Engineering functional recombinant proteins based on antibody domains and fragments of C. botulinum neurotoxin

Jolanta E Tolkacz (2010)

Oxford Brookes University

School of Life Sciences

Engineering functional recombinant proteins based on antibody domains and fragments of *C. botulinum* neurotoxin

Jolanta Eliza Tolkacz

A thesis submitted in partial fulfilment of the requirements of Oxford Brookes University for the Degree of Master of Philosophy

December 2010

Supported by Syntaxin Ltd, Abingdon

Oxford Brookes University

Syntaxin Limited

ABSTRACT

SCHOOL OF LIFE SCIENCES

Master of Philosophy

ENGINEERING FUNCTIONAL RECOMBINANT PROTEINS BASED ON ANTIBODY DOMAINS AND FRAGMENTS OF *C. botulinum* NEUROTOXIN

By Jolanta Eliza Tolkacz

Genetic engineering allows the construction of many diverse antibody-related molecules. The focus of this study is single chain Fv (Fragment variable) and single domain antibodies. Recombinant botulinum neurotoxin fragments derived from serotype A, B, C and D are selected as backbones for an investigation into the feasibility and properties of novel recombinant targeted multidomain proteins.

Single chain Fv's consist of a V_H region of heavy chain linked by a stretch of synthetic peptide to a V_L region of the light chain. Fv is the region for binding to antigens as determined by immunoglobulin Ig hypervariable domains called complementary-determining regions (CDR'S). The selected scFv binds to CD117 that is a membrane tyrosine kinase receptor present on the surface of various tumour cells. The variable part of chains shows great variability in amino acid sequence among the chains. The unique sequence of amino acid residues for variable fragments leads to the large diversity of structure, which counts for antibody specificity.

For the creation of single domain antibodies (sdAbs), where the antibody domain is engineered to comprise only the variable heavy or variable light chain (and therefore consist of three CDR's), six of the anti-EGFR sdAbs have been chosen in order to obtain variety of data.

All botulinum neurotoxin backbones include an activation site, which is cleaved by a protease such as EK (Enterokinase) or FXa (Factor Xa) at specific residues within the sequence. The contrast in molecule activation will distinguish the best candidates.

It was decided that targeting the tyrosine kinase receptor family was most suited to this research because (i) there is significant knowledge in the literature regarding receptor biology; (ii) ligands such as epidermal growth factor (EGF) which bind to EGFR were of importance to Syntaxin Ltd, (iii) growth factors (including EGF) are involved in tumour maintenance and growth, therefore many tumour cells express high levels of EGFR; and (iv) antibodies to EGFR are promising leads in cancer biology. There are now 26 successfully cloned clostridial endopeptidase-antibody fragment/ antibody domain constructs and six full-length high quality proteins available for further investigation.

"A fact is a simple statement that everyone believes.

It is innocent, unless found guilty.

A hypothesis is a novel suggestion that no one wants to believe.

It is guilty, until found effective. "

Edward Teller

To my parents, a special mention for all of their great encouragement during my study.

Acknowledgements

I would like to thank my supervisors Dr. John Chaddock and Professor Chris Hawes for their help and guidance during the last three years, and Syntaxin Limited for funding this research.

I would like to thank members of the research group at Syntaxin limited with particular thanks to Dr Aimee Cossins who assisted greatly in establishing this project, Dr Ian Birch-Machin, Dr Matthew Beard, Dr Teena Puri and David Burgin for guidance during my study.

I would like to thank Sarah Cooper, Malgorzata Olszowka and Dr Helen Ludlow for their technical assistance.

I would also like to thank Professor Ravi Acharya for initiating crystallographic studies at Bath University.

Finally, I would like to thank all my friends and family who have supported me through this work and believed in me.

Post viva note

Thank you to both examiners Professor Lynne Roberts and Dr David Meredith for the valuable remarks.

Table of Contents

ABSTRACT 2
Acknowledgements 4
Table of Contents 5
Table of Figures 8
Content of Tables 13
Abbreviations 14
Chapter 1. Introduction 18
1.1 Syntaxin Ltd. Strategy for recombinant chimera construction 18
1.2 Botulinum neurotoxin 19
1.2.1 Discovery of Botulinum neurotoxin action and its application 20
1.2.2 Biochemical structure of BoNT 22
1.2.3 Process of Intoxication 25
1.2.4 SNARE complex 29
1.3 The Immunoglobulin 32
1.3.1 Antibody structure 32
1.3.2 Immunoglobulins and their fragments 35
1.4 Therapeutic use of clostridia neurotoxin fragments 41
1.5 Chimeric protein conjugates with toxin 41
1.6 Functionality of novel molecules 44
1.7 Aims of the thesis 45
Chapter 2. Materials and Methods 48
2.1. Chemicals 48
2.2 Design of recombinant antibody based constructs 48
2.2.1 Antibody selection 48
2.2.1.1 Single chain antibody fragments 48
2.2.1.2 Single domain antibodies 49
2.2.2 Antibody design and optimisation 50
2.2.2.1 Back-translation of amino acids to DNA 50
2.2.2.2 Optimisation of DNA 51
2.2.3 Gene synthesis 52
2.2.4 Archive of newly synthesised DNA 53
2.2.5 DNA stock and microbanks 53
2.2.5.1 Microbank clone stocks 53
2.2.5.2 DNA clone stocks 54
2.2.5.3 Transformation to TOP10 cells 54
2.3 Cloning of antibody-LH_N chimeras 55
2.3.1 Restriction digest 55
2.3.2 Agarose gel electrophoresis 56
2.3.2.1 Agarose gel preparation 56
2.3.2.2 Sample preparation 56
2.3.2.3 1kb DNA Ladder preparation 57

2.3.2.4 Agarose gel run *57*
2.3.2.5 The extraction and purification of DNA from the agarose gel in TAE buffer *58*
2.3.3 Ligation *58*
2.3.4 Grow overnight and DNA purification of newly designed molecules *58*
2.3.5 The screening test for purified DNA *61*
2.3.6 DNA sequencing *61*
2. 4 Expression of recombinant antibody based constructs 62
2.4.1 Modified Terrific Broth *62*
2.4.2 Expression strain stock *63*
2.4.3 Standard expression protocol *63*
2.4.4 Changes applied to the standard expression protocol for purpose of study *64*
2.5 Purification of recombinant antibody based constructs 64
2.5.1 Cleavage enzymes *65*
2.5.2 Metal Ion affinity chromatography (IMAC) using AktaExpress *67*
2.5.3 Metal Ion affinity chromatography (IMAC) using AktaPurifier *68*
2.5.4 Hydrophobic Interaction Chromatography (HIC) using AktaPurifier *69*
2.5.5 Anion Exchange Chromatography – Q Sepharose FF *69*
2.5.6 Protein Analysis *70*
2.5.6.1 SDS Polyacrylamide Gel Electrophoresis (SDS-PAGE) *70*
2.5.6.2 Spectrophotometric determination of protein concentration *71*
2.5.6.4 Western blot system *73*
2.5.6.4.1 Electroblotting procedure *74*
2.5.6.4.2 Immunodetection *74*
2.5.6.5 N-terminal protein sequencing *75*
2.6 Activation studies *76*
2.6.1 Reagents *76*
2.6.2 Preparation of EGF *76*
2.6.3 Dilution of EGF in media *76*
2.6.4 Preparation of plates for activation experiment *76*
2.6.5.1 BCA assay *78*
2.6.5.2 Western blots *78*
Chapter 3. Design of recombinant antibody based constructs 81
3.1 Single chain variable fragments 81
3.1.2 Single chain variable fragment antibody structure *82*
3.1.3 Selection of scFv *83*
3.1.4 Results *84*
3.1.4.1 Choice of light and heavy chain variable regions *84*
3.1.4.1.1 The kappa light chain of SR-1 IgG *84*
3.1.4.1.2 The heavy chain of SR-1 IgG *85*
3.1.4.2 Design of scFVCD117I *85*
3.1.4.3 Optimisation of DNA *89*
3.1.4.3.1 Each triplet position vs. usage table *89*
3.1.4.3.2 Each codon vs. usage table *89*
3.1.4.4 Back-translated and optimised DNA *90*
3.1.4.4.1 Back-translated *90*
3.1.4.4.2 Optimized DNA *91*

3.2 Single domain antibodies 91
3.2.1 Single domain antibodies..... 92
3.2.3 Selection of sdAbs..... 95
3.2.4.1 Design of single domain antibodies 97
Chapter 4. Cloning of recombinant antibody based constructs 104
4.1 Re-engineering botulinum neurotoxin 104
4.2 Results 108
Chapter 5. Expression of recombinant antibody-based constructs 129
5.1 Overview of expression system 129
5.2 Escherichia coli expression system 131
5.3 pET system 134
5.4 Results 135
Chapter 6. Purification of recombinant antibody based constructs..... 140
6.1 Overview of protein purification..... 140
6.2 Introduction to the purification of BoNT 141
6.3 Purification of recombinant antibody based constructs 143
6.4 Results 144
6.4.1 Purification of single chain variable antibody fragments-LH_N chimera ... 145
6.4.2 Purification of single domain antibodies-LH_N chimera..... 162
Chapter 7. Activation Studies 180
7.1 Principle of the activation study 180
7.2 Ligand binding 180
7.3 Receptor Tyrosine Kinases 181
7.4 Epidermal growth factor (EGF) 183
7.5 Results 184
Chapter 8. Internalisation Studies 188
8.1 Endocytosis 188
8.2 Exocytosis 190
8.3.1 Internalization assay..... 192
Chapter 9. Conclusions 197
Chapter 10.Discussion 201
Appendix 206
References..... 247

Table of Figures

Figure 1 Protein domain structure of recombinant botulinum neurotoxin 18
Figure 2 Tertiary structure of botulinum neurotoxin type A (3BTA) (adapted from Turton, Chaddock and Acharya, 2002)..... 23
Figure 3 Clostridial Endopeptidase: Building Blocks for Engineering (adapted from Foster, 2004) 24
Figure 4 Model for the entry of BoNTs into nerve cells (adapted from Binz and Rummel, 2009). 26
Figure 5 Stages of botulinum neurotoxin action: Inhibiting cell secretion (adapted from Turton et. al., 2002). 28
Figure 6 Illustrating the disulphide bonds (red) that link the light (green) and heavy (blue) protein subunits of Immunoglobulin G (IgG) molecules. 33
Figure 7 The position of six CDRs in the light and heavy chain of an antibody. 34
Figure 8 Main classes of immunoglobulin IgG, IgA, IgM, IgD and IgE. 37
Figure 9 Schematic of Ig domains (A); Use of pepsin and papain to generate $F(ab')_2$ and Fab fragment (B) respectively. 39
Figure 10 Various Antibodies and derived fragments from them (adapted from Holt, 2003). 40
Figure 11 Antibodies as Cancer Therapeutics. 43
Figure 12 Sequences of single domain antibodies. 50
Figure 13 Codon Usage in *E. coli* genes. 52
Figure 14 1 kb DNA Ladder visualized by ethidium bromide staining on a 0.8% agarose gel. Mass values are for 0.5 µg/lane. 57
Figure 15 Flow diagram of plasmid DNA isolation and purification using the Wizard® Plus SV Miniprep DNA Purification System. 60
Figure 16 Protein Standards for electrophoresis. 71
Figure 17 Western blot system. 73
Figure 18 A single chain variable fragments (scFv) molecule. 82
Figure 19 The most common SNAP-ON restriction enzymes within LHN backbone. 88
Figure 20 Schematic diagram of the VHH domain of a camelid heavy chain antibody and disulphide bond framework regions (adapted from Wesolowski *et al.,* 2009) 94
Figure 21 The di-chain structure of a Clostridial neurotoxin –botulinum neurotoxin A (BoNT/A). (Adapted from Turton *et. al.*, 2002) 105
Figure 22 LH_N domain of botulinum neurotoxin type C and epidermal growth factor (EGF) (adapted from Foster, 2005). 106
Figure 23 Design of LH_N fragment with ligand. 107
Figure 24 Design of LH_N fragment with antibody. 107
Figure 25 Design of LH_N with single chain antibody fragment. 108
Figure 26 Design of LH_N with single domain antibody. 109
Figure 27 Design of LH_N/B fragment with activation side EK (Enterokinase) in SXN101327. 110
Figure 28 Design of LH_N/A fragment with activation side EK (Enterokinase) in SXN100532. 110
Figure 29 Design of LH_N/D fragment with activation side by EK (Enterokinase) in SXN101832. 110

Figure 30 Design of LH_N/B fragment with activation side FXa (factor Xa) in SXN101641. 111
Figure 31 Design of LH_N/C fragment with activation side FXa (factor Xa) in SXN100736. 111
Figure 32 Cloning of DNAs encoding single variable antibody fragments (scFvs) and the single domain antibodies (sdAbs) with restriction endonucleases: *XbaI* on the (5′ end) and *HindIII* on the (3′ end) into modified pET25b vector. 112
Figure 33 Modified pET26b vector with inserted scFv or sdAb. 113
Figure 34 DNA sequencing of LH_N/B-EGFR6sdAb=SXN101861. 115
Figure 35 Linear map (SeqBuilder) of molecule LN_N/A-EGFR7sdAb=SXN101869. 116
Figure 36 0.8% agarose gel of purified cut plasmid DNA. 117
Figure 37 Products obtained following *HindIII* and *HindIII/XbaI* digests on plasmids LHN/AscFvCD117I=SXN101623 and LHN/BscFvCD117I=SXN101624 sent to Geneservice for sequencing. 118
Figure 38 0.8% agarose gel of purified uncut and cut plasmid DNA of scFvCD117I. 119
Figure 39 0.8% agarose gel of purified uncut and cut plasmid DNA of SXN101641. 119
Figure 40 Product obtained following *HindIII* and *HindIII/XbaI* digest on plasmid LH_N/BscFvCD117I=SXN101776 sent to Geneservice for sequencing. 120
Figure 41 0.8% agarose gel of purified uncut plasmid DNA of sdAbs. 120
Figure 42 0.8% agarose gel of purified cut plasmid DNA of sdAbs. 121
Figure 43 0.8% agarose gel of purified cut plasmid DNA. 122
Figure 44 Products obtained following *HindIII* and *HindIII/PstI* digests on LH_N/sdAbs plasmids sent to Geneservice for sequencing. 123
Figure 45 Products obtained following *HindIII* and *HindIII/PstI* digests on LH_N/sdAbs plasmids sent to Geneservice for sequencing. 124
Figure 46 Products obtained following *HindIII* and *HindIII/PstI* digests on LH_N/sdAb plasmid and LH_N/C-scFvCD117I=SXN101882 sent to Geneservice for sequencing. 125
Figure 47 Structure of *Escherichia coli* cell envelope. (adapted from Weisser and Hall, 2009) 132
Figure 48 Protein production in pET expression system. 134
Figure 49 SafeStain NuPage 4-12% Bis-Tris gel (A) and Western blot against light chain (B) showing expression of LH_N/A with a single domain antibody. 136
Figure 50 SafeStain NuPage 4-12% Bis-Tris gel (A) and Western blot against light chain (B) showing expression of LH_N/B with a single domain antibody. 136
Figure 51 SafeStain NuPage 4-12% Bis-Tris gel (A) and Western blot against light chain (B) showing expression of LH_N/D with a single domain antibody. 137
Figure 52 SafeStain NuPage 4-12% Bis-Tris gel showing expression of LH_N/C with a single chain antibody fragment. 138
Figure 53 SafeStain NuPage 4-12% Bis-Tris gel (A) and Western blot against light 138
Figure 54 NuPage 4-12% Bis-Tris gel of LH_N/B-scFvCD117I=SXN101624 produced from a nickel column. 145
Figure 55 NuPage 4-12% Bis-Tris gel of LH_N/B-scFvCD117I=SXN101624 produced from a cleavage process. 147
Figure 56 Quality control of the target protein LH_N/B-scFvCD117I=SXN101624 represent on the NuPage 4-12% Bis-Tris gel (A) and confirmed by Western blot against light chain (B). 148
Figure 57 NuPAGE 4-12% Bis-Tris gel of LH_N/B-scFvCD117I=SXN101776 produced from a nickel column. 149

Figure 58 NuPAGE 4-12% Bis-Tris gel of LH_N/B-scFvCD117I=SXN101776 cleaved by FXa. 150
Figure 59 Quality control of the target protein LH_N/B-scFvCD117I=SXN101776 represent on the NuPage 4-12% Bis-Tris gel (A) and confirmed by Western blot against His tag (B) and light chain of B (C). 152
Figure 60 NuPAGE 4-12% Bis-Tris gel of LH_N/B-scFvCD117I=SXN101776 cleaved by FXa New England BioLabs and R&D System. 153
Figure 61 NuPage 4-12% Bis-Tris gel of LH_N/B-scFvCD117I=SXN101776 produced from a Q Sepharose column. 154
Figure 62 Quality control of the target protein LH_N/B-scFvCD117I=SXN101776 represent on the NuPage 4-12% Bis-Tris gel (A) and confirmed by Western blot against His tag (B) and light chain of B (C). 155
Figure 63 NuPage 4-12% Bis-Tris gel of LH_N/B-scFvCD117I=SXN101776 produced from a Q Sepharose column as first purification step. 156
Figure 64 NuPage 4-12 % Bis-Tris gel of LH_N/B-scFvCD117I=SXN101776 produced from a (HIC) Hydrophobic Interaction Chromatography column as first purification step. 157
Figure 65 NuPAGE 4-12 % Bis-Tris gel of LH_N/C-scFvCD117I=SXN101882 produced from a nickel column. 158
Figure 66 NuPAGE 4-12 % Bis-Tris gel of LH_N/C-scFvCD117I=SXN101882 cleaved by FXa. 158
Figure 67 Quality control of the target protein LH_N/C-scFvCD117I=SXN101882 represent on the NuPage 4-12% Bis-Tris gel (A) and confirmed by Western blot against His tag (B) and light chain of C (C). 159
Figure 68 Protein sequence report from Alta Bioscience. 160
Figure 69 Protein sequence report from Alta Bioscience. 161
Figure 70 Protein sequence report from Alta Bioscience. 161
Figure 71 Protein sequence report from Alta Bioscience. 161
Figure 72 NuPage 4-12% Bis-Tris gel of LH_N/B-EGFR31sdAb=SXN101865 produced from a nickel column. 163
Figure 73 NuPage 4-12% Bis-Tris gel of LH_N/B-EGFR6sdAb=SXN101861 produced from a nickel column. 163
Figure 74 NuPage 4-12% Bis-Tris gel of LH_N/B-EGFR31sdAb=SXN101865 cleaved by FXa. 164
Figure 75 NuPage 4-12% Bis-Tris gel of LH_N/B-EGFR6sdAb=SXN101861 cleaved by FXa. 164
Figure 76 Quality control of the target protein LH_N/B-EGFR31sdAb=SXN101865 represent on the NuPage 4-12% Bis-Tris gel (A) and confirmed by Western blot against His tag (B) and light chain of B (C). 165
Figure 77 Quality control of the target protein LH_N/B-EGFR6sdAb=SXN101861 represent on the NuPage 4-12% Bis-Tris gel. 166
Figure 78 NuPage 4-12% Bis-Tris gel of LH_N/B-EGFR43sdAb=SXN101866 produced from a nickel column. 167
Figure 79 NuPage 4-12% Bis-Tris gel of LH_N/B-EGFR43sdAb=SXN101866 cleaved by FXa. 167
Figure 80 NuPage 4-12% Bis-Tris gel of LH_N/A-EGFR6sdAb=SXN101868 produced from a nickel column. 168
Figure 81 NuPage 4-12% Bis-Tris gel of LH_N/A-EGFR6sdAb=SXN101868 cleaved by EK. 169

Figure 82 NuPage 4-12% Bis-Tris gel of LH_N/A-EGFR31sdAb=SXN101872 produced from a nickel column. 169
Figure 83 NuPage 4-12% Bis-Tris gel of LH_N/A-EGFR31sdAb=SXN101872 cleaved by EK. 170
Figure 84 NuPage 4-12% Bis-Tris gels of LH_N/A-EGFR31sdAb=SXN101872 produced from a HIC column. 171
Figure 85 Quality control of the target protein LH_N/A-EGFR6sdAb=SXN101868 represent on the NuPage 4-12% Bis-Tris gel (A) and confirmed by Western blot against His tag (B) and light chain of A (C). 172
Figure 86 Quality control of the target protein LH_N/A-EGFR31sdAb=SXN101872 represent on the NuPage 4-12 % Bis-Tris gel. 173
Figure 87 NuPage 4-12% Bis-Tris gel of LH_N/D-EGFR28sdAb=SXN101877 produced from a nickel column. 174
Figure 88 NuPage 4-12% Bis-Tris gel of L_N/D-EGFR31sdAb=SXN101878 produced from a nickel column. 174
Figure 89 NuPage 4-12% Bis-Tris gel of LH_N/D-EGFR28sdAb=SXN101877 cleaved by EK. 175
Figure 90 NuPage 4-12% Bis-Tris gel of LH_N/D-EGFR31sdAb=SXN101878 cleaved by EK. 175
Figure 91 Quality control of the target protein LH_N/D-EGFR28sdAb=SXN101877 represent on the NuPage 4-12% Bis-Tris gel. 176
Figure 92 Quality control of the target protein LH_N/D-EGFR31sdAb=SXN101878 represent on the NuPage 4-12 % Bis-Tris gel. 177
Figure 93 The key components of the MAPK/ERK pathway (adapted from JWSchmidt). 182
Figure 94 Western blot against β actin. 185
Figure 95 Western blot against MAPK. 185
Figure 96 Western blot against phosphorylated MAPK. 186
Figure 97 Western blot against EGFR. 186
Figure 98 Endocytosis pathways (adapted from Cell Biology, Wikipedia 2009). 188
Figure 99 Pathway of Clathrin depended Endocytosis (adapted from Grant and Sato, 2005). 190
Figure 100 Secretion SNARE mediated fusion. 191
Figure 101 Method for internalization assay. 193
Figure 102 Internalization of biotinylated EGF in A549 cells. 194
Figure 103 Internalization of biotinylated LHC-EGF in A549 cells. 195
Figure 104 Each triplet position usage table sdAb EGF28. 207
Figure 105 Each codon usage table sdAb EGF28. 208
Figure 106 Each triplet position usage table sdAb EGF28 final adjustments. 209
Figure 107 Each codon usage table sdAb EGF28 final adjustments. 210
Figure 108 Each triplet position usage table sdAb EGF10. 211
Figure 109 Each codon usage table sdAb EGF10. 212
Figure 110 Each triplet position usage table sdAb EGF10 final adjustments. 213
Figure 111 Each codon usage table sdAb EGF10 final adjustments. 214
Figure 112 Each triplet position usage table sdAb EGF6. 215
Figure 113 Each codon usage table sdAb EGF6. 216
Figure 114 Each triplet position usage table sdAb EGF6 final adjustments. 217
Figure 115 Each codon usage table sdAb EGF6 final adjustments. 218

Figure 116 Each triplet position usage table sdAb EGF7. 219
Figure 117 Each triplet position usage table sdAb EGF7 final adjustments. 220
Figure 118 Each codon usage table sdAb EGF7 final adjustments. 221
Figure 119 Each triplet position usage table sdAb EGF43. 222
Figure 120 Each codon usage table sdAb EGF43. 223
Figure 121 Each triplet position usage table sdAb EGF43 final adjustments. 224
Figure 122 Each codon usage table sdAb EGF43 final adjustments. 225
Figure 123 Each triplet position usage table sdAb EGF31. 226
Figure 124 Each codon usage table sdAb EGF31. 227
Figure 125 Each triplet position usage table sdAb EGF31 final adjustments. 228
Figure 126 Each codon usage table sdAb EGF31 final adjustments. 229
Figure 127 Each triplet position usage table scFv CD117I. 231
Figure 128 Each codon usage table scFv CD117I. 232
Figure 129 Each triplet position usage table scFv CD117I final adjustments. 234
Figure 130 Each codon usage table scFv CD117I final adjustments. 235
Figure 131 Entelechon report for single domain antibodies 236
Figure 132 Entelechon report for single domain antibodies. 237
Figure 133 Entelechon report for single chain antibody fragment. 238
Figure 134 pCR4TOPO vector for cloned molecules by Entelechon. 239
Figure 135 Entelechon CD117I gene report with quality control. 240
Figure 136 Entelechon EG6 gene report and quality control. 241
Figure 137 Entelechon EG7 gene report and quality control. 242
Figure 138 Entelechon EG10 gene report and quality control. 243
Figure 139 Entelechon EG28 gene report and quality control. 244
Figure 140 Entelechon E31 gene report and quality control. 245
Figure 141 Entelechon EG43 gene report and quality control. 246

Content of Tables

Table 1 Serotypes and SNARE cleavage.....29

Table 2 WO/2007/127317 Patent data for CD117I and CD117I.....47

Table 3 Reaction mixture for double digestion (50 µl).....53

Table 4 Reaction mixture for ligation (20µl).....56

Table 5 Reaction mixture for sequential digest.....58

Table 6 Reaction mixture for double digest.....58

Table 7 mTB composition.....59

Table 8 Activation enzymes.....63

Table 9 SNAP ON Restriction enzymes in LH_N open reading frame.....86

Table 10 The full-length antibody-LH_N chimeras cloned for purpose of this study.....123

Table 11 LH_N control chimeras available at Syntaxin Ltd.....124

Table 12 The full-length fusion proteins of LH_N antibody based chimeras obtained in this study.....177

Abbreviations

aa	Amino acid
Ab	Antibody
A_X	Absorbance at wavelength X nm
BBB	blood-brain barrier
BCA	Bicinchoninic acid
Bis-scFv	Bispecific Single-chain variable fragment
BLE	blepharospasm
BoNT	Botulinum neurotoxin
BoNT/A	Botulinum neurotoxin serotype A
BoNT/B	Botulinum neurotoxin serotype B
BSA	Bovine serum albumin
BTX-A	Botulinum toxin type A
C	Protein concentration
CAPS	3-(Cyclohexylamino)-1-propanesulfonic acid
CDR	complementary-determining region
CDR-H3	Variable antibody hypervariable loop of heavy chain
CDR-L3	Variable antibody hypervariable loop of light chain
C_H	Constant domain of immunoglobulin heavy chain
C-kit	Cytokine receptor (CD117)
C_L	Constant domain of immunoglobulin light chain
dH_2O	Distilled water
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
E. coli	Escherichia coli
ECL	*Erythrina cristagalli* lectin
EDTA	Ethylenediamineteraacetic acid
EGF	Epidermal growth factor
EGFR	Epidermal Growth factor receptor
EGFR	epidermal growth factor receptor
EGTA	Ethylene Glycol-bis(β-aminoethyl ether)
EK	Enterokinase
Fab	Antigen binding fragment
FBS	Fetal Bovine Serum
Fc	Crystalline fragment
FDA	Food and Drug administration
Fv	Fragment variable
Fwd	Forward
FXa	Factor Xa
GST	glutathione S-transferase
H_2L	fragment of BoNT comprising the light chain and N-terminal domain of the heavy chain
H_c	Heavy chain
hcAbs	heavy chain antibodies
hcAbs	Heavy chain antibodies

H_{CC}	Heavy chain C-terminal ganglioside binding domain
H_{CN}	Heavy chain N-terminal translocation domain
HIC	Hydrophobic interaction chromatography
His tags	Histidine tag
HRP	Horseradish Peroxidase
Ig	Immunoglobulin
IgA	Immunoglobulin A
IgD	Immunoglobulin D
IgE	Immunoglobulin E
IgG	Immunoglobulin G
IgM	Immunoglobulin M
IM	Inner membrane
IMAC	Immobilized metal affinity chromatograph
IPTG	Isopropyl-B-galactosidase
K. lactis	Kluyveromyces Lactis
Kan	Kanamycin
K_{DS}	dissociation constants
L	Path length
LB	Luria Bertani
LC	Light Chain
LDS	Lithium Dodecyl Sulphate
LHA	Modified recombinant Botulinum neurotoxin serotype A
LHB	Modified recombinant Botulinum neurotoxin serotype B
LHD	Modified recombinant Botulinum neurotoxin serotype D
LHn	Modified recombinant Botulinum neurotoxin
mAb	monoclonal antibodies
MAPK/ ERK	Mitogen activated protein kinase
MBP	maltose-binding protein
MEK	Mitogen-activated protein kinase
mTB	modified Terrific Broth
MW	Molecular weight
NAD	Nicotinamide adenine dinucleotide
NEB	New England Bioscience
NMJ	Neuromuscular junction
OD_X	Optical density at wavelength X nm
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered Saline
PVDF	Polyvinylidene fluoride
Qa	syntaxins
Qb	SNAP N-terminal
QC	Quality Control
Qc	SNAP C-terminal
rAb	recombinant antibody
RAF	Proto-oncogene serine/threonine-protein kinase
RAS	Abbreviation of Rat Sarcoma (protein subfamily of GTPases
Rev	Reverse
RME	Receptor-mediated endocytosis
RNA	Ribonucleic acid

RTK	tyrosine kinase receptors
S. cerevisiae	Saccharomyces cerevisiae
ScFv	Single-chain variable fragment
SdAbs	Single domain antibodies
SDS	Sodium dodecyl sulphate
SNAP	synaptosomal – associated protein
SNAP-ON	set of restriction enzymes at Syntaxin Ltd.
SNARE	soluble N-ethylmaleimide – sensitive factor attachment protein
SV	synaptic vesicle
TAE	Tris-Acetate-EDTA
Tris	Tris (hydroxymethyl) aminomethane
TSI	Targeted secretion inhibitor
t-SNAREs	SNARE located in the plasma membrane target compartment
Tween	Polyoxyethylenesorbitan monolaurate
UV	Ultra violet
VAMP	vesicle-associated membrane protein
V_H	Variable domain of immunoglobulin heavy chain
VHH	Heavy chain antibody (Nanobody)
V_L	Variable domain of immunoglobulin light chain
v-SNAREs	SNARE located in the vesicle
α -chain	Alpha heavy chain of immunoglobulin
γ -chain	Gamma heavy chain of immunoglobulin
δ -chain	Delta heavy chain of immunoglobulin
ε –chain	Epsilon heavy chain of immunoglobulin
ε	Molar extinction coefficient
κ -chain	Kappa light chain of immunoglobulin
λ -chain	Lambda light chain of immunoglobulin
μ –chain	Mu heavy chain of immunoglobulin

Chapter 1

Introduction

Chapter 1. Introduction

The main aspect of this research is to engineer novel recombinant proteins that will deliver fragments of botulinum neurotoxin to cells of therapeutic interest. The application of this research will be to increase the efficacy of biotherapeutics, for example novel agents for the treatment of solid cancers.

The focus of this research is to create antibody-botulinum toxin molecules using publicly available DNA sequences for the targeting antibodies. These novel antibody-botulinum toxin molecules are expressed in an *E. coli* expression system and purified using affinity chromatography. The antibody-botulinum toxin molecules were applied to cells to determine the functionality of the component domains, for example receptor activation and internalization of the warhead.

1.1 Syntaxin Ltd. Strategy for recombinant chimera construction

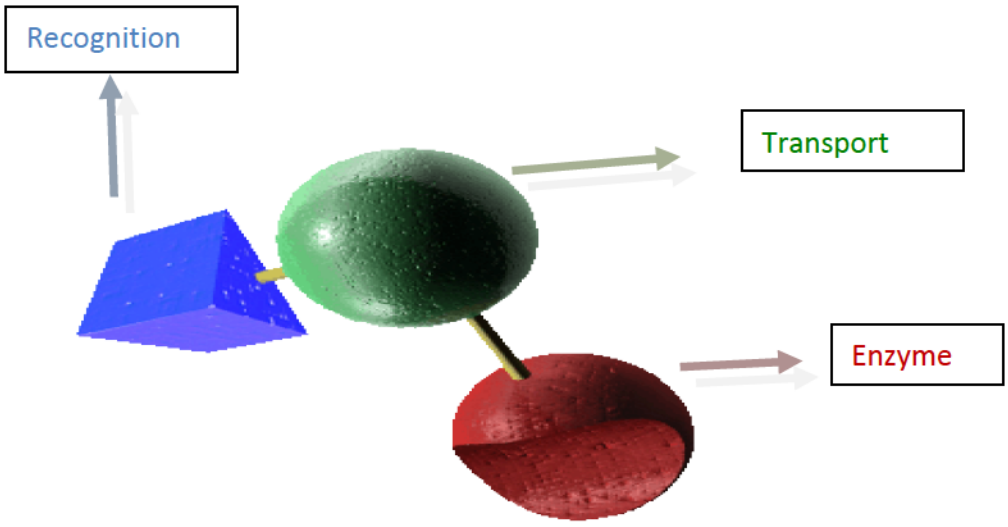


Figure 1 Protein domain structure of recombinant botulinum neurotoxin

The replacement of heavy chain (H_C) domain with alternative binding ligands is a possibility for targeting the botulinum neurotoxin endopeptidases to new target cells. This approach forms the basis of Syntaxin Ltd's strategy for development of novel recombinant therapeutics (Foster, 2004) (Fig. 1).

In furtherance to this idea, there are many potential avenues to engineer new molecules for therapeutic use; such as vaccine components, modification of intracellular process by delivery of enzymes into cells, manipulation of cell interactions with the extracellular environment and inhibition of secretion for extended periods within specific cells (Chaddock and Marks, 2006).

The use of a targeting monoclonal antibody fused to the tetanus toxin light chain to block transmitter release from targeted neurones provides a useful approach for inducible and reversible control of synaptic transmission in specific neuronal types in the brain. (Kobayashi *et*. al., 2008). However, there is no evidence for the similar production of recombinant BoNT re-targeted with small antibody fragments. The antibody-botulinum toxin conjugates could potentially bind to relevant antigens that would result in efficient internalisation into the cells and suppression of transmitter release from the target neurons via the proteolytic cleavage of protein from the SNARE complex. Used within the area of biotherapeutics this could increase the efficacy of biopharmaceuticals for example in the treatment of solid cancers. The novel molecules could improve accessibility with effective transport across the cell barrier, block specific signal transduction receptors to prevent cancer cells from growing or dividing and promote a local immune response known as an antibody-dependent cell-mediated cytotoxicity against the tumour.

The re-targeting of recombinant botulinum neurotoxin with antibodies and antibody fragments is the focus of this study. Therefore the characteristics of both antibody and botulinum neurotoxin features will be discussed in detail.

1.2 Botulinum neurotoxin

Botulinum neurotoxin (BoNT) is the most potent neurotoxic protein known to man and is produced by strains of neurotoxigenic clostridia, including *Clostridium botulinum*, and more rarely by strains of *Clostridium baratti*, *Clostridium butyricum* and *Clostridium argentinense*. The toxin produced by a number of species of *Clostridia* has been classified into seven serologically different Botulinum neurotoxins (BoNTs) named A, B, C, D, E, F and G. All of these BoNTs differ from one to another in the specific intracellular target proteins that they target, their characteristics of activation

and their potencies. However, they cannot be distinguished from each other based upon the toxins inhibition of acetycholine release from nerve terminals and their primary structure (Dressler *et*. al., 2005).

1.2.1 Discovery of Botulinum neurotoxin action and its application

Botulinum neurotoxin causes botulism, which is characterised by flaccid paralysis leading to suffocation and death in severe cases when no adequate treatment is administered (Johnson, 2005).

Human botulism occurs in three forms; food borne, wound and intestinal. Based upon a number of historical sources food borne botulism was the most prevalent disease before the 19th Century.

Medical officer Justinus Kerner published the first well-documented descriptions of food borne botulism symptoms caused by botulinum toxin (called, sausage poisoning') which occurred between 1817 and 1822 in Southern Germany. He is the pioneer of the idea to use botulinum neurotoxin as a therapeutic agent in neurologic diseases because it paralyzed skeletal muscle (Erbguth and Naumann, 1999).

In 1895, Emile Pierre van Ermengem discovered the pathogen of Botulinum toxin during an investigation of a botulism outbreak after a funeral dinner with smoked ham. He proposed the name *Bacillus botulinus* after the Latin *botulus* (sausage). Just over 30 years later in 1928, P. Tessmer Snipe and Hermann Sommer purified the toxin for the first time. This was followed, in 1949, by the discovery by Burgen's group that botulinum toxin blocked neuromuscular transmission.

The engineering of BoNT as a therapeutic, using the toxin's ability to relax tense muscles, originates in the work of two researchers. In the late 1960s, Alan Scott and Edward Schantz were the first to work on a standardized botulinum toxin preparation for therapeutic purposes (Crouch, 2006).

Alan Scott was an ophthalmologist who, in 1973, pioneered the preclinical use of botulinum toxin type A (BTX-A) in experiments on monkeys and in 1980 the subsequent clinical use in patients with strabismus (crossed eye) and blepharospasm (uncontrollable blinking) using the crystallized botulinum neurotoxin type A produced by Edward Schantz, a protein chemist (Jankovic, 2004). Scott proposed that botulinum

toxin could be used as a denervating agent in humans for the treatment of many conditions involving neuromuscular activity. Treatment of strabismus was the first therapeutic application of botulinum toxin developed by Alan Scott. He successfully demonstrated that the loss of vision associated with strabismus did not occur in any cases during 22 years of treatment (Crouch, 2006).

In the same period, Drachman's studies at John Hopkins University using material provided by Schantz provided the supportive data that botulinum toxin A induced selective denervation, muscle weakening and muscle atrophy in developing chicken embryos. The experiments indicated that botulinum toxin did not have a generalized toxin effect on cardiac muscles or a cytotoxic effect. In 1971, Drachman proposed botulinum toxin as a model of an "ideal (nerve) blocking agent." He emphasised the following criteria for the blocking agent:

(i) mode of action in selectively blocking cholinergic transmission,

(ii) specificity in blocking only cholinergic transmission,

(iii) reversibility in not permanently impairing function or structure of nerve or muscle,

(iv) generality of action in blocking exocytosis from all motor neuron terminals of striated skeletal muscle,

(v) convenience of use in requiring simple injection of soluble toxin preparation into desired muscle regions,

(vi) safety, if used with appropriate precautions and proper doses, and

(vii) absence of systematic or CNS effects (Johnson, 2005).

The possibility of using botulinum toxin as a therapeutic agent was further investigated in the early 1980's by groups of university-based ophthalmologists scattered throughout the U.S.A. and Canada. By 1985, a scientific protocol of injection sites and dosage had been determined for treatment of blepharospasm (BLE) and strabismus. Side effects had been estimated as a mild and treatable. The beneficial effects of the injection lasted only 4-6 months, so that BLE patients had to return to the clinics for re-injection two or three times a year.

In December 1989, the U.S. Food and Drug Administration (FDA) approved BTX-A (BOTOX®) a crystalline botulinum toxin type A produced in 1979 by Schantz, for the treatment of strabismus (crossed eyes), blepharospasm (uncontrollable blinking), and hemifacial spasm (twitching) in patients over 12 years old (Johnson, 2005). It was subsequently approved for the treatment of cervical dystonia and later for the treatment of glabellar wrinkles (frown wrinkles) and hyperhidrosis (abnormal sweating) (Barnes, 2007).

The cosmetic effect of BTX-A on wrinkles was originally documented by a plastic surgeon from Sacramento, Dr. Richard Clark, and published in Plastic and Reconstructive Surgery in 1989. On April 15, 2002, the FDA announced the approval of botulinum toxin type A (BOTOX Cosmetic®) for the reduction of frown lines between the eyebrows. BTX-A was later approved for the treatment of excessive underarm sweating.

The acceptance of BTX-A for the treatment of spasticity and muscle pain disorders is pending in many European countries. In addition, there are many studies currently involving the use of BTX-A for the treatment of headaches (including migraine), prostatic symptoms, asthma, obesity and many other possible indications. There are a number of producers of botulinum neurotoxin for therapeutic and cosmetic use which include; Allergan, Inc. the manufacturers of BOTOX® which is used for both therapeutic and cosmetic use; Ipsen who manufacture Dysport® which is a therapeutic formulation of the type A toxin developed and manufactured in Ireland and which is licensed for the treatment of focal dystonias and certain cosmetic worldwide and Medy-Tox Inc. of South Korea who introduced another BTX-A product, Neuronox® onto the market.

Finally, botulinum Toxin Type B (BTX-B) received FDA approval for the treatment of cervical dystonia and is marketed under the trade names of Myobloc® in the United States, and Neurobloc® in the European Union.

1.2.2 Biochemical structure of BoNT

Botulinum neurotoxins are composed of a heavy and a light chain linked together by a single disulphide bond. The BoNTs are synthesised as an inactive single chain polypeptide with a molecular mass of approximately 150kDa. This polypeptide chain

become active, when proteolytically cleaved into a di-chain molecule consisting of a 100kDa heavy chain and a 50kDa light chain (Dressler, 2005) (Fig 2).

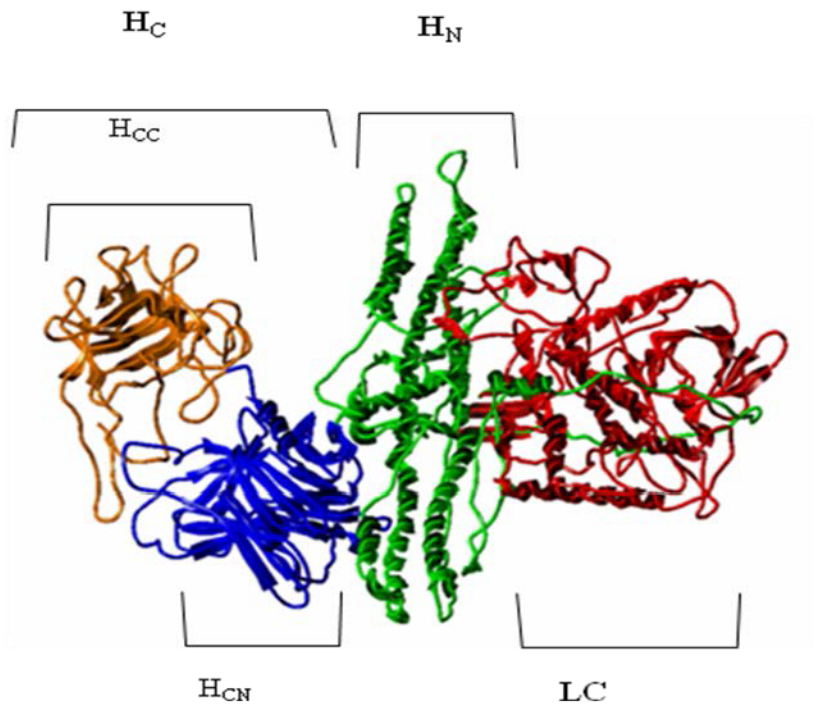


Figure 2. Tertiary structure of botulinum neurotoxin type A (3BTA) (adapted from Turton, Chaddock and Acharya, 2002)

The light chain (LC) is a globular structure of α-helix and β-strands, which acts as a zinc –dependent endopeptidase (Turton, Chaddock and Acharya, 2002). This catalytic subunit has a His-Glu-Xaa-Xaa-His (HEXXH) motif (Matthews, 1988) responsible for cleavage of the substrate SNARE protein and inhibition of neurotransmitter release (Foster, 2004).

The heavy chain consists of two functional domains, each of 50kDa. The N- terminal part of the heavy chain (H_n) forms low pH-induced ion channels in lipid bilayers in the intracellular endosome membrane and therefore acts as a translocation domain (Shone *et*. al., 1985). This cylindrical shaped heavy chain domain delivers the light chain (LC) into the cytosol to access the SNARE substrate. It also contains a belt region, an unstructured polypeptide, which wraps around the LC to protect the active site cleft preventing access to the zinc atom (Chaddock and Marks, 2006). The C-terminal part of

the heavy chain (H_C) is a ganglioside binding domain which is responsible for the high affinity neuronal binding (Halpern and Loftus 1993, Shone *et. al.*, 1985) to the target cell membrane and consequent internalisation of the toxin molecules into cholinergic neurons. The H_C domain contains two subdomains, H_{CN} and H_{CC}. They enclose highly important residues responsible for the ganglioside binding of the neurotoxins (Schiavo *et. al.*, 2000). The N-terminal subdomain (H_{CN}) is a structure of two seven-stranded antiparallel β sheets sandwiched together in a jelly-roll motif. The C-terminal subdomain (H_{CC}) has a β trefoil fold (Umland *et. al.*, 1997).

The crystal structure of BoNT/A (Lacy *et. al.*, 1998) and BoNT/B (Swaminathan and Eswaramoorthy, 2000) have been determined and published. The crystal structures shows the three distinctive structural domains, which correspond to the three steps process of intoxication: binding to the target cells, translocations, and catalytic activity, which leads to inhibition of neurotransmitter, release (Hanson, Stevens, 2002, Swaminathan and Eswaramoorthy, 2002). The three structurally individual domains are organized sequentially; therefore, there is no interaction between the binding and catalytic domains (Foster, 2004) (Fig. 3).

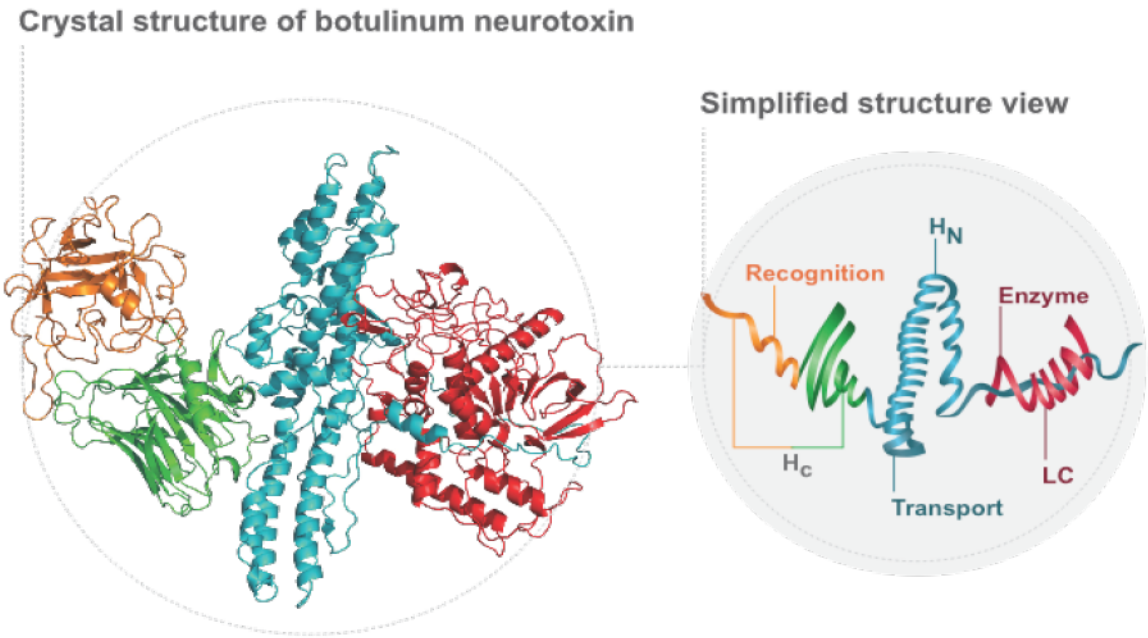


Figure 3. Clostridial Endopeptidase: Building Blocks for Engineering (adapted from Foster, 2004)

1.2.3 Process of Intoxication

The neurointoxication process is a series of biochemical events, in which five distinct steps can be identified:

1. Transcytosis of the neurotoxin from the gut to the blood stream
2. Binding of the neurotoxin to the pre-synaptic membrane of the neuromuscular junction (NMJ)
3. Uptake of the toxin into the motor neurons by endocytosis
4. Translocation of the LC into the cytosol
5. Cleavage of the target substrate, a key component of the vesicular fusion machinery

The last step of this process leads to inactivation of the neuronal vesicular release mechanism (Chaddock and Melling, 2001) and inhibition of acetylcholine release from motor neurons or sensory nerves (Foster, 2004). The blockade of neurotransmitter release from peripheral cholinergic nerve terminals (Rossetto *et. al.*, 2001) is regulated by proteolytic cleavage of the soluble N-ethylmaleimide – sensitive factor attachment protein (SNARE) proteins (Foster, 2004). This route prevents the stimulation of the muscle fibre at the neuromuscular junction, which subsequently leads to the flaccid paralysis known as botulism (Chaddock and Melling, 2001).

The introduction of BoNT (botulinum neurotoxin) into a target tissue acts as a signal to initiate a cascade of reactions. This cascade involves a concentration step via complex polysialo-gangliosides at the plasma membrane and the uptake into recycling synaptic vesicles initiated by binding to a specific protein receptor (Binz and Rummel, 2009). The H_C-fragment of the heavy chain of the botulinum neurotoxin binds to a glycoprotein structure found on cholinergic nerve terminals at the neuromuscular junction (Dressler, Saberi and Barbosa, 2005). This binding requires the presence of polysialo-gangliosides complex such as GD1a, GD1b, GT1b, as well as glycosphingolipids that are found particularly in the outer leaflet of neuronal cell membranes (van Heyningen and Miller, 1961; Simpson and Rapport, 1971). Ganglioside distribution and affinity binding studies contributed to an understanding of the binding process that leads to the dual-receptor-model (Montecucco, 1986). This

model involves two sequential binding steps: the toxin binds initially to the abundant polysialo gangliosides, which act as low affinity receptors that accumulate CNTs and then the neurotoxins linger in that position until they get access to protein receptors. Simultaneous interaction with ganglioside and protein receptor may be considered as high affinity binding and be mandatory for the specific Endocytosis step (Binz and Rummel, 2009) (Fig. 4).

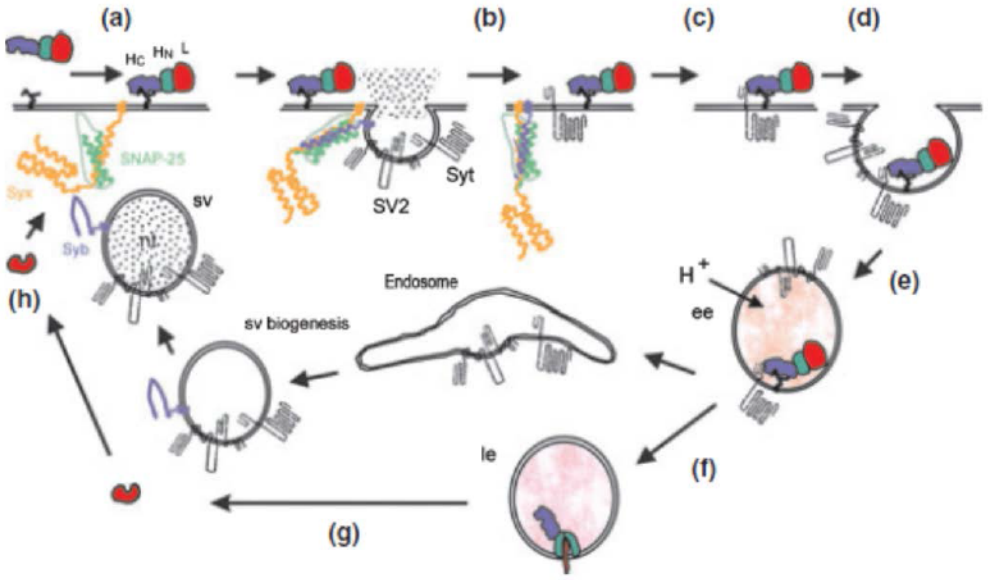


Figure 4 Model for the entry of BoNTs into nerve cells (adapted from Binz and Rummel, 2009).

(a) BoNTs bind initially to a ganglioside molecule on the cell surface.
(b) The intra-vesicular parts of synaptic vesicle proteins become exposed on the plasma membrane upon SNARE mediated neurotransmitter release.
(c) BoNTs access their protein receptor (exemplified here for Syt) possibly by means of lateral diffusion while bound to the ganglioside receptor.
(d) Having bound the protein receptor, the neurotoxins become endocytosed via retrieval of synaptic vesicles.
(e) The lumen of the recycling vesicle becomes acidified via the action of the vesicular proton pump.
(f) Acidification provokes a structural rearrangement in the neurotoxins, whereby the H_N-domain forms a channel through the vesicular membrane.
(g) The LCs pass the channel subsequent to partial unfolding and are released to the cytosol following reduction of the disulfide bond.
(h) Ultimately, the LCs cleave their target SNARE(s), Synaptobrevin (Syb, light blue), Syntaxin (Syx, orange), or SNAP-25 (light green), and thus block the synaptic vesicle cycle which they had exploited for cell entry. ee, early endosome; le, late endosome; sv, synaptic vesicle.

In this theory, it was proposed that the binding of BoNT to a protein receptor is required before internalization commences. In a number of cases, the details of these

protein receptors are unknown, however there is some evidence based on crystallographic data that provide structural insights into receptor recognition by BoNT/B. BoNT/B co-crystallized with synaptotagmin (Syt II) shows binding in crevice at the tip of the H_C-fragment (aa 44-60) in the neighbourhood of ganglioside binding pocket (Chai *et. al.*, 2006). This study provides a molecular basis for studying other potential neurotoxin-receptor interactions and support dual-receptor concept. BoNT/B uses two homologues of synaptotagmin (Syt-I and Syt-II) (Dong *et. al.,* 2003), as well as BoNT/G also binds Syt-I and Syt-II as their protein receptors (Rummel *et. al.*, 2004b). BoNT/A use all three isoforms of synaptic vesicle 2 (SV2A, SV2B and SV2C) (Dong *et. al.*, 2006; Mahrhold *et. al.*, 2006); BoNT/E relies on SV2A and SV2B, but not on SV2C (Dong *et. al.*, 2008). SV2 and Syt-I/Syt-II are synaptic vesicle membrane proteins. Their toxin-recognition domains are located in the vesicle lumen and are exposed to the extracellular milieu transiently after synaptic vesicles fuse with the plasma membrane (Chai *et. al.*, 2006).

Following the binding of the heavy chain to the surface of the neuron, the clostridium neurotoxin internalizes into acidic compartments through an endocytic process (Turton, Chaddock and Acharya, 2002). Research indicates that the structural arrangement of botulinum neurotoxin inside an acidic compartment is a pH-dependent (Montecucco et. al., 1994) and most probably, a pH reduction activates the structural change in BoNT (Turton, Chaddock and Acharya, 2002). The toxin light chain probably unfolds at low pH and refolds after interchain disulphide reduction in the cytosol following its transport through a pore that is formed in the membrane through a structural change in the heavy chain (Rossetto *et. al.*, 2001).

After translocation to the cytosol, the light chain of botulinum neurotoxin binds with high specificity to the SNARE protein complex, (Rizo and Sudhof, 1998). SNAREs comprise of three proteins: synaptosomal – associated proteins of 25kDa (SNAP – 25), syntaxin, and vesicle-associated membrane protein (VAMP /synaptobrevin/) (Montecucco, Schiavo and Pantano, 2005). SNARE complexes mediate fusion of synaptic vesicles with the plasma membrane (Bajjalieh, 1999), which regulate secretion of neurotransmitter in cholinergic neurons (Montecucco, Schiavo and Pantano, 2005).

SNARE complex formation continues even if BoNT cleaves the individual SNARE protein. Nevertheless, these complexes are non-functional, as the coupling between Ca^{2+} influx and fusion is broken (Humeau *et. al.*, 2000). It has been observed that Ca^{2+} has a particular role in the inhibition of neurotransmitter release, because increasing Ca^{2+} concentration in the synaptic terminal reverses the effect of BoNT/A (Meunier *et. al.,* 2002). The proteolytic cleavage of the SNARE protein complex prevents the docking of the acetylcholine vesicle on the inner surface of the cellular membrane. This results in the blockade of vesicle fusion. This process of acetylcholine secretion inhibition by clostridial neurotoxins is terminated by restoration of a functional SNARE protein complex (Dressler, 2005).

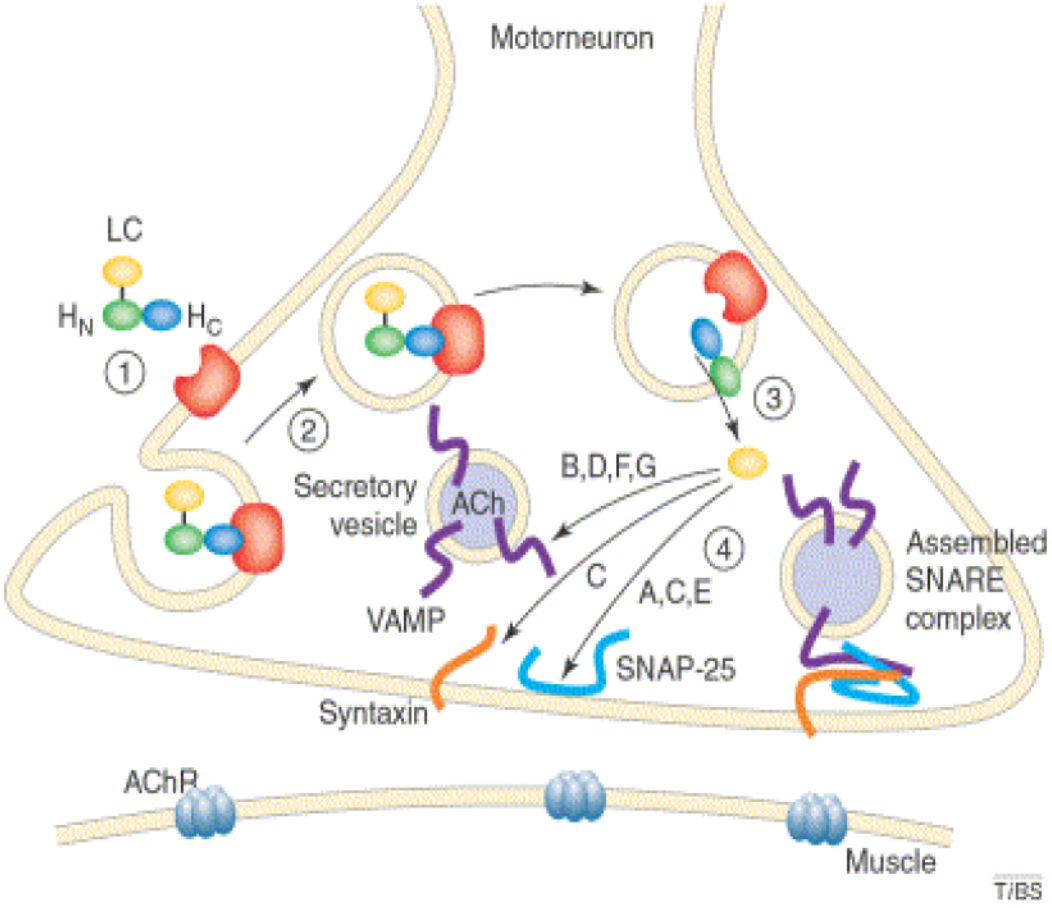


Figure 5 Stages of botulinum neurotoxin action: Inhibiting cell secretion (adapted from Turton et. al., 2002).

Legend to figure 5
Mechanism of action of botulinum neurotoxin (BoNT). Each of the structural elements of botulinum neurotoxin, the light chain (LC) endopeptidase (yellow), the HN translocation domain (green) and the HC cell-binding domain (blue) have a role in the mechanism of action of neurotoxin.
The first stage in the intoxication process (interaction with gangliosides and an, as yet unidentified,
protein receptor) (1) is followed by stage two, internalization of the toxin-receptor complex into an
intracellular vesicle (2). The third stage (translocation) is characterized by release of the light chain
endopeptidase from an acidic intracellular compartment into the cytosol (3). Once liberated from the
vesicle, the light chain performs the final stage of intoxication; highly specific proteolytic cleavage of
one of the proteins of the SNARE complex (4). BoNTs B, D, F and G cleave proteins of the VAMP family
(purple) and BoNTs A, C and E cleave SNAP-25 (light blue). BoNT/C also has the capacity to cleave
syntaxin (orange). Cleaved SNARE proteins are competent for facilitating docking of the secretory
vesicle with the synaptic membrane, but fusion of the vesicle is compromised. Thus neurotransmitter
release is inhibited. Abbreviations: HN, heavy chain N-terminal subdomain; HC, heavy chain
C-terminal subdomain; SNARE, soluble NSF-attachment protein receptors; VAMP, vesicle-associated
membrane protein; SNAP-25, synaptosomal-associated protein of 25 kDa; ACh, acetylcholine; AChR, acetylcholine receptor.

1.2.4 SNARE complex

SNAREs are distinguished by structural and functional classifications. The functional organization describes them as v-SNAREs and t-SNAREs. The SNAREs incorporated into the membranes of transport vesicles, where they are required as a component of the vesicle, are v-SNAREs. The SNAREs located in the membranes of target compartments are t-SNAREs (Sollner *et. al.*, 1993). Individual sets of SNAREs build the core for different intracellular membrane fusion processes. V-SNAREs have homology with neuronal synaptobrevin; t-SNAREs have homology with neuronal syntaxin and SNAP-25. These proteins have a helical structure and the interaction between v-SNAREs and t-SNAREs leads to the formation of a *trans*-SNARE complex. This complex consists of four SNARE motifs in parallel, in which the four-helical bundles catalyse the docking and fusion of the vesicle with the target compartment (Weber and Zemelman, 1998 Fukuda et. al., 2000). The zippering of such SNAREs into parallel four-helix coiled-coil

bundles provide the driving force for membrane to merge (Sutton *et. al.,* 1998; Jahn and Scheller 2006), 'squeezing out' water until the lipid bilayers are within 1.5nm of each other. This allows lipid to achieve hemifusion and then full fusion.

When the three cognate partners meet, SNARE motifs form four intertwined parallel helices that have extraordinary stability (these are two helices from SNAP-25, 1 helix from the Syntaxin t-SNARE and 1 from the v-SNARE). These larger hydrophobic interactions give rise to the newer nomenclature relating to the single hydrophilic residue lying in the middle of each helix (at the so called 0 layer) where there are 3 Gln (Q) and an Arg (R) (Fasshauer and Sutton, 1998). The Q group can be further divided into three sub-groups according to their overall homology within the SNARE domain: Qa (or syntaxins), Qb (or SNAP N-terminal) and Qc (or SNAP C-terminal) R refers to synaptobrevin-like SNAREs (Bock, 2001).

VAMP/ synaptobrevin 1 and 2 are 13kDa, large dense-core granular membrane proteins of SV (synaptic vesicle). Syntaxins 1 and 2 are bound to presynaptic membrane of neurons by a TM segment linked to a short C-terminal domain exposed to the cell surface and a large cytosolic portion. Synaptosomal - associated membrane protein of 25kDa (SNAP-25) is localized on the cytosolic face of nerve membranes (Montecucco, 2005).

The different serotypes of BoNT differ in their functional characteristics, particularly regarding to specific cleavage of target SNARE, potency and duration of action. For instance, BoNT/A and BoNT/C have the longest inhibitory effect in comparison to the other BoNT's; where BoNT/E causes the shortest on-set time for paralysis (Eleopra *et. al*,. 1997, Eleopra *et. al*., 1998, Foran *et. al*., 2003). BoNT/A and BoNT/B are the most widely used serotypes of botulinum neurotoxin. The various BoNTs and their specific substrates are given in Table 1.

Table 1 Serotypes and SNARE cleavage

Endopeptidase Serotype	SNARE Substrate	Duration (*in vivo* / rat)	Duration (clinic / man)
A	SNAP25	30 days	3-4 mo
B	VAMP	18 days	2-3 mo
C	SYNTAXIN	-	~3 mo
D	VAMP	-	-
E	SNAP25	4.5 days	<30 days
F	VAMP	7.5 days	30 days
G	VAMP	-	-

(-) lack of data

1.3 The Immunoglobulin

Immunoglobulins are glycoproteins produced by B cells and plasma cells in the immune system in the blood or tissues in response to a specific antigen, such as a bacterium or a toxin. They can neutralize or destroy bacteria by binding to critical portion of the toxin or blocking proteins necessary for attachment to cells.

These proteins, known also as an antibodies or gamma globulins (abbreviated Ig) are products of the white blood cells called B-lymphocytes, where they function as a receptor for antigen. They comprise the principle component of the adaptive humoral immune response (Lucas, 2001). The term "antibody" dates to 1901. Prior to that time, an "antibody" referred to any of a host of different substances that served as "bodies" (foot soldiers) in the fight against infection and its ill effects.

Antibodies have been found in the immune system of all vertebrates. They represent a strong evolutionary solution to recognize and eliminate microbial pathogens to which they can bind with high specificity and affinity (Du Pasquier and Flajnik, 1999). These characteristics have enabled the use of antibodies as molecular tools for the purpose of diagnostic and therapeutic applications (Auf der Maur, 2002).

1.3.1 Antibody structure

The structure of the monomeric immunoglobulin molecule consists of four polypeptide chains: two identical heavy (H) chains and two identical light (L) chains covalently bonded by several intra- and inter- chain disulfide bridges between cysteine residues (Fig 6) (Edelman et al., 1969).

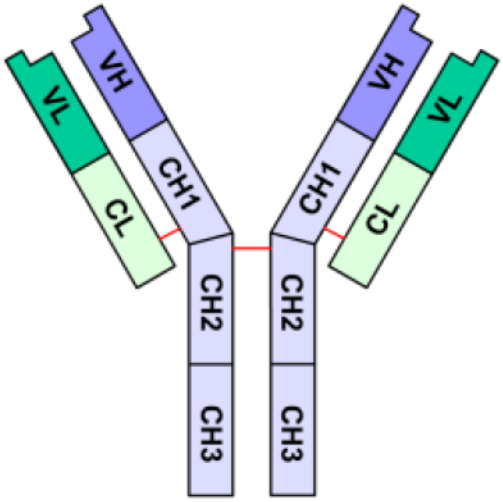


Figure 6 Illustrating the disulphide bonds (red) that link the light (green) and heavy (blue) protein subunits of Immunoglobulin G (IgG) molecules.
This diagram also illustrates the relative positions of the variable (V) and constant (C) domains of an IgG molecule. The heavy and light chain variable regions come together to form antigen binding sites at the end of the two symmetrical arms of the antibody. V_L-variable light chain domain; V_H-variable heavy chain domain; CH1, CH2 and CH3-constant domains of heavy chain; CL-constant domain of light chain (adapted from Edelman et al., 1969).

In addition, the chains are arranged in the way that the H and L chains form pairs (Lucas, 2001). The approximate molecular weight of antibody is estimated to be 150 kDa. However there are exceptions from this rule, and for example pentameric IgM antibodies have a high molecular weight, which can be 900 kDa. Each Ig H chain is about 440 amino acid long and each L chain is about 220 amino acid long. Stabilized by disulphide bonds the H and L chains form 110 amino acid domains, which are a common feature of many soluble molecules and membrane bound receptors.

All antibodies have one of two types of light chain (L): kappa (κ) or lambda (λ) and one of five types of heavy chains (H): alpha (α), gamma (γ), delta (δ), epsilon (ε) and mu (μ) (Gottlieb et. al., 1968). The H and L chains are divided into two regions: the variable V region, which is located at the N –terminal end of polypeptide chains and the constant C region at the C terminus. The amino acid sequence of N-terminal of each chain varies considerably between different antibodies, and is known as the variable region. The amino acid sequence of the carboxyl half of the light chain (~110 residues) and 3/4 of the heavy chain (~330 residues) has limited variation and therefore is called the

"constant region". The VH (heavy chain) and VL (light chain) domains distinguish the variable regions. Each light chain has one V_L domain and each heavy chain has one V_H domain. The variable regions of L+H chains consist of hypervariable regions with high sequence variability and framework regions with insignificant sequence variability. Each antibody molecule has two identical binding sites (Wu and Kabat, 1970) because the hypervariable regions during folding produces the antigen binding pockets, which are composed of one V_H and one V_L domain.

The vast antigen binding capacity of the antibody structure is achievable by the association of polymorphic V_H and V_L regions (Lucas, 2001). These hypervariable regions, which represent the closest contact between antibody and antigen, are the complementary determining regions (CDRs) (Gauci and Alderton, 2001). There are six CDR's that produce the binding pocket (Weisser and Hall, 2009 Kortt *et. al.*, 2001), three in the hypervariable regions of the variable light chain and three in the variable heavy chain (Fig. 7). The antigen-binding site accommodates approximately four to seven amino acids or sugar residues. An antibody binds its specific antigens using a combination of hydrogen bonds, ionic bonds, hydrophobic interactions, and Van der Waals interactions. This binding is reversible, as there are no covalent bonds formed between antigen and antibody (Montero, 2003).

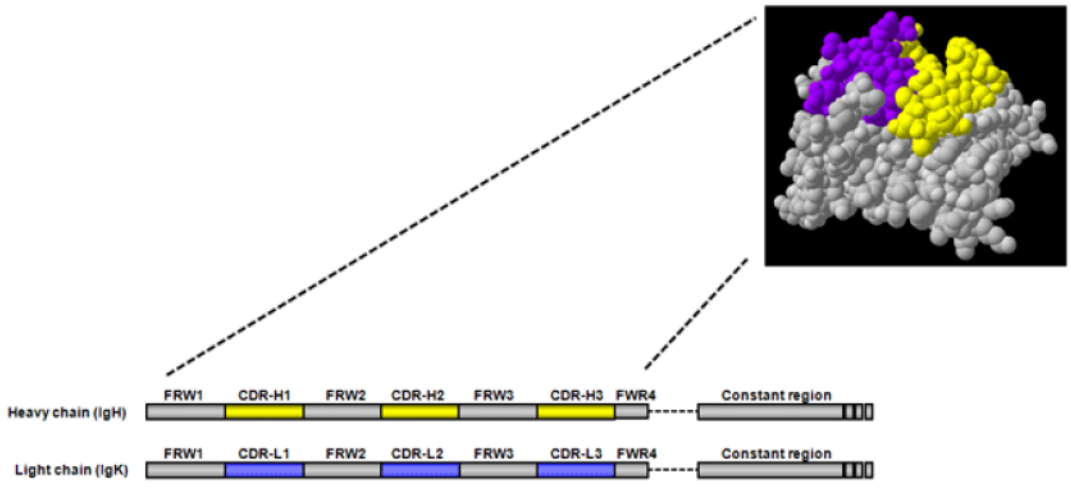


Figure 7 The position of six CDRs in the light and heavy chain of an antibody.
Heavy chain CDR's are depicted in yellow and light chain CDR's are in purple. Regions in grey provide the framework upon which the CDR's are supported. The space-filling model illustrates the way in which CDR's from both heavy and light combine in three dimensions to create a binding pocket (adapted from Vaccinex, 2007).

The constant region of the antibody has one C_L domain for each light chain and three (α, γ and δ chains) or four (ε and μ chains) C_H domains for each heavy chain (Montero, 2003). The function of C_H domains is to mediate biological effector function (Lucas, 2001) such as complement activation, F_C receptor binding, avidity and serum half-life (Ravetch and Kinet, 1991).

The only region, which does not fold into a domain, is a hinge region between H chain C_H1 and C_H2, which is, connected via intrachain disulphide bridges (Gauci and Alderton, 2001). The hinge region allows independent movements of two antigen-binding regions to bind antigen (Montero, 2003). It is also responsible for the separation of antigen-binding sites from the Fc (crystalline fragment) segment, where proteolytic enzymes preferentially target this region (Gauci and Alderton, 2001).

An epitope is a portion of antigen recognized and bound by antibody domain. Each epitope has its own antibody(s) directed against it. The strength of the interaction between a single antigen-binding site on the antibody and its specific epitope is called the binding affinity of the antibody. The affinity measures the association between antigen and antibody, the strength of which depends on affinity levels. A high affinity guarantees the presence of antigen in the binding side (Ofek *et. al.*, 2010).

The beta-sheet and loop topology of the immunoglobulin fold has a core of four beta – strands covered in a curled beta-sheet sandwich. On the edge of this structure are additional beta-strands, which exhibit high structural flexibility. The infinite potential for protein – protein interactions between antibodies and antigens occur as the loops between beta-strands ensure sequence specific topology (Hoogenboom, 2005).

1.3.2 Immunoglobulins and their fragments

Antibody molecules belong to groups or classes according to their heavy chains. In humans, there are five classes of immunoglobulin: IgM, IgD, IgG, IgA and IgE (Fig. 8). The biological properties vary between the classes (Lucas, 2001). Immunoglobulin G (IgG) has been identified as the most suitable antibody molecule for use studying research and therapeutic purposes, as it has higher affinity for antigen in comparison to other immunoglobulin classes (Guci and Malcolm, 2001). After 35 years of study on the use of monoclonal antibodies (mAb) as a potential therapeutic this is able to

recognize a specific molecular target (Köhler and Milstein, 1975). The USA Food and Drug Administration (FDA) have approved 30 therapeutic monoclonal antibodies or antibody-derivatives (Fab fragments, radioimmunoconjugates, immuneconjugates and F_C fusion proteins) until 2008 (Beck *et. al.*, 2008). They are directed against cancer, treatment of transplant rejection and to combat autoimmune diseases (Reichert, 2001). In addition they are being used to treat rheumatoid arthritis (Centocor's Remicade and Abbott Laboratories Humira), non-Hodgkin's lymphoma (Genentech's Rituxan and IDEC's Zevalin) and respiratory syncytial virus infection (Medimmune's Synagis) (Holt *et. al.*, 2003). FDA approved five mAbs until 2010 and the additional two mAbs in 2011. Although this success is significant, there is a need to produce smaller antibody fragments that will avoid the formulation and manufacturing issues that are present in the production of conventional antibodies. These include limited expression systems (Holt *et. al.*, 2003), inability to trigger human effector functions and the danger of repeated administrations due to a reaction by the immune system against murine Ab domains (Sanz, Blanco and Alvarez-Vallina, 2004). These drawbacks partly have been resolved by chimerization (fusion of mouse variable regions to human constant regions) and humanization (grafting complementary determining regions onto human acceptor Ab frameworks) (Hudson and Souriau, 2003), however the use of high volume bacterial or yeast cell culture (Holt *et. al.*, 2003) of small antibody fragments has higher potential for new class of drugs.

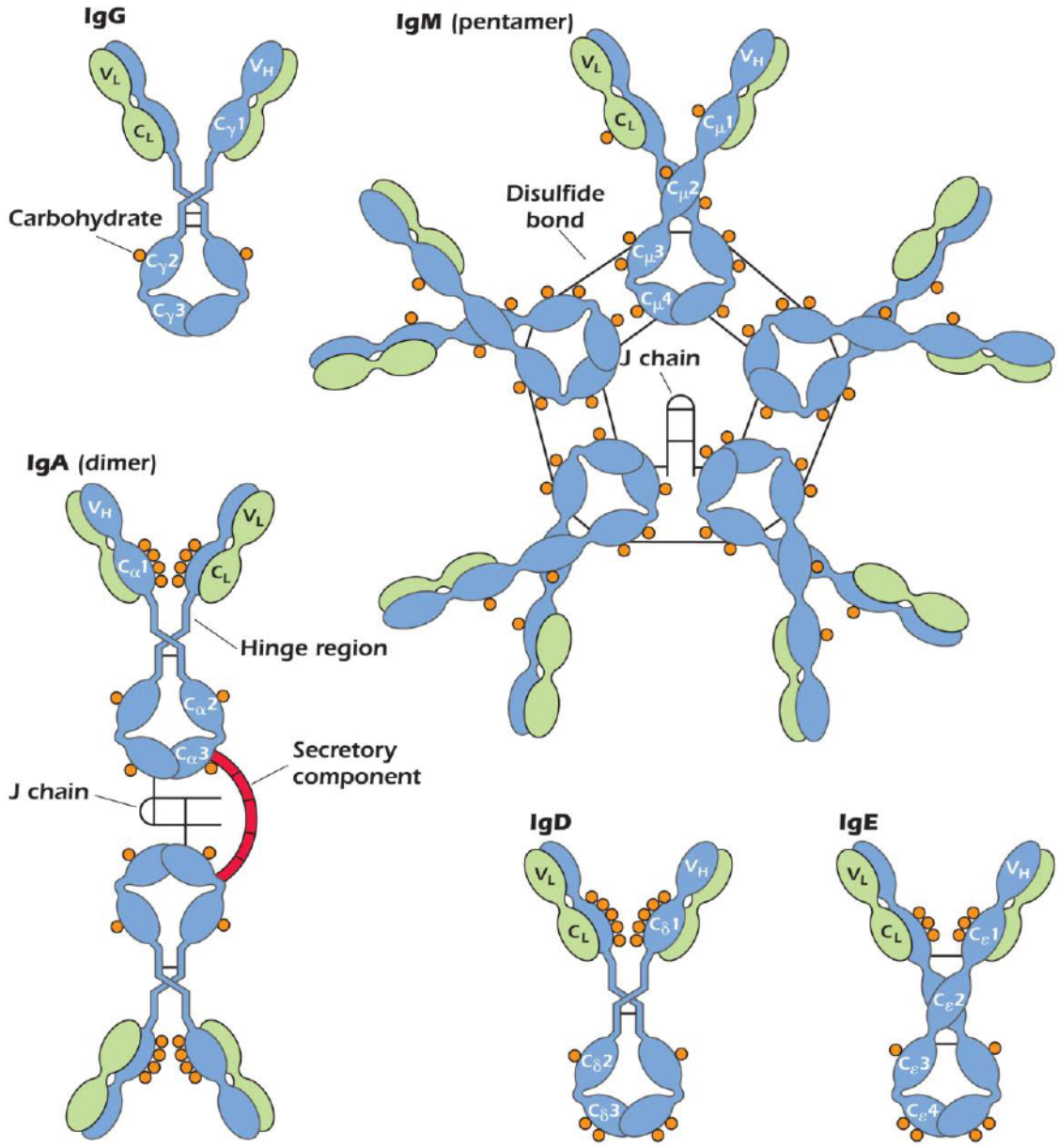


Figure 8 Main classes of immunoglobulin IgG, IgA, IgM, IgD and IgE.
The heavy chains are depicted as blue, and the light chains are depicted in green. Orange circles denote areas of glycosylation. The polymeric IgM and IgA molecules contain a polypeptide known as the joining (J) chain, which is disulfide-linked to the tail pieces and stabilizes the multimeric structure. The dimeric IgA molecule shown includes the secretory component in red. The parts labelled V represent the variable regions, and the parts labelled C represent the constant regions (adapted from Goldsby et. al., 2003).

Early studies of antibody structure showed that antibody molecules could be digested with enzymes to generate different functional fragments that demonstrated the whole molecule is not required to maintain specificity (Porter, 1973). The discovery of antibodies by Parventjev (Parventjev, 1936), and subsequent studies on horse antibody

performed by Petermann and Pappenheimer (Petermann and Pappenheimer, 1941) and follow-up experiments with papain by Petermann (Petermann, 1946) to generate functional fragments has opened up new opportunities for the development of new therapeutic antibodies. The specific cleavage sites of pepsin and papain are shown in figure 9. These studies, combined with investigations of antigenic specificity of protein antigens (Landsteiner, 1942) suggested that antigenic sites and antibody combining sites did not include the complete antibody molecule (Porter, 1973). Although various experiments in different conditions and enzymes were performed (Porter, 1950), only papain gave active products (Porter, 1973).

The digestion of immunoglobulin IgG with papain generates three fragments: two Fab fragments (antigen binding) and an Fc fragment (crystallizable) (Fig. 9). Fab fragments are composed of entire L chain and the V_H and C_H1 domains of heavy chain. The Fc fragment poses effector functions, such as complement fixation and binding to the surface of phagocytic cells (Lucas, 2001). The isolated antibody heavy (Utsumi and Karush, 1964) and light chains (Yoo *et. al.*, 1967) are able to retain antigen specificity, but their affinity and solubility is normally reduced (Ward *et. al.,* 1989).

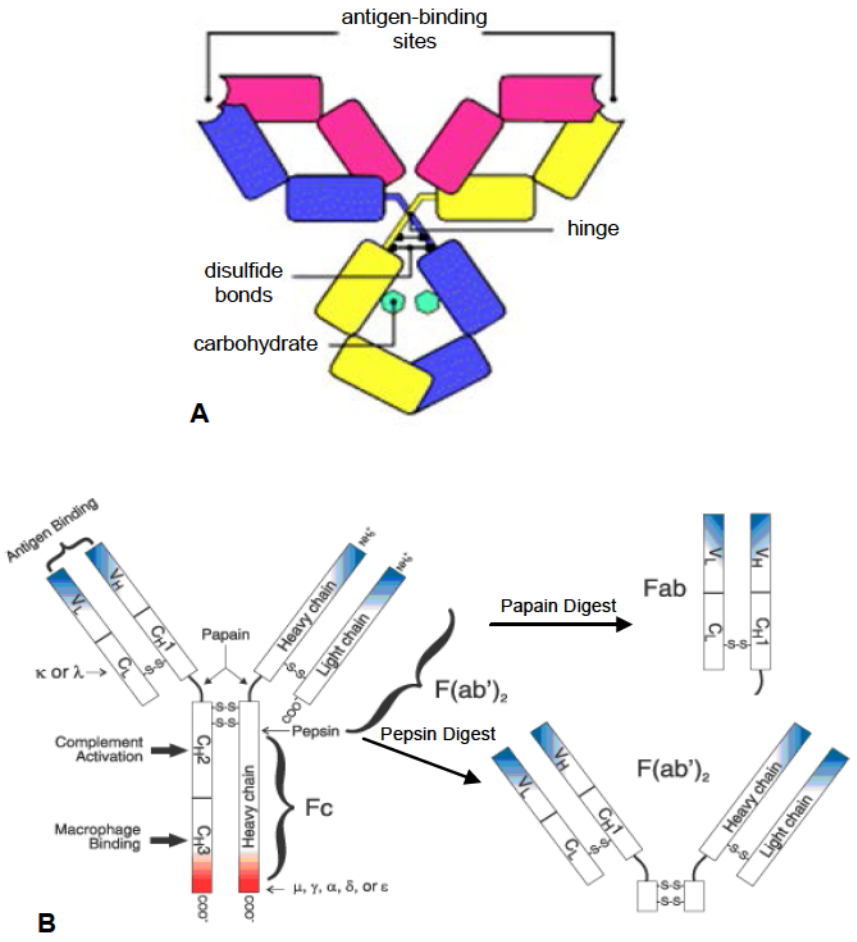


Figure 9 Schematic of Ig domains (A); Use of pepsin and papain to generate F(ab')$_2$ and Fab fragment (B) respectively.

A) The antigen binding sites formed by the heavy and light chain variable regions are labelled, so too are the hinge region and disulfide bonds (adapted from Janeway and Travers, 1996). **B)** The sites where two well used proteases, pepsin and papain, cleave IgG and the resulting fragments Fab and F(ab')$_2$ (adapted from Molecular Probes, 2005).

The paired N terminal variable domains of heavy and light chains are sufficient for antigen binding (Sundberg and Mariuzza, 2002). This type of antibody structure can be produced as a monovalent (Fab) fragment or a single chain antibody fragment (scFv) where just the V_H and V_L domains are fused by a peptide linker (Harmsen and Haard, 2007). The recombinant Fabs (~57 kDa) and single chain Fv fragments (scFv ~27 kDa) have two important advantages over the whole immunoglobulin: expression is possible in a bacterial host and isolation can occur without the need for animal immunisation (Jung, 2009, Sollner et. al., 1993). Another breakthrough was the discovery of single domain antibodies (sdAbs), especially those produced by camels, with fragments ranging from 11 kDa to 15 kDa. They are highly stable and soluble and can be formulated into larger molecules to create drugs with prolonged serum half-life (Holt *et. al.*, 2003) (Fig 10). After taking into consideration the smaller size of scFvs and sdAbs, in comparison to Fabs fragments; the higher yields in bacterial culture and the increased ease of combining them with another protein of biopharmaceutical potential they were chosen for the purpose of this MPhil study. The sdAbs and scFv fragments were used for experimental work; therefore, the more detailed description of properties is given in the appropriate chapters.

