCl_2$, 200 µM of each dNTP and 0.2 units of PlatinumTaq DNA (Invitrogen) were used. The

conditions were: an initial Denaturation step of 95° C for 2 minutes, followed by 30 amplification cycles of 95° C for 45 seconds, 55° C for 45 seconds and 72° C for 60 seconds, final extension was 7 minutes at 72° C and a final cooling step at 4° C. The amplification products were visualised by 1.2 % agarose gel electrophoresis in 1X TBE at 80V for 90 minutes in the presence of a molecular-weight size marker and stained with ethidium bromide. The self-ligating control of the pET vector was included.

3.2.9 PCR to verify that the insert is an VHH

To corroborate that each insert present in the recombinant colonies is the VHH sequence, a PCR was performed using the VHHSfi-(*Nco*I) and ALLVHHR2-*Not* primers that specifically recognise the VHH sequence. Each reaction contained 0.5 µM of each primer, 1X PCR reaction buffer (Invitrogen), 2mM $MgCl_2$, 200 µM of each dNTP and 0.2 units of PlatinumTaq DNA (Invitrogen). The conditions were: an initial denaturation step of 94° C for 2 minutes, followed by 30 amplification cycles of 94° C for 30 seconds, 58° C for 30 seconds and 72° C for 30 seconds, the final extension was 10 minutes at 72° C and a final cooling step at 4° C. The products were visualised by 1.2% agarose gel electrophoresis in 1X TBE buffer at 80V for 90 minutes in the presence of a molecular-weight size marker and stained with ethidium bromide.

3.2.10 Plasmid DNA extraction

The Wizard plus SV Minipreps DNA Purification System Kit (PROMEGA) was used following alkaline lysis protocol of Sambrook *et al*[44], established by the manufacture. The DNA obtained was visualised by 1.5% agarose gel electrophoresis in 1X TBE buffer, 80 V for 90 minutes in the presence of a standard molecular-weight size marker and stained with ethidium bromide. DNA was quantified by Nanodrop at 260 nm and stored at -20° C until use.

3.2.11 Restriction enzyme digestion to verify correct insert location

The location of the insert in the correct position within the pET-22b(+) expression vector was verified by enzymatic cleavages with *Bgl*II and *Not*I. In recombinant plasmids with the insert positioned in the correct reading frame, the fragments obtained after double digestion would be approximately 572 bp and 5266 bp while in those whose insert was not in the correct position the fragments would be 172 bp and 5666 bp. For a 20 ul reaction: 0.5 ug of recombinant plasmid, 1X enzymatic digestion buffer (NEB 3:1), 1 Unit of *Not*I (NEB) and 1 Unit of *Bgl*II (NEB) and 1 Unit of RNAase (Promega) were used.

The products were visualised by 1.5 % agarose gel electrophoresis in 1X TBE buffer at 80 V for 120 minutes in the presence of molecular-weight size marker and stained with ethidium bromide.

3.2.12 Expression induction of recombinant VHH protein

Each selected recombinant colony was inoculated in 2ml of LB supp and grown at 37° C with shaking at 250 rpm overnight. 500ul of the culture were added to 4.5 ml of fresh LB supp and incubated at 37° C with shaking at 250 until the culture reached an OD_{600nm} between 0.5 and 0.530 (1 hour 50 minutes approximately). IPTG was added to a final

concentration of 1mM, after removal of 1 ml of culture for non-induction control. Samples were incubated at 37° C for 3 hours with shaking at 250 rpm and then centrifuged at 7650 g for 15 minutes at 4° C. The pellet was resuspended in 150 ul of 100 mM TrisHCl and the supernatant was filtered with 0.2 uM pore. The pellet and supernatant were stored at -20° C until use. The pellet of induced and non-induced samples were visualised by SDS-PAGE in 1 mm thick vertical gels, 15% polyacrylamide, at 80 V, for 3 hours in the presence of a protein standard and using the BioRad system according to the method described by Laemmli[70]. After electrophoresis, the gel was fixed (40% methanol, 10% acetic acid) and stained (Coomassie blue R-250 0.1%, 40% methanol and 10% acetic acid) and finally decoloured with an isopropanol-acetic acid-water solution (25:20:55).

3.2.13 anti-CD105 ELISA

100 ul of CD105 dissolved in bicarbonate carbonate buffer pH 9.6 at 0.50 ng/ul (50 ng in each well) were added in wells of an ELISA plate and incubated at RT° for 2 hours with gentle shaking for fixation. The contents of the wells were discarded and washed 5 times by adding 200 ul of ELISA wash buffer, shaking gently for 4.5 minutes and then left to stand for 30 seconds each time. 200 ul of ELISA blocking solution were added to each well and incubated at 37° C for 1 hour. Samples were washed 5 times with washing solution as described above. 100 ul of the filtered supernatant of each induced recombinant VHH, non-induced recombinant VHH (negative control) and the positive control (EngVHH17 20ug/ml) was added and incubated for 2 hours at 37° C. Samples were washed 5 times by adding 200 ul wash solution and shaking 30 seconds each time. 100 ul of Penta His HRP conjugated antibody (Qiagen) diluted 1: 2000 in blocking solution were added to each well and incubated at 37° C for one hour. Samples were washed 5 times with wash solution for 30 seconds each time and 100 µl of a fresh TMB substrate solution were added to each well. Samples were incubated for 10 minutes in the dark. The reaction was stopped with 50 µl of 2N sulfuric acid. The absorbance was read in an ELISA plate reader at 450 nm. For the blank control, 50 ng of antigen was fixed but no the recombinant VHH was added, only the blocking solution. All reactions were performed in triplicate. Positive reactions were those with an absorbance 30% higher than the negative control, the unrelated protein Fas2 or 3-fold higher than the blank control.

3.2.14 Expression and purification of recombinant VHH proteins by IMAC

3.2.14.1 Induction of Expression

The selected recombinant colony was seeded in 5ml of LB supp and grown at 37° C with shaking at 250 rpm overnight. 2ml of the culture were inoculated to 48 ml of LB supp and incubated at 37° C with shaking at 250 rpm until the culture reached an OD_{600nm} between 0.5 and 0.530 (approximately 3 hours). IPTG was added to a final concentration of 1mM, after removal of 1 ml of culture for non-induction control. It was incubated at 37° C for 3 hours with shaking at 250 rpm.

3.2.14.2 Purification of recombinant VHH proteins

Purification protocols were adapted from those established by Qiagen according to the product used[71,72]. The proteins were visualised by SDS-PAGE in a 15% gel at 80 V for 3 hours in the presence of a protein molecular-weight size marker. The concentration of proteins was quantified by nanodrop at 280 nm considering a molecular weight of 15000 daltons and an average molar extinction coefficient of 25121 $M^{-1}.cm^{-1}$ as predetermined for its *in silico* sequence (Chapter II, figure 2.10b).

To determine the purification efficiency, the concentration of each of the protein eluates from 5 ml (Spin) and 50 ml (batch and column) of culture was first obtained by nanodrop and then the mg of protein were obtained by summing the amount of protein from all eluates,

3.2.14.2.1 Spin column purification

Bacterial cells were collected from 5 ml of the IPTG-induced culture by centrifugation at 3500g for 5 minutes. The bacterial pellet was resuspended in 630 ul of lysis buffer (NPI-10, Qiagen), 70 ul lysozyme (10 mg/ml) were added and incubated for 30 minutes on ice. The cell lysate was centrifuged at 12000g for 20 minutes and the supernatant was added to a Ni-NTA spin minicolumn that was previously equilibrated with NPI-10 buffer. The spin-minicolumn was centrifuged for 5 minutes at 270g and the flow-through was collected. The spin-minicolumn was washed twice with 600 ul of NPI-20 buffer (Qiagen) by centrifugation at 890 g for 2 minutes. The protein was eluted twice by 300 ul of NPI-500 buffer (Qiagen) and centrifugation at 890g for 2 minutes. Regarding the supernatant of the induced culture, it was added to a Ni-NTA spin-minicolumn and proceeded as described above for the pellet.

3.2.14.2.2 Batch purification

The induced culture was centrifuged at 7650g for 30 minutes at 4° C and the supernatant was separated and filtered with 0.22 uM millipore filter. Then, this was equilibrated to 20mM NaH_2PO_4, 300 mM NaCl and 15 mM Imidazole at pH 7.4 and incubated with 5 ml Ni-NTA Agarose matrix (Qiagen) for 1 hour at 4° C with gentle agitation. The matrix was assembled within a minicolumn and wash buffer was added until obtaining fractions with an OD_{280nm} of 0. Elution buffer was added and 0.5 ml aliquots were collected.

3.2.14.2.3 Column Purification

The supernatant of the induced culture was added to a column containing the Ni-NTA Agarose matrix previously equilibrated with the equilibration buffer. The effluent supernatant was collected and added back to the minicolumn to ensure that no protein remained in the effluent. The minicolumn was washed until obtaining fractions with an $OD_{280\ nm}$ of 0. Protein was eluted by collecting 0.5 ml aliquots.

3.2.15 EngVHH17 positive control

The nucleotide sequence of a CD105-specific VHH, EngVHH17, previously obtained by phage display technique and contained in the recombinant phagemid pHEN2[73], was

amplified by PCR with primers (VHHSfi(*Nco*I) and ALLVHHR2-*Not*), and sub-cloned into pET-22b(+) (Novagen) for its expression in *Escherichia coli* strain BL21. The recombinant VHH protein of approximately 15 kDa was visualised in 15 % SDS-PAGE gel and purified by affinity cromatography in Ni-NTA minicolumns (Qiagen) and quantified in a Fluorometer (QubitTMFluorometer). The specificity of the protein obtained was assessed by anti-CD105 ELISA and the identity of the recombinant protein as an anti-CD105 VHH was verified. EngVHH17 was used to establish the conditions for expression of the recombinant VHH proteins and the conditions for the anti-CD105 ELISA.

3.2.15.1 Determination of recombinant VHH expression conditions

Expression conditions were obtained by testing 3 different concentrations of inducer IPTG (1mM, 1.5 mM and 2mM), 2 concentrations of the antibiotic Ampicillin (100ug/ml and 150 ug/ml) and 3 induction times (3h, 6h and O/N). Recombinant VHH expression was evidenced by comparison of the protein pattern of induced and non-induced cells visualised in a 15% SDS-PAGE gel. The electrophoresis was performed in a BioRad system at 80 V for 1.5 hours and in the presence of a protein molecular-weight size marker.

3.2.15.2 Determination of anti-CD105 ELISA conditions

3.2.15.2.1 Antigen amount

The minimum required amount of CD105 and ELISA conditions were established by testing: 4 differents antigen amounts (50 ng, 100 ng, 200 ng and 300 ng), 3 sera of alpaca previously immunised with CD105 by Gushiken (pre-immunisation, 2nd and 3rd immunisation)[73] and 2 dilutions of Protein A HRP conjugated antibody (Sigma) (1:5000 and 1:10000). CD105 was diluted in carbonate-bicarbonate buffer pH 9.6 to established concentrations and each was added to wells of ELISA plate. ELISA was performed as previously described in anti-CD105 ELISA. The absorbance was measured at $OD_{450\ nm}$ to quantify the response, Positive reactions were considered to be those that showed an absorbance 3-fold or more higher than the blank control[74] (well without fixed antigen).

3.2.15.2.2 Recombinant VHH protein amount

To verify the identity and specificity of the recombinant VHH protein, an ELISA was performed by testing 3 amounts of antigen, 25 ng, 50 ng and 100 ng, 3 dilutions of the VHH eluate, 1:10, 1:50 and 1:100. The absorbance was read on an ELISA plate reader at $OD_{450\ nm}$. For the blank control, antigen was fixed in each well but no recombinant VHH protein eluate was added, only the blocking solution; in addition, an unrelated recombinant VHH protein was used as negative control (CN-). All reactions were performed in duplicate. Positive reactions were those with an absorbance 30% higher than the negative control.

3.3 Results and Discussion

The 19 anti-CD105 VHH sequences were amplified by PCR using the M13R and AHis primers flanking the VHH-specific primer sequences, VHHSfi-(*Nco*I) and ALLVHHR2-*Not* in pHEN2-VHH (Figure 3.3a). The amplification products of approximately 630 bp had a concentration between 150 and 165 ng/ul for each VHH. (Figure 3.3b).

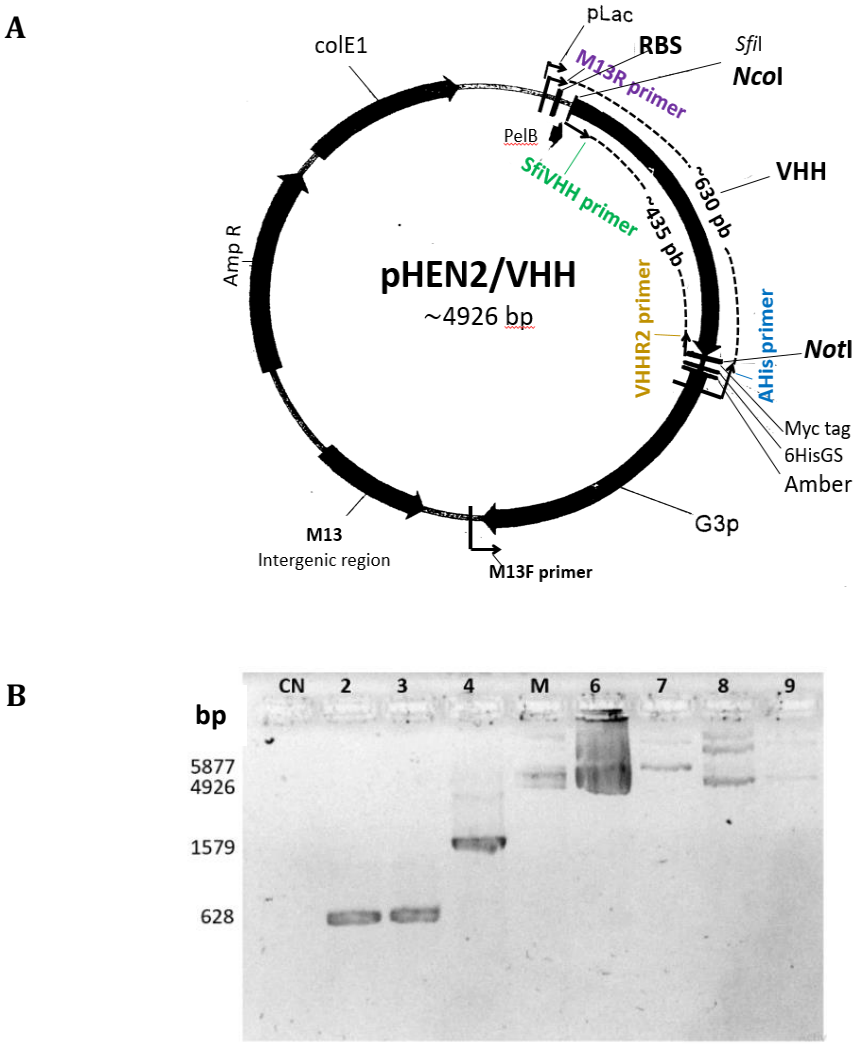


Figure 3.3- PCR amplification of the VHH in pHEN2/VHH with AHis and M13R primers. a) pHEN2/VHH recombinant phagemid vector, b) Electrophoresis gel in 1.2 % agarose, TBE 1X, 80 V, 1 hour. 2 and 3 PCR of VHH, 4 PCR of pHEN2, 6 pHEN2 (4ug/ul), 7 pHEN2 (10 ng/ul), 8 pHEN2-VHH (300 ng/ul), 9 pHEN2-VHH (10 ng/ul), **CN** Negative reaction control, **M** DNA ladder 123 bp (Sigma)

The improved ability and better capacity to obtain and purify recombinant proteins with the requirements for their application in microarrays made the pET-22b(+)(Novagen) system the best option for the expression of VHH proteins. For this, it was necessary to subclone the VHH nucleotide sequences into the expression vector. The subcloning scheme is shown in figure 3.4.

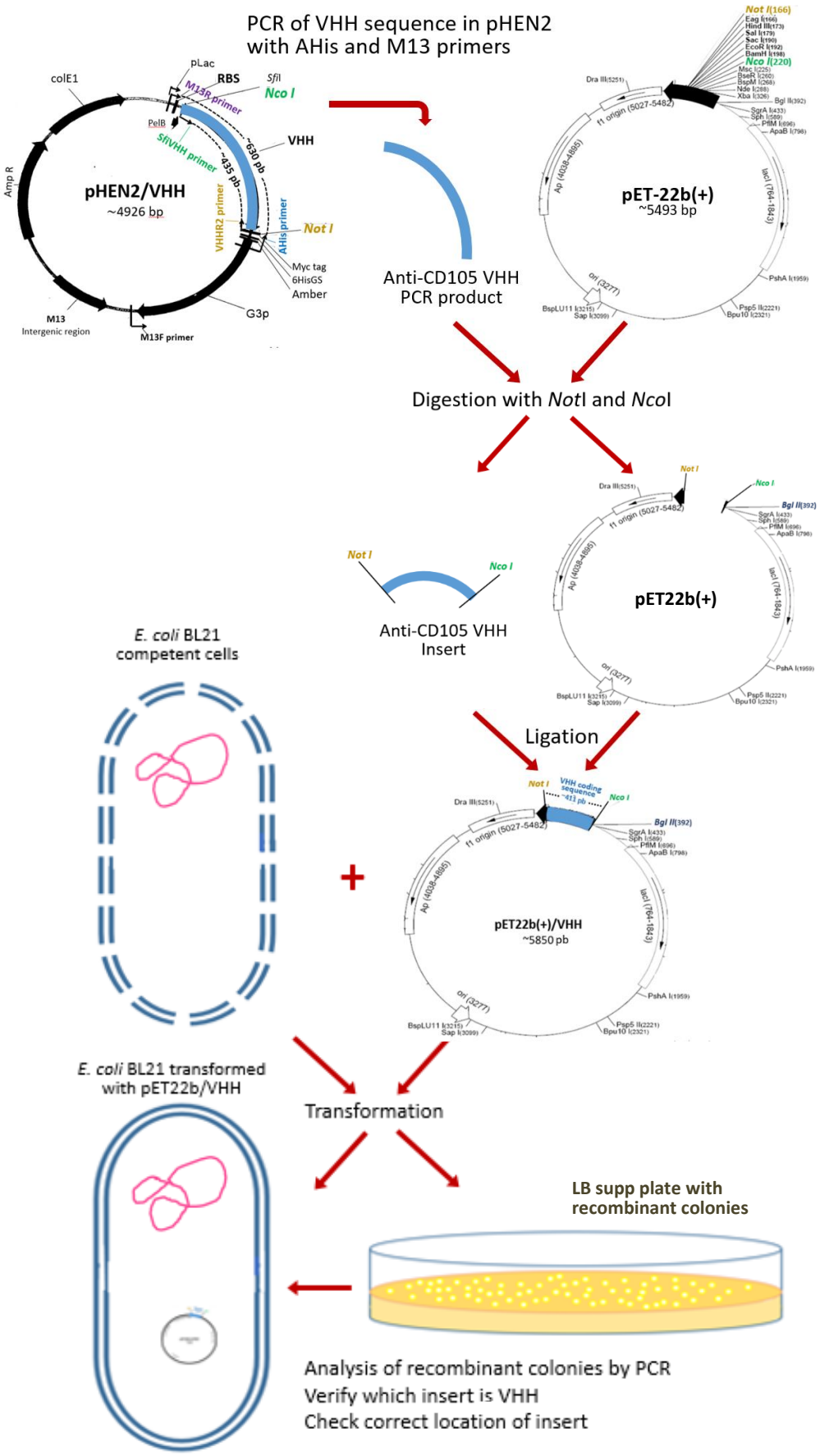


Figure 3.4- Schematic diagram of the subcloning of the VHH nucleotide sequence amplified by PCR with specific primers from pHEN2 into the expression vector pET-22b(+) (Novagen)[39]

For the subcloning to obtain the recombinant VHH proteins, both the amplification products (Figure 3.5a) and the vector (Figure 3.5b) were cleaved. *Nco*I and *Not*I enzymes were used, whose cleavage recognition sites were contained in the specific amplification primers of VHH sequence (between Region Framework 1, RF1 and the hinge region)[75]. The fragments were recovered from the gel (Figure 3.5) and ligated into pET-22b(+) (Novagen) to obtain pET22/VHH plasmids (Figure 3.6).

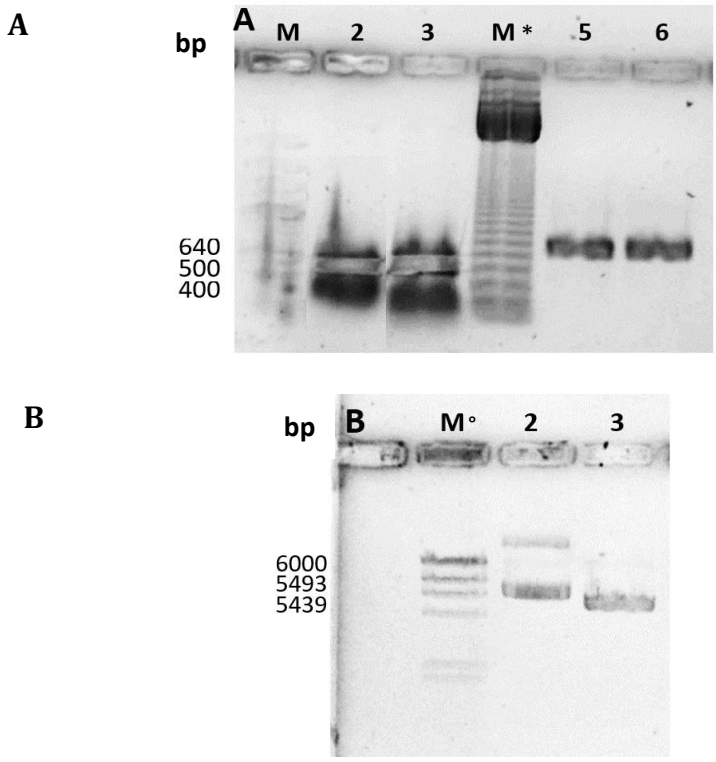


Figure 3.5- Double digestion with *Not*I and *Nco*I. 1.2 % and 0.8 % agarose gel electrophoresis, TBE 1X, 80V, 2 hours. a) Gel A: 2 and 3 VHH/*Not*I+*Nco*I (Fragment of interest recovered from gel), 5 and 6 VHH amplified, b) Gel B: 2 pET-22b(+)(Novagen), 3 pET-22/*Not*I+*Nco*I. **M** GeneRuler 100bp Plus DNA ladder (Thermo Scientific), **M*** DNA ladder 123bp (Sigma), **M°** λ/HindIII.

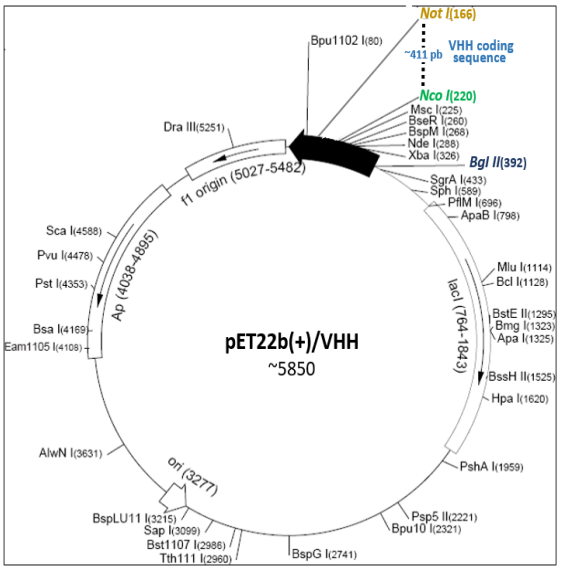


Figure 3.6- Recombinant plasmid vector pET-22b(+)/VHH for protein expression.

Recombinant plasmids were initially introduced into *E. coli* strain DH5α for analysis of such recombinants, verification of insert identity and its position in the correct reading frame for expression. The transformation efficiency was approximately $5.06x10^6$ CFU (colony forming units). Between 8 and 16 recombinant colonies were analysed for each of the 19 VHH. Some recombinants of the VHH 9 and VHH 10 showed variability in the size of the PCR product (Figure 3.7). After a few repetitions with the same result, the presence of a cleavage site for the *Nco*I enzyme within these VHH sequences between positions 103-109 was noticed, so that the fragment obtained was smaller than expected, 102 bp less (34 aa) (Figure 3.8). Therefore, great care was taken to select only recombinants with the appropriate fragment size. On average out of 24 colonies tested, 20 colonies were considered positive for the presence of the recombinant plasmid, that is, presenting the PCR product with the expected size, approximately 700 bp when amplified with primers T7F and T7R (Figure 3.9) and approximately 450 bp when amplified with primers VHHSfi-(*Nco*I) and ALLVHHR2-*Not* (Figure 3.10).

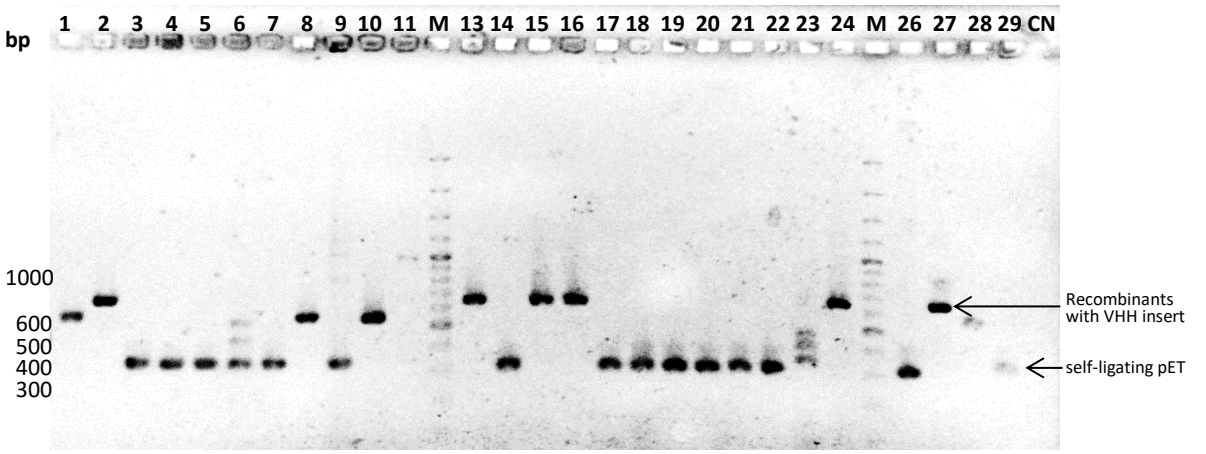


Figure 3.7- PCR with T7 primers for analysis of recombinant colonies. 1.2% agarose gel electrophoresis, 1X TBE, 80 V, 1.5 hour. 1-11, 13-24 Recombinant VHH, 26 and 29 pET self-ligation control, 27. Positive control, 28. EngVHH17, **M** GeneRuler 100 bp Plus DNA ladder (ThermoScientific), **CN** Negative reaction control.

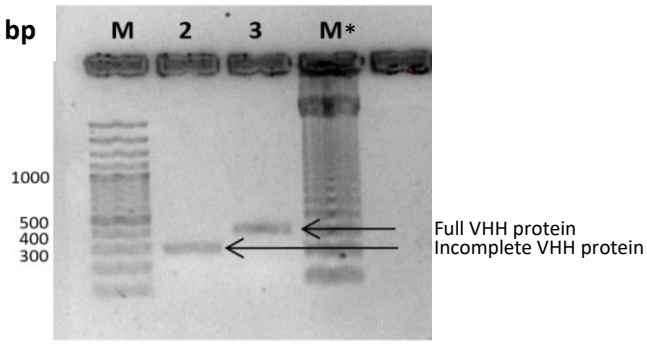


Figure 3.8- Variability in the size of the PCR products of recombinant VHH colonies. 1.2% agarose gel electrophoresis, 1X TBE, 80 V, 1.5 hour, 2 VHH/*Not*I + *Nco*I not complete, 3 VHH/*Not*I + *Nco*I complete. **M** GeneRuler100 bp Plus DNA ladder (ThermoScientific), **M*** DNA ladder 123bp (Sigma).

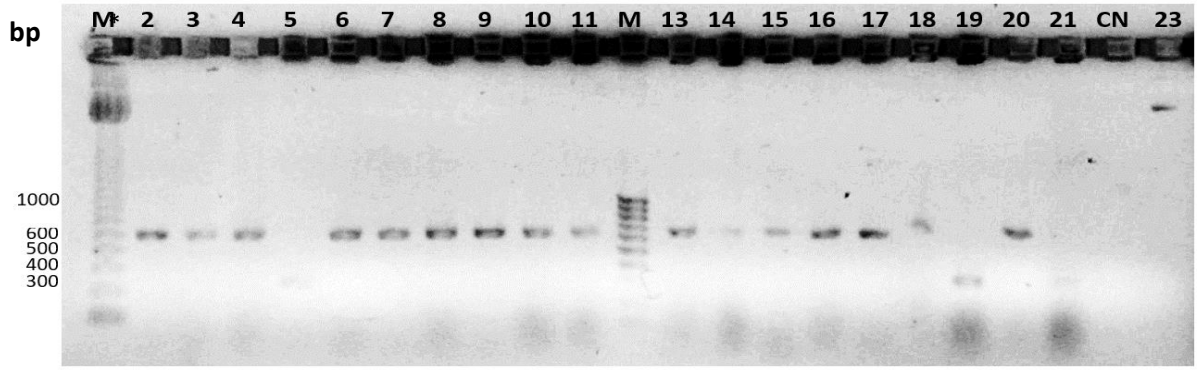


Figure 3.9- PCR with T7 primers for analysis of recombinant VHH colonies. 1.2% agarose gel electrophoresis, 1X TBE, 80 V, 1.5 h, 2-11, 13-18 Recombinant VHH, 19 and 21 pET22 self-ligation control, 20 Recombinant EngVHH17, 23 pET22. **M** GeneRuler 100 bp Plus DNA ladder (ThermoScientific), **M*** DNA ladder 123 bp (Sigma), **CN** Negative reaction control.

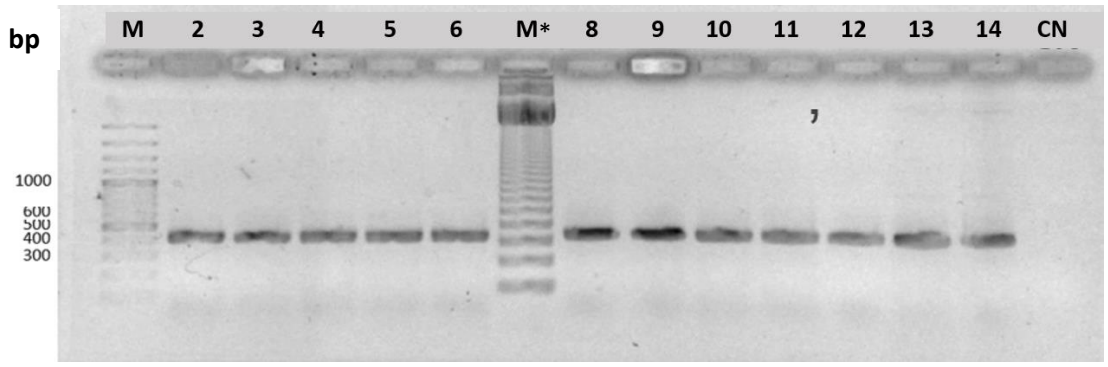


Figure 3.10- PCR with primers VHHSfi- (*Nco*I) and ALLVHHR2-*Not* to verify that the insert is a VHH. 1.2% agarose gel electrophoresis, 1X TBE, 80 V, 1.5 hours, 2-6, 8-14 Recombinant VHH, **M** GeneRuler 100 bp Plus DNA ladder (ThermoScientific), **M*** DNA ladder 123 bp (Sigma), **CN** Negative reaction control.

Plasmid DNA was extracted from each colony selected as positive, that is, the recombinant plasmids pET/VHH of approximately 5850 bp (Figure 3.11, 3.6) and they were digested with the enzymes *Bgl*II and *Not*I to confirm that the insert (of each plasmid recombinant selected) was positioned in the correct reading frame. The fragments obtained would be 583 bp and 5267 bp approximately if the insert was correctly positioned while in those whose insert was not in the correct position (inverted) the fragments would be 172 bp and 5678 bp (Figure 3.12)

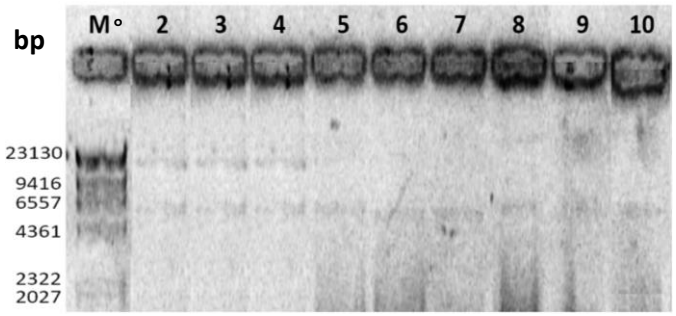


Figure 3.11- Plasmid DNA from recombinant colonies. 1% agarose gel electrophoresis, 1X TBE, 80 V, 1.5 hour, 2-9 Recombinant pET/VHH plasmids. **M°** λ/*Hind*III.

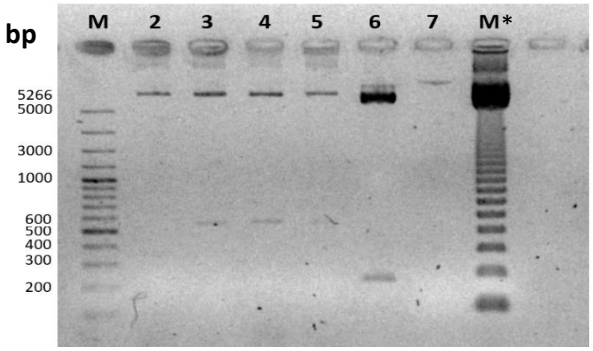


Figure 3.12- Digestion with restriction enzymes *Bgl*II and *Not*I to verify the correct location of the insert. 2-5 Digestion of recombinant VHH plasmids, 6 Digestion of pET-22b(+) (Novagen), 7 pET-22b(+) (Novagen). **M** GeneRuler100 bp Plus DNA ladder (ThermoScientific), **M*** DNA ladder 123bp (Sigma).

The *E. coli* BL21 strain was transformed with the recombinant plasmids selected for protein expression induction, with a transformation efficiency of 9.6 x10^6. Five recombinant clones for each VHH were selected to choose those with the highest protein expression capacity.

The conditions for the induction of VHH expression were previously determined with the positive control EngVHH17. For this purpose, different concentrations of inducer, antibiotic and induction times were tested and these conditions were set as 1mM IPTG, 150 ug/ml ampicillin and 3 hours of induction. 3 hours because a longer induction time did not mean an increase in protein expression so it was unnecessary to take more time taking into consideration the half-life of the protein, which even being longer than 10 hours, it was preferred to avoid any risk of degradation and maintain as much functional protein as possible. Increasing the antibiotic concentration from 100 ug/ml to 150 ug/ml ensured the permanence of the recombinant plasmid and prevented the growth of satellites competing with the bacterial population producing the protein and the consequent decrease in production.
Initially, 3 ml of culture were used to identify and select the colonies with the highest expression efficiency and then scale up protein production to a larger volume depending on the requirement. Expression of recombinant VHHs was verified by comparison of the protein patterns obtained from the pellets (Figures 3.13 a, b, c) and supernatants (Figure 3.13 d) of the induced and non-induced cells visualised on a 15% SDS-PAGE gel.

On average out of 10 selected clones that were induced, 7 showed the protein with the expected size (approximately 15 kDa). Generally, more than one clone was obtained containing each of the 19 different VHH inserts correctly placed in the vector for expression (Figure 3.14).

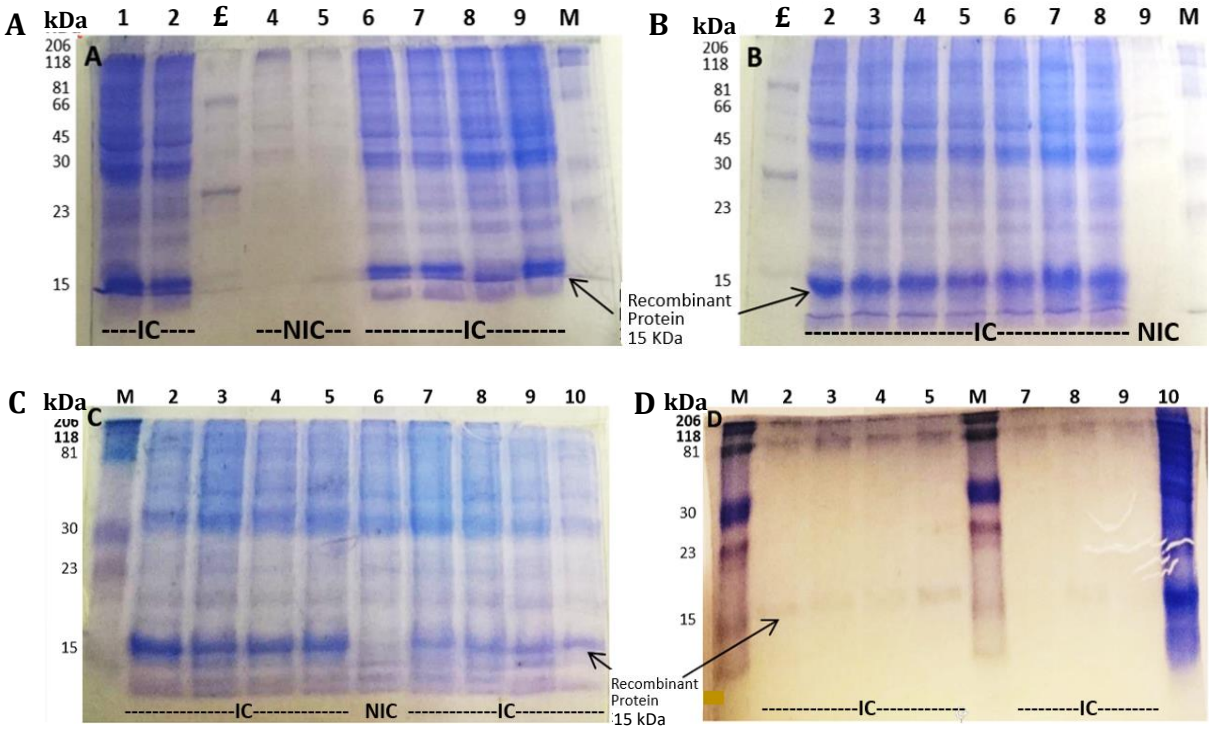
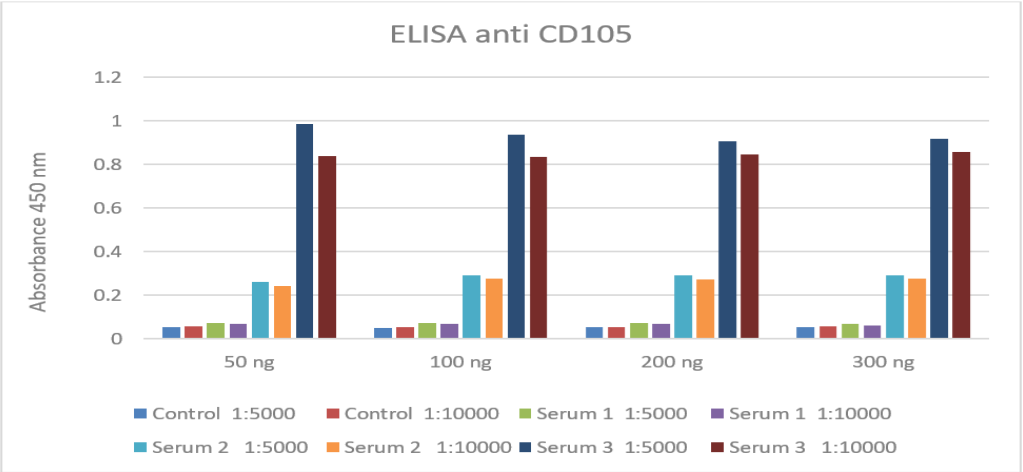


Figure 3.13- Expression induction of recombinant VHH protein from selected clones. SDS-PAGE 15% polyacrylamide, 60 V for 3.5 hours. a) Gel A Pellet: 1 1-II, 2 2-I, 4 1-II, 5 2-I, 6 3-II, 7 5-II, 8 9-II, 9 4-IV, b) Gel B Pellet: 2 1-II, 3 6-II, 4 7-II, 5 8-II, 6 11-V, 7 12-I, 8 13-I, 9 4-IV, c) Gel C Pellet: 2 19-II, 3 18-V, 4 17-III, 5 16-IV, 6 2-I, 7 15-II, 8 14-II, 9 10-III, 10 EngVHH17 VI, d) Gel D Supernatant: 2 EngVHH17 VI, 3 1-II, 4 2-I, 5 19-II, 7 18-V, 8 17-III, 9 6-IV, 10 EngVHH17 VI, (pellet). **IC** Induced Clone, **NIC** Non Induced Clone. **M** KaleidoscopeTM Prestained Standards (BioRad), **£** Protein Standard Albumin 66kDa, Ovalbumin 45kDa, Glyceraldehyde 3P 36kDa, Carbonic Anhydrase 29kDa, a-lacto albumin 14.2kDa.

TG1/pHEN2-VHH anti CD105 clones		Recombinant clones for each VHH	BL21/pET22-VHH anti CD105 clones	
1	D2-I	1	**1**-II	**1**
2	H6-I	1	**2**-I	**2**
3	F6-II	1	**3**-II	**3**
4	H8-II	4	**4**-II	**5**
			4-V	**30**
			4-VI	**31**
			4-VII	**32**
5	A1-II	1	**5**-II	**7**
6	D4-II	1	**6**-II	**8**
7	A4-II	1	**7**-II	**9**
8	A3-II	2	**8**-I	**10**
			8-II	**11**
9	D2-III	1	**9**-II	**12**
10	B4-III	2	**10**-III	**33**
			10-IV	**34**
11	H2-III	1	**11**-V	**19**
12	C2-III	1	**12**-I	**14**
13	E1-III	1	**13**-I	**15**
14	F3-III	1	**14**-II	**16**
15	A1-IV	1	**15**-II	**17**
16	E11-IV	2	**16**-IV	**20**
			16-III	**21**
17	E1-IV	2	**17**-III	**22**
			17-IV	**23**
18	D6-IV	2	**18**-IV	**25**
			18-V	**26**
19	E7-IV	1	**19**-II	**18**

Figure 3.14- Number of recombinant BL21colonies per sequence VHH.

To confirm that the recombinant proteins obtained were VHH and to determine the VHH specificity for its antigen, human CD105, an ELISA was performed. The conditions for this anti-CD105 ELISA were previously established using first the immunised alpaca serum and then the positive control EngVHH17 with different amounts of the antigen. For this, it was tried to decrease 2/3 (200 ng), 1/3 (100 ng), and 1/6 (50 ng) the antigen amount used initially (300 ng). No significant difference was evident between the different amounts of antigen with respect to the initial amount (Figure 3.15), so it was determined that using 50 ng of antigen would be sufficient to obtain a response similar to if 300 ng were used. The dilution of the conjugated antibody was also evaluated and 1:5000 was chosen. Regarding the serum response, pre-immunisation, second and third immunisation, it was as expected according to the immunisation criteria.



Antigen amount	Blank Control		Serum 1		Serum 2		Serum 3	
	1:5000	1:10000	1:5000	1:10000	1:5000	1:10000	1:5000	1:10000
50 ng	0.052	0.055	0.073	0.067	0.259	0.242	0.985	0.836
100 ng	0.05	0.054	0.072	0.069	0.29	0.274	0.934	0.833
200 ng	0.053	0.054	0.07	0.066	0.29	0.271	0.904	0.844
300 ng	0.051	0.056	0.067	0.059	0.29	0.276	0.915	0.856

Figure 3.15- Determination of the conditions for ELISA anti-CD105 with serum from immunised alpacas, not with recombinant VHH proteins

To determine the amount of VHH providing the best signal, the EngVHH17 protein was tested. This control VHH protein was expressed and purified in Ni-NTA Agarose columns (Qiagen, 2012) from the supernatant of two selected recombinant clones (I and VI). Approximately 4.5 mg of protein (5 ml at 0.885 mg/ml) were obtained from clone I and 2.5 mg of protein (5 ml at 0.5 mg/ml) from clone VI as quantified by Fluorometer (Qubit™Fluorometer) (Figure 3.16). Three different dilutions of the EngVHH17 eluate from clone I were used (because it had a higher concentration than the eluate from clone VI, not because of higher affinity), with the 1:10 dilution showing the best signal (almost 20 times higher than blank control). It was tested using 3 different amounts of antigen (25 ng, 50 ng and 100 ng), half and double the amount determined in the previous assay (50 ng) to corroborate if this concentration of antigen is sufficient, or even if it could be lower, since a purified VHH would be used and not the total immunized serum. However, although the signal at 25 ng was good (almost 12 times higher than blank control), it was notably lower than that obtained at 50 ng and 100 ng, which were almost twice as high (Figure 3.17).

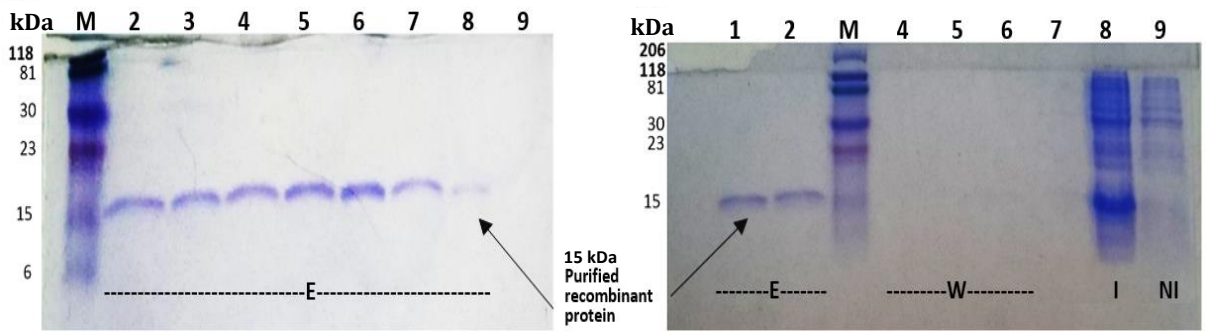
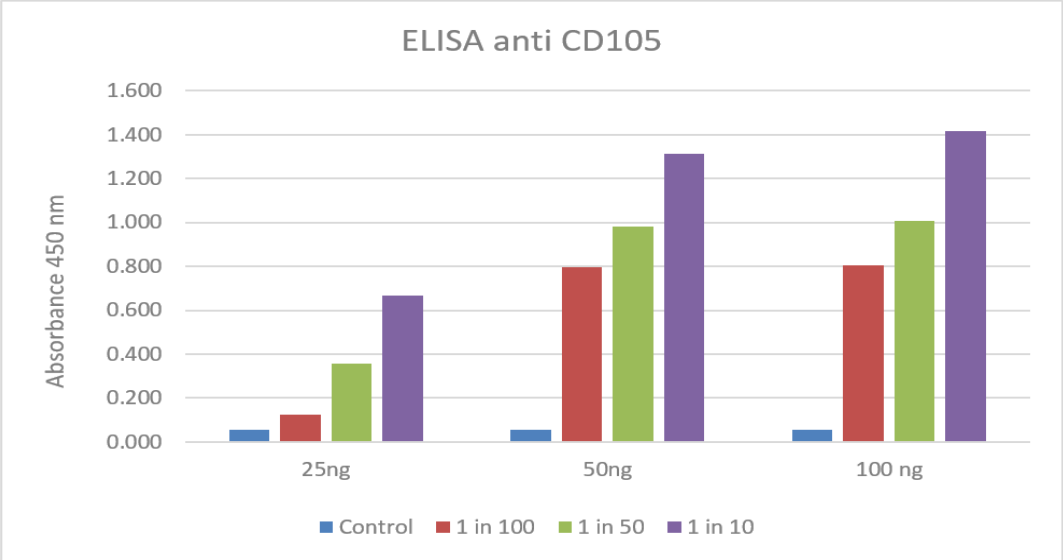


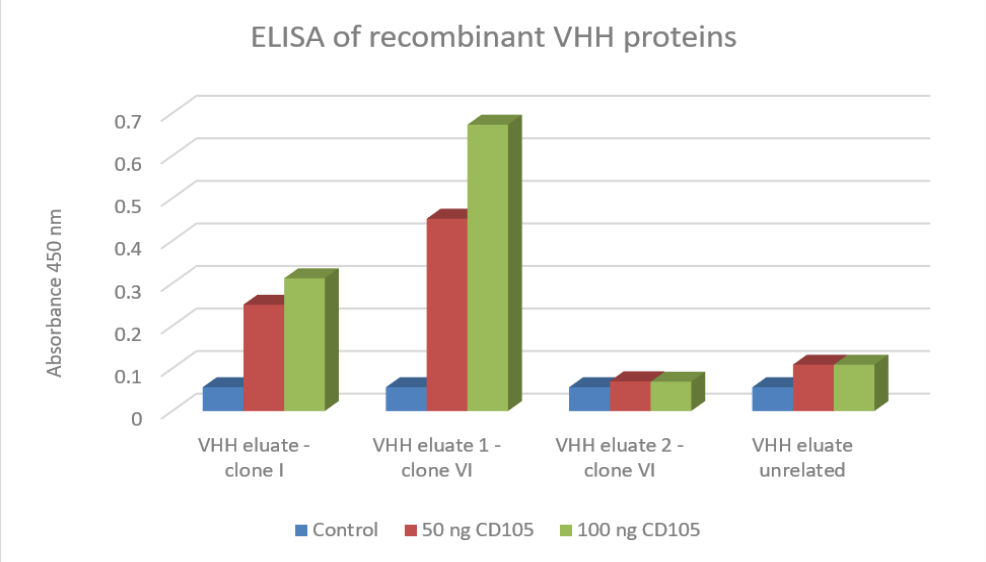
Figure 3.16- Purification of recombinant EngVHH17 protein. SDS-PAGE 15% polyacrylamide, 60 V for 3.5 hours. **M** Kaleidoscope Prestained Standards (BioRad), **E** EngVHH17 protein eluates, **W** wash, **I** induced colony pellet, **NI** non-induced colony pellet.



CD105	Blank Control	1 in 100	1 in 50	1 in 10
25ng	0.054	0.126	0.357	0.666
50ng	0.057	0.798	0.981	1.313
100 ng	0.056	0.806	1.009	1.416

Figure 3.17 Determination of the antigen amount and the recombinant VHH amount for ELISA. Control without VHH

Therefore, to test the elutions of recombinant proteins obtained from different clones, these two amounts of antigen (50 ng and 100 ng) were still used in the ELISA. This anti-CD105 ELISA used one eluate of the recombinant VHH protein from clone I, two eluates from clone IV and one eluate of an unrelated recombinant VHH protein. At both antigen concentrations, the clone I eluate showed an absorbance approximately 5 times higher than the blank control, while clone VI eluates, only the first one showed a positive reaction with approximately 10 times higher absorbance than the blank control. The unrelated recombinant VHH protein eluate was a negative control (Figure 3.18). This not only verified that the recombinant protein obtained was a VHH, and an anti-CD105 VHH, but also that the anti-CD105 ELISA conditions for recombinant VHH proteins were standardized.



