



Terms and Conditions of Use of Digitised Theses from Trinity College Library Dublin

Copyright statement

All material supplied by Trinity College Library is protected by copyright (under the Copyright and Related Rights Act, 2000 as amended) and other relevant Intellectual Property Rights. By accessing and using a Digitised Thesis from Trinity College Library you acknowledge that all Intellectual Property Rights in any Works supplied are the sole and exclusive property of the copyright and/or other IPR holder. Specific copyright holders may not be explicitly identified. Use of materials from other sources within a thesis should not be construed as a claim over them.

A non-exclusive, non-transferable licence is hereby granted to those using or reproducing, in whole or in part, the material for valid purposes, providing the copyright owners are acknowledged using the normal conventions. Where specific permission to use material is required, this is identified and such permission must be sought from the copyright holder or agency cited.

Liability statement

By using a Digitised Thesis, I accept that Trinity College Dublin bears no legal responsibility for the accuracy, legality or comprehensiveness of materials contained within the thesis, and that Trinity College Dublin accepts no liability for indirect, consequential, or incidental, damages or losses arising from use of the thesis for whatever reason. Information located in a thesis may be subject to specific use constraints, details of which may not be explicitly described. It is the responsibility of potential and actual users to be aware of such constraints and to abide by them. By making use of material from a digitised thesis, you accept these copyright and disclaimer provisions. Where it is brought to the attention of Trinity College Library that there may be a breach of copyright or other restraint, it is the policy to withdraw or take down access to a thesis while the issue is being resolved.

Access Agreement

By using a Digitised Thesis from Trinity College Library you are bound by the following Terms & Conditions. Please read them carefully.

I have read and I understand the following statement: All material supplied via a Digitised Thesis from Trinity College Library is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of a thesis is not permitted, except that material may be duplicated by you for your research use or for educational purposes in electronic or print form providing the copyright owners are acknowledged using the normal conventions. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone. This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.



Expression of antibody fragments in tobacco using transient and stable expression systems

A dissertation submitted to the University of Dublin
for the Degree of Doctor of Philosophy

Yi Hui (Audrey) Teh

School of Genetics and Microbiology
Trinity College Dublin, April 2011



Thesis 9573

Declaration

I declare that this thesis has not been submitted as an exercise for a degree at this or any other university and it is entirely my own work.

I agree to deposit this thesis in the University’s open access institutional repository or allow the library to do so on my behalf, subject to Irish Copyright Legislation and Trinity College Library conditions of use and acknowledgement.



Dedication

This thesis is dedicated to everyone who has supported me along the way. Thank you for your encouragement when I am in the deepest of troughs and for cheering me on when I am riding the crest of life.

Acknowledgements

I would like to take this opportunity to express my heart-felt gratitude to many people who has guided and assisted me through this journey. It has certainly been very challenging and I deeply appreciate all the support that I have been given to make the road that much smoother.

First and foremost, I would like to thank my supervisor Tony for his tremendous patience and insight. Thank you for taking me on board, for your guidance and input from the start of the project, for the countless times you have steered me back on course when I am stuck and for the all the valuable feedback. I am deeply honoured to have you as my supervisor.

Many thanks also goes to members of the lab, past and present: to Nga for your invaluable help and your treasure trove of knowledge; to Alan Maloney for showing me the ropes in the very beginning; to Liina, friend and sounding board; and also to Alan Magee, Marie and Eva. I would also like to thank the Biochemistry department for allowing me to use the plate reader. To Lillian, Paola, David, Kevin, Gavin, Joern, Inga, Nora, Emmanuelle, Sam, Diarmuid, Kasha, Francesca, Francizka, Andrea, Letal, Dave Noone, Sebastian, Eric, Leigh, Aoife, Ciara, Sahar, Claire, Clare, Anh, Sean and countless others, thank you for your insights, your tips and tricks as well as your cakes, chocolates and biscuits. To Takashi and family, who were affected by the recent Japanese earthquake and tsunami, my thoughts are with you. A big thank you also goes to Olive whose tea and coffee signals a welcome break and helps me feel re-energized.

To my parents, thank you for your unconditional love and support. I will always have time for your constant bombardment of questions in your effort to understand my work. Many thanks too to my brothers, confidants and a constant source of news, gossip and jokes from home. An Qiao, congratulations on your certification and your new bride. An Liang, all the best in your final year project. To Mei Ee, my new sister-in-law, welcome to the family.

Last but not least, I would like to say a big thank you to my friends in Dublin and around the world for your support and your company, be it in person or online. To Geoff, a big thank you for all the support you have given me. Many thanks go to

your wonderful family too. To Gain Loon and Fang Yin, thanks for all the encouragement. Last but not least, a big shout out to an extraordinary group of swing dancers whom I have the privilege of calling them my friends: Karen, Sarah, Bobb, Siamsa, Colm, Rory, Dave, Gill, Gerré, Helen, Caroline, Dan, Gordon, Anita, Chris, Paul and Norma, Marie-George, Sabine, Paola, Michael, Sandra, Julie, Zsé, Carl, Rodney, Linda, Sven, Alejandro, Catherine, Ronan, Anton, Peter, Amy, Joe and countless others. No matter what happens, keep on smiling, keep on dancing.

I have gained a lot from this experience which I will cherish and carry with me as I go on with life. I want to extend my most sincere thanks and warmest wishes to all of you.

Table of Contents

Declaration i

Dedication..... ii

Acknowledgementsiii

Table of Contents..... v

List of Tables..... x

List of Figures..... xi

Summaryxv

List of Abbreviations..... xvii

Chapter 1. General Introduction..... 1

 1.1 Antibodies 1

 1.2 IgG derivatives and their applications 3

 1.3 Heavy chain-only antibodies (HCAb) in the antibody repertoire of Camelids..... 4

 1.4 Heavy chain-only variable fragment (VHH) 6

 1.5 Advantages of Camel VHH antibodies..... 6

 1.6 Applications of Camel VHH antibodies 9

 1.7 The search for a ‘universal’ VHH framework and the ‘humanization’ of VHH 11

 1.8 Production platforms for derivatives of conventional and Camel antibodies 12

 1.9 Transgenic plants as an expression platform..... 12

 1.10 Plant tissue and cell culture 14

 1.11 Strategies for gene expression 16

 1.12 Research Aims and Objectives..... 18

Chapter 2. Materials and Methods 19

 2.1 Expression vector modules 19

 2.1.1 ICON Genetics system provector modules 19

 2.1.2 The CPMV expression vector, pEAQ-HT 22

2.1.3	TRBO expression vectors (pJL-TRBO).....	22
2.2	scFvs and VHHs investigated	25
2.2.1	Design and synthesis of CLDSyn and CBESyn	25
2.2.2	Design and synthesis of PS11, CPS, RN5 and CRN.....	25
2.3	Cloning of GOIs into vectors for transient expression.....	28
2.3.1	PCR amplification of GOIs.....	28
2.3.2	Cloning of synthetic genes into pICH11599.....	31
2.3.3	Cloning of SEKDEL-tagged genes into pICH11599	31
2.3.4	Cloning of a CBESyn-GFP fusion construct into pICH11599	33
2.3.5	Cloning of GOIs into the pEAQ-HT and pJL-TRBO vectors	33
2.3.6	Swapping of framework regions between CLDSyn and CBESyn	35
2.4	Cloning of GOIs into stable transformation vectors	37
2.4.1	Amplification of GOIs using DNA from leaves infiltrated with ICON Genetics transient expression vectors	37
2.4.2	Amplification of GOIs using mRNA from leaves infiltrated with ICON vectors	38
2.5	Plant growth and transformation	39
2.5.1	Transient transformation.....	39
2.5.2	Stable transformation	40
2.6	Detection of GOI mRNA in tobacco leaves	40
2.7	Detection of protein expression in Tobacco leaves	41
2.8	Protein purification	42
2.8.1	Purification of His-tagged proteins.....	42
2.8.2	Purification of Strep-tagged proteins.....	42
2.9	Immunoabsorbent assays (ELISA).....	43
2.9.1	Direct ELISA for His-tagged proteins.....	43
2.9.2	Direct ELISA assay for Strep-tagged proteins	43
2.9.3	Measurement of VHH affinity by competitive ELISA	44
2.9.4	Measurement of the heat stability of VHHs	44
2.10	Bacterial expression of VHHs.....	45

2.10.1 Cloning of GOIs into the expression vector pET-22b	45
2.10.2 Bacterial growth and induction of protein expression	45
Chapter 3. Expression of human scFvs and Camel VHHs in <i>Nicotiana</i> <i>benthamiana</i>	46
3.1 Introduction	46
3.2 Results	48
3.2.1 Cloning and expression of scFv and VHH constructs.....	48
3.2.2 Accumulation of scFvs and VHHs in different cellular compartments	49
3.2.3 Northern blot analyses of mRNA accumulation	53
3.2.4 Design of the CL3 variant CLD	55
3.2.5 Purification and characterization of scFvs and VHHs.....	56
3.2.6 Expression of the human anti- β -galactosidase scFv ABGST over time at the protein and mRNA level	60
3.3 Discussion.....	63
3.3.1 Comparison of yields obtained by the ICON system to alternative expression systems in plants and other production platforms.....	63
3.3.2 The effect of protein targeting to various cellular compartments on yield.....	65
3.3.3 The effect of C-terminal purification tags on yield.....	66
Chapter 4. Investigations of codon optimization and framework changes on VHH accumulation in plants	67
4.1 Introduction	67
4.2 Results	69
4.2.1 Design of modified anti-HEWL VHH constructs	69
4.2.2 Expression and purification of modified anti-HEWL VHHs	73
4.2.3 Expression of anti-HEWL VHH constructs over time	75
4.2.4 Characterisation of anti-HEWL VHHs	83
4.2.5 Cloning and expression of a CBESyn-GFP fusion protein	88
4.2.6 Accumulation of CBESyn-GFPST targeted to different cellular compartments	90

4.3	Discussion.....	92
4.3.1	Accumulation of the three versions of anti-HEWL VHH constructs in <i>N. benthamiana</i>	92
4.3.2	The effect of codon optimization/amino acid changes on yield.....	94
4.3.3	The effect of framework changes on yield.....	95
Chapter 5. Exploring the importance of Framework Regions for VHH accumulation and yield in plants		97
5.1	Introduction	97
5.2	Results	102
5.2.1	Design of anti-PSA and anti-RNase A VHH constructs.....	102
5.2.2	Expression and purification of anti-PSA and anti-RNase A VHHs	104
5.2.3	Expression of anti-PSA and anti-RNase VHH constructs over time....	109
5.2.4	Characterisation of anti-RNase A VHHs	111
5.2.5	Construction of chimeric anti-HEWL VHHs containing combinations of Framework Regions from CLDSyn and CBESyn.....	115
5.2.6	Expression and purification of chimeric anti-HEWL VHHs	119
5.2.7	Characterization of anti-HEWL chimeric VHHs	123
5.2.8	Expression of an anti-HEWL and anti-RNase A VHH in <i>Escherichia coli</i> : a comparison.....	128
5.3	Discussion.....	132
5.3.1	The use of alternative Framework Regions from cAbBCII10 to increase the yield of Camel VHHs <i>in planta</i>	132
5.3.2	Comparison of Camel VHHs obtained from <i>E. coli</i> and transiently- transformed plants (<i>N. benthamiana</i>)	134
Chapter 6. Investigating the use of different expression vector systems (ICON, pEAQ-HT, pJL-TRBO, pBIN2 and pBIN8) for transient VHH expression ...		136
6.1	Introduction	136
6.2	Results	139
6.2.1	Comparison of GFP expression from three virus-based transient expression vectors	139
6.2.2	Comparison of apoplast-targeted CBESynHS expression from three virus-based transient expression vectors	142

6.2.3	Expression, over time, of Ap-CBE from the pEAQ-HT vector at the protein and mRNA level.....	144
6.2.4	Comparison of PSAHS and CPSHS expression from three viral-based transient expression vectors	147
6.3	Discussion.....	150
6.3.1	Comparison of mRNA and protein expression levels of plant transformed with ICON, pEAQ-HT and pJL-TRBO vectors.....	150
6.3.2	Expression of PSAHS and CPSHS using ICON, pEAQ-HT and pJLTRBO	152
Chapter 7. VHH transgene expression in stably transformed <i>N. tabacum</i> and <i>N. benthamiana</i> lines.....		154
7.1	Introduction	154
7.2	Results	156
7.2.1	Construction of an Ap-CBE transgene by RT-PCR, using RNA from infiltrated leaves as the template	156
7.2.2	Stable transformation of <i>N. tabacum</i> and <i>N. benthamiana</i> using the pEAQ-HT vector.....	160
7.2.3	Expression of the Ap-CBE cDNA transgene in <i>N. tabacum</i> and <i>N. benthamiana</i>	163
7.2.4	Construction of a genomic version of the Ap-CBE transgene using DNA from ICON vector-infiltrated leaves as the PCR template	166
7.2.5	Expression of the LF and SF transgenes in <i>N. tabacum</i> and <i>N. benthamiana</i>	169
7.3	Discussion.....	174
7.3.1	Difficulties in the construction of transgenes generated from amplified cDNA or recombined DNA from plants infiltrated with ICON provector modules	174
7.3.2	Expression of VHH transgenes in stably transformed <i>N. tabacum</i> and <i>N. benthamiana</i> plants.....	175
Chapter 8. Conclusion and Future Work		178
References		183