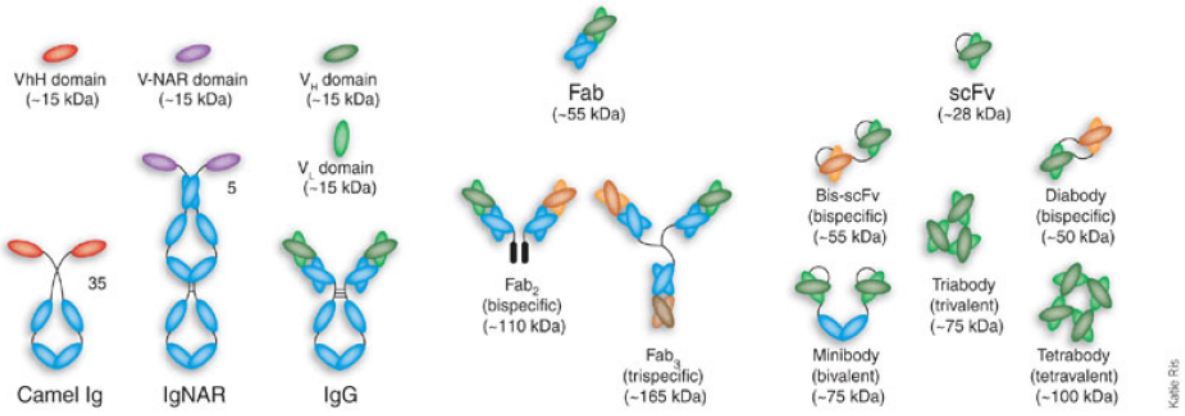


Figure 10 Various Antibodies and derived fragments from them (adapted from Holt, 2003).

1.4 Therapeutic use of clostridia neurotoxin fragments

Research of the biochemical neurotoxin action and structure lead to the realisation that it could be possible to design novel molecules with the ability to inhibit secretion (Foster, 2004). BoNTs have an effect on the cholinergic neurones at the neuromuscular junctions and on acetylcholine – releasing neurons in the autonomic nervous system (Bhidayasiri and Truong, 2005). Lately identification of the effects on non-cholinergic pathways of BoNTs has extended the spectrum of diseases against which the botulinum neurotoxin can be used (Foster, 2004). Each domain of botulinum neurotoxin is independently functional and therefore there is an opportunity to utilize them individually. The H_C, H_N and LC are responsible for different stages of cellular intoxication (Foster and Chaddock, 2007). It has been reported that a molecule named LH_N/A is a representative of the LC and H_N domains of type A neurotoxin coupled by single disulphide bond (Shone, Hambleton and Melling, 1985). The LH_N/A fragment is nontoxic because it lacks the H_C domain responsible for binding to receptors on the neuronal surface. The substitution of H_C domain with novel ligands gives an opportunity to retarget the LH_N/A fragment into different neuronal and non-neuronal cells. This subsequently leads to inhibition of those cells by cleavage of specific SNARE proteins, in this case SNAP 25 (Chaddock, Purkiss and Friis, 2000, Chaddock, Purkiss and Duggan, 2000). The first crystal structure of LH_N/A fragment without the (H_C) binding domain of botulinum neurotoxin has been determined. This structure does not show any apparent structural differences between the full-length of BoNT/A and the recombinant fragment lacking H_C domain. It therefore opens up an opportunity for the development of optimized proteins for the medical industry (Masuyer *et. al.*, 2009).

1.5 Chimeric protein conjugates with toxin

Within the last 30 years, many attempts have been made to create targeted cytotoxic agents that have the potential to specifically kill target cells, for example in the treatment of cancer. Some of these agents have used existing proteins as the targeting component, for example, the therapeutic agent, which contains human interleukin-2 and FDA approves truncated diphtheria toxin for use in cutaneous T-cell lymphoma

(Kreitman, 2006). Additionally, small proteins such as genetically fused growth factors and cytokines to protein toxin were useful for this type of study (Cawley *et. al.*, 1980).

A second class of retargeted agent are Immunotoxins. Immunotoxins are proteins that contain a toxin with an antibody that binds specifically to target cells. The successful action of immunotoxin takes place when it binds to and is internalized into the target cells to release the warhead (Kreitman, 2006). Originally, immunotoxins were created by chemical conjunction of an antibody to a whole protein toxin or by using a protein toxin devoid of its natural binding domain to achieve selective activity (Moolten and Cooperband, 1970, Krolick *et. al.*, 1980). The most common use of immunotoxins has been applied for investigating new approaches to cancer treatment (Fig 11).

There are the following advantages of using immunotoxin:

- Increases the efficacy of chemotherapeutic agents in the treatment of solid cancers;
- Improves accessibility through effective transport across cell barriers;
- Targets a specific cell type;
- Recognizes tumour specific-antigens;
- Binds to the cell and internalizes;
- Assists toxin to kill the target cell;

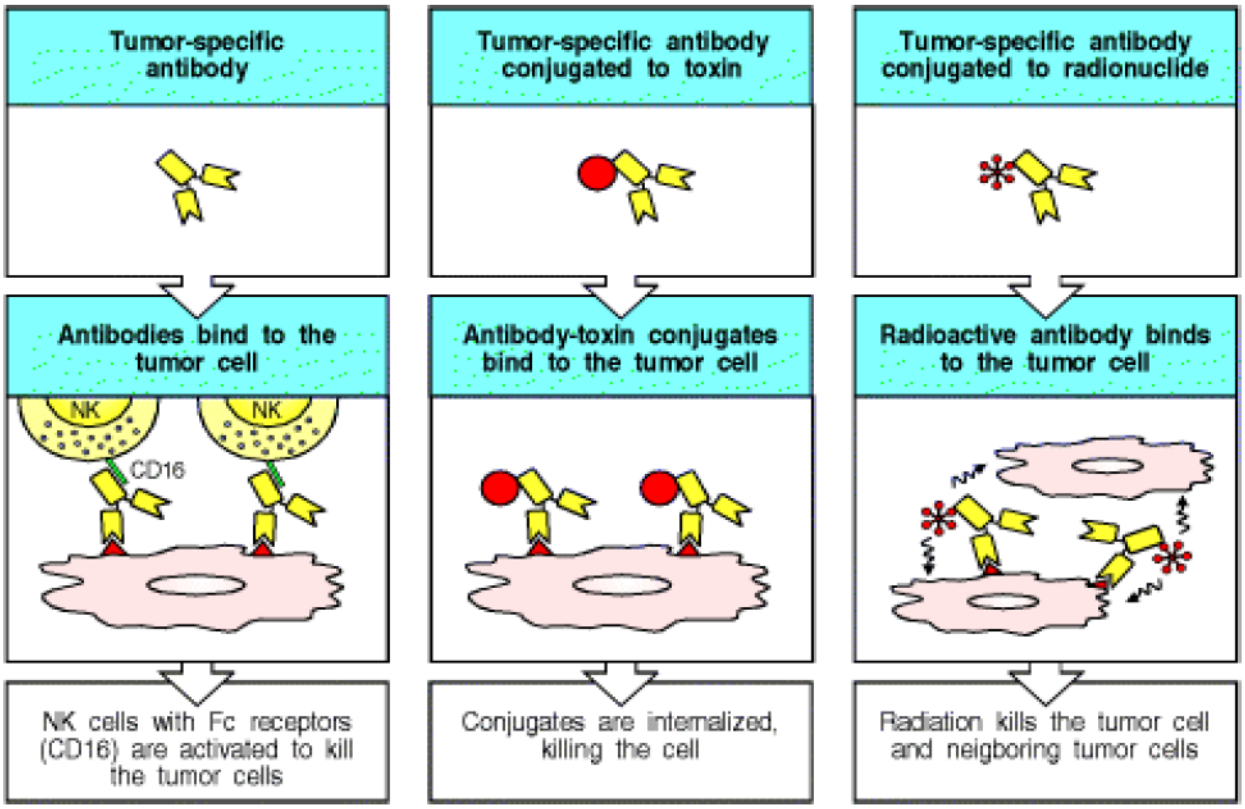


Figure 11 Antibodies as Cancer Therapeutics.

There are the following disadvantages of using immunotoxin:

- May show lack of translocation within cells;
- VLS (vascular leak syndrome);
- Immune response to foreign toxin;
- Short Half-life in the circulation;
- Hepatotoxicity implies chemical driven liver damage;
- Lack of the target effects;

1.6 Functionality of novel molecules

The construction of recombinant neurotoxin chimera with antibody molecules, such as scFv (single chain antibodies) or sdAbs (single chain antibodies) gives an opportunity to develop new class of biopharmaceuticals.

However, in order to generate novel drugs is important to know principles of signalling mechanism in which this drug is involved. The choice of the antibody molecule against the particular receptor directs which cell pathway would be activated and controlled; for example, EGF activates the tyrosine kinase pathway initiated from EGFR. The strength of the signal depends on the molecule characteristics, such as size, penetration time, and level of recognition of the target cell. The measure of activation and internalization of novel antibody-LH_N chimeras was not a part of this study, however the primary research towards this concept have been initiated. All single chain and single domain antibody fragments designed for this research are directed against EGF (Epidermal Growth Factor) receptor. The EGF belongs to Tyrosine Kinase family. The binding of EGFR on cells leads to phosphorylation and activation of kinases family, which are commonly known as a MAP (Mitogen Activated Protein) kinases. EGF is a growth factor that plays an important role in cell growth, proliferation, differentiation by binding to its receptor EGFR. Growth factors (including EGF) are involved in tumour maintenance and growth and many tumour cells therefore express high levels of EGF-R. Therefore, recombinant antibodies to such receptors are promising leads in cancer biology. A more detailed description about EGF, its activation and internalization can be found in the Chapter 6 and 7.

Stem Cell Factor (SCF) is a glycoprotein that signals through the cells and membrane tyrosine kinase receptor defined as a proto-oncogene C-KIT (CD117). Altered forms of this receptor may be associated with some types of cancer. The design of humanized C-KIT antibodies, play an important role in treatment of C-KIT, associated inflammatory and cancerous diseases. GIST (gastrointestinal stromal tumour) is a tumour that grows in the connective tissue between muscle layers in the digestive tract. All GISTs can become cancerous. The tumour usually starts when a gene called *kit* develops a mutation (defect) and forms a protein called Kit. The abnormal Kit protein sends out a

signal that causes the cell to multiply out of control, forming a tumour. The production of recombinant neurotoxin chimera with antibody molecules to CD117 has the potential to be therapeutically beneficial where secretion from the particular cell plays a causative role in a disease or medical condition.

1.7 Aims of the thesis

The initial aims of the project:

- Design of antibodies used for the construction of recombinant molecules. The antibody can be tumour specific that will efficiently target the tumour and transport the drug conjugates into and even across the tumour endothelium for direct delivery to the underlying tumour cells (Chapter 3).
- Construction of botulinum based antibody-related molecules. The delivery of antibody molecule into cells to target cancer can has greater potential with botulinum endopeptidase translocation domain than on its own. The antibody-botulinum toxin molecule binds to the relevant antigens and efficiently internalize into the cells, where can suppress transmitter release from target neurones via proteolytic cleavage of specific SNARE protein (Chapter 4).
- Expression of recombinant chimeras; standard expression *E. coli* system was explored for suitability for expression of the recombinants (Chapter 5).
- Purification of recombinant chimeras. The purification system based on nickel metal affinity chromatography matrices for protein tagged with ten consecutive Histidine residues is the most suitable one for recombinant proteins produce within Syntaxin Ltd. Company (Chapter 6).

Once these aims had been met, further aims were established that made use of the tools developed in the initial stages of the project.

- The standard expression protocol in *E. coli* system was changed to reach higher purity of purified molecules (Chapter 5).

- Other purification systems were introduced to the project to achieve better results than this provided by affinity purification. New systems were incorporated to aid an unspecific cleavage within LH_N single chain antibody fragment molecules (Chapter 6).
- The activation assay for family ligands of tyrosine kinase such as EGF (Epidermal growth factor) was established that would allow testing novel molecules (Chapter 7).
- Internalization experiments were performed to check if the internalization of chimeras is possible and the same time the secretion pathway can be inhibited (Chapter 8).

Chapter 2

Materials and Methods

Chapter 2. Materials and Methods

The studies on the novel recombinant antibody LH_N chimeras required a selection of techniques on the level of DNA, microbiology, and protein production and analysis. This chapter describes common techniques use at Syntaxin Ltd. as well as techniques developed for the purpose of this study or personal interest, such as assay for process of EGFR activation and internalization.

2.1. Chemicals

General chemicals were purchased from Fisher Scientific UK; Loughborough UK; Sigma Aldrich, Poole, UK; Melford Laboratories, Ipswich, UK; Anachem, Luton, UK; Bio-Rad Laboratories, Hemel Hempstead, UK; Invitrogen by Life Technologies, Paisley, UK; Merck, NJ, USA; Millipore, Billerica, MA, USA; Prebio/Pierce, Rockford, IL, USA; Sartorius, Gottingen, Germany; VWR, East Grinstead, UK; Corning, NY, USA;
Specialized chemicals, kits and services were described under each chapter that required their use.

2.2 Design of recombinant antibody based constructs

2.2.1 Antibody selection

2.2.1.1 Single chain antibody fragments

Single chain antibody fragments scFvs sequences were obtained from the patent (WO/2007/127317) Humanized C-KIT antibody, as seen in the table 2. Inventors of this patent are NG, Gordon; (US); SHEN, Wenyan; (US). This invention relates to compositions and methods for treating c-Kit associated disorders such as fibrosis, and more particularly, to compositions containing humanized c-Kit antibodies.

Table 2 WO/2007/127317 Patent data for CD117I and CD117II

PATENT DATA	humanized light chain	humanized SR-I IgG2 heavy chain	humanized SR-I MULC IgG2 heavy chain
SEQ ID NO	2	4	6
Length (amino acids)	20 - 248	19 - 467	19 - 463
CDR1	43 - 58	50 - 54	26 - 35
CDR2	74 - 80	69 - 85	50 - 66
CDR3	113 - 121	118 - 125	99 - 106
CD117I	YES	YES	NO
CD117II	YES	NO	YES

2.2.1.2 Single domain antibodies

Sequences for single domain antibodies were obtained from the patent (WO 2008/141449 A1) Single domain antibodies and heavy chain antibody against EGFR and uses thereof. Figure 12 represents amino acids sequences that were used for design of the single domain antibodies. Inventors of this patent are ZHANG, Jianbing; (CA). MACKENZIE, Colin Roger; (CA). BELL, Andrea; (CA). The present invention relates to the field of antibodies directed towards epidermal growth factor receptor (EGFR). More particularly, the present invention relates to anti-EGFR polypeptides (e.g. single-domain antibodies: sdAb) and nucleic acid sequences encoding same, directed towards and clones thereof, which target EGFR. The invention also concerns an sdAb, which is fused with a crystallizable fragment (Fc) of an immunoglobulin protein in order to generate a chimeric protein. These anti-EGFR proteins can thus be used in the targeting of tumors presenting EGFR on their surface, as well as for diagnosing and treating several types of cancer associated with cells over-expressing EGFR on their surface.

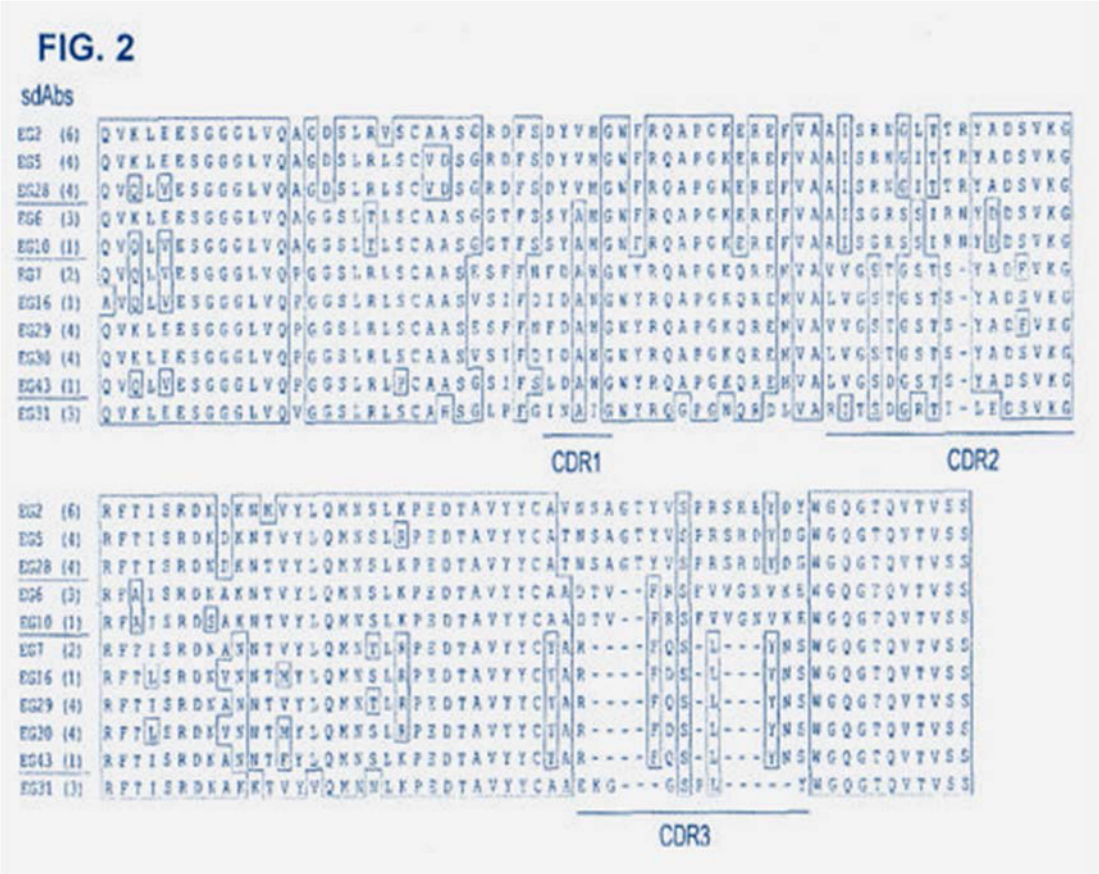


Figure 12 Sequences of single domain antibodies.

2.2.2 Antibody design and optimisation

2.2.2.1 Back-translation of amino acids to DNA

Using Entelechon back-translation tool at

http://www.entelechon.com/2008/10/backtranslation-tool/ the following steps in the number of title tabs were completed:

Protein: Paste in **aa sequence**

Genetic code: Should be set to **Standard**

Codon usage: Under ‚Download codon usage' set to ***Escherichia coli*** and download cut

Edit cut: Leave setting as they are

Optimize: Under ‚General' **discard codons below 10% of theoretical ratio** and Select: distribute to other codons, and then Apply

Motifs: on right hand panel find all **SNAP-ON restriction sites** and add them to **bad motifs**

Output: Under ‚Sequence' select **Start**

The raw DNA sequence is codon optimised for expression in E. coli. The insertion of SNAP-ON restriction sites were blocked and should not appear during synthesis.

2.2.2.2 Optimisation of DNA

The raw DNA sequence from back-translation tool was past into the Word Document. The changes to DNA were applied by the Graphical Codon Usage Analyser at [http://www.gcua.schoedl.de/]

It contains two programmes:

1. Each triplet position vs. usage table – this shows each amino acid in the sequence and gives visual representation of the frequency of the codon usage of each codon. This was utilised to reduce the usage of rare codons or runs of less frequently used codons to prevent pausing or stalling of translation.
2. Each codon usage vs. table – this shows the codons that can be used for each amino acid and compares how often it should be used in a sequence (%) against how often it is used in optimised sequence. This was utilised to choose which codons to use for replacement in the sequence. Figure 13 represents codon usage in *E. coli* system.

The originating organism for each of these programs was set to ‚Synthetic' with *Escherichia coli* codon usage. Each novel designed molecule has a number of iterations in order to get the optimal sequence. The changes of DNA nucleobases were made in the way that the common codons are not overloaded; therefore, the spread generally matches the usage table. The rare codons were avoided as well as less common codons. The optimised sequences were dropped into SeqBuilder (DNASTAR Lasergene 8) program to check if any of SNAP-ON restriction sites have not been introduced during redesign.

The redesign sequences were submitted in the chapter 3. The copies of the each triplet position vs. usage table and each codon vs. usage table can be seen in the appendix.

CODON USAGE IN *E. COLI* GENES[1]

	Codon	Amino acid[2]	%[3]	Ratio[4]	Codon	Amino acid	%	Ratio	Codon	Amino acid	%	Ratio	Codon	Amino acid	%	Ratio	
U	UUU	Phe (F)	1.9	0.51	UCU	Ser (S)	1.1	0.19	UAU	Tyr (Y)	1.6	0.53	UGU	Cys (C)	0.4	0.43	U
	UUC	Phe (F)	1.8	0.49	UCC	Ser (S)	1.0	0.17	UAC	Tyr (Y)	1.4	0.47	UGC	Cys (C)	0.6	0.57	C
	UUA	Leu (L)	1.0	0.11	UCA	Ser (S)	0.7	0.12	UAA	STOP	0.2	0.62	UGA	STOP	0.1	0.30	A
	UUG	Leu (L)	1.1	0.11	UCG	Ser (S)	0.8	0.13	UAG	STOP	0.03	0.09	UGG	Trp (W)	1.4	1.00	G
C	CUU	Leu (L)	1.0	0.10	CCU	Pro (P)	0.7	0.16	CAU	His (H)	1.2	0.52	CGU	Arg (R)	2.4	0.42	U
	CUC	Leu (L)	0.9	0.10	CCC	Pro (P)	0.4	0.10	CAC	His (H)	1.1	0.48	CGC	Arg (R)	2.2	0.37	C
	CUA	Leu (L)	0.3	0.03	CCA	Pro (P)	0.8	0.20	CAA	Gln (Q)	1.3	0.31	CGA	Arg (R)	0.3	0.05	A
	CUG	Leu (L)	5.2	0.55	CCG	Pro (P)	2.4	0.55	CAG	Gln (Q)	2.9	0.69	CGG	Arg (R)	0.5	0.08	G
A	AUU	Ile (I)	2.7	0.47	ACU	Thr (T)	1.2	0.21	AAU	Asn (N)	1.6	0.39	AGU	Ser (S)	0.7	0.13	U
	AUC	Ile (I)	2.7	0.46	ACC	Thr (T)	2.4	0.43	AAC	Asn (N)	2.6	0.61	AGC	Ser (S)	1.5	0.27	C
	AUA	Ile (I)	0.4	0.07	ACA	Thr (T)	0.1	0.30	AAA	Lys (K)	3.8	0.76	AGA	Arg (R)	0.2	0.04	A
	AUG	Met (M)	2.6	1.00	ACG	Thr (T)	1.3	0.23	AAG	Lys (K)	1.2	0.24	AGG	Arg (R)	0.2	0.03	G
G	GUU	Val (V)	2.0	0.29	GCU	Ala (A)	1.8	0.19	GAU	Asp (D)	3.3	0.59	GGU	Gly (G)	2.8	0.38	U
	GUC	Val (V)	1.4	0.20	GCC	Ala (A)	2.3	0.25	GAC	Asp (D)	2.3	0.41	GGC	Gly (G)	3.0	0.40	C
	GUA	Val (V)	1.2	0.17	GCA	Ala (A)	2.1	0.22	GAA	Glu (E)	4.4	0.70	GGA	Gly (G)	0.7	0.09	A
	GUG	Val (V)	2.4	0.34	GCG	Ala (A)	3.2	0.34	GAG	Glu (E)	1.9	0.30	GGG	Gly (G)	0.9	0.13	G
	U				C				A				G				

[1] The data shown in this table is from the Arabidopsis Research Companion on the World Wide Web (//weeds/mgh.harvard.edu). Codon frequencies for many other bacteria can be found at http://morgan.angis.su.oz.au/Angis/Tables.html.

[2] The letter in parenthesis represents the one-letter code for the amino acid.

[3] % represents the average frequency this codon is used per 100 codons.

[4] Ratio represents the abundance of that codon relative to all of the codons for that particular amino acid.

Figure 13 Codon Usage in *E. coli* genes.

2.2.3 Gene synthesis

Entelechon GmbH (synthetic genes company) based in Regensburg, Germany, synthesized single domain antibodies as well as single chain antibody fragments: [http://www.entelechon.com]. The genes were received in the tubes containing 5-10 µg of lyophilised DNA. The exact concentrations of genes are stated on the gene synthesis report.

The DNA was dissolved in 50 µl distilled water by carefully shaking for 10 minutes. The final DNA concentration was approximately 100-200 ng/µl.

2.2.4 Archive of newly synthesised DNA

The newly synthesized DNA has a summary of design process and sequence selection recorded in the lab book belongs to Syntaxin Ltd. When DNA was delivered, it was archived using DNA and microbank archive protocol. The lab book references and an annotated sequence of that DNA were recorded in the microbank databases on the BIX home page at http://192.168.0.6/BIX_Home.html. Each newly synthesized DNA has a unique clone number SXN10xxxx. Each newly created an annotate sequence saved as a SeqBuilder file is named with the clone number A00xxxx. Clones were then archived using the following Protocol for microbank and DNA stock.

2.2.5 DNA stock and microbanks

2.2.5.1 Microbank clone stocks

The delivered synthesised DNA or sequence confirmed new chimera DNA was transformed with TOP10 cells. 5 µl plasmid DNA with 25 µl TOP10 cells were treated as described in the TOP10 cell method in the section 2.2.5.3

For each clone green, yellow and red microbank tubes was labelled with the clone number (A00xxxx) and the construct name.

A streak from the transformation plate was then taken and added to each of the microbank tubes. The tubes were shaken to coat the beads with cells, left for 10-15 minutes and the liquid was removed.

The completed microbank tubes were placed into the appropriate boxes in the -80°C freezers.

2.2.5.2 DNA clone stocks

40 µl of the plasmid DNA (at a concentration of approximately 100 ng/µl, i.e. the standard concentration from a plasmid miniprep) from the delivered synthesised DNA or the sequence confirmed new chimera DNA was placed in a tube labelled with the clone number and construct name.

The tube was placed in the designated box and stored at -20°C.

The DNA stock and the yellow and red microbanks are archive stock and should not be used as working stocks.

The green microbank is the working stock and was used for preparation of any plasmid DNA required.

2.2.5.3 Transformation to TOP10 cells

One Shot® TOP10 Chemically Competent *E.coli* (Cat. No. C4040-06) were purchased from Invitrogen, Inchinnan Business Park, Paisley, UK.

TOP10 *E.coli* is provided at a transformation efficiency of 1 x 10e9 cfu/µg supercoiled DNA and are ideal for high-efficiency cloning and plasmid propagation. It allows stable replication of high-copy number plasmids. The genotype of TOP10 cells is similar to the DH10B™strain. Each new batch of TOP 10 cells is check by transformation of pUC19 control plasmid DNA supplied with the kit to verify the efficiency of competent cells. The test transformation was prepared according to the protocol provided in the kit.

Calculating Transformation Efficiency

Use the following formula to calculate the transformation efficiency as transformants (in cfu) per µg of plasmid DNA. The total volume of the transformation mixture is 300 µl.

Transformation efficiency (# transformants/µg DNA) =

$$\frac{\#ofcolonies}{10pgpCU19DNA} \times \frac{10^6\,pg}{\mu g} \times \frac{300\,\mu ltotalvolume}{X\ldots\mu lplated} \times dilution\ factor$$

For example, if transformation of 10 pg of pUC19 DNA yields 100 colonies when 30 µl of a 1:10 dilution is plated, then the transformation efficiency is:

$$\frac{100 colonies}{10 pgDNA} \times \frac{10^6 pg}{\mu g} \times \frac{300 \mu l total volume}{30 \mu l plated} \times 10 = 1 \times 10^9$$

Transformation of synthesized DNA or new chimera DNA

Thaw on the ice one-shot competent TOP 10 cells for each transformation

Add 5 µl of ligation mixture to 50 µl Top 10 cells and mix by flicking the tube

Incubate vials on ice for 20 minutes

Heat shock the mixture by immersing in 42˚C water bath for 45 seconds.

Return the mixture to the ice for 10 minutes (be careful not to agitate the mixture as this could shear the DNA being taken up by the cells)

Add 150 µl of LB to the mixture and incubate at 37˚C for 45-60 minutes at 225 rpm in a shaking incubator

Plate the mixture onto an agar plate with kanamycin to dry for 30 minutes

Place the plate into the 37˚C incubator and leave overnight. (Alternatively leave the plate on the bench over the weekend.)

2.3 Cloning of antibody-LH_N chimeras

2.3.1 Restriction digest

The inserts and backbones were digested by restriction enzymes *XbaI* and *Hind III* at 37˚C for 3 hours.

Table 3 Reaction mixture for double digestion (50 µl)

Reagent	Volume (µl)
DNA	22
BSA (10×)	5
NEB buffer	5
Xba I	3
Pst I	3
dH_20	12
Total volume	50

The successfully cleaved a DNA substrate with two restriction endonucleases (double digestion) depends on selecting the best NEB buffer to provide reaction conditions that optimize enzyme activity as well as avoid star activity associated with some enzymes.

NEB provides four standard colour-coded NEB buffers (1, 2, 3 or 4) for optimal activity. The NEB buffer 2 (blue, cat. no. B7002S) was use for purpose of this cloning.

The composition of NEB buffer 2 (1×)

10 mM Tris-HCl, 10 mM $MgCl_2$, 50 mM NaCl, 1 mM dithiothreitol, (pH 7.9 @25˚C)

The best enzyme combination and choice of appropriate buffer can be found in the Double Digest Finder at www.neb.com

To obtain 100 % activity the bovine serum albumin (BSA×100, cat. no: B9001S, NEB, UK) was added to reaction to a final concentration of 100 µg/ml.

2.3.2 Agarose gel electrophoresis

The digested DNA samples were run on 0.8 % agarose gel.

2.3.2.1 Agarose gel preparation

The gel tank Bio-Rad with appropriate size comb was assembled.

1.2 g Agarose (Agarose MB 1200, cat no: 9012-36-6, Melford Laboratories) and 150 ml with 1x TAE buffer (Tris-Acetate-EDTA 50× buffer, cat no: BP1332-20, Fisher Scientific) was mixed in a conical flask. It made up 0.8% Agarose gel. The mixture was heated in the microwave for 4-5 minutes until Agarose completely dissolved and cool in water down for about 20-30 seconds.

Subsequently, 15 µl of SYBR® (Safe DNA gel stain, cat no: S33102, Invitrogen) was added to 0.8% agarose gel poured into the sledge. The gel was ready to use after 30-45 minutes.

2.3.2.2 Sample preparation

20 µl Orange G (x5) (cat no: O3756, Sigma Aldrich) added to each digest. Orange G is a tracking dye in nucleic acid gel electrophoresis, running approximately at the size of a 50 base pair (bp) DNA molecule.

Added 10 µl of Orange G to each digested sample, therefore samples in the volume of 60 µl were loaded onto the agarose gel.

2.3.2.3 1kb DNA Ladder preparation

1 kb DNA ladder (cat no: N3232L, NEB) yields 10 bands suitable for use as molecular weight standards for agarose gel electrophoresis. The digested DNA includes fragments ranging from 0.5-10.0 kilobases (kb). The 3.0 kb fragment has increased intensity to serve as a reference band. The 1 kb DNA ladder is not design for precise quantification of DNA mass, but it can be use for approximating the mass of DNA in comparably intense samples of similar size.

Add 2 µl of 1Kb DNA ladder to 4 µl Orange G and 14 µl dH_2O.

1 kb DNA ladder in the volume of 20 µl was loaded onto the Agarose gel.

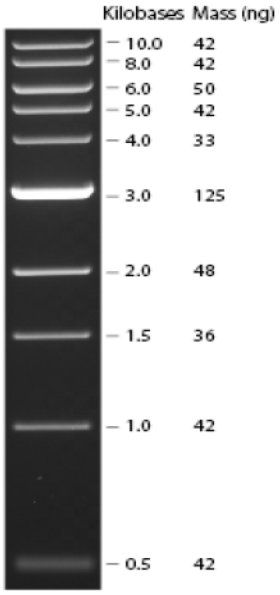


Figure 14 1 kb DNA Ladder visualized by ethidium bromide staining on a 0.8% agarose gel. Mass values are for 0.5 µg/lane.

2.3.2.4 Agarose gel run

The wells were set at the negative end of the tank and the agarose gel run for 1 hour at 120 V.

2.3.2.5 The extraction and purification of DNA from the agarose gel in TAE buffer.

The bands on the gel were visualized under U.V. light in the Gene Flash from the Syngene Bio Imaging. The desire bands were cut from the gel out on the Safe Imager from the Invitrogen and the taken picture of agarose gel confirmed correct sizes of bands.The extraction and purification of DNA was prepared according to the QIAquick Gel Extraction protocol provided with kit (www.QIAGEN.com, cat. no. 28706)
The 30 µl of eluted DNA was used for the ligation.

2.3.3 Ligation

The reaction mix shown in table 4 was incubated at 16˚C overnight or at 21˚C for 3 hours and then entire ligation product was transformed to TOP 10 cells (refer to section 2.2.5.3). The T4 DNA ligase together with buffer (10×, with 10 mM ATP) was purchased from (cat. no. B0202S) NEB, UK,

Table 4 Reaction mixture for ligation (20 µl)

Reagent	Volume (µl)
Insert (Abs domain or fragment)	14
Backbone (LH_N)	2
T4 DNA Ligase buffer	2
T4 DNA Ligase (400,000 U/ml)	2
Total volume	20

The concentration of insert and backbone is approximately 100 ng/µl. U stands for Unit.

2.3.4 Grow overnight and DNA purification of newly designed molecules

Preparation of agar plates

The 200 ml of solid agar (cat. no. 44391, BioMerieux, Basingstoke, UK) was dissolved in the microwave for 10 minutes. Subsequently, it was cooled down to the appropriate temperature, therefore 200 µl of 30 µl/mg kanamycin (Kanamycin sulphate from *Streptomyces kanamyceticus,* cat. no. K4000-5G, Sigma-Aldrich, UK) was added and mixed with agar. The ready agar was poured into Petri dishes.

E. coli Growth overnight

The transformed to TOP 10 cells ligation mixtures were plated on the agar plate with kanamycin that were left at 37˚C incubator, overnight. Two colonies for each molecule were picked from the plates up and inoculated in 10 ml LB (Luria Bertani, cat. no. 44393, BioMerieux, Basingstoke, UK) with 10 µl of 30 µl/mg kanamycin. The inoculated cultures were grown at 37˚C, overnight on the shaking platform.

Minipreps DNA purification

The DNA was purified from 10ml overnight cultures using Wizard® Plus SV Minipreps DNA Purification System provided with kit from the Promega, UK (cat. no. A1460).

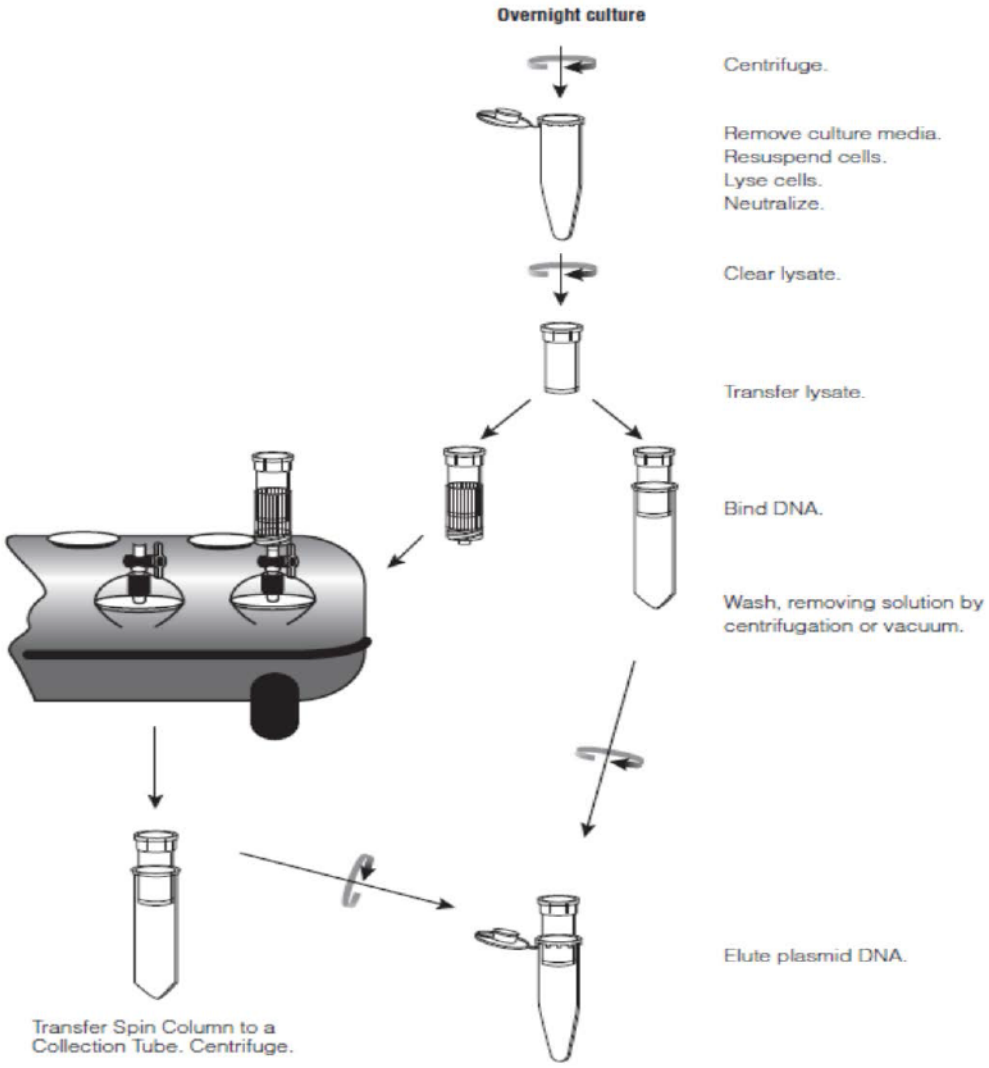


Figure 15 Flow diagram of plasmid DNA isolation and purification using the Wizard® Plus SV Miniprep DNA Purification System.

2.3.5 The screening test for purified DNA

The newly cloned molecules were tested by digestion with *Xba I* and *Hind III* restriction enzymes in order to confirm correct insertion of new genes.

Table 5 Reaction mixture for sequential digest

Digest mixture for sequential digest

Reagent	Volume (µl)
BSA (10×) Stock conc. 10 mg/ml	1
NEBuffer 2	1
DNA (100 ng/µl)	4
Hind III 20,000 U/ml	1
dH_2O	3
Total volume	10

Table 6 Reaction mixture for double digest

Digest mixture for double digest

Reagent	Volume (µl)
BSA (10×) Stock conc. 10 mg/ml	1
NEBuffer 2	1
DNA (100 ng/µl)	4
Hind III 20,000 U/ml	1
Xba I 20,000 U/ml	0.5
dH_20	2.5
Total volume	10

2.3.6 DNA sequencing

Tested molecules in the volume of 20 µl were sent to Sources Bioscience (former Geneservice) based in the Oxford, UK, http://www.lifesciences.sourcebioscience.com/. Molecules were sequenced using Geneservice universal primers and custom primers designed in house. Universal primers that are suitable for sequencing are T7 forward

and T7 terminator. The choice of custom primers was based on the type of serotype to be screened.

The DNA sequencing data were obtained in the DNA Star Lasergene 8 format and they are stored on the Syntaxin Ltd. database. MegAlign programme was used to align Sources Bioscience sequences with predicted ones created on the SeqBuilder programme.

The copy of all alignments can be found in the lab books. Each new clone was then archived and microbanked according to the protocol (refer to section 10.2.4 and 10.2.5).

2. 4 Expression of recombinant antibody based constructs

2.4.1 Modified Terrific Broth

The expression were performed using modified Terrific Broth (mTB) purchased from Melford, UK. The modified Terrific Broth is supplied in a dry form in the capsules.

Table 7 mTB composition

Modified Terrific Broth Media composition

Components	Formula for one capsule (g/l)
Tryptone	12
Yeast extract	24
di-potassium phosphate	9.4
Mono-potassium phosphate	2.2

Preparation of 1L expression media

2 capsules, 4 ml of glycerol and 1 litre of warm distilled water. The water was warmed up to 37˚C, so gelatine capsules can be easily dissolved. The capsules need to be dissolved prior to autoclaving of mTB media. The media was autoclaved in sterile conditions at 121˚C for 15 minutes.

2.4.2 Expression strain stock

Chimeras transformed with an expression *E. coli* strain for protein purification were set up from transformation plate into a single blue microbank tube. A unique expression stock numbers E00xxxx were obtained from the expression microbank databases on the BIX Syntaxin Ltd. home page. All tubes were labelled with the unique expression stock number, the construct name and cell strain.

BL21 (DE3) Competent cells are T7 expression host strain that were purchased from Novagen, UK (cat. no. 69450-4). They are supplied in 0.2 ml aliquots and can be used for 10 transformations.

Bacterial Strain Genotype of BL21 (DE3) is F– *ompT hsdS*B(rB– mB–) *gal dcm* (DE3)

The guaranteed transformation efficiency is 2 × 106 cfu/µg of Test Plasmid.

Transformation to BL21 (DE3) cells

5 µl DNA was added to 20 µl BL21 (DE3) cells. The transformation procedure was the same as described in the section 10.2.5.3. However the efficiency of BL21 (DE3) is higher than TOP 10 cells, therefore 50 µl of transformation mixture was plated onto the agar plate. The lower volume on the plate should give single colonies and prevent the plate from over growing. The streak of colonies was taken on the loop and mixed with microbank broth of the blue tube. The tube was shaken to disperse cells and left for 10-15 minutes. Eventually the broth was removed and microbank was stored at -80˚C freezer labelled with number E00xxxx. The expression numbers are available at BIX Syntaxin Ltd. home page.

2.4.3 Standard expression protocol

The expressions were prepared on the shaking platform in the incubator from the New Brunswick Scientific, US.

100 ml modified TB + 0.2% glucosamine + 100 µg/ml ampicillin/ or 30 µg/ml kanamycin in 250 ml flasks was inoculated using a microbank bead from the LH_N-ligand expression strains or a single colony from a streaked agar plate. The cultures grew at 37˚C, 225 rpm for 16 hours. The OD_{600nm} was checked to ensure that readings are

between 0.1 and 0.5 OD by dilution with fresh mTB, so a record of the growth conditions can be maintained.

After 16 hours 1-litre cultures of modified TB + 0.2% glucosamine + 100 μg/ml ampicillin/ or 30 μg/ml kanamycin in 2-litre flasks were inoculated with 10 ml of overnight culture for each DNA. The cultures grew at 37˚C, and 250 rpm until an approximate OD_{600nm} of 0.5 was reached at which point the temperature was turned down to 16˚C while continuing to shake at 250 rpm. 1 hour later the OD_{600nm} was recorded to ensure reading between 0.1 and 0.5 OD by dilution with fresh broth. On that point, the cultures were induced with 1 mM IPTG. The cultures grew for approximately 20 hours at 16˚C, 250 rpm.

The final OD_{600nm} was measured to ensure that readings are between 0.1 and 0.5 OD diluted in TB, so a comparison of the final cell density of the cultures could be made.

The cultures were spun down at 5000 rpm, 4˚C for 20 minutes. The supernatant was removed as much as possible and each litre of cell paste was re-suspended in 20 ml room temperature lysis buffer (50 mM HEPES pH 7.2, 200 mM NaCl). The lysed cell pastes were placed in 50 ml falcon tubes labelled with construct, expression strain, weight, date, and cell expression number.

The cell pastes were weighed by subtracting the pre-weighed 20 ml of buffer and falcon tube from the total mass of the cell paste mixed with buffer and frozen at -80˚C.

2.4.4 Changes applied to the standard expression protocol for purpose of study

To obtain higher enrichment of expressed proteins the harvested cell pastes were re-suspended in the lysis buffer 50 mM HEPES pH 7.2, 500 mM NaCl and then they were spun down at 4˚C, 12500 rpm for 30 min to remove as many contaminates and debris as possible. The supernatant was removed and cell pastes were frozen dry at -80˚C freezer.

2.5 Purification of recombinant antibody based constructs

The production of target proteins was prepared using two purification systems AktaExpress and AktaPurifier bought from the GE Healthcare, UK & Ireland. The resins

and columns were purchased from GE Healthcare, UK and all product codes can be find on this website:
http://www.gelifesciences.com/aptrix/upp01077.nsf/content/protein_purification.
A full description of the purification protocol is found below.

2.5.1 Cleavage enzymes

New England BioLabs, USA supplies cleavage enzymes: EK (Enterokinase) and FXa (Factor Xa) upon request. Each new batch is tested at Syntaxin Ltd. and acceptance to general use is made on the outcome of successful activation of molecule.
The Recombinant Human Coagulation Factor X bought from the R&D System was used for purpose of this study.

Table 8 Activation enzymes

Enzyme	Source	Molecular Mass (kDa)
Factor Xa (FXa)	Purified from bovine plasma and activated by treatment with the activating enzyme from Russell' viper venom	**43** (consists of two disulfide-linked chains of approximately **27** and **16**) On SDS-PAGE the reduced chains have apparent MW of **30** and **20**
Recombinant Human Coagulation Factor X (rhFX)	Expressed with a C-terminal 10× His tag in an insect cell lines, Sf21; the secreted pro enzyme was purified and activated	Consists of heavy chain/catalytic domain and two fragments of light chain, On SDS-PAGE under reducing conditions migrates at **33-36** and **13-14**
Enterokinase (EK)	Purified from *K.lactis* containing a clone of the light chain of the bovine Enterokinase gene	**26.3** On SDS-PAGE gel is **31**

2.5.2 Metal Ion affinity chromatography (IMAC) using AktaExpress

The appropriate cell paste was thawed from -80˚C and re-suspended in the buffer 50 mM Hepes pH 7.2, 200 mM NaCl. The re-suspended cell paste was homogenized at 4˚C, 15000 psi on the Constant System homogenizer and it was spun down at 13000 rpm at 4˚C for at least 1 hour. The lysate was clarified by filtration using 0.2 µm filter units on the Watson-Marlow peristaltic pump. The supernatant was loaded at 3 ml/min onto pre-packed 5 ml His Trap HP column (GE Healthcare) pre-charged with Ni Sepharose High Performance and equilibrated with 50 mM Hepes pH 7.2, 200 mM NaCl, 20 mM Imidazole.

After loading lysate onto the column, the column was washed with 15 CV (column volumes) of equilibration buffer. The fusion protein was eluted with different concentrations of Imidazole. The 40 mM Imidazole wash was set up to 15 CV, while the 250 mM Imidazole elution step was set up to 12 CV. The protein eluted at 4 ml/min. The buffer 50 mM Hepes pH 7.2, 200 mM NaCl was used during purification process. The eluted protein at 250 mM Imidazole was stored in loops on the AktaExpress system and passed through desalt column for the buffer exchange to 50 mM Hepes pH 7.2, 150 mM NaCl. Consequently, the fusion protein was activated on the same day as the capture purification step.

5.2 µl of EK was added per 1 mg of fusion protein and activated at 25˚C for approximately 10 hours or 1.5U of FXa was added per 1 mg of fusion protein and activated at 4˚C for 10 hours. A Nanodrop measured the concentration of fusion protein at A_{280nm}.

The next day the presence of the precipitations in the fraction was visible. It was required to spin the sample down in the centrifuge for 15 min, at 4˚C, at 4000 rpm. In order to estimate the loss in yield the concentration of protein was checked in the sample before and after spin.

The second His column has the same parameters and buffer conditions as first His column. Eluted fusion protein was dialysed at 4˚C, o/n in 5L of 50 mM Hepes, and 150 mM NaCl buffer.

On the third day of the purification process, the dialyzed fusion protein was concentrated to approximately 1 mg/ml in the VivaSpin concentrators and frozen at -20˚C freezer. Each new purified molecule has a batch number made of initials, reverse date. All information regarding to process and outcome of purified molecule are stored on the BIX home page at Syntaxin Ltd.

2.5.3 Metal Ion affinity chromatography (IMAC) using AktaPurifier

The purification process using AktaPurifier was slightly different from that using AktaExpress. The cell pastes used for this purification process were washed in the buffer and frozen dry at -80˚C. The thawed cell pastes were re-suspended in the buffer 50 mM Hepes pH 7.2, 500 mM NaCl and homogenized at 4˚C, 15000 Psi on the Constant System homogenizer, then it was spun down at 12500 rpm at 4˚C for at least 1 hour. Lysate was clarified by filtration using 0.2 µm filter units on the Watson-Marlow peristaltic pump. The XK16 (GE Healthcare) column was packed with 10 ml Chelating Sepharose resin and charged with a 0.1 M $NiSO_4$. The equilibrated column with 50 mM Hepes pH 7.2, 500 mM NaCl was ready to use for purification. The clarified lysate by filtration using 0.2 µm filter units was loaded onto a column at 4 ml/min. The fusion protein was eluted with step-wise gradient of Imidazole (40 mM, 80 mM and 250 mM) in the buffer 50 mM Hepes pH 7.2, 500 mM NaCl at 4 ml/min.

The target protein eluted at 250 mM Imidazole concentration. The AktaPurifier system can be use only for one column at the time; therefore, buffer exchange in the fusion protein was prepared by dialysis at 4˚C, for 5 hours. The composition of buffer depends on the activation enzyme used and 5L of 50 mM Hepes pH 7.2, 50 mM NaCl was applied when activated with EK or 5L of 50 mM Hepes pH 7.2, 150 mM NaCl buffer was used while activated with FXa.

To activate the toxin 3.2 µl of EK per 1 mg of fusion protein was added and incubated at 4˚C for 10 hours with an additional 1.6 µl of EK per 1 mg of protein added for 4 hours prior to purification alternatively, 1.5 U of FXa per 1 mg of fusion protein was added and incubated at 4˚C for 10 hours. After 3 to 4 hours, activation of Enterokinase is lost. After the activation process, the fusion protein was purified by Hydrophobic Interaction Chromatography (HIC), which is described below.

2.5.4 Hydrophobic Interaction Chromatography (HIC) using AktaPurifier

The XK16 (GE Healthcare) column was packed with 30 ml Phenyl Sepharose 6 Fast Flow resin and equilibrated with 50 mM Hepes pH 7.2, 1 M ammonium sulphate. 1 M ammonium sulphate in the solution was added to activate the fusion protein during the polishing step or to a clarified lysate for initial capture step. The fusion protein was eluted at 3 ml/min with a linear gradient of increasing buffer B concentration; where buffer B was 50 mM Hepes pH 7.2 and buffer A was 50 mM Hepes pH 7.2, 1 M ammonium sulphate.

The samples were analyzed on the SDS-PAGE gel under reducing conditions by addition of 0.1 M DTT (dithiothreitol) and non-reducing conditions if the activation process was completed. This way the only well activated samples were chosen for further purification. The eluted fusion protein was dialysed against 5 L of 50 mM Hepes pH 7.2, 150 mM NaCl at 4˚C, overnight. Subsequently, the fusion protein was concentrated in the VivaSpin concentrators to approximately 1 mg/ml and stored at -20˚C or -80˚C freezer. The recording process of newly purified protein is the same as described in the chapter 10.5.2.

Throughout the initial capture process, the analyzed fusion protein was discarded, as only this stage was important for the purpose of study.

2.5.5 Anion Exchange Chromatography – Q Sepharose FF

The anion exchange chromatography was used in the initial capture step as a possibility to avoid unspecific cleavage of the single chain antibody fragments - LH_N chimeras. Anion-exchange resins are positively charged and bind (and exchange) negatively charged ions (anions). The column XK16 (GE Healthcare) was packed with 20 ml of Q Sepharose Fast Flow resin. The column was equilibrated with two buffers. Initially, in order to charge column it was equilibrated with 50 mM Hepes pH 7.2, 1 M NaCl and next with 50 mM Hepes pH 7.2. The clarified lysate was re-suspended in 50 mM Hepes pH 7.2, 50 mM NaCl buffer, loaded onto a column at 3 ml/min, and eluted with decreasing concentration of 1 M NaCl in 50 mM Hepes pH 7.2. The sample was

analyzed by the SDS-PAGE. The purification was completed at this stage and fusion protein was discarded.

2.5.6 Protein Analysis

2.5.6.1 SDS Polyacrylamide Gel Electrophoresis (SDS-PAGE)

Separation of protein to determine content and purity was carried out using a discontinuous Tris-glycine SDS-PAGE (Sodium Dodecyl Sulphate) system. 10 or 12 well 4-12 % NuPAGE Bis-Tris commercial gel was placed in the BioRad gel tank and filled up with 1× MOPS SDS Running Buffer (20×, cat. no: NP0001-02, Invitrogen) diluted with dH_2O.

Sample under non-reducing conditions

50 µl of tested protein sample

25 µl of NuPAGE LDS sample buffer (4×, cat. no: NP0007, Invitrogen)

25 µl of dH_2O

Sample under reducing conditions

50 µl of tested protein sample

25 µl of NuPAGE LDS sample buffer (4×, cat. no: NP0007, Invitrogen)

15 µl of dH_2O

10 µl of 1M DTT (100 mM in reduce sample)

10 µl of sample was loaded onto a 4-12% NuPAGE Bis-Tris stacking gel. The gel was run at 200 V (volts) for 45 minutes. It was stained with SimplyBlue™ SafeStain buffer (cat. no: LC6065, Invitrogen). After electrophoresis, the gel was placed in 100 ml of deionized water and microwaved for 2 minutes to remove SDS and buffer salts. This step was repeated again and then gel was covered with 20 ml of SimplyBlue™ SafeStain buffer and microwaved for the additional 1 minute. The gel was shaking on the orbital shaker until bands were visualized. After detection of bands, the SafeStain buffer was removed with dH_2O and G: BOX from Syngene Bio Imagine took a picture of gel.

Protein ladders

5 µl of Benchmark ladder (cat. no: 10747-012, Invitrogen) was used for story gels and His tag Western blots. 3 µl of Magic Mark (cat. no: LC5602, Invitrogen) was used for any another blots prepared for experiments related to this study. The only exception is Biotinylated protein ladder (cat. no: 7727L Cell Signalling) used for internalization study in the volume of 5 µl.

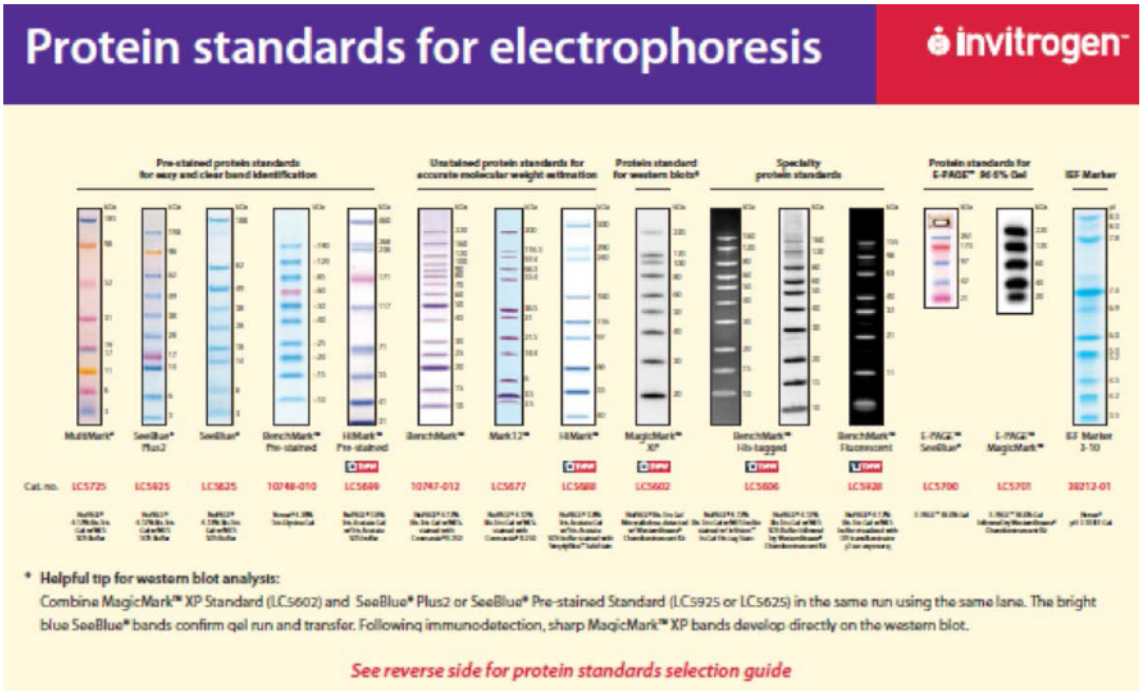


Figure 16 Protein Standards for electrophoresis.

2.5.6.2 Spectrophotometric determination of protein concentration

The concentration of total protein in a sample was determined using the Beer Lambert law.

$A=\varepsilon C l$

Where A – Absorbance

ε – Molar extinction coefficient

C – Protein concentration

L – Path length

The A_{280nm} reading was taken by a Nanodrop. The molecular weight of the protein and the molar extinction coefficient were calculated on the ExPASy proteomics server ProtParam tool: http://www.expasy.org/tools/protparam.htlm.

2.5.6.3 Purity and activation assessment

The purity and activation assessments were prepared using densitometry and recorded in percentage. The final sample at 0.1 mg/ml loaded in 10 µl on the SDS PAGE gel was used to determined purity. The sample at 0.1 mg/ml in reducing conditions (100 mM DTT) loaded in 10µl on the SDS PAGE gel was used to determined activation.

2.5.6.4 Western blot system

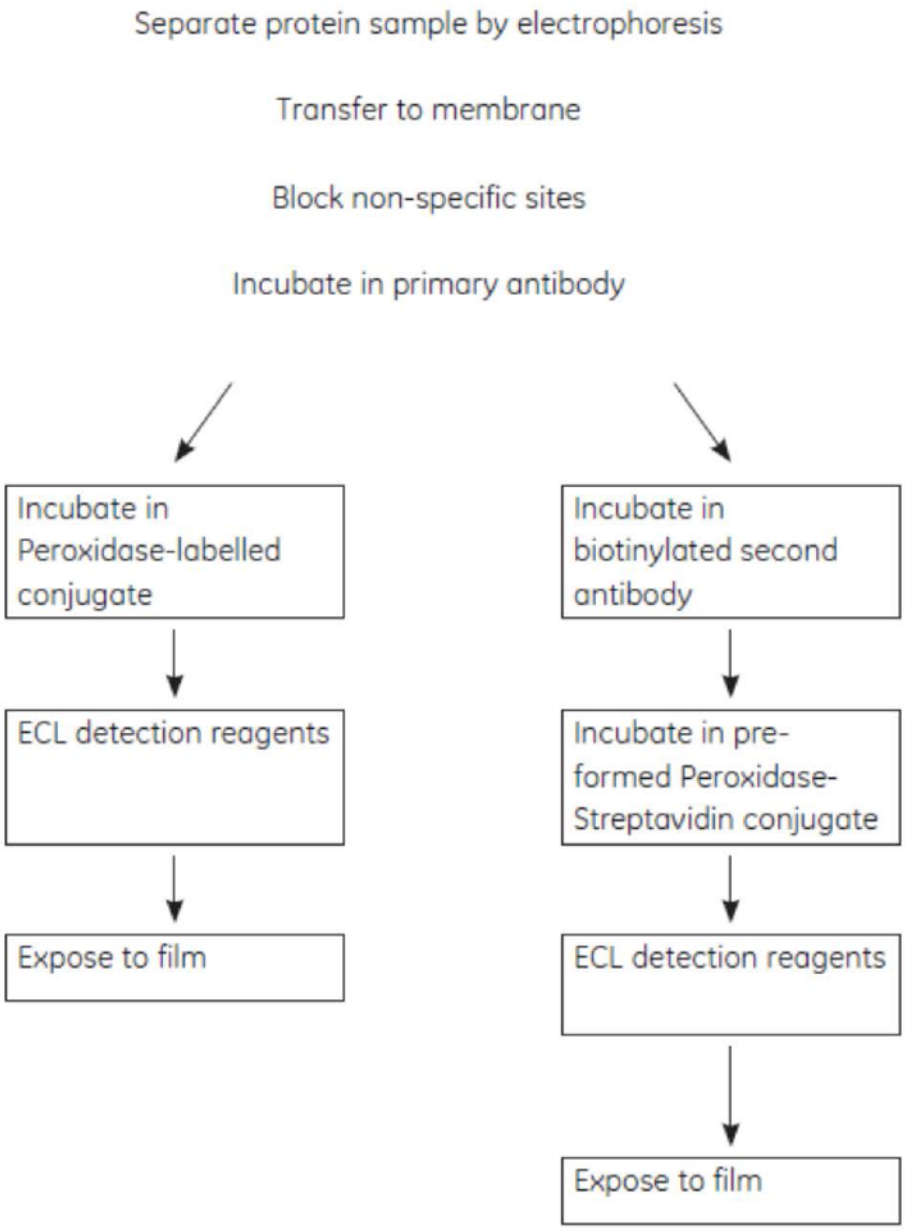


Figure 17 Western blot system.

2.5.6.4.1 Electroblotting procedure

The electroblotting procedure was used for transferring proteins to membranes. The protein was separated by gel electrophoresis. Next, the gel was soaked in the protein transfer buffer.

Protein transfer buffer

200 ml Tris – Glycine SDS-Running Buffer (10×, cat. no: LC2675-5, Invitrogen)

400 ml Methanol

1400 ml Distilled water

The total volume for one transfer tank is 2000 ml

The protein was transferred to pre-cut Nitrocellulose Membrane (cat. no: LC2000, Invitrogen). The electroblotting cassette was assembled and placed between electrodes in the blocking unit in the way that a protein was migrating through electric current from cathode to anode. The wet transfer took 1 hour at 400 mA with variable volts.

2.5.6.4.2 Immunodetection

Blocking the membrane

Non-specific binding sites were blocked by immersing the membrane in the blocking reagent: 0.5% BSA (Albumin from bovine serum, cat. no: A8022-500G, Sigma) or 3 % Marvel (dried skimmed milk, TESCO) in PBS -T (1× PBS – Phosphate Buffered Saline (10×), 0.1 % Tween (20%)) for at least 1 hour or overnight at 4˚C on the shaking platform.

Washing membrane

Prior to application of antibody the membrane was washed for 30 minutes in the PBS – Tween changed three times during this time.

Primary and secondary antibody

Purification sample were tested using:

1/2500 dilution His Tetra Antibody (cat. no: 34670, QIAGEN) – 1/1000 dilution anti-Mouse

1/1000 dilution Light chain antibody produce against specific serotype of LH_N molecule – 1/1000 dilution anti – Rabbit

Internalization was tested using:

1/2000 dilution Streptavidin-HRP (Sigma)

Activation was tested using:

1/1000 dilution phospho-MAPK (#9101Cell Signalling Technology)-1:1000 anti-Rabbit
1/1000 dilution MAPK (#9102 Cell Signalling Technology)-1:1000 anti-Rabbit
1/1000 dilution phospho-EGFR (#2236 Cell Signalling Technology)-1:1000 anti-Mouse
1/1000 dilution EGFR (#2232 Cell Signalling Technology)-1:1000 anti-Rabbit
1/500 dilution β-actin (Abcam ab8227)-1:1000 anti-Rabbit

Secondary antibodies were purchased from Sigma.

Anti-Rabbit IgG (whole molecule), Peroxidase developed in Goat, cat. no: A6154-1ML

Anti-Mouse IgG (γ-chain specific), Peroxidase developed in Goat, cat. no: A3673-1ML

After primary and secondary antibody, the blots were washed three times in PBS-T for 10 minutes.

Detection

1 ml of ECL detection reagent (Super Signal west duration substrate) was applied (0.5 ml reagent A + 0.5 ml reagents B) on each blot. After 5 minutes, each blot was developed inside the GENOME (Syngene Bio Imagine) for 3 to 5 minutes what depends on the used antibody.

2.5.6.5 N-terminal protein sequencing

Alta Bioscience at University of Birmingham prepared the order of amino acids at N-terminus of sample. The laboratory uses the Edman N-terminal sequencing method. The detection of protein sequencing is only possible by transferring the protein to a PVDF membrane. PVDF membranes were blotted in 10 mM CAPS (3-(Cyclohexylamino)-1-propanesulfonic acid) pH 11.0 and stained by Ponceau S solution (cat. no: P7170-1L, Sigma) to visualized bands on the membrane. The membranes were prepared at Syntaxin Ltd.

2.6 Activation studies

2.6.1 Reagents

Media – Gibco 31966 (should be left 1 h at the room temperature and half an hour in the water bath at 37˚C before actual experiment).

FBS- 50 ml takes half an hour to be defrosted; it should be filtered using 20 ml syringe and 0.2 µm filter,

A549 (Human cells) - EGF receptor activation,

EGF-SIGMA E9644-2MG

2.6.2 Preparation of EGF

Used human, recombinant EGF, expressed in *S. Cerevisiae* 50 ng/ml, (6 kDa) purchased from Sigma; cat. no. E1264.

Reconstitute the contest to 1 mg/ml using filtered 10 mM acetic acid. Dilution to lower concentrations (not less than 10 pg/ml) will require addition of 0.1% BSA.

1mg/ml was reconstituted in 200 µl 10mM acetic acid (166.7 µM = 166667 nM)

2.6.3 Dilution of EGF in media

32 ml of media was added to 19.2 µl of EGF stock (1 mg/ml) **to get 100 nM stock**,

9 ml of media was added to 1 ml of 100 nM stock **to get 10 nM stock**,

27 ml of media was added to 3 ml of 10 nM stock **to get 1 nM stock,**

The 10 nM stock solution will be discarded.

All dilutions were prepared under fume cupboard.

2.6.4 Preparation of plates for activation experiment

Seven plates were prepared for this experiment:

Plate nr 1- serum - free control

Plate's nr 2, 3 – 1 nM, and 100 nM keep for 5 min. in the 37˚C incubator to detect phospho-EGFR

Plate's nr 4, 5 – 1 nM and 100 nM keep for 20 min. in the 37˚C incubator to detect phospho-ERK ½

Plate's nr 6, 7 – 1 nM and 100 nM keep for 45 min. in the 37˚C incubator to detect phospho- ERK $_{1/2}$

Plates were prepared two days before experiment; amount of plates depends on time of experiment and dilutions of ligand or fusion protein, the media should be changed in the plates each day,

1. Plates were removed (serum-free media) from the incubator 37˚C and cell grow was checked under microscope, positive results suggested about the continuation of the experiment,
2. Media were discarded from the plates under fume cupboard.
3. Plates were numerate and 10 ml of media + EGF was added to each plate (the best start with plate of the lowest to the highest concentration). It is important to keep pipette on the side of Petri dish when adding media + EGF to avoid damage of cells.
4. Plates were placed at the 37˚C in the incubator and time was counted from that moment.
5. In the meantime in the fume cupboard the following reagents and products were placed : bucket with ice to stop phosphorylation, 50 ml of dH_2O in the 50 ml falcon tube, seven eppendorfs (amount depends on the amount of plates), 10 ml sterile pipettes, sterile scraper, piece of tissue, 200 µl and 500 µl pipettes and sterile tips for them, squeezing pipettes, bijou. PBS 1% from the fridge was placed on the ice under fume cupboard as well as 500 µl of lysis buffer (10×, cat. no: 9803, Cell Signalling Technology).
6. Preparation of Cell Lysis Buffer: 4.5 ml of dH_2O add 500 µl of Lysis Buffer (10×).
7. After first 5 min, two plates were removed to stop phosphorylation on ice.
8. Petri-dishes were washed twice with 10 ml of PBS-1% applied on the edge of plate avoid damage of the cells. Plates were place in the vertical position on the side of bucket and squeezed all PBS-1% out.
9. 200 µl of Lysis Buffer was applied on the plate for a few minutes. All cells were scraped from the edge of plate and next towards the middle of the plate surface. Cells were moved to one point of plate and transferred to eppendorfs using sterile pipette.

10. Steps from number 7 were repeated for another nine plates.
11. Control plate was prepared on the end of experiment.
12. All lysates were kept in the freezer -20˚C.

2.6.5 BCA assay and Western blots

2.6.5.1 BCA assay

The concentration of protein was determined by BCA™ Protein assay according to Pierce protocol. The microplate procedure was used with Maxisorp Nunc-Immunoplate.

The BCA standards were 0.1 mg/ml; 0.2 mg/ml; 0.4 mg/ml; 0.6 mg/ml; 0.8 mg/ml; 1.0 mg/lm; 2.0 mg/ml.

The lysate were thawed from -20˚C and spun in the micro centrifuge at 4˚C for 10 min at 14000 xg.

10 µl of each lysate was diluted 1:5 (can be lower or higher dilution depending on estimated concentration). 1:5 dilutions of lysate were prepared by addition 10 µl of lysate to 40 µl of dH_2O.

2.6.5.2 Western blots

Western Blots were prepared against β-actin, EGFR, phospho-EGFR, MAPK, phospho-MAPK

The Western blots were prepared according to the chapter 10.5.6.4 with some exceptions described below.

Samples; lysate + LB + DTT, volume of each will depend on BCA calculations.

<u>Wet Transfer</u>

Phospho-MAPK and MAPK, β-actin -40 V for 45 min followed by 60 V for 30 min.

Phospho-EGFR and EGFR- 60 V for 45 min followed by 100 V for 30 min.

Chilled transfer buffer: (200 ml of Tris - Glycine, 400 ml of Methanol and up to 2000 ml of dH_2O).

<u>Blotting membranes</u>

Phospho-MAPK, MAPK, phospho-EGFR, EGFR were blotted in 5% BSA in PBST.

β-actin was blotted in 5% Marvel milk in PBST.

2.7 Internalization studies

Cells were grown in the flasks, when semi-confluent was harvested in warmed HANKS based dissociation buffer, pelleted and re-suspended in ice-cold DMEM with 1 mg/ml BSA. The cells were counted, so that the same numbers of cells were used for each experimental point (1-5 x 10^6). The biotinylated protein (300-600 nM) was added on this stage and incubates for 1 h on ice.

The appropriate tubes were transferred to 37°C and incubated for the desired time (2-180 min), followed by rapid chilling on ice and centrifugation for 2 min at 1200 x *g*.

Cells were washed in a cold PBS containing 1 mM $MgCl_2$ and 0.1 mM $CaCl_2$ three times. Cells were left on the ice without acid stripping until step 8.

Some of cells were acid stripped by re-suspending in acid buffer (50 mM Glycine pH 5.0, 450 mM NaCl) and left on ice for 2 min. After that time, cells were spun down and washed once in PBS. Next, cell were re-suspended in Proteinase K (1 µg/ml) and left on ice for 5 min. Subsequently, they were washed twice with PBS. On this stage, the lysis was checked using a haemocytometer.

Cells were spun down and re-suspended in 250 µl of lysis buffer *(Cell signalling technology: 20 mM Tris-HCl (pH 7.5), 150 mM NaCl, 1mM Na_2EDTA, 1 mM EGTA, 1% Triton, 2.5 mM sodium pyrophosphate, 1 mM ß-glycerphosphate, 1 mM Na_3VO_4, 1 µg/ml leupeptin) with cocktail protease inhibitors (Roche*) and left for 15 min on ice.

The lysate was cleared by 15 min centrifugation at 16,000 x g. The supernatant was adsorbed with Streptavidin magnetic particles (Pierce), overnight at 4°C or 2 h at room temperature.

The particles were washed three times in lysis buffer at 4°C, followed by one wash in the water. The biotinylated protein was eluted by boiling in 100 µl of LDS-PAGE sample buffer with 100 mM DTT. The beads were separated from sample using the magnet.

The samples were run on a 4-12% SDS-PAGE gel (50 min at 200 V) and subsequently were blotted onto nitrocellulose membrane by the wet transfer method (400 mV, 100 V for 1 h at 4°C). Blots were blocked in 5% BSA in PBS-Tween for 1 h. The membranes were probed with streptavidin-HRP (Sigma) 1/2000 dilution (1 h at RT or overnight at 4°C), washed and developed using Super-signal West Dura substrate (Pierce).

Chapter 3

Design of recombinant antibody based constructs

Chapter 3. Design of recombinant antibody based constructs

3.1 Single chain variable fragments

A variety of recombinant antibody (rAb) formats has been tailored for specific applications. One of most common types of rAb are scFvs (single chain variable fragments) which are superior to their Fab and IgG counterparts due to their higher affinity to the target antigens, successful modification into a number of different antibody formats, and uncomplicated expression by several expression systems. For these reasons, the scFvs were the choice for this study.

3.1.1 Single chain variable fragment antibody

Single chain variable fragments, developed independently by Huston *et al.* (1988) and Bird *et al.* (1988) (Weisser and Hall, 2009), are one of the most widely used antibody formats engineered into larger, multivalent and conjugated forms which are used for many clinical applications.

A single chain variable fragment antibody consists of antibody variable light (V_L) and heavy (V_H) fragment linked together by peptide linker (Raag and Whitlow, 1995). Also known as single chain variable fragments (scFvs), their size range is from 27,000 to 30,000 Da. Single chain antibodies are the smallest antibody fragments encoded by a single gene to carry the entire antigen-binding region (Bird *et. al.*, 1988). They retain the binding properties of their parent antibodies (Glockshuber *et. al.,* 1990) and can be expressed in functional form in *Escherichia coli* expression systems (Hu, O'Dwyer and Wall, 2005). Single chain Fv fragments have demonstrated intracellular properties, such as binding to a specific target in the cell (Tanaka, Lobato and Rabbitts, 2003). Since they have been effective against target proteins *in vivo* (Rondon and Marasco, 1997, Biocca, Pierandrei-Amaldi and Cattaneo, 1993, Tavladoraki *et. al.,* 1993) they can be use in the treatment of human diseases and in functional genomics (Rondon and Marasco, 1997, Cattaneo and Biocca, 1999) in which the target protein interactions

can be found only inside the cell (Tanaka, Lobato and Rabbitts, 2003). ScFvs are suitable for tumour targeting (Owens and Young, 1994) due to rapid tissue penetration and fast blood clearance. In summary the single –chain variable fragment (scFv) are a well characterised format of recombinant antibody (rAb) and they have been engineered into larger, multivalent and conjugated forms for many therapeutic applications (Weisser and Hall, 2009). Single chain variable fragments were selected as potentially favourable candidates for engineering novel recombinant proteins based on the botulinum neurotoxin chimera.

3.1.2 Single chain variable fragment antibody structure

The V_H and V_L domains of the single chain antibody fragments are tethered together with a peptide linker (Weisser and Hall, 2009). The common structure of scFv monomers is designed with the C –terminal end of V_H domain joined by a peptide linker to the N-terminal residue of V_L; and in the reverse orientation, the C-terminal end of the V_L domain is linked to the N-terminal residue of V_H (Malby *et. al.*, 1993, 1998). The peptide linker is typically 10-25 amino acid residues in length and includes hydrophilic amino acids that connect the carboxyl terminus of one variable fragment with the amino group of the other variable fragment (Fig 18).

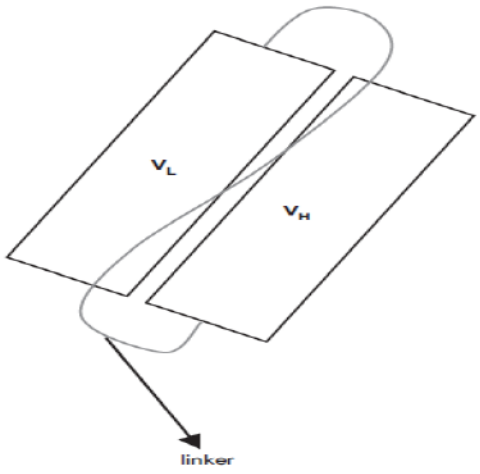


Figure 18 A single chain variable fragments (scFv) molecule.

The choice of peptide linker is based on factors, such as the suitability of the linker conformation and length (Takkinen *et. al.*, 1991) and the intended application of the scFv (Raag and Whitlow, 1995). The selection of linkers should not influence major obstructions to variable chains structure (Takkinen *et. al.*, 1991). The non-specific interactions between long linkers and variable fragments need to be avoided in order to achieve the increased antigen binding selectivity (Leong and Chen, 2007). Proteolytic stability of the linkers is another critical feature in order to minimise protein aggregation, which will result in function and yield losses (Arndt *et. al.*, 1998, Trinh *et. al.*, 2004). Two practical considerations in scFv design are (i) maintaining solubility and (ii) optimizing linker sequences to minimize the possibilities of protease attack and minimize scFv oligomerization.

3.1.3 Selection of scFv

Mast/stem cell growth factor receptor (SCFR) also known as proto-oncogene c-kit or CD117 is a protein that in humans is encoded by the Kit gene. The overexpression or mutation of this protein can lead to cancer. CD117 is of great importance for the classification of tumours, therefore the single chain antibody against CD117 was generated. Recombinant scFv CD117 sequence was obtained from the patent WO/ 2007/ 127317. The patent covered two different sequences for the heavy chain of whole immunoglobulin (SR-1 IgG) and only one sequence for the light chain of whole immunoglobulin (SR-1 IgG) against CD117. In this instance, the variable light chain and one of the variable heavy chains were used for design of scFv against CD117. The variable part of chains show great variability in amino acid sequence among the chains and has areas called hypervariable regions that form the antigen-binding sites. The unique sequence of amino acid residues for variable fragments leads to the large diversity of structure, which counts for antibody specificity. Single chain antibody (scFv CD117) targets C-kit receptors belonging to the family of tyrosine kinase receptors (RTK), which have a critical role in the development and progression of many types of cancer (Nygren *et. al.,* 1994). The tumour usually starts when a gene called *kit* develops a mutation. The abnormal Kit protein sends out a signal that causes the cell to multiply out of control, which causes the formation of a tumour. This happens in gastrointestinal stromal tumours (GIST). It is has already been established that the

humanized C-kit antibodies play an important role in the treatment of cancerous diseases. Imatinib (Gleevec) is a multi-target tyrosine kinase receptor inhibitor that targets C-kit signalling activity. Gleevec is a unique drug developed to fight a cancer by turning off the enzyme that causes the cells to become cancerous (LaVallie and McCoy, 1995). This successful product from Novartis increased the anticipation for the possibility of generating a new class of drugs based on antibody structure. This evidence of the possibility to produce novel drugs with scFvs gives a better opportunity to develop LH_N-scFv molecules. The choice of recombinant scFv enabled the targeting of EGFR (epidermal growth factor receptor) which would allow experiments to be performed on the activation and internalization of the receptor.

3.1.4 Results

Recombinant scFvCD117 was designed with the sequences provided in the patent WO/2007/127317 (taken from a PATENTSCOPE website http://www.wipo.int/patentscope/search/en/WO2007127317).

A peptide linker $(Gly_4Ser)_3$ called GS15 links the variable heavy chains to the variable light chains was introduced. The new scFv was recorded in the central stock under name scFvCD117I.

3.1.4.1 Choice of light and heavy chain variable regions

3.1.4.1.1 The kappa light chain of SR-1 IgG

The nucleic acid encoding the humanized kappa light chain SEQ ID NO: 2 from the patent, includes amino acids 20 to 248 of SEQ ID NO: 2.

The following amino acids sequence in the SEQ ID NO: 2 represents the CDR's (complementarity determining regions): CDR1 is amino acids 43 to 58, CDR2 is amino acids 74 to 80 and CDR3 is amino acids 113 to 121.

The CDR's sequences are in blue. The variable part of light chain is in black. The constant part of light chain is in grey.

DIVMTQSPDSLAVSLGERATINCRASESVDIYGNSEMHWYQQKPGQPPKLLIYIASNIESGVPDR FSGSGSGTDFTLTISSLQAEDVAVYYCQQNNEDPYTFGGGTKVEIKR

TVAAPSVFIFPPSDEQLKSGTASWCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYS
LSSTLTLSKADYEKHKVYACEVTHQGLSSPVTFNRGEC

The coding DNA sequence of the light chain is prepared by back-translation of the protein sequence. This will allow the scFv molecule to be synthesized.

3.1.4.1.2 The heavy chain of SR-1 IgG

The nucleic acid encoding the humanized heavy chain SR-1 IgG2 SEQ ID NO: 4 is from the patent. The following amino acid sequence in the SEQ ID NO: 4 represent the CDR's: CDR1 is amino acids 50 to 54, CDR2 is amino acids 69 to 85 and CDR3 is amino acids 118 to 125. The raw SEQ ID NO: 4 pg (attached Heavy chain – raw from patent)

The CDR's sequences are in blue. The variable part of heavy chain is in black. The constant part of heavy chain is in grey.

QVQLVQSGAEVKKPGASVKVSCKASGYTFTSYNMHWVRQAPGQGLEWMGVIYSGNGDTSY
NQKFKGRVTITADKSTSTAYMELSSLRSEDTAVYYCARERDTRFGNWGQGTLVTVSSAS

TKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSWT
VPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSCDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMIS
RTPEVTCVWDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYQSTYRWSVLTVLHQDWLNGKE
YKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNG
QPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK

The coding DNA sequence of the heavy chain is prepared by back-translation of the protein sequence. This will allow the scFv molecule to be synthesized.

3.1.4.2 Design of scFVCD117I

The gene organization of scFvs represents the variable regions connected in either the V_{H}- linker-V_{L} or V_{L}-linker-V_{H} orientation, where linker can be $(Gly_4Ser)_3$. However, the most common form is where the scFv gene starts with the N terminus of the variable heavy chain sequence, which run through to 13-18 amino acids past CDR-H3 (CDR3 of the heavy chain). The final typical set of VTVSA/S amino acids is maintained within variable heavy chain antibody domain. This variable heavy chain amino acid sequence is joined by peptide linker for instance $(Gly_4Ser)_3$ to the N terminus of the variable light

chain. The scFv gene ends on or one residue after TK motif 8 residues from end of CDR-L3 CDR3 of the light chain.

The choice of linker was based on a number of considerations. The length of the linker according to the information within the patent indicates that it should be around 10 residues in length, however according to the data collected in-house a linker that has 20 residues shows the best results for solubility and folding of the purified molecules. On the other hand, most of literature provides information that the scFvs peptide linker can vary between 10 to 25 residues. The proposed linker in the patent is EEGEFSEAR, but after taking to account the above suggestions it was decided to use a GS15 $(Gly_4Ser)_3$ linker, which combines the length characteristics.