CD105	Blank Control	Eluate clone I	Eluate 1 clone IV	Eluate 2 clone IV	Unrelated eluate (CN-)
50 ng	0.056	0.2495	0.4515	0.0705	0.1085
100 ng	0.056	0.3115	0.6725	0.069	0.1095

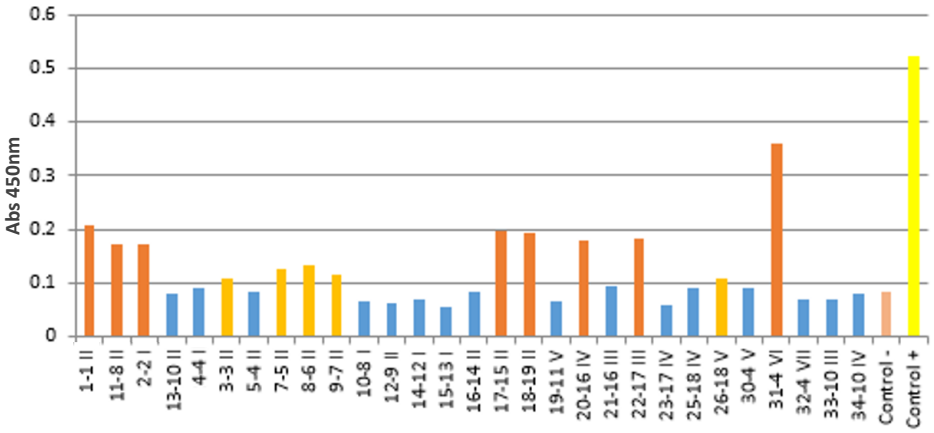
Figure 3.18- ELISA assay to verify the identity and specificity of recombinant EngVHH17 protein.

The anti-CD105 ELISA of the recombinant clones containing the 19 VHH inserts (Figures 3.14 and 3.19) were performed using the supernatant of each bacterial culture. This is in knowledge that the pET22 system secretes the expressed protein into the periplasm and then into the external medium in sufficient quantity for detection by ELISA (between 0.01 and 0.09 mg/ml as previously determined with the EngVHH17 control). The expressed proteins were in the medium at an average concentration of 0.05 mg/ml.

Thirteen of the nineteen recombinant VHH proteins maintained their ability to recognise CD105 (Figure 3.19). Of these 13 VHH proteins (Figure 3.20), 8 of them had a strong signal. Clone 31 (VHH4) had the strongest signal, more than 4 times the negative control, followed by clones 1 (VHH1), 17 (VHH 15), 18 (VHH 19), 22 (VHH 17), 20 (VHH 16), 11(VHH 8) and 2 (VHH 2), with almost 3 times the signal.

Six recombinant VHH proteins did not have sufficient signal to be considered as positive, with values very close to or lower than the negative control, the VHH-protein not related to CD105, Fas2 protein (Figure 3.19). The fact that these proteins lost their ability to recognise CD105 can be explained because they were obtained after a process of subcloning of their sequence that involved, among other things, cleaving with restriction enzymes and ligation. In addition, the control and induction of protein expression were carried out in a prokaryotic system, which meant more processing in the synthesis and therefore, more possibilities for change or error at critical points for their recognition function.

	1	2	3	4	5	6	7	8	9	10	11	12	Abs
A	0,242	0,172	0,093	0,091	0,194	0,191	0,08	0,1	0	0	0	0	450
B	0,142	0,102	0,081	0,076	0,066	0,066	0,344	0,378	0	0	0	0	450
C	0,205	0,138	0,068	0,064	0,122	0,117	0,066	0,077	0	0	0	0	450
D	0,1	0,06	0,067	0,061	0,057	0,055	0,067	0,07	0	0	0	0	450
E	0,117	0,063	0,074	0,069	0,17	0,155	0,07	0,089	0	0	0	0	450
F	0,097	0,072	0,056	0,056	0,063	0,058	0,074	0,095	0	0	0	0	450
G	0,091	0,079	0,072	0,094	0,089	0,091	0,559	0,085	0	0	0	0	450
H	0,102	0,088	0,173	0,123	0,123	0,094	0,49	0,107	0	0	0	0	450



VHH clone		Abs_{450nm}
1-1 II	VHH 1 Seq 1	0.207
11-8 II	VHH 11 Seq 8	0.172
2-2 I	VHH 2 Seq 2	0.172
13-10 II	VHH 13 Seq 10	0.080
4-4 I	VHH 4 Seq 4	0.090
3-3 II	VHH 3 Seq 3	0.110
5-4 II	VHH 5 Seq 4	0.085
7-5 II	VHH 7 Seq 5	0.125
8-6 II	VHH 8 Seq 6	0.132
9-7 II	VHH 9 Seq 7	0.115
10-8 I	VHH 10 Seq 8	0.066
12-9 II	VHH 12 Seq 9	0.064
14-12 I	VHH 14 Seq 12	0.072
15-13 I	VHH 15 Seq 13	0.056
16-14 II	VHH 16 Seq 14	0.083
17-15 II	VHH 17 Seq 15	0.198
18-19 II	VHH 18 Seq 19	0.193
19-11 V	VHH 19 Seq 11	0.066
20-16 IV	VHH 20 Seq 16	0.180
21-16 III	VHH 22 Seq 16	0.096
22-17 III	VHH 22 Seq 17	0.183
23-17 IV	VHH 23 Seq 17	0.061
25-18 IV	VHH 25 Seq 18	0.090
26-18 V	VHH 26 Seq 18	0.109
30-4 V	VHH 30 Seq 4	0.090
31-4 VI	VHH 31 Seq 4	0.361
32-4 VII	VHH 32 Seq 4	0.072
33-10 III	VHH 33 Seq 10	0.069
34-10 IV	VHH 32 Seq 10	0.080
Control -	VHH Fas 2 (CN)	0.084
Control +	EngVHH17 (CP)	0.525
Blank Control (no CD105)		0.036

Figure 3.19- ELISA for anti-CD105 VHH recognition. In dark orange: VHH proteins recognising CD105 with strong signal, in orange: VHH proteins recognising CD105, in pink: negative control, unrelated protein Fas2, in yellow: positive control EngVHH17, in light blue: VHHs not specifically recognising CD105.

VHH en pET		VHH en pHEN2	
1	1-II	**1**	D2-I
2	2-I	**2**	H6-I
3	3-II	**3**	F6-II
31	4-VI	**4**	H8-II
7	5-II	**5**	A1-II
8	6-II	**6**	D4-II
9	7-II	**7**	A4-II
11	8-II	**8**	A3-II
17	15-II	**15**	A1-IV
20	16-IV	**16**	E11-IV
22	17-III	**17**	E1-IV
26	18-V	**18**	D6-IV
18	19-II	**19**	E7-IV

Figure 3.20-The 13 recombinant anti-CD105 VHH proteins. In pink: nomenclature of *E. coli* BL21 colonies containing VHH sequences in pET22 (expression), in light blue, nomenclature of *E. coli* TG1 colonies containing VHH sequences in pHEN2 (sequencing)

In the case of clones 12 (VHH9), 33 and 34 (VHH10), their sequences had an extra recognition site for the *Nco*I enzyme inside, which could result in an incomplete protein, unable to properly recognise its antigen due to the absence of the complete recognition site or of the correct folding. The sequences of clones 11 (VHH 19), 13 (VHH 15) and 14 (VHH 16) encoded unstable proteins (Chapter II, figure 2.8b) so, it is possible that the amount of protein may have been too low, or mostly without the correct folding for interaction with the antigen. On the other hand, the sequence of clone 12 (VHH 14) differed by one amino acid from the sequence of clone 18 (VHH 26) and by two amino acids from the sequence of clone 4 (VHH31) (chapter II, figure 2.5a) so, it would have been expected that, like them, it encoded a VHH protein able to recognise its antigen with high specificity. However, according to anti-CD105 ELISA, this was not the case, showing a very low signal value, even lower than the negative control. It seems that it is the substitution at position 50 (A50G) that somehow influences the recognition of the antigen, possibly by loss of adequate folding due to increased hydrophobicity of the region.

Thirteen recombinant proteins were expressed and purified by IMAC with Ni-NTA Agarose. The Ni-NTA ligand covalently bound to a cross-linked agarose matrix allowed the selective purification of the recombinant VHH proteins with 6X His-tag, due to the coordination of the imidazole groups of the His residues with the Ni^{2+} ion (that replaced the bound water molecules).

VHH proteins were recovered by elution with imidazole, thanks to their ability to coordinate with the Ni^{2+} ion and displace the bound protein. For VHH purification, the culture volume of the 13 selected recombinant clones was scaled up, first from 3 ml to 5-10 ml and then to 50 - 100 ml. The Qiagen Ni-NTA Spin kit, which employs Ni-NTA mini-column centrifugation for protein purification, was used initially. Highly pure proteins were obtained (Figure 3.21) but the small volume of sample that could be used (5 ml of the culture) did not allow obtaining the amount of protein needed for the next application in the biochips. The amount of protein obtained was measured by Nanodrop at 280 nm, obtaining an average of 0.092 mg of protein from the pellet and 0.09 mg from the supernatant per 5 ml of culture.

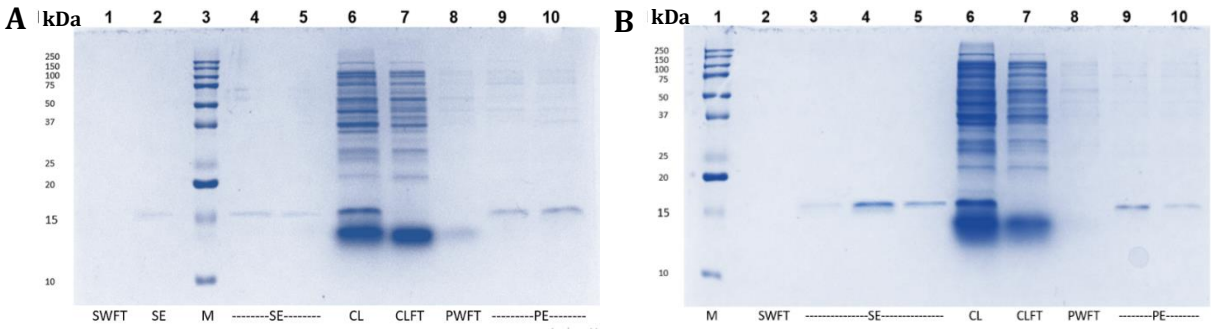


Figure 3.21- Purification of recombinant VHH proteins under native conditions with Ni-NTA Spin Column (Qiagen, 2016), from cell lysate and supernatant. SDS-PAGE 15% polyacrylamide, 80 V for 1.5 hours. a) Colony 4-VI VHH 31 (VHH 4), b) Colony 2-1 VHH 2 (VHH 2). M Kaleidoscope Prestained Standards (BioRad); **SWFT** Supernatant Wash Flow through; **SE** Protein Eluate from Supernatant; **CL** Cell lysate; **PWFT** Pellet Wash Flow through; **PE** Protein Eluate from Pellet.

The conventional protocol established by Qiagen for protein purification was modified and the culture supernatant was used instead of the bacterial pellet as the protein is preferentially excreted to the medium. The purification scheme is shown in figure 3.22. The VHH proteins remain stable for more than 24 hours at RT°; however, purification was performed immediately after induction of expression. 50 ml of culture were used for batch purification and another 50 ml for column purification, one followed after the other, using the same column.

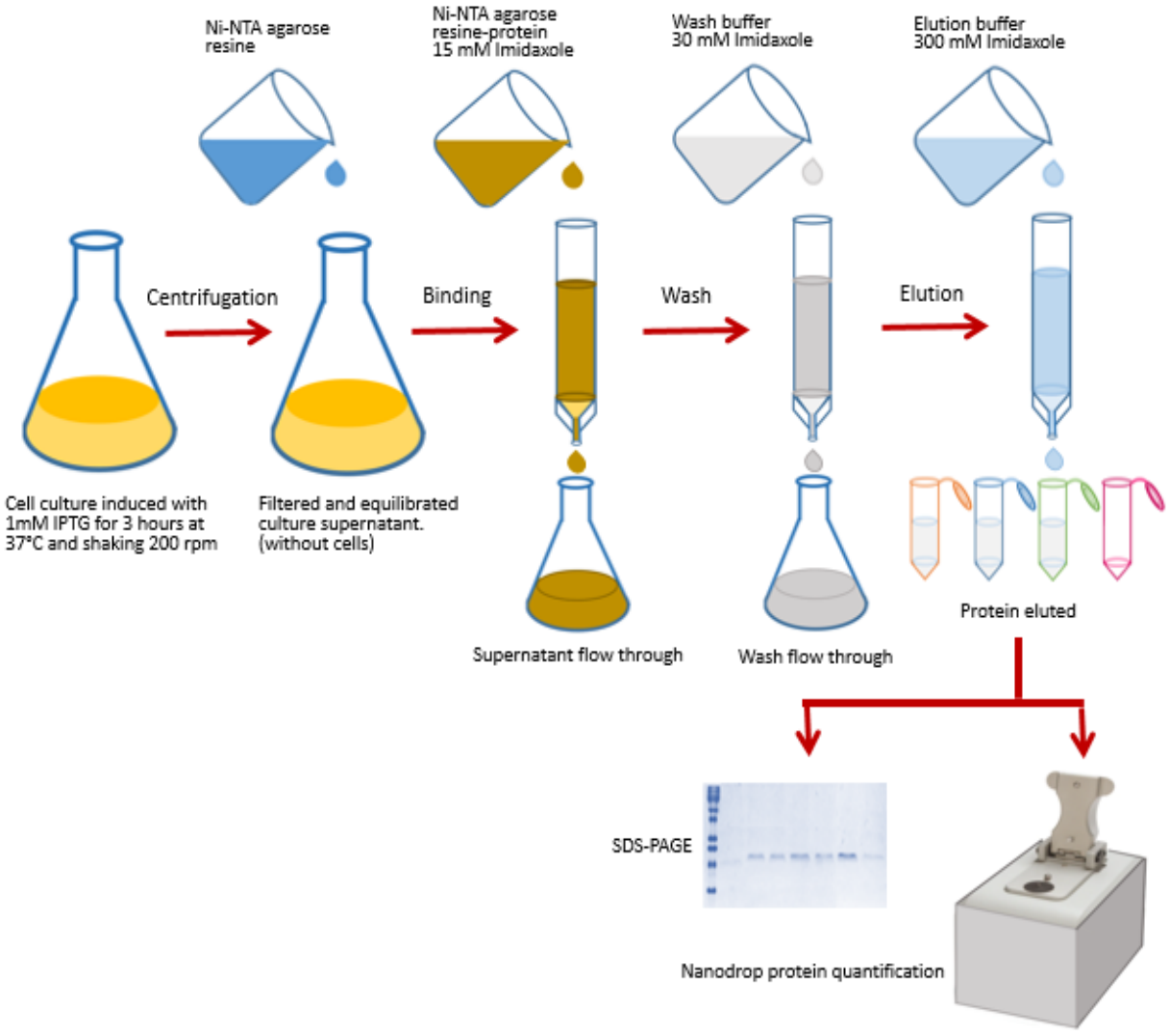


Figure 3.22- Schematic diagram of purification of recombinant VHH proteins by immobilized metal affinity chromatography (IMAC) with batch procedure.

To verify the size and purity of the purified proteins in batch (Figure 3.23) and in column (Figure 3.24), they were visualised in gels at 15% SDS-PAGE as a single band of approximately 15kDa.

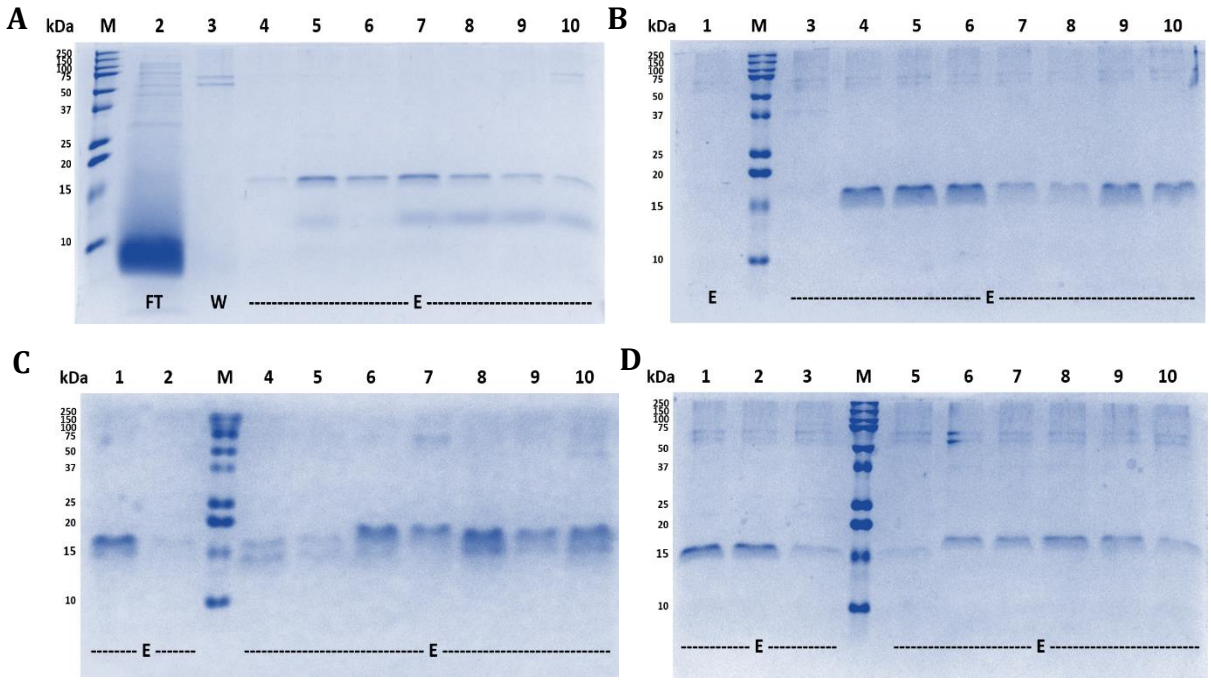


Figure 3.23- Purification of recombinant VHH proteins under native conditions with Ni-NTA agarose in batch procedure from supernatant. SDS-PAGE 15% polyacrylamide, 80 V for 2.5 hours. a) Gel A: 4-7 EngVHH17 (CP), 8-10 VHH 11, b) Gel B: 1, 3 and 4 VHH 1, 5 and 6 VHH 2, 7 and 8 VHH 3, 9 and 10 VHH 31, c) Gel C: 1 and 2 VHH18, 4 and 5 VHH 8, 6 and 7 VHH 7, 8-10 VHH 17, d) Gel D: 1-3 and 5 VHH 9, 3 VHH 22, 7 and 8 VHH 20, 9 and 10 VHH 26, **M** Kaleidoscope Prestained Standards (BioRad), **E** Protein Eluates, **FT** Flow through, **W** Wash.

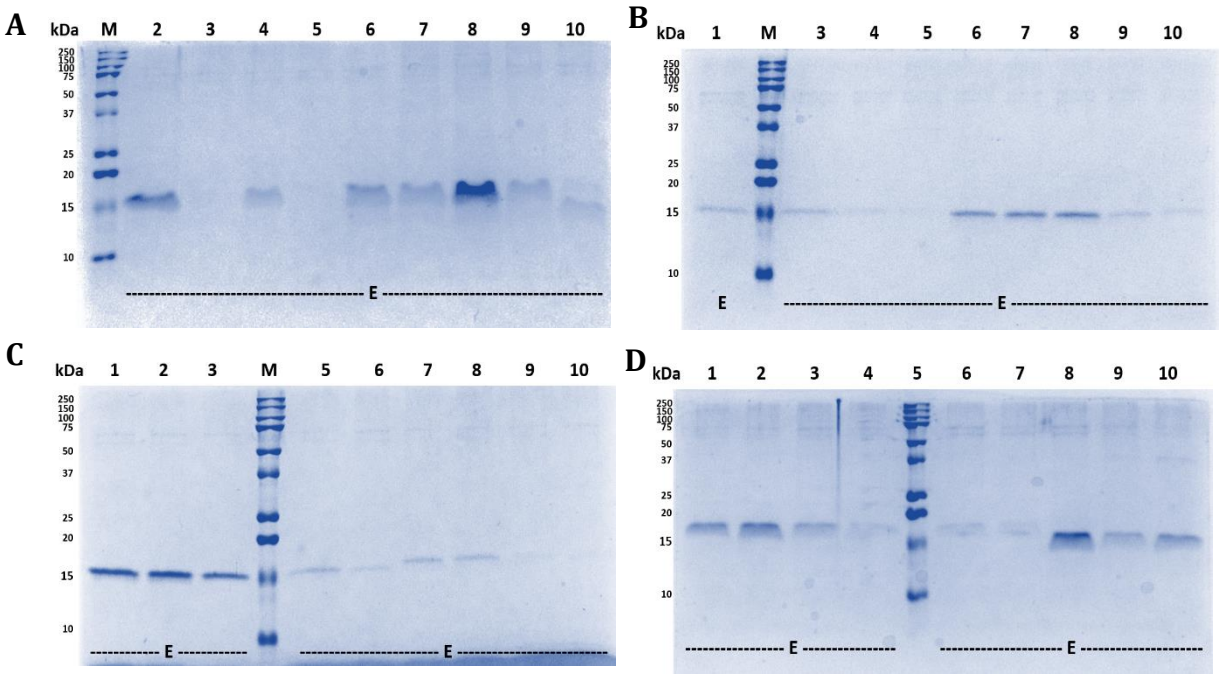


Figure 3.24 - Purification of recombinant VHH proteins under native conditions with Ni-NTA agarose in column procedure from supernatant. SDS-PAGE 15% polyacrylamide, 80 V for 2.5 hours. a) Gel A: 2 EngVHH17 (CP), 3 and 4 VHH 1, 5 and 6 VHH 2, 7 and 8 VHH 31, 9 and 10 VHH 3, b) Gel B: 1, 3 and 4 VHH 8, 5 -7 VHH 8 , 8-10 VHH 18, c) Gel C: 1-2 VHH 17, 3, 5 and 6 VHH 22, 7-9 VHH 20, d) Gel D: 1 and 2 VHH 26, 3, 4 and 6 VHH 11, 7-9 VHH 9, 10 EngVHH17 (CP), **M** Kaleidoscope Prestained Standards (BioRad), **E** Protein eluates,

The average protein concentration obtained by batch purification was 0.225 mg.ml^{-1} and by column purification 0.22 mg.ml^{-1}, as there were no significant difference, both methods were used consecutively to simplify the process and reuse the Ni-NTA Agarose matrix VHH. Approximately 4.9 mg of protein was obtained per 100 ml of culture. A total amount of more than 5 mg of each VHH was purified (Figure 3.25), sufficient for functionalisation and application assays on biochips for the recognition of CD105 on human cells by SPRi.

VHH		mg
1	(VHH 1)	5.316
2	(VHH 2)	5.176
3	(VHH 3)	5.472
31	(VHH 4)	5.284
7	(VHH 5)	5.341
8	(VHH 6)	5.563
9	(VHH 7)	5.501
11	(VHH 8)	5.218
17	(VHH 15)	5.430
20	(VHH 16)	5.019
22	(VHH 17)	5.511
26	(VHH 18)	5.388
18	(VHH 19)	5.168

Figure 3.25 - Amount in mg of the 13 anti-CD105 VHH proteins obtained by purification in Ni-NTA agarose matrix under native conditions from culture supernatant of selected recombinant bacterial colonies.

Conclusions

Nineteen anti-CD105 VHH sequences were subcloned by PCR with specific primers and restriction enzymes from the phagemid vector pHEN2 into the plasmid vector pET-22b(+) in the correct reading frame for protein expression. Then, out of 19 recombinant proteins expressed in *E. coli*, 13 were positive for anti-CD105 ELISA, thereby verifying that the expressed proteins were recombinant anti-CD105 VHH. The conditions for the anti-CD105 ELISA were previously determined using the recombinant VHH protein EngVHH17 and it was established that the minimum amount of antigen necessary for a good detection signal was 50 ng, the dilution of the conjugated antibody 1:5000 and the dilution of the eluate 1:10. These 13 anti-CD105 VHH proteins were then expressed and purified. For this, the expression conditions were scaled up from a culture volume of 3 ml to 5-10 ml and then to 50-100 ml and set as 3 hours of induction with 1mM IPTG in bacterial cultures of 3 hours of growth (OD_{600nm} of 0.5) in LB supp at 37° C and shaking at 200 rpm. While for their purification under native conditions, the IMAC method in Ni-NTA Agarose matrix proposed by the manufacturer Qiagen was adapted to obtain them from the culture medium and not from the cells. Furthermore, the appropriate amount of Imidazole was determined for each step, establishing that 15 mM was required for equilibration, 25 mM for washing and 300 mM for elution. The average protein obtained was 0.225 mg/ml by batch purification and 0.220 mg/ml by column purification. Thus, purified anti-CD105 VHH proteins were obtained in sufficient quantity and quality, that is, on average 4.9 mg of protein per 100 ml of culture, for their functionalisation and application in protein microarrays for characterisation in SPRi assays.

References

1) Ritchie, C. (2021) Protein purification. *Materials and Methods 2*, 134. Last modified: 2021-04-30; original version: 2012-11-17.

2) Rosano, G. L., Morales, E. S. & Ceccarelli, E. A. (2019) New tools for recombinant protein production in *Escherichia coli*: A 5-year update. *Protein Science 28*, 1412–1422.

3) Gileadi, O. (2017) Recombinant protein expression in *E. coli*: a historical perspective. *Methods Mol Biol 1586*, 3–10.

4) García, J., Santana, Z., Zumalacárregui, L., Quintana, M., González, D., Furrazola, G. & Cruz, O. (2013) Estrategias de obtención de proteínas recombinantes en *Escherichia coli. Vacci Monitor 22(2),* 30-39.

5) Cai, D., Rao, Y., Zhan, Y., Wang Q. & Chen, S. (2019) Engineering *Bacillus* for protein production. *Journal of Applied Microbiology 126,* 1632-1642.

6) Terpe, K. (2006) Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems. *Appl Microbiol Biotechnol 72,* 211–222.

7) Flores, J. (2015) *Sobreproducción de la proteína viral HN-RVP en la levadura Schizosaccharomyces pombe*. Benemerita Universidad Autónoma de Puebla, Mexico: Thesis to obtain the title of Biologo.

8) Hunter, M., Yuan, P., Vavilala, D. & Fox, M. (2019) Optimization of Protein Expression in Mammalian Cells. *Curr Protoc Protein Sci 95(1),* e77.

9) Khun, K. H. (2013) Gene Expression in Mammalian Cells and its Applications. *Advanced Pharmaceutical Bulletin 3(2),* 257-263.

10) Drago, M. & Sainz, T. (2006) Sistemas de expresión para proteínas terapéuticas recombinantes. *Revista Mexicana de Ciencias Farmacéuticas 37(1)*, 38-44.

11) Ferrer-Miralles, N., Domingo, J., Corchero, J., Vázquez, E. & Villaverde, A. (2009) Microbial factories for recombinant pharmaceuticals. *Microbial Cell Factories 8(1)*, 17-18.

12) Jia, B. & Jeon, C. O. (2016) High-throughput recombinant protein expression in Escherichia coli: current status and future perspectives. *Open Biol 6(8),* 6160196.

13) Baeshen, M. N., Al-Hejin, A. M., Bora, R. S., Ahmed, M. M., Ramadan, H. A., Saini, K. S., . . . Redwan, E. M. (2015) Production of biopharmaceuticals in *E. coli*: current scenario and future perspectives. *J Microbiol Biotechnol 25,* 953–962.

14) Lara, A. (2011) Producción de proteínas recombinantes en *Escherichia coli*. *Revista Mexicana de Ingeniería Química 10(2),* 209-223.

15) Choi, J. H., Keum, K. C. & Lee, S. Y. (2006) Production of recombinant proteins by high cell density culture of *Escherichia coli*. *Chemical Engineering Science 61,* 876-885.

16) Lebendiker, M. & Danieli, T. (2014) Production of prone-toaggregate proteins. *FEBS Lett 588,* 236–246.

17) Rosano, G. L. & Ceccarelli, E. A. (2014) Recombinant protein expression in *Escherichia coli*: advances and challenges. *Front Microbiol 5,* 172.

18) Suzuki, T., Kitajima, K., Inoue, S., Inoue, Y. (1995) N-glycosylation/deglycosylation as a mechanism for the post-translational modification/remodification of proteins. *Glycoconjugate Journal 12(3),* 189-193.

19) Sai Sree, L, Reddy, M. & Gowrishankar, J. (2001) IS186 insertion at a hot spot in the lon promoter as a basis for lon protease deficiency of *Escherichia coli* B: identification of a consensus target sequence for IS186 transposition. *J Bacteriol 183,* 6943–6946.

20) Grodberg, J. & Dunn, J. J. (1988) ompT encodes the Escherichia coli outer membrane protease that cleaves T7 RNA polymerase during purification. *J Bacteriol 170,* 1245–1253

21) Bessette, P. H., Aslund, F., Beckwith, J. & Georgiou, G. (1999) Efficient folding of proteins with multiple disulfide bonds in the *Escherichia coli* cytoplasm. *Proc Natl Acad Sci USA 96,* 13703-13708.

22) Lehmann, K., Hoffmann, S., Neudecker, P., Suhr, M., Becker, W. M. & Rosch, P. (2003) High-yield expression in *Escherichia coli*, purification, and characterization of properly folded major peanut allergen Ara h2. *Protein Expr Purif 31,* 250-259.

23) Schlegel, S., Genevaux, P. & de Gier, J. W. (2017) Isolating *Escherichia coli* strains for recombinant protein production. *Cell Mol Life Sci 74,* 891–908.

24) Blount, Z.D. (2015) The unexhausted potential of *E. coli*. *Elife 4*, e05826.

25) Baumgarten, T., Schlegel, S., Wagner, S., Low, M., Eriksson, J., Bonde, I., . . . de Gier, J. W. (2017) Isolation and characterization of the *E. coli* membrane protein production strain Mutant56(DE3). *Sci Rep 7,* 45089.

26) Jonasson, P., Liljeqvist, S., Nygren, P. Å. & Sthål S. (2002) Genetic design for facilitated production and recovery of recombinant proteins in *Escherichia coli. Biotechnology and Applied Biochemistry, 35(2),* 91-105.

27) Baneyx, F. (1999) Recombinant protein expression in *Escherichia coli*. *Curr Opin Biotechnol 10,* 411–421.

28) Schmidt, F. R. (2004) Recombinant expression systems in the pharmaceutical industry. *Applied Microbiology and Biotechnology 65(4),* 363-372.

29) Choi, J.H. & Lee, S.Y. (2004) Secretory and extracellular production of recombinant proteins using *Escherichia coli. Appl Microbiol Biotechnol 64,* 625–635.

30) Rathore, A.S., Sobacke, S.E., Kocot, T.J., Morgan, D.R., Dufield, R.L. & Mozier, N.M. (2003) Analysis for residual host cell proteins and DNA in process streams of a recombinant protein product expressed in *Escherichia coli* cells. *Journal of Pharmaceutical and Biomedical Analysis 32(6),* 1199-1211.

31) Correa, A., Ortega, C., Obal, G., Alzari, P., Vincentelli, R. & Oppezzo, P. (2014) Generation of a vector suite for protein solubility screening. *Front Microbiol 5,* 67.

32) Jeong, H., Barbe, V., Lee, C. H., Vallenet, D., Yu, D. S., Choi, S. H., . . . Kim, J. F. (2009) Genome sequences of *Escherichia coli* B strains REL606 and BL21(DE3). *J Mol Biol 394,* 644-652.

33) Studier, F. W., Daegelen, P., Lenski, R. E., Maslov, S & Kim, J.F. (2009) Understanding the differences between genome sequences of *Escherichia coli* B strains REL606 and BL21(DE3) and comparison of the. *E. coli* B and K-12 genomes. *J Mol Biol 394,* 653–680.

34) Burdette, L. A., Leach, S. A., Wong, H. T. & Tullman-Ercek, D. (2018) Developing gram-negative bacteria for the secretion of heterologous proteins. *Microb Cell Fact 17,* 196.

35) Yoon, S. H., Jeong, H., Kwon, S-K. & Kim, J. F. (2009) Genomics, biological features, and biotechnological applications of *Escherichia coli* B: "is B for better?!". In: Lee, S. Y.,

Systems biology and biotechnology of Escherichia coli (pp 1–17). Dordrecht: Springer Netherlands.

36) Shamlou, P. A. (2003) Scaleable processes for the manufacture of therapeutic quantities of plasmid DNA. *Biotechnology and Applied Biochemistry 37(3),* 207-218.

37) Eguia, F. A. P., Ramos, H. R., Kraschowetz, S., Omote, D., Ramos, C. R. R., Ho, P. L., . . . Goncalves, V. M. (2018) A new vector for heterologous gene expression in *Escherichia coli* with increased stability in the absence of antibiotic. *Plasmid 98,* 22–30.

38) Ali, S. A., Chew, Y. W., Omar, T. C. & Azman, N. (2015) Use of FabV-Triclosan plasmid selection system for efficient expression and production of recombinant proteins in *Escherichia coli. PloS One 10,* e0144189.

39) Novagen (2003) pET System Manual. Obtained from https://lifewp.bgu.ac.il/wp/zarivach/wp-content/uploads/2017/11/Novagen-pET-system-manual-1.pdf. Revised April 2018

40) Kim, S., Jeong, H., Kim, E. Y., Kim, J. F., Lee, S. Y. & Yoon, S. H, (2017) Genomic and transcriptomic landscape of *Escherichia coli* BL21(DE3). *Nucleic Acids Res 45,* 5285–5293.

41) Studier, F. W., Rosenberg, A. H., Dunn, J. J. & Dubendorff, J. W. (1990) Use of T7 RNA polymerase to direct expression of cloned genes. *Meth. Enzymol 185,* 60–89.

42) Rosenberg, A. H., Lade, B. N., Chui, D. S., Lin, S. W., Dunn, J. J. & Studier, F. W. (1987) Vectors for selective expression of cloned DNAs by T7 RNA polymerase. *Gene 56(1),* 125–135.

43) Studier, F. W. & Moffatt, B. A. (1986) Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J Mol Biol 189(1),* 113–130.

44) Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular cloning. A laboratory manual*. New York: Cold Spring Harbor Laboratory Press.

45) Purifying Challenging Proteins. (2007) *Principles and Methods*. Uppsala, Sweden: GE Healthcare Bio-Sciences.

46) Fischer B., Perry, B., Sumner, I., Goodenough, P. (1992) A novel sequential procedure to enhance the renaturation of recombinant protein from Escherichia coli inclusion bodies. *Protein engineering 5(6),* 593-596.

47) Middelberg, A. P. J. (2002) Preparative protein refolding. *Trends in Biotechnology 20(10),* 437-443.

48) Geng, X. & Wang, C. (2007) Protein folding liquid chromatography and its recent developments. *Journal of Chromatography B 849(1-2*), 69–80.

49) Wurm, D. J., Quehenberger, J., Mildner, J., Eggenreich, B., Slouka, C., Schwaighofer, A., . . . Spadiut, O. (2018) Teaching an old pET new tricks: tuning of inclusion body formation and properties by a mixed feed system in *E. coli. Appl Microbiol Biotechnol 102*, 667–676.

50) Ramon, A., Senorale-Pose, M. & Marin, M. (2014) Inclusion bodies: not that bad. *Front Microbiol 5*, 56.

51) Jong, W. S. P., Vikström, D., Houben, D., van den Berg van Saparoea, H. B., de Gier, J-W. & Luirink, J. (2017) Application of an *E. coli* signal sequence as a versatile inclusion body tag. *Microb Cell Fact 16,* 50.

52) Paraskevopoulou, V. & Falcone, F. H. (2018) Polyionic tags as enhancers of protein solubility in recombinant protein expression. *Microorganisms 6(2),* 47.

53) Vandemoortele, G., Eyckerman, S. & Gevaert, K. (2019) Pick a tag and explore the functions of your pet protein. *Trends Biotechnol 37(10),* 1078-1090.

54) Smith, D. B. & Johnson, K. S. (1988) Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. *Gene 67,* 31–40.

55) Raines, R. T., McCormick, M., Van Oosbree, T. R. & Mierendorf, R. C. (2000) The S.Tag fusion system for protein purification. *Methods Enzymol 326*, 362–376

56) Jayanthi, S., Gundampati, R. K. & Kumar, T. K. S. (2017) Simple and efficient purification of recombinant proteins using the heparin-binding affinity tag. *Curr Protoc Prot Sci 90,* 6.16.11–6.16.13.

57) Morris, J., Jayanthi, S., Langston, R., Daily, A., Kight, A., McNabb, D. S., . . . Kumar, T. K. S. (2016) Heparinbinding peptide as a novel affinity tag for purification of recombinant proteins. *Prot Express Purif 126*, 93–103.

58) Kimple, M. E., Allison, L., Brill, A. L., & Pasker, R. L. (2015) Overview of Affinity Tags for Protein Purification. *Curr Protoc Protein Sci 73,* 9.9.1-9.9.23.

59) Malhotra, A. (2009) Chapter 16: Tagging for Protein Expression. *Methods in Enzymology 463,* 239-258.

60) Qiagen. (2003) *The QIAexpressionist™. A handbook for high-level expression and purification of 6xHis-tagged proteins.* Obtained from https://www.qiagen.com/es/resources/download.aspx?id=79ca2f7d-42fe-4d62-8676-4cfa948c9435&lang= en. (Revised April 2019).

61) Nguyen, T. K. M., Ki, M. R., Son, R. G. & Pack, S. P. (2019) The NT11, a novel fusion tag for enhancing protein expression in *Escherichia coli*. *Appl Microbiol Biotechnol 103,* 2205–2216.

62) Han, Y., Guo, W., Su, B., Guo, Y., Wang, J., Chu, B. & Yang, G. (2018) High-level expression of soluble recombinant proteins in *Escherichia coli* using an HE-maltotriose-binding protein fusion tag. *Prot Express Purif 142,* 25–31.

63) Ward, W. W. & Swiatek, G. (2009) Protein Purification. *Current Analytical Chemistry 5(2),* 85-105.

64) Zhang, J. Z., Lu, T-W., Stolerman L. M., Tenner, B., Yang, J. R., Zhang J-F., . . . Zhang, J. (2020) Phase Separation of a PKA Regulatory Subunit Controls cAMP Compartmentation and Oncogenic Signaling. *Cell 182,* 1531-1544.

65) Bangaru, S., Ozorowski, G., Turner, H., Antanasijevic, A., Huang, D., Wang, X, . . . Ward, A. B. (2020) Structural analysis of full-length SARS-CoV-2 spike protein from an advanced vaccine candidate. *Science 370,* 1089–1094

66) Arakawa, T., Philo. J., Tsumoto, K., Yumioka, R. & Ejima, D. (2004) Elution of antibodies from a Protein-A column by aqueous arginine solutions. *Protein Expr Purif 36*, 244-248.

67) Magdeldin, S. & Moser, A. (2012) Affinity chromatography: Principles and applications. In: Magdeldin, S., *Affinity chromatography* (pp 3-28). Rijeka, Croatia: inTech.

68) Qiagen. (2020) *QIAquick® Spin Handbook.* Obtained from https://www.qiagen.com/us/resources/download.aspx?id=95f10677-aa29-453d-a 222-0e19f01ebe17&lang=en. (Revised April 2019).

69) Qiagen. (2018) *Quick-Start protocol. QIAquick® Gel Extraction Kit QIAquick® PCR & Gel Cleanup Kit*. Obtained from https://www.qiagen.com/us/resources /download.aspx?id=a72e2c07-7816-436f-b920-98a0ede5159a&lang=en. (Revised January 2019).

70) Laemmli, U. K. (1970) Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature 227,* 680–685.

71) Qiagen (2008) *Ni-NTA Spin Kit Handbook*. Obtained from https://www.qiagen.com /us/resources/download.aspx?id=3fc8c76d-6d21-4887-9bf8-f35f78fcc2f2&lang= en. (Revised January 2019).

72) Qiagen. (2016) *Ni-NTA Agarose Purification of 6xHis-tagged Proteins from E. coli under Native Conditions.* Obtained from https://www.qiagen.com/us/resources/ download.aspx?id=dc89c299-75d3-4120-adf0-35251f16a7af& lang=en. (Revised December 2018).

73) Gushiken, E. (2016) *Generación de anticuerpos de dominio único específicos para CD105 humano.* Universidad Peruana Cayetano Heredia, Perú. Thesis to obtain the degree of Magister en Ciencias, mention Biochemistry and Molecular Biology.

74) Lamdan, H., Ayala, M., Rojas, G., Munoz, Y., Morera, Y., Guirola, O., . . . Gavilondo, J. (2011) Isolation of a novel neutralizing antibody fragment against human vascular endothelial growth factor from a phage-displayed human antibody repertoire using an epitope disturbing strategy. *J Biotechnol 151(2),* 166–174.

75) Maass, D. R., Sepulveda, J., Pernthaner, A. & Shoemaker, C. B. (2007).Alpaca (*Lama pacos*) as a convenient source of recombinant camelid heavy chain antibodies (VHHs). *J Immunol Methods 324(1–2),* 13–25.

Chapter IV: Recombinant VHH protein microarrays for the detection of CD105 on cells by SPR imaging

Abstract

The role of CD105 as key molecule in the regulation of angiogenesis make it a target for developing CD105-binding VHHs both as a marker for the diagnosis of several diseases and for biological anti-tumour therapy. In this scenario, the developing biochips to study the interaction of the VHHs to CD105 in normal and transformed cells *in vivo* as well as the capacity to detect, bind and capture the target with nanoprobes is promising. Purified recombinant anti-CD105 VHHs were functionalised for application in biochips to assess the affinity and specificity binding to CD105 expressed on cells by Surface Plasmon Resonance imaging (SPRi) in real-time. The microarrayed VHH on the biochip were incubated with SC cells, which express CD105 on their surface, and THP-1 cells, which do not. The analysis of the SPRi images and the variations in reflectivity (SPRi signal) showed that most of the VHH were able to bind SC cells, but very little or none to THP-1 cells, suggesting their binding specificity. The cell density for each VHH spot was obtained by counting cells in a spot area of 2.28 x10^5 μm^2 (477.09 µm x 477.09 µm) from differential images at 2 hours after the addition of cells in the assay and using grey level shifts of 0.1-0.15 for SC cells and 0.3 for THP-1 cells. The grey levels were chosen to provide cell density values in accordance with the reflectivity variations (SPRi signals) of each VHH in the same assay time (2 hours). The reflectivity variation average among the VHHs at 2 hours of assays with SC cells was 4.59 and with the THP-1 cells was 2.69. Cell density average with SC cells was 578 cells/area and with THP-1 cells was 367 cells/area and approximation to the apparent association constant average with SC cells was 1.1019 and with THP-1 cells was 1.0626. The VHHs were divided into 4 groups by their specificity, cell density and approximation to the apparent affinity constant. The top six anti-CD105 VHH in the ranking, VHH 17, VHH 7, VHH31 in group 1 and VHH22, VHH 3 and VHH 2 in group 2, had higher average values among the VHHs and are proposed as nanoprobes for real-time capture and detection of CD105 (soluble and membrane-bound) for diagnostic or therapeutic research studies.

4.1 Scientific context

4.1.1 Microarrays

Microarray technology is a powerful research tool for the simultaneous analysis of multiple analytes from different biological samples in a single assay[1]. Since the first report in 1983 on protein microarrays[2] and in 1995 on DNA microarrays[3], the application of microarrays has spread widely in various fields due to its versatility, robustness and automatization. DNA and protein microarrays have been used for a variety of research as described in multiple reviews[4-39].

Some of these applications include: dose-response studies to establish differential gene expression profiles in different experimental conditions (diseases, treatments, etc.)[40], to quantify RNA molecules[41], to measure the number of DNA copies present in a genome[42], to compare allele expression levels with polymorphisms[43], to assess point mutations, evolutionary studies, to perform pharmacological analyses[44], in kinetics for

transcriptional analysis[45], to determine presence of methylations. To design better cancer treatment schemes and to make predictions of response to chemotherapy treatment[46-48], to identify and diagnose some types of cancer that are difficult to establish clinically[49]. To study different diseases such as colorectal cancer[50], breast cancer[51], oesophageal cancer[52], lung cancer[53], marginal zone lymphoma[54], multiple sclerosis[55], osteoporosis[56]; to elucidate mechanisms inherent to diabetes mellitus[57], to discover new drugs[58,59], to identify new drug targets such as in spinal muscular atrophy[60]. To characterise and identify microorganism strains of clinical interest[61-63], to diagnose diseases[64-66], to study protein functions[67-69], to develop analytical sensors[70-72], recognition studies of proteins capable of interacting with other proteins such as the new calmodulin- and phospholipid-interacting proteins[73]. To test antigen production in tissues such as antigens obtained from oral cavity cancer biopsies[74], or autoantigens in rheumatoid arthritis[75] or lupus erythematosus[76], to discover biomarkers[77-80] among many others. Some studies even combine the use of DNA microarrays with protein/tissue microarrays, complementing gene expression information with protein-protein interactions and antigen production in tissues.

One of the main features and strengths of protein microarrays is that they can provide relevant information on protein-protein interactions[25]. Antibody microarrays have been used to detect and analyse antigens of interest from cells or tissues of different origins[30]. In terms of antibodies as recombinant probes, high-affinity scFvs have proven especially useful for protein microarrays[81] enabling biomarker profiling of cancers[82,83] and autoimmune diseases[84].

Proteomics seek to describe and understand the functions of individual proteins in living matter as part of an extended network of interacting molecules[85]. The assessment of biomolecular interactions in microarrays occurs by two main mechanisms: one based on labelling technology, which can be fluorescent[86,87] or radiative[88-90], and the other based on non-labelling methods such as waveguide technology[91], piezoelectric techniques based on mass shifts[92] and surface plasmon resonance[93-95]. The former are more sensitive than the latter (from pg/ml to ng/ml), however, they can also present cross-sensitivity with other analytes than those of interest in the evaluation[96,97], and they are more costly and laborious than non-labelling techniques.

The development of a proteomic platform, which applies microarray technology for the screening of VHHs with higher specificity and affinity to their antigen, offers a promising prospect for improving the VHH selection process by increasing the repertoire evaluated and accelerating the process. Previously selected VHHs expressed by subcloning their nucleotide sequences into a bacterial expression vector can be structurally analysed to gain a more detailed understanding of the antigen-antibody interaction using biosensor microarray technology based on Surface Plasmon Resonance imaging (SPRi).

4.1.2 Biosensors

A biosensor or biological sensor is a system with three components: a biological compound that recognize an analyte, a transducer and a signal-processing unit. The

biological compound can be a protein, an antibody[98], an enzyme[99], inhibitors[100], DNA, organelles, cells, living microorganisms, plant or animal tissues[101], aptamers[102], etc. The biosensor is a device that selectively measure a specific analyte in a medium with complex mixtures of analytes through a physicochemical detector. The device translates the chemical change produced by the binding of the biological compound to the analyte into an analytical signal that can be processed within minutes[103]. Detection can be done by electrical, mechanical, electro-chemical, thermal or optical signals[104]. The first biosensors were enzymatic electrodes developed for glucose by Leland C. Clark in 1962[105,106]. Since then, they have been further developed to be more sensitive, specific and reliable. Biosensors are widely used by researchers, whether in medicine[107-109], food industry, environmental monitoring[110], industrial and technological process control[111], so, it is necessary to have an understanding of the various types of biosensors that can be used, the principles behind them, as well as their advantages and limitations.

Biosensors are classified according to the type of transducer, the type of biorecognition element employed, or the type of analyte detected. Depending on the type of transducer they can be: electrochemical[109], thermal/ calorimetric[112,113], acoustic[114], optical[115], etc.

4.1.2.1 SPR-based optical biosensors

Optical biosensors are based on the measurement of changes in optical characteristics caused by the analyte-receptor reaction. In unlabelled detection the assessment of the analyte is direct and in labelled detection (such as fluorescence) the effect is observed in the labelled analyte compared to the surrounding medium[116]. The transducer can detect changes in absorbance, luminescence, polarisation, or refractive index[117,118]. Many optical biosensors employ techniques based on the surface plasmon resonance (SPR) phenomenon[119] or its modified version, Surface Plasmon Resonance imaging (SPRi)[96,120-123], multiparametric surface plasmon resonance (MP-SPR)[124] or localised surface plasmon resonance (LSPR)[125].

SPR-based optical biosensors have gained attention due to their speed of detection, high specificity, high sensitivity, small size, good analytical parameters, cost-effectiveness and possibility of real-time analysis, and have gained recognition as a valuable tool for investigating biological interactions. Use of SPR for biosensors, for example, those that need to evaluate antigen-antibody recognition interactions, ensures affinity, specificity and reliability of detection. It is a direct measurement, so no labelling of the molecules to be analysed is necessary[121,126-128].

4.1.3. Basis of the SPR method

The SPR method is an optical technique that detects and measures minute changes in the refractive index of a metallic surface caused by molecules bound to that surface. This highly sensitive detection is based on a collective excitation of surface plasmons in a metal film on a glass or crystal support, leading to the total absorption of light at a particular angle of incidence, which is dependent on the refractive indices of either side of the metal film[103]. Plasmons, free oscillating electrons, are present in metals. Plasmons, like photons, are quanta of electromagnetic vibrations, so the role they play in the optical properties of

metals is important. The plasmon resonance phenomenon occurs when the energy of the photon incident on the surface of the metal is equal to the energy of the oscillating plasmon vibrations. The plasmon energy inside the metal is higher than the photon energy in the visible light range, only plasmons in the metal surface layer (with a thickness of 300-400 nm) can enter into resonance. This is why this phenomenon is called Surface Plasmon Resonance (SPR)[116]. The SPR effect is observable under specific conditions[116]. First, it is necessary to place the thin metallic film between two media of different optical density, e.g. glass/water, glass/air, etc[129], and the evanescent field needs to be generated, so the metal used on the surface must allow the electrons to behave freely. Suitable metals include silver, gold, platinum, copper and aluminium; which gold and silver (less chemically stable than gold) are the most commonly used. Silver provides a sharp SPR resonance peak and gold because of its high chemical stability provides a stable peak[103,130,131].

Two configuration schemes (Otto and Kretschmann) are the most popular to use the surface plasmon resonance phenomenon. The Kretschmann configuration is used in most practical applications and is based on the total reflection phenomenon[97,131-133]. When a monochromatic, p-polarised light beam penetrates the system: glass (with high refractive index n1)/thin metal film/dielectric interface (with low refractive index n2), the evanescent wave will interact with plasmons on the surface of the metal film[134]. The evanescent wave is a wave that is generated at the boundary of the interface between two media where a light beam is reflected and its intensity decays exponentially with penetration into the medium[135]. Therefore, when surface plasmon resonance occurs, the energy of the incident light is lost in the metal film resulting in a decrease of the reflected light intensity[116]. The phenomenon of resonance where the reduction of the intensity of reflected light is produced, occurs at a particular angle called resonance angle, SPR angle or angle of minimum reflectivity. This angle is sensitive to the composition of the coating on the metal surface and dependent on the refractive index of the medium near the surface[133,136-138].

