Development of *Escherichia coli* cell surface display for selection of single domain antibodies from immune libraries

Doctoral Thesis

Valencio Francis Salema

Madrid 2014
Tesis presentada por D. Valencio Francis Salema para optar al grado de Doctor en Ciencias por la Universidad Autónoma de Madrid.

Director de la Tesis Doctoral:

Dr. Luis Ángel Fernández Herrero, Investigador Científico del Consejo Superior de Investigaciones Científicas.

Este trabajo ha sido realizado en el Departamento de Biotecnología Microbiana del Centro Nacional de Biotecnología (CSIC) gracias a la financiación recibida de la Fundación "La Caixa", el Ministerio de Economía y Competitividad (MINECO), la Comunidad Autónoma de Madrid y el European Research Council (ERC).
ACKNOWLEDGEMENTS

First and foremost, I thank my Thesis director, Dr Luis Angel Fernandez, for the excellent supervision and support extended to me during my PhD work. You have facilitated my integration into the lab, and I am very grateful for the time and effort you put in the process. Through everything, you have been patient and supportive, always valuing people above science. You are a brilliant scientist with fantastic ideas and I have been very fortunate to learn as much as I could, from you.

Secondly, I thank all my labmates, both current and former members, for the excellent working environment and friendly atmosphere I have experienced during my stay. It was a joy to come work in everyday, especially knowing how everyone is willing to help each other. In particular: Elvira, Carlos, Carmen Palomino, Gustavo, David and Ana for all the discussions and guidance. You all have helped me think more critically and I wouldn’t be where I am today without you. Thank you for all for the help you have offered me through the years, especially teaching me all the basics – like running a gel! Special thanks to Carlos, Carmen Mañas, Lorenza, Massiel, Ariadna, Beatriz and Yago, for always being there to discuss science, but more importantly to discuss everything else. I know I wouldn’t have survived the PhD without our morning and afternoon coffee breaks!

I also extend my sincere gratitude to my PhD tutor, Dr José Lopez Carrascosa, for the help, guidance and ideas received during my PhD work, and to Dr Paolo Natale for the review of my thesis manuscript and suggestions for improvement.

Performance of a lot of my PhD work would not have been possible without collaboration with other laboratories in the CNB and outside. First and foremost, I wish to thank the Flow cytometry department, especially Dr Maria del Carmen and Sara, for the help you offered during setting up of the flow cytometry protocols. I also acknowledge the help received from the Protein Tools unit, especially Dr Leonor Kremer and Mayte Martin during the BiaCore experiments. Special thanks to Dr José M. Pingarrón and group, Electroanalysis and Electrochemical Biosensors.
Group, UCM for the collaborative work that lead to the development of a biosensor for fibrinogen detection.

I thank all my flatmates, especially Tribhuwan, Praveen, Acharya and Mayank, and special friends, especially Carlos, Gorjana and Daniela. You have all touched my life and made me a better person, and I hope that our friendship lasts the test of time.

I thank God for this wonderful opportunity to carry out my PhD work in Spain. It has been a fantastic experience filled with challenges and learning. I have enjoyed travelling and exploring what this great country has to offer. I have come across so many wonderful people, many of whom have touched my life.

Lastly, I thank my Father (Alex), Mother (Norma) and Brother (Valmiki), for all the love, support and help that I have received all these years. It has been difficult living away from you, and hence I would not have achieved what I have, without your constant love, encouragement and prayers.
# TABLE OF CONTENTS

TABLE OF CONTENTS .................................................................................................................. 6
ABBREVIATIONS ......................................................................................................................... 12
SUMMARY .................................................................................................................................. 16
PRESENTACIÓN .......................................................................................................................... 20
INTRODUCTION ......................................................................................................................... 26

1. Antibodies and Recombinant antibody fragments ............................................................... 28
   1.1 Antibodies and their application in human therapy ......................................................... 28
   1.2 Recombinant fragments from conventional antibodies ............................................... 31
   1.3 Heavy-chain only antibodies, single domain antibodies and nanobodies ................. 32

2. Selection of antibodies from libraries using surface display ........................................... 35
   2.1 Libraries of antibody genes ......................................................................................... 35
   2.2 Surface display methods for screening of Ab libraries ................................................. 37
       2.2.1 Phage display ........................................................................................................ 37
       2.2.2 Cell surface display .............................................................................................. 40
       2.2.2.1 Yeast Display .................................................................................................. 40
       2.2.2.2 Bacterial cell surface display .......................................................................... 42
   2.3 *E. coli* cell surface display of Ab libraries ............................................................... 42

OBJECTIVES ............................................................................................................................ 48

MATERIALS AND METHODS ..................................................................................................... 52

  1. Bacterial strains .................................................................................................................. 54
  2. Conditions of bacterial growth ......................................................................................... 54
  3. Cell lines and Growth conditions ................................................................................... 55
  4. DNA constructions ........................................................................................................... 55
  5. Protein electrophoresis and Western blot ......................................................................... 57
  6. Protease accessibility assays ............................................................................................ 59
  7. ELISA .................................................................................................................................. 59
  8. Purification of TirM<sub>EHEC</sub> ......................................................................................... 60
  9. Magnetic Cell Sorting (MACS) of V<sub>HH</sub> immune libraries .......................................... 60
 10. Flow cytometry of V<sub>HH</sub> clones or V<sub>HH</sub> immune libraries ........................................ 61
 11. Affinity Determination by Flow cytometry ................................................................... 62
12. Purification of V_{HHs} from the periplasm of *E. coli* ........................................62
13. Purification of V_{HHs} from the periplasm as fusions with Maltose binding protein (MBP) ........................................................................................................63
14. Surface plasmon resonance (SPR) ........................................................................64
15. Selection of V_{HH} immune libraries on cells ..........................................................65
16. Immunofluorescence microscopy (IFM) ....................................................................65

RESULTS ....................................................................................................................68

Chapter 1: Display of single domain Ab clones and immune libraries against *TirM_{EHEC}* and Human fibrinogen on the surface of *E. coli* cells ..........................70

1.1 Comparison of EhaA and Intimin \( \beta \)-domains for display of sdAbs on *E. coli* ..........70
1.2 Construction of immune libraries against *TirM_{EHEC}* and Fib, analysis of expression, display levels and cellular toxicity ..................................................................................74

Chapter 2: Selection and characterization of single domain Abs against *TirM_{EHEC}* from immune libraries displayed on the surface of *E. coli* ................80

2.1 Translocated Intimin Receptor domain M (TirM) ..................................................80
2.2 Selection of single domain Abs against *TirM_{EHEC}* by magnetic sorting of *E. coli* bacteria displaying the immune library ........................................................................81
2.3 Characterization of V_{TIR1} sdAb and determination of its affinity by Surface Plasmon Resonance (SPR) .................................................................................................86
2.4 Estimation of the affinity of V_{TIR1} by *E. coli* display ........................................88

Chapter 3: Selection and Characterization of high affinity single domain Abs against Human Fibrinogen from Immune libraries displayed on the surface of *E. coli* cells .............................................................................................................91

3.1 Human Fibrinogen (Fib). .........................................................................................91
3.2 Selection of high affinity single domain Abs from the immune library against Fib by MACS .........................................................................................................................92
3.3 Purification of V_{FIB1} and V_{FIB2} V_{HHs} from the periplasm of *E. coli* as fusions to MBP .........................................................................................................................95
   3.3.1 Expression and solubility of the MBP-V_{HH_{His6}} fusions ................................95
   3.3.2 Purification of MBP-V_{HH_{His6}} fusions and antigen binding activity ..........97
   3.3.3 Purification of monomeric V_{HH_{His6}} from MBP-V_{HH_{His6}} fusions ..........99
3.4 Determination of the affinity of V_{FIB1} and V_{FIB2} by Surface Plasmon Resonance (SPR) ................................................................................................................102
Chapter 4: Cell Selection and characterization of high affinity single domain Abs against human EGFR from immune libraries displayed on the surface of E. coli cells

4.1 Epidermal Growth Factor Receptor (EGFR) ................................................................. 105
4.2 Construction of immune sdAb library against Human EGFR .................................... 106
4.3 Selection of high affinity sdAbs from the anti-EGFR immune library by MACS ............. 107
4.4 Direct cell selection of high affinity sdAbs against EGFR using E. coli display ....... 108
4.5 Binding of selected clones to EGFR+ tumor cells by immunofluorescence microscopy .................................................................................................................. 112
4.6 Binding of selected clones to biotinylated rhEGFR-Fc by flow cytometry ................. 114
4.7 Estimation of the affinity of V_{EGFR1} and V_{EGFR2} by E. coli display ................... 116
4.8 Binding of the selected clones to EGFR in the presence of EGF ............................... 116

DISCUSSION.................................................................................................................. 118
CONCLUSIONS ............................................................................................................. 128
REFERENCES ............................................................................................................. 134
ANNEXURES .............................................................................................................. 148
LIST OF FIGURES

Figure 1. Structure of conventional antibodies and antibody derived fragments. ..............28
Figure 2. Chimeric and humanized antibodies.................................................................30
Figure 3. Structure of camelid heavy chain only antibodies and nanobodies. ...............32
Figure 4. Comparison between conventional V_{H}/V_{L} and V_{HH} antibody fragments. ....34
Figure 5. Phage display of antibodies..............................................................................39
Figure 6. Yeast display of antibodies............................................................................41
Figure 7. *E. coli* display of antibodies. .................................................................44
Figure 8. Proposed mechanism of the biogenesis of protein members belonging to the type V secretion system (T5SS). .................................................................46
Figure 9. Modular organization and structure of AT and Intimin proteins.......................47
Figure 10. *E. coli* cell surface display of V_{HH}S with EhaA and Intimin beta domains. .....72
Figure 11. Growth of *E. coli* cultures expressing VgfpA and NVgfp fusions................73
Figure 12. *E. coli* cell surface display and antigen-binding activity of VgfpA and NVgfp. ....74
Figure 13. *E. coli* cell surface display of anti-TirMEHEC and anti-Fib immune libraries.....76
Figure 14. Quantification of the number of molecules of V_{HH}A and NV_{HH} fusions expressed in *E. coli*.................................................................77
Figure 15. Growth of *E. coli* cultures expressing V_{HH}A and NV_{HH} libraries against anti-TirMEHEC. .................................................................77
Figure 16. The Intimin-Tir interaction between EPEC/EHEC bacteria and host cell plasma membrane.........................................................................................80
Figure 17. Schematic illustration of Magnetic cell sorting (MACS). .................................81
Figure 18. Magnetic cell sorting of V_{HH}A and NV_{HH} *E. coli* display libraries with biotinylated antigen, TirMEHEC. .................................................................83
Figure 19. Binding of *E. coli* cells displaying selected clones from V_{HH}A and NV_{HH} libraries to biotinylated TirMEHEC. .................................................................84
Figure 20. *E. coli* cell surface display levels of V_{TIR1}A and NV_{TIR1} clones.....................85
Figure 21. ELISA of sdAbs selected by *E. coli* display against TirMEHEC. ......................86
Figure 22. Monomeric behaviour and binding activity of the purified sdAb, V_{TIR1}. ........87
Figure 23. Determination of the equilibrium dissociation constant (K_{D}) of V_{TIR1} by SPR. .....88
Figure 24. Estimation of the equilibrium dissociation constant (K_{D}) of V_{TIR1} by *E. coli* display. ........................................................................................................89
Figure 25. Structure of Fibrinogen. .............................................................................91
Figure 26. Magnetic cell sorting of VHH A and NVHH E. coli display libraries with biotinylated antigen, Fib. .......................................................... 93

Figure 27. Binding of E. coli cells displaying selected clones from VHH A and NVHH libraries to biotinylated Fib. .......................................................... 94

Figure 28. Schematic representation of protein expression vector. ......................... 95

Figure 29. Induction and solubility of the MBP-fusions. ................................... 96

Figure 30. Purification of the MBP-fusions by Amylose affinity chromatography and their functional analysis by ELISA. ........................................ 98

Figure 31. Expression and purification of VHHs in E. coli. .............................. 99

Figure 32. Protease digestion of MBP-VHH fusions, Immobilized Metal Affinity chromatography (IMAC) and Gel filtration. ........................................ 100

Figure 33. Monomeric VHHs and assay of functional activity by ELISA. ............ 101

Figure 34. Determination of the equilibrium dissociation constant (Kd) of clones VFIB1 and VFIB2 by SPR. .......................................................... 102

Figure 35. Estimation of the equilibrium dissociation constant (Kd) of clones VFIB1 and VFIB2 against Fib by E. coli display. .............................. 103

Figure 36. Schematic of EGFR and the states in which it exists. ....................... 106

Figure 37. Magnetic cell sorting of NVHH E. coli display libraries with biotinylated rhEGFR-Fc antigen. .......................................................... 108

Figure 38. Direct selection of NVHH E. coli display libraries on cells. .............. 109

Figure 39. Flow cytometry analysis of anti-EGFR NVHH E. coli display libraries selected on cells. ............................................................................ 110

Figure 40. Bright field and immunofluorescence microscopy images of E. coli bacteria displaying the indicated VHH clone to HER14 and NIH-3T3 2.2 cells. ........... 113

Figure 41. Bright field and immunofluorescence microscopy images of E. coli bacteria displaying the indicated VHH clone to HER14 and NIH-3T3 2.2 cells (continued). .................................................................................. 114

Figure 42. Binding of E. coli cells displaying selected clones from anti-EGFR NVHH libraries to biotinylated rhEGFR-Fc. ........................................ 115

Figure 43. Estimation of the equilibrium constant (KD) of VEGFR1 and VEGFR2 by E. coli display. .......................................................... 116

Figure 44. Flow cytometry analysis of binding to EGFR by the clones selected by E. coli display in the presence of EGF. ........................................ 117
<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Benefits and Limitations of Phage display and Yeast Display</td>
<td>42</td>
</tr>
<tr>
<td>2</td>
<td>Bacterial strains used in this work</td>
<td>54</td>
</tr>
<tr>
<td>3</td>
<td>List of plasmids used in this work</td>
<td>56</td>
</tr>
<tr>
<td>4</td>
<td>List of oligonucleotides</td>
<td>57</td>
</tr>
<tr>
<td>5</td>
<td>Summary of MACS with <em>E. coli</em> display libraries against <em>TirM</em> of <em>EHEC</em>.</td>
<td>82</td>
</tr>
<tr>
<td>6</td>
<td>CDR3 sequences of anti-<em>TirM</em>-<em>EHEC</em> clones selected by <em>E. coli</em> display.</td>
<td>85</td>
</tr>
<tr>
<td>7</td>
<td>Summary of MACS with <em>E. coli</em> display libraries against <em>Fib</em>.</td>
<td>92</td>
</tr>
<tr>
<td>8</td>
<td>CDR3 sequences of anti-<em>Fib</em> clones selected by <em>E. coli</em> display.</td>
<td>94</td>
</tr>
<tr>
<td>9</td>
<td>Protein yields after purification from the periplasm of <em>E. coli</em> HM140.</td>
<td>101</td>
</tr>
<tr>
<td>10</td>
<td>Summary of MACS with <em>E. coli</em> display libraries against rhEGFR.</td>
<td>107</td>
</tr>
<tr>
<td>11</td>
<td>Summary of Cell selections with NVHH <em>E. coli</em> display library against rhEGFR.</td>
<td>110</td>
</tr>
<tr>
<td>12</td>
<td>CDR3 sequences of anti-<em>EGFR</em> clones selected by <em>E. coli</em> display.</td>
<td>111</td>
</tr>
</tbody>
</table>
ABBREVIATIONS

2-ME 2-mercaptoethanol
aa  amino-acid
Ab  antibody
ADCC Antibody dependent cellular cytotoxicity
Ap  Ampicillin
APEx Anchored periplasmic expression
AT  Autotransporter
bp  base pair
BSA Bovine Serum Albumin
CDR Complementarity determining region
C_h Constant domain of the heavy chain of an immunoglobulin
C_l Constant domain of the light chain of an immunoglobulin
Cm Chloramphenicol
Da  daltons
DAPI 4',6-diamidino-2-phenilindol
DMEM Dulbecco’s modified Eagle’s medium
DNA Deoxyribonucleic acid
DTT Dithiotreitol
EDTA Ethylenediamino tetracetic acid
EGFR Epidermal growth factor receptor
EHEC Enterohemorrhagic Escherichia coli
ELISA Enzyme-linked immunosorbent assay
EPEC Enteropathogenic Escherichia coli
Fab Antigen binding fragment
FACS Fluorescence activated cell sorting
Fc  Crystallizable fragment of an Ab
Fib Human Fibrinogen
Fv  Variable fragment of an Ab
GFP Green Fluorescent Protein
h  hour

HCAb  Heavy chain only Ab

His$_{6}$  Hexahistidine sequence

Ig  Immunoglobulin

IM  Inner membrane

IPTG  Isopropylthio-β-D-galactoside

kb  kilobase

kDa  kiloDalton

Km  Kanamycin

LB  Luria-Bertani medium

LPS  Lipopolysaccharide

mAb  Monoclonal antibody

MACS  Magnetic cell sorting

MBP  Maltose binding protein

µl  microlitre

µM  micromolar

ml  millilitre

mM  millimolar

min  minute

MOI Multiplicity of infection

MW  Molecular weight

Nb  Nanobody

nm  nanometer

nM  nanomolar

OD$_{600}$  Optical density at 600 nm

OD$_{490}$  Optical density at 490 nm

OM  Outer membrane

ORF  Open reading frame

PBS  Phosphate buffered saline

PCR  Polymerase chain reaction

PG  Peptidoglycan

POD  Peroxidase
PPS  Periplasmic space
rpm  revolutions per minute
rAb  Recombinant antibody
scFv  Single chain fragment variable
sdAb  Single domain antibody
SDS  Sodium dodecyl sulphate
SPR  Surface plasmon resonance
T5SS  Type V secretion system
Tir  Translocated intimin receptor
TirM  Extracellular domain of Tir
V_H  Variable domain of the heavy chain of an Ig
V_HH  Variable domain of the heavy chain of a HCAb, also called nanobody
V_L  Variable domain of the light chain of an Ig
wt  Wild type
Screening of antibody (Ab) libraries by direct display on the surface of E. coli cells is hampered by the presence of the outer membrane (OM). In this work, we demonstrate that the native β-domains of EhaA autotransporter and Intimin, two proteins from enterohemorrhagic E. coli O157:H7 (EHEC) with opposite topologies in the OM, are effective systems for the display of immune libraries of single domain Abs (sdAbs) from camelids (nanobodies or V_{HH}) on the surface of E. coli K-12 cells and for the selection of high affinity sdAbs using magnetic cell sorting (MACS). We analyzed the capacity of EhaA and Intimin β-domains to display individual sdAbs and sdAb libraries obtained after immunization of camelids with different proteins of biomedical interest, i.e. the extracellular domain of the translocated intimin receptor from EHEC (TirM_{EHEC}), human fibrinogen (Fib) and A431 cells displaying human epidermal growth factor receptor (EGFR). We demonstrated that both systems displayed functional sdAbs on the surface of E. coli cells with little proteolysis and cellular toxicity, although E. coli cells displaying sdAbs with the β-domain of Intimin showed higher antigen-binding capacity. The sdAb libraries against TirM_{EHEC} and Fib cloned in both E. coli display platforms were screened for antigen binding clones by MACS using purified biotinylated antigen. High affinity binders were selected by both display systems, although more efficiently with the Intimin β-domain. The specificity of the selected clones against their respective antigen was demonstrated by flow cytometry of E. coli cells, along with ELISA and surface plasmon resonance with purified sdAbs. In addition, we employed the E. coli cell display systems to provide an estimation of the affinity of the selected sdAb by flow cytometry analysis under equilibrium conditions. Further, we used the β-domain of Intimin to display a sdAb library against human EGFR and developed a method for the direct selection of this sdAb library on cells. Bacteria displaying the anti-EGFR sdAb immune library were subjected to consecutive rounds of selection on EGFR-negative cells (i.e. murine fibroblast cell line, 3T3 2.2) and EGFR-positive cells HER14 (i.e. 3T3 2.2 transfected cells expressing human EGFR) to enrich for EGFR-specific binding clones. EGFR-specific clones that bound HER14 cells and not 3T3 2.2 cells were identified and their specific binding to EGFR confirmed by flow
cytometry analysis using biotinylated EGFR-Fc fusion. In addition, we used *E. coli* display to characterize the affinity and binding of these sdAbs to EGFR-Fc in the presence or absence of an excess of EGF, a natural ligand of EGFR. Our data demonstrates that *E. coli* display allows the selection of sdAbs against relevant tumor-associated antigens from libraries generated by cell immunization and performance of direct selection on tumor cells without the need for purified antigens.
Introducción
Los anticuerpos (Abs, del inglés “antibodies”) o inmunoglobulinas (Igs) son proteínas globulares plasmáticas producidas por los linfocitos B y utilizadas por el sistema inmune para identificar moléculas ajenas al organismo llamadas antígenos (Ags). Los anticuerpos convencionales están formados por dos cadenas pesadas (H, del inglés “heavy”) y dos cadenas ligeras (L, del inglés “light”) idénticas. Cada una de estas cadenas se componen de regiones variables (V_H y V_L) y regiones constantes (C_H y C_L). La especificidad de cada anticuerpo se encuentra definida por las regiones determinantes de complementariedad (CDRs, del inglés “complementarity determining regions”) localizadas en las regiones VH y VL, mientras que la región Fc media funciones efectoras. Los Abs han sido ampliamente utilizados tanto para detectar antígenos así como en terapias. Los Abs convencionales monoclonales (mAbs) son producidos mediante hibridomas utilizando ratones, pero estos mAbs producen una respuesta inmune no deseada en humanos lo que restringe su potencial terapéutico. Por ello, se han desarrollado estrategias alternativas para producir mAbs completamente humanos como son la producción de mAbs en animales transgénicos con loci de Ig humanos y la utilización de genotecas de Abs que se expresan en la superficie de bacteriófagos de *E. coli* y de células de levadura. Además, se han dedicado grandes esfuerzos en el desarrollo de reactivos de tipo anticuerpo basados en fragmentos de estos. Los fragmentos de Ab más comúnmente utilizados incluyen los fragmentos Fv monocadena (scFv, consisten en los dominios V_H y V_L unidos con un conector peptídico), los fragmentos Fab (generados mediante la fusión del dominio C con dominios V), y los anticuerpos monodominio (sdAbs, del inglés “single domain antibodies”) que son fragmentos de pequeño tamaño (peso molecular de aproximadamente 12-15 kDa), totalmente capaces de unirse a antígenos y formados por un único dominio variable (V). El tipo más común de sdAbs, también conocidos como “nanobodies” o V_Hs, es generado a partir de dominios V_H de anticuerpos que poseen únicamente cadenas pesadas (HCAbs, del inglés “heavy-chain-only antibodies”) y que se encuentran de forma natural en camélidos (llamas,
dromedarios, etc.). Su tamaño pequeño, sus propiedades biofísicas y su capacidad de unión a los antígenos hacen de los “nanobodies” unas moléculas muy atractivas para múltiples aplicaciones. Además, las secuencias de sdAbs son muy similares a las secuencias de los dominios V_{H3} humanos, por lo que pueden ser utilizados en terapia y en diagnóstico *in vivo* con un riesgo menor de efectos adversos.

Las genotecas de anticuerpos pueden ser obtenidas después de una inmunización del organismo donante con el antígeno de interés (en inglés “immune libraries”), a partir de donantes no inmunizados (colecciones naïve) o generadas artificialmente mediante el ensamblaje *in vitro* de los genes V (colecciones sintéticas). La técnica de la presentación de anticuerpos en fagos (en inglés “phage display”) es el método más común de presentación y selección de anticuerpos. Una de las principales ventajas del “phage display” es la posibilidad de realizar colecciones de gran tamaño (hasta $10^{10}$ clones) debido a la alta eficiencia del clonaje de genes V en vectores fágicos en *E. coli*. Sin embargo, este sistema tiene algunas limitaciones como la alta probabilidad de aislar clones no específicos de unión al antígeno por la tendencia de los fagos de adherirse inespecíficamente, la necesidad de infectar bacterias para generar partículas fago-Ab y la imposibilidad de analizar los clones obtenidos mediante citometría de flujo (FACS, del inglés “fluorescence activated cell sorting”) debido al pequeño tamaño de las partículas víricas. Un sistema alternativo utiliza células de levadura para la exposición de anticuerpos en la superficie celular y su posterior selección. Entre las ventajas de este método se encuentran la posibilidad de utilizar la técnica de citometría de flujo para la selección y caracterización de los anticuerpos aislados, la capacidad de las levaduras de realizar modificaciones post-traduccionales que pueden ser relevantes para la funcionalidad de los anticuerpos y la posibilidad de determinar directamente la afinidad de los anticuerpos expuestos. Sin embargo, las levaduras tienen menor eficiencia de transformación y tasa de crecimiento en comparación con *E. coli*. 
La presentación de anticuerpos en *E. coli* ofrece una alternativa atractiva a los sistemas de exposición de anticuerpos anteriormente descritos, ya que se obtienen altas eficiencias de transformación, tasas de crecimiento elevadas y la posibilidad de utilizar métodos de citometría de flujo para la selección y caracterización de los clones de anticuerpos aislados. A pesar de todas estas ventajas, la presencia de la membrana externa (OM, del inglés “outer membrane”) en *E. coli* interfiere con la correcta exposición del anticuerpo y ha dificultado el desarrollo de sistemas eficientes de presentación de anticuerpos en este microorganismo. Estudios realizados anteriormente en nuestro laboratorio con modelos de V_{HH} demostraron que los dominios β de anclaje a la OM pertenecientes al autotransportador EhaA y de la intimina de la cepa enterohemorrágica *E. coli* O157:H7 (EHEC) son sistemas muy prometedores para la exposición de sdAbs en la superficie de la cepa *E. coli* K-12.