3.1.4.2.1 ScFv –CD117I

QVQLVQSGAEVKKPGASVKVSCKASGYTFTSYNMHWVRQAPGQGLEWMGVIYSGNGDTSY
NQKFKGRVTITADKSTSTAYMELSSLRSEDTAVYYCARERDTRFGNWGQGTLVTVSSASGGGG
SGGGGSGGGGSDIVMTQSPDSLAVSLGERATINCRASESVDIYGNSEMHWYQQKPGQPPKLLI
YIASNIESGVPDRFSGSGSGTDFTLTISSLQAEDVAVYYCQQNNEDPYTFGGGTKVEIR

```
        10         20         30         40         50         60
QVQLVQSGAE VKKPGASVKV SCKASGYTFT SYNMHWVRQA PGQGLEWMGV IYSGNGDTSY

        70         80         90        100        110        120
NQKFKGRVTI TADKSTSTAY MELSSLRSED TAVYYCARER DTRFGNWGQG TLVTVSSASG

       130        140        150        160        170        180
GGGSGGGGSG GGGSDIVMTQ SPDSLAVSLG ERATINCRAS ESVDIYGNSE MHWYQQKPGQ

       190        200        210        220        230        240
PPKLLIYIAS NIESGVPDRF SGSGSGTDFT LTISSLQAED VAVYYCQQNN EDPYTFGGGT

KVEIKR
```

Number of amino acids: 246

Molecular weight: 26180.8

The conversions of amino acid sequence to the coding DNA sequences for CD117ScFv was prepared using Entelechon's backtranslation tool. This tool is available at the website http://www.entelechon.com/backtranslation-tool.

The sequence was entered into the Entelechon programme and the codon usage parameters were set to *Escherichia coli*. The SNAP-ON restriction sites (Table 9) belonging to the open reading frame of the LH_N backbone, therefore they needed to be identified as bad motifs to avoid their presence within the sequence in order to clone them into the correct position of molecule. The existence of these restriction sites within the sequence can be a reason of unsuccessful cloning of scFvs to recombinant botulinum chimeras. The sequence from Entelechon back-translation tool is codon-optimised for the expression in *E. coli* system and it lack insertion of restriction sites. This sequence was further optimised across the whole sequence.

Table 9 SNAP ON Restriction enzymes in LH_N open reading frame

Restriction enzyme	Cleavage site
Ava I	C↓(C/T)CG(G/A)G
Avr II	C↓CTAGG
BamH I	G↓GATCC
Bgl II	A↓GATCT
EcoR I	G↓AATTC
Hind III	A↓AGCTT
Nde I	C↓ATATG
Nhe I	G↓CTAGC
Pst I	CTGCA↓G
Sal I	G↓TCGAC
Spe I	A↓CTAGT
Xba I	T↓CTAGA
Xma I	C↓CCGGG

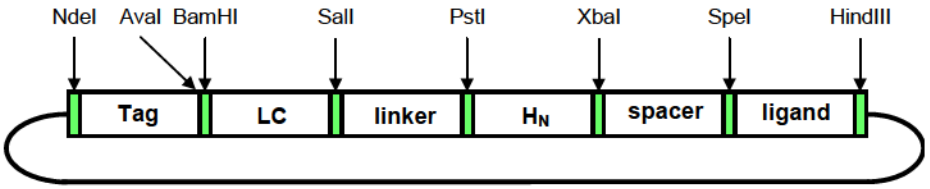


Figure 19 The most common SNAP-ON restriction enzymes within LHN backbone.

3.1.4.3 Optimisation of DNA

The optimised sequence was then copied from the results of the back-translation tool and pasted into the word document, so the changes in the DNA sequence could be identified.

The optimisation of DNA was checked by a second software package called the graphical codon usage analyser. The programme is available at http://www.gcua.schoedl.de/ website. This tool analyses two parameters and the results are displayed in appendix each triplet position vs. usage table and each codon vs. usage table.

3.1.4.3.1 Each triplet position vs. usage table

The each triplet position vs. usage table shows each amino acid in the sequence and gives a visual representation of the frequency of the codon usage of each codon. This was utilized to reduce the usage of rare codons or runs of less frequently used codons to preventing pausing or stalling of translation.

3.1.4.3.2 Each codon vs. usage table

The each codon vs. usage table shows the codons that can be used for each amino acid and compares how often it should be used in a sequence (in %) against how often it is used in designed sequences. This was utilized to choose which codons to use as replacements in the designed sequences.

The back-translated sequences were optimized according to *Escherichia coli* codon usage table. The iterations were introduced avoiding very rare codons and overloading very common codons.

3.1.4.4 Back-translated and optimised DNA

3.1.4.4.1 Back-translated

The substitutions made within the sequence are highlighted in blue.

CD117IScFv

CAAGTCCAGCTGGTACAGAGTGGCGCTGAGGTCAAAAAGCCAGGAGCAAGCGTTAAAGT
TTCCTGTAAGGCGTCGGGTTACACTTTCACAAGTTATAATATGCATTGGGTACGCCAGG
CGCCTGGACAAGGGCTTGAATGGATGGGTGTCATTTATTCGGGTAACGGTGATACCAGC
TATAACCAGAAATTCAAAGGGCGTGTGACGATAACTGCTGATAAAAGTACCTCTACAGC
TTACATGGAGCTTTCTTCGTTGCGTTCAGAAGACACGGCAGTGTACTATTGTGCACGGG
AGCGCGATACGCGTTTTGGGAATTGGGGACAAGGCACTCTGGTGACAGTATCCAGCGCG
TCAGGGGGCGGAGGTTCTGGGGGAGGCGGATCAGGGGGAGGTGGTTCCGATATTGTCAT
GACTCAGTCTCCGGACTCCCTGGCCGTTTCGCTGGGTGAAAGAGCTACGATAAACTGCC
GAGCCAGCGAGAGCGTAGACATCTATGGCAACAGTGAAATGCACTGGTATCAACAGAAA
CCGGGTCAACCCCCTAAGCTCTTAATCTATATAGCATCGAATATCGAGTCTGGAGTTCC
AGATCGCTTTTCCGGCAGTGGGTCAGGCACAGACTTTACCTTGACAATAAGTTCATTAC
AGGCGGAGGACGTGGCAGTTTACTACTGCCAACAAAATAATGAAGATCCCTACACTTTC
GGAGGCGGGACCAAGGTAGAAATTAAGAGA

The prepared sequence of scFv was checked using SeqBuilder programme of the DNASTAR Lasergene package to confirm that no unwanted changes had been introduced during the redesign, such the SNAP-ON restriction sites.

The final predicted sequence was inserted into a new DNA folder in the SeqBuilder programme. Then the restriction site *XbaI* was added to the N-terminal of sequence with addition of "A" base after TCTAGA motif to avoid a frame shift. Therefore, after digestion it will leave CTA GAA (LE) at the start of the sequence.

The restriction site *HindIII* (AAGCTT) was added on the C-terminus of the sequence. *HindIII* will leave one bp "A" at the end of the sequence. The sequence of ScFv has two stop codons: TAA and TAG from the C-terminus side of the molecule.

The predicted sequence of CD117IScFvs was prepared. The drop out fragment between the restriction sites: *XbaI* and *HindIII* from ScFv, separately was inserted into the backbone construct SXN101327 (pK7-10HT-LB-EK-HB (K191A)-GS20-EGF). The insertion of CD117IScFv in place of the EGF ligand was performed using the SeqBuilder program to ensure that there were no unintentional changes in the sequence.

3.1.4.4.2 Optimized DNA

CD117IScFv

tctagaaCAAGTCCAGCTGGTACAGAGTGGCGCTGAGGTCAAAAAGCCAGGAGCAAGCGTTAA
AGTTTCCTGTAAGGCGTCGGGTTACACTTTCACAAGTTATAATATGCATTGGGTACGCCAGGC
GCCTGGACAAGGGCTTGAATGGATGGGTGTCATTTATTCGGGTAACGGTGATACCAGCTATA
ACCAGAAATTCAAAGGGCGTGTGACGATAACTGCTGATAAAAGTACCTCTACAGCTTACATGG
AGCTTTCTTCGTTGCGTTCAGAAGACACGGCAGTGTACTATTGTGCACGGGAGCGCGATACGC
GTTTTGGGAATTGGGGACAAGGCACTCTGGTGACAGTATCCAGCGCGTCAGGGGGCGGAGG
TTCTGGGGGAGGCGGATCAGGGGGAGGTGGTTCCGATATTGTCATGACTCAGTCTCCGGACT
CCCTGGCCGTTTCGCTGGGTGAAAGAGCTACGATAAACTGCCGAGCCAGCGAGAGCGTAGAC
ATCTATGGCAACAGTGAAATGCACTGGTATCAACAGAAACCGGGTCAACCCCCTAAGCTCTTA
ATCTATATAGCATCGAATATCGAGTCTGGAGTTCCAGATCGCTTTTCCGGCAGTGGGTCAGGC
ACAGACTTTACCTTGACAATAAGTTCATTACAGGCGGAGGACGTGGCAGTTTACTACTGCCAA
CAAAATAATGAAGATCCCTACACTTTCGGAGGCGGGACCAAGGTAGAAATTAAGAGAtaataga
agctt

The synthesized scFv gene was subsequently cloned into LH_N backbones derived from different serotypes of botulinum neurotoxin.

Entelechon prepared the synthesis of 757bp scFv sequence. The Entelechon reports are attached in the appendix.

3.2 Single domain antibodies

The naturally occurring single domain antibodies that lack L-chain were discovered in the serum of camels. This, the smallest known antigen binding fragments, attracted a lot of attention from many research groups all over the world and become a one of the most important molecule in the creation of novel drugs. It was advantageous to

introduce the repertoire of these antibodies as a part of LH_N molecules because of their small size, which may increase expression levels of the full-length protein.

3.2.1 Single domain antibodies

Single domain antibodies (sdAbs), also known as domain antibodies (dAbs) or nanobodies are the smallest antibody fragments with a size of 11-15 kDa (Saerens, Ghassabeh and Muyldermans, 2008) and llamas, other camelids (Genst *et. al.*, 2006) and sharks, naturally produce these heavy chain antibodies (hcAbs) that lack light chain altogether. Single domain antibodies have been also isolated from either heavy chain or light chain variable regions of human antibodies (Wesolowski *et. al.,* 2009).

The single domains of camelids are more often used for biotechnological applications than those from sharks because the immunization process is more simple (Harmsen and Haard, 2007). Some of the properties of single domain antibodies are noteworthy. They remain functional with an active antigen-binding specificity after prolonged incubation at temperatures from 80 to 92˚C (Vranken *et. al.*, 2002, Tanha *et. al.*, 2002, Ewert *et. al.*, 2003, van der Linden *et. al.*, 1999, Dumoulin *et. al.*, 2002), they show an increased resistance against denaturation (Perez *et. al.,* 2001, Ewert *et. al.*, 2003), they have rapid blood clearance and fast tissue penetration. This makes them exceptional useful for targeting VHHs coupled to toxic substances to tumours (Cortez-Retamozo *et. al.,* 2004). The short serum half- live of dAbs has been extended by their targeting to long-lived proteins, such as albumin (Coppieters *et. al.*, 2006, Roovers *et. al.*, 2007) and immunoglobulin (Harmsen *et. al.*, 2005a). They could be used for many therapies including the treatment of cancer. It has been recorded that sAbs binds to epidermal growth factor receptor (EGFR), so they block EGF binds to this receptor. This approach is currently being investigated in treatment of solid cancers (Roovers *et. al.*, 2007). Further experiments showed that they could be used in targeting drugs across the blood-brain barrier (BBB) into the brain (Muruganandam *et. al.*, 2002) what subsequently could be used for treatment of neurological disorders (Gueorguieva *et. al.*, 2006). Overall, the biophysical properties of VHH antibodies seem to be good candidates for the construction of recombinant antibody – botulinum toxin hybrid molecules.

3.2.2 Single domain antibody structure

Camelids and sharks produce functional antibodies composed only of heavy chains (Brekke and Sandlie, 2003, Hamers-Casterman *et. al.*, 1993). They are devoid of the light chains (Hamers-Casterman *et. al.*, 1993) and in the case of camelids antibodies; the CH1 domain is absent (Holliger and Hudson, 2005). The CH1 domain in the conventional antibody associates with the light chain and to some degree interacts with V_H domain (Harmsen and Haard, 2007). Their single domain of hcAbs binds antigen without requiring domain pairing and is linked directly via a hinge region to the Fc domain (Wesolowski *et. al.*, 2009). The antigen-binding domain is designated VHH for camelids hcAbs and VNAR for shark hcAbs (Holliger and Hudson, 2005). Each single domain antibody therefore contains three of the six naturally occurring complementarily determining regions (CDRs) from an antibody (Holt *et. al.,* 2003). A comparison of the structure of CDR1 and CDR2 with the set of loops found in human antibodies exposed significant differences in their main chain conformation (Decanniere *et. al.*, 2000). The CDR3 regions of these antibodies have an extraordinary capacity to form long extensions that can extend into cavities present on antigens (Wesolowski et. al., 2009). For this reason, they are able to recognize enzyme active sites and conserved cryptic epitopes (Lauwereys *et. al.,* 1998). In addition, the CDR3 regions of VHHs and VNARs are much longer in comparison to the conventional V_H domains (Wu, Johnson and Kabat, 1993, Harmsen and Haard, 2007) being up to 17 residues in length compared the 12 residues in human or 9 residues in mouse (Holt *et. al.*, 2003). An additional disulphide bond connecting the CDR3 to CDR1 loop or to CDR2 loop occasionally stabilizes the extended CDR3 (Wesolowski *et. al.,* 2009) (Fig 20). The CDR2 region in the shark single domain antibodies does not contribute to antigen binding (Streltsov *et. al.,* 2004).

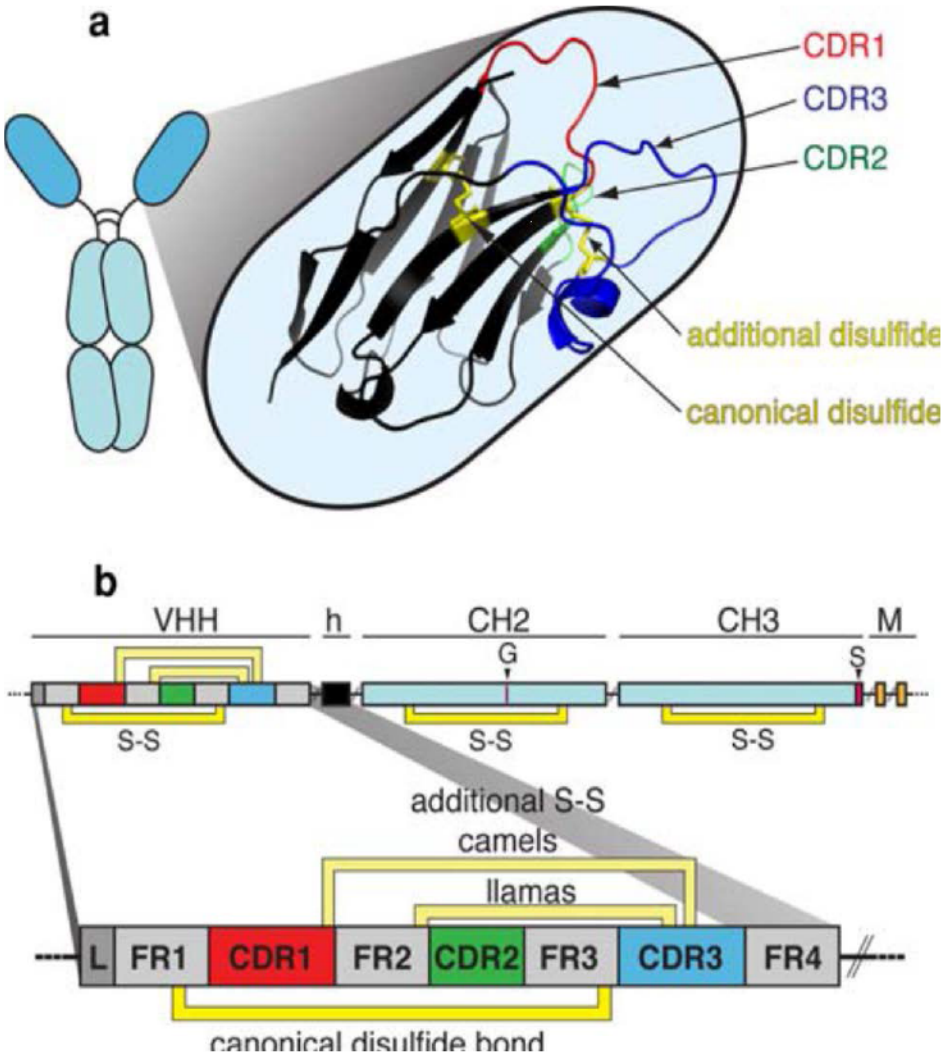


Figure 20 Schematic diagram of the VHH domain of a camelid heavy chain antibody and disulphide bond framework regions (adapted from Wesolowski *et al.*, 2009)

a The three complementarity determining regions (CDRs) of the antigen-binding paratope are depicted as coloured loops: CDR1 *red*, CDR2 *green*, and CDR3 *blue*.

b The canonical disulfide bond connecting framework regions 1 and 3 (FR1 and FR3) in the two *β*-sheets of the immunoglobulin domain is indicated in *yellow*. Many camelid antibodies contain an additional disulfide bond (S–S) connecting the CDR3 with the end of the CDR1 (camels) or the beginning of the CDR2 (llamas). *h* Hinge, *M* transmembrane domain of membrane isoform, *G* glycosylation site, *S* stop codon of secretory isoform.

3.2.3 Selection of sdAbs

For the creation of single domain antibodies (sdAbs), where the antibody domain is engineered to comprise only the variable heavy or variable light chain (and therefore consist of three CDR's), the amino acid sequences of 11 anti-EGFR sdAbs have been obtained from patent WO 2008/141449 A1 (taken from a PATENTSCOPE website http://www.wipo.int/patentscope/search/en/WO2008141449).

They are named EGX, where X symbolizes a number of each single domain antibody. Six of the anti-EGFR sdAbs have been chosen from the patent for purpose of this study. The selection of sdAbs was based on the experimental description in the patent WO 2008/141449 A1. The scope of the patent relates to the field of antibodies directed towards epidermal growth factor receptor (EGFR). The anti-EGFR proteins in this case single domain antibodies (sdAb) represented in this patent can be used in the targeting of tumours presenting EGFR on their surface and for the diagnosis and treatment of certain types of cancer associated with cells over-expressing EGFR on their surface.

Based on the sequence identity of their CDR's, the sdAbs has been divided into four groups:

Group 1: EG2; EG5; EG28

Group 2: EG6; EG10

Group 3: EG7; EG16; EG29; EG30

Group 4: EG31

The amino acid sequences of 11 sdAbs specific for EGFR-extra cellular domain (ECD) with complementarily determining regions CDR1, CDR2, CDR3 have been illustrated (Fig 12 in chapter 2).

The single domain antibodies were selected according to the description of experiments in the patent WO 2008/141449 A1, a brief explanation of which is given below.

The data recorded on the sensorgrams from the reading by the surface plasmon resonance show the binding of 0.5 µM EG2, EG10, EG31; EG43 to EGFR-extracellular domain (ECD). The EG2 shows the strongest binding in the longest period. The EG10 and the EG43 have the same ability of binding; however, the EG10 lasts for the longer phase than the EG43. The EG31 has the lowest binding capability in the shortest period.

One of the sdAb gene from each of four groups was chosen and subcloned into an *E. coli* periplasmic expression vector (pSJF2) to generate four clones: pEG2, pEG10, pEG31, pEG43. The four sdAb, each tagged with a 6 Histidine (His) at their C-termini were produced in *E. coli* and purified using IMAC (Immobilized Metal Affinity Chromatography). The yields of EG2, EG10, EG31, and EG43 were 11, 19.4, 7.8, 43 mg per litre of TG1 culture, respectively.

The same four anti-EGFR sdAbs were analyzed for their binding to EGFR-ECD by surface plasmon resonance. The dissociation constants (K_{DS}) of the sdAbs ranged from 55 nM to 440 nM and are shown below

EGF2: 55 nM

EG10: 126 nM

EG31: 440 nM

EG43: 316 nM

A high value of K_{DS} indicates that the anti-EGFR sdAbs have a low binding to EGFR-ECD. Therefore, the best binding abilities of the sdAbs are as follows in the following order: EG2, EG10, EG43, and EG31.

The data obtained from the size exclusion chromatography of EG2, V2C-EG2, and EG2-hFc following EG2 and V2C-EG2 expression in *E. coli* and EG2hFc expression in HEK293 cells gave another characteristic to consider.

The elution position of V2C-EG2 and EG2-hFc are almost identical, whereas that of EG2 is different. This suggests that all sdAbs will behave similar to EG2 during purification, as the expression system was the same for EG2 and V2C-EG, but they varied during purification.

According to the given results in the patent the following sdAbs were chosen for the cloning of sdAbs-LH_N

Group 1: EG28

Group 2: EG10, EG6

Group 3: EG7, EG43

Group 4: EG31 (It should represents a weak example)

3.2.4. Results

3.2.4.1 Design of single domain antibodies

The sequences of six sdAbs were obtained from the table (Fig 12 in the chapter 2) provided in the patent WO 2008/141449 A1. The CDR's: CDR1, CDR2, and CDR3 are highlighted in blue. Introduced changes to the DNA sequence after optimisation by the graphic codon usage analyser are in grey (Attached documents for graphic codon usage analyser).

As previously described in the section ‚Design of CD117l' it is important to check for the presence of SNAP-ON restriction enzyme recognition sites within the sequence. The back-translated and optimised sdAbs were therefore checked using the SeqBuilder programme of the DNASTAR Lasergene package.

The final predicted sequences of six sdAbs were inserted into a new DNA file in the SeqBuilder programme and the restriction enzyme recognition site for *XbaI* was added to the 5'-terminal of the sequences with the addition of an "A" base after TCTAGA motif to avoid a frame shift. Therefore, after digestion it will leave the sequence CTA GAA (coding for amino acid residues LE) at the start of the coding sequence.

The restriction enzyme recognition site for *HindIII* with the sequence AAGCTT was added onto the C-terminus of the sequences. The inserted *HindIII* recognition site will leave one bp "A" at end of the sequence. Both sequences of ScFvs have two stop codons: TAA and TAG from the C-terminus side of the molecule. The *XbaI* and *HindIII* restriction enzymes are coloured in green, where additional 'A' base residue and STOP codons are coloured in red. Entelechon synthesized the DNA.

Group 1:

EG28 with identified CDRs

QVQLVESGGGLVQAGDSLRLSCVDSGRDFSDYVMGWFRQAPGKEREFVAAISRNGITTRYADS
VKGRFTISRDNDKNTVYLQMNSLKPEDTAVYYCATNSAGTYVSPRSRDYDGWGQGTQVTVSS

Back-translated and optimised EG28

CAAGTTCAGCTGGTCGAATCGGGAGGTGGGTTAGTTCAAGCCGGAGATTCTCTTCGCCTGAG
CTGTGTAGACTCCGGAAGAGACTTTTCAGACTACGTAATGGGGTGGTTTCGCCAAGCGCCTGG
CAAGGAGCGTGAGTTCGTTGCTGCCATCTCACGCAACGGTATTACCACGCGGTATGCGGACA
GTGTCAAAGGCCGTTTCACAATAAGTAGAGATAATGACAAGAACACAGTATACTTGCAGATAA
TAGCCTGAAACCGGAAGATACGGCTGTCTATTACTGCGCAACCAATAGTGCAGGTACTTATGT
GTCGCCACGGAGCCGTGATTATGATGGCTGGGGACAGGGGACACAAGTGACTGTGTCCTCT

Final EG28 synthesized by Entelechon

tctagaaCAAGTTCAGCTGGTCGAATCGGGAGGTGGGTTAGTTCAAGCCGGAGATTCTCTTCGC
CTGAGCTGTGTAGACTCCGGAAGAGACTTTTCAGACTACGTAATGGGGTGGTTTCGCCAAGC
GCCTGGCAAGGAGCGTGAGTTCGTTGCTGCCATCTCACGCAACGGTATTACCACGCGGTATGC
GGACAGTGTCAAAGGCCGTTTCACAATAAGTAGAGATAATGACAAGAACACAGTATACTTGC
AGATGAATAGCCTGAAACCGGAAGATACGGCTGTCTATTACTGCGCAACCAATAGTGCAGGT
ACTTATGTGTCGCCACGGAGCCGTGATTATGATGGCTGGGGACAGGGGACACAAGTGACTGT
GTCCTCTtaatagaagctt

Group 2:

EG10 with identified CDRs

QVQLVESGGGLVQAGGSLTLSCAASGGTFSSYAMGWFRQAPGKEREFVAAISGRSSIRNYDDS
VKGRFAISRDSAKNTVYLQMNSLKPEDTAVYYCAADTVFRSFVVGNVKEWGQGTQVTVSS

Back-translated and optimised EG10

CAGGTGCAATTAGTTGAGTCGGGTGGGGGACTGGTGCAAGCTGGTGGATCACTGACCTTGTC
CTGTGCGGCCAGCGGTGGCACATTCAGTAGCTATGCCATGGGCTGGTTTAGACAGGCTCCAG
GGAAAGAACGCGAATTTGTTGCTGCCATAAGTGGGCGCTCTTCCATTCGGAATTACGATGACT
CTGTCAAGGGACGTTTTGCAATCAGCAGAGACTCGGCGAAGAATACGGTCTATCTTCAGATGA
ACTCACTGAAACCTGAGGATACTGCGGTCTACTATTGCGCAGCAGATACTGTATTCCGTTCTTT
CGTAGTGGGTAACGTTAAAGAATGGGGCCAAGGAACACAAGTAACGGTATCAAGT

Final EG10 synthesized by Entelechon

tctagaaCAGGTGCAATTAGTTGAGTCGGGTGGGGGACTGGTGCAAGCTGGTGGATCACTGAC
CTTGTCCTGTGCGGCCAGCGGTGGCACATTCAGTAGCTATGCCATGGGCTGGTTTAGACAGGC
TCCAGGGAAAGAACGCGAATTTGTTGCTGCCATAAGTGGGCGCTCTTCCATTCGGAATTACGA
TGACTCTGTCAAGGGACGTTTTGCAATCAGCAGAGACTCGGCGAAGAATACGGTCTATCTTCA
GATGAACTCACTGAAACCTGAGGATACTGCGGTCTACTATTGCGCAGCAGATACTGTATTCCG
TTCTTTCGTAGTGGGTAACGTTAAAGAATGGGGCCAAGGAACACAAGTAACGGTATCAAGTta
atagaagctt

Group 2:

EG6 with identified CDRs

QVKLEESGGGLVQAGGSLTLSCAASGGTFSSYAMGWFRQAPGKEREFVAAISGRSSIRNYDDS
VKGRFAISRDNAKNTVYLQMNSLKPEDTAVYYCAADTVFRSFVVGNVKEWGQGTQVTVSS

Back-translated and optimised EG6

CAAGTTAAACTGGAAGAATCAGGAGGAGGCCTGGTACAGGCCGGTGGTTCCCTTACTTTAAG
TTGTGCTGCCAGTGGTGGAACCTTTAGCTCGTATGCTATGGGTTGGTTCCGGCAGGCCCCAGG
GAAAGAGAGAGAGTTCGTGGCGGCAATAAGCGGACGTTCGTCTATCCGTAATTATGATGACA
GTGTAAAGGGGCGTTTTGCAATTTCACGCGATAATGCGAAAAACACGGTATATCTGCAAATGA
ACTCATTGAAACCTGAAGACACTGCTGTGTACTACTGCGCGGCAGATACAGTTTTTAGATCCTT
CGTCGTTGGCAATGTCAAGGAATGGGGCCAAGGGACACAGGTGACGGTCAGC TCT

Final EG6 synthesized by Entelechon

tctagaaCAAGTTAAACTGGAAGAATCAGGAGGAGGCCTGGTACAGGCCGGTGGTTCCCTTACT
TTAAGTTGTGCTGCCAGTGGTGGAACCTTTAGCTCGTATGCTATGGGTTGGTTCCGGCAGGCC
CCAGGGAAAGAGAGAGAGTTCGTGGCGGCAATAAGCGGACGTTCGTCTATCCGTAATTATGA
TGACAGTGTAAAGGGGCGTTTTGCAATTTCACGCGATAATGCGAAAAACACGGTATATCTGCA
AATGAACTCATTGAAACCTGAAGACACTGCTGTGTACTACTGCGCGGCAGATACAGTTTTTAG
ATCCTTCGTCGTTGGCAATGTCAAGGAATGGGGCCAAGGGACACAGGTGACGGTCAGCTCTta
atagaagctt

Group 3:

EG7 with identified CDRs

QVQLVESGGGLVQPGGSLRLSCAASESFFNFDAWGWYRQAPGKQREMVAVVGSTGSTSYAD
FVKGRFTISRDNANNTVYLQMNTLRPEDTAVYYCYARFQSLYNSWGQGTQVTVSS

Back-translated and optimised EG7

CAAGTGCAGCTGGTTGAGTCCGGTGGGGGCCTTGTGCAGCCAGGTGGGTCGTTGCGGCTG
TCTTGCGCAGCCAGCGAATCTTTTTTCAATTTTGATGCGTGGGGCTGGTATCGTCAGGCGCCT
GGAAAACAACGTGAAATGGTTGCAGTCGTGGGAAGTACGGGATCGACGAGTTACGCTGACTT
TGTAAAGGGGAGATTCACAATAAGTAGAGACAACGCAAATAACACCGTTTATCTGCAAATGA
ACACTTTACGCCCAGAGGATACAGCTGTATACTATTGTTATGCCCGCTTCCAGAGCTTATACAA
TTCATGGGGCCAAGGTACACAAGTAACTGTCTCCTCA

Final EG7 synthesized by Entelechon

tctagaaCAAGTGCAGCTGGTTGAGTCCGGTGGGGGCCTTGTGCAGCCAGGTGGGTCGTTGCG
GCTGTCTTGCGCAGCCAGCGAATCTTTTTTCAATTTTGATGCGTGGGGCTGGTATCGTCAGGC
GCCTGGAAAACAACGTGAAATGGTTGCAGTCGTGGGAAGTACGGGATCGACGAGTTACGCTG
ACTTTGTAAAGGGGAGATTCACAATAAGTAGAGACAACGCAAATAACACCGTTTATCTGCAAA
TGAACACTTTACGCCCAGAGGATACAGCTGTATACTATTGTTATGCCCGCTTCCAGAGCTTATA
CAATTCATGGGGCCAAGGTACACAAGTAACTGTCTCCTCAtaatagaagctt

Group 3:

EG43 with identified CDRs

QVQLVESGGGLVQPGGSLRLPCAASGSIFSLDAWGWYRQAPGKQREMVALVGSDGSTSYADS
VKGRFTISRDNANNTFYLQMNSLKPEDTAVYYCYARFQSLYNSWGQGTQVTVSS

Back-translated and optimised EG43

CAGGTACAATTGGTGGAATCTGGAGGGGGCTTAGTTCAACCCGGTGGATCTCTCCGTCTGCCA
TGTGCCGCAAGTGGCTCGATTTTCAGCCTGGACGCTTGGGGATGGTATCGCCAGGCGCCTGG
CAAACAACGGGAGATGGTCGCACTTGTAGGTTCCGATGGGTCCACAAGTTATGCTGACAGTG
TTAAGGGGAGATTCACCATAAGCCGCGATAATGCGAATAACACTTTTTACCTGCAAATGAACT
CGTTGAAACCGGAAGATACGGCAGTGTATTACTGCTATGCCCGTTTTCAGTCATTATACAATA
GCTGGGGACAGGGTACACAAGTAACTGTCTCATCA

Final EG43 synthesized by Entelechon

tctagaaCAGGTACAATTGGTGGAATCTGGAGGGGGCTTAGTTCAACCCGGTGGATCTCTCCGT
CTGCCATGTGCCGCAAGTGGCTCGATTTTCAGCCTGGACGCTTGGGGATGGTATCGCCAGGC
GCCTGGCAAACAACGGGAGATGGTCGCACTTGTAGGTTCCGATGGGTCCACAAGTTATGCTG
ACAGTGTTAAGGGGAGATTCACCATAAGCCGCGATAATGCGAATAACACTTTTTACCTGCAAA
TGAACTCGTTGAAACCGGAAGATACGGCAGTGTATTACTGCTATGCCCGTTTTCAGTCATTATA
CAATAGCTGGGGACAGGGTACACAAGTAACTGTCTCATCAtaatagaagctt

Group 4:

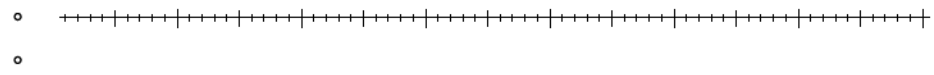
EG31 with identified CDRs

QVKLEESGGGLVQVGGSLRLSCAHSGLPFGINAIGWYRQGPGNQRDLVARITSDGRTILEDSVK
GRFTISRDNAKKTVYVQMNNLKPEDTAVYYCAAEKGGSPLYWGQGTQVTVSS

Back-translated and optimised EG31

CAGGTGAAGTTAGAGGAAAGTGGTGGGGGACTTGTTCAAGTGGGGGGAAGTCTCCGTTTGT
CTTGCGCTCATAGCGGCTTGCCCTTTGGCATTAACGCAATAGGATGGTACAGACAGGGTCCGG
GTAACCAGCGTGACTTAGTAGCGAGAATTACATCGGATGGACGCACTATCCTGGAAGATTCGT
AAAAGGGCGGTTCACTATATCTCGCGATAATGCGAAGAAAACCGTTTATGTCCAAATGAATAA
TCTGAAACCTGAGGACACGGCCGTCTATTACTGTGCTGCAGAAAAGGGCGGTTCACCACTGTA
TTGGGGCCAAGGGACACAAGTTACGGTAAGCTCC

Pst I site identified in DNA sequence (see below)



and therefore changed bp291 from T to G (lower case below) to give the sequence below.

CAGGTGAAGTTAGAGGAAAGTGGTGGGGGACTTGTTCAAGTGGGGGGAAGTCTCCGTTTGT
CTTGCGCTCATAGCGGCTTGCCCTTTGGCATTAACGCAATAGGATGGTACAGACAGGGTCCGG
GTAACCAGCGTGACTTAGTAGCGAGAATTACATCGGATGGACGCACTATCCTGGAAGATTCA
GTAAAAGGGCGGTTCACTATATCTCGCGATAATGCGAAGAAAACCGTTTATGTCCAAATGAAT
AATCTGAAACCTGAGGACACGGCCGTCTATTACTGTGCTGCgGAAAAGGGCGGTTCACCACTG
TATTGGGGCCAAGGGACACAAGTTACGGTAAGCTCC

Final EG31 synthesized by Entelechon

tctagaaCAGGTGAAGTTAGAGGAAAGTGGTGGGGGACTTGTTCAAGTGGGGGGAAGTCTCCG
TTTGTCTTGCGCTCATAGCGGCTTGCCCTTTGGCATTAACGCAATAGGATGGTACAGACAGGG
TCCGGGTAACCAGCGTGACTTAGTAGCGAGAATTACATCGGATGGACGCACTATCCTGGAAG
ATTCAGTAAAAGGGCGGTTCACTATATCTCGCGATAATGCGAAGAAAACCGTTTATGTCCAAA
TGAATAATCTGAAACCTGAGGACACGGCCGTCTATTACTGTGCTGCgGAAAAGGGCGGTTCAC
CACTGTATTGGGGCCAAGGGACACAAGTTACGGTAAGCTCCtaatagaagctt

The six synthesized sdAbs genes were subsequently cloned into plasmids encoding LH_N backbones derived from different serotypes of botulinum neurotoxin. Entelechon reports are submitted in the appendix.

Chapter 4

Cloning of recombinant antibody based constructs

Chapter 4. Cloning of recombinant antibody based constructs

The synthesized two single chain variable antibody fragments and six single domain antibodies were cloned into the LH_N backbone. As mentioned in Chapter 1 there are seven (named A through G) distinct botulinum neurotoxins produced by *Clostridium botulinum* (BoNT) of which four (A, B, C and D) were used with this study.

4.1 Re-engineering botulinum neurotoxin

The ability of botulinum neurotoxin to disrupt transmitter release from the target cells, often for prolonged periods, has become beneficial for the therapeutic use to treat a variety of diseases (Turton, Chaddock and Acharya, 2002) for example acromegaly, pain, endocrine. In order to develop novel cell secretion inhibitor drugs it was important to re-engineer botulinum neurotoxin to become a non-toxic molecule, whilst at the same time maintaining the ability to block secretion.

In order for this to be achieved, a single chain polypeptide of ˜150kDa, which is cleaved to form a di-chain molecule linked by a disulphide bond, has to be produced (Turton, Chaddock and Acharya, 2002). The toxic BoNT structure consists of heavy chain (HC, ˜100kDa) covalently joint to light chain (LC, ˜50kDa) (Simpson, 1986) (Fig 21). The HC contains two domains each of ˜50kDa. C-terminal domain (H_C) responsible for binding to target cells (Shone *et. al*., 1985, Halpern and Loftus, 1993) and an N-terminal domain (H_N) involved in intracellular membrane translocation (Shone *et. al.,* 1987, Koriazova and Montal, 2003). The LC domain is a zinc-depended endopeptidase that cleaves SNARE proteins that are essential for neurotransmitter release (Humeau *et. al*., 2000).

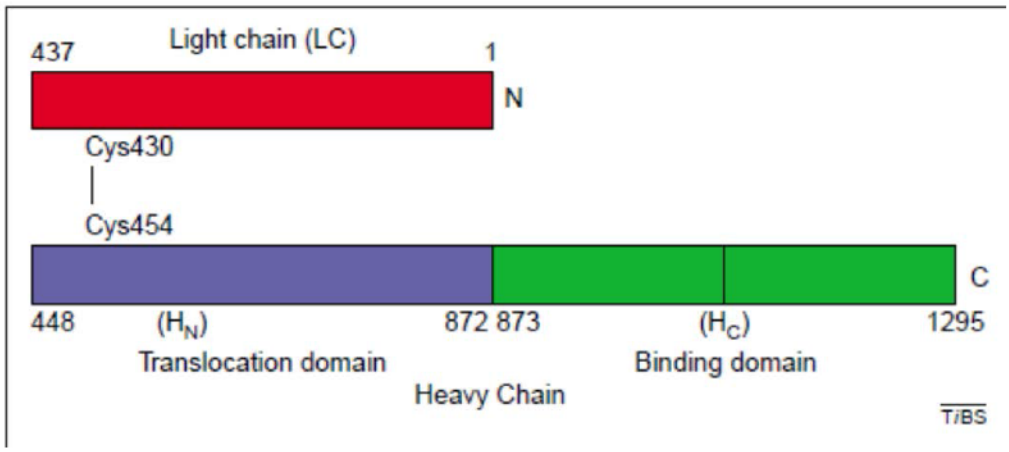


Figure 21 The di-chain structure of a Clostridial neurotoxin –botulinum neurotoxin A (BoNT/A). (Adapted from Turton *et. al.*, 2002)

Clostridial neurotoxins are ~150-kDa proteins, synthesized as single-chain polypeptides and post-translationally nicked to form di-chain molecules. They share the same domain architecture and overall structure. The light and heavy chains of BoNT/A are linked by a single disulfide bond, Cys430–Cys454. The light chain (LC), shown in red, functions as zinc-dependent endopeptidase and contains the catalytic zinc atom and HExxH motif associated with zinc-dependent proteases. The heavy chain comprises two functional domains of roughly equal size. The N-terminal section (H_N), shown in blue, is the translocation domain, which forms ion channels spanning endosomal membranes and is thought to be involved in translocation and activation of the LC. The C-terminal section (H_C), shown in green, is the binding domain.

Experiments based on BoNT/A proteolytic cleavage by trypsin gives an opportunity to produce active, non-toxic and non-cell binding molecules. These molecules are derived from BoNT/A and are ˜100kDa polypeptides called LH_N/A, which represent the LC and the H_N domains of type A neurotoxin coupled together by a disulphide bond (Shone, Hambleton and Melling, 1987). The same LH_N/A domains have been expressed and purified from a heterologous expression host (Chaddock *et. al.*, 2002).

Consequently, re-engineering clostridium neurotoxin molecules by the replacement of their native binding domain (H_C) with an alternative ligand has created functional LH_N/A fragments to target cells of interest (Chaddock *et. al.*, 2000, Chaddock *et. al.,* 2000). The first experiments were performed using chemical conjugates, which have been prepared by coupling LH_N/A fragments to a lectin obtained from *Erythrina cristagalli (*ECL) (Duggan *et. al.,* 2002). Lectins are non-immunoglobulin proteins that recognize and bind selectively to specific carbohydrates and can therefore be used to differentiate between cell types (Lis and Sharon, 1986, Sharon 1993, Streit *et. al.,* 1985, Streit *et. al.,* 1986). Therefore, the lectins are appropriate ligands for selectively targeting LH_N/A endopeptidase to nociceptive afferents. Additional experiments with

LH_N/A-ECL conjugates showed that they are able to complete all stages of the intoxication process and are able to cause the inhibition of neurotransmitter release from cultured eDRG neuronal cell types in preference to spinal cord neurons (Duggan *et. al.*, 2002).

Nevertheless, although these molecules were able to inhibit acetylcholine exocytosis their construction by the chemical coupling of the LH_N fragment and the targeting ligand was not a suitable method for producing pharmaceutical agents. Therefore, a new approach has been developed to avoid the inherent problem of the production of heterogeneous products observed with the chemical coupling approach. The production of a fully recombinant fusion protein from the recombinant gene encoding both the LH_N domain of a clostridial neurotoxin and a specific targeting domain has been achieved to solve this problem. One of the first constructed molecules was composed of the LH_N domain of botulinum neurotoxin type C and epidermal growth factor (EGF), which was designed to inhibit secretion of mucus from epithelial cells (Foster *et. al.*, 2005) (Fig 22).

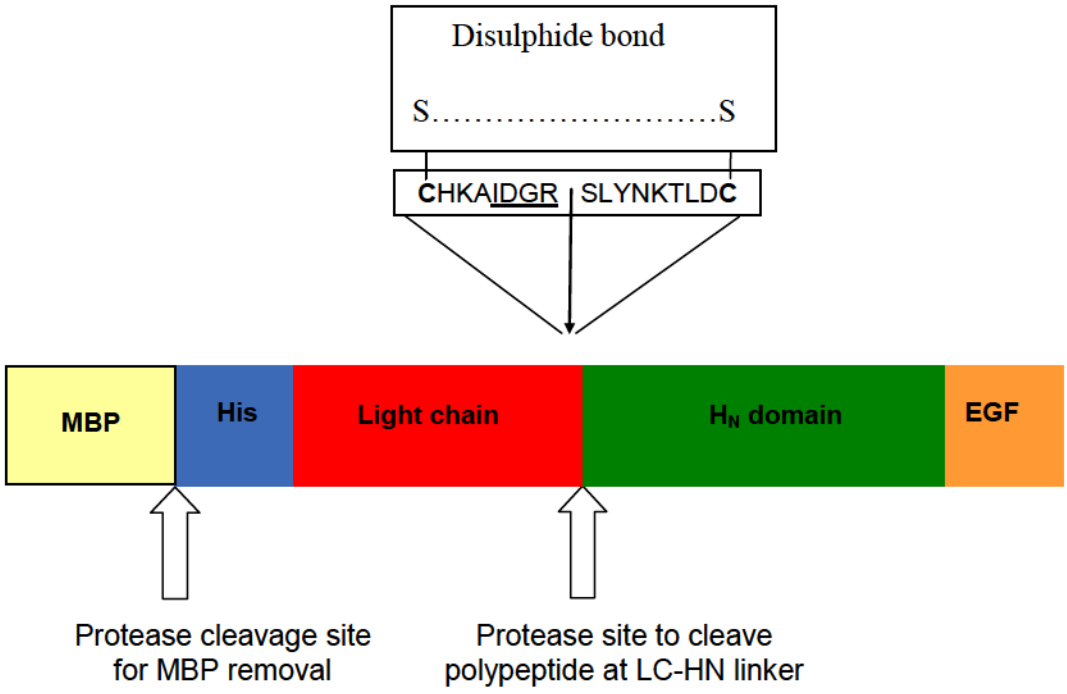


Figure 22 LH_N domain of botulinum neurotoxin type C and epidermal growth factor (EGF) (adapted from Foster, 2005).

Legend: **EGF** (Epidermal growth factor); **His** (Histidine tag); **MBP** (Maltose-binding protein); **IDGR** (Residues of the Factor Xa cleavage site)

The work on the re-engineering of botulinum neurotoxin molecules has opened up new opportunities for the creation of active LH_N fragments fused to a targeting ligand (Fig 23). These novel recombinant fusion proteins target clostridial endopeptidase activity to a variety of cell types and at the same time achieve prolonged inhibition of exocytosis within the target cells. This same approach was used to generate LH_N-fragments with antibody domains, as ligands.

Legend applies to figures from 22 to 30: **LC** (Light chain of LHN); **H_N** (Heavy chain of LHN); **linker** (cleavage site); arrow indicates the cleavage site of EK or FXa; **spacer** (peptide linker between H_N and ligand); **rAb** (recombinant antibody); **scFv** (single chain antibody)

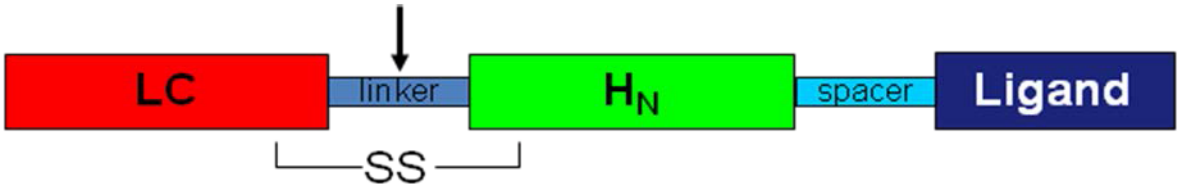


Figure 23 Design of LH_N fragment with ligand.

The recent work on antibody targeting of recombinant proteins produces an opportunity to produce modified Abs-BoNT. In some cases, the targeting ligand is not suited to expression as a recombinant protein; because it would lack essential post-translational modifications or that there may not be a ligand available with sufficient specificity. The use of antibody domain fragments or domain antibodies expressed in *E.coli* can therefore be used to target LH_N to specific cells (Fig 24). The spacer was removed from the original backbone (Fig 23) to simplify structure of novel molecules.

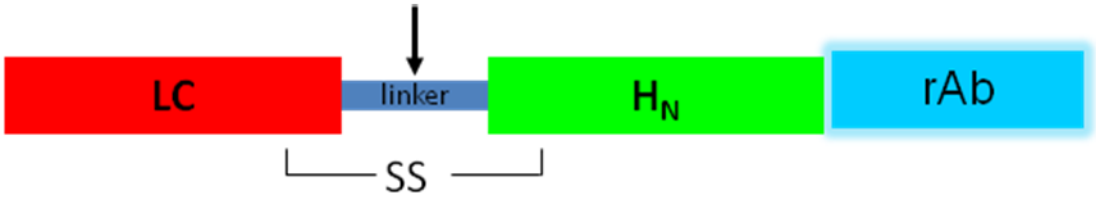


Figure 24 Design of LH_N fragment with antibody.

4.2 Results

As mentioned in Chapter 3 the DNA ScFvCD117I was synthesised for the production of LH_N molecules with single chain variable antibody fragments; and six DNA: EGFR6sdAb, EGFR7sdA, EGFR10sdAb, EGFR28sdAb, EGFR31sdAbs and EGFR43sdAb for the construction of LH_N with single domain antibodies. The DNA (5-10 µg) for all of the genes were sent in a lyophilised form and was stored at -20°C in the freezer until it was dissolved in 50 µl of distilled water by gentle shaking for 10 min. The final DNA concentration was approximately 100-200 ng/µl. Subsequently the DNA of the ScFv and sdAbs genes was microbanked according to the protocol (submitted in Chapter 2 Material and Methods). Each gene was assigned an SXNxxxxxx.

The following number was given for single chain variable antibody fragment for records in the BIX central home page at Syntaxin Ltd:

SXN101561 – pCR4TOPO – ScFvCD117I referred to throughout as CD117I

The ScFv gene present in SXN101561 was then cloned into the LH_N domain of botulinum neurotoxin type A, B and C with a His tag to facilitate affinity purification. Recombinant scFv-LH_N molecule produced possessed a 10 His tag on the N-terminus (Fig 25).

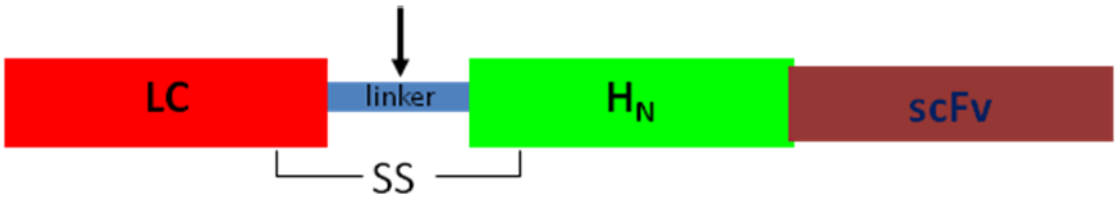


Figure 25 Design of LH_N with single chain antibody fragment.

The following numbers were given for the single domain antibodies for records in the BIX central home page at Syntaxin Ltd:

SXN101841 - pCR4TOPO- EGFR6sdAbs

SXN101842 – pCR4TOPO- EGFR7sdAbs

SXN101843 - pCR4TOPO- EGFR10sdAbs

SXN101844 - pCR4TOPO- EGFR28sdAbs

SXN101845 - pCR4TOPO- EGFR31sdAbs

SXN101846 - pCR4TOPO- EGFR43sdAbs

The above single domain antibody genes have been cloned into the LH_N domain of botulinum neurotoxin type A, B and D with His tags to facilitate affinity purification. All sdAb-LH_N molecules produced possess 10 His tag on the N-terminus (Fig 26). Their nomenclature is given as LC (for light chain N –terminal) A or B or D for type of BoNT.

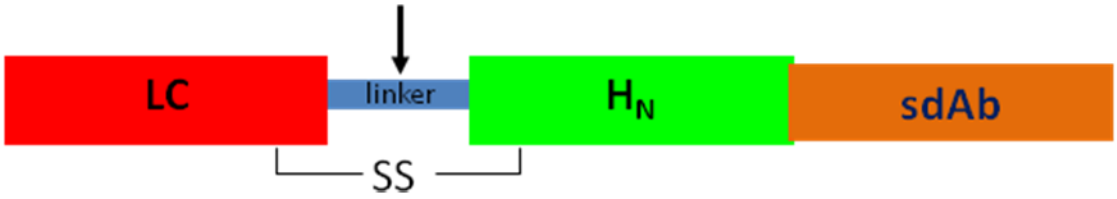


Figure 26 Design of LH_N with single domain antibody.

Two types of linkers were used in order to activate the fusion proteins: EK (Enterokinase) and FXa (Factor Xa). Enterokinase is a specific protease that cleaves after K Lysine (Lys) at its cleavage site Asp-Asp-Asp-Asp-Lys (Aspartic Acid - Aspartic Acid - Aspartic Acid – Aspartic Acid - Lysine), although it will sometimes cleave at other basic residues, depending on the conformation of the protein substrate. Enterokinase will not cleave at a site followed by P Proline (Pro). Factor Xa cleaves after the R Arginine (Arg) residue in its preferred cleavage site Ile - (Glu or Asp) - Gly – Arg (Isoleucine – (Glutamic Acid or Aspartic Acid) – Glycine – Arginine) although it will sometimes cleave at other basic residues, depending on the conformation of the protein substrate. The most common secondary site, among those that have been

sequenced, are Gly - Arg (Glycine – Arginine). Factor Xa will not cleave a site followed by Proline (Pro) or Arginine (Arg) residues.

The construct SXN101327 (Fig 27) has been selected as a backbone for the cloning of recombinant botulinum neurotoxin type B fragment fused to single chain variable antibody fragments.

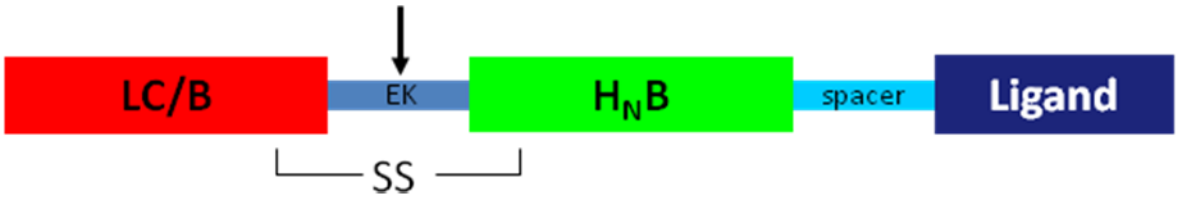


Figure 27 Design of LH_N/B fragment with activation side EK (Enterokinase) in SXN101327.

The construct SXN100532 (Fig 28) has been selected as a backbone for the cloning of recombinant botulinum neurotoxin type A fragment fused to single chain variable antibody fragments and single domain antibodies.

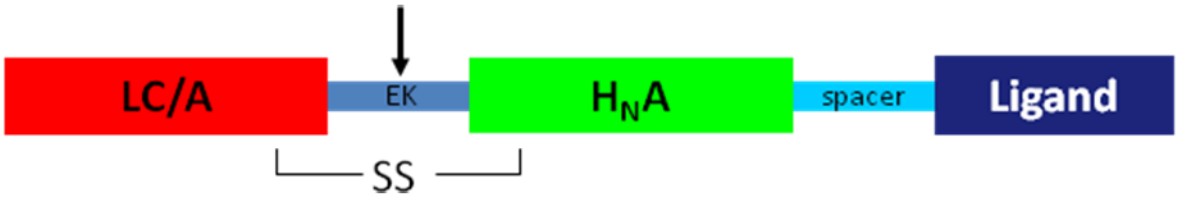


Figure 28 Design of LH_N/A fragment with activation side EK (Enterokinase) in SXN100532.

The construct SXN101832 (Fig 29) (pK7-10HT-LD-EK-HD-GS20-GRP-R) has been selected as a backbone for the cloning of recombinant botulinum neurotoxin type D fragment fused to single domain antibody.

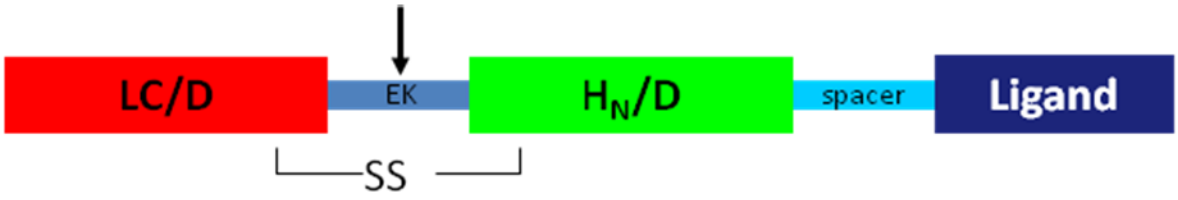


Figure 29 Design of LH_N/D fragment with activation side by EK (Enterokinase) in SXN101832.

All backbones described above include an activation site, which is cleaved between the LC and H_N domains by Enterokinase as indicated above.

The construct SXN101641 (Fig 30) has been selected as a backbone for the cloning of the recombinant botulinum neurotoxin type B fused to single chain variable antibody fragments and single domain antibodies.

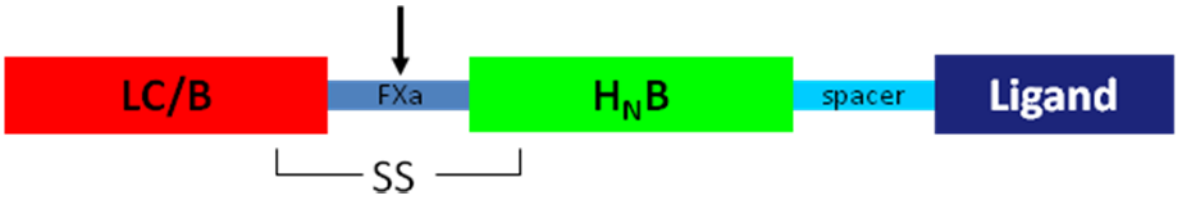


Figure 30 Design of LH_N/B fragment with activation side FXa (factor Xa) in SXN101641.

The construct SXN100736 (Fig 31) has been selected as a backbone for the cloning of recombinant botulinum neurotoxin type C fragment fused to single chain antibody.

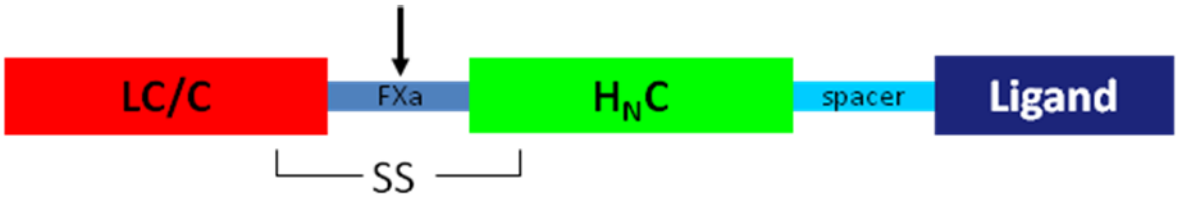


Figure 31 Design of LH_N/C fragment with activation side FXa (factor Xa) in SXN100736.

Backbones derived from molecules SXN101641 and SXN100736 include an activation site, which is cleaved between the LC and H_N domains by Factor Xa as indicated above.

Plasmid DNA was digested with *XbaI* (TCTAGA) and *HindIII* (AAGCTT) restriction endonucleases in order to insert ScFv or sdAbs in the place of the natural H_C binding domain within the vector backbone pK7 that has been modified from pET26b vector. The figures below demonstrate the cloning steps.

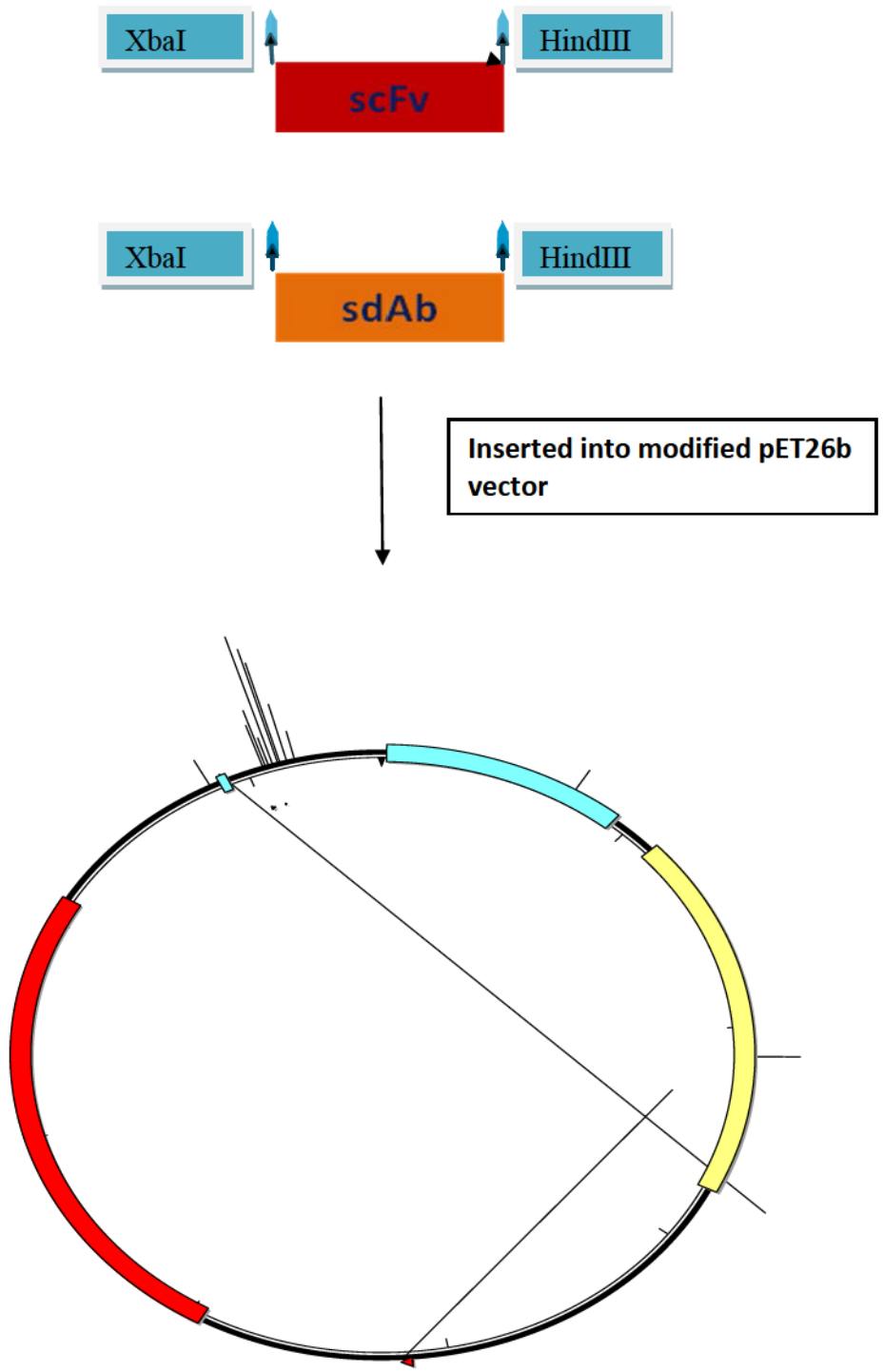


Figure 32 Cloning of DNAs encoding single variable antibody fragments (scFvs) and the single domain antibodies (sdAbs) with restriction endonucleases: *XbaI* on the (5' end) and *HindIII* on the (3' end) into modified pET25b vector.

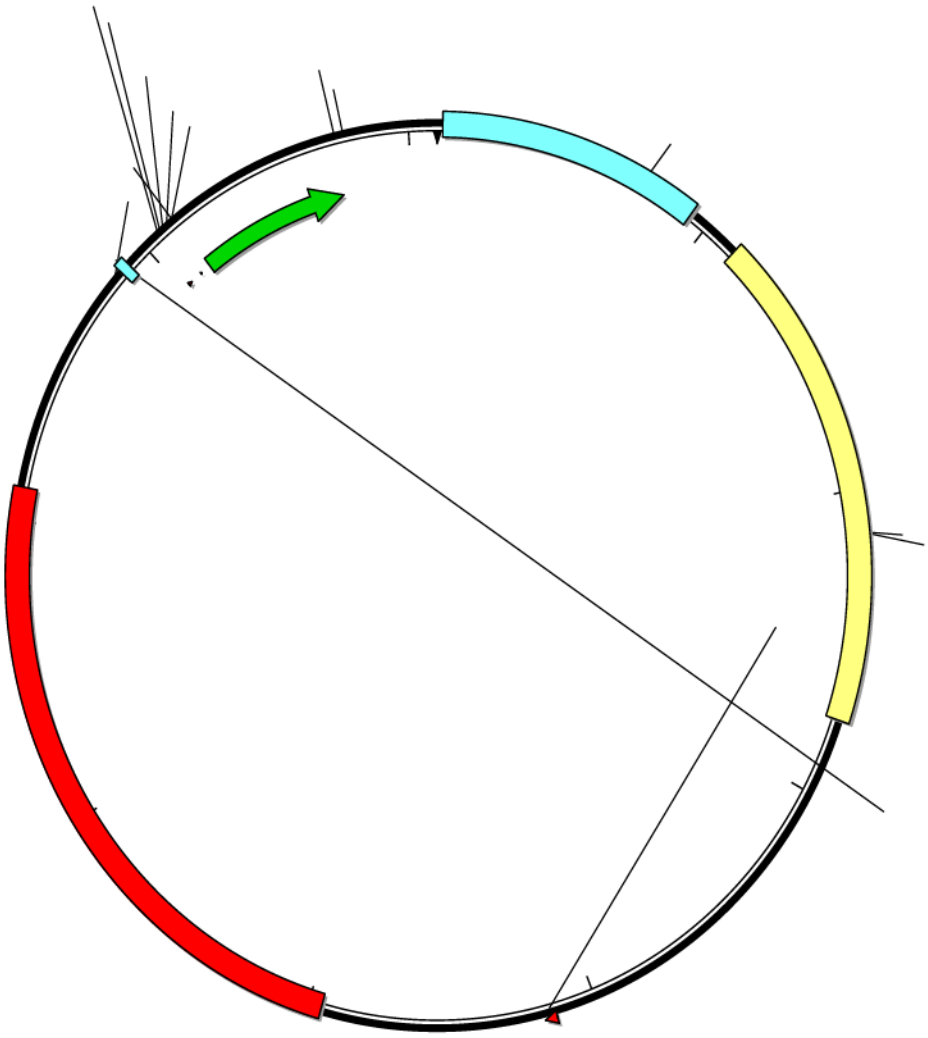


Figure 33 Modified pET26b vector with inserted scFv or sdAb.

All DNA manipulations involving cloning were performed according to standard procedures. Molecular Cloning: A Laboratory Manual, 2nd Edition (Cold Spring Harbor Laboratory Press), Sambrook, Fritsch and Maniatis (eds), 1990.

The extraction and purification of DNA from agarose gels and the purification of plasmid DNA were prepared according to standard protocols (Chapter 2 Method and Materials). The correct sizes of the cloned constructs were confirmed by using a 1 kb DNA ladder (NEB) in a 0.8 % agarose gel during electrophoresis (Fig 14, pg 55). Figures 36, 38, 39, 41, 42 and 43 illustrate cut and uncut plasmids DNA used for the cloning of scFvCD117I gene into backbones A, B and C, as well as sdAbs genes cloned into backbones A, B and D. Figures 37, 40, 44, 45 and 46 illustrate products obtained from the cloning of scFvCD117I and sdAbs genes inserted into backbones. In the end, the cloning was confirmed by submitting the DNA for sequencing to Source BioScience Lifescience former Geneservice http://www.lifesciences.sourcebioscience.com/. A large quantity of sequencing data was obtained and analysed using SeqMan (DNASTAR Lasergene 8 program); only one is presented as an example for molecule SXN101861 (LH_N/B-EGFR6sdAb) (Fig 34).

The DNA of all the constructs was transformed into TOP10 cells (Invitrogen) and BL21 (DE3) cells (Novagen) according to the protocols provided by the suppliers. The novel constructs were microbanked according to the standard protocol (Chapter 2 Methods and Materials) and stored at the -80°C in the freezer, an SXNxxxxxx number was allocated for the DNA microbank stocks, and an Exxxxxx number was given for the expression microbank stocks. DNA sequence maps were generated for 26 molecules listed in the table 10, pg 123-124. As noted above, many linear maps were created using SeqBuilder DNASTAR Lasergene 8 program; only one is presented as an example for molecule SXN101869 (LN_N/A-EGFR7sdAb) (Fig 35).

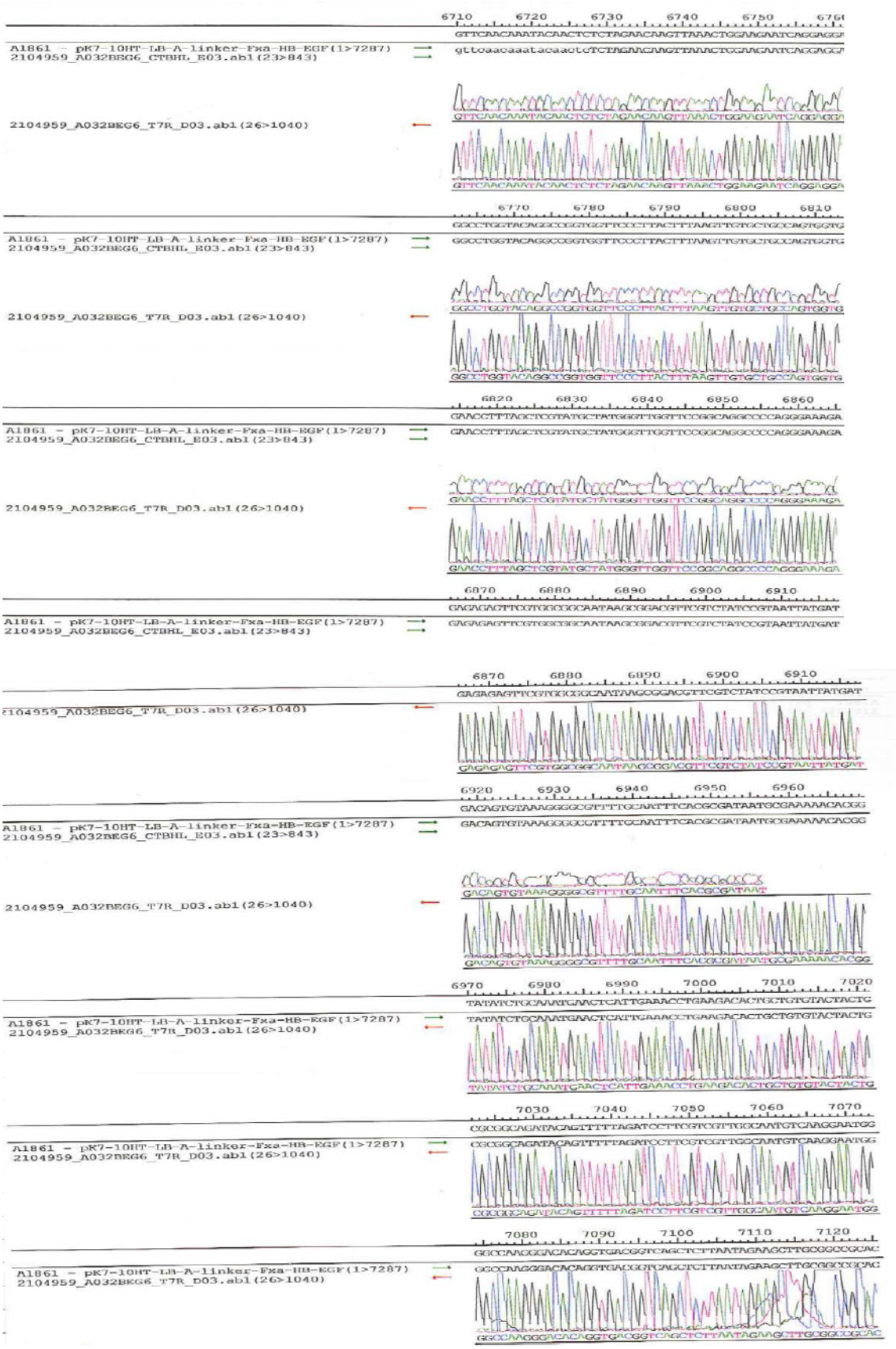


Figure 34 DNA sequencing of LH_N/B-EGFR6sdAb=SXN101861.

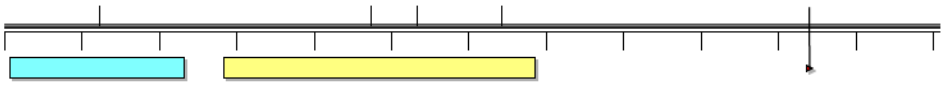


Figure 35 Linear map (SeqBuilder) of molecule LN_N/A-EGFR7sdAb=SXN101869.

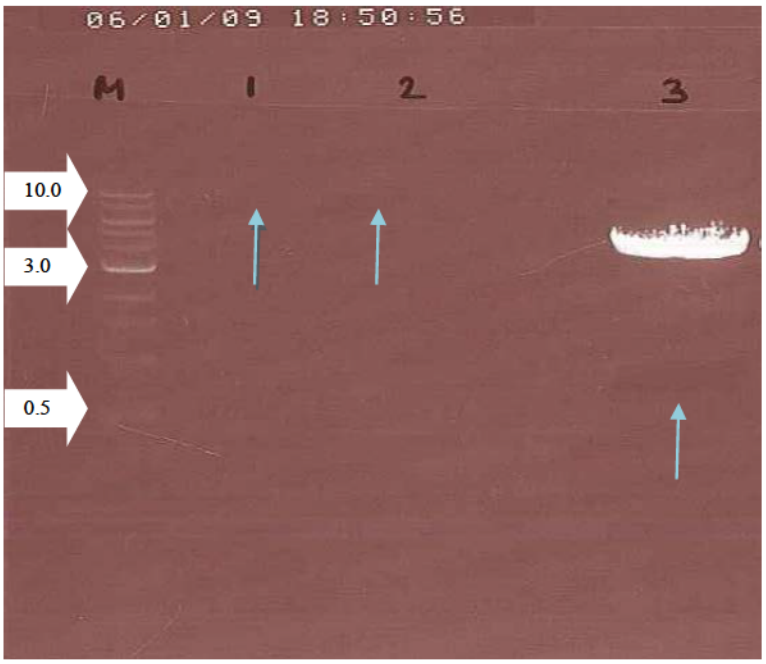


Figure 36 0.8% agarose gel of purified cut plasmid DNA.

Lane M is a 1kb DNA ladder (Fig 14, Chapter 2 Material and Methods);
Lane 1 shows the position of the excised backbone **SXN101327** (Fig 27) digested with *XbaI* and *HindIII* to obtain the vector backbone pK7 containing the LH_N/B (indicated by arrow);
Lane 2 shows the position of the excised backbone **SXN100532** (Fig 28) digested with *XbaI* and *HindIII* to obtain the vector backbone pK7 containing the LH_N/A (indicated by arrow);
Lane 3 shows the **pCR TOPO vector** digested with *XbaI* and *HindIII* and the position of the excised scFvCD117I gene obtained from the digest (indicated by arrow);

The plasmid SXN101327 is 7.11 kb, with a 6.89 kb backbone pK7 containing the LH_N/B.
The plasmid SXN100532 is 7.48 kb, with a 6.87 kb backbone pK7 containing the LH_N/A.
The pCR4 TOPO vector is a total of 4.73 kb, with a 0.75 kb of excised scFvCD117I gene.

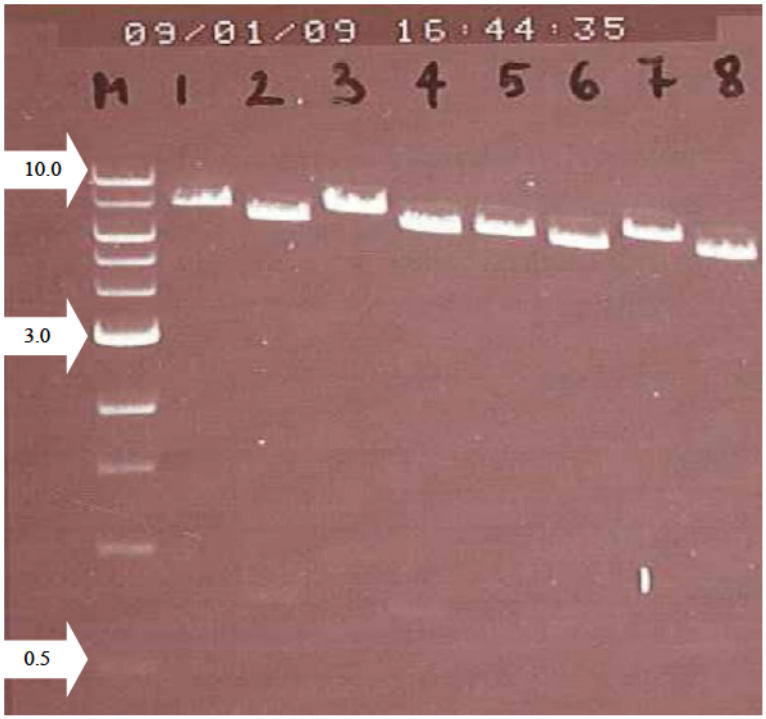


Figure 37 Products obtained following *HindIII* and *HindIII/XbaI* digests on plasmids LHN/AscFvCD117I=SXN101623 and LHN/BscFvCD117I=SXN101624 sent to Geneservice for sequencing.

Lane M is a 1 kb DNA ladder (Fig 14, Chapter 2 Material and Methods);
Lanes 1 and 3 show plasmid DNA in duplicate for **LHN/AscFvCD117I=SXN101623** (Table 10) digested with *HindIII* to unsure the correct products have been produced from ligation that was confirmed by Geneservice;
Lanes 2 and 4 show plasmid DNA in duplicate for **LHN/AscFvCD117I=SXN101623** (Table 10) digested with *HindIII/XbaI* to ensure that correct size gene (scFvCD117I) was inserted to the backbone LH_N/A;
Lanes 5 and 7 show plasmid DNA in duplicate for **LHN/BscFvCD117I=SXN101624** (Table 10) digested with *HindIII* to unsure the correct products have been produced from ligation that was confirmed by Geneservice;
Lanes 6 and 8 show plasmid DNA in duplicate for **LHN/BscFvCD117I=SXN101624** (Table 10) digested with *HindIII/XbaI* to ensure that correct size gene (scFvCD117I) was inserted to the backbone LH_N/B;

The plasmid LHN/AscFvCD117I=SXN101623 is 7.63 kb.
The plasmid LHN/BscFvCD117I=SXN101624 is 7.64 kb.

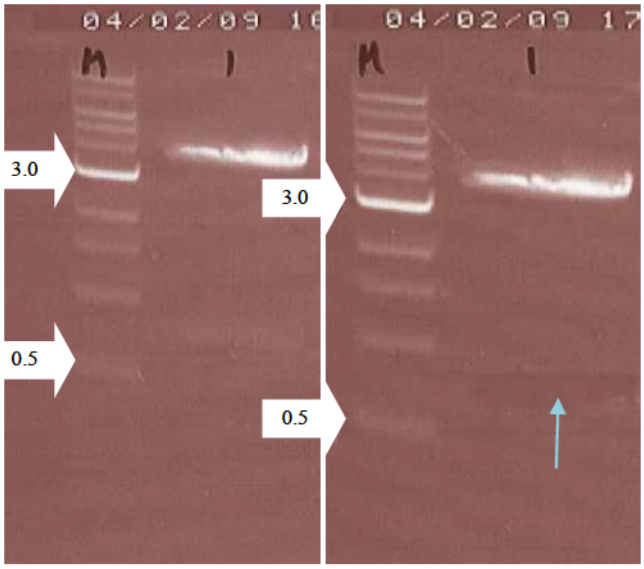


Figure 38 0.8% agarose gel of purified uncut and cut plasmid DNA of scFvCD117I.

Lane M is a 1 kb DNA ladder (Fig 14, Chapter 2 Material and Methods); lane 1 shows the pCR4 TOPO vector digested with *XbaI* and *HindIII* and the position of the excised scFvCD117I gene obtained from the digest (indicted by arrow);

The pCR4 TOPO vector is a total of 4.73 kb, with a 0.75 kb of excised scFvCD117I gene.

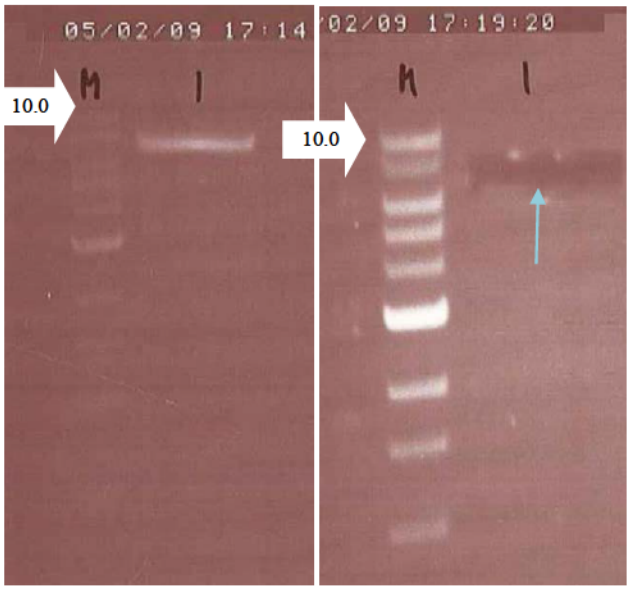
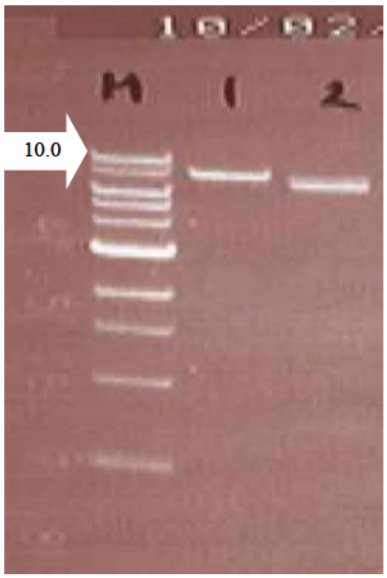


Figure 39 0.8% agarose gel of purified uncut and cut plasmid DNA of SXN101641.

Lane M is a 1 kb DNA ladder (Fig 14, Chapter 2 Material and Methods); lane 1 shows the SXN101641 gene digested with *XbaI* and *HindIII* and the position of the excised vector backbone obtained from the digest (indicted by arrow);

The plasmid SXN101641 is 6.99 kb, with a 6.82 kb backbone pK7 containing the LH_N/B.



Lane M is a 1 kb DNA ladder (Fig 14, Chapter 2 Material and Methods);
Lane 1 shows plasmid DNA **LH_N/BscFvCD117I=SXN101776** (Table 10) digested with *HindIII* to unsure the correct products have been produced from ligation that was confirmed by Geneservice;
Lane 2 shows plasmid DNA **LH_N/BscFvCD117I=SXN101776** (Table 10) digested with *HindIII/XbaI* to ensure that correct size gene (scFvCD117I) was inserted to the backbone LH_N/B;

The plasmid LH_N/BscFvCD117I=SXN101776 is 7.65 kb.

Figure 40 Product obtained following *HindIII* and *HindIII/XbaI* digest on plasmid LH_N/BscFvCD117I=SXN101776 sent to Geneservice for sequencing.

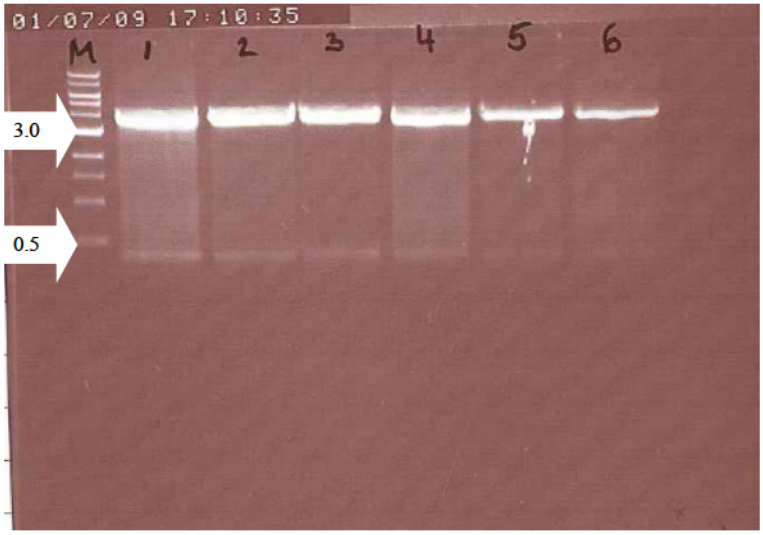


Figure 41 0.8% agarose gel of purified uncut plasmid DNA of sdAbs.