To explain the SPR principle in a simplified way, Rickling[133] considered the properties of photons and surface plasmons (which are both a form of electromagnetic energy and can only be fully described by quantum physics) as vector quantities. Thus, the momentum of the light photon at the interface can be decomposed into two vector components (parallel and perpendicular to the interface). The magnitude of these incident light vectors (ilv) depends directly on the light angle. The surface plasmon wave can similarly be described as a vector, which depends on several factors (metal properties, layer thickness, surrounding medium). Only when the energy and momentum of the incident light vector correspond exactly to those of the surface plasmon vector (spv), a resonance phenomenon occurs and the photons are converted into plasmons. Otherwise, there is no such conversion and the light is reflected (Figure 4.1).

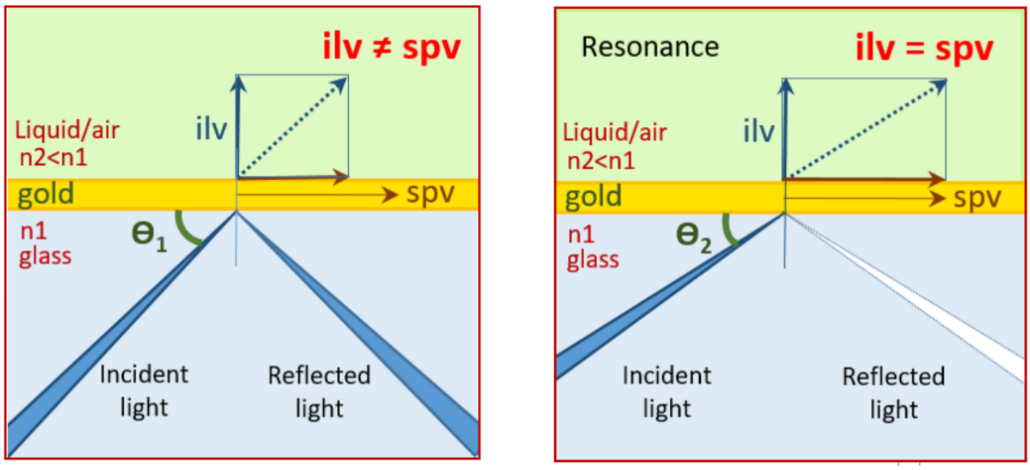


Figure 4.1- SPR principle on a gold surface explained with vector quantities. a) When the incident light vectors (ilv) have a different value than the surface plasmon vector (spv), the light is reflected, b) At a single specific angle there is coincidence of the vectors, there is resonance. **θ:** angle of incidence, **n1** or **n2:** refractive index. (Adapted from Rickling [133] and Biacore[130]).

The SPR technique consists of measuring the intensity of the reflected light. The value of the SPR angle or resonance angle depends on the refractive index of the medium with a lower refractive index at the metal surface. The binding of the molecules to the surface causes a change in the refractive index at the vicinity of the metal surface, which results in a variation of the SPR angle that is detected and quantified by the instrument. There is a linear relationship between molecule mass and SPR angle shift[131]. SPR angle shifts can be determined by varying the angle of incidence and recording the reflected light intensity during the binding interactions of the molecules[119].

The difference of the initial and final SPR angles is proportional to the adsorption and the positive gradient of the SPR adsorption curve gives the rate of the reaction[132]. The unit of the response signal is called the resonance unit (RU) and represents a shift in the resonance angle of approximately 10^{-4}°[139], i.e. a change in signal equal to 1000 RU corresponds to a shift in the SPR angle of 0.1°. Since the mass of the molecules directly influences the refractive index, SPR biosensors are often referred to as mass detectors. In the case of proteins, the correlation between sensor signal and mass increase was experimentally determined: 1 RU = 1 pg/mm^2 [136]. For globular proteins with 1000 RU is equivalent to a surface concentration of 1 ng/mm^2. The total range covered by the SPR detector is 3°, or 30,000 RU[103]. The resonance signal at any given time is the sum of the contributions from the sensor surface, the interacting molecules, and the amount of solution[140].

4.1.4 SPRi method

Conventional SPR sensors measure the dependence of reflectivity on the angle of incidence, whereas in the SPR imaging version the complexity of the scanning angle is eliminated, i.e. measurements are performed at a single particular angle of incidence of the light. In addition, SPR sensors performs single-point analysis while the SPRi performs multiple point analysis. In SPRi, the reflected light is collected through a charge-coupled device (CCD) and presented as an image[116,120,141]. The angle at which the measurement is

made is in the linear region of the reflectivity decrease. Thus, the changes in light intensity are proportional to the successively bound molecules on the detection surface. After adsorption of the molecules, higher resonance occurs at larger angles of incidence. More light is reflected from the reaction site than from the gold-covered areas. The SPR shift curve occurs when analyte molecules interact with reaction sites, which increases the degree of reflection (Figure 4.2).

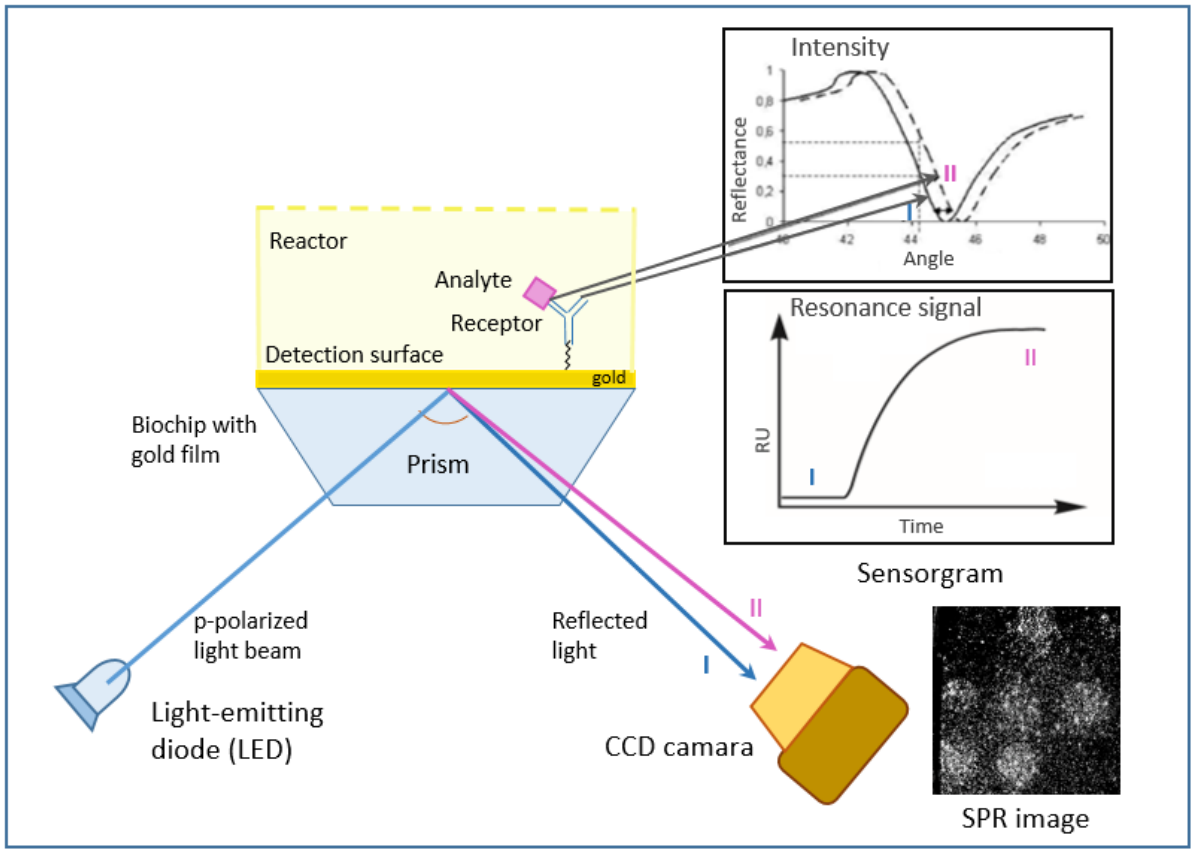


Figure 4.2- Schematic diagram of Surface Plasmon Resonance imaging (SPRi) setup, detection and readout. (Adapted from Sankiewicz *et al*[116], Suraniti *et al*[120] and Biacore[130]).

SPRi instrument is generally composed of an optical system, a transduction medium that interrelates the optical with the biochemical and an electronic system for data processing. The first consists of: a) the optical radiation source which can be a He-Ne laser, Ar laser or light-emitting diode (LED), b) the optical system of lenses, polarisers and mirror, and c) a prism with a metal film on which the surface plasmons are excited[142,143]. The second transforms changes in the metal surface into changes in the refractive index that can be determined by optically examining the SPR. The third generates and processes an electronic signal during the SPR examining process. A CCD (Charge-Coupled Device) camera is used to capture the reflected light from the entire chip surface[116].

In addition, it is necessary to generate the biochip with a sensitive recognition element[144]. Usually, they are glass chips coated with an inert metal (e.g. gold with a thickness of 50 nm). The unmodified gold layer is not a suitable surface environment for biomolecular interaction. Therefore, the gold is modified, e.g. with a special monolayer such as thiol[102,131,144-146]. The immobilisation of the biocomponents on the chemically modified gold layer is done by adsorption, covalent bonds, hydrophobic interaction or capture

particles by their relative specificity[147]. This is an important step in the production of a biosensor because it influences its efficiency.

The development of an SRPi microarray that allows the detection of changes in the optical properties produced by an antigen-antibody (VHH) interaction between the receptor previously deposited on the metal surface (anti-CD105 VHH), and the complementary analyte (CD105), will allow the characterisation of each anti-CD105 VHH and real-time measurements of the interactions without the need for labelling. In addition, it will explore the use of VHHs as tools for CD105 recognition and capture, not only in free form but also on cells that express it, and perhaps establish a basis for future biomedical research. Research to aid in the detection, diagnosis or treatment of diseases or syndromes where CD105 plays an important role, or for studies that seek to better understand the mechanisms and the way in which CD105 participates.

4.2 Procedures

4.2.1 Functionalization of the VHH

The functionalization of the proteins was performed following the protocol established in CREAB laboratory, SyMMES for antibody functionalization (IgG) and adapting it for the nanobodies (VHH). The anti-CD105 VHH proteins were taken up to 500 µg/ml concentration in PBS by centrifugation in 3,000 MWCO (3K) VivaSpin minicolumns. In order to achieve this, minicolumns were washed with 200 µl of PBS pH7.4 and centrifuged at 15,000 g for 15 minutes at RT°. The 13 different VHH proteins were added, each one in a different microcolumn and with the required volume to obtain 100 µg of protein, then they were centrifuged at 15,000 g for 15 minutes at RT°. They were washed twice with 100 µl of PBS before suspending them in 200 µl of PBS.

For the coupling reaction, the SH-C_{12}-N-Hydroxysuccinimide (Thiol-NHS) linker was added in a ratio 10:1 in relevance with the molar concentration of the protein. That is, 333.3 µM of NHS linker since in 200 ul of each VHH with an average molecular weight of 15 kDa (15000g/mol) at 0.5 mg/ml there are approximately 6.6 nmoles, i.e. 33.3 µM. The coupling reaction was performed in PBS pH 7.4, at RT° for 1 hour. After this time, the conjugated proteins were purified in VivaSpin minicolumns previously washed with PBS pH 7.4 by centrifugation at 15,000 g for 15 minutes at RT°. The conjugated proteins were washed twice with 200 µl PBS and finally recovered in 200 µl PBS pH 7.4. The concentration of each VHH protein was measured by absorbance at 280 nm in Nanodrop.

4.2.2 Preparation of the biochips

To generate the microarray carrying the anti-CD105 VHHs to be evaluated, commercially obtained glass prisms with a gold surface were used as biochips. 24 hours before their use, the biochips were placed in a low-pressure plasma generator system for Sputtering by gas flow, 75% oxygen, 25% argon for 3 minutes at 0.6 mBars according to the parameters stablished at the laboratory following the protocol of Guedon *et al*[148]. 5% glycerol was added to the functionalised VHH protein samples at 0.1 µg/ul for their application over the biochips. Using a micropipette, a drop of 0.1 ul of each sample was

placed over the gold surface of the biochip without the tip touching such surface. The biochips with the VHH proteins distributed over them (microarrays) were placed in a humid chamber for 24 hours. After this time, the microarrays were rinsed with PBS 1X pH 7.4 and were kept immersed in this buffer at 4° C until use.

4.2.3 Cell cultures

4.2.3.1 Cell line SC (ATCC® CRL-9855™)

SC is a cellular line of the monocytes/macrophages type obtained from mononuclear cells of human peripheral blood. They grow as free cells in suspension with a granular appearance[149]. This cell line was deposited originally in support of a patent (United States patent 5.447.861). For the cell culture the complete growth medium recommended by ATCC was employed, consisting of 90% of Iscove's Modified Dulbecco's Medium (IMDM) (ATCC 30-2005TM) supplemented with 0.05 mM of 2-mercaptoethanol, 0.1 mM of hypoxanthine, 0.016mM of thymidine, 1X PenStrep and 10% of Fetal Bovine Serum (FBS)

4.2.3.2 Cell line THP-1(ATCC® TIB-202™)

THP-1 is a cell line of the monocyte type obtained from mononuclear cells of human peripheral blood[150]. They are phagocytic cells that lack surface and cytoplasmic immunoglobulin. They can be differentiated by induction with the ester of phorbol 12-O-tetradecanoylphorbol-13-acetate (TPA)[151]. For the cell culture, the complete growth medium composed by 90% Roswell Park Memorial Institute 1640 (RPMI 1640) medium (ATCC® 30-2001™) supplemented with 1 X PenStrep and 10% of Fetal Bovine Serum (FBS) was used.

4.2.3.3 Culture

The complete medium was prepared immediately prior to use and previously incubated at 37° C for 15 minutes to reach a pH between 7.0 and 7.6 and avoid its excessive alkalinity during the recovery of the cells. The cryopreserved cells were thawed and transferred to a complete fresh medium for viable cell counting. After centrifugating at 130 g for 15 minutes, the cells were suspended in the complete medium to 2.5 to 3 x 10^5 cells/ml. The culture flasks were incubated at 37° C with 85% relative humidity and 5% CO_2 in the atmosphere. The cells were sub-cultivated by replacement of the complete growth medium every 2 to 3 days depending on the cell density, which was maintained between 2.5 x 10^5 and 1 x 10^6 cells/mL.

4.2.3.4 Cryopreservation

The cell cultures were centrifuged at 125g for 5 minutes and the cellular pellet was suspended in cryopreservation medium (complete growth medium supplemented with 5% v/v of DMSO) until a cell density of 5 x 10^6 cells/mL was reached. Aliquots of 1ml of the cellular suspension were distributed into cryo-vials, which were placed in isopropanol-based freezing chamber overnight at -80° C. Finally, the carefully labelled cryovials were stored in liquid nitrogen vapor phase.

4.2.4 SPRi

The SPRi-based assays extended the principle of the method used by Cherif *et al*[121] and Bouguelia *et al*[126]. This optical technique measured the changes in the refraction index of the detection surface and its vicinity in response to the interaccions of recognition antigen – antibody (CD105 – VHH), at the same time that images were generated throughout the entire process. The anti-CD105 VHH (capture agents) immobilized on the gold surface of the biochip, conformed this detection surface that came into contact with the cell suspension, (SC and THP-1 cell lines at 1 x 10^6 cel/ml) that carry (or not) the CD105 antigen (molecule to be detect). The reactions were made in the cell culture medium, IMDI when working with SC and RPMI when working with THP-1, at 37° C. To achieve it, the VHH microarray was placed immediately covered with the cell-free culture medium tempered to avoid drying out. BSA (Bovine Serum Albumin) 1% as a blocking agent was added and washed three times with the medium before adding the corresponding cell suspension.

The reactions were performed in a PEEK (polyether ether ketone) reactor in combination with a SPRi PlexII instrument and a CCD camera inside an incubator set at 37° C. The signal measurements were made using a SPRi Lab™ system. The SPR signal was monitored and the changes in reflectivity of each ROI (region of interest) were tracked and plotted at the same time. The data were collected and processed in spreadsheets. The assays included a positive control (EngVHH17), a control with an antigen other than CD105 (anti-Fas2 VHH, unrelated protein) and a control with no antigen present (THP-1 cells) control.

4.2.5 Image analysis

The images were captured with an 8-bit camera and processed with the Image J program version FIJI (http://fiji.sc/Downloads) to visualize the differential images obtained throughout the whole process. For cell counting, the Analize/Analyze Particles tool of the Image J program and the command im2bw from MATLAB program[152] were employed. The particle size was 15.39 um, which is approximately the average cell size. Each image consisted of a two-dimensional array or matrix of 1040 x 696 pixels and typically contained 15 VHH spots distributed across it. In each spot, a region of 31 x 31 pixels or 477.09 µm x 477.09 µm (a correspondence was established between the number of pixels and the real length, 1 pixel = 15.39 µm) located approximately in the centre of each spot was selected. Once this area was cropped, the values of each pixel, which must be in a grayscale, were read. That is, each pixel had a numerical value ranging from 0 to 255 representing the luminance (light intensity per unit area) of the image. Thus, the value 0 was equivalent to no luminance or zero luminance (black), while 255 is the highest luminance (white). This is known as a greyscale image. A criterion was established to discriminate these pixels by binarisation, that is, to reassign to each luminance value a new value which can be either 0 or 1. If the value is 0, the particle is considered not to be present and therefore not counted, but if the value is 1, a particle is added to the count. To achieve this, BW= im2bw(I,level) converts the greyscale image I to binary image BW, replacing all pixels of the input image with a luminance greater than the preset level by

the value 1 (white) and replacing all other pixels by the value 0 (black). The most commonly used discrimination criterion is the one that gives a value of 0 to pixels whose luminance ranges from 0 to 128, while a value of 1 is assigned to the remaining pixels whose luminance ranges from 129 to 255, corresponding to level 0.5 (50%) (intermediate between the minimum and the maximum luminance value). Thus, this command allows to set a percentage of pixels that will be considered as having a value of 0 or 1. For example, at a level 0.15 (15%), the command run in the subsequent manner: >>Imbw=im2bw(Im,0.15); criteria to binarize the area 31 x 31. That is, setting a level of 0.15, meant that, first of all, the Matlab command set a lower limit corresponding to the pixel with the lowest luminance value in the area and an upper limit corresponding to the pixel with the highest value. Within this range, it only selected the pixels with the value in the highest 15% of luminance and assigned them a value of 1, while the other pixels were assigned a value of 0. It then counted how many 1's there are in each region to obtain as a result the number of pixels (particles) per area. Levels of value 0.1, 0.15, 0.2, 0.3, 0.4, 0.6 and 0.8 were used with this programme and the data was collected for analysis in a spreadsheet. The level that provided a number of particles (cells) per área (cell density) proportional to the SPRi signal in a given time (2 hours) for each VHH, was selected.

The values of the approximations to the apparent association constant of the VHHs were calculated by dividing the values of the SPRi signal (reflectivity variation) 2 hours after the cells were added to the assay over the values when the cells were added to the assay (initial values). The values for the calculations are shown in annexes (section A2).

4.3 Results and Discussion

To make the microarray of proteins or proteomic platform, the 13 purified recombined VHH proteins were functionalised for their immobilization over the biochip. For this purpose, the Thiol-NHS linker was employed. The crosslinkers contain reactive groups that are often activated to react with certain functional groups such as amines and sulfhydryl[153]. The esters of NHS, like Thiol-NHS, react with primary amines (-NH2) making them, therefore, one of the most employed linkers to functionalize proteins. The primary amines at the N-terminal and the side chain of the lysine residues located in the exterior of the tertiary structure of the native protein can be easily accessible to the conjugation reagents introduced in the aqueous medium. The thiol-NHS linker reacts with the primary amines of the proteins under physiologic to lightly alkaline conditions (pH 7,2 to 9) to produce stable amide bonds. The reaction releases N-hydroxysuccinimide (NHS). The outline of the expected reaction in the functionalization of the VHH is shown in figure 4.3.

The thiol-VHH (VHH-SH) were immobilized to the biochip by direct adsorption in a chemisorption process by gold-thiol interactions. This reaction is possible since gold can act as a weak electron pair acceptor (Lewis's weak acid) and has a strong affinity for weak electron pair donors (Lewis's weak base) such as thiols (R-SH), disulfides (RS-SR) and thioethers (R-S-R)[154,155].

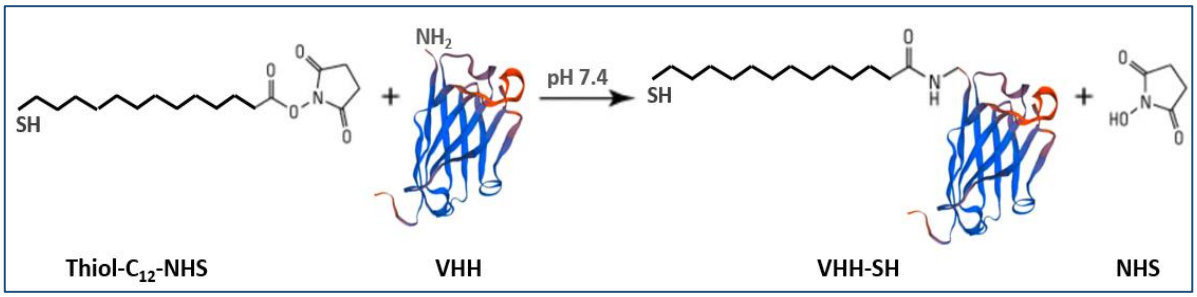


Figure 4.3- Scheme of the chemical conjugation reaction between Thiol NHS and the primary amines of the VHH proteins for their functionalization.

Since the functionalization of the VHH proteins occurs at the level of the primary amines, including the side-chain amine (ε amine) of the lysine residues and that according to the analysis of VHH primary structure (Chapter II, Figure 2.8b), they possess between 3 to 7 lysine residues (Figure 4.4), there exist several possibilities for positioning the protein over the biochip. One or some of these positions would not be suitable for the antigen recognition, as they would not have the correct exposure to it. In fact, the direct adsorption processes are more recommended for peptides or oligonucleotides, where the conformational changes will not affect their functionality[145,156]. The VHH are small and conformationally very stable and to diminish the loss of some of these capture agents by an inconvenient positioning, the concentration of VHH was increased from 10 ng to 50 ng per volume in each spot, meaning 5 times more.

VHH sequence in pHEN2	VHH 1	VHH 2	VHH 3	VHH 4	VHH 5	VHH 6	VHH 7	VHH 8	VHH 15	VHH 16	VHH 17	VHH 18	VHH 19
Amount of Lys (K) present	7	7	5	6	3	7	6	6	7	5	5	6	9
Recombinant VHH in pET22b(+)	VHH 1	VHH 2	VHH 3	VHH 31	VHH 7	VHH 8	VHH 9	VHH 11	VHH 17	VHH 20	VHH 22	VHH 26	VHH 18

Figure 4.4- Amount of Lysines residues (K) present in the composition of VHH determined from their nucleotide sequences (chapter II)

Glycerol was added to the solutions with the VHH to give them viscosity and prevent their dispersion during their application over the biochip surface, just the same to diminish the evaporation of the microdroplets by reducing the vapor pressure of the solution[157].

Initially microarrays of proteins with three replicates of the 13 VHH and positive and negative controls were made on a single biochip employing an automatic injection system that would allow the distribution of the15 samples in triplicates, that is, 45 spots. However, the area of distribution for each individual VHH (spot) on the biochip was too small for the interaction with the antigen on the cells, which are considerably larger when compared to the VHH (15 μm approximately), to have enough space to unfold and occur properly. Therefore, the size of the spot was increased adding manually the samples and the repetitions were made on different biochips. Two or three ROIs (regions of interest) were chosen inside each spot of every microarray (Figure 4.5).

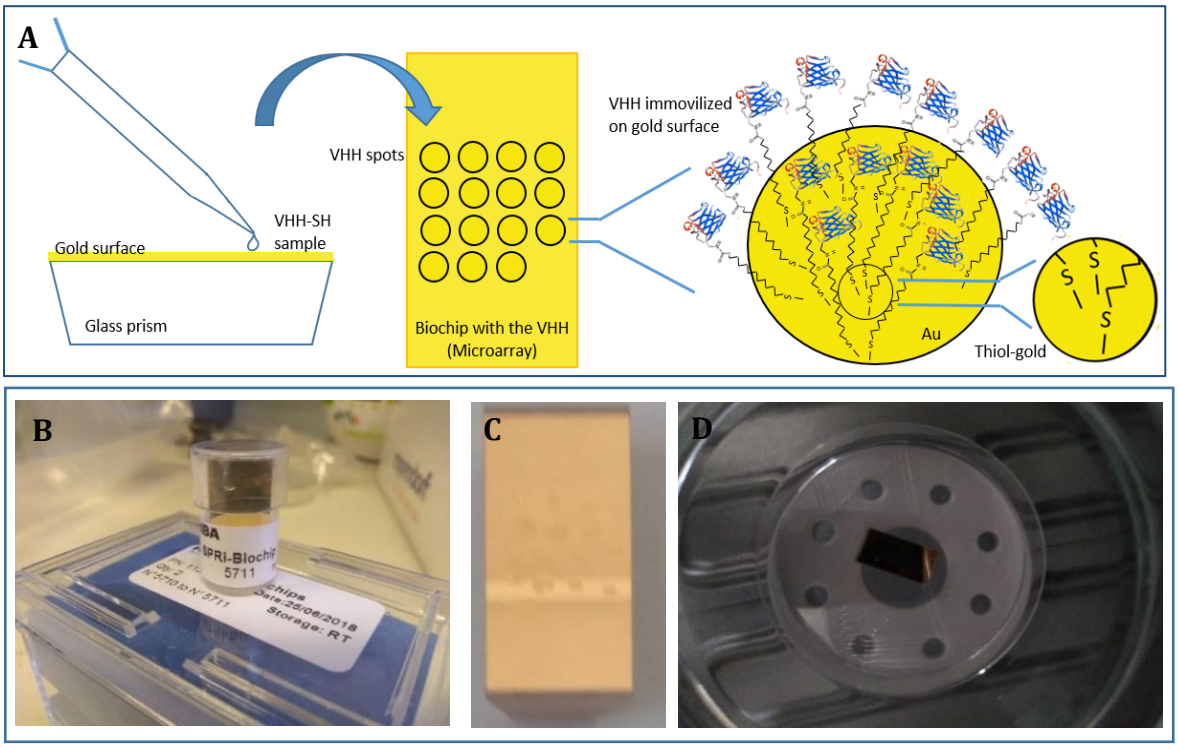


Figure 4.5- Preparation of biochip. a) Scheme of the immobilization of the VHH-SH over the gold surface of the biochip by the gold-thiol interactions. b) Commercial biochip. c) Biochip arrayed with the different anti-CD105 VHHs. d) BIochip in a chamber

The antigen, the human CD105 membrane protein, is expressed mainly on endothelial cells, in fact, first studies that reported its existence and importance were on the endothelium[158]. However, they are not the only cells that express it. For the assays, a non-adherent cell type was needed, meaning cells that stay in suspension, like blood cells. Monocytes/macrophages cells of different tissues and in different stages of cellular differentiation express CD105 but the monocytes of peripheral blood do not, unless their differentiation is induced *in vitro* with phorbol esters[159]. The SC cell line (ATCC® CRL-9855™) of monocyte/macrophage type was stablished from monocellular cells of human peripheral blood and grow as granular looking cells that are maintained free in suspension. These characteristics fit with what was required for the assays to recognise the antigen over cells that express it, and determine the capacity of the VHH immobilized on the biochip as capture agents and thus detection agents. The THP-1 cell line are monocyte type cells that do not express CD105, (except for induced differentiation) and are widely used in the laboratory due to their ease of culture, versatility and extensive knowledge of their handling, they grow in suspension. For these reasons, this cell line were selected as antigen-free control for the assays that evaluated the specificity of the VHH.

Each VHH proved to be efficient in recognizing recombinant CD105 in their free form (soluble CD105) with high specificity, as demonstrated by the previous ELISA (Chapter III, figure 3.19). However, recognizing CD105 over cells that naturally express it, as native non-free CD105 (membrane CD105), and doing the recognition immobilised in dependent form to its structure, leads to another level of efficiency that brings them closer to a real application in biosensors for the detection of CD105. To determine this or these VHH

according to their capabilities to recognise and capture their antigen while immobilised on a biochip, the 13 VHH were evaluated by SPRi. The SPRi instrument measured the antigen-antibody (CD105 over SC cells–VHH) interactions by detection of the minimal changes in the refractive index on the microarray surface and its immediate vicinity producing a sensogram at the same time the camera was capturing images during the entire process, which lasted approximately 3 hours.

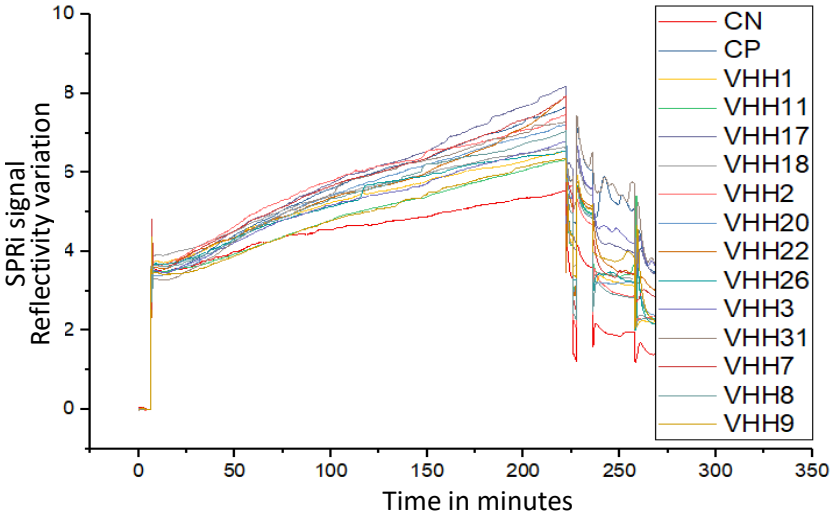
The SPRi curves of the VHH were plotted from the averages of the SPRi signals (light reflectivity variations) of each VHH throughout the assay (Figures 4.6a, 4.9a, 4.12a, 4.15a and 4.18a). The differential images, (images obtained after "subtracting" a reference image from the images captured during the process using the Image J program) were visualized and displayed for their analysis and cell counting (Figures 4.6b, 4.9b, 4.12b, 4.15b and 4.18b). At the cell counting, the grey level that provided a cell density (cells/area) according to the SPR signal (reflectivity variation) for each VHH at a given time was chosen. For this, in each assay, cell counting was performed for each VHH using different grey levels ranging from 0 to 1 from the differential images captured at 30, 60, 90, 90, 120 and 150 minutes after cell addition. The 120 minute (2 hour) differential image was selected as it allowed the best cell count and, the cell densities of each VHH (Figures 4.7a and b, 4.10 a and b, 4.13 a and b, 4.16 a and b, 4.19 a and b) were compared with their SPRi signals (reflectivity variation) obtained at the same time of assay.

First, visual comparison of the SPRi signal with the cell density obtained at grey levels 0.1, 0.15, 0.2, 0.3 and 0.4 was made individually (Figures 4.7c, 4.10c, 4.13c, 4.16c, 4.19c) and then collectively (Figures 4.7d, 4.10d, 4.13d, 4.16d, 4.19d). In addition, Pearson's correlation coefficient was used. This coefficient is a measurement of linear dependence between two quantitative random variables and is independent of the measurement scale of the variables unlike the covariance. Pearson´s coefficient measures the degree of relationship of two variables if they are quantitative and continuous. Values above 0 and below 1 reflect a positive correlation[160]. For the selection of the grey level to be used, Pearson's coefficient values equal to or greater than 0.8 were considered (Figures 4.7b, 4.10b, 4.13b, 4.16b and 4.19b). To determine the cell density for each VHH in the assays with SC cells, grey levels of 0.1 and 0.15 were chosen and for the assays with THP-1 cells was 0.3, according to visual comparative analysis and Pearson's coefficient (Figures 4.8, 4.11, 4.14, 4.17, 4.20).

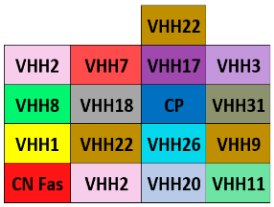
In the assays with cell line THP-1 the cells were counted from the differential image at 2.5 hours since that the differential image at 2 hours had values too low or 0 in order to be compared (Figure 4.21). This could be explained by the fact that THP-1 cells, which do not overexpress CD105 (negative antigen control) would have minimal interactions with the VHH and more time was required to capture an image that allowed distinction of the cells and this as a consequence of cell sedimentation by the course of time and not necessarily due to cell capture. In a similar way, the cell count of "control gold", i.e. the control without an immobilized capture agent (VHH), the gold surface free of VHH) few cells or none cell were counted in images at the 2 hour but at 2.5h or 3h there is a increase due to cell sedimentation.

SPRi assay I with SC cells

A



B



C

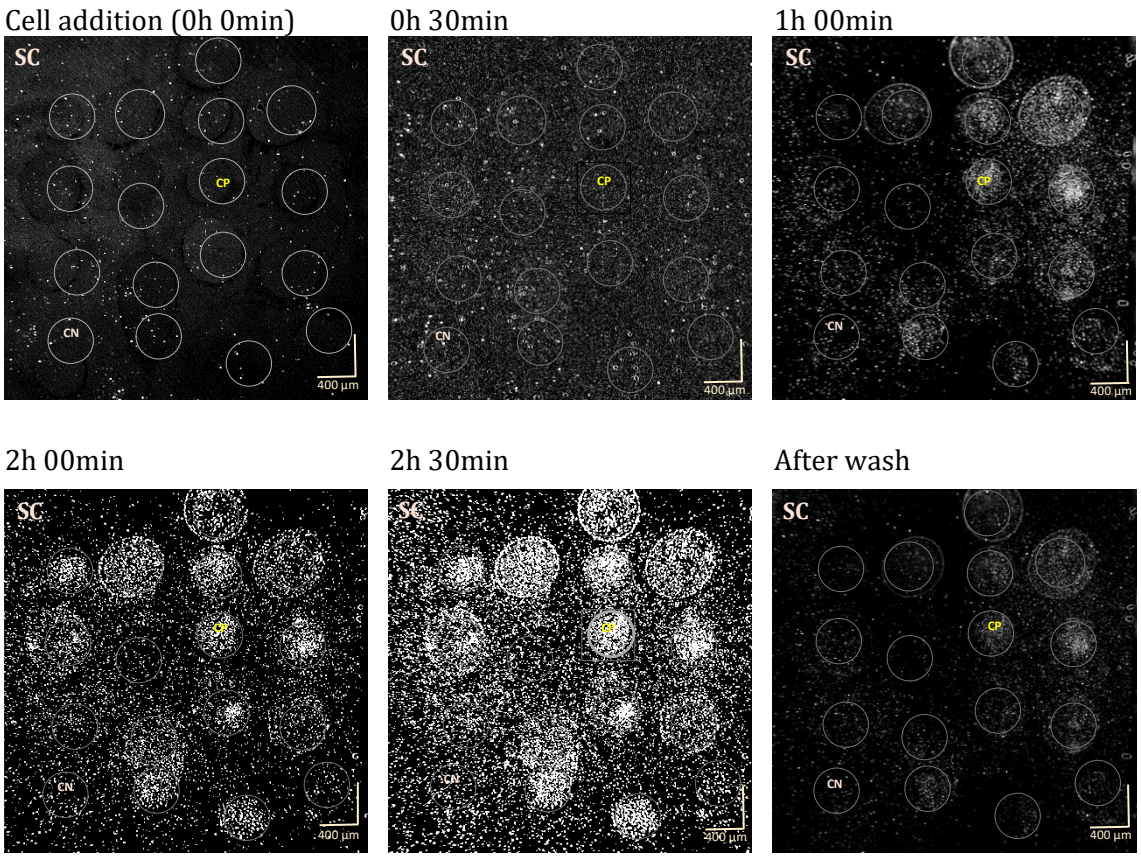
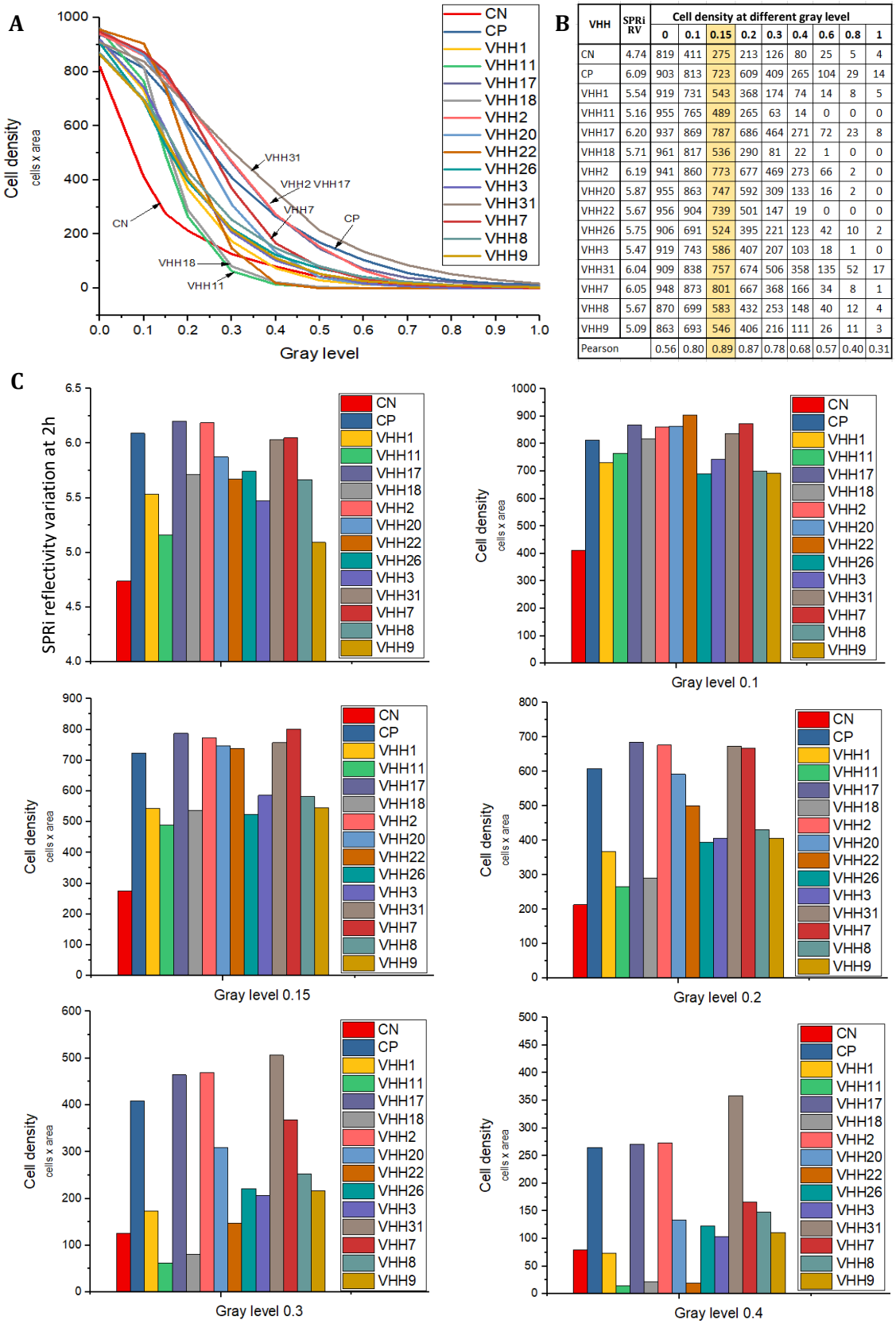


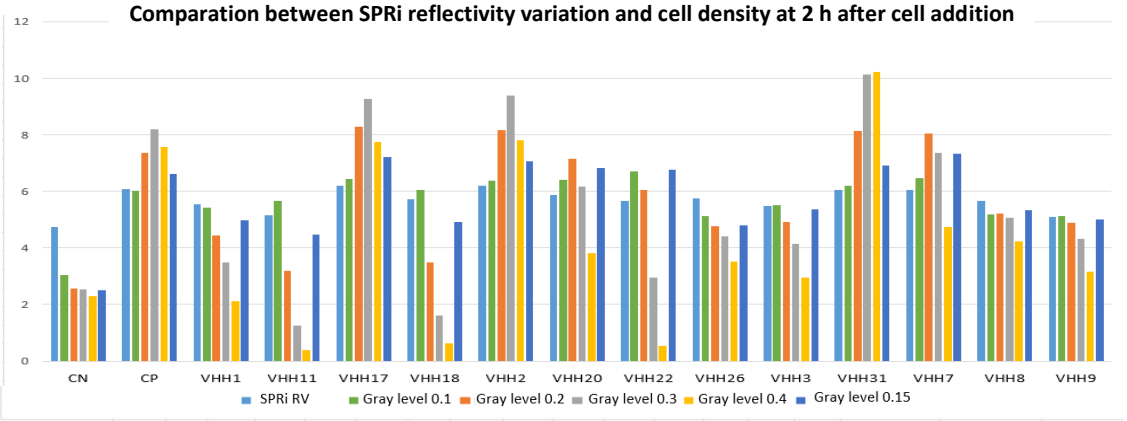
Figure 4.6- SPRi assay I of the 13 recombinant anti-CD105 VHH proteins (100 ng/µl or 6.6 µM) at different incubation times with the SC cell line (CD105 (+)) (2 x 10^6 cells). a) SPRi curves monitoring reflectivity variations upon the incubation time of each VHH with the cells. b) Distribution of the VHH proteins on the biochip (each VHH shows a different colour). c) Differential images at different times (0, 30, 60, 120 and 150 minutes) of the assay . 0.5 ul of each VHH sample was applied on the biochip, that is. 50 ng or 3.33 µmoles of each VHH per spot. **CN**: negative control (VHH anti Fas2), **CP**: positive control (EngVHH17).



VHH	SPRi RV	Cell density at different gray level								
		0	0.1	0.15	0.2	0.3	0.4	0.6	0.8	1
CN	4.74	819	411	275	213	126	80	25	5	4
CP	6.09	903	813	723	609	409	265	104	29	14
VHH1	5.54	919	731	543	368	174	74	14	8	5
VHH11	5.16	955	765	489	265	63	14	0	0	0
VHH17	6.20	937	869	787	686	464	271	72	23	8
VHH18	5.71	961	817	536	290	81	22	1	0	0
VHH2	6.19	941	860	773	677	469	273	66	2	0
VHH20	5.87	955	863	747	592	309	133	16	2	0
VHH22	5.67	956	904	739	501	147	19	0	0	0
VHH26	5.75	906	691	524	395	221	123	42	10	2
VHH3	5.47	919	743	586	407	207	103	18	1	0
VHH31	6.04	909	838	757	674	506	358	135	52	17
VHH7	6.05	948	873	801	667	368	166	34	8	1
VHH8	5.67	870	699	583	432	253	148	40	12	4
VHH9	5.09	863	693	546	406	216	111	26	11	3
Pearson		0.56	0.80	0.89	0.87	0.78	0.68	0.57	0.40	0.31

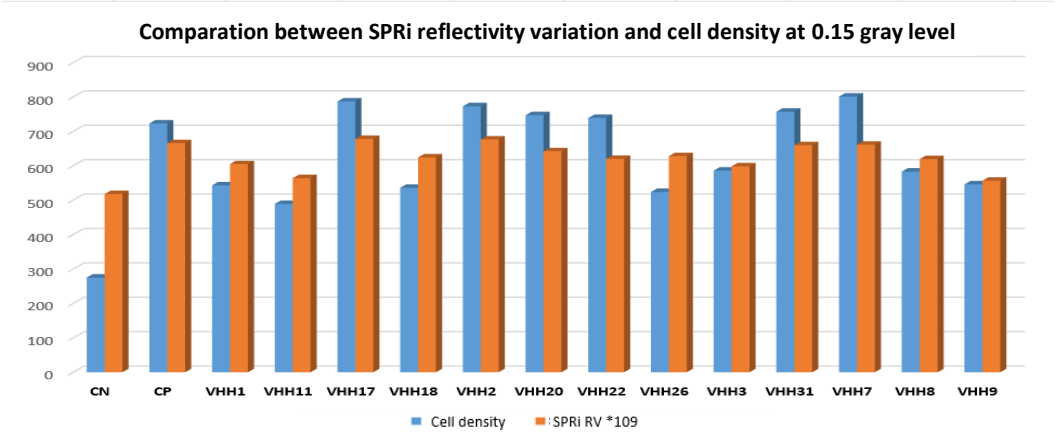
Figure 4.7- Determination of cell density for each VHH by comparison with SPRi reflectivity variation (SPRi RV) 2 hours after adding the cells in SPRi assay I with SC cells. a) Cell density curve of each VHH at different gray levels, b) Pearson's correlation coefficient, c) y d) Visual comparison between SPRi RV and cell density different at gray levels (0.1 to 0.4), individually (c) in group (d). The area used was 477.09 µm x 477.09 µm (31x31 pixeles). **CN**: negative control (VHH anti Fas2), **CP**: positive control (EngVHH17)

D



	CN	CP	VHH1	VHH11	VHH17	VHH18	VHH2	VHH20	VHH22	VHH26	VHH3	VHH31	VHH7	VHH8	VHH9
SPRi RV	4.74	6.09	5.54	5.16	6.20	5.71	6.19	5.87	5.67	5.75	5.47	6.04	6.05	5.67	5.09
Gray level 0.1/135	3.0	6.0	5.4	5.7	6.4	6.1	6.4	6.4	6.7	5.1	5.5	6.2	6.5	5.2	5.1
Gray level 0.2/83	2.6	7.3	4.4	3.2	8.3	3.5	8.2	7.1	6.0	4.8	4.9	8.1	8.0	5.2	4.9
Gray level 0.3/50	2.5	8.2	3.5	1.3	9.3	1.6	9.4	6.2	2.9	4.4	4.1	10.1	7.4	5.1	4.3
Gray level 0.4/35	2.3	7.6	2.1	0.4	7.7	0.6	7.8	3.8	0.5	3.5	2.9	10.2	4.7	4.2	3.2
Gray level 0.15/109	2.5	6.6	5.0	4.5	7.2	4.9	7.1	6.8	6.8	4.8	5.4	6.9	7.3	5.3	5.0

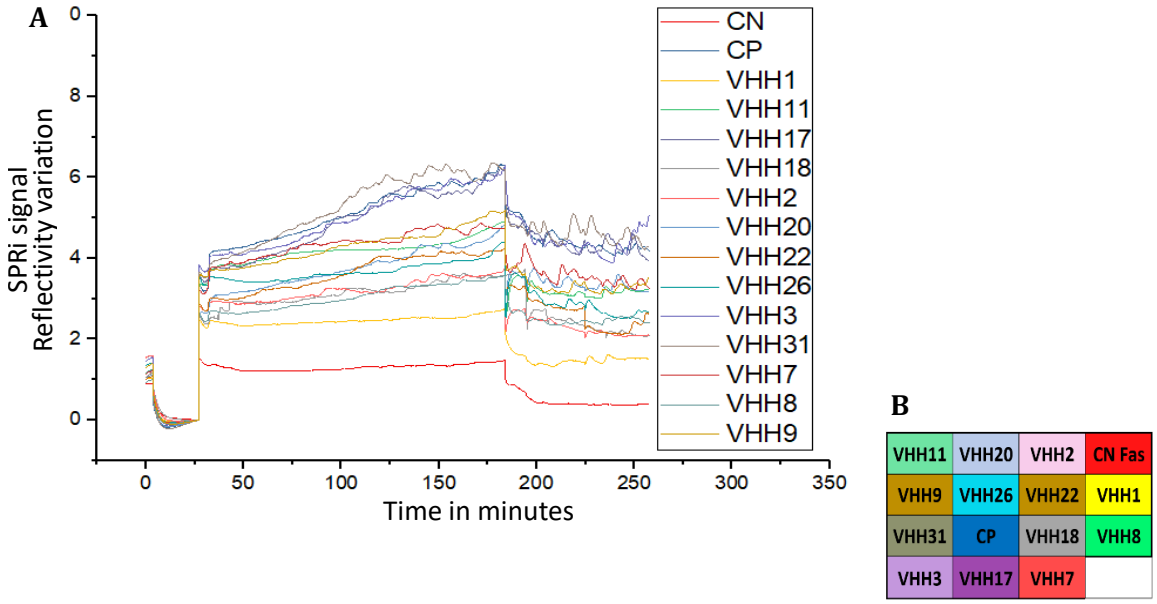
Figure 4.7- (continuation)



VHH	SPRi RV	Cell density at 0.15 gray level	SPRi RV *109
CN	4.74	275	517.86
CP	6.09	723	665.67
VHH1	5.54	543	605.06
VHH11	5.16	489	564.08
VHH17	6.20	787	677.97
VHH18	5.71	536	624.30
VHH2	6.19	773	676.27
VHH20	5.87	747	642.10
VHH22	5.67	739	619.96
VHH26	5.75	524	627.95
VHH3	5.47	586	598.32
VHH31	6.04	757	659.71
VHH7	6.05	801	661.11
VHH8	5.67	583	619.30
VHH9	5.09	546	556.66

Figure 4.8- Cell density for each VHH obtained from differential images at 0.15 grey level in the SPRi assay I with SC cells.