**Objetivos**

Los objetivos del presente trabajo son:

1) Evaluar y comparar los dominios β de EhaA e intimina como sistemas de exposición de sdAbs funcionales en la superficie de *E. coli*.

2) Utilizar los sistemas de presentación de anticuerpos basados en los dominios β de EhaA e intimina para la expresión de genotecas de sdAbs de animales inmunizados y posterior selección de clones con alta afinidad por antígenos proteicos de origen bacteriano y humano.

3) Caracterizar los sdAbs seleccionados mediante la técnica de presentación en la superficie *E. coli*, determinar su afinidad de unión y especificidad mediante citometría de flujo, y comparar los resultados con aquellos obtenidos con tecnologías bioquímicas estándar como el ensayo por inmuoabsorción ligado a enzimas (ELISA, del inglés “enzyme-linked immunosorbent assay”) y la resonancia de plasmones de superficie (SPR, del inglés “surface plasmon resonance”).

4) Utilizar la plataforma de presentación de anticuerpos en *E. coli* para la selección directa de sdAbs utilizando células tumorales vivas que expresan un antígeno tumoral relevante en su superficie, como el receptor de factor
de crecimiento epidérmico (EGFR, del inglés “epidermal growth factor receptor”).

Resultados
En primer lugar analizamos la capacidad de los dominios β de EhaA e intimina (pertenecientes a EHEC) para exponer sdAbs individuales y colecciones post-inmunización de sdAbs originados a partir de V_{HH}s de camellos dirigidos contra el dominio extracelular del receptor translocado de intimina de EHEC (TirM_{EHEC}) y contra fibrinogéno humano (Fib). Demostramos que ambos sistemas pueden ser utilizados para exponer sdAbs funcionales en la superficie de células de E. coli con poca proteolisis y toxicidad celular, aunque las células de E. coli que exponían los sdAbs con el dominio β de intimina mostraron una mayor capacidad de unión al antígeno. Se examinaron ambos tipos de sistemas de presentación de E. coli con colecciones de sdAbs para buscar clones específicos de antígeno mediante separación magnética de células (MACS, del inglés “magnetic cell sorting”) utilizando antígenos biotinilados. Se aislaron clones con alta afinidad por sus respectivos antígenos utilizando los dos tipos de presentación de sdAbs, aunque el proceso fue más eficiente con el sistema del dominio β de intimina. Se determinó la especificidad de los clones seleccionados contra TirM_{EHEC} y Fib mediante análisis por citometría de flujo de células de E. coli, y con las técnicas ELISA y SPR utilizando sdAbs purificados. A continuación, utilizamos los sistemas de presentación de sdAbs en células de E. coli para estimar la afinidad de los sdAbs seleccionados mediante análisis de citometría de flujo en condiciones de equilibrio.

Adicionalmente, la plataforma de presentación de sdAbs con el dominio β de intimina fue empleada para la selección de V_{HH}s a partir de una genoteca inmune generada contra EGFR mediante la inmunización de llamas (Lama glama) con la línea celular humana de origen tumoral A431. Las bacterias que presentaban la colección de sdAbs en su superficie fueron sometidas a rondas sucesivas de selección sobre células negativas para EGFR (línea celular fibroblástica murina, NIH-3T3 2.2) y células positivas para EGFR, como HER14 (células NIH-3T3 2.2 transfectadas que expresan EGFR humano) para el enriquecimiento de clones que
unían específicamente EGFR. Se identificaron clones EGFR-específicos que se unían a células HER14 pero no a células NIH-3T3 2.2, y se confirmó su capacidad de unión específica a EGFR a través de análisis por citometría de flujo utilizando fusiones biotiniladas EGFR-Fc. Además, utilizamos el sistema de presentación de sdAbs en *E. coli* para determinar la afinidad de la unión de estos sdAbs a EGFR-Fc en presencia de un exceso de EGF, el ligando natural de EGFR. Nuestros datos demuestran que el sistema de presentación de sdAbs en *E. coli* permite la selección de sAbs contra antígenos relevantes asociados a tumores a partir de colecciones generadas mediante inmunización con células. Esta selección se puede realizar directamente sobre células tumorales sin la necesidad de utilizar el antígeno purificado.

Por último, desarrollamos un método alternativo para la producción y purificación de sdAbs del periplasma de *E. coli* basado en fusiones de sdAbs con la proteína de unión a maltosa (MBP, del inglés “maltose binding protein”) en el extremo N-terminal y con una cola de histidinas (His6) en el extremo C-terminal (MBP-V\textsubscript{HH}-His6). Se produjeron consistentemente niveles elevados de fusiones solubles MBP-V\textsubscript{HH}^\textsuperscript{His6} (>12 mg/L de cultivo inducido en matraces) en el periplasma de *E. coli* HM140, una cepa que carece de varias proteasas periplásmicas. Se recuperaron fusiones MBP-V\textsubscript{HH}^\textsuperscript{His6} de gran pureza y fusiones libres V\textsubscript{HH}^\textsuperscript{His6} (después de una proteolisis específica de sitio de las fusiones), tras realizar dos procesos de cromatografía de afinidad a amilosa y a metales. Se determinó la naturaleza monomérica de los V\textsubscript{HH}^\textsuperscript{His6} mediante cromatografía de filtración en gel. También demostramos mediante ELISA que las fusiones monoméricas V\textsubscript{HH}^\textsuperscript{His6} y MBP-V\textsubscript{HH}^\textsuperscript{His6} retenían su capacidad de unión al antígeno y su especificidad, facilitando su utilización directa en ensayos de reconocimiento de antígeno.
INTRODUCTION
1. Antibodies and Recombinant antibody fragments

1.1 Antibodies and their application in human therapy

Antibodies (Abs) or Immunoglobulins (Igs) are proteins that play a very important role in the immune system of humans and animals. These proteins are produced by B-lymphocytes and are deployed by the immune system to identify and target foreign or non-self molecules (Nelson et al., 2000; Ohlin and Zouali, 2003). Every human being has on an average, about one to two billion different Abs continuously flowing through his or her bloodstream, patrolling and fighting infections or diseases. The basic structure of an Ab consists of two identical light and heavy chains, linked by disulfide bonds. Each heavy and light chain contains a variable sequence (\(V_H\) and \(V_L\) respectively) in the amino terminal and constant sequences (\(C_H\) and \(C_L\)) in the remaining portion of the chain. Variability of the Abs, which accounts for their different specificities, is located in the \(V_H\) and \(V_L\), clustered in several hypervariable regions: the complementarity determining regions (CDRs). The CDRs determine the specificity of the Ab for its cognate antigen, thus allowing the detection of millions of different antigens present in nature. The Fc region of the Ab, on the other hand, is essential for mediating effector functions like antibody-dependent cytotoxicity (ADCC), Ab-dependent cellular phagocytosis, complement mediated lysis and regulation of cell activation or proliferation. Thus, the structure of the Ab (Figure 1A) determines its binding specificity and biological activity (Birch and Racher, 2006).

![Figure 1. Structure of conventional antibodies and antibody derived fragments.](image)

**Figure 1. Structure of conventional antibodies and antibody derived fragments.**

(A) Structure of conventional IgGs with heavy chain (H) and light (L) chains indicating the crystallizable fragment (Fc) region and antigen binding sites or CDRs. (B) Common antibody-derived antigen-binding fragments: single-chain Fv (scFv) and Fab are shown along with single domain \(V_H\) and \(V_L\) fragments.
Abs have been widely used for antigen detection as well as in therapeutics, and their specificity combined with low toxicity make them promising pharmaceutical molecules (Berger et al., 2002). At present, they comprise the second-largest category of biological medicines in clinical development, after vaccines (Seymour, 2004). However, therapeutic use of antibodies was initially limited by methodological constraints in raising them. Conventional monoclonal antibodies (mAbs) were commonly obtained by immunization of experimental animals, usually mice, with target antigens. Fusions of B-cells from immunized animals with myeloma cell lines produced hybridomas able to grow in culture and that could be screened for specific Ab-producing clones (Kohler and Milstein, 1975). Although well established, this technology is laborious and biased by the immune system of the experimental animal, which limits the ability to obtain high affinity Abs against conserved mammalian proteins of interest. Additionally, the heterologous character of these murine mAbs make them immunogenic to humans and elicit “Human anti-mouse antibody” (HAMA) response, which restricts their use in therapy (Schroff et al., 1985). Hence, human antibodies are of particular interest due to a lower immunogenic response. Human hybridomas were initially evaluated, but were found to be unstable, difficult to prepare and secreted low levels of mAbs of the IgM class with low affinities. Two other approaches to produce fully human mAbs from phage display libraries (McCafferty et al., 1990) or transgenic animals (Bruggemann et al., 1991) have been possible since the early 1990s. At present, about 30 mAbs produced by these methods have already been approved by the Food and Drug Administration (FDA) for therapy (Scott et al., 2012; Pandey and Mahadevan, 2014).

In order to be successful as therapeutic molecules, Abs need to have properties like high binding affinities, reduced immunogenicity, increased half-life in the human body and the ability to adequately recruit human effector functions. These factors combined with the knowledge that the segmented structure of the Ab molecule allows functional domains (carrying antigen-binding or effector functions) to be exchanged, has led to the construction of either, chimeric antibodies i.e. by coupling of the animal antigen-binding variable (V) region with the human constant domains or humanized antibodies, i.e. by grafting of the CDRs derived from murine antibodies with desired specificity onto carefully chosen human V_H and V_L.
Introduction

frameworks (FRs) (Figure 2). Six chimeric, fifteen humanized and ten fully human mAbs are currently on the market, including some Fabs.

**Figure 2. Chimeric and humanized antibodies.**

(From left to right) A murine antibody (blue), a human antibody (yellow), a chimeric antibody (with murine variable domains and human constant domains) and a humanized antibody (a human antibody with CDRs of murine origin).

Despite the general success of whole IgGs in therapy, the large size of these molecules leads to practical drawbacks such as lower tumor or tissue penetration, long serum half-life and subsequent low tumor-to-blood ratio; all or some of which makes difficult their use in applications like radioimmunotherapy and *in vivo* diagnostic imaging. Secondly, conventional mAbs have specificity for a single antigen epitope, and since most diseases are multifactorial, involving multiple ligands, receptors and signalling cascades, it would be beneficial to have bivalent or multivalent molecules that could target two or more epitopes. Lastly, the structural properties of a full-length Ab require sophisticated folding mechanisms as well as an oxidizing environment for the generation of disulfide bonds; since most traditional expression hosts do not provide such an environment, the industrial production of full length IgGs is mostly restricted to mammalian cells leading to slow production at high cost.

To circumvent these limitations of mAbs, substantial efforts have gone in to the development of the next wave of therapeutic and diagnostic reagents based on Ab fragments. In many cases, efforts have been made to improve or even delete some characteristics. For example, to achieve better tumor penetration or a better tumor-to-blood ratio for visualizing metastases, it would be preferable to have a relatively small Ab fragment with a fairly short half-life. On the other hand, the molecule should not be too small, in order to avoid rapid clearance from the body.
immediately after its application. In addition, it would be advantageous in certain applications to improve the effector functions, while in others, such as in radioimmunotherapy or in vivo imaging, it is imperative that Fc mediated cellular effects and prolonged half life do not exist.

1.2 Recombinant fragments from conventional antibodies

It is striking to note that the use of Ab fragments started as early as in the late 1950s after the pioneering work of Porter (Porter, 1958), which made possible the better understanding of the topology of IgGs and the functions associated with the different Ab fragments obtained after controlled proteolytic cleavage. Conversion of monovalent Ab fragments into multivalent formats increases functional affinity; decreases dissociation rates when bound to cell-surface receptors or polyvalent antigens and enhances biodistribution (Holliger and Hudson, 2005). These molecules are particularly useful in applications, where epitope binding is sufficient for the desired effect, including therapeutic applications such as virus neutralization or receptor blocking. The smallest antigen-binding fragment of IgGs maintaining its complete antigen-binding site is the Fv fragment, which consists only of V regions. However, the expression of an individual variable domain of a Fv, either V\text{H} or V\text{L} domain, exposes a large hydrophobic side to the aqueous environment that render these types of molecules difficult to handle. This shortcoming is overcome by introducing a soluble and flexible amino acid peptide linker linking the V\text{H} and V\text{L} domains, to form a scFv fragment, which leads to stabilization of the molecule (Bird et al., 1988). Also common is to add the constant (C) domains to the V regions to obtain a Fab fragment.

Till date, scFvs and Fabs are the most widely used Ab fragments produced in bacteria (Figure 1B). Other formats that have been produced in prokaryotic and eukaryotic cells include disulfide-bond stabilized scFvs (ds-scFv) (Schmiedl et al., 2000), single chain Fab fragments (scFab) (Hust et al., 2007) which combines properties of scFv and Fab, as well as di- and multimeric Ab formats like dia-, tria-, tetra-bodies (Hudson and Kortt, 1999; Carter, 2006), minibodies (Hu et al., 1996; Holliger and Hudson, 2005) or trimerbodies (Cuesta et al., 2012).
1.3 Heavy-chain only antibodies, single domain antibodies and nanobodies

In the early 1990s, a novel type of Abs lacking light chains was discovered in the serum of camelids (Hamers-Casterman et al., 1993). This unique class of heavy chain only antibodies (HCAbs) are apparently the outcome of recent adaptive changes occurring in conventional antibodies within the Camelidae lineage (Flajnik et al., 2011) and play a role in the immune response of these animals. A second type of HCAbs is found in sharks (Greenberg et al., 1995). The structure of camelid HCAbs is illustrated in Figure 3A. The homodimeric HCAbs lack light chains and thus antigen recognition is possible solely through the variable domain of the heavy chain, referred to as V_{HH} for V_{H} of HCAbs. The recombinant expression of V_{HH}s yields a soluble single domain Ab (sdAb) fragment with dimensions of 2-4 nm, and has been referred to as Nanobody (Nb). Nbs have a molecular weight of ~15 kDa, which is at least half the size of the intact antigen-binding site of a conventional Ab (i.e, the V_{H}-V_{L} pair) and are the smallest, intact antigen-binding fragments derived from a functional immunoglobulin (Figure 3A and 3B).

![Figure 3. Structure of camelid heavy chain only antibodies and nanobodies.](image)

(A) Schematic representation of a heavy chain only antibody with the structure of the V_{HH} domain highlighted. The CDRs are labelled in different colours: CDR1 in yellow, CDR2 in green, CDR3 in blue and the framework regions (FR) are indicated in red. (B) The linear structural diagram of a V_{HH} domain. CDRs are labelled with the same colour scheme as in part A and also the disulfide bonds in the molecule, i.e. canonical and additional are illustrated.
These molecules have naturally acquired important adaptations to remain soluble and functional in the absence of the associated light chain variable domain. Further, they have evolved long CDRs with novel conformations for antigen recognition, subtle amino acid adaptations including substitutions of conserved hydrophobic residues in classical V_Hs (V37, G44, L45, and F47/W47) for more hydrophilic amino acids (Y37/F37; E44/Q44; R45/C45; G47/R47/L47/S47) (Muyldermans, 2013), that confer them strict monomeric behaviour, reversible folding properties, resistance to proteolysis and thermal degradation when compared with the V_H from conventional antibodies (van der Linden et al., 1999; Dumoulin et al., 2002). Furthermore, V_HHs often contain, besides the canonical disulfide bond of Ig domains, an extra-disulfide bond connecting CDR3 and CDR1 (in camels) or CDR3 with CDR2 (in llamas) that assist in stabilizing the conformation of these CDRs and the overall stability of the domain (Figure 3B) (Govaert et al., 2012).

The small size and long CDRs of V_HHs allow them to recognize epitopes located in clefts and protein cavities (Figure 4A) such as active sites of enzymes, conserved inner regions of surface proteins from pathogens, which are frequently less accessible to large molecules and conventional Abs (Desmyter et al., 1996; Lauwereys et al., 1998; Stijlemans et al., 2004). In addition, they share high sequence identity with human V_H sequences of family 3 (Figure 4B), which opens the possibility of their use in human therapy and in in vivo diagnosis (Wesolowski et al., 2009; Vaneycken et al., 2011). Apart from these advantages, the recombinant expression of V_HHs turned out to be favourable as well, as is validated in various expression systems. They are preferentially expressed in E. coli, where they can be produced economically as soluble and non-aggregating recombinant proteins with an incorporated histidine tag for easy purification (Arbabi-Ghahroudi et al., 2005), though the levels of soluble V_HHs produced are highly variable (Alvarez-Rueda et al., 2007). Higher production levels of V_HH can be obtained using high-density cultures in yeast systems like Pichia pastoris and Saccharomyces cerevisiae (Frenken et al., 2000) or in tobacco plants (Rajabi-Memari et al., 2006).
Figure 4. Comparison between conventional VH/VL and VH₃ antibody fragments.

(A) Three-dimensional structure of chicken lysozyme recognized by an scFv (VH/VL domains) (left; PDB code 1mlc) or by a VH domain of camelid HCAb (right; PDB code 1mel). The CDRs of VH, VL and VH₃ are shown in different colors (CDR1 in orange, CDR2 in green, CDR3 in cyan). The long CDR3 loop of the VH₃ protrudes as a finger-like structure toward an epitope located in the cavity of the active site of lysozyme (right), whereas the flat VH/VL paratope binds to an epitope located at the exposed surface of lysozyme (left).

(B) Cartoon showing the polypeptide chains of conventional VH and camelid VH₃s, indicating the amino acids of framework region 2 of classical VHs (V37, G44, L45, W47) that are substituted by more hydrophilic amino acids (e.g. Y37, E44, R45, G47) in VH₃s. Other distinct features shown are the different length of CDR1 and CDR3 in VHs and VH₃s, the presence of the conserved canonical disulfide bond between residues C22 and C92, and the extra-disulfide bond between Cys residues of CDR1 and CDR3 in VH₃s.

Because of their single domain nature, VH₃s, offer several advantages for biotechnological applications. Libraries of VH₃s from immunized camels and llamas can be generated through a straightforward cloning procedure in various display systems and selected using antigen immobilized in plates or on cells (Arbabi Ghahroudi et al., 1997; Roovers et al., 2007; Ahmadvand et al., 2009; Ryckaert et al., 2010; Fleetwood et al., 2012). VH₃s within these libraries retain full functional diversity resulting in isolation of high-affinity antigen-binding clones. The short serum-life due to renal clearance may limit the efficacy of VH₃s in therapeutic applications. Therefore, bispecific VH₃s recognizing long-lived serum proteins like albumin or immunoglobulins and the therapeutic target have been generated, thus resulting in increased half-lives (Tijink et al., 2008; Vosjan et al., 2012). The binding and stability properties portrayed by natural VH₃s has led to considerable interest in the development of human sdAbs based on human VH (or VL) sequences that could
mimic natural V\textsubscript{HH}s. Murine V\textsubscript{H} sdAbs against lysozyme were isolated from phage display in \textit{E. coli} before the discovery of natural HCAbs in camels (Ward et al., 1989). However, most V\textsubscript{H} sequences from conventional Abs tend to aggregate in solution and show low affinity for antigens. Introduction of ‘camel’ mutations G44E, L45R, and W47G, in a human V\textsubscript{H}3 sequence significantly increased its solubility (Davies and Riechmann, 1994). Human V\textsubscript{H} clones with low tendency to aggregate were isolated by repeated cycles of heating and cooling of phages with displayed V\textsubscript{H} sequences (Jespers et al., 2004). Based on these findings, synthetic libraries of human V\textsubscript{H}s have been developed by randomizing CDRs in ‘camelized’ and/or selected human V\textsubscript{H} sequences and used for isolation of functional sdAbs with nanomolar affinity (Davies and Riechmann, 1994; Dumoulin et al., 2002; Holt et al., 2008; Arbabi-Ghahroudi et al., 2009). sdAbs based on human V\textsubscript{L} domains have also been reported and expressed in \textit{E. coli}. These include V\textsubscript{L} domains from conventional IgGs and naïve scFv libraries (Colby et al., 2004; Martsev et al., 2004; Cossins et al., 2007; Schiefner et al., 2011) and sdAbs isolated from synthetic libraries (van den Beucken et al., 2001; Holt et al., 2008).

2. Selection of antibodies from libraries using surface display.

2.1 Libraries of antibody genes

Combinatorial Ab engineering or directed evolution is an efficient approach for generation of Abs with new or improved properties (e.g. affinity) and is based on the generation of molecular libraries containing up to hundreds of billions of different Abs, from which specific binders might be isolated by display and high-throughput screenings or selections. One of the considerations to be taken in to account when constructing a library is the size. A larger library covers a larger part of the theoretical repertoire and results in a higher probability of containing and isolating high-affinity clones. This correlation has been supported by experimental data generated in different studies (Bradbury and Marks, 2004). Equally important is that the Abs in the library are functional. Ab libraries can be classified into three main types, namely immune, naïve and synthetic libraries.
Immune libraries are generated from the V genes of B-cells isolated from the lymph nodes, spleen, or peripheral blood of immunized animals or human donors (Clackson et al., 1991) and takes advantage of the diversity created in vivo by the immune system of the donor. These types of libraries are enriched in antigen-specific Abs, some of which have already been affinity matured by the immune system of the donor, and thus serve as a rich source of Abs with higher affinities than those obtained from other sources. In addition, large-sized immune libraries are not necessary in order to obtain high-affinity antibodies. Multiple immune libraries have been constructed and reported, including mouse (Clackson et al., 1991; Chester et al., 1994; Kettleborough et al., 1994), human (Barbas et al., 1993), chicken (Yamanaka et al., 1996), rabbit (Lang et al., 1996) and camel (Arbabi Ghahroudi et al., 1997). The amplified V genes of an immune library are then cloned into a display vector for selection of binders with high affinity and specificity to the target antigen. Despite these advantages, immune libraries also have some limitations. Firstly, active immunization is not always possible due to ethical constraints, or is not effective due to toxicity to or tolerance mechanisms towards the antigen used. Secondly, immune libraries require repeated antigen immunization and library construction for each antigen used. Ideally, universal and antigen-unbiased libraries should be screened to select high affinity Abs to any chosen antigen independent of the immune history. At present, a few such libraries have already been described, either naïve or synthetic (depending upon the source of the immunoglobulin genes), and are particularly useful for the selection of human Abs.

A naïve library, if sufficiently large and diverse, can be used to generate Abs for a large panel of antigens, including self, non-immunogenic and relatively toxic antigens. The murine naïve repertoire has been estimated to contain <5x10^8 different B-lymphocytes, while the human repertoire maybe a hundred to a thousand times bigger (Winter et al., 1994). The affinity of Abs selected from a naïve library is proportional to the size of the library, ranging from 10^{-6} to 10^{-7} M for a library with 10^7 clones (Griffiths et al., 1994), to 10^{-8} to 10^{-10} M for a large repertoire of about 10^{10} clones (Vaughan et al., 1996). Synthetic Ab repertoires contain Abs built artificially by the in vitro assembly of V-gene segments and D/J segments. V-genes may be assembled by introducing a predetermined level of CDR randomization into
germline V-gene segments or re-arranged V-genes. The regions and degree of diversity may be chosen to correspond to areas of highest natural diversity of the Ab repertoire, like in the loop central to the antigen-combining site i.e. the CDR3 of the heavy chain. Several synthetic libraries have been constructed and been used to select Abs against different antigens like haptens (Barbas et al., 1992; Hoogenboom and Winter, 1992), proteins (Nissim et al., 1994) and cell-surface markers (de Kruif et al., 1995), but their affinities are typically around $10^{-6}$ to $10^{-7}$ M.

2.2 Surface display methods for screening of Ab libraries

In vitro display technology is based on the linkage of phenotype to genotype and is currently one of the major technologies for creating mAbs for human therapy, in addition to the use of transgenic mice carrying human Ig genes (Jakobovits et al., 2007) and the humanization of mAbs (Lee et al., 2014). It involves the display and screening of Ab gene libraries on the surface of a biological entity e.g. bacteriophages, yeasts or ribosomes (Hoogenboom, 2005). Selection systems include the in vitro display technologies as well as protein complementation-based strategies. The function of a selection system is to enable isolation of specific affinity proteins from the majority of non-binding background proteins present in the library, as well as to co-select the encoding DNA in order to simplify amplification and identification. It is important that selection pressure is dominated by high affinity for the antigen, while potential biases like differences in expression levels or amplification rate are minimized. Many different selection methods have been widely used, including selection on antigen immobilized on polystyrene plates or columns, selection on biotinylated antigen in solution, selection on mammalian cells expressing a surface antigen (Cai and Garen, 1995; de Kruif et al., 1995), enrichment on tissue sections (Van Ewijk et al., 1997) and selection in vivo using living animals (Zou et al., 2004; Deramchia et al., 2012).