List of Tables

Table 2.1	Names and descriptions of ICON provectors modules	20
Table 2.2	The scFvs and VHs investigated and their respective antigens	26-27
Table 2.3	Template, forward and reverse primers used for amplification of GOI	29-30
Table 2.4	Forward and reverse primers used for amplification of KDEL-tagged CBESyn constructs	32
Table 2.5	Forward and reverse primers used for amplification of GOIs to be cloned into pEAQ-HT or pJL-TRBO	34
Table 2.6	Template, forward and reverse primers used for amplification of the second and third Framework Regions (FR) of different chimeric genes	36
Table 3.1	List of the scFvs and Camel VHs initially investigated, with and without different purification tags	51
Table 3.2	Table showing the compartment with the highest accumulation and the yield after purification for ABGHS, 163ST, ABGST, CL3ST and CLDST	58
Table 4.1	Table showing for each construct, the compartment with the highest accumulation, the estimated yield before purification and the yield after purification	81
Table 5.1	Table showing estimated yields for the anti-PSA and anti-RNase A VHs, the compartment with the highest accumulation, the estimated yield before purification and the yield after purification	108
Table 5.2	Table showing, for the 6 chimeric VHs as well as the parental controls CLDSyn and CBESyn, the compartment with the highest accumulation, the estimated yield before purification and the yield after purification	121
Table 5.3	Table showing, for the 6 chimeric VHs as well as the parental controls CLDSyn and CBESyn, their respective dissociation constant (K_D) as determined by competitive ELISA (Friguet <i>et al.</i> , 1985)	126
Table 5.4	Table showing a comparison of the estimated yield before purification for His-tagged CLDSyn, CBESyn, RN5 and CRN obtained from the <i>E. coli</i> expression system and from <i>N. benthamiana</i> plant infiltrated with the ICON system	131

List of Figures

Figure 1.1	Schematic diagram depicting the structure of conventional IgG along with derivatives including scFv, Fab and diabody	2
Figure 1.2	Differences between conventional antibodies and HCABs	5
Figure 1.3	Derivatives of HCAB including VHH, bivalent VHH and pentabody	8
Figure 2.1	ICON Genetics provector modules	21
Figure 2.2	The pEAQ-HT vector	23
Figure 2.3	The pJL-TRBO vector	24
Figure 3.1	Predicted N-terminal amino acid sequences of proteins encoded by a GOI when expressed in the cytoplasm or targeted to different cellular compartments	50
Figure 3.2	The accumulation of CL3ST, 163ST, ABGST, ABGHS and ABG (no tag) in the cytoplasm, apoplast and Chloroplast at 7dpi as detected by Western Blot or PAGE gel	52
Figure 3.3	Detection of mRNA transcripts in plants infiltrated with the TT2ST or 163HS constructs which encode either a cytoplasmically-expressed antibody fragment, or variants targeted to the apoplast or the chloroplast; The ICON expression system showing the assembly and processing of the 5' and 3' provector modules <i>in planta</i> ; The position of subgenomic promoters in the TMV expression vector	54
Figure 3.4	PAGE gel showing the crude extract and eluted fraction of 163ST, ABGST, CL3ST and ABGHS.	57
Figure 3.5	Antigen recognition specificity of the scFvs and Camel VHHs as determined by ELISA	59
Figure 3.6	The accumulation of mRNA transcripts in plants infiltrated with a construct directing ABGST expression in the cytoplasm at 0, 3, 4, 5, 7, 9, and 11 dpi compared with non-infiltrated leaves.	61
Figure 3.7	Coomassie blue-stained gel showing accumulation of ABGST in the cytoplasm at 0, 2, 5, 7, 9, 11, 13, and 15 dpi compared with uninfiltrated leaves.	62
Figure 4.1	Comparison of the nucleotide sequences of the original His-tagged CLD (CLDHS) and codon-optimized version (CLDSynHS)	71
Figure 4.2	A comparison of the GC content of His-tagged CLD (CLDHS) and CLDSyn (CLDSynHS)	72

Figure 4.3	Comparison of the amino acid sequences of CLD and CBE (excluding the C-terminal tag)	72
Figure 4.4	Accumulation of anti-HEWL antibodies targeted to different cellular compartments	76
Figure 4.5	PAGE gel showing VHH abundance in 3 randomly sampled leaves infiltrated with the three different His-tagged and Strep-tagged anti-HEWL constructs	77
Figure 4.6	Schematic representation of the different SEKDEL-tagged CBESyn constructs	78
Figure 4.7	Accumulation of anti-HEWL antibodies targeted to the ER	79
Figure 4.8	Purification and yield of His-tagged CLD, CLDSyn and CBESyn	80
Figure 4.9	Accumulation of apoplast-targeted anti-HEWL antibodies over time	82
Figure 4.10	Specificity of the anti-HEWL camel antibodies as determined by ELISA	85
Figure 4.11	Affinity of His-tagged CLDSyn and CBESyn measured according to Friguet <i>et al.</i> (1985)	86
Figure 4.12	Comparison of thermal stability of CLDHS and CBESynHS	87
Figure 4.13	Schematic diagram of the CBESyn-GFP fusion constructs	89
Figure 4.14	Fluorescence of 7 dpi leaves of <i>N. benthamiana</i> accumulating CBESyn-GFPST in cytoplasm, targeted to the apoplast or Chloroplast	89
Figure 4.15	PAGE gel and Western blot of CBESyn-GFPST accumulating in cytoplasm apoplast and chloroplast at 7 dpi	91
Figure 5.1	Schematic diagram showing the differences between VH and VHH based on sequence comparisons of cDNA clones	100
Figure 5.2	Ribbon representation of the X-ray structure of the VHH cAb-Lys3 in complex with its antigen hen egg-white lysozyme (HEWL)	101
Figure 5.3	Comparison of the amino acid sequences of His-tagged anti-PSA, anti-RNase A and anti-HEWL VHHS	103
Figure 5.4	Accumulation of anti-PSA and anti-RNase A camel antibodies in various cellular compartments	105-106
Figure 5.5	Purification and yield of His-tagged anti-PSA and anti-RNase A VHH	107
Figure 5.6	Accumulation of anti-PSA VHH and apoplast-targeted anti-RNase A VHH over time	110

Figure 5.7	Specificity of the anti-RNase A camel antibodies as determined by ELISA	112
Figure 5.8	Affinity of His-tagged anti-RNase VHH with original framework and alternative framework measured according to Friguet <i>et al.</i> (1985)	113
Figure 5.9	Comparison of thermal stability of RN5HS and CRNHS	114
Figure 5.10	Relative positions of the nucleotide differences between the anti-HEWL VHHs CLDSyn and CBESyn	116
Figure 5.11	Combinations of forward primers and VHH templates used to generate fragments with different combinations of the FR2 and FR3 regions	117
Figure 5.12	Structure of the 6 chimeric anti-HEWL VHH genes	118
Figure 5.13	PAGE gel and Western blot showing accumulation of the six different anti-HEWL camel antibody chimeras DBB, DBD, DDB, BDD, BBD and BDB when targeted to the apoplast	120
Figure 5.14	Purification and yield of 6 anti-HEWL VHH chimeras	122
Figure 5.15	Specificity of the anti-HEWL camel antibodies chimeras as determined by ELISA.	124
Figure 5.16	Affinity of the anti-HEWL camel antibodies chimeras measured according to Friguet <i>et al.</i> (1985)	125
Figure 5.17	Comparison of thermal stability of the anti-HEWL VHH chimeras	127
Figure 5.18	Accumulation of anti-HEWL and anti-RNase A VHH in <i>E. coli</i> over time	130
Figure 6.1	Comparison of GFP fluorescence levels in leaves of <i>N. benthamiana</i> infiltrated with the pICON-GFP, pEAQ-GFP and pTRBO-GFP	140
Figure 6.2	PAGE gel and Western blot of GFP accumulating in the cytoplasm of 7 dpi <i>N. benthamiana</i> leaves infiltrated with the pICON-GFP	141
Figure 6.3	PAGE gel and Western blot of CBE accumulation in <i>N. benthamiana</i> leaves at various days post-infiltration with pEAQ:Ap-CBE, pBIN2:Ap-CBE, pBIN8:Ap-CBE and pTRBO:Ap-CBE, compared to leaves transformed with empty vector or the pICON:Ap-CBE vector	143
Figure 6.4	Northern blot showing CBE mRNA accumulation at 0 – 6 dpi in leaves of <i>N. benthamiana</i> infiltrated with pEAQ:Ap-CBE, compared to leaves transformed with empty vector or pICON:Ap-CBE	145

Figure 6.5	PAGE gel and Western blot showing CBE protein accumulation at 0 - 6 dpi in <i>N. benthamiana</i> leaves infiltrated with pEAQ:Ap-CBE compared to leaves transformed with empty vector or pICON:Ap-CBE	146
Figure 6.6	PAGE gel and Western blot showing PSAHS or CPSHS protein accumulating at 0-11 days post-infiltration (dpi) in the cytoplasm of <i>N. benthamiana</i> leaves infiltrated with the pICON constructs, the pEAQ constructs or pTRBO constructs compared to leaves transformed with the corresponding empty vectors	149
Figure 7.1	Nucleotide and amino acid sequence of an Ap-CBE clone in pUC19 starting from the <i>lacZa</i> translational start site	158
Figure 7.2	Comparison of three constructs for stable transformation derived from cDNA of transiently-transformed <i>N. benthamiana</i> leaves	159
Figure 7.3	Green fluorescence in leaves of <i>N. tabacum</i> and <i>N. benthamiana</i> plants stably transformed with pEAQ-GFP; and <i>N. benthamiana</i> stably transformed with pEAQ:Ap-GFP compared with a wild-type control	161
Figure 7.4	GFP accumulation in stably transformed <i>N. tabacum</i> cv. Samsun and <i>N. benthamiana</i> plants	162
Figure 7.5	PCR screening of putative <i>N. tabacum</i> and <i>N. benthamiana</i> transformants for the presence of the Ap-CBE transgene	164
Figure 7.6	mRNA transcripts of <i>N. tabacum</i> cv Samsun and <i>N. benthamiana</i> transformants containing the Ap-CBE transgene.	165
Figure 7.7	Diagram showing the predicted structure of the <i>in planta</i> recombined ICON 5' and 3' vector modules, and the regions contained within the PCR-amplified fragments.	168
Figure 7.8	PCR screening of putative <i>N. tabacum</i> and <i>N. benthamiana</i> transformants containing Short Fragment (SF) constructs.	170
Figure 7.9	PCR screening of putative <i>N. tabacum</i> and <i>N. benthamiana</i> transformants containing Long Fragment (LF) constructs	171
Figure 7.10	Northern blot analysis of total RNA from <i>N. tabacum</i> cv Samsun and <i>N. benthamiana</i> transformants containing Short Fragment (SF) constructs.	172
Figure 7.11	Northern blot analysis of total RNA from <i>N. tabacum</i> cv Samsun and <i>N. benthamiana</i> transformants containing Long Fragment (LF) constructs.	173

Summary

Plants are an attractive alternative to mammalian and microbial cell culture systems as a potential production platform for recombinant proteins. Apart from avoiding ethical problems associated with transgenic animals, plant systems have various advantages such as low cultivation/production costs, ease of scalability via standard agricultural practices, the possibility of using edible plant material expressing therapeutics (e.g. vaccines) as food or animal feed, and of using whole plants for bioremediation purposes. The use of plants also reduces the risk of product contamination by mammalian viruses, blood-borne pathogens, and bacterial toxins.

The primary aim of this thesis was to explore strategies aimed at increasing the accumulation and yield of antibody fragments in *Nicotiana benthamiana* and *N. tabacum*. The antibody fragments of interests included human single-chain variable fragments (scFvs) derived from conventional immunoglobulin G (IgG), and heavy-chain variable fragments (VHHs or nanobodies) derived from Camelid heavy chain-only antibodies (HCAbs). VHHs are especially suitable for many biotechnological applications due to their small size, monomeric behaviour, high stability and solubility, and their ability to bind epitopes not accessible to conventional antibodies.

We investigated transient expression of scFvs and VHHs in *N. benthamiana* using 'deconstructed virus'-based transient expression vectors. Initially, two versions of a human anti β -galactosidase scFv and three versions of a camel anti-hen egg white lysozyme (HEWL) VHH were expressed. The anti β -galactosidase scFvs include the original scFv and a mutant selected to fold correctly and accumulate at a high level in the reducing environment of the bacterial cytoplasm. The anti-HEWL nanobodies included the original VHH from an immunized library (cAbLys3), a plant codon-optimized derivative, and a codon-optimized hybrid nanobody comprising the CDRs of cAbLys3 grafted onto an alternative 'universal' VHH framework. Both 6xHis- and Strep-tagged derivatives of each scFv and VHH were expressed in different subcellular compartments, including the cytoplasm, chloroplast, apoplast and ER using different targeting pre-sequences.

We found that antibody fragment tertiary structure had a major impact not only on protein accumulation and overall yield, but also on the subcellular compartment in which fragment showed the highest accumulation. For example, overall accumulation and yield of the mutant anti- β -galactosidase scFv was greater

compared with its wild-type counterpart. Moreover, the two scFvs accumulated most highly in different subcellular compartments: the mutant in the cytoplasm, the wild-type when targeted to chloroplasts. For the anti-HEWL nanobodies, intact functional protein accumulated at an exceptionally high level (up to 30% total leaf protein) when the original VHH framework was replaced with the alternative 'universal' framework. Fusion of this hybrid VHH with GFP changed the subcellular compartment in which it showed the highest accumulation from apoplast to cytoplasm. Also, different purification tags affected the accumulation and the integrity of the VHHs. On the other hand, when the yield of the plant codon-optimized VHH construct was compared with that of the original mammalian codon version, no increase in protein accumulation was observed.

Subsequently, it was found that the use of the 'universal' framework increased the accumulation of two other nanobodies (anti-PSA and anti-RNase A VHHs), relative to their original-framework counterparts. This was observed not only when different transient expression vectors were used to direct their expression *in planta*, but also when the constructs were expressed in *E. coli*. In addition, 6 anti-HEWL nanobody chimeras were constructed that contained different combinations of framework regions (FRs) from the 'universal' and the native VHH framework. When transiently expressed in *N. benthamiana*, the 6 chimeras showed different levels of accumulation and differential susceptibility to proteolytic degradation.