SPRi assay II with SC cells



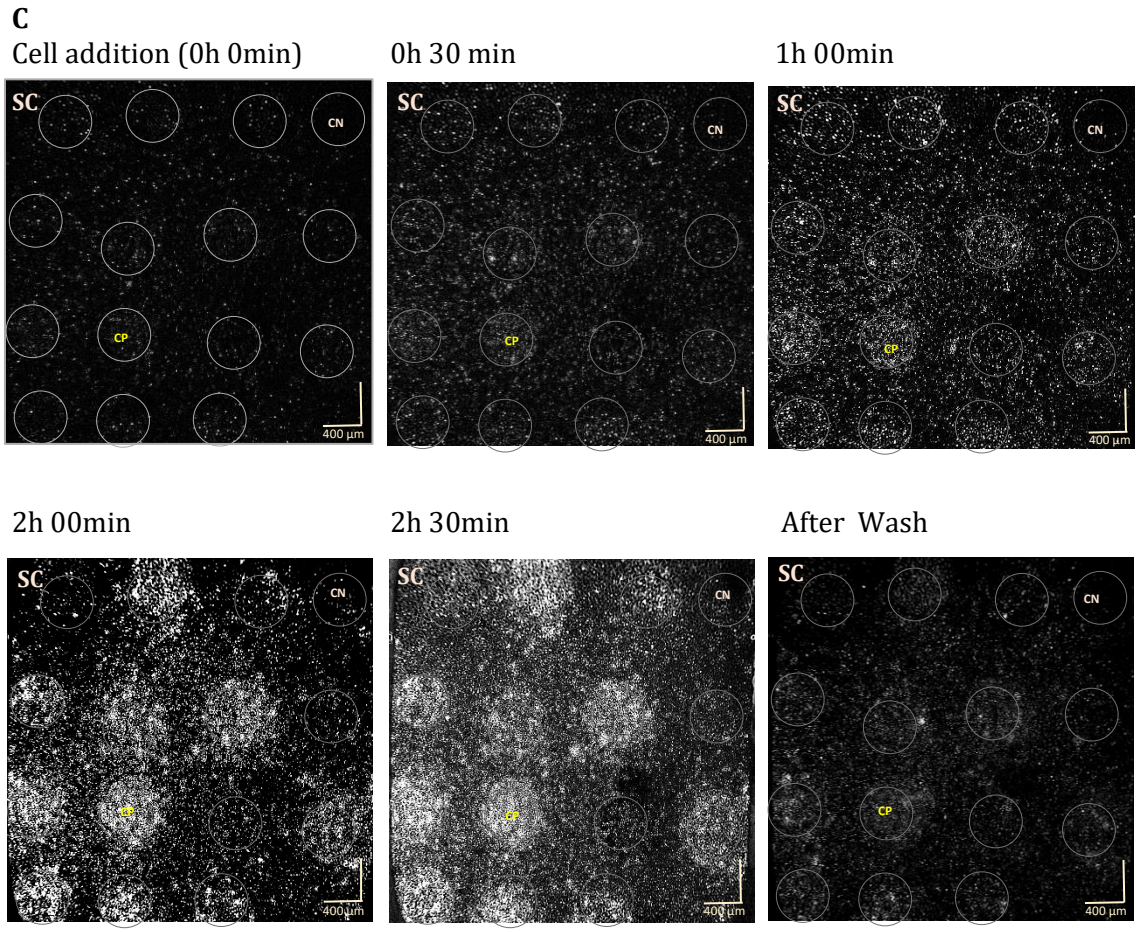
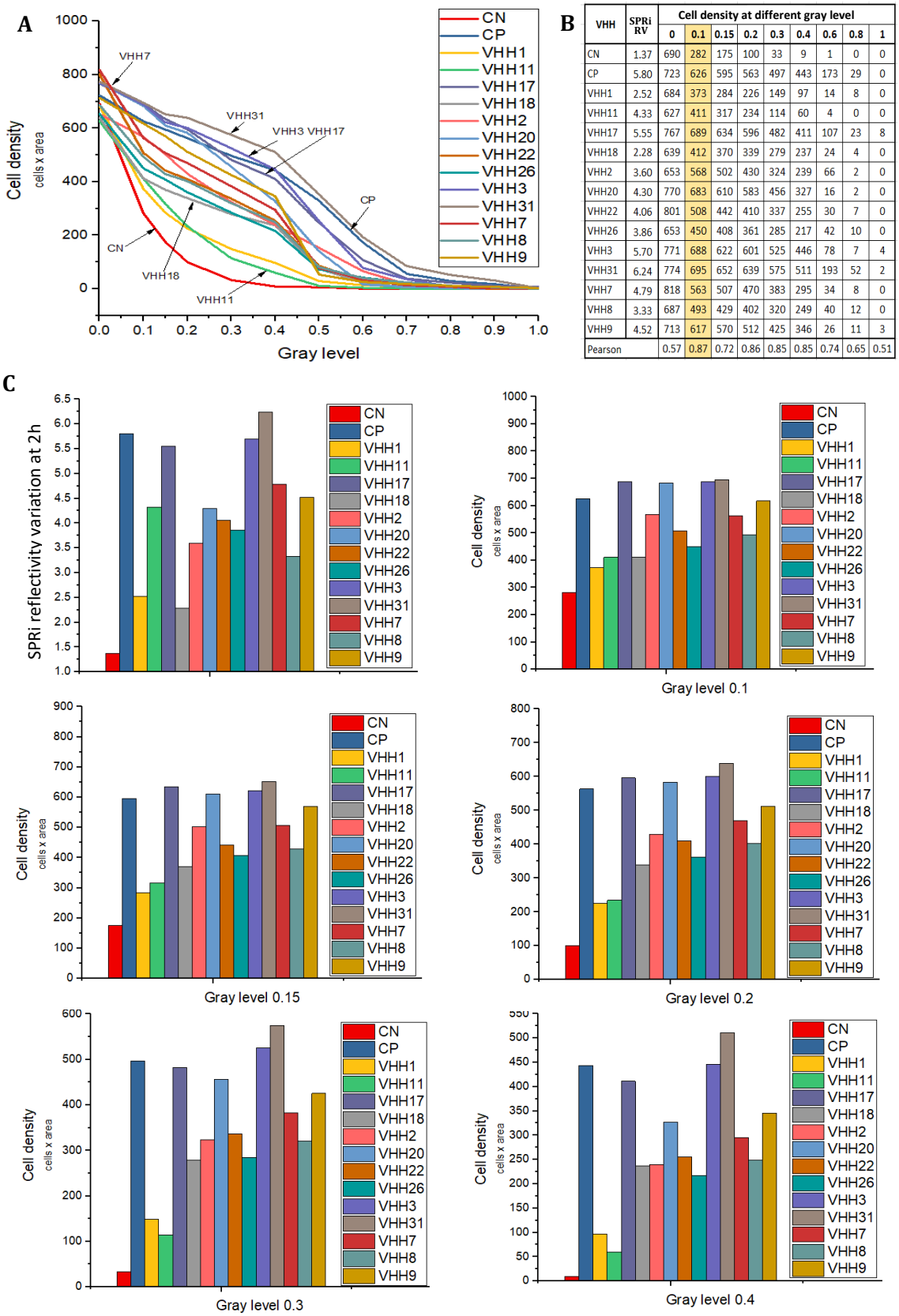
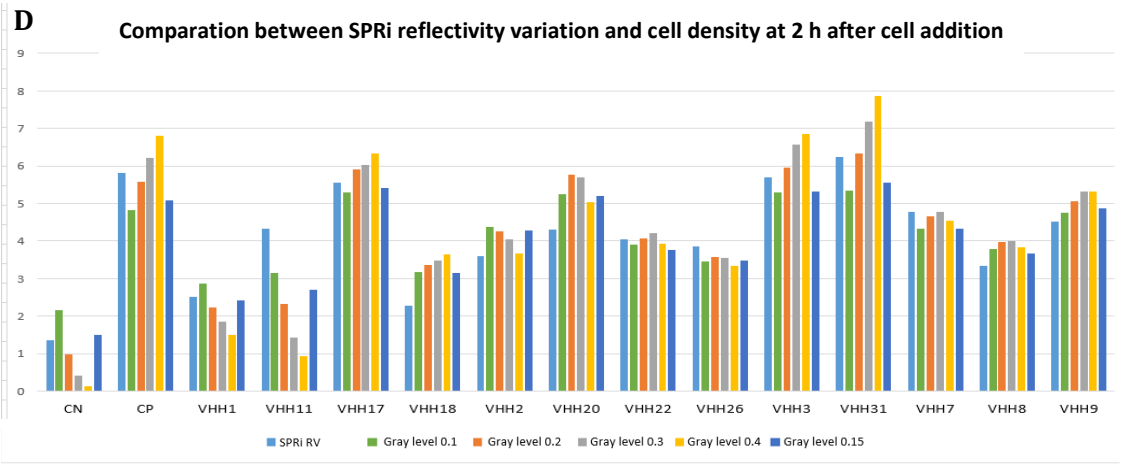


Figure 4.9- SPRi assay II of the 13 recombinant anti-CD105 VHH proteins (100 ng/µl or 6.6 µM) at different incubation times with the SC cell line (CD105 (+)) (2 x 10^6 cells). a) SPRi curves monitoring reflectivity variations upon the incubation time of each VHH with the cells. b) Distribution of the VHH proteins on the biochip (each VHH shows a different colour). c) Differential images at different times (0, 30, 60, 120 and 150 minutes) of the assay . 0.5 ul of each VHH sample was applied on the biochip, that is. 50 ng or 3.33 µmoles of each VHH per spot. **CN**: negative control (VHH anti Fas2), **CP**: positive control (EngVHH17).



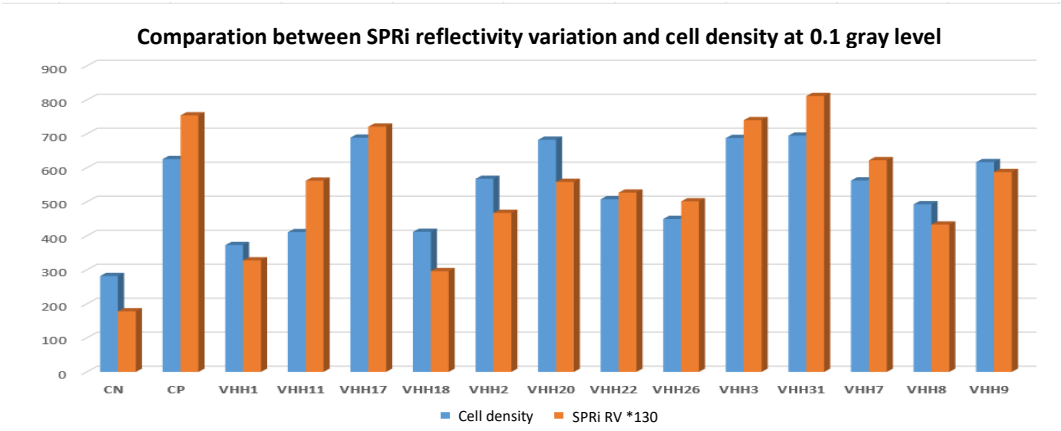
VHH	SPRi RV	Cell density at different gray level								
		0	0.1	0.15	0.2	0.3	0.4	0.6	0.8	1
CN	1.37	690	282	175	100	33	9	1	0	0
CP	5.80	723	626	595	563	497	443	173	29	0
VHH1	2.52	684	373	284	226	149	97	14	8	0
VHH11	4.33	627	411	317	234	114	60	4	0	0
VHH17	5.55	767	689	634	596	482	411	107	23	8
VHH18	2.28	639	412	370	339	279	237	24	4	0
VHH2	3.60	653	568	502	430	324	239	66	2	0
VHH20	4.30	770	683	610	583	456	327	16	2	0
VHH22	4.06	801	508	442	410	337	255	30	7	0
VHH26	3.86	653	450	408	361	285	217	42	10	0
VHH3	5.70	771	688	622	601	525	446	78	7	4
VHH31	6.24	774	695	652	639	575	511	193	52	2
VHH7	4.79	818	563	507	470	383	295	34	8	0
VHH8	3.33	687	493	429	402	320	249	40	12	0
VHH9	4.52	713	617	570	512	425	346	26	11	3
Pearson		0.57	0.87	0.72	0.86	0.85	0.85	0.74	0.65	0.51

Figure 4.10- Determination of cell density for each VHH by comparison with SPRi reflectivity variation (SPRi RV) 2 hours after adding the cells in SPRi assay II with SC cells. a) Cell density curve of each VHH at different gray levels, b) Pearson's correlation coefficient, c) y d) Visual comparison between SPRi RV and cell density different at gray levels (0.1 to 0.4), individually (c) in group (d). The area used was 477.1 µm x 477.1µm (31x31 pixeles). **CN**: negative control (VHH anti Fas2), **CP**: positive control (EngVHH17)



	CN	CP	VHH1	VHH11	VHH17	VHH18	VHH2	VHH20	VHH22	VHH26	VHH3	VHH31	VHH7	VHH8	VHH9
SPRi RV	1.37	5.80	2.52	4.33	5.55	2.28	3.60	4.30	4.06	3.86	5.70	6.24	4.79	3.33	4.52
Gray level 0.1/130	2.2	4.8	2.9	3.2	5.3	3.2	4.4	5.3	3.9	3.5	5.3	5.3	4.3	3.8	4.7
Gray level 0.2/101	1.0	5.6	2.2	2.3	5.9	3.4	4.3	5.8	4.1	3.6	6.0	6.3	4.7	4.0	5.1
Gray level 0.3/80	0.4	6.2	1.9	1.4	6.0	3.5	4.1	5.7	4.2	3.6	6.6	7.2	4.8	4.0	5.3
Gray level 0.4/65	0.1	6.8	1.5	0.9	6.3	3.6	3.7	5.0	3.9	3.3	6.9	7.9	4.5	3.8	5.3
Gray level 0.15/117	1.5	5.1	2.4	2.7	5.4	3.2	4.3	5.2	3.8	3.5	5.3	5.6	4.3	3.7	4.9

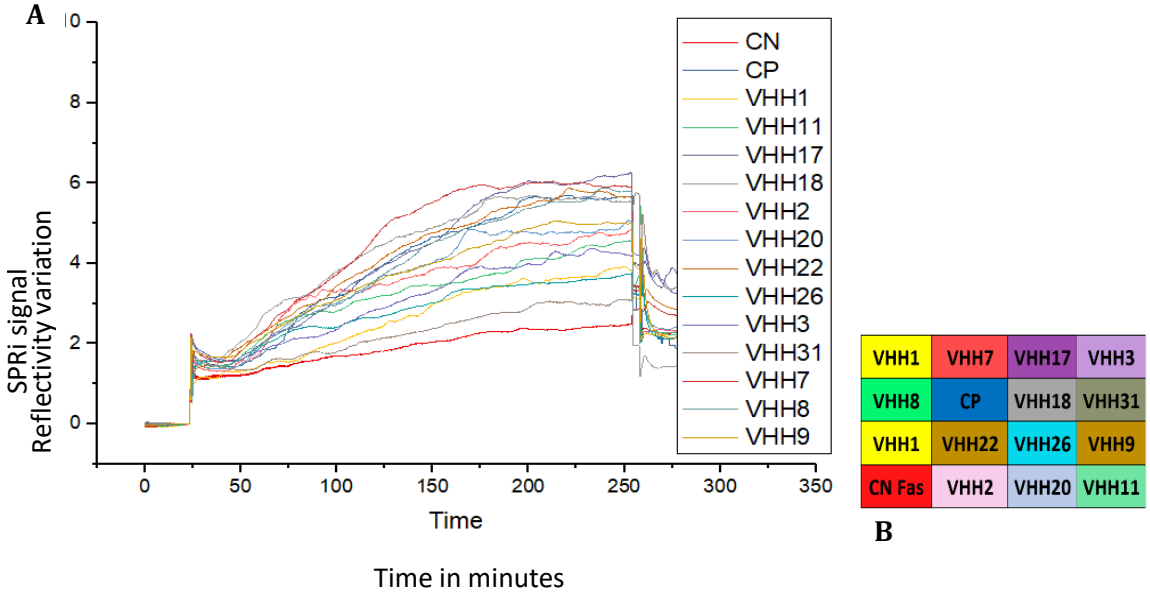
Figura 4.10- (continuation)



VHH	SPRi RV	Cell density at 0.1 gray level	SPRi RV *130
CN	1.37	282	178.01
CP	5.80	626	754.56
VHH1	2.52	373	328.09
VHH11	4.33	411	562.46
VHH17	5.55	689	721.40
VHH18	2.28	412	296.26
VHH2	3.60	568	467.66
VHH20	4.30	683	558.73
VHH22	4.06	508	527.29
VHH26	3.86	450	501.64
VHH3	5.70	688	740.52
VHH31	6.24	695	811.46
VHH7	4.79	563	622.56
VHH8	3.33	493	433.28
VHH9	4.52	617	587.86

Figure 4.11- Cell density for each VHH obtained from differential images at 0.1 grey level in the SPRi assay II with SC cells

SPRi assay III with SC cells



C

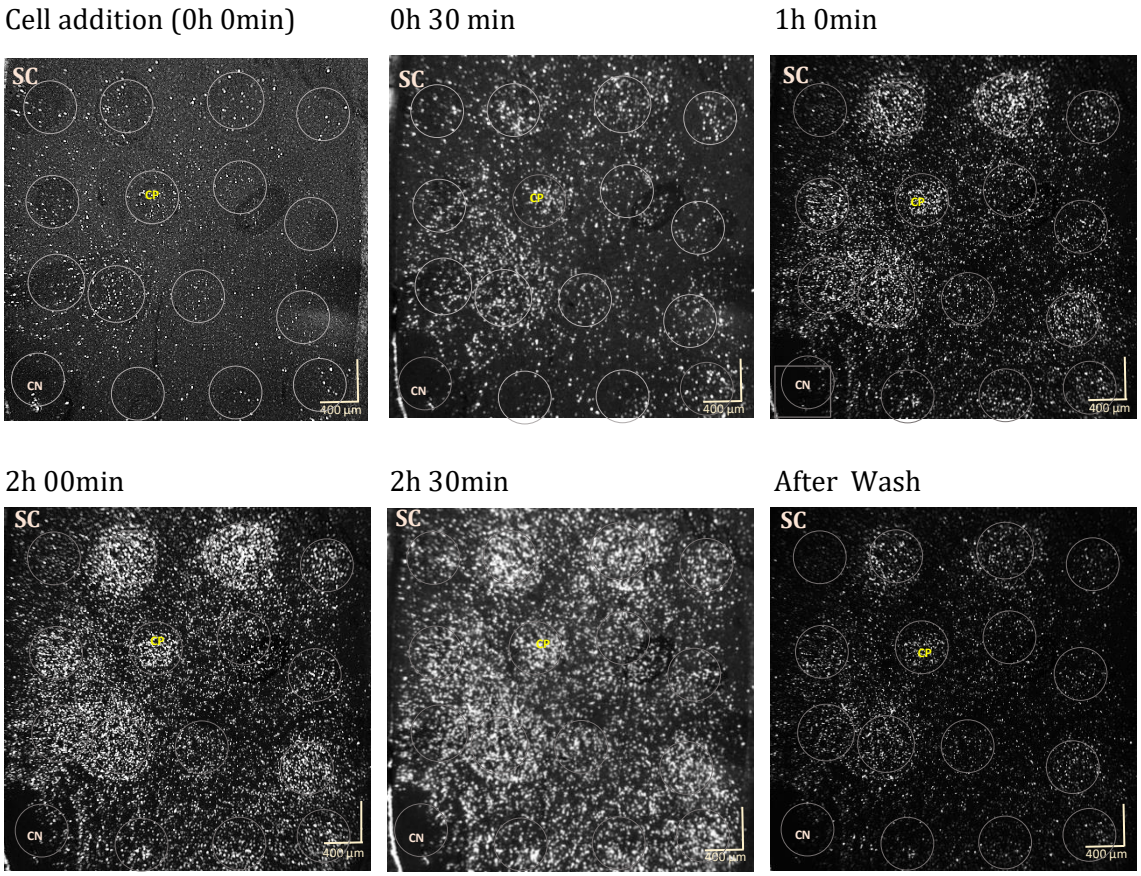
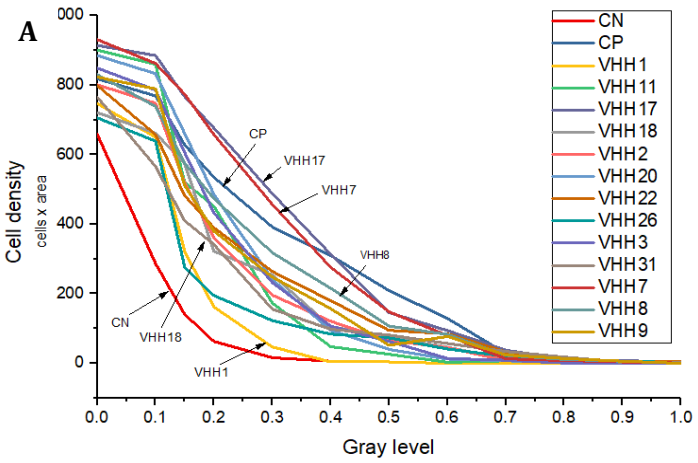


Figure 4.12- SPRi assay III of the 13 recombinant anti-CD105 VHH proteins (100 ng/µl or 6.6 µM) at different incubation times with the SC cell line (CD105 (+)) (2 x 10^6 cells). a) SPRi curves monitoring reflectivity variations upon the incubation time of each VHH with the cells. b) Distribution of the VHH proteins on the biochip (each VHH shows a different colour). c) Differential images at different times (0, 30, 60, 120 and 150 minutes) of the assay. 0.5 ul of each VHH sample was applied on the biochip, that is. 50 ng or 3.33 µmoles of each VHH per spot. **CN**: negative control (VHH anti Fas2), **CP**: positive control (EngVHH17).



B

VHH	SPRi RV	Cell density at different gray level								
		0	0.1	0.15	0.2	0.3	0.4	0.6	0.8	1
CN	2.00	659	287	141	63	16	6	0	0	0
CP	4.74	818	769	628	535	392	310	129	12	0
VHH1	2.71	747	650	321	162	47	4	0	0	0
VHH11	3.44	901	860	520	452	173	48	3	3	0
VHH17	4.37	914	885	767	677	488	314	94	11	3
VHH18	4.79	721	662	574	322	254	106	46	2	0
VHH2	3.71	801	748	519	361	196	121	45	2	0
VHH20	3.93	885	833	661	486	241	95	12	2	0
VHH22	4.61	799	658	482	390	264	180	85	16	4
VHH26	2.94	706	639	277	195	123	84	41	12	2
VHH3	3.22	849	786	606	433	232	107	14	1	0
VHH31	2.30	766	566	410	342	156	98	57	18	2
VHH7	5.37	931	862	774	658	457	276	80	8	1
VHH8	4.33	830	739	575	474	318	216	83	11	0
VHH9	3.91	823	789	511	380	248	157	77	14	0
Pearson		0.62	0.65	0.81	0.75	0.85	0.77	0.65	0.21	0.13

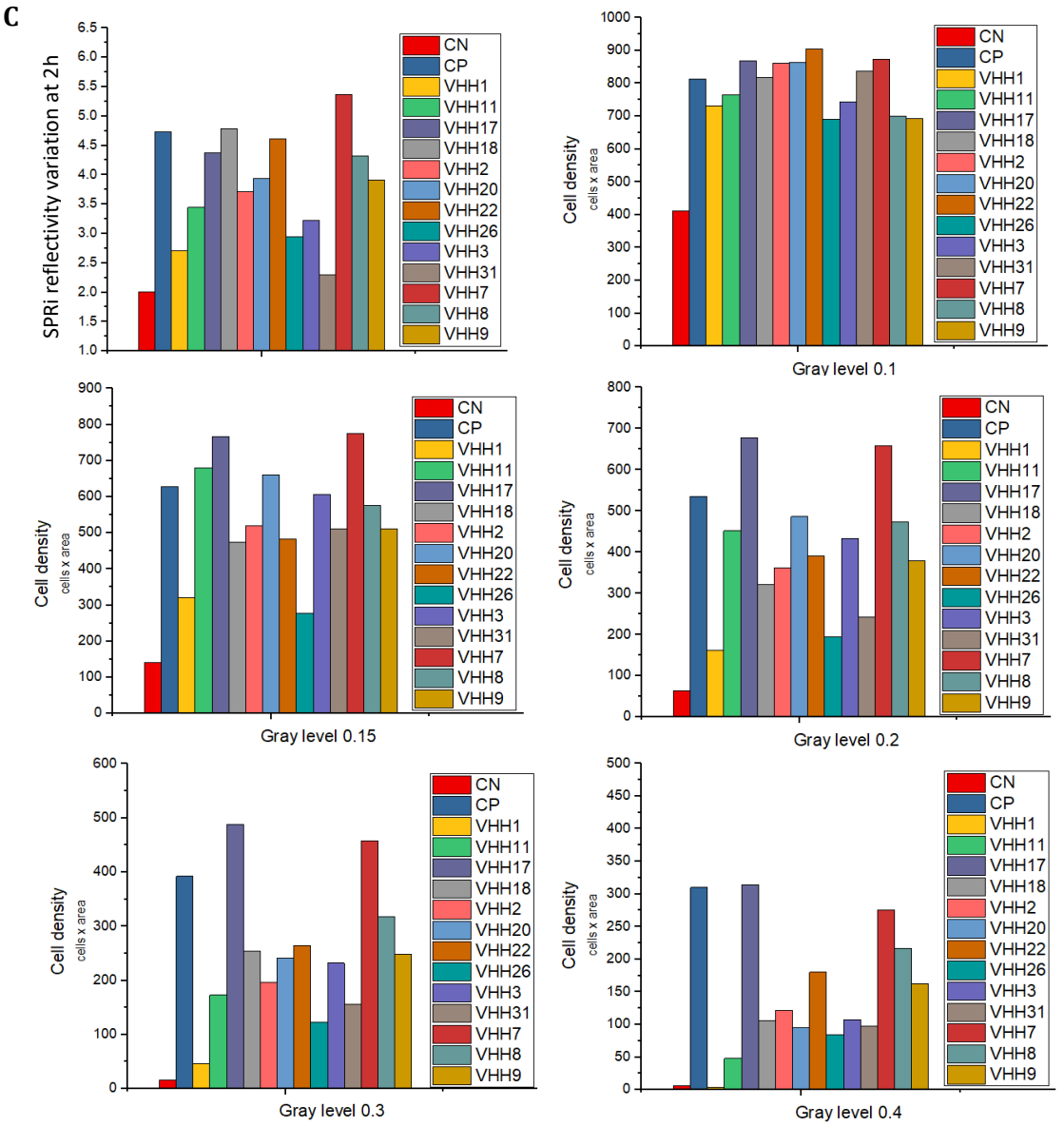
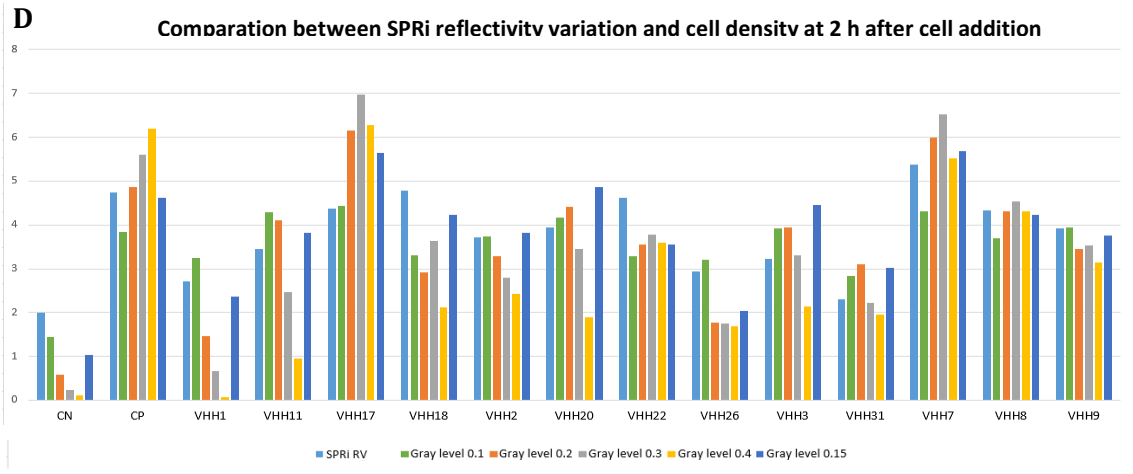
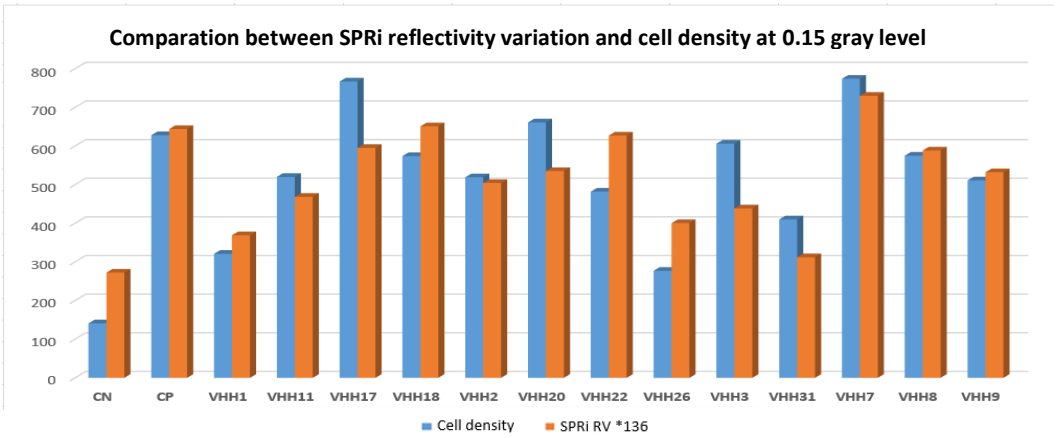


Figure 4.13- Determination of cell density for each VHH by comparison with SPRi reflectivity variation (SPRi RV) 2 hours after adding the cells in SPRi assay III with SC cells. a) Cell density curve of each VHH at different gray levels, b) Pearson's correlation coefficient, c) y d) Visual comparison between SPRi RV and cell density different at gray levels (0.1 to 0.4), individually (c) in group (d). The area used was 477.09 um x 477.09 um (31x31 pixeles). **CN**: negative control (VHH anti Fas2), **CP**: positive control (EngVHH17)



	CN	CP	VHH1	VHH11	VHH17	VHH18	VHH2	VHH20	VHH22	VHH26	VHH3	VHH31	VHH7	VHH8	VHH9
SPRi RV	2.00	4.74	2.71	3.44	4.37	4.79	3.71	3.93	4.61	2.94	3.22	2.30	5.37	4.33	3.91
Gray level 0.1/200	1.4	3.8	3.3	4.3	4.4	3.3	3.7	4.2	3.3	3.2	3.9	2.8	4.3	3.7	3.9
Gray level 0.2/110	0.6	4.9	1.5	4.1	6.2	2.9	3.3	4.4	3.5	1.8	3.9	3.1	6.0	4.3	3.5
Gray level 0.3/70	0.2	5.6	0.7	2.5	7.0	3.6	2.8	3.4	3.8	1.8	3.3	2.2	6.5	4.5	3.5
Gray level 0.4/50	0.1	6.2	0.1	1.0	6.3	2.1	2.4	1.9	3.6	1.7	2.1	2.0	5.5	4.3	3.1
Gray level 0.15/136	1.0	4.6	2.4	3.8	5.6	4.2	3.8	4.9	3.5	2.0	4.5	3.0	5.7	4.2	3.8

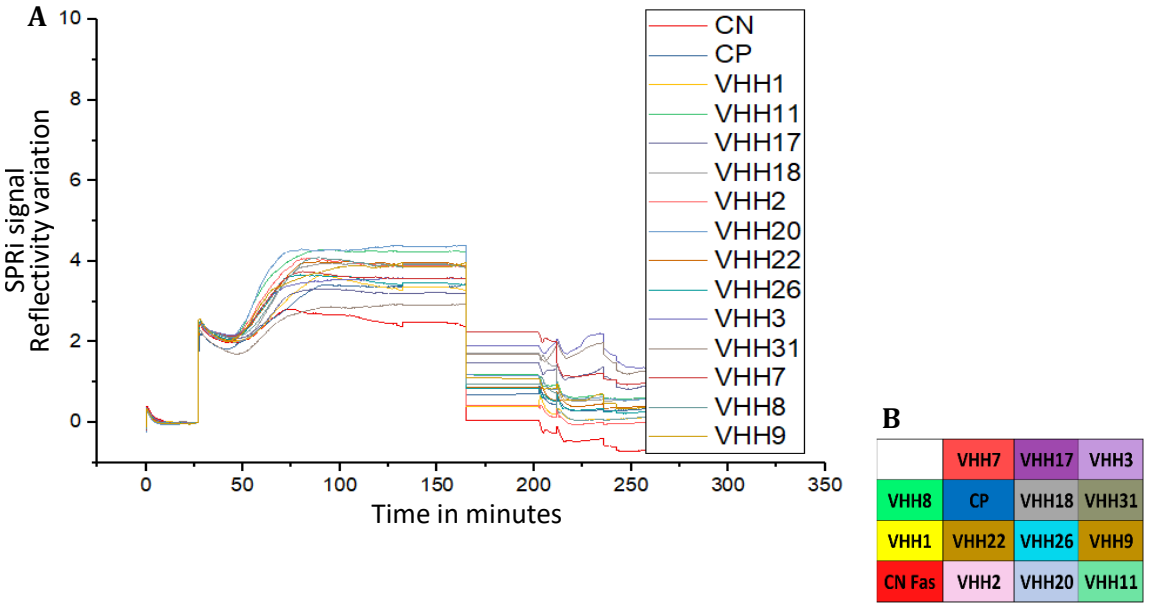
Figure 4.13- (continuation)



VHH	SPRi RV	Cell density at 0.15 gray level	SPRi RV *136
CN	2.00	141	272.35
CP	4.74	628	643.97
VHH1	2.71	321	369.02
VHH11	3.44	520	468.47
VHH17	4.37	767	594.96
VHH18	4.79	574	650.98
VHH2	3.71	519	504.26
VHH20	3.93	661	535.13
VHH22	4.61	482	627.11
VHH26	2.94	277	400.38
VHH3	3.22	606	438.46
VHH31	2.30	410	312.16
VHH7	5.37	774	729.99
VHH8	4.33	575	588.46
VHH9	3.91	511	532.20

Figure 4.14- Cell density for each VHH obtained from differential images at 0.15 grey level in the SPRi assay III with SC cells

SPRi assay I with THP-1 cells



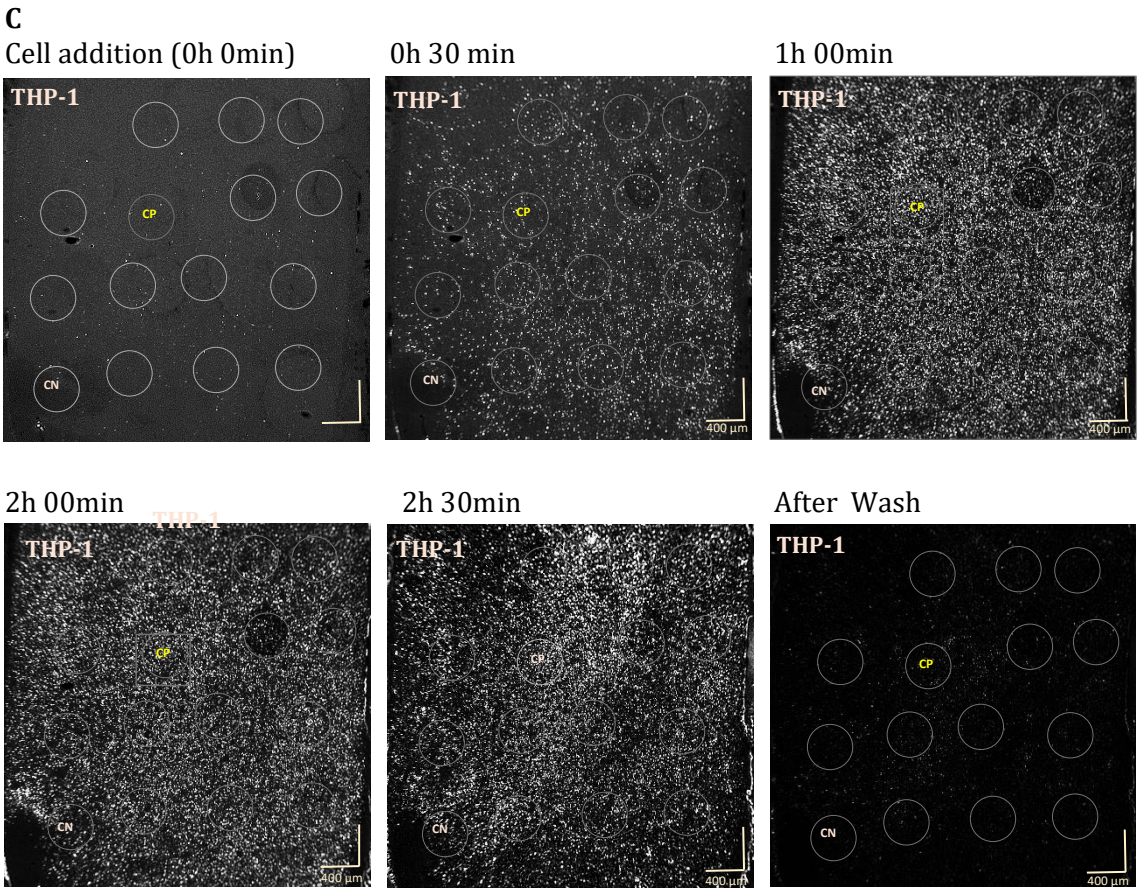
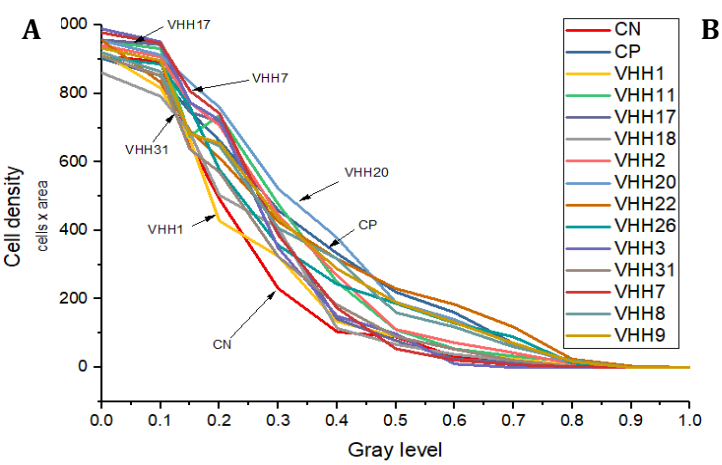


Figure 4.15- SPRi assay I of the 13 recombinant anti-CD105 VHH proteins (100 ng/µl or 6.6 µM) at different incubation times with the THP-1 cell line (CD105 (-)) (2 x 10^6 cells). a) SPRi curves monitoring reflectivity variations upon the incubation time of each VHH with the cells. b) Distribution of the VHH proteins on the biochip (each VHH shows a different colour). c) Differential images at different times (0, 30, 60, 120 and 150 minutes) of the assay. 0.5 ul of each VHH sample was applied on the biochip, that is. 50 ng or 3.33 µmoles of each VHH per spot. **CN**: negative control (VHH anti Fas2), **CP**: positive control (EngVHH17)..



B

VHH	SPRi RV	Cell density at different gray level								
		0	0.1	0.15	0.2	0.3	0.4	0.6	0.8	1
CN	2.27	910	893	656	492	231	104	32	7	0
CP	3.54	903	852	747	663	459	333	160	14	0
VHH1	3.46	919	816	664	428	325	134	53	5	0
VHH11	4.35	955	931	676	736	477	249	54	17	0
VHH17	3.43	957	943	746	718	403	142	24	3	0
VHH18	4.31	861	792	686	503	408	114	38	2	0
VHH2	3.75	941	907	774	707	445	269	72	11	0
VHH20	4.76	955	913	833	760	522	379	140	13	0
VHH22	3.95	956	833	692	613	426	317	184	24	0
VHH26	3.42	906	886	760	578	357	243	129	13	0
VHH3	3.88	989	951	775	723	349	150	10	0	0
VHH31	3.25	909	853	637	571	324	183	53	3	0
VHH7	4.16	978	947	808	742	387	174	21	2	0
VHH8	4.01	920	864	686	647	407	317	118	21	0
VHH9	3.89	933	896	682	656	436	288	131	17	0
Pearson		0.74	0.03	0.06	0.41	0.82	0.48	0.22	0.1	0

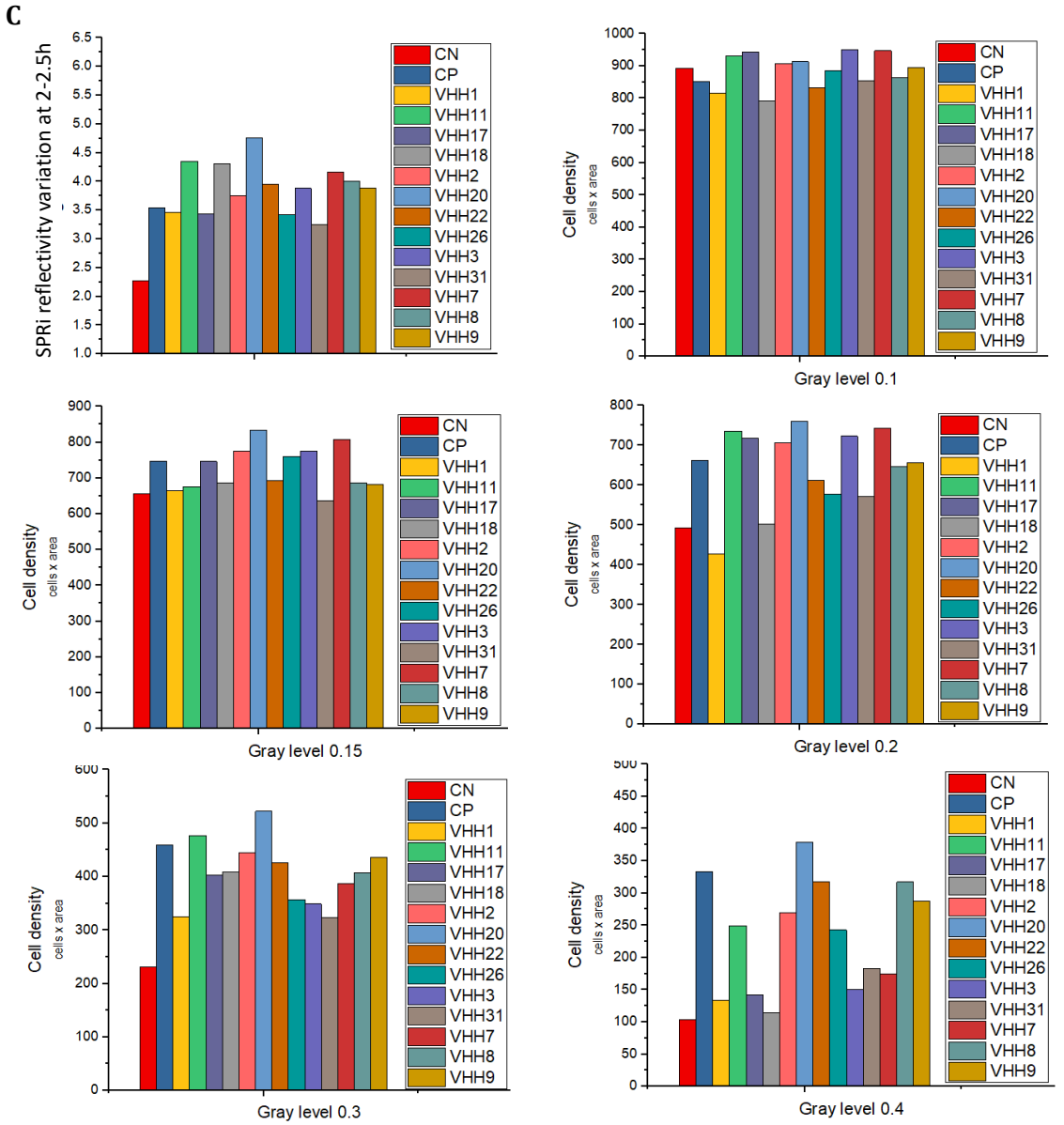
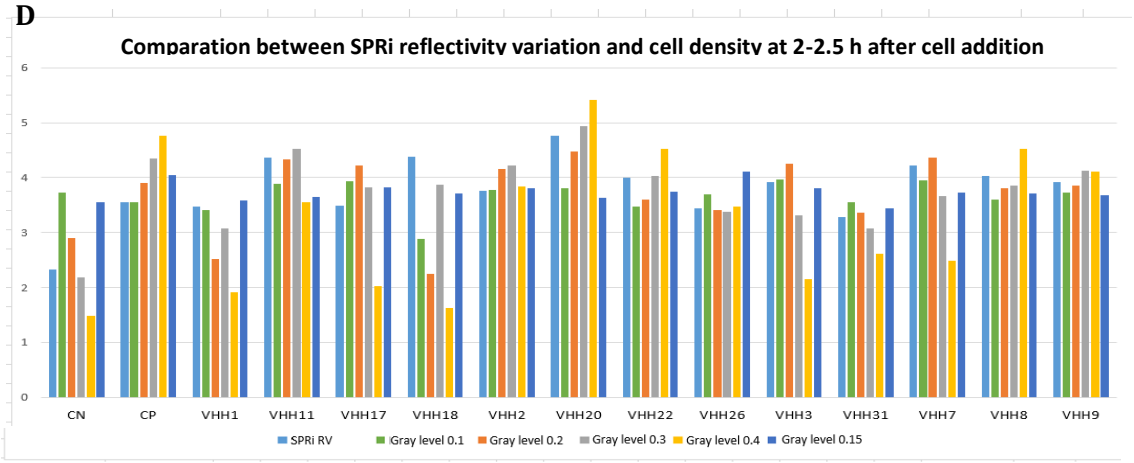


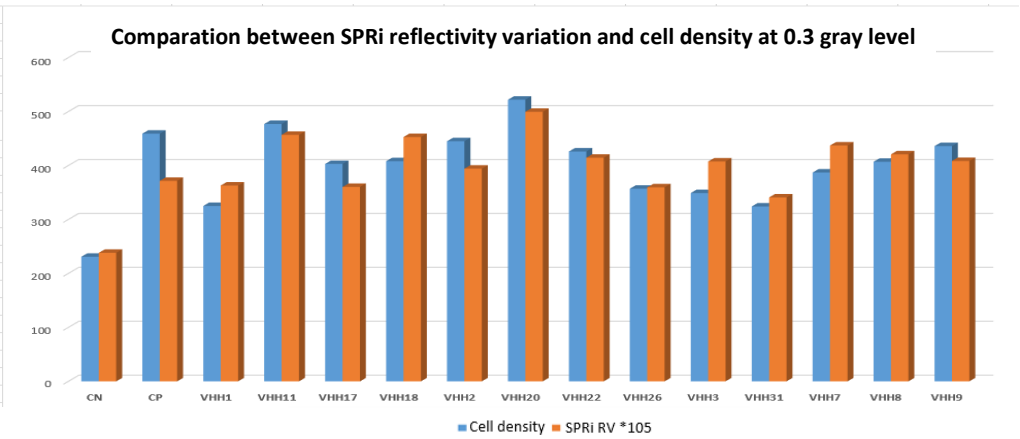
Figure 4.16- Determination of cell density for each VHH by comparison with SPRi reflectivity variation (SPRi RV) 2 hours after adding the cells in SPRi assay I with THP-1 cells. a) Cell density curve of each VHH at different gray levels, b) Pearson's correlation coefficient, c) y d) Visual comparison between SPRi RV and cell density different at gray levels (0.1 to 0.4), individually (c) in group (d). The area used was 477.09 µm x 477.09 µm (31x31 pixeles). **CN**: negative control (VHH anti Fas2), **CP**: positive control (EngVHH17).

D



	CN	CP	VHH1	VHH11	VHH17	VHH18	VHH2	VHH20	VHH22	VHH26	VHH3	VHH31	VHH7	VHH8	VHH9
SPRi RV	2.32	3.55	3.47	4.37	3.49	4.38	3.76	4.76	4.00	3.44	3.91	3.28	4.22	4.02	3.91
Gray level 0.1/240	3.7	3.6	3.4	3.9	3.9	3.3	3.8	3.8	3.5	3.7	4.0	3.6	3.9	3.6	3.7
Gray level 0.2/170	2.9	3.9	2.5	4.3	4.2	3.0	4.2	4.5	3.6	3.4	4.3	3.4	4.4	3.8	3.9
Gray level 0.3/105	2.2	4.3	3.1	4.5	3.8	3.9	4.2	4.9	4.0	3.4	3.3	3.1	3.7	3.9	4.1
Gray level 0.4/70	1.5	4.8	1.9	3.6	2.0	1.6	3.8	5.4	4.5	3.5	2.1	2.6	2.5	4.5	4.1
Gray level 0.15/185	3.5	4.0	3.6	3.7	4.0	3.7	4.2	4.5	3.7	4.1	4.2	3.4	4.4	3.7	3.7

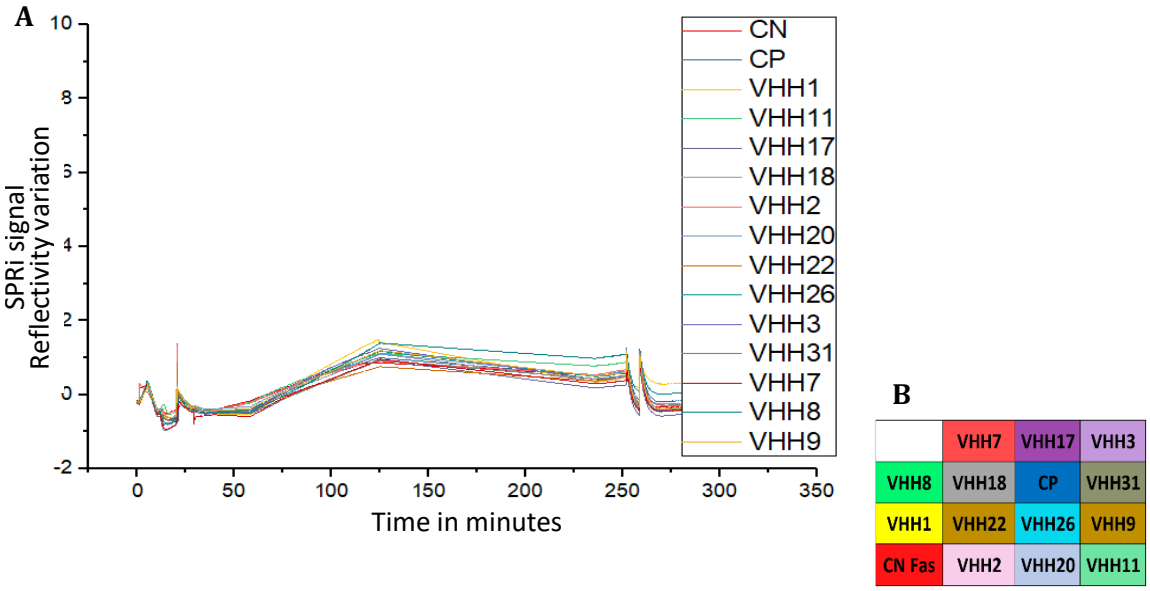
Figure 4.16- (continuation)



VHH	SPRi RV	Cell density at 0.3 gray level	SPRi RV *105
CN	2.27	231	238.26
CP	3.54	459	371.61
VHH1	3.46	325	362.88
VHH11	4.35	477	456.90
VHH17	3.43	403	360.32
VHH18	4.31	408	452.83
VHH2	3.75	445	394.18
VHH20	4.76	522	499.55
VHH22	3.95	426	414.49
VHH26	3.42	357	359.61
VHH3	3.88	349	407.55
VHH31	3.25	324	340.87
VHH7	4.16	387	437.07
VHH8	4.01	407	420.76
VHH9	3.89	436	408.46

Figure 4.17- Cell density for each VHH obtained from differential images at 0.3 grey level in the SPRi assay I with THP-I cells.