Lane M is a 1kb DNA ladder (Fig 14, Chapter 2 Material and Methods);
Lane 1 shows the pCR4 TOPO vector containing the EGFR43 sdAb gene (SXN101846);
Lane 2 shows the pCR4 TOPO vector containing the EGFR31 sdAb gene (SXN101845);
Lane 3 shows the pCR4 TOPO vector containing the EGFR28 sdAb gene (SXN101844);
Lane 4 shows the pCR4 TOPO vector containing the EGFR10 sdAb gene (SXN101843);
Lane 5 shows the pCR4 TOPO vector containing the EGFR7 sdAb gene (SXN101842);
Lane 6 shows the pCR4 TOPO vector containing the EGFR6 sdAb gene (SXN101841);

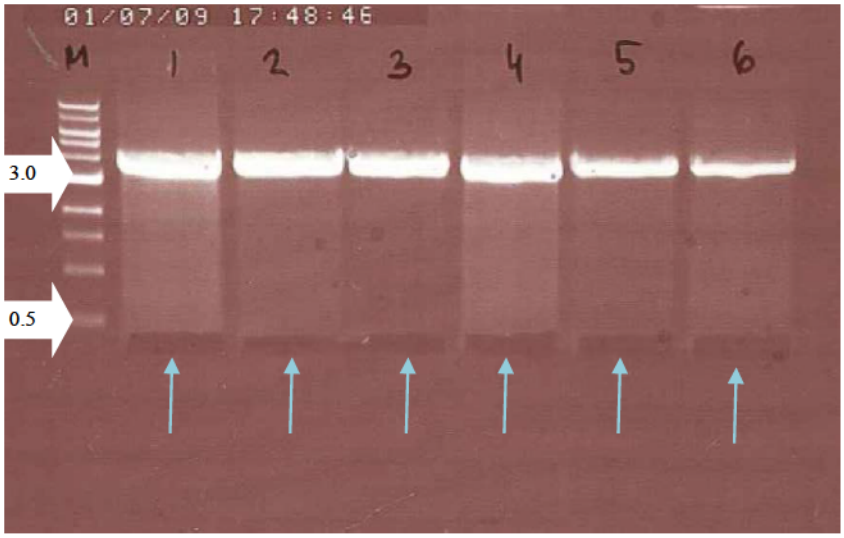


Figure 42 0.8% agarose gel of purified cut plasmid DNA of sdAbs.

Lane M is a 1 kb DNA ladder (Fig 14, Chapter 2 Material and Methods);
Lane 1 shows the pCR4 TOPO vector containing the EGFR43 sdAb gene (SXN101846);
Lane 2 shows the pCR4 TOPO vector containing the EGFR31 sdAb gene (SXN101845);
Lane 3 shows the pCR4 TOPO vector containing the EGFR28 sdAb gene (SXN101844);
Lane 4 shows the pCR4 TOPO vector containing the EGFR10 sdAb gene (SXN101843);
Lane 5 shows the pCR4 TOPO vector containing the EGFR7 sdAb gene (SXN101842);
Lane 6 shows the pCR4 TOPO vector containing the EGFR6 sdAb gene (SXN101841);

The **SXN101846** plasmid is a total of 4.3 kb, with a 3.94 kb vector and 0.36 kb of excised EGFR43 sdAb gene (indicated by arrow).
The **SXN101845** plasmid is a total of 4.3 kb, with a 3.94 kb vector and 0.36 kb of excised EGFR31 sdAb gene (indicated by arrow).
The **SXN101844** plasmid is a total of 4.3 kb, with a 3.91 kb vector and 0.39 kb of excised EGFR28 sdAb gene (indicated by arrow).
The **SXN101843** plasmid is a total of 4.3 kb, with a 3.92 kb vector and 0.38 kb of excised EGFR10 sdAb gene (indicated by arrow).
The **SXN101842** plasmid is a total of 4.3 kb, with a 3.94 kb vector and 0.36 kb of excised EGFR7 sdAb gene (indicated by arrow).
The **SXN101841** plasmid is a total of 4.3 kb, with a 3.92 kb vector and 0.38 kb of excised EGFR6 sdAb gene (indicated by arrow).

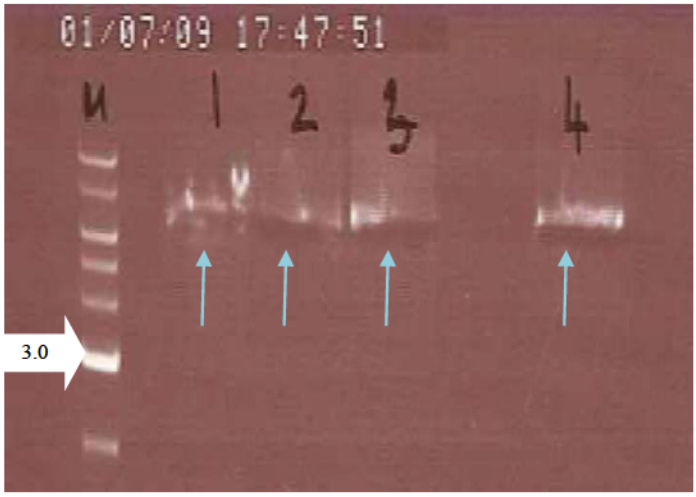


Figure 43 0.8% agarose gel of purified cut plasmid DNA.

Lane M is a 1 kb DNA ladder (Fig 14, Chapter 2 Material and Methods);
Lane 1 shows the position of the excised **backbone SXN101832** (Fig 29) digested with *XbaI* and *HindIII* to obtain the vector backbone pK7 containing the **LH_N/D** (indicated by arrow);
Lane 2 shows the position of the excised **backbone SXN101641** (Fig 30) digested with *XbaI* and *HindIII* to obtain the vector backbone pK7 containing the **LH_N/B** (indicated by arrow);
Lane 3 shows the position of the excised **backbone SXN100736** (Fig 31) digested with *XbaI* and *HindIII* to obtain the vector backbone pK7 containing the **LH_N/C** (indicated by arrow);
Lane 4 shows the position of the excised **backbone SXN100532** (Fig 28) digested with *XbaI* and *HindIII* to obtain the vector backbone pK7 containing the **LH_N/A** (indicated by arrow);

The plasmid SXN101832 is 7.05 kb, with a 6.89 kb backbone pK7 containing the LH_N/D.
The plasmid SXN101641 is 6.99 kb, with a 6.82 kb backbone pK7 containing the LH_N/B.
The plasmid SXN100736 is 7.04 kb, with a 6.88 kb backbone pK7 containing the LH_N/C.
The plasmid SXN100532 is 7.48 kb, with a 6.87 kb backbone pK7 containing the LH_N/A.

Figure 44 Products obtained following *HindIII* and *HindIII/PstI* digests on LH_N/sdAbs plasmids sent to Geneservice for sequencing.

Lane M is a 1 kb DNA ladder (Fig 14, Chapter 2 Material and Methods);
Lane 1 shows plasmid DNA **LH_N/A-EGFR6sdAb=SXN101868** (Table 10) digested with *HindIII* to unsure the correct products have been produced from ligation;
Lane 2 shows plasmid DNA **LH_N/A-EGFR6sdAb=SXN101868** (Table 10) digested with *HindIII/PstI* to ensure that correct size gene (EGFR6sdAb) was inserted to the backbone LH_N/A;
Lane 3 shows plasmid DNA **LN_N/A-EGFR7sdAb=SXN101869** (Table 10) digested with *HindIII* to unsure the correct products have been produced from ligation;
Lane 4 shows plasmid DNA **LN_N/A-EGFR7sdAb=SXN101869** (Table 10) digested with *HindIII/PstI* to ensure that correct size gene (EGFR7sdAb) was inserted to the backbone LH_N/A;
Lane 5 shows plasmid DNA **LH_N/A-EGFR10sdAb=SXN101870** (Table 10) digested with *HindIII* to unsure the correct products have been produced from ligation;
Lane 6 shows plasmid DNA **LH_N/A-EGFR10sdAb=SXN101870** (Table 10) digested with *HindIII/PstI* to ensure that correct size gene (EGFR10sdAb) was inserted to the backbone LH_N/A;
Lane 7 shows plasmid DNA **LH_N/A-EGFR28sdAb=SXN101871** (Table 10) digested with *HindIII* to unsure the correct products have been produced from ligation;
Lane 8 shows plasmid DNA **LH_N/A-EGFR28sdAb=SXN101871** (Table 10) digested with *HindIII/PstI* to ensure that correct size gene (EGFR28sdAb) was inserted to the backbone LH_N/A;
Lane 9 shows plasmid DNA **LH_N/A-EGFR31sdAb=SXN101872** (Table 10) digested with *HindIII* to unsure the correct products have been produced from ligation;
Lane 10 shows plasmid DNA **LH_N/A-EGFR31sdAb=SXN101872** (Table 10) digested with *HindIII/PstI* to ensure that correct size gene (EGFR31sdAb) was inserted to the backbone LH_N/A;
Lane 11 shows plasmid DNA **LH_N/A-EGFR43sdAb=SXN101873** (Table 10) digested with *HindIII* to unsure the correct products have been produced from ligation;
Lane 12 shows plasmid DNA **LH_N/A-EGFR43sdAb=SXN101873** (Table 10) digested with *HindIII/PstI* to ensure that correct size gene (EGFR43sdAb) was inserted to the backbone LH_N/A;
Lane 13 shows plasmid DNA **LH_N/B-EGFR6sdAb=SXN101861** (Table 10) digested with *HindIII* to unsure the correct products have been produced from ligation;
Lane 14 shows plasmid DNA **LH_N/B-EGFR6sdAb=SXN101861** (Table 10) digested with *HindIII/PstI* to ensure that correct size gene (EGFR6sdAb) was inserted to the backbone LH_N/B;
Lane 15 shows plasmid DNA **LH_N/B-EGFR7sdAb=SXN101862** (Table 10) digested with *HindIII* to unsure the correct products have been produced from ligation;
Lane 16 shows plasmid DNA **LH_N/B-EGFR7sdAb=SXN101862** (Table 10) digested with *HindIII/PstI* to ensure that correct size gene (EGFR7sdAb) was inserted to the backbone LH_N/B;
Lane 17 shows plasmid DNA **LH_N/B-EGFR10sdAb=SXN101863** (Table 10) digested with *HindIII* to unsure the correct products have been produced from ligation;
Lane 18 shows plasmid DNA **LH_N/B-EGFR10sdAb=SXN101863** (Table 10) digested with *HindIII/PstI* to ensure that correct size gene (EGFR10sdAb) was inserted to the backbone LH_N/B;

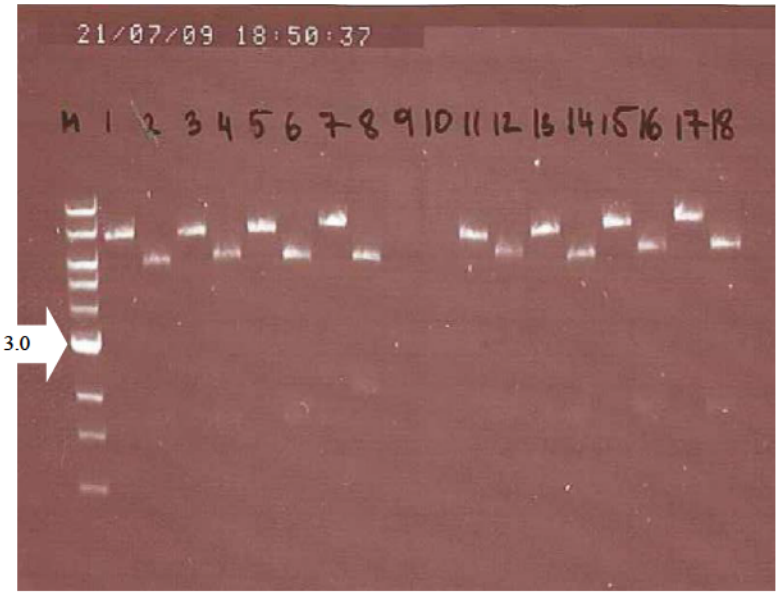


Figure 45 Products obtained following *HindIII* and *HindIII/PstI* digests on LH_N/sdAbs plasmids sent to Geneservice for sequencing.

Lane M is a 1 kb DNA ladder (Fig 14, Chapter 2 Material and Methods);
Lane 1 shows plasmid DNA **LH_N/B-EGFR28sdAb=SXN101864** (Table 10) digested with *HindIII* to unsure the correct products have been produced from ligation;
Lane 2 shows plasmid DNA **LH_N/B-EGFR28sdAb=SXN101864** (Table 10) digested with *HindIII/PstI* to ensure that correct size gene (EGFR28sdAb) was inserted to the backbone LH_N/B;
Lane 3 shows plasmid DNA **LH_N/B-EGFR31sdAb=SXN101865** (Table 10) digested with *HindIII* to unsure the correct products have been produced from ligation;
Lane 4 shows plasmid DNA **LH_N/B-EGFR31sdAb=SXN101865** (Table 10) digested with *HindIII/PstI* to ensure that correct size gene (EGFR31sdAb) was inserted to the backbone LH_N/B;
Lane 5 shows plasmid DNA **LH_N/B-EGFR43sdAb=SXN101866** (Table 10) digested with *HindIII* to unsure the correct products have been produced from ligation;
Lane 6 shows plasmid DNA **LH_N/B-EGFR43sdAb=SXN101866** (Table 10) digested with *HindIII/PstI* to ensure that correct size gene (EGFR43sdAb) was inserted to the backbone LH_N/B;
Lane 7 shows plasmid DNA **LH_N/D-EGFR6sdAb=SXN101874** (Table 10) digested with *HindIII* to unsure the correct products have been produced from ligation;
Lane 8 shows plasmid DNA **LH_N/D-EGFR6sdAb=SXN101874** (Table 10) digested with *HindIII/PstI* to ensure that correct size gene (EGFR6sdAb) was inserted to the backbone LH_N/D;
Lane 9 is blank;
Lane 10 is blank;
Lane 11 shows plasmid DNA **LH_N/D-EGFR7sdAb=SXN101875** (Table 10) digested with *HindIII* to unsure the correct products have been produced from ligation;
Lane 12 shows plasmid DNA **LH_N/D-EGFR7sdAb=SXN101875** (Table 10) digested with *HindIII/PstI* to ensure that correct size gene (EGFR7sdAb) was inserted to the backbone LH_N/D;
Lane 13 shows plasmid DNA **LH_N/D-EGFR10sdAb=SXN101876** (Table 10) digested with *HindIII* to unsure the correct products have been produced from ligation;
Lane 14 shows plasmid DNA **LH_N/D-EGFR10sdAb=SXN101876** (Table 10) digested with *HindIII/PstI* to ensure that correct size gene (EGFR10sdAb) was inserted to the backbone LH_N/D;
Lane 15 shows plasmid DNA **LH_N/D-EGFR28sdAb=SXN101877** (Table 10) digested with *HindIII* to unsure the correct products have been produced from ligation;
Lane 16 shows plasmid DNA **LH_N/D-EGFR28sdAb=SXN101877** (Table 10) digested with *HindIII/PstI* to ensure that correct size gene (EGFR28sdAb) was inserted to the backbone LH_N/D;
Lane 17 shows plasmid DNA **LH_N/D-EGFR31sdAb=SXN101878** (Table 10) digested with *HindIII* to unsure the correct products have been produced from ligation;
Lane 18 shows plasmid DNA **LH_N/D-EGFR31sdAb=SXN101878** (Table 10) digested with *HindIII/PstI* to ensure that correct size gene (EGFR31sdAb) was inserted to the backbone LH_N/D;

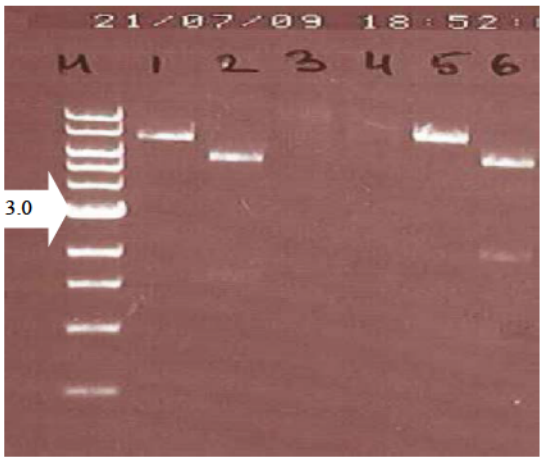


Figure 46 Products obtained following *HindIII* and *HindIII/PstI* digests on LH_N/sdAb plasmid and LH_N/C-scFvCD117I=SXN101882 sent to Geneservice for sequencing.

Lane M is a 1 kb DNA ladder (Fig 14, Chapter 2 Material and Methods);
Lane 1 shows plasmid DNA **LH_N/D-EGFR43sdAb=SXN101879** (Table 10) digested with *HindIII* to unsure the correct products have been produced from ligation;
Lane 2 shows plasmid DNA **LH_N/D-EGFR43sdAb=SXN101879** (Table 10) digested with *HindIII/PstI* to ensure that correct size gene (EGFR43sdAb) was inserted to the backbone LH_N/D;
Lane 3 is blank;
Lane 4 is blank;
lane 5 shows plasmid DNA **LH_N/C-scFvCD117I=SXN101882** (Table 10) digested with *HindIII* to unsure the correct products have been produced from ligation that was confirmed by Geneservice;
Lane 6 shows plasmid DNA **LH_N/C-scFvCD117I=SXN101882** (Table 10) digested with *HindIII/XbaI* to ensure that correct size gene (scFvCD117I) was inserted to the backbone LH_N/C;

The plasmid **LH_N/A-EGFR6sdAb=SXN101868** is 7.26 kb.
The plasmid **LH_N/A-EGFR7sdAb=SXN101869** is 7.24 kb.
The plasmid **LH_N/A-EGFR10sdAb=SXN101870** is 7.26 kb.
The plasmid **LH_N/A-EGFR28sdAb=SXN101871** is 7.26 kb.
The plasmid **LH_N/A-EGFR31sdAb=SXN101872** is 7.24 kb.
The plasmid **LH_N/A-EGFR43sdAb=SXN101873** is 7.24 kb.
The plasmid **LH_N/B-EGFR6sdAb=SXN101861** is 7.28 kb.
The plasmid **LH_N/B-EGFR7sdAb=SXN101862** is 7.26kb.
The plasmid **LH_N/B-EGFR10sdAb=SXN101863** is 7.29kb.
The plasmid **LH_N/B-EGFR28sdAb=SXN101864** is 7.29 kb.
The plasmid **LH_N/B-EGFR31sdAb=SXN101865** is 7.26 kb.
The plasmid **LH_N/B-EGFR43sdAb=SXN101866** is 7.26 kb.
The plasmid **LH_N/D-EGFR6sdAb=SXN101874** is 7.28 kb.
The plasmid **LH_N/D-EGFR7sdAb=SXN101875** is 7.26 kb.
The plasmid **LH_N/D-EGFR10sdAb=SXN101876** is 7.28 kb
The plasmid **LH_N/D-EGFR28sdAb=SXN101877** is 7.29 kb
The plasmid **LH_N/D-EGFR31sdAb=SXN101878** is 7.26 kb
The plasmid **LH_N/D-EGFR43sdAb=SXN101879** is 7.26 kb
The plasmid **LH_N/C-scFvCD117I=SXN101882** is 7.64 kb

Table 10 LH_N antibody based chimeras cloned for purpose of this study.

SXN number	E number	Vector	Tag	LH_N	Activation side	Type of recombinant antibody	MW (molecular weight) kDa	Domain-wise order used in a text
SXN101623	E001219	pK7	10HT	LH_N /A	EK	**ScFvCD117I**	128534.5	**LH_N/A-scFvCD117I= SXN101623**
SXN101624	E001220	pK7	10HT	LH_N /B	EK	**ScFvCD117I**	129741.2	LH_N/B-scFvCD117I= SXN101624
SXN101776	E001342	pK7	10HT	LH_N /B	FXa	**ScFvCD117I**	129883.7	**LH_N/B-scFvCD117I= SXN101776**
SXN101882	E001427	pK7	10HT	LH_N /C	FXa	**ScFvCD117I**	129173.8	LH_N/C-scFvCD117I= SXN101882
SXN101861	E001406	pK7	10HT	LH_N/B	FXa	**EGFR6sdAb**	116880.5	**LH_N/B-EGFR6sdAb= SXN101861**
SXN101862	E001407	pK7	10HT	LH_N/B	FXa	**EGFR7sdAb**	116653.2	LH_N/B-EGFR7sdAb= SXN101862
SXN101863	E001408	pK7	10HT	LH_N/B	FXa	**EGFR10sdAb**	116986.6	**LH_N/B-EGFR10sdAb= SXN101863**
SXN101864	E001409	pK7	10HT	LH_N/B	FXa	**EGFR28sdAb**	117604.1	LH_N/B-EGFR28sdAb= SXN101864
SXN101865	E001410	pK7	10HT	LH_N/B	FXa	**EGFR31sdAb**	115531.3	**LH_N/B-EGFR31sdAb= SXN101865**
SXN101866	E001411	pK7	10HT	LH_N/B	FXa	**EGFR43sdAb**	129565.3	LH_N/B-EGFR43sdAb= SXN101866
SXN101868	E001413	pK7	10HT	LH_N/A	EK	**EGFR6sdAb**	116345.0	**LH_N/A-EGFR6sdAb= SXN101868**
SXN101869	E001414	pK7	10HT	LH_N/A	EK	**EGFR7sdAb**	115140.8	LN_N/A-EGFR7sdAb= SXN101869
SXN101870	E001415	pK7	10HT	LH_N/A	EK	**EGFR10sdAb**	115474.2	**LH_N/A-EGFR10sdAb= SXN101870**

SXN10 1871	E001416	pK7	10HT	LH_N/A	EK	**EGFR28sdAb**	116091.7	LH_N/A-EGFR28sdAb= SXN101871
SXN10 1872	E001417	pK7	10HT	LH_N/A	EK	**EGFR31sdAb**	114832.6	**LH_N/A-EGFR31sdAb= SXN101872**
SXN10 1873	E001418	pK7	10HT	LH_N/A	EK	**EGFR43sdAb**	114957.6	LH_N/A-EGFR43sdAb= SXN101873
SXN10 1874	E001419	pK7	10HT	LH_N/D	EK	**EGFR6sdAb**	116458.1	**LH_N/D-EGFR6sdAb= SXN101874**
SXN10 1875	E001420	pK7	10HT	LH_N/D	EK	**EGFR7sdAb**	116067.5	LH_N/D-EGFR7sdAb= SXN101875
SXN10 1876	E001421	pK7	10HT	LH_N/D	EK	**EGFR10sdAb**	116401.0	**LH_N/D-EGFR10sdAb= SXN101876**
SXN10 1877	E001422	pK7	10HT	LH_N/D	EK	**EGFR28sdAb**	117018.5	LH_N/D-EGFR28sdAb= SXN101877
SXN10 1878	E001423	pK7	10HT	LH_N/D	EK	**EGFR31sdAb**	115759.4	**LH_N/D-EGFR31sdAb= SXN101878**
SXN10 1879	E001424	pK7	10HT	LH_N/D	EK	**GFR43sdAb**	115884.4	LH_N/D-EGFR43sdAb= SXN101879

Table 11 LH_N control chimeras available at Syntaxin Ltd.

SXN number	E number	Vector	Tag	LH_N	Activation side	MW (molecular weight) kDa	Domain-wise order used in a text
SXN100590	E000445	pK7	10HT	LH_N/A	EK	101070.5	LH_N/A=SXN100590
SXN101687	E001283	pK7	10HT	LH_N/B	FXa	102419.7	LH_N/B=SXN101687
SXN101654	E001250	pK7	10HT	LH_N/C	EK	102535.7	LH_N/C=SXN101654
SXN101655	E001251	pK7	10HT	LH_N/D	EK	101997.2	LH_N/D=SXN101655

Chapter 5

Expression of recombinant antibody based constructs

Chapter 5. Expression of recombinant antibody-based constructs

Successful production of novel antibody-botulinum toxin molecules has been obtained in an *E. coli* expression system using an expression vector modified in-house called pK7 that is based upon pET26b.

5.1 Overview of expression system

Expression systems are designed to produce large amounts of the desired protein within the host cell. In order to accomplish this, the gene encoding the required protein is cloned into an expression vector, which is then transformed into a host cell. Typically, the expression vector contains genetic elements to control protein expression which include a promoter sequence appropriate to the host cell, a DNA sequence that terminates transcription and a DNA sequence that codes for ribosome binding (Heller *et. al.,* 2001).

Recombinant antibody fragments (rAb) have been expressed in wide variety of hosts including prokaryotes, such as *E. coli* (Martin *et. al.*, 2006), eukaryotes, including *S. cerevisiae* (Hackel *et. al.*, 2006), *Pichia pastoris* (Ren *et. al.,* 2008), insect cells (Bruenke *et. al.,* 2004), plant cells (Mayfield *et. al.,* 2003), plants (Makvandi-Nejad *et. al.,* 2005) and mammalian cells (Natsume *et. al.*, 2006). The large-scale production of antibody fragments for therapeutic purposes has been made possible through the development of expression systems that are effective and cost efficient (Andersen and Reilly, 2004). The optimal expression system depends on the type of recombinant antibody fragments, the precise sequence of the individual antibody being expressed, as well as the required purity and quantity of the final product (Verma, Boleti and George, 1998). The glycosylation process is critical for monoclonal antibody production (Jefferis, 2005). N-linked glycosylation at N Asparagine (Asn) -297 of the C_H2 domain of IgG has functional and structural consequences because it affects the interaction of the F_C region with cellular F_C. Therefore, the mammalian cell expression system is the most suitable for the production of monoclonal antibodies (Filpula, 2007). The antibody fragments, such as sdAbs or ScFvs do not require glycosylation process, as they are lack

the C_H2 domain. Therefore bacterial or yeast fermentation can be used for the large-scale production of recombinant protein. In the case of single chain antibodies, a few expression systems have been assessed. The successful production in yeast and fungi have been achieved in *S. cerevisiae* (Shusta *et. al.*, 1998, Hackel *et. al.*, 2006) *Pichia pastoris* (Ren *et. al.,* 2008, Gurkan *et. al.*, 2004, Shi *et. al.,* 2003, Hellwing *et. al.*, 2001, Cunha *et. al.,* 2004, Eldin 1997, Boado, Ji and Pardrige, 2000) *Aspergillus awamori* (Sotiriadis *et. al.*, 2001) and *Kluyveromyces lactis* (Swennen *et. al.,* 2002). This expression systems offers cost-effectiveness and scale up benefits combined with advantages of eukaryotic expression, including protein processing, folding and post translation modifications. Additionally purified protein does not contain endotoxins which are a contaminant present in proteins purified following expression in *E.coli* (Weisser and Hall, 2009). Although the recombinant antibody production in fungi and especially in yeast is practicable, there is still a need to shorten the process time to increase the volumetric productivity (Gasser and Mattanovich, 2007). Recent research based on the compression of *S. cerevisiae, Pichia pastoris* and *E. coli* expression systems indicated that *E. coli* was the fastest and most consistent way to obtain and characterize purified scFvs (Miller *et. al.*, 2005). Another method of expressing single chain antibodies is using transgenic plants (Makvandi-Nejad *et. al.*, 2005). The advantage of producing rAbs in plants is that they lack any mammalian pathogens and contain no endotoxin contamination; however, the long initial time for production, regulatory uncertainties and questions about the suitability of plant glycans in human therapeutics (Andersen and Reilly, 2004) has postponed their implementation for commercial production (Pujol *et. al.,* 2005).

Single domain antibodies are often produced in *E. coli* (Arbabi-Ghahroudi *et. al.,* 1997, Rahbarizadeh *et. al.*, 2005). There is only one example of VHH sdAb production in filamentous fungi, which resulted in limited proteolytic degradation of the secreted product (Joosten *et. al.,* 2005), because of high levels of proteases secreted by this host (Gerngross, 2004). The single domain antibodies have been also produced in the yeast expression host *S. cerevisiae* (Frenken *et. al.*, 2000, Thomassen *et. al.*, 2002, Van der Vaart, 2002) and *Pichia pastoris* (Rahbarizadeh *et. al.*, 2006). However occasionally yeast produced, are N-glycosylated sdAbs (Frenken *et. al.*, 2000, Harmsen *et. al.*, 2005) which can reduce antigen binding (Van der Vaart et. al., 2006). In addition, the

presence of yeast specific oligosacchariades can result in high immunogenicity and decreased serum half-life of the recombinant protein (Sethuraman and Stadheim 2006).

E. coli is the most suitable bacterial expression system for production of small non-glycosylated recombinant antibody fragments (Wang *et. al.,* 2008). For the purpose of this study the single domain or single chain antibodies were fused to recombinant botulinum neurotoxin fragments. It has been reported that one of the best approach to stabilize proteins is to express them as a fusion proteins (Uhlen *et. al.*, 1992, Nygren *et. al*., 1994, LaVallie and McCoy, 1995). There are several fusion tags to be considered as solubilising agents; that is when they fused with insoluble partner the product is a fusion protein with enhanced solubility (Shaki-Loewenstein *et. al.,* 2005). The most popular solubilising fusion tags include glutathione S-transferase (GST) (Tudyka and Skerra, 1997), thioredoxin (LaVallie *et. al*., 1993) and maltose-binding protein (MBP) (Kapust and Waugh, 1999). The successful expression of the LH_N fragment of serotype A was first reported in 2002. The presence of GST as a fusion partner on the N-terminal of light chain allowed the expression of soluble LH_N/A protein in the cytosol of *Escherichia coli*. The expression of these molecules opened up an opportunity to produce a non-toxic 100kDa endopeptidase species of clostridia neurotoxin, which are able to cleave the relevant SNARE protein (Foster and Chaddock, 2007). The *E. coli* expression system has been used for the production of antibody targeting LH_N molecules of different types of endopeptidases.

5.2 *Escherichia coli* expression system

The bacterial expression system is commonly used for the production of immunoglobulin fragments. The main advantage of this system is the ability to produce high quantities of protein. The transformation of *E. coli* cells with the foreign DNA is straightforward and can be accomplished with small amount of DNA. This inexpensive expression system is usually a first choice for protein production in many biotech companies. It has been employed for antibody production, where glycosylation process is not required. The antibody fragments expressed in *E. coli* are stable, highly soluble and react specifically and with high affinity to the antigens (Rahbarizadeh *et. al.,* 2005). The well-established knowledge of the basic cellular

processes of *E. coli* opens the way for easy genetic manipulation and characterization of protein expression and a determination of what factors could improve the yield and quality of purified proteins (Leong and Chen, 2007). The expression efficiency of recombinant molecules depend on the efficiency of transcription, translation and correct folding or refolding, which usually is due to the characteristics of the individual recombinant protein and their interaction with the expression system. Various strategies have been developed for targeting the expression of active rAbs to distinct compartments of *E. coli* cells. Secretion during the bacterial expression can occur into extracellular and periplasmic spaces, the inner and outer membrane or within the cytoplasm. (Weisser and Hall, 2009) (Fig 34).

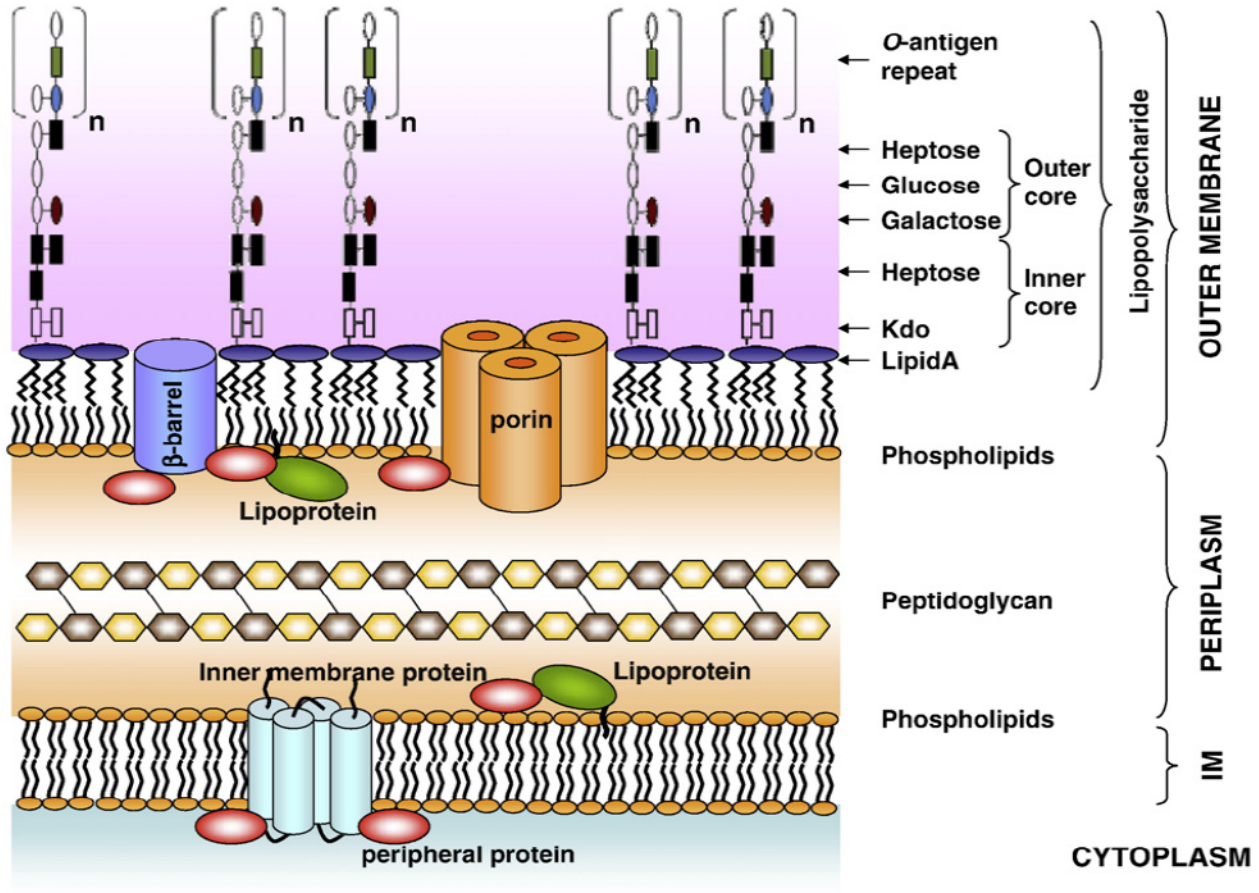


Figure 47 Structure of *Escherichia coli* cell envelope. (adapted from Weisser and Hall, 2009)

The expression of the recombinant antibody fragments in the reducing environment of the cytoplasm usually results in the formation of insoluble inclusion bodies that contain unfolded protein (Martineau *et. al.*, 1998). The purification of recombinant products from inclusion bodies has recently been improved by the combination of genetic engineering techniques and metal-affinity chromatography. Periplasmic secretion increases the probability of correct protein folding due to the more oxidizing cellular environment that favours correct disulphide bond formation. (Skerra and Pluckthun, 1988). The formation of intra-domain disulphide bonds that are essential for correct folding and functionality of the antibody fragments design for this research are more likely to be stabilized in a presence of LH_N molecule. Up to date the production of LH_N molecules fused to various ligands has been successfully employed in *E. coli* and therefore this expression system may be appropriate for the production of novel antibody-botulinum toxin molecules. In addition, the production of recombinant immunotoxin in *Escherichia coli* has been reported to be stabilized by a disulphide bond (Brinkmann *et. al.*, 1993).

In order to achieve successful expression, the gene encoding the recombinant molecule must be placed in the context of appropriate control sequences that allow transcription and translation of the protein. Inducible promoters are normally used to control expression of the protein. This is important to prevent loss or mutation of the gene should expression of the recombinant protein prove to be toxic to the bacteria. There are a few well-known promoters, but for this research, the T7 RNA promoter has been used to obtain tightly controlled, high level expression (Tabor and Richardson, 1985, Studier and Moffatt, 1986). A second important factor for efficient production in *E. coli* is the presence of a prokaryotic ribosome-binding site (Gold *et. al.*, 1981). The last control element is the transcription terminator, which prevents transcription beyond the desired gene and adds stability to the mRNA (Verma, Boleti and George, 1998). Taking to account all considerations the pET system was specifically developed for the cloning and expression of recombinant proteins in *E. coli.*

5.3 pET system

W. F. Studier and B. A. Moffatt developed the pET expression system in 1986 for tightly regulated and high expression of recombinant protein in *E.coli,* which used the highly specific bacteriophage T7 RNA polymerase (Studier and Moffatt, 1986) (Fig 35). The in-house expression vector, pK7, used in this system is a vector derived from the pET26b vector Syntaxin Ltd modified pET26b to remove the pel B leader sequence, to introduce a multiple cloning site to include specific restriction endonuclease sites and to convert the vector mobilisation status from deficient to defective.

For protein production, the recombinant plasmid (pK7) is transferred into an *E. coli* strain containing a chromosomal copy of the gene for T7 RNA polymerase. The host used is BL21 (DE3) bacteria. The (DE3) designation indicates that the host is a lysogen of λ bacteriophage DE3, which carries a chromosomal copy of T7 RNA polymerase gene under the control of the *lacUV5* promoter and DNA fragment containing the *lacI* gene (Studier et. al., 1986). Once a DE3 lysogen is produced, the only promoter which will direct transcription of the T7 RNA polymerase gene is the *lacUV5* promoter, which is induced by addition of isopropyl-B-galactosidase (IPTG) to the culture medium (Weisser and Hall, 2009) (Fig 35).

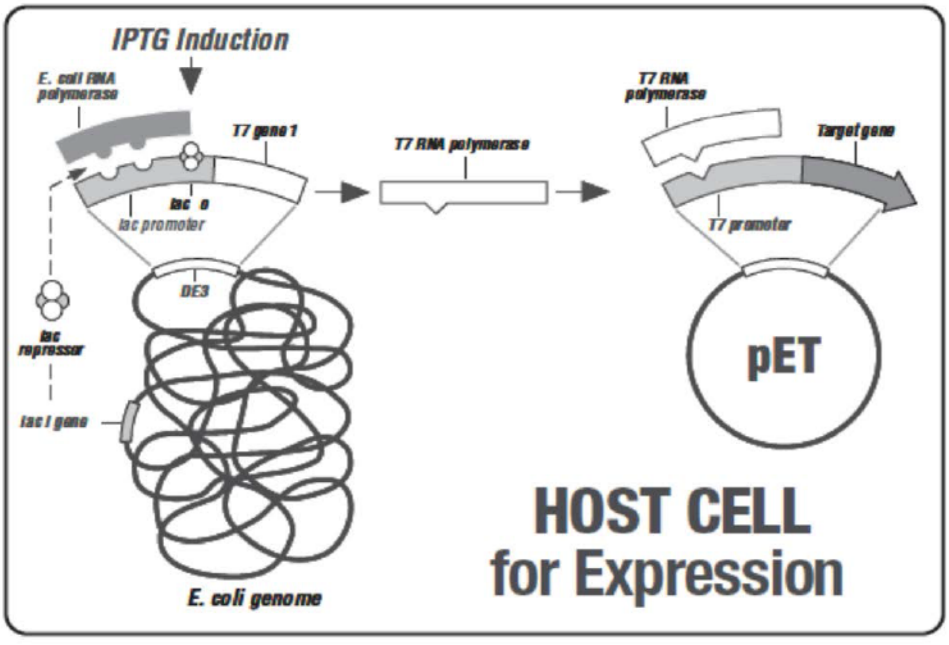


Figure 48 Protein production in pET expression system.
(Figure adapted from Novagen pET System Manual)

5.4 Results

The successful expressions of novel molecules in a bacterial expression system have been advantage to this study. Expressed scFv is usually insoluble; however, one method that can improve its solubility is fusion with another protein (Zengh *et al*., 2003). In addition, the purification of His-tagged scFv with an increased concentration of imidazole also has shown improved solubility of these molecules (Hamilton *et al*., 2003). These considerations could potentially explain the lack of ineffective production of scFv fused to LH_N backbone and purified on an affinity column. The successful soluble expression of the single domain antibodies was more likely expected, as they a high stability and solubility (Harmsen and Haard, 2007).

E.coli BL21 (DE3) cells transformed with the expression plasmids were grown in the shaking culture (220 rpm, 37°C) in LB (Luria Bertani) broth supplemented with 30 µg/ml kanamycin and 20% glucosamine to an OD of 0.5-0.6 at 600 nm. Protein expression was induced by the addition of 1 mM isopropyl-B-D-thiogalactopyranoside (IPTG, Sigma) for 20 h, at 16°C, 220 rpm. The harvested cell pellets were re-suspended in 50 mM Hepes pH 7.2, 200 mM NaCl and stored at – 80°C until required. Another method used for the purification of these molecules differed in the harvest of the cell pellets. The harvested cells were re-suspended in the 50 mM Hepes pH 7.2, 500 mM NaCl buffer and centrifuged at 12500 rpm for 30 min. The supernatant was removed and the cell pellets were kept frozen at -80˚C. This additional re-suspension of the cell pellets in the high salt concentration buffer should help to remove *E. coli* contaminants and unwanted debris in order to obtain high purity of the expressed protein. The expressions of soluble LH_N single domain antibodies (Figure 49, 50 and 51) and LH_N single chain antibody fragment (Figure 52 and 53) have been accomplished.

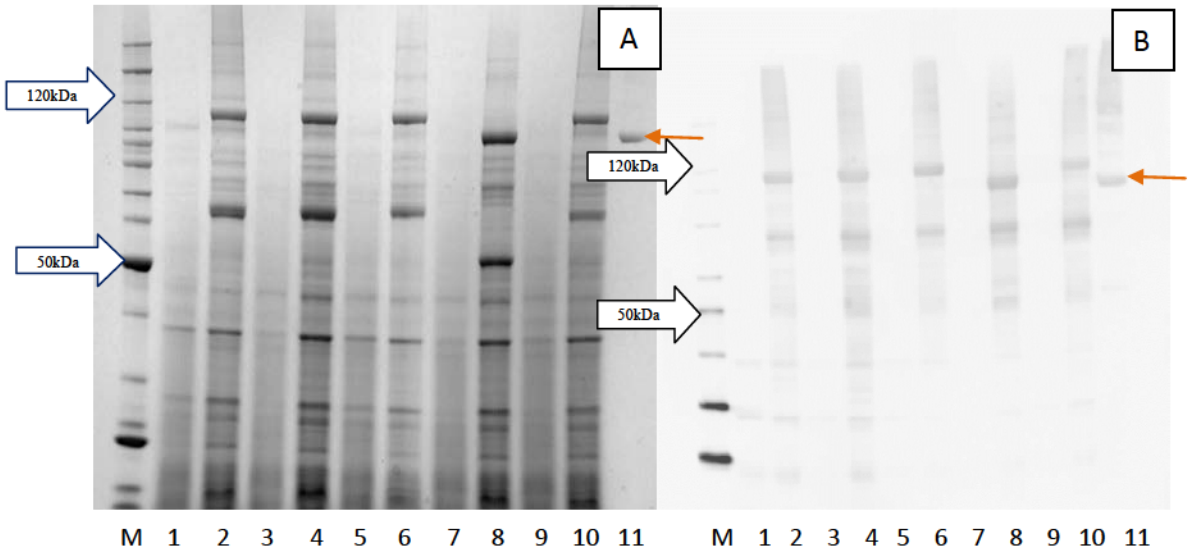


Figure 49 SafeStain NuPage 4-12% Bis-Tris gel (A) and Western blot against light chain (B) showing expression of LH_N/A with a single domain antibody.

Lane M on the SDS PAGE gel (A) is a Benchmark ladder and on Western blots (B) is a Magic Mark; sizes in kDa (Fig 16 in chapter 2 Material and Methods);
Lanes 1 **(LN_N/A-EGFR7sdAb=SXN101869)**, 3 **(LN_N/A-EGFR10sdAb=SXN101870)**,
5 **(LN_N/A EGFR28sdAb=SXN101871)**, 7 **(LN_N/A-EGFR31sdAb=SXN101872)** and
9 **(LN_N/A-EGFR43sdAb=SXN101873)** show the *E. coli* extract containing the expression plasmid prior to induction;
Lanes 2 **(LN_N/A-EGFR7sdAb=SXN101869)**, 4 **(LN_N/A-EGFR10sdAb=SXN101870)**, 6 **(LN_N/A-EGFR28sdAb=SXN101871)**, 8 **(LN_N/A-EGFR31sdAb=SXN101872)** and 10 **(LN_N/A-EGFR43sdAb=SXN101873)** show the E. coli extract containing the expression plasmid after induction;
Lane 11 shows the control **LH_N/A=SXN100590** indicated by orange arrow (Table11);

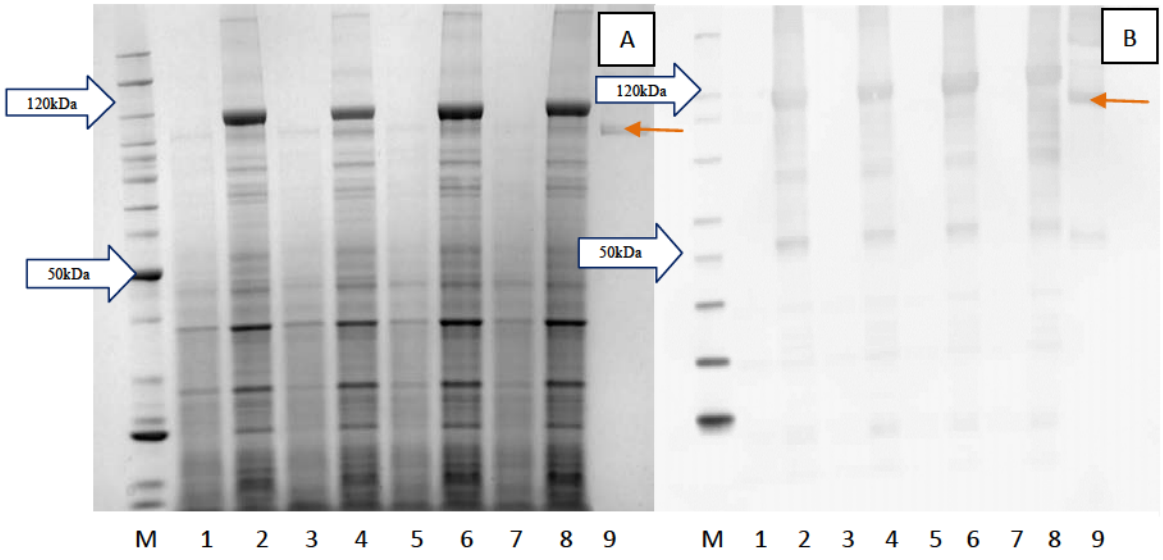


Figure 50 SafeStain NuPage 4-12% Bis-Tris gel (A) and Western blot against light chain (B) showing expression of LH_N/B with a single domain antibody.

Lane M on the SDS PAGE gel (A) is a Benchmark ladder and on Western blots (B) is a Magic Mark; sizes in kDa (Fig 16 in chapter 2 Material and Methods);
Lanes 1 (**LH_N/B-EGFR6sdAb=SXN101861**), 3 (**LH_N/B-EGFR7sdAb=SXN101862**), 5 (**LH_N/B-EGFR10sdAb=SXN101863**) and 7(**LH_N/B-EGFR28sdAb=SXN101864**) show the *E. coli* extract containing the expression plasmid prior to induction;
Lanes 2(**LH_N/B-EGFR6sdAb=SXN101861**), 4(**LH_N/B-EGFR7sdAb=SXN101862**), 6(**LH_N/B-EGFR10sdAb=SXN101863**) and 8(**LH_N/B-EGFR28sdAb=SXN101864**) show the E. coli extract containing the expression plasmid after induction;
Lane 9 shows the control **LH_N/B=SXN101687** indicated by orange arrow (Table11);

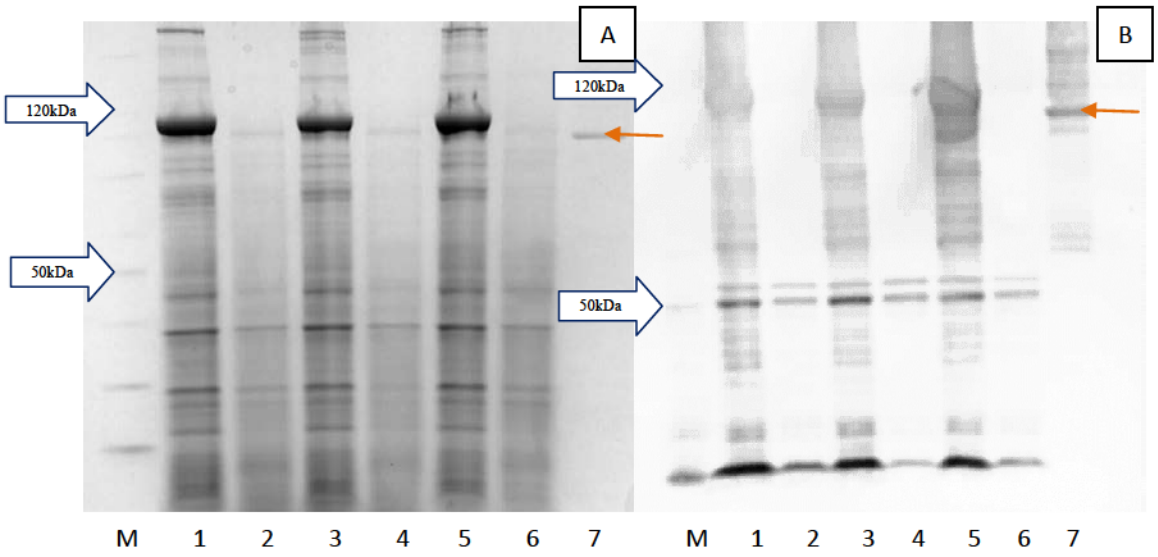


Figure 51 SafeStain NuPage 4-12% Bis-Tris gel (A) and Western blot against light chain (B) showing expression of LH_N/D with a single domain antibody.

Lane M on the SDS PAGE gel (A) and on Western blots (B) is a Magic Mark; sizes in kDa (Fig 16 in chapter 2 Material and Methods);
Lanes 2 (**LH_N/D-EGFR28sdAb=SXN101877**), 4 (**LH_N/D-EGFR31sdAb=SXN101878**) and 6 (**LH_N/D-EGFR43sdAb=SXN101879**) show the *E. coli* extract containing the expression plasmid prior to induction;
Lanes 1(**LH_N/D-EGFR28sdAb=SXN101877**), 3 (**LH_N/D-EGFR31sdAb=SXN101878**) and 5 (**LH_N/D-EGFR43sdAb=SXN101879**) show the E. coli extract containing the expression plasmid after induction;
Lane 11 shows the control **LH_N/D=SXN101655** indicated by orange arrow (Table11);

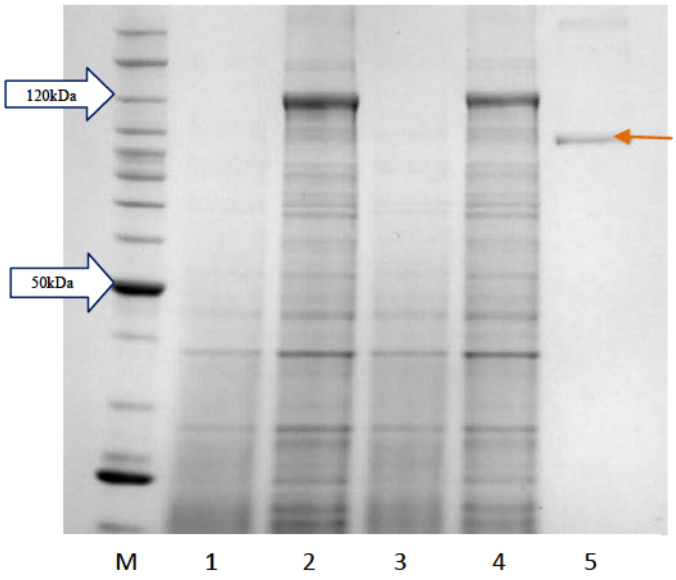


Figure 52 SafeStain NuPage 4-12% Bis-Tris gel showing expression of LH_N/C with a single chain antibody fragment.

Lane M on the SDS PAGE gel is a Benchmark ladder; sizes in kDa (Fig 16 in chapter 2 Material and Methods);
Lanes 1 (**LH_N/C-scFvCD117I=SXN101882)** and lane 3 (**LH_N/C-scFvCD117I=SXN101882)** show the *E. coli* extract containing the expression plasmid prior to induction;
Lanes 2 (**LH_N/C-scFvCD117I=SXN101882)**, 3 (**LH_N/C-scFvCD117I=SXN101882)** show the E. coli extract containing the expression plasmid after induction;
Lane 11 shows the control **LH_N/C=SXN101654** indicated by orange arrow (Table11);

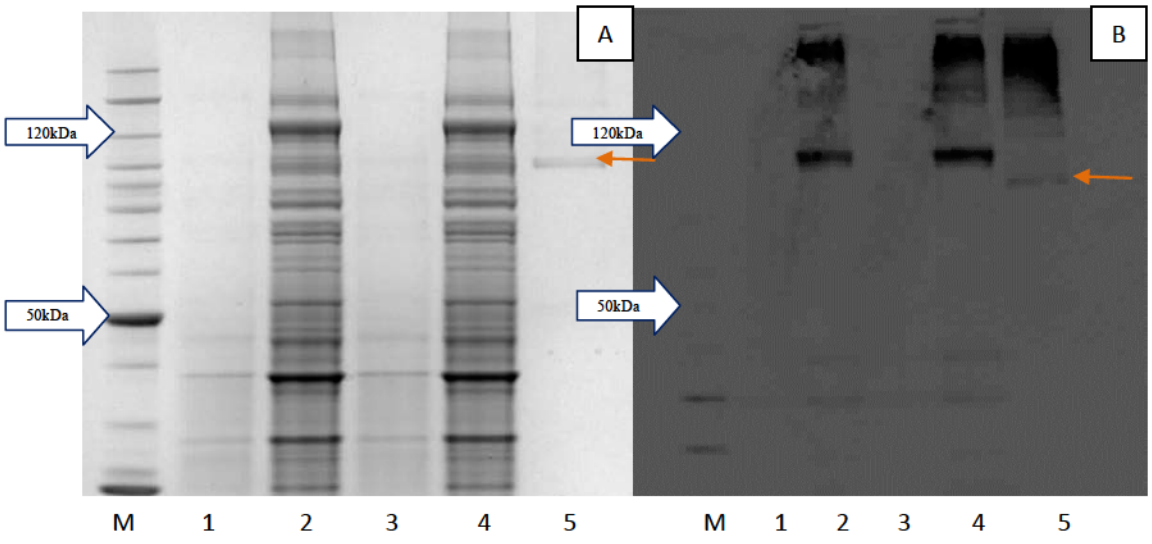


Figure 53 SafeStain NuPage 4-12% Bis-Tris gel (A) and Western blot against light chain (B) showing expression of LH_N/B with a single chain antibody fragment.

Lane M on the SDS PAGE gel is a Benchmark ladder; sizes in kDa (Fig 16 in chapter 2 Material and Methods);
Lanes 1 (**LH_N/B-scFvCD117I=SXN101776)** and lane 3 (**LH_N/B-scFvCD117I=SXN101776)** show the *E. coli* extract containing the expression plasmid prior to induction;
Lanes 2 **(LH_N/B-scFvCD117I=SXN101776)**, 3 (**LH_N/B-scFvCD117I=SXN101776)** show the E. coli extract containing the expression plasmid after induction;
Lane 11 shows the control **LH_N/B=SXN101687** indicated by orange arrow (Table11);

Chapter 6

Purification of recombinant antibody based constructs

Chapter 6. Purification of recombinant antibody based constructs

The purification was one of the most fundamental elements of this research. The ability to purify soluble, full-length products gives opportunities to perform cell and pharmacology tests that will provide data regarding to functionality of these molecules. The stability of purified single domain antibodies has been confirmed during the concentration to over 10 mg/ml. The molecules show neither signs of aggregation nor change of colour.

6.1 Overview of protein purification

Proteins play a crucial role in many biochemical processes, such as binding, catalysis, transport. These various possibilities derive from chemical diversity, the flexibility of the polypeptide chain and folding ability of polypeptide chains into many three-dimensional structures. Only proteins that retain functional are useful components to study biochemical interactions, what can subsequently lead to discovery of new therapeutic. In order to analyze protein it preferably needs to be purified, optimally to high standards. The quality of purified protein depends on the purification process, which often need to be developed specifically for the target protein. In order to develop accurate method it is important to establish the purpose of the protein of interest. The study based on activity of protein will require less pure compound in comparison to the protein used for study of the structure of molecule. However, the presence of active contaminants, which interfere with the activity of enzyme, cannot be acceptable by any circumstances and need to be removed. In case of proteins applied for therapeutic use, all associated contaminates need to be separated from target proteins. In this instance, it is necessary to introduce additional steps to remove contaminants, such as endotoxin, aggregates, and DNA. The quality of final product is always considered in the first place, but there are also different aspects, which need to be taken to account, while developing purification process. The quantity, cost and time have usually a crucial impact on the establishment of the final purification process. The protein source is the significant part of the development of purification procedure, as these impacts on the yield and contaminant profile. Currently, growing cells in

culture, such as *Escherichia coli*, *Saccharomyces cerevisiae* or mammalian cells are the most common approach for protein production. The more detailed description of these cells can be found in the expression of novel antibody-botulinum chimera's chapter. The application of suitable expression system for molecules will subsequently have influence on the purification method. Though recombinant systems can enhance the yield of target protein, host cell can bring some disadvantages to expressed proteins, such as presence of contaminants, lack of a protein glycosylation, insolubility or unfolding of protein. These factors can bring some limitations to purified protein and in order to avoid them it is essential to possess knowledge of chemical and physical properties of the protein to utilize robust purification techniques. The purification methods of botulinum toxins and their fragments have been improved together with reengineering this molecule during the years.

6.2 Introduction to the purification of BoNT

All natural botulinum neurotoxins are synthesised as complexes with non-toxic proteins, which under conditions of high ionic strength and high pH dissociate (Sugiyama, 1980). Combining these procedures with anion exchange chromatography and gel filtration allowed the first isolation of pure type A neurotoxin from its complex (Dasgupta *et al.*, 1966; Dasgupta and Boroff 1967, 1968). Afterwards, the purification process of neurotoxins was improved by introducing ammonium sulphate precipitations to enrich the toxin from culture extracts (Dasgupta *et al.*, 1970; Tse *et al.*, 1982). This step has been added to the purification of *Clostridia* neurotoxins and it is a routine enrichment step for many proteins.

Further research brought a nickel column to capture His-tagged BoNT's into the development of purification process (Moberg and Sugiyama 1978). This modification together with introduction of DEAE-Sephacel anion-exchange step generated the procedure used routinely for the large-scale purification of type A or B neurotoxin (Tse *et al.*, 1982; Williams *et al.*, 1983; Evans *et al.*, 1986). The procedure established for purification of C_1, D (Shone, 1987) and E was completed at the same time as the rapid method for purification of type F neurotoxin was in the process of development (Wadsworth *et al.*, 1990).

First attempts to isolate neurotoxin polypeptides (HC and LC) was challenging, as the neurotoxin shows a remarkable decrease in solubility upon reduction of its intra- and inter- chain disulfide bonds. In this case, the purification of BoNT type B subunits involved maintaining solubility by reducing the disulfide bridges in the presence of 1-2 M urea prior to isolation by gel filtration (Kozaki and Sakaguchi 1975; Kozaki *et al.*, 1977). On another hand, the extremely low solubility of the LC of BoNT type A was used to purify pure HC in the presence of 0.5 M NaCl and 1% (w/v) nicotinamide adenine dinucleotide (NAD) (Krysinski and Sugiyama, 1980). Later on, the purification of both chains from BoNT type A were obtained by adsorbing the toxin on the column of QAE-Sephadex followed by thiol reduction with dithiothreitol (DTT) in the presence of 2 M urea and 1 M NaCl (Kozaki *et al.*, 1981). The modification of this method during subsequent years developed procedure for isolation of the chains from BoNT types A and B (Maisey et al., 1988). The digestion of the neurotoxin type A with trypsin has led to isolation of the H_2L fragment. These experiments were a good starting point for the purification of the amino terminal moiety of the heavy chain H_2 (Shone et al., 1985, 1987). The techniques for the toxin purification have become more robust during years, but it become easier while producing recombinant non-toxic version of this molecule. Various techniques have been used to purify recombinant fragments of BoNT's.

The LH_NA-ECL (*Erythrina cristagalli)* non-toxic molecule mentioned in the chapter 4 ‚Cloning' was successfully purified by others in two step process, involving size exclusion chromatography (SEC) to remove unconjugated ECL followed by affinity chromatography to remove unconjugated LH_N/A. The final product detected on the SDS-PAGE gel had a molecular mass at approximately 220 kDa (Chaddock et al., 2004). The demonstration of catalytic activity of LH_N/A was confirmed to be maintained by using an *in vitro* synaptosomal-associated protein cleavage assay (Hallis, B., James, B. A. F., Shone, C. C., 1996). The purified LH_N/A –ECL product was able to inhibit the release of the neurotransmitter glutamate and the neuropeptide substance P from multiple vesicle population (Duggan et al., 2002). The more highly developed non-toxic version of LH_N molecule EGF-LH_N/C was purified using a nickel chelating Sepharose affinity column where fusion protein was eluted with 100 mM Imidazole. The purified recombinant EGF-LH_N/C has shown activity of a functional endopeptidase and access

to the NCI-H292 cells (Foster *et al.,* 2005). The introduction of nickel column to aid purification of His tagged EGF-LH_N/C molecule opened a new way of producing soluble fusion protein containing an active LH_N fragment. During recent years, the purification method has been slightly modified and improved, but affinity column continues as a first step capture for purification of non-toxic tagged LH_N fragments. The purification of novel antibody-botulinum toxin molecules was prepared according to standard two-step affinity chromatography, as well as using methods developed during this study to achieve high purity of protein.

6.3 Purification of recombinant antibody based constructs

Purification of recombinant antibody based constructs was performed using different columns and techniques. The standard protein purification techniques can be applied to various proteins. However, it is not always the case that proteins belonging to the same class and exhibiting close structural and functional relationship can be suitably purified by the same methodology. A combination of methods has been developed to purified antibodies, such as ion-exchange chromatography or affinity chromatography. As the purification process of recombinant botulinum neurotoxin is a very well established it helped to develop the upstream process for the recombinant antibody-toxin chimeras. It is important to distinguish purification of the single chain antibodies from the single domain antibodies, as these structures are differ one from another. As mentioned before the single chain antibodies (scFvs) has a linker between light and heavy chain, which potentially can have an effect on the isolation of the final product. The single domain antibodies (Abs) are simpler in structure; this can make purification process more successful.

The development of techniques and the selection of an appropriate purification process for the production of recombinant antibody-botulinum toxin molecules were mainly based on the requirements in relation to functionality and purity of the final product. The molecules designed for this study carry polyhistidine affinity tags, which consist of ten consecutive Histidine residues on the N-terminus. In this case, the most common method used for purification of this kind of molecules is immobilized metal affinity chromatography (IMAC), which allows differentiation of recombinant proteins

expressed in bacterial system from the background of total cellular proteins. Immobilized metals such as Co^{2+}, Cu^{2+}, Ni^{2+},Zn^{2+}, Ca^{2+} and Fe^{3+} can be used to purify polyhistidine fusion proteins.

6.4 Results

In the capture step, chelating Sepharose (Amersham Bioscience) charged with immobilized Ni^{2+} ions was used for the purifications of novel chimeras. The cell paste was homogenized using a Constant Systems homogenizer and clarified by centrifugation. The clarified lysate was loaded onto the affinity column. The bound proteins were eluted from the column using an imidazole step-wise gradient (from 40 mM to 250 mM). Fractions were dialyzed against a 50 mM Hepes pH 7.2, 150 mM NaCl buffer or a 50 mM Hepes pH 7.2, 50 mM NaCl. The samples were analyzed by SDS-PAGE on 4-12% acrylamide gels. The choice of dialyzing buffer depends on the type of enzyme used for activation. All molecules require an activation process with a proteolytic enzyme, such as Enterokinase (EK BioLabs) or FXa (Factor Xa BioLabs) after the first affinity step. The enzyme selection is determined by enzyme specific recognition sites within the primary structure of the molecule. The choice of second column was dictated by the purity of the cleaved molecule. Affinity chromatography or hydrophobic interaction chromatography (HIC) was used in the final polishing step of this purification process. In every case, proteins were purified using a two-column method: two sequential nickel columns or an initial nickel column follow by HIC. The purified protein was filtered and stored in aliquots at -20°C or -80°C. On occasions, the purification process between single chain variable antibody fragments-LH_N and single domain antibody-LH_N chimeras was significantly diverse due to different observations recorded during the purification process, these differences are described in detail in each section below.

6.4.1 Purification of single chain variable antibody fragments-LH$_N$ chimera

Purification of **LH$_N$/B-scFvCD117I=SXN101624** (Table 10) was attempted a number of times in order to improve activation and minimize truncations of full-length product. The products obtained following purification with nickel column through a step elution are shown in figure 54.

When subjected to SDS-PAGE under non-reducing conditions the pure proteins migrated as single bands of molecular weight 129741.2 Da (MW). However, non-specific cleavage was observed, seen as two bands with 117516.7 Da (MW) and 103578.3 Da (MW) on SDS-PAGE (Fig 54).

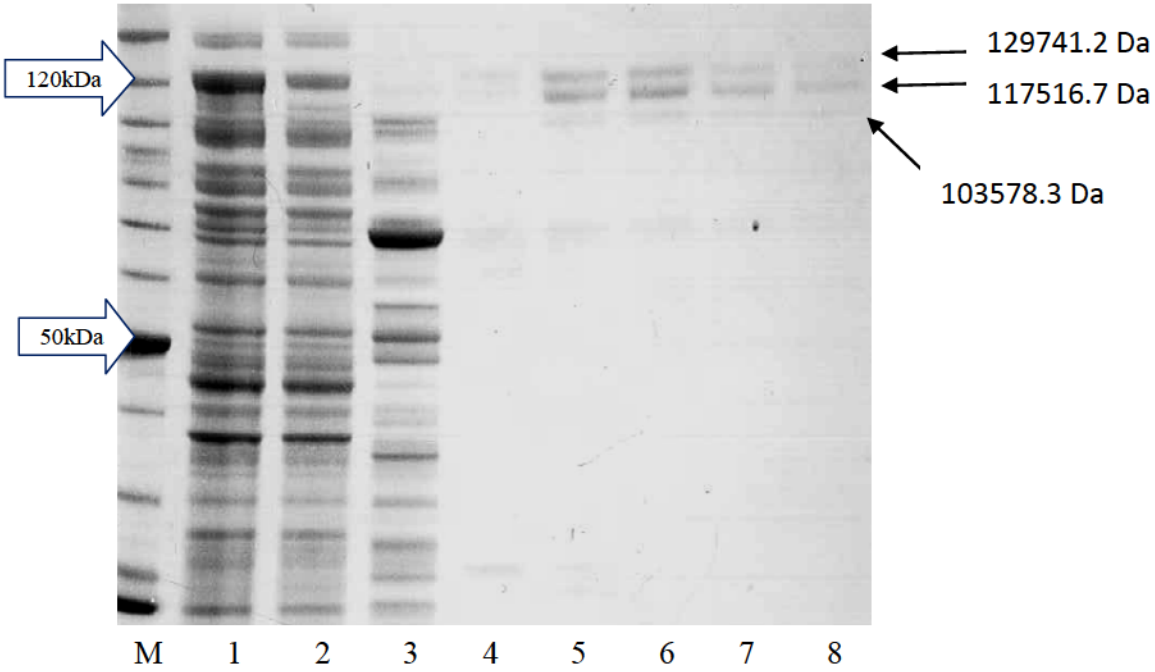


Figure 54 NuPage 4-12% Bis-Tris gel of LH$_N$/B-scFvCD117I=SXN101624 produced from a nickel column.

Lane M is a Benchmark ladder, sizes in kDa (Fig 16 in chapter 2 Material and Methods);
Lane 1 is a clarified lysate;
Lane 2 is a flow through after His column;
Lane 3 is a 40mM Imidazole wash; Lane 4 is a 100mM Imidazole wash; Lanes 5, 6, 7, 8 are a 250mM Imidazole washes;
The arrows indicate a truncated products of LH$_N$/B-scFvCD117I=SXN101624;

According to the interpretation of bands on the SDS-PAGE (Fig 54), the truncation within a full-length structure took place on the affinity column, as it is not visible in the lane 1 total lysate or lane 2 flows through. Truncated bands as indicated in the lane 8 started to appear during elution of the fusion protein by Imidazole. Also it was confirmed by Western Blot against the light chain of B (Fig 56). The sizes of truncated products were estimated by ProtoPram programme. The band at around 117516.7 Da (MW) probably represents the lack of scFv light chain and the band around 103578.3 Da (MW) most likely represents LHBK191A-ScFv that lacks ScFv ligand. It was noted that the truncation of **LH_N/B-scFvCD117I=SXN101624** was less prominent around the lower band of 103578.3 Da (MW).

Subsequently, the molecule was cleaved by EK at 25°C. The cleavage is detectable under reducing conditions (+DTT) on the SDS-PAGE gel (Fig 55, lane 6). The intensity of bands in lanes 2 and 4 are higher in comparison to lanes 1, 3 and 5 (indicated by purple arrow) due to the presence of DTT that has changed the accessibility of the Safe Stain to protein resulting in an increase in signal strength of target protein **LH_N/B-scFvCD117I=SXN101624**. Furthermore bands representing two distinguish domains, one H_N-CD117I at 74744.9 Da and the other 10HT-LC at 55151.3 Da have a noticeable higher intensity to the band of full-length product of **LH_N/B-scFvCD117I=SXN101624** at 129741.2 Da in lane 5. These observations could suggest that there is a presence of truncated product cleaved by EK or there are forms of aggregates that correspond to the same full-length compound. The control samples without addition of Enterokinase were unstable in both temperatures, as under reducing conditions (Fig 55, lanes 2 and 4) presence of band at 74744.9 Da corresponds to the molecular weight of the H_N-CD117I product visible in lane 6.

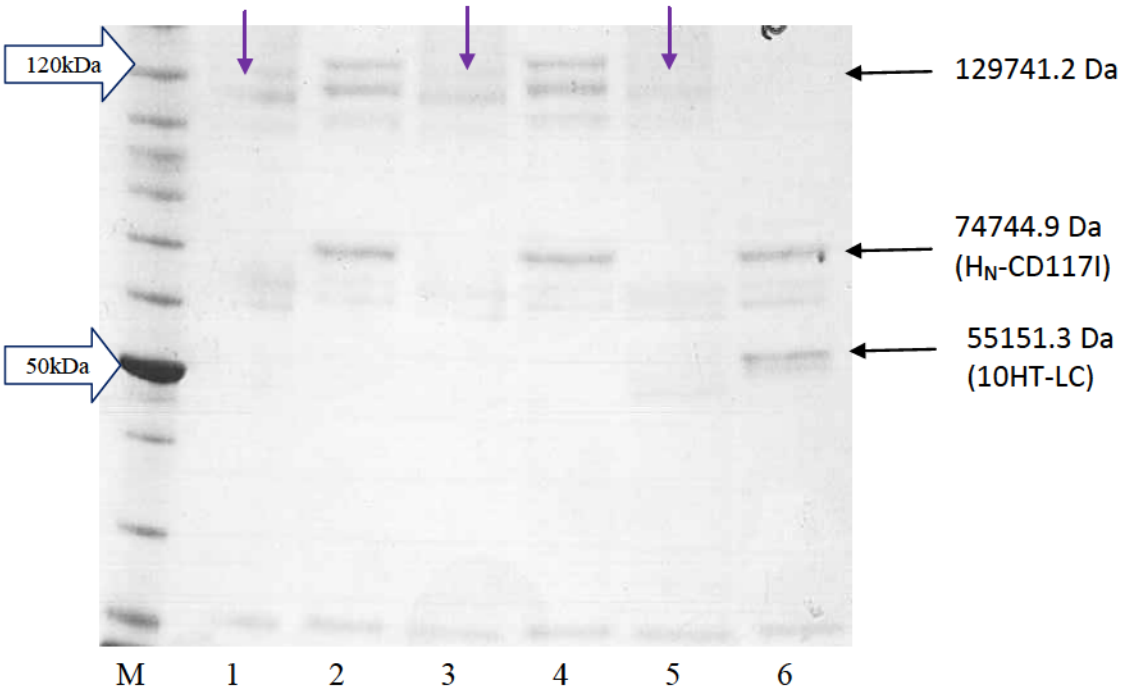


Figure 55 NuPage 4-12% Bis-Tris gel of LH_N/B-scFvCD117I=SXN101624 produced from a cleavage process.

Lane M is a Benchmark ladder, sizes in kDa (Fig 16 in chapter 2 Material and Methods);
Lane 1 is a control sample **LH_N/B-scFvCD117I=SXN101624** incubated at -20˚C;
Lane 2 is a control sample **LH_N/B-scFvCD117I=SXN101624** incubated at -20˚C under reducing condition with addition of DTT;
Lane 3 is a control sample **LH_N/B-scFvCD117I=SXN101624** incubated at +25˚C;
Lane 4 is a control sample **LH_N/B-scFvCD117I=SXN101624** incubated at +25˚C under reducing condition with addition of DTT;
Lane 5 is a cleaved **LH_N/B-scFvCD117I=SXN101624** incubated at +25˚C;
Lane 6 is a cleaved **LH_N/B-scFvCD117I=SXN101624** incubated at +25˚C under reducing condition with addition of DTT;

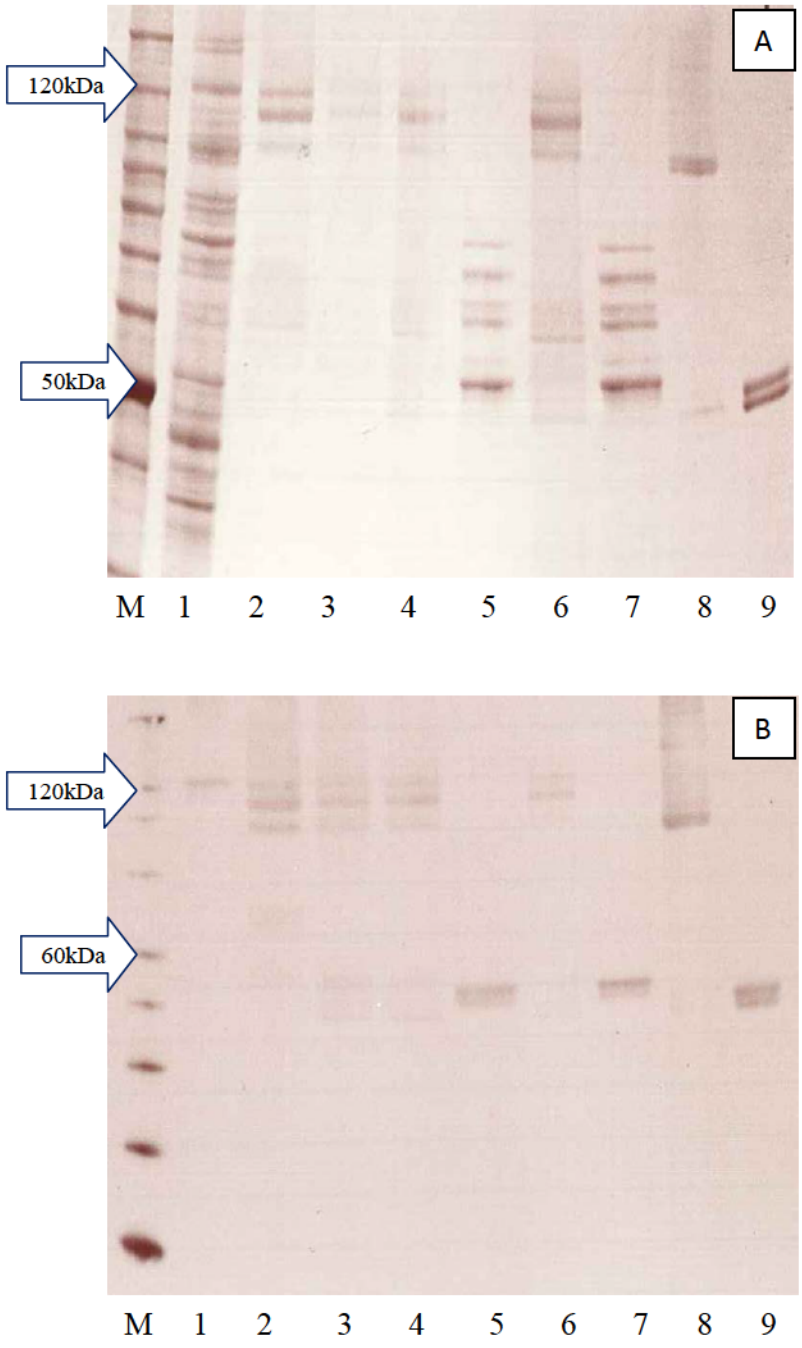


Figure 56 Quality control of the target protein LH_N/B-scFvCD117I=SXN101624 represent on the NuPage 4-12% Bis-Tris gel (A) and confirmed by Western blot against light chain (B).

Lane M on the gel (A) is a Benchmark ladder and on Western blot (B) is a Magic Mark; sizes in kDa (Fig 16 in chapter 2 Material and Methods);
Lane 1 shows clarified lysate of **LH_N/B-scFvCD117I=SXN101624**;
Lane 2 is a -20˚C control;
Lane 3 is a +25˚C control;
Lane 4 is a final target fusion protein **LH_N/B-scFvCD117I=SXN101624** at 0.1 mg/ml loaded at 5µl;
Lane 5 is a final target fusion protein **LH_N/B-scFvCD117I=SXN101624** at 0.1 mg/ml loaded at 5µl (+DTT);
Lane 6 is a final target fusion protein **LH_N/B-scFvCD117I=SXN101624** at 0.1 mg/ml loaded at 10µl;
Lane 7 is a final target fusion protein **LH_N/B-scFvCD117I=SXN101624** at 0.1 mg/ml loaded at 10µl (+DTT);
Lane 8 is an unliganded control **LH_N/B=SXN101687** (Table 11) at 0.1mg/ml loaded at 10µl;
Lane 9 is an unliganded control **LH_N/B=SXN101687** (Table 11) at 0.1mg/ml loaded at 10µl (+DTT);

The fragment of ScFvCD117I molecule was cloned into the LHB backbone activated by FXa (**LH_N/B-scFvCD117I=SXN101776)** and purified using different purification techniques to check if the cleavage process can be improved and truncations could be reduced. The cleavage step was performed at the following temperatures: 25°C and 4°C. There are not significant dissimilarities in the outcome of cleavage between these temperatures. The elimination of non-specific cleavage was not successful in this instance; therefore, purification was repeated by alternative methods.

The **LH_N/B-scFvCD117I=SXN101776** was expressed in *E. coli* bacterial system and captured on the affinity column charged with Ni^{2+} (Fig 56).

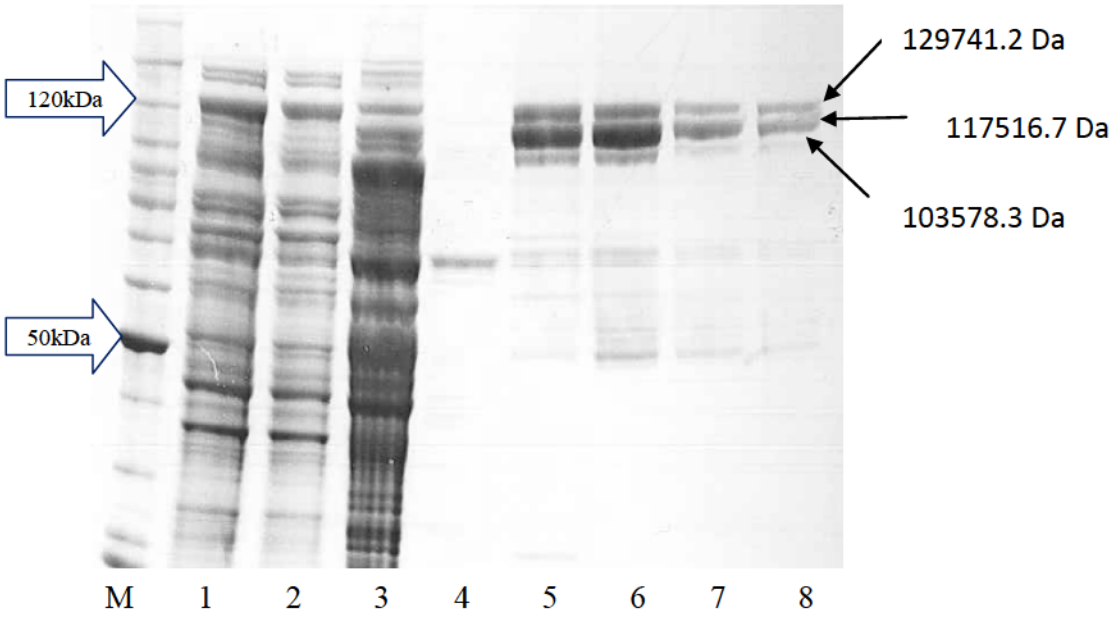


Figure 57 NuPAGE 4-12% Bis-Tris gel of LH_N/B-scFvCD117I=SXN101776 produced from a nickel column.

Lane M is a Benchmark ladder, sizes in kDa (Fig 16 in chapter 2 Material and Methods);
Lane 1 is a clarified lysate;
Lane 2 is a flow through from His column;
Lane 3 is a 40mM Imidazole wash; Lane 4 is a 100mM Imidazole wash; Lanes 5, 6, 7, 8 are a 250mM Imidazole washes;
The arrows indicate a truncated products of **LH_N/B-scFvCD117I=SXN101776**;

The fusion protein **LH_N/B-scFvCD117I=SXN101776** was cleaved by FXa at +25˚C (Fig 57).