Finally, investigations aimed at achieving stable expression of apoplast-targeted anti-HEWL nanobodies (both with the native and the 'universal' FRs) in *N. tabacum* and *N. benthamiana* are described. Several strategies were evaluated, including the use of different promoters (CaMV35S and *rbcS* promoters), the inclusion of an intron in the coding region, and the use of 5' and 3' untranslated region (UTRs) from Tobacco Mosaic Virus (TMV) and Cowpea Mosaic Virus (CPMV).

List of Abbreviations

μ	micro-
$^{\circ}\text{C}$	Degrees Celsius
AMEF	Antibody-mediated enzyme formation
<i>A. tumefaciens</i>	<i>Agrobacterium tumefaciens</i>
BAP	Benzylaminopurine
BeYDV	Bean yellow dwarf virus
BSA	Bovine Serum Albumin
BY2	Bright Yellow 2
CaMV	Cauliflower mosaic virus
cDNA	Complementary Deoxyribonucleic acid
CDR	Complementarity Determining Region
CH	Constant heavy domain
CHO	Chinese Hamster Ovary
CL	Constant light domain
CP	Coat protein
CPMV	Cowpea mosaic virus
Cr-TMV	Crucifer-infecting Tobacco mosaic virus
CV	Column volume
dpi	Days post-infiltration
DIG	Digoxigenin
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
<i>E. coli</i>	<i>Escherichia coli</i>
ECL	Enhanced Chemiluminescence
EDTA	Ethylenediaminetetraacetic acid
ELISA	Enzyme-linked immunosorbent assay
ELP	Elastin-like peptide
ER	Endoplasmic reticulum
ERAD	Endoplasmic reticulum-associated degradation pathway
EGFR	Epidermal Growth Factor Receptor
Fab	Fragment, antigen binding
Fc	Fragment crystallizable (antibody region)
FL	Free linker
FR	Framework Region

g	gram
GFP	Green Fluorescent Protein
GOI	Gene of Interest
H ₂ SO ₄	Sulphuric acid
HBcAg	Hepatitis B Core Antigen
HCAb	Heavy-chain only antibody
HCl	Hydrochloric acid
HEWL	Hen egg-white lysozyme
HIV	Human Immunodeficiency Virus
HRP	Horseradish peroxidase
HT	Hypertranslatable
Id	Idiotype
Ig	Immunoglobulin
IPTG	Isopropyl-D-thiogalactopyranoside
IS	Infection Solution
IL	Interleukin
kb	kilobase
kDa	kilodalton
K _D	Dissociation constant
l	litre
LB	Luria-Bertani
LiCl	Lithium chloride
m	metre
M	mole
mAb	Monoclonal antibody
MCS	Multiple Cloning Site
MgSO ₄	Magnesium sulphate
MP	Movement protein
mRNA	Messenger Ribonucleic acid
MS	Murashige and Skoog
MUC1	Cancer associated mucin
n	nano-
NAA	α-Naphthaleneacetic acid
NaCl	Sodium chloride
NaH ₂ PO ₄	Sodium phosphate
<i>N. benthamiana</i>	<i>Nicotiana benthamiana</i>
<i>N. tabacum</i>	<i>Nicotiana tabacum</i>
Nef	Negative Regulatory Factor

NLS	Nuclear localizing signal
NTA	Nitrilotriacetic acid
NTR	Non-translated region
ORF	Open reading frame
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
PEG	Polyethyleneglycol
PMSF	Phenylmethylsulfonyl fluoride
PSA	Prostate Specific Antigen
PVX	Potato Virus X
<i>rbcS</i>	Rubisco small subunit
RdRP	RNA-dependent RNA polymerase
Rep	Replication initiator protein
RNA	Ribonucleic acid
RT-PCR	Reverse transcription Polymerase Chain Reaction
RUBISCO	Ribulose-1,5-bisphosphate carboxylase oxygenase
SBEA	Starch branching enzyme A
scFv	Single-chain variable fragment
SD	Standard deviation
SDS	Sodium Dodecyl Sulphate
sIg	Secretory Immunoglobulin
SIM	Shoot Inducing Medium
SSU	Small subunit
TBSV	Tomato bushy stunt virus
T-DNA	Transfer Deoxyribonucleic acid
TNF	Tumour necrosis factor
TRBO	TMV RNA-based Overexpression vector
TSP	Total soluble protein
TMV	Tobacco mosaic virus
TVCV	Turnip vein-clearing virus
UTR	Untranslated region
UV	Ultraviolet
VH	Variable heavy domain
VHH	Heavy-chain only variable fragment
VL	Variable light domain
v	Volume
WT	Wild-type

w

weight

WT	Wild-type
V	Volume
VL	Variable light domain
VHL	Heavy-chain only variable domain
VH	Variable heavy domain
UV	Ultraviolet
ULR	Untranslated region
TVCV	Tamoxifen-vascular-endothelial
TMV	Tobacco mosaic virus
TSP	Total soluble protein
TRBO	TMV RNA-binding protein
TRF	Tumor necrosis factor
T-DNA	T-DNA
TRSV	Tobacco etch virus
ESU	Small subunit
STF	Short tandem repeat
slg	Secretory immunoglobulin
SDS	Sodium dodecyl sulfate
SD	Standard deviation
stf	Single-chain antibody fragment
SSEA	Starch-binding domain
RUBISCO	Ribulose 1,5-bisphosphate carboxylase/oxygenase
RT-PCR	Reverse transcription-polymerase chain reaction
RNA	Ribonucleic acid
Rsp	Repetitive DNA sequence
RdL	Ribonuclease L
hpc	Ribonuclease H
PVX	Potato virus X
PFA	Paraformaldehyde
PMSE	Phosphorylated protein
PEG	Polyethylene glycol
PCR	Polymerase chain reaction
PS	Phosphatidylserine
PSA	Prostate-specific antigen
WT	Wild-type

Chapter 1. General Introduction

1.1 Antibodies

Antibodies are immunoglobulins produced by B lymphocytes of the vertebrate immune system. They participate in the recognition and elimination of non-self antigens by binding to unique molecular epitopes on antigenic proteins, glycoproteins, carbohydrates and lipids. Antigens that have been coated with antibodies are marked for destruction by effector cells of the immune system (Janeway *et al.*, 2001).

The main antibody isotypes found in mammalian serum are IgG, IgA, IgM, IgD and IgE. Of these, IgG is the most common type (Figure 1.1). It consists of two pairs of identical polypeptides, the heavy and light chains. Each heavy chain consists of 4 domains: 3 constant domains (CH) and 1 variable domain (VH). The two heavy chains join to form a 'Y' shaped structure stabilized by disulphide bonds in the hinge region (Woof and Burton, 2004). The two light chains, each comprising 1 constant domain (CL) and 1 variable domain (VL), are connected to the 'arms' of the heavy chain by disulphide bonds. The N-terminal ends of the 'arms' of the IgG molecule, which contain the variable regions of both the heavy and the light chain, interact with antigen epitopes via 3 Complementarity Determining Regions (CDRs), variable loops present at the ends of each variable fragment. The variability of IgG is achieved through the combination of different VH and VL domains. Therefore, each variable region is bivalent and capable of recognizing two different antigen molecules.

The ability of antibodies to recognize and bind an extraordinarily diverse array of antigens with high specificity and affinity has generated intense interest in exploiting them for biotechnological applications based on molecular recognition. As a result, major advances in antibody selection, the engineering of smaller antigen binding derivatives, and the development of recombinant expression systems have been made during the past two decades.

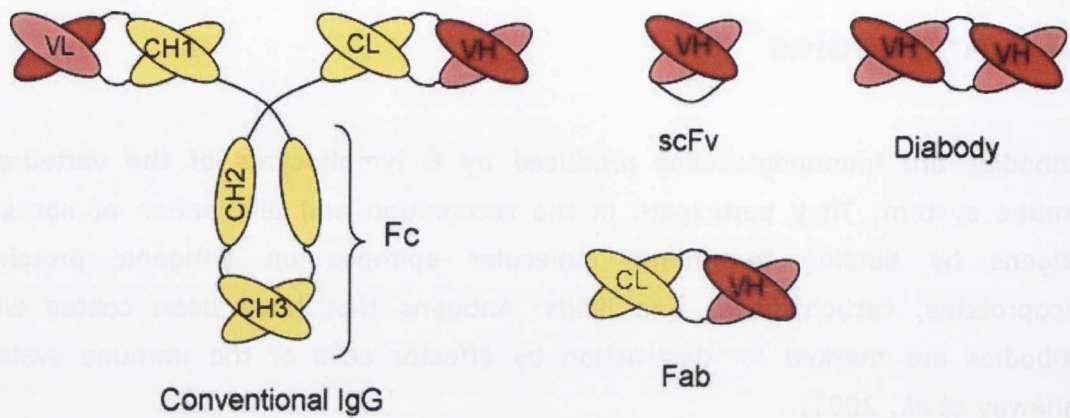


Figure 1.1 Schematic diagram depicting the structure of conventional IgG along with derivatives including scFv, Fab and diabody. Variable heavy domains (VH) are coloured in red, variable light domains (VL) are coloured pink while constant domains (CH and CL) are coloured yellow (adapted from Harmsen and De Haard, 2007). For each of these formats, high affinity antigen binding requires both a VH and a VL domain.

1.2 IgG derivatives and their applications

Although the constant regions of IgG play a role in modulating immune responses, they are not required for antigen recognition and can in some instances be undesirable. Therefore, smaller derivatives of IgG which retain antigen binding specificity have been engineered, firstly by removing the constant domains by enzyme-mediated proteolysis, then later through the use of recombinant DNA technology. Examples of IgG derivatives include monovalent fragments such as Fab (Fragment, antigen binding), single-chain variable fragments (scFv) and single variable VH and VL domains; and bivalent formats such as F(ab')₂ fragments, diabodies and minibodies, as well as multivalent formats (Holliger and Hudson, 2005). Fabs contain the sequences distal to the hinge region while scFvs consist of the VH and VL domains joined by a flexible peptide chain. These derivatives usually show improved tissue penetration capabilities while retaining the specific antigen binding affinity of the parent IgG (Wu and Senter, 2005)

Antibody fusion proteins can also be generated combining the antibody variable domain with other antibodies to generate bispecific molecules (Hoffmann *et al.*, 2005; Schlereth *et al.*, 2005) or with proteins such as toxins, biological response modifiers or enzymes (Halin *et al.*, 2002; Bang *et al.*, 2005; Sharma *et al.*, 2005). It is therefore not surprising that recombinant antibodies and their derivatives are increasingly being developed for use as imaging and diagnostic reagents (Wu and Yazaki, 2000; Michalet *et al.*, 2005; Wei *et al.*, 2008), as therapeutic agents (Adams and Weiner, 2005; Wu and Senter, 2005; Reichert and Valge-Archer, 2007), and for many other biotechnological applications (Ziegler and Torrance, 2002).

1.3 Heavy chain-only antibodies (HCAb) in the antibody repertoire of Camelids

Almost 18 years ago, Hamers-Casterman *et al.* (1993) identified a unique class of antibodies in the serum of members of the *Camelidae* family (camel, dromedary, llama, alpaca and vicuna). These antibodies, called heavy chain-only antibodies (HCAb), lack the light chain normally found in conventional antibodies. They also lack the first constant heavy domain (CH1) which normally associates with the light chain and to a lesser degree, interacts with the variable heavy (VH) domain (Figure 1.2). This is due to removal of the CH1 coding region caused by a point mutation at the 5' end of the CH1 hinge region (Nguyen *et al.*, 1999). Thus HCABs have lost the ability to interact with light chains.

In camels, approximately half of the total IgG repertoire, which has a concentration of approximately 5-10 mg/ml in serum, consists of HCAb (Hamers-Casterman *et al.*, 1993) while in llamas this number is lower (approximately 25-45%) (van der Linden *et al.*, 1999). Approximately 50 – 75% of the IgG repertoire comprises conventional IgG, depending on the species. Nine γ genes have been identified in camels of which five are functional. These five genes code for distinct hinge regions and are considered to correspond to distinct subclasses. A similar pattern was also found for llamas (De Genst *et al.*, 2006). At the protein level, Hamers-Casterman *et al.* (1993) divided the IgG in camels into three different subclasses based on their molecular weight and differential binding to protein A and protein G affinity columns. The IgG1 fraction consists of conventional antibodies while the IgG2 and IgG3 fractions contain HCAb which binds to protein A (IgG2) or both protein A and G (IgG3).

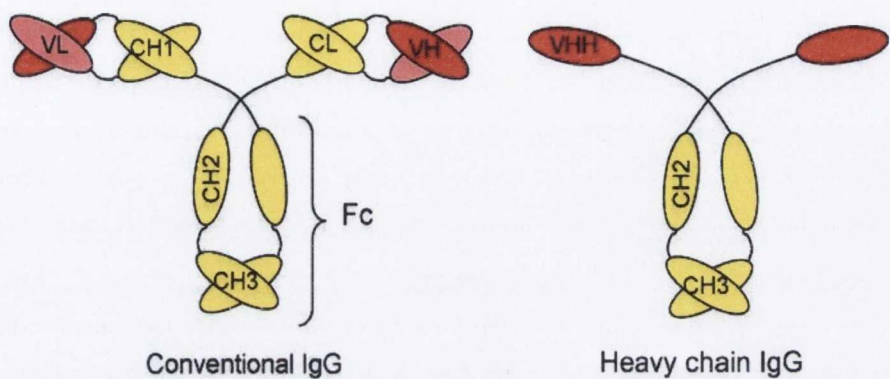


Figure 1.2 Differences between conventional antibodies and HCAs. Variable heavy domains (VH and VHH) are coloured in red, variable light domains (VL) are coloured pink while constant domains (CH and CL) are coloured yellow (adapted from Harmsen and De Haard, 2007). While high affinity antigen binding by conventional IgG requires a VH and a VL domain, HCAs bind antigens via a single VHH domain.