2.2.1 Phage display

In 1985, George Smith reported the use of recombinant DNA technology for display of heterologous peptides on the surface of filamentous phages, laying the ground for the “phage display” technology (Smith, 1985). This technology has proven to be a very powerful technique to display libraries containing billions of different peptides.
and proteins, including Abs and is the most common Ab display method in use. DNA encoding Ab fragments is cloned in phagemids as a fusion to the phage coat proteins (e.g. pIII, pVIII). Upon expression of the Abs, the coat protein fusion is incorporated into new phage particles that are assembled in bacteria. Expression of the phage coat protein fusion product, and its subsequent incorporation into the mature phage coat results in the Ab or Ab fragment being presented on the phage surface. Bacteriophages displaying relevant Abs or Ab fragments on the surface (Figure 5A and 5B) are retained during selection methods, while non-adherent phages are washed away (Figure 5C). Bound phages are recovered by elution, reinfected into E. coli and re-grown for further enrichment and eventually analysis of binding. The most commonly used phage for display is the non-lytic filamentous phage, M13 and the most extensively used phage coat protein is the pIII (present in three to five copies per phage particle), which is involved in recognition of F-pilus and Tol-proteins in the bacterial cell envelope during infection (Winter et al., 1994; Vodnik et al., 2011). The vast majority of reports using phage display utilize scFv libraries (Hust and Dubel, 2004), since the smaller size and the composition from a single polypeptide facilitate production and folding in E. coli, thus allowing a more efficient presentation of the Ab repertoire on the phage surface. More complex proteins like Fab fragments or full-length Abs have also been expressed, but the necessity to produce two different polypeptide chains, has led to construction of bicistronic expression vectors with increased complexity (Hust and Dubel, 2005; Mazor et al., 2010). The major advantages of phage display include the large library sizes that can be attained in E. coli, i.e. $10^{10}$ members, and the high titres that can be produced ($>10^{13}$ phage particles/ml) thus accounting for a good representation of all clones in the library. Secondly, purification of soluble Ab fragments is possible directly, after selection by phage display, without the need for re-cloning in a different protein expression vector. However, phage display suffers limitations such as high background binding due to the sticky nature of phages, the need for bacterial infection to produce phage-Ab particles in each round of selection, including titre of the phage particles produced, leading to longer experimental procedures, and the unfeasibility of fluorescence activated cell sorting (FACS) or flow cytometry for analysis due to the small size of the phage particles.
Figure 5. Phage display of antibodies.

(A) Soluble Ab fragments, depicted as a scFv, are exported to the periplasm with an N-terminal signal peptide (e.g. PelBss). The Ab fragment can also be fused at its C-terminus to the minor coat protein III (pIII) of filamentous phages for phage display. In this case, the Ab fragment is exposed to the periplasm but tethers to the inner membrane (IM) of E. coli by pIII. (B) Phage-scFV particles assembled upon infection of E. coli cells expressing Ab-pIII fusions with a helper phage. The Ab-pIII fusion is displayed on the phage capsid (usually one copy per phage), which contains the phagemid DNA with the Ab gene. (C) Scheme of the steps involved during panning of phage display libraries on plates coated with antigen.
2.2.2 Cell surface display

Cell surface display systems are effective alternative methods for Ab engineering by directed evolution, and show great promise in both medical and industrial applications. Various cell surface display systems have been developed in a range of host organisms, including Gram negative bacteria (Daugherty, 2007; Mazor et al., 2007), Gram positive bacteria (Wernerus and Stahl, 2004; Fleetwood et al., 2012), yeast (Boder and Wittrup, 2000; Feldhaus and Siegel, 2004; Wen et al., 2011; Boder et al., 2012), insect cells/baculovirus (Makela and Oker-Blom, 2008) and mammalian cells (Ho et al., 2006; Beerli et al., 2008). These systems complement phage display and cell-free protein engineering systems. Currently, yeast and bacteria are the most commonly used cell surface display platforms for protein engineering.

2.2.2.1 Yeast Display

The eukaryotic secretory pathway, coupled with protein folding and quality control mechanisms of yeast allows a wide variety of proteins to be displayed, including many proteins with complex folds like single-chain T-cell receptors (scTCRs) (Richman et al., 2009; Aggen et al., 2011), human major histocompatibility (MHC) class II molecules (Wen et al., 2011), human epidermal growth factor (EGF), cytokines, extracellular domains of epidermal growth factor receptor (EGFR), scFvs (Feldhaus and Siegel, 2004) and Fabs (Blaise et al., 2004; Weaver-Feldhaus et al., 2004; Lin et al., 2012). Proteins are displayed on the yeast Saccharomyces cerevisiae cell surface via fusion to the α-agglutinin yeast adhesion receptor, which is located in the yeast cell wall (Figure 6A). The display level on the cell is variable (on an average about $3 \times 10^4$ fusions per cell for a scFv), but the intrinsic avidity of this display system is counteracted by the power of cell sorting. By staining cells with both fluorescently labelled antigen and anti-epitope tag reagent, the yeast cells can be sorted according to the level of antigen binding and Ab expression on the surface (Figure 6B). Non-immune human scFv libraries with over $10^9$ members have been generated and efficiently selected using magnetic beads (Yeung and Wittrup, 2002; Feldhaus et al., 2003; van den Beucken et al., 2003) and then sorted by flow cytometry to yield Abs with nanomolar affinities (Feldhaus et al., 2003).
Figure 6. Yeast display of antibodies.

(A) Ab fragments, depicted as a scFv in this scheme, are fused at its C-terminus to the Yeast agglutinin receptor (Aga) present on the yeast surface for display. (B) A schematic of Fluorescence assisted cell sorting (FACS) using yeast-displayed Ab fragments is depicted.

Yeast display has been used with a much more diverse range of proteins than prokaryotic hosts. A comparative study carried out using phage and yeast display; with a HIV-1 immune scFv library against the HIV-1 gp120 envelope glycoprotein showed that although both libraries yielded the same six scFvs, yeast display identified an additional twelve clones. Only five of these twelve clones could be displayed on the phage particle, likely due to the greater protein folding capabilities of yeast (Bowley et al., 2007). Limiting factors of yeast display include the smaller library sizes due to the low transformation efficiency of yeast, longer experimental time due to slower culture growth and low speed of fluorescent cell sorting (FACS).

A table comparing the two most common Ab display methods in use i.e phage display and yeast display is illustrated in Table 1.
### Table 1. Benefits and Limitations of Phage display and Yeast Display.

<table>
<thead>
<tr>
<th>Phage Display</th>
<th>Yeast Display</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benefits</td>
<td>Limitations</td>
</tr>
<tr>
<td>1) A highly robust Ab display system. Phages can be produced and handled in high densities (upto 10^11 phages/ml)</td>
<td>1) High background levels due to sticky nature of phage particles.</td>
</tr>
<tr>
<td>2) Large library sizes can be commonly obtained (upto 10^9 members).</td>
<td>2) Sticky nature of phages makes difficult the use of complex antigenic surfaces such as cells for selection.</td>
</tr>
<tr>
<td>3) Use of stringent conditions of wash and elution is possible.</td>
<td>3) Use of high throughput methods based on flow cytometry is not possible due to the small size of phages.</td>
</tr>
</tbody>
</table>

#### 2.2.2.2 Bacterial cell surface display

Although both phage display and yeast display are good Ab display systems, they suffer from some limitations (Table 1). Hence, it would be beneficial to develop an Ab display system that could combine the major advantages of both methods, while improving on the limitations. Bacterial display could be an alternative to these existing platforms of Ab display. Firstly, the larger size of bacteria than phages allows the use of flow cytometry based methods for selection, screening and characterization of clones. Secondly, bacteria are less sticky than phages, could result in lower background levels and selection could be carried out more easily using complex antigenic surfaces such as cells. Lastly, the faster growth of bacteria such as *E. coli* compared to that of yeast, greatly reduces experimental times and this fact combined with the highly versatile protein expression systems and methods available in *E. coli*, makes this method of Ab display very attractive. *Escherichia coli* is currently the most commonly used host for protein engineering by bacterial surface display, though a few reports used Gram-positive bacteria for display of enzymes and Ab fragments (Kronqvist et al., 2008; Fleetwood et al., 2012).

#### 2.3  *E. coli* cell surface display of Ab libraries

The expression of Abs in *E. coli*, both full-length IgGs molecules and smaller antigen-binding fragments, provides a set of powerful technologies for the
generation of Abs with novel specificities and improved properties (Holliger and Hudson, 2005; Beck et al., 2010). A major advantage of this display platform is that the relatively high transformation efficiencies of *E. coli* allow Ab libraries with up to $10^{10}$ members to be constructed. Moreover, *E. coli* remains a suitable microorganism for the generation, amplification and maintenance of large Ab repertoires owing to its versatile protein expression and secretion systems.

Although *E. coli* lacks eukaryotic protein folding and post-translational modification machineries, it allows expression of many eukaryotic proteins or protein domains in functional form, including the V domains of Iggs. *E. coli* cells have an inner membrane (IM) and an outer membrane (OM) separated by periplasmic space. The presence of the OM in *E. coli* has hindered the development of effective cell surface display methods for Ab selection. In addition, display of heterologous proteins on the surface of *E. coli* often affects cell physiology, resulting in growth arrest, cell lysis and poor display performance. In order to be displayed on the surface, the polypeptides produced in the cytoplasm have to traverse the IM, the periplasm and be inserted into the OM. Thus, OM proteins native to the organism have served as convenient targeting and anchoring portions and are widely used for display, though heterologous anchor proteins have also been evaluated. Several *E. coli* surface display systems have been described for display of peptides, enzymes and other proteins (Bessette et al., 2004; Rice et al., 2006; Hall et al., 2007; Dane et al., 2009; Kenrick and Daugherty, 2010; Little et al., 2011; van Bloois et al., 2011), but very few of these are focussed on the surface display of Abs, either full length or Ab fragments (Figure 7).

Georgiou and co-workers demonstrated in a proof-of-principle experiment that scFvs could be functionally displayed on the surface of *E. coli* and that FACS could be employed to isolate target-binding bacteria spiked in the background of non-binding cells (Francisco et al., 1993), and performed the first cell sorting of a library of scFv mutants. Both these experiments displayed anti-digoxin scFv libraries containing up to $4 \times 10^6$ scFvs on *E. coli* as N-terminal fusions to a chimeric lipoprotein (Lpp-OmpA'). Lpp-OmpA' is composed of the N-terminal signal peptide and the first nine residues of the mature Lpp fused to the residues 46 to 159 of
OmpA, which is a truncated fragment of its native 8-stranded β-barrel. The Lpp-OmpA’ format was the first bacterial display system to be combined with flow cytometry for library screening and used in protein engineering applications (Daugherty et al., 1998). scFv variants that bound digoxin with a $K_D$ of $<1$ nM were identified. However, Lpp-OmpA’ fusions were found to be toxic to the host cells primarily due to membrane disruption, and resulted in lysed cultures following protein induction (Christmann et al., 1999; Daugherty et al., 1999).

Figure 7. *E. coli* display of antibodies. (A) Lpp-OmpA’ display (B) Anchored Periplasmic Expression (APEx) of scFvs and IgGs.

Displaying proteins of interest on the periplasmic side of the IM of *E. coli* is also a valuable display system for protein engineering. Advantages of this system include the lack of complex carbohydrates like LPS, that could sterically interfere with protein binding to large targets, and that the protein of interest is required to traverse only one membrane, which avoids factors that may restrict export. The anchored periplasmic expression (APEx) system anchors the protein of interest to the inner membrane, either as an N-terminal fusion to a six-residue sequence derived from the native *E. coli* lipoprotein NlpA or as a C-terminal fusion to the pIII gene that codes for the minor coat protein of M13 (Harvey et al., 2004). A slight modification of the APEx system enabled the display of full-length Abs. Abs are secreted into the bacterial periplasm, where they are captured by an inner membrane protein that binds their Fc domain. After permeabilization of the outer membrane, the displayed antibodies could interact with fluorescently labeled antigen, and binding clones were selected by flow cytometry. As the anchoring protein is not encoded by the plasmid carrying the antibody genes, soluble, full-length, aglycosylated IgG molecules can subsequently be produced by transformation of a bacterial strain devoid of the Fc-binding anchor protein (Mazor
et al., 2007). However, library screening involves the permeabilization of the OM (to permit access of the fluorescently labelled antigen to the periplasm) generating spheroplasts followed by sorting using FACS of these spheroplasts. Disruption of the OM and spheroplast production decreases dramatically the viability of bacteria and the isolated clones after FACS cannot be amplified by simple growth, thus requiring amplification of the enriched Ab genes by PCR and subcloning for the next round of sorting. Hence, although both the Lpp-OmpA⁻ and APEx are good technologies for Ab selection, alternative methods that enable direct display of Abs or Ab libraries on the outer surface of *E. coli* cells, with little cellular toxicity and without the need for generation of Phabs or spheroplasts, would be of great interest.

Integral OM proteins and secreted proteins with OM-anchoring domains have also been used for display of Abs on the surface of *E. coli* (Lofblom, 2011), of which the autotransporters (AT) and Intimin families stand out by their apparent simplicity and modularity as attractive systems (Wentzel et al., 2001; Rutherford and Mourez, 2006; Jose and Meyer, 2007; Wilhelm et al., 2011; Leo et al., 2012; Leyton et al., 2012b; Nicolay et al., 2013; Zude et al., 2013). These proteins play important roles in the virulence and survival of both pathogenic and environmental Gram-negative bacteria. Protein members of the AT and Intimin families are a part of the type V secretion system of Gram-negative bacteria (Leo et al., 2012). Intimins and ATs are large, secreted polypeptides that contain three functional regions: i) a N-terminal SP, that drives their Sec-dependent translocation across the IM; ii) a β-domain, that is anchored into the OM by a 12-stranded β-barrel with an internal peptide linker; and iii) a *passenger* region, that is secreted to the extracellular milieu (Leyton et al., 2012a). The passenger domains are highly diverse and carry the specific AT functions which maybe enzymatic, proteolytic, cytotoxic or adhesive (in the case of Intimin), and contribute to colonization, immune evasion and biofilm formation. Although, their mechanism of secretion remains uncertain, both AT and Intimin appear to use the β-barrel assembly machine (BAM) complex of the OM for insertion and translocation of the *passengers* to the cell surface (Figure 8) (Bernstein, 2007; Bodelon et al., 2009; Leyton et al., 2012a; Gruss et al., 2013; Noinaj et al., 2013).
Autotransporters and Intimin are members of the T5SS and their secretion to the extracellular medium shares a common mechanism. Secretion of Intimin is shown in the Figure. Intimin is translocated across the IM via the Sec-translocon. In the periplasm, DsbA catalyzes the formation of the disulfide bond (S-S) present in the D3 domain of Intimin. Intimin mostly follows the SurA pathway for its delivery to the BAM complex, which is responsible for folding and OM insertion of its β-barrel domain. The alternate DegP/Skp pathway is used less frequently and is depicted with a dashed line. The LysM domain is predicted to bind the peptidoglycan (PG) layer. Translocation of the secreted domains D0 (green oval), D1, D2 and D3 is believed to take place concomitantly with insertion of the β-domain in the OM via the BAM complex. Based on (Bodelon et al., 2009; Fairman et al., 2012)

Despite their similarities, AT and Int/Inv proteins have important differences. They have opposite topological organization in the OM, the exposed passenger being located in the N-terminal region in AT, whereas is found at the C-terminus in the case of Int/Inv proteins (Figure 9). The distinct topologies are also reflected in their β-domains. In the case of AT, the first strand of the β-barrel is preceded by an α-helix linker that fills the lumen and connects its N-terminus with the passenger region (Oomen et al., 2004; Barnard et al., 2007). In contrast, the last strand of the β-barrel of Int/Inv proteins is followed by a peptide linker that runs through the lumen in an extended conformation connecting its C-terminus to the passenger region (Fairman et al., 2012). In addition, the natural passenger domains of AT and Int/Inv have distinct structures, i.e. β-helical rods in most ATs and tandems of Ig-like domains in Int/Inv proteins (Luo et al., 2000; Otto et al., 2005).
Figure 9. Modular organization and structure of AT and Intimin proteins.

(A) Scheme of AT proteins, showing the N-terminal SP, a secreted passenger domain, and an OM anchored C-terminal \( \beta \)-domain. Crystal structures of the passenger domain of Hbp (an AT protein from \textit{E. coli}) (Otto et al., 2005) and \( \beta \)-domain of NalP (an AT protein from \textit{N. meningitidis}) are shown.

(B) Scheme of Intimin, showing the N-terminal SP, an OM anchored \( \beta \)-domain and a C-terminal passenger domain. The passenger consists of three Ig-like domains (D0, D1 and D2) and a lectin-like domain (D3). Crystal structures of the passenger domain (Luo et al., 2000) and the OM \( \beta \)-barrel (Fairman et al., 2012) of Intimin are shown.

Previous studies from our laboratory showed that the AT protein, IgA protease (IgAP) from \textit{Neisseria gonorrhoeae} could display model scFv and sdAb clones on the surface of \textit{E. coli} fused to its \( \beta \)-domain (Veiga 2004). In addition, improved display levels of a model sdAb on \textit{E. coli} cells were observed when the \( \beta \)-domain of EhaA, an AT from enterohemorrhagic \textit{E. coli} (EHEC) O157:H7 was used (Marin et al., 2010). The displayed sdAb was correctly folded with the canonical disulfide bond of V domains formed by the action of DsbA (Veiga et al., 2004; Marin et al., 2010). In a different study, the expression of Intimin from EHEC and enteropathogenic \textit{E. coli} (EPEC) strains was demonstrated in \textit{E. coli} K-12 cells with the display of its native Ig-like and lectin-like domains on the bacterial surface with a disulfide bond formed by DsbA (Bodelon et al., 2009). However, the utility of the \( \beta \)-domains of EhaA and Intimin for the display of Ab libraries had not been investigated. Hence, in this work we have evaluated these systems for display and selection of sdAbs from immune libraries raised against diverse antigens of biomedical importance, viz. a protein involved in the infection of enteropathogenic and enterohemorrhagic \textit{E. coli} strains (i.e. Tir); a human serum protein that plays an important role in the blood coagulation (i.e. fibrinogen); and a tumor antigen overexpressed on the cell surface of many epithelial tumors, i.e. Epidermal growth factor receptor 1 (EGFR/HER1).
OBJECTIVES
The objectives of this work were as follows:

1) To evaluate and compare the β-domains of Intimin and EhaA proteins from EHEC for the display of functional sdAbs on the surface of *E. coli*.

2) To utilize the *E. coli* display systems based on the β-domains of Intimin and EhaA proteins for expression of immune libraries of sdAbs and selection of high-affinity clones binding protein antigens of bacterial and human origin.

3) To characterize sdAbs selected by *E. coli* display and determine their binding affinities and specificity using flow cytometry of *E. coli* cells, as well as compare the data with that obtained with standard biochemical technologies (ELISA and SPR).

4) To apply the *E. coli* display platform for direct selection of sdAbs on live tumor cells expressing a relevant antigen such as EGFR on their surface.
MATERIALS AND METHODS
1. Bacterial strains

The bacterial strains used in this work are described in Table 2.

Table 2. Bacterial strains used in this work.

<table>
<thead>
<tr>
<th>Name</th>
<th>Genotype / Characteristics</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>BL21 (DE3)</td>
<td>F- ompT hsdS6(rB-, mB-) gal dcm lon λ(DE3) [lacI lacUV5-T7 gene1 ind1 sam7 nin5])</td>
<td>Novagen</td>
</tr>
<tr>
<td>DH10B-T1\textsuperscript{T}</td>
<td>F- mcrA Δmrr-hsdRMS-mcrBC φ80lacZDM15 ΔlacX74 recA1 araD139 Δ(ara,leu) 7697 galU galK rpsL (StrR) nupG tonA λ-</td>
<td>Invitrogen</td>
</tr>
<tr>
<td>MG1655</td>
<td>K-12 (F- λ-)</td>
<td>(Blattner, 1997)</td>
</tr>
<tr>
<td>EcM1</td>
<td>MG1655ΔfimA-H</td>
<td>(Blomfield et al., 1991)</td>
</tr>
<tr>
<td>UT5600</td>
<td>K-12 (F- λ-) Δ(lac-proAB), galE, strA, nal, F' [lacZDM15, proAB]</td>
<td>(Grodberg and Dunn, 1988)</td>
</tr>
<tr>
<td>WK6</td>
<td>Δ (lac-proAB), galE, strA, nal, F' [lacZDM15, proAB]</td>
<td>(Zell and Fritz, 1987)</td>
</tr>
<tr>
<td>HB2151</td>
<td>K-12 Δ(lac-proAB), ara, nal\textsuperscript{T}, thi, F' [lacZDM15, proAB]</td>
<td>(Carter et al., 1985)</td>
</tr>
<tr>
<td>HM140</td>
<td>F- ΔlacX74 galE galK thi rpsL (StrA) ΔPhoA (PvuII) degP ptr ompT eda tsp rpoH15</td>
<td>(Meerman and Georgiou, 1994)</td>
</tr>
</tbody>
</table>

2. Conditions of bacterial growth

The \textit{E. coli} DH10B-T1\textsuperscript{T} strain was used in the preparation of electrocompetent cells for use in the cloning experiments. Bacteria carrying plasmids with V\textsubscript{HH} were grown at 30 °C in Luria–Bertani (LB) liquid medium or on agar plates with the appropriate antibiotic for plasmid selection. LB plates and pre-inoculum media prior to induction contained 2% (w/v) glucose for repression of the \textit{lac} promoter. The preinocula cultures were started from individual colonies (for single clones) or from a mixture of clones (in case of libraries), freshly grown and harvested from plates, diluted to an initial OD\textsubscript{600} of 0.5, and grown overnight (o/n) under static conditions. For induction, bacteria (corresponding to an OD\textsubscript{600} of 0.5) were harvested by centrifugation (4000
Materials and Methods

xg, 5 min), and grown in the same media with 0.05 mM isopropylthio-β-D-galactoside (IPTG), but without glucose for 3 h with agitation (160 rpm), unless indicated otherwise. For over-expression of soluble V_{HH} in the periplasm, E. coli WK6 cells with the corresponding pCANTAB6-V_{HH} plasmid (Ap^{R}) were induced with 0.3 mM IPTG for 3 h at 30 ºC. Secretion of V_{HH} into the culture media was performed using the hemolysin (Hly) secretion system of E. coli HB2151 cells carrying pVDL9.3 (HlyBD; Cm^{R}) and the corresponding pEHlyA4SD-VHH (Ap^{R}) plasmids, induced with 0.3 mM IPTG for 6 h. Over-expression of soluble TirM_{EHEC} with N-terminal His-tag was induced in E. coli BL21 (DE3) cells carrying the pET28a-TirM_{EHEC} plasmid (Km^{R}) and grown at 37 ºC in LB medium containing 1.0 mM IPTG for 2 h. For over-expression of soluble V_{HH} as MBP fusions in the periplasm, E. coli HM140 strain with the corresponding pMAL1-V_{HH} plasmid (Ap^{R}) was induced with 0.3 mM IPTG for 3 h with agitation (250 rpm).

3. **Cell lines and Growth conditions**

Frozen aliquots of mouse fibroblast cells, viz. NIH 3T3 2.2 and HER14 were thawed and seeded into culture flasks (BD Falcon) containing 25 ml DMEM (Sigma) supplemented with 10% Fetal bovine serum (FBS, Sigma), 2 mM Glutamine, and grown in an incubator (Thermo Scientific) at 37º C with 5% CO_{2}. Cells were subcultured periodically, when the cultures attained a confluency of ~80%. In experiments of selection of V_{HH} libraries on cells, growing cells in culture were trypsinized and seeded in 6-well plates (Falcon) containing growth media to a confluency of ~20%, the day before the experiment. For Immunofluorescence microscopy (IFM) experiments, cells were seeded and grown on circular coverslips (previously irradiated with UV for 30 mins) present in 24-well plates (Falcon), prior to infection with bacteria, staining and analysis under the microscope.