SPRi assay II with THP-1 cells



C

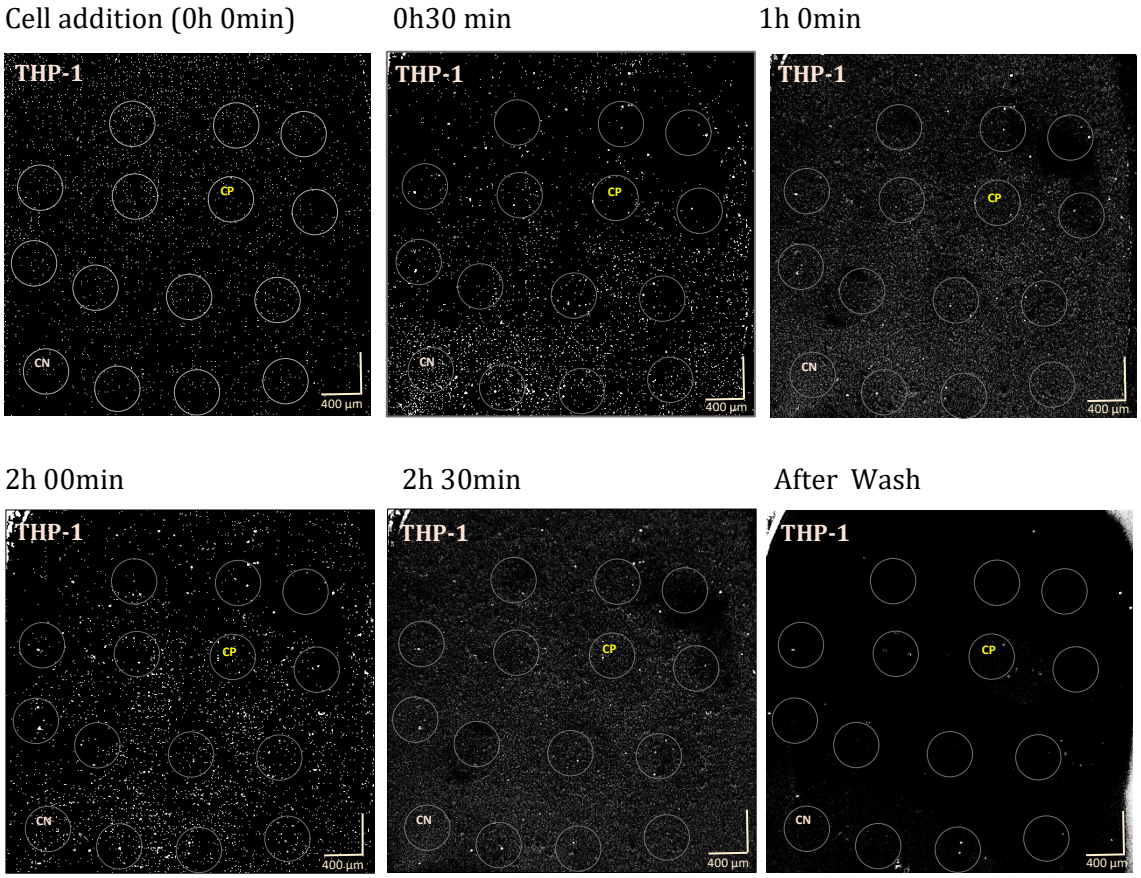
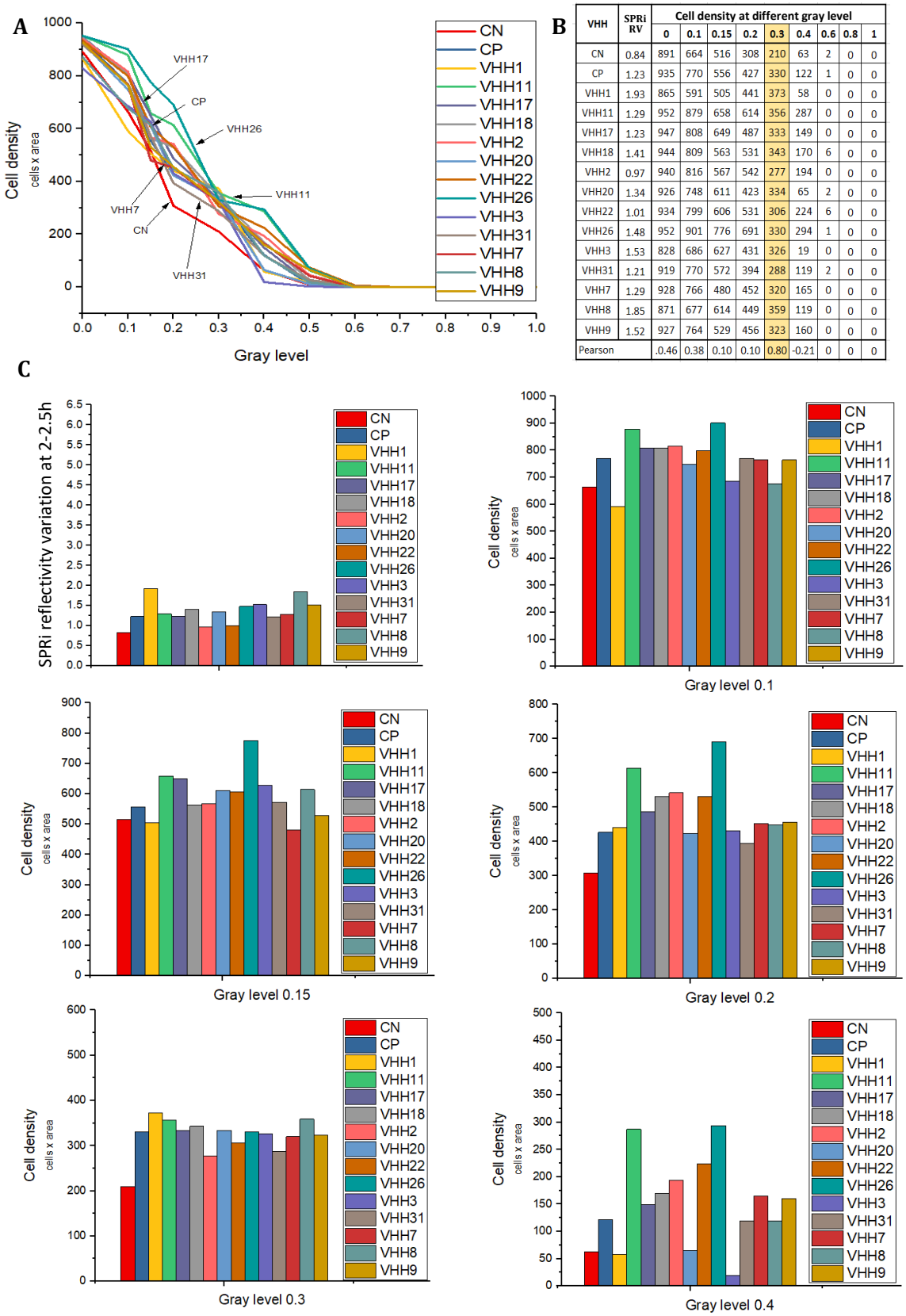


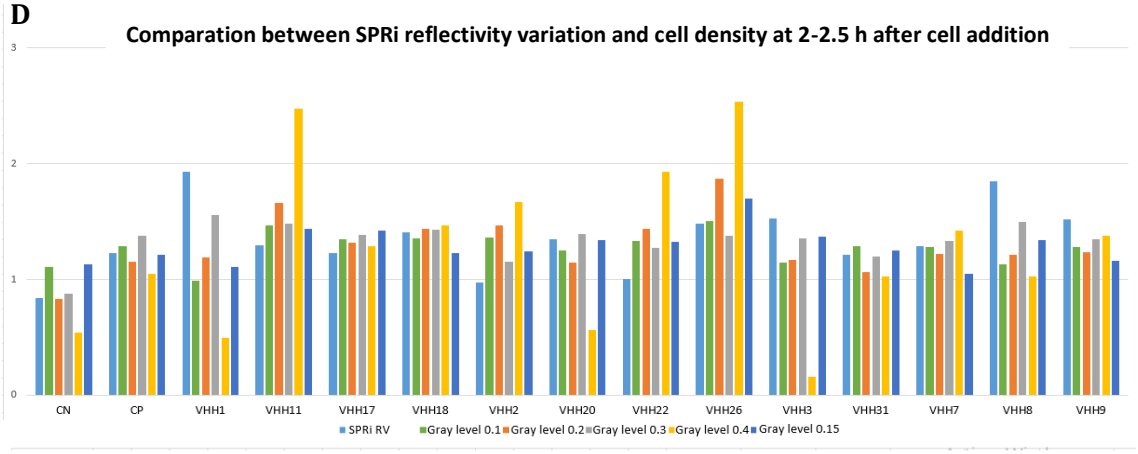
Figure 4.18- SPRi assay II of the 13 recombinant anti-CD105 VHH proteins (100 ng/µl or 6.6 µM) at different incubation times with the THP-1 cell line (CD105 (-)) (2 x 10^6 cells). a) SPRi curves monitoring reflectivity variations upon the incubation time of each VHH with the cells. b) Distribution of the VHH proteins on the biochip (each VHH shows a different colour). c) Differential images at different times (0, 30, 60, 120 and 150 minutes) of the assay. 0.5 µl of each VHH sample was applied on the biochip, that is. 50 ng or 3.33 µmoles of each VHH per spot. **CN**: negative control (VHH anti Fas2), **CP**: positive control (EngVHH17).



VHH	SPRi RV	Cell density at different gray level								
		0	0.1	0.15	0.2	0.3	0.4	0.6	0.8	1
CN	0.84	891	664	516	308	210	63	2	0	0
CP	1.23	935	770	556	427	330	122	1	0	0
VHH1	1.93	865	591	505	441	373	58	0	0	0
VHH11	1.29	952	879	658	614	356	287	0	0	0
VHH17	1.23	947	808	649	487	333	149	0	0	0
VHH18	1.41	944	809	563	531	343	170	6	0	0
VHH2	0.97	940	816	567	542	277	194	0	0	0
VHH20	1.34	926	748	611	423	334	65	2	0	0
VHH22	1.01	934	799	606	531	306	224	6	0	0
VHH26	1.48	952	901	776	691	330	294	1	0	0
VHH3	1.53	828	686	627	431	326	19	0	0	0
VHH31	1.21	919	770	572	394	288	119	2	0	0
VHH7	1.29	928	766	480	452	320	165	0	0	0
VHH8	1.85	871	677	614	449	359	119	0	0	0
VHH9	1.52	927	764	529	456	323	160	0	0	0
Pearson		.0.46	0.38	0.10	0.10	0.80	-0.21	0	0	0

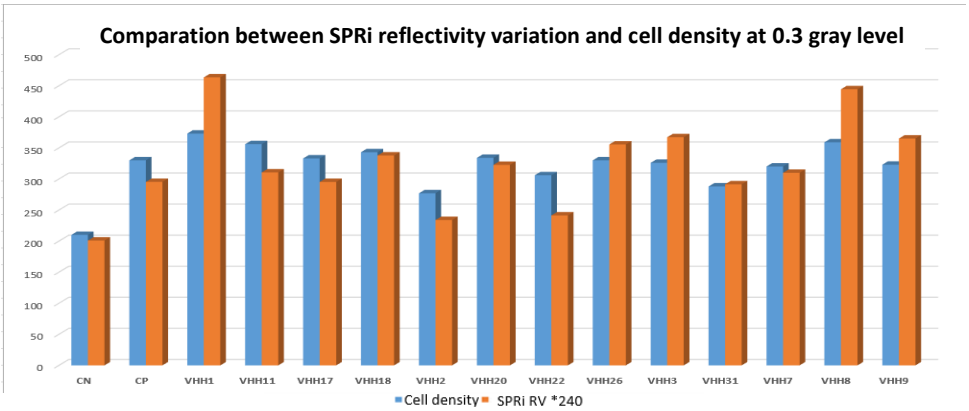
Figure 4.19- Determination of cell density for each VHH by comparison with SPRi reflectivity variation (SPRi RV) 2 hours after adding the cells in SPRi assay II with THP-I cells. a) Cell density curve of each VHH at different gray levels, b) Pearson's correlation coefficient, c) y d) Visual comparison between SPRi RV and cell density different at gray levels (0.1 to 0.4), individually (c) in group (d). The area used was 477.09 µm x 477.09 µm (31x31 pixeles). **CN**: negative control (VHH anti Fas2), **CP**: positive control (EngVHH17).

D



	CN	CP	VHH1	VHH11	VHH17	VHH18	VHH2	VHH20	VHH22	VHH26	VHH3	VHH31	VHH7	VHH8	VHH9
SPRi RV	0.84	1.23	1.93	1.29	1.23	1.41	0.97	1.34	1.01	1.48	1.53	1.21	1.29	1.85	1.52
Gray level 0.1/598	1.1	1.3	1.0	1.5	1.4	1.4	1.4	1.3	1.3	1.5	1.1	1.3	1.3	1.1	1.3
Gray level 0.2/370	0.8	1.2	1.2	1.7	1.3	1.4	1.5	1.1	1.4	1.9	1.2	1.1	1.2	1.2	1.2
Gray level 0.3/240	0.9	1.4	1.6	1.5	1.4	1.4	1.2	1.4	1.3	1.4	1.4	1.2	1.3	1.5	1.3
Gray level 0.4/116	0.5	1.1	0.5	2.5	1.3	1.5	1.7	0.6	1.9	2.5	0.2	1.0	1.4	1.0	1.4
Gray level 0.15/457	1.1	1.2	1.1	1.4	1.4	1.2	1.2	1.3	1.3	1.7	1.4	1.3	1.1	1.3	1.2

Figure 4.19- (continuation)



VHH	SPRi RV	Cell density at 0.3 gray level	SPRi RV *240
CN	0.84	210	200.86
CP	1.23	330	295.30
VHH1	1.93	373	463.49
VHH11	1.29	356	310.49
VHH17	1.23	333	295.18
VHH18	1.41	343	337.82
VHH2	0.97	277	233.95
VHH20	1.34	334	322.66
VHH22	1.01	306	241.30
VHH26	1.48	330	355.49
VHH3	1.53	326	367.25
VHH31	1.21	288	291.43
VHH7	1.29	320	309.91
VHH8	1.85	359	444.43
VHH9	1.52	323	365.11

Figure 4.20- Cell density for each VHH obtained from differential images at 0.3 grey level in the SPRi assay II with THP-I cells.

THP-1 cells (I)												
VHH	SPR signal 2 hours	Cell density at gray level					SPR signal 2.5 hours	Cell density at gray level				
		0.1	0.15	0.2	0.3	0.4		0.1	0.15	0.2	0.3	0.4
CN	2.27	0	0	0	0	0	2.32	893	656	492	231	104
CP	3.54	0	2	0	0	0	3.55	852	747	663	459	333
VHH1	3.46	0	0	0	0	0	3.47	816	664	428	325	134
VHH11	4.35	3	0	0	0	0	4.37	931	676	736	477	249
VHH17	3.43	4	1	0	0	0	3.49	943	746	718	403	142
VHH18	4.31	0	0	0	0	0	4.38	792	686	503	408	114
VHH2	3.75	2	1	0	0	0	3.76	907	774	707	445	269
VHH20	4.76	3	0	0	0	0	4.76	913	833	760	522	379
VHH22	3.95	0	0	0	0	0	4.00	833	692	613	426	317
VHH26	3.42	0	1	0	0	0	3.44	886	760	578	357	243
VHH3	3.88	2	2	0	0	0	3.91	951	775	723	349	150
VHH31	3.25	0	0	0	0	0	3.28	853	637	571	324	183
VHH7	4.16	3	2	0	0	0	4.22	947	808	742	387	174
VHH8	4.01	0	0	0	0	0	4.02	864	686	647	407	317
VHH9	3.89	0	0	0	0	0	3.91	896	682	656	436	288

THP-1 cells (II)												
VHH	SPR signal 2 hours	Cell density at gray level					SPR signal 2.5 hours	Cell density at gray level				
		0.1	0.15	0.2	0.3	0.4		0.1	0.15	0.2	0.3	0.4
CN	0.84	0	1	1	0	0	0.84	664	516	308	210	63
CP	1.23	1	1	0	0	0	1.27	770	556	427	330	122
VHH1	1.93	0	0	0	0	0	1.99	591	505	441	373	58
VHH11	1.29	0	0	0	0	0	1.31	879	658	614	356	287
VHH17	1.23	0	0	0	0	0	1.28	808	649	487	333	149
VHH18	1.41	0	0	0	0	0	1.46	809	563	531	343	170
VHH2	0.97	1	0	0	0	0	0.98	816	567	542	277	194
VHH20	1.34	1	0	0	0	0	1.37	748	611	423	334	65
VHH22	1.01	0	0	0	0	0	1.04	799	606	531	306	224
VHH26	1.48	0	1	0	0	0	1.53	901	776	691	330	294
VHH3	1.53	0	0	0	0	0	1.57	686	627	431	326	19
VHH31	1.21	0	1	0	0	0	1.26	770	572	394	288	119
VHH7	1.29	1	1	0	0	0	1.34	766	480	452	320	165
VHH8	1.85	0	0	0	0	0	1.92	677	614	449	359	119
VHH9	1.52	0	0	0	0	0	1.58	764	529	456	323	160

Figure 4.21- Cell density counted in differential images at 2h and 2.5h of the SPRi assays I and II with THP-1 cells. The SPRi signals (reflectivity variations) at 2h and 2.5h showed no significant differences. However, there were differences in the cell densities counted at various grey levels in images at 2.5h compared to those counted in images at 2h, where there were very few or no cells and no distinctions were observed between the VHH; therefore, images at 2.5h were considered for cell count.

The average quantity of cells captured by each VHH in an area of 2.28 x10^5 um^2 (477.09 µm x 477.09 µm), that is, the cell density average and the SPR signal (reflectivity variation) average of each VHH in the assays with SC (CD105(+)) and THP-1 (CD105(-)) cells at 2 hours after cell addition, are shown in figures 4.22 and 4.23.

The VHH were listed according to cell density values considering those with the highest values as those with the most capacity to capture and retain cells, meaning, those that not only have a large number of interactions but also keep the largest number of captured cells.

In the SC cell assays, the values ranged between 748 and 233 cells/area, being the VHH 17 (VHH sequence 15) and VHH 7 (VHH sequence 5) those who presented the highest values, 748 and 717 cells/area respectively, even above the positive control (CP) with 659 cells/area. Meanwhile the VHH 1 (VHH sequence 1) and VHH 26 (VHH sequence 18) presented the lowest values, with 412 and 417 cells/area respectively.

Regarding the nine VHH remaining, VHH 20 (VHH sequence 16), VHH 3 (VHH sequence 3), VHH 31 (VHH sequence 4) and VHH 2 (VHH sequence 2) showed values superior to 620 cells/area. The VHH 22 (VHH sequence 17), VHH 9 (VHH sequence 7) and VHH 8 (VHH sequence 6) showed values superior to 550 cells/area. Meanwhile, the VHH 18 (VHH sequence 19) and 11 (VHH sequence 8) values of 507 and 473 cells/area respectively.

A

SC cells (CD105 (+))											
VHH	SPRi signal (reflectivity variation)					VHH	Cell density (cell/area)				
	Assay I	Assay II	Assay III	SPRi average	St. Dev.		Assay I	Assay II	Assay III	Cell dens. average	St. Dev.
CP	6.09	5.80	4.74	5.5	0.7	CP	723	626	628	659	55.4
VHH17	6.20	5.55	4.37	5.4	0.9	VHH17	787	689	767	748	51.8
VHH7	6.05	4.79	5.37	5.4	0.6	VHH7	801	563	774	713	130.3
VHH20	5.87	4.30	3.93	4.7	1.0	VHH20	747	683	661	697	44.7
VHH3	5.47	5.70	3.22	4.8	1.4	VHH3	586	688	606	627	54.0
VHH31	6.04	6.24	2.30	4.9	2.2	VHH31	757	695	410	621	185.1
VHH2	6.19	3.60	3.71	4.5	1.5	VHH2	773	568	519	620	134.7
VHH22	5.67	4.06	4.61	4.8	0.8	VHH22	739	508	482	576	141.5
VHH9	5.09	4.52	3.91	4.5	0.6	VHH9	546	617	511	558	54.0
VHH8	5.67	3.33	4.33	4.4	1.2	VHH8	583	493	575	550	49.8
VHH18	5.71	2.28	4.79	4.3	1.8	VHH18	536	412	574	507	84.7
VHH11	5.16	4.33	3.44	4.3	0.9	VHH11	489	411	520	473	56.2
VHH26	5.75	3.86	2.94	4.2	1.4	VHH26	524	450	277	417	126.8
VHH1	5.54	2.52	2.71	3.6	1.7	VHH1	543	373	321	412	116.1
CN	4.74	1.37	2.00	2.7	1.8	CN	275	282	141	233	79.5
CG	4.44	1.35	1.96	2.58	1.6						

B

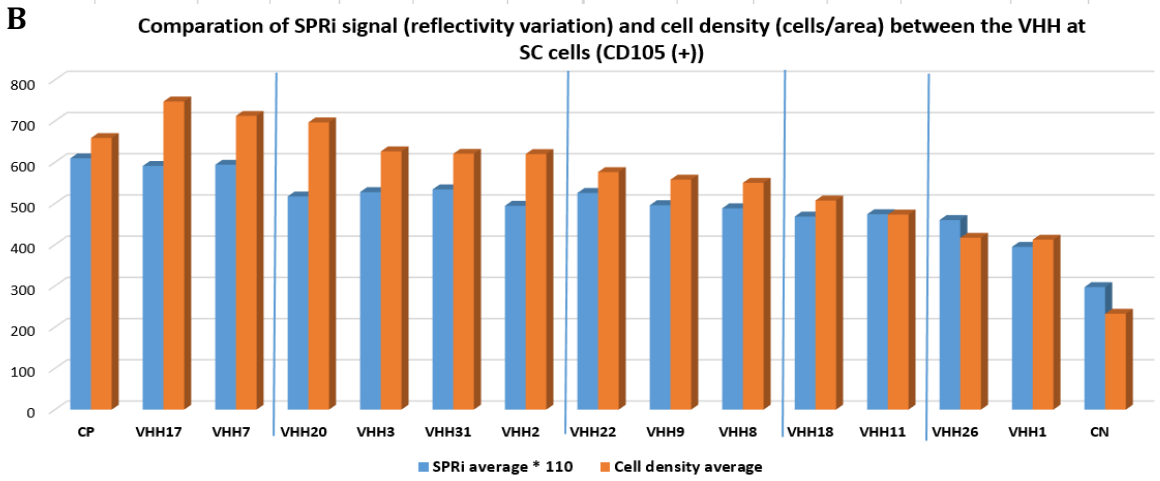


Figure 4.22- Average SPRi signal (reflectivity variation) and average cell density (cells in area of 2.28 x10^5 um^2) of the 13 anti-CD105 VHH evaluated 2 hours after cell addition in assays with SC cells (CD105 (+)). a) Average SPRi signal and cell density values for each VHH in the SPRi assays with SC cell, b) Comparison between the 13 VHH. SPRi signal values (reflectivity variations) were multiplied by a factor of 110 for comparison with the number of cells obtained at the cell densities CP: Positive Control, CN: Negative Control, CG: Gold Control.

In the assays with THP-1 cells, both the SPRi signal (reflectivity variation) and the cell density values did not show a significant difference between the VHH and the negative control. The VHH 20 and VHH 11 had the highest values, 428 and 417 cells/area respectively, while the VHH 31 and VHH 3 with 306 and 338 cells/area had the lowest values. The cell density values without significant differences between the VHH and close to the value of the negative control (anti-Fas VHH, non-antigen related protein) support the specificity of the VHH to their antigen, CD105, as THP-1 does not express CD105.

A

THP-1 cells (CD105 (-))									
VHH	SPRi signal (reflectivity variation)				VHH	Cell density (cell/area)			
	Assay I	Assay II	SPRi average	St. Dev.		Assay I	Assay II	Cell dens. average	St. Dev.
CP	3.54	1.23	2.4	1.6	CP	459	330	395	91.2
VHH17	3.43	1.23	2.3	1.6	VHH17	403	333	368	49.5
VHH7	4.16	1.29	2.7	2.0	VHH7	387	320	354	47.4
VHH20	4.76	1.34	3.1	2.4	VHH20	522	334	428	132.9
VHH3	3.88	1.53	2.7	1.7	VHH3	349	326	338	16.3
VHH31	3.25	1.21	2.2	1.4	VHH31	324	288	306	25.5
VHH2	3.75	0.97	2.4	2.0	VHH2	445	277	361	118.8
VHH22	3.95	1.01	2.5	2.1	VHH22	426	306	366	84.9
VHH9	3.89	1.52	2.7	1.7	VHH9	436	323	380	79.9
VHH8	4.01	1.85	2.9	1.5	VHH8	407	359	383	33.9
VHH18	4.31	1.41	2.9	2.1	VHH18	408	343	376	46.0
VHH11	4.35	1.29	2.8	2.2	VHH11	477	356	417	85.6
VHH26	3.42	1.48	2.5	1.4	VHH26	357	330	344	19.1
VHH1	3.46	1.93	2.7	1.1	VHH1	325	373	349	33.9
CN	2.27	0.84	1.6	1.0	CN	231	210	221	14.8
CG	2.28	0.49	1.38	1.3					

B

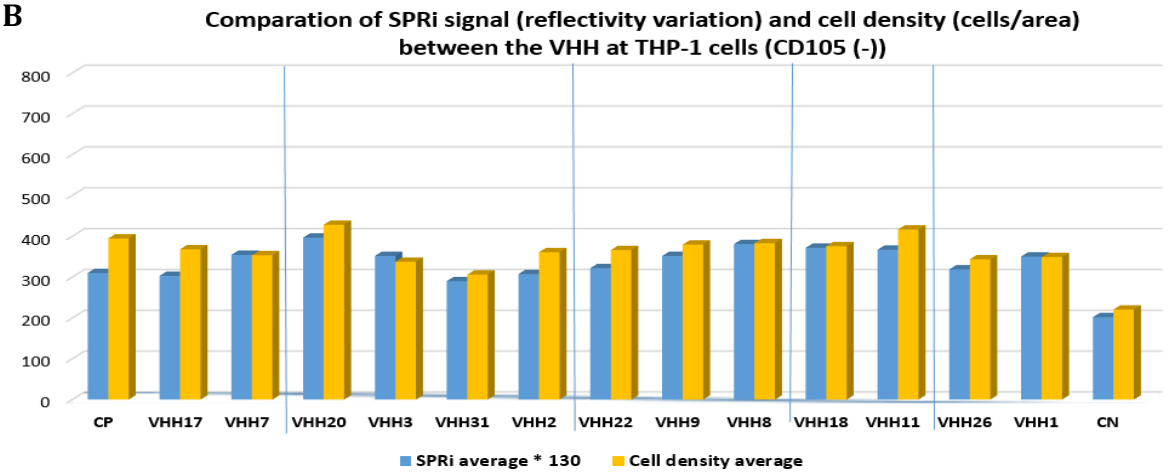


Figure 4.23- Average SPRi (reflectivity variation) and average cell density (cells in area of 2.28 x10^5 um^2) of the 13 anti-CD105 VHH evaluated 2-2.5 hours after cell addition in assays with THP-1 cells (CD105 (-)). a) Average SPRi signal and cell density values for each VHH in the SPRi assays with THP-1 cells, b) Comparison between the 13 VHH. SPRi signal (reflectivity variations) values were multiplied by a factor of 130 for comparison with the number of cells obtained at the cell densities CP: Positive Control, CN: Negative Control, CG: Gold Control.

The specificity is the capacity that an antibody possesses to bind to the antigen that stimulated it through the epitope (portion of the antigen recognized and to which antibodies bind) by weak and no covalent interactions, such as electrostatic interactions, hydrogen bonds, Van der Waals forces and hydrophobic interactions[161]. This binding is due to the specific chemical constitution of each antibody, which makes it very precise and distinguishes between chemical groups with minimal differences, therefore each antibody, in this case each VHH, can bind only one specific antigen. Those VHH with less nonspecific interactions, such as those that could occur in the absence of antigen (CD105 (-) cells), were therefore the most specific.

A preliminary segregation of five groups was made by comparison of the cell density of each VHH in the assays with SC cells (CD105 (+)) and THP-1 cells (CD105 (-)), that is, the VHH were grouped by their capacity to specifically capture CD105 on cells. (Figure 4.24).

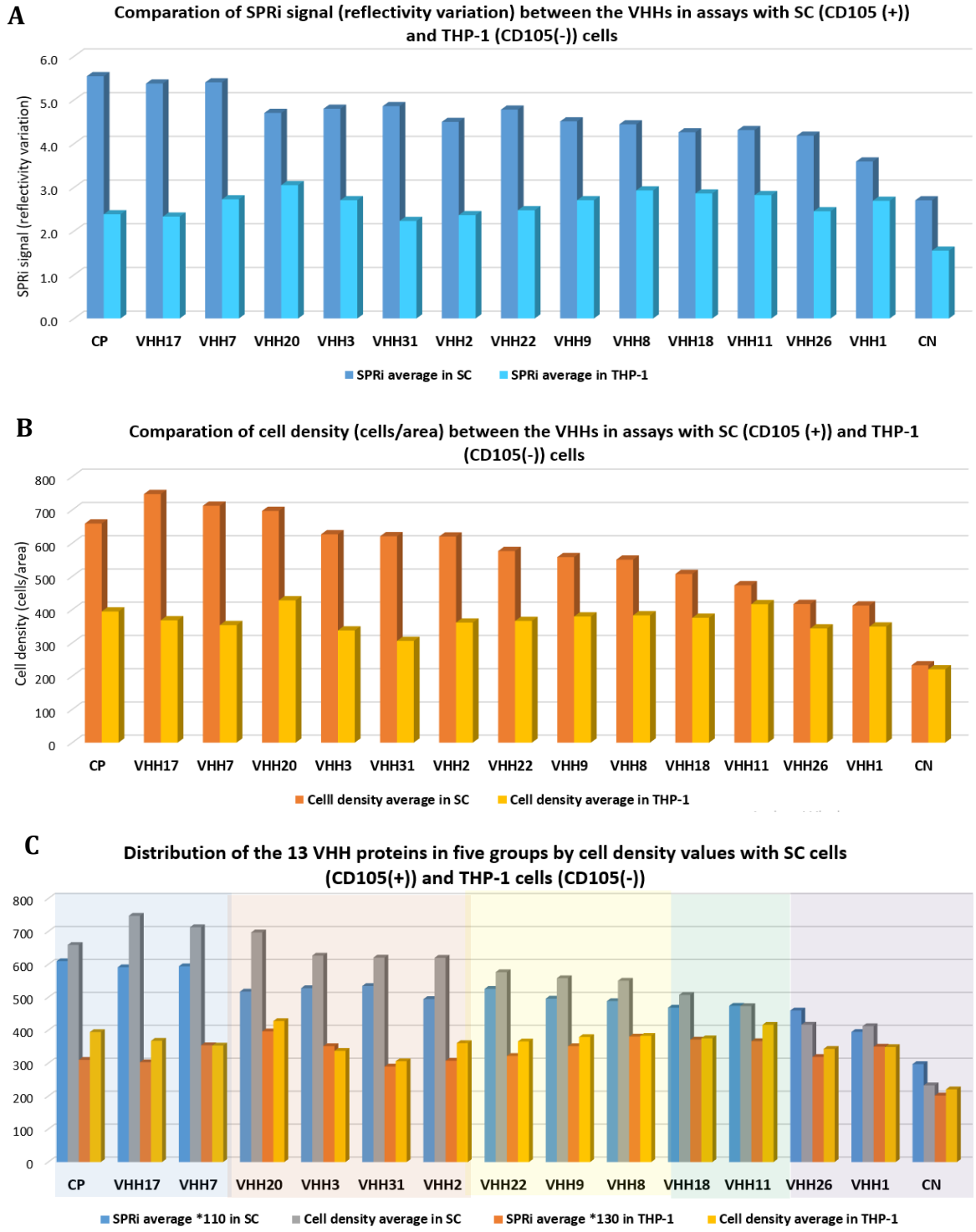


Figure 4.24-. Comparison of cell density and SPRi signal (reflectivity variation) between the VHH in the assays with the SC (CD105 (+)) and THP-1 (CD105 (-)) cell lines at 2-2.5 hours after cell addition. a) SPRi signals (reflectivity variations) comparison between the VHH in the assays with the SC (CD105 (+)) and THP-1 (CD105 (-)) cells, b) Cell density comparison between the VHH in the assays with the SC (CD105 (+)) and THP-1 (CD105 (-)) cells, c) VHH distribution in 5 groups by their specificity to CD105 on cells. The groups are shown in colours: group 1 in light blue, group 2 in pink, group 3 in yellow, group 4 in green, group 5 in light grey.

Thus, for group 1, the VHHs with cell density values higher than 700 cells/area were considered, for group 2 values higher than 600 cells/area, for group 3 values higher than 550 cells/area, for group 4 values higher than 450 cells/area and for group 5 values higher than 400 cells/area. The affinity is the measurement of the strength of the interaction between the epitope and the antibody-binding site. Therefore, a high affinity will be the result of a strong interaction with more attractive forces between the epitope and the binding site and less dissociation between them[162]. The interaction of the VHH immobilized on the gold film with the antigen expressed on the cells resulted in changes in the intensity of the reflected light (reflectivity variations), that could be detected and monitored in real time by SPRi. SPRi reflectivity is directly proportional to the mass that is immobilized.

The analysis of the rhythm of change in the reflectivity (SPRi signal) allows obtaining apparent rhythm constants for the association and dissociation phases of the reaction. The ratio of the reflectivity values in association and in dissociation approximates an apparent equilibrium constant (association constant or "affinity constant") at a given time[103,135,163]. Approximations to the apparent association constants of the VHH at a given time (the same time considered for cell counting), were obtained from the variations of reflectivity (SPRi signal shifts) at 2 hours after addition of the cells (association), with respect to the initial reflectivity, that is, upon addition of the cells (dissociation). For this purpose, the quotients of the 2-hour reflectivity values by the initial reflectivity values were calculated for each VHH. (Figure 4.25a).

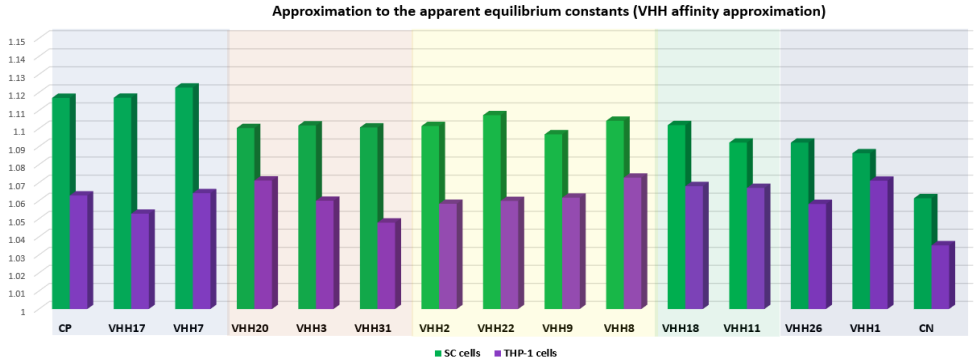
Considering the higher or lower affinity approximation and the specificity of each VHH, a segregation of the VHH into only 4 groups was performed (Figure 4.25b). In the group 1 formed by the VHH 17 and VHH 7, VHH 31 was incorporated. VHH31 was initially considered within the group 2, together with VHH 20, VHH 3 and VHH 2 (Figure 4.24).

VHH 20 showed a higher cell density than VHH 3, VHH 31 and VHH 2 in the SC cell assays, however, it also showed a higher cell density in THP-1 cell assays, (figure 4.23 a), meaning it had the lowest specificity in the group. While VHH 31 turned out to be the most specific, not only of the group but of all the VHH evaluated, with the cell density value in the THP-1 cell assays being the lowest of all the VHH, where only the negative control was inferior (Figure 4.23).

VHH 22 was also considered as part of the group 2 due to its apparent association constant approximation value, which showed that it was a higher affinity VHH than those VHHs in the group 3. Thus, group 2 consisted of VHH 2, VHH 3 and VHH 22. VHH 20 was considered as part of group 3 with the VHH 18, VHH 8, VHH 9 and VHH 26, this last one with cell density and apparent affinity constant approximation values not very high but was the third with the highest specificity of the 13 VHHs, after VHH 31 and VHH 3. VHH 11 had the lowest affinity and VHH 1 the lowest specificity so, they were considered in the last group, the group 4.

A

Apparent equilibrium constant (Affinity)										
SC cells (CD105 (+))						THP-1 cells (CD105 (-))				
VHH	Assay I	Assay II	Assay III	Affinity average	St. Dev.	VHH	Assay I	Assay II	Affinity average	St. Dev.
CP	1.1303	1.1067	1.1140	1.1170	0.01	CP	1.1001	1.0256	1.0629	0.05
VHH17	1.1354	1.1052	1.1107	1.1171	0.02	VHH17	1.0786	1.0271	1.0528	0.04
VHH7	1.1373	1.0875	1.1433	1.1227	0.03	VHH7	1.0989	1.0295	1.0642	0.05
VHH20	1.1240	1.0746	1.1023	1.1003	0.02	VHH20	1.1124	1.0300	1.0712	0.06
VHH3	1.1172	1.1114	1.0766	1.1017	0.02	VHH3	1.0863	1.0337	1.0600	0.04
VHH31	1.1271	1.1201	1.0547	1.1006	0.04	VHH31	1.0698	1.0261	1.0479	0.03
VHH2	1.1426	1.0633	1.0984	1.1014	0.04	VHH2	1.0936	1.0230	1.0583	0.05
VHH22	1.1253	1.0715	1.1253	1.1074	0.03	VHH22	1.0959	1.0239	1.0599	0.05
VHH9	1.1089	1.0859	1.0957	1.0968	0.01	VHH9	1.0892	1.0342	1.0617	0.04
VHH8	1.1335	1.0577	1.1215	1.1043	0.04	VHH8	1.1024	1.0433	1.0728	0.04
VHH18	1.1419	1.0395	1.1244	1.1019	0.05	VHH18	1.1027	1.0336	1.0681	0.05
VHH11	1.1108	1.0803	1.0856	1.0922	0.02	VHH11	1.1044	1.0298	1.0671	0.05
VHH26	1.1288	1.0709	1.0769	1.0922	0.03	VHH26	1.0819	1.0345	1.0582	0.03
VHH1	1.1350	1.0434	1.0804	1.0863	0.05	VHH1	1.0931	1.0490	1.0711	0.03
CN	1.1014	1.0221	1.0606	1.0614	0.04	CN	1.0486	1.0220	1.0353	0.02
CG	1.1033	1.0234	1.0460	1.0575	0.04	CG	1.0548	1.0122	1.0335	0.03



B

VHH	Kte. apparent equilibrium	
	SC (CD105(+))	THP-1(CD105(-))
CP	111.70×10^{-2}	106.29×10^{-2}
VHH17	111.71×10^{-2}	105.28×10^{-2}
VHH7	112.27×10^{-2}	106.42×10^{-2}
VHH31	110.06×10^{-2}	104.79×10^{-2}
VHH22	110.74×10^{-2}	105.99×10^{-2}
VHH2	110.14×10^{-2}	105.83×10^{-2}
VHH3	110.17×10^{-2}	106.00×10^{-2}
VHH9	109.68×10^{-2}	106.17×10^{-2}
VHH18	110.19×10^{-2}	106.81×10^{-2}
VHH26	109.22×10^{-2}	105.82×10^{-2}
VHH8	110.43×10^{-2}	107.28×10^{-2}
VHH20	110.03×10^{-2}	107.12×10^{-2}
VHH11	109.22×10^{-2}	106.71×10^{-2}
VHH1	108.63×10^{-2}	107.11×10^{-2}
CN	106.14×10^{-2}	103.53×10^{-2}

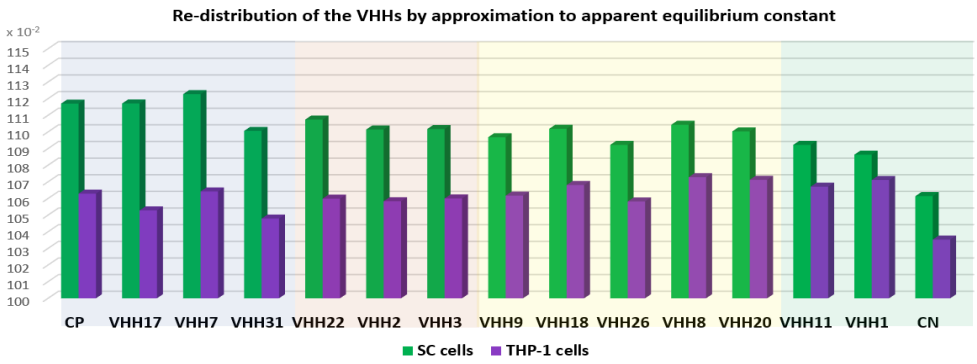


Figure 4.25- Segregation of the VHH according to the apparent association constant approximations a) Values of the apparent affinity constant approximations of each VHH with initial VHH distribution in 5 groups, b) Average values of the apparent affinity constant approximations of each VHH and re-distribution of the VHH in 4 grupos according these values. The groups are shown in colours: group 1 in light blue, group 2 in pink, group 3 in yellow, group 4 in green, group 5 in light grey.

Since the antigen-antibody binding is dependent of the conformation and folding of the antibody (VHH) and this in turn is dependent of the composition of the amino acids (Figure 4.26), it was important to review the structure and characteristics of the predetermined VHH proteins from their sequences (chapter II, figure 2.8b). VHH 17 (VHH sequence 15), VHH 7 (VHH sequence 5) and VHH 31 (sequence 4) which lead the ranking with density values superior to the positive control, would have the most appropriate conformation for the interaction with their antigen even after immobilization on the microarray.

Their sequences are not very similar to each other. The first sequence determines a protein of 144 amino acid residues, the second a larger protein of 150 and the third a protein of 147. However, all sequences determine stable, hydrophilic proteins, even when VHH 17 and VHH 31 are visibly more stable than VHH7 (which without the 6His tale would be unstable, chapter II, figure 2.8a).

The pI of VHH 17 (7.8) is closer to the pH of the assay medium, pH 7.4 (physiological pH) than the pI of VHH 7 and VHH 31 which are more alkaline, 9.4 and 9.6, respectively. However, as all three VHH proteins would present a net positive charge at the pH of the assay, they would be soluble, even though the former may be less soluble than the latter.

VHH 7 possesses three lysin residues, which means that structurally it has fewer conformational possibilities that may affect the functionality of the protein in comparison with VHH 17 and 31, which posses seven and six lysin residues respectively (Figures 4.4 and 4.26). However, this situation does not seem to affect greatly the adequate conformation for the interaction with the antigen over the cells.

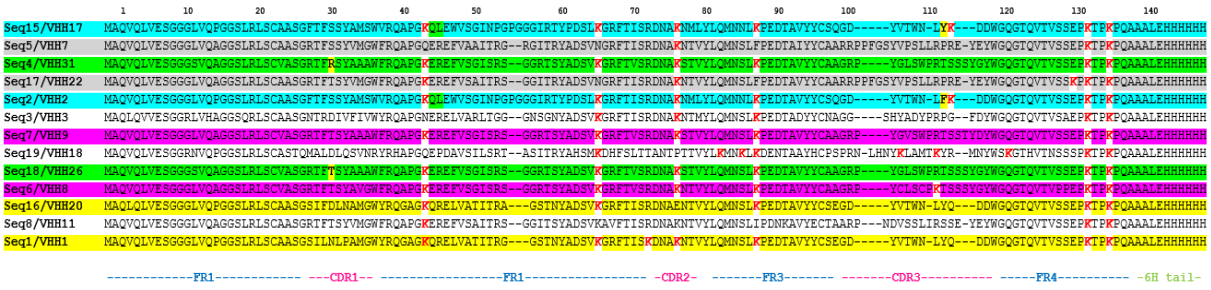


Figure 4.26- Amino acid sequences of the 13 anti-CD105 VHH. In red, the Lysine residues (K). Similar sequences show the same colour. In yellow background, the R30T substitution in VHH 31 and VHH 26; and the Y107F substitution in VHH17 and VHH2.

Perhaps, because the positionings of most of these lysin residues are in regions that do not drastically influence the protein folding, or the linker added to the protein for its immobilization on the biochip (SH-C12-NHS) provides enough flexibility for it to maintain the appropriate conformation whatever the binding point (amine). VHH 31 had the highest specificity according to ELISA with the immobilized CD105 antigen (chapter III, figure 3.16) and SPR assays with SC cells (CD105(+)) and THP-1 cells (CD105(-)) confirmed this specificity. Similarly, VHH 17 was the third in specificity according to ELISA.

VHH 17 and VHH2 are very similar to each other, with only one difference, Y107F, so, it was expected that both would also have similar behavior. Effectively, VHH 2 was among those which presented great capacity to interact specifically and with high affinity with the antigen, even if it was not as high as VHH 17. Perhaps because this single (and seemingly simple) change was of an apolar to a polar amino acid residue, which seems to help VHH17 not only in increasing the hydrophilicity and stability of the protein but also in improving the interaction with the antigen and the permanence of the binding. These VHH did not present the substitutions G44E and L45R that distinguish them from the VH regions of the antibodies and diminish the hydrophobicity of that region (chapter II, figure 2.5b) so, any change that favour the hydrophilicity and stability of the protein could aid to improve its functionality.

Other sequences that differed at a single amino acid residue were VHH 31 and VHH 26 (VHH sequence 18), which was R30T. VHH 31 is very specific and VHH 26 is also, but the latter has less capacity to capture cells with CD105 (with lower cell density value). As in the previous case, the change of a residue seems to contribute to the better functionality of the protein, in this case, it is the change of a charged polar amino acid residue (R), by a polar residue (T), which could have helped VHH 31 improve its cell capture capacity.

The sequences VHH 7 and VHH 22 (sequence 17) are also similar, so it was expected that VHH 22 would also has a good capacity to interact with the antigen, even though having more lysin residues in its composition (five against three) could imply more conformations unsuitable for interactions and that could affect its functionality. As VHH 22 does not differ significantly from VHH7 in its interaction with the antigen (similar as VHH 17 and VHH 31), it seems that the location of the anchorage points of the VHH (Ɛ amine of K) on the biochip does not greatly affect the conformation that allows the recognition and interaction with the antigen. So the greater capacity for interaction falls mainly in the composition of amino acids which better stabilises the folding of the protein.

Other sequences similar to each other were VHH 9 (VHH sequence 7) with VHH 8 (VHH sequence 6), and VHH 1 (VHH sequence 1) with VHH 20 (sequence 16). VHH 9 and VHH 8, along with VHH 18 (VHH sequence 19), performed similarly in the assays, so the former was considered to have better capacity than the others, but only by a small margin. VHH 18 sequence was the most different among all the VHH sequences obtained (Chapter II, figure 2.5) and with the highest quantity of lysin in its composition (nine), nevertheless, it had the stability and adequate folding for a good interaction with CD105 as the other VHH that conform its group (Figure 4.25).

Another VHH that did not show much similarity with the other sequences was the VHH 3. In regard of VHH 1 and VHH 20, they were the only ones that presented a pI below physiological pH and therefore, their net charge during the assays was negative. In contrast to VHH1, VHH 20 had a good capacity to capture SC cell, with a cell density value above positive control. However, VHH 20 was also good at capturing THP-1 cells so it is not very specific. VHH 1 and VHH 11 (VHH sequence 8) had the last places in the ranking. VHH 11, as VHH 20, is one of the least specific VHH, but unlike VHH 20, its capacity to

capture cells is not very high, but rather close to that of the VHH considered in the group 3. VHH 1 was the least specific and had the lowest capacity to capture cells with results close to the negative control.

Finally, the average values of the reflectivity variations (SPRi signal), cell density and apparent association constant approximations obtained in the THP-1 (CD105 (-)) cell assays were subtracted from those obtained in the SC (CD105 (+)) cell assays to obtain the difference for each VHH. Furthermore, an average of these values among all VHH was made, resulting in VHH SPRi reflectivity average of 4.59 with SC cells and 2.69 with THP-1 cells, VHH cell density average of 578 with SC cells and 367 with THP-1 cells and VHH apparent association constant approximation average of 1.1019 with SC cells and 1.0626 with THP-1 cells.

In the assays with SC cells, most of the VHH in the groups 1 and 2 had above-average values meanwhile the values of the VHH in the group 3 were closer to the average. On the other hand, in the assays with THP-1 cells, the VHH had values closer to the average, means that there was not much variability (low standard deviation) even though VHH 31 and VHH 11 showed the highest and furthest values from the average (Figure 4.27a).

The difference obtained for each VHH coming from the subtraction of the results with SC cells minus the results with THP-1 cells allowed to determine the values obtained only for their interactions with CD105, removing the possible non-specific interactions. The VHH were sorted according to these difference values (reflectivity, cell density and association approximation) and were segregated into groups that matched the previously established distribution of four groups. The VHH with values well above and above the average were placed in the first two groups respectively (group 1 and group 2), those with values equal to or very close to the average in the third group (group 3) and those with values closer to the negative control in the last group (group 4) (figure 4.27b)

With this confirmation, the ranking of the anti-CD105 VHH according to their capacity and ability to recognise and interact with their antigen expressed over cells, and with high specificity and affinity while they were immobilized over a microarray for SPRi, was stablished (Figure 4.27b).

A

SPRi signal (reflectivity variations)				
VHH	SC (CD105(+))	THP-1 (CD105(-))	Difference	Diff. X 120
CP	5.54	2.38	3.16	379.0
VHH17	5.38	2.33	3	365.4
VHH7	5.40	2.73	2.67	321.0
VHH31	4.86	2.23	2.63	315.3
VHH22	4.78	2.48	2.30	276.4
VHH2	4.50	2.36	2.13	256.0
VHH3	4.80	2.71	2.09	251.1
VHH9	4.51	2.71	1.80	216.4
VHH18	4.26	2.86	1.40	167.9
VHH26	4.18	2.45	1.73	207.6
VHH8	4.44	2.93	1.51	181.5
VHH20	4.70	3.05	1.65	198.2
VHH11	4.31	2.82	1.49	178.6
VHH1	3.59	2.69	0.90	107.7
CN	2.70	1.55	1.15	138.0
Average	4.59	2.64	1.95	234.1
St.D	0.5	0.2	0.6	64.4
CG	2.58	1.38	1.20	

Cell density (cells/area)			
VHH	SC (CD105(+))	THP-1 (CD105(-))	Difference
CP	659	395	265
VHH17	748	368	380
VHH7	713	354	359
VHH31	621	306	315
VHH22	576	366	210
VHH2	620	361	259
VHH3	627	338	289
VHH9	558	380	179
VHH18	507	376	132
VHH26	417	344	74
VHH8	550	383	167
VHH20	697	428	259
VHH11	473	417	57
VHH1	412	349	63
CN	233	221	12
Average	578	367	211
St.D	95.7	23.4	100.7

Apparent affinity constant approximations				
VHH	SC (CD105(+))	THP-1 (CD105(-))	Difference	Diff. x 6000
CP	1.117	1.0629	0.0541	324.6
VHH17	1.1171	1.0528	0.0643	385.8
VHH7	1.1227	1.0642	0.0585	351
VHH31	1.1006	1.0479	0.0527	316.2
VHH22	1.1074	1.0599	0.0475	285
VHH2	1.1014	1.0583	0.0431	258.6
VHH3	1.1017	1.06	0.0417	250.2
VHH9	1.0968	1.0617	0.0351	210.6
VHH18	1.1019	1.0681	0.0338	202.8
VHH26	1.0922	1.0582	0.0340	204
VHH8	1.1043	1.0728	0.0315	189
VHH20	1.1003	1.0712	0.0291	174.6
VHH11	1.0922	1.0671	0.0251	150.6
VHH1	1.0863	1.0711	0.0152	91.2
CN	1.0614	1.0353	0.0261	156.6
Average	1.1019	1.0626	0.0394	236.1
St.D	0.0091	0.0071	0.0113	67.98
CG	1.0575	1.0335	0.0241	

B

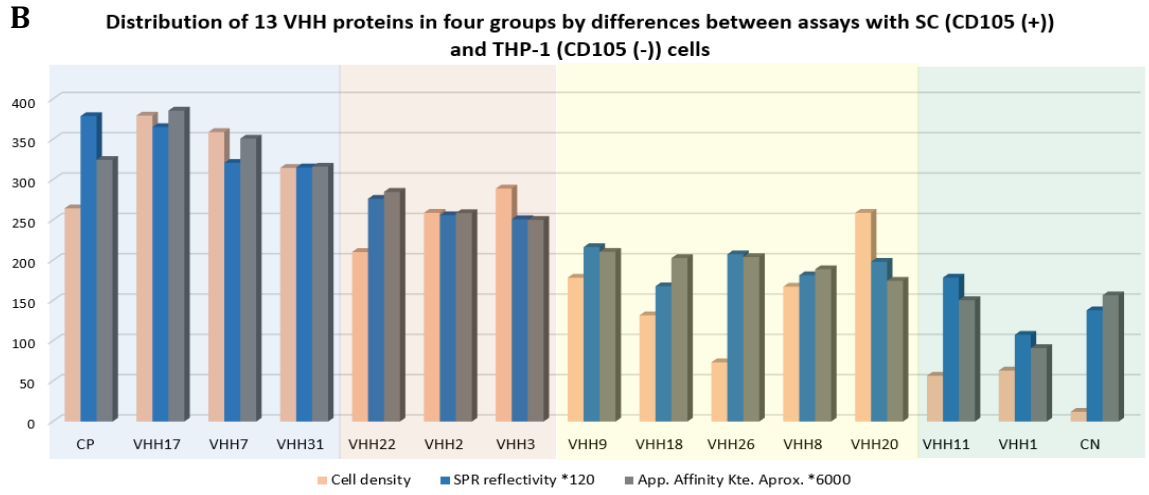


Figure 4.27- Comparison of the 13 VHHs by difference in SPRi reflectivity, cell density and apparent affinity constant approximation values in assays with SC (CD105 (+)) and THP-1 (CD105 (-)) cells to rank the VHH. a) Difference values obtained by subtracting the average values of SPRi reflectivity, cell density and apparent affinity constant approximation of each VHH in assays with THP-1 cells from the average values in assays with SC cells, b) Distribution of VHH in four groups by difference of the values obtained in assays with SC and THP-1 cells. Ranking of the anti-CD105 VHH by their antigen-interacting capacity with higher specificity and affinity. Groups are shown in colours: group 1 in light blue, group 2 in pink, group 3 in yellow, group 4 in green. For comparison with cell density, the SPRi reflectivity and apparent affinity constant approximation values were multiplied by factors of 120 and 6000 respectively.