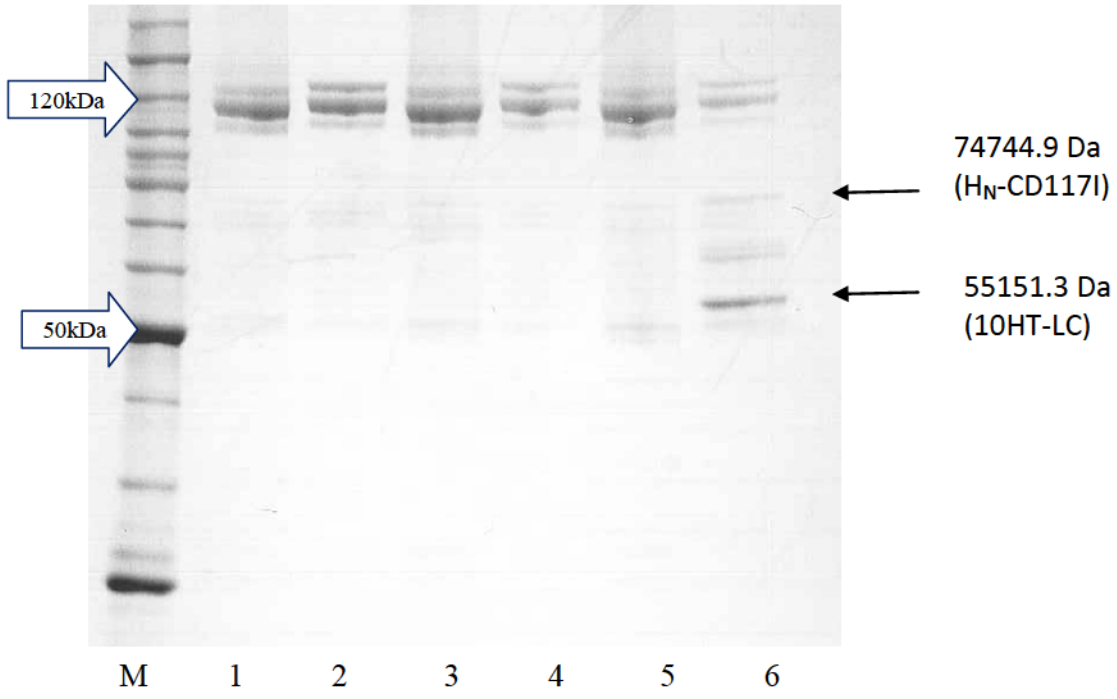
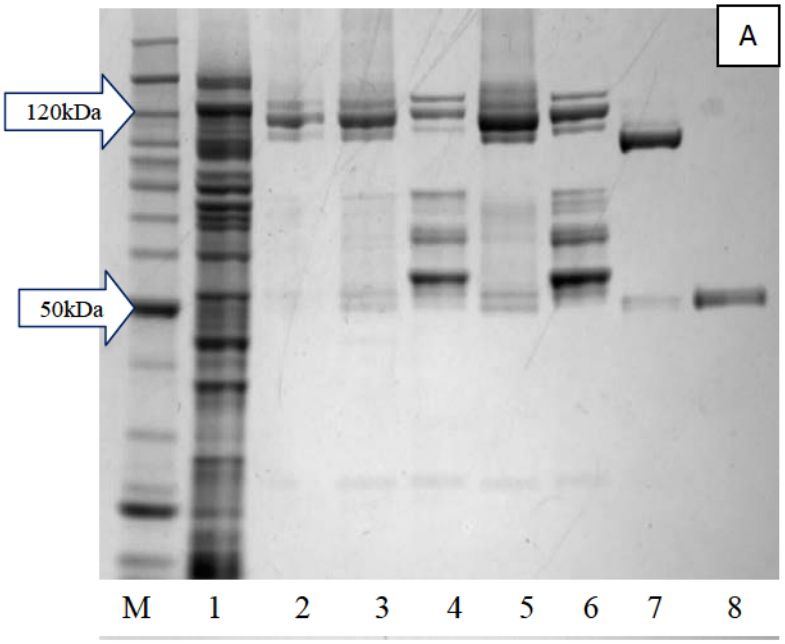
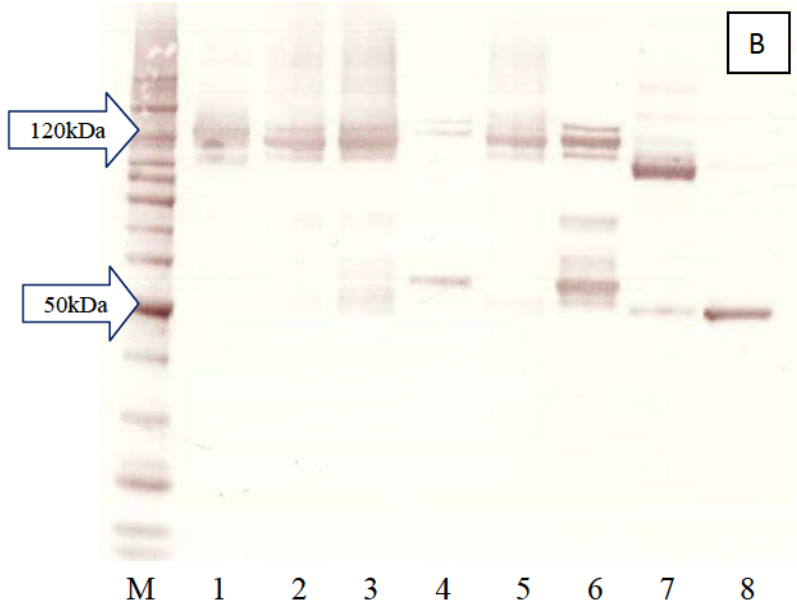


Figure 58 NuPAGE 4-12% Bis-Tris gel of LH_N/B-scFvCD117I=SXN101776 cleaved by FXa.

Lane M is a Benchmark ladder, sizes in kDa (Fig 16 in chapter 2 Material and Methods);
Lane 1 is a control sample LH_N/B-scFvCD117I=SXN101776 at -20˚C without addition of FXa;
Lane 2 is a control sample LH_N/B-scFvCD117I=SXN101776 at -20˚C without addition of FXa (+DTT);
Lane 3 is a control sample LH_N/B-scFvCD117I=SXN101776 at +25˚C without addition of FXa;
Lane 4 is a control sample LH_N/B-scFvCD117I=SXN101776 at +25˚C without addition of FXa (+DTT);
Lane 5 is cleaved LH_N/B-scFvCD117I=SXN101776 by FXa at +25˚C;
Lane 6 is cleaved LH_N/B-scFvCD117I=SXN101776 by FXa at +25˚C (+DTT);

The outcome of this purification is shown on SDS-PAGE story gel (A), where the presence of the full-length **LH_N/B-scFvCD117I=SXN101776** target protein was detected by Western blots against His tag (B) and light chain of B (C) (Fig 58).





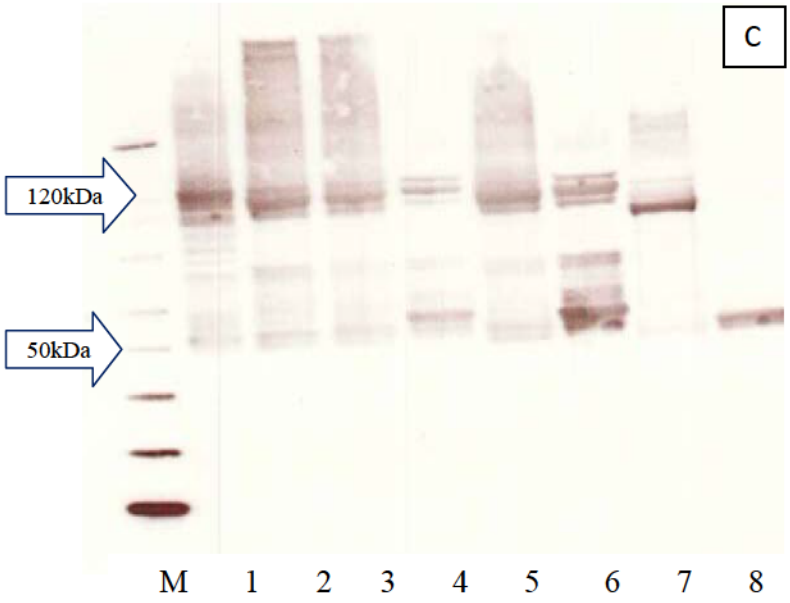


Figure 59 Quality control of the target protein LH_N/B-scFvCD117I=SXN101776 represent on the NuPage 4-12% Bis-Tris gel (A) and confirmed by Western blot against His tag (B) and light chain of B (C).

Lane M on the gel (A) is a Benchmark ladder and on Western blots (B and C) is a Magic Mark; sizes in kDa (Fig 16 in chapter 2 Material and Methods);
Lane 1 shows clarified lysate of **LH_N/B-scFvCD117I=SXN101776**;
Lane 2 is a -20˚C control;
Lane 3 is a final target fusion protein **LH_N/B-scFvCD117I=SXN101776** at 0.1 mg/ml loaded at 5µl;
Lane 4 is a final target fusion protein **LH_N/B-scFvCD117I=SXN101776** at 0.1 mg/ml loaded at 5µl (+DTT);
Lane 6 is a final target fusion protein **LH_N/B-scFvCD117I=SXN101776** at 0.1 mg/ml loaded at 10µl;
Lane 7 is a final target fusion protein **LH_N/B-scFvCD117I=SXN101776** at 0.1 mg/ml loaded at 10µl (+DTT);
Lane 8 is an unliganded control **LH_N/B=SXN101687** (Table 11) at 0.1mg/ml loaded at 10µl;
Lane 9 is an unliganded control **LH_N/B=SXN101687** (Table 11) **at** 0.1mg/ml loaded at 10µl (+DTT);

In a repeated experiment using **LH_N/B-scFvCD117I=SXN101776** molecule, the alternative FXa enzyme purchased from R&D System was tested during the cleavage process (B) (Fig 59). The test was set up at + 25˚C, overnight for Recombinant Human Coagulation Factor X (R&D System) (B) and Factor X (BioLabs) (A). Additional sample was prepared with 4 mM $CaCl_2$ to check if it could enhance the activity of the FXa enzyme. The FXa (R&D System) did not cleave protein even in the presence of 4 mM $CaCl_2$ (B); therefore, FXa (BioLabs) was used to cleave fusion protein (A) (Fig 59).

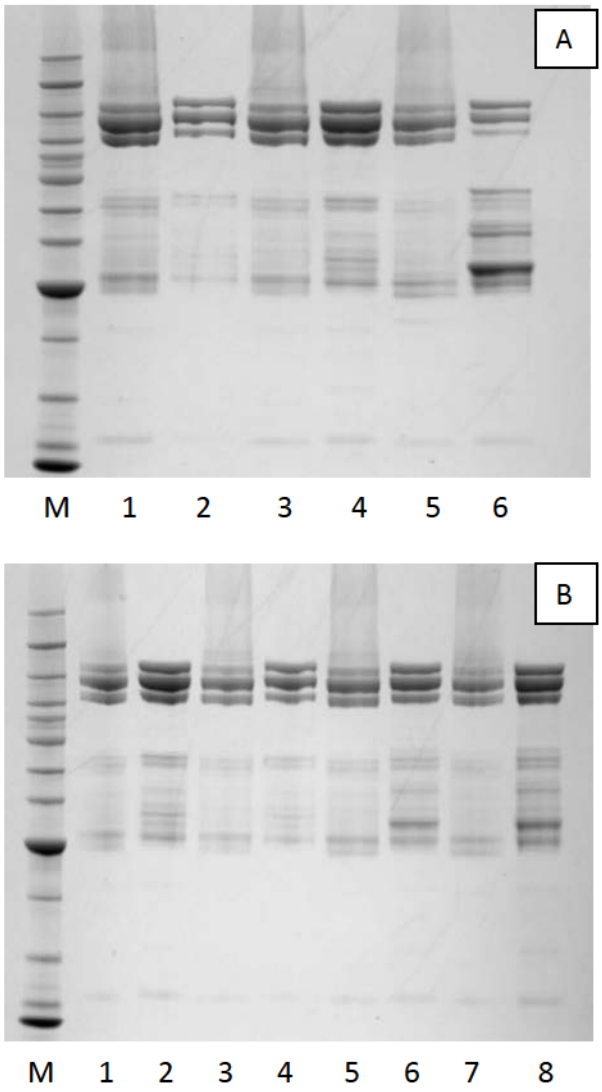


Figure 60 NuPAGE 4-12% Bis-Tris gel of LH_N/B-scFvCD117I=SXN101776 cleaved by FXa New England BioLabs and R&D System.

Lane M is a Benchmark ladder, sizes in kDa (Fig 16 in chapter 2 Material and Methods);
Lane 1 is a control sample LH_N/B-scFvCD117I=SXN101776 at -20˚C without addition of FXa;
Lane 2 is a control sample LH_N/B-scFvCD117I=SXN101776 at -20˚C without addition of FXa (+DTT);
Lane 3 is a control sample LH_N/B-scFvCD117I=SXN101776 at +25˚C without addition of FXa;
Lane 4 is a control sample LH_N/B-scFvCD117I=SXN101776 at +25˚C without addition of FXa (+DTT);
Lane 5 on the picture A is cleaved LH_N/B-scFvCD117I=SXN101776 by FXa New England BioLabs at +25˚C;
Lane 6 on the picture A is cleaved LH_N/B-scFvCD117I=SXN101776 by FXa New England BioLabs at +25˚C (+DTT);
Lane 5 on the picture B is cleaved LH_N/B-scFvCD117I=SXN101776 by FXa R&D System at +25˚C;
Lane 6 on the picture A is cleaved LH_N/B-scFvCD117I=SXN101776 by FXa R&D System at +25˚C (+DTT);

Subsequently, in the final step, a His affinity nickel column was substituted with Q column for purification of **LH_N/B-scFvCD117I=SXN101776**. The molecule was purified using a linear gradient. The Q Sepharose column partially separated the band at 103578.3 Da (MW) (indicated by red arrow) from the full-length product 129741.2 Da (indicated by green arrow), however the band at 117516.7 Da (MW) (indicated by blue arrow) still remains and dominates over the full-length product (Fig 60).

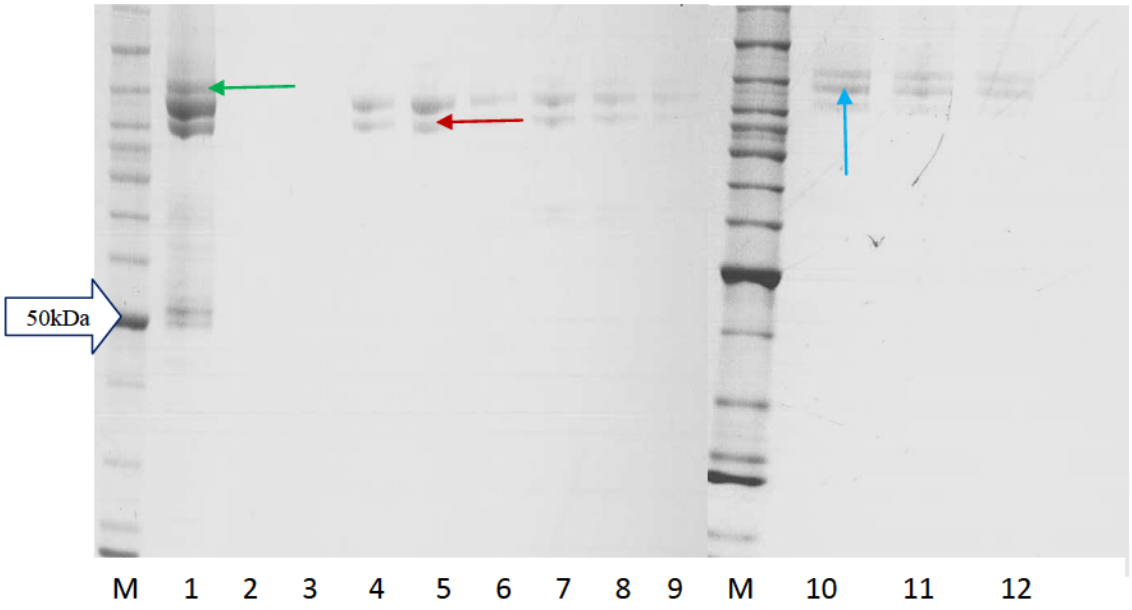


Figure 61 NuPage 4-12% Bis-Tris gel of LH_N/B-scFvCD117I=SXN101776 produced from a Q Sepharose column.

Lane M is a Benchmark ladder, sizes in kDa (Fig 16 in chapter 2 Material and Methods);
Lane 1 is a cleaved LH_N/B-scFvCD117I=SXN101776 by FXa;
Lanes 2, 3,4,5,6,7,8,9, 10, 11 and 12 show a eluted fractions during linear gradient on a Q column;

SDS-PAGE (A) (Fig 61) corresponds to the Western blots against His tag (B) and light chain of B (C) which detected evidence of truncation in purified protein (Fig 61).

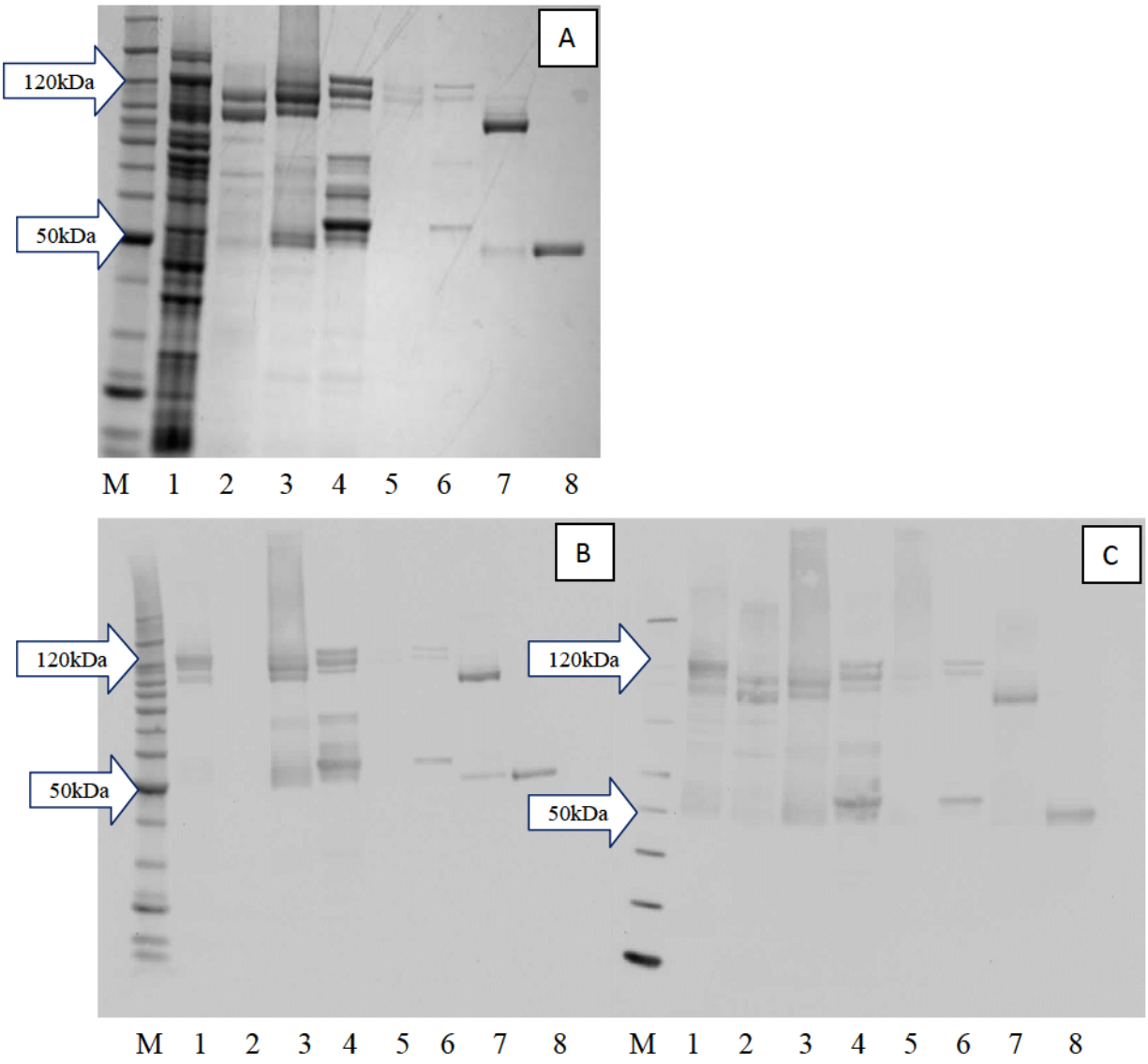


Figure 62 Quality control of the target protein LH_N/B-scFvCD117I=SXN101776 represent on the NuPage 4-12% Bis-Tris gel (A) and confirmed by Western blot against His tag (B) and light chain of B (C).

Lane M on the gel (A) and on Western blots (B) is a Benchmark ladder on the Western blot (C) is a Magic Mark; sizes in kDa (Fig 16 in chapter 2 Material and Methods);
Lane 1 shows clarified lysate of **LH_N/B-scFvCD117I=SXN101776**;
Lane 2 is a -20˚C non-cleaved control **LH_N/B-scFvCD117I=SXN101776**;
Lane 3 is a final target fusion protein **LH_N/B-scFvCD117I=SXN101776** at 0.1 mg/ml loaded at 10µl;
Lane 4 is a final target fusion protein **LH_N/B-scFvCD117I=SXN101776** at 0.1 mg/ml loaded at 10µl (+DTT);
Lane 6 is a final target fusion protein **LH_N/B-scFvCD117I=SXN101776** at 0.1 mg/ml loaded at 5µl;
Lane 7 is a final target fusion protein **LH_N/B-scFvCD117I=SXN101776** at 0.1 mg/ml loaded at 5µl (+DTT);
Lane 8 is an unliganded control **LH_N/B=SXN101687** (Table 11) at 0.1mg/ml loaded at 10µl;
Lane 9 is an unliganded control **LH_N/B=SXN101687** (Table 11) at 0.1mg/ml loaded at 10µl (+DTT);

The unsuccessful purification of **LH_N/B-scFvCD117I=SXN101776** molecule was decided to be rescued by using the Q Sepharose column as a capture step for purification. In addition, the HIC column was considered as a good alternative to prevent truncations of target protein. The Q Sepharose (Fig 62) and HIC hydrophobic Interaction Chromatography (Fig 63) columns substituted the affinity column in the first purification step of **LH_N/B-scFvCD117I=SXN101776**.

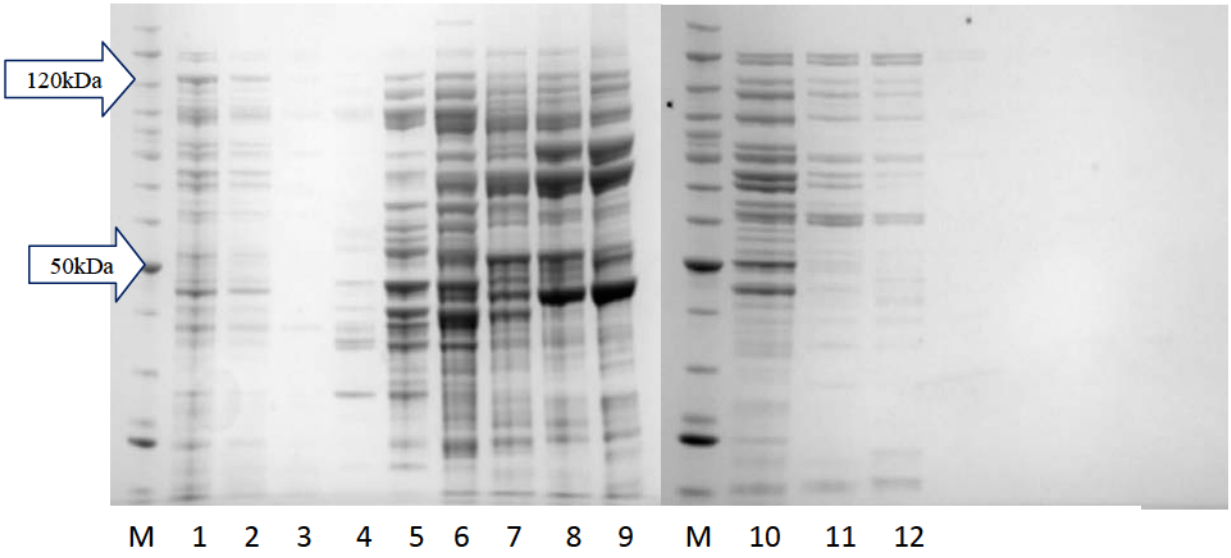


Figure 63 NuPage 4-12% Bis-Tris gel of LH_N/B-scFvCD117I=SXN101776 produced from a Q Sepharose column as first purification step.

Lane M is a Benchmark ladder, sizes in kDa (Fig 16 in chapter 2 Material and Methods);
Lane 1 is a clarified lysate of LH_N/B-scFvCD117I=SXN101776;
Lane 2 is a flow through collected after a Q Sepharose column run;
Lanes 3,4,5,6,7,8,9,10,11 and 12 show a eluted fractions during linear gradient on a Q column;

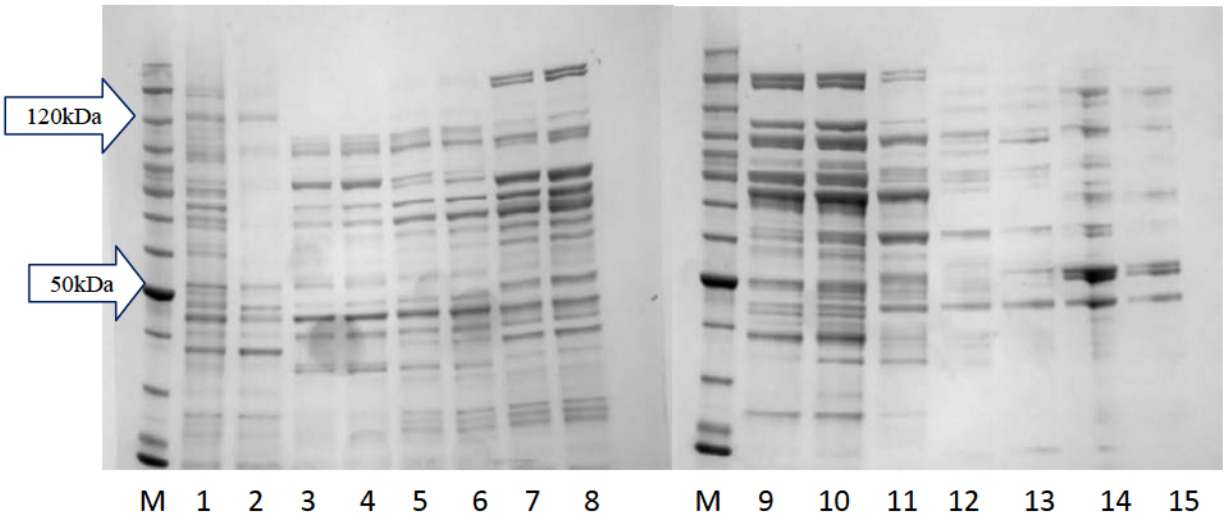


Figure 64 NuPage 4-12 % Bis-Tris gel of LH_N/B-scFvCD117I=SXN101776 produced from a (HIC) Hydrophobic Interaction Chromatography column as first purification step.

Lane M is a Benchmark ladder, sizes in kDa (Fig 16 in chapter 2 Material and Methods);
Lane 1 is a clarified lysate of LH_N/B-scFvCD117I=SXN101776;
Lane 2 is a flow through collected after a HIC column run;
Lanes 3,4,5,6,7,8,9,10,11,12,13,14 and 15 show a eluted fractions during linear gradient on a Q column;

The application of Q Sepharose or HIC column resulted in non-specific cleavage of the eluted proteins, so these methods of purification were not routinely used.

The CD117IScFv was cloned into LH_NC backbone with an FXa cleavage site. The lysate extract containing soluble recombinant **LH_N/C-scFvCD117I=SXN101882** molecule was subject to IMAC purification and desalted on the same day in order to reduce non-specific cleavage. The fusion protein eluted at 250 mM Imidazole (Fig 64) was cleaved with 1.5 Unit of FXa per 1 mg of protein at 25 °C, overnight (Fig 65), but the truncations remained as previously observed for the **LH_N/B-scFvCD117I=SXN101624** or **LH_N/BscFvCD117I=SXN101776** molecule.

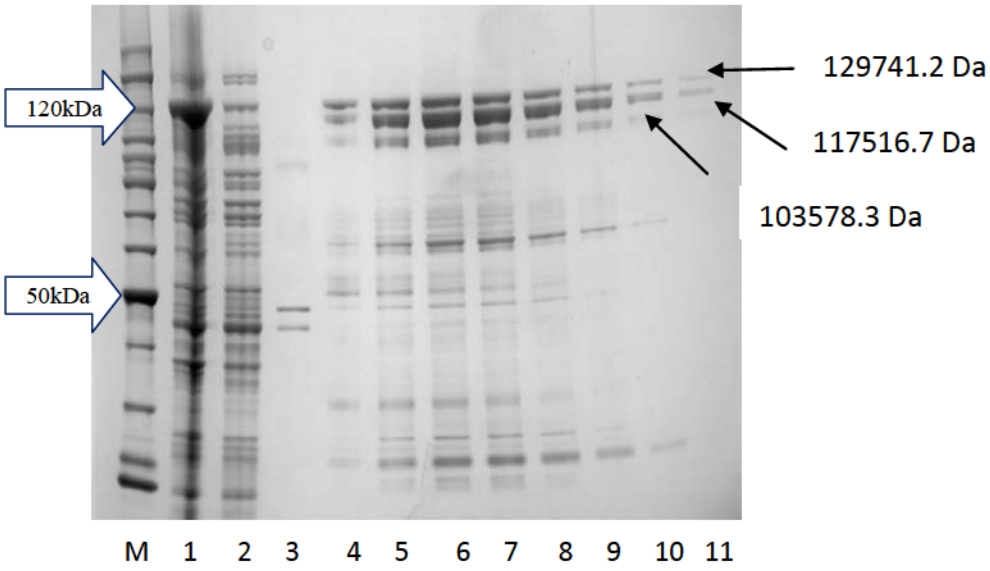


Figure 65 NuPAGE 4-12 % Bis-Tris gel of LH_N/C-scFvCD117I=SXN101882 produced from a nickel column.

Lane M is a Benchmark ladder, sizes in kDa (Fig 16 in chapter 2 Material and Methods);
Lane 1 is a clarified lysate;
Lane 2 is a flow through from His column;
Lane 3 is a 40mM Imidazole wash;
Lane 4, 5, 6, 7, 8, 9, 10 and 11 are desalted 250mM Imidazole washes;
The arrows indicate a truncated products of **LH_N/C-scFvCD117I=SXN101882**;

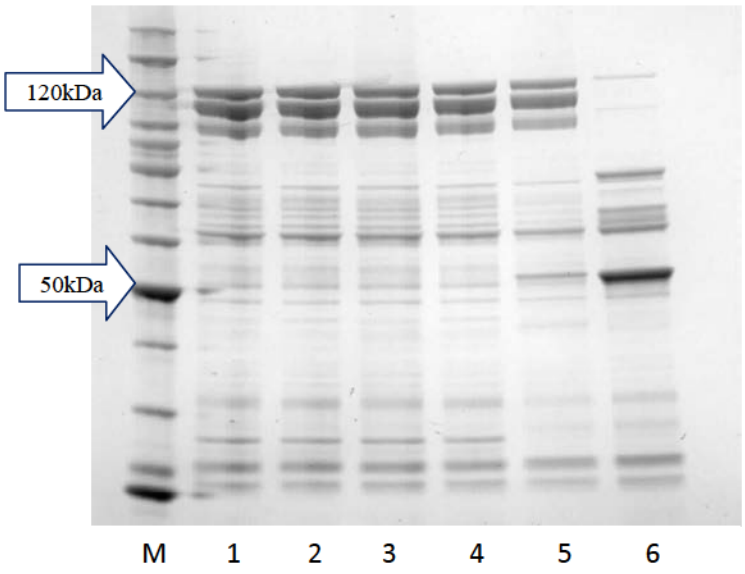


Figure 66 NuPAGE 4-12 % Bis-Tris gel of LH_N/C-scFvCD117I=SXN101882 cleaved by FXa.

Lane M is a Benchmark ladder, sizes in kDa (Fig 16 in chapter 2 Material and Methods);
Lane 1 is a control sample LH_N/C-scFvCD117I=SXN101882 at -20˚C without addition of FXa;
Lane 2 is a control sample LH_N/C-scFvCD117I=SXN101882 at -20˚C without addition of FXa (+DTT);
Lane 3 is a control sample LH_N/C-scFvCD117I=SXN101882 at +25˚C without addition of FXa;
Lane 4 is a control sample LH_N/C-scFvCD117I=SXN101882 at +25˚C without addition of FXa (+DTT);
Lane 5 is cleaved LH_N/C-scFvCD117I=SXN101882 by FXa at +25˚C;
Lane 6 is cleaved LH_N/C-scFvCD117I=SXN101882 by FXa at +25˚C (+DTT);

The fusion protein of **LH_N/C-scFvCD117I=SXN101882** molecule is shown on the SDS-PAGE gel (A). Western blots against His tag (B) and light chain of C (C) detected evidence of truncation in purified protein (Fig 66).

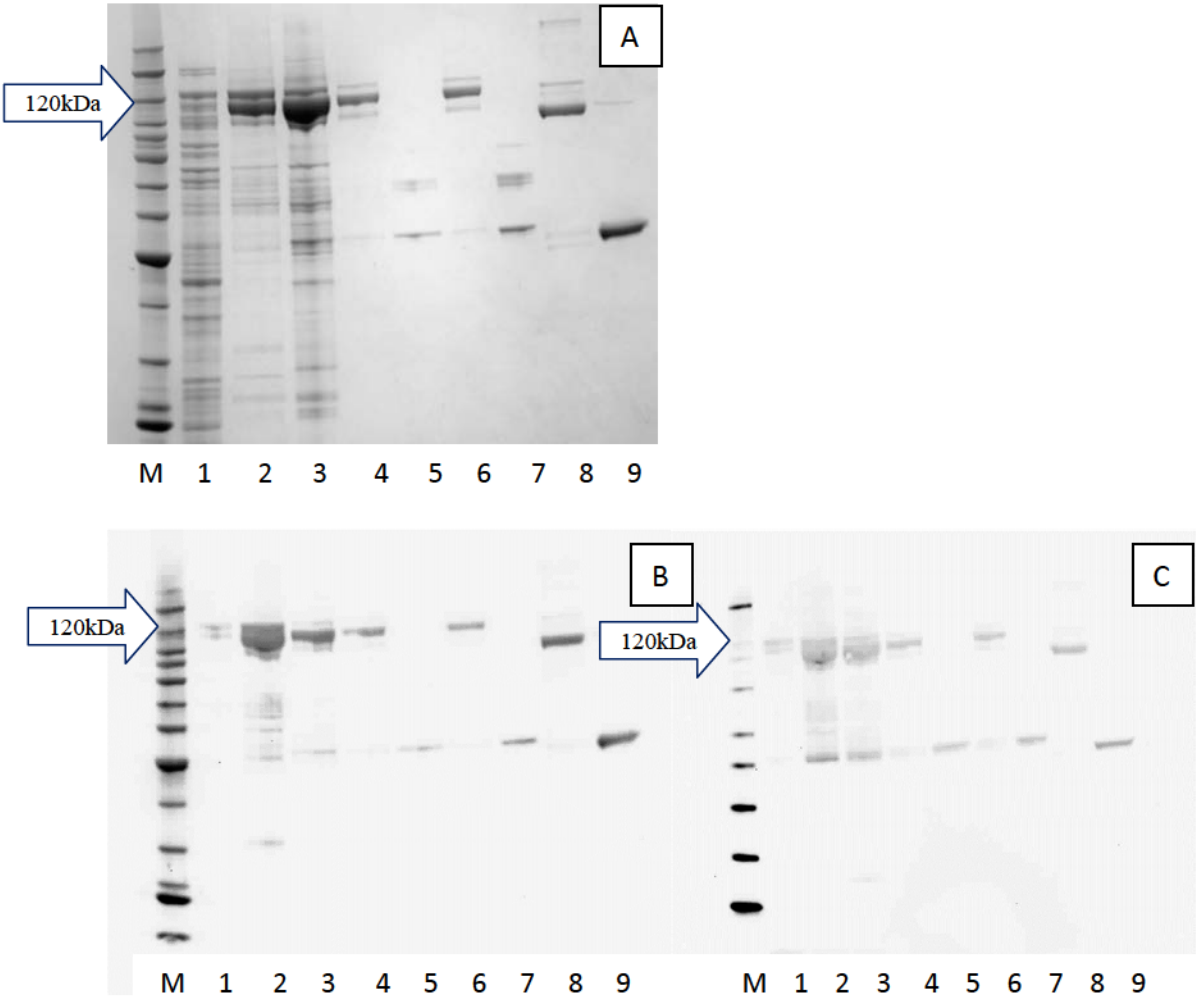


Figure 67 Quality control of the target protein LH_N/C-scFvCD117I=SXN101882 represent on the NuPage 4-12% Bis-Tris gel (A) and confirmed by Western blot against His tag (B) and light chain of C (C).

Lane M on the gel (A) and on Western blots (B) is a Benchmark ladder on the Western blot (C) is a Magic Mark; sizes in kDa (Fig 16 in chapter 2 Material and Methods);
Lane 1 shows clarified lysate of **LH_N/C-scFvCD117I=SXN101882**;
Lane 2 is a -20˚C non-cleaved control **LH_N/C-scFvCD117I=SXN101882**;
Lane 3 is a cleaved **LH_N/C-scFvCD117I=SXN101882 by FXa;**
Lane 3 is a final target fusion protein **LH_N/C-scFvCD117I=SXN101882**at 0.1 mg/ml loaded at 5µl;
Lane 4 is a final target fusion protein **LH_N/C-scFvCD117I=SXN101882**at 0.1 mg/ml loaded at 5µl (+DTT);
Lane 6 is a final target fusion protein **LH_N/C-scFvCD117I=SXN101882**at 0.1 mg/ml loaded at 10µl;
Lane 7 is a final target fusion protein **LH_N/C-scFvCD117I=SXN101882**at 0.1 mg/ml loaded at 10µl (+DTT);
Lane 8 is an unliganded control **LH_N/C=SXN101654** (Table 11) at 0.1mg/ml loaded at 10µl;
Lane 9 is an unliganded control **LH_N/C=SXN101654** (Table 11) **at** 0.1mg/ml loaded at 10µl (+DTT);

The molecule **LH_N/C-scFvCD117I=SXN101882** was sent for a N-terminal sequencing to Alta Bioscience at Birmingham University, in order to establish the accurate position of truncations within ScFv molecules. An N-terminal sequencing was performed on the full-length and reduced proteolytically digested protein. The results indicated that the Mr band 129173.8 is a full-length product starting with the primary sequence of MH (Methionine-Histidine) indicated by the blue arrow (Fig 68) this corresponds to the sequence of **LH_N/C-scFvCD117I=SXN101882**. The FXa cleaved protein in reducing conditions represents the correct band indicated by purple arrow in the figure 69, which starts at **SLYNKTL** motif after **IDGR** activation site. One of the truncated products possibly begins just after the heavy chain of LH_n/C backbone at **NLE (Mr** 102654.6**)** motif (indicated by red arrow in figure 70) and the second one at **GTL (Mr** 115371.8**)** motif (indicated by green arrow), which separates the heavy chain of ScFv from light chain of ScFv just before the **GS15** linker. The introduction of an **EEGEFSAR** linker, which is shorter than the GGGGSGGGGSGGGGS linker used in the existing ScFvs, could be a first approach to resolve the nonspecific cleavage within these molecules.

N terminus

Residue			
1	M ←		
2	H ←	L?	I?
3	H ←	D?	T?
4	N	I?	
5	L?		
6	N	F?	T?
7	---		

Figure 68 Protein sequence report from Alta Bioscience.

N terminus

Residue			
1	S?		
2	L		
3	N?	I?	
4	N		
5	K?		
6	G?	L?	T?
7	L?		

Figure 69 Protein sequence report from Alta Bioscience.

N terminus

Residue			
1	---		
2	N?		
3	L?		
4	T?		
5	E?		
6	L?		
7	N?		

Figure 70 Protein sequence report from Alta Bioscience.

N terminus

Residue			
1	S		
2	L		
3	---		
4	N?		
5	G		
6	T		
7	L		

Figure 71 Protein sequence report from Alta Bioscience.

6.4.2 Purification of single domain antibodies-LH_N chimera

Purification of recombinant single domain antibodies as a part of LH_N chimeras seemed to be more promising than single chain variable antibody fragments-LH_N chimeras, as their structure is less complicated in comparison to recombinant single chain antibodies. AktaExpress System was used to purify molecules **LH_N/B-EGFR31sdAb=SXN101865, LH_N/B-EGFR43sdAb=SXN101866** and **LH_N/A-EGFR6sdAb=SXN101868** (Table 10) on the IMAC column. The fusion proteins eluted at 250 mM Imidazole were activated with FXa and subsequently passed through the IMAC column once more. The buffer was exchanged by a desalting column, which shortens the purification process to two days.

The molecules **LH_N/B-EGFR6sdAb=SXN101861, LH_N/A-EGFR31sdAb=SXN101872, LH_N/D-EGFR28sdAb=SXN101877** and **LH_N/D-EGFR31sdAb=SXN101878** (Table 10) were purified by the affinity chromatography in the capture step and on the HIC (Hydrophobic Interaction Chromatography) during second step. The affinity chromatography method is based on the buffer with higher salt concentration, which is 50 mM Hepes pH 7.2, 500 mM NaCl. The higher salt concentration should enhance production of soluble protein and disassociate contaminants from the fusion protein. An HIC phenyl Sepharose column was introduced as a second purification step to increase purity of the target protein, so they can be used for crystallography study at Bath University.

All full-length fusion proteins of LH_N single domain antibodies chimeras obtained in this study are listed in the table 12.

Both LH_N/ B based molecules the **LH_N/B-EGFR31sdAb=SXN101865** (Fig 72) and the **LH_N/B-EGFR6sdAb=SXN101861** (Fig 73) were purified with high purity and yield.

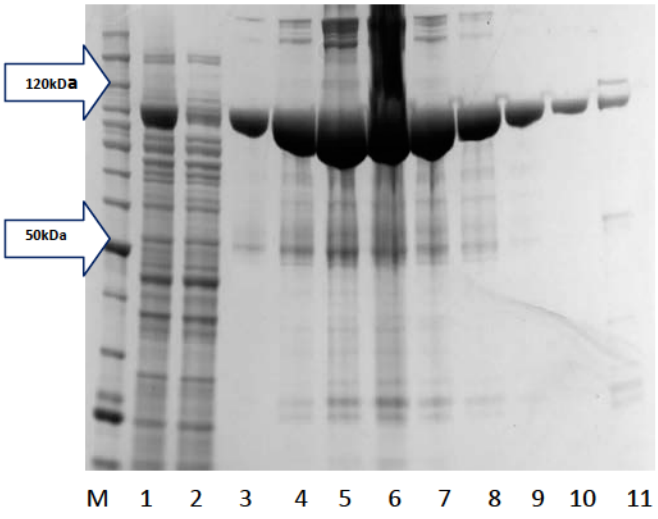


Figure 72 NuPage 4-12% Bis-Tris gel of LH_N/B-EGFR31sdAb=SXN101865 produced from a nickel column.

Lane M is a Benchmark ladder, sizes in kDa (Fig 16 in chapter 2 Material and Methods);
Lane 1 is a clarified lysate;
Lane 2 is a flow through from His column;
Lane 3, 4, 5, 6, 7, 8, 9, 10 and 11 show a desalted 250mM Imidazole elution;

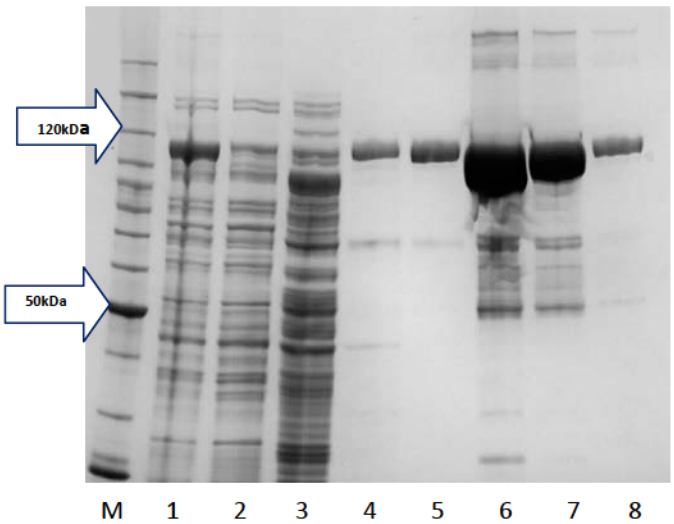


Figure 73 NuPage 4-12% Bis-Tris gel of LH_N/B-EGFR6sdAb=SXN101861 produced from a nickel column.

Lane M is a Benchmark ladder, sizes in kDa (Fig 16 in chapter 2 Material and Methods);
Lane 1 is a clarified lysate;
Lane 2 is a flow through from His column;
Lane 3 is a 40mM Imidazole wash; Lanes 4, 5, 6, 7 and 8 are a 250mM Imidazole washes;

The outcome of activation was better for **LH_N/B-EGFR31sdAb=SXN101865** (Fig 74) than **EGFR6sdAb=SXN101861** (Fig 75) molecule; however, both represent high quality data according to the QC procedure at Syntaxin Ltd (Fig 76 and 77).

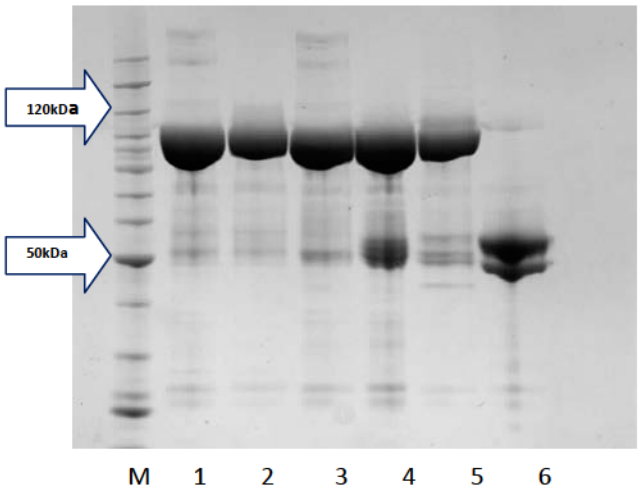


Figure 74 NuPage 4-12% Bis-Tris gel of LH_N/B-EGFR31sdAb=SXN101865 cleaved by FXa.

Lane M is a Benchmark ladder, sizes in kDa (Fig 16 in chapter 2 Material and Methods);
Lane 1 is a control sample LH_N/B-EGFR31sdAb=SXN101865 at -20˚C without addition of FXa;
Lane 2 is a control sample LH_N/B-EGFR31sdAb=SXN101865 at -20˚C without addition of FXa (+DTT);
Lane 3 is a control sample LH_N/B-EGFR31sdAb=SXN101865 at +25˚C without addition of FXa;
Lane 4 is a control sample LH_N/B-EGFR31sdAb=SXN101865 at +25˚C without addition of FXa (+DTT);
Lane 5 is cleaved LH_N/B-EGFR31sdAb=SXN101865 by FXa at +25˚C;
Lane 6 is cleaved LH_N/B-EGFR31sdAb=SXN101865 by FXa at +25˚C (+DTT);

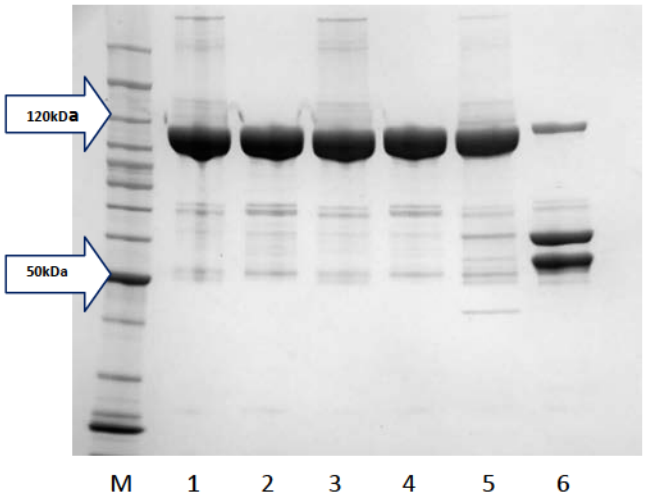


Figure 75 NuPage 4-12% Bis-Tris gel of LH_N/B-EGFR6sdAb=SXN101861 cleaved by FXa.

Lane M is a Benchmark ladder, sizes in kDa (Fig 16 in chapter 2 Material and Methods);
Lane 1 is a control sample LH_N/B-EGFR6sdAb=SXN101861 at -20˚C without addition of FXa;
Lane 2 is a control sample LH_N/B-EGFR6sdAb=SXN101861 at -20˚C without addition of FXa (+DTT);
Lane 3 is a control sample LH_N/B-EGFR6sdAb=SXN101861 at +25˚C without addition of FXa;
Lane 4 is a control sample LH_N/B-EGFR6sdAb=SXN101861 at +25˚C without addition of FXa (+DTT);
Lane 5 is cleaved LH_N/B-EGFR6sdAb=SXN101861 by FXa at +25˚C;
Lane 6 is cleaved LH_N/B-EGFR6sdAb=SXN101861 by FXa at +25˚C (+DTT);

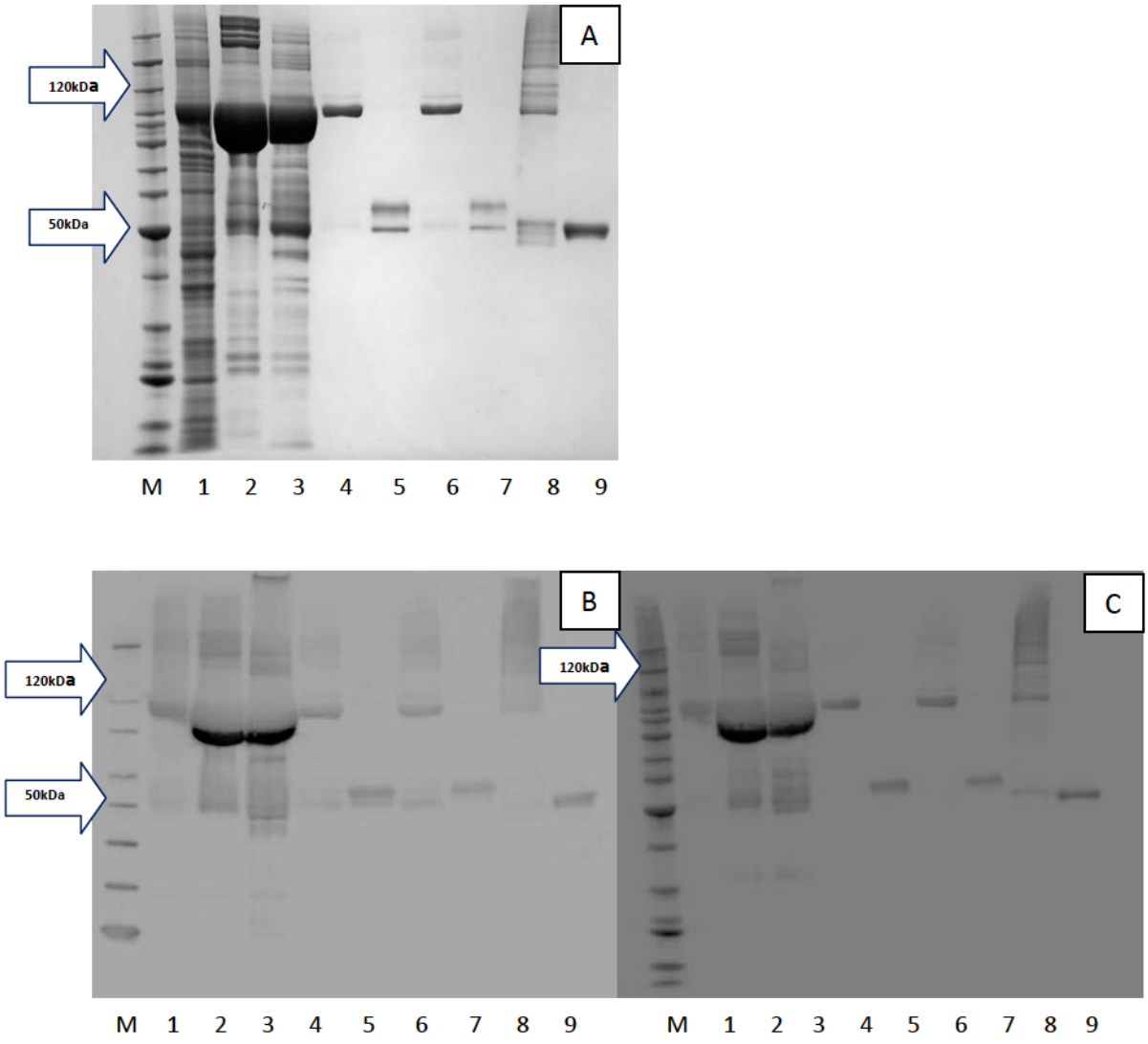


Figure 76 Quality control of the target protein LH_N/B-EGFR31sdAb=SXN101865 represent on the NuPage 4-12% Bis-Tris gel (A) and confirmed by Western blot against His tag (B) and light chain of B (C).

Lane M on the SDS PAGE gel (A) and on Western blot (C) is a Benchmark ladder on the Western blot (B) is a Magic Mark; sizes in kDa (Fig 16 in chapter 2 Material and Methods);
Lane 1 shows clarified lysate of LH_N/B-EGFR31sdAb=SXN101865;
Lane 2 is a -20˚C control of LH_N/B-EGFR31sdAb=SXN101865;
Lane 3 is a cleaved LH_N/B-EGFR31sdAb=SXN101865 by FXa;
Lane 4 is a final target fusion protein LH_N/B-EGFR31sdAb=SXN101865 at 0.1 mg/ml loaded at 10µl;
Lane 5 is a final target fusion protein LH_N/B-EGFR31sdAb=SXN101865 at 0.1 mg/ml loaded at 10µl (+DTT);
Lane 6 is a final target fusion protein LH_N/B-EGFR31sdAb=SXN101865 at 0.1 mg/ml loaded at 5µl;
Lane 7 is a final target fusion protein LH_N/B-EGFR31sdAb=SXN101865 at 0.1 mg/ml loaded at 5µl (+DTT);
Lane 8 is an unliganded control **LH_N/B=SXN101687** (Table 11) at 0.1mg/ml loaded at 10µl;
Lane 9 is an unliganded control **LH_N/B=SXN101687** (Table 11) **at** 0.1mg/ml loaded at 10µl (+DTT);

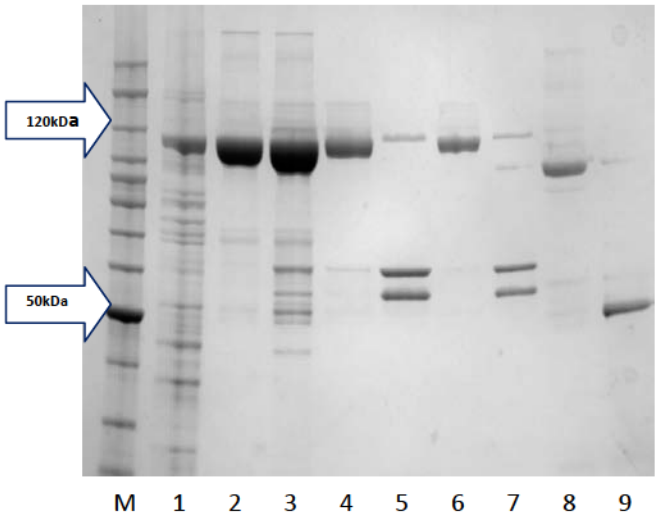


Figure 77 Quality control of the target protein LH_N/B-EGFR6sdAb=SXN101861 represent on the NuPage 4-12% Bis-Tris gel.

Lane M is a Benchmark; sizes in kDa (Fig 16 in chapter 2 Material and Methods);
Lane 1 shows clarified lysate of LH_N/B-EGFR6sdAb=SXN101861;
Lane 2 is a -20˚C control of LH_N/B-EGFR6sdAb=SXN101861;
Lane 3 is a cleaved LH_N/B-EGFR6sdAb=SXN101861 by FXa;
Lane 4 is a final target fusion protein LH_N/B-EGFR6sdAb=SXN101861 at 0.1 mg/ml loaded at 10µl;
Lane 5 is a final target fusion protein LH_N/B-EGFR6sdAb=SXN101861 at 0.1 mg/ml loaded at 10µl (+DTT);
Lane 6 is a final target fusion protein LH_N/B-EGFR6sdAb=SXN101861at 0.1 mg/ml loaded at 5µl;
Lane 7 is a final target fusion protein LH_N/B-EGFR6sdAb=SXN101861at 0.1 mg/ml loaded at 5µl (+DTT);
Lane 8 is an unliganded control **LH_N/B=SXN101687** (Table 11) at 0.1mg/ml loaded at 10µl;
Lane 9 is an unliganded control **LH_N/B=SXN101687** (Table 11) **at** 0.1mg/ml loaded at 10µl (+DTT);

On another instance the molecule **LH_N/B-EGFR43sdAb=SXN101866** was purified with the low yield (Fig 78), did not activate and in addition the unspecific cleavage was observed on the SDS-PAGE gel picture (indicated by turquoise arrow) (Fig 79). This fusion protein was lost during the second part of purification.

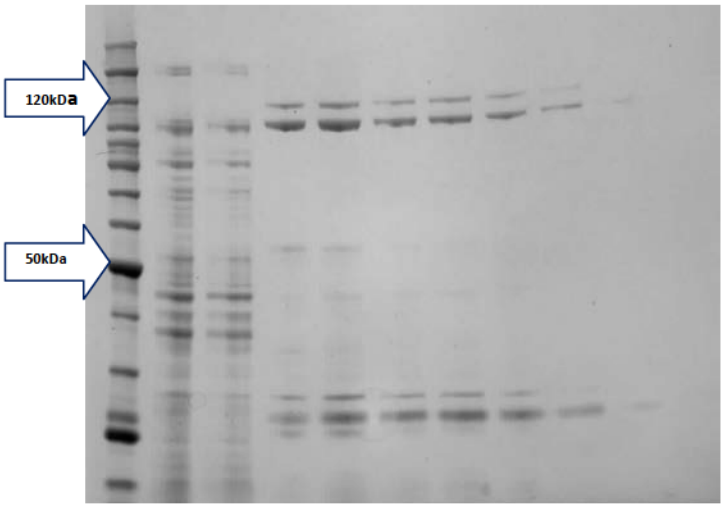


Figure 78 NuPage 4-12% Bis-Tris gel of LH_N/B-EGFR43sdAb=SXN101866 produced from a nickel column.

Lane M is a Benchmark ladder, sizes in kDa (Fig 16 in chapter 2 Material and Methods);
Lane 1 is a clarified lysate;
Lane 2 is a flow through from His column;
Lane 3, 4, 5, 6, 7, 8 and 9 show a desalted 250mM Imidazole elution;

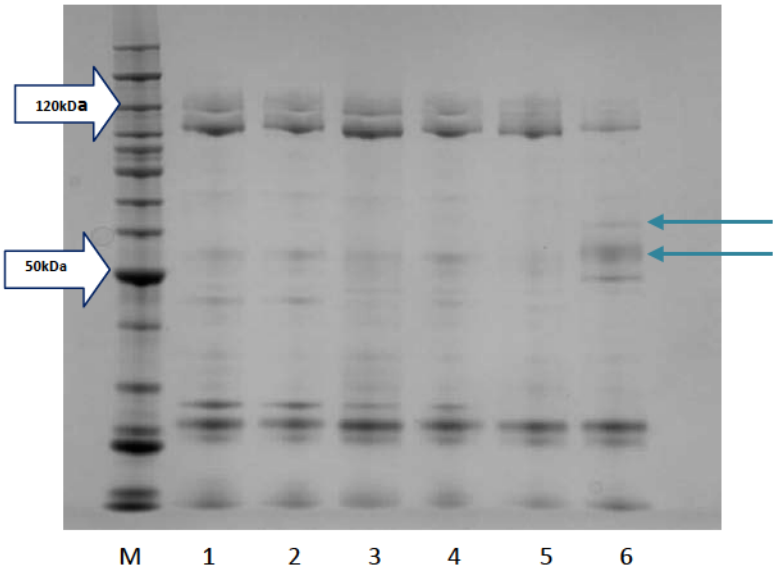


Figure 79 NuPage 4-12% Bis-Tris gel of LH_N/B-EGFR43sdAb=SXN101866 cleaved by FXa.

Lane M is a Benchmark ladder, sizes in kDa (Fig 16 in chapter 2 Material and Methods);
Lane 1 is a control sample LH_N/B-EGFR43sdAb=SXN101866 at -20˚C without addition of FXa;
Lane 2 is a control sample LH_N/B-EGFR43sdAb=SXN101866 at -20˚C without addition of FXa (+DTT);
Lane 3 is a control sample LH_N/B-EGFR43sdAb=SXN101866 at +25˚C without addition of FXa;
Lane 4 is a control sample LH_N/B-EGFR43sdAb=SXN101866 at +25˚C without addition of FXa (+DTT);
Lane 5 is cleaved LH_N/B-EGFR43sdAb=SXN101866 by FXa at +25˚C;
Lane 6 is cleaved LH_N/B-EGFR43sdAb=SXN101866 by FXa at +25˚C (+DTT);

As mentioned in the chapter two ‚Choice of sdAb' according to the tests described in the patent, the EGFR-43sdAb have the highest yield per litre of culture and ability of binding to EGFR-ECD. The EGFR-31sdAb compared to EGFR-43sdAb has five time's lower yield and reduced binding capability. The conclusion made in the patent did not correspond to observations made after purification of single domain antibodies as a part of recombinant botulinum chimeras. This could indicate that single domain antibodies in the presence of recombinant molecules show different characteristics, which could be a result of changes within the structure.

Similar observations were made while purifying **LH_N/A-EGFR6sdAb=SXN101868** (Fig 80) and **LH_N/A-EGFR31sdAb=SXN101872** (Fig 82) molecules. Both molecules were purified to high purity; however, **LH_N/A-EGFR6sdAb=SXN101868** was fully activated (Fig 81), while **LH_N/A-EGFR31sdAb=SXN101872** still carried some inactivated full-length material (Fig 83).

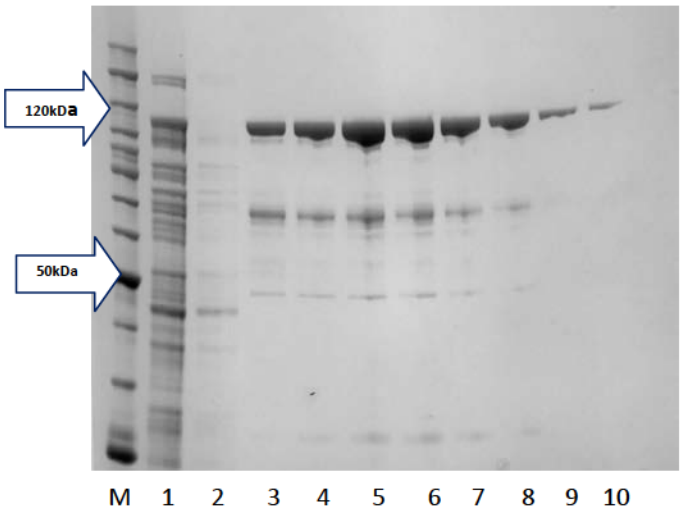


Figure 80 NuPage 4-12% Bis-Tris gel of LH_N/A-EGFR6sdAb=SXN101868 produced from a nickel column.

Lane M is a Benchmark ladder, sizes in kDa (Fig 16 in chapter 2 Material and Methods);
Lane 1 is a clarified lysate;
Lane 2 is a flow through from His column;
Lane 3, 4, 5, 6, 7, 8, 9 and 10show a desalted 250mM Imidazole elution;

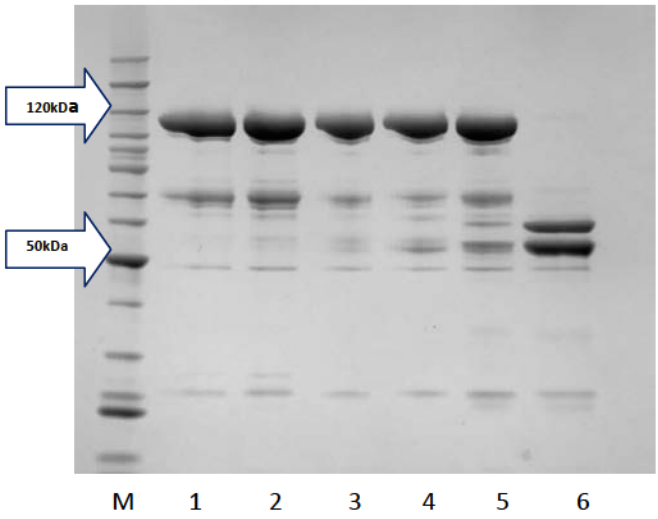


Figure 81 NuPage 4-12% Bis-Tris gel of LH_N/A-EGFR6sdAb=SXN101868 cleaved by EK.

Lane M is a Benchmark ladder, sizes in kDa (Fig 16 in chapter 2 Material and Methods);
Lane 1 is a control sample LHN/A-EGFR6sdAb=SXN101868 at -20˚C without addition of EK;
Lane 2 is a control sample LHN/A-EGFR6sdAb=SXN101868 at -20˚C without addition of EK (+DTT);
Lane 3 is a control sample LHN/A-EGFR6sdAb=SXN101868 at +25˚C without addition of EK;
Lane 4 is a control sample LHN/A-EGFR6sdAb=SXN101868 at +25˚C without addition of EK (+DTT);
Lane 5 is cleaved LHN/A-EGFR6sdAb=SXN101868 by EK at +25˚C;
Lane 6 is cleaved LHN/A-EGFR6sdAb=SXN101868 by EK at +25˚C (+DTT);

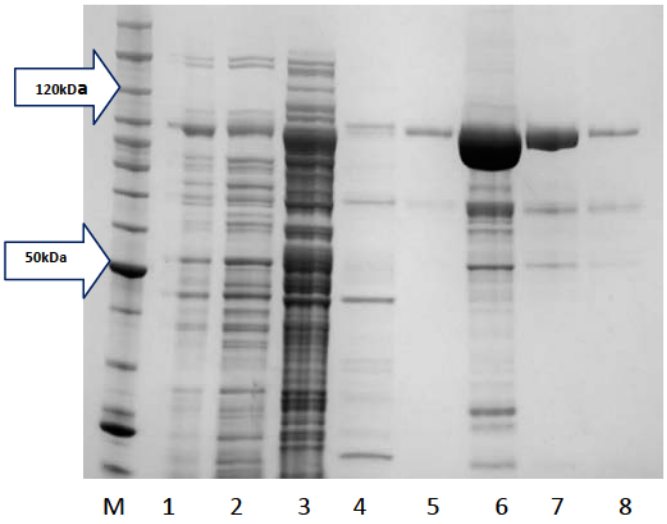


Figure 82 NuPage 4-12% Bis-Tris gel of LH_N/A-EGFR31sdAb=SXN101872 produced from a nickel column.

Lane M is a Benchmark ladder, sizes in kDa (Fig 16 in chapter 2 Material and Methods);
Lane 1 is a clarified lysate;
Lane 2 is a flow through from His column;
Lane 3 is a 40mM Imidazole wash; Lane 4 is a 80mM Imidazole wash; Lanes 5, 6, 7 and 8 are a 250mM Imidazole washes;

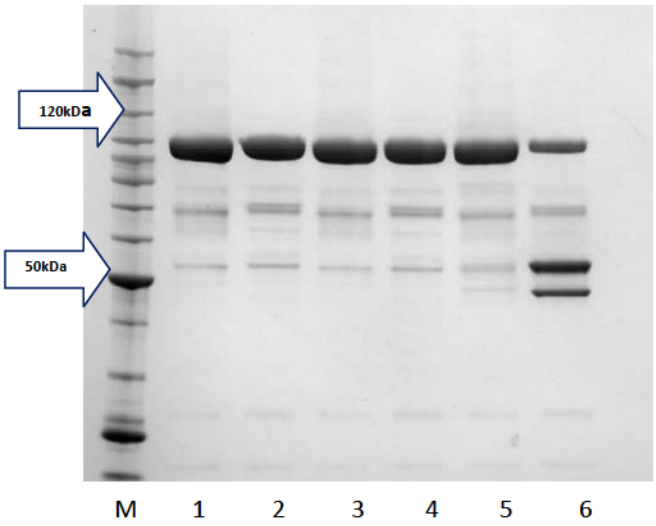


Figure 83 NuPage 4-12% Bis-Tris gel of LH_N/A-EGFR31sdAb=SXN101872 cleaved by EK.

Lane M is a Benchmark ladder, sizes in kDa (Fig 16 in chapter 2 Material and Methods);
Lane 1 is a control sample LH_N/A-EGFR31sdAb=SXN101872 at -20˚C without addition of EK;
Lane 2 is a control sample LH_N/A-EGFR31sdAb=SXN101872 at -20˚C without addition of EK (+DTT);
Lane 3 is a control sample LH_N/A-EGFR31sdAb=SXN101872 at +25˚C without addition of EK;
Lane 4 is a control sample LH_N/A-EGFR31sdAb=SXN101872 at +25˚C without addition of EK (+DTT);
Lane 5 is cleaved LH_N/A-EGFR31sdAb=SXN101872 by EK at +25˚C;
Lane 6 is cleaved LH_N/A-EGFR31sdAb=SXN101872 by EK at +25˚C (+DTT);

Regarding the data available in house, the LH_N/A based constructs do not activate as well as an **LH_N/A-EGFR6sdAb=SXN101868**, therefore the hypothesis that the EGFR-6sdAbs improve cleavage and enhance purity could be accepted. It is also important to emphasize that the **LH_N/A-EGFR6sdAb=SXN101868** was purified only on the two-step IMAC columns, whereas an **LH_N/A-EGFR31sdAb=SXN101872** was passed on the HIC phenyl Sepharose in the second purification step, which further improves the purity of this molecule (Fig 84). Nevertheless, the **LH_N/A-EGFR6sdAb=SXN101868** compound was of higher quality in comparison to compound **LH_N/A-EGFR31sdAb=SXN101872**, despite of less advanced purification process being implemented.

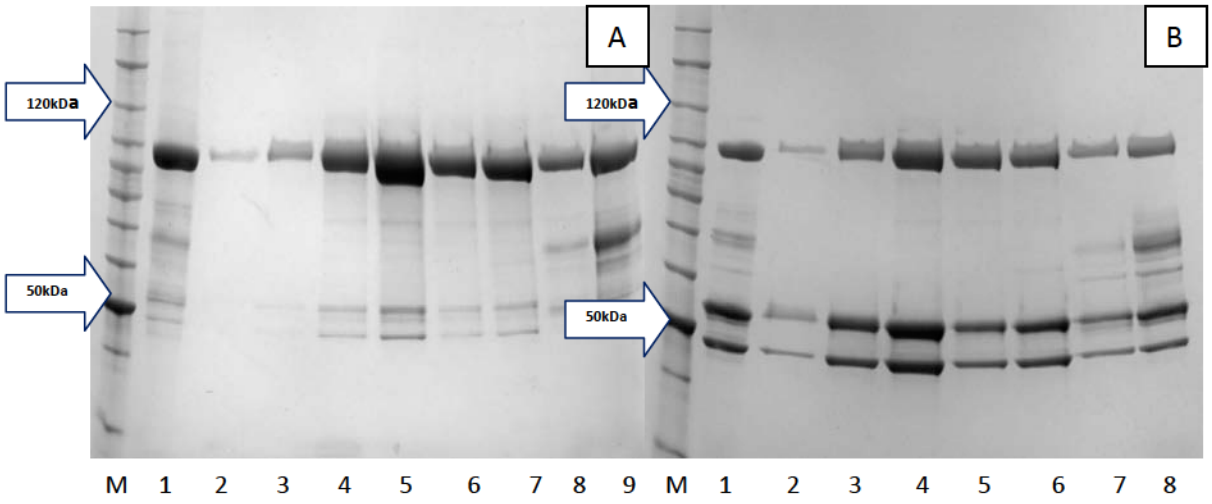


Figure 84 NuPage 4-12% Bis-Tris gels of LH_N/A-EGFR31sdAb=SXN101872 produced from a HIC column.

Lane M is a Benchmark ladder, sizes in kDa (Fig 16 in chapter 2 Material and Methods);
Lane 1 is a cleaved LH_N/A-EGFR31sdAb=SXN101872;
Lanes 2, 3, 4, 5, 6, 7, 8 and 9 are elution fractions after hydrophobic interaction chromatography column;
The samples on the SDS PAGE gel (A) were run in oxidizing condition;
The samples on the SDS PAGE gel (B) were run in reducing condition (+DTT);

On this instance, the data shown in the patent regarding to the molecule EGFR-31sdAbs has been confirmed, where an EGFR-31sdAb molecule out of four tested, has the lowest final yield as well as an **LH_N/A-EGFR31sdAb=SXN101872** molecule cloned for this study.

The molecule EGFR-6sdAb was not included in the experiments described in the patent, but it belongs to group 2 together with EGFR-10sdAb, which has reasonable ability of binding in the longest period in comparison to other tested molecules, also it has shown a reasonable yield after purification.

In overall, the high quality of data can be obtained by a combination of appropriate choice of the single domain antibody and the LH_N structure of the different serotype (Fig 85 and 86).

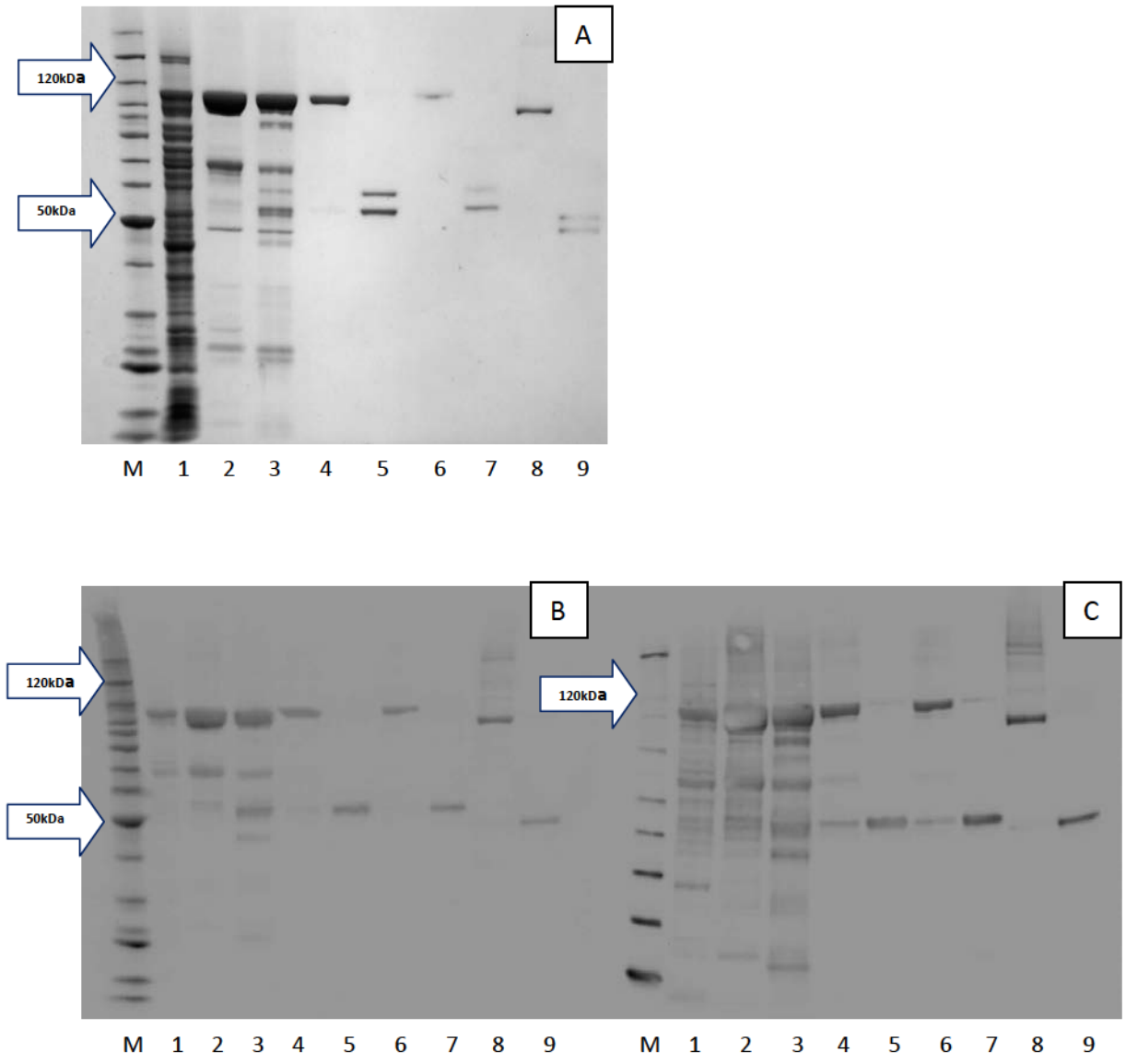


Figure 85 Quality control of the target protein LH_N/A-EGFR6sdAb=SXN101868 represent on the NuPage 4-12% Bis-Tris gel (A) and confirmed by Western blot against His tag (B) and light chain of A (C).

Lane M on the SDS PAGE gel (A) and on Western blot (B) is a Benchmark ladder on the Western blot (C) is a Magic Mark; sizes in kDa (Fig 16 in chapter 2 Material and Methods);
Lane 1 shows clarified lysate of LH_N/A-EGFR6sdAb=SXN101868;
Lane 2 is a -20˚C control of LH_N/A-EGFR6sdAb=SXN101868;
Lane 3 is a cleaved LH_N/A-EGFR6sdAb=SXN101868by EK;
Lane 4 is a final target fusion protein LH_N/A-EGFR6sdAb=SXN101868at 0.1 mg/ml loaded at 10µl;
Lane 5 is a final target fusion protein LH_N/A-EGFR6sdAb=SXN101868at 0.1 mg/ml loaded at 10µl (+DTT);
Lane 6 is a final target fusion protein LH_N/A-EGFR6sdAb=SXN101868at 0.1 mg/ml loaded at 5µl;
Lane 7 is a final target fusion protein LH_N/A-EGFR6sdAb=SXN101868at 0.1 mg/ml loaded at 5µl (+DTT);
Lane 8 is an unliganded control **LH_N/A=SXN100590** (Table 11) at 0.1mg/ml loaded at 10µl;
Lane 9 is an unliganded control **LH_N/A=SXN100590** (Table 11) **at** 0.1mg/ml loaded at 10µl (+DTT);

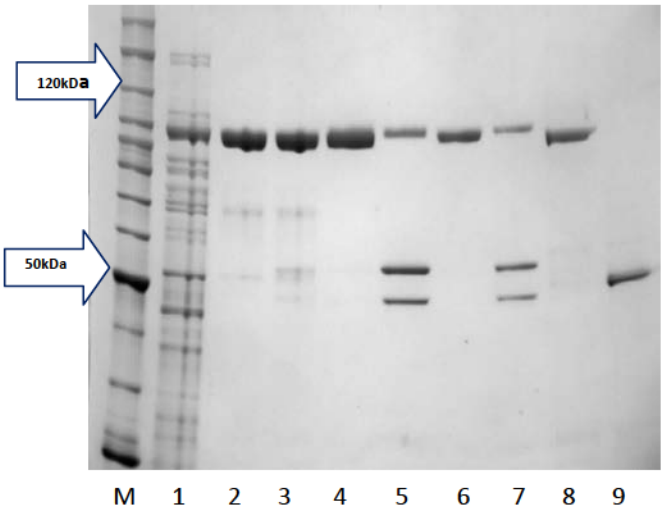


Figure 86 Quality control of the target protein LH_N/A-EGFR31sdAb=SXN101872 represent on the NuPage 4-12 % Bis-Tris gel.

Lane M is a Benchmark ladder; sizes in kDa (Fig 16 in chapter 2 Material and Methods);
Lane 1 shows clarified lysate of LH_N/A-EGFR31sdAb=SXN101872;
Lane 2 is a -20˚C control of LH_N/A-EGFR31sdAb=SXN101872;
Lane 3 is a cleaved LH_N/A-EGFR31sdAb=SXN101872by EK;
Lane 4 is a final target fusion protein LH_N/A-EGFR31sdAb=SXN101872 at 0.1 mg/ml loaded at 10µl;
Lane 5 is a final target fusion protein LH_N/A-EGFR31sdAb=SXN101872 at 0.1 mg/ml loaded at 10µl (+DTT);
Lane 6 is a final target fusion protein LH_N/A-EGFR31sdAb=SXN101872 at 0.1 mg/ml loaded at 5µl;
Lane 7 is a final target fusion protein LH_N/A-EGFR31sdAb=SXN101872 at 0.1 mg/ml loaded at 5µl (+DTT);
Lane 8 is an unliganded control **LH_N/A=SXN100590** (Table 11) at 0.1mg/ml loaded at 10µl;
Lane 9 is an unliganded control **LH_N/A=SXN100590** (Table 11) at 0.1mg/ml loaded at 10µl (+DTT);

In addition to these, the LH_N/D based molecules **LH_N/D-EGFR28sdAb=SXN101877** (Fig 87) and **LH_N/D-EGFR31sdAb=SXN101878** (Fig 88) were purified, therefore the data were compared to establish more robust conclusion about the single domain antibodies in the combination of LH_N structure.

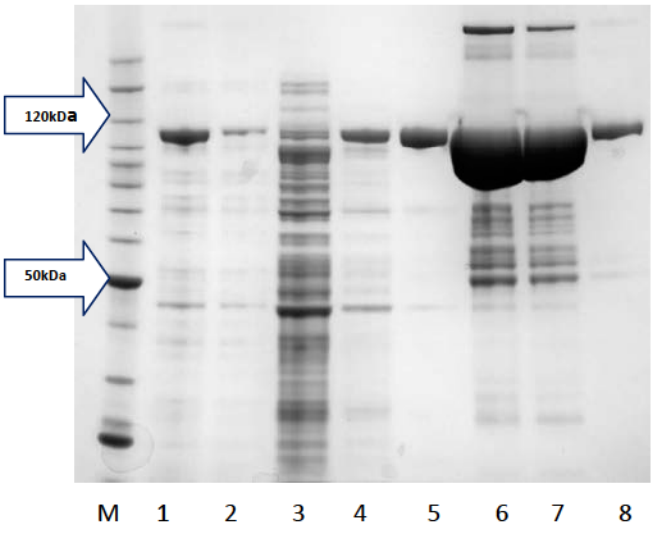


Figure 87 NuPage 4-12% Bis-Tris gel of LH_N/D-EGFR28sdAb=SXN101877 produced from a nickel column.

Lane M is a Benchmark ladder, sizes in kDa (Fig 16 in chapter 2 Material and Methods);
Lane 1 is a clarified lysate;
Lane 2 is a flow through from His column;
Lane 3 is a 40mM Imidazole wash; Lane 4 is a 80mM Imidazole wash; Lanes 5, 6, 7 and 8 are a 250mM Imidazole washes;

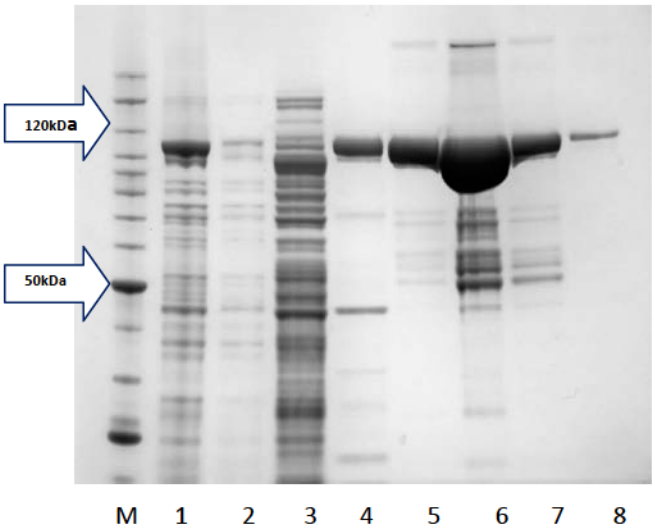


Figure 88 NuPage 4-12% Bis-Tris gel of L_N/D-EGFR31sdAb=SXN101878 produced from a nickel column.