4. **DNA constructions**

The plasmid constructions used in this work were carried out using standard techniques of DNA amplification and modification as described by Ausubel et al, 1997.
<table>
<thead>
<tr>
<th>Name</th>
<th>Characteristics</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>pAK-Not</td>
<td>(Cm&lt;sup&gt;+&lt;/sup&gt;), lac&lt;sup&gt;+&lt;/sup&gt;-Plac promoter, pBR322 ori</td>
<td>(Veiga et al., 1999)</td>
</tr>
<tr>
<td>pHEA</td>
<td>pAK-Not derivative; for fusions to C-EhaA (pelB-His-EhaA&lt;sub&gt;989-1327&lt;/sub&gt;)</td>
<td>(Marin et al., 2010)</td>
</tr>
<tr>
<td>pVgfpa</td>
<td>pHEA-derivative; Vgfpa fused to C-EhaA (pelB-Vgfpa-E-tag-EhaA&lt;sub&gt;989-1327&lt;/sub&gt;)</td>
<td>(Salema et al., 2013)</td>
</tr>
<tr>
<td>pVTir&lt;sub&gt;n&lt;/sub&gt;A</td>
<td>pHEA-derivative; VTir&lt;sub&gt;n&lt;/sub&gt; (n clone) fused to C-EhaA (pelB-VTir&lt;sub&gt;n&lt;/sub&gt;-E-tag-EhaA&lt;sub&gt;989-1327&lt;/sub&gt;)</td>
<td>(Salema et al., 2013)</td>
</tr>
<tr>
<td>pVFib&lt;sub&gt;n&lt;/sub&gt;A</td>
<td>pHEA-derivative; VFib&lt;sub&gt;n&lt;/sub&gt; (n clone) fused to C-EhaA (pelB-VFib&lt;sub&gt;n&lt;/sub&gt;-E-tag-EhaA&lt;sub&gt;989-1327&lt;/sub&gt;)</td>
<td>This work</td>
</tr>
<tr>
<td>pNeae</td>
<td>pAK-Not derivative; Neae[Intimin&lt;sub&gt;EHEC&lt;/sub&gt;(1-659)]-E-His-tag</td>
<td>(Bodelon et al., 2009)</td>
</tr>
<tr>
<td>pNeae2</td>
<td>pNeae-derivative; for fusions to Neae-myc [Intimin&lt;sub&gt;EHEC&lt;/sub&gt;(1-659)-E-His-myc tag]</td>
<td>(Salema et al., 2013)</td>
</tr>
<tr>
<td>pNVgfpa</td>
<td>pNeae-myc-derivative; NVgfpa fusion [Intimin&lt;sub&gt;EHEC&lt;/sub&gt;(1-659)-E-Vgfpa-myc tag]</td>
<td>(Salema et al., 2013)</td>
</tr>
<tr>
<td>pNV&lt;sub&gt;Tir&lt;/sub&gt;n&lt;sub&gt;i&lt;/sub&gt;</td>
<td>pNeae-myc-derivative;NV&lt;sub&gt;Tir&lt;/sub&gt;n&lt;sub&gt;i&lt;/sub&gt; (n clone) fusion [Intimin&lt;sub&gt;EHEC&lt;/sub&gt; (1-659)-E-V&lt;sub&gt;Tir&lt;/sub&gt;n&lt;sub&gt;i&lt;/sub&gt;-myc tag]</td>
<td>(Salema et al., 2013)</td>
</tr>
<tr>
<td>pNVFib&lt;sub&gt;n&lt;/sub&gt;</td>
<td>pNeae-myc-derivative; NVFib&lt;sub&gt;n&lt;/sub&gt; (n clone) fusion [Intimin&lt;sub&gt;EHEC&lt;/sub&gt; (1-659)-E-VFib&lt;sub&gt;n&lt;/sub&gt;-myc tag]</td>
<td>This work</td>
</tr>
<tr>
<td>pNV&lt;sub&gt;E&lt;/sub&gt;GFR&lt;sub&gt;n&lt;/sub&gt;</td>
<td>pNeae-myc-derivative; NVEGFR&lt;sub&gt;n&lt;/sub&gt; (n clone) fusion [Intimin&lt;sub&gt;EHEC&lt;/sub&gt; (1-659)-E-VEGFR&lt;sub&gt;n&lt;/sub&gt;-myc tag]</td>
<td>This work</td>
</tr>
<tr>
<td>pCANTAB6&lt;sub&gt;-V&lt;/sub&gt;Tir1</td>
<td>pCANTAB6-derivative; for expression of sdAb VTir1 with His-myc-tags</td>
<td>(Salema et al., 2013)</td>
</tr>
<tr>
<td>pCANTAB6-Vgfpa</td>
<td>pCANTAB6-derivative; for expression of sdAb Vgfpa with His-myc-tags</td>
<td>(Salema et al., 2013)</td>
</tr>
<tr>
<td>pCANTAB6-VFib&lt;sub&gt;n&lt;/sub&gt;</td>
<td>pCANTAB6-derivative; for expression of sdAb VFib&lt;sub&gt;n&lt;/sub&gt; (n clone) with His-myc-tags</td>
<td>This work</td>
</tr>
<tr>
<td>pET28-a</td>
<td>(Km&lt;sup&gt;+&lt;/sup&gt;), pBR322-ori, T7 promoter; for N-terminal His-tag protein fusions</td>
<td>Novagen</td>
</tr>
<tr>
<td>pET28-a-Tir&lt;sub&gt;M&lt;/sub&gt; &lt;sup&gt;EHEC&lt;/sup&gt;</td>
<td>pET28-a derivative; His-tagged TirM&lt;sub&gt;EHEC&lt;/sub&gt; (residues 252 to 360 of Tir&lt;sub&gt;EHEC&lt;/sub&gt;)</td>
<td>(Salema et al., 2013)</td>
</tr>
<tr>
<td>pVDL9.3</td>
<td>(Cm&lt;sup&gt;+&lt;/sup&gt;), pSC101-ori, lac promoter, for production of HlyB and HlyD transporters</td>
<td>(Fernandez et al., 2000)</td>
</tr>
<tr>
<td>pEHlyA2SD</td>
<td>(Ap&lt;sup&gt;+&lt;/sup&gt;), pUC-ori, lac promoter, C-terminal E-tagged HlyA signal</td>
<td>(Fernandez et al., 2000)</td>
</tr>
<tr>
<td>pEHlyA4SD</td>
<td>pEHlyA2SD derivative with modified polylinker having unique Sfi and NotI sites</td>
<td>(Salema et al., 2013)</td>
</tr>
<tr>
<td>pEHlyA4SD-V&lt;sub&gt;Tir&lt;/sub&gt;n</td>
<td>pEHlyA4SD derivative; VTir&lt;sub&gt;n&lt;/sub&gt; (n clone) fused to C-terminal E-tagged HlyA signal</td>
<td>(Salema et al., 2013)</td>
</tr>
</tbody>
</table>
**Materials and Methods**

<table>
<thead>
<tr>
<th>Name</th>
<th>Nucleotide sequence (5’-3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RSP -48</td>
<td>AGCGGATAAACAATTTACACACAGGA</td>
</tr>
<tr>
<td>pAK-directo</td>
<td>GTGACGCGTAGCGGTAACACGGCAGAC</td>
</tr>
<tr>
<td>eae5</td>
<td>GACTTCAGCACTTAATGCCAGTGCGG</td>
</tr>
<tr>
<td>VHH-Sfi2</td>
<td>GTCCTCGAACTCGCCGACCGCCATGGTGCAGCTGGTGG</td>
</tr>
<tr>
<td>VHH-Not2</td>
<td>GGACTAGTGCGGCCGAACTAGGACGACGAC</td>
</tr>
<tr>
<td>VH1-CDR3_up</td>
<td>GAGAGTGAGGGATGGGC3TG</td>
</tr>
<tr>
<td>VH2-CDR3_up</td>
<td>GTAGTTATACTTCCCATATCGTAC</td>
</tr>
<tr>
<td>VH3-CDR3_up</td>
<td>GACTTCAGCACTTAATGCCAGTGCGG</td>
</tr>
<tr>
<td>VH4-CDR3_up</td>
<td>GCCAGCGGTCGTCAGACTGTCGATG</td>
</tr>
<tr>
<td>pMAL-up</td>
<td>TTGAAAACGACCCGAGTGCAAGGAC</td>
</tr>
<tr>
<td>EcoRI-PreScission-SfiI</td>
<td>CCGGAATTCCTGGAAGTTCTGTTCCAGGCCGAGGTGCGGAGGTGCGGAGGTGCGGAGGTGCGGAGGTGCGGAGGTGCGGAGGTGCGGAGGTGCGGAGGTGCGGAGGTGCGGAGGTGCGGAGGTGCGGAGGTGCGGAGGTGCGGAGGTGCGGAGGTGCGGAGGTGCGGAGGTGCGGAGGTGCGGAGGTGCGGAGGTGCGGAGGTGCGGAGGTGCGGAGGTGCGGAGGTGCGGAGGTGCGGAGGTGCGGAGGTGCGGAGGTGCGGAGGTGCGGAGGTGCGGAGGTGCGGAGGTGCGGAGGTGCGGAGGTGCGGAGGTGCGGAGGTGCGGAGGTGCGGAGGTGCGGAGGTGCGGAGGTGCGGAGGTGCGGAGGTGCGGAGGTGCGGAGGTGCGGAGGTGCGGAGGTGCGGAGGTGCGGAGGTGCGGAGGT</td>
</tr>
<tr>
<td>HindIII-stop-myc</td>
<td>CGGGCAGAGAAGCTTTTACTATGCGGCCCATTCAGATC</td>
</tr>
</tbody>
</table>

**Table 4. List of oligonucleotides.**

### 5. Protein electrophoresis and Western blot

Whole cell protein extracts were prepared by harvesting bacteria after induction (1 ml of OD<sub>600</sub> 1.5), resuspended in 50 µl of 10 mM Tris HCl pH 8.0, mixed with the
Materials and Methods

same volume of SDS-sample buffer (2X) or urea-SDS sample buffer (2X) and boiled for 10 min (pHEA constructs) or 30 min (pNeae constructs). The SDS-sample buffer (1X) consists of 60 μM Tris-HCl pH 6.8, 1% w/v SDS, 5% (v/v) glycerol, 0.005% (w/v) bromophenol blue and 1% (v/v) 2-mercaptoethanol (2-ME). The urea-SDS-sample buffer (1X) contains 60 μM Tris-HCl pH 6.8, 2% w/v SDS, 4 M urea, 5 mM EDTA, 5% (v/v) glycerol, 0.005% (w/v) bromophenol blue and 1% (v/v) 2-mercaptoethanol (2-ME). The boiled samples were sonicated (5 sec; Labsonic B Braun), centrifuged (14,000 xg, 5 min) to pellet insoluble material, loaded onto 8% or 10% SDS-PAGE gels and run using a Miniprotean III electrophoresis system (Bio-Rad). For Western blot, the gels were transferred to a polyvinylidene difluoride membrane (PVDF, Immobilon-P, Millipore) using a semi-dry electrophoresis transfer apparatus (Bio-Rad) and the membranes were blocked in phosphate buffered saline (PBS; 8 mM Na₂HPO₄, 1.5 mM KH₂PO₄, 3 mM KCl, 137 mM NaCl, pH 7.0) with 3% (w/v) skimmed milk (Milk-PBS) for 1 hour at RT. For immunodetection of the E-tagged proteins, membranes were incubated for 1 h at room temperature (RT) in the same buffer with anti-E-tag mAb (Phadia), while for immunodetection of the myc-tagged proteins anti-c-myc-POD mAb (clone 9E10; Roche) was used. Membranes were washed three times with PBS containing 0.1% Tween 20 to remove unbound Abs. Bound anti-E-tag mAb was developed using anti-mouse IgG conjugated with peroxidase (POD) (Sigma). Streptavidin-POD conjugate (Roche) was employed to detect the biotinylated broad range SDS-PAGE protein markers (Bio-Rad). All mAbs and POD conjugates were used in a 1:5000 dilution. For developing, a chemiluminescence reaction was prepared using a mixture of 1.25 mM luminol (Sigma) and 200 μM p-coumaric acid (Sigma) in 10 ml of 100 mM Tris–HCl (pH 8.0). Following a rapid rinse in PBS, the membranes were soaked in the chemiluminiscence mixture and H₂O₂ (Sigma) was added at 0.02% (v/v), followed by one minute incubation in the dark and the PVDF-membranes were either exposed to an X-ray film (Curix, Agfa) or scanned in a Chemi-Doc XRS (Bio-Rad). To quantify the total number of VH₁H₁ fusions expressed in E. coli, Western blots of whole cell protein extracts and dilutions of a purified E-tagged VH₁H₁ of known concentration (hereafter referred to as “unknowns” and “standard”, respectively) were visualized on a ChemiDoc XRS and analyzed using the Quantity One software (Bio-Rad). The
intensity of bands from standard and unknown samples was measured and the local background was corrected. The total number of molecules in the various dilutions of the standard protein was calculated and was plotted against the corresponding values of density of its band (Intensity/mm²) to generate a standard curve. Based upon the standard curve plot, the number of molecules of V_{HH} fusions in the unknown samples was estimated, assuming that \( \sim 1.5 \times 10^8 \) bacteria (0.15 units of OD_{600}) were loaded per lane. Two independent experiments were done with each sample in duplicates.

6. **Protease accessibility assays**

Induced bacteria (1 ml, OD_{600}=1.5) were harvested by centrifugation (4000 xg, 3 min) and resuspended in 100 ml of 10 mM Tris HCl pH 8.0. This bacterial suspension was incubated with trypsin (10 mg/ml; Sigma) or with proteinase K (ProtK; 40 mg/ml; Roche) as indicated, for 20 min at 37 °C. Next, the trypsin inhibitor (5 mg/ml; Sigma) or the serine proteases inhibitor (PMSF 1 mM; Sigma) was added to stop further proteolysis. The cell suspension was centrifuged (14,000 xg, 1 min), the cell pellet resuspended in 50 ml of 10 mM Tris HCl pH 8.0, lysed with one volume of SDS-sample buffer (2X) or urea-SDS-sample buffer (2X), boiled and analyzed by Western blot.

7. **ELISA**

TirM_{EHEC} or BSA (Sigma) proteins were adsorbed at 4°C o/n onto 96-well immunoplates (Maxisorb; Nunc) at a concentration of 5 \( \mu \)g/ml in PBS. Next, immunoplates were washed in PBS and blocked by incubation with 200 \( \mu \)l of 3% (w/v) Milk-PBS for 2 h at RT. The sdAbs (secreted or purified) were diluted in 3% (w/v) Milk-PBS, added at the indicated concentrations (0.1-100 nM) in duplicates and incubated for 1 h at RT. After incubation, the wells were washed three times with PBS (Immunowash 1575, Bio-Rad) and the bound sdAbs was detected by the addition of anti-c-myc-POD mAb (clone 9E10; Roche; 1:1000), or anti-E-tag mAb (Phadia; 1:1000) followed by anti-mouse-POD (Sigma; 1:1000) for E-tagged sdAb, and incubation of the plates for 1 h at RT. The plates were washed three times with PBS and developed with \( \mathrm{H}_2\mathrm{O}_2 \) and o-phenylenediamine (OPD; Sigma) as previously
described (Jurado et al., 2002). The plates were read at 490 nm using the iMark ELISA plate reader (Bio-Rad).

8. **Purification of TirM<sub>EHEC</sub>**

500 ml cultures of induced *E. coli* BL21(DE3) cells carrying pET28a-TirM<sub>EHEC</sub> were centrifuged (4000 xg, 15 min, 4°C) and the pellet was resuspended in 25 ml of PN3 buffer (50 mM sodium phosphate pH 7.4, 300 mM NaCl) containing DNase (0.1 mg/ml; Roche) and a cocktail of protease inhibitors (Complete EDTA-free; Roche). The cells were lysed in a French Press at 1200 psi (3 cycles) and the whole cell lysate was ultracentrifuged (40000 xg, 1 h, 4°C). The supernatant was loaded (flow rate of 1 ml/min) onto a column packed with 3 ml of a Cobalt-containing resin (Talon, Clontech) pre-equilibrated with 30 ml PN3 buffer. The column was subsequently washed with 10 ml of PN3 buffer, followed by a wash with 10 ml of PN3 buffer containing 5 mM imidazole and the protein was finally eluted in 1 ml fractions with the PN3 buffer containing 150 mM imidazole. The eluted fractions were dialyzed against HEPES-buffer (20 mM HEPES pH 7.4, 200 mM NaCl, sterile filtered and degassed). Dialyzed fractions were concentrated 10 fold in a 3-kDa centrifugal filter unit (Amicon Ultra-15) and loaded onto a gel filtration column (HiLoad 16/600 Superdex 75 preparative grade, GE Healthcare), pre-equilibrated with HEPES-buffer and calibrated with protein markers (Gel Filtration Standards, Bio-Rad) and Blue dextran (for exclusion volume V<sub>0</sub>; Sigma). Fractions of 1 ml containing TirM<sub>EHEC</sub> were collected and checked for purity by SDS-PAGE. Protein concentration was estimated using the Bicinchoninic acid (BCA) Pierce protein assay kit (Thermo Scientific).

9. **Magnetic Cell Sorting (MACS) of V<sub>H</sub>H immune libraries**

Induced *E. coli* cells (equivalent to a final OD<sub>600</sub> of 5.0) were harvested by centrifugation (4000 xg, 3 min), washed three times with 2 ml PBS (sterile filtered and degassed), and resuspended in a final volume of 1 ml of PBS. Biotinylated antigen (TirM<sub>EHEC</sub>, hFib or rhEGFR, at concentrations indicated) was added to 100 µl of bacteria, the final volume was adjusted to 200 µl with PBS-BSA (PBS supplemented with 0.5% w/v BSA, sterile filtered and degassed), and incubation was carried out for 1 h at RT. After incubation, bacteria were washed three times
with 1 ml of PBS-BSA, resuspended in 100 ml of the same buffer containing 20 µl of anti-biotin paramagnetic beads (Miltenyi Biotec) and incubated at 4 ºC for 20 min. Next, bacteria were washed three times with 1 ml of PBS-BSA, resuspended in 500 µl of the same buffer, of which 10 µl was kept aside to calculate the input bacteria before the procedure, while the rest (490 µl) was applied onto a MACS MS column (Miltenyi Biotec), previously equilibrated with 500 µl of PBS-BSA and placed on the OctoMACS Separator (Miltenyi Biotec). The flow through of unbound cells was collected and the column was washed three times with 500 µl of PBS-BSA. The wash was combined with the flow-through as “Unbound fraction”. Next, the column was removed from the OctoMACS Separator and placed onto a new collection tube, 2 ml of LB was added and the cells were eluted out. This fraction was labeled as the “Bound fraction”. Serial dilutions of Unbound and Bound fractions were plated to determine CFU and to harvest the bound bacteria.

10. Flow cytometry of VHH clones or VHH immune libraries

For standard flow cytometry, induced bacterial cells (equivalent to a final OD$_{600}$ of 1.0; ~10$^9$ CFU) were harvested by centrifugation (4,000 xg, 3 min), washed twice with 500 µl of PBS (filter-sterilized) and resuspended in a final volume of 400 µl of PBS. Next, 190 µl of this cell suspension (~3x10$^8$ CFU) was incubated with the primary Ab or antigen (as indicated) and PBS was added to adjust the total volume to 200 µl. The primary Abs (for assay of expression levels) were anti-E-tag mAb (1:200; Phadia) or anti-c-myc mAb (1:200; 9B11 clone; Cell Signalling), while biotinylated antigens (GFP, TirM$_{EHEC}$, hFib, rhEGFR or BSA) were used at 50 nM for assay of antigen binding, unless otherwise indicated. The samples were incubated at RT for 1h. After incubation, the cells were washed once with 500 µl of PBS, and resuspended either in 500 µl of PBS containing 1 µl of anti-mouse-IgG1 conjugated to Alexa 488 Fluor (2 mg/ml, Invitrogen) or in 200 µl of PBS containing 30 µl of 1:200 dilution of Streptavidin-phycocerythrin (PE) (0.5 mg/ml, Beckman Coulter). The mixture was incubated 30 min at 4 ºC in the dark. The cells were washed once with 500 µl of PBS and resuspended in a final volume of 1 ml in PBS. For each experiment at least 100,000 cells were analyzed in a cytometer (Gallios, Beckman Coulter).
11. **Affinity Determination by Flow cytometry**

Induced *E. coli* cells (equivalent to final OD$_{600}$ of 1) were centrifuged (4000 xg, 3 min), washed twice with 1 ml of PBS (filter-sterilized) and resuspended in a final volume of 1 ml of PBS. Next, 50 µl of this cell suspension (~3x10$^7$ CFU) was incubated at room temperature for 90 min with a fixed amount of biotinylated TirM$_{EHEC}$ and hFib (2 pmols) or rhEGFR (1 pmol) and increasing volumes of PBS (from 0.1 to 1.5 ml) to attain a final concentration range between 20 nM to 1 nM or 100 nM to 0.8 nM respectively. After incubation, cells were centrifuged (4000 xg, 3 min), washed twice with 1 ml of PBS (filter-sterilized) and labeled with Streptavidin-PE as described for standard flow cytometry. After a final washing step with PBS, the mean fluorescence intensity (MFI) of Phycoerythrin (PE) was quantified in a cytometer (Gallios, Beckman Coulter). Data of MFI (relative values to maximum MFI) obtained from the cytometer were plotted against the concentration of biotinylated antigen to obtain the dissociation constant (K$_D$). Curve was fitted according to nonlinear least squares regression method and one site-specific binding saturation kinetics model using the data analysis tool in Prism software (GraphPad).

12. **Purification of VHHs from the periplasm of *E. coli***

Soluble VHHs with His$_6$ and myc tags in their C-termini were induced in *E. coli* WK6 cells carrying pCANTAB6-V$_{TIR1}$ or pCANTAB6-Vgfp. Cells were pelleted by centrifugation (4000 xg, 12 min, 4°C) from 1 L cultures, resuspended in 22.5 ml Periplasmic Extraction buffer [50 mM Sodium phosphate pH 7.4, 200 mM NaCl, 5 mM EDTA and 1 mg/ml polymyxin B sulphate (Sigma)] and stirred at 4°C for 2 h using a magnetic stirrer. The periplasmic extract was obtained by ultracentrifugation (40000 xg, 30 min, 4°C) and dialyzed o/n at 4°C against 5 L of PN2 buffer (50 mM sodium phosphate pH 7.4, 200 mM NaCl). Dialyzed extract was loaded onto a Cobalt-containing affinity resin (Talon, Clontech), washed, and bound protein eluted in PN2 with 150 mM imidazole. Eluted sdAb was dialyzed, concentrated, and loaded onto a calibrated gel filtration column (HiLoad 16/600 Superdex 75 preparative grade, GE Healthcare) as described previously for TirM$_{EHEC}$. The fractions corresponding to the monomeric sdAb were collected and concentrated in a 3-kDa centrifugal filter unit (Amicon Ultra-15). Protein concentration was
estimated using the Bicinchoninic acid (BCA) Pierce protein assay kit (Thermo Scientific).

13. Purification of VHHs from the periplasm as fusions with Maltose binding protein (MBP)

Frozen pellets of *E. coli* cells (ca. 3 grams of wet weight), harvested from induced 1 L cultures, were thawed in 25 ml of buffer NaPi (50 mM sodium phosphate pH 7.4, 200 mM NaCl) containing DNaseI (0.1 mg/ml, Roche) and a cocktail of protease inhibitors (Complete EDTA-free, Roche). Bacteria were lysed using the French Press at 1200 psi (3 cycles). All of the following steps were carried out at 4°C. The cell lysate was clarified by a low-speed centrifugation (5000 xg, 15 min), to pellet non-lysed bacteria and cell debris, followed by a high-speed centrifugation (20000 xg, 60 min) using the supernatant obtained from the previous step. The final supernatant was filtered through a 0.22 µm syringe filter (PVDF, Durapore, Merck Millipore) and loaded (at a flow rate of 1 ml/min) onto a chromatography column packed with 2 ml of amylose resin (New England Biolabs) pre-equilibrated with 20 ml (10 column volumes) of buffer NaPi. All steps were performed in a BioLogic LP chromatography system (Bio-Rad). The column was washed with 20 ml of buffer NaPi, followed by elution of the MBP-fusion protein with 20 ml of buffer NaPi containing 10 mM D-maltose (Sigma), in 1-ml fractions. The collected fractions were analyzed for purity by SDS-PAGE and total protein content by the Bicinchoninic assay (BCA, Pierce) with BSA as protein standard. Selected fractions of high protein content and purity were pooled together and dialyzed o/n against 5 L of buffer NaPi. Dialyzed protein samples were loaded at a flow rate of 1 ml/min onto a chromatography column packed with 2 ml IMAC resin (Talon, Clontech), pre-equilibrated with 20 ml of buffer NaPi. The column was washed with 10 ml of buffer NaPi followed by 10 ml of buffer NaPi containing 10 mM imidazole (Sigma). MBP-fusions were eluted with 20 ml of buffer NaPi containing 150 mM imidazole in 1-ml fractions. The eluted protein fractions were analyzed for protein content by BCA assay and purity by SDS-PAGE analysis. Selected fractions of high purity and protein content were concentrated 10-fold and the buffer was exchanged to NaPi in a 3-kDa centrifugal filter unit (Amicon Ultra-15, Merck Millipore) to remove imidazole.
In cases when V$_{HH}$ free from the MBP was desired, enzymatic cleavage of the MBP-V$_{HH}$His$_6$ fusions was carried out using dialyzed protein samples (from above). Fifty units of PreScission protease® (GE Healthcare) were added to 5 mg samples of MBP- V$_{HH}$His$_6$ fusions in buffer NaPi and incubated at 4 °C for 16 h on a rolling wheel. Digested samples were loaded onto a chromatography column packed with 2 ml amylose resin (New England Biolabs) in order to remove free MBP and undigested MBP- V$_{HH}$His$_6$ fusions. The V$_{HH}$s with C-terminal His$_6$- and myc- tags were collected in the flow-through fraction and loaded directly onto a chromatography column packed with 2 ml IMAC resin (Talon, Clontech), as described above with MBP-fusions. The column was washed and V$_{HH}$s were eluted as previously described with MBP-fusions. All steps were performed in a BioLogic LP chromatography system (Bio-Rad) at 4 °C. The eluted protein fractions were analyzed for protein content by BCA assay and purity by SDS-PAGE analysis. Selected fractions of high purity and protein content were concentrated 10-fold and the buffer was exchanged to 20 mM HEPES pH 7.4, 200 mM NaCl in a 3-kDa centrifugal filter unit (Amicon Ultra-15, Merck Millipore). The concentrated protein was loaded at a flow rate of 1 ml/min onto a Gel-filtration column (HiLoad 16/600 Superdex 75 pg, GE Healthcare Life Sciences) pre-equilibrated with 360 ml of 20 mM HEPES pH 7.4, 200 mM NaCl and pre-calibrated with Gel-filtration protein standards (Bio-Rad) and Blue Dextran (Sigma). The fractions corresponding to the size of monomeric V$_{HH}$s were collected, concentrated 10-fold in a 3-kDa centrifugal filter unit (Amicon Ultra-15, Merck Millipore) and analyzed for purity by Western blot and SDS-PAGE. Total protein concentration was determined by the microBCA protein assay (Piercenet) with BSA as the protein standard.