4.4 Conclusions

Recombinant anti-CD105 VHH proteins highly specific for their antigen in free form (soluble CD105) showed their recognition capacity on CD105-expressing cells (membrane CD105) by real-time evaluation with SPRi. The conditions of functionalisation of Ig G proteins were adapted for application to VHH proteins. The 13 purified recombinant VHH were coupled to the linker SH-C_{12}-N-Hydroxysuccinimide (Thiol-NHS) in a 1:10 ratio with respect to the molar concentration of the protein, for their immobilisation on the SPRi biochip by direct adsorption (chemisorption by gold-thiol interactions). The VHH microarrayed on the biochip were manually grafted to obtain spots large enough to allow the antigen-antibody interaction to be properly displayed given the size difference of the VHH (2.4 x 4.2 nm) with the antigen-expressing cells (~15 uM). The SC (ATCC® CRL-9855™) and THP-1 (ATCC® TIB-202™) non-adherent cell lines were chosen for the SPRi assays. SC because they are monocyte-macrophage type cells and overexpress CD105 on their surface and THP-1 because they are monocyte type cells that do not express CD105. SPRi imaging showed that most of the VHH were capable to bind to SC cells, but not (or very little) to THP-1 cells, demonstrating their specificity. Cell counting was performed on differential images at 2 hours of incubation with the cells and the grey level that provided a cell density (cells in an area of 477.09 µm x 477.09 µm) concordant with the variation in reflectivity (SPRi signal) of each VHH at the same assay time was chosen. That is, grey level 0.1- 0.15 for SC cells and 0.3 for THP-1 cells. Approximations to the apparent association constant were obtained and the VHH were divided into 4 groups according to specificity and affinity. VHH 17, VHH 7 and VHH 31 formed group 1, VHH 22, VHH 2 and VHH 3 formed group 2, VHH 9, VHH 18, VHH 26, VHH 8 and VHH 20 formed group 3 and VHH 11 and VHH1 formed group 4. The average of the SPRi reflectivity variation between the VHHs at 2 hours was 4.59 with SC cells and was 2.69 with the THP-1 cells. The average of the cell density was 578 with SC cells and 367 with THP-1 cells and the average of the approximation to the apparent association constant was 1.1019 with SC cells and 1.0626 with THP-1 cells. The difference between the values of cell density, SPRi reflectivity variation and approximation to the apparent association constant obtained in assays with SC and THP-1 cells corroborated the distribution of the 13 VHHs in four groups. In the group 1 were VHH 7, VHH 17 and VHH 31 with values well above average among the VHH and in the group 2 were the VHH 22, VHH 2 and VHH 3 with values above average. In the group 3 were the VHH 9, VHH 18, VHH 26, VHH 8 and VHH 20 with values equal or very close to average and in the group 4 were the VHH 11 and VHH 1 with values below average and close to the negative control. The six highest ranked recombinant anti-CD105 VHHs (groups 1 and 2) would have the most efficient conformations and folding for antigen interaction and are therefore proposed as nanoprobes for CD105, not only in free form (soluble CD105), but also on cells (membrane CD105), for research studies in the diagnosis or therapeutics of implicated diseases.

References

1) Medina-Torres, E. A. & Espinosa-Rosales, F. (2009) Microarreglos: Tecnología con aplicaciones en el campo de la salud humana. *Alergia Asma Inmunologia Pediátricas 18(2)*, 52-59.

2) Chang, T. W. (1983). Binding of cells to matrixes of distinct antibodies coated on solid surface. *J Immunol Methods 65(1–2),* 217–223.

3) Schena, M., Shalon, D., Davis, R. & Brown, P. (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science 270*, 467-470.

4) Schena, M. (1999) DNA Microarrays: A Practical Approach. Oxford, UK: Oxford University Press.

5) Call, D. (2001) DNA microarrays–their mode of action and possible applications in molecular diagnostics. *Veterinary Sciences Tomorrow 3,* 1-9.

6) Gabig, M., & Wegrzyn, G. (2001) An introduction to DNA Chips: principles, technology, applications and analysis. *Acta Biochimica Polonica, 48(3),* 615-622.

7) Heller, M. J. (2002) DNA microarray technology: devices, systems, and applications. *Annu Rev Biomed Eng 4(1),* 129-153.

8) Fadiel, A. & Naftolin, F. (2003) Microarray applications and challenges: a vast array of possibilities. *Int Arch Biosci 1,* 1111-1121.

9) Zhou, J. & Thompson, D. K. (2004) Microarray technology and applications in environmental microbiology. *Advances in Agronomy 82,* 183-270

10) Majtan, T., Bukovska, G. & Timko, J. (2004) DNA microarrays—techniques and applications in microbial systems. *Folia Microbiol (Praha) 49(6),* 635-664.

11) Hoheisel, J. D. (2006) Microarray technology: beyond transcript profiling and genotype analysis. *Nature, 7*, 200-210.

12) Sassolas, A., Leca-Bouvier, B. D. & Blum, L. J. (2008) DNA biosensors and microarrays. *Chem Rev 108,* 109-139.

13) Gresham, D., Dunham, M. J. & Botstein, D. (2008) Comparing whole genomes using DNA microarrays. *Nat Rev Genet 9(4),* 291-302

14) Teng, X. & Xiao, H. (2009) Perspectives of DNA microarray and nextgeneration DNA sequencing technologies. *Sci China C Life Sci 52(1),* 7-16.

15) Patirupanusara, P. & Suthakorn, J. (2012) Introduction of an active DNA microarray fabrication for medical applications. *International Conference on Advances in Electrical and Electronics Engineering (ICAEEE'2012) April 13-15*, 75-79.

16) Bumgarner, R. (2013) Overview of DNA microarrays: Types, applications, and their future. *Curr Protoc Mol Biol 101(1),* 22.1.

17) Singh, A. & Kumar, N. (2013) A review on DNA microarray technology. *Int J Curr Res Rev 5(22),* 1.

18) Nsofor, C. A. (2014) DNA microarrays and their applications in medical microbiology. *Biotechnol Molecul Biol Rev 9(1),* 1-11.

19) Patil, M. & Bhong, C. (2015) Veterinary diagnostics and DNA microarray technology. *Int J Livest Res 5,* 1-9

20) Dadkhah, K., Anijdan, S. H. M., Karimi, M., Abolfazli, M. K., Rezaei, F., Adel, F., & Ataei, G. (2015) DNA microarray, types and its application in medicine. *Sch Acad J Biosci 3(7),* 598-602.

21) Beena, V., Pawaiya, R., Gururaj, K., Shivasharanappa, S. & Karikalan, M. (2016) Application of microarray in animal disease pathogenesis and diagnosis. *J Vet Sci Technol 7(6),* 1000402.

22) Murtaza, I., Laila, O. & Ubaid-ullah, S. (2016) DNA microrray: A miniaturized high throughput technology. *DNA 12(2),* 9-15.

23) Fesseha, H. & Tilahun, H. (2020) Principles and applications of deoxyribonucleic acid microarray: A Review. *Pathol Lab Med Open J 1(1),* 54-62.

24) Haab, B. B., Dunham, M. J. & Brown, P. O. (2001) Protein microarrays for highly parallel detection and quantitation of specific proteins and antibodies in complex solutions. *Genome biology 2(2),* RESEARCH0004.1

25) MacBeath, G. & Schreiber, S. L. (2000) Printing proteins as microarrays for high-throughput function determination. *Sci (New York NY) 289(5485),* 760–763.

26) Kodadek, T. (2001) Protein microarrays: Prospects and problems. *Chem Biol 8(2),* 105–115.

27) Chen, C. S. & Zhu, H. (2006) Protein microarrays. *BioTechniques 40(4),* 423. 425.

28) Sutandy, F. X, Qian, J., Chen, C. S. & Zhu, H. (2013) Overview of protein microarrays. *Curr Protoc Protein Sci Chapter 27(1): Unit 27.1,* 27.1.1–27.1.16

29) Abel, L., Kutschki, S., Turewicz, M., Eisenacher, M., Stoutjesdijk, J., Meyer, H. E., . . . May, C. (2014) Autoimmune profiling with protein microarrays in clinical applications. *Biochim Biophys Acta (BBA) - Proteins Proteomics 1844(5),* 977–987.

30) Wilson, J. J., Burgess, R., Mao, Y. Q., Luo, S., Tang, H., Jones, V. S., . . . Huang, R. P. (2015) Antibody arrays in biomarker discovery. *Adv Clin Chem 69,* 255–324.

31) Ayoglu, B., Schwenk, J. M. & Nilsson, P. (2016) Antigen arrays for profiling autoantibody repertoires. *Bioanalysis 8(10),* 1105–1126.

32) Moore, C. D., Ajala, O. Z. & Zhu, H. (2016) Applications in high-content functional protein microarrays. *Curr Opin Chem Biol 30,* 21–27.

33) Gupta, S., Manubhai, K. P., Kulkarni, V. & Srivastava, S. (2016) An overview of innovations and industrial solutions in protein microarray technology. *Proteomics 16(8),* 1297–1308.

34) Duarte, J. G. & Blackburn, J. M. (2017) Advances in the development of human protein microarrays. *Expert Rev Proteomics 14(7)*, 627–641.

35) Chen, Z., Dodig-Crnković, T., Schwenk, J. M., & Tao, S. C. (2018) Current applications of antibody microarrays. *Clin Proteomics 15,* 7.

36) Szymczak, L. C., Kuo, H. Y. & Mrksich, M. (2018) Peptide arrays: Development and application. *Analytical Chem 90(1),* 266–282.

37) Qi, H., Wang, F. & Tao, S. C. (2019) Proteome microarray technology and application: higher, wider, and deeper. *Expert Rev Proteomics 16(10),* 815–827.

38) Syu, G. D., Dunn, J. & Zhu, H. (2020) Developments and applications of functional protein microarrays. *Mol Cell Proteomics MCP 19(6*), 916–927

39) Li, S., Song, G., Bai, Y., Song, N., Zhao, J., Liu, J. & Hu, C. (2021) Applications of protein microarrays in biomarker discovery for autoimmune disease. *Front. Immunol 12,* 1518.

40) Tarca, A. L., Romero, R. & Draghici, S. (2006) Analysis of microarray experiments of gene expression profiling. *Am J Obstet Gynecol. 195(2),* 373-388.

41) Orntoft, T. & Kruhøffer. (2006) DNA Chips and microarrays. *Encyclo Life Scien 10*, 1038.

42) Snijders, A., Nowak, N., Segraves, R., Blackwood, S., Brown, N., Conroy, J., . . . Albertson, D. (2001) Assembly of microarrays for genome-wide measurement of DNA copy number. *Nat Genet 29*, 263-264.

43) Liljedahl, U., Fredriksson, M., Dahlgren, A. & Syvänen, A. (2004) Detecting imbalanced expression of SNP alleles by minisequencing on microarrays. *BMC Biotechnology 4*, 24.

44) Kasukabe, T., Kado, J., Kato, N., Sass, T., & Honma, Y. (2005). Effects of combined treatment with rapamycin and cotylenin A, a novel differentiation-inducing agent, on human breast carcinoma MCF-7 cells and xenografts. *Breast Cancer Res 7*, 1097-1110.

45) Mantripragada, K., Buckley, P., Stahl, T. & Dumanski, J. (2004) Genomic microarrays in the spotlight. *Trends Genet 20*, 87-94.

46) Van't Veer, L., Dai, H., Van de Vijver, M., Witteveen, A., Schreiber, G., Kerkhoven, R., . . . Friend, S. (2002) Gene expression profiling predicts clinical outcome of breast cancer. *Nature 415*, 530-536

47) Fillies, T., Werkmeister, R., Van Diest, P., Brandt, B., Joos, U. & Buerger, H. (2005) HIF1-alpha overexpression indicates a good prognosis in early stage squamous cell carcinomas of the oral floor. *BMC Cancer 5*, 84.

48) Abramovitz, M. & Leyland-Jones, B. (2006) A systems approach to clinical oncology: Focus on breast cancer. *Proteome Science 4*, 5.

49) Alizadeh, A., Eisen, M., Davis, R., Ma, C., Lossos, I., Rosenwald, A., . . . Staudt, L. (2000) Distinct types of diffuse of large B- cell Iymphoma identified by gene expression profoling. *Nature 403*, 503-511.

50) Graudens, E., Boulanger, V., Mollard, C., Mariage-Samson, R., Barlet, X., Gremy, G., . . . Imbeaud, S. (2006). Deciphering cellular states of innate tumor drug responses. *Genome Biology 7*, R19.

51) Nagasaki, K. & Miki, Y. (2006) Gene expression profiling of breast cancer. *Breast Cancer 13*, 2-7.

52) Hu, N., Wang, C., Hu, Y., Yang, H., Kong, L., Lu, N., . . . Lee, M. (2006). Genome-wide loss of heterozygosity and copy number alteration in esophageal squamous cell carcinoma using the affymetrix genechip mapping 10 K array. *BMC Genomics 7*, 299.

53) Pan, J., Song, G., Chen, D., Li, Y., Liu, S., Hu, S., ... Huang, Yi. (2017) Identification of serological biomarkers for early diagnosis of lung cancer using a protein array-based approach. *Mol Cell Proteomics MCP 16(12),* 2069–2078.

54) Trøen, G., Nygaard, V., Jenssen, T., Ikonomou, I., Tierens, A., Matutes, E., . . . Delabie, J. (2004) Constitutive expression of the AP-1 transcription factors c-jun, junD, junB, and c-fos and the marginal zone B-Cell transcription factor Notch2 in splenic marginal zone lymphoma. *J Mol Diagnos 6(4)*, 297-307.

55) Lock, C., Hermans, G., Pedotti, R., Brendolan, A., Schadt, E., Garren, H., . . . Steinman, L. (2002) Gene-microarray analysis of multiple sclerosis lesions yields new targets validated in autoimmune encephalomyelitis. *Nat Med 8*, 500-508.

56) Liu, Y., Hui, S., Peng, X., Dong-Hai, X., Li-Hua, L., Recker, R. & Deng, H. (2004) Molecular genetic studies of gene identification for osteoporosis: A 2004 Update. *J Bone Miner Res 21(10)*, 1511-1535.

57) Shashkin, P., Jain, N., Miller, Y., Rissing, B., Huo, Y., Keller, S., . . . McIntyre, T. (2006) Insulin and glucose play a role in foam cell formation and function. *Cardiovascular Diabetology 5*, 13.

58) Lal, S., Chrisopherson, R. & Dos Remedios, C. (2002) Antibody arrays: an embryonic but rapidly growing technology. *Drug Discovery Today 7(18Suppl)*, S143-149.

59) Chen, R., Pan, S., Brentnall, T. & Aebersold, R. (2005) Proteomic profiling of pancreatic cancer for biomarker discovery. *Mol Cell Proteomics 4 (4)*, 523-525.

60) Lee, S., Sayin, A., Grice, S., Burdett, H., Baban, D. & Van den Heuvel, M. (2008) Genome-wide expression analysis of a spinal muscular atrophy model: towards discovery of new drug targets. *PLoS ONE 3(1)*, e1404.

61) Chizhikov, V., Rasooly, A., Chumakov, K. & Levy, D. (2001) Microarray analysis of microbial virulence factors. *Applied and Enviromental Microbiology 67*, 3258-3263.

62) Peterson, G, Bai, J, Nagaraja, T. & Narayanan, S. (2010) Diagnostic microarray for human and animal bacterial diseases and their virulence and antimicrobial resistance genes. *J Microbiol Methods 80(3),* 223-230.

63) Shallom, S. J., Weeks, J. N., Galindo, C. L., McIver, L., Sun, Z., McCormick, J., . . . Garner, H. R. (2011) A species independent universal bio-detection microarray for pathogen forensics and phylogenetic classification of unknown microorganisms. *BMC Microbiol. 11(1),* 132.

64) Yoo, S. M., Choi, J. H., Lee, S. Y. & Yoo, N. C. (2009) Applications of DNA microarray in disease diagnostics. *J Microbiol Biotechnol 19(7),* 635-646.

65) Singh, S. & Bedekar, M. K. (2012) DNA microarray and its applications in disease diagnosis. *International Journal of Science and Research 3,* 1572-1573.

66) Hu, C. J., Pan, J. B., Song, G., Wen, X. T., Wu, Z. Y., Chen, S., . . . Li, Y. Z. (2017) Identification of novel biomarkers for Behcet disease diagnosis using human proteome microarray approach. *Mol Cell Proteomics: MCP 16(2),* 147–156.

67) Hall, D.A., Ptacek, J. & Snyder, M. (2012) Protein Microarray Technology. *Mech Ageing Dev 128 (1),* 161–167.

68) Ramachandran, N., Hainsworth, E., Bhullar, B., Eisenstein, S., Rosen, B., Lau, A.Y.,… LaBaer, J. (2004) Self-Assembling Protein Microarrays. *Science 305 (5680)*, 86-90.

69) Albeck, J., MacBeath, G., White, F., Sorger, P., Lauffenburger, D. & Gaudet, S. (2006) Collecting and organizing systematic sets of protein data. *Nature Reviews Molecular Cell Biology 7*, 803.

70) Hood, L., Heath, J., Phelps, M. & Lin, B. (2004) Systems Biology and New Technologies Enable Predictive and Preventative Medicine. *Science 306 (5696)*, 640-643.

71) Christodoulides, N., Floriano, P., Acosta, S., Michael Ballard, K., Weigum, S., Mohanty, S., . . . McDevitt, J. (2005) Toward the development of a lab-on-a-chip dual-function leukocyte and C-reactive protein analysis method for the assessment of inflammation and cardiac risk. *Clin Chem 51 (12)*, 2391-2395.

72) Nimse, S. B., Song, K., Sonawane, M. D., Sayyed, D. R. & Kim, T. (2014) Immobilization techniques for microarray: Challenges and applications. *Sensors (Basel) 14,* 22208-22229.

73) Zhu, H., Bilgin, M., Bangham, R., Hall, D., Casamayor, A., Bertone, P., . . . Snyder, M. (2001) Global analysis of protein activities using proteome chips. *Science 293(5537)*, 2101-2105.

74) Huang, R., Huang, R., Fan, Y. & Lin, Y. (2001) Simultaneous detection of multiple cytokines from conditioned media and patient's sera by an antibody-based protein array system. *Anal Biochem 294(1)*, 55-62.

75) Poulsen, T. B. G., Damgaard, D., Jørgensen, M. M., Senolt, L., Blackburn, J. M., Nielsen, C. H. & Stensballe, A. (2020) Identification of novel native autoantigens in rheumatoid arthritis. *Biomedicines 8(6),* 141.

76) Huang, W., Hu, C., Zeng, H., Li, P., Guo, L., Zeng, X., . . . Wu, L. (2012) Novel systemic lupus erythematosus autoantigens identified by human protein microarray technology. *Biochem Biophys Res Commun 418(2),* 241–246.

77) Beyer, N. H., Lueking, A., Kowald, A., Frederiksen, J. L. & Heegaard, N. H. (2012) Investigation of autoantibody profiles for cerebrospinal fluid biomarker discovery in patients with relapsing-remitting multiple sclerosis. *J Neuroimmunol 242(1-2)*, 26–32.

78) Bombaci, M., Pesce, E., Torri, A., Carpi, D., Crosti, M., Lanzafame, M., . . . Grifantini, R. (2019) Novel biomarkers for primary biliary cholangitis to improve diagnosis and understand underlying regulatory mechanisms. *Liver Int:* Off *J Int Assoc Study Liver 39(11),* 2124–2135.

79) Vanarsa, K., Soomro, S., Zhang, T., Strachan, B., Pedroza, C., Nidhi, M., . . . Mohan, C. (2020) Quantitative planar array screen of 1000 proteins uncovers novel urinary protein biomarkers of lupus nephritis. *Ann Rheum Dis 79(10),* 1349–1361.

80) Dang, K., Zhang, W., Jiang, S., Lin, X. & Qian, A. (2020) Application of lectin microarrays for biomarker discovery. *Chemistry Open 9(3),* 285–300.

81) Wingren, C., Ingvarsson, J., Dexlin, L., Szul, D. & Borrebaeck, C. (2007) Design of recombinant antibody microarrays for complex proteome analysis: choice of sample labeling-tag and solid support. *Proteomics 7*, 3055–3065.

82) Ingvarsson, J., Wingren, C., Carlsson, A., Ellmark, P., Wahren, B., Engström, G., . . . Borrebaeck, C. (2008) Detection of pancreatic cancer using antibody microarray-based serum protein profiling. *Proteomics 8*, 2211–2219.

83) Carlsson, A., Persson, O., Ingvarsson, J., Widegren, B., Salford, L., Borrebaeck, C. & Wingren, C. (2010) Plasma proteome profiling reveals biomarker patterns associated with prognosis and therapy selection in glioblastoma multiforme patients. *Proteomics Clin Appl 4*, 591–602.

84) Carlsson, A., Wuttge, D. M., Ingvarsson, J., Bengtsson, A. A., Sturfelt, G., Borrebaeck, C. A. & Wingren, C. (2011). Serum protein profiling of systemic lupus erythematosus and systemic sclerosis using recombinant antibody microarrays. *Mol Cell Proteomics 10(5),* M110.005033, 1-14.

85) Eisenberg, D., Marcotte, E., Xenarios, I. & Yeates, T. (2000) Protein function in the post-genomic era. *Nature 405 (6788)*, 823-826.

86) Angenendt, P., Glokler, J., Murphy, D., Lehrach, H. & Cahill, DJ. (2002) Toward optimized antibody microarrays: a comparison of current microarray support materials. *Anal Biochem 309*, 253-260.

87) Zhu, H. & Snyder, M. (2003) Protein chip technology. *Curr Opin Chem Biol 7 (1)*, 55-63.

88) Kambhampati, D. (2003) Protein Microarrays: From Fundamental Screening to Clinical Diagnostics. In: Kambhampati, D. *Protein Microarray Technology* (pp 1-9). Weinheim: Wiley VCH, Verlag GmbH & Co. KGaA.

89) Gong, P., Harbers, G. & Grainger, D. (2006) Multi-technique Comparison of Immobilized and Hybridized Oligonucleotide Surface Density on Commercial Amine-Reactive Microarray Slides *Anal Chem 78(7)*, 2342-2351.

90) Steel, A., Levicky, R., Herneand, T. & Tarlov, M. (2000) Immobilization of nucleic acids at solid surfaces: effect of oligonucleotide length on layer assembly. *Biophys J 79(2)*, 975-981.

91) Pawlak, M., Schick, E., Bopp, M., Schneider, M., Oroszlan, P. & Ehrat, M. (2002) Zeptosens' protein microarrays: A novel high performance microarray platform for low abundance protein analysis. *Proteomics 2(4)*, 383-393.

92) Wu, G., Datar, R., Hansen, K., Thundat, T., Cote, R. & Arun Majumdar, A. (2001) Bioassay of prostate-specific antigen (PSA) using microcantilevers. *Nature Biotechnol 19*, 856-860.

93) Stenlund, P., Babcock, G., Sodroski, J. & Myszka, D. (2003) Capture and reconstitution of G protein-coupled receptors on a biosensor surface. *Anal Biochem 316 (2)*, 243-250.

94) Goodrich, T., Wark, A., Corn, R. & Lee, H. (2006) Surface plasmon resonance imaging measurements of protein interactions with biopolymer microarrays. *Methods Mol Biol 328*, 113-130.

95) Lee, H., Nedelkov, D. & Corn, R. (2006) Surface plasmon resonance imaging measurements of antibody arrays for the multiplexed detection of low molecular weight protein biomarkers. *Anal Chem 78(18)*, 6504-6510.

96) Homola, J. (2003) Present and Future of surface plasmon resonance biosensors. *Anal Bioanal Chem 377*, 528–539.

97) Altuzara, V., Mendoza-Barrera, C., Muñoz, M., Mendoza-Alvarez, J. & Sanchez-Sinencio, F. (2010) Analisis cuantitativo de interacciones moleculares proteína-proteína mediante la combinacion de microarreglos y un lector óptico basado en el fenómeno de resonancia de plasmones superficiales. *Revista Mexicana de Física 56(2)*, 147–154.

98) Shabani, A. & Tabrizian, M. (2013) Design of a universal biointerface for sensitive, selective, and multiplex detection of biomarkers using surface plasmon resonance imaging. *Analyst 138*, 6052–6062.

99) Gorodkiewicz, E. & Łuszczyn, J. (2011) Surface Plasmon Resonance Imaging (SPRI) Sensor for Cystatin Determination Based on Immobilized Papain. *Prot Pept Lett 18*, 23–29.

100) Gorodkiewicz, E., Ostrowska, H. & Sankiewicz, A. (2011) SPR imaging biosensor for the 20S proteasome; sensor development and application to measurement of proteasomes in human blood plasma. *Microchim Acta 175*, 177–184.

101) Vo-Dinh, T. & Cullum, B. (2000) Biosensors and biochips: Advances in biological and medical diagnostics. *Fresenius J Anal Chem 366*, 540–551.

102) Daniel, C., Mélaïne, F., Livache, T. & Buhot, A. (2013) Real time monitoring of thrombin interactions with its aptamers: insights into the sandwich complex formation. *Biosens Bioelectron 40(1)*, 186–192.

103) Ramírez, N. (2005) Biosensores: Un Acercamiento a La Resonancia del Plasmon Superficial. *Revista CENIC Ciencias Biológicas 36()*. Obtained from https://www.redalyc.org/articulo.oa?id=181220525017. Revised January 2021

104) Thevenot, D., Toth, K., Durst, R. & G.S., W. (1999) Electrochemical Biosensors: Recommended Definition and Classification. *Pure Appl Chem 7*, 2333–2348.

105) Clark, L. & Lyons, C. (1962) Electrode Systems for Continuous Monitoring in Cardiovascular Surgery. *Annals of the New York Academy of Sciences 102(1)*, 29–45.

106) Heineman, W. R. & Jensen, W. B. (2006) Leland C. Clark Jr. (1918–2005). *Biosensors and Bioelectronics 21*, 1403–1404.

107) Tothill, I. E. (2009) Biosensors for cancer markers diagnosis. *Semin Cell Dev Bio 20*, 55–62.

108) Gorodkiewicz, E., Regulska, E. & Roszkowska –Jakimiec, W. (2010) Determination of the active form concentration of cathepsin D and B by SPRI biosensor. *Journal of Laboratory diagnostics 46*, 107–109.

109) Monošík, R., Streďanský, M. & Šturdík, E. (2012) Application of Electrochemical Biosensors in Clinical Diagnosis. *J Clin Lab Anal 26(1)*, 22–34.

110) Rodriguez-Mozaz, S., López de Alda, M. J., Marco, M. & Barcelo, D. (2005) Biosensors for environmental monitoring A global perspective. *Talanta 65(2)*, 291-297.

111) Katrlik, J., Mastihuba, V., Vostiar, I., Sefćovićova, J., Stefuca, V. & Gemeinar, P. (2006) Amperometric biosensors based on two different enzyme systems and their use for glycerol determination in samples from biotechnological fermentation process. *Anal Chim Acta 566*, 11–18.

112) Ramanathan, K. & Danielson, B. (2001) Principles and applications of thermal biosensors. *Biosens Bioelectron 16*, 417–423.

113) Zheng, Y., Liu, J., Ma, Y., Xu, Y., Xu, F. & Hua, T. (2012) Temperature effects on enzyme activity of chicken liver esterase used in calorimetric biosensor. *Artif Cells Blood Substit Immobil Biotechnol 40*, 125–131.

114) Janshoff, A., Galla, H., & Steinem, C. (2000). Piezoelectric mass-sensing devices as biosensors-An alternative to optical biosensors? *Angew Chem Int 39*, 4004–4032.

115) Geschwindner, S., Carlsson, J. & Knecht, W. (2012) Application of Optical Biosensors in Small-Molecule Screening Activities. *Sensors 12*, 4311–4323.

116) Sankiewicz, A, Puzan, B. & Gorodkiewicz, E. (2014) Biosensors SPRi as a diagnostic tool in the future. *CHEMIK 68(6)*, 528–535.

117) Fan, X., White, I., Shopova, S., Zhu, H., Suter, J. & Sun, Y. (2008) Sensitive optical biosensors for unlabeled targets: A review. *Anal Chim Acta 620*, 8–16.

118) Long, F., Zhu, A. & Shi, H. (2013) Recent Advances in Optical Biosensors for Environmental Monitoring and Early Warning. *Sensors 13(10)*, 13928–13948.

119) Nguyen, H. H., Park, J., Kang, S. & Kim, M. (2015) Surface plasmon resonance: a versatile technique for biosensor applications. *Sensors 15,* 10481–10510.

120) Suraniti, E., Sollier, E., Calemczuk, R., Livache, T., Marche, P. N., Villiers, M-B & Roupioz, Y. (2007) Real-time detection of lymphocytes binding on an antibody chip using SPR Imaging. *Lab Chip 7,* 1206–1208.

121) Cherif, B., Roget, A., Villiers, C. L., Calemczuk, R., Leroy, V., Marche, P. N., . . . Villiers, M. B. (2006) Clinically related protein-peptide interactions monitored in real time on novel peptide chips by surface plasmon resonance imaging. *Clin Chem 52*, 255–262

122) Homola, J., Vaisocherova, H., Dostalek, J. & Piliarik, M. (2005) Multi-analyte surface plasmon resonance biosensing. *Methods 37*, 26-36.

123) Helmerhorst, E., Chandler, D., Nussio, M. & Mamotte, C. (2012) Real-time and label-free bio-sensing of molecular interaction by Surface Plasmon Resonance: a laboratory Medicine perspective. *Clin Biochem Rev 33*, 161–173.

124) Robinson, M., Kuncova-Kallio, J., Grangvist, N. & Sadowski, J. (2012) Multi-Parametric Surface Plasmon Resonance – A new technique to determine thickness and refractive index of thin and thick layers. *Nanotech 1*, 42–44.

125) Zhao, J., Zhang, X., Yonzon, C., Haes, A. & Van Duyne, R. (2006) Localized surface plasmon resonance biosensors. *Nanomedicine 1(2)*, 219–228.

126) Bouguelia, S., Roupioz, Y., Slimani, S., Mondani, L., Casabona, M., D. C., . . . Livache, T. (2013). On-chip microbial culture for the specific detection of very low levels of bacteria. *Lab Chip 13*, 4024–4032.

127) Milgram, S., Cortes, S., Villiers, M. B., Marche, P., Buhot, A., Livache, T. & Roupioz, Y. (2011). On chip real time monitoring of B-cells hybridoma secretion of immunoglobulin. *Biosensors & bioelectronics* 26(5), 2728–2732.

128) Sepúlveda, B., Prieto, F., Calle, A. & Lechuga, L. (1999) Prototipo de Biosensor Óptico basado en la Resonancia de Plasmón Superficial con Sistema de referencia. *Proyecto AMB98-1048-C04-02 financiado por la Comisión Interministerial de Ciencia y Tecnología y por la Comunidad de Madrid*.

129) Steiner, G., Sablinskas, V., Hubner, A., Kune, C. & Salzer, R. (1999) Surface plasmon resonance imaging of microstructured monolayers. *J. Mol. Structure 509*, 265–273.

130) Biacore, A. (2001) Surface Plasmon Resonance. *Technology Note 1*, 23.

131) Homola, J. (2006) *Surface plasmon resonance based sensors*. *Springer Series on Chemical Sensors and Biosensor 4* (pp 26–43). Berlin: Springer-Verlag, Heidelberg Ed.

132) Green, R., Frazier, R., Shakesheff, K., Davies, M., Roberts, C. & Tendler, S. (2000) Surface plasmon resonance analysis of dynamic biological interactions with biomaterials. *Biomaterials 21*, 1823-1835.

133) Ricklin, D. (2005) *Surface Plasmon Resonance Applications in Drug Discovery with an Emphasis on Small Molecule and Low Affinity Systems.* Faculty of Philosophy and Natural Sciences of the University of Basel. Swiss: Thesis to obtain the degree of Doctor of Philosophy.

134) O'Shannessy, D., Burke, M., Soneson, K., Hensley, P. & Brooks, I. (1993) Determination of rate and equilibrium binding constants for macromolecular interactions using surface plasmon resonance: use of nonlinear least squares analysis methods. *Analytical Biochemistry 212*, 457-468.

135) Velasquez, A. (2010) *Acopladores de fibra óptica con recubrimiento de azopolímeros.* Universidad Nacional autonoma de Mexico: Thesis to obtain the title of ingeniero electrico electronico.

136) Sternberg, E., Persson, B., Roos, H. & Urbaniczky, C. (1991) Quantitative determination of surface concentration of protein with surface plasmon resonance by using radiolabelled proteins. *Journal of Colloid and Interface Science 143*, 513-526.

137) Mirabella, F. M. & Harrick, N. J (1985) *Internal Reflection Spectroscopy Review and Supplement*. Ossining, New York: *Harrick Scientific Corporation.*

138) de Mello, A. J. (1996) In: Davies, J., *Surface Analytical Techniques for Probing Biomaterial Processes*. Boca Raton, New York: CRC Press.

139) Jonsson, U., Fagerstam, L., Ivarsson, B., Johnsson, B., Karlsson, R., Lundh, K., . . . al., e. (1991). Real-time biospecific interaction analysis using surface plasmon resonance and a sensor chip technology. *Biotechniques 11*, 620-627.

140) Matsubara, K., Kawata, S., & Minami, S. (1988) A compact surface plasmon resonance sensor for water in process. *Applied Spectroscopy 42*, 1375-1379.

141) Steiner, G. (2004) Surface plasmon resonance imaging. *Anal Bioanal Chem 379*, 328–331.

142) Gorodkiewicz, E., Sankiewicz, A. & Laudański, P. (2014) Surface plasmon resonance imaging biosensors for aromatase based on a potent inhibitor and a specific antibody: Sensor development and application for biological material. *Cent Eur J Chem. 12(5)*, 557–587.

143) Abbas, A., Linman, M. J. & Cheng, Q. (2011) New trends in instrumental design for surface plasmon resonance-based biosensors. *Biosens Bioelectron 26(5)*, 1815–1824.

144) Bombera, R. (2011) *Développement de biopuces dédiées au tri d'échantillons cellulaires.* Université de Grenoble, France: Thesis to obtain the degree of Doctor, Spécialité Chimie-Biologie.

145) Melaine, F., Roupioz, Y. & Buhot, A. (2015) Gold Nanoparticles Surface Plasmon Resonance Enhanced Signal for the Detection of Small Molecules on Split-Aptamer Microarrays (Small Molecules Detection from Split-Aptamers). *Microarrays 4,* 41–52.

146) Melaine, F., Roupioz, Y. & Buhot, A. (2017) Small molecule SPR imaging detection from split aptamer microarrays. *Procedia Technology 27*, 6 – 7.

147) Sassolas, A., Blum, L.J. & Leca-Bouvier, B. (2012) Immobilization strategies to develop enzymatic biosensors. *Biotechnol Adv 30*, 489–511.

148) Guedon, P., Livache, T., Martin, F., Lesbre, F., Roget, A. & Bidan, G. (2000) Characterization and optimization of a real-time, parallel, label free, polypyrrole-based DNA sensor by surface plasmon resonance imaging. *Anal Chem 72*, 6003-6009.

149) Collins, GW. & Largen, M. (1995) Continuous mammalian cell lines having monocyte/macrophage characteristics and their establishment *in vitro*. *US Patent 5,447,861 dated Sep 5*.

150) Tsuchiya, S., Yamabe, M., Yamaguchi, Y., Kobayashi, Y., Konno, T. & Tada, K. (1980) Establishment and characterization of a human acute monocytic leukemia cell line (THP-1). *Int J Cancer 26*, 171-176.

151) Tsuchiya, S., Kobayashi, Y., Goto, Y., Okumura, H., Nakae, S., Konno, T. & Tada, K. (1982) Induction of maturation in cultured human monocytic leukemia cells by a phorbol diester. *Cancer Res. 42*, 1530-1536.

152) MATLAB. (2010) *version 7.10.0 (R2010a)*. Natick, Massachusetts: The MathWorks Inc.

153) ThermoScientific. (2012) *Crosslinking Technical Handbook.* Obtained from https://www.thermofisher.com/pierce. Revised January 2021

154) Gedig, E. (2017) *Handbook of Surface Plasmon Resonance.* Royal Society of Chemistry.

155) Alvarado, R. (2018) *Stratégies de fonctionnalisation pour le développement de biopuces innovantes.* Communaute Université Grenoble Alpes, France: Thesis to obtain the degree of Doctor, Spécialité Chimie-Biologie.

156) Lombana, A., Raja, Z., Casale, S., Pradier, C. M., Foulon,T., Ladram, A. & Humblot, V. (2014) Temporin-SHa peptides grafted on gold surfaces display antibacterial activity. *Journal of Peptide Science 20*, 563–569.

157) Carr, A. R., Townsend, R. E. & Badger, W. L. (1925) Vapor Pressures of Glycerol-Water and Glycerol-Water-Sodium Chloride Systems. *Industrial & Engineering Chemistry 17*, 643–646.

158) Gougos, A. & Letarte, M. (1988) Identification of a human endothelial cell antigen with monoclonal antibody 44G4 produced against a pre-B leukemic cell line. *J Immunol 141*, 1925-1933.

159) Lastres, P., Bellón, T., Cabañas, C., Sánchez-Madrid, F., Acevedo, A., Gougos, A., . . . Bernabeu, C. (1992) Regulated expression on human macrophages of endoglin, an Arg-Gly-Asp-containing surface antigen. *Eur J Immunol 22*, 393-397.

160) Restrepo, L. & González, J. (2007) De Pearson a Spearman. *Rev Col Cienc Pec 20*, 183-192.

161) Janeway, C. J., Travers, P., Walport, M. & Shlomchik, M. (2001) Immunobiology: The Immune System. En *Health and Disease .* New York: Garland Science (5ta edition).

162) Valdés, Y. A. & Hernández, A. (2001) Procedimientos para la detección e identificación de anticuerpos eritrocitarios. *Rev Cubana Hematol Inmunol Hemoter 17(2)*, 98-107.

163) *Astbury Centre for Structural Molecular Biology.* (2003) Obtained from http://www.astbury.leeds.ac.uk. Revised October 2020

Chapter V: Conclusions and perspectives

V.1 Generals conclusions

Endoglin (CD105) is an integral membrane protein that acts as a co-receptor for Transforming Growth Factor-β (TFGF-β) and is up regulated in proliferating endothelial cells. CD105 is suggested as an appropriate marker for neovascularisation and tumour-related angiogenesis. Antibodies anti-CD105 have been studied for application in diagnosis, prognostic and treatment of cancer (as well as other diseases) both alone and concomitantly with antibodies specific to other targets. It is of interest the generation of novel biomolecules that specifically bind CD105 that allow targeting, detection and capture of CD105 expressing cells, as angiogenesis is a key process in the growth and development of normal and transformed cells in tissues.

In the present work is presented the evaluation of several anti-CD105 single variable domain of heavy chain (VHH) antibodies that were screened from PBMC cDNA library generated from an alpaca immunized with a lysate of a bladder cancer cell line cell that express CD105. Nineteen recombinant VHHs that binds soluble CD105 (sCD105) with high affinity and specificity were obtained after the screening of this cDNA library by phage display and phage-ELISA. The anti-CD105 VHHs were analysed *in silico* to verify their identity and uniqueness as well as to predict the physico-chemical characteristics and the three-dimensional structure of the recombinant proteins expressed in the pET22b(+)/*E. coli* system. The 19 sequences analysed were different and shown homology to VHH sequences in *Gen Bank* and Protein Data Bank, they had substitutions in the VHH residues that are not present in the VH of conventional antibodies. All 19 sequences were subcloned and expressed in *E. coli*; however, only 13 VHH proteins retained their high specificity and affinity for binding to sCD105 in ELISA, so only the 13 VHH proteins were expressed and purified in sufficient quantity for evaluation.

In order to evaluate the capacity of the 13 VHHs to bind CD105 expressed in cells, a direct and simultaneous analysis of multiple VHHs in real time by SPRi was developed by using CD105 expressing cells that were non-adherent and can be assayed in suspension. CD105 is mainly expressed on endothelial cells but is not exclusive to them; it is also expressed in macrophages. The SPRi binding assays was performed with the SC monocyte-macrophage cells that express CD105 and THP-1 monocyte cells that do not express CD105. The SPRi assays were performed with cells incubated in microarrayed VHHs on a biochip and SPRi data showed that anti-CD105 VHHs bind to SC cells but not to THP-1 cells. The SPRi assay shown the feasibility to study the binding of anti-CD105 VHHs to cells on a biochip.

The average of SPRi reflectivity variation among the 13 VHHs at 2 hours of assays was 4.59 with SC cells and was 2.69 with the THP-1 cells. Cell density average was 578 cells/area with SC cells and 367 cells/area with THP-1 cells and the average of the approximation to the apparent association constant was 1.1019 with SC cells and 1.0626 with THP-1 cells. The SPRi assays results suggested that the anti-CD105 VHHs could be distributed in 4 groups according to their specificity and affinity approximation. In the first group, VHH 7, VHH 17 and VHH 31 were ranked because they had the highest affinity

and specificity of all the VHH, well above the average among VHHs, even above the positive control. In the second group, VHH 22, VHH 2 and VHH 3 were placed for having higher affinity and specificity than the average VHH. In the third group, VHH 9, VHH 18, VHH 26, VHH 8 and VHH 20 were placed with affinity and specificity equal or very close to the average VHH. The fourth group included VHH 11 and VHH 1, which had low specificity and affinity.

Since the antigen-antibody binding is dependent of the conformation and folding of the antibody, the VHHs of the first two groups would have the most efficient conformation for interaction with CD105. The six top-ranked anti- CD105 VHHs, that bind with high affinity and specificity to CD105 in solution and in membrane bound cells, can be used both in studies that seek a better understanding of the CD105 function *in vitro* and *in vivo* and in application as nanoprobes for the diagnosis and therapeutics of diseases where angiogenesis plays an important role.

V.2 Perspectives

The present work shows the potential of anti-CD105 VHHs for application in biochips to study the function of this co-receptor in cells in vitro and in vivo. Nonetheless, it is interesting that the capacity of the immunization with a complex preparation of antigens can be a source to generate as cDNA library with a collection of VHHs that bind different cellular antigens, making the library amenable of further screening for other antigens of interest. The VHH cDNA library representation, *i.e*. the population of VHHs that recognize and bind to different proteins expressed by T24 cells, was not evaluated yet but suggests the feasibility of using VHH in microarrays to screen with high processivity a large number of clones, which might be eventually developed to identify other cellular targets. In this context, the automation of biochips microarrayed with VHHs from a library of high complexity as prepared by immunization with a lysate of cells and simultaneous assays for binding by SPRi might speed up the identification of VHHs that bind to over-expressed proteins in cells. This is particularly important for the discovery of both novel targets by comparing transformed vs normal cells, cells responding to pathogen, drugs or other substances and as well to environmental changes, and the corresponding target-binding antibodies.

Anti-CD105 VHHs studied in the present work can be evaluated for the design of biosensors that allow the detection of soluble or cell-bound CD105 for the diagnosis or prognosis of diseases involving CD105 increased expression on cells or high levels of the protein in the bloodstream such as pre-eclampsia, hypertension, and some types of cancer or metastasis. VHHs are advantageous as are amenable of gene editing by site directed mutagenesis, Crisp-Casp9 or other related technologies that can improve the binding, half-life, stability and enrich the function in a cellular context by incorporating sequences in the VHH that encode signals, effectors or other functional domains. Furthermore, the attachment of drugs or toxins to anti-CD105 VHH can assess their application in cell-target therapeutics, diminishing collateral damage to other cells that do not require treatment.

The regulation of angiogenesis by using anti-CD105 VHHs in Human Umbilical Vein Endothelial Cells (HUVEC) in vitro would provide information about the potential aplication of these VHHs in CD105 cancer therapy by targeting highly proliferative endothelial cells, to inhibit metastasis or induce tumour shrinkage by preventing vital nutrient supply and waste exchange of tumour cells. In addition, anti-CD105 VHHs assayed on endothelial cells by SPRi shows the potential of analysis of different VHHs against different targets in simultaneous and parallel assays providing information of the effect of both single and multiple VHHs raised against different targets on normal and transformed cells. Useful information as the treatment of cancer requires the concerted action of a combination of biomolecules acting on different targets in transformed cells,

On the other hand, it is possible to explore the application of such anti-CD105 VHH in bone tissue engineering as attachment molecules for mesenchymal stem cells (MSCs) to the scaffold. Tissue engineering is aimed at repairing or partially or totally replacing tissues damaged by trauma or disease. Basically, cells, scaffold, growth factors and extracellular matrix are needed[1]. The most commonly used cells for bone regeneration are MSCs because of their multipotentiality to differentiate or transdifferentiate into a variety of specialised cells *in vivo* and *in vitro*[2,3]. The scaffold must be biodegradable and similar or better to the tissue to be repaired so that it can be combined with living cells[4]. Thus, bone cells must be cultured on three-dimensional scaffolds that allow the formation of 3D adhesions and resemble the morphology of the tissue to be replaced. To capture and fix the MSCs in the scaffold so that they can differentiate into bone tissue, antibodies specific for MSC cell markers are attached to the scaffold.

The size of conventional antibodies may hinder cell capture at inaccessible locations on the scaffold, so the application of VHH as bridges between the cells to be fixed and the scaffold is an alternative. Since CD105 is present on the surface of MSCs[5,6] it is an important cell marker to identify and isolate MSCs[7]. The scaffold-bound anti-CD105 VHHs would specifically recognise and bind CD105 on MSCs, so that with growth factors for osteogenesis the cells can differentiate into bone tissue to be repaired. Therefore, the use of anti-CD105 VHHs as a cell attachment tool would aid the development of three-dimensional scaffolds suitable for bone regeneration.

A: Annexes

A.1 Materials

A.1.1 Microorganisms

A.1.1.1 Bacterial strains

- ***E. coli* TG1** (Sigma Aldrich)

  Used for antibody phage display or peptides / proteins phage display creation library. This strain has an amber codon suppression (*supE*)) important for cloning and preparation of phage libraries. TG1 features include: [F' *traD36 proAB lacIqZ ΔM15] supE thi-1 Δ(lac-proAB) Δ(mcrB-hsdSM)5(rK - mK -)*

- ***E. coli* DH5α** (Thermo Scientific™)

  Used for plasmid maintenance and propagation. Cells are suitable for high efficiency transformation in applications such as plasmid isolation, cloning, and subcloning. DH5α features include: *fhuA2 argF-lac(del)U169 phoA glnV44 Φ80' lacZ(del)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17*[8].

- ***E. coli* BL21 (DE3)-Novagen** (Sigma Aldrich)

  Host for expression of genes donated on pET plasmids. Lysogenic for bacteriophage ƛDE3 carrying the phage T7 RNA polymerase gene linked to the IPTG-inducible promoter under the control of lacUVS, for use with any expression plasmid containing the T7 promoter. Its features include *fhuA2 [lon]* F – *ompT gal (λ DE3) [dcm] ΔhsdS, λ DE3 = λ sBamHIo ΔEcoRI-B int::(lacI::PlacUV5::T7 gene1) i21 Δnin5*[9].

A.1.1.2 VCS-M13 Helper Phage (Antibody Design Labs)

It is a derivative of M13KO7 originally marketed by Stratagene and often used in phage display. M13KO7 is a derivative of M13, with the origin of replication of P15A and the kanamycin resistance gene of Tn903, inserted in domain B of the M13 origin of replication. The kanamycin resistance gene allows the selection of cells superinfected by the helper. It is able to replicate in the absence of phagemid DNA. In the presence of a phagemid bearing an M13 or wild-type f1 origin, the single-stranded phagemid is preferentially packaged and secreted into the culture medium. It is a remarkable auxiliary phage with a high level production of particles containing ssDNA phagemid for DNA sequencing, mutagenesis and phage display[10]. M13 is a non-lytic filamentous bacteriophage of circular single-stranded DNA encapsulated in a protein coat of approximately 2700 copies of the larger coating protein P8, and hooded at the ends by five copies of two different minor coat proteins (P9, P6, P3). The minor coating protein P3 allows the phage to bind to a receptor present at the tip of the host pilli *E. coli*. The p3 protein consists of N-terminal N1 and N2 domains involved in phage infectivity and a C-terminal (CT) domain essential for phage assembly. Importantly, fusion of proteins with the N-terminus of N1 does not suppress the functional activity of the p3 protein.

A.1.2 Cell lines

A.1.2.1 THP-1 (ATCC® TIB-202™)

Cell type with monocyte morphology obtained from human peripheral blood of a male patient with acute monocytic leukaemia. Cells grow in suspension as non-adherent cells. Antigen expression HLA A2, A9, B5, DRw1, DRw2[11]. The cells are phagocytic (for both latex beads and sensitized erythrocytes) and lack surface and cytoplasmic immunoglobulin.

A.1.2.2 SC (ATCC® CRL-9855 TM)

Cell type monocyte, macrophage morphology. The SC cell line was established from human peripheral blood mononuclear cells. As non-adherent cells, they grow as single cells in suspension and have a granular appearance. Expression of macrophage surface proteins such as CD105, vascular epithelial growth factor A and TGF-β1[12,13].