1.4 Heavy chain-only variable fragment (VHH)

The minimal antigen binding site of a HCAb consists of a single domain variable fragment (VHH), also known as a nanobody. Nanobodies have a molecular weight of 14-15 kDa and a structure that is highly similar to the human family III VH domain (Muyldermans *et al.*, 1994; Nguyen *et al.*, 2000). Similar to VH, nanobodies possess four Framework Regions (FRs) that form a β -barrel core or scaffold as well as three Complementarity Determining Regions (CDRs) which are involved in antigen recognition and binding (please refer to Section 5.1 for detailed explanations of the structure of VHHs). Camel VHHs can be subdivided into 7 subclasses (VHH1, VHH2 etc) based on the presence/absence, and the position of a cysteine residue in CDR1 as well as the presence of certain amino acid sequences in their framework regions (Nguyen *et al.*, 1998; Harmsen *et al.*, 2000; Nguyen *et al.*, 2000). VHH2 is the most frequently occurring of the seven VHH subclasses accounting for 75% of all isolated VHHs (Nguyen *et al.*, 2000). For llama, different subfamily divisions are in place (Harmsen *et al.*, 2000; Achour *et al.*, 2008).

1.5 Advantages of Camel VHH antibodies

VHHs have several advantages over their conventional VH counterparts (De Genst *et al.*, 2006; Harmsen and De Haard, 2007). Firstly, they are smaller (≈ 15 kDa) compared to scFvs (≈ 30 kDa), Fab fragments (≈ 60 kDa) and whole IgG (≈ 150 kDa). This advantage, coupled with an extended and highly flexible CDR3, enables VHHs to reach antigenic sites inaccessible to conventional antibodies (Lauwereys *et al.*, 1998; Stijlemans *et al.*, 2004; De Genst *et al.*, 2006). Nanobodies are also more soluble due to their hallmark substitutions, as shown by Davies and Riechmann (1994), who observed an increase in solubility when these hallmark residues were grafted onto conventional VH. Furthermore, hydrophobic patches of the VH-VL interface equivalent are shielded by the CDR3. This, together with the increased solubility, renders the VHHs more homogeneous and less prone to aggregation. Hence, VHH display strict monomeric behaviour unlike VH derivatives which tend to aggregate (Ward *et al.*, 1989).

Another important additional advantage of VHHs is the ease with which they can be obtained. Genes encoding VHH can be easily obtained from immunized camels (or llamas or alpacas) by amplifying the VHH gene repertoire using specific primers,

cloning the genes into phagemid vectors and selecting antigen-specific VHH via phage display technology. VHH libraries from immunized camelids retain full functional diversity since they have been matured *in vivo* against the antigen. In contrast, conventional libraries lose their diversity due to reshuffling of the VL and VH domains during the construction of the library (De Genst *et al.*, 2006). Also, their single domain nature and small size eases genetic manipulation, and antibody diversity can be generated through gene variations instead of through the combination of light and heavy chains. It also enables the generation of multivalent formats (Figure 1.3) to increase avidity or to produce bispecific antibodies (Coppieters *et al.*, 2006; Roovers *et al.*, 2007).

VHHs have been found to be very stable and can retain their functionality at high temperatures (up to 70°C) (Goldman *et al.*, 2006; Ladenson *et al.*, 2006), following exposure to high temperatures (Perez *et al.*, 2001), in the presence of denaturants such as sodium dodecyl sulphate, urea and guanidine-HCl (Dolk *et al.*, 2005), and even in the protease-rich, pH-extreme environment of the gastrointestinal tract (Harmsen *et al.*, 2006; Harmsen and De Haard, 2007; Winichayakul *et al.*, 2009). Their single domain nature means that they only need to refold in order to reacquire functionality.

From an immunotherapeutic perspective, VHH exhibit low immunogenicity because they share a high degree of sequence similarity with the homologous human VH framework (Harmsen and De Haard, 2007). Their small size, enhanced stability and strictly monomeric behaviour facilitate efficient penetration of tissues, conferring in certain cases the ability to cross the blood-brain barrier (Muruganandam *et al.*, 2002). However, their small size also results in rapid renal clearance and, in consequence, a reduced half-life in the blood stream measured in hours, unlike conventional antibodies which can persist for weeks (Harmsen *et al.*, 2005). Fortunately, this potentially disadvantageous trait can be ameliorated by increasing their size by, for example, engineering multivalent formats (Hmila *et al.*, 2008) or by fusing nanobodies to various effector molecules (Cortez-Retamozo *et al.*, 2004; Harmsen *et al.*, 2005).

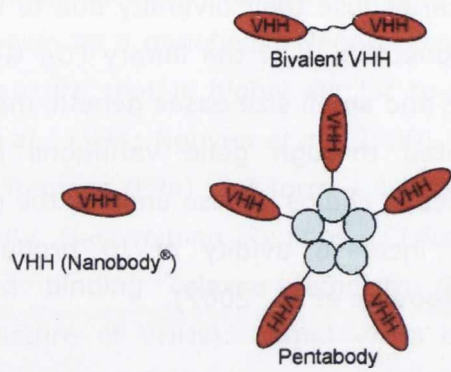


Figure 1.3 Derivatives of HCAb including VHH (nanobody), bivalent VHH and pentabody. VHH domains are coloured red (adapted from Harmsen and De Haard, 2007).

1.5 Advantages of Camel VHH antibodies

From an immunological perspective, VHH exhibit low immunogenicity because they are small and lack the sugar chains that are often found in mammalian antibodies. This makes them ideal for use in diagnostic and therapeutic applications. Additionally, VHH are highly stable and can function in a wide range of environments, including high temperatures and high salt concentrations. This stability is due to the unique structure of the VHH domain, which is highly resistant to denaturation. Furthermore, VHH can be easily modified and fused to other proteins, making them a versatile tool for biotechnology. The small size of VHH also allows them to penetrate tissues more effectively than larger antibodies, making them valuable for drug delivery and imaging. Finally, the high specificity of VHH for their targets makes them ideal for diagnostic and therapeutic applications.

As we have seen, VHH antibodies have many advantages over other types of antibodies. Their small size, high stability, and high specificity make them a valuable tool for a wide range of applications. Additionally, their low immunogenicity makes them ideal for use in diagnostic and therapeutic applications. Finally, their ease of modification and fusion to other proteins makes them a versatile tool for biotechnology.

1.6 Applications of Camel VHH antibodies

Given their unique properties, VHHs have been developed for various biotechnological, diagnostic and therapeutical applications, either as single units or in bivalent or multivalent formats (Saerens *et al.*, 2008). For example, they can be used as capture agents for immunoaffinity purification (Verheesen *et al.*, 2003). Rothbauer *et al.* (2008) fused an anti-GFP VHH to a monovalent matrix which allows for fast and efficient isolation of *in vivo*-expressed GFP fusion proteins and their interacting factors for biochemical and proteomic analyses. Schornack *et al.* (2009) found that such fusion proteins can also be used in plants to mislocalize or 'trap' a protein of interest in the plant cytosol, thereby altering its accumulation profile, and ultimately the phenotype mediated by the protein of interest.

The small size and highly specific binding of VHH also makes them excellent probes for recognizing biomarkers in proteomic studies (Saerens *et al.*, 2008). Pleschberger *et al.* (2004) developed a prostate-specific antigen assay by using a VHH fusion protein to generate a nanopatterned sensing layer to enable the detection of prostate-specific antigen by surface plasmon resonance technology. Others, such as Omidfar *et al.* (2004) and Gottlin *et al.* (2009), explored the use of VHHs in tumour imaging and diagnostics through the use of a VHH recognizing Epidermal Growth Factor Receptor (EGFR), a protein that is normally overexpressed or mutated in a high percentage of cancers.

The fact that VHH can be stable even in the most extreme of conditions has not been overlooked. Dolk *et al.* (2005) isolated a VHH that binds to a cell wall protein of a dandruff-causing fungus even in the presence of shampoo detergents and is stable in the presence of high urea concentrations. Ladenson *et al.* (2006) isolated a thermostable, caffeine-specific llama VHH which can be used in a biosensor application to analyse caffeine content in both hot and cold beverages. Several VHHs have also been shown, in some instances, to be fully functional in the reducing environment of the cytoplasm, thus making them attractive candidates as intrabodies or as immunomodulators of enzyme function. Tanaka and Rabbitts (2003) isolated a VHH against the oncogenic protein RAS and when transfected into NIH3T3 cells, it inhibited mutant RAS-induced oncogenic transformation in the cells. Serruys *et al.* (2009) isolated a VHH against the envelope protein of Hepatitis B Virus (HBV) and when expressed and retained in the ER, it reduced the levels of secreted Hepatitis B surface antigen particles both *in vivo* and *in vitro*. In plants,

Jobling *et al.* (2003) demonstrated that VHH can be used for *in vivo* immunomodulation of enzyme activity. In this case, the activity of potato starch branching enzyme A (SBEA) was inhibited by the expression of an anti-SBE A VHH targeted into tobacco chloroplasts and the level of inhibition was found to be more efficient than antisense-RNA technology.

VHHs also have the potential to be used as therapeutics. For example, a camelid anti-human lysozyme VHH was found to inhibit the formation of amyloid fibrils by human lysozyme (Dumoulin *et al.*, 2003; Chan *et al.*, 2008). Baral *et al.* (2006) on the other hand created a trivalent bispecific conjugate by fusing a VHH targeting the conserved cryptic epitopes of African trypanosomes that cause sleeping sickness, with human trypanolytic factor Tr-*apoL*-I which lyses most trypanosomes. Treatment with this conjugate cured or alleviated acute and chronic trypanosome infections in mice. Other conjugates of VHH with potential medical applications include a bivalent molecule consisting of both anti-human and anti-mouse Tumour Necrosis Factor (TNF)- α with an anti-TNF VHH domain which was found to be 500 times more potent than monovalent molecules and the current antibodies used in the treatment of rheumatoid arthritis (Coppieters *et al.*, 2006).

Furthermore, Arbabi-Ghahroudi *et al.* (1997) isolated VHHs which bind to and neutralize toxins from *Clostridium tetani*. Another group isolated an anti-HIV VHH which neutralizes HIV-1 primary isolates of subtype B and C, thus inhibiting HIV-1 infection (Forsman *et al.*, 2008). VHH has also the potential of being used in passive immunotherapy. Pant *et al.* (2006) expressed a VHH which neutralizes diarrhoea-inducing rotavirus on the surface of lactobacilli and it was found that oral administration of the lactobacilli shortened the disease duration, severity and viral load in mouse. VHH also can be used as a prophylactic against dental caries. Krüger *et al.* (2006) isolated a llama VHH against *Streptococcus mutans* and its daily administration significantly reduced the development of smooth surface cavities.

1.7 The search for a 'universal' VHH framework and the 'humanization' of VHH

It has always been the goal of antibody engineering to minimize the size of antigen-binding molecules. Hence the reduction of multidomain proteins such as monoclonal IgG or HCAb to minimal sized fragments that retain sufficiently high antigen specificity and affinity such as scFv and VH. Additional desired criteria include increased shelf-life and functionality even after exposure to harsh environments such as high concentrations of denaturant (e.g. after regeneration of biosensors), elevated pH (e.g. in the stomach or in the reducing environment of the cell cytoplasm) or elevated temperatures. In the case of nanobodies, Saerens *et al.* (2005) identified a highly stable VHH framework which they propose as a 'universal' scaffold onto which novel antigen-binding specificities can be grafted via CDR loop exchange (please refer to Section 4.1 for details).

'Humanization' of camelid-specific amino acid sequences in the framework regions to their human VH equivalents is a likely required step if camel antibodies are to be accepted as human therapeutics. Vincke *et al.* (2009) assessed the biochemical properties of several VHHs after the humanization process. They chose cAbBCII10, the scaffold Saerens *et al.* (2005) found to have the greatest potential as a universal CDR graft acceptor and cABHuL6, an anti-human lysozyme VHH (Dumoulin *et al.*, 2002). They then demonstrated that conserved VHH residues at positions 37 and 47 (Kabat numbering; Kabat *et al.*, 1991) have a major impact on the affinity of the VHH for its antigen due to repositioning of the CDR3 loop. Surprisingly, humanizing the residues at positions 44 and 45 stabilized the VHH without dramatically affecting solubility and antigen affinity. Moreover, the substitution of 11 'non-human' residues outside FR2 did not significantly affect the antigen-binding capacity of the VHH but caused a decrease in expression, yield and stability of the VHH. This is consistent with the finding that some of the substituted residues were previously found to affect the stability of llama VHH fragments (Tanha *et al.*, 2002). However, humanization of residues 44 and 45 was sufficient to compensate for this loss in stability. Furthermore, they successfully grafted the CDRs of different VHHs onto the humanized framework of cAbBCII10 (h-NbBCII10_{FGLS}) resulting in transfer of the antigen specificity of the loop donor. This showed that the scaffold could also be used as a humanized universal framework.

1.8 Production platforms for derivatives of conventional and Camel antibodies

Many different expression systems have been used to produce monoclonal antibodies (mAbs) as well as other recombinant proteins used for diagnostic, research and therapeutic applications. Single celled prokaryotes and yeast are the most commonly used platforms. mAbs, for example, have been produced in prokaryotic systems such as *Escherichia coli* (Martineau *et al.*, 1998; Müller-Schiffmann *et al.*, 2009). In fact, *E. coli* is the platform commonly used for the production of antibody fragments lacking the Fc region which do not require N-glycosylation (Andersen and Reilly, 2004; Arbabi-Ghahroudi *et al.*, 2005). Furthermore, mAbs have also been expressed in cell free systems (Hoffmann *et al.*, 2004; Galeffi *et al.*, 2006), in yeasts such as *Saccharomyces cerevisiae* (Shusta *et al.*, 1998; Miller *et al.*, 2005) and *Pichia pastoris* (Shi *et al.*, 2003; Gurkhan *et al.*, 2004).