14. **Surface plasmon resonance (SPR)**

SPR measurements were performed using a Biacore 3000 instrument (GE Healthcare). All proteins solutions were dialyzed against HEPES-buffer [20 mM HEPES 200 mM NaCl (pH 7.4) sterile filtered and degassed] at 4°C o/n. Biotinylated TirM$_{EHEC}$ (0.1 µg/ml) or Human fibrinogen (60 µg/ml) was immobilized on a Streptavidin SA chip (GE Healthcare) at 150 response units (RU) or 18000 response units (RU) at a flow rate of 10 µl/min in HEPES-buffer containing 0.005% (v/v) of the
Materials and Methods

65

surfactant Polysorbate 20 (P20, GE Healthcare). For determination of binding kinetics, dilutions of purified sdAb (analyte) from 32 nM to 200 pM (in case of V_TIR1), from 40 nM to 2.5 nM (in case of V_FIB1) or from 80 nM to 5 nM (in case of V_FIB2) were flown at 30 µl/min in HEPES-buffer and sensograms were generated. The biotinylated antigen surface on the Streptavidin SA chip was regenerated after every cycle using three injections (10 µl) of 10 mM Glycine-HCl (pH 1.7) or 10 mM Glycine-HCl (pH 2.5) in case of TirM_EHEC and Human fibrinogen respectively. Sensograms with the different concentrations of analyte were overlaid, aligned and analyzed with BIAevaluation 4.1 software (GE Healthcare). All data were processed using a double-referencing method.

15. Selection of V_HH immune libraries on cells

Mouse fibroblast cell lines i.e. NIH-3T3 2.2 (EGFR⁺) and HER14 (EGFR⁺), growing in monolayers in 6-well culture plates (BD Falcon) containing culture media and at a confluency of ~40-50% i.e. ~6X10^5 cells were washed with Hank’s balanced salt solution (HBS, Sigma) solution. Induced E. coli cells (equivalent to a final OD_{600} of 1) were harvested by centrifugation (4000 xg, 3 min), washed with 2 ml HBS. 300 µl of washed bacteria (containing ~6X10^7 cells) were added to wells with NIH-3T3 2.2 cells and incubation was carried out for 1 h at 37°C. After incubation, the unbound bacteria were recovered, added to wells containing HER14 cells and further incubation was carried out for 15 min at 37°C. Next, HER14 cells were washed three times with 1 ml of HBS, to remove any non-specific binders and lysed by addition of HBS supplemented with 0.2% SDS and 0.1% DNAse. Serial dilutions of the cell lysate containing bacteria were plated to determine the CFU of bacteria recovered after the procedure. All experiments were carried out in duplicates.

16. Immunofluorescence microscopy (IFM)

~5X10^4 cells of the mouse fibroblast cell lines viz. NIH-3T3 2.2 and HER14 (in DMEM supplemented with FBS and Glutamine), were seeded onto circular coverslips (previously UV irradiated for 30 mins, placed in 24 well tissue culture treated plates), and grown in an incubator at 37°C and 5% CO₂ for 36 hours. Induced E. coli cells (1 OD_{600}) were washed in HBSS and used to infect the cells at a Multiplicity of Infection (MOI) of 100:1 for 20 mins at 37°C. The coverslips were
washed by immersion in a beaker containing PBS (1X). Washed cells were fixed with 4% paraformaldehyde (PFA, Sigma) for 20 mins and subsequently washed ten times in PBS to remove excess of PFA. The coverslips were incubated at room temperature with the primary Ab for 45 mins (as described below), washed ten times in PBS and incubated with the secondary Ab (as described below) for 30 mins. Upon incubation, the coverslips were washed ten times in PBS, mounted on slides using ProLong Gold anti-fade reagent (Life technologies, Ref: P36930) and visualized using a Zeiss Axioimager immunofluorescence microscope.

Different primary and secondary Abs, diluted in PBS and 10% goat serum, were used to stain the eukaryotic cell and the bacteria respectively. EGFR present on the surface of eukaryotic cells was stained with an anti-EGFR mAb (EMD Millipore, Ref: GR01, dilution 1:750), while *E. coli* cells were stained with a rabbit anti-*E. coli* all antigens polyclonal Ab (Amsbio, Ref: B65001R, dilution 1:1500). In case of secondary Abs used, a rabbit anti-mouse mAb conjugated to the fluorophore Alexa 488 (Life technologies, Ref: A-11059, dilution 1:500) stained the EGFR in green, a goat anti-rabbit Ab conjugated to the fluorophore Alexa 594 (Life technologies, Ref: A-11012, dilution 1:500) stained the bacteria in red and DAPI (4',6-diamidino-2-phenilindole, Life technologies, Ref: D3571) stained the DNA in the microscopy preparations in blue.
RESULTS
Chapter 1: Display of single domain Ab clones and immune libraries against \( \text{TirM}_{\text{EHEC}} \) and Human fibrinogen on the surface of \( E. \ coli \) cells

1.1 Comparison of EhaA and Intimin \( \beta \)-domains for display of sdAbs on \( E. \ coli \)

To compare the potential of Intimin and EhaA \( \beta \)-domains for the display of sdAbs we employed camelid \( \text{V}_{\text{HH}} \) (nanobodies) (Muyldermans et al., 2009). Initial assays were conducted using a model \( \text{V}_{\text{HH}} \) clone binding GFP (Vgfp) (Rothbauer et al., 2008) fused to the \( \beta \)-domains of EhaA and Intimin for comparison of the two systems before cloning of an immune library of sdAbs. Vgfp was cloned in pHEA vector (Table 2) in frame with the N-terminal SP of PelB (Keen and Tamaki, 1986) and the C-terminal fragment of EhaA (residues 989-1327; named as C-EhaA), bearing its native \( \beta \)-barrel with \( \alpha \)-helix linker, and including the E-tag epitope between the \( \text{V}_{\text{HH}} \) and C-EhaA (Figure 1A) (Marin et al., 2010). The Vgfp sequence was cloned in pNeae2 (Table 2) in frame with the N-terminal fragment of Intimin (residues 1-659; named as Neae), comprising its N-terminal SP, periplasmic LysM domain (expected to bind the peptidoglycan), native \( \beta \)-barrel with C-terminal linker, the first Ig-like domain (D0); and in addition, the E-tag and myc-tag (EQKLISEEDL) epitopes flanking the \( \text{V}_{\text{HH}} \) for detection of expression and binding (Figure 1B). Both \( E. \ coli \) display vectors contain unique \( \text{SfiI} \) and \( \text{NotI} \) restriction sites flanking the \( \text{V}_{\text{HH}} \) in the same frame as those of conventional phagemids (e.g. pHEN6, pCANTAB6) (McCafferty et al., 1994). The resulting fusion proteins were referred to as \( \text{V}_{\text{HH}} \text{A} \) (fusions to C-EhaA) and \( \text{NV}_{\text{HH}} \) (fusions to Neae).

The expression of VgfpA and NVgfp fusions in \( E. \ coli \) K-12 cells (strain UT5600; Table 1) was analyzed by Western blot after induction with 0.05 mM IPTG at 30 °C for 3 h (see Materials and Methods). Discrete protein bands corresponding to VgfpA and NVgfp were detected with anti-E or anti-myc mAbs in whole cell protein extracts from the induced cells (Figure 10C and Figure 10D). Both fusion proteins, i.e. VgfpA and NVgfp, showed a shift in their electrophoretic mobility (Figure 10C
and Figure 10D), a characteristic of native OMPs with correctly folded β-barrels, which makes them resistant to SDS denaturation at low temperatures and hence migrate faster than the unfolded polypeptides due to the compact structure of the β-barrel (Schnaitman, 1973; Koebnik et al., 2000).

Interestingly, NVgfp was resistant to 2% SDS and 4 M urea at low temperatures (i.e. 22 °C) and required boiling in this buffer to unfold, as previously reported for full-length Intimin and its β-domain (Bodelon et al., 2009). The major protein bands detected with anti-E mAb in the boiled samples corresponded to full-length VgfpA and NVgfp fusions (Figure 10C and 10D; labelled with arrows). Detection of NVgfp with anti-myc mAb confirmed the integrity of its C-terminal end (Figure 10D). Minor bands of lower MW were also detected in Western blot with anti-E mAb, which likely represent proteolytic fragments of the full-length fusions (Figure 10C and 10D; labelled with asterisks).

The accessibility of VgfpA and NVgfp fusions to the external milieu was initially compared by incubation of intact E. coli cells with externally added proteases. Trypsin digested full-length VgfpA leaving some weakly detectable proteolytic fragments with a size similar to C-EhaA and Vgfp domains (Figure 10C, lane 3). The NVgfp fusion was resistant to Trypsin (data not shown) and sensitive to Proteinase K (ProtK) (Figure 10D, lane 3) but ProtK digestion of NVgfp fusion left a resistant fragment comprising Neae.

The induced E. coli cultures expressing VgfpA or NVgfp fusions showed only a slight decrease in their growth rate compared to control cultures having the empty vector (pAK-Not), or expressing the β-domains C-EhaA (pHEA) or Neae (pNeae2), and reached final optical densities at 600 nm (OD<sub>600</sub>) identical to controls (Figure 11A and 11B). Hence, both the β-domains display the sdAb to the extracellular milieu, in a way that is accessible to externally added proteases but C-EhaA fusions are more sensitive to digestion than the Neae fusions. Resistance to proteolysis was also previously observed for full-length Intimin (Bodelon et al., 2009).
Results

Figure 10. *E. coli* cell surface display of VHHs with EhaA and Intimin beta domains.

(A) Scheme of EhaA autotransporter and VHH fusions (left), showing N-terminal SP, secreted passenger or VHH domain, and C-terminal β-domain. Model of VHH fusion in the OM (right), with N-terminal VHH domain exposed to the extracellular milieu and with C-EhaA β-barrel inserted in the OM. These domains are connected with the E-tag epitope and the internal α-helical linker of the β-barrel. (B) Scheme of Intimin and NVHH fusions (left), showing N-terminal SP, LysM and β-domains, and secreted D0-D3 Ig-like and lectin-like domains, or VHH domain replacing D1-D3 in NVHH fusions. Model of NVHH fusion in the OM (right), with N-terminal LysM domain in the periplasm, β-barrel with linker in the OM, and connecting with C-terminal D0 and VHH domains exposed to the extracellular milieu. The E-tag and myc-tag epitopes flanking the VHH domain are indicated. (C) and (D) Western blots of whole-cell protein extracts from induced *E. coli* UT5600 harbouring pVgfpA (C) or pNVgfp (D). Intact *E. coli* cells were incubated with (+) or without (-) the indicated protease, Trypsin or Proteinase-K (ProtK), before lysis. Protein extracts were prepared in SDS (C) or SDS-urea (D) sample buffers and boiled (+) or not boiled (-) before SDS-PAGE. Western blots were developed with anti-E or anti-myc mAb, as indicated. The positions of full-length VgfpA and NVgfp fusions are labeled with arrows. Asterisks indicate protein bands detected in protease-treated samples. The mass of protein markers (in kDa) is shown on the left.
Figure 11. Growth of *E. coli* cultures expressing VgfpA and NVgfp fusions. 

A) Growth curve of LB cultures of *E. coli* UT5600 cells carrying plasmids pVgfpA, pHEA (expressing C-EhaA), or pAK-Not (empty vector). (B) Growth curve of LB cultures of *E. coli* UT5600 cells carrying plasmids pNVgfp, pNeae2 (expressing Neae), or pAK-Not (empty vector). The cultures were incubated at 30 °C with agitation (160 rpm) and induced with 0.05 mM IPTG at the time indicated by an arrow. The optical density at 600 nm (OD<sub>600</sub>) of the cultures was monitored at the time points shown.

Surface display of VgfpA and NVgfp was assessed by flow cytometry (Figure 12). Induced *E. coli* cells harboring pVgfpA, pNVgfp, or pAK-Not (control) were stained with anti-E or anti-myc mAbs followed by anti-mouse IgG-Alexa488 (Figure 12, left panels). *E. coli* cells expressing VgfpA or NVgfp were positively bound by anti-E mAb, though cells expressing NVgfp were also positively bound with the anti-myc mAb. Control *E. coli* cells with pAK-Not were negative for both mAbs. Importantly, the presence of a single peak in the flow cytometry histograms indicated that most *E. coli* cells were expressing a homogenous level of the fusion proteins. The mean fluorescence intensity (MFI) of cells with anti-E-tag mAb suggested a higher expression and display level of NVgfp than VgfpA (~3-fold). The antigen-binding activity of the surface displayed Vgfp was compared by flow cytometry after incubation of *E. coli* cells expressing these fusions with 50 nM biotin-labeled GFP (positive antigen) or biotin-labeled BSA (negative antigen), followed by incubation with Streptavidin-Phycocerythrin (PE) conjugate (Streptavidin-PE) (Figure 12, right panels). This analysis showed the specific binding of the *E. coli* cells expressing VgfpA and NVgfp fusions to GFP, whereas control *E. coli* cells did not bind GFP. No significant binding to BSA was observed. Hence, the β-domains of EhaA and Intimin allow the functional display of sdAb on the surface of *E. coli* cells. However, the MFI of GFP binding was clearly higher in *E. coli* cells expressing NVgfp than in those with VgfpA (ca. 8-fold) (Figure 12, right panels). The higher expression level of NVgfp...
Results

does not appear to be sufficient to account for this difference in binding, suggesting that the sdAb may have a higher antigen-binding activity when fused to the β-domain of Intimin.

Figure 12. *E. coli* cell surface display and antigen-binding activity of VgfpA and NVgfp.
Fluorescent flow cytometry analysis of induced *E. coli* UT5600 cells bearing the indicated plasmids: pAK-Not (control), pVgfpA, and pNVgfp. Histograms show the fluorescence intensity of bacteria stained with anti-E or anti-myc mAbs (as indicated) and secondary anti-mouse IgG-Alexa 488 (left panels) or incubated with biotinylated antigens (GFP or BSA, as labeled) and secondary Streptavidin-phycoerythrin (PE) (right panels).

1.2 Construction of immune libraries against TirMEHEC and Fib, analysis of expression, display levels and cellular toxicity

Two distinct immune libraries of V_{HH} genes from dromedaries immunized with purified antigens were used to demonstrate the effectiveness of *E. coli* Intimin and EhaA display systems with large sdAb repertoires. An immune library of V_{HH} was generated against the soluble extracellular fragment of the translocated intimin receptor (tir) from EHEC (named TirMEHEC, corresponding to residues 252-360 of
full-length TirEHEC) (Frankel and Phillips, 2008) was cloned in vectors pHEA and pNeae2. Similarly, another immune library against Human Fibrinogen (Fib) was also cloned in vectors pHEA and pNeae2. Both VHH libraries were obtained by immunization of a dromedary with purified antigens, i.e. TirMEHEC-His6 and Fib, and subsequent amplification of the VHH gene segments from ~2x10^7 lymphocytes isolated from a peripheral blood sample (see Materials and Methods). The amplified VHH gene segments were cloned into the SfiI and NotI sites of pHEA and pNeae2 vectors, generating two E. coli display immune libraries of similar size (~2-3x10^6 clones). The E. coli strain EcM1 (Table 2) was used as host for cell display. This strain is derived from the reference wild type K-12 strain (MG1655) with a deletion in the operon encoding type 1 fimbriae (ΔfimA-H) (Munera et al., 2008). Sequencing of 40 clones picked randomly from each of the two libraries in both display systems confirmed the cloning of different VHH sequences in frame with the β-domains of EhaA and Intimin respectively (data not shown). The expression and display of both libraries with the β-domains of EhaA (VHH A) and Intimin (NVH H) were analyzed by flow cytometry with anti-E and anti-myc mAbs, revealing a fairly homogeneous expression of both libraries in E. coli EcM1 (Figure 13A and 13C).

The MFI with anti-E mAb indicated a similar expression level of VHH A and NVH H libraries, in contrast to the significantly lower expression of VgfpA observed previously. Western blot analysis of whole-cell protein extracts from induced cultures revealed major protein bands with the expected size for full-length VHH A and NVH H fusions, upon boiling in SDS or SDS-urea buffer, respectively, and which have heat-modifiable electrophoretic mobility indicating the correct folding of their β-barrels (Figure 13B and 13D).

Quantification of the Western blot signals with anti-E mAb using ~1.5x10^8 bacteria (0.15 units of OD_600) expressing VHH A or NVH H fusions, was carried out using a standard curve generated with a purified E-tagged VHH of known concentration (Figure 14), and allowed an estimation of ~6.5x10^3 molecules of VHH A and ~7.8x10^3 molecules of NVH H per bacterium.
Figure 13. *E. coli* cell surface display of anti-TirMEHEC and anti-Fib immune libraries.

Fluorescent flow cytometry analysis of induced *E. coli* EcM1 cells expressing either (A) VHH\(\text{A}\) or NV\(\text{HH}\) immune libraries anti-TirMEHEC (as indicated) or (C) VHH\(\text{A}\) or NV\(\text{HH}\) immune libraries anti-Fib (as indicated). Control cells carried the empty vector pAK-Not. Histograms show the fluorescence intensity of bacteria stained with anti-E or anti-myc mAbs (as labeled) and secondary anti-mouse IgG-Alexa 488. Western blots of whole-cell protein extracts from induced *E. coli* EcM1 cells expressing either (B) VHH\(\text{A}\) or NV\(\text{HH}\) immune libraries anti-TirMEHEC (as indicated) or (D) VHH\(\text{A}\) or NV\(\text{HH}\) immune libraries anti-Fib (as indicated). Protein extracts were prepared in SDS (VHH\(\text{A}\) library) or SDS-urea (NV\(\text{HH}\) library) sample buffers and boiled (+) or not boiled (-) before SDS-PAGE. Western blots were developed with anti-E mAb. Positions of full-length fusions are labeled with arrows and mass of protein markers (in kDa) is shown on the left.
The plot shows the intensity of protein bands from Western blots developed with anti-E-tag mAb and quantified on a ChemiDoc XRS using the Quantity One software (Bio-Rad). Samples analyzed were whole-cell protein extracts from ~1.5x10⁸ bacteria (0.15 units of OD₆₀₀) of induced E. coli EcM1 cells carrying the pVHH A or pNVHH anti-TirMEHEC libraries. The standard curve was generated with the values of band intensities (Intensity/mm²) of a purified E-tagged VHH of known concentration. Two independent experiments were done with similar results.

The growth of E. coli cultures expressing VHH A or NVHH libraries against TirMEHEC was only slightly delayed compared to a control with pAK-Not and the cultures reached similar OD₆₀₀ after induction (Figure 15).

Plating of the induced cultures to determine the number of colony forming units (CFU) per OD₆₀₀ gave ~1.0x10⁹ CFU/OD₆₀₀ in the control (pAK-Not), ~0.9 x10⁹ CFU/OD₆₀₀ in the anti-TirMEHEC VHH A library, and ~0.6x10⁹ CFU/OD₆₀₀ and anti-
TirM\textsubscript{EHEC} NVHH library. While for the anti-Fib libraries we found $\sim 0.83 \times 10^9$ CFU/OD\textsubscript{600} in VHHA library and $\sim 0.8 \times 10^9$ CFU/OD\textsubscript{600} in the NVHH library. Hence, growth of bacteria is not overly affected due to expression or transport of the fusions, though expression of some NVHH fusions appear to slightly reduce the viability of \textit{E. coli} cells when compared to control cultures, especially with the anti-TirM\textsubscript{EHEC} library. Nevertheless, since the CFU/OD\textsubscript{600} after expression of NVHH fusions remains within the same order of magnitude as the control, the diversity of the sdAb library is not compromised. Cell toxicity during the expression of Intimin constructs has been reported previously in some \textit{E. coli} K-12 strains (Wentzel et al., 2001).
Chapter 2: Selection and characterization of single domain Abs against TirM_{EHEC} from immune libraries displayed on the surface of *E. coli*

2.1 Translocated Intimin Receptor domain M (TirM)

The translocated intimin receptor (Tir) of enteropathogenic and enterohemorrhagic *Escherichia coli* strains (EPEC and EHEC, respectively) is an important protein for the infection of these enteric pathogens, which can result in severe diarrhea, hemorrhagic colitis and the life-threatening hemolytic uremic syndrome (in the case of EHEC). Tir is translocated through the bacterial type III secretion system (Kenny et al., 1997; Gauthier and Finlay, 2003) into the host cell, where it becomes inserted in the plasma membrane through two hydrophobic helices that leave an extracellular region exposed to the surface of the host cell (Figure 16).

![Figure 16. The Intimin-Tir interaction between EPEC/EHEC bacteria and host cell plasma membrane.](image)

The model is based on structural data of the complex of the C-terminal fragment of intimin (domains D1, D2 and D3) and the extracellular domain of Tir (TirM) (Luo et al., 2000). Intimin is shown in green with its domains labelled. The transmembrane domain folds as a β-barrel (Fairman et al., 2012) and is inserted in the outer membrane of EPEC/EHEC. The Ig-like domains D0, D1, D2, and the lectin-like domain D3, which binds to the TirM, are shown as ovals. Tir is shown as a dimer (in pink and purple) in the host-cell membrane, with extracellular domain (residues 252-360) flanked by the two predicted transmembrane (TM) domains. The N-terminal domain of Tir anchors host cytoskeletal components (such as actin) that are needed to form the characteristic attaching and effacing (A/E) lesion on the host-cell surface upon bacterial adhesion.
This extracellular region of Tir, named TirM, which corresponds to amino acid residues 232-360 for Tir of EHEC, serves as anchor for the bacterial adhesin Intimin, responsible for the intimate attachment of these pathogens to the enterocyte. Intimin is an integral OM protein whose extracellular C-terminal domains are displayed on the surface of EPEC and EHEC bacteria and bind TirM region (Kenny et al., 1997). The Intimin-Tir interaction ultimately leads to F-actin polymerization within the epithelial cell at the sites of bacterial attachment, and the formation of actin pedestals beneath the bacterium (Lai et al., 2013). It is postulated that blocking Intimin-Tir interaction could prevent the stable attachment of the bacteria to the intestinal epithelium and thus, reduce the infection process and the disease.

2.2 Selection of single domain Abs against TirM_{EHEC} by magnetic sorting of E. coli bacteria displaying the immune library

*E. coli* EcM1 cells expressing VHH and NVHH libraries were screened to isolate clones binding to TirM_{EHEC} by magnetic cell sorting (MACS) (Figure 17).

---

**Figure 17.** Schematic illustration of Magnetic cell sorting (MACS).

General scheme summarizing the steps followed during MACS of an *E. coli* display library of sdAb with a biotinylated antigen. *E. coli* cells binding the biotinylated antigen are captured in a MACS column held in a magnet, while *E. coli* cells that do not bind the antigen are washed out of the column. Elution of bound bacteria is done with fresh LB media upon column removal from the magnet. The CFU in the Washed and Bound fractions are determined by plating.
The MACS conditions for capturing *E. coli* cells expressing VHH or NVHH fusions were established using biotin-labeled anti-E mAb. An initial input of 0.1 units of OD$_{600}$ (~6-9x10$^7$ CFU of each culture) was incubated with 50-250 nM of biotinylated anti-E mAb allowing the recovery of a total of ~2-4x10$^7$ CFU in the Wash and Bound fractions. The Bound fractions contained ~95-99% of the CFU in the VHH and NVHH libraries, and only ~0.2-0.5% in control *E. coli* cells with pAK-Not. Using these conditions, the VHH and NVHH libraries were incubated with biotinylated TirM$_{EHEC}$ (250 nM) for the first selection step (MACS1) and ~0.3-0.6% of the total CFU from both libraries were collected in the Bound fractions (Table 5). The colonies grown from the Bound fractions of each library were pooled independently, their plasmids purified and electroporated into fresh *E. coli* EcM1 cells to obtain VHH and NVHH sublibraries (~2x10$^6$ transformants). Next, the VHH and NVHH sublibraries were subjected to a new round of selection with biotinylated TirM$_{EHEC}$ using conditions identical to those used in MACS1. Bacteria harvested from Bound fractions were pooled, their plasmids purified and transformed for the following rounds of MACS. Antigen concentration was reduced to 50 nM in the following MACS. The percentage of *E. coli* bacteria recovered in the Bound fractions showed a significant increase from the initial 0.3-0.6% to over 70% in MACS3 of NVHH and MACS4 of VHH (Table 5), suggesting an enrichment of antigen binding clones in both libraries.

**Table 5. Summary of MACS with *E. coli* display libraries against TirM$_{EHEC}$.**

<table>
<thead>
<tr>
<th>Round</th>
<th>Antigen conc (TirM$_{EHEC}$)</th>
<th>VHH library %</th>
<th>NVHH library %</th>
</tr>
</thead>
<tbody>
<tr>
<td>MACS1</td>
<td>250 nM</td>
<td>0.3</td>
<td>0.6</td>
</tr>
<tr>
<td>MACS2</td>
<td>250 nM</td>
<td>2.6</td>
<td>12.5</td>
</tr>
<tr>
<td>MACS3</td>
<td>50 nM</td>
<td>12.5</td>
<td>75.6</td>
</tr>
<tr>
<td>MACS4</td>
<td>50 nM</td>
<td>70.5</td>
<td>-</td>
</tr>
</tbody>
</table>

(％): Percentage of bacteria in Bound fractions with reference to total bacteria in Wash+Bound fractions (ca. 1-5 x 10$^7$ CFU)

Bacteria from the different rounds of selection were analyzed by flow cytometry to test their binding to biotinylated TirM$_{EHEC}$ (50 nM) (Figure 18), which demonstrated an enrichment of *E. coli* cells binding to TirM$_{EHEC}$ along the selection rounds, from ~0.2 % positives in the original libraries to ~45% after MACS4 of the VHH library and more than 75% positives after MACS2 and MACS3 of the NVHH library.
significant binding to biotinylated BSA was detected by flow cytometry in these populations. The expression levels of the V_HH_A and NV_HH_fusions in the bacterial pools after the MACS steps were similar to those of the original libraries (data not shown).