Lane M is a Benchmark ladder, sizes in kDa (Fig 16 in chapter 2 Material and Methods);
Lane 1 is a clarified lysate;
Lane 2 is a flow through from His column;
Lane 3 is a 40mM Imidazole wash; Lane 4 is a 80mM Imidazole wash; Lanes 5, 6, 7 and 8 are a 250mM Imidazole washes;

The molecules **LH_N/D-EGFR28sdAb=SXN101877** and **LH_N/D-EGFR31sdAb=SXN101878** purified to high standards were cleaved effectively by Enterokinase, as shown on SDS-PAGE in figure 89 and 90.

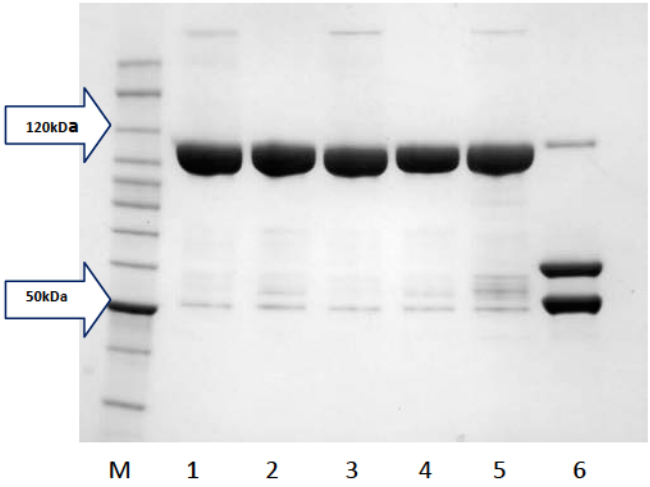


Figure 89 NuPage 4-12% Bis-Tris gel of LH_N/D-EGFR28sdAb=SXN101877 cleaved by EK.

Lane M is a Benchmark ladder, sizes in kDa (Fig 16 in chapter 2 Material and Methods);
Lane 1 is a control sample LH_N/D-EGFR28sdAb=SXN101877 at -20˚C without addition of EK;
Lane 2 is a control sample LH_N/D-EGFR28sdAb=SXN101877 at -20˚C without addition of EK (+DTT);
Lane 3 is a control sample LH_N/D-EGFR28sdAb=SXN101877 at +25˚C without addition of EK;
Lane 4 is a control sample LH_N/D-EGFR28sdAb=SXN101877 at +25˚C without addition of EK (+DTT);
Lane 5 is cleaved LH_N/D-EGFR28sdAb=SXN101877 by EK at +25˚C;
Lane 6 is cleaved LH_N/D-EGFR28sdAb=SXN101877 by EK at +25˚C (+DTT);

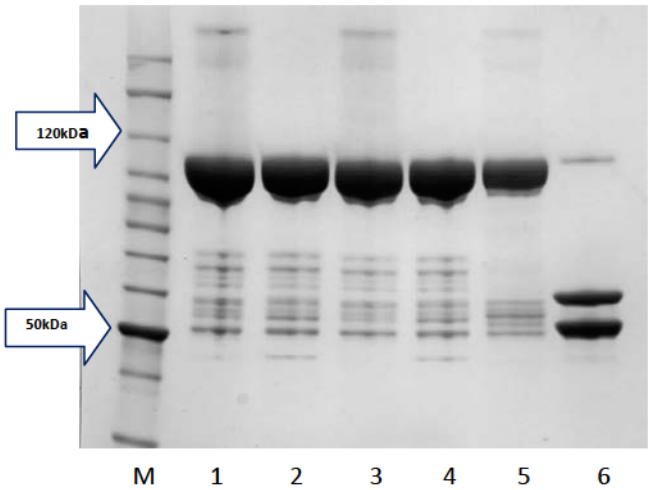


Figure 90 NuPage 4-12% Bis-Tris gel of LH_N/D-EGFR31sdAb=SXN101878 cleaved by EK.

Lane M is a Benchmark ladder, sizes in kDa (Fig 16 in chapter 2 Material and Methods);
Lane 1 is a control sample LH_N/D-EGFR31sdAb=SXN101878 at -20˚C without addition of EK;
Lane 2 is a control sample LH_N/D-EGFR31sdAb=SXN101878 at -20˚C without addition of EK (+DTT);
Lane 3 is a control sample LH_N/D-EGFR31sdAb=SXN101878 at +25˚C without addition of EK;
Lane 4 is a control sample LH_N/D-EGFR31sdAb=SXN101878 at +25˚C without addition of EK (+DTT);
Lane 5 is cleaved LH_N/D-EGFR31sdAb=SXN101878 by EK at +25˚C;
Lane 6 is cleaved LH_N/D-EGFR31sdAb=SXN101878 by EK at +25˚C (+DTT);

However, an **LH_N/D-EGFR31sdAb=SXN101878** (Fig 92) molecule has higher purity than an **LH_N/D-EGFR28sdAb=SXN101877** (Fig 91) fusion protein; the yield of **LH_N/D-EGFR28sdAb=SXN101877** is almost four times higher in comparison to the molecule **LH_N/D-EGFR31sdAb=SXN101878**.

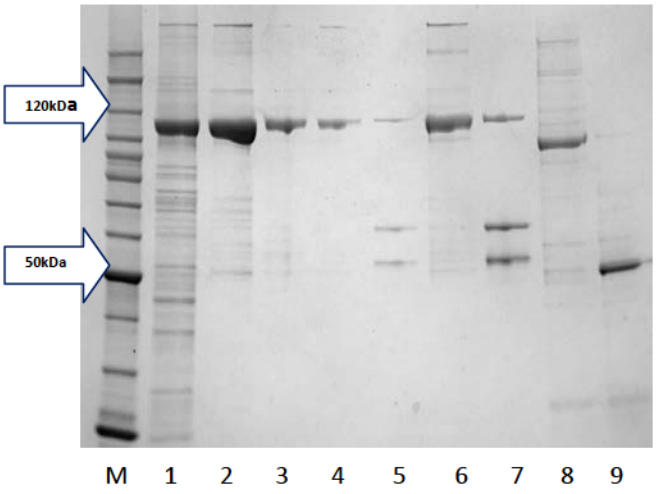


Figure 91 Quality control of the target protein LH_N/D-EGFR28sdAb=SXN101877 represent on the NuPage 4-12% Bis-Tris gel.

Lane M is a Benchmark ladder; sizes in kDa (Fig 16 in chapter 2 Material and Methods);
Lane 1 shows clarified lysate of LH_N/D-EGFR28sdAb=SXN101877;
Lane 2 is a -20˚C control of LH_N/D-EGFR28sdAb=SXN101877;
Lane 3 is a cleaved LH_N/D-EGFR28sdAb=SXN101877 by EK;
Lane 4 is a final target fusion protein LH_N/D-EGFR28sdAb=SXN101877 at 0.1 mg/ml loaded at 5µl;
Lane 5 is a final target fusion protein LH_N/D-EGFR28sdAb=SXN101877 at 0.1 mg/ml loaded at 5µl (+DTT);
Lane 6 is a final target fusion protein LH_N/D-EGFR28sdAb=SXN101877 at 0.1 mg/ml loaded at 10µl;
Lane 7 is a final target fusion protein LH_N/D-EGFR28sdAb=SXN101877 at 0.1 mg/ml loaded at 10µl (+DTT);
Lane 8 is an unliganded control **LH_N/D=SXN101655** (Table 11) at 0.1mg/ml loaded at 10µl;
Lane 9 is an unliganded control **LH_N/D=SXN101655** (Table 11) at 0.1mg/ml loaded at 10µl (+DTT);

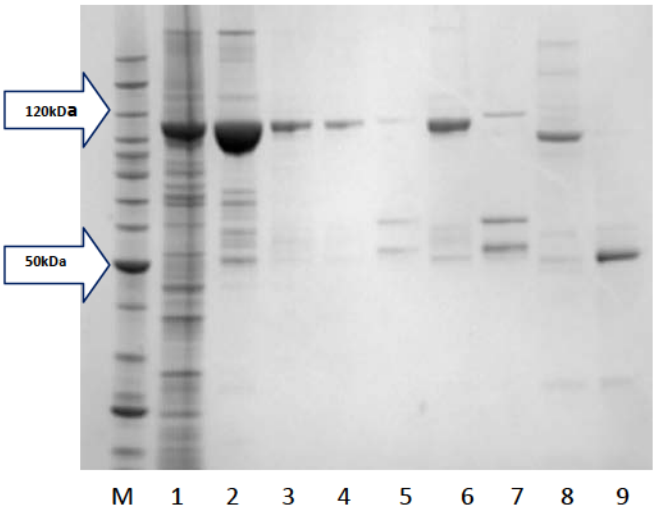


Figure 92 Quality control of the target protein LH_N/D-EGFR31sdAb=SXN101878 represent on the NuPage 4-12 % Bis-Tris gel.

Lane M is a Benchmark ladder; sizes in kDa (Fig 16 in chapter 2 Material and Methods);
Lane 1 shows clarified lysate of LH_N/D-EGFR31sdAb=SXN101878;
Lane 2 is a -20˚C control of LH_N/D-EGFR31sdAb=SXN101878;
Lane 3 is a cleaved LH_N/D-EGFR31sdAb=SXN101878 by EK;
Lane 4 is a final target fusion protein LH_N/D-EGFR31sdAb=SXN101878 at 0.1 mg/ml loaded at 5µl;
Lane 5 is a final target fusion protein LH_N/D-EGFR31sdAb=SXN101878 at 0.1 mg/ml loaded at 5µl (+DTT);
Lane 6 is a final target fusion protein LH_N/D-EGFR31sdAb=SXN101878 at 0.1 mg/ml loaded at 10µl;
Lane 7 is a final target fusion protein LH_N/D-EGFR31sdAb=SXN101878 at 0.1 mg/ml loaded at 10µl (+DTT);
Lane 8 is an unliganded control **LH_N/D=SXN101655** (Table 11) at 0.1mg/ml loaded at 10µl;
Lane 9 is an unliganded control **LH_N/D=SXN101655** (Table 11) at 0.1mg/ml loaded at 10µl (+DTT);

The data obtained from purification of the **LH_N/A-EGFR31sdAb=SXN101872** and the **LH_N/D-EGFR31sdAb=SXN101878** in a LH_N/A and a LH_N/D backbone are comparable to the data obtained in the patent WO 2008/141449 A1 on the EGFR31sdAb. The final yield of purified molecules is lower in comparison to the yield of other LH_N single domain antibodies and the same time similar to the yield of EGFR31sdAb. The exception was observed for the compound based on a LH_N/B backbone, where compound **LH_N/B-EGFR31sdAb=SXN101865** gave outstanding results in comparison to the other molecules, the same time it becomes so far the best choice to achieve high quality of data.

It is also important to compare the data obtained from the purification of **LH_N/D-EGFR28sdAb=SXN101877** molecule to the data of EGFR28sdAb obtained from the patent. This molecule EGFR28sdAb has not been tested by assays established in the patent. However, the EGFR2-sdAb that belongs to the same group as the EGFR28sdAb showed the strongest binding ability in the longest time and it did not give high yields

of the final expressed product. The yield of purified **LH_N/D-EGFR28sdAb=SXN101877** is comparable to the yield of EGFR2-sdAb; therefore it seems to be comparable to data obtained from the patent. This conclusion can be only accepted if I assume that the grouping system of the single domain antibodies in the patent is analogous.

Table 12. The fusion proteins of the full-length antibody-LH_N chimeras obtained in this study.

Name	SXN number	E number	MW Molecular Weight Da	Purification method	Purity %	A280nm mg/ml	Yield mg	BCA assay mg/ml
10HT-LD-EK-HD-EGFR31sdAb	SXN101878	E001423	115759.4	His-EK-HIC	90	9.8	6.8	5.0
10HT-LB-A-linker-Xa-HB-EGFR6sdAb	SXN101861	E001406	116880.5	His-EK-HIC	96	5.7	18.9	4.1
10HT-LB-A-linker-Xa-HB-EGFR31sdAb	SXN101865	E001410	116345.0	His-Desalt-Fxa-His	96	6.1	18.3	4.2
10HT-LA-EK-HA-EGFR6sdAb	SXN101868	E001413	115531.3	His-Desalt-Fxa-His	97	3.6	3.3	3.0
10HT-LA-EK-HA-EGFR31sdAb	SXN101872	E001417	114832.6	His-EK-HIC	100	7	9.5	5.0
10HT-LD-EK-HD-EGFR28sdAb	SXN101877	E001422	117018.5	His-EK-HIC	74	11.8	24.8	N/A

Chapter 7

Activation Studies

Chapter 7. Activation Studies

Engineering functional recombinant proteins based on antibody domains and fragments of *C. botulinum* neurotoxin was the concept of the project that is described in chapters 3 to 6. The successful production of the molecules gave opportunity to perform additional experiments, which were not required for purpose of this project, but they covered the area of personal interest. The area of interest includes the investigation of the functionality of produced molecules by studying activation and internalization of constructs of this type and other Syntaxin TSIs (Targeted Secretion Inhibitors). The preparation of activation and internalization study also enabled to develop new techniques.

7.1 Principle of the activation study

The purpose of experiment was to measure a receptor activation and phosphorylation on cells. The experiments were based on the tyrosine kinase receptor family, as the produced Syntaxin Ltd molecules carried ligand against this receptor. This is EGF the (Epidermal growth factor) ligand of the tyrosine kinase family of receptors.

7.2 Ligand binding

A ligand is a substance that is able to bind and form a complex with biomolecules to serve a biological purpose. A ligand in this study is a single domain or a single chain ant-EGFR antibody fragment that is a part of recombinant botulinum neurotoxin molecule. The biomolecule that forms a complex with ligand is a receptor in the plasma membranes.

Receptors possess sites accessible to the extracellular environment, where binding of soluble ligands can take place. The amino acids, which make up the binding site, determine specificity and affinity for the ligand. John Newport Langley discovered the concept of specific binding site for a ligand (Langley, 1878).

Receptors bind ligands via weak noncovalent interactions, such as ionic interactions, van der Waals forces, hydrogen bonding or hydrophobic interactions. As the

irreversible covalent binding between ligand and receptor does not take place, so the docking is reversible. Therefore, the receptor can be recycled for subsequent use. The strength of binding ligand to receptor is its affinity. The concentration of ligand has an impact on the detection of the ligand by receptor and subsequently an affinity. Low concentration of ligand present on the ligand-binding site results in high affinity ligand binding, while high concentration of ligand implies to low affinity of ligand.

The Docking is a method that predicts the preferred orientation of one molecule to a second when bound to each other to form a stable complex. Molecular docking is described as a ‚*best-fit*' orientation of a ligand that binds to a particular protein of interest. Molecular definition of docking represents two analogy: ‚*lock-and-key*' and ‚*hand-in-glove*'. The second analogy seems to be more appropriate since the ligand and the receptor are flexible to achieve the ‚*best fit*' (Hancock, 1997).

The orientation of the two interacting molecules (receptor and ligand) may affect a type of signal produced. A ligand is called antagonist or agonist based on the signal produced. An antagonist is a ligand that binds to receptor, but does not result in the activation of it. However if the ligand binds to a receptor which then is activated it is called an agonist. This binding stimulates the intracellular response via a complex signal-transduction pathway. The cell needs to respond to a variety of signals, therefore a range of receptors are involved in this role.

Signalling receptors are divided into five classes: ion channel linked, G protein linked, containing intrinsic enzymatic activity, tyrosine kinase linked, intracellular. I have been interested in the tyrosine kinase pathway, as EGFR belongs to this family.

7.3 Receptor Tyrosine Kinases

Receptor tyrosine kinases (RTK)s are the high affinity cell surface receptors. These receptors possess an extracellular ligand-binding domain, a single transmembrane domain, and a cytoplasmic domain that contains the kinase activity. The binding of the ligand into extracellular binding site of the receptor activates the kinase activity on the cytoplasmic side of the membrane. EGF (epidermal growth factor) is the most common extracellular protein ligand that initiates the MAPK/ERK pathway, which tests were attempted in these experiments.

In the diagram, "P" represents phosphate. Note EGF at the very top.

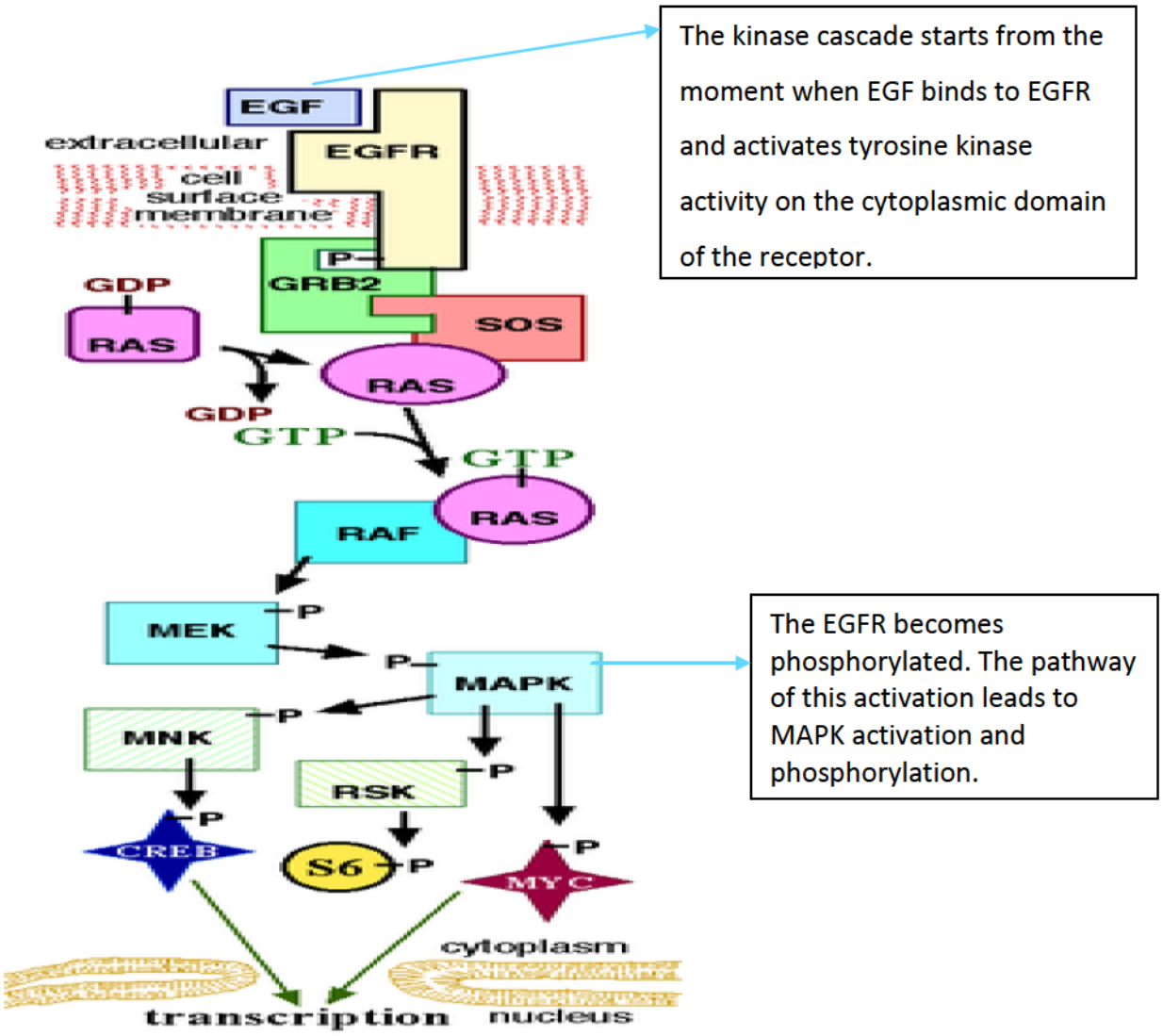


Figure 93 The key components of the MAPK/ERK pathway (adapted from JWSchmidt).

Beginning of the MAPK/ERK pathway cascade requires a signal from the receptor on the surface of the cell to be transduced to the DNA in the nucleus. The signal starts when EGF binds to the epidermal growth factor receptor (EGFR) on the cell surface and ends when the DNA in the nucleus expresses a protein and produces some change in the cell. Binding of epidermal growth factor to EGFR activates the tyrosine kinase on the cytoplasmic surface. On this stage, two processes occur, phosphorylation and dimerization of the receptor. Phosphorylation is the signal for binding other cytoplasmic proteins. Where the dimerization occurs, it allows the formation of high-

affinity, binding sites important for efficient phosphorylation of the protein that can be in low abundance within the cell.

On occupation by its ligand EGF, the EGF-R forms a dimer and this induces a change in the confirmation of the cytoplasmic domain that reveals its latent tyrosine protein kinase activity. This allows auto-phosphorylation of certain tyrosine residues on the dimerized receptor molecule. The dimerized phosphorylated molecule constitutes the catalytically active receptor. The Grb2 protein attaches to the tyrosine-phosphorylated receptor (EGFR) through its SH2 domain, while SH3 domain is involved in the activation of RAS protein. Activated RAS associates with protein kinase RAF (proto-oncogene serine/threonine-protein kinase), and subsequently activates MEK (Mitogen-activated protein kinase). The MEK kinase phosphorylates ERK (MAPK) on tyrosine and threonine residue. Dimerization of ERK (MAPK) exposes a signal that allows this MAP kinase to interact with proteins that promote its translocation into nucleus. The result is altered transcription.

7.4 Epidermal growth factor (EGF)

Stanley Cohen of Vanderbilt University along with Rita Levi-Montalcini received the Noble Prize in 1986 for the discovery of EGF described as a contaminating factor that causes precocious growth in epidermis (Cohen, 1958, 1964).

The EGF that binds and activates EGF receptors (EGFR) belongs to class of first messengers, as it initiates intracellular activity. It stimulates growth, initiates apoptosis, differentiation, and gene expression. Growth factors (including EGF) are involved in tumour maintenance and growth and many tumour cells therefore express high levels of EGF-R. Therefore, recombinant antibodies to such receptors are promising leads in cancer biology. If an activation is detected it can suggest that the protein will be able to internalize into cells and any attached cargo will have an effect in the cell. The experiments were designed to check and measure the phosphorylation of EGFR binding and EGFR-Ab BoNT fusions.

7.5 Results

The aim of this study was to measure activation and phosphorylation of EGF. EGF was biotin labelled and become biotinylated to reduce its aggregation in solution. A biotin was brought from Pierce, UK. Molecule was tested on A549 cells, which carry receptor for EGF. A549 cells are adenocarcinomic human alveolar basal epithelial cells.
The A549 cells were grown on Petri dishes at 37°C in the incubator. EGF was diluted in media to 100 nM and 1 nM. The dilutions were prepared in series to obtain accurate values. 10 ml of media for each dilution of EGF was added to three plates. All plates were incubated at 37°C for: 5, 10, and 45 minutes. An additional control plate with 10 ml of media only was incubated for 45 minutes. The control was a plate with cells lacking an EGF ligand. The phosphorylation process was stopped on ice after each incubation stage for each dilution. Subsequently a Cell Lysis Buffer was applied to collect cells for analysis. The concentration of protein was determined by BCA assay to ensure the same amount of protein was loaded onto the gel for each sample. A control sample with no added ligand was included on each gel.

The outcome of an activation experiment using EGF was interpreted by Western blots against β-actin, EGFR, phospho-EGFR, MAPK, and phospho-MAPK. The β-actin Western Blot (Fig 94) detected equal amount of the protein loaded onto the gel. The MAPK (Mitogen activated protein kinase) (Fig 95) and p-MAPK (phosphorylated Mitogen activated protein kinase) (Fig 96) blots detected bands with expected size of 42-44 kDa. However, there was detection of undesirable MAPK and p-MAPK in the control sample that indicates the presence of EGF in the cells or suggests that the assay had not worked. Both the concentration of EGF added to the cells and the length of incubation has impact on the signal. The highest concentration of EGF 100 nM kept for 45 min gave the lowest signal. It can be due to cell death, therefore the cell did not respond to the presence of ligand. The Western blot against EGFR (Fig 97) should reveal a band around 175 kDa, which was noticeable in reduced visibility. The same experiment was set up using LH_N/ C-EGF fusion protein (product of Syntaxin Ltd). The data were not shown, as they similar to that of EGF.

The presence of EGF in the control sample did not allow me to validate activation and phosphorylation process, even if detection of MAPK/p-MAPK and EGF/p-EGF in the course of Western blots was possible. The experiments will need to be repeated with media free of EGF to establish robust assay for testing of EGFR-Ab BoNTs.

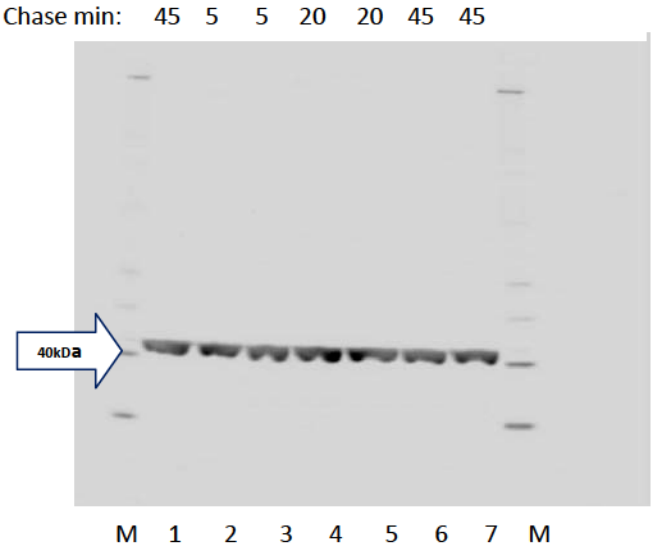


Figure 94 Western blot against β actin.

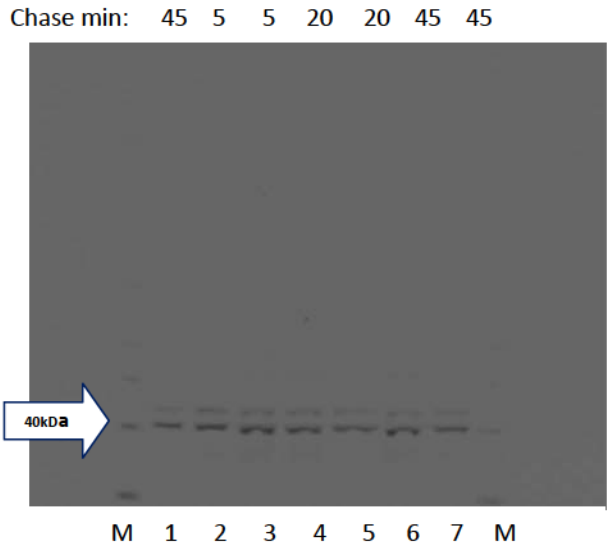


Figure 95 Western blot against MAPK.

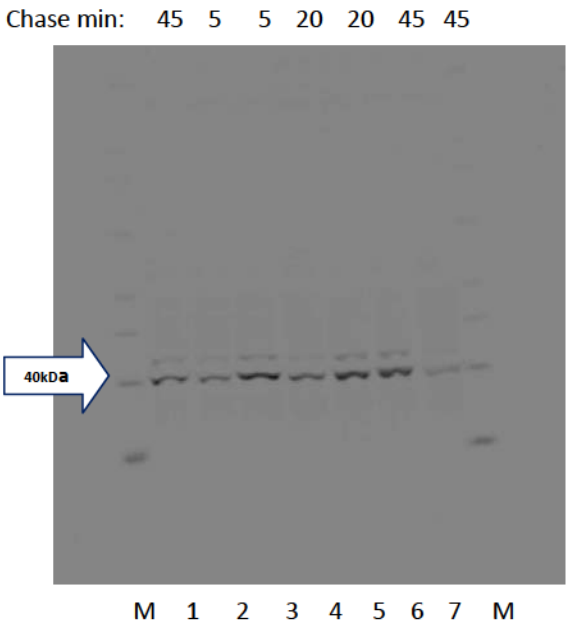


Figure 96 Western blot against phosphorylated MAPK.

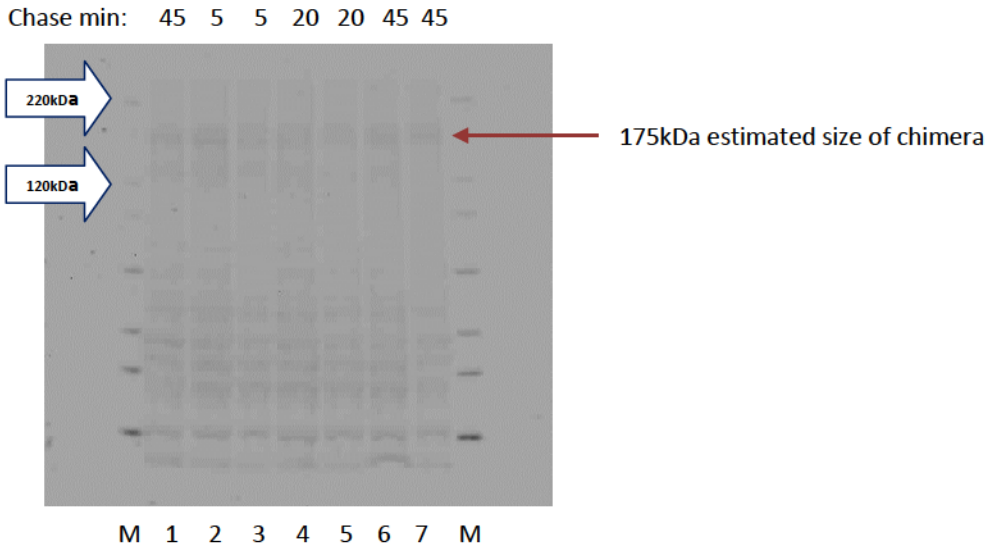


Figure 97 Western blot against EGFR.

This legend applies to figure 94, 95, 96 and 97.
10^6 A549 cells were added to each Petri dishes and 1nm or 100nm of EGF was added for 5, 20 and 45 minutes before samples were taken and extracts prepared for SDS PAGE and Western blotting. A BCA assay was used so that, equivalent amounts of protein were loaded per lane (40µg). The developed β actin blot revealed similar loadings.
Lane 1 is a control sample without addition of EGF; Lanes 2, 4 and 6 show test sample with 1nM of EGF taken in the different duration of time; Lanes 3, 5 and 7 show test samples with 100nM of EGF taken in the different duration of time;

Chapter 8

Internalisation Studies

Chapter 8. Internalisation Studies

8.1 Endocytosis

Internalization (endocytosis) of macromolecules into a cell can occur by a variety of mechanisms. Pathways and can be subdivided into four categories: clathrin-mediated endocytosis, caveolae, macropinocytosis, and phagocytosis (Fig 98).

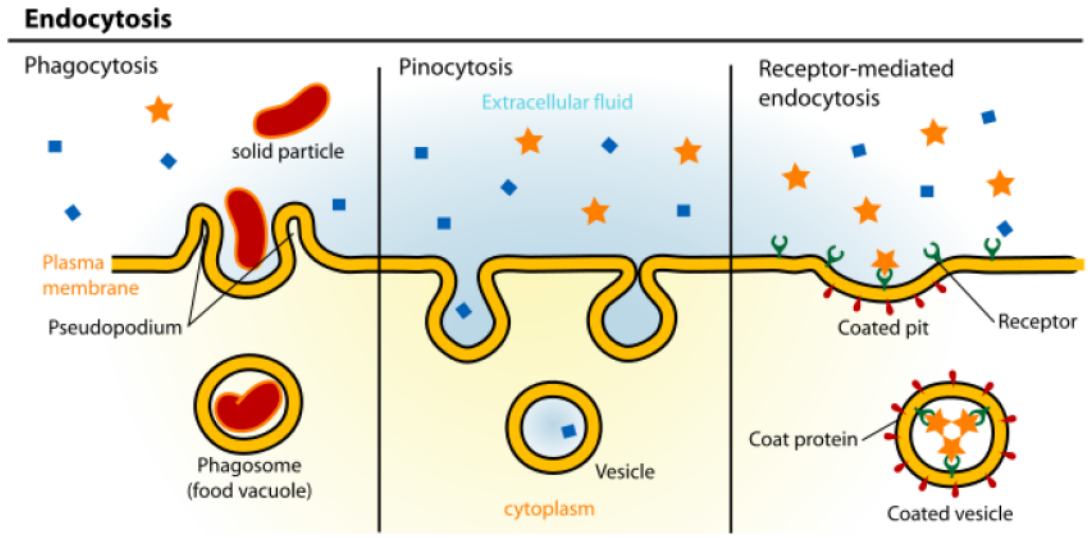


Figure 98 Endocytosis pathways (adapted from Cell Biology, Wikipedia 2009).

However, for this study I am most interested in the process of clathrin-dependent endocytic pathway for eukaryotic cells, because BoNTs are known to enter this way. Receptor-mediated endocytosis (RME) can be clathrin-dependent or clathrin independent at high ligand doses. RME is a process by which cells internalize proteins in clathrin-mediated endocytosis. The internalization takes place only upon the formation of endocytic clathrin-coated vesicles, which occur through the interactions of cytosolic adaptors and clathrin proteins with cytosolic domains of ligand-bound receptors on the plasma membrane. The biogenesis of clathrin-coated vesicle is best characterized in nerve terminals, where it serves as the major pathway for recycling of synaptic vesicle components after release of neurotransmitter in response to action potentials. The correct sorting of endocytic ligands and receptors is essential for many cellular activities. A primary function of the endocytic system is to sort internalized ligands and

receptors to different destinations, where they can be targeted to different organelles for processing or directed to specific cellular locations for function (Lakadamyali, 2006).

Epidermal growth factor (EGF) primarily enters cells via clathrin-mediated endocytosis when added at low concentration (Carpenter and Cohen, 1979). At high concentrations, some EGF molecules are diverted to other pathways such as cholesterol dependent caveolae and clathrin independent pathway (Jiang and Sorkin, 2003; Sigismund et al., 2005). Most probably, EGF first enters a common pool of early endosomes where sorting takes place (Dickson et al., 1982; Ghosh et al., 1994). Receptor tyrosine kinases such as the EGF receptor possess intrinsic protein tyrosine kinase activity. Ligand binding causes a change in the conformation of the monomeric receptor that leads to non-covalent dimerization of two receptors (Lakadamyali, 2006). Dimerization facilitates tyrosine auto-phosphorylation, as described in Chapter 7. Ligand-induced internalization of the EGF receptor has been shown to depend on intrinsic tyrosine kinase activity, but receptor auto-phosphorylation is not required for endocytic activity because mutant receptors lacking any auto-phosphorylation sites can undergo ligand-induced internalization (Lakadamyali, 2006). Other sites in EGFR interact with AP-2 adaptors. These in turn recruit clathrin to form a coat to promote budding. Following clathrin unseating, the vesicles fuse with endosomes. Even at low pH, the receptors do not dissociate from ligand. The complexes can be recycled if they do not become trapped in the intralumenal vesicles of multivesicular bodies (MVB), the MVB is degraded (Sorkin and Goh, 2008).

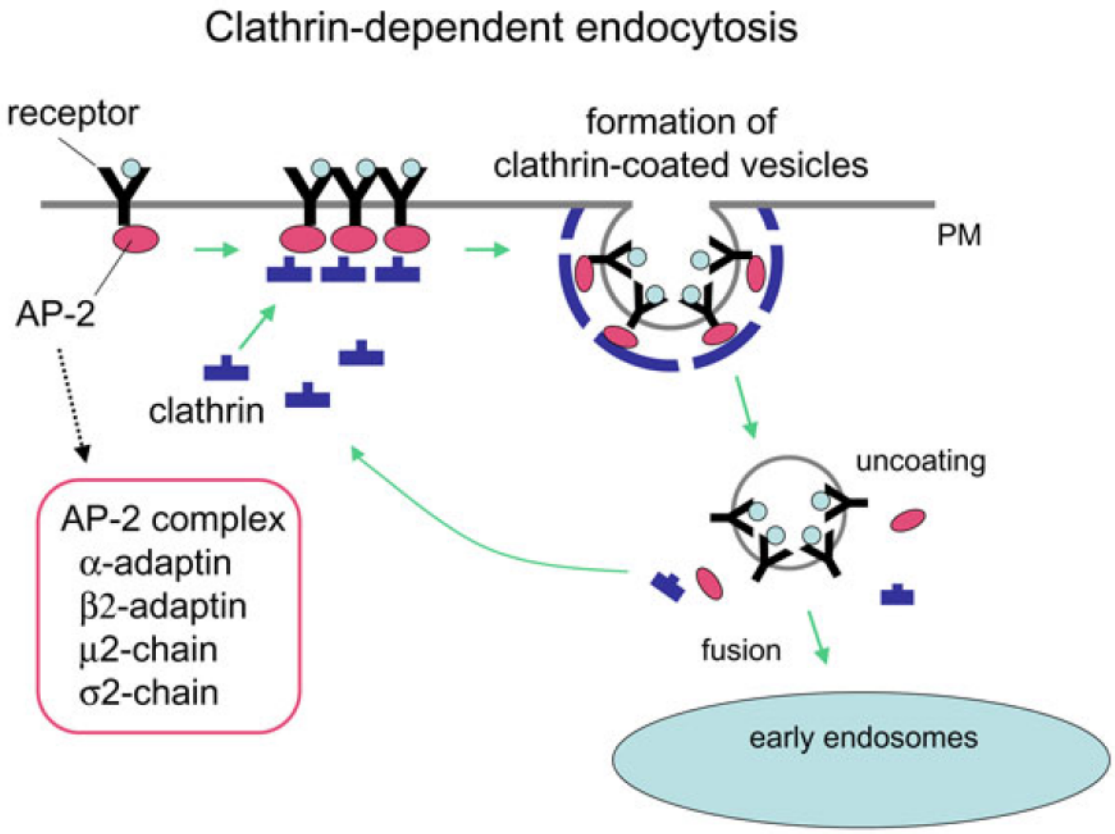


Figure 99 Pathway of Clathrin depended Endocytosis (adapted from Grant and Sato, 2005).

Mechanism of clathrin-dependent endocytosis. Clathrin and cargo molecules are assembled into clathrin-coated pits on the plasma membrane together with an adaptor complex called AP-2 that links clathrin with transmembrane receptors, concluding in the formation of mature clathrin-coated vesicles (CCVs). CCVs are then actively uncoated and transported to early/sorting endosomes.

8.2 Exocytosis

The successful internalization of LH_N based chimeras lead to the inhibition of synaptic vesicles to blocks exocytosis that is the opposite process to endocytosis.

As mentioned Chapter 1 specific Neuronal complexes termed soluble *N*-ethylmaleimide-sensitive factor (NSF) attachment protein (SNAP) receptor or (SNARE) proteins are essential to mediate the release of neurotransmitter contents, such as acetylcholine from synaptic vesicles (Montecucco, 2005). The cytosolic synaptic vesicle (SV) binds and fuses with presynaptic nerve terminal membrane to release its

neurotransmitter contents into the synaptic cleft (Hua and Scheller, 2001, Han *et. al.,* 2004).

The pre synaptic SNARE proteins are the specific substrates of the seven-botulinum neurotoxins: BoNT/A-G. The assembly of specific multimeric SNARE complexes is crucial for all stages of exocytosis. Delivery of zinc-endopeptidase N-terminal domain LC (light chain) of BoNT to the cytosol prevents neurosecretion by cleavage of one of SNARE proteins that are essential for formation of a functional SNARE complex between synaptic vesicle and plasma membrane (Montecucco, 2005).

The three SNARE proteins associate into SNARE complex through interaction of coil domains between four helices. The coiling process continues until the formation of ~12nm long four helices bundle with a parallel orientation is accomplished and ready to membrane fusion (Fig 100) (Lin and Scheller, 1997). The SNARE complex has a *trans* configuration with charged surface, what enable specific protein-protein interaction. (Jahn *et. al*., 2003). This complex brings the vesicle and plasma membrane into close contact to initiate membrane fusion, which in the case of neuronal cell of synaptic vesicles indicates the final step of neuroexocytosis (Weber *et. al.,* 1998).

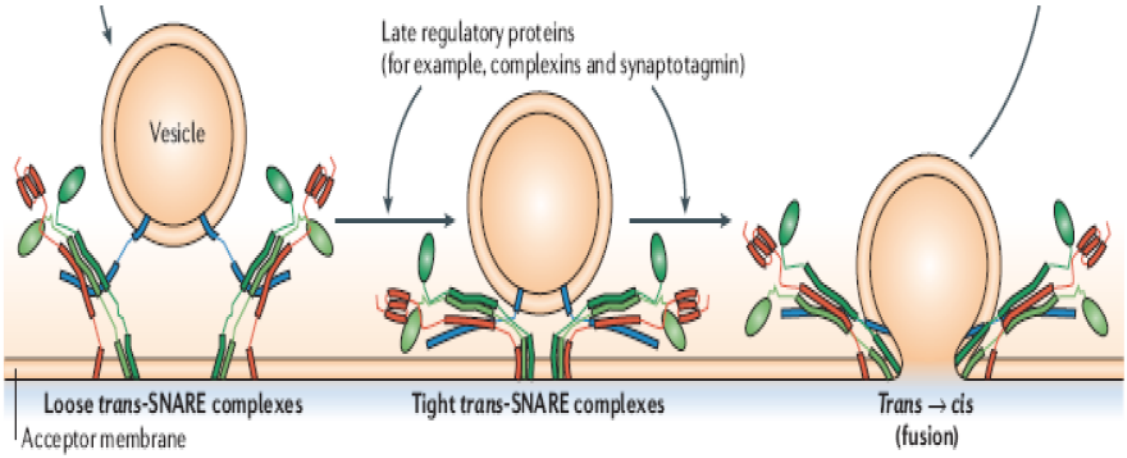


Figure 100 Secretion SNARE mediated fusion.

The interaction of H_C domain with specific receptors (polysialyl, gangliosides and protein receptor) on the neuronal surface permits BoNTs to bind to target motor neurons. The conformational change in the toxin and formation of ion channels by H_N domain takes a place in an endosomal compartment after internalization (Foster and

Chaddock, 2007). The ion channels permit the light chain (LC) to reach its substrates (Oblatt-Montal *et. al.*, 1995).

There are numerous opportunities for the development of therapeutics based on botulinum neurotoxin fragments; therefore, inhibition of secretion for extended periods by retargeted LC domain is one of main approaches (Chaddock and Marks, 2005). The inhibition of secretion can only take place when the light chain successfully internalized into the cytosol.

8.3 Results

8.3.1 Internalization assay

The purpose of this experiment was to measure internalization of biotinylated EGF and LH_N/ C-EGF (SXN100501) fusion proteins. The internalization assay using biotinylated proteins was based on the information obtained from two sources of experiments:

- *Irie et* al. Nat Cell Biol 2005 – showed an increase in the internalization of biotinylated transferrin in the presence of ephrinB2 in a EphB2R transfected cell line
- *Bronfman et* al. J Neuroscience 2003 – showed internalization of biotinylated NGF into PC12 and nnr5 cells

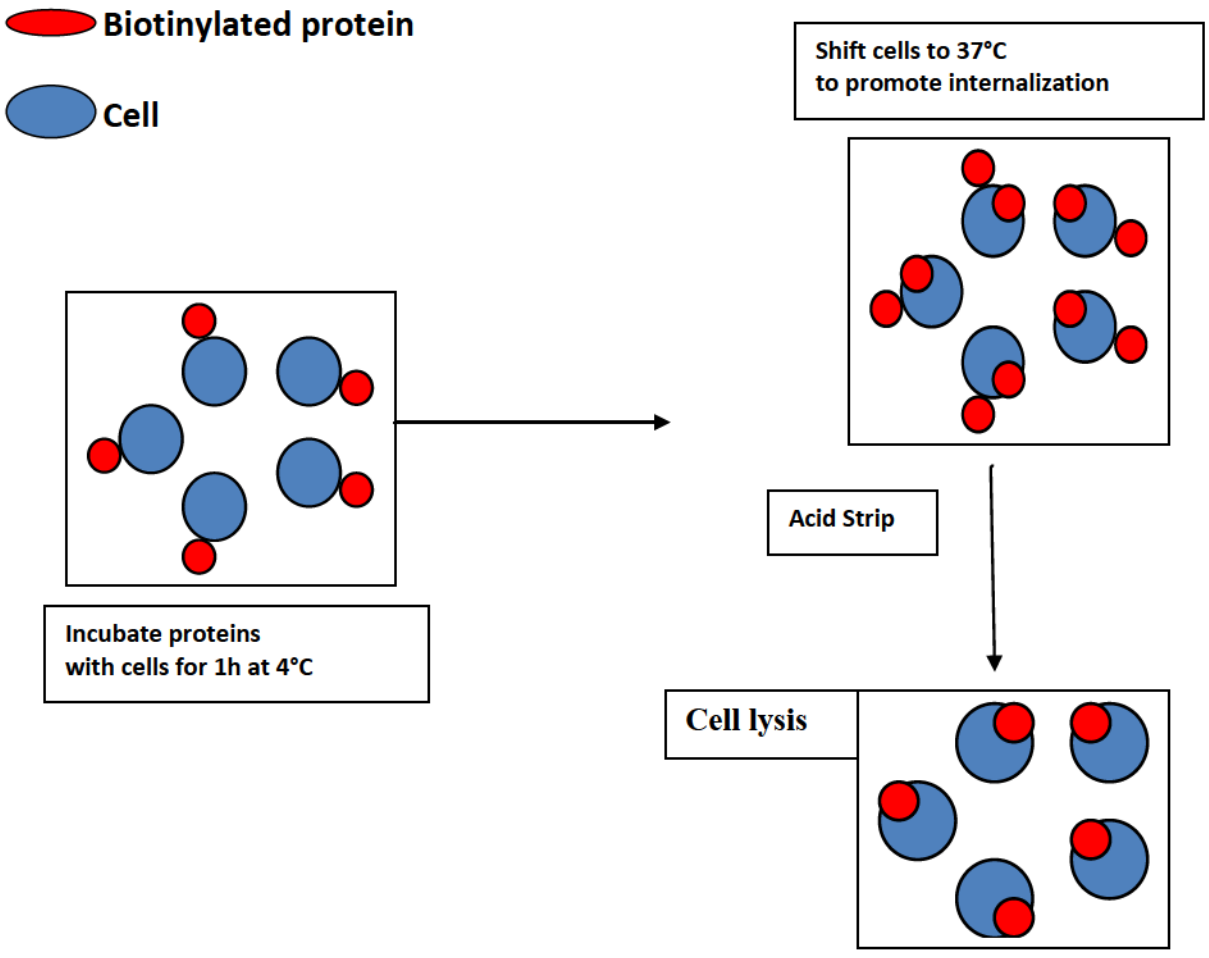


Figure 101 Method for internalization assay.

The cells used for this experiment are (1-5 × 10^6) A549 cells that were treated for 2-180 min hour with either biotinylated EGF or biotinylated LH_N/ C-EGF fusion protein (300-600 nM). The experiment was performed at 37˚C in a water bath or at 0˚C on the ice. Four samples in triplicate were prepared for each tested construct. Additionally two controls lacking the addition of construct were kept at both temperatures. All samples in triplicates were run on SDS PAGE. The control samples at 37˚C and 0˚C as well as untreated test samples were run on each gel. The biotin marker was used for determination of bands size. The graphical representation of quantified data was generated using a Prism programme (Fig 102 and 103). The internalization of biotinylated EGF and LH_N/C-EGF was not detected; the represented data can suggest that EGF or LH_N/C-EGF become bound to the cells. It will be necessary to optimize the signal for internalized molecules. The graphs below represent an example of internalization experiments prepared with the biotinylated EGF and LH_N/C-EGF.

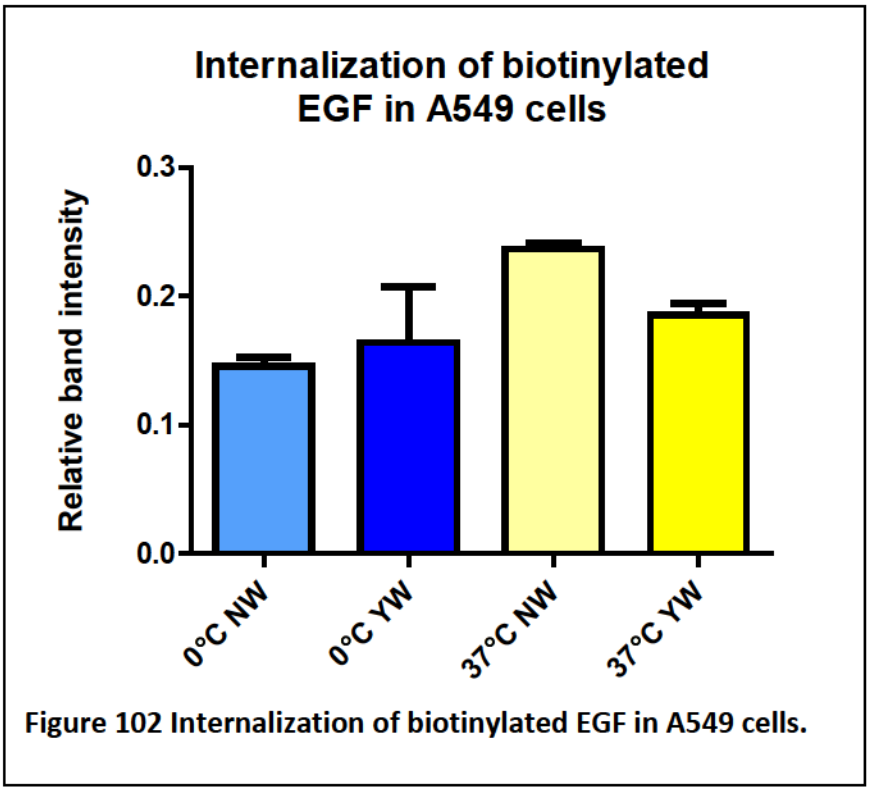


Figure 102 Internalization of biotinylated EGF in A549 cells.

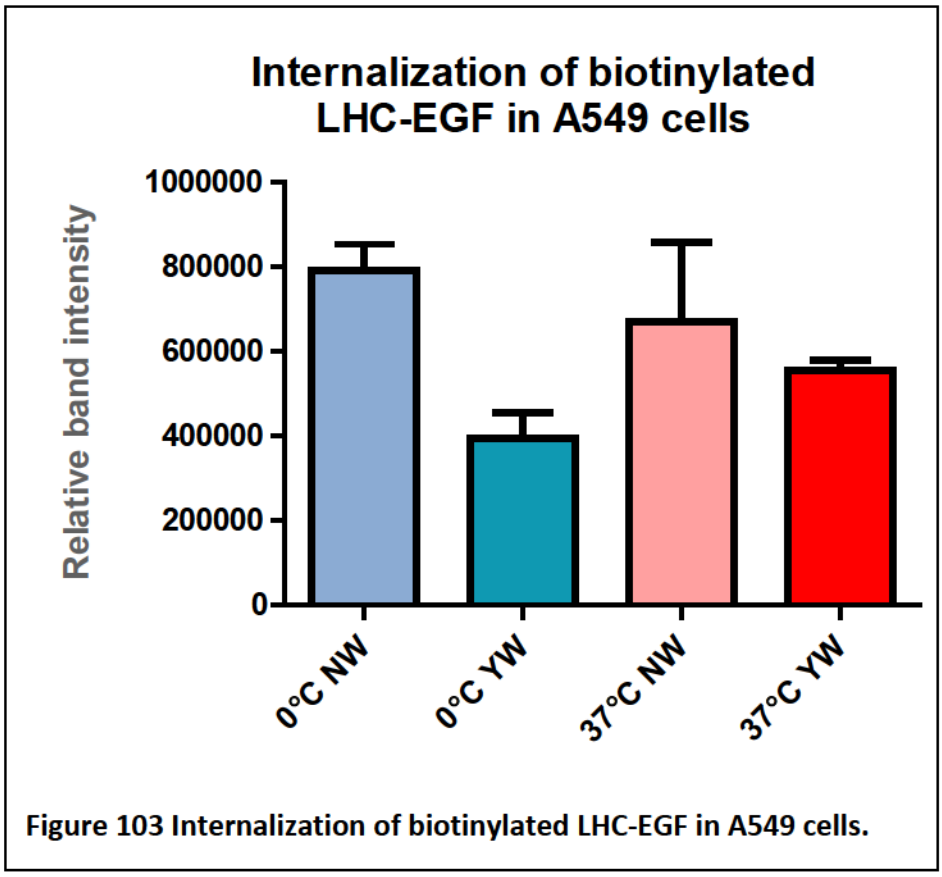


Figure 103 Internalization of biotinylated LHC-EGF in A549 cells.

Legend below corresponds to graphs represented above; (Fig 102 and Fig 103)

The 300-600nM was added to 1-5 × 10^6 A549 cells;

The samples were divided to two groups and incubated at 37˚C or 0˚C for 2-180 min;

Cell were washed in cold PBS containing 1 mM $MgCl_2$ and 0.1 mM $CaCl_2$ three times to remove all unbound biotinylated fusion protein LHC-EGF or biotinylated EGF. On this stage, cells from each temperature group were divided on half. One-half was acid stripped by 50 mM Glycin pH 5.0, 450 mM NaCl (YW – wash with acid) and the other was not (NW – no wash with acid). The acid strip should remove all biotinylated fusion protein LHC-EGF or biotinylated EGF bound to the surface of the cells, so a biotinylated fusion protein LHC-EGF or biotinylated EGF that were internalized inside the cells will remain present. Subsequently, the lysis buffer with cocktail protease inhibitors lysed cells and supernatant was pooled with Streptavidin magnetic particles. The samples were resolved by SDS PAGE and blotted against Streptavidin-HRP in 1/2000 dilution. The relative band intensity corresponds to total bound biotinylated fusion protein LHC-EGF or biotinylated EGF at 0˚C or 37˚C without stripping.

Chapter 9

Conclusions

Chapter 9. Conclusions

This study was set out to investigate whether antibodies and antibody fragments could be successfully engineered onto botulinum toxin fragments to create novel targeted proteins. The studies have focussed on six key areas, which Chapters 3 to 8 have described in detail. The conclusions from this work are listed below.

DNA encoding twenty-six novel recombinant multi-domain proteins have been successfully designed, cloned and transformed into a suitable *E. coli* expression host (Table 10, pg 125-126).

The *E. coli* bacterial expression system selected for this study (T7 promoter; BL21 (DE3) host) has been successfully used for the generation of novel, soluble recombinant fusion proteins comprising fragments of botulinum neurotoxin and antibodies (Chapter 5).

Where scFv were cloned onto the LH_N backbone (Chapter 4), this study focussed on the use of a Gly-Ser repeat spacer (GS15) between the variable light and variable heavy domains (Section 3.1.4). This has resulted in the successful expression of full-length fusion protein (Fig 52 and 53), of which ones truncations were observed during the imidazole elution phase of the IMAC affinity chromatography (Refer to figure 54, 57, 59, 61, 62, 65, 66 and 67 in the section 6.4.1 ‚Purification of single chain variable antibody fragments-LH_N chimera'). Unfortunately, attempts to improve the recovery of full-length protein using alternative chromatographic procedures such as an ion-exchange chromatography (Q-Sepharose) (Figure 61 and 63) or hydrophobic interaction chromatography (HIC-phenyl Sepharose) (Figure 64) were unsuccessful.

To explore the potential impact of serotype and activation enzyme on the ability to create single chain antibody-LH_N molecules, four novel constructs were designed incorporating scFvCD117I with one of LHA, LHB or LHC and with one of FXa or EK activation site (Table 10). Most often, the activation of the single chain antibody-LH_N was unsuccessful despite the use of different conditions. However, the best results were achieved from the activation of LHC-scFvCD117I with 1.5U of Factor Xa per 1 mg of fusion protein at 25˚C, which achieved almost completed activation (Figure 66 and 67).

Truncations of LHC-scFvCD117I were investigated by an N-terminal sequencing (Figure 68, 67, 69, 70 and 71). The truncation appeared to be in two places within the primary structure of the protein. One was after the heavy chain of the LH_N/C backbone but before the start of the single chain antibody and another one was after the heavy chain of scFv but before the start of the intra-domain GS15 linker. It is important to note that a full-length product was detected using an N-terminal sequencing.

Expression of the single domain antibody fragments with LH_N A (Figure 49), B (Figure 50) and D (Figure 51) was successful. The purification of six single domain antibody-LH_N chimeras was completed successfully (Table 12). Full-length material of high purity and yield was achieved by using affinity chromatography, incorporating a step gradient elution scheme. This study investigated the use of an additional HIC column to improve purity of the final material. However, the data indicated that the purity of LH_N/A-EGFR31sdAb=SXN101872 purified on HIC column was not significantly enhanced (Figure 84) in comparison to LH_N/A-EGFR6sdAb=SXN101868 purified on His column. The quality control SDS PAGE gels of these molecules shows high purity (Fig 85 and 86). In one case (Figure 78 and 79), the quality of the protein did not meet criteria established within Syntaxin Ltd for *in vitro* cellular assay. This was due to non-specific cleavage within the LH_N component.

The choice of the successful candidate depends on many aspects, but one of the most important is that activation should be above 85%. The following molecules, **LH_N/A-EGFR6sdAb=SXN101868,** **LH_N/B-EGFR31sdAb=SXN101865**, **LH_N/D-EGFR31sdAb=SXN101878**, **LH_N/B-EGFR6sdAb=SXN101861**, **LH_N/D-EGFR28sdAb=SXN101877**, and **LH_N/A-EGFR31sdAb=SXN101872** were categorized in descending percentage of molecule cleavage, respectively: 100%, 100%, 88%, 88%, 76% and 66%. Although the cleavage of the last two molecules is not sufficient according to the required criteria the chimera **LH_N/A-EGFR31sdAb=SXN101872** exhibits 100% of purity, therefore if the effectiveness of activation would be increased this construct could be a potential candidate.

The antibody sequences for the construction of the recombinant chimera comprising single antibody domains were obtained from patent WO/2007/127317. In a number of cases, the protein characteristics of the chimera have not matched the predicted

properties of the domain antibody alone, which were described in the patent. It appears that a compromise must be reached between the botulinum neurotoxin serotype and choice of the antibody fragment in order to find a high quality of target protein.

An assay for detection EGFR activation through tyrosine kinase was attempted. Activation of EGFR in E549 cells was demonstrated using EGF (Sigma) (Figure 94, 95, 96 and 97), but the presence of pMAPK in control cells suggests that further optimisation is required for example by use of growth factor free media. It was also possible to detect MAPK phosphorylation on the Western Blots but it was impossible to confirm that EGF bound to EGFR on the cells. This assay could be used in the future to investigate whether the antibody-LH_N molecules can activated the receptor once the controls are satisfactory.

Chapter 10

Discussion

Chapter 10.Discussion

Antibody fragments and their recombinant derivatives have important practical applications in research, diagnostic and therapy. Utilization of antibodies as drug delivery vehicles has become a successful application. There are, however significant manufacturing implications with conventional antibodies. For most applications, high-yield production, solubility, stability and small size are critical factors. Attempts to reduce size of antibodies, while retaining its antigen-binding properties have been reported. It resulted in a production of antibody fragments constructs, such as single chain antibody (scFv). New opportunities arose with the discovery of heavy chain antibodies (hcAbs) in camels in 1993. The use of single domain antibodies has been rapidly growing for biotechnological applications. The deep penetration into tissues and rapid elimination via the kidney make antibody fragments and single domain antibody favourable tools for the delivery of cytotoxic agents (Holt *et. al.*, 2003).

This study serves as the first demonstration that the antibody fragments and single domain antibodies can be engineered into the LH_N domains of recombinant BoNTs, subsequently expressed in an *E. coli* system and purified. In order to choose the best antibody candidate it will be important to perform additional experiments on cells and pharmacology assays in order to establish a functional characterization of novel compounds.

During this study there have been a number of observations and occurrences that are worthy of further discussion. For example, during the cloning phase, the recombinant scFv and sdAbs DNA sequences obtained from Entelechon were incorporated into the LH_N backbone DNA using restriction enzymes XbaI and *HindIII*, though successful method is susceptible to frame shift. It was required to introduce 'A, base after XbaI side (TCTAGA) to avoid the frame shifting during cloning.

Truncation of the scFv-containing constructs was a major observation of this study. It could be hypothesised that the simplicity of single domain antibody structure compared to scFv enhanced production of a full-length product, as the stability of the scFv depends on the VH-VL domains and their interaction with the flexible linker. A

number of the scFv sequences were obtained from patent WO/2007/127317, in which the linker was reported to be EEGEFSAR. In this study, it was decided to replace this linker with a flexible GS15 (GGGGSGGGGSGGGGS) linker as this had been found to give the best results during screening tests done on the LH_N chimeras with different linkers lengths (Data obtained from Syntaxin Ltd databases). It is possible that the truncations observed within scFvs could be due to the different linker design than that proposed in the patent WO/2007/127317 (linker EGGEFSAR), but this hypothesis needs to be checked by performing the additional experiments. One experiment would be to substitute the GS15 linker with the linker EEGEFSAR proposed in the patent. In addition, removal of the spacer from LH_N-ligand TSIs (Fig 23 and 24) between the heavy chain and the ligand could potentially cause truncation of the full-length protein. The reorganization of the single chain antibody structure is a necessity in order to progress any additional purification experiments since attempts to resolve truncation issues during the purification process were ineffective, even after substitution of the first His chromatography column for an ion exchange or a hydrophobic interaction chromatography. From these results, it was noted that the detailed engineering of the scFv antibody construct requires further exploration in order to be optimise in the context of an LH_N fusion.

These studies have also given valuable insight into the insights of recombinant LH_N fusions. For example, the single domain antibody - LH_N molecules have shown higher purity than many of the routinely produced LH_N-ligand (TSI) molecules. In order to purify routinely made TSIs of high purity the additional changes have been made to the standard method (attached in the chapter 2 methods and materials), one of them is use of the 50 mM Hepes pH 7.2, 500 mM NaCl buffer during cell lysis and column steps. The buffer with higher salt concentration eliminates non-specific contaminants from *E. coli* expression system. The use of high salt for the purification of LH_N-single domain antibody molecules was not required, as the final purity in presence or without 50 mM Hepes pH 7.2; 500 mM NaCl buffer gave the same results. It looks like during the expression of the novel designed molecules the occurrence of protein contaminant from *E. coli* system is lower in compare to purification of LH_N molecules in the presence of ligands other than the single domain antibody fragments. The production can be similar, but removal of contaminants during the beginning of purification seems

to be much more straightforward. There can be a number of reasons for this occurrence; however, the element, which is missing from the typical C terminal structure, is the spacer between the heavy chain of recombinant BoNT (Fig 24) and the presence of single domain antibody, which simplifies the structure of novel molecule. It has been reported that some *E. coli* proteins form aggregates with recombinantly expressed target protein and such a complex can be carried throughout purification or can be disassociated by different methods. The orientation of novel molecules could have an impact on the level of complexes aggregating, which are carried over during the purification. This hypothesis could be checked through purification experiments based on the novel LH_N single antibody chimeras and new design LH_N molecules with ligand lacking spacer from the side of heavy chain.

Only the production of the single domain antibody-LH_N molecule **LH_N/B-EGFR43sdAb=SXN101866,** so far was unsuccessful. This molecule was purified with a very low yield after the first affinity column and did not cleave with protease (Fig 78 and 79). According to the patent WO 2008/141449 A1, EGFR43sdAbs gave the best results after production in *E. coli* and purification on the IMAC column, the final yield being 42mg per 1-litre of culture. There can be a number of reasons for this dissimilarity of data. There are two aspects worth considering. Firstly, the single domain antibody EGFR43 fragment in the presence of LH_H molecule can obstruct some of the properties, which influence the protein production. On the other hand, the substitution of the expression vector from pSJF2 to pK7 can be a reason for the lower yield of the soluble protein, which did not apply to other tested molecules, however it is less likely a case, especially that it did not have an effect on other single domain antibody molecules. It will be worth to consider preparing the N-terminal sequencing to confirm correct sequence start point of protein and purifying EGFR43sdAb in the presence of serotype A and D.

The single domain antibody LH_N molecules ranked in order of purity from the highest to the lowest in the following order, **LH_N/A-EGFR31sdAb=SXN101872**, **LH_N/A-EGFR6sdAb=SXN101868**, **LH_N/B-EGFR31sdAb=SXN101865**, **LH_N/B-EGFR6sdAb=SXN101861**, **LH_N/D-EGFR31sdAb=SXN101878**, and **LH_N/D-EGFR28sdAb=SXN101877**. Their outstanding purity qualified five of them to be used

for a crystal generation that leads to X-ray structural studies. The idea was incorporated from the article: A Crystal Structure of the anti-His tag antibody 3D5 single-chain fragment complexed to its antigen (Kaufmann et. al., 2002), where anti-His scFv crystals are proposed to act as a framework for the crystallization of His-tagged target proteins difficult to crystallize. The authors used a scFv antibody fragment for the purpose of this study, but since the single domain antibody represents high quality, it was worth the attempt to determine the single domain antibody-LH_N structure. The successful outcome of crystallization is a combination of luck, total protein concentration and high purity, therefore protein purification and concentration need to be accurate as possible. The reading A_{280nm} measures tyrosine and tryptophan, in the presence of unwanted proteins often this reading is inaccurate. The reliable data can be obtain by BCA or Bradford assay. According to data provided in chapter six, there was not much discrepancy in the purity of molecules between two assays (Table 12). This indicates that the novel molecules have high purity and the presence of *E. coli* proteins is not significant. Experimental studies are ongoing, but the evidence that crystals have been formed is confirmation for the quality of the proteins prepared in this study.

One of the aspects, which will be considered during the selection of antibody fragments, is functionality in cell assays, such as activation and internalization. The activation assay was well established, however needs to be improved in order to be used for the purpose of future studies, while the internalization studies performed as well as activation assays on EGF require introduction of different type experiments to make reliable conclusions. Initially novel molecules should be tested in an activation assay to show binding to the desired receptor to represent the detection of an activated signalling pathway within cells. If obtained, the protein can be tested in an internalization assay.

A number of valuable outcomes have emerged from a study that can be applied across Syntaxin's TSI (Targeted Secretion Inhibitors) platform. These include: (i) methods for improved activation of TSIs; (ii) improved yield following incorporation of an N-terminal His tag; (iii) the difference in stability behaviour between serotypes means

that multiple serotypes should be used in the initial screening phase; (iv) improved purity of purified molecules.

In summary, the ability to engineer the antibody-LH_N chimeras opens new prospects of research areas for Syntaxin Ltd. The novel molecules can target a wider spectrum of cells and therefore broaden the applications to different areas of diseases. Specific and selective targeting is a major goal in today's research, where one of the leading aims is to improve therapeutics against cancer. The identification of the antibody-LH_N chimeras effectively targeting EGFR could potentially treat solid tumors and other diseases. However, a number of obstacles still need to be resolved before clinical application of novel antibody-LH_N chimeras become possible.

List of obstacles:

- Redesign of single chain antibody in order to avoid truncations;
- Optimisation of activation assay on EGFR;
- Development of robust internalization assay;
- In view of this approach together with attributes of recombinant botulinum-neurotoxin chimeras the efficient delivery of therapeutic to target cells could be developed;

Appendix

Engineering functional recombinant proteins based on antibody domains and fragments of *C. botulinum* neurotoxin

The analyzes of triplet position vs. usage table and each codon vs. usage table (Chapter 3).

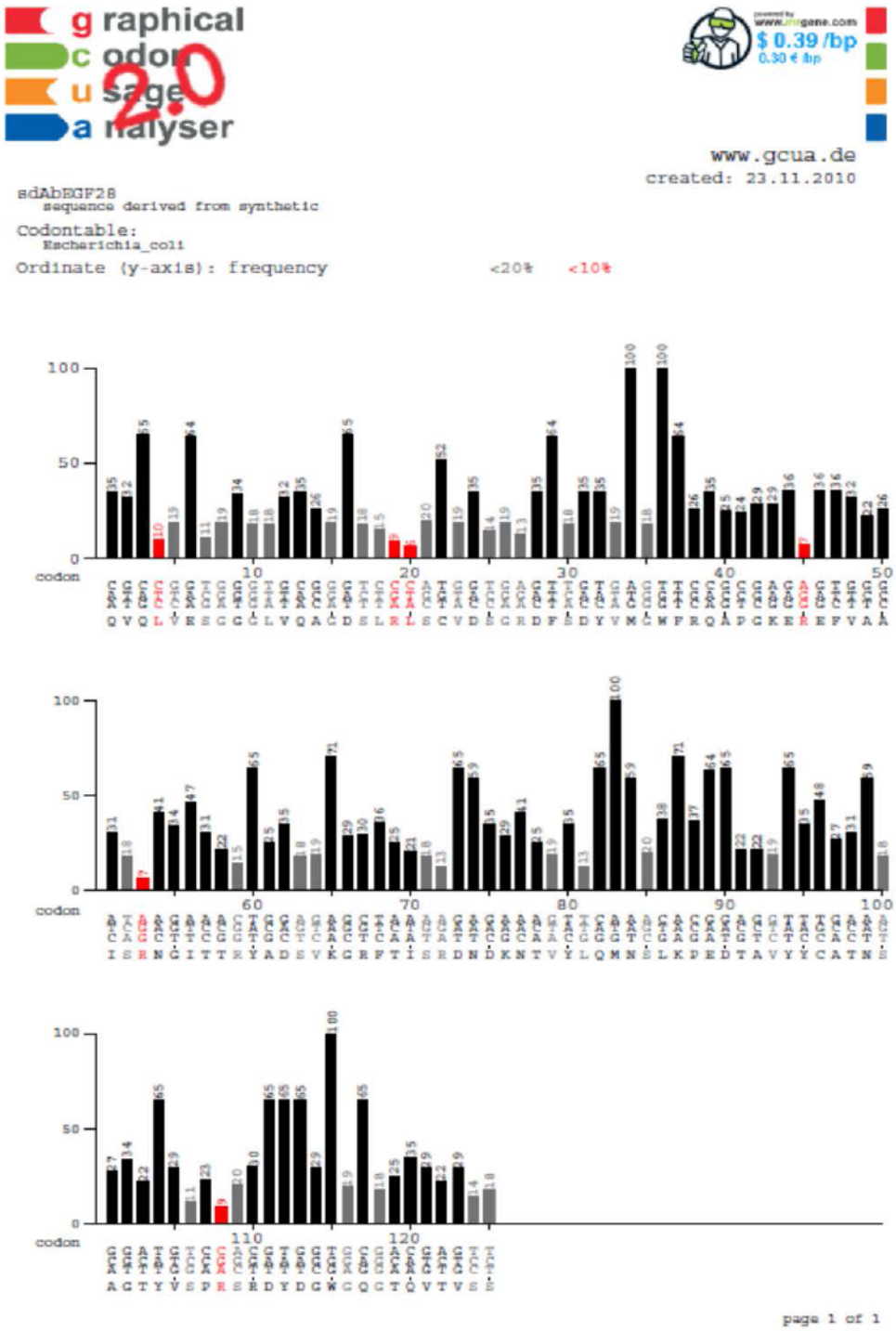


Figure 104 Each triplet position usage table sdAb EGF28.

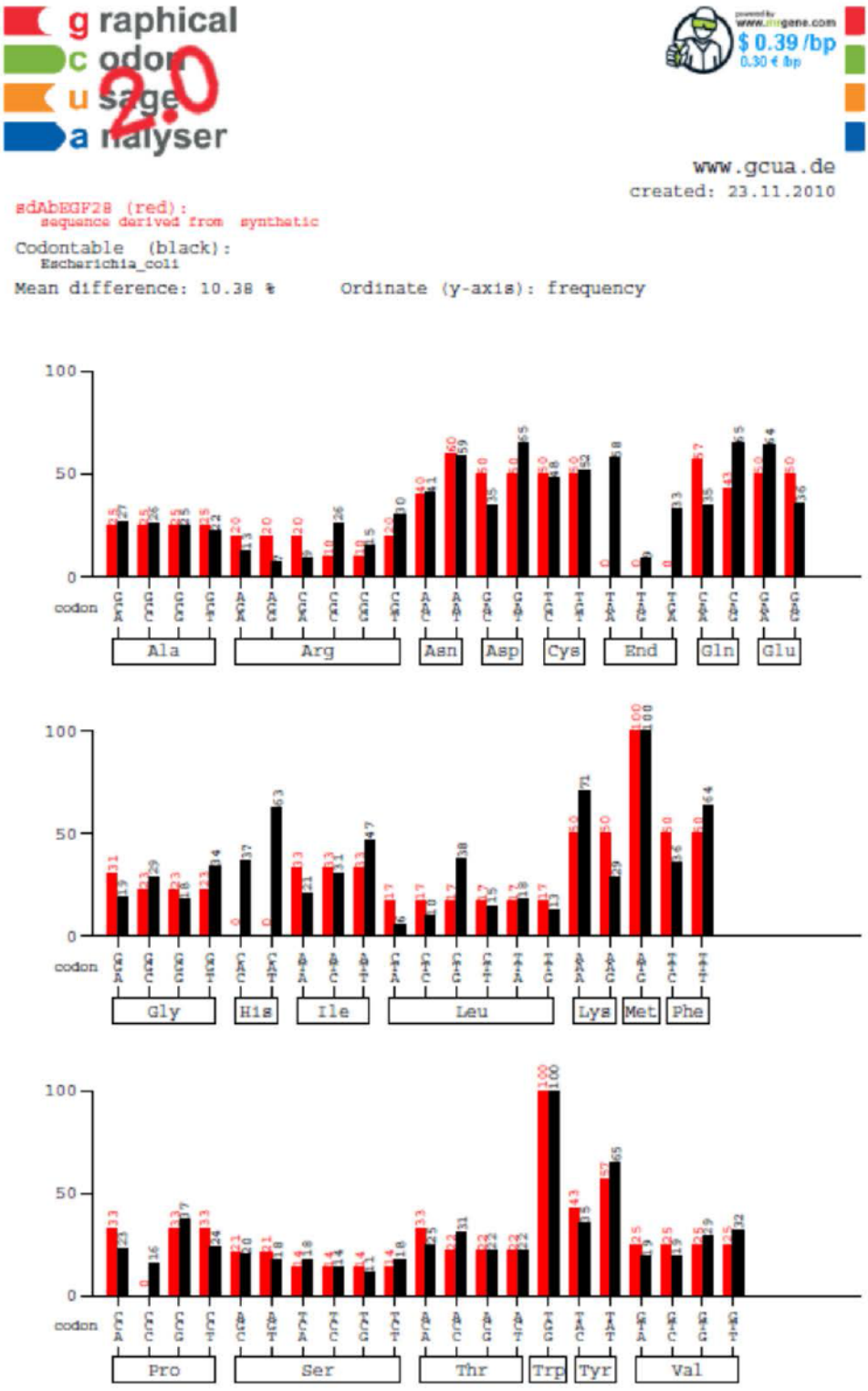


Figure 105 Each codon usage table sdAb EGF28.

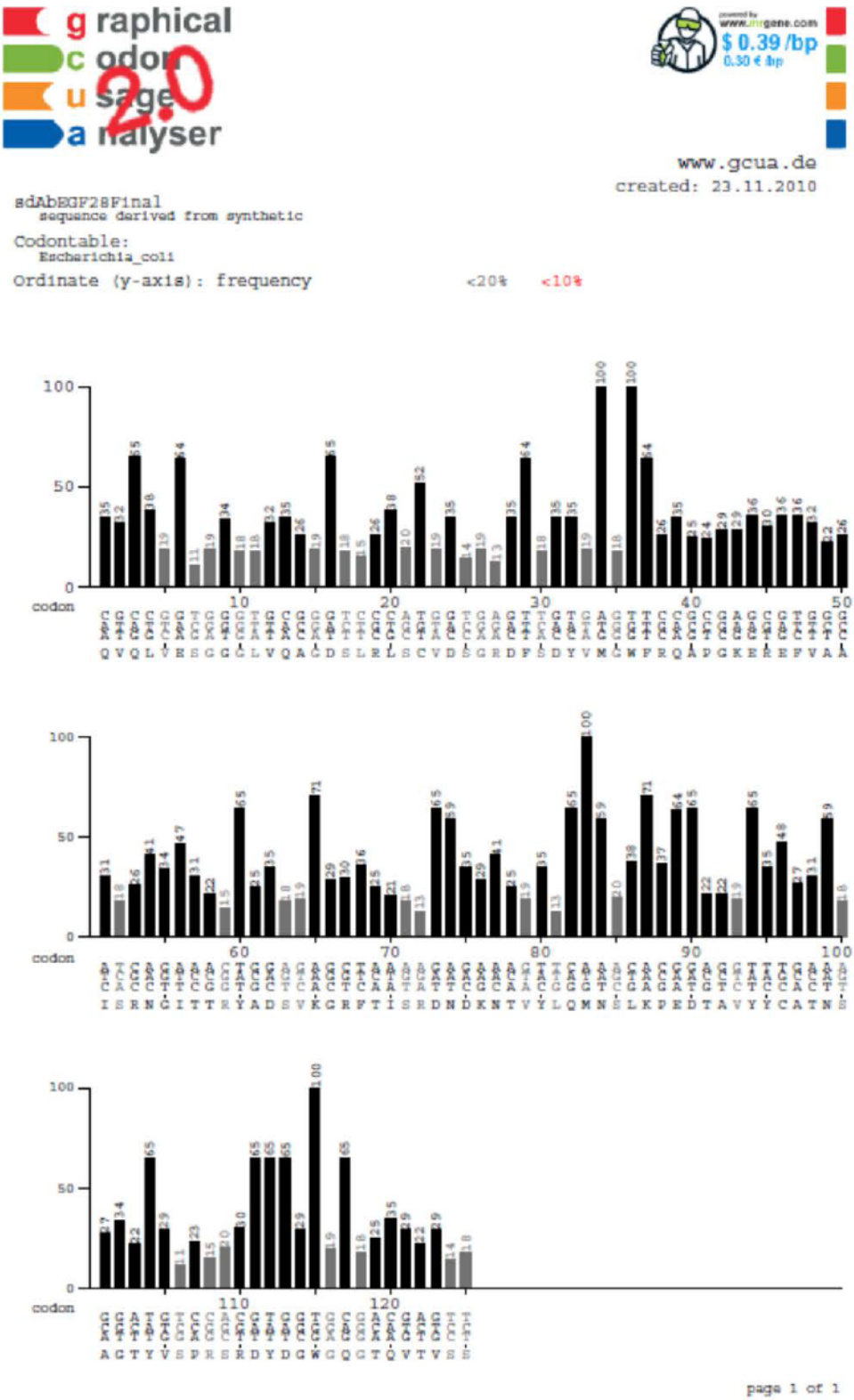


Figure 106 Each triplet position usage table sdAb EGF28 final adjustments.

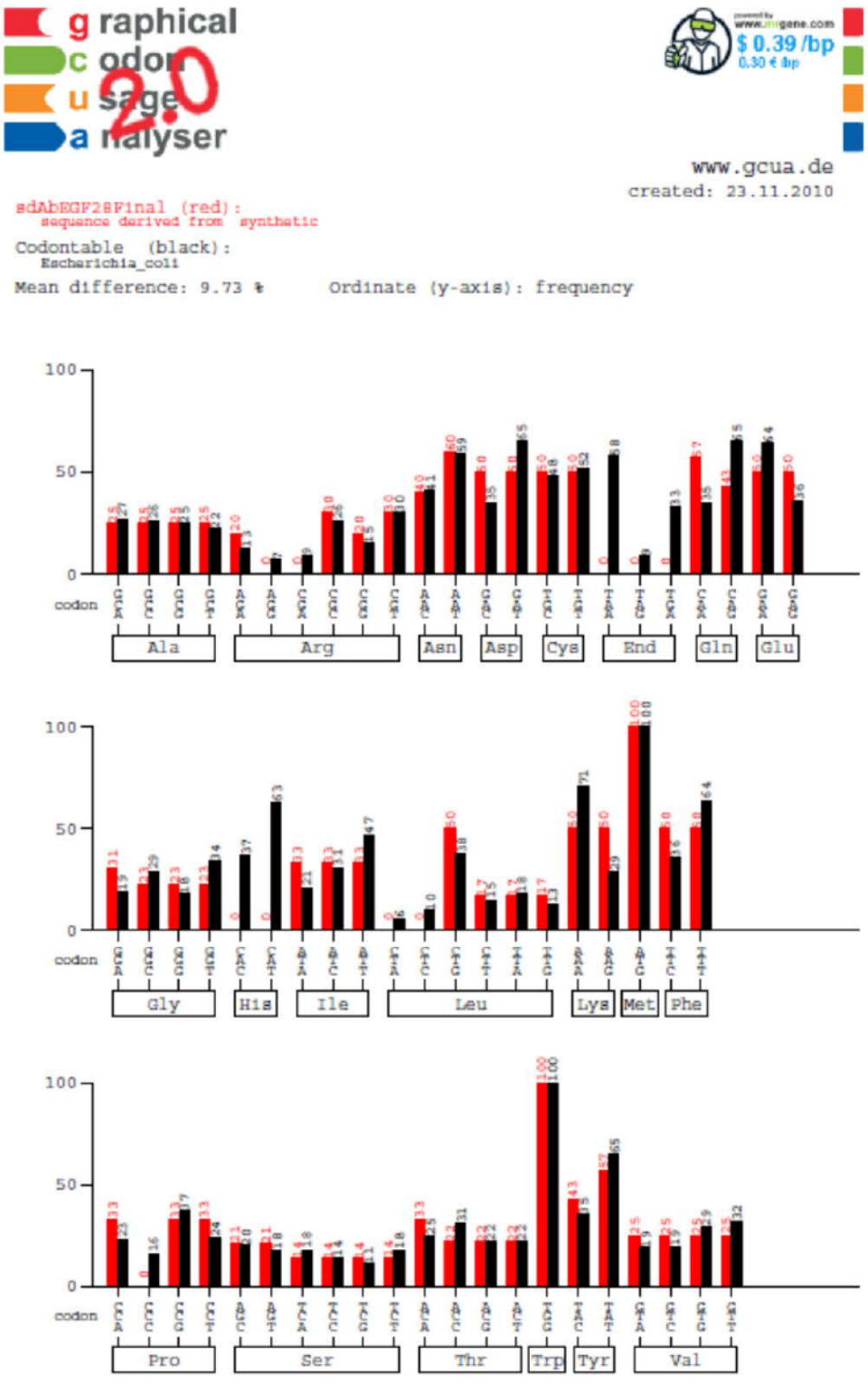


Figure 107 Each codon usage table sdAb EGF28 final adjustments.