However, eukaryotic systems such as Chinese Hamster Ovary (CHO) cells, baculovirus-infected insect cells, and transgenic plants or plant cell cultures are increasingly being used as alternative production platforms because they are capable of assembling complex proteins and, moreover, provide post-translational modifications like glycosylation (Kamarck, 2006). Various mAbs, including scFvs and VHs, have been produced in eukaryotic systems including insect cells (Kretzschmar *et al.*, 1996; Johansson *et al.*, 2007), fungi (Joosten *et al.*, 2005), cultured mammalian cells (Cardinale *et al.*, 1998; Bazl *et al.*, 2007; Braren *et al.*, 2007) and plants (Ma *et al.*, 1995; Schouten *et al.*, 1997; Fischer *et al.*, 1999; Ismaili *et al.*, 2007; Van Droogenbroeck *et al.*, 2007; Winichayakul *et al.*, 2009; Conrad and Floss, 2010; Conrad *et al.*, 2011).

1.9 Transgenic plants as an expression platform

Plants have generated great interest as an attractive alternative to mammalian and microbial cell culture platform for heterologous protein production (Twyman *et al.*, 2003; Bock, 2007). Apart from avoiding the ethical problems associated with transgenic animals and the expense of mammalian cell culture systems, factors in favour of using plants as a production system include low cultivation/production

costs and the ease of scalability via standard agricultural practices. Other important advantages include the greatly reduced risk of product contamination by mammalian viruses, blood-borne pathogens, and bacterial toxins, since plants are not hosts for the relevant organisms. Furthermore, plants are capable of synthesizing correctly folded and assembled multimeric therapeutic proteins of mammalian origin. In fact, plants have been found to be more suitable for producing properly folded, active recombinant proteins compared to yeast (*Saccharomyces cerevisiae*) or *Escherichia coli* cells (Popescu *et al.*, 2007).

In the past, the difference between the N-glycosylation patterns of proteins produced by mammalian and plant cells raised the concern that plant-produced therapeutics might elicit an undesirable immune response in humans. However, *Nicotiana benthamiana* has been genetically modified to efficiently produce fully functional proteins with human-like glycosylation patterns, thus solving a major hurdle to using plant systems as a production platform (Strasser *et al.*, 2008). Furthermore, Vézina *et al.* (2009) successfully incorporated a chimeric human β 1,4-galactosyltransferase into a transient transformation vector and successfully produced antibodies with human-like N-glycans at high levels and in a short time. There is also the possibility of directly using edible plant material expressing therapeutics (e.g. vaccines and prophylactics) as food or animal feed, hence reducing downstream processing procedures (Walmsley and Arntzen, 2000; Mason *et al.*, 2002). Verma *et al.* (2010) bioencapsulated coagulation factor IX in tobacco chloroplasts and showed that oral delivery of this protein effectively blocked inhibitory antibody formation and ultimately eliminated fatal anaphylactic reactions in Haemophilia B mice. Transgenic plants can also be generated for the purpose of phytoremediation. In fact, recombinant tobacco plants expressing membrane-bound or rhizosecreted antibodies against the environmental pollutant microcystin-LR have been produced. Microcystin-LR can either be taken up by the plants and be bound to the antibody fused to the plasma membrane (Barbi *et al.*, 2010) or the antibodies can be secreted from the roots to the surrounding environment to bind to the pollutant (Drake *et al.*, 2010).

There are several established protocols for introducing a gene of interest (GOI) into plants using either *Agrobacterium* or biolistic methods and it is possible to introduce additional (and multiple) transgenes via sexual crosses (gene stacking; Halpin and Boerjan, 2003). Transgenes introduced into plants typically encode proteins responsible for either input traits or output traits. Input traits include drought, herbicide or insect resistance, designed to improve the agronomic performance of

the plant. Output traits, on the other hand, change the way the plants are subsequently used. These include, for example, modification of nutritional value, as has been achieved in the case of vitamin-enriched Golden rice (Ye *et al.*, 2000; Datta *et al.*, 2003); biofuel production (Andrianov *et al.*, 2010), and the production of recombinant therapeutic proteins.

The production of recombinant proteins of diagnostic, therapeutic and biotechnological interest has been extensively investigated in transgenic plants. These include candidate vaccine antigens (Mason *et al.*, 2002; Streatfield and Howard, 2003; Pogrebnyak *et al.*, 2005; Azhakanandam *et al.*, 2007; McCabe *et al.*, 2008; Peters and Stoger, 2011); proteins providing prophylaxis against organophosphorous nerve agents and pesticides (Geyer *et al.*, 2010) and the inhalation of anthrax spores (Wycoff *et al.*, 2011); biopharmaceuticals such as human serum albumin (Fernandez-San *et al.*, 2003), human growth hormone (Staub *et al.*, 2000), and collagen (Merle *et al.*, 2002); proteins for industrial uses (Somerville and Dario Bonetta, 2001) such as silk-like proteins (Yang *et al.*, 2005); and antibodies. The latter include various antibody formats (please see Ma *et al.*, 2003; Schillberg *et al.*, 2003; Ko and Koprowski, 2005; Conrad and Floss, 2010; Peters and Stoger, 2011 for reviews), such as full-length conventional IgG (Hiatt *et al.*, 1989; Ma *et al.*, 1994) and sIgA antibodies (Hiatt *et al.*, 1989; Ma *et al.*, 1995; Voss *et al.*, 1995), as well as scFvs (Fiedler *et al.*, 1997; Fischer *et al.*, 1999; Schillberg *et al.*, 2000; Van Droogenbroeck *et al.*, 2007), Fabs (De Wilde *et al.*, 2002), and more recently, VHHs (Ismaili *et al.*, 2007; Winichayakul *et al.*, 2009). However, it should be noted that in stably transformed plants, the expression levels achieved for recombinant proteins encoded by nuclear transgenes do not, in general, exceed 1-2% total soluble protein (tsp).

1.10 Plant tissue and cell culture

Plant tissue culture is the *in vitro* multiplication and/or regeneration of plant tissues under sterile conditions after the original organisms has been removed. This is possible due to the totipotent nature of plant cells which can give rise to a whole fertile plant without the loss of genetic information (Razdan, 2003). Through the *in vitro* use of phytohormones, tissue explants can be induced to dedifferentiate, forming callus tissues, which continue to propagate in this state. Sometimes, as in

the case of *Agrobacterium*-mediated transformation, the callus tissues can be redifferentiated to form whole plants by changing the ratio of the phytohormones auxin and cytokinin.

There are several approaches which can be used for *in vitro* cultivation of plant cells and tissues, including the generation and propagation of 'hairy' roots, shooty teratomas, immobilized cells and cell suspension cultures (Hellwig *et al.*, 2004). The major focus tends to be on suspension cultures and hairy root cultures because they are the most amenable to GMP procedures and can be cultivated easily in large-scale bioreactors (Lee *et al.*, 2004). Plant cell suspension cultures are prepared by the agitation of friable callus tissue in shaker flasks or fermenters to form single cell and small aggregate suspensions. The resulting cells suspensions are grown in liquid culture medium containing the same hormones to promote rapid growth and prevent redifferentiation. The most commonly used plants for protein production by this route are tobacco, in particular the cell lines referred to as Bright Yellow2 (BY2) (e.g. Fischer *et al.*, 1999b; Fujiyama *et al.*, 2007). Soybean (Höhl *et al.*, 1988) and tomato cell cultures (Kwon *et al.*, 2003) have also been investigated. Hairy root cultures on the other hand are generated by infection of leaf or shoot material with *Agrobacterium rhizogenes*. This induces the production of genetically transformed adventitious roots (so-called 'hairy' roots) which are capable of unlimited propagation in hormone-free media without showing any further differentiation into shoot/leaf tissues.

Plant cell cultures and hairy root cultures have been applied in many fields ranging from molecular biology to crop improvement. They have also been used to produce various recombinant proteins including mAbs and antibody fragments (Fischer *et al.*, 1999b; Ramírez *et al.*, 2000; Semenyuk *et al.*, 2003; Francesca and Lorenza, 2008), enzymes (Kurata *et al.*, 1998; Verdelhan Des Molles *et al.*, 1999; Huang *et al.*, 2002) as well as proteins of therapeutic value such as human interleukin (IL) (Magnuson *et al.*, 1998; Kwon *et al.*, 2003b), human α 1-antitrypsin (Huang *et al.*, 2001) and human serum albumin (Sijmons *et al.*, 1990). The recombinant protein can be produced in suspension cultures either by using transgenic plants as the initial starting material for callus generation, or by transforming wild-type cell suspensions with the plasmid harbouring the GOI via *Agrobacterium*-mediated transformation or particle bombardment. Then, the expressed protein can be harvested from the culture biomass or culture medium or both.

1.11 Strategies for gene expression

There are two general strategies for expressing a gene of interest (GOI) in plant cells: expression via the production of stable transformants (generally referred to as transgenic plants), or transient expression.

A GOI can be stably incorporated into the plant nuclear genome by *Agrobacterium*-mediated delivery of T-DNA (Gelvin, 2003), by biolistics (Sanford, 1988; Sanford *et al.*, 1993), by polyethyleneglycol (PEG)-mediated gene transfer (Uchimiya *et al.*, 1986), electroporation of protoplasts and intact cells (Fromm *et al.*, 1985; Lörz *et al.*, 1985; De la Pena *et al.*, 1987; Arencibia *et al.*, 1995), or by microinjection (De la Pena *et al.*, 1987). Integration of a GOI, delivered by biolistics or via PEG-mediated transfer, into the chloroplast genome typically occurs via homologous recombination between sequences flanking the GOI and the corresponding endogenous sequences within the chloroplast. Stably transformed plants, regenerated from the initially transformed cells, offer the advantage of stable expression and inheritance of the GOI. Furthermore, crossing different transgenic lines can be employed to combine or stack multiple transgenes in the same plant (Halpin, 2005). Stable transformation is further discussed in Section 7.1.

In contrast, in the transient expression strategy the GOI is expressed for only a short time period following its introduction into the cell and regeneration of whole plants from the initially transformed cells is not pursued (Janssen and Gardner, 1990; Kapila *et al.*, 1997). In this scenario, the GOI cannot be passed on to the next generation. Transient expression methods are, however, easier and less costly to use. Moreover, they also have a much faster timeline, measured in days rather than months and also circumvent the problem of highly variable expression, because expression of the GOI is not dependent on stable incorporation into the plant genome.

The simplest strategy for transient expression consists of inserting the GOI, flanked by appropriate 5' and 3' regulatory sequences, into the T-DNA of an *Agrobacterium* Ti plasmid (for example Popescu *et al.*, 2007). More sophisticated transient expression vectors contain part or the entire genome of a plant RNA virus, such as the tobamovirus Turnip Vein Clearing Virus (TVCV) or Tobacco mosaic virus (TMV), or alternatively, Potato Virus X (PVX) (Chapman *et al.*, 1992) or Cowpea Mosaic

Virus (CPMV) (Sainsbury and Lomonossoff, 2008). The earliest viral vectors were based on infectious cDNA clones of the wild-type viral genome delivered into the plant by mechanical inoculation. The GOI was typically introduced as an additional ORF into the viral genome or, alternatively, it replaced a viral gene such as that encoding the coat protein (CP) (Scholthof *et al.*, 1993; Scholthof *et al.*, 1996). Expression of the GOI was driven by the addition of an extra viral subgenomic promoter (Dawson *et al.*, 1989; Donson *et al.*, 1991), or the GOI was expressed as a translational fusion with the viral coat protein (Werner *et al.*, 2006). These various expression strategies are known as the 'full virus' strategy. Disadvantages of using this approach, however, include size restrictions on the sequences that can be inserted and are compatible with viability of the virus (Avesani *et al.*, 2007), the high mutation rate to which the GOI is subjected during the process of viral RNA replication, and the tendency of the viral genome to delete the GOI (Dawson *et al.*, 1989).

The disadvantages of the 'full virus' vector strategy led to the development of novel 'deconstructed virus' strategies. Here the viral genome is 'deconstructed', leaving only those genes required for expression of the GOI and amplification of the corresponding mRNA transcript. The missing functions such as those controlling initial infection of the host plant and DNA integration are replaced with non-viral components. Instead of delivery of the GOI by virus-mediated infection, *Agrobacterium* is used to deliver T-DNA containing the GOI flanked by regulatory elements derived from the viral genome. Following T-DNA delivery, the transcribed mRNA is able to self-replicate in the cytoplasm due to the presence of the viral replicase complex. Therefore, these vectors not only take advantage of the presence of the RNA-dependent RNA replication functions of the virus to further amplify the initially transcribed mRNA, they also make use of the efficient manner with which *Agrobacterium* transfer T-DNA into the plant host cell nucleus. Examples of 'deconstructed virus' vectors include the tobamovirus-based ICON Genetics system (Marillonnet *et al.*, 2004; Gleba *et al.*, 2007; please refer to Section 3.1 for detailed explanation), the TMV RNA-based Overexpression (TRBO) vector (Lindbo, 2007) and the CPMV RNA2-based pEAQ-HT vector (Sainsbury and Lomonossoff, 2008; please refer to Section 6.1 for detailed explanations of TRBO and pEAQ-HT vectors).

1.12 Research aims and objectives

As discussed above, plants represent an attractive alternative to mammalian and microbial cell culture systems as a potential production platform for recombinant proteins, including antibodies.