![Graph showing flow cytometry results for V_HH_A and NV_HH_fusions](image)

**Figure 18.** Magnetic cell sorting of V_HH_A and NV_HH E. coli display libraries with biotinylated antigen, TirMEHEC.

Fluorescent flow cytometry analysis of IPTG-induced E. coli EcM1 cells expressing V_HH_A (top panel) or NV_HH (bottom panel) immune libraries, or their respective sublibraries enriched after the indicated round of MACS with biotinylated TirMEHEC. Histograms show the fluorescence intensity of bacteria incubated with biotinylated TirMEHEC and secondary Streptavidin-PE.

Fifty colonies from the final round of selection of each library were randomly picked for plasmid isolation and DNA sequencing. A V_HH sequence, named as V_TIR1, was found in all NV_HH clones and in 36 V_HH_A clones, while the rest were different V_HH sequences. Flow cytometry analysis confirmed the specific binding of biotinylated TirMEHEC (50 nM) by E. coli cells displaying V_TIR1 fused to EhaA and Intimin β-domains whereas these cells did not bind to biotinylated BSA (Figure 19A and 19B).
Results

Figure 19. Binding of *E. coli* cells displaying selected clones from VHH\textsubscript{A} and NVHH\textsubscript{A} libraries to biotinylated TirM\textsubscript{EHEC}.

Fluorescent flow cytometry analysis of induced *E. coli* EcM1 cells bearing the indicated plasmids selected from (A) the VHH\textsubscript{A} library: pVTIR1\textsubscript{A}, pVTIR2\textsubscript{A}, pVTIR3\textsubscript{A}; and (B) from the NVHH\textsubscript{A} library: pNV\textsubscript{TIR1}, pNV\textsubscript{TIR4}, pNV\textsubscript{TIR5}. Histograms show the fluorescence intensity of bacteria incubated with biotinylated antigens (TirM\textsubscript{EHEC} or BSA, as labeled) and secondary Streptavidin-PE.

The MFI of *E. coli* cells displaying VTIR1 was higher in the NVTIR1 fusion than with VTIR1\textsubscript{A} fusion (Figure 20), although both were expressed at similar levels (Figure 20). We also conducted a flow cytometry screening of the remaining 14 non-VTIR1 clones from MACS4 of VHH\textsubscript{A} library, identifying three other VHH\textsubscript{A} sequences that bound to biotinylated TirM\textsubscript{EHEC}, referred to as VTIR2, VTIR3, VTIR4 (Figure 20A, and data not shown). VTIR4 did not bind biotinylated BSA, but VTIR2 and VTIR3 clones exhibited non-specific binding to biotinylated BSA and were not further analyzed (Figure 19A and 19B).
Results

We sought for additional \textit{V}_{HH} \textit{sequences} binding Tir\textit{MEHEC} in the NV_{HH} library by screening 96 colonies picked randomly after the first round of selection (MACS1), in which a higher diversity of binders with low and high affinities are expected. PCR screening with a specific primer hybridizing the CDR3 of \textit{VTIR1} enabled us to identify 17 \textit{VTIR1} clones out of these 96 colonies. This number fits with the percentage of clones recovered from this population after MACS2 (Table 5). Flow cytometry screening of the remaining clones allowed the identification of two additional \textit{V}_{HH} sequences that specifically bound biotinylated Tir\textit{MEHEC}, \textit{VTIR4} (2 clones), which was found previously in the \textit{V}_{HH}A library, and \textit{VTIR5} (1 clone). The MFI of binding of these clones to 50 nM biotinylated Tir\textit{MEHEC} was low compared to \textit{VTIR1} (Figure 19B) and increased at higher antigen concentrations (200 nM) (data not shown) suggesting that these clones had a lower affinity for Tir\textit{MEHEC}. The CDR3 amino acid sequences of the selected \textit{V}_{HH} clones are shown in Table 6.

Table 6. CDR3 sequences of anti-Tir\textit{MEHEC} clones selected by \textit{E. coli} display.

<table>
<thead>
<tr>
<th>Clone Name</th>
<th>Amino acid sequence of CDR3</th>
<th>(\beta)-domain system</th>
</tr>
</thead>
<tbody>
<tr>
<td>VTIR1</td>
<td>GTAPYWHTPIPTLSEDKYFY</td>
<td>Neae, C-EhaA</td>
</tr>
<tr>
<td>VTIR2</td>
<td>GNSGSRGFDY</td>
<td>C-EhaA</td>
</tr>
<tr>
<td>VTIR3</td>
<td>AKGPRRCNQGFDY</td>
<td>C-EhaA</td>
</tr>
<tr>
<td>VTIR4</td>
<td>PDLSTNCDTVLTNSGALNY</td>
<td>Neae, C-EhaA</td>
</tr>
<tr>
<td>VTIR5</td>
<td>PKYGGTWRWRVEEEKTI</td>
<td>Neae</td>
</tr>
</tbody>
</table>
The \( V_{HH} \)s encoded by \( V_{TIR1} \), \( V_{TIR4} \), \( V_{TIR5} \) and one unrelated \( V_{HH} \) binding \( \alpha \)-amylase as a control, were secreted into \( E. \ coli \) culture media as soluble fragments with the hemolysin system (Fernandez et al., 2000; Fraile et al., 2004) and used in ELISA against \( \text{TirMEHEC} \) and BSA. This experiment showed that soluble \( V_{TIR1} \), \( V_{TIR4} \), and \( V_{TIR5} \) sdAbs, bound specifically to \( \text{TirMEHEC} \) (Figure 21). According to ELISA data, \( V_{TIR1} \) was the clone with an apparent higher affinity, as could be also inferred from its strong enrichment in both the \( E. \ coli \) display libraries.

![ELISA of sdAbs selected by \( E. \ coli \) display against \( \text{TirMEHEC} \).](image)

**Figure 21.** ELISA of sdAbs selected by \( E. \ coli \) display against \( \text{TirMEHEC} \).

ELISA against \( \text{TirMEHEC} \) of sdAbs secreted into culture media as E-tagged HlyA fusions from the indicated \( VTIR \) clones and one a negative control (Vamy) (Fraile et al., 2004). The plot shows the average OD values at 490 nm with standard error from duplicate experimental samples obtained with the secreted sdAbs at the indicated concentrations. ELISAs were developed with anti-E-tag mAb and anti-mouse-POD. ELISA signals against a control antigen (BSA) are subtracted from the represented values.

### 2.3 Characterization of \( VTIR1 \) sdAb and determination of its affinity by Surface Plasmon Resonance (SPR)

The \( VTIR1 \) clone was produced in the periplasm of \( E. \ coli \) WK6 cells as soluble sdAb with C-terminal His- and myc-tags and purified by metal-affinity chromatography followed by gel-filtration chromatography (Materials and Methods). As a control, Vgfp was also expressed and purified in the same manner. Both sdAbs behave as monomers with an apparent mass of ~15 kDa in gel filtration chromatography (Figure 22A). The binding activity of the purified \( VTIR1 \) was confirmed in ELISA (Figure 22B).
Results

Figure 22. Monomeric behaviour and binding activity of the purified sdAb, VTIR1.
(A) Gel-filtration chromatograms of sdAbs VTIR1 and Vgfp purified from the periplasm of *E. coli* WK6 cells (carrying the corresponding pCANTAB6-derivative) after a metal-affinity chromatography step. Gel-filtration chromatography was performed in a HiLoad 16/600 Superdex 75 column calibrated with protein markers (labeled in kDa) and Blue dextran (for exclusion volume Vo). Both sdAbs have major peaks of ~15 kDa corresponding to their monomeric forms. (B) ELISA of purified monomeric VTIR1 and Vgfp (control) against TirM\_EHEC and BSA. The plot represents the OD values at 490 nm obtained with the indicated concentrations of sdAbs. ELISA developed with anti-myc mAb-POD as secondary.

In order to determine the apparent equilibrium dissociation constant (K_D) between VTIR1 and TirM\_EHEC, their interaction was studied in surface plasmon resonance (SPR) experiments with a Streptavidin (SA) sensor chip coated with biotinylated TirM\_EHEC (Materials and Methods). The change in resonance units (RU) was recorded with time at different concentrations of purified VTIR1 from 0.2 to 32 nM showing a clear binding to TirM\_EHEC that reached the steady state equilibrium in ~220 s for the two highest concentrations used, but not for the lower concentrations (Figure 23A). No binding was observed when Vgfp (40 nM) was flown over this sensor surface, or
when V₇₁ (40 nM) was flown over a SA cell lacking biotinylated TirM₃₁ (data not shown). Injection of buffer to evaluate the dissociation of V₇₁ (labeled with an arrow in Figure 23A) showed no loss of RU for > 200 s, indicating that V₇₁ remained stably bound to TirM₃₁ over long periods of time. Because of the very slow dissociation of V₇₁ from TirM₃₁, we were not able to calculate the value of Kᵩ directly from the curves. Hence, kinetic constants were calculated at a time point when the curves reached equilibrium (i.e. 220 s). From the graph, Kᵩ was calculated as 2.2 X 10⁻⁹ M (Figure 23B).

Figure 23. Determination of the equilibrium dissociation constant (Kₒ) of V₇₁ by SPR.

(A) SPR sensograms monitoring real-time association and dissociation of purified sdAb V₇₁ (at the indicated concentrations) to biotinylated TirM₃₁ immobilized onto a Streptavidin-SA sensor chip. The increase in resonance units (RU) is recorded along time (in seconds). Dissociation of V₇₁ is evaluated by injection of buffer at the time indicated with an arrow. (B) RU values at 220 seconds (labeled with a rectangle in A) are plotted versus the different concentrations of V₇₁. The curve was fitted by non-linear least squares regression.

2.4 Estimation of the affinity of V₇₁ by E. coli display

Flow cytometry analysis under equilibrium conditions has been used to estimate the apparent Kᵩ of Abs and anticalins displayed on the surface of yeast and E. coli cells (Daugherty et al., 1998; Boder and Wittrup, 2000; Binder et al., 2010). Thus, we tested whether the affinity of V₇₁ could be estimated by flow cytometry analysis of E. coli cells with this sdAb on their surface and incubated with biotinylated TirM₃₁ under conditions expected to be close to the equilibrium. We chose the Intimin display system given its superior MFI signals in flow cytometry with the antigen. E. coli EcM1 cells displaying NV₇₁ (~3x10⁷ CFU) were incubated for 90 min with a fixed amount of biotinylated TirM₃₁ (2 pmols) in two-fold increasing volumes of PBS (from 0.1 to 2 ml) to reach a final concentration range from 20 nM to 1 nM.
After this incubation, cells were washed and labeled with Streptavidin-PE as previously described.

The relative MFI of the cells was plotted against the antigen concentration used and the curve fitted by non-linear least squares regression, giving an estimated apparent $K_D$ of $1.7 \times 10^{-9}$ M (Figure 24). This value was consistent with the apparent $K_D$ determined by SPR analysis ($K_D$ of $\sim 2.2 \times 10^{-9}$ M, Figure 23B), and indicates that the *E. coli* display systems could also be used to estimate the $K_D$ of selected sdAbs before purification.

![Figure 24. Estimation of the equilibrium dissociation constant (K_D) of V_TIR1 by E. coli display.](image)

The $K_D$ of $V_{TIR1}$ was estimated by flow cytometry analysis of *E. coli* cells expressing $NV_{TIR1}$ incubated with different concentrations of biotinylated TirM_{EHEC} under equilibrium conditions. The mean fluorescent intensities (MFI) of bacteria, after labeling with Streptavidin-PE, were plotted versus the concentration of TirM_{EHEC} used in the assays. The curve was fitted by non-linear least squares regression.
Chapter 3: Selection and Characterization of high affinity single domain Abs against Human Fibrinogen from Immune libraries displayed on the surface of *E. coli* cells

3.1 **Human Fibrinogen** (Fib).

Fibrinogen (Fib, Factor I) is a 340 kDa glycoprotein produced by the liver and circulated in plasma at concentrations ranging between 1.5 and 4.5 mg/ml (Lowe et al., 2004). It plays a key role in the hemostatic system, being involved in the final step of blood coagulation. Abnormal Fib concentration in blood has been reported to be associated with cardiovascular diseases, venous thrombosis or myocardial infarction (Smith et al., 1998). There is also increasing evidence that Fib is a biomarker of oxidative stress (Selmeci et al., 2010) and metabolic syndrome (Onat et al., 2009) in human plasma. Lower protein concentrations indicate the risk of bleeding and maybe related to liver diseases, whereas higher Fib concentrations reveal an important risk for ischemic vascular or coronary accidents (Kamath and Lip, 2003), (Lowe et al., 2004), (Dudek et al., 2010). The structure of fibrinogen is illustrated in Figure 25.

![Structure of Fibrinogen](image)

**Figure 25. Structure of Fibrinogen.**

Fibrinogen is a dimeric glycoprotein composed of three pairs of peptide chains, i.e. α, β and γ chains. The two sub-units are composed of a D domain containing a globular region and an E domain, which contains a disulfide bond which links the two subunits. Fibrinogen is an important acute phase protein that is an important part of coagulation cascade of proteins. Thrombin rapidly hydrolyzes fibrinogen into fibrin and releases fibrinopeptides A and B. The fibrin is further stabilized by Factor Xlla, leading to aggregation of fibrin and blood platelets that block the damaged blood vessel preventing further bleeding. Plasmin is involved in Fibrin degradation into Fragment D and E.
Fib is also an important determinant of the metastatic potential of circulating tumor cells (Palumbo et al., 2000) and a prognostic blood marker for survival in patients with ovarian (Ma et al., 2007), (Polterauer et al., 2009) and gastric cancer (Yamashita et al., 2005).

3.2 Selection of high affinity single domain Abs from the immune library against Fib by MACS

Similar to previous work carried out with the immune libraries against TirM\textsubscript{EHEC}, MACS was used to isolate high affinity clones binding to biotinylated Fib from \textit{E. coli} EcM1 bacteria expressing \textit{V\textsubscript{HH}A} and \textit{NV\textsubscript{HH}} libraries, generated from dromedaries immunized with purified Fib. Initially, the \textit{V\textsubscript{HH}A} and \textit{NV\textsubscript{HH}} libraries were incubated with biotinylated Fib (250 nM) for the first selection step (MACS1) and \textasciitilde0.3-0.4\% of the total CFU from both libraries were collected in the Bound fractions (Table 7). The colonies grown from the Bound fractions of each library were pooled independently and their plasmids purified and electroporated into fresh \textit{E. coli} EcM1 cells to obtain \textit{V\textsubscript{HH}A} and \textit{NV\textsubscript{HH}} sublibraries ($\geq$2x$10^6$ transformants). Next, the \textit{V\textsubscript{HH}A} and \textit{NV\textsubscript{HH}} sublibraries were subjected to a new round of selection with 100 nM biotinylated Fib and maintaining the other conditions identical to those used in MACS1. Bacteria harvested from Bound fractions were pooled, their plasmids purified and transformed for the following rounds of MACS. In total, 5 rounds of MACS were performed with both the \textit{NV\textsubscript{HH}} and \textit{V\textsubscript{HH}A} libraries.

\textbf{Table 7. Summary of MACS with \textit{E. coli} display libraries against Fib.}

<table>
<thead>
<tr>
<th>Round</th>
<th>Antigen conc (Fib)</th>
<th>\textit{V\textsubscript{HH}A} library %</th>
<th>\textit{NV\textsubscript{HH}} library%</th>
</tr>
</thead>
<tbody>
<tr>
<td>MACS1</td>
<td>250 nM</td>
<td>0.3</td>
<td>0.4</td>
</tr>
<tr>
<td>MACS2</td>
<td>100 nM</td>
<td>0.8</td>
<td>0.7</td>
</tr>
<tr>
<td>MACS3</td>
<td>100 nM</td>
<td>1.7</td>
<td>3.3</td>
</tr>
<tr>
<td>MACS4</td>
<td>100 nM</td>
<td>2.9</td>
<td>9.6</td>
</tr>
<tr>
<td>MACS5</td>
<td>100 nM</td>
<td>26.1</td>
<td>24.5</td>
</tr>
</tbody>
</table>

(\%): Percentage of bacteria in Bound fractions with reference to total bacteria in Wash+Bound fractions (ca. 1-5 x 10\textsuperscript{7} CFU)

The percentage of \textit{E. coli} bacteria recovered in the Bound fractions showed a very gradual increase from the initial 0.3-0.4\% to over 50\% in MACS5 of \textit{NV\textsubscript{HH}} and \textit{V\textsubscript{HH}A} (Table 7), suggesting an enrichment of antigen binding clones in both libraries.
Bacteria from the different rounds of selection were analyzed by flow cytometry to test their binding to biotinylated Fib (50 nM) (Figure 26A and 26B), which demonstrated an enrichment of *E. coli* cells binding to Fib along the selection rounds, from ~0.25 % positives in the original libraries to ~40% and ~60% after MACS5 of the VHH-A library and NVHH library respectively, with no significant binding to biotinylated BSA. The expression levels of the VHH-A and NVHH fusions in the bacterial pools obtained after MACS was similar to those of the original libraries (Figure 26A and 26B).

**Figure 26.** Magnetic cell sorting of VHH-A and NVHH *E. coli* display libraries with biotinylated antigen, Fib.

Fluorescent flow cytometry analysis of IPTG-induced *E. coli* EcM1 cells expressing (A) VHH-A or (B) NVHH immune libraries, or their respective sub-libraries enriched after the indicated round of MACS with biotinylated Fib. Histograms show the fluorescence intensity of bacteria incubated with biotinylated Fib and secondary Streptavidin-PE or anti-E tag mAb and anti-mouse conjugated to Alexa 488 fluorophore.
Fifty colonies from the fifth round of MACS selection from both libraries were randomly picked for plasmid isolation and DNA sequencing. Upon DNA sequencing, three different VH₃ sequences, named as V₁FIB₁, V₁FIB₂ and V₁FIB₃ were identified in both systems. The remaining clones were negative in binding to biotinylated fibrinogen (data not shown). The screening is summarized in Table 8. Flow cytometry analysis confirmed the specific binding of biotinylated Fib (50 nM) by E. coli cells displaying each of the three VH₃ sequences fused to the Intimin β-domains, with negligible binding to biotinylated BSA. The MFI of V₁FIB₁ and V₁FIB₃ were the best and at similar levels (Figure 27).

![Figure 27. Binding of E. coli cells displaying selected clones from VH₃A and NVH₃ libraries to biotinylated Fib.](image)

Fluorescent flow cytometry analysis of induced E. coli EcM1 cells bearing the indicated plasmids selected from the NVH₃ library: pNV₁FIB₁, pNV₁FIB₂ and pNV₁FIB₃. Histograms show the fluorescence intensity of bacteria incubated with biotinylated antigens (Fib or BSA, as labeled) and secondary Streptavidin-PE or anti-E mAb and secondary anti-mouse Alexa 488.

**Table 8. CDR3 sequences of anti-Fib clones selected by E. coli display.**

<table>
<thead>
<tr>
<th>Clone Name</th>
<th>Amino acid sequence of CDR3</th>
<th>Display system</th>
</tr>
</thead>
<tbody>
<tr>
<td>V₁FIB₁</td>
<td>RWGWASSSNWYDMGKYNY</td>
<td>Neae 15/50</td>
</tr>
<tr>
<td>V₁FIB₂</td>
<td>RCAPES</td>
<td>C-EhaA 26/50</td>
</tr>
<tr>
<td>V₁FIB₃</td>
<td>KYYRSCTSLGDRNY</td>
<td>06/50</td>
</tr>
</tbody>
</table>
3.3 Purification of V<sub>FIB1</sub> and V<sub>FIB2</sub> V<sub>HH</sub>s from the periplasm of <i>E. coli</i> as fusions to MBP

The two more frequent clones against Fib isolated by <i>E. coli</i> display, V<sub>FIB1</sub> and V<sub>FIB2</sub>, were initially produced in the periplasm of <i>E. coli</i> WK6 cells as soluble V<sub>HH</sub>s with C-terminal His<sub>6</sub>- and myc-tags, but poor yields were attained (~0.1-0.2 mg/L; data not shown). Similar problems of variability in protein expression have been previously reported during the production of V<sub>HH</sub>s against methotrexate, (Alvarez-Rueda et al., 2007). Another common issue is that the single IMAC step is generally not sufficient to avoid the presence of protein contaminants in V<sub>HH</sub>s preparations, thus requiring further purification by less efficient chromatographic steps (e.g. gel filtration, ion-exchange). Hence, in order to circumvent these issues, we developed a purification method for V<sub>HH</sub>s based on their fusion to the <i>E. coli</i> maltose binding protein (MBP) in the N-terminus and a His<sub>6</sub>-tag in the C-terminus.

3.3.1 Expression and solubility of the MBP-V<sub>HH</sub><sub>His6</sub> fusions

In order to investigate whether fusion to MBP could have a positive effect on expression levels of V<sub>HH</sub>s and on their purification from the <i>E. coli</i> periplasm, we constructed an expression vector, named pMAL1-V<sub>HH</sub> (Figure 28).

![Figure 28. Schematic representation of protein expression vector.](image-url)
pMAL1-\(V_{HH}\) is based on the pMAL-p2E backbone (New England Biolabs), and contains in addition, the cleavage site for the site-specific protease PreScission (LEVLFQ/GP) in the linker connecting MBP with the \(V_{HH}\), a C-terminal His\(_6\) and myc tags. The MBP-fusions are under the control of the IPTG-inducible \(P_{\text{lac}}\) promoter, while the vector also encodes the LacI repressor and contains unique \(SfiI\) and \(NotI\) restriction sites flanking the \(V_{HH}\) sequence encoding the \(V_{HH}\). Three characterized \(V_{HH}\) genes with different complementarity determining region (CDR) sequences, encoding \(V_{HH}\)s that specifically recognized human fibrinogen (\(V_{FIB1}\) and \(V_{FIB2}\)), or GFP (\(V_{GFP}\)) were cloned and the corresponding pMAL1-\(V_{HH}\) derivative was generated. Expression of MBP-\(V_{HH}\)His\(_6\) fusions was initially tested in \(E.\ coli\) BL21 strain (Studier and Moffatt, 1986) but high proteolysis of the full-length fusions was observed (data not shown). Hence, we tested protein production in an \(E.\ coli\) strain deficient in several periplasmic proteases, \(E.\ coli\) HM140 (Meerman and Georgiou, 1994), and high-level accumulation of the full-length MBP-\(V_{HH}\)His\(_6\) (~60 kDa) was obtained after 3 h induction at 30ºC with 0.3 mM IPTG (Figure 29A).

**Figure 29.** Induction and solubility of the MBP-fusions.  

(A) 10% SDS-PAGE gel loaded with pre- and post-induction Whole Cell Extract (WCE) samples to check the induction and expression of the MBP-fusion protein. The MBP-\(V_{HH}\) fusion has a size of about 60 kDa (as labelled with the arrow).  

(B) 10% SDS-PAGE gel loaded with (Lane 1) Whole Cell Extract (WCE) obtained after lysis of IPTG-induced \(E.\ coli\) HM140 cells expressing the MBP fusion proteins. The WCE sample was centrifuged at 20000 X g for 60 mins in a tabletop centrifuge, and the pellet after centrifugation (Insoluble fraction, Lane 2) and the supernatant (Soluble fraction, Lane 3) were loaded onto the 10% SDS-polyacrylamide gel. The MBP-\(V_{HH}\) fusions are found in the soluble fraction of the WCE after centrifugation. Mass of protein standards is shown on the left (in kDa).
Approximately 15-20 mg of MBP-V_{HH}{\text{His6}} fusions were produced per litre of induced culture in shake flasks with a final OD_{600} \sim 1.0. Importantly, in all cases the majority of the accumulated MBP-V_{HH}{\text{His6}} fusion remained in the soluble fraction after cell lysis and high-speed centrifugation (20000 \times \text{g}, 1 \text{ h}), as indicated by SDS-PAGE analysis (Figure. 29B).

### 3.3.2 Purification of MBP-V_{HH}{\text{His6}} fusions and antigen binding activity

The supernatants containing the MBP-V_{HH}{\text{His6}} fusions obtained after high-speed centrifugation of the cell lysate were passed through a chromatography column packed with amylose resin and the bound proteins were eluted with 10 mM maltose (Figure 30A). SDS-PAGE analysis showed a pre-dominant band of \sim 60 \text{kDa} in the eluted fractions that corresponds to MBP-V_{HH}{\text{His6}} fusions (Figure 29A). The fractions containing the MBP-V_{HH}{\text{His6}} fusions were pooled and loaded onto cobalt-containing columns for IMAC. Bound MBP-V_{HH}{\text{His6}} fusions were eluted with 150 mM imidazole, pooled and analyzed by SDS-PAGE showing the presence of a single major band corresponding to the full-length MBP-V_{HH}{\text{His6}} fusions (Figure 30B). In all cases, a final yield of \sim 12-16 mg of purified MBP-V_{HH}{\text{His6}} fusions were obtained per L of induced culture (Table 9).