A.1.3 DNA

A.1.3.1 Vectors

A.1.3.1.1 Phagemid pHEN2

Donated by Dr. Daniel Baty through Dr. Kastelic (Protein Biotechnology Group, Babraham Bioscience Technologies, UK). It is a phagemid vector, a plasmid containing a phage origin of replication and a morphogenetic signal, both necessary for phagemid packaging in phage particles. It has the sequence of alkaline phosphatase between the *Sfi*I and *Not*I restriction sites, this place is where the recombinant proteins are inserted. Plasmids bearing the intergenic region of filamentous phage (oriF1) can be packaged as ssDNA in viral particles in the presence of a phage[14]. When wild-type phages are used, the interference of plasmid with phage replication leads to a reduction in phage copy number and a drastic decrease in virion production[15]. Helper phages are designed to overcome interference, maximize virion production and maintain packaging of their own ssDNA at a low level. After transformation with a phagemid library, the bacteria need to be infected by phage helper such as VCS-M13, which provides all the genes required for phage assembly.

A.1.3.1.2 pET22b(+)(Novagen)

Expression vector in *E. coli* host cells that contain on their chromosome a copy of the T7 RNA polymerase gene under control of IPTG inducible lacUV5 promoter. It has a lactose operator associated with the T7 promoter and a sequence encoding the lac repressor capable of repressing lacUVS and T7 promoters in the absence of IPTG. It carries a signal sequence that fuses to the sequence of the protein under study, so that, once expressed, it allows its displacement towards the periplasm to facilitate purification and the formation of disulfide bonds, and

a sequence that expresses a hexahistidine tail, which is added to the protein for purification by affinity chromatography[9].

A.1.3.2 Primers

- AHis 5´ CAACTTTCAACAGTCTAGCTCCC 3´
- M13R 5´ AGCGGATAACAATTTCACACAGG 3´
- T7F 5´ TAATACGACTCACTATAG 3´
- T7R 5´ GCTAGTTATTGCTCAGCG 3´
- VHHSfi(*Nco*I)
  5´TCGCGGCCCAGCCGGCCATGGCKCAGKTGCAGCTCGTGGAGTCNGG 3´
- ALLVHHR2-Not 5´ TTGCGGCCGCTTGTGGTTTTGGTGTCTTGGG 3´

A.1.3.3 DNA molecular weight size marker

- Gene Ruler 100 bp Plus DNA ladder (ThermoScientific),
- DNA ladder 123 bp (Sigma)
- λ/HindIII (Fermentas)

A.1.4 Proteins and enzyme-conjugated antibodies

A.1.4.1 Proteins

- CD105 (Endoglin) was purchased from R&D Systems as Recombinant Human Endoglin/CD105 Protein (Catalogue Number 1097 EN). The lyophilised protein (25 ug) was resuspended in PBS 1X pH 7.4 to a concentration of 25 ug/ml.
- Recombinant Fas2 protein produced in the UBM laboratory, LID, UPCH.
- BSA (Bovine Serum Albumin) lyophilized powder, suitable for cell culture (Sigma- Aldrich).
- FBS (Fetal Bovine Serum), sterile-filtered, suitable for cell culture (Sigma-Aldrich).
- Whole milk powder (Sigma- Aldrich).

A.1.4.2 Enzymes

Each enzyme used its corresponding buffers

- *Not*I (New England Biolabs)
- *Nco*I (New England Biolabs)
- *BgI*II (New England Biolabs)
- Platinum Taq DNA (Invitrogen).
- Recombinant bacteriophage T4 ligase (New England Biolabs)

A.1.4.3 Enzyme-conjugated antibodies

- M13 phage coat protein g8p Monoclonal Antibody (RL-ph2) (Invitrogen ThermoFisher Scientific)

- Protein A-HRP conjugate (Sigma)

A.1.4.4 Protein molecular weight size marker

- Kaleidoscope™ Prestained Standards (BioRad).
- Protein Standard 66kDa Albumin, 45kDa Ovalbumin, 36kDa Glyceraldehyde 3P, 29kDa Carbonic Anhydrase, 14.2kDa a-lacto albumin.

A.1.5 Growth media and solutions

A.1.5.1 Growth media

A.1.5.1.1 Microbiological culture

They were performed as established by Sambrook *et al*[16].

- **2xYT medium:** 16 g of Tryptone, 10 g of yeast extract, 5 g of NaCl were dissolved in 1 liter of distilled water. It was autoclaved at 121° C for 20 minutes. It was stored at 4° C.
- **2xYT medium supplemented (2xYTsupp):** 2xYT medium supplemented with 100 ug ml^{-1} ampicillin and 2% Glucose: 100 µl of filtered Ampicillin 100 $mg.ml^{-1}$ and 10 ml of sterile 20% Glucose were added to 100 ml of sterile 2xYT medium at RT°.
- **2xYT supp plates:** 2xYT plates supplemented with 100 $ug.ml^{-1}$ ampicillin and 2% Glucose: 15 g of agar per liter of 2xYT medium was added. It was autoclaved for 20 minutes. 1 ml of ampicillin 100 mg ml^{-1} and 100 ml of 20% glucose were added before distributing the medium in the plates (when the temperature was between 50° to 55° C).
- **LB medium:** 10 g of Tryptone, 5 g of yeast extract, 10 g of NaCl were dissolved in 1 liter of distilled water. It was autoclaved at 121° C for 20 minutes. It was stored at 4° C.
- **SOB medium:** 20 g of Tryptone, 5 g of yeast extract, 0.5 g of NaCl and 10 ml of 0.5M KCl were dissolved in 1 liter of distilled water. It was autoclaved at 121° C for 20 minutes. It was stored at 4° C.
- **SOC medium:** In 975 ml of sterile SOB medium, 20 ml of 1M Glucose and 5 ml of $MgCl_2$ were added immediately before use (the solutions were previously sterilised by filtration).
- **LB medium supplemented (LB supp):** 100 ul of freshly prepared and filtered Ampicillin 100 $mg.ml^{-1}$ was added to 100 ml of sterile LB medium at RT° to obtain LB medium supplemented with ampicillin 100 $ug.ml^{-1}$ and 150 ul of Ampicillin 10 $mg.ml^{-1}$ was added to obtain LB medium supplemented with ampicillin 150 $ug.ml^{-1}$.
- **LB plates and LB supp plates:** 15 g of agar was added per litre of LB medium. The medium was autoclaved at 121° C for 20 minutes and plated when the temperature was between 50 and 55° C. For LB supp plates, ampicillin was

added just before distributing the medium on the plates according to the required concentration.

A.1.5.1.2 Cell culture

- Iscove's Modified Dulbecco's Medium (IMDM) (ATCC 30®-2005™) containing 4 mM L-glutamine, 4.5 $g.L^{-1}$ glucose and 1.5 $g.L^{-1}$ sodium bicarbonate.
- IMDM complete growth medium: 90% of IMDM supplemented with 0.05 mM 2- mercaptoethanol, 0.1 mM hypoxanthine, 0.016 mM thymidine, 1X PenStrep and 10% of FBS.
- IMDM cryopreservation medium: IMDM complete growth medium with 10% (v/v) DMSO
- Roswell Park Memorial Institute (RPMI 1640) medium (ATCC 30®-2001™) containing 2 mM L-glutamine, 10 mM HEPES, 1 mM sodium pyruvate, 4.5 $g.L^{-1}$ glucose and 1.5 $g.L^{-1}$ sodium bicarbonate.
- RPMI complete growth medium: 90% of RPMI supplemented with 1X PenStrep and 10% of FBS.
- RPMI cryopreservation medium: RPMI complete growth medium with 10% (v/v) DMSO

A.1.5.2 Solutions

A.1.5.2.1 Antibiotics

- 100 $mg.ml^{-1}$Ampicillin: 1 g of Ampicillin (Sigma Aldrich) in 10 ml of ultrapure water. It was sterilised by filtration through a 0.22 µM millipore filter and distributed in 1 ml aliquots that were stored at −20° C.
- 25 $mg.ml^{-1}$ Kanamycin: 0.25 g of Kanamycin (Sigma Aldrich) in 10 ml of ultrapure water. It was sterilised by filtration through a 0.22 µM millipore filter and distributed in 1 ml aliquots that were stored at −20° C.
- 100 X Penicillin Streptomycin solution (PenStrep) (Sigma –Aldrich): 10,000 units penicillin and 10 mg $streptomycin.mL^{-1}$ solubilized in a proprietary citrate buffer.

A.1.5.2.2 1M IPTG (Isopropyl-β-D-1-thiogalactopyranoside)

2.38 g of IPTG (Sigma-Aldrich) in 10 ml ultrapure water. It was sterilised through 0.22 µM millipore filter and distributed in 1 ml aliquots. Stored at −20° C.

A.1.5.2.3 20% glucose

2 g of D-glucose (Sigma-Aldrich) in 10 ml ultrapure water. It was sterilised through a 0.4 µM millipore filter and distributed in 1 ml aliquots that were stored at −20° C.

A.1.5.2.4 0.1 M hypoxanthine

136.11 mg of hypoxanthine (Sigma-Aldrich) in 10 ml ultrapure water. It was sterilised by filtration through a 0.22 µM millipore filter and distributed in 1 ml aliquots that were stored at −20° C.

A.1.5.2.5 0.016 M thymidine

38.76 mg of thymidine (Sigma-Aldrich) in 10 ml ultrapure water. It was sterilised by filtration through a 0.22 µM millipore filter and distributed in 1 ml aliquots that were stored at −20° C.

A.1.6. Buffers

They were performed as established by Sambrook *et al*[16].

A.1.6.1 Electrophoresis

- Buffer TE: 10 mM Tris-HCl, 1 mM EDTA. pH 8.0
- Buffer TBE 1X: 89 mM Tris-borate, 2 mM EDTA, pH 8.0.
- Buffer Tris-Glycine: 25 mM Tris-HCl, 125 mM Glycine, 0.1%, SDS pH 8.3.

A.1.6.2 ELISA

- Bicarbonate carbonate buffer: 0.015 M Na_2CO_3, 0.035 M $NaHCO_3$, pH 9.6
- Phosphate-Buffered Saline (PBS): 137mM NaCl, 2.7mM KCl, 10mM Na_2HPO_4, 1.8mM KH_2PO_4 pH 7.4
- Washing solution: PBS, 0.05% Tween
- Blocking solution: PBS, 0.05% Tween, 2% BSA
- Phage-ELISA blocking solution: PBS, 0.05% Tween, 2% Milk
- Concentrated phage-ELISA blocking solution: PBS, 0.05% Tween, 4% Milk

A.1.6.3 Plasmid DNA extraction (Wizard plus SV Minipreps DNA Purification System (PROMEGA)

- Cell Resuspension Solution: 50mM Tris-HCl pH 7.5, 10mM EDTA, 100µg/ml RNasa A
- Cell Lysis Solution: 1% SDS, 0.2 N NaOH
- Neutralization Solution: 4.09M guanidine hydrochloride(Gu-HCl), 0.759M potassium acetate, 2.12M glacial acetic acid

A.1.6.4 DNA gel extraction (QIAquick Gel Extraction Kit[17])

- Buffer QG: 5.5 M guanidine thiocyanate (GuSCN), 20 mM Tris HCl pH 6.6
- Buffer PB: 5 M Gu-HCl, 30% isopropanol
- Buffer PE: 10 mM Tris-HCl pH 7.5, 80% ethanol
- Buffer EB: 10 mM Tris·HCl, pH 8.5)

A.1.6.5 Protein purification

pH was adjusted to 8.0 using NaOH and sterilised with filter 0.2 or 0.45 µm.

A.1.6.5.1 Spin column procedure (Ni-NTA Spin Kit [18])

- NPI-10 (Binding/lysis buffer for native conditions): 50 mM NaH_2PO_4, 300 mM NaCl, 10 mM imidazole
- NPI-20 (Wash buffer for native conditions): 50 mM NaH_2PO_4, 300 mM NaCl, 20 mM imidazole μm).
- NPI-500 (Elution buffer for native conditions) 50 mM NaH_2PO_4, 300 mM NaCl, 500 mM imidazole

A.1.6.5.2 Batch and column procedures

They were performed as established by Qiagen[19]

- Equilibrium solution: 20mM NaH_2PO_4, 300mM NaCl and 15mM imidazole
- Wash solution: 20 mM NaH_2PO_4, 300 mM NaCl and 25 mM imidazole
- Elution solution: 20mM NaH2PO4, 300mM NaCl and 300mM imidazole

A.1.7 Kits and reagents

A.1.7.1 Kits

- Wizard plus SV Minipreps DNA Purification System (PROMEGA)
- QIAquick Gel Extraction Kit (Qiagen)
- Quant-iT™Protein Assays (Thermo Fisher Scientific)
- Penta His HRP Conjugate Kit (Qiagen)
- Ni-NTA Spin Kit (Qiagen)

A.1.7.2 Reagents

- Ni-NTA Agarose resine (Qiagen): IMAC matrix for purifying recombinant proteins carrying a His tag. Histidine residues in the His tag bind to the vacant positions in the coordination sphere of the immobilized Ni^{+2}.

A.2 SPRi data for the calculation of the apparent association constant

Assay I (SC cells)

Time (min)	SPRi signal*	CN	CP	VHH1	VHH11	VHH17	VHH18	VHH2	VHH20	VHH22	VHH26	VHH3	VHH31	VHH7	VHH8	VHH9
6.35	initial (Si)	46.7170	46.7512	41.0087	46.5862	45.7984	40.2522	43.3958	47.3902	45.2543	44.6032	46.7202	47.5032	44.0666	42.4324	46.7668
127.60	2h (Sf)	51.4550	52.8415	46.5445	51.747	52.0012	45.9640	49.5831	53.2649	50.9264	50.3484	52.1943	53.5390	50.1152	48.0985	51.8598
	(Sf-Si)	4.738	6.0903	5.5358	5.1608	6.2028	5.7118	6.1873	5.8747	5.6721	5.7452	5.4741	6.0358	6.0486	5.6661	5.0930
App. Assoc. Kte° (Sfi/Si)		1.1014	1.1303	1.1350	1.1108	1.1354	1.1419	1.1426	1.1240	1.1253	1.1288	1.1172	1.1271	1.1373	1.1335	1.1089

Assay II (SC cells)

Time (min)	SPRi signal*	CN	CP	VHH1	VHH11	VHH17	VHH18	VHH2	VHH20	VHH22	VHH26	VHH3	VHH31	VHH7	VHH8	VHH9
27.0	initial (Si)	61.9520	54.3772	58.1242	53.8767	52.7608	57.6454	56.8246	57.5850	56.7104	54.4628	51.1390	51.9870	54.7522	57.7379	52.6298
147.02	2h (Sf)	63.3213	60.1815	60.6480	58.2033	58.3100	59.9243	60.4220	61.8829	60.7665	58.3216	56.8353	58.2290	59.5411	61.0708	57.1518
	(Sf-Si)	1.3693	5.8043	2.5238	4.3266	5.5492	2.2789	3.5974	4.2979	4.0561	3.8589	5.6963	6.2420	4.7889	3.3329	4.5220
App. Assoc. Kte° (Sfi/Si)		1.0221	1.1067	1.0434	1.0803	1.1052	1.0395	1.0633	1.0746	1.0715	1.0709	1.1114	1.1201	1.0875	1.0577	1.0859

Assay III (SC cells)

Time (min)	SPRi signal*	CN	CP	VHH1	VHH11	VHH17	VHH18	VHH2	VHH20	VHH22	VHH26	VHH3	VHH31	VHH7	VHH8	VHH9
23.3	initial (Si)	33.0348	41.5532	33.7518	40.2228	39.5014	38.4785	37.6639	38.4496	36.7921	38.2719	42.0693	41.9939	37.4463	35.6009	40.9111
143.31	2h (Sf)	35.0374	46.2883	36.4652	43.6674	43.8761	43.2651	41.3717	42.3844	41.4032	41.2159	45.2933	44.2892	42.8139	39.9278	44.8243
	(Sf-Si)	2.0026	4.7351	2.7134	3.4446	4.3747	4.7866	3.7078	3.9348	4.6111	2.9440	3.2240	2.2953	5.3676	4.3269	3.9132
App. Assoc. Kte° (Sfi/Si)		1.0606	1.1140	1.0804	1.0856	1.1107	1.1244	1.0984	1.1023	1.1253	1.0769	1.0766	1.0547	1.1433	1.1215	1.0957

*reflectivity variation

°approximation to the apparent association constant

Assay I (THP-1 cells)

Time (min)	SPRi signal*	CN	CP	VHH1	VHH11	VHH17	VHH18	VHH2	VHH20	VHH22	VHH26	VHH3	VHH31	VHH7	VHH8	VHH9
27	initial (Si)	46.7336	35.3431	37.1042	41.6704	43.6732	42.0124	40.1035	42.3415	41.1779	41.8077	45.0003	46.4832	42.1051	39.1357	43.6002
145.32	2h (Sf)	49.0027	38.8822	40.5602	46.0218	47.1048	46.3251	43.8576	47.0991	45.1254	45.2326	48.8817	49.7296	46.2677	43.1429	47.4903
	(Sf-Si)	2.2691	3.5391	3.4560	4.3514	3.4316	4.3127	3.7541	4.7576	3.9475	3.4249	3.8814	3.2464	4.1626	4.0072	3.8901
App. Assoc. Kte° (Sfi/Si)		1.0486	1.1001	1.0931	1.1044	1.0786	1.1027	1.0936	1.1124	1.0959	1.0819	1.0863	1.0698	1.0989	1.1024	1.0892

Assay II (THP-1 cells)

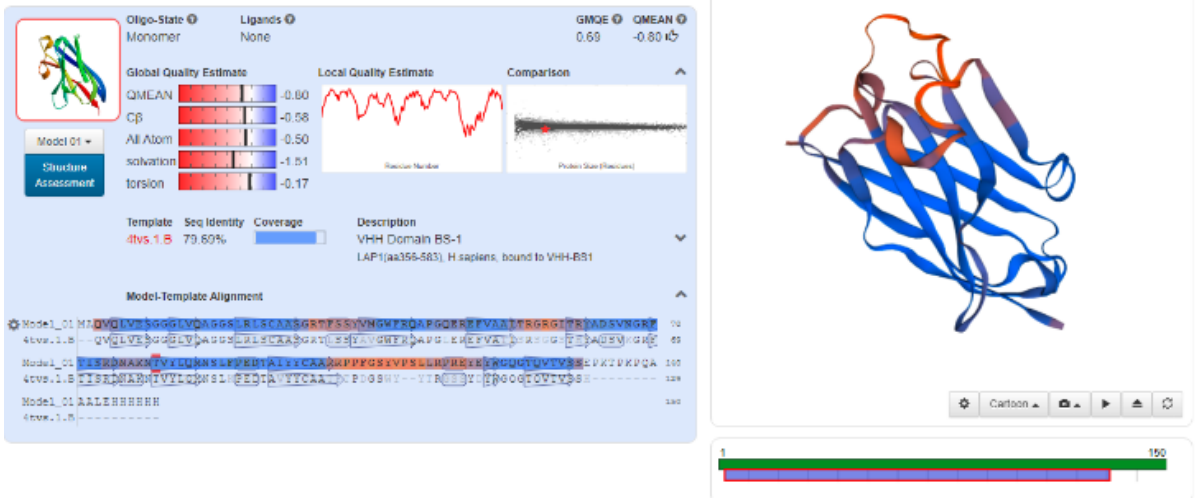
Time (min)	SPRi signal*	CN	CP	VHH1	VHH11	VHH17	VHH18	VHH2	VHH20	VHH22	VHH26	VHH3	VHH31	VHH7	VHH8	VHH9
20.25	initial (Si)	38.0633	48.0582	39.3748	43.4767	45.4243	41.9064	42.3092	44.8741	42.0664	42.8787	45.4188	46.6113	43.7177	42.7719	44.4516
140.25	2h (Sf)	38.9002	49.2886	41.3060	44.7704	46.6542	43.3140	43.2840	46.2185	43.0718	44.3599	46.9490	47.8256	45.0090	44.6237	45.9729
	(Sf-Si)	0.8369	1.2304	1.9312	1.2937	1.2299	1.4076	0.9748	1.3444	1.0054	1.4812	1.5302	1.2143	1.2913	1.8518	1.5213
App. Assoc. Kte° (Sfi/Si)		1.0220	1.0256	1.0490	1.0298	1.0271	1.0336	1.0230	1.0300	1.0239	1.0345	1.0337	1.0261	1.0295	1.0433	1.0342

*reflectivity variation
°approximation to the apparent association constant

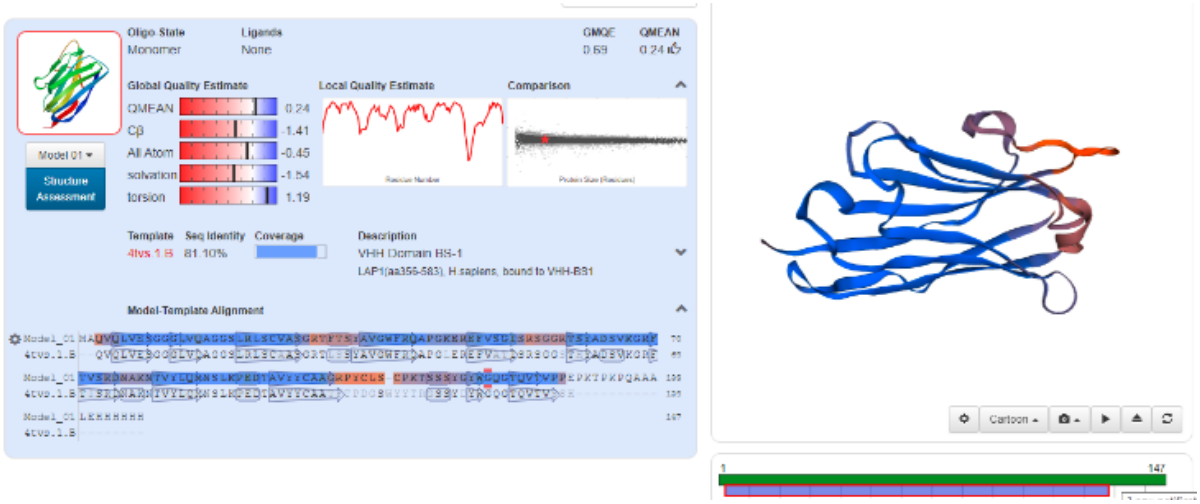
A.3 Images of the three-dimensional structures of the VHHs modeled by Swiss Model

- **SeqVHH-4/H8-II**

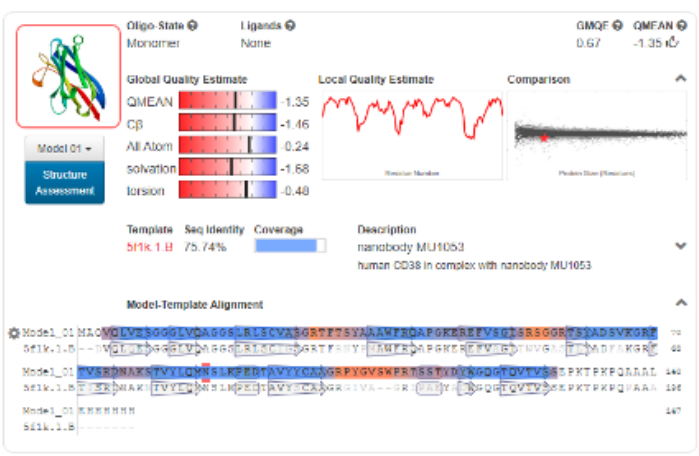
- **SeqVHH-5/A1-II**



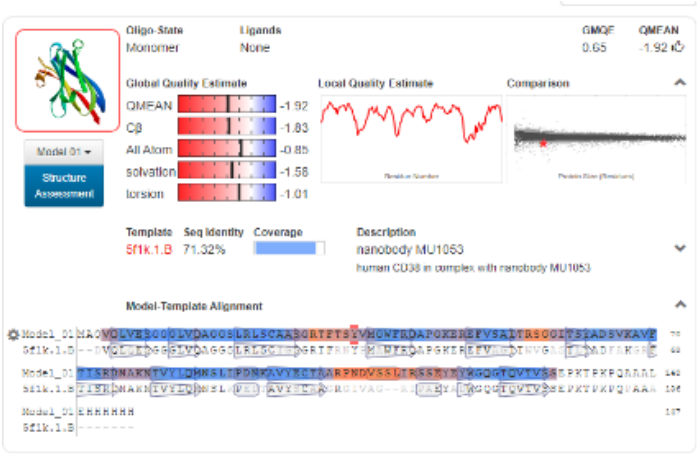
- **SeqVHH-6/D4-II**

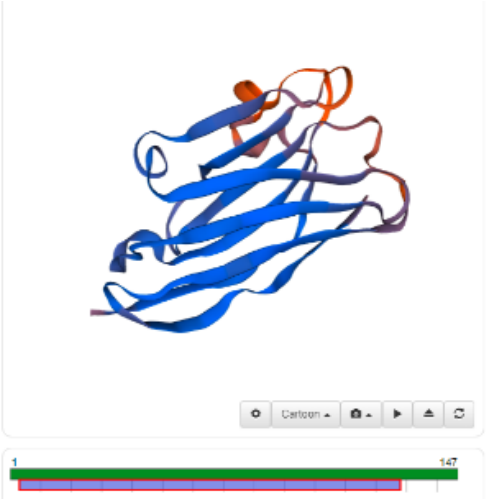


- **SeqVHH-7/A4-II**

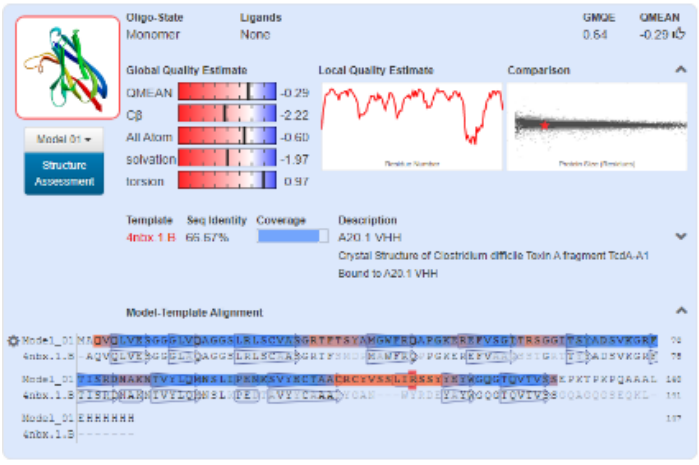


- **SeqVHH-8/A3-II**

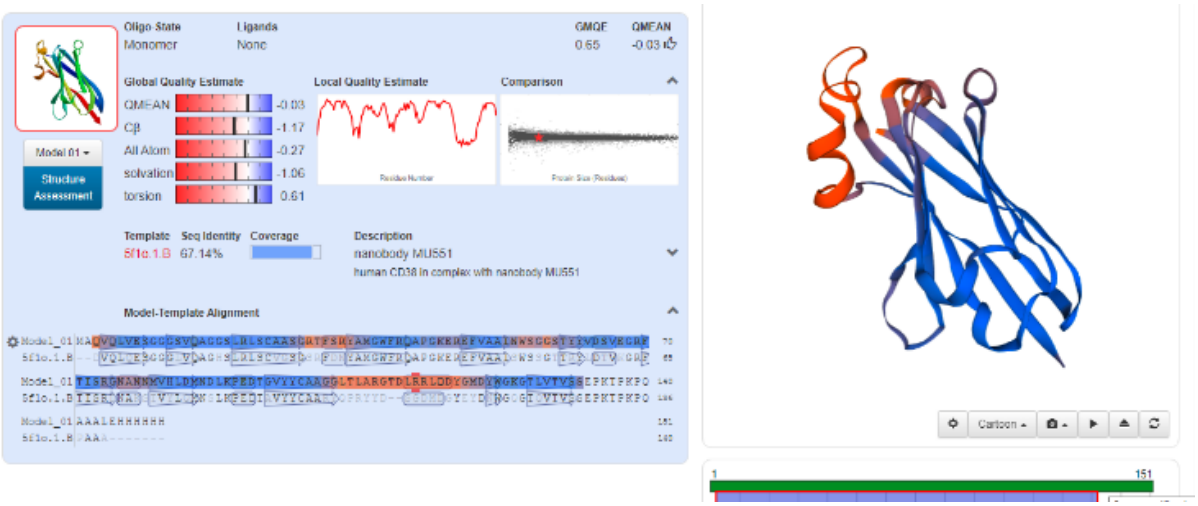




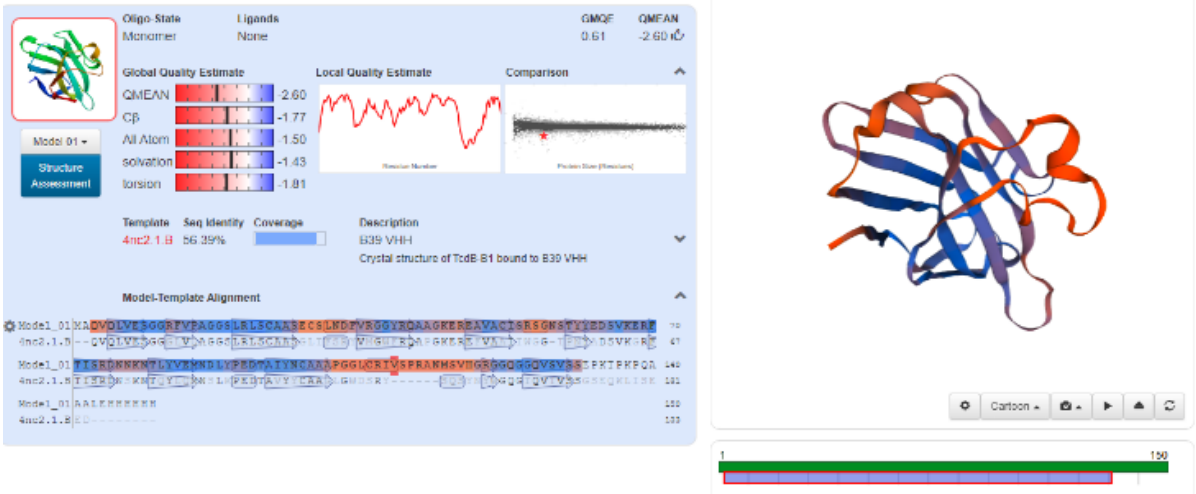
- **SeqVHH-9/D2-III**



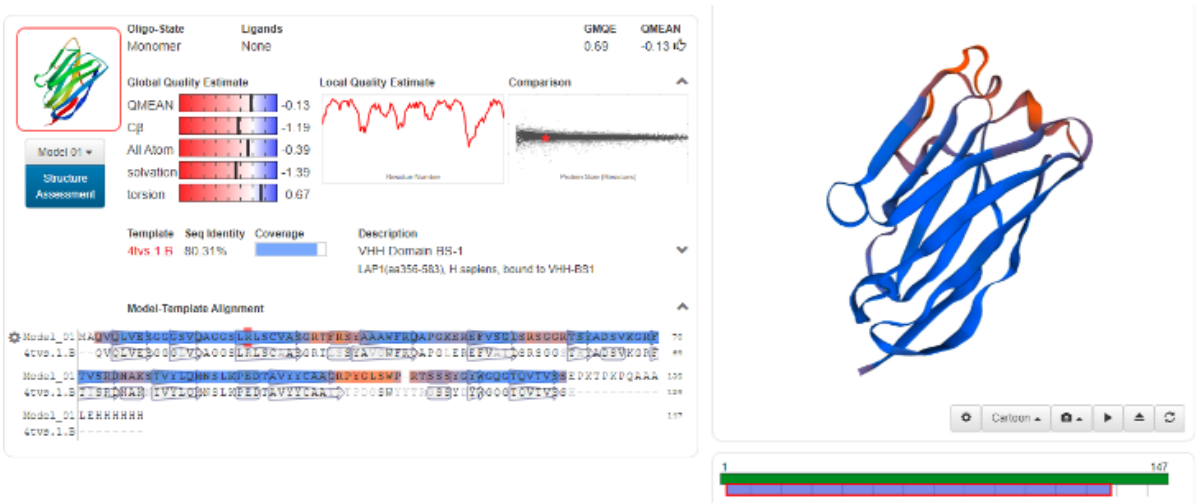
- **SeqVHH-10/B4-III**



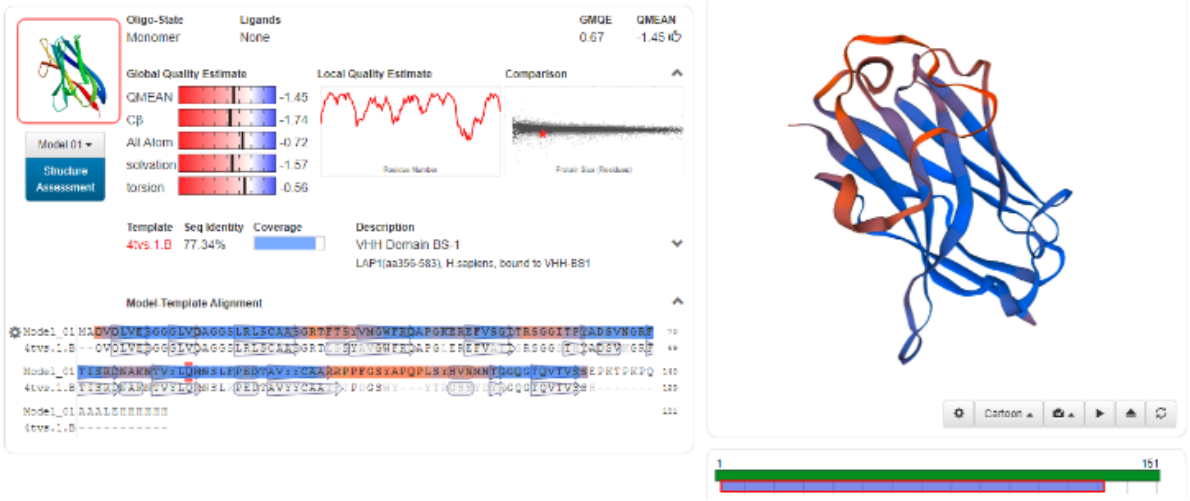
- **SeqVHH-11/H2-III**



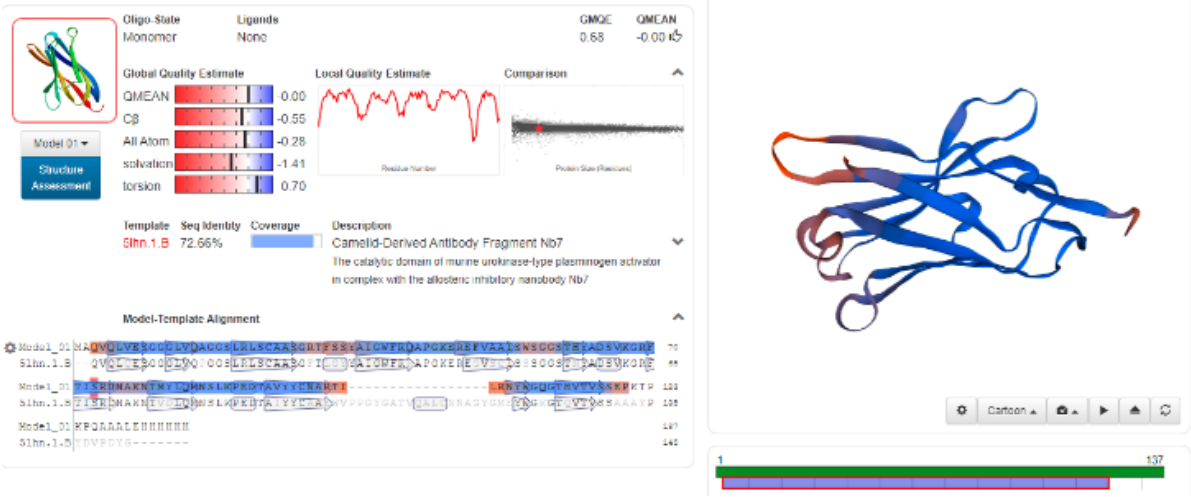
- **SeqVHH-12/C2-III**



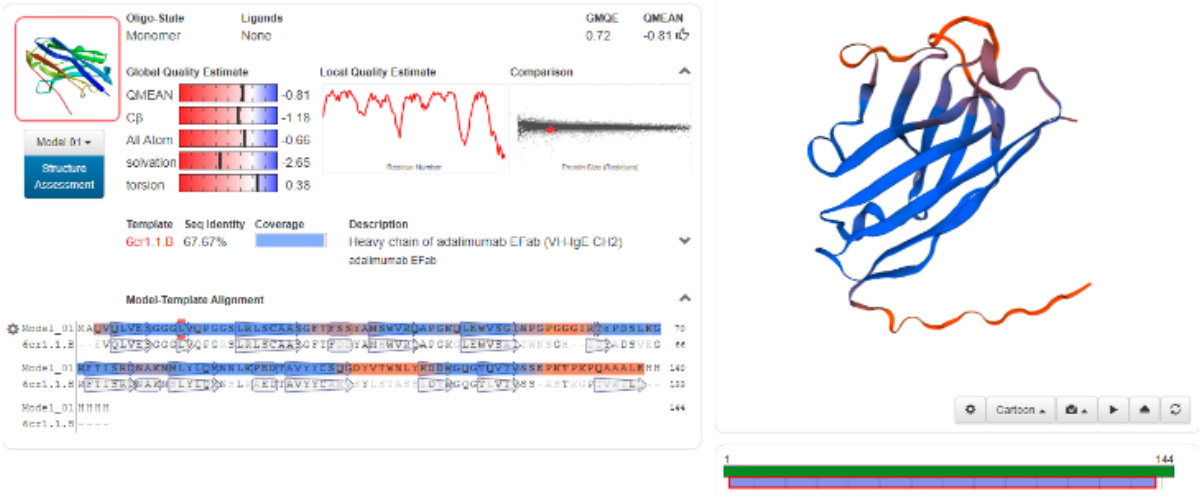
- **SeqVHH-13/E1-III**



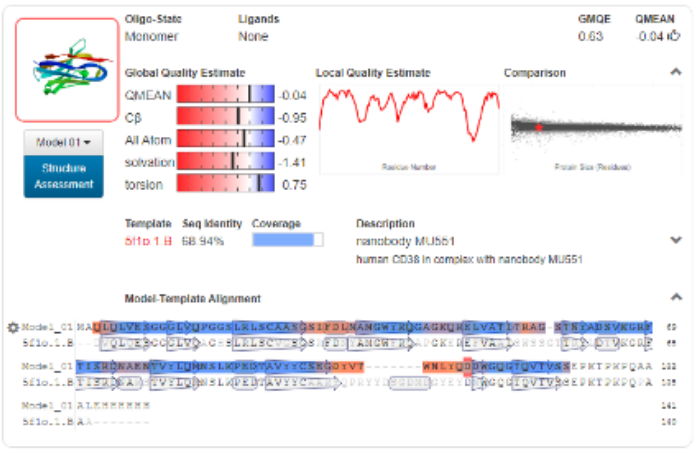
- **SeqVHH-14/F3-III**

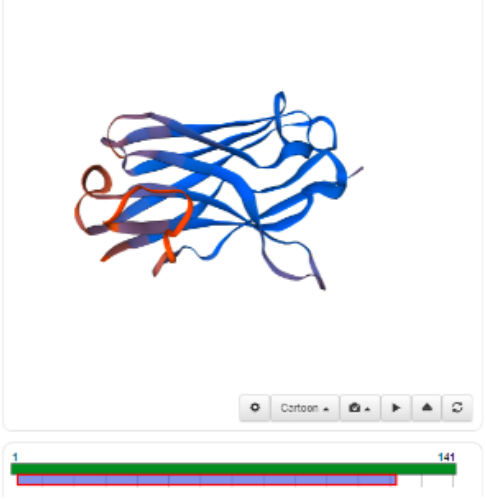


- **SeqVHH-15/A1-IV**

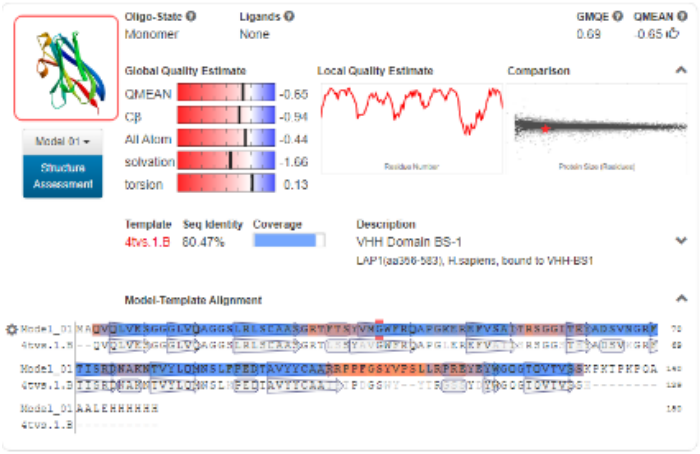


- **SeqVHH-16/E11-IV**

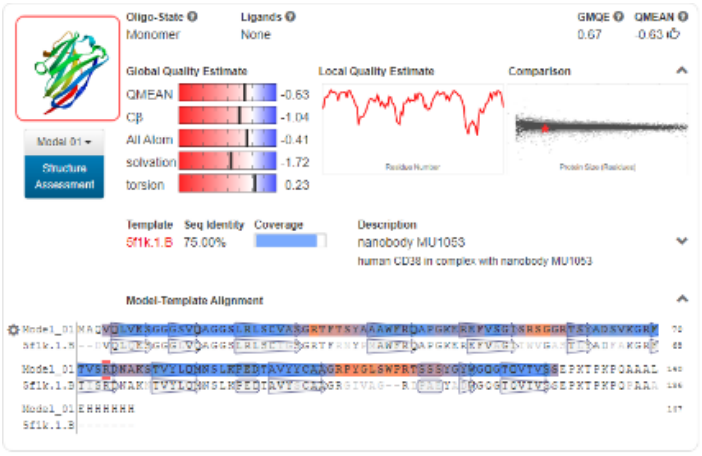




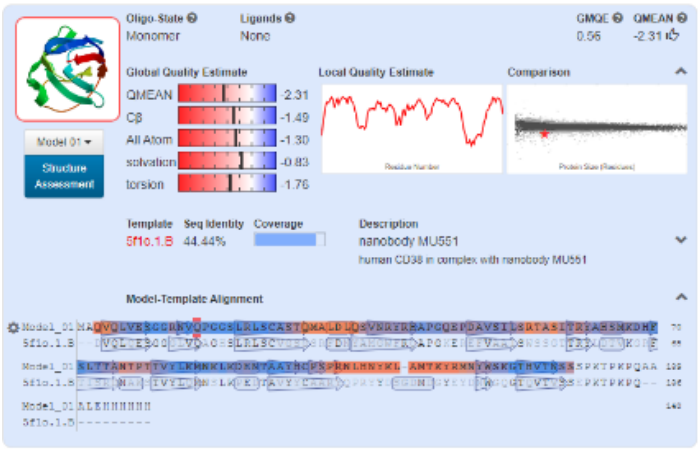
- **SeqVHH-17/E1-IV**

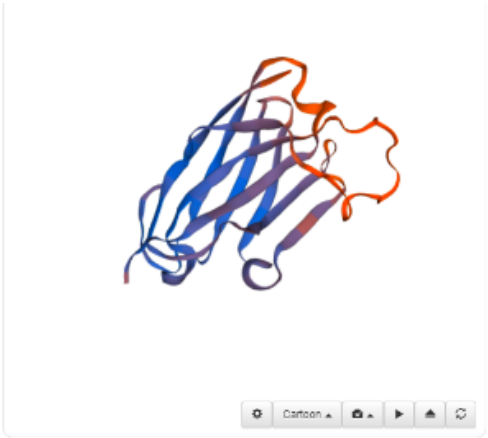


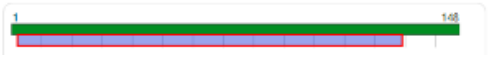
- **SeqVHH-18/D6-IV**



- **SeqVHH-19/E7-IV**







B: Version en français

Chapitre I: Généralités

La fabrication de protéines conçues pour se lier à leur cible avec une affinité et une spécificité élevées s'est inspirée des anticorps naturels. Les camélidés et les requins produisent des anticorps inhabituels, dépourvus de chaînes légères et d'un domaine constant de la chaîne lourde (CH1), pour lesquels ils sont plus petits (95 kDa environ) que les anticorps classiques. Ces anticorps fonctionnels qui ne possèdent que des chaînes lourdes sont appelés anticorps à chaîne lourde (HCAbs)[20]. Les HCAbs sont de type IgG et présentent une structure similaire à celle des chaînes lourdes des anticorps conventionnels. Le domaine variable des HCAbs, VHH (avec un autre H pour le différencier du domaine variable de la chaîne lourde des anticorps conventionnels) est l'aspect le plus important de ces anticorps à chaîne lourde.

Le domaine VHH est un polypeptide d'environ 110 à 140 acides aminés (12-15 kDa) à l'extrémité N-terminale de HCAb où se produit la liaison de l'antigène. Le domaine VHH peut se replier indépendamment du reste de la protéine tout en conservant la capacité de lier l'antigène avec affinité et spécificité[21-23]. Par conséquent, la protéine VHH indépendante peut être produite par recombinaison sans perte significative de ses propriétés de reconnaissance de l'antigène. Les protéines VHH recombinantes (VHHs) sont également connues sous le nom de fragments d'anticorps à domaine unique, de nanoanticorps ou simplement de nanocorps, "antibodies", comme ils étaient initialement appelés.

Les VHHs possèdent diverses propriétés uniques qui les définissent comme des outils de recherche puissants et prometteurs, distincts des anticorps, offrant une alternative plus pratique. Premièrement, les VHHs sont plus petits que les anticorps (presque 10 fois moins que les IgG), ils ont donc la capacité de pénétrer dans des sites inaccessibles aux anticorps, un attribut attrayant pour une utilisation en médecine. Deuxièmement, leur spécificité et leur affinité élevées peuvent être évaluées à l'aide d'un large éventail d'approches allant de l'exposition sur phage (phage display) à l'imagerie de la résonance des plasmons de surface (SPRi). Troisièmement, les VHHs sont structurellement et biochimiquement stables et très solubles, avec une faible immunoréactivité croisée et une forte homologie avec la famille des gènes VH3 humains. Enfin, les VHHs peuvent être clonés, modifiés génétiquement ou chimiquement, et produits par recombinaison dans diverses cellules avec un coût de production relativement faible. Les systèmes d'expression bactérienne permettent de produire des VHHs purifiés en quantités de l'ordre de milligrammes par litre de culture, offrant ainsi un système d'approvisionnement illimité[23]. À cet égard, les bibliothèques de VHH sont produites par amplification directe de l'ARNm des lymphocytes et par amplification de l'ADNc avec des amorces spécifiques du VHH, puis sélectionnées par phage display. Ainsi, un large panel de peptides ou de protéines peut être évalué grâce à l'utilisation des VHHs. Ces attributs des VHHs ont augmenté leur utilisation dans la recherche fondamentale et médicale, ainsi que dans les applications biotechnologiques[24-26].

CD105, également connu sous le nom d'endogline, est une glycoprotéine membranaire de type I située à la surface des cellules. C'est un récepteur accessoire de type III du facteur de croissance transformant β (TGF-β) qui se lie aux isoformes TGF-β1 et TGF-β3, ainsi qu'à d'autres membres de la famille TGF- β, sur les cellules endothéliales humaines. Il peut former des complexes avec les récepteurs de type I et II (TGFβR-I et -II), et agit comme un modulateur des interactions du TGF-β avec ses récepteurs de signalisation[27-29]. En 1993, lors du 5e Atelier International sur les Antigènes de Différenciation des Leucocytes Humains (HLDA) à Boston, cette protéine a été désignée sous le nom de Cluster of Differentiation 105 (CD105), en raison de l'augmentation de son expression lors de la différenciation du monocyte en macrophage[30].

La structure de CD105 consiste en un complexe homodimère transmembranaire de 180 kDa dont les sous-unités sont liées par des ponts disulfures. Chaque monomère est constitué de 658 résidus d'acides aminés comprenant un peptide signal de 25 résidus d'acides aminés, un grand domaine extracellulaire de 561 résidus d'acides aminés, une région transmembranaire de 25 résidus d'acides aminés et un domaine cytoplasmique de 47 résidus d'acides aminés riches en sérine (S) et en thréonine (T)[31-33].

Deux isoformes de CD105 sont générées par épissage alternatif, l'isoforme longue principalement exprimée, L-CD105 (658 aa), et l'isoforme courte, S-CD105 (625 aa), toutes deux ont été détectées dans les tissus humains et murins[34]. En outre, l'action protéolytique de la métalloprotéinase MMP-14 (MT1-MMP) sur CD105 complet, donne lieu à une forme soluble de 65 kDa de CD105 qui est libérée dans le milieu extracellulaire et appelée CD105 soluble (sCD105)[27,35,36].

À la fin des années 1980, CD105 a été identifié comme un antigène marqueur des cellules endothéliales humaines[37]. Par la suite, on a signalé que CD105 était exprimé sur les cellules hématopoïétiques (précurseurs des cellules B[38], proérythroblastes[39], cellules de la lignée myéloïde, macrophages[40], cellules stromales de la moelle osseuse[41]). syncytiotrophoblastes placentaires[42], fibroblastes[43], chondrocytes du cartilage[44], cellules mésangiales du rein[45], cellules stellaires du foie[46], kératinocytes épidermiques des follicules pileux et des glandes sudoripares[47].

CD105 est un récepteur auxiliaire de la famille des protéines TGF-β qui jouent un rôle clé dans différents processus physiologiques tels que le développement, la prolifération cellulaire, la synthèse de la matrice extracellulaire, l'angiogenèse ou les réponses immunitaires et leur dérégulation peut conduire au développement de tumeurs. L'expression de CD105 augmente en réponse à certains processus physiologiques ou pathologiques, tels que : la transition du monocyte au macrophage[40], le développement de la prééclampsie[48], dans les processus tumoraux et métastatiques (mélanome[49,50], cancer de l'ovaire[51] et de la prostate[52]) l'athérosclérose et la réparation des plaies dans les cellules musculaires lisses[53], la réparation vasculaire médiée par les cellules sanguines[54], la différenciation myogénique dans le développement embryonnaire[55], l'érythropoïèse dans les cellules progénitrices hématopoïétiques de la moelle osseuse[56], l'ischémie-reperfusion dans le rein et le cœur[57], etc.

La présence de CD105 sur presque tous les types de cellules vasculaires et son expression élevée soulignent le rôle important de CD105 dans la régulation du remodelage vasculaire et de l'angiogenèse dépendant du TGF-β[27,29,33]. L'expression accrue de CD105 sur les cellules endothéliales est associée à: des maladies auto-immunes[58], des lésions épidermiques[59], des dommages endothéliaux[53,60], des processus d'angiogenèse active[61,62], l'embryogenèse[56] et des processus tumoraux[63]. CD105 est actuellement considéré comme un modulateur de la réponse TGF-β avec des fonctions importantes dans le cancer.