The aim of the present study is to identify and investigate key factors influencing the expression, yield, functionality, stability, and recovery of antibody fragments in plants; specifically, murine scFv and camelid VHH (with major emphasis on the latter).

The research programme will therefore investigate:

1. Gene-intrinsic factors: i.e. the effects of (i) codon optimization (ii) changes in VHH framework regions, and (iii) purification tags.
2. The influence of the subcellular compartment in which the antibodies accumulate i.e. cytoplasm, chloroplast, ER, and apoplast.
3. The use of an alternative 'universal' framework regions, derived from cAbBCII10 to increase the yield of VHHs in general; and the effects of individual framework region changes on yield.
4. The use of different virus-based transient expression systems, as well as *E. coli* expression system, for expression of VHHs in *N. benthamiana*.
5. The expression system: high-level transient expression systems vs. expression in stable nuclear transformants.
6. Different strategies to improve expression in stable nuclear transformants: i.e. (i) the use of different promoters, (ii) the inclusion of an intron, and (iii) the use of 5' and 3' untranslated regions (UTRs) from Tobacco Mosaic Virus (TMV) and Cowpea Mosaic Virus (CPMV).

Chapter 2. Materials and Methods

2.1 Expression vector modules

2.1.1 ICON Genetics system provector modules

The tobamovirus-based expression system developed by the biotechnology company ICON Genetics (Marillonnet *et al.*, 2004; Santi *et al.*, 2006) for transient expression of genes of interest (GOIs) comprises 3 separate components: 1) A 5' provector module encoding the Turnip vein-clearing virus (TVCV) RNA-dependent RNA polymerase gene (RdRP) driven by an *Arabidopsis Actin2* promoter, the crucifer-infecting strain of Tobacco mosaic virus (Cr-TMV) movement protein gene (MP), subcellular targeting signals (for pICH8420 and pICH10530), a 5' splice donor site and half intron, and the phage C31 attP recombination site; 2) A 3' provector module consisting of an attB site, half intron and 3' splice acceptor site, and a Multiple Cloning Site (MCS); 3) An integrase expression vector encoding phage PhiC31 integrase fused to a nuclear localizing signal (NLS) and driven by an *Arabidopsis* heat shock protein Hsp81.1 promoter. The GOI is cloned into the MCS of the 3' provector module.

Figure 2.1 shows schematic diagrams of the provector constructs and the functions of individual modules are summarized in Table 2.1.

Table 2.1 Names and descriptions of ICON provectors modules

Name	Description
pICH10881	Encodes phage Phi C31 integrase which joins a 5' provector module to the 3' module via site-specific recombination.
pICH10570	A 5' provector module that lacks subcellular protein targeting signals; the protein encoded by the GOI therefore accumulates by default in the cytoplasm
pICH8420	A 5' provector module which contains an apoplast targeting presequence derived from tobacco calreticulin; the protein encoded by the GOI accumulates extracellularly in the apoplast
pICH10530	A 5' provector module which contains a chloroplast targeting presequence; the protein encoded by the GOI accumulates in chloroplasts.
pICH11599	3' provector module; Contains a multiple cloning site (MCS) for insertion of a GOI

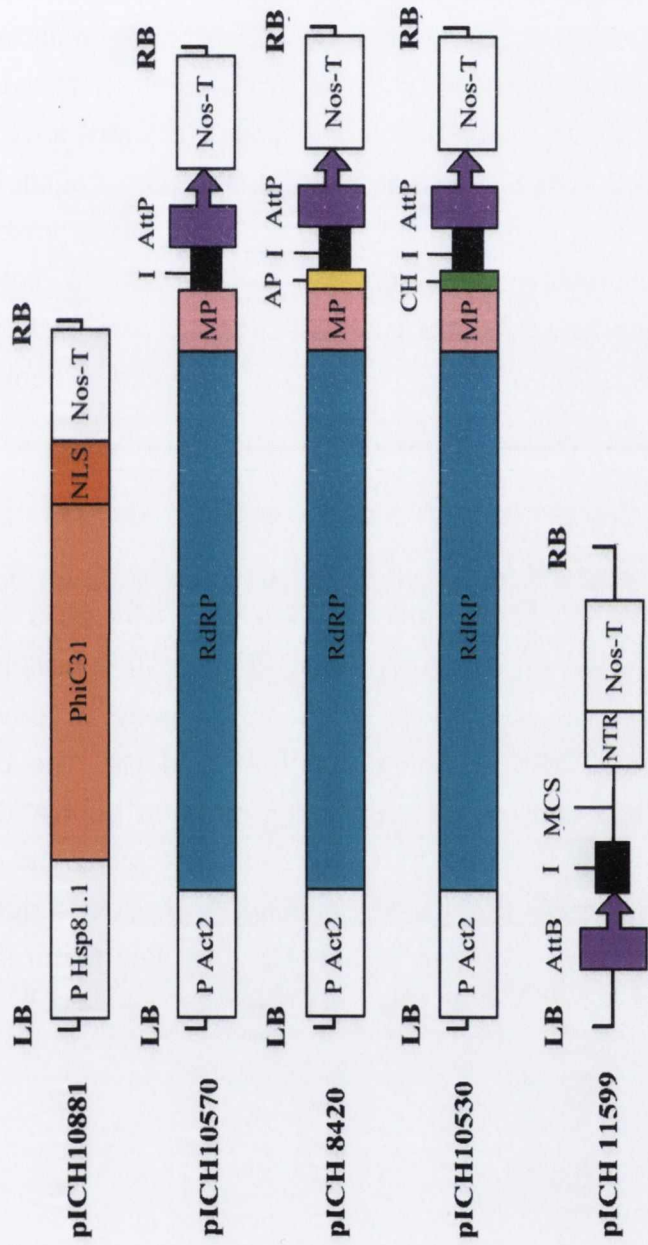


Figure 2.1 ICON Genetics provector modules. LB – Left border; RB – Right border; MCS – Multiple cloning site for inserting a gene of interest (GOI); NTR – Non-translated region; PhiC31 – Integrase; NLS – Nuclear localising signal; P Act2 – *Arabidopsis* actin 2 promoter; RdRP – RNA dependent RNA polymerase; MP – Movement Protein; I – Intron; AP – Tobacco calreticulin apoplast targeting presequence; CH – Synthetic chloroplast targeting presequence; AttP/AttB – recombination sites. PhiC31-mediated site-specific recombination between AttB and AttP, which occurs in *planta*, results in the joining of the 3'-module, which contains the GOI, to one of 3 possible 5'-modules (adapted from Marillonnet *et al.*, 2004; Santi *et al.*, 2006).

2.1.2 The CPMV expression vector, pEAQ-HT

The Cowpea Mosaic Virus (CPMV)-based expression system pEAQ-HT (Sainsbury and Lomonossoff, 2008) facilitates transient expression of a GOI. It comprises the following components: 1) The 5' untranslated region (UTR) of CPMV RNA-2 with a U161 $\square$ C mutation to remove an in-frame start codon upstream of the main start codon at position 512. This region is known as the hypertranslatable (HT) leader (Sainsbury and Lomonossoff, 2008); 2) A multiple cloning site (MCS) with a 6x His tag; 3) The 3' UTR of CPMV RNA-2. The whole cassette is driven by a Cauliflower Mosaic Virus (CaMV) 35S promoter. In addition, pEAQ-HT contains an expression cassette encoding Tomato Bushy Stunt Virus (TBSV) P19 protein, a potent suppressor of RNA silencing (Voinnet *et al.*, 2003). The P19 gene is flanked by a CaMV35S promoter and CaMV35S terminator. Figure 2.2 shows a schematic diagram of the pEAQ-HT vector.

2.1.3 TRBO expression vectors (pJL-TRBO)

The TMV-based expression system pJL-TRBO (Lindbo, 2007) also facilitates transient expression of a GOI. It consists of: 1) A CaMV35S promoter driving transcription of the TMV genome region encoding the RdRP and MP. The coat protein (CP) gene has been deleted and the CP start codon mutated from ATG to AGA; 2) A MCS with sites for the enzymes PacI, AvrII, NotI was inserted at the end of the MP; 3) A ribozyme gene and the CaMV 3' polyA signal/transcription terminator were placed immediately downstream of the MCS. The key features of pJL-TRBO are shown in Figure 2.3.

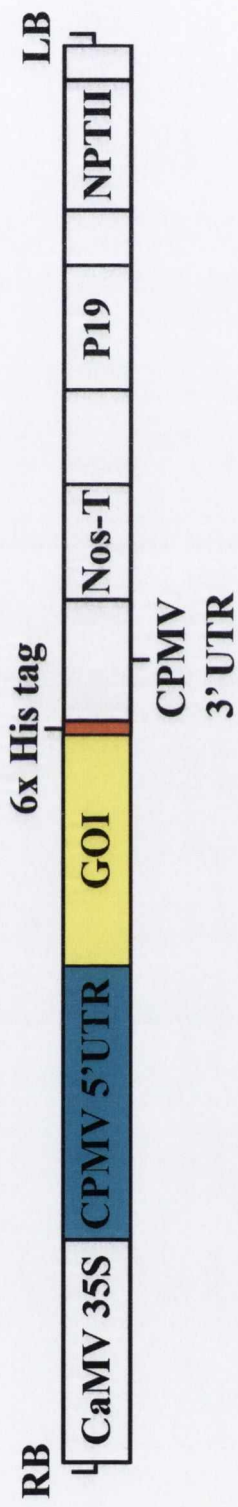


Figure 2.2 The pEAQ-HT vector. RB - Right border; LB - Left border; CaMV 35S - Cauliflower Mosaic Virus 35S promoter; GOI - Gene of interest; UTR - untranslated region; Nos-T: Nopaline synthase terminator; P19 - Tomato Bushy Stunt Virus p19 suppressor of silencing (Voinnet *et al.*, 2003) flanked by CaMV35S promoter and CaMV35S terminator.

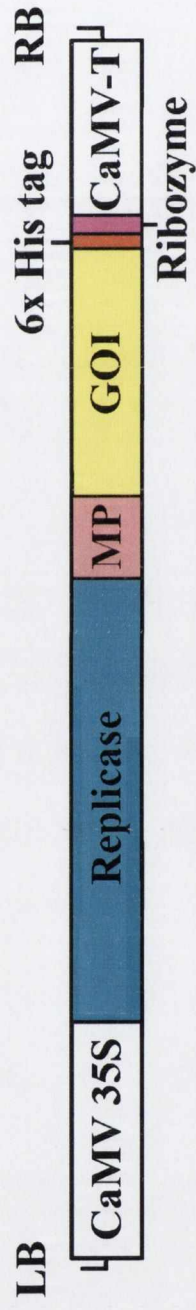


Figure 2.3 pJL-TRBO vector. LB-Left border; RB-Right border; CaMV 35S- Cauliflower Mosaic Virus duplicated 35S promoter; Replicase - TMV RdRP; MP - Movement protein; GOI - Gene of interest; Ribozyme gene; CaMV-T: CaMV 3' polyA signal/transcription terminator (Adapted from Lindbo, 2007).

2.2 scFvs and VHHs investigated

Two genes encoding human single chain variable fragment antibodies (scFvs) and 8 genes encoding camel heavy chain antibody fragments (VHHs) were investigated in this study. The antibodies and their respective antigens are detailed in Table 2.2

2.2.1 Design and synthesis of CLDSyn and CBESyn

For CLDSyn and CBESyn, codon optimization was performed using the web-based software GENEius Lite (Koehler and Wagner, 2007), using the sequences cAB-Lys3 (Arbabi-Ghahroudi *et al.*, 1997) and cAbBCII10-L-L-L (Saerens *et al.*, 2005) as a basis for CLDSyn and CBESyn respectively. The codon usage frequencies for *N. benthamiana* were obtained from the Codon Usage Database (Nakamura *et al.*, 2000). Any undesirable sequences such as rare codons, direct or indirect repeats, potential hairpin structures, polyadenylation signals (AATGAA, AATAAA, AATAAT, AATATT, AATTAA, GATAAA), destabilizing sequences (ATTTA), and 5' intron splice signals (GTAAGT and GTACGT) were also removed. Priority was given to the removal of rare unwanted sequences while maintaining codon diversity. Furthermore, restriction sites were introduced into the gene sequence (3 for CLDSyn and 4 for CBESyn) to facilitate subsequent manipulations. A His-tagged version of both genes with an NcoI site at the 5' end and a BamHI site at the 3' end were synthesized by Genscript, USA.

2.2.2 Design and synthesis of PS11, CPS, RN5 and CRN

The sequence of PS11 and RN5 are based on cAb-PSA11 and cAbRn05 (Saerens *et al.*, 2005) respectively. On the other hand, CPS and CRN are based on cAbBCII10-P-P-P and cAbBCII10-R-R-R (Saerens *et al.*, 2005). The genes were optimized as per Section 2.2.1 with the codons kept as close as possible to CLSynHS (for PS11 and RN5) or CBEHS (for CPS and CRN). Codons which are different from CLSynHS and CBEHS were replaced with the most frequently used *N. benthaminana* codons obtained from the Codon Usage Database (Nakamura *et al.*, 2000) wherever possible without introducing undesirable sequences. An NcoI site was added at the 5' end while a 6xHis tail and a BamHI site were added at the 3' end. The genes were synthesized by Genscript, USA.

Table 2.2 The scFvs and VHHs investigated and their respective antigens.

Name	Abbreviation	Antibody class	Antigen
Anti β -galactosidase	163	scFv	β -galactosidase
Anti β -galactosidase AMEF mutant	ABG	scFv	β -galactosidase
Anti hen egg white lysozyme	CLD	Class II VHH	Hen egg white lysozyme (HEWL)
Anti tetanus toxoid	CTT2	Class II VHH	Tetanus toxoid
Codon optimized synthetic gene of CLD (See part 2.2.1 for details)	CLDSyn	Class II VHH	Hen egg white lysozyme (HEWL)
Codon optimized synthetic gene with cBCII10 framework and cAb-Lys3 CDRs (See part 2.2.1 for details)	CBESyn	Class II VHH	Hen egg white lysozyme (HEWL)

Table 2.2 (continued) The scFvs and VHHs investigated and their respective antigens.