ELISA was subsequently used to test the antigen-binding activity of the purified MBP-V_{HH}{\text{His6}} fusions and different concentrations of the purified MBP-V_{HH}{\text{His6}} fusions were incubated with their respective antigens (GFP or Fib) and with a negative control antigen (BSA). Bound MBP-V_{HH}{\text{His6}} fusions were developed with anti-c-myc-mAb-POD. The generated binding curves (Figure 30C) demonstrate that these fusions recognized their specific antigens, indicating that the presence of MBP did not prevent the binding of the antigen by the V_{HH}, which is in agreement with previous reports that MBP-scFv fusions maintain the antigen binding activity of scFvs (Bach et al., 2001). Thus, our data showed that MBP-V_{HH}{\text{His6}} fusions can be purified in high yields by two steps of affinity chromatography and they can be directly used for antigen binding or applications thereof.
Results

Figure 30. Purification of the MBP-fusions by Amylose affinity chromatography and their functional analysis by ELISA.

(A) Chromatogram of the purification of MBP-fusion proteins by Amylose Affinity Chromatography. The flowthrough fraction contains non-specific proteins and other impurities that do not bind the column and the elute fraction contains the purified MBP-VHH fusions. (B) 10% SDS-PAGE gel loaded with each of the three purified MBP-VHH fusions i.e. MBP-VFIB1 and MBP-VFIB2 are the fusions of MBP with either of two different VHHs against human fibrinogen and MBP-VGFP is a fusion of MBP with a VHH against GFP isolated from VHH libraries against GFP. (C) ELISA to determine the antigen binding activity of the purified MBP-VHH fusions. Different wells were coated with 5 µg/ml of GFP and human fibrinogen (positive antigens) or BSA (negative control) proteins, followed by addition of MBP-VHH solution (0.05-100 nM) to all coated wells. The presence of bound MBP-VHH fusion was detected by incubation with anti-c-myc-POD and development using a chromogenic substrate. The absorbance values reflect the specific binding of each of the MBP-VHHs to their respective antigens and have been adjusted by subtraction of the respective blank values (non-specific binding of each MBP-VHH to BSA).
3.3.3 Purification of monomeric VHH_{His6} from MBP-VHH_{His6} fusions

To purify monomeric VHH_{His6}s free of MBP, we evaluated the following purification scheme summarized in Figure 31.

**Figure 31. Expression and purification of VHHs in *E. coli*.**

The scheme followed for purification of VHHs from MBP-VHH fusion proteins is illustrated.

Briefly, fractions containing the MBP-VHH_{His6} fusions eluted from the amylose resin were dialyzed to remove maltose and digested with the PreScission protease at 4 °C for 16 h. The products of the proteolysis were analyzed on a 12% SDS-PAGE gel (Figure 32A). About 80-90% of the initial full-length MBP-VHH_{His6} fusion present in the reaction was digested into free MBP (~45 kDa) and the corresponding VHH_{His6} (~17 kDa) in all cases. The proteolytic products were subsequently loaded onto an amylose resin in order to capture the undigested MBP-fusion and free MBP, while collecting the VHH_{His6} in the flow-through. The flow-through fractions for each of the three VHHs were loaded onto a cobalt-containing column for IMAC. PreScission protease and residual free MBP, along with other impurities, were washed from the IMAC column and the purified VHH_{His6} clones were eluted with 150 mM imidazole (Figure 32B). Some residual amounts of the full-length MBP-VHH_{His6} fusion could be
detected in some cases, due to the presence of His$_6$-tag. The final yields of purified VHH$_{\text{His6}}$ were ~2-3 mg per L of induced culture (Table 9).

![Protease digestion of MBP-VHH fusions, Immobilized Metal Affinity chromatography (IMAC) and Gel filtration.](image)

**Figure 32.** Protease digestion of MBP-VHH fusions, Immobilized Metal Affinity chromatography (IMAC) and Gel filtration.

**A)** SDS-PAGE of MBP-VHH fusions digested overnight with PreScission protease at 4°C for 16 h. Upon observation of the gel, it can be inferred that majority of the MBP-VHH fusion was digested and separated from the VHH. Mass of protein standards is shown on the left (in kDa). **(B)** IMAC purification of VHH using Talon resin. The chromatogram of the chromatographic run is illustrated in this figure. Residual impurities like MBP and PreScission protease do not bind to the Talon resin and are found in the flowthrough, while the VHH binds to the resin and is eluted with 200 mM imidazole in the elution peak. **(C)** The VHH was loaded on to a Gel filtration column to analyze the protein for presence of multimers or dimers and separate the monomeric VHH from other forms. The exclusion volume (Vo) and the molecular weights of the gel filtration standards are indicated (in kDa). Peak 1 corresponds to the VHH in monomeric form with a molecular weight of ~16 kDa.
Table 9. Protein yields after purification from the periplasm of E. coli HM140.

<table>
<thead>
<tr>
<th>Fusion protein</th>
<th>MBP- V_{HH}^{His6} (mg/litre culture)</th>
<th>MBP-V_{HH}^{His6} (mg/litre culture)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MBP-V_{HH1}</td>
<td>12.5</td>
<td>2.2</td>
</tr>
<tr>
<td>MBP-V_{HH2}</td>
<td>16.3</td>
<td>3.2</td>
</tr>
<tr>
<td>MBP-V_{HH3}</td>
<td>14.6</td>
<td>2.5</td>
</tr>
</tbody>
</table>

Gel filtration chromatography through a precalibrated HiLoad 16/600 Superdex 75 pg column was used to analyze the composition of the eluted V_{HH}^{His6} after IMAC and to purify the fractions corresponding to monomeric V_{HH}^{His6}. In each of the three cases, ≥95% of the protein eluted corresponded to the monomeric V_{HH}^{His6} as depicted by the single major peak with an apparent molecular weight of ca. 15 kDa obtained by gel filtration (Figure 32C).

Further, each purified monomeric V_{HH}^{His6} (namely V_{FIB1}, V_{FIB2} and V_{GF}) was analyzed by SDS-PAGE and revealed a single protein band with molecular weights of ca. 15 -17 kDa (Figure 33A) and was found to be functional by ELISA (Figure 33B).

**Figure 33. Monomeric V_{HH}s and assay of functional activity by ELISA.**

(A) Monomeric V_{HH}s after gel filtration were loaded on to a 12% SDS-polyacrylamide gel and stained with Coomassie Brilliant Blue solution. The purified V_{HH}s (Lane 1 and Lane 2 – V_{HH}s against Fibrinogen, Lane 3 – V_{HH} against GFP) have a molecular weight between 15 and 18 kDa. Mass of protein standards is shown on the left (in kDa). (B) ELISA to determine the antigen binding activity of purified monomeric V_{HH}s. Wells were coated with 5 mg/ml of GFP, Human Fibrinogen (positive antigen) or BSA (negative control) proteins, followed by addition of V_{HH} solution (0.05-100 nM) to the wells. The signal was developed by incubation with anti-c-myc-POD.
3.4 Determination of the affinity of V\textsubscript{FIB1} and V\textsubscript{FIB2} by Surface Plasmon Resonance (SPR)

In order to determine the apparent equilibrium dissociation constant ($K\textsubscript{D}$) of V\textsubscript{FIB1} and V\textsubscript{FIB2} against Fib, their interaction was studied by SPR experiments with a Streptavidin (SA) sensor chip coated with biotinylated Fib (Materials and Methods). The change in resonance units (RU) was recorded with time at different concentrations of purified V\textsubscript{FIB1} (from 2.5 to 40 nM) or V\textsubscript{FIB2} (from 5 to 80 nM), and showed clear binding to Fib that reached the steady state equilibrium in ~180s (Figure 34A and 34B). No binding was observed when Vgfp (40 nM) was flown over this sensor surface, or when V\textsubscript{FIB1}, V\textsubscript{FIB2} (40 nM) was flown over a SA flow-cell lacking biotinylated Fib (data not shown). Injection of buffer to evaluate the dissociation of V\textsubscript{FIB1} or V\textsubscript{FIB2} showed gradual loss of RU with time and the kinetic constants ($k\textsubscript{on}$ and $k\textsubscript{off}$) were calculated from the sensograms. V\textsubscript{FIB1} had a $K\textsubscript{D}$ of 3.2 x 10$^{-9}$ M, with $k\textsubscript{on} = 6.62 \times 10^5$ Ms$^{-1}$ and $k\textsubscript{off} = 2.1 \times 10^{-3}$ s$^{-1}$. V\textsubscript{FIB2}, on the other hand, had a $K\textsubscript{D}$ of 24.7 x 10$^{-9}$ M, with $k\textsubscript{on} = 1.02 \times 10^6$ Ms$^{-1}$ and $k\textsubscript{off} = 2.52 \times 10^{-2}$ s$^{-1}$.

![Figure 34. Determination of the equilibrium dissociation constant (K\textsubscript{D}) of clones V\textsubscript{FIB1} and V\textsubscript{FIB2} by SPR.](image)

SPR sensograms monitoring real-time association and dissociation of purified sdAbs (A) V\textsubscript{FIB1} and (B) V\textsubscript{FIB2} (at the indicated concentrations) to biotinylated Fib immobilized onto a Streptavidin-SA sensor chip. The increase in resonance units (RU) is recorded along time (in seconds). Dissociation of V\textsubscript{FIB1} and V\textsubscript{FIB2} was evaluated by injection of buffer. The curve was fitted by non-linear least squares regression.
3.5 Estimation of the affinity of \textit{VFIB1} and \textit{VFIB2} by \textit{E. coli} display

Flow cytometry analysis was used to estimate the apparent $K_D$ of the selected \textit{VFIB} sdAbs as previously performed with the \textit{VTIR1} sdAb. It was known that the sdAbs, \textit{VFIB1} had higher apparent affinity than \textit{VFIB2}, therefore the final concentration range of biotinylated Fib in the assay for \textit{VFIB1} was maintained between 30 nM and 0.3 nM, while in the case of \textit{VFIB2}, this range was between 300 nM and 3 nM. Briefly, \textit{E. coli} cells ($\sim3 \times 10^7$ CFU) with the respective sdAb on their surface were incubated with a fixed amount of biotinylated Fib (2 pmols) for 90 mins in two-fold increasing volumes of PBS (from 0.01 to 2 ml) to reach the respective final concentration ranges (as mentioned above). After this incubation, cells were washed and labeled with Streptavidin-PE as previously described.

The relative MFI of the cells was plotted against the antigen concentration used and the curve fitted by non-linear least squares regression, giving an estimated apparent $K_D$ of $3 \times 10^{-9}$ M for \textit{VFIB1} and $35 \times 10^{-9}$ M for \textit{VFIB2} (Figure 35A and 35B). These values were consistent with the $K_D$ determined by SPR (Figure 34A and 34B).

![Figure 35. Estimation of the equilibrium dissociation constant (KD) of clones VFIB1 and VFIB2 against Fib by E. coli display.](image)

The $K_D$ of the selected clones (as indicated) was estimated by flow cytometry analysis of \textit{E. coli} cells expressing NVFIB1 or NVFIB2 incubated with different concentrations of biotinylated Fib under equilibrium conditions. The mean fluorescent intensities (MFI) of bacteria, after labeling with Streptavidin-PE, were plotted \textit{versus} the concentration of Fib used in the assays. The curve was fitted by non-linear least squares regression.
Chapter 4: Cell Selection and characterization of high affinity single domain Abs against human EGFR from immune libraries displayed on the surface of \textit{E. coli} cells

4.1 Epidermal Growth Factor Receptor (EGFR)

The epidermal growth factor receptor (EGFR/Her1/ErbB1) is a transmembrane receptor belonging to the ErbB family of tyrosine kinases and is implicated in many human cancers (especially on squamous cancer cells), being an important target for several classes of therapeutic agents, including Ab-based drugs. These receptors are abnormally expressed in many epithelial tumors such as colorectal, lung, brain, breast, head and neck tumors etc (Gullick, 1991; Baselga and Arteaga, 2005; Huang et al., 2009) and influence their growth and survival in malignant states (Zaczek et al., 2005). The receptor comprises a transmembrane domain, an extracellular ligand-binding domain and an intracellular domain with tyrosine kinase activity (Figure 36A). Several mammalian ligands are known to activate EGFR, including EGF, transforming growth factor-\(\alpha\) (TGF-\(\alpha\)), heparin-binding EGF-like growth factor (HB-EGF), amphiregulin (AR), betacellulin (BTC), epiregulin (EPR) and epigen (EPI). EGFR signalling promotes angiogenesis, proliferation, invasion and metastasis of the tumor cells.

It is well established that Ab binding to the extracellular domain of EGFR can inhibit ligand-induced receptor activation and tumor growth (Sato et al., 1983; Gill et al., 1984). Various EGFR agents have been developed as anti-cancer drugs, for example, cetuximab, panitumumab (both are mAbs), gefitinib and erlotinib (small-molecule inhibitors). The mAbs target the extracellular domain of the receptor and inhibit ligand-dependent EGFR signal transduction, while the small-molecule inhibitors target the intracellular tyrosine kinase. Inactive EGFR can exist as both monomers and dimers, suggesting that the mechanism regulating EGFR activity may be subtle. Dimerization-driven activation of the intracellular kinase domains of the epidermal growth factor receptor (EGFR) upon extracellular ligand binding is crucial to cellular pathways regulating proliferation, migration, and differentiation. Molecular dynamic simulations of the membrane with the embedded EGFR
(Arkhipov et al., 2013) suggest that in ligand-bound dimers, the extracellular domains assume conformations favoring dimerization of the transmembrane helices near their N termini, dimerization of the juxtamembrane segments, and formation of asymmetric (active) kinase dimers. In ligand-free dimers, by holding apart the N termini of the transmembrane helices, the extracellular domains instead favor C-terminal dimerization of the transmembrane helices, juxtamembrane segment dissociation and membrane burial, and formation of symmetric (inactive) kinase dimers (Figure 36B).

![Figure 36. Schematic of EGFR and the states in which it exists.](image)

**A** Scheme of EGFR domains (left), showing extracellular domain, transmembrane (TM) α-helix, juxtamembrane domains A&B, kinase domain, and C-terminal tail. **B** EGFR is present in the monomeric form or as a pre-formed dimer (inactive). Encounter of its ligand i.e. EGF, leads to phosphorylation of its C-terminal tail (active dimer) (Arkhipov et al., 2013), initiation of signalling cascades and stimulation of cell cycle machinery. Leads to secretion of growth, angiogenic and motility factors.

### 4.2 Construction of immune sdAb library against Human EGFR

An immune sdAb library against Human EGFR cloned in a phage display vector (Roovers et al., 2007) was a kind gift from Dr Rob Roovers and Paul van Bergen en Henegouwen, University of Utrecht, Netherlands. This library contains ca. $10^7$
independent clones and was constructed from $V_{HH}$ gene segments amplified from lymphocytes isolated from two llamas (*Llama glama*) immunized with intact human A431 cells, which overexpress EGFR (ca. $2 \times 10^6$ molecules of EGFR/cell). The $V_{HH}$ repertoire of this library was cloned into the *SfiI* and *NotI* sites of the pNeae2 vector and transformed into EcM1 cells for *E. coli* display. A library of $\sim 2 \times 10^7$ independent clones was generated.

### 4.3 Selection of high affinity sdAbs from the anti-EGFR immune library by MACS

To test the *E. coli* display NV$_{HH}$ library anti-EGFR we employed first MACS with purified rhEGFR-Fc, a recombinant protein carrying the extracellular domain of human EGFR fused to an Ig Fc region. Similar to previous selections with TirM and Fib, $2 \times 10^8$ bacteria were incubated with biotinylated rhEGFR-Fc (100 nM) for the first selection step (MACS1) and $\sim 0.6-0.7\%$ of the total CFU were collected in the Bound fraction (Table 10). The colonies grown from the Bound fraction were pooled, their plasmids purified and electroporated into fresh *E. coli* EcM1 cells ($\geq 10^7$ transformants). This sublibrary was subjected to a new round of selection with 100 nM biotinylated rhEGFR-Fc under identical conditions.

**Table 10. Summary of MACS with *E. coli* display libraries against rhEGFR.**

<table>
<thead>
<tr>
<th>Round</th>
<th>Antigen conc (rhEGFR-Fc)</th>
<th>NV$_{HH}$ library%</th>
</tr>
</thead>
<tbody>
<tr>
<td>MACS1</td>
<td>100 nM</td>
<td>0.3</td>
</tr>
<tr>
<td>MACS2</td>
<td>100 nM</td>
<td>2.6</td>
</tr>
</tbody>
</table>

(\%): Percentage of bacteria in Bound fractions with reference to total bacteria in Wash+Bound fractions (ca. $1-5 \times 10^7$ CFU)

The percentage of *E. coli* bacteria recovered in the Bound fractions increased to $2.6\%$ in MACS2 (Table 10) and were analyzed by flow cytometry to test their binding to biotinylated rhEGFR-Fc (50 nM) (Figure 37), which demonstrated an enrichment of *E. coli* clones binding to rhEGFR-Fc with more than 70\% positives after MACS2. No significant binding to biotinylated BSA was detected by flow cytometry in these populations. The expression levels of the NV$_{HH}$ fusions in the bacterial pools obtained after MACS were similar to those in the original library (Figure 37).
4.4 Direct cell selection of high affinity sdAbs against EGFR using *E. coli* display.

We wanted to determine whether *E. coli* display could allow the direct selection of sdAbs from immune libraries against EGFR on live tumor cells overexpressing EGFR on its surface, instead of purified recombinant EGFR protein, previously used during MACS. Hence, we designed a simple cell selection (CellS) procedure consisting of an initial incubation step of the *E. coli* display library with a monolayer of the murine tumor cell line, NIH-3T3 2.2, which lacks expression of endogenous EGFR. This allows subtractive selection of binders recognizing other cell surface antigens. Next, clones that did not bind NIH-3T3 2.2 cells were incubated with a monolayer of the murine tumor cell line HER14, which is a stably transfected NIH-3T3 2.2 clone expressing hEGFR (Honegger et al., 1987). Bacteria binding to HER14 were recovered on LB plates for plasmid isolation and subjected to subsequent rounds of selection. A schematic illustration of the methodology proposed is shown in Figure 38.
Results

Figure 38. Direct selection of NVHH E. coli display libraries on cells.

E. coli cells are initially bound to eukaryotic cells not expressing EGFR i.e. NIH-3T3 2.2, and any bacteria binding non-specifically to these cells are removed. The bacteria that do not bind are recovered and incubated with eukaryotic cells that express EGFR i.e. HER14. The HER14 cells are washed to remove non-specific or loosely bound E. coli cells, followed by lysis of eukaryotic cells and elution of bound bacteria. The CFU of bound bacteria in the cell lysate is determined by plating.

Both cell lines were seeded in 6-well cell culture plates to attain a confluence of 50% in 24 h (=6X10^5 cells per well). Firstly, depletion of bacteria displaying VHH clones binding to cell surface antigens but not to EGFR was carried out by incubation of induced bacteria (=6X10^7) with NIH-3T3 2.2 cells (EGFR-) at a MOI of 100, for 1h at 37°C. Secondly, the bacteria that were not bound to the negative cell line were recovered and incubated with HER14 cells (EGFR+) for 15 mins at 37°C. Further, the cells were washed three times with HBSS to remove unbound bacteria and lysed by addition of HBSS supplemented with 0.2% SDS and 0.1% DNase. The resulting cell lysate containing bacteria was plated.

The colonies grown in plates were pooled, their plasmids purified and electroporated into fresh E. coli EcM1 cells. The sublibrary was subjected to a additional round of CellS with NIH-3T3 2.2 and HER14 cell lines using identical conditions. The percentage of E. coli bacteria recovered in the HER14 cell-bound fractions showed an increase from the initial 0.3% in CellS1 to 1.2% in CellS2 (Table 11), suggesting an enrichment of clones binding specifically to HER14.
Results

Table 11. Summary of Cell selections with NV\textsubscript{HH} \textit{E. coli} display library against rhEGFR.

<table>
<thead>
<tr>
<th>Round</th>
<th>Bacteria in input</th>
<th>Bacteria in output</th>
<th>% bacteria recovered</th>
</tr>
</thead>
<tbody>
<tr>
<td>CellS1</td>
<td>$6 \times 10^7$</td>
<td>$2 \times 10^5$</td>
<td>0.3</td>
</tr>
<tr>
<td>CellS2</td>
<td>$6 \times 10^7$</td>
<td>$7 \times 10^5$</td>
<td>1.2</td>
</tr>
</tbody>
</table>

A total of $6 \times 10^6$ cells were seeded in 6 well culture plates. Bacteria were added at a MOI =100

Bacterial pools from each of the different rounds of cell selection were analyzed by flow cytometry to test their binding to biotinylated rhEGFR (50 nM) (Figure 39), which demonstrated an enrichment of \textit{E. coli} cells binding to rhEGFR along the selection rounds, from ~0.2 % positives in the original libraries to ~35% after CellS2. No significant binding to biotinylated BSA was detected by flow cytometry in these populations. The expression levels of the NV\textsubscript{HH} fusions in the bacterial pools obtained after cell selection was similar to those of the original libraries (Figure 39).

Figure 39. Flow cytometry analysis of anti-EGFR NV\textsubscript{HH} \textit{E. coli} display libraries selected on cells.

Fluorescent flow cytometry analysis of IPTG-induced \textit{E. coli} EcM1 cells expressing NV\textsubscript{HH} immune libraries, or their respective sub-libraries enriched after the indicated round of cell selection with biotinylated rhEGFR. Histograms show the fluorescence intensity of bacteria incubated with biotinylated rhEGFR-Fc (50 nM) and secondary Streptavidin-PE or anti-myc tag mAb and anti-mouse conjugated to Alexa 488 fluorophore.
Ninety-six colonies from the second round of selection of the NV_{HH} library by MACS (i.e MACS2) were randomly picked for plasmid isolation and DNA sequencing. In the screening of the MACS2, a V_{HH} sequence, named as V_{EGFR1}, was found in 72 NV_{HH} clones, while other clones, namely V_{EGFR2} (10 clones), V_{EGFR3} (1 clone), V_{EGFR4} (3 clones) and V_{EGFR5} (1 clone) were also found, the rest being non-EGFR binding V_{HH} sequences.

Similarly, ninety-six clones from CellS2 were screened by visual observation under a light microscope of their specific binding to HER14 cells (EGFR^+), but not to NIH-3T3 2.2 cells (EGFR^−). Briefly, induced bacteria were added to wells containing either HER14 cells or NIH-3T3 2.2 cells at a MOI of 100 and the infection was monitored after 20 minutes by visual inspection in an inverted microscope. Clones that did not bind either cell line were classified as “Non-binders”, those that bound both NIH-3T3 2.2 cells and HER14 cells were classified as “non-specific binders” and those that bound only HER14 cells, but not NIH-3T3 2.2 cells were called “specific binders”. Out of the 96 clones screened by this method, 45 clones were non-binders, 26 clones were non-specific binders and 25 clones were specific binders. DNA sequences of all 25 specific binders were isolated and sequenced. Like previous screening of MACS2 population, V_{EGFR1} was found 18 times and was the major clone. V_{EGFR2} (3 clones), V_{EGFR4} (2 clones) were also found, while one new clone, V_{EGFR6} (2 clones) was also identified. The CDR3 sequences of all identified clones are illustrated in Table 12.

**Table 12. CDR3 sequences of anti-EGFR clones selected by E. coli display.**

<table>
<thead>
<tr>
<th>Clone Name</th>
<th>Amino acid sequence of CDR3</th>
<th>Selection system</th>
</tr>
</thead>
<tbody>
<tr>
<td>V_{EGFR1}</td>
<td>DKWSSRRSVDYDYSR</td>
<td>MACS 72/96</td>
</tr>
<tr>
<td>V_{EGFR2}</td>
<td>TYNPYSRDHYFPRMTTEYDY</td>
<td>CellS 18/96</td>
</tr>
<tr>
<td>V_{EGFR3}</td>
<td>RYSVIFLTLPFRY</td>
<td>MACS 10/96</td>
</tr>
<tr>
<td>V_{EGFR4}</td>
<td>STYSRDSVFTKWANYNY</td>
<td>CellS 03/96</td>
</tr>
<tr>
<td>V_{EGFR5}</td>
<td>GPIGSSSYRSQAWR</td>
<td>MACS 01/96</td>
</tr>
<tr>
<td>V_{EGFR6}</td>
<td>DLNFIGIVTTSEEKDY</td>
<td>CellS -</td>
</tr>
</tbody>
</table>
4.5 Binding of selected clones to EGFR+ tumor cells by immunofluorescence microscopy

All anti-EGFR clones isolated by *E. coli* display were further characterized for their specific binding to HER14 cells (EGFR⁺) but not to NIH-3T3 2.2 cells (EGFR⁻) by immunofluorescence microscopy, staining the samples with DAPI (nuclei) and fluorescent antibodies against EGFR and *E. coli*. This experiment revealed that all positive anti-EGFR *E. coli* obtained by MACS and CellS (Table 12) bound specifically to HER14 cells but not to NIH-3T3 2.2 cells (Figures 40 and 41), while one non-specific clone (number 37 from CellS, VCLONE37) from CellS screening bound to both cell lines and a negative control *E. coli* displaying NVFIB1 bound to neither of the cell lines (Figure 40).
Figure 40. Bright field and immunofluorescence microscopy images of \textit{E. coli} bacteria displaying the indicated VHH clone to HER14 and NIH-3T3 2.2 cells.