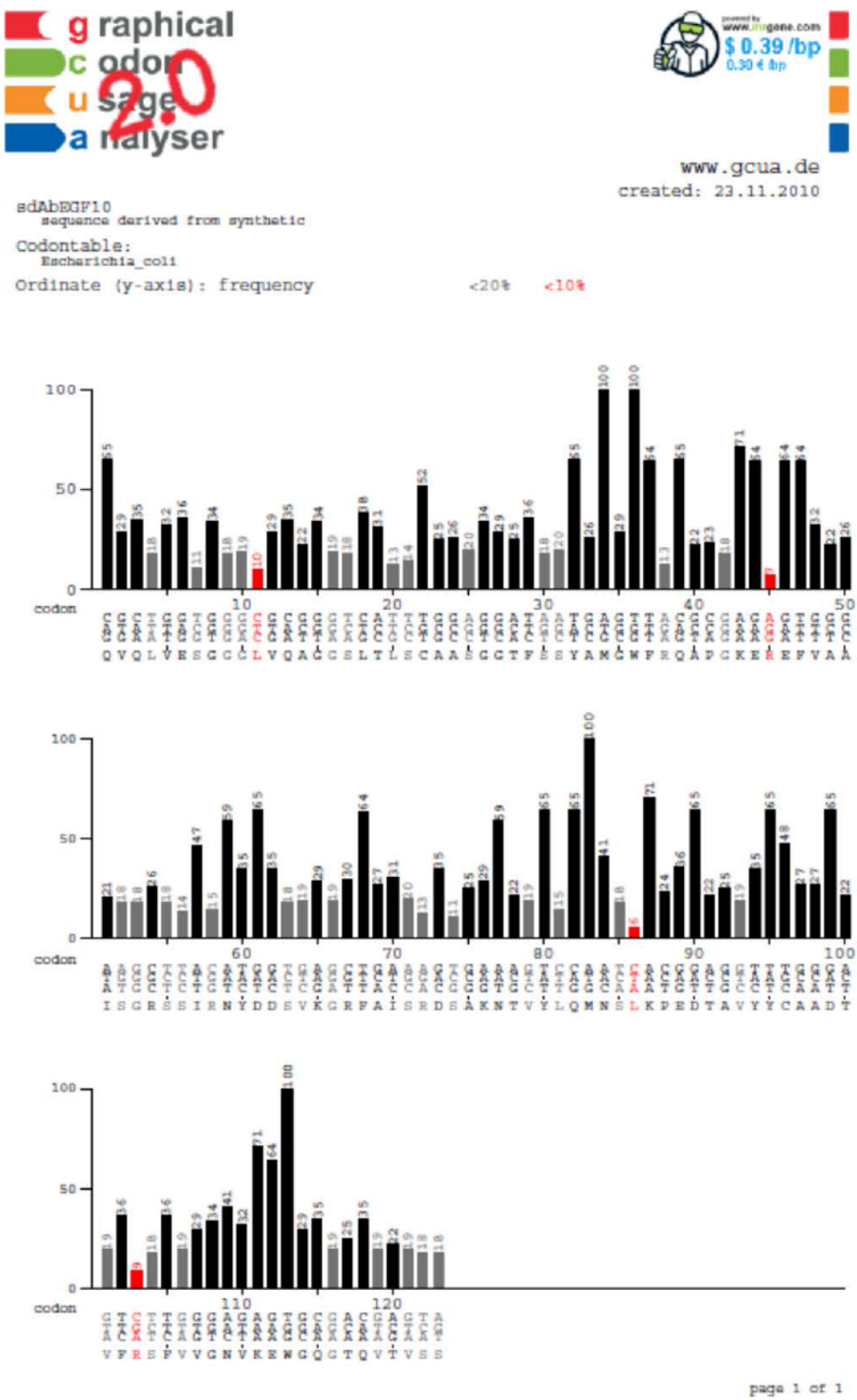


Figure 108 Each triplet position usage table sdAb EGF10.

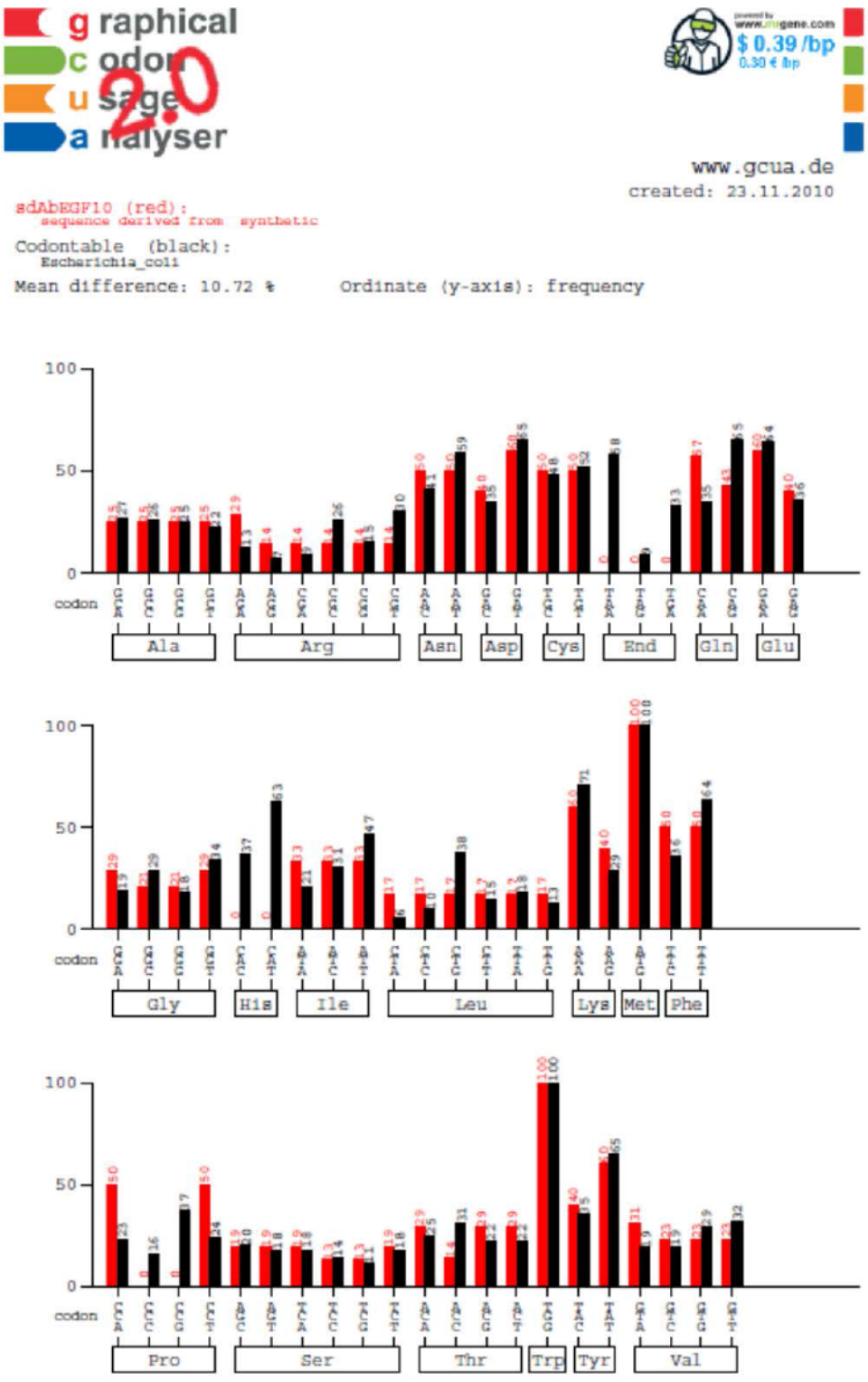


Figure 109 Each codon usage table sdAb EGF10.

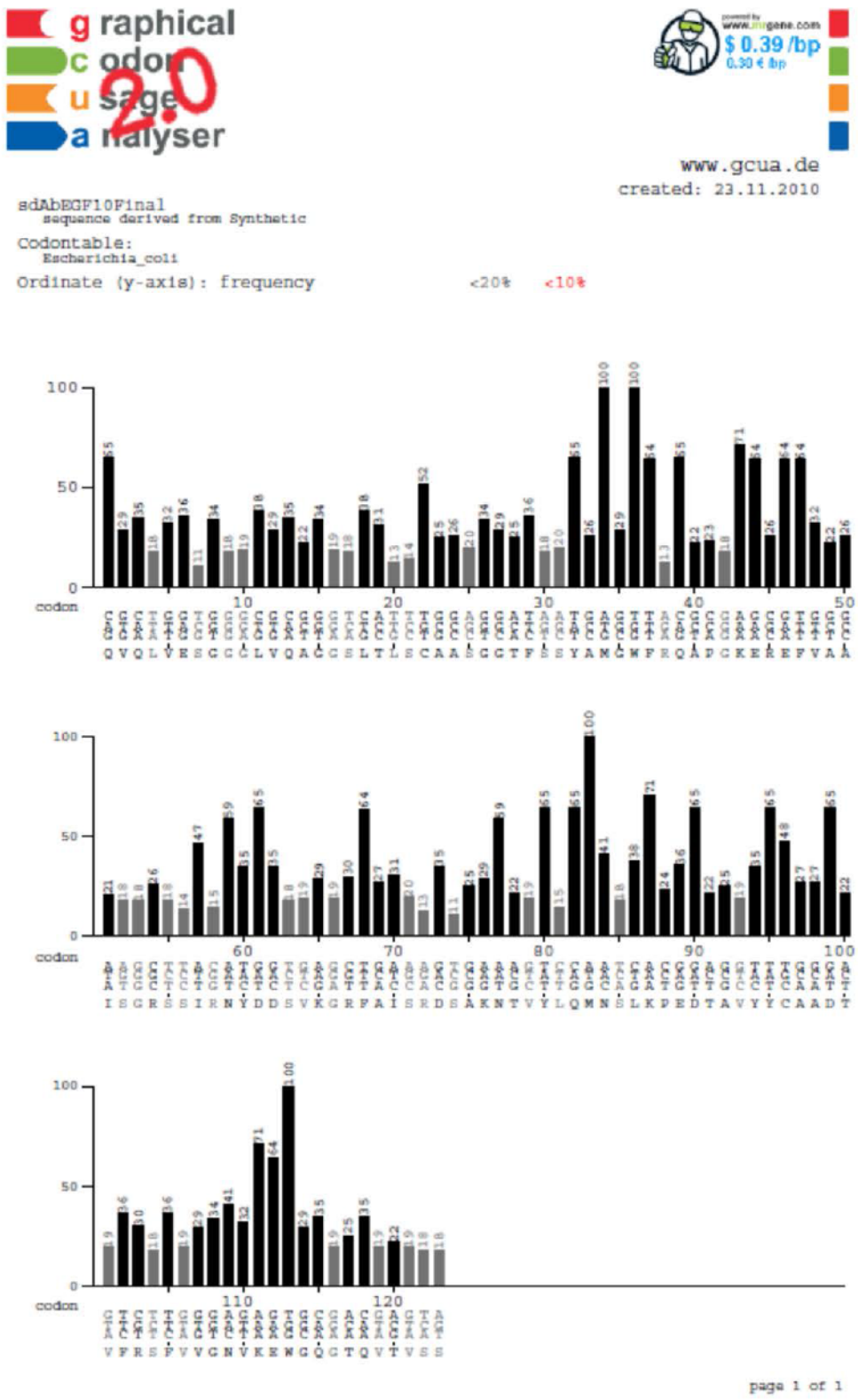


Figure 110 Each triplet position usage table sdAb EGF10 final adjustments.

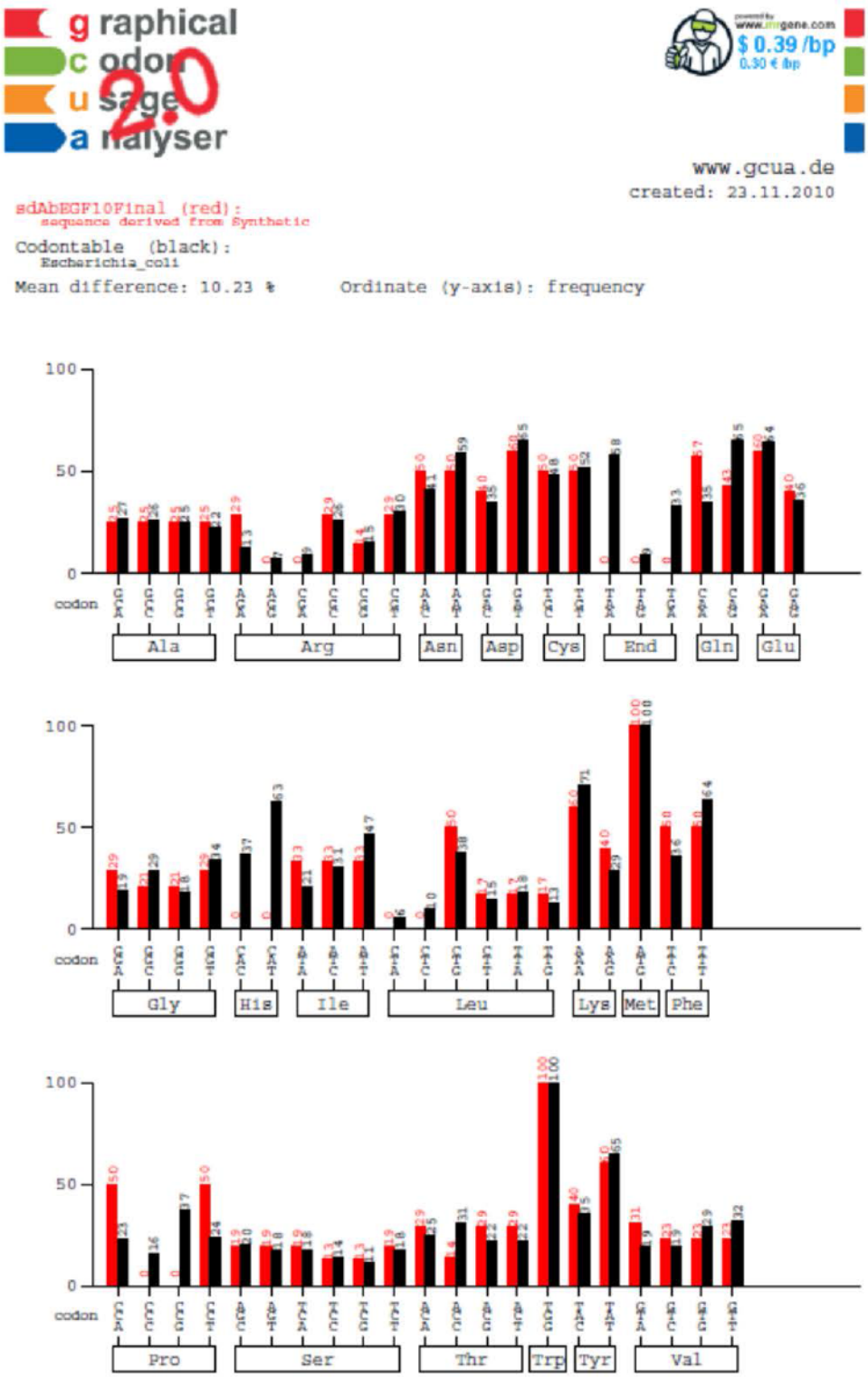


Figure 111 Each codon usage table sdAb EGF10 final adjustments.

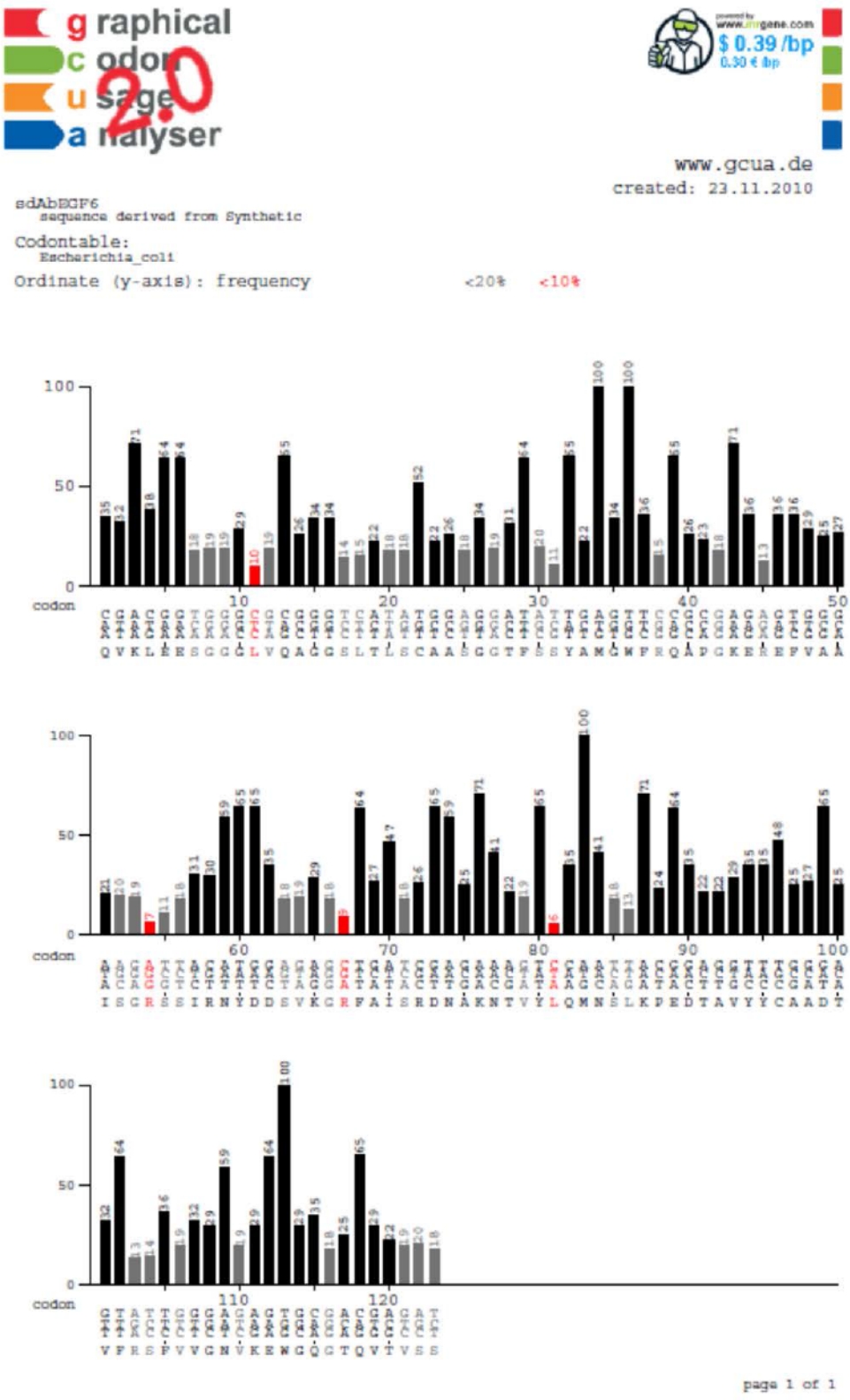


Figure 112 Each triplet position usage table sdAb EGF6.

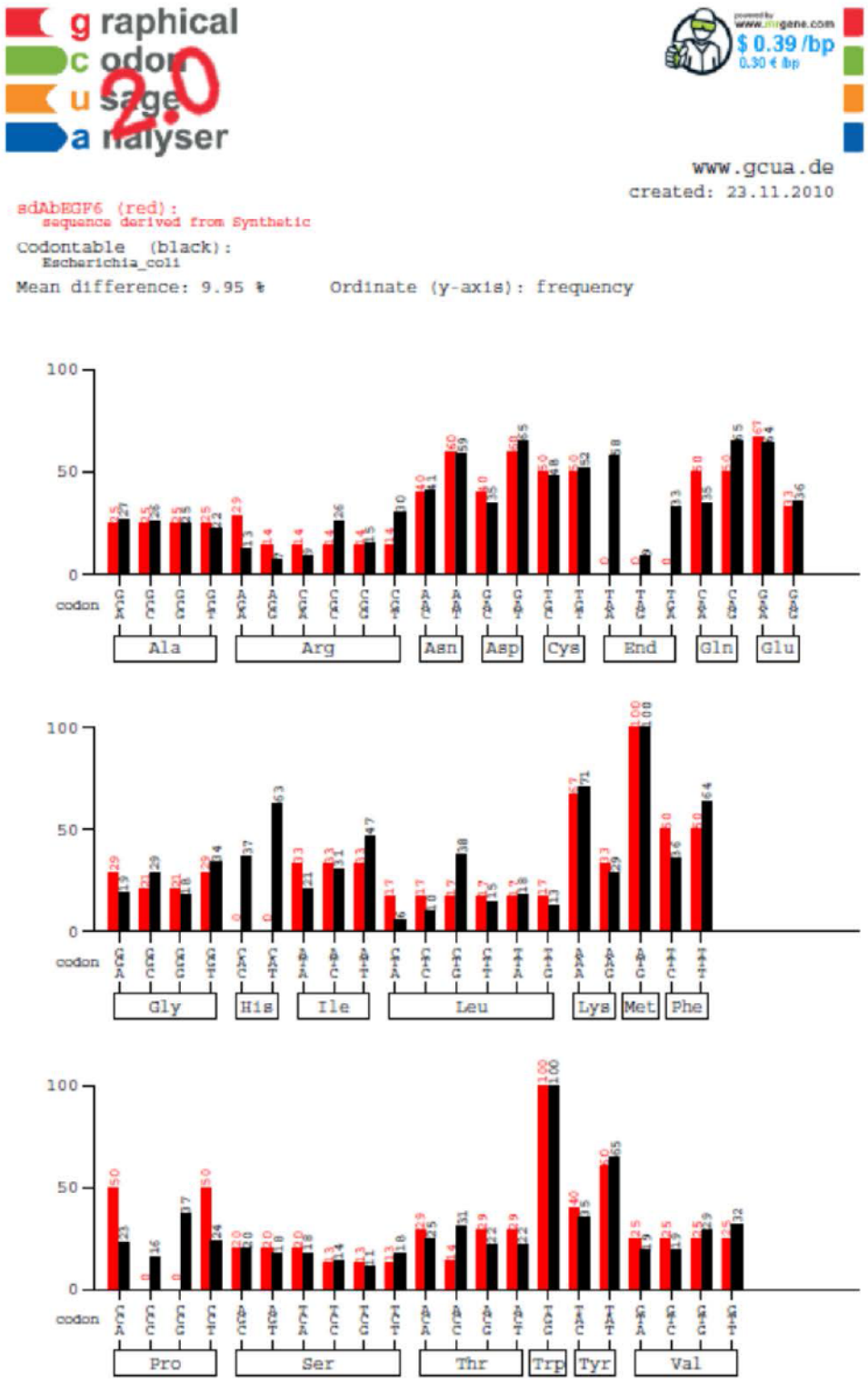


Figure 113 Each codon usage table sdAb EGF6.

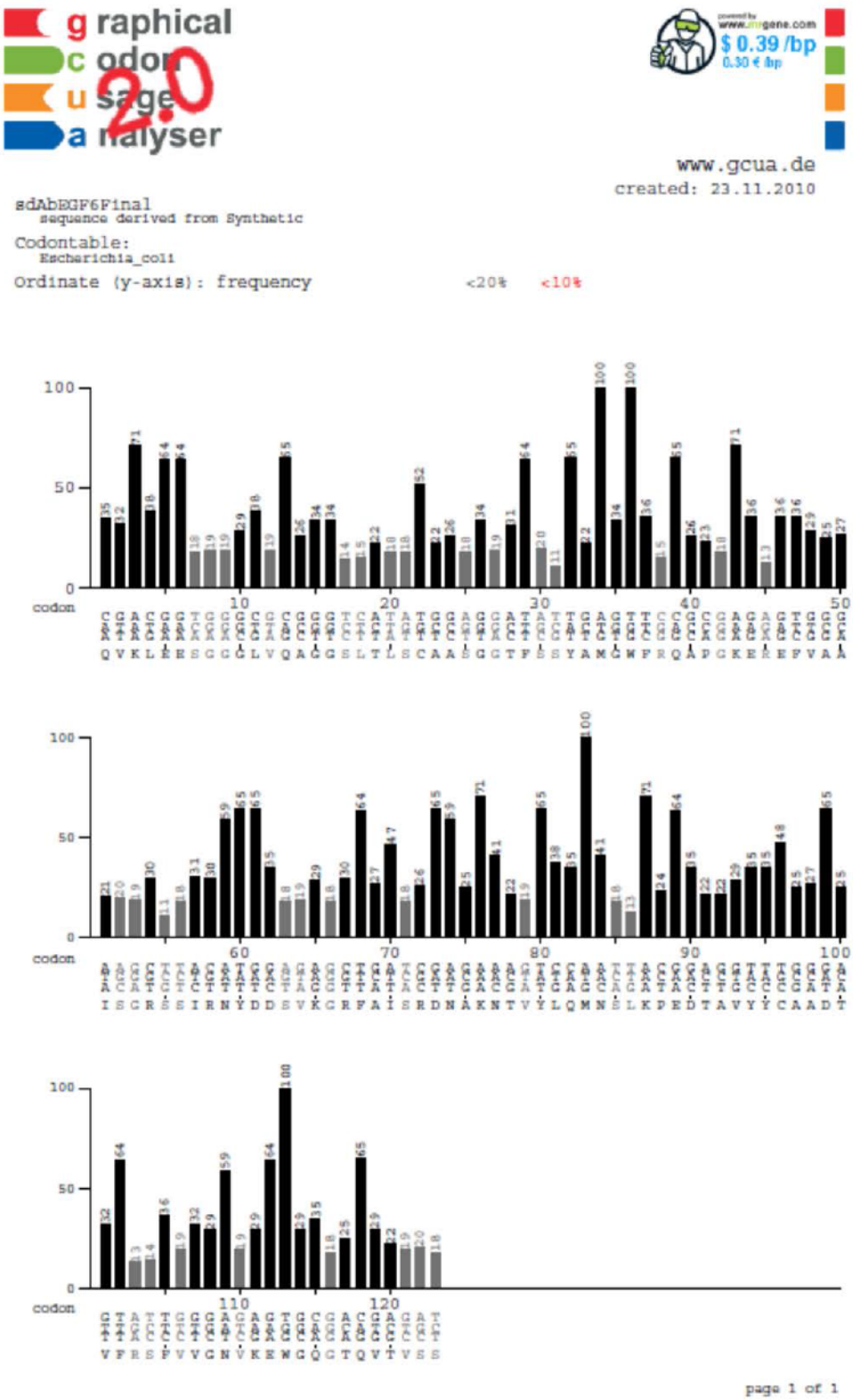


Figure 114 Each triplet position usage table sdAb EGF6 final adjustments.

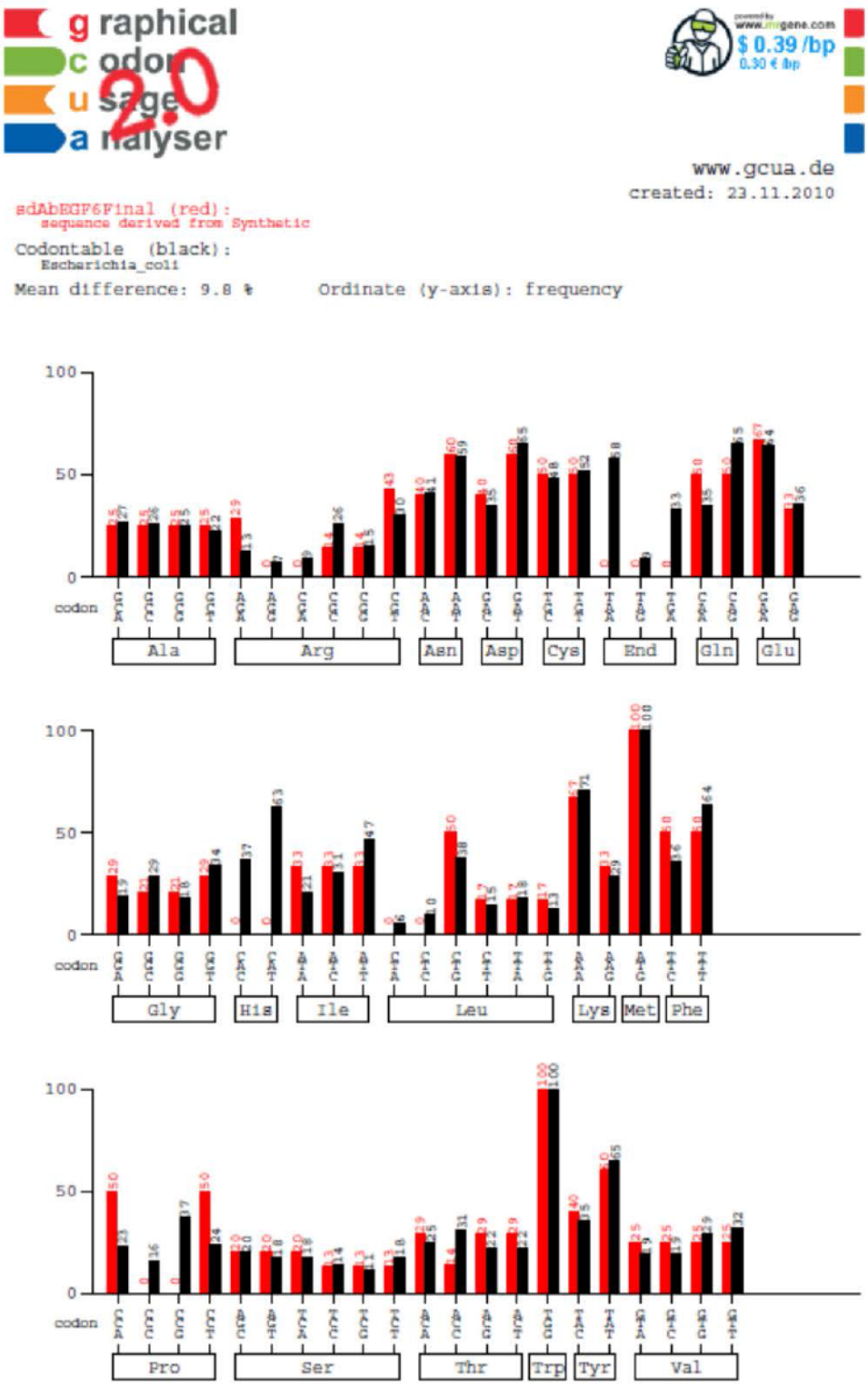


Figure 115 Each codon usage table sdAb EGF6 final adjustments.

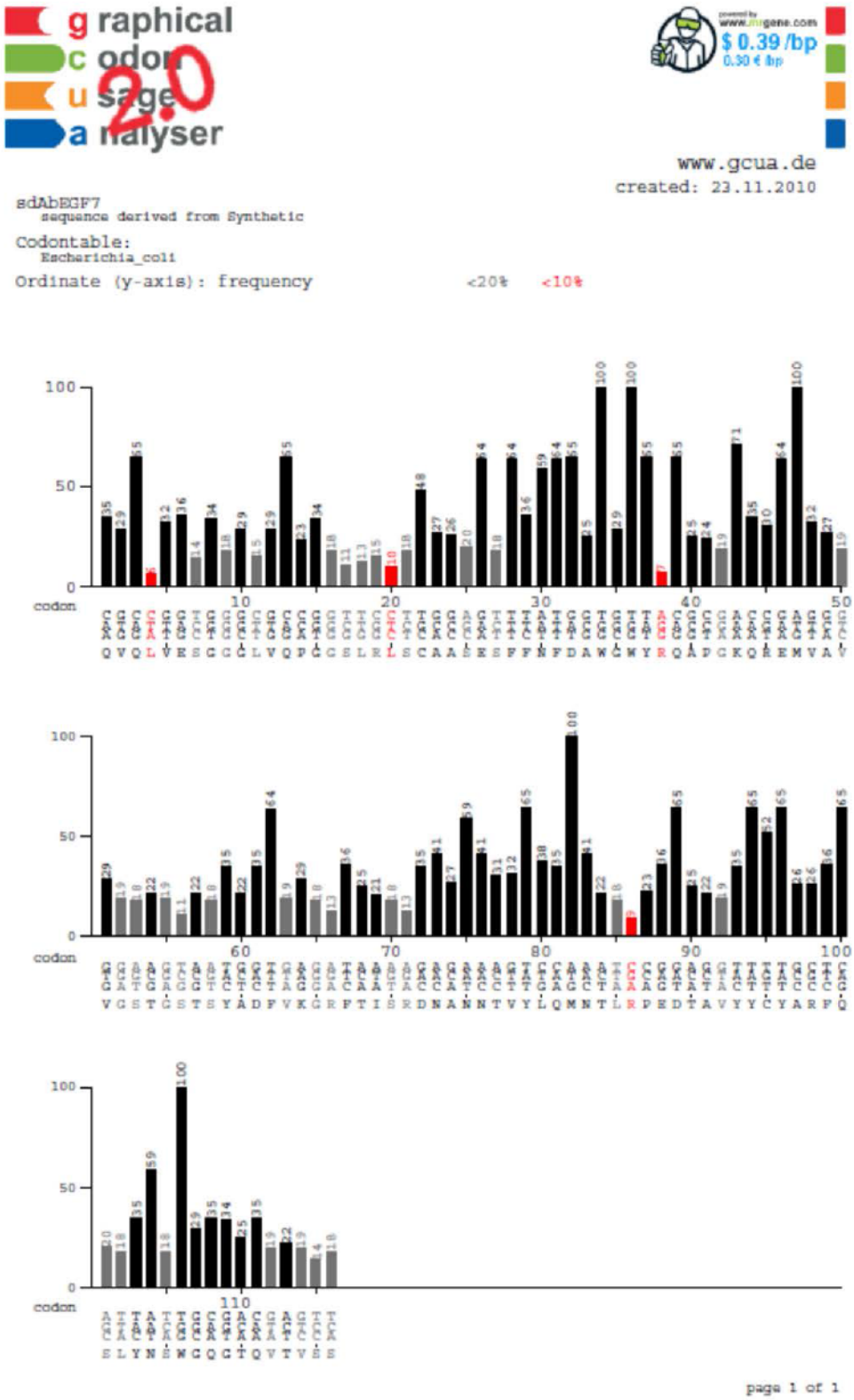


Figure 116 Each triplet position usage table sdAb EGF7.

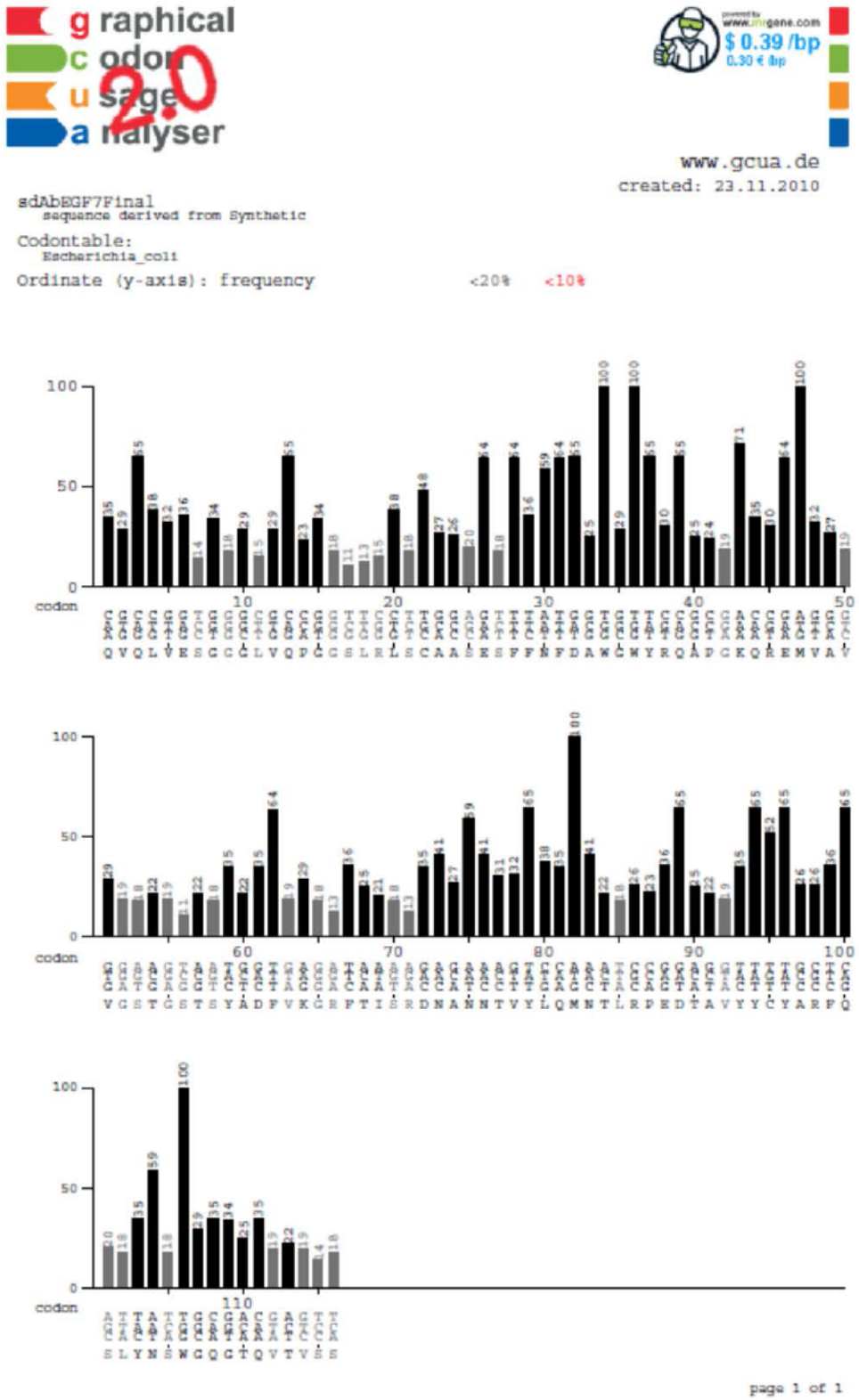


Figure 117 Each triplet position usage table sdAb EGF7 final adjustments.

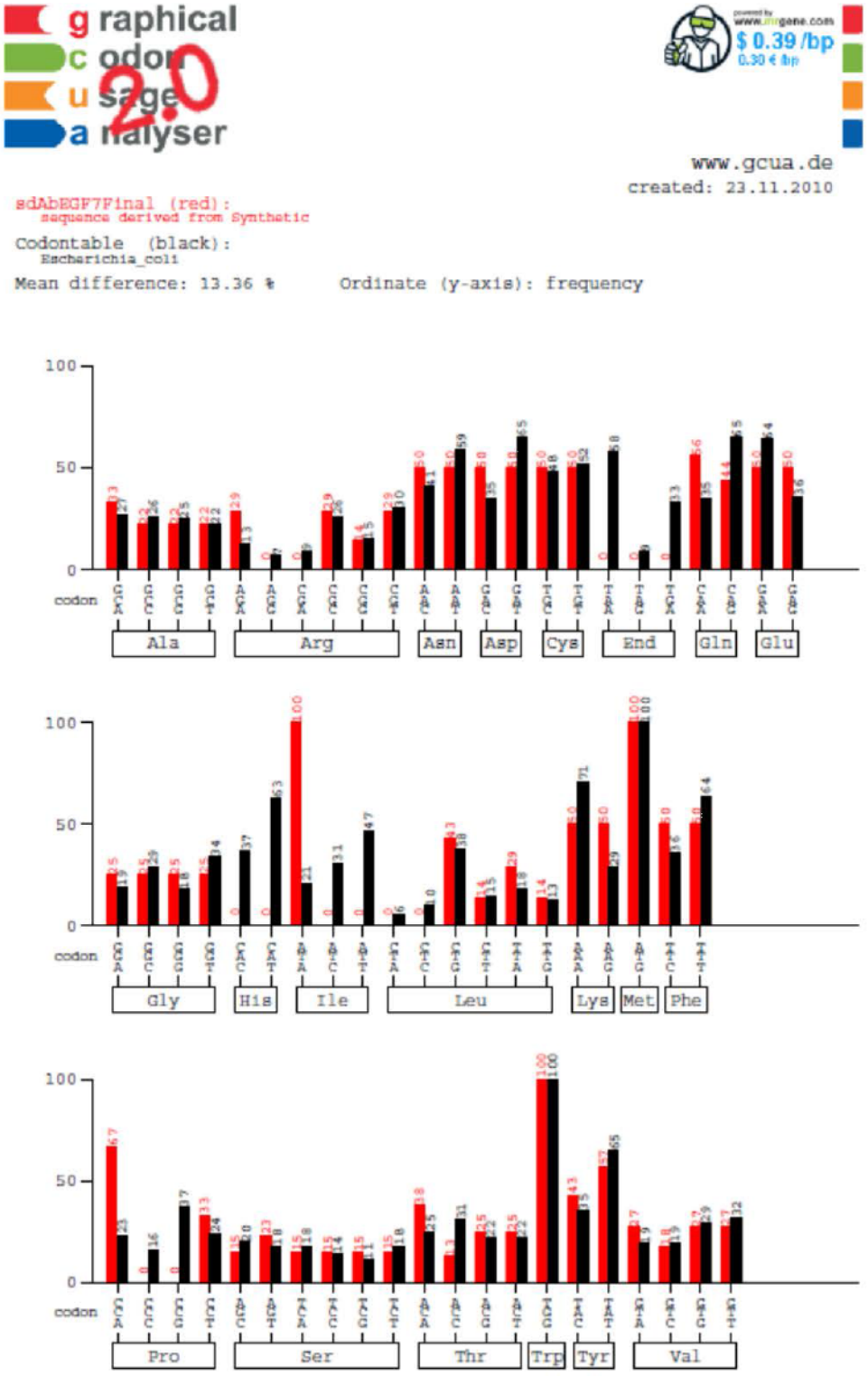


Figure 118 Each codon usage table sdAb EGF7 final adjustments.

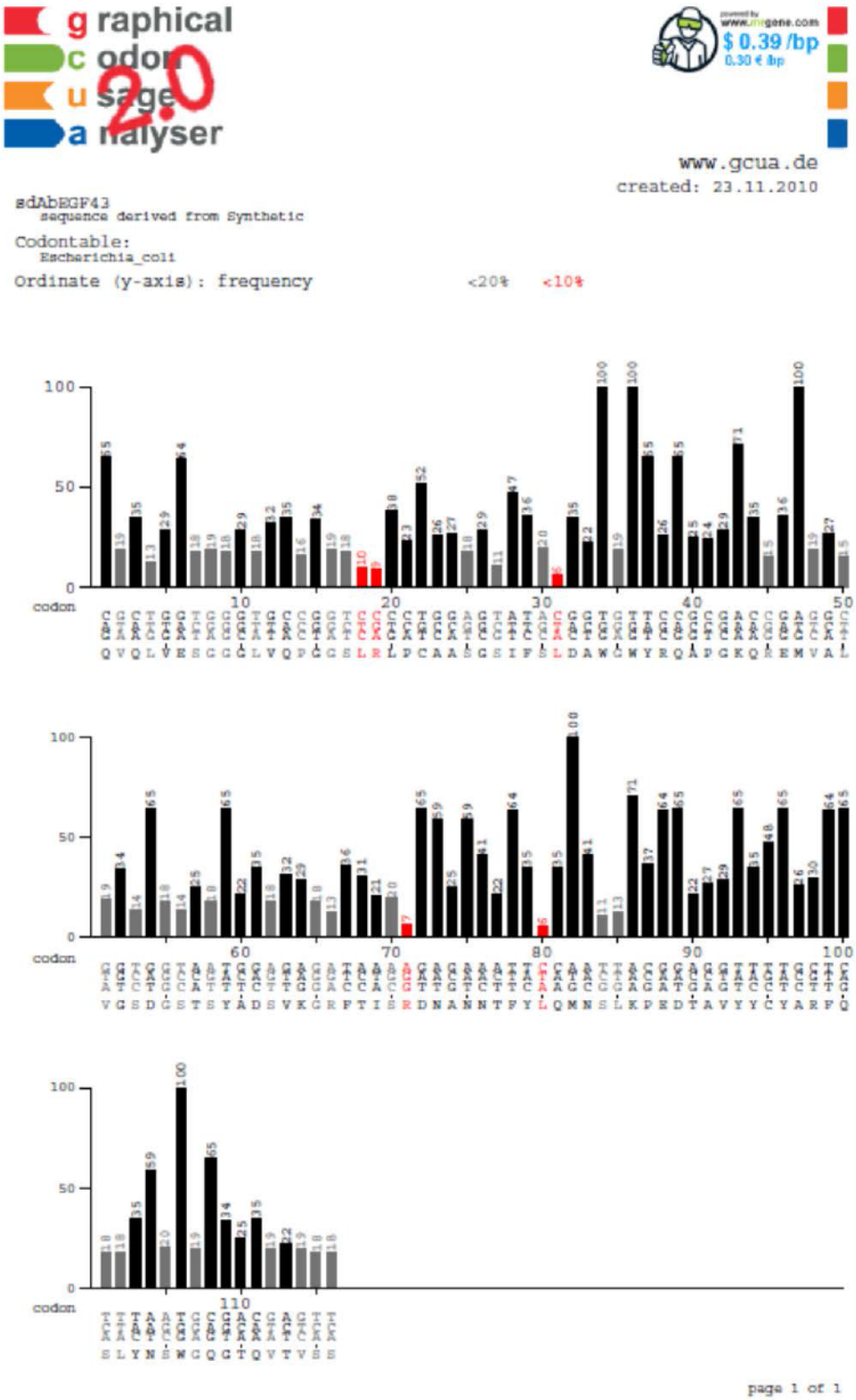


Figure 119 Each triplet position usage table sdAb EGF43.

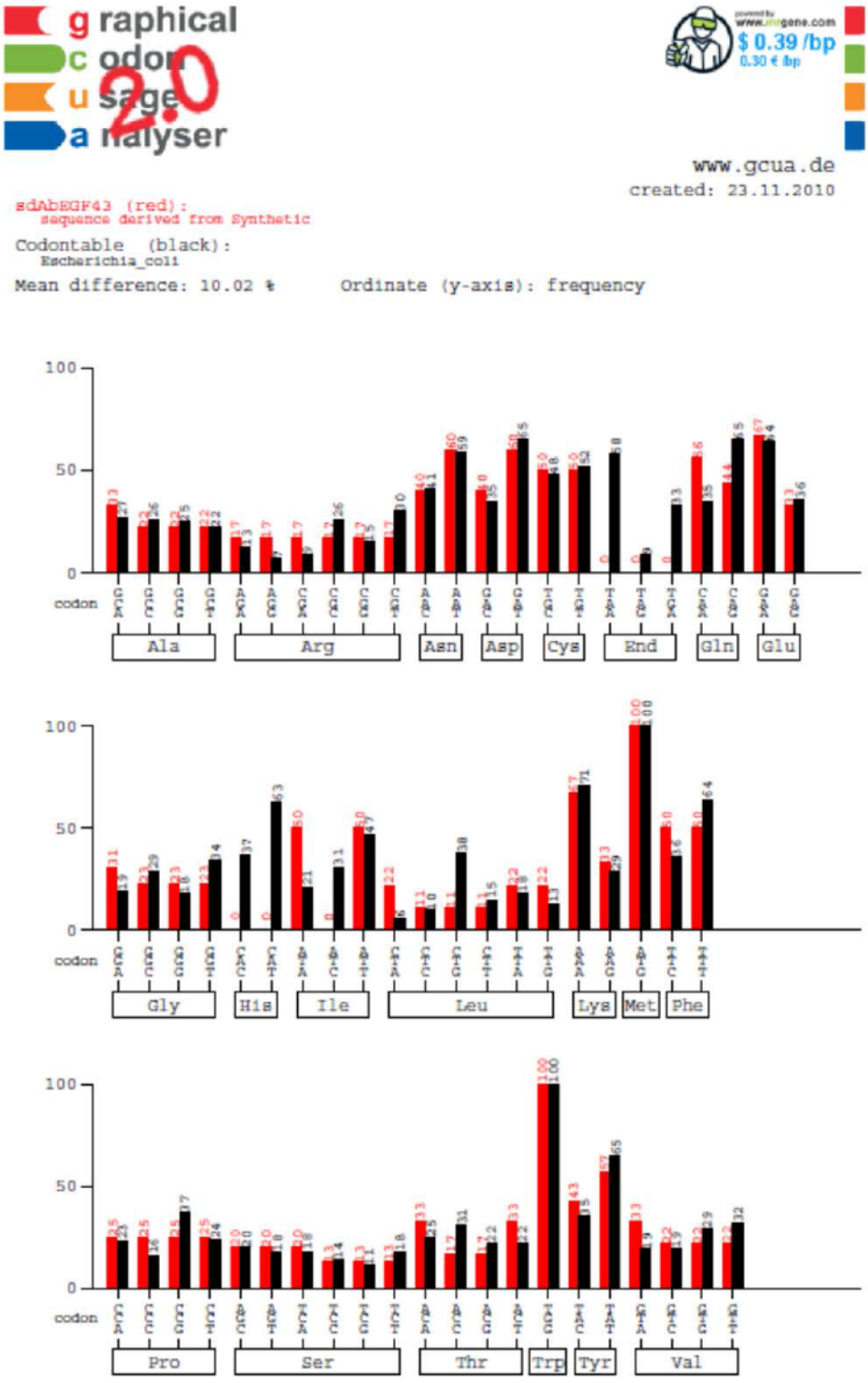


Figure 120 Each codon usage table sdAb EGF43.

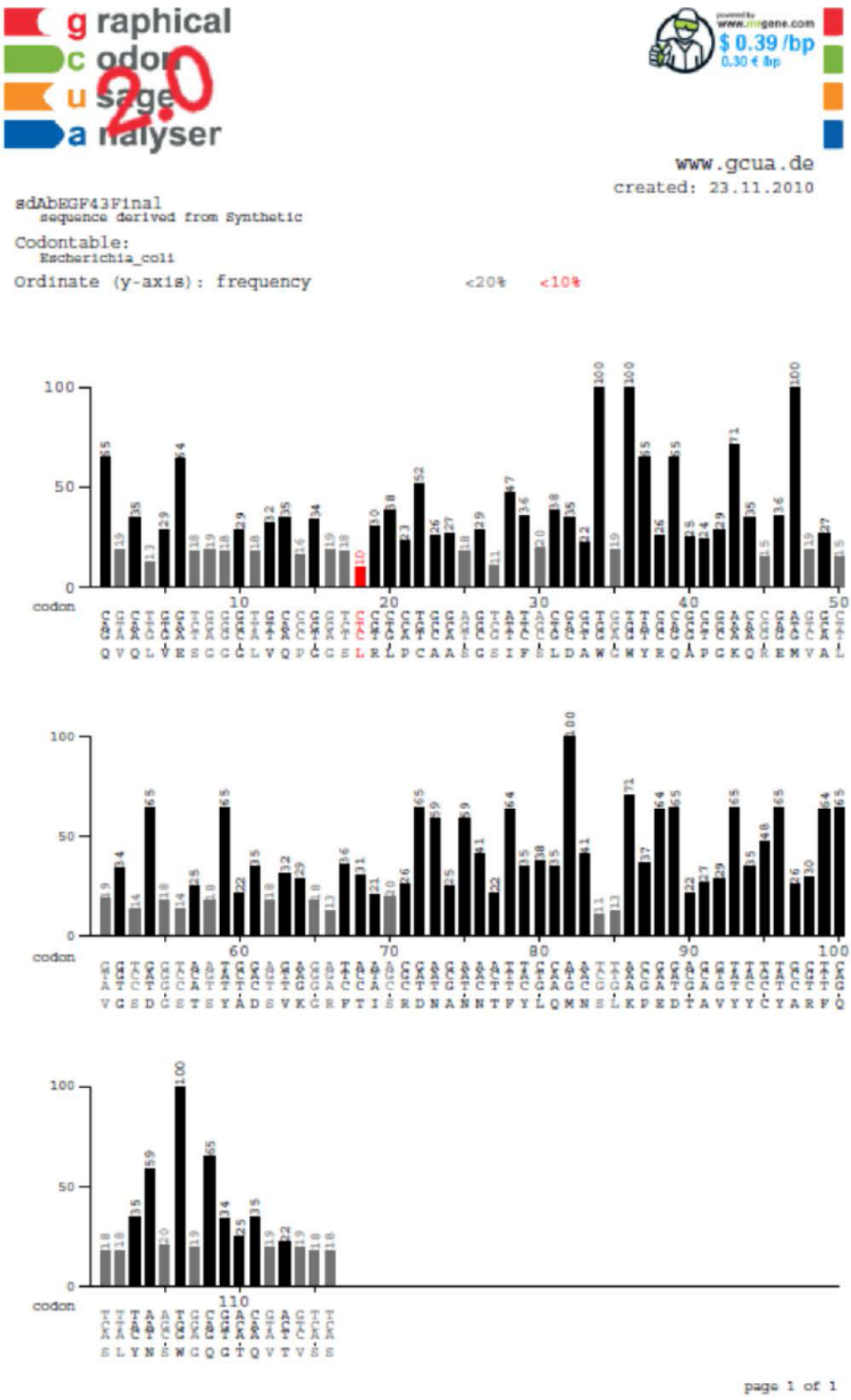


Figure 121 Each triplet position usage table sdAb EGF43 final adjustments.

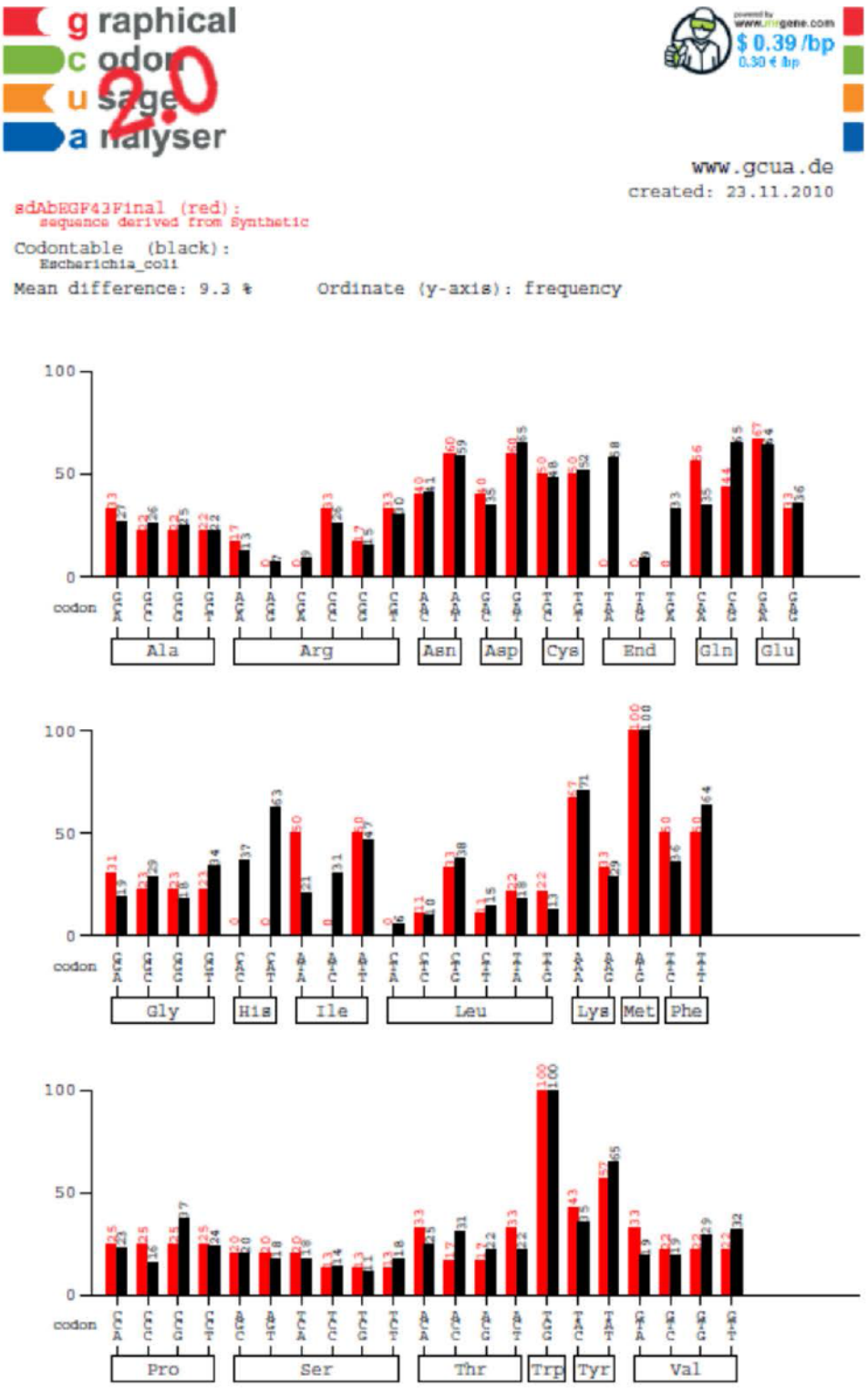


Figure 122 Each codon usage table sdAb EGF43 final adjustments.

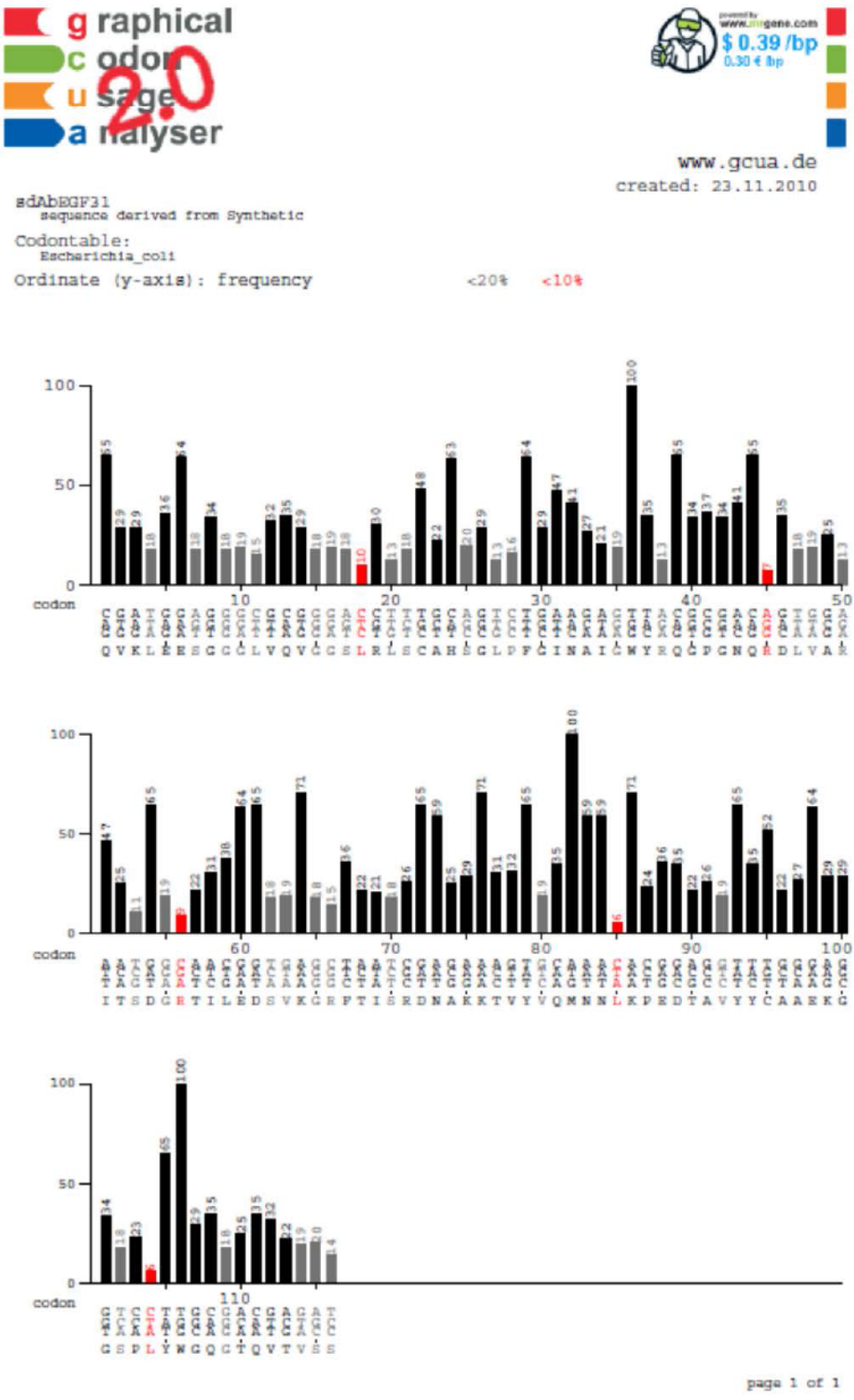


Figure 123 Each triplet position usage table sdAb EGF31.

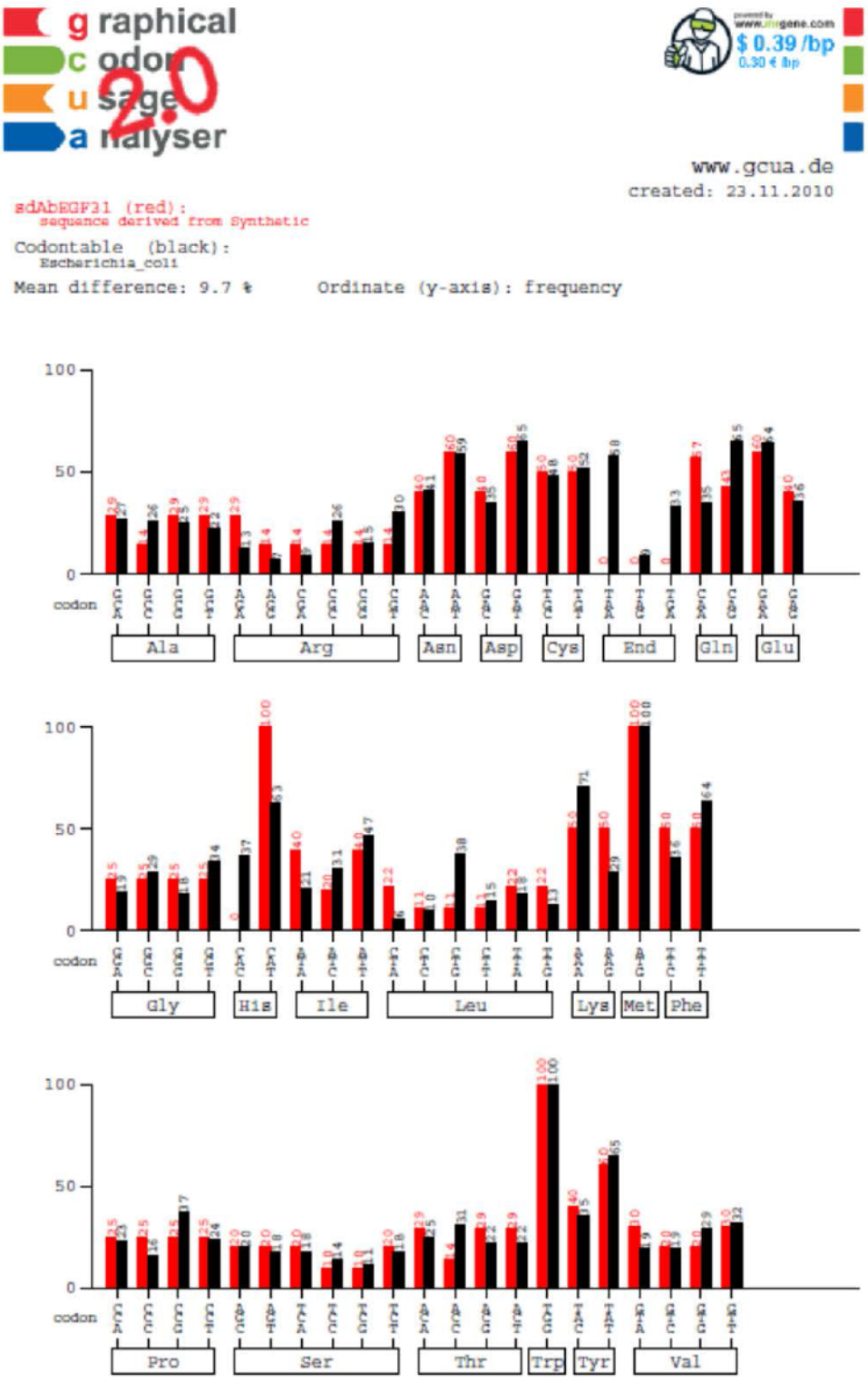


Figure 124 Each codon usage table sdAb EGF31.

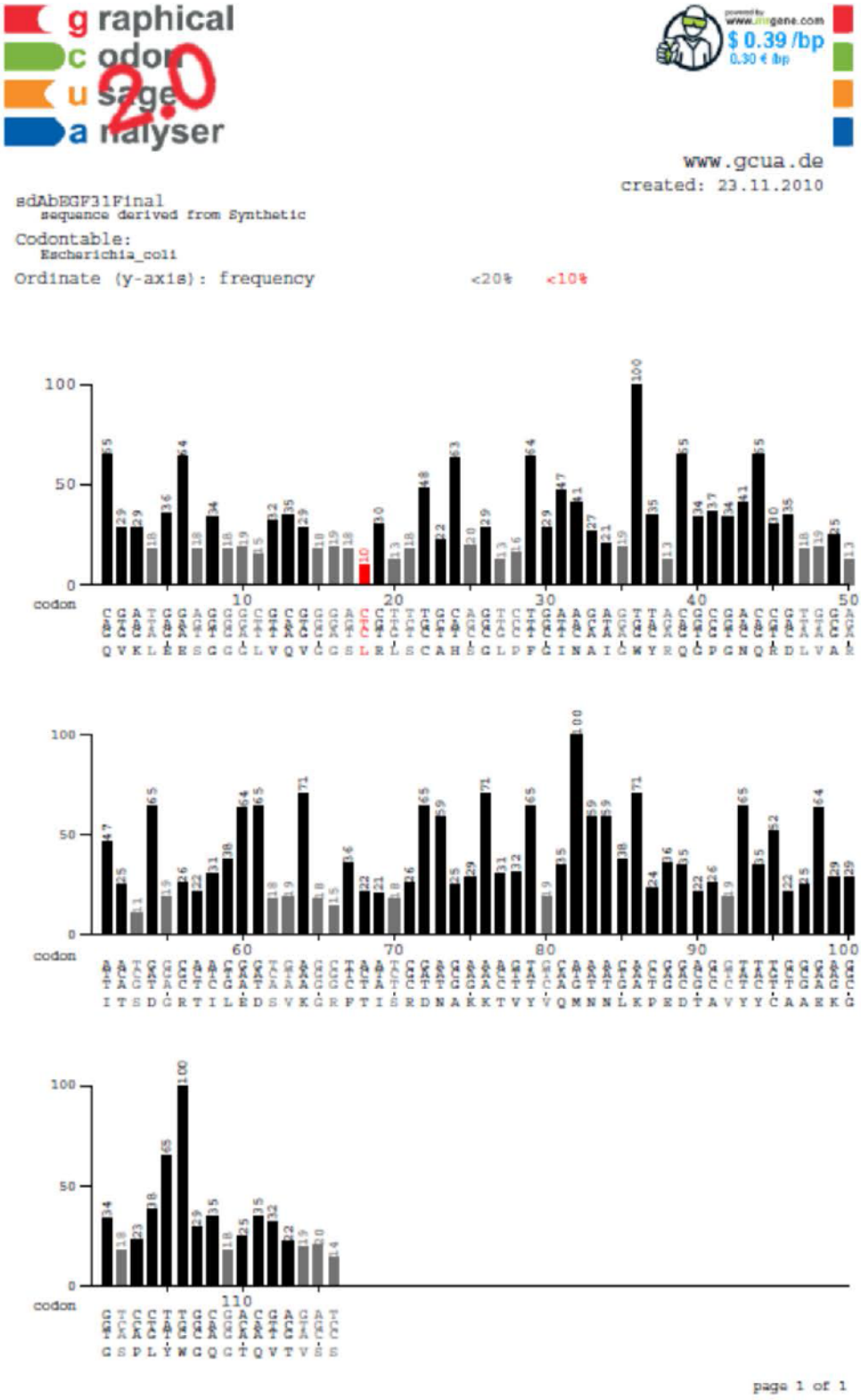


Figure 125 Each triplet position usage table sdAb EGF31 final adjustments.

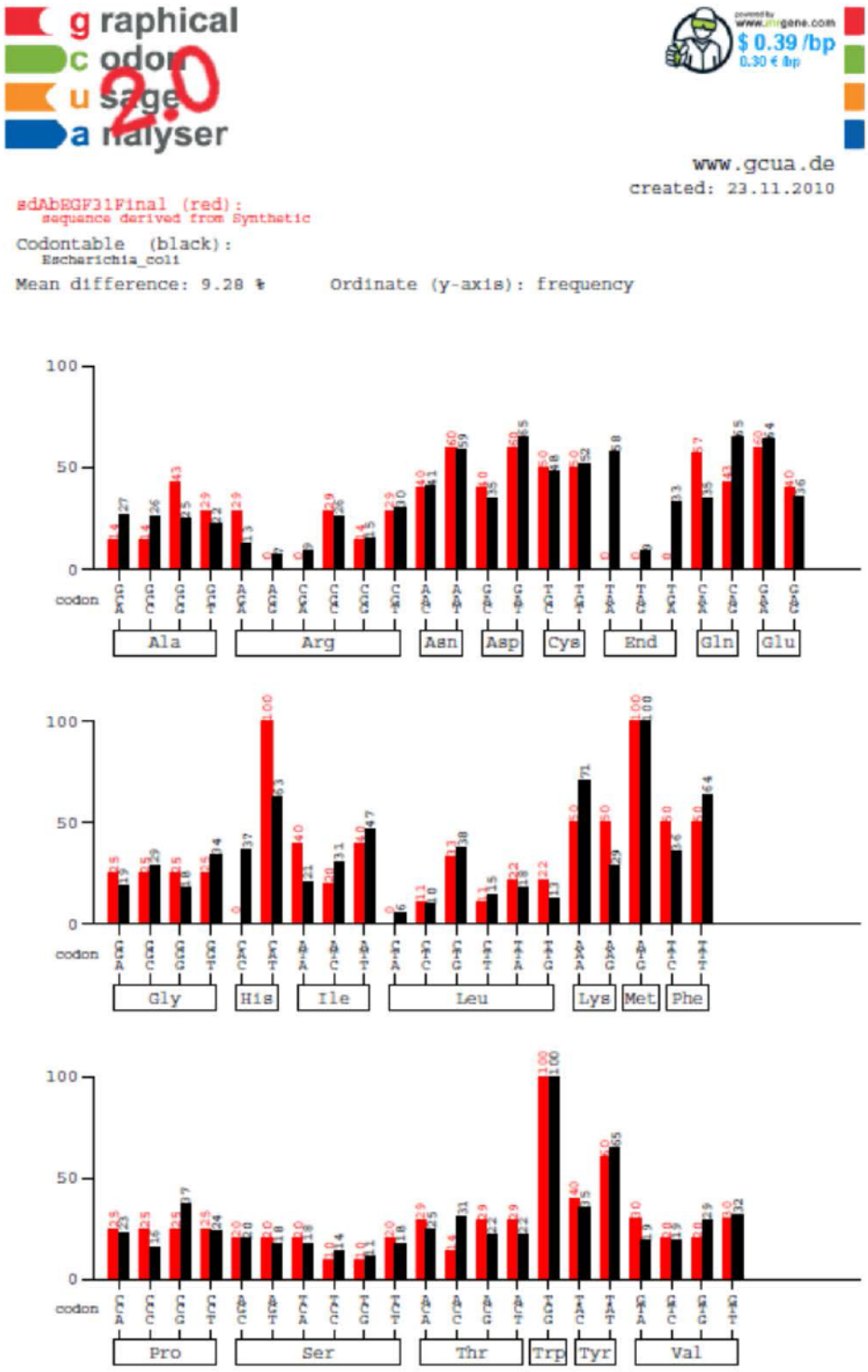
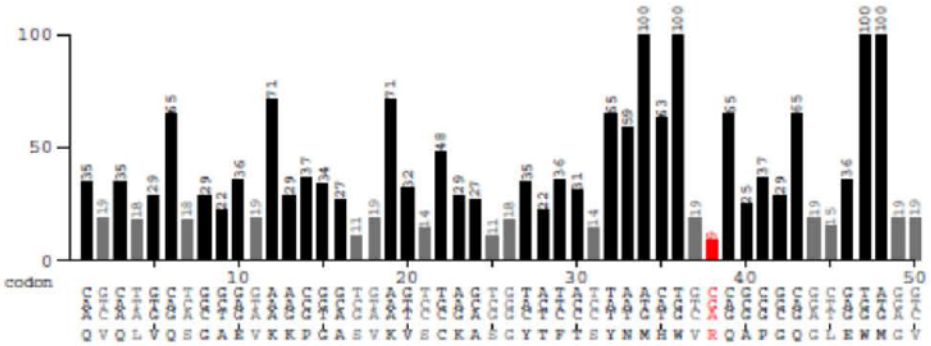


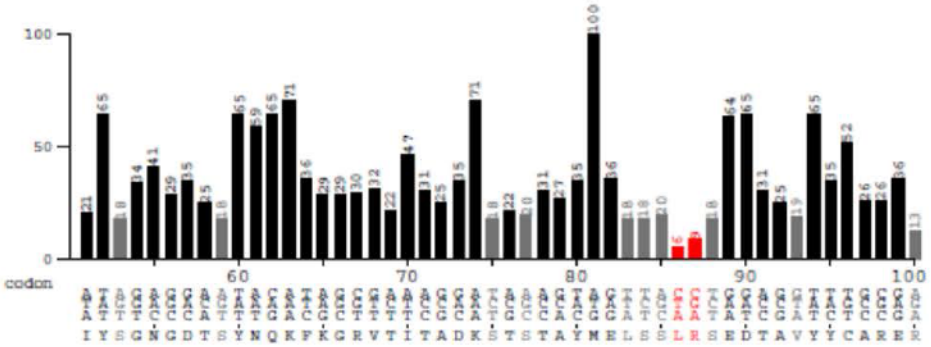
Figure 126 Each codon usage table sdAb EGF31 final adjustments.

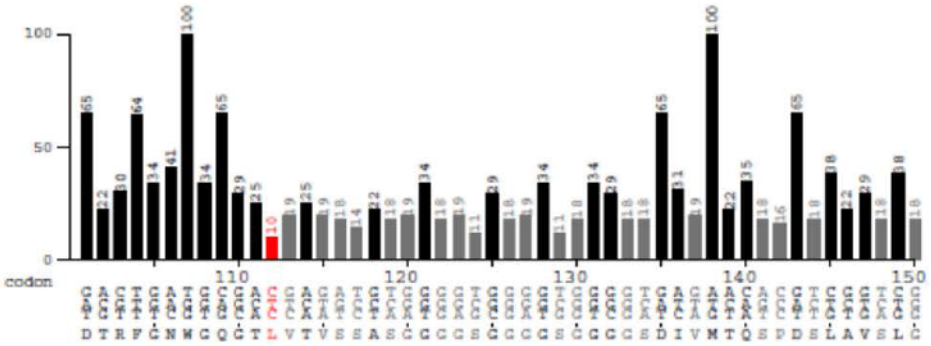
CD117I
sequence derived from Synthetic

Codontable:
Escherichia_coli

Ordinate (y-axis): frequency <20% <10%







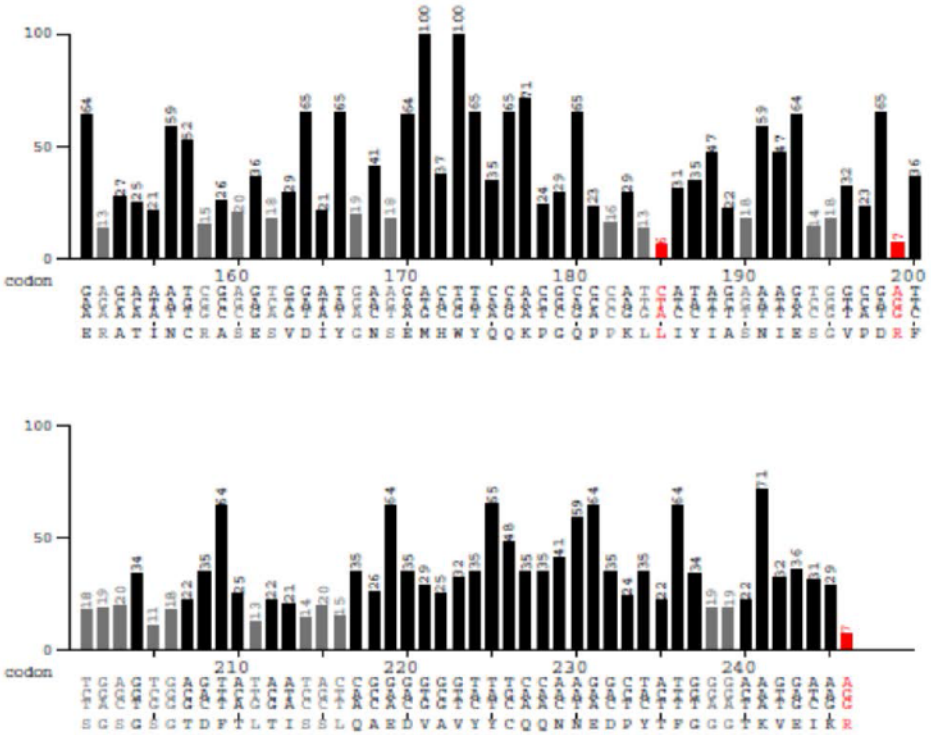


Figure 127 Each triplet position usage table scFv CD117I.

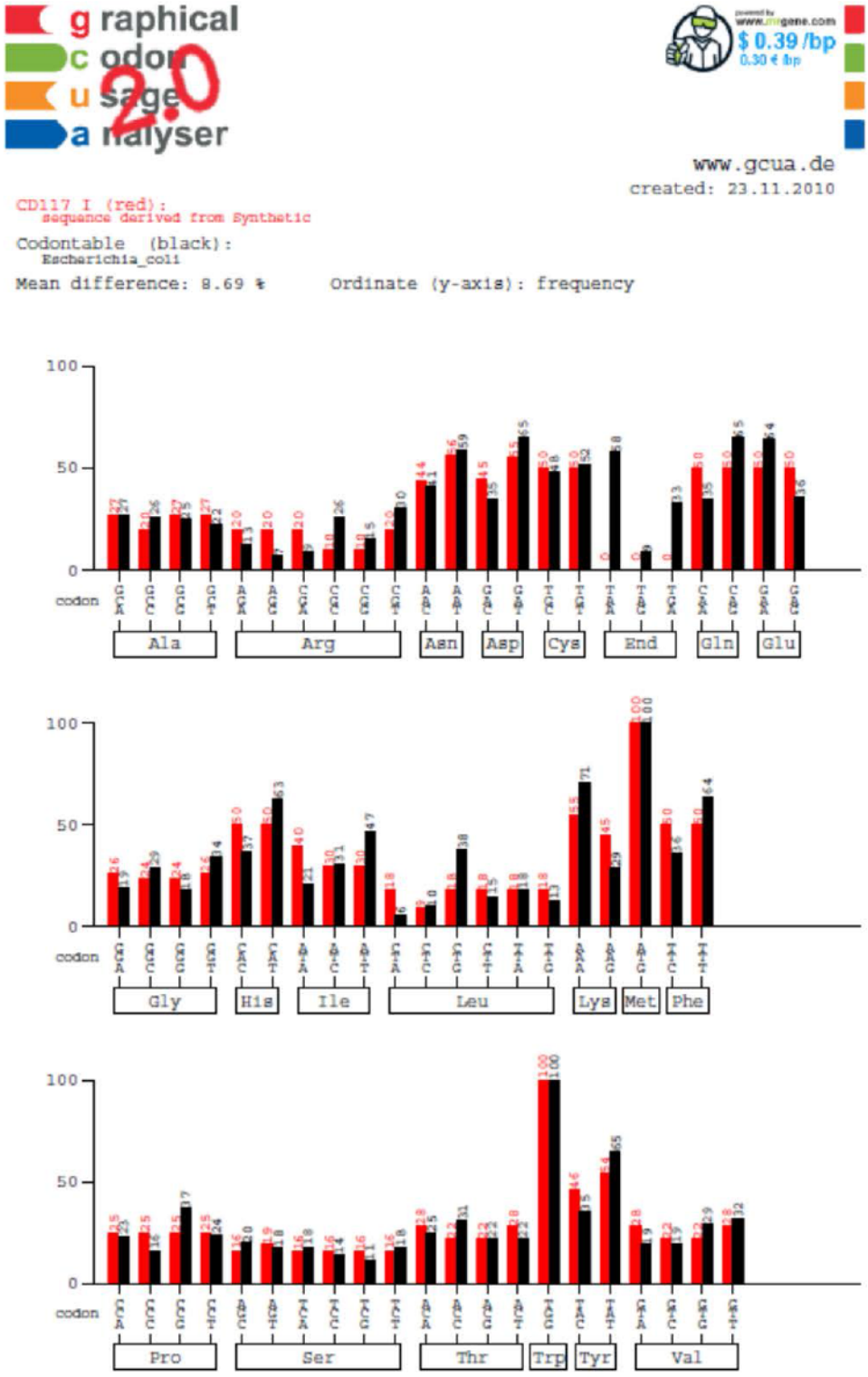


Figure 128 Each codon usage table scFv CD117I.