Les mutations du gène CD105 (ENG), situé sur le bras long du chromosome 9 humain (9q33- q34.1) sont responsables de la télangiectasie hémorragique héréditaire de type 1 (HHT1)[64], un syndrome caractérisé par une altération génétique des vaisseaux sanguins entraînant des structures vasculaires anormales qui produisent: des télangiectasies (petits vaisseaux sanguins dilatés), des malformations artério-veineuses (AVM), des épistaxis, des saignements, des hémorragies et une anémie[65].

L-CD105 a les effets généralement attribués à CD105 et S-CD105 l'inverse. L'expression de S-CD105 est augmentée dans la sénescence des cellules endothéliales. L'augmentation de S-CD105 entraînerait une réduction de la capacité de dilatation des vaisseaux sanguins en raison de la diminution de l'expression de la eNOS (enzyme oxyde nitrique synthase) et de l'augmentation de la COX2 (cyclooxygénase 2) et contribuerait à la pathologie vasculaire du vieillissement[66].

CD105 soluble (sCD105) semble s'opposer à CD105 lié à la membrane (L-CD105), car il inhibe l'angiogenèse *in vitro* et peut induire une perméabilité vasculaire et une hypertension *in vivo*[35,67-69]. L'augmentation des niveaux de sCD105 dans le sang est associée à des pathologies vasculaires telles que la prééclampsie, l'hypertension, la sclérose systémique, la cardiomyopathie dilatée, l'infarctus aigu du myocarde et l'athérosclérose[69-72]. De même, sCD105 est associé aux pathologies inflammatoires et auto-immunes telles que la sclérose en plaques, le psoriasis et la polyarthrite rhumatoïde[59,73].

Les recherches visant à révéler de nouvelles façons de comprendre les mécanismes impliqués dans les pathologies associées à CD105 sont en développement constant. L'attrait croissant de la protéine CD105 est évident. Il est donc essentiel de créer des outils permettant de l'étudier.

Objectifs du doctorat et contexte coopératif

Au sein de l'Unité de Biologie Moléculaire de l'Universidad Peruana Cayetano Heredia, le groupe Nanobodies dirigé par le Dr. Espinoza développe des recherches visant à la génération de protéines VHH recombinantes à partir d'alpagas immunisés avec des antigènes d'intérêt biomédical, tels que les protéines Fas2 de *Fasciola hepatica* et CD105 humain. Pour obtenir les VHHs, des bibliothèques de phages portant l'ADNc de VHH ont été réalisées par phage display. Ainsi, Gushiken en 2016 a obtenu une bibliothèque de 2,7 x 10^7 phages présentant des VHHs anti-CD105[74]. Pour cela, un alpaga a été immunisé avec un lysat de cellules surexprimant CD105, la lignée cellulaire T24 provenant d'une tumeur de la vessie urinaire, produisant une réponse immunitaire polyclonale contre le lysat T24

et CD105. A partir de l'ARN extrait du sang de l'animal, une bibliothèque d'ADNc de 1,2 x 10^8 séquences de VHH a d'abord été produite puis, par phage display, la bibliothèque de phages portant des VHHs anti-CD105 a été obtenue. Deux VHHs anti-CD105 ont été sélectionnés par phage-ELISA et les structures tridimensionnelles des protéines ont été prédites à partir de leurs séquences.

La nécessité de réaliser d'autres essais permettant un criblage exhaustif de la bibliothèque d'ADNc, de manière à ce que davantage de séquences VHH puissent être capturés pour leur expression en tant que protéines VHH anti-CD105 qui élargissent le répertoire des protéines potentiellement utilisables dans les biocapteurs, a motivé la première partie de cette étude.

L'application des biopuces à protéines basés sur l'imagerie de la résonance des plasmons de surface (SPRi) dans la recherche ont fortement augmenté ces dernières années en raison de leurs avantages, tels que la rapidité et le temps réel, la spécificité et la sensibilité élevées, l'absence de marquage, les petits volumes et la compatibilité avec de nombreux types de matériaux biologiques[75,76]. Et plus particulièrement, pour cette étude, parce qu'il s'agit d'une technologie de biodétection appropriée pour tester les essais d'interaction biomoléculaire tels que l'interaction antigène-anticorps (dans le cas présent, antigène-VHH) et pour la détection d'analytes[77,78] en utilisant un format biopuce[79]. Des biopuces d'anticorps ont été générées pour déterminer le profil des protéines présentes dans des échantillons de fluides et de tissus d'origine variée. Les anticorps imprimés agissent comme des sondes spécifiques pour la détection et la capture de protéines biomarqueurs, telles que celles associées à de nombreuses pathologies comme le cancer, le rhumatisme psoriasique, les agents infectieux, la signalisation, les pathogènes, les virus, etc., dans des réactions à l'échelle du microlitre[80]. La construction de ces biopuces présente des points critiques tels que la conception des biopuces, la stœchiométrie des protéines, l´analyse du signal de détection et le traitement des données, facteurs qui doivent être adaptés aux objectifs des essais et à la nature de l'échantillon[81,82]. Des recherches supplémentaires sont donc nécessaires dans ce domaine.

Le développement de biopuces à protéines pour l'analyse cellulaire par la technologie SPRi prend en compte d'autres aspects pertinents. Le groupe Chimie pour la reconnaissance et l'étude des architectures biologiques (CREAB) de l'unité de recherche Systèmes moléculaires et nanomatériaux pour l'énergie et la santé (SyMMES) du CEA-Grenoble, Université Grenoble Alpes et CNRS en France, où travaille par le Dr Roupioz, a consacré de nombreuses recherches au développement de biopuces et de biocapteurs SPRi en tant que nouveaux outils d'analyse, de diagnostic et d'investigation des processus cellulaires pertinents et de la détection des cellules et bactéries.

La polyvalence de la détection cellulaire par SPRi ouvre plusieurs champs d'application et permet la conception de nouveaux systèmes miniaturisés pour l'analyse de la détection ou les études cinétiques de l'adhésion entre cellules et surfaces. D'autre part, l'attractivité croissante de CD105 est évidente, il est donc essentiel de générer des outils permettant sa recherche et son étude. Le développement d'une biopuce SRPi pour la détection de

l'interaction antigène-anticorps (VHH) entre le récepteur préalablement déposé sur la surface d'or (VHH anti-CD105) et l'analyte complémentaire (en présence de CD105), permettrait d'évaluer en temps réel et sans besoin de marquage, la capacité de reconnaissance et la spécificité de chaque VHH anti-CD105 vis à vis de son antigène exprimé sur les cellules et pas seulement sous forme libre. Cela pourrait déterminer l'application de ces VHHs comme outils pour cribler la capacité des candidats VHHs anti-CD105 disposés sur une biopuce SPRi pour détecter et capturer CD105 sur les cellules qui l'expriment, c'est-à-dire comme nanosondes anti-CD105.

La génération d'un biocapteur SPRi pour la détection de CD105, qui contribuera à l'établissement des bases de futures recherches biomédicales pour aider à la détection, au diagnostic ou au traitement de maladies ou de syndromes où CD105 joue un rôle important ou pour des études qui cherchent à mieux comprendre les mécanismes et la manière dont CD105 participe, a été la motivation de la deuxième partie de cette étude.

Objectif général

Développer une plateforme protéomique de type "biopuce" qui permet l'évaluation rapide en temps réel et la caractérisation des VHHs d'alpaga spécifiques contre CD105 sur les cellules par SPRi pour la sélection des VHHs les plus efficaces dans la détection et la capture de CD105 et l'application de ces VHHs comme nanosondes dans les études de recherche diagnostique et thérapeutique.

Objectifs spécifiques

- Criblage d'une bibliothèque d'ADNc de VHH provenant d'alpagas immunisés avec un lysat de cellules humaines de cancer de la vessie (lignée cellulaire T24 surexprimant CD105) pour la sélection de VHHs spécifiques de CD105.
- Séquençage des ADNc de VHH anti-CD105 pour l'analyse des séquences et la caractérisation *in silico* des protéines VHH recombinantes.
- Sous-clonage des séquences VHH anti-CD105 dans le vecteur pET22b(+) (Qiagen) pour l'expression et la purification des protéines recombinantes par chromatographie d'affinité sur matrice d'Agarose Ni-NTA.
- Fonctionnalisation des protéines VHH anti-CD105 recombinantes sélectionnées pour les placer sur une biopuce afin de permettre la conception et la génération de la plateforme protéomique (protein microarray) pour la détection rapide et en temps réel de CD105.
- Reconnaissance et caractérisation des interactions antigène-anticorps des VHHs anti-CD105 sur la biopuce avec l'antigène sur les cellules qui l'expriment par SPRi. Cela permettrait de déterminer les VHHs avec une affinité et une spécificité plus élevées pour leur application dans les études de recherche, le diagnostic ou le traitement des maladies qui impliquent l'expression de CD105.

Chapitre II: Sélection des VHHs anti-CD105

Une bibliothèque d'ADNc de VHH a été synthétisée à partir d'ARN extrait de cellules mononucléaires du sang périphérique (PBMC) d'alpagas immunisés avec un lysat de cellules humaines du cancer de la vessie (lignée cellulaire T24). La bibliothèque a été criblée avec la protéine sCD105 par phage display, ce qui a donné 376 phages monoclonaux recombinants qui ont été évalués par phage-ELISA.

Dix-neuf phages VHH étaient positifs à sCD105: neuf phages ont montré un signal fort (jusqu'à 35 fois plus élevé que le contrôle) et dix phages ont montré un signal juste au-dessus du seuil pour être considéré comme positif (3 fois plus élevé que le contrôle). Les neuf phages VHH anti-CD105 présentant un signal fort sont: n° 12 (D2-I), n° 48 (H6-I), n° 140 (F6-II), n° 158 (H8-II), n° 200 (D2-III), n° 214 (B4-III), n° 283 (A1-IV), n° 335 (E7-IV) et n° 367 (E11-IV). Les dix phages VHH anti-CD105 présentant un signal avec le ratio minimum sont: n° 95 (A1-II), n° 111 (A3-II), n° 119 (A4-II), n° 122 (D4-II), n° 193 (E1-III), n° 199 (C2-III), n° 204 (H2-III), n° 210 (F3-III), n° 287 (E1-IV) et n° 326 (D6-IV). Cette variation des valeurs de phage-ELISA suggère que le criblage a donné lieu à différents clones de VHH anti-CD105.

Les ADNc des 19 VHHs anti-CD105 ont été amplifiés et séquencés avec des amorces spécifiques pour l´analyse et l´identification. Les séquences d'ADN ont été éditées, traduites et analysées avec les outils d¨ExPASY et GeneDoc. Les caractéristiques physico-chimiques théoriques de chaque protéine VHH prédites *in silico* ont été établies pour mieux les connaître et pouvoir ainsi améliorer la production de protéines VHH recombinantes. En plus, les séquences ont été soumises à BLAST pour déterminer leur identité et leur unicité et SwissModel pour la modélisation de la structure tridimensionnelle de chaque protéine.

L'analyse BLAST a déterminé que les séquences correspondaient en majorité à des séquences d'anticorps de camélidés (principalement lamas et vigognes) avec des pourcentages d'identité entre 73% et 85% et a également corroboré que toutes les séquences codaient pour des protéines différentes, même si elles ne différaient que par un seul acide aminé. Ainsi, les 19 séquences d'ADNc analysées se sont révélées être différentes, présentant des caractéristiques structurelles des VHHs telles que quatre régions charpente (FRs), trois régions déterminant la complémentarité (CDRs), des substitutions dans FR2 qui augmentent l'hydrophilie de la región, un CD3 de 16-18 acides aminés et une structure secondaire conservée consistant en 9 feuillets β et 8 boucles comme modélisé par SwissModel.

Les séquences VHH ont été réparties en groupes en fonction de la plus ou moins grande similarité entre elles. Deux groupes dans lesquels les séquences VHH ne diffèrent que par un seul acide aminé, un groupe composé de VHH2 et VHH15 et l'autre de VHH4, VHH12 et VHH18. Les autres groupes étaient les suivants: VHH 1 et 16 ; VHH 6 et VHH7 ; VHH 5, VHH 17 et VHH13 ; VHH 8 et VHH9; VHH 3, VHH11 et VHH14. Ce dernier groupe a montré une similarité entre eux et le moins de similarité avec le reste des séquences VHH. VHH 10 et VHH 19 étaient les moins similaires parmi toutes les séquences VHH.

Conclusions

La modélisation SwissModel de la structure tridimensionnelle des séquences protéiques basées sur les protéines VHH rapportées dans la Protein Data Bank (PDB), en plus de l'analyse BLAST des séquences et de la caractérisation *in silico* des protéines prédites, a permis d'identifier les 19 séquences différentes comme des ADNc codant pour des protéines VHH.

Chapitre III: Sous-clonage, expression et purification des VHHs anti-CD105

Les dix-neuf ADNc de VHH anti-CD105 différents clonés dans le vecteur phagemid pHEN2 ont été sous-clonés par PCR avec des amorces spécifiques et des enzymes de restriction dans le vecteur plasmidique pET-22b(+) (Novagen) pour être exprimés en tant que protéines recombinantes dans la souche BL21 d'*Escherichia coli*. L'insertion des séquences VHH dans le cadre de lecture correct dans le vecteur a été vérifiée par digestion avec des enzymes de restriction et par amplification par PCR en utilisant deux paires d'amorces spécifiques pour le vecteur et la séquence VHH.

Les conditions d'expression des protéines VHH recombinantes ont été standardisées en faisant varier le temps d'induction, la concentration en inducteur et la concentration en antibiotique. L'expression des protéines VHH par induction avec 1mM IPTG a été visualisée en SDS-PAGE à 15%. Treize protéines VHH étaient positifs à CD105 par ELISA sur un total de dix-neuf protéines VHH recombinantes analysées. Les conditions pour l'ELISA anti-CD105 ont été déterminées précédemment en utilisant la protéine VHH recombinante EngVHH17 et il a été établi que la quantité minimale d'antigène nécessaire pour un bon signal de détection était de 50 ng, la dilution de l'anticorps conjugué 1:5000 et la dilution de l'éluat 1:10.

Ces 13 VHHs anti-CD105 ont ensuite été exprimés et purifiés. Pour cela, les conditions d'expression ont été augmentées d'un volume de culture de 3 ml à 5-10 ml puis à 50-100 ml et fixées à 3 heures d'induction avec 1mM IPTG dans des cultures bactériennes pendant 3 heures de croissance (DO600nm de 0.5) dans un milieu LB avec 150 ug/ml d'ampicilline à 37° C et agitation à 200 rpm. Alors que pour leur purification en conditions natives, la méthode IMAC en matrice d'Agarose Ni-NTA proposée par le fabricant Qiagen a été adaptée pour les obtenir à partir du surnageant de culture bactérienne et non des cellules. De plus, la quantité appropriée d'Imidazole a été déterminée pour chaque étape, établissant que 15 mM étaient nécessaires pour l'équilibrage, 25 mM pour le lavage et 300 mM pour l'élution.

La concentration moyenne de protéine obtenue était de 0,225 mg/ml par purification par lots et de 0,220 mg/ml par purification sur colonne. Environ 4,9 mg de chaque protéine VHH ont été purifiés à partir de 100 ml de culture. Les protéines VHH ont été visualisées sous forme de bandes uniques de 15 kDa dans une SDS-PAGE à 15 % colorée au bleu de coomassie. Ces protéines VHH ont été utilisées pour la fonctionnalisation et l'application sur des biopuces afin de déterminer l'affinité et la spécificité de liaison par l´imagerie de la résonance des plasmons de surface (SPRi).

Conclusions

Sur les 19 protéines recombinantes exprimées dans *E. coli*, 13 étaient positives pour l'ELISA anti-CD105, vérifiant ainsi que les protéines exprimées étaient des VHHs anti-CD105 recombinantes. Les conditions d'expression de la protéine VHH ont été déterminées comme suit : 1 mM IPTG pendant 3 heures à 37° C sous agitation à 200 rpm dans des cultures bactériennes avec 150 $ug.ml^{-1}$ d'ampicilline. Ces 13 protéines VHH anti-CD105 ont été purifiées par chromatographie d'affinité dans des colonnes d'agarose Ni-NTA à partir du surnageant des cultures bactériennes, ce qui a permis d'obtenir environ 5 mg de chaque protéine par 100 ml de culture. Ainsi, les protéines VHH anti-CD105 purifiées ont été obtenues en quantité et qualité suffisantes pour leur fonctionnalisation et leur application dans des biopuces à protéines pour leur caractérisation dans les essais SPRi.

Chapitre IV: Biopuces de protéines VHH recombinantes pour la détection de CD105 sur les cellules par l´imagerie SPR (SPRi)

Le rôle de CD105 en tant que molécule clé dans la régulation de l'angiogenèse en fait une cible pour le développement des VHHs qui se lient à CD105 à la fois comme marqueur pour le diagnostic de plusieurs maladies et pour la thérapie biologique anti-tumorale. Dans ce scénario, le développement de biopuces pour étudier l'interaction des VHHs avec CD105 dans les cellules *in vivo* ainsi que la capacité de détecter, lier et capturer la cible sous forme de nanosondes est prometteur.

Les VHHs anti-CD105 recombinants purifiés ont été fonctionnalisés pour être appliqués dans des biopuces afin d'évaluer l'affinité et la spécificité de la liaison à CD105 exprimé sur les cellules par l´imagerie de la résonance des plasmons de surface (SPRi) en temps réel. Les conditions de fonctionnalisation des protéines Ig G telles que développées au CREAB, ont été adaptées pour une application aux protéines VHH.

Les 13 VHHs recombinantes purifiées ont été couplées au linker HS-C_{12}-N-Hydroxysuccinimide (Thiol-NHS) dans un ratio de 1:10 par rapport à la concentration molaire de la protéine, pour leur immobilisation sur la biopuce SPRi par adsorption directe (chimisorption par interactions or-thiol). Les VHHs ont été greffés manuellement sur la biopuce pour obtenir des spots suffisamment grands et permettre une visualisation correcte de l'interaction antigène-anticorps étant donné la différence de taille des VHHs (~2,4 x 4,2 nm) avec les cellules exprimant l'antigène (~15 uM).

Les lignées cellulaires non adhérentes SC (ATCC® CRL-9855™) et THP-1 (ATCC® TIB-202™) ont été choisies pour les essais SPRi. SC parce que ce sont des cellules de type monocyte-macrophage et qu'elles surexpriment CD105 à leur surface et THP-1 parce que ce sont des cellules de type monocyte qui n'expriment pas CD105. Les VHHs immobilisés sur la biopuce ont été incubés avec des cellules SC et des cellules THP-1. L'analyse des images SPRi et des variations de réflectivité (signal SPRi) a montré que la plupart des VHHs étaient capables de se lier aux cellules SC, mais très peu ou pas du tout aux cellules THP-1, ce qui suggère leur spécificité de liaison.

Le comptage des cellules afin d'obtenir la densité cellulaire (cellules dans une superficie de 477,09 µm x 477,09 µm) pour chaque spot VHH, a été effectué sur les images différentielles 2 heures après l'ajout des cellules dans l´essai et en utilisant des niveaux de gris (de 0 à 1) de 0,1-0,15 pour les cellules SC et de 0,3 pour les cellules THP-1. Le niveau de gris a été choisi comme un reflet de la densité cellulaire cohérente avec la variation de réflectivité (signal SPRi) de chaque VHH au même temps du dosage (2 heures).

Une répartition des VHHs en cinq groupes a été effectuée en comparant la densité cellulaire de chaque VHH dans des essais avec des cellules SC (CD105 (+)) et des cellules THP-1 (CD105 (-)), c'est-à-dire qu'ils ont été regroupés selon leur capacité à capturer spécifiquement le CD105 sur les cellules. Les approximations de la constante d'affinité apparente ont été obtenues en calculant les quotients des valeurs de réflectivité á 2 heures d´essai par les valeurs de réflectivité initiales et les VHHs ont été divisés en seulement 4 groupes selon la spécificité et l'affinité.

La moyenne des variations de réflectivité parmi les VHHs à 2 heures d´essai avec les cellules SC était de 4,59 et avec les cellules THP-1 était de 2,69, la moyenne de la densité cellulaire avec les cellules SC était de 578 cellules/superficie de référence et avec les cellules THP-1 était de 367 cellules/superficie de référence et la moyenne de l'approximation de la constante d'affinité apparente avec les cellules SC était de 110,19 x 10^{-2} et avec les cellules THP-1 était de 106,26 x 10^{-2}.

La différence entre les valeurs de la densité cellulaire, la variation de la réflectivité SPRi et l'approximation de la constante d'affinité apparente obtenues dans les essais avec les cellules SC et THP-1 a corroboré la distribution des 13 VHHs en 4 groupes. Dans le groupe 1 se trouvaient les VHH 7, VHH 17 et VHH 31 avec des valeurs bien supérieures à la moyenne parmi les VHH, dans le groupe 2 se trouvaient les VHH 22, VHH 2 et VHH 3 avec des valeurs supérieures à la moyenne, dans le groupe 3 se trouvaient les VHH 9, VHH 18, VHH 26, VHH 8 et VHH 20 avec des valeurs égales ou très proches de la moyenne et dans le groupe 4 se trouvaient les VHH 11 et VHH 1 avec des valeurs inférieures à la moyenne et proches du contrôle négatif.

Les six VHHs anti-CD105 recombinants les mieux classés (groupes 1 et 2) auraient les conformations et les repliements les plus efficaces pour l'interaction avec l'antigène et ont été donc proposés comme nanosondes pour CD105, non seulement sous forme libre (CD105 soluble), mais aussi sur les cellules (CD105 membranaire), pour des études de recherche dans le diagnostic ou la thérapeutique des maladies impliquées.

Conclusions

Les protéines VHH anti-CD105 recombinantes spécifiques de leur antigène sous forme libre (CD105 soluble) ont montré leur capacité de reconnaissance sur des cellules exprimant CD105 (CD105 membranaire) par évaluation en temps réel avec SPRi. L'imagerie SPRi a montré que la plupart des VHHs étaient capables de se lier aux cellules SC qui expriment CD105 à leur surface, mais pas (ou très peu) aux cellules THP-1 qui ne l'expriment pas, démontrant ainsi leur spécificité.

La densité cellulaire pour chaque spot VHH a été obtenue en comptant les cellules dans une zone de spot de 477,09 µm x 477,09 µm sur les images différentielles à 2 heures d'incubation avec les cellules. Le niveau de gris a été choisi comme un reflet de la densité cellulaire cohérente avec la variation de réflectivité (signal SPRi) de chaque VHH au même temps d´ essai. Les VHHs ont été divisés en 4 groupes en fonction de leur spécificité, de la densité cellulaire et de l'approximation de la constante d'affinité apparente. Les six premiers VHHs anti-CD105 du classement, VHH 17, VHH 7, VHH31 dans le groupe 1 et VHH22, VHH 3 et VHH 2 dans le groupe 2, avaient des valeurs moyennes plus élevées parmi les VHHs et ont été proposés comme nanosondes pour la capture et la détection en temps réel de CD105 (soluble et lié à la membrane) pour des études de recherche diagnostique ou thérapeutique.

Chapitre V: Conclusions et perspectives

Conclusions générales

L'endogline (CD105) est une protéine membranaire intégrale qui agit comme un co-récepteur du facteur de croissance transformant-β (TGF-β) et est régulée à la hausse dans les cellules endothéliales en prolifération. CD105 est suggéré comme un marqueur approprié pour la néovascularisation et l'angiogenèse liée aux tumeurs. Les anticorps anti-CD105 ont été étudiés pour leur application dans le diagnostic, le pronostic et le traitement du cancer (ainsi que d'autres maladies), à la fois seuls et en concomitance avec des anticorps spécifiques d'autres cibles. Il est intéressant de générer de nouvelles biomolécules qui se lient spécifiquement à CD105 et qui permettent de cibler, de détecter et de capturer les cellules exprimant CD105, car l'angiogenèse est un processus clé dans la croissance et le développement des cellules normales et transformées dans les tissus.

Ce travail présente l'évaluation de dix-neuf VHHs recombinants qui lient sCD105, obtenus après le criblage par phage display et phage-ELISA d'une bibliothèque d'ADNc de VHH générée à partir d´alpagas immunisés avec un lysat d'une lignée de cellules cancéreuses exprimant CD105. Les 19 séquences de VHH anti-CD105 analysées *in silico* présentaient une homologie avec les séquences VHH de la Gen Bank et de la Protein Data Bank et avaient des substitutions typiques des VHH par rapport aux séquences VH des anticorps conventionnels, prouvant ainsi leur identité. Sur les dix-neuf protéines VHHs qui ont été exprimées dans E. coli, treize ont conservé leur haute spécificité et affinité pour se lier à sCD105 selon le test ELISA,

Afin d'évaluer la capacité des 13 VHHs à se lier à CD105 exprimé dans les cellules, les essais de liaison SPRi ont été réalisés avec les cellules monocytes/macrophages SC qui expriment CD105 et les cellules monocytes THP-1 qui n'expriment pas CD105. Les données SPRi ont montré que les VHHs anti-CD105 se lient aux cellules SC mais pas ou peu aux cellules THP-1. Les essais SPRi ont montré la faisabilité d'étudier la liaison des VHHs anti-CD105 aux cellules sur une biopuce.

Les résultats des essais SPRi ont suggéré que les VHHs anti-CD105 peuvent être distribués en 4 groupes selon leur spécificité et leur approximation d'affinité. Dans le premier groupe, VHH 7, VHH 17 et VHH 31 ont été classés parce qu'ils avaient la plus grande

affinité et spécificité de tous les VHHs, bien au-dessus de la moyenne des VHHs, même au-dessus du contrôle positif. Dans le deuxième groupe, VHH 22, VHH 2 et VHH 3 ont été placées pour avoir une affinité et une spécificité plus élevées que la moyenne des VHHs. Dans le troisième groupe, ont été placés les cinq VHHs avec une affinité et une spécificité égales ou très proches de la moyenne des VHHs. Le quatrième groupe comprenait les deux VHHs qui avaient une spécificité et une affinité faibles.

Comme la liaison antigène-anticorps dépend de la conformation et du repliement de l'anticorps, les VHHs des deux premiers groupes auraient la conformation la plus efficace pour interagir avec CD105. Les six VHHs anti-CD105 les mieux classés, qui se lient avec une affinité et une spécificité élevées à CD105 en solution et dans les cellules liées à la membrane, peuvent être utilisés à la fois dans des études qui cherchent à mieux comprendre la fonction de CD105 *in vitro* et *in vivo* et dans des applications comme nanosondes pour le diagnostic et la thérapie de maladies où l'angiogenèse joue un rôle important.

Perspectives

Les VHHs anti-CD105 étudiés dans le présent travail peuvent être évalués pour la conception de biocapteurs qui permettent la détection de CD105 soluble ou lié aux cellules pour le diagnostic ou le pronostic de maladies impliquant une expression accrue de CD105 sur les cellules ou des niveaux élevés de la protéine dans la circulation sanguine comme la pré-éclampsie, l'hypertension, et certains types de cancer ou de métastases. Les VHHs sont avantageux car ils se prêtent à l'édition de gènes par mutagenèse dirigée, Crisp-Casp9 ou d'autres technologies connexes qui peuvent améliorer la liaison, la demi-vie, la stabilité et enrichir la fonction dans un contexte cellulaire en incorporant des séquences dans le VHH qui codent pour des signaux, des effecteurs ou d'autres domaines fonctionnels. En outre, la fixation de médicaments ou de toxines aux VHHs anti-CD105 peut permettre d'évaluer leur application dans des thérapeutiques ciblant les cellules, en diminuant les dommages collatéraux sur d'autres cellules qui ne nécessitent pas de traitement.

La régulation de l'angiogenèse par l'utilisation des VHHs anti-CD105 dans les cellules endothéliales de la veine ombilicale humaine (HUVEC) *in vitro* fournirait des informations sur l'application potentielle de ces VHHs dans la thérapie du cáncer avec CD105 en ciblant les cellules endothéliales hautement prolifératives, afin d'inhiber les métastases ou d'induire la rétraction de la tumeur en empêchant l'approvisionnement en nutriments vitaux et l'échange de déchets des cellules tumorales. De plus, le dosage des VHHs anti-CD105 sur les cellules endothéliales par SPRi montre le potentiel de l'analyse de différents VHH contre différentes cibles dans des essais simultanés et parallèles fournissant des informations sur l'effet des VHHs uniques ou multiples dirigés contre différentes cibles sur les cellules normales et transformées. Il s'agit d'une information utile car le traitement du cancer nécessite l'action concertée d'une combinaison de biomolécules qui agissent sur différentes cibles dans les cellules transformées,

D'autre part, il est possible d'explorer l'application de ces VHHs anti-CD105 dans l'ingénierie tissulaire osseuse comme molécules d'attachement des cellules souches mésenchymateuses (MSC) à la structure on squelette ("scaffold"). L'ingénierie tissulaire vise à réparer ou à remplacer partiellement ou totalement les tissus endommagés par un traumatisme ou une maladie. Fondamentalement, des cellules, un "scaffold", des facteurs de croissance et une matrice extracellulaire sont nécessaires[1]. Les cellules les plus couramment utilisées pour la régénération osseuse sont les MSCs en raison de leur multipotentialité pour se différencier ou se transdifférencier en une variété de cellules spécialisées *in vivo* et *in vitro*[2,3]. Le "scaffold" doit être biodégradable et similaire au tissu à réparer pour pouvoir être associé à des cellules vivantes[4]. Ainsi, les cellules osseuses doivent être cultivées sur des "scaffolds" tridimensionnels qui permettent la formation d'adhérences 3D et ressemblent à la morphologie du tissu à remplacer. Pour capturer et fixer les MSCs dans le "scaffold" afin qu'elles puissent se différencier en tissu osseux, des anticorps spécifiques des marqueurs cellulaires des MSCs sont fixés à le "scaffold".

La taille des anticorps conventionnels peut entraver la capture des cellules à des endroits inaccessibles sur le "scaffold", l'application des VHHs comme ponts entre les cellules à fixer et le "scaffold" est donc une alternative. Comme CD105 est présent à la surface des MSCs[5,6], c'est un marqueur cellulaire important pour identifier et isoler les MSC[7]. Les VHHs anti-CD105 liés à le "scaffold" reconnaîtraient et lieraient spécifiquement CD105 sur les MSC, de sorte qu'avec des facteurs de croissance pour l'ostéogenèse, les cellules peuvent se différencier en tissu osseux à réparer. Par conséquent, l'utilisation des VHHs anti-CD105 comme outil de fixation des cellules faciliterait le développement des "scaffolds" tridimensionnels adaptés à la régénération osseuse.

References

1) Kim, M. & Evans, D. (2005) Tissue Engineering: The Future of Stem Cells. *Top Tissue Eng 2,* 1–22.

2) Bunnell, B. A., Flaat, M., Gagliardi, C., Patel, B. & Ripoll, C. (2008) Adipose-derived stem cells: Isolation, expansion and differentiation. *Methods 45(2),* 115–120.

3) Wakitani, S., Nawata, M., Tensho, K., Okabe, T., Machida, H. & Ohgushi, H. (2007) Repair of articular cartilage defects in the patello-femoral joint with autologous bone marrow mesenchymal cell transplantation: three case reports involving nine defects in five knees. *J Tissue Eng Regen Med 1(1),* 74-79.

4) Stock, U. A. & Vacanti, J. P. (2001) Tissue engineering: Current state and prospects. *Annu Rev Med 52(1),* 443–451.

5) St Jacques, S., Cymerman, U., Pece, N. & Letarte, M. (1994) Molecular characterization and *in situ* localization of murine endoglin reveal that it is a transforming growth factor-β binding protein of endothelial and stromal cells. *Endocrinology 134,* 2645-2657.

6) Rokhlin, O. W., Cohen, M. B., Kubagawa, H., Letarte, M. & Cooper, M. D. (1995) Differential expression of endoglin on fetal and adult hematopoietic cells in human bone marrow. *J Immunol 154,* 4456–4465.

7) Godino Izquierdo, M. (2020) Capacitación osteogénica *in vitro* de células madre mesenquimales de médula ósea para su aplicación en resecciones segmentarias de hueso. *Rev Esp Cir Ortop Traumatol 64(4),* 236-243

8) Hanahan, D. (1983) Studies on transformation of *Escherichia coli* with plasmids. *J Mol Biol 166(4),* 557-580.

9) Novagen (2003) pET System Manual. Obtained from https://lifewp.bgu.ac.il/wp/zarivach/wp-content/uploads/2017/11/Novagen-pET-system-manual-1.pdf. Revised April 2018

10) Kramer, R. A., Cox, F., van der Horst, M., van der Oudenrijn, S., Res, P. C., Bia, J., . . . de Kruif, J. (2003) A novel helper phage that improves phage display selection efficiency by preventing the amplification of phages without recombinant protein. *Nucleic acids research 31(11),* e59.

11) Tsuchiya, S., Yamabe, M., Yamaguchi, Y., Kobayashi, Y., Konno, T. & Tada, K. (1980) Establishment and characterization of a human acute monocytic leukemia cell line (THP-1). *Int. J. Cancer 26,* 171-176.

12) Hesketh, M., Sahin, K. B., West, Z. E. & Murray, R. Z. (2017) Macrophage phenotypes regulate scar formation and chronic wound healing. *Int J Mol Sci 18(7),* 1545.

13) Collins, GW. & Largen, M. (1995) Continuous mammalian cell lines having monocyte/macrophage characteristics and their establishment in vitro. *US Patent 5,447,861 dated Sep 5*.

14) Dotto, G. P., Enea, V. & Zinder, N. D. (1981) Functional analysis of bacteriophage f1 intergenic region. *Virology 114(2),* 463-473.

15) Enea, V. & Zinder, N. D. (1982) Interference resistant mutants of phage f1. *Virology. 122(1),* 222-226.

16) Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular cloning. A laboratory manual*. New York: Cold Spring Harbor Laboratory Press.

17) Qiagen (2018) *Quick-Start protocol. QIAquick® Gel Extraction Kit QIAquick® PCR & Gel Cleanup Kit.* Obtained from https://www.qiagen.com/us/knowledge-and-support/knowledgehub/search/resources/?categories=RESOURCES&page=0&filters= % 7B%7D&query=Quick-Start%20protocol%20QIAquick. (Revised July 2018)

18) Qiagen (2008) *Ni-NTA Spin Kit Handbook*. Obtained from https://www.qiagen.com/us/resources/download.aspx?id=3fc8c76d-6d21-4887-9bf8-f35f78fcc2f2&lang=en. (Revised January 2019).

19) Qiagen. (2016) *Ni-NTA Agarose Purification of 6xHis-tagged Proteins from E. coli under Native Conditions.* Obtained from https://www.qiagen.com/us/resources/download.aspx?id=dc89c299-75d3-4120-adf0-35251f16a7af& lang=en. (Revised December 2018).

20) Hamers-Casterman C., Atarhouch T, Muyldermans S, Robinson G, Hamers C, Bajyana Songa, E . . . Hammers, R (1993) Naturally occurring antibodies devoid of light chains. Nature 363(6428), 446–448.

21) Harmsen, M. M. & De Haard, H. J. (2007) Properties, production, and applications of camelid single-domain antibody fragments. *Applied Microbiology and Biotechnology 77(1)*, 13-22.

22) Muyldermans, S., Atarhouch, T., Saldanha, J., Barbosa, J. & Hamers, R. (1994) Sequence and structure of VH domain from naturally occurring camel heavy chain immunoglobulins lacking light chains. *Protein Eng 7(9),* 1129–1135.

23) Dmitriev, O. Y., Lutsenko, S. & Muylderman, S. (2016) Nanobodies as Probes for Protein Dynamics *in Vitro* and in Cells. *J Biological Chemistry 291(8),* 3767–3775

24) Wesolowski, J., Alzogaray, V., Reyelt, J., Unger, K., Juárez, K., Cauerhff, A.,. . . Koch-Nolte, F. (2009) Single domain antibodies: promising experimental and therapeutic tools in infection and immunity. *Med Microbiol Immunol 198(3),* 157-174.

25) Helma, J., Cardoso, M. C., Muyldermans, S. & Leonhardt, H. (2015) Nanobodies and recombinant binders in cell biology. *J Cell Biol 209(5),* 633–644.

26) Yang, E. Y. & Shah, K. (2020) Nanobodies: Next Generation of Cancer Diagnostics and Therapeutics. *Frontiers in Oncology 10,* 1182.

27) Lopez-Novoa, J. M. & Bernabéu, C. (2012) ENG (endoglin). *Atlas Genet Cytogenet Oncol Haematol*. http://atlasgeneticsoncology.org/Genes/ENGID40452ch9q34.html. Revised July 2020

28) Santibañez, J. F., Quintanilla, M. & Bernabéu, C. (2011) TGF-β/TGF-β receptor system and its role in physiological and pathological conditions*. Clin Sci (Lond) 121(6),* 233-251

29) Jerkic, M., Rivas, J. V., Carrón, R., Sevilla, M. A., Rodríguez-Barbero, A., Bernabéu, C., . . . López Novoa, J. M. (2002) Endoglina, un componente del complejo de receptores de TGF-β, es un regulador de la estructura y función vascular. *Nefrología XX11,* 2.

30) Letarte, M., Greaves, A. & Vera S. (1995) CD105 (endoglin) cluster report. In Schlossman, S.F. *et al* editors. Leukocyte typing V: white cell differentiation antigens. (pp1756–1759). Oxford, UK: Oxford University Press.

31) Gougos, A. & Letarte, M. (1990) Primary structure of endoglin, an RGD-containing glycoprotein of human endothelial cells. *J Biol Chem 265,* 8361-8364.

32) Bellón, T., Corbi, A., Lastres, P., Cales, C., Cebrian, M., Vera, S., . . . Bernabéu, C. (1993) Identification and expression of two forms of the transforming growth factor beta-

binding protein endoglin with distinct cytoplasmic regions. *Eur J Immunol 23,* 2340-2345

33) Gallardo-Vara, E. M. (2018) *Endoglina soluble: mecanismo de generación y función en células endoteliales y su efecto en el remodelado vascular*. Universidad Complutense de Madrid, Facultad de Ciencias Quimicas, Departamento de Bioquímica y Biología Molecular: Thesis to obtain the degree of Doctor en Ciencias Biológicas.

34) Pérez-Gómez, E., Eleno, N., López-Novoa, J. M., Ramirez, J. R., Velasco, B., Letarte, M., . . . Quintanilla, M. (2005). Characterization of murine S-endoglin isoform and its effects on tumor development. *Oncogene. 24(27),* 4450-4461.

35) Valbuena-Díez, A. C., Blanco, F. J., Oujo, B., Langa, C., Gonzalez-Nuñez, M., Llano, E., . . . Bernabéu, C. (2012) Oxysterol-induced soluble endoglin release and its involvement in hypertension. *Circulation 126(22),* 2612-2624.

36) Gallardo-Vara, E., Blanco, F. J., Roque, M., Friedman, S. L., Suzuki, T., Botella, L. M. & Bernabéu, C. (2016) Transcription factor KLF6 upregulates expression of metalloprotease MMP14 and subsequent release of soluble endoglin during vascular injury. *Angiogenesis 19,* 155–171.

37) Gougos, A. & Letarte, M. (1988a) Identification of a human endothelial cell antigen with monoclonal antibody 44G4 produced against a pre-B leukemic cell line. *J Immunol 141,* 1925-1933

38) Zhang, H., Shaw, A. R., Mak, A. & Letarte, M. (1996) Endoglin is a component of the transforming growth factor (TGF) beta-receptor complex of human pre-B leukemic cells*. J Immunol 156,* 564–573.

39) Bürhing, H. J., Müller, C. A., Letarte, M., Gougos, A., Saalmüller, A., von Agthoven, A. J. & Busch, F. W. (1991) Endoglin is expressed on a subpopulation of inmature erythroid cells of normal human bone marrow. *Leukemia 5,* 841-847

40) Lastres, P., Bellón, T., Cabañas, C., Sánchez-Madrid, F., Acevedo, A., Gougos, A., . . . Bernabéu, C. (1992) Regulated expression on human macrophages of endoglin, an Arg-Gly-Asp-containing surface antigen. *Eur J Immunol 22,* 393-397.

41) Robledo, M. M., Hidalgo, A., Lastres, P., Arroyo, A. G., Bernabéu, C., Sánchez-Madrid, F. & Teixidó, J. (1996) Characterization of TGF-beta 1-binding proteins in human bone marrow stromal cells. *Br J Haematol 93(3),* 507-514.

42) Gougos, A., St Jacques, S., Greaves, A., O'Connell, P.J., d'Apice, A.J.F., Bürhing, H.J., . . . Letarte, M. (1992) Identification of distinct epitopes of endoglin, an RGD-containing glycoprotein of endothelial cells, leukemic cells, and syncytiotrophoblasts. *Int Immunol 4(1),* 83-92.

43) Bernabéu, C., Conley, B. A. & Vary, C. P. (2007) Novel biochemical pathways of endoglin in vascular cell physiology. *J Cell Biochem 102(6),* 1375-1388.

44) Parker, W. L., Goldring, M. B. & Philip, A. (2003) Endoglin is expressed on human chondrocytes and forms a heteromeric complex with betaglycan in a ligand and type II TGFbeta receptor independent manner. *J Bone Min. Res 18,* 289–302.

45) Rodriguez-Barbero, A., Obreo, J., Eleno, N., Rodriguez-Pena, A., Duwel, A., Jerkic, M., . . . Lopez-Novoa, J. M. (2001) Endoglin expression in human and rat mesangial cells and its upregulation by TGF-beta1. *Biochem Biophys Res Commun 282,* 142–147.

46) Meurer, S. K., Tihaa, L., Lahme, B., Gressner, A. M. & Weiskirchen, R. (2005) Identification of endoglin in rat hepatic stellate cells: new insights into transforming growth factor beta-receptor signaling. *J Biol Chem 280,* 3078–3087.

47) Quintanilla, M., Ramirez, J. R., Pérez-Gómez, E., Romero, D., Velasco, B., Letarte, M., . . . Bernabéu, C. (2003) Expression of the TGF-beta coreceptor endoglin in epidermal keratinocytes and its dual role in multistage mouse skin carcinogenesis. *Oncogene 22(38),* 5976-5985.

48) St-Jacques, S., Forte, M., Lye, S. J. & Letarte, M. (1994b) Localization of endoglin, a transforming growth factor-beta binding protein, and of CD44 and integrins in placenta during the first trimester of pregnancy. *Biol Reprod 51,* 405–413 .

49) Altomonte, M., Montagner, R., Fonsatti, E., Colizzi, F., Cattarossi, I., Brasoveanu, L. I., . . . Maio, M. (1996) Expression and structural features of endoglin (CD105), a transforming growth factor beta1 and beta3 binding protein, in human melanoma. *Br J Cancer 74(10),* 1586-1591.

50) Quintanilla, M., Ramirez, J. R., Pérez-Gómez, E., Romero, D., Velasco, B., Letarte, M., . . . Bernabéu, C. (2003) Expression of the TGF-beta coreceptor endoglin in epidermal keratinocytes and its dual role in multistage mouse skin carcinogenesis. *Oncogene 22(38),* 5976-5985.

51) Henriksen, R., Gobl, A., Wilander, E., Oberg, K., Miyazono, K. & Funa, K. (1995) Expression and prognostic significance of TGF-b isotypes, latent TGF-b1 binding protein, TGF-b type I and type II receptors, and endoglin in normal ovary and ovarian neoplasms. *Lab Invest 73(2),* 213-220.

52) Craft, C. S., Romero, D., Vary, C. P. & Bergan, R. C. (2007) Endoglin inhibits prostate cancer motility via activation of the ALK2-Smad1 pathway. *Oncogene 26(51),* 7240-7250.

53) Botella, L. M., Sánchez-Elsner, T., Sanz-Rodriguez, F., Kojima, S., Shimada, J., Guerrero-Esteo, M., . . . Bernabéu, C. (2002) Transcriptional activation of endoglin and transforming growth factor-beta signaling components by cooperative interaction between Sp1 and KLF6: their potential role in the response to vascular injury. *Blood 100(12),* 4001-4010.

54) van Laake, L. W., van den Driesche, S., Post, S., Feijen, A., Jansen, M. A., Driessens, M. H., . . . Mummery, C. L. (2006) Endoglin has a crucial role in blood cell-mediated vascular repair. *Circulation 114(21),* 2288-2297.

55) Mancini, M. L., Verdi, J. M., Conley, B. A., Nicola, T., Spicer, D. B., Oxburgh, L. H. & Vary, C. P. (2007) Endoglin is required for myogenic differentiation potential of neural crest stem cells. *Dev Biol. 308(2),* 520-533.

56) Perlingeiro, R. C. (2007) Endoglin is required for hemangioblast and early hematopoietic development. *Development 134(16),* 3041-3048.

57) Docherty, N. G., López-Novoa, J. M., Arevalo, M., Düwel, A., Rodriguez-Peña, A., Pérez-Barriocanal, F., Bernabéu, C. & Eleno, N. (2006) Endoglin regulates renal ischaemia-reperfusion injury. *Nephrol Dial Transplant 21(8),* 2106-2119.

58) Marazuela, M., Sánchez-Madrid, F., Acevedo, A., Larrañaga, E., de Landázuri, M. O. (1995) Expression of vascular adhesion molecules on human endothelial in autoimmune thyroid disorders. *Clin Exp Immunol 102(2),* 328-334.

59) van de Kerkhof, P. C., Rulo, H. F., van Pelt, J. P., van Vlijmen-Willems, I. M., De Jong, E. M. (1998) Expression of endoglin in the transition between psoriatic uninvolved and involved skin. *Acta Derm Venereol 78(1),* 19-21.

60) Ma, X., Labinaz, M., Goldstein, J., Miller, H., Keon, W. J., Letarte, M. & O'Brien, E. (2000) Endoglin is overexpressed after arterial injury and is required for transforming growth factor-beta-induced inhibition of smooth muscle cell migration. *Arterioscler Thromb Vasc Biol 20(12*), 2546-2552.

61) Fonsatti, E., Altomonte, M., Nicotra, M. R., Natali, P. G., Maio, M. (2003) Endoglin (CD105): a powerful therapeutic target on tumor-associated angiogenetic blood vessels. *Oncogene 22(42),* 6557-6563.

62) Duff, S. E., Li, C., Garland, J. M., Kumar, S. (2003) CD105 is important for angiogenesis: Evidence and potential applications. *FASEB J 17,* 984-992.

63) Bauman, T. M., Huang, W., Lee, M. H. & Abel, E. J. (2016) Neovascularity as a prognostic marker in renal cell carcinoma. *Hum Pathol 57,* 98-105.

64) McDonald, M. T., Papenberg, K. A., Ghosh, S., Glatfelter, A. A., Biesecker, B. B., Helmbold, E. A., . . . Marchuk, D. A. (1994) A disease locus for hereditary haemorrhagic telangiectasia maps to chromosome 9q33-34. *Nat Genet 6,* 197-204.

65) Di Cosola, M., Cazzolla, A. P., Scivetti, M., Testa, N. F., Lo Muzio, L., Favia, G., . . . Bascones, A. (2005) Síndrome de Rendu-Osler-Weber o Telangiectasia Hemorrágica Hereditaria (HHT) Descripción de dos casos y revisión de la literatura. *Av Odontoestomatol 21(6),* 297-303.

66) Blanco, F. J., Grande, M. T., Langa, C., Oujo, B., Velasco, S., Rodriguez-Barbero, A., . . . Bernabéu, C. (2008) S-endoglin expression is induced in senescent endothelial cells and contributes to vascular pathology. *Circ Res 103(12),* 1383-1392.

67) Hawinkels, L. J., Kuiper, P., Wiercinska, E., Verspaget, H. W., Liu, Z., Pardali, E., . . . ten Dijke, P. (2010) Matrix metalloproteinase-14 (MT1-MMP)-mediated endoglin shedding inhibits tumor angiogenesis. *Cancer Res 70(10),* 4141-4150.

68) Venkatesha, S., Toporsian, M., Lam, C., Hanai, J., Mammoto, T., Kim, Y. M., . . . Karumanchi, S. A. (2006). Soluble endoglin contributes to the pathogenesis of preeclampsia. *Nat Med 12(6),* 642-649.

69) Blázquez-Medela, A. M., García-Ortiz, L., Gómez-Marcos, M. A., Recio-Rodríguez, J. I., Sánchez-Rodríguez, A., López-Novoa, J. M. & Martínez-Salgado, C. (2010) Increased plasma soluble endoglin levels as an indicator of cardiovascular alterations in hypertensive and diabetic patients. *BMC Med 8,* 86.

70) Cruz-González, I., Pabón, P., Rodríguez-Barbero, A., Martín-Moreiras, J., Pericacho, M., Sánchez, P. L., . . . López-Novoa, J. M. (2008) Identification of serum endoglin as a novel prognostic marker after acute myocardial infarction. *J Cell Mol Med 12(3),* 955-961

71) Nachtigal, P., Zemankova, Vecerova, L., Rathouska, J. & Strasky, Z. (2012) The role of endoglin in atherosclerosis. *Atherosclerosis 224(1),* 4-11.

72) Fujimoto, M., Hasegawa, M., Hamaguchi, Y., Komura, K., Matsushita, T., Yanaba, K., . . . Sato, S. (2006) A clue for telangiectasis in systemic sclerosis: elevated serum soluble endoglin levels in patients with the limited cutaneous form of the disease. *Dermatology 213(2),* 88-92.

73) Karampoor, S., Zahednasab, H., Ramagopalan, S., Mehrpour, M. & Keyvani, H. (2016) Angiogenic factors are associated with multiple sclerosis. *J Neuroimmunol 301*, 88-93.

74) Gushiken, E. (2016) *Generación de anticuerpos de dominio único específicos para CD105 humano.* Universidad Peruana Cayetano Heredia, Perú. Thesis to obtain the degree of Magister en Ciencias, mention Biochemistry and Molecular Biology.

75) Cherif, B., Roget, A., Villiers, C. L., Calemczuk, R., Leroy, V., Marche, P. N., . . . Villiers, M. B. (2006) Clinically related protein-peptide interactions monitored in real time on novel peptide chips by surface plasmon resonance imaging. *Clin Chem 52,* 255–262.

76) Nguyen, H. H., Park, J., Kang, S. & Kim, M. (2015) Surface plasmon resonance: a versatile technique for biosensor applications. *Sensors 15,* 10481–10510.

77) Kodoyianni, V. (2011) Label-free analysis of biomolecular interactions using SPR imaging. *Biotechniques 50,* 32–40.

78) Abadian, P. N., Kellev, C. P. & Goluch, E. D. (2014) Cellular analysis and detection using Surface plasmon resonance techniques. *Anal Chem 86,* 2799–2812.

79) Mallevre, F., Temp,er, V., Mathey, R., Leroy, L., Roupioz, Y., Fernandes, T. F., . . . Livache, T. (2016) Real-time toxicity testing of silver nanoparticles to *Salmonella enteritidis* using surface plasmon resonance imaging: A proof of concept. *NanoImpact 1,* 55–59.

80) Li, S., Song, G., Bai, Y., Song, N., Zhao, J., Jian Liu, J. & Hu, C (2021) Applications of Protein Microarrays in Biomarker Discovery for Autoimmune Diseases. *Front. Immunol 12*, 645632.

81) Arenkov, P., Kukhtin, A., Gemmell, A., Voloshchuk, S., Chupeeva, V. & Mirzabekov, A. (2000) Protein microchips: use for immunoassay and enzymatic reactions. *Anal Biochem 278,* 123–131.

82) MacBeath, G. & Schreiber, S. L. (2000) Printing proteins as microarrays for high-throughput function determination. *Science 289(1),* 60-63.