These images show the binding specificity of \textit{E. coli} bacteria displaying the selected anti-EGFR clones \textit{VEGFR1} and \textit{VEGFR2}, to HER14 and not to NIH-3T3 2.2 cells. A negative clone (VfIB1) and a non-specific clone from CellS (VCLONE 37) are also included as controls. Bacteria are labelled with anti-\textit{E. coli} (red), EGFR is labelled with anti-EGFR mAb (green), DNA and cell nuclei labelled with DAPI (blue).
Results

4.6 Binding of selected clones to biotinylated rhEGFR-Fc by flow cytometry.

Flow cytometry analysis confirmed the specific binding of biotinylated rhEGFR-Fc (50 nM) by *E. coli* cells displaying each of the six VHH sequences, with negligible binding to biotinylated BSA. The MFI of VEGFR1 and VEGFR2 was the best and at similar levels (Figure 42A), which suggested that these clones had a high apparent
affinity for rhEGFR, while the MFI of VEGFR3-6 was found to be lower, suggesting that it had a lower apparent affinity for rhEGFR. An overlay of the binding of all selected clones to rhEGFR in flow cytometry is illustrated in Figure 42B.

Figure 42. Binding of *E. coli* cells displaying selected clones from anti-EGFR NVHH libraries to biotinylated rhEGFR-Fc.

(A) Fluorescent flow cytometry analysis of induced *E. coli* EcM1 cells bearing the indicated plasmids selected from the anti-EGFR NVHH library. (B) Overlay of histograms of binding of the selected clones to biotinylated EGFR. Histograms show the fluorescence intensity of bacteria incubated with biotinylated antigens (rhEGFR or BSA, as labeled) and secondary Streptavidin-PE.
4.7 Estimation of the affinity of VEGFR\textsubscript{1} and VEGFR\textsubscript{2} by \textit{E. coli} display

Flow cytometry analysis was used to estimate the apparent $K_D$ of two $V_{\text{EGFR}}$ clones with highest binding signals. The final concentration range of biotinylated rhEGFR-Fc in the assay for was between 100 nM and 0.8 nM. Briefly, \textit{E. coli} cells (~3x10\textsuperscript{7} CFU) with the respective sdAb on their surface were incubated with a fixed amount of biotinylated rhEGFR-Fc (1 pmol) for 90 mins in two-fold increasing volumes of PBS (from 0.01 to 2 ml) to reach the desired final concentration range (as above). After this incubation, cells were washed and labeled with Streptavidin-PE as previously described. The relative MFI of the cells was plotted against the antigen concentration used and the curve fitted by non-linear least squares regression, giving an estimated apparent $K_D$ of 30 x10\textsuperscript{-9} M for VEGFR\textsubscript{1} (Figure 43A) and 12 x10\textsuperscript{-9} M for VEGFR\textsubscript{2} (Figure 43B).

![Figure 43. Estimation of the equilibrium constant (KD) of VEGFR1 and VEGFR2 by \textit{E. coli} display.](image)

The $K_D$ of the selected clones (as indicated) was estimated by flow cytometry analysis of \textit{E. coli} cells expressing NV\textsubscript{VEGFR}1 or NV\textsubscript{VEGFR}2 incubated with different concentrations of biotinylated EGFR (0.8-100 nM) as specified, under equilibrium conditions. The mean fluorescent intensities (MFI) of bacteria, after labeling with Streptavidin-PE, were plotted versus the concentration of EGFR used in the assays. The curve was fitted by non-linear least squares regression.

4.8 Binding of the selected clones to EGFR in the presence of EGF

All the clones selected by \textit{E. coli} display were further characterized for their ability to bind EGFR in the presence of EGF using flow cytometry. It was observed that clones VEGFR\textsubscript{2}, VEGFR\textsubscript{3}, VEGFR\textsubscript{4}, VEGFR\textsubscript{6} and to a lesser extent VEGFR\textsubscript{1}, had reduced...
Results

binding signals to rhEGFR-Fc (50 nM) when 25 µM EGF was present, while the binding of VEGFR5 was almost the same in the presence or absence of EGF (Figure 44).

Figure 44. Flow cytometry analysis of binding to EGFR by the clones selected by E. coli display in the presence of EGF.

Fluorescent flow cytometry analysis of binding of the selected anti-EGFR clones to rhEGFR-Fc in the presence or absence of EGF. Histograms show the fluorescence intensity of bacteria incubated with biotinylated rhEGFR (50 nM) in the presence or absence of 25 µM EGF.
DISCUSSION
In this thesis, we have evaluated and described the application of two $\beta$-domains of EhaA and Intimin from EHEC, for the surface display of $V_{HH}$ clones and immune libraries against 3 different antigens of biomedical importance i.e. the extracellular domain of the translocated intimin receptor from EHEC (TirM$\text{EHEC}$), human fibrinogen (Fib) and epidermal growth factor receptor (EGFR). The anti-TirM$\text{EHEC}$ and anti-Fib $V_{HH}$ libraries were initially used to setup the selection conditions by MACS and to compare the effectiveness of the two $\beta$-domains in selection of $V_{HH}$ thereof. Further, the more effective $\beta$-domain from Intimin was used to display an immune library raised using tumor cells overexpressing hEGFR and select specific high affinity binders against hEGFR using the recombinant extracellular domain of hEGFR fused to Fc and murine tumor cell lines transfected with or without hEGFR.

We have found that the $\beta$-domains of EhaA and Intimin are effective platforms for the surface display of $V_{HH}$ libraries on $E. \ coli$ K-12 cells, and allow for selection of high affinity $V_{HH}$s from immune libraries using biotinylated antigen and MACS (Salema et al., 2013). Despite their opposite topologies, both systems express stable fusion proteins with the native $\beta$-barrel properly folded in the OM and the $V_{HH}$ domain displayed to the extracellular milieu, retaining functionality and antigen binding capacity. Most clones in the immune libraries were displayed at good levels on $E. \ coli$ with both EhaA and Intimin $\beta$-domains, with an average of 6000-8000 molecules/bacterium. From each of the immune libraries used in our studies, we obtained about 3-6 different camelid $V_{HH}$ sequences with different binding affinities (in the nM range) to the antigen used in the immunization, which is one of the advantages of using immune libraries. These results demonstrate that both the $E. \ coli$ display systems can be used to retrieve high-affinity binders from immune libraries of $V_{HH}$ sequences. The small molecular weight, high protein solubility and stability of camelid $V_{HH}$ domains are likely to aid their effective translocation across the OM fused to the $\beta$-domains of either EhaA and Intimin. Similar properties have been reported with other natural sdAbs, like $\text{V}_{\text{NAR}}$s from sharks (Dooley, 2003), which have been engineered in synthetic libraries based on human $V_{\text{H}}$s and $V_{\text{L}}$s. Thus, it is likely that other types of sdAbs could also be efficiently displayed on $E. \ coli$ cells with the EhaA and Intimin $\beta$-domains. Display of larger rAb fragments
based on a single polypeptide such as scFvs could be also possible, although the tendency to aggregate of many scFv molecules may difficult their translocation across the OM (Veiga et al., 2004). Nevertheless, this could be advantageous for the selection of highly stable and soluble scFvs from libraries. Lastly, Ab molecules with separate H and L polypeptide chains (e.g Fabs, IgGs) cannot be displayed with these β-domains (atleast in their current conformation), and either phage display or APEX (Harvey et al., 2004; Mazor et al., 2007; Mazor et al., 2010) should be used instead for their selection in E. coli.

Comparison of EhaA and Intimin display systems led to the observation of some important differences between them. Firstly, NV\textsubscript{HH} fusions were found to be more stable than V\textsubscript{HH}A fusions \textit{in vivo} (to \textit{E. coli} proteases) and \textit{in vitro} (to externally added proteases), which could be attributed to the natural resistance of Intimin to proteolysis and denaturation (Bodelon et al., 2009); and/or due to the susceptibility of certain ATs to endogenous bacterial proteases as a part of their secretion mechanism (Bernstein, 2007; Leyton et al., 2012b). Secondly, expression of V\textsubscript{HH}A clones appeared to be more variable than NV\textsubscript{HH} fusions, with some clones showing significantly lower expression levels. This suggests that the N-terminal fragment of Intimin could have a positive effect on the expression of V\textsubscript{HH}S, as observed with other N-terminal fusion partners (i.e. MBP, Trx1, GST) used routinely for the production of recombinant proteins and Ab fragments (Jurado et al., 2006). Thirdly, the antigen binding levels determined by flow cytometry with NV\textsubscript{HH} library were found to be atleast 3 fold higher than the respective V\textsubscript{HH}A library. The lower antigen binding signals of \textit{E. coli} cells displaying V\textsubscript{HH}A fusions could not be solely attributed to lower expression levels, and indicated the existence of additional factors. Eventhough partial misfolding of V\textsubscript{HH}A fusions cannot be totally excluded, this seems unlikely given that both EhaA and Intimin β-domains use a common secretion pathway, exposing the V\textsubscript{HH}S to the same periplasmic folding factors (e.g. DsbA) and chaperones (e.g. FkpA and Skp) that are known to participate in the folding of Ig domains (Bodelon et al., 2012). An alternative explanation could be that the longer linker region in NV\textsubscript{HH} fusions (with the D0 domain), could make the V\textsubscript{HH} domain more accessible for binding to the antigen in solution by increasing its
distance from the OM. The improved stability, higher expression levels and superior antigen binding activity could explain the reason why selection of high affinity binders from the anti-TirM_{EHEC} and anti-Fib libraries was more efficient using the Intimin system than the EhaA system, with a higher percentage of positive antigen-binding clones being achieved in fewer selection rounds. The superior antigen binding signals with the Intimin β-domain also facilitates the estimation of the affinity of the selected NV_{HH} clones to their antigen by flow cytometry. From our data, the only possible limitation of Intimin display could be a slight decrease in the viability of *E. coli* cultures expressing some NV_{HH} fusions, which was observed for anti-TirM_{EHEC} library. However, this reduction in viability of NV_{HH} was not observed during the expression of anti-Fib library. We have recently tested the viability of anti-EGFR library (data not shown), and found that the viability is not affected, being ~0.9x10^9 CFU/OD_{500}. Hence, some specific attributes of the anti-TirM_{EHEC} library could cause the observed toxicity in NV_{HH} fusions. In any case, the observed reduction does not have a significant effect on the representation of immune libraries with a diversity of ~10^7 clones, since an excess of input bacteria over the library size is always used during MACS (≥10^8 bacteria).

The number of antigen-specific high affinity binders obtained from each immune library could be more connected to attributes of the library itself, such as the immunogenicity of antigen, number of animals used in the immunization, or size and complexity of the library (Bradbury and Marks, 2004), rather than on the effectiveness of the surface display system used for selection. In the case of the anti-TirM_{EHEC} library, we obtained a single high affinity binder i.e. VTIR1 and four other binders i.e. VTIR2 to VTIR5, of lower affinities. However, previous selection of the same anti-TirM_{EHEC} library by phage display had failed to retrieve specific high-affinity binders other than VTIR1 isolated by *E. coli* display (data not shown). Similarly, selections of the anti-Fib and anti-EGFR libraries using *E. coli* display led to identification of a similar number of high affinity binders and essentially the same clones that were also isolated by phage display of these V_{HH} libraries (Roovers et al., 2007); and data not shown. Thus, it appears that the actual diversity of binders
retrieved from immune libraries is more dependent on intrinsic factors of the library rather than on the display system used for selection.

Abs are in huge demand for use in diagnosis and therapy and so is the need for inexpensive Ab expression systems for their selection and production. *E. coli* is an attractive expression system since it requires inexpensive media, has fast growing times, high efficiency of transformation (>10\(^{10}\) transformants/µg of plasmid DNA), well-characterized, versatile cloning and expression vectors. Although the purification of Nbs from the periplasm of *E. coli* is relatively straightforward owing to their high solubility and simplicity in structure, variable yields are commonly obtained among different clones (e.g. 0.1-0.5 mg/L of induced culture for Nbs). To alleviate this issue and increase the total yield of Nb produced, we investigated their production fused to the periplasmic maltose binding protein (MBP) of *E. coli*. In addition, we included a His\(_6\) tag in the C-terminus of the fusions to allow the use of a double affinity chromatography during the purification, with amylose and immobilized metal resins. Expression of the constructs in standard *E. coli* strains such as BL21 resulted in high levels of proteolysis, and so an *E. coli* strain deficient in periplasmic proteases i.e. HM140 was used. Soluble MBP-V\(_{\text{HH}}\)His\(_6\) fusions were consistently expressed at high levels (>12 mg/L of induced culture) in the periplasm of *E. coli*. Highly pure MBP-V\(_{\text{HH}}\)His\(_6\) fusions were recovered by amylose and immobilized metal affinity chromatography steps. Good yields of soluble V\(_{\text{HH}}\)His\(_6\) (between 2-3 mg/L of induced culture) were obtained after site-specific proteolysis of the fusions and their monomeric nature was confirmed by gel filtration. Both, MBP-V\(_{\text{HH}}\)His\(_6\) fusions and free V\(_{\text{HH}}\)His\(_6\) retained antigen binding activity and specificity (Salema and Fernandez, 2013). The anti-Fib V\(_{\text{HH}}\), i.e. V\(_{\text{FIB1}}\) purified in this manner was successfully applied in the development of a disposable magnetoimmunosensor for detection of fibrinogen in plasma (Campuzano et al., 2014).

We chose MACS to select and recover *E. coli* cells bound to the antigen since this technology does not require expensive equipment like the fluorescent cell sorter used in FACS, and multiple samples can be processed in parallel. While we used a
Discussion

Manual MACS system (holding up to 8 mini-MACS columns, each with a capacity of $\sim 10^8$ bacteria), MACS is easily scalable to columns with higher loading and capturing capacities, with the possibility of being automated, thus making screening of large Ab libraries faster and more efficient than FACS.

Although MACS is an easy and fast way to select high affinity binders from an Ab library, this method relies on the availability of purified antigen, which is not always practical or feasible. Direct selection on cells expressing the antigen(s) of interest maybe required when an antigen is produced at low yields, has low solubility, or does not maintain its native conformation upon purification. It may also be an attractive approach to identify Abs against unknown antigens expressed on the surface of target cells (Ahmadvand et al., 2009; Pavoni et al., 2014). This may be challenging for antigens with low levels of expression, but it is clearly realistic with surface antigens overexpressed under certain conditions (e.g. tumor cells). In our work we used an anti-EGFR library that was obtained by immunizing llamas with whole A431 cells (EGFR+)(Roovers et al., 2007). This library was cloned into the Intimin system and selected directly on cells using a combination of negative selection cycles on cells that did not express EGFR (NIH-3T3 2.2) and positive selection cycle on cells expressing EGFR (HER14). The selection was very effective and we were able to enrich the anti-EGFR libraries in just two rounds of selection on cells. Out of the CellS, we obtained four specific binders with good affinities to EGFR (one of which was specific to this method of selection), as validated by flow cytometry analysis of binding to these clones to recombinant EGFR antigen and BSA (control). All these clones also bound specifically to HER14 cells but not to NIH-3T3 2.2 cells and competed with EGF for binding to EGFR at different levels (as demonstrated in flow cytometry assays). Further, the $K_D$ of the two best clones (i.e $V_{EGFR1}$ and $V_{EGFR2}$) was estimated by flow cytometry and was found to be 30 nM and 12 nM respectively.

In this work we have demonstrated that *E. coli* display is suitable for use with immune libraries with $10^6$-$10^7$ clones. Nonetheless, this method is not limited to these types of libraries and could also be used with larger naïve and synthetic
libraries of $\sim 10^9$ clones (Arbabi-Ghahroudi et al., 2009; Monegal et al., 2009; Hussack et al., 2012). As mentioned previously, MACS can be scaled up with larger columns than those used in our work, hence allowing the screening of $\sim 5 \times 10^9$ bacteria/column. Several of these MACS columns of large capacity can be run in parallel covering the total diversity of libraries $\sim 10^9$-$10^{10}$ clones. In these type of MACS columns, larger volumes and densities of *E. coli* bacteria can be loaded. Importantly, *E. coli* can be manipulated with ease at a density of is $\sim 1 \times 10^{10}$ CFU/ml, which is similar to the maximum size of most naïve and synthetic libraries of Ab fragments. For large naïve and synthetic Ab libraries, phage display may appear to be more advantageous than *E. coli* display, given that bacteriophages can be produced and handled in higher densities than *E. coli* (i.e. $10^{13}$ PFU/ml). However, the diversity of phage display libraries is also strongly limited by intrinsic *E. coli* factors like transformation efficiency and the ability of phages to infect cultures with sufficient amount of bacteria. In addition, the percentage of phage particles actually displaying at least a single Ab molecule on the surface could be as low as less than 10% of the total infecting particles when conventional helper phages with wild type pIII (e.g. M13K07) are used to rescue phagemids (Rondot et al., 2001; Oh et al., 2007; Soltes et al., 2007). Hence, the actual density of phage-antibody particles is likely to be at least one order of magnitude lower than the titer of phage solutions and the actual number of clones is limited by *E. coli* transformation, efficiency, as it happens with *E. coli* display. Therefore, *E. coli* display and phage display libraries are appropriate for libraries containing $\sim 10^9$-$10^{10}$ different clones. Screening of larger Ab libraries with higher diversities will require the use of cell-free display systems (e.g. ribosome display) that avoid transformation of plasmid vectors to *E. coli* (Pluckthun, 2012; Sun et al., 2012; Kanamori et al., 2014). One of the major benefits of *E. coli* display over phage display is the use of flow cytometry for the direct determination of the expression levels, antigen-binding specificity and affinity of the selected clones. In addition, eventhough *E. coli* cells are more sensitive than bacteriophages to extremes of pH, temperature and other strong denaturants, *E. coli* bacteria having their OM intact can be washed with most common buffers as well as tolerate significant concentrations of detergents (e.g.
0.1-0.4% w/v of TX-100, SDS or deoxycholate) marking an improvement over the washing conditions tolerated by spheroplasts in APEx. The multivalent Ab display and the less sticky properties of *E. coli* cells compared to bacteriophages makes selections to complex antigenic surfaces e.g. mammalian cells, tissues and organs, with reduced background binding possible.

We have envisaged some putative biomedical applications for the antigen-specific high affinity V\textsubscript{HH}s selected from the each of the different V\textsubscript{HH} libraries. The V\textsubscript{HH} against Tir\textsubscript{EHEC}, i.e. \textit{VTIR1} is a high affinity binder and has a very slow rate of dissociation. Such a binder could be useful in recognition and binding to Tir, thus preventing the association of EHEC to the intestinal surface and further pathogenesis of the bacteria. The V\textsubscript{HH}s against Fib can be used for diagnostic applications such as detection of fibrinogen levels in the blood using ELISA or other detection methods. In this regard, we have used \textit{V\textsubscript{FIB1}} in electrochemical biosensor for the measurement of fibrinogen levels in the blood plasma (Campuzano et al., 2014). Finally, the V\textsubscript{HH}s against human EGFR need further evaluation of their ability to compete and block the EGFR binding of natural ligands such as EGF, TGF-\textalpha{} etc., thus mimicking the mode of action of some mAbs currently in therapy e.g. Erbitux or Vectibix. This would result in the switching off of signal transduction cascades that lead to uncontrolled growth and proliferation of cells, as occurring in cancer.
CONCLUSIONS
Conclusions

The following are the conclusions of the doctoral thesis:

1. The *E. coli* display systems based on the β-domains of Intimin and EhaA from EHEC represent a good alternative to phage display and other cell display platforms (e.g. yeast display), for the selection of high affinity sdAb binders from immune libraries of V_{HH}s. The ability to display and select sdAbs without the need for permeabilization of the OM, along with the lower toxicity of the Neae, EhaA fusions are advantages of the two β-domains tested over other existing *E. coli* display formats.

2. *E. coli* bacteria displaying two different V_{HH} libraries (obtained after immunization with TirM_{EHEC} or Fib antigens) using both β-domains, were successfully screened with their respective biotinylated antigens using MACS, and facilitated the isolation of specific binders with high affinity against TirM_{EHEC} and Fib. These studies also revealed that *E. coli* display with the Intimin β-domain is a more effective platform for selection of binders in fewer rounds. In addition, *E. coli* bacteria displaying V_{HH}s with Intimin β-domain have higher antigen-binding signals in flow cytometry.

3. Flow cytometry of *E. coli* bacteria displaying V_{HH} clones and libraries was used for direct determination of the expression levels, antigen binding specificity and affinity of the selected sdAbs by flow cytometry, representing a major benefit of *E. coli* display. Apparent K_D values determined by flow cytometry of *E. coli* bacteria displaying sdAbs against TirM_{EHEC} and Fib were very similar to those obtained with the purified sdAbs and state-of-the-art technologies such as SPR.

4. The multivalent nature of *E. coli* display and the less sticky properties of *E. coli* cells compared to filamentous bacteriophages, reduces background and non-specific binding and, thus makes *E. coli* display an optimal platform for selections not only with purified antigens but also on complex antigenic surfaces and cells (e.g. tumor cells). High-affinity sdAbs binding human EGFR were selected from an *E. coli* display library of V_{HH}s generated by immunization of llamas with human tumor cells overexpressing EGFR i.e. A431, using selections with either a
recombinant extracellular domain of hEGFR or murine tumor cell lines transfected with hEGFR.

5. Soluble sdAbs can be purified in high amounts from the periplasm of *E. coli* strain HM140 as N-terminal fusions to MBP. MBP-\(V_{HH}\) fusions bind their cognate antigen specifically and allow purification of the soluble \(V_{HH}\) moiety with a C-terminal His-tag after cleavage of the fusion with a site-specific protease.
Del trabajo presentado en esta tesis podemos obtener las siguientes conclusiones:

1. Los sistemas de presentación de sdAbs basados en los dominios β de intimina y EhaA de EHEC representan una buena alternativa a la presentación en fagos y otras plataformas de presentación en células (por ejemplo la presentación en levaduras), para la selección de sdAbs con alta afinidad por antígeno a partir de genotecas inmunes de V_{HH}s. La capacidad de exponer y seleccionar sdAbs sin la necesidad de permeabilizar la OM, junto con la baja toxicidad de las fusiones con Neae (intimina) o con EhaA son ventajas de estos sistemas en relación a otros formatos de presentación de sdAbs en E. coli actualmente existentes.

2. Se han seleccionado con éxito clones de alta afinidad a antígenos específicos desde genotecas de sdAbs presentados en la superficie de bacterias E. coli. Dos genotecas de V_{HH} diferentes (obtenidas después de inmunizaciones con los antígenos TirM_{EHEC} o Fib) se presentaron en la superficie de E. coli con ambos dominios β y fueron rastreadas mediante MACS con sus respectivos antígenos marcados con biotina para seleccionar clones que se unían específicamente y con gran afinidad a estos antígenos. Estos estudios también revelaron que el sistema de presentación con el dominio β de la intimina es una plataforma más eficaz para el aislamiento de clones específicos realizando pocas rondas de selección. Asimismo, las bacterias E. coli que presentaban V_{HH}s con el dominio β de la intimina presentan señales de fluorescencia más elevadas en los análisis por citometría de flujo.

3. Hemos utilizado la técnica de citometría de flujo con bacterias E. coli que presentaban clones de V_{HH} individuales y colecciones de V_{HH}s para la determinación directa de los niveles de expresión de los V_{HH}s, la especificidad de unión al antígeno y la afinidad de los sdAbs seleccionados, lo que representa una ventaja importante de los sistemas de presentación de sdAbs en E. coli. Los valores de K_{D} estimados mediante análisis por
citometría de flujo de bacterias *E. coli* que presentan sdAbs frente a Tir\textsubscript{EHEC} y Fib fueron muy similares a los valores obtenidos con sdAbs purificados aplicando tecnologías de última generación como SPR.

4. La naturaleza multivalente de la presentación de sdAbs en *E. coli* y la tendencia menor a adherirse inespecíficamente de las células de *E. coli* en comparación con bacteriófagos filamentosos, reduce la probabilidad de obtener clones inespecíficos. Por ello, la presentación de sdAbs en *E. coli* es una plataforma óptima para la selección de clones específicos, no sólo utilizando antígenos purificados sino también sobre superficies antigénicamente complejas y células (ej. células tumorales). Hemos seleccionado sdAbs de afinidad frente al EGFR humano a partir de una genoteca de sdAbs presentada en *E. coli* y generada mediante inmunización con células tumorales humanas que sobreexpresaban EGFR (ej. A431). Las selecciones se realizaron con éxito utilizando tanto un dominio extracelular de EGFR humano y una línea celular murina transfectada con EGFR humano.

5. Los sdAbs pueden ser purificados en grandes cantidades y de forma soluble como fusiones amino terminal a MBP producidas en el periplasma de la cepa de *E. coli* HM140 deficiente en proteasas periplásmicas. Las fusiones MBP-V\textsubscript{HH} se unen específicamente a sus antígenos respectivos y permiten la purificación del dominio V\textsubscript{HH} con una cola de histidinas (His\textsubscript{6}) en su extremo C-terminal, tras la escisión de la MBP utilizando una proteasa específica de sitio.


surface display of conformationally constrained peptides. Protein engineering 12, 797-806.


Fairman, J.W., Dautin, N., Wojtowicz, D., Liu, W., Noinaj, N., Barnard, T.J., Udho, E., Przytycka, T.M., Cherezov, V. and Buchanan, S.K., 2012, Crystal structures of the outer membrane domain of intimin and invasin from enterohemorrhagic E. coli and


References


ANNEXURES