Name	Abbreviation	Antibody class	Antigen
Codon-optimized synthetic Anti Prostate Specific Antigen (PSA) (See part 2.2.2 for details)	PSA	Class II VHH	Human Prostate Specific Antigen (PSA)
Codon-optimized synthetic gene with cBCII10 framework and cAb-PSA11 CDRs (See part 2.2.2 for details)	CPS	Class II VHH	Human Prostate Specific Antigen (PSA)
Codon-optimized synthetic Anti RNase (See part 2.2.2 for details)	RN5	Class II VHH	Bovine RNase (Type XII-A)
Codon-optimized synthetic gene with cBCII10 framework and cAb-RN05 CDRs (See part 2.2.2 for details)	CRN	Class II VHH	Bovine RNase (Type XII-A)

2.3 Cloning of GOIs into vectors for transient expression

2.3.1 PCR amplification of GOIs

The scFvs and VHHs were amplified using the required forward and reverse primers (as detailed in Table 2.3). The template was amplified using Pfu or Taq DNA polymerase, 0.2 mM each dNTP, 1 μ M forward and reverse primer and 10 ng template. The PCR conditions included an initial denaturation step (2 minutes at 90°C), followed by 30 consecutive cycles comprising: denaturation (1 minute at 90°C), annealing (1 minute at 57°C) and elongation (45 seconds at 72°C); and a final elongation (10 minutes at 72°C) step.

Table 2.3 Template, forward and reverse primers used for amplification of GOI. Start and stop codons are in bold; Restriction sites (NcoI and BamHI) are underlined; His- or Strep-tag sequences are in italics.

Name	Template	Forward primer	Reverse primer
163	Anti-β-galactosidase (Martineau <i>et al.</i> , 1998)	5'ATTCC <u>ATG</u> GGCCGAGGTG CAGCT-3'	His tag-containing: 5'TAGGTGCGGCCGCAGGCCCATCACCATCACCT <u>TAAGGATCCGCA</u> -3' Strep tag-containing: 5'TAGGTGCGGCCGCAGGCCGGTTGGTCTCATCTCAATTTGAAAAA <u>TAAGGATCC</u> GCA-3'
ABG	Anti-β-galactosidase AMEF mutant (Martineau <i>et al.</i> , 1998)	Similar to 163	Similar to 163
CL3	Anti-HEWL VHH (cAb-Lys3) (Arbabi-Ghahroudi <i>et al.</i> , 1997)	5'ATCC <u>ATG</u> GGCCCGAGGTG AG-3'	His tag-containing: 5'GTCTCCTCAGGCCCATCACCATCACCTGAGGATCCGCA-3' Strep tag-containing: 5'GTCTCCTCACACGGCGCGGTTGGTCTCATCTCAATTTGAAAAA <u>TAAGGATCCGCA</u> -3'

Table 2.3 (continued)

Template, forward and reverse primers used for amplification of GOI. Start and stop codons are in bold;
An Nco I restriction site is underlined; 6xHis- or Strep-tag sequences are in italics.

Name	Template	Forward primer	Reverse primer
CLD	Similar to CL3	5'ATCC ATG <u>GACGTGCAGCTGCAG</u> -3'	Similar to CL3
CTT2	Anti-tetanus toxoid VHH (cAb-TT2) (Arbabi-Ghahroudi <i>et al.</i> , 1997)	Similar to CLD	Similar to CLD
CLDSyn	Codon optimized synthetic version of CLD	(See part 2.3.2 for details)	
CBESyn	Codon optimized synthetic gene with the framework of cBCII10 and the CDRs of cAb-Lys3 (Saerens <i>et al.</i> , 2005)	(See part 2.3.2 for details)	

2.3.2 Cloning of synthetic genes into pICH11599

The synthetic genes CLDSyn, CBESyn, PSA, CPS, RN5 and CRN were designed to include an NcoI site at the 5' end and a 6xHis-tag followed by a BamHI site at the 3' end. The genes were cloned into the Multiple Cloning Site of pICH11599 using the NcoI and BamHI restriction sites following gel purification of the digested insert. To make the corresponding Strep-tagged constructs, the synthetic genes were amplified with the forward primer 5'-ATCCATGGACGTTTCAGTTG-3' (NcoI site underlined) and the reverse primer: 5'-GTAACAGTCTCTTCCGGCGCCGGTTGGTCTCATCCTCAATTTGAAAAATGAGGATCCGCA-3' (BamHI site underlined).

2.3.3 Cloning of SEKDEL-tagged genes into pICH11599

Three different versions of SEKDEL-tagged CBESyn were made: 1) CBESyn with a 3' SEKDEL sequence (CBEKDEL); 2) CBESyn with a 5' 6xHis tag and 3' SEKDEL sequence (HSCBEKDEL); 3) CBESyn with a 3' 6xHis tag followed by the SEKDEL sequence (CBEHSKDEL). Please refer to Figure 4.6 for a schematic diagram of the different SEKDEL-tagged constructs.

All 3 constructs were amplified by PCR using pUC57-CBESyn as the template and PCR conditions as detailed in Section 2.3.1. The respective forward and reverse primers used to amplify the inserts are detailed in Table 2.4. For CBEKDEL and HSCBEKDEL, the genes were cloned into the MCS of pICH11599 using the NcoI and HindIII sites. For CBEHSKDEL, the PCR product was cloned into the gel-purified NcoI-BamHI large vector-containing fragment of pICH11599 CBEKDEL.

Table 2.4 Forward and reverse primers used for amplification of KDEL-tagged CBESyn constructs. Start and stop codons are in bold; Restriction sites (NcoI, BamHI, HindIII) are underlined; SEKDEL sequences are in italics; 6xHis- sequence in lower case

Name of Construct	Forward primer	Reverse primer
CBEKDEL	5'ATCC <u>ATG</u> GACGTTTCAGTTG-3'	5'CAGGTAACAGTCTCTTCCGGATCCGAGAAAGGATGAGTTG <u>TGAA</u> AGCTT-3'
HSCBEKDEL	5'ATCC <u>ATG</u> GGGAcaccatcacca tcacCTGCAGGCTGACGTTTCAGT TGGT-3'	5'CAGGTAACAGTCTCTTCCGGATCCGAGAAAGGATGAGTTG <u>TGAA</u> AGCTT-3'
CBEHSKDEL	5'ATCC <u>ATG</u> GACGTTTCAGTTG-3'	5'CAGGTAACAGTCTCTTCCGGAcaccatcaccaccac <u>GGATC</u> CTG-3'

2.3.4 Cloning of a CBESyn-GFP fusion construct into pICH11599

A CBESyn-GFPST fusion construct (Figure 4.13) was created by fusing the synthetic gene CBESyn via a free linker (FL) region to a gene coding for Strep-tagged Green Fluorescent Protein (GFP). The free linker was added to the 3' end of CBESyn by amplifying pUC57 CBESynHS with the forward primer 5'-ATCCATGGACGTTTCAGTTG-3' and the reverse primer 5'-CAGGTAACAGTCTCTTCCACTAGTGGAGCTGGAGCTGAATTCTA-3' (restriction sites NcoI, SpeI and EcoRI, respectively, are underlined; FL sequences are in italics). The amplified fragment was digested with NcoI and EcoRI and cloned into a version of pICH11599 which already contained a gene coding for the Strep-tagged GFP cloned into the EcoRI and BamHI sites.

2.3.5 Cloning of GOIs into the pEAQ-HT and pJL-TRBO vectors

CBESyn, PSA and CPS were cloned into pEAQ-HT (Sainsbury and Lomonosoff, 2008) and pJL-TRBO (Lindbo, 2007) transient expression vectors. For CBESyn, the inserts were amplified using PCR conditions detailed in Section 2.3.1 with pUC19-ApoCBE as the template. The vector pUC19-ApoCBE contains the CBESynHS gene with a tobacco calreticulin apoplast-targeting presequence located at the 5' end, cloned into pUC19 (please refer to Section 2.4.2 for its construction). The respective forward and reverse primers used to amplify the inserts are detailed in Table 2.5.

For PSA and CPS, the genes were amplified using pUC57-PSAHS and pUC57-CPSHS as the templates and the primers detailed in Table 2.5. The amplified fragments were then cloned into pUC19 (Yanisch-Perron *et al.*, 1985) using the restriction sites HindIII and EcoRI. The constructs were subsequently digested with AgeI and XhoI (for cloning into pEAQ-HT) or PacI and NotI (for cloning into pJL-TRBO). The digested fragments were gel-purified and cloned into pEAQ-HT or pJL-TRBO.

Table 2.5 Forward and reverse primers used for amplification of GOIs to be cloned into pEAQ-HT or pJL-TRBO. Start and stop codons are in bold; Restriction sites (AgeI, PacI, SacI, EcoRI) are underlined or in lower case; 6xHis- sequence in italics;

Name of Construct	Forward primer	Reverse primer
pEAQ-HT CBESyn	5'TATAagctt <u>ACCGGT</u> ATGG GCTACTC	5' <u>CATCACCA</u> <u>TCA</u> <u>TCACCACTGATAA</u> ctcgagCTCgcggccgcGAATTCATA-3' AACGAAAGGGCAAAC-3'
pJL-TRBO CBESyn	5'TATAagcttAATTAA CA ATGG GCTACT	5' <u>CATCACCA</u> <u>TCA</u> <u>TCACCACTGATAA</u> ctcgagCTCgcggccgcGAATTCATA-3' CAACGAAAGGGCAAAC-3'
pEAQ-HT PSA/pEAQ-HT CPS	5'TATAagctt <u>ACCGGT</u> ATGG GACGTT	5' <u>CATCACCA</u> <u>TCA</u> <u>TCACCACTGATAA</u> ctcgagCTCgcggccgcGAATTCATA-3' CAGTTG-3
pJL-TRBRO PSA/ pJL-TRBO CPS	5'TATAagcttAATTAA CA ATGG GACGTT	5' <u>CATCACCA</u> <u>TCA</u> <u>TCACCACTGATAA</u> ctcgagCTCgcggccgcGAATTCATA-3' CAGTT-3'

2.3.6 Swapping of framework regions between CLDSyn and CBESyn

pUC57-CLDSynHS and pUC57-CBESynHS which contain VHH genes with flanking 5' NcoI and 3' HindIII sites were digested with XmaI (internal site) and HindIII. The longer fragments were gel-purified. Both constructs were also used as templates to amplify 4 different inserts with either the forward primer CLD-FWC-for (5' GCTCCCGGGAAGGAACGAGAGGGTGTGCTGCA-3') or CBE-FWC-for (5'GCTCCCGGGCAAGAACGAGAGGCTGTGCTG-3') (SmaI = XmaI sites underlined) and the reverse primer FWC-rev (5'CATCAC CACTGATAAAAGCTTGCA-3') (HindIII site underlined) (detailed in Table 2.6). The inserts possess Framework Regions (FR) 2 and 3 of either CLDSyn, CBESyn or a combination of both. The digested fragments were ligated into the gel-purified pUC CLDSynHS or pUC CBESynHS vectors to create 6 chimeras which contain different permutations of FR1, FR2 and FR3 from CLDSyn and CBESyn (please refer to Section 5.2.5 for a detailed explanation of the chimeras).

Table 2.6 Template, forward and reverse primers used for amplification of the second and third Framework Regions (FR) of different chimeric genes. The fragments possess Framework Regions 2 and 3 of either CLDSyn, CBESyn or a combination of both.

Fragments	Framework regions		Template	Forward primer	Reverse primer
	FR2	FR3			
D2D3	CLDSyn	CLDSyn	pUC57 CLDSyn	CLD FWC for	FWC rev
D2B3	CLDSyn	CBESyn	pUC57 CBESyn	CLD FWC for	FWC rev
B2B3	CBESyn	CBESyn	pUC57 CBESyn	CBE FWC for	FWC rev
B2D3	CBESyn	CLDSyn	pUC57 CLDSyn	CBE FWC for	FWC rev

2.4 Cloning of GOIs into stable transformation vectors

2.4.1 Amplification of GOIs using DNA from leaves infiltrated with ICON Genetics transient expression vectors

N. benthamiana leaves were infiltrated with a combination of *Agrobacterium* strains harbouring the ICON Genetics vectors pICH10881, pICH8420 and either pICH11599-CLDSyn or pICH11599-CBESyn (please refer to Section 2.5.1 for detailed explanation of *Agrobacterium* transformation). DNA from 7 dpi leaves were homogenized in liquid nitrogen and extracted with 1 volume of lysis buffer (0.1M Tris-HCl, 50mM EDTA, 50mM NaCl, 2% SDS, 0.75% β -mercaptoethanol) followed by phenol/chloroform and RNase treatments. The extracted DNA (100 ng/ μ l) was boiled for 10 minutes, cooled on ice for 5 minutes and used as a template for amplification of a Short Fragment (SF) and a Long Fragment (LF) containing either CLDSyn or CBESyn (Please refer to Section 7.2.4 for detailed description of the components of SF and LF).