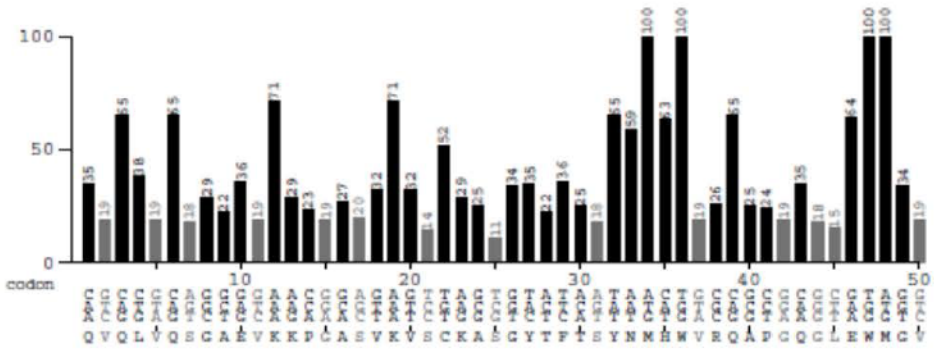
graphical codon usage analyser 2.0

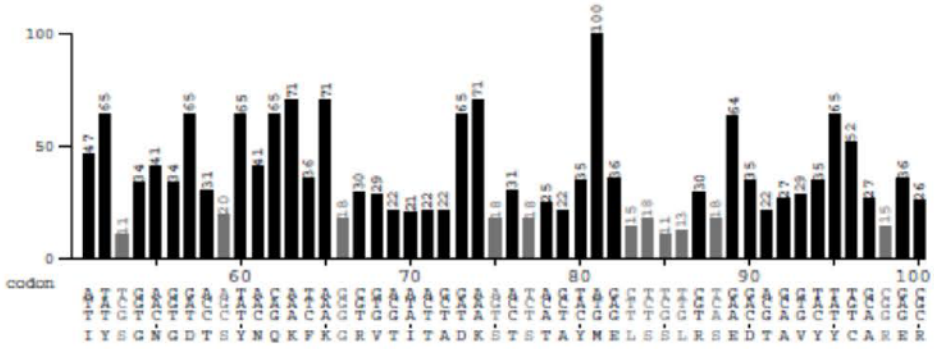
www.gcua.de
created: 15.06.2010

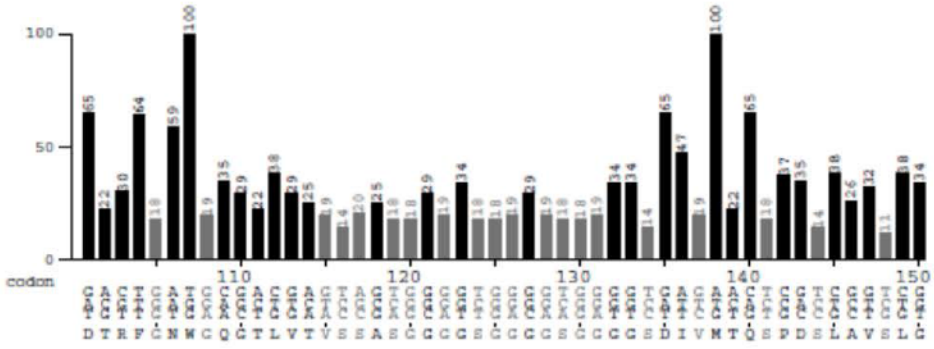
CD117I
sequence derived from SYNTHETIC

Codontable:
Escherichia_coli

Ordinate (y-axis): frequency <20% <10%







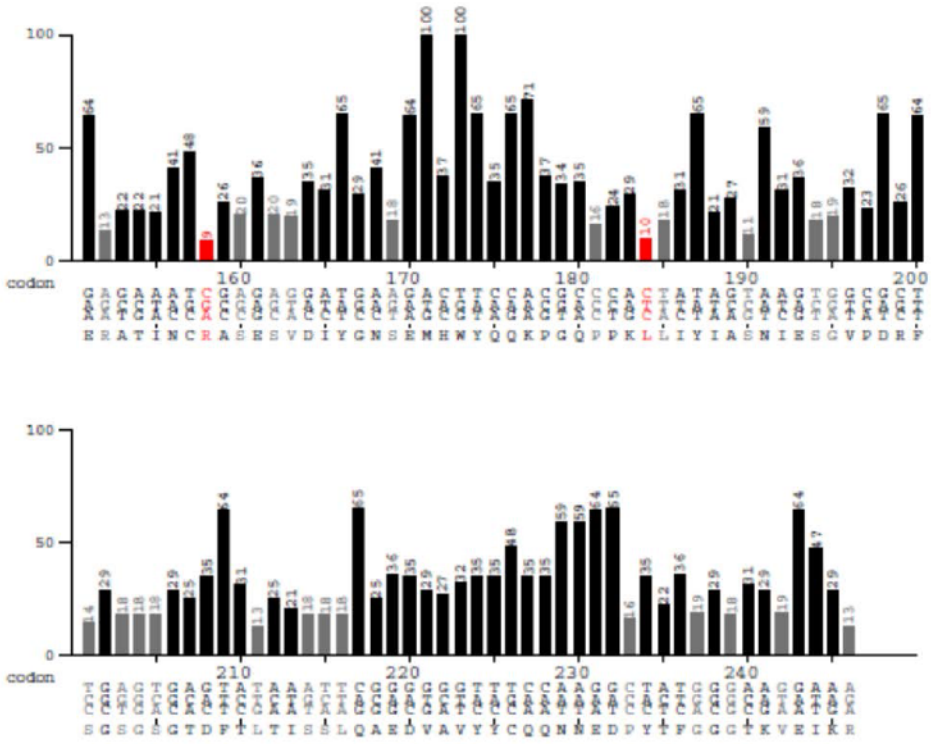


Figure 129 Each triplet position usage table scFv CD117I final adjustments.

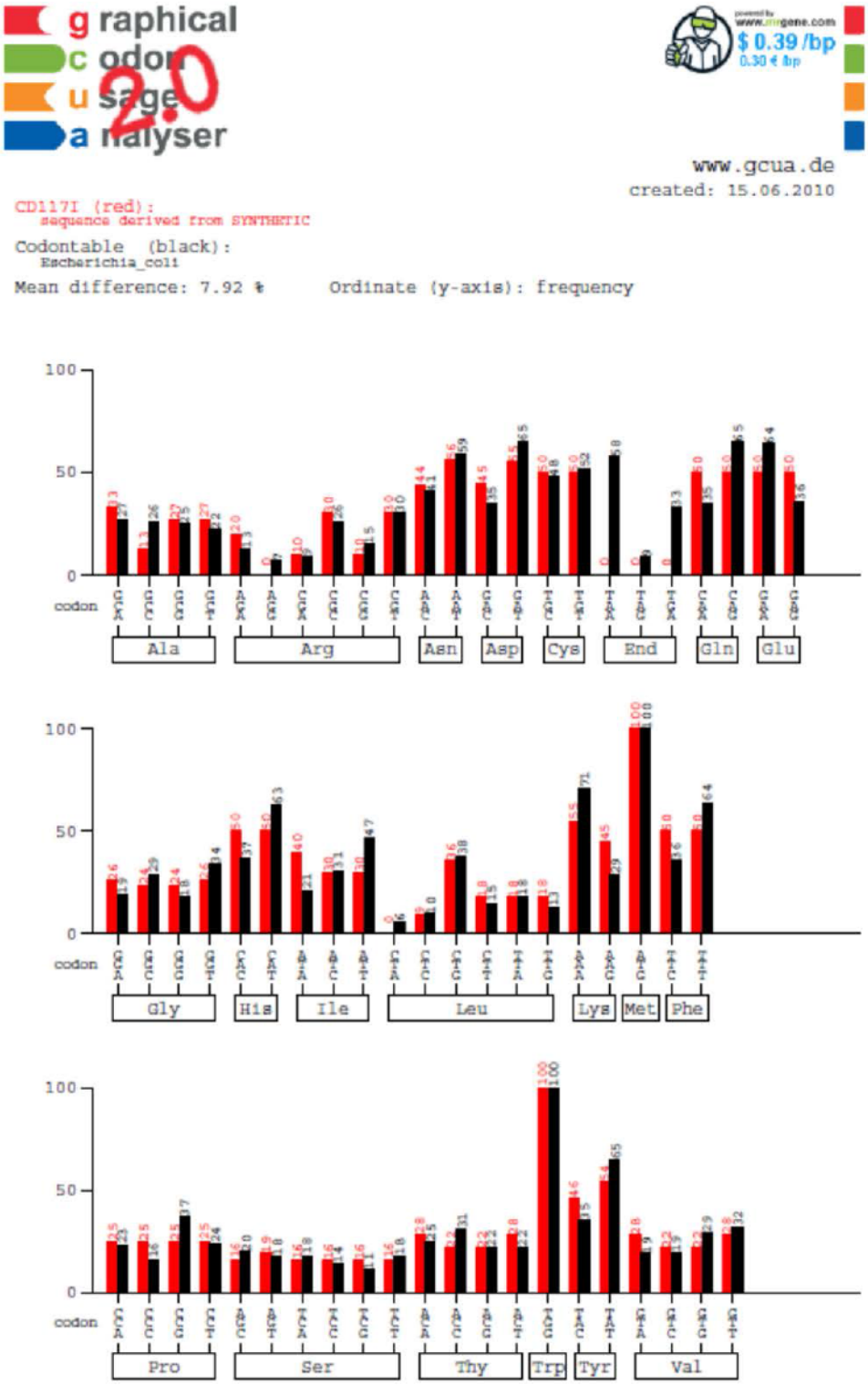


Figure 130 Each codon usage table scFv CD117I final adjustments.

Gene synthesis received from Entelechon.

Gene synthesis report

Date: 25.05.2009
Ordered by: Dr. Jolanta Tolkacz
Date of order: 27.04.2009
Order: synthetic genes cloned into E.coli vector

Ordered sequences:

EG6

TCTAGAACAAGTTAAACTGGAAGAATCAGGAGGAGGCCTGGTACAGGCCGGTGGTTCCCTTACTTTAAGTTGTGCTGCCAGTG
GTGGAACCTTTAGCTCGTATGCTATGGGTTGGTTCCGGCAGGCCCCAGGGAAAGAGAGAGTTCGTGGCGGCAATAAGCGGA
CGTTCGTCTATCCGTAATTATGATGACAGTGTAAAGGGGCGTTTTGCAATTTCACGCGATAATGCGAAAAACACGGTATATCT
GCAAATGAACTCATTGAAACCTGAAGACACTGCTGTGTACTACTGCGCGGCAGATACAGTTTTTAGATCCTTCGTCGTTGGCA
ATGTCAAGGAATGGGGCCAAGGGACACAGGTGACGGTCAGCTCTTAATAGAAGCTT

Length: 388 bp

EG7

TCTAGAACAAGTGCAGCTGGTTGAGTCCGGTGGGGGCCTTGTGCAGCCAGGTGGGTCGTTGCGGCTGTCTTGCGCAGCCAGCG
AATCTTTTTTCAATTTTGATGCGTGGGGCTGGTATCGTCAGGCGCCTGGAAAACAACGTGAAATGGTTGCAGTCGTGGGAAGT
ACGGGATCGACGAGTTACGCTGACTTTGTAAAGGGGAGATTCACAATAAGTAGAGACAACGCAAATAACACCGTTTATCTGCA
AATGAACACTTTACGCCCAGAGGATACAGCTGTATACTATTGTTATGCCCGCTTCCAGAGCTTATACAATTCATGGGGCCAAG
GTACACAAGTAACTGTCTCCTCATAATAGAAGCTT

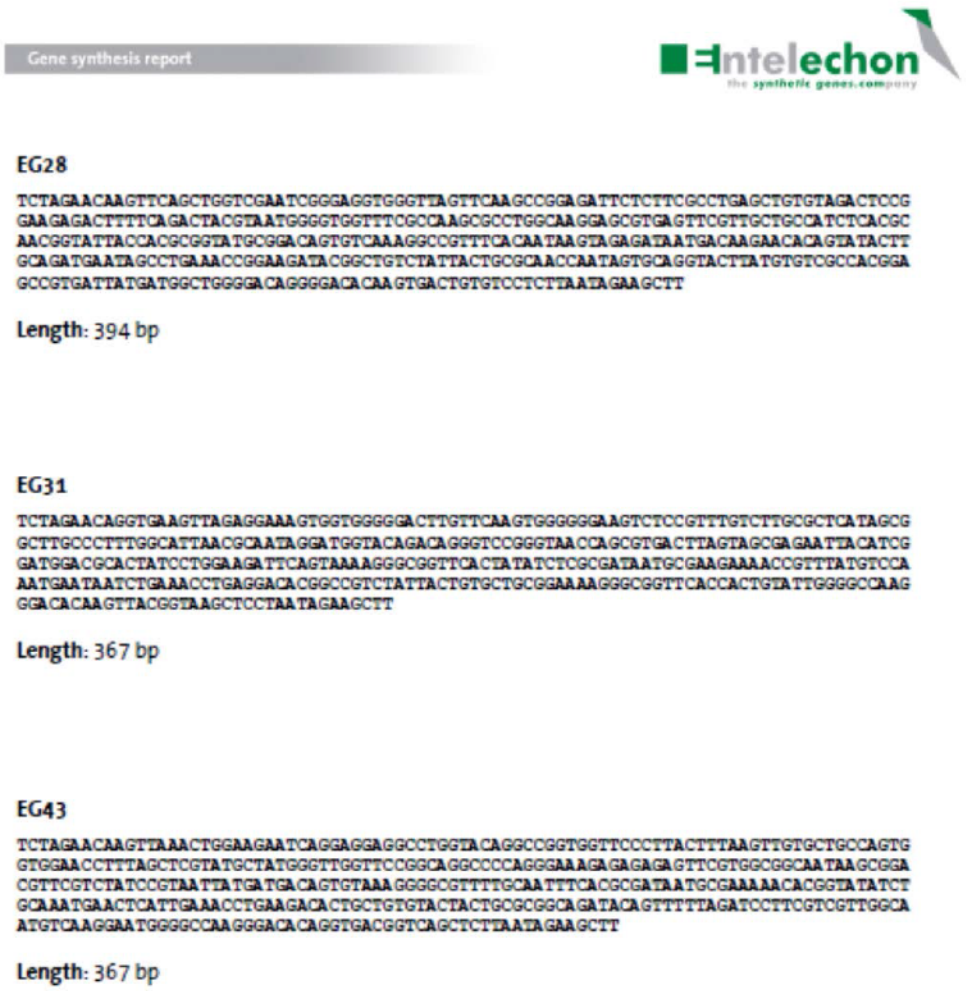
Length: 367 bp

EG10

TCTAGAACAGGTGCAATTAGTTGAGTCGGGTGGGGGACTGGTGCAAGCTGGTGGATCACTGACCTTGTCCTGTGCGGCCAGCG
GTGGCACATTCAGTAGCTATGCCATGGGCTGGTTTAGACAGGCTCCAGGGAAAGAACGCGAATTTGTTGCTGCCATAAGTGGG
CGCTCTTCCATTCGGAATTACGATGACTCTGTCAAGGGACGTTTTGCAATCAGCAGAGACTCGGCGAAGAATACGGTCTATCT
TCAGATGAACTCACTGAAACCTGAGGATACTGCGGTCTACTATTGCGCAGCAGATACTGTATTCCGTTCTTTCGTAGTGGGTA
ACGTTAAAGAATGGGGCCAAGGAACACAAGTAACGGTATCAAGTTAATAGAAGCTT

Length: 388 bp

Figure 131 Entelechon report for single domain antibodies



Gene synthesis report

Entelechon
the synthetic genes company

EG28

TCTAGAACAAGTTCAGCTGGTCGAATCGGGAGGTGGGTTAGTTCAAGCCGGAGATTCTCTTCGCCTGAGCTGTGTAGACTCCG
GAAGAGACTTTTCAGACTACGTAATGGGGTGGTTTCGCCAAGCGCCTGGCAAGGAGCGTGAGTTCGTTGCTGCCATCTCACGC
AACGGTATTACCACGCGGTATGCGGACAGTGTCAAAGGCCGTTTCACAATAAGTAGAGATAATGACAAGAACACAGTATACTT
GCAGATGAATAGCCTGAAACCGGAAGATACGGCTGTCTATTACTGCGCAACCAATAGTGCAGGTACTTATGTGTCGCCACGGA
GCCGTGATTATGATGGCTGGGGACAGGGGACACAAGTGACTGTGTCCTCTTAATAGAAGCTT

Length: 394 bp

EG31

TCTAGAACAGGTGAAGTTAGAGGAAAGTGGTGGGGGACTTGTTCAAGTGGGGGGAAGTCTCCGTTTGTCTTGCGCTCATAGCG
GCTTGCCCTTTGGCATTAACGCAATAGGATGGTACAGACAGGGTCCGGGTAACCAGCGTGACTTAGTAGCGAGAATTACATCG
GATGGACGCACTATCCTGGAAGATTCAGTAAAAGGGCGGTTCACTATATCTCGCGATAATGCGAAGAAAACCGTTTATGTCCA
AATGAATAATCTGAAACCTGAGGACACGGCCGTCTATTACTGTGCTGCGGAAAAGGGCGGTTCACCACTGTATTGGGGCCAAG
GGACACAAGTTACGGTAAGCTCCTAATAGAAGCTT

Length: 367 bp

EG43

TCTAGAACAAGTTAAACTGGAAGAATCAGGAGGAGGCCTGGTACAGGCCGGTGGTTCCCTTACTTTAAGTTGTGCTGCCAGTG
GTGGAACCTTTAGCTCGTATGCTATGGGTTGGTTCCGGCAGGCCCCAGGGAAAGAGAGAGAGTTCGTGGCGGCAATAAGCGGA
CGTTCGTCTATCCGTAATTATGATGACAGTGTAAAGGGGCGTTTTGCAATTTCACGCGATAATGCGAAAAACACGGTATATCT
GCAAATGAACTCATTGAAACCTGAAGACACTGCTGTGTACTACTGCGCGGCAGATACAGTTTTTAGATCCTTCGTCGTTGGCA
ATGTCAAGGAATGGGGCCAAGGGACACAGGTGACGGTCAGCTCTTAATAGAAGCTT

Length: 367 bp

Figure 132 Entelechon report for single domain antibodies.

Date of order: 25.11.2008
Order: synthetic genes cloned into E.coli vector

Ordered sequences:

CD117l

TCTAGAACAAGTCCAGCTGGTACAGAGTGGCGCTGAGGTCAAAAAGCCAGGAGCAAGCGTTAAAGTTTCCTGTAA
GGCGTCGGGTTACACTTTCACAAGTTATAATATGCATTGGGTACGCCAGGCGCCTGGACAAGGGCTTGAATGGAT
GGGTGTCATTTATTCGGGTAACGGTGATACCAGCTATAACCAGAAATTCAAAGGGCGTGTGACGATAACTGCTGA
TAAAAGTACCTCTACAGCTTACATGGAGCTTTCTTCGTTGCGTTCAGAAGACACGGCAGTGTACTATTGTGCACG
GGAGCGCGATACGCGTTTTGGGAATTGGGGACAAGGCACTCTGGTGACAGTATCCAGCGCGTCAGGGGGCGGAGG
TTCTGGGGGAGGCGGATCAGGGGGAGGTGGTTCCGATATTGTCATGACTCAGTCTCCGGACTCCCTGGCCGTTTC
GCTGGGTGAAAGAGCTACGATAAACTGCCGAGCCAGCGAGAGCGTAGACATCTATGGCAACAGTGAAATGCACTG
GTATCAACAGAAACCGGGTCAACCCCCTAAGCTCTTAATCTATATAGCATCGAATATCGAGTCTGGAGTTCCAGA
TCGCTTTTCCGGCAGTGGGTCAGGCACAGACTTTACCTTGACAATAAGTTCATTACAGGCGGAGGACGTGGCAGT
TTACTACTGCCAACAAAATAATGAAGATCCCTACACTTTCGGAGGCGGGACCAAGGTAGAAATTAAGAGATAATA
GAAGCTT

Length: 757 bp

Figure 133 Entelechon report for single chain antibody fragment.

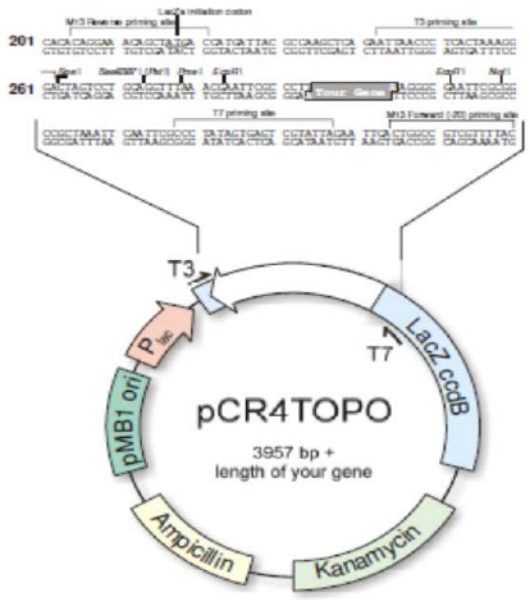
Quality Assurance Documentation

Entelechon
the synthetic genes company

Quality Assurance Documentation

Designation: *E. coli* K12
Vectorbackbone: pCR4-TOPO (Invitrogen)
Host strain: TOP10

F¯ *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) ϕ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *deo*R *ara*D139 Δ(*ara-leu*)7697 *gal*U *gal*K *rps*L (Str^R) *end*A1 *nup*G



The synthetic genes were assembled from HPLC- and PAGE-purified oligonucleotides. The fragments were cloned in pCR4-TOPO (Invitrogen).
The final constructs were verified by sequencing using plasmid specific M13rev and M13 fwd primers. The sequence congruence was 100%. See the accompanying data sheet for sequences.

Note

Plasmid-DNA was isolated from an E.coli strain containing methylases therefore methylation-sensitive restriction enzymes are blocked.

Product of synthesis according to sequence chromatograms:
Sequencing showed correct sequence of synthesis products.

Quantification:

We send you approximately 10 µg of each lyophilized plasmid.

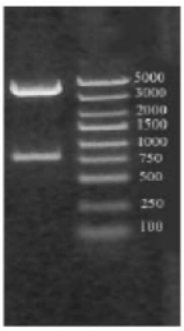
We included the ABI chromatograms and sequence files online under your account. A link to a freeware program with which to view these files can be found on our website at www.entelechon.com in the menu under "Products and Services > Bioinformatics".

Figure 134 pCR4TOPO vector for cloned molecules by Entelechon.

gene CD117I 757 bp
cloned in pCR4 TOPO
orientation M13rev - gene - M13fwd

```
                1          11         21         31         41         51         61
     CD117I.txt: tctagaacaa gtccagctgg tacagagtgg cgctgaggtc aaaaagccag gagcaagcgt taaagtttcc
K263.1_M13f.scf: ---------- ---------- ---------- --------TC AAAAAGCCAG GAGCAAGCGT TAAAGTTTCC
K263.1_M13r.scf: TCTAGAACAA GTCCAGCTGG TACAGAGTGG CGCTGAGGTC AAAAAGCCAG GAGCAAGCGT TAAAGTTTCC

                71         81         91        101        111        121        131
     CD117I.txt: tgtaaggcgt cgggttacac tttcacaagt tataatatgc attgggtacg ccaggcgcct ggacaagggc
K263.1_M13f.scf: TGTAAGGCGT CGGGTTACAC TTTCACAAGT TATAATATGC ATTGGGTACG CCAGGCGCCT GGACAAGGGC
K263.1_M13r.scf: TGTAAGGCGT CGGGTTACAC TTTCACAAGT TATAATATGC ATTGGGTACG CCAGGCGCCT GGACAAGGGC

                141        151        161        171        181        191        201
     CD117I.txt: ttgaatggat gggtgtcatt tattcgggta acggtgatac cagctataac cagaaattca aagggcgtgt
K263.1_M13f.scf: TTGAATGGAT GGGTGTCATT TATTCGGGTA ACGGTGATAC CAGCTATAAC CAGAAATTCA AAGGGCGTGT
K263.1_M13r.scf: TTGAATGGAT GGGTGTCATT TATTCGGGTA ACGGTGATAC CAGCTATAAC CAGAAATTCA AAGGGCGTGT

                211        221        231        241        251        261        271
     CD117I.txt: gacgataact gctgataaaa gtacctctac agcttacatg gagctttctt cgttgcgttc agaagacacg
K263.1_M13f.scf: GACGATAACT GCTGATAAAA GTACCTCTAC AGCTTACATG GAGCTTTCTT CGTTGCGTTC AGAAGACACG
K263.1_M13r.scf: GACGATAACT GCTGATAAAA GTACCTCTAC AGCTTACATG GAGCTTTCTT CGTTGCGTTC AGAAGACACG

                281        291        301        311        321        331        341
     CD117I.txt: gcagtgtact attgtgcacg ggagcgcgat acgcgttttg ggaattgggg acaaggcact ctggtgacag
K263.1_M13f.scf: GCAGTGTACT ATTGTGCACG GGAGCGCGAT ACGCGTTTTG GGAATTGGGG ACAAGGCACT CTGGTGACAG
K263.1_M13r.scf: GCAGTGTACT ATTGTGCACG GGAGCGCGAT ACGCGTTTTG GGAATTGGGG ACAAGGCACT CTGGTGACAG

                351        361        371        381        391        401        411
     CD117I.txt: tatccagcgc gtcagggggc ggaggttctg ggggaggcgg atcaggggga ggtggttccg atattgtcat
K263.1_M13f.scf: TATCCAGCGC GTCAGGGGGC GGAGGTTCTG GGGGAGGCGG ATCAGGGGGA GGTGGTTCCG ATATTGTCAT
K263.1_M13r.scf: TATCCAGCGC GTCAGGGGGC GGAGGTTCTG GGGGAGGCGG ATCAGGGGGA GGTGGTTCCG ATATTGTCAT

                421        431        441        451        461        471        481
     CD117I.txt: gactcagtct ccggactccc tggccgtttc gctgggtgaa agagctacga taaactgccg agccagcgag
K263.1_M13f.scf: GACTCAGTCT CCGGACTCCC TGGCCGTTTC GCTGGGTGAA AGAGCTACGA TAAACTGCCG AGCCAGCGAG
K263.1_M13r.scf: GACTCAGTCT CCGGACTCCC TGGCCGTTTC GCTGGGTGAA AGAGCTACGA TAAACTGCCG AGCCAGCGAG

                491        501        511        521        531        541        551
     CD117I.txt: agcgtagaca tctatggcaa cagtgaaatg cactggtatc aacagaaacc gggtcaaccc cctaagctct
K263.1_M13f.scf: AGCGTAGACA TCTATGGCAA CAGTGAAATG CACTGGTATC AACAGAAACC GGGTCAACCC CCTAAGCTCT
K263.1_M13r.scf: AGCGTAGACA TCTATGGCAA CAGTGAAATG CACTGGTATC AACAGAAACC GGGTCAACCC CCTAAGCTCT

                561        571        581        591        601        611        621
     CD117I.txt: taatctatat agcatcgaat atcgagtctg gagttccaga tcgcttttcc ggcagtgggt caggcacaga
K263.1_M13f.scf: TAATCTATAT AGCATCGAAT ATCGAGTCTG GAGTTCCAGA TCGCTTTTCC GGCAGTGGGT CAGGCACAGA
K263.1_M13r.scf: TAATCTATAT AGCATCGAAT ATCGAGTCTG GAGTTCCAGA TCGCTTTTCC GGCAGTGGGT CAGGCACAGA

                631        641        651        661        671        681        691
     CD117I.txt: ctttaccttg acaataagtt cattacaggc ggaggacgtg gcagtttact actgccaaca aaataatgaa
K263.1_M13f.scf: CTTTACCTTG ACAATAAGTT CATTACAGGC GGAGGACGTG GCAGTTTACT ACTGCCAACA AAATAATGAA
K263.1_M13r.scf: CTTTACCTTG ACAATAAGTT CATTACAGGC GGAGGACGTG G

                701        711        721        731        741        751
     CD117I.txt: gatcoctaca ctttcggagg cgggaccaag gtagaaatta agagataata gaagctt
K263.1_M13f.scf: GATCCCTACA CTTTCGGAGG CGGGACCAAG GTAGAAATTA AGAGATAATA GAAGCTT
```

Quality control:



Gene name: CD117I
Clone ID#:K263.1
res:*Eco*RI

Figure 135 Entelechon CD117I gene report with quality control.

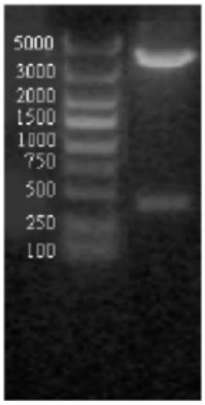
gene EG6 388 bp

cloned in pCR4 TOPO

orientation M13fwd – gene – M13rev

```
                1         11         21         31         41         51         61
       EG6.txt: tctagaacaa gttaaactgg aagaatcagg aggaggcctg gtacaggccg gtggttcoct tactttaagt
K322.1_M13f.scf: TCTAGAACAA GTTAAACTGG AAGAATCAGG AGGAGGCCTG GTACAGGCCG GTGGTTCOCT TACTTTAAGT
K322.1_M13r.scf: TCTAGAACAA GTTAAACTGG AAGAATCAGG AGGAGGCCTG GTACAGGCCG GTGGTTCOCT TACTTTAAGT

                71         81         91        101        111        121        131
       EG6.txt: tgtgctgcca gtggtggaac ctttagctcg tatgctatgg gttggttccg gcaggccoca gggaaagaga
K322.1_M13f.scf: TGTGCTGCCA GTGGTGGAAC CTTTAGCTCG TATGCTATGG GTTGGTTCCG GCAGGCCOCA GGGAAAGAGA
K322.1_M13r.scf: TGTGCTGCCA GTGGTGGAAC CTTTAGCTCG TATGCTATGG GTTGGTTCCG GCAGGCCOCA GGGAAAGAGA

                141        151        161        171        181        191        201
       EG6.txt: gagagttogt ggcggcaata agcggacgtt cgtctatccg taattatgat gacagtgtaa aggggcgttt
K322.1_M13f.scf: GAGAGTTCGT GGCGGCAATA AGCGGACGTT CGTCTATCCG TAATTATGAT GACAGTGTAA AGGGGCGTTT
K322.1_M13r.scf: GAGAGTTCGT GGCGGCAATA AGCGGACGTT CGTCTATCCG TAATTATGAT GACAGTGTAA AGGGGCGTTT

                211        221        231        241        251        261        271
       EG6.txt: tgcaatttca cgcgataatg cgaaaaacac ggtatatctg caaatgaact cattgaaacc tgaagacact
K322.1_M13f.scf: TGCAATTTCA CGCGATAATG CGAAAAACAC GGTATATCTG CAAATGAACT CATTGAAACC TGAAGACACT
K322.1_M13r.scf: TGCAATTTCA CGCGATAATG CGAAAAACAC GGTATATCTG CAAATGAACT CATTGAAACC TGAAGACACT

                281        291        301        311        321        331        341
       EG6.txt: gctgtgtact actgcgcggc agatacagtt tttagatcct tcgtcgttgg caatgtcaag gaatggggcc
K322.1_M13f.scf: GCTGTGTACT ACTGCGCGGC AGATACAGTT TTTAGATCCT TCGTCGTTGG CAATGTCAAG GAATGGGGCC
K322.1_M13r.scf: GCTGTGTACT ACTGCGCGGC AGATACAGTT TTTAGATCCT TCGTCGTTGG CAATGTCAAG GAATGGGGCC

                351        361        371        381
       EG6.txt: aagggacaca ggtgacggtc agctcttaat agaagctt
K322.1_M13f.scf: AAGGGACACA GGTGACGGTC AGCTCTTAAT AGAAGCTT
K322.1_M13r.scf: AAGGGACACA GGTGACGGTC AGCTCTTAAT AGAAGCTT
```

Quality control:



Gene name: EG6
Clone ID#:K322.1
res:*Eco*RI

Figure 136 Entelechon EG6 gene report and quality control.

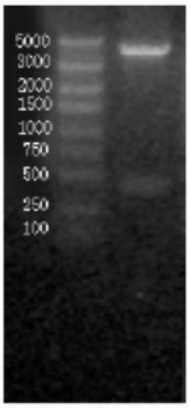
gene EG7 367 bp

cloned in pCR4 TOPO

orientation M13rev – gene – M13fwd

```
                1         11         21         31         41         51         61
       EG7.txt: tctagaacaa gtgcagctgg ttgagtccgg tgggggcctt gtgcagccag gtgggtcgtt gcggctgtct
K323.1_M13f.scf: TCTAGAACAA GTGCAGCTGG TTGAGTCCGG TGGGGGCCTT GTGCAGCCAG GTGGGTCGTT GCGGCTGTCT
K323.1_M13r.scf: TCTAGAACAA GTGCAGCTGG TTGAGTCCGG TGGGGGCCTT GTGCAGCCAG GTGGGTCGTT GCGGCTGTCT

               71         81         91        101        111        121        131
       EG7.txt: tgcgcagcca gcgaatcttt tttcaatttt gatgcgtggg gctggtatcg tcaggcgcct ggaaaacaac
K323.1_M13f.scf: TGCGCAGCCA GCGAATCTTT TTTCAATTTT GATGCGTGGG GCTGGTATCG TCAGGCGCCT GGAAAACAAC
K323.1_M13r.scf: TGCGCAGCCA GCGAATCTTT TTTCAATTTT GATGCGTGGG GCTGGTATCG TCAGGCGCCT GGAAAACAAC

              141        151        161        171        181        191        201
       EG7.txt: gtgaaatggt tgcagtcgtg ggaagtacgg gatcgacgag ttacgctgac tttgtaaagg ggagattcac
K323.1_M13f.scf: GTGAAATGGT TGCAGTCGTG GGAAGTACGG GATCGACGAG TTACGCTGAC TTTGTAAAGG GGAGATTCAC
K323.1_M13r.scf: GTGAAATGGT TGCAGTCGTG GGAAGTACGG GATCGACGAG TTACGCTGAC TTTGTAAAGG GGAGATTCAC

              211        221        231        241        251        261        271
       EG7.txt: aataagtaga gacaacgcaa ataacaccgt ttatctgcaa atgaacactt tacgcccaga ggatacagct
K323.1_M13f.scf: AATAAGTAGA GACAACGCAA ATAACACCGT TTATCTGCAA ATGAACACTT TACGCCCAGA GGATACAGCT
K323.1_M13r.scf: AATAAGTAGA GACAACGCAA ATAACACCGT TTATCTGCAA ATGAACACTT TACGCCCAGA GGATACAGCT

              281        291        301        311        321        331        341
       EG7.txt: gtatactatt gttatgcccg cttccagagc ttatacaatt catggggcca aggtacacaa gtaactgtct
K323.1_M13f.scf: GTATACTATT GTTATGCCCG CTTCCAGAGC TTATACAATT CATGGGGCCA AGGTACACAA GTAACTGTCT
K323.1_M13r.scf: GTATACTATT GTTATGCCCG CTTCCAGAGC TTATACAATT CATGGGGCCA AGGTACACAA GTAACTGTCT

              351        361
       EG7.txt: cctcataata gaagctt
K323.1_M13f.scf: CCTCATAATA GAAGCTT
K323.1_M13r.scf: CCTCATAATA GAAGCTT
```

Quality control:



Gene name: EG7
Clone ID#:K323.1
res:*Eco*RI

Figure 137 Entelechon EG7 gene report and quality control.

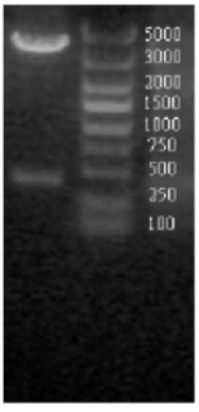
gene EG10 388 bp

cloned in pCR4 TOPO

orientation M13fwd – gene – M13rev

```
                1          11         21         31         41         51         61
      EG10.txt: tctagaacag gtgcaattag ttgagtcggg tgggggactg gtgcaagctg gtggatcact gaccttgtcc
K324.1_M13f.scf: TCTAGAACAG GTGCAATTAG TTGAGTCGGG TGGGGGACTG GTGCAAGCTG GTGGATCACT GACCTTGTCC
K324.1_M13r.scf: TCTAGAACAG GTGCAATTAG TTGAGTCGGG TGGGGGACTG GTGCAAGCTG GTGGATCACT GACCTTGTCC

                71         81         91        101        111        121        131
      EG10.txt: tgtgcggcca gcggtggcac attcagtagc tatgccatgg gctggtttag acaggctcca gggaaagaac
K324.1_M13f.scf: TGTGCGGCCA GCGGTGGCAC ATTCAGTAGC TATGCCATGG GCTGGTTTAG ACAGGCTCCA GGGAAAGAAC
K324.1_M13r.scf: TGTGCGGCCA GCGGTGGCAC ATTCAGTAGC TATGCCATGG GCTGGTTTAG ACAGGCTCCA GGGAAAGAAC

               141        151        161        171        181        191        201
      EG10.txt: gcgaatttgt tgctgccata agtgggcgct cttccattcg gaattacgat gactctgtca agggacgttt
K324.1_M13f.scf: GCGAATTTGT TGCTGCCATA AGTGGGCGCT CTTCCATTCG GAATTACGAT GACTCTGTCA AGGGACGTTT
K324.1_M13r.scf: GCGAATTTGT TGCTGCCATA AGTGGGCGCT CTTCCATTCG GAATTACGAT GACTCTGTCA AGGGACGTTT

               211        221        231        241        251        261        271
      EG10.txt: tgcaatcagc agagactcgg cgaagaatac ggtctatctt cagatgaact cactgaaacc tgaggatact
K324.1_M13f.scf: TGCAATCAGC AGAGACTCGG CGAAGAATAC GGTCTATCTT CAGATGAACT CACTGAAACC TGAGGATACT
K324.1_M13r.scf: TGCAATCAGC AGAGACTCGG CGAAGAATAC GGTCTATCTT CAGATGAACT CACTGAAACC TGAGGATACT

               281        291        301        311        321        331        341
      EG10.txt: gcggtctact attgcgcagc agatactgta ttccgttctt tcgtagtggg taacgttaaa gaatggggcc
K324.1_M13f.scf: GCGGTCTACT ATTGCGCAGC AGATACTGTA TTCCGTTCTT TCGTAGTGGG TAACGTTAAA GAATGGGGCC
K324.1_M13r.scf: GCGGTCTACT ATTGCGCAGC AGATACTGTA TTCCGTTCTT TCGTAGTGGG TAACGTTAAA GAATGGGGCC

               351        361        371        381
      EG10.txt: aaggaacaca agtaacggta tcaagttaat agaagctt
K324.1_M13f.scf: AAGGAACACA AGTAACGGTA TCAAGTTAAT AGAAGCTT
K324.1_M13r.scf: AAGGAACACA AGTAACGGTA TCAAGTTAAT AGAAGCTT
```

Quality control:



Gene name: EG10
Clone ID#:K324.1
res:*Eco*RI

Figure 138 Entelechon EG10 gene report and quality control.

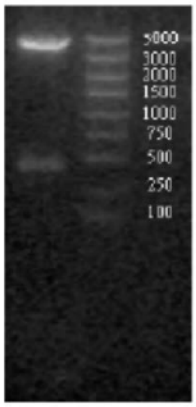
gene EG28 394 bp

cloned in pCR4 TOPO

orientation M13fwd – gene – M13rev

```
                1          11         21         31         41         51         61
      EG28.txt: tctagaacaa gttcagctgg tcgaatcggg aggtgggtta gttcaagccg gagattctct tcgcctgagc
K325.1_M13f.scf: TCTAGAACAA GTTCAGCTGG TCGAATCGGG AGGTGGGTTA GTTCAAGCCG GAGATTCTCT TCGCCTGAGC
K325.1_M13r.scf: TCTAGAACAA GTTCAGCTGG TCGAATCGGG AGGTGGGTTA GTTCAAGCCG GAGATTCTCT TCGCCTGAGC

                71         81         91        101        111        121        131
      EG28.txt: tgtgtagact ccggaagaga cttttcagac tacgtaatgg ggtggtttcg ccaagcgcct ggcaaggagc
K325.1_M13f.scf: TGTGTAGACT CCGGAAGAGA CTTTTCAGAC TACGTAATGG GGTGGTTTCG CCAAGCGCCT GGCAAGGAGC
K325.1_M13r.scf: TGTGTAGACT CCGGAAGAGA CTTTTCAGAC TACGTAATGG GGTGGTTTCG CCAAGCGCCT GGCAAGGAGC

               141        151        161        171        181        191        201
      EG28.txt: gtgagttcgt tgctgccatc tcacgcaacg gtattaccac gcggtatgcg gacagtgtca aaggccgttt
K325.1_M13f.scf: GTGAGTTCGT TGCTGCCATC TCACGCAACG GTATTACCAC GCGGTATGCG GACAGTGTCA AAGGCCGTTT
K325.1_M13r.scf: GTGAGTTCGT TGCTGCCATC TCACGCAACG GTATTACCAC GCGGTATGCG GACAGTGTCA AAGGCCGTTT

               211        221        231        241        251        261        271
      EG28.txt: cacaataagt agagataatg acaagaacac agtatacttg cagatgaata gcctgaaacc ggaagatacg
K325.1_M13f.scf: CACAATAAGT AGAGATAATG ACAAGAACAC AGTATACTTG CAGATGAATA GCCTGAAACC GGAAGATACG
K325.1_M13r.scf: CACAATAAGT AGAGATAATG ACAAGAACAC AGTATACTTG CAGATGAATA GCCTGAAACC GGAAGATACG

               281        291        301        311        321        331        341
      EG28.txt: gctgtctatt actgcgcaac caatagtgca ggtacttatg tgtcgccacg gagccgtgat tatgatggct
K325.1_M13f.scf: GCTGTCTATT ACTGCGCAAC CAATAGTGCA GGTACTTATG TGTCGCCACG GAGCCGTGAT TATGATGGCT
K325.1_M13r.scf: GCTGTCTATT ACTGCGCAAC CAATAGTGCA GGTACTTATG TGTCGCCACG GAGCCGTGAT TATGATGGCT

               351        361        371        381        391
      EG28.txt: ggggacaggg gacacaagtg actgtgtcct cttaatagaa gctt
K325.1_M13f.scf: GGGGACAGGG GACACAAGTG ACTGTGTCCT CTTAATAGAA GCTT
K325.1_M13r.scf: GGGGACAGGG GACACAAGTG ACTGTGTCCT CTTAATAGAA GCTT
```

Quality control:



Gene name: EG28
Clone ID#:K325.1
res:*Eco*RI

Figure 139 Entelechon EG28 gene report and quality control.

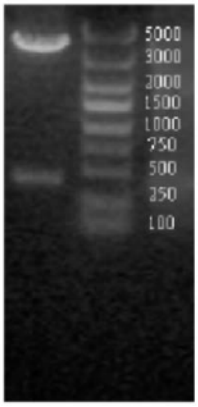
gene EG31 367 bp

cloned in pCR4 TOPO

orientation M13fwd – gene – M13rev

```
                1          11         21         31         41         51         61
      EG31.txt: tctagaacag gtgaagttag aggaaagtgg tgggggactt gttcaagtgg ggggaagtct cogtttgtct
K326.1_M13f.scf: TCTAGAACAG GTGAAGTTAG AGGAAAGTGG TGGGGGACTT GTTCAAGTGG GGGGAAGTCT CGGTTTGTCT
K326.1_M13r.scf: TCTAGAACAG GTGAAGTTAG AGGAAAGTGG TGGGGGACTT GTTCAAGTGG GGGGAAGTCT CGGTTTGTCT

                71         81         91        101        111        121        131
      EG31.txt: tgcgctcata gcggcttgcc ctttggcatt aacgcaatag gatggtacag acagggtccg ggtaaccagc
K326.1_M13f.scf: TGCGCTCATA GCGGCTTGCC CTTTGGCATT AACGCAATAG GATGGTACAG ACAGGGTCCG GGTAACCAGC
K326.1_M13r.scf: TGCGCTCATA GCGGCTTGCC CTTTGGCATT AACGCAATAG GATGGTACAG ACAGGGTCCG GGTAACCAGC

                141        151        161        171        181        191        201
      EG31.txt: gtgacttagt agcgagaatt acatcggatg gacgcactat cctggaagat tcagtaaaag ggcggttcac
K326.1_M13f.scf: GTGACTTAGT AGCGAGAATT ACATCGGATG GACGCACTAT CCTGGAAGAT TCAGTAAAAG GGCGGTTCAC
K326.1_M13r.scf: GTGACTTAGT AGCGAGAATT ACATCGGATG GACGCACTAT CCTGGAAGAT TCAGTAAAAG GGCGGTTCAC

                211        221        231        241        251        261        271
      EG31.txt: tatatctcgc gataatgcga agaaaacgt ttatgtccaa atgaataatc tgaaacctga ggacacggcc
K326.1_M13f.scf: TATATCTCGC GATAATGCGA AGAAAACCGT TTATGTCCAA ATGAATAATC TGAAACCTGA GGACACGGCC
K326.1_M13r.scf: TATATCTCGC GATAATGCGA AGAAAACCGT TTATGTCCAA ATGAATAATC TGAAACCTGA GGACACGGCC

                281        291        301        311        321        331        341
      EG31.txt: gtctattact gtgctgcgga aaagggcggt tcaccactgt attggggcca agggacacaa gttacggtaa
K326.1_M13f.scf: GTCTATTACT GTGCTGCGGA AAAGGGCGGT TCACCACTGT ATTGGGGCCA AGGGACACAA GTTACGGTAA
K326.1_M13r.scf: GTCTATTACT GTGCTGCGGA AAAGGGCGGT TCACCACTGT ATTGGGGCCA AGGGACACAA GTTACGGTAA

                351        361
      EG31.txt: gctcctaata gaagctt
K326.1_M13f.scf: GCTCCTAATA GAAGCTT
K326.1_M13r.scf: GCTCCTAATA GAAGCTT
```

Quality control:



Gene name: EG31
Clone ID#:K326.1
res:*Eco*RI

Figure 140 Entelechon E31 gene report and quality control.

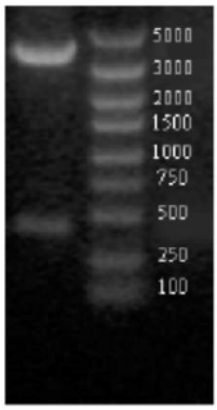
gene EG43 367 bp

cloned in pCR4 TOPO

orientation M13fwd – gene – M13rev

```
                1         11         21         31         41         51         61
      EG43.txt: tctagaacag gtacaattgg tggaatctgg agggggctta gttcaaccog gtggatctct cogtctgcca
K327.1_M13f.scf: TCTAGAACAG GTACAATTGG TGGAATCTGG AGGGGGCTTA GTTCAACCCG GTGGATCTCT CCGTCTGCCA
K327.1_M13r.scf: TCTAGAACAG GTACAATTGG TGGAATCTGG AGGGGGCTTA GTTCAACCCG GTGGATCTCT CCGTCTGCCA

                71         81         91        101        111        121        131
      EG43.txt: tgtgccgcaa gtggctcgat tttcagcctg gacgcttggg gatggtatcg ccaggcgcct ggcaaacaac
K327.1_M13f.scf: TGTGCCGCAA GTGGCTCGAT TTTCAGCCTG GACGCTTGGG GATGGTATCG CCAGGCGCCT GGCAAACAAC
K327.1_M13r.scf: TGTGCCGCAA GTGGCTCGAT TTTCAGCCTG GACGCTTGGG GATGGTATCG CCAGGCGCCT GGCAAACAAC

                141        151        161        171        181        191        201
      EG43.txt: gggagatggt cgcacttgta ggttccgatg ggtccacaag ttatgctgac agtgttaagg ggagattcac
K327.1_M13f.scf: GGGAGATGGT CGCACTTGTA GGTTCCGATG GGTCCACAAG TTATGCTGAC AGTGTTAAGG GGAGATTCAC
K327.1_M13r.scf: GGGAGATGGT CGCACTTGTA GGTTCCGATG GGTCCACAAG TTATGCTGAC AGTGTTAAGG GGAGATTCAC

                211        221        231        241        251        261        271
      EG43.txt: cataagccgc gataatgcga ataacacttt ttacctgcaa atgaactcgt tgaaaccgga agatacggca
K327.1_M13f.scf: CATAAGCCGC GATAATGCGA ATAACACTTT TTACCTGCAA ATGAACTCGT TGAAACCGGA AGATACGGCA
K327.1_M13r.scf: CATAAGCCGC GATAATGCGA ATAACACTTT TTACCTGCAA ATGAACTCGT TGAAACCGGA AGATACGGCA

                281        291        301        311        321        331        341
      EG43.txt: gtgtattact gctatgcccg ttttcagtca ttatacaata gctggggaca gggtacacaa gtaactgtct
K327.1_M13f.scf: GTGTATTACT GCTATGCCCG TTTTCAGTCA TTATACAATA GCTGGGGACA GGGTACACAA GTAACTGTCT
K327.1_M13r.scf: GTGTATTACT GCTATGCCCG TTTTCAGTCA TTATACAATA GCTGGGGACA GGGTACACAA GTAACTGTCT

                351        361
      EG43.txt: catcataata gaagctt
K327.1_M13f.scf: CATCATAATA GAAGCTT
K327.1_M13r.scf: CATCATAATA GAAGCTT
```

Quality control:



Gene name: EG43
Clone ID#:K327.1
res:*Eco*RI

Figure 141 Entelechon EG43 gene report and quality control.

References

Andersen, D. C. and D. E. Reilly (2004). "Production technologies for monoclonal antibodies and their fragments." Curr Opin Biotechnol 15(5): 456-62.

Arbabi Ghahroudi, M., A. Desmyter, et al. (1997). "Selection and identification of single domain antibody fragments from camel heavy-chain antibodies." FEBS Lett 414(3): 521-6.

Arndt, K. M., K. M. Muller, et al. (1998). "Factors influencing the dimer to monomer transition of an antibody single-chain Fv fragment." Biochemistry 37(37): 12918-26.

Bach, H., Y. Mazor, et al. (2001). "Escherichia coli maltose-binding protein as a molecular chaperone for recombinant intracellular cytoplasmic single-chain antibodies." J Mol Biol 312(1): 79-93.

Bajjalieh, S. M. (1999). "Synaptic vesicle docking and fusion." Curr Opin Neurobiol 9(3): 321-8.

Balbas, P., X. Soberon, et al. (1986). "Plasmid vector pBR322 and its special-purpose derivatives--a review." Gene 50(1-3): 3-40.

Bhidayasiri, R. and D. D. Truong (2005). "Expanding use of botulinum toxin." J Neurol Sci 235(1-2): 1-9.

Beck, A., Wurch, T., Corvaia, N. (2008). "Therapeutic antibodies and derivatives: from the bench to the clinic". Curr Pharm Biotechnol 9(6): 412-2.

Binz, T. and A. Rummel (2009). "Cell entry strategy of clostridial neurotoxins." J Neurochem 109(6): 1584-95.

Biocca, S., P. Pierandrei-Amaldi, et al. (1993). "Intracellular expression of anti-p21ras single chain Fv fragments inhibits meiotic maturation of xenopus oocytes." Biochem Biophys Res Commun 197(2): 422-7.

Bird, R. E., K. D. Hardman, et al. (1988). "Single-chain antigen-binding proteins." Science 242(4877): 423-6.

Birtalan, S., Y. Zhang, et al. (2008). "The intrinsic contributions of tyrosine, serine, glycine and arginine to the affinity and specificity of antibodies." J Mol Biol 377(5): 1518-28.

Boado, R. J., A. Ji, et al. (2000). "Cloning and expression in Pichia pastoris of a genetically engineered single chain antibody against the rat transferrin receptor." J Drug Target 8(6): 403-12.

Bock, J. B., H. T. Matern, et al. (2001). "A genomic perspective on membrane compartment organization." Nature 409(6822): 839-41.

Bond, C. J., C. Wiesmann, et al. (2005). "A structure-based database of antibody variable domain diversity." J Mol Biol 348(3): 699-709.

Brinkmann, U., Y. Reiter, et al. (1993). "A recombinant immunotoxin containing a disulfide-stabilized Fv fragment." Proc Natl Acad Sci U S A 90(16): 7538-42.

Bronfman, F. C., M. Tcherpakov, et al. (2003). "Ligand-induced internalization of the p75 neurotrophin receptor: a slow route to the signaling endosome." J Neurosci 23(8): 3209-20.

Bruenke, J., B. Fischer, et al. (2004). "A recombinant bispecific single-chain Fv antibody against HLA class II and FcgammaRIII (CD16) triggers effective lysis of lymphoma cells." Br J Haematol 125(2): 167-79.

Cattaneo, A. and S. Biocca (1999). "The selection of intracellular antibodies." Trends Biotechnol 17(3): 115-21.

Chaddock, J. A., M. H. Herbert, et al. (2002). "Expression and purification of catalytically active, non-toxic endopeptidase derivatives of Clostridium botulinum toxin type A." Protein Expr Purif 25(2): 219-28.

Chaddock, J. A. and P. M. Marks (2006). "Clostridial neurotoxins: structure-function led design of new therapeutics." Cell Mol Life Sci 63(5): 540-51.

Chaddock, J. A., J. R. Purkiss, et al. (2004). "Retargeted clostridial endopeptidases: inhibition of nociceptive neurotransmitter release in vitro, and antinociceptive activity in in vivo models of pain." Mov Disord 19 Suppl 8: S42-7.

Chaddock, J. A., J. R. Purkiss, et al. (2000). "A conjugate composed of nerve growth factor coupled to a non-toxic derivative of Clostridium botulinum neurotoxin type A can inhibit neurotransmitter release in vitro." Growth Factors 18(2): 147-55.

Chaddock, J. A., J. R. Purkiss, et al. (2000). "Inhibition of vesicular secretion in both neuronal and nonneuronal cells by a retargeted endopeptidase derivative of Clostridium botulinum neurotoxin type A." Infect Immun 68(5): 2587-93.

Chai, Q., J. W. Arndt, et al. (2006). "Structural basis of cell surface receptor recognition by botulinum neurotoxin B." Nature 444(7122): 1096-100.

Coloma, M. J., A. Hastings, et al. (1992). "Novel vectors for the expression of antibody molecules using variable regions generated by polymerase chain reaction." J Immunol Methods 152(1): 89-104.

Coppieters, K., T. Dreier, et al. (2006). "Formatted anti-tumor necrosis factor alpha VHH proteins derived from camelids show superior potency and targeting to inflamed joints in a murine model of collagen-induced arthritis." Arthritis Rheum 54(6): 1856-66.

Cortez-Retamozo, V., N. Backmann, et al. (2004). "Efficient cancer therapy with a nanobody-based conjugate." Cancer Res 64(8): 2853-7.

Crouch, E. R. (2006). "Use of botulinum toxin in strabismus." Curr Opin Ophthalmol 17(5): 435-40.

Cunha, A. E., J. J. Clemente, et al. (2004). "Methanol induction optimization for scFv antibody fragment production in Pichia pastoris." Biotechnol Bioeng 86(4): 458-67.

DasGupta, B. R. and D. A. Boroff (1967). "Chromatographic isolation of hemagglutinin-free neurotoxin from crystalline toxin of Clostridium botulinum type A." Biochim Biophys Acta 147(3): 603-5.

DasGupta, B. R. and D. A. Boroff (1968). "Separation of toxin and hemagglutinin from crystalline toxin of Clostridium botulinum type A by anion exchange chromatography and determination of their dimensions by gel filtration." J Biol Chem 243(5): 1065-72.

DasGupta, B. R., D. A. Boroff, et al. (1966). "Chromatographic fractionation of the crystalline toxin of Clostridium botulinum type A." Biochem Biophys Res Commun 22(6): 750-6.

De Genst, E., D. Saerens, et al. (2006). "Antibody repertoire development in camelids." Dev Comp Immunol 30(1-2): 187-98.

De Genst, E., D. Saerens, et al. (2006). "Antibody repertoire development in camelids." Dev Comp Immunol 30(1-2): 187-98.

Decanniere, K., S. Muyldermans, et al. (2000). "Canonical antigen-binding loop structures in immunoglobulins: more structures, more canonical classes?" J Mol Biol 300(1): 83-91.

Dong, M., H. Liu, et al. (2008). "Glycosylated SV2A and SV2B mediate the entry of botulinum neurotoxin E into neurons." Mol Biol Cell 19(12): 5226-37.

Dong, M., D. A. Richards, et al. (2003). "Synaptotagmins I and II mediate entry of botulinum neurotoxin B into cells." J Cell Biol 162(7): 1293-303.

Dong, M., F. Yeh, et al. (2006). "SV2 is the protein receptor for botulinum neurotoxin A." Science 312(5773): 592-6.

Dressler, D., F. A. Saberi, et al. (2005). "Botulinum toxin: mechanisms of action." Arq Neuropsiquiatr 63(1): 180-5.

Druker, B. J. and N. B. Lydon (2000). "Lessons learned from the development of an abl tyrosine kinase inhibitor for chronic myelogenous leukemia." J Clin Invest 105(1): 3-7.

Duggan, M. J., C. P. Quinn, et al. (2002). "Inhibition of release of neurotransmitters from rat dorsal root ganglia by a novel conjugate of a Clostridium botulinum toxin A endopeptidase fragment and Erythrina cristagalli lectin." J Biol Chem 277(38): 34846-52.

Duggan, M. J., C. P. Quinn, et al. (2002). "Inhibition of release of neurotransmitters from rat dorsal root ganglia by a novel conjugate of a Clostridium botulinum toxin A endopeptidase fragment and Erythrina cristagalli lectin." J Biol Chem 277(38): 34846-52.

Dumoulin, M., K. Conrath, et al. (2002). "Single-domain antibody fragments with high conformational stability." Protein Sci 11(3): 500-15.

Edelman, G. M., B. A. Cunningham, et al. (1969). "The covalent structure of an entire gammaG immunoglobulin molecule." Proc Natl Acad Sci U S A 63(1): 78-85.

Eldin, P., M. E. Pauza, et al. (1997). "High-level secretion of two antibody single chain Fv fragments by Pichia pastoris." J Immunol Methods 201(1): 67-75.

Eleopra, R., V. Tugnoli, et al. (1998). "Different time courses of recovery after poisoning with botulinum neurotoxin serotypes A and E in humans." Neurosci Lett 256(3): 135-8.

Eleopra, R., V. Tugnoli, et al. (1997). "Botulinum neurotoxin serotype C: a novel effective botulinum toxin therapy in human." Neurosci Lett 224(2): 91-4.

Erbguth, F. J. and M. Naumann (1999). "Historical aspects of botulinum toxin: Justinus Kerner (1786-1862) and the "sausage poison"." Neurology 53(8): 1850-3.

Evans, D. M., R. S. Williams, et al. (1986). "Botulinum neurotoxin type B. Its purification, radioiodination and interaction with rat-brain synaptosomal membranes." Eur J Biochem 154(2): 409-16.

Ewert, S., T. Huber, et al. (2003). "Biophysical properties of human antibody variable domains." J Mol Biol 325(3): 531-53.

Fasshauer, D., R. B. Sutton, et al. (1998). "Conserved structural features of the synaptic fusion complex: SNARE proteins reclassified as Q- and R-SNAREs." Proc Natl Acad Sci U S A 95(26): 15781-6.

Filpula, D. (2007). "Antibody engineering and modification technologies." Biomol Eng 24(2): 201-15.

Foran, P. G., N. Mohammed, et al. (2003). "Evaluation of the therapeutic usefulness of botulinum neurotoxin B, C1, E, and F compared with the long lasting type A. Basis for

distinct durations of inhibition of exocytosis in central neurons." J Biol Chem 278(2): 1363-71.

Foster, K. A. (2004). "The analgesic potential of clostridial neurotoxin derivatives." Expert Opin Investig Drugs 13(11): 1437-43.

Foster, K. A., E. J. Adams, et al. (2006). "Re-engineering the target specificity of Clostridial neurotoxins - a route to novel therapeutics." Neurotox Res 9(2-3): 101-7.

Foster, K. A., H. Bigalke, et al. (2006). "Botulinum neurotoxin - from laboratory to bedside." Neurotox Res 9(2-3): 133-40.

Frenken, L. G., R. H. van der Linden, et al. (2000). "Isolation of antigen specific llama VHH antibody fragments and their high level secretion by Saccharomyces cerevisiae." J Biotechnol 78(1): 11-21.

Frenken, L. G., R. H. van der Linden, et al. (2000). "Isolation of antigen specific llama VHH antibody fragments and their high level secretion by Saccharomyces cerevisiae." J Biotechnol 78(1): 11-21.

Fukuda, R., J. A. McNew, et al. (2000). "Functional architecture of an intracellular membrane t-SNARE." Nature 407(6801): 198-202.

Gasser, B. and D. Mattanovich (2007). "Antibody production with yeasts and filamentous fungi: on the road to large scale?" Biotechnol Lett 29(2): 201-12.

Gerngross, T. U. (2004). "Advances in the production of human therapeutic proteins in yeasts and filamentous fungi." Nat Biotechnol 22(11): 1409-14.

Glockshuber, R., M. Malia, et al. (1990). "A comparison of strategies to stabilize immunoglobulin Fv-fragments." Biochemistry 29(6): 1362-7.

Gold, L., D. Pribnow, et al. (1981). "Translational initiation in prokaryotes." Annu Rev Microbiol 35: 365-403.
Gottlieb, P. D., B. A. Cunningham, et al. (1968). "Variable regions of heavy and light polypeptide chains of the same gammaG-immunoglobulin molecule." Proc Natl Acad Sci U S A 61(1): 168-75.

Gueorguieva, D., S. Li, et al. (2006). "Identification of single-domain, Bax-specific intrabodies that confer resistance to mammalian cells against oxidative-stress-induced apoptosis." FASEB J 20(14): 2636-8.

Gura, T. (2002). "Therapeutic antibodies: magic bullets hit the target." Nature 417(6889): 584-6.

Gurkan, C., S. N. Symeonides, et al. (2004). "High-level production in Pichia pastoris of an anti-p185HER-2 single-chain antibody fragment using an alternative secretion expression vector." Biotechnol Appl Biochem 39(Pt 1): 115-22.

Hackel, B. J., D. Huang, et al. (2006). "Production of soluble and active transferrin receptor-targeting single-chain antibody using Saccharomyces cerevisiae." Pharm Res 23(4): 790-7.

Hallis, B., B. A. James, et al. (1996). "Development of novel assays for botulinum type A and B neurotoxins based on their endopeptidase activities." J Clin Microbiol 34(8): 1934-8.

Halpern, J. L. and A. Loftus (1993). "Characterization of the receptor-binding domain of tetanus toxin." J Biol Chem 268(15): 11188-92.

Halpern, J. L. and A. Loftus (1993). "Characterization of the receptor-binding domain of tetanus toxin." J Biol Chem 268(15): 11188-92.

Hamers-Casterman, C., T. Atarhouch, et al. (1993). "Naturally occurring antibodies devoid of light chains." Nature 363(6428): 446-8.

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

Harmsen, M. M., C. B. Van Solt, et al. (2005). "Prolonged in vivo residence times of llama single-domain antibody fragments in pigs by binding to porcine immunoglobulins." Vaccine 23(41): 4926-34.

Hellwig, S., F. Emde, et al. (2001). "Analysis of single-chain antibody production in Pichia pastoris using on-line methanol control in fed-batch and mixed-feed fermentations." Biotechnol Bioeng 74(4): 344-52.

Holliger, P. and P. J. Hudson (2005). "Engineered antibody fragments and the rise of single domains." Nat Biotechnol 23(9): 1126-36.

Holt, L. J., C. Herring, et al. (2003). "Domain antibodies: proteins for therapy." Trends Biotechnol 21(11): 484-90.

Hoogenboom, H. R. (2005). "Selecting and screening recombinant antibody libraries." Nat Biotechnol 23(9): 1105-16.

Hu, X., R. O'Dwyer, et al. (2005). "Cloning, expression and characterisation of a single-chain Fv antibody fragment against domoic acid in Escherichia coli." J Biotechnol 120(1): 38-45.

Humeau, Y., F. Doussau, et al. (2000). "How botulinum and tetanus neurotoxins block neurotransmitter release." Biochimie 82(5): 427-46.

Humeau, Y., F. Doussau, et al. (2000). "How botulinum and tetanus neurotoxins block neurotransmitter release." Biochimie 82(5): 427-46.

Irie, F., M. Okuno, et al. (2005). "EphrinB-EphB signalling regulates clathrin-mediated endocytosis through tyrosine phosphorylation of synaptojanin 1." Nat Cell Biol 7(5): 501-9.

Jankovic, J. (2004). "Botulinum toxin in clinical practice." J Neurol Neurosurg Psychiatry 75(7): 951-7.

Jefferis, R. (2005). "Glycosylation of recombinant antibody therapeutics." Biotechnol Prog 21(1): 11-6.

Joosten, V., M. S. Roelofs, et al. (2005). "Production of bifunctional proteins by Aspergillus awamori: llama variable heavy chain antibody fragment (V(HH)) R9 coupled to Arthromyces ramosus peroxidase (ARP)." J Biotechnol 120(4): 347-59.

Jung, C. H., Y. S. Yang, et al. (2009). "Inhibition of SNARE-driven neuroexocytosis by plant extracts." Biotechnol Lett 31(3): 361-9.

Kapust, R. B. and D. S. Waugh (1999). "Escherichia coli maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused." Protein Sci 8(8): 1668-74.

Kaufmann, M., P. Lindner, et al. (2002). "Crystal structure of the anti-His tag antibody 3D5 single-chain fragment complexed to its antigen." J Mol Biol 318(1): 135-47.

Kobayashi, T., N. Kai, et al. (2008). "Transient silencing of synaptic transmitter release from specific neuronal types by recombinant tetanus toxin light chain fused to antibody variable region." J Neurosci Methods 175(1): 125-32.

Kohler, G. and C. Milstein (1975). "Continuous cultures of fused cells secreting antibody of predefined specificity." Nature 256(5517): 495-7.

Koriazova, L. K. and M. Montal (2003). "Translocation of botulinum neurotoxin light chain protease through the heavy chain channel." Nat Struct Biol 10(1): 13-8.

Kortt, A. A., O. Dolezal, et al. (2001). "Dimeric and trimeric antibodies: high avidity scFvs for cancer targeting." Biomol Eng 18(3): 95-108.

Kozaki, S., S. Miyazaki, et al. (1977). "Development of antitoxin with each of two complementary fragments of Clostridium botulinum type B derivative toxin." Infect Immun 18(3): 761-6.

Kozaki, S., S. Sugii, et al. (1975). "Proceedings: Clostridium botulinum type A, B, E and F 12S toxins." Jpn J Med Sci Biol 28(1): 70-2.

Kozaki, S., S. Togashi, et al. (1981). "Separation of Clostridium botulinum type A derivative toxin into two fragments." Jpn J Med Sci Biol 34(2): 61-8.

Krysinski, E. P. and H. Sugiyama (1980). "Purification and some properties of H chain subunit of type A botulinum neurotoxin." Toxicon 18(5-6): 705-10.

Lacy, D. B., W. Tepp, et al. (1998). "Crystal structure of botulinum neurotoxin type A and implications for toxicity." Nat Struct Biol 5(10): 898-902.

Landsteiner, K. (1942). "Serological Reactivity of Hydrolytic Products from Silk." J Exp Med 75(3): 269-76.

LaVallie, E. R., E. A. DiBlasio, et al. (1993). "A thioredoxin gene fusion expression system that circumvents inclusion body formation in the E. coli cytoplasm." Biotechnology (N Y) 11(2): 187-93.

LaVallie, E. R. and J. M. McCoy (1995). "Gene fusion expression systems in Escherichia coli." Curr Opin Biotechnol 6(5): 501-6.

Lilley, G. G., O. Dolezal, et al. (1994). "Recombinant single-chain antibody peptide conjugates expressed in Escherichia coli for the rapid diagnosis of HIV." J Immunol Methods 171(2): 211-26.

Lis, H. and N. Sharon (1986). "Lectins as molecules and as tools." Annu Rev Biochem 55: 35-67.

Mahrhold, S., A. Rummel, et al. (2006). "The synaptic vesicle protein 2C mediates the uptake of botulinum neurotoxin A into phrenic nerves." FEBS Lett 580(8): 2011-4.

Maina, C. V., P. D. Riggs, et al. (1988). "An Escherichia coli vector to express and purify foreign proteins by fusion to and separation from maltose-binding protein." Gene 74(2): 365-73.

Maisey, E. A., J. D. Wadsworth, et al. (1988). "Involvement of the constituent chains of botulinum neurotoxins A and B in the blockade of neurotransmitter release." Eur J Biochem 177(3): 683-91.

Makvandi-Nejad, S., M. D. McLean, et al. (2005). "Transgenic tobacco plants expressing a dimeric single-chain variable fragment (scfv) antibody against Salmonella enterica serotype Paratyphi B." Transgenic Res 14(5): 785-92.

Martin, C. D., G. Rojas, et al. (2006). "A simple vector system to improve performance and utilisation of recombinant antibodies." BMC Biotechnol 6: 46.

Martineau, P., P. Jones, et al. (1998). "Expression of an antibody fragment at high levels in the bacterial cytoplasm." J Mol Biol 280(1): 117-27.

Masuyer, G., N. Thiyagarajan, et al. (2009). "Crystal structure of a catalytically active, non-toxic endopeptidase derivative of Clostridium botulinum toxin A." Biochem Biophys Res Commun 381(1): 50-3.

Mayfield, S. P., S. E. Franklin, et al. (2003). "Expression and assembly of a fully active antibody in algae." Proc Natl Acad Sci U S A 100(2): 438-42.

Miller, K. D., J. Weaver-Feldhaus, et al. (2005). "Production, purification, and characterization of human scFv antibodies expressed in Saccharomyces cerevisiae, Pichia pastoris, and Escherichia coli." Protein Expr Purif 42(2): 255-67.

Moberg, L. J. and H. Sugiyama (1978). "Affinity chromatography purification of type A botulinum neurotoxin from crystalline toxic complex." Appl Environ Microbiol 35(5): 878-80.

Montecucco, C., E. Papini, et al. (1994). "Bacterial protein toxins penetrate cells via a four-step mechanism." FEBS Lett 346(1): 92-8.

Montecucco, C., G. Schiavo, et al. (2005). "SNARE complexes and neuroexocytosis: how many, how close?" Trends Biochem Sci 30(7): 367-72.

Montero, C. (2003). "The antigen-antibody reaction in immunohistochemistry." J Histochem Cytochem 51(1): 1-4.

Muruganandam, A., J. Tanha, et al. (2002). "Selection of phage-displayed llama single-domain antibodies that transmigrate across human blood-brain barrier endothelium." FASEB J 16(2): 240-2.

Muyldermans, S. (2001). "Single domain camel antibodies: current status." J Biotechnol 74(4): 277-302.

Natsume, A., M. Wakitani, et al. (2006). "Fucose removal from complex-type oligosaccharide enhances the antibody-dependent cellular cytotoxicity of single-gene-encoded bispecific antibody comprising of two single-chain antibodies linked to the antibody constant region." J Biochem 140(3): 359-68.

Nishiki, T., Y. Kamata, et al. (1994). "Identification of protein receptor for Clostridium botulinum type B neurotoxin in rat brain synaptosomes." J Biol Chem 269(14): 10498-503.

Nygren, P. A., S. Stahl, et al. (1994). "Engineering proteins to facilitate bioprocessing." Trends Biotechnol 12(5): 184-8.

Ofek, G., F. J. Guenaga, et al. "Elicitation of structure-specific antibodies by epitope scaffolds." Proc Natl Acad Sci U S A 107(42): 17880-7.

Owens, R. J. and R. J. Young (1994). "The genetic engineering of monoclonal antibodies." J Immunol Methods 168(2): 149-65.

Perez, J. M., J. G. Renisio, et al. (2001). "Thermal unfolding of a llama antibody fragment: a two-state reversible process." Biochemistry 40(1): 74-83.

Petermann, M. L. (1946). "The splitting of human gamma globulin antibodies by papain and bromelin." J Am Chem Soc 68: 106-13.

Petermann, M. L. and A. M. Pappenheimer, Jr. (1941). "The Action of Crystalline Pepsin on Horse Anti-Pneumococcus Antibody." Science 93(2419): 458.

Porter, R. R. (1950). "The formation of a specific inhibitor by hydrolysis of rabbit antiovalbumin." Biochem J 46(4): 479-84.

Porter, R. R. (1973). "Structural studies of immunoglobulins." Science 180(87): 713-6.

Pujol, M., N. I. Ramirez, et al. (2005). "An integral approach towards a practical application for a plant-made monoclonal antibody in vaccine purification." Vaccine 23(15): 1833-7.

Raag, R. and M. Whitlow (1995). "Single-chain Fvs." FASEB J 9(1): 73-80.

Rahbarizadeh, F., M. J. Rasaee, et al. (2006). "Over expression of anti-MUC1 single-domain antibody fragments in the yeast Pichia pastoris." Mol Immunol 43(5): 426-35.

Rahbarizadeh, F., M. J. Rasaee, et al. (2005). "High expression and purification of the recombinant camelid anti-MUC1 single domain antibodies in Escherichia coli." Protein Expr Purif 44(1): 32-8.

Ravetch, J. V. and J. P. Kinet (1991). "F_C receptors". Annu Rev Immunol 9: 457-92.

Reichert, J. M. (2001). "Monoclonal antibodies in the clinic." Nat Biotechnol 19(9): 819-22.

Ren, F., B. C. Li, et al. (2008). "Expression, purification and characterization of anti-BAFF antibody secreted from the yeast Pichia pastoris." Biotechnol Lett 30(6): 1075-80.

Rizo, J. and T. C. Sudhof (1998). "Mechanics of membrane fusion." Nat Struct Biol 5(10): 839-42.

Rondon, I. J. and W. A. Marasco (1997). "Intracellular antibodies (intrabodies) for gene therapy of infectious diseases." Annu Rev Microbiol 51: 257-83.

Roovers, R. C., T. Laeremans, et al. (2007). "Efficient inhibition of EGFR signaling and of tumour growth by antagonistic anti-EFGR Nanobodies." Cancer Immunol Immunother 56(3): 303-317.

Rossetto, O., M. Seveso, et al. (2001). "Tetanus and botulinum neurotoxins: turning bad guys into good by research." Toxicon 39(1): 27-41.

Rummel, A., T. Karnath, et al. (2004). "Synaptotagmins I and II act as nerve cell receptors for botulinum neurotoxin G." J Biol Chem 279(29): 30865-70.

Saerens, D., K. Conrath, et al. (2008). "Disulfide bond introduction for general stabilization of immunoglobulin heavy-chain variable domains." J Mol Biol 377(2): 478-88.

Saerens, D., G. H. Ghassabeh, et al. (2008). "Single-domain antibodies as building blocks for novel therapeutics." Curr Opin Pharmacol 8(5): 600-8.

Saerens, D., G. H. Ghassabeh, et al. (2008). "Single-domain antibodies as building blocks for novel therapeutics." Curr Opin Pharmacol 8(5): 600-8.

Saerens, D., M. Pellis, et al. (2005). "Identification of a universal VHH framework to graft non-canonical antigen-binding loops of camel single-domain antibodies." J Mol Biol 352(3): 597-607.

Sanz, L., B. Blanco, et al. (2004). "Antibodies and gene therapy: teaching old 'magic bullets' new tricks." Trends Immunol 25(2): 85-91.

Schiavo, G., M. Matteoli, et al. (2000). "Neurotoxins affecting neuroexocytosis." Physiol Rev 80(2): 717-66.

Schmiedl, A., F. Breitling, et al. (2000). "Effects of unpaired cysteines on yield, solubility and activity of different recombinant antibody constructs expressed in E. coli." J Immunol Methods 242(1-2): 101-14.

Sethuraman, N. and T. A. Stadheim (2006). "Challenges in therapeutic glycoprotein production." Curr Opin Biotechnol 17(4): 341-6.

Shaki-Loewenstein, S., R. Zfania, et al. (2005). "A universal strategy for stable intracellular antibodies." J Immunol Methods 303(1-2): 19-39.

Sharon, N. (1993). "Lectin-carbohydrate complexes of plants and animals: an atomic view." Trends Biochem Sci 18(6): 221-6.

Shi, X., T. Karkut, et al. (2003). "Optimal conditions for the expression of a single-chain antibody (scFv) gene in Pichia pastoris." Protein Expr Purif 28(2): 321-30.

Shone, C. C., P. Hambleton, et al. (1985). "Inactivation of Clostridium botulinum type A neurotoxin by trypsin and purification of two tryptic fragments. Proteolytic action near the COOH-terminus of the heavy subunit destroys toxin-binding activity." Eur J Biochem 151(1): 75-82.

Shone, C. C., P. Hambleton, et al. (1985). "Inactivation of Clostridium botulinum type A neurotoxin by trypsin and purification of two tryptic fragments. Proteolytic action near the COOH-terminus of the heavy subunit destroys toxin-binding activity." Eur J Biochem 151(1): 75-82.

Shone, C. C., P. Hambleton, et al. (1987). "A 50-kDa fragment from the NH2-terminus of the heavy subunit of Clostridium botulinum type A neurotoxin forms channels in lipid vesicles." Eur J Biochem 167(1): 175-80.

Shone, C. C. and H. S. Tranter (1995). "Growth of clostridia and preparation of their neurotoxins." Curr Top Microbiol Immunol 195: 143-60.

Sidhu, S. S., B. Li, et al. (2004). "Phage-displayed antibody libraries of synthetic heavy chain complementarity determining regions." J Mol Biol 338(2): 299-310.

Simpson, L. L. (1986). "Molecular pharmacology of botulinum toxin and tetanus toxin." Annu Rev Pharmacol Toxicol 26: 427-53.

Simpson, L. L. and M. M. Rapport (1971). "Ganglioside inactivation of botulinum toxin." J Neurochem 18(7): 1341-3.

Skerra, A. and A. Pluckthun (1988). "Assembly of a functional immunoglobulin Fv fragment in Escherichia coli." Science 240(4855): 1038-41.

Sollner, T., S. W. Whiteheart, et al. (1993). "SNAP receptors implicated in vesicle targeting and fusion." Nature 362(6418): 318-24.

Sollner, T., S. W. Whiteheart, et al. (1993). "SNAP receptors implicated in vesicle targeting and fusion." Nature 362(6418): 318-24.

Sorkin, A. and L. K. Goh (2008). "Endocytosis and intracellular trafficking of ErbBs." Exp Cell Res 314(17): 3093-106.

Sotiriadis, A., T. Keshavarz, et al. (2001). "Factors affecting the production of a single-chain antibody fragment by Aspergillus awamori in a stirred tank reactor." Biotechnol Prog 17(4): 618-23.

Streit, W. J., B. A. Schulte, et al. (1985). "Histochemical localization of galactose-containing glycoconjugates in sensory neurons and their processes in the central and peripheral nervous system of the rat." J Histochem Cytochem 33(10): 1042-52.

Streit, W. J., B. A. Schulte, et al. (1986). "Evidence for glycoconjugate in nociceptive primary sensory neurons and its origin from the Golgi complex." Brain Res 377(1): 1-17.

Streltsov, V. A., J. N. Varghese, et al. (2004). "Structural evidence for evolution of shark Ig new antigen receptor variable domain antibodies from a cell-surface receptor." Proc Natl Acad Sci U S A 101(34): 12444-9.

Studier, F. W. and B. A. Moffatt (1986). "Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes." J Mol Biol 189(1): 113-30.

Sugiyama, S., T. Ozawa, et al. (1980). "Effects of verapamil and propranolol on ventricular vulnerability after coronary reperfusion." J Electrocardiol 13(1): 49-54.

Sundberg, E. J. and R. A. Mariuzza (2002). "Molecular recognition in antibody-antigen complexes." Adv Protein Chem 61: 119-60.

Swaminathan, S. and S. Eswaramoorthy (2000). "Structural analysis of the catalytic and binding sites of Clostridium botulinum neurotoxin B." Nat Struct Biol 7(8): 693-9.

Swennen, D., M. F. Paul, et al. (2002). "Secretion of active anti-Ras single-chain Fv antibody by the yeasts Yarrowia lipolytica and Kluyveromyces lactis." Microbiology 148(Pt 1): 41-50.

Tabor, S. and C. C. Richardson (1985). "A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes." Proc Natl Acad Sci U S A 82(4): 1074-8.

Takkinen, K., M. L. Laukkanen, et al. (1991). "An active single-chain antibody containing a cellulase linker domain is secreted by Escherichia coli." Protein Eng 4(7): 837-41.

Tanaka, T., M. N. Lobato, et al. (2003). "Single domain intracellular antibodies: a minimal fragment for direct in vivo selection of antigen-specific intrabodies." J Mol Biol 331(5): 1109-20.

Tanha, J., G. Dubuc, et al. (2002). "Selection by phage display of llama conventional V(H) fragments with heavy chain antibody V(H)H properties." J Immunol Methods 263(1-2): 97-109.

Tavladoraki, P., E. Benvenuto, et al. (1993). "Transgenic plants expressing a functional single-chain Fv antibody are specifically protected from virus attack." Nature 366(6454): 469-72.

Todorovska, A., R. C. Roovers, et al. (2001). "Design and application of diabodies, triabodies and tetrabodies for cancer targeting." J Immunol Methods 248(1-2): 47-66.

Trinh, R., B. Gurbaxani, et al. (2004). "Optimization of codon pair use within the (GGGGS)3 linker sequence results in enhanced protein expression." Mol Immunol 40(10): 717-22.

Tse, C. K., J. O. Dolly, et al. (1982). "Preparation and characterisation of homogeneous neurotoxin type A from Clostridium botulinum. Its inhibitory action on neuronal release of acetylcholine in the absence and presence of beta-bungarotoxin." Eur J Biochem 122(3): 493-500.

Tudyka, T. and A. Skerra (1997). "Glutathione S-transferase can be used as a C-terminal, enzymatically active dimerization module for a recombinant protease inhibitor, and functionally secreted into the periplasm of Escherichia coli." Protein Sci 6(10): 2180-7.

Uhlen, M., G. Forsberg, et al. (1992). "Fusion proteins in biotechnology." Curr Opin Biotechnol 3(4): 363-9.

Umland, T. C., L. M. Wingert, et al. (1997). "Structure of the receptor binding fragment HC of tetanus neurotoxin." Nat Struct Biol 4(10): 788-92.

Utsumi, S. and F. Karush (1964). "The Subunits of Purified Rabbit Antibody." Biochemistry 3: 1329-38.

van der Vaart, J. M. (2002). "Expression of VHH antibody fragments in Saccharomyces cerevisiae." Methods Mol Biol 178: 359-66.

van der Vaart, J. M., N. Pant, et al. (2006). "Reduction in morbidity of rotavirus induced diarrhoea in mice by yeast produced monovalent llama-derived antibody fragments." Vaccine 24(19): 4130-7.

Van Heyningen, W. E. and P. A. Miller (1961). "The fixation of tetanus toxin by ganglioside." J Gen Microbiol 24: 107-19.

Verma, R., E. Boleti, et al. (1998). "Antibody engineering: comparison of bacterial, yeast, insect and mammalian expression systems." J Immunol Methods 216(1-2): 165-81.

Vranken, W., D. Tolkatchev, et al. (2002). "Solution structure of a llama single-domain antibody with hydrophobic residues typical of the VH/VL interface." Biochemistry 41(27): 8570-9.

Wadsworth, J. D., M. Desai, et al. (1990). "Botulinum type F neurotoxin. Large-scale purification and characterization of its binding to rat cerebrocortical synaptosomes." Biochem J 268(1): 123-8.

Wang, H., J. Dai, et al. (2008). "Expression, purification, and characterization of an immunotoxin containing a humanized anti-CD25 single-chain fragment variable antibody fused to a modified truncated Pseudomonas exotoxin A." Protein Expr Purif 58(1): 140-7.

Ward, E. S., D. Gussow, et al. (1989). "Binding activities of a repertoire of single immunoglobulin variable domains secreted from Escherichia coli." Nature 341(6242): 544-6.

Weber, T., B. V. Zemelman, et al. (1998). "SNAREpins: minimal machinery for membrane fusion." Cell 92(6): 759-72.

Weisser, N. E. and J. C. Hall (2009). "Applications of single-chain variable fragment antibodies in therapeutics and diagnostics." Biotechnol Adv 27(4): 502-20.

Weisser, N. E. and J. C. Hall (2009). "Applications of single-chain variable fragment antibodies in therapeutics and diagnostics." Biotechnol Adv 27(4): 502-20.

Wesolowski, J., V. Alzogaray, et al. (2009). "Single domain antibodies: promising experimental and therapeutic tools in infection and immunity." Med Microbiol Immunol 198(3): 157-74.

Williams, R. S., C. K. Tse, et al. (1983). "Radioiodination of botulinum neurotoxin type A with retention of biological activity and its binding to brain synaptosomes." Eur J Biochem 131(2): 437-45.

Worn, A. and A. Pluckthun (2001). "Stability engineering of antibody single-chain Fv fragments." J Mol Biol 305(5): 989-1010.

Wu, T. T., G. Johnson, et al. (1993). "Length distribution of CDRH3 in antibodies." Proteins 16(1): 1-7.

Wu, T. T. and E. A. Kabat (1970). "An analysis of the sequences of the variable regions of Bence Jones proteins and myeloma light chains and their implications for antibody complementarity." J Exp Med 132(2): 211-50.

Yoo, T. J., O. A. Roholt, et al. (1967). "Specific binding activity of isolated light chains of antibodies." Science 157(789): 707-9.

Zwick, E., J. Bange, et al. (2001). "Receptor tyrosine kinase signalling as a target for cancer intervention strategies." Endocr Relat Cancer 8(3): 161-73.