The template was amplified using Platinum[®] *Taq* DNA polymerase High Fidelity (Invitrogen, USA), 2 mM MgSO₄, 0.2 mM of each dNTP, 1 μ M forward and reverse primer and 10 ng template. The primers used were the forward primer 5'CATGGATCCAACTCGAGAAAC**ATG**GCTACTCAA-3' (BamHI site underlined, ATG in bold) and the reverse primer 5'CATCACCATCATCACC**ACTGATA**ACTCGAGCTCATA-3' (SacI site underlined; Stop codons in bold) for amplification of SF, while the forward primer 5'CATAGATCTGTCGATTGGTTGCAGCATTT-3' (BglII site underlined) and the reverse primer 5'ATCCTCCCTAACCCCGGGTAGGGGGAGCTCATA-3' (SacI site underlined) were used for amplification of LF. The PCR conditions included an initial denaturation step (2 minutes at 94°C), followed by 30 consecutive cycles comprising: denaturation (30 seconds at 94°C), annealing (30 seconds at 57°C) and elongation (1 minute at 68°C); and a final elongation (10 minutes at 68°C) step.

The PCR product was digested with BamHI and SacI for SF fragments or BglII and SacI for LF fragments. The fragments were cloned into pROK-219 (pUC19-based vector containing CaMV35S promoter and Nos terminator; Baulcombe *et al.*, 1986) digested with BamHI and SacI. For pROK-219 containing SF inserts, the plasmid

was digested with HindIII and EcoRI and the CaMV35S-SF-Nos fragment was cloned into the binary vector pBINPLUS/ARS (Belknap *et al.*, 2008). For pROK-219 containing LF inserts, the plasmid was digested separately with HindIII and KpnI as well as KpnI and EcoRI. A 1.3 kb [CaMV35S-5'LF-partial] fragment (from the HindIII/KpnI digest) and a 0.9 kb [3'partial LF-Nos] fragment (from the KpnI/EcoRI digest) was gel purified and cloned into pBINPLUS/ARS by triple ligation.

2.4.2 Amplification of GOIs using mRNA from leaves infiltrated with ICON vectors

N. benthamiana leaves were infiltrated with a combination of *Agrobacterium* strains harbouring the ICON Genetics vectors pICH10881, pICH8420 and pICH11599-CBESyn (please refer to Section 2.5.1 for detailed explanation of *Agrobacterium* transformation). Total RNA was isolated from 7dpi leaves by first homogenizing samples to a powder in liquid nitrogen and then dispersing the powder in 1 volume lysis buffer (0.1M Tris-HCl, 50mM EDTA, 50mM NaCl, 2% SDS, 0.75% β -mercaptoethanol). The homogenate was extracted twice with buffer-saturated phenol/chloroform and total RNA precipitated by addition of 0.5 volume 8 M LiCl followed by incubation at -20°C for 30 min. The recovered RNA pellet was washed with 75% ethanol, air-dried and re-suspended in 100 μ l sterile water at 0.1 μ g/ μ l. SuperScript™ II Reverse Transcriptase (Invitrogen, USA) was used to synthesize first strand cDNA according to the manufacturer's instructions and the cDNA was then amplified with the forward primer 5'CGGGATCCAAC**ATGG**CTACTCAACGAAGG-3' (BamHI site underlined) and the reverse primer 5'**CATCACC**ATCATCACC**ACTGAT**AA**ctcgagCTCgcgccgcGAATTC**A-3' (SacI and EcoRI sites, respectively, underlined; Stop codons in bold) using the conditions detailed in Section 2.3.1. The amplified products were digested with BamHI and SacI and cloned into pROK-219 or pROK-819 (pUC19-based vector with an *rbcS* promoter (the *rbcS* gene encodes the small subunit of Rubisco) and the Nos terminator). The resulting constructs were digested with HindIII and EcoRI and the released inserts cloned into the binary vector pBINPLUS/ARS (Belknap *et al.*, 2008).

2.5 Plant growth and transformation

2.5.1 Transient transformation

For the ICON Genetics vectors, agroinfiltration was performed as described in Marillonnet *et al.* (2004). *Agrobacterium tumefaciens* strain GV3101 transformed with individual constructs harbouring the GOI was grown overnight in Luria-Bertani (LB) broth, 0.02 mM acetosyringone, 0.01 mM MES and 100 µg/mL Rifampicin at 28°C. After the removal of the medium by centrifugation at 8,000 rpm, 10 min at 12°C, the pellet was resuspended in Infection Solution (IS) containing 0.1 mM acetosyringone, 0.01 mM MES, 0.01 mM MgCl₂ and 100 µg/mL Rifampicin. The final OD₆₀₀ of the bacterial suspensions were adjusted to 0.8 using IS, and equal volumes of the three constructs were mixed. Fully expanded leaves of 3.5-week-old *Nicotiana benthamiana* plants grown in a greenhouse were infiltrated on the abaxial side with 200 - 1000 µl bacterial suspension in IS using a 1 mL syringe (without a needle). The plants were then further grown in the greenhouse.

For CPMV and TRBO vectors, agroinfiltration was performed as described above with slight modifications. Firstly, CPMV vectors were transformed into *Agrobacterium tumefaciens* strain LBA4404 while TRBO vectors were transformed into *Agrobacterium tumefaciens* strain GV3101. *Agrobacterium* were grown overnight at 28°C in Luria-Bertani (LB) broth, 0.02 mM acetosyringone, 0.01 mM MES, 100 µg/mL Rifampicin with an additional 50 µg/mL Kanamycin as a selection agent for GV3101 or Streptomycin for LBA4404. After removal of the medium by centrifugation, the final OD₆₀₀ of the bacterial suspensions were adjusted to 1.2 for CPMV constructs and 1.0 for TRBO constructs using the Infection Solution. Fully expanded leaves of 3 - 3.5 week-old *N. benthamiana* plants grown in a greenhouse were infiltrated as described above.

2.5.2 Stable transformation

Stable transformation of tobacco (*N. tabacum* cv Samsun and *N. benthamiana*) leaf discs was performed as described in Horsch *et al.* (1985). Briefly, *Agrobacterium tumefaciens* strain GV3101 or LBA4404 transformed with binary vector constructs harbouring the GOI was grown for 24 hours at 28°C in 10 mL Luria-Bertani (LB) broth with 100 µg/mL antibiotic for *Agrobacterium* selection (Rifampicin for GV3101 or Streptomycin for LBA4404) and 50 µg/mL antibiotic for plasmid selection (Kanamycin for the pBIN2, pBIN8, CPMV and TRBO constructs). After the removal of the medium by centrifugation at 4,000 rpm for 10 min at 12°C, the pellet was resuspended in 10 mL liquid MS medium (Murashige and Skoog, 1962) supplemented with vitamins (Gamborg *et al.*, 1968). Leaf discs (1cm²) of 4 week-old *N. tabacum* cv Samsun or *N. benthamiana* (approximately 50 discs per construct) were immersed in the *Agrobacterium* suspension for 20 minutes. The leaf discs were then incubated abaxial side upwards for 2 days on Shoot Inducing Medium (SIM) consisting of MS medium with vitamins, 0.1 µg/mL α-Naphthaleneacetic acid (NAA) and 1 µg/mL Benzylaminopurine (BAP) in a tissue culture room at 23°C with 16/8 hours light/dark cycle. The leaf discs were then transferred to SIM with 500 µg/mL Carbenicillin to kill off the *Agrobacterium* and 100 µg/mL Kanamycin to select for transformed cells. Regenerated shoots were rooted on MS medium. Rooted plantlets were transferred to soil and allowed to grow in the greenhouse.

2.6 Detection of GOI mRNA in tobacco leaves

Total RNA was prepared from transformed *N. tabacum* cv Samsun or *N. benthamiana* leaves as described above (Section 2.4.2). Denatured total RNA samples were resolved in 1% agarose gels containing 2.2 M formaldehyde (Sambrook *et al.*, 1989), transferred onto a Biodyne® B Membrane (Pall Corporation, USA) by capillary transfer, and GOI-specific mRNAs were detected by hybridizing the blots with Digoxigenin (DIG)-labelled DNA probes (Yamaguchi *et al.*, 1994). Bound hybridization probes were detected using an alkaline phosphatase-conjugated anti-DIG antibody followed by addition of the CSPD® light-emitting substrate (Applied Biosystems, USA), which was visualized using the LAS-3000 imaging system (FujiFilm, Japan) or by exposure to X-ray film (FujiFilm, Japan).

2.7 Detection of protein expression in Tobacco leaves

Transformed *N. tabacum* cv Samsun or *N. benthamiana* leaves were homogenized with liquid nitrogen and the powder extracted with 3 volumes (for every 100 mg) of 50mM NaH₂PO₄ (pH8.0), 300 mM NaCl, 10 mM Imidazole and 1μM PMSF for His-tagged proteins or 100 mM Tris (pH8.0), 150 mM NaCl, 1 mM EDTA and 1 μM PMSF for Strep-tagged[®] proteins. Protein concentration was determined via the Bradford assay using known concentration of Bovine Serum Albumin (BSA) as standards. Approximately 10 μg of total soluble protein extract was denatured in Laemmli buffer and resolved by SDS-PAGE (5% stacking gel and 15% running gel). Proteins were directly visualized by Coomassie R-250 staining or by Western Blot. Estimation of protein concentration on SDS-PAGE gels was performed by comparing the relative density of sample protein bands to bands of known concentration in the Perfect Protein[™] Marker (Novagen, USA) sample, using the image analysis software MultiGauge (FujiFilm, Japan).

For Western blot analyses, protein samples were resolved by SDS-PAGE and transferred onto a nitrocellulose membrane. The membrane was blocked with PBS containing 3% skimmed milk powder. For detection of His-tagged proteins, the membrane was incubated overnight with a 1:5,000 dilution of mouse anti-His primary antibody (Sigma, USA), followed by a 2-hour incubation of 1:10,000 dilution of goat anti-mouse IgG-HRP conjugate. For detection of Strep-tagged proteins, the membrane was incubated for 1 hour with a 1:5,000 dilution of Strep-tactin-HRP conjugate (Sigma, USA). For detection of Green Fluorescent Protein (GFP), the membrane was incubated overnight with a 1:1,000 dilution of rabbit anti-GFP primary antibody (Sigma, USA), followed by a 1.5 hour incubation of 1:75,000 goat anti-rabbit IgG-HRP conjugate. Detection was performed using the ECL system (Pierce, USA) and visualized using LAS-3000 imaging system (FujiFilm, Japan) or by exposure on X-ray film (FujiFilm, Japan).

2.8 Protein purification

2.8.1 Purification of His-tagged proteins

Transformed leaves were harvested at 11 days post-infiltration (dpi) for CLDHS, CLDSynHS and 7dpi for CBESynHS and ABGHS constructs. They were homogenized in liquid nitrogen and the resulting powder extracted with 3 volumes (for every 100 mg) of extraction buffer (50 mM NaH_2PO_4 , pH 8.0, 300 mM NaCl, 10 mM Imidazole and 1 μM PMSF). The clarified crude extracts were purified using a Nickel-NTA (Qiagen, USA) column made from 50% slurry. The column was washed with 5 column volumes (CV) of wash buffer (similar to extraction buffer but without PMSF) and the bound protein eluted with Elution buffer (50 mM NaH_2PO_4 , pH8.0, 300 mM NaCl, 250 mM Imidazole). Excess imidazole was dialyzed out of the sample by ultrafiltration with a Nanosep[®] Centrifugal device (Pall, USA) which had a molecular weight cutoff of 3 kDa for the Camel VHHs and 10 kDa for the scFvs. The final solution which contained 50 mM NaH_2PO_4 , pH 8.0 and 300 mM NaCl was adjusted to contain 1 μM PMSF, 10% (v/v) glycerol and 0.01% (v/v) Thimerosal.

2.8.2 Purification of Strep-tagged proteins

Transformed leaves were harvested at 11 dpi for CL3ST, CLDST, and 7 dpi for 163ST and ABGST. They were homogenized in liquid nitrogen and extracted with 3 volumes (for every 100mg) of extraction buffer (100 mM Tris pH8.0, 150 mM NaCl, 1 mM EDTA and 1 μM PMSF). The crude extracts were purified using a Strep-Tactin[®] (IBA, Germany) column made from 50% suspension. The column was washed with 2 column volumes (CV) of wash buffer (similar to extraction buffer without PMSF), followed by 2 CVs of high salt wash buffer (100 mM Tris pH8.0, 300 mM NaCl and 1 mM EDTA) and 2 CVs of wash buffer. Bound protein was eluted with Elution buffer (100 mM Tris pH8.0, 150 mM NaCl, 1 mM EDTA and 2.5 mM D-desthiobiotin). Excess D-desthiobiotin was dialyzed out of the sample by ultrafiltration with a Nanosep[®] Centrifugal device (Pall, USA) which had a molecular weight cutoff of 3 kDa. The final solution, containing 100 mM Tris pH8.0, 150 mM NaCl and 1 mM EDTA was adjusted to include 1 μM PMSF, 10% (v/v) glycerol and 0.01% (v/v) Thimerosal.

2.9 Immunosorbent assays (ELISA)

2.9.1 Direct ELISA for His-tagged proteins

Antigens (β -galactosidase, HEWL, RNase and PSA; all from Sigma, USA), were coated onto ELISA plates at a concentration of 20 $\mu\text{g/mL}$ (for β -galactosidase), 5 $\mu\text{g/mL}$ (for HEWL), 20 mg/mL (for RNase) and 20 $\mu\text{g/mL}$ (for PSA) in 100 mM carbonate/bicarbonate buffer (pH 9.6) overnight at 4°C. The plates were then washed three times with sterile distilled water. The plates were blocked with PBS containing 5% skimmed milk powder or 1% ELISA grade BSA (Sigma, USA) at 25°C for 2 hours. After 3 washes with water, the antibody samples (as crude plant extracts or pure protein) were incubated at 25°C for 2 hours (for crude extract) or at 4°C overnight (for pure protein). All dilutions were made in blocking solution. The plates were washed three times with PBS containing 0.05% (v/v) Tween 20 (PBSt). A 1:5000 dilution of HRP-conjugated Ni-NTA (Qiagen, USA) was then added and the plates incubated overnight at 4°C. Following washing with PBSt solution, the bound HRP-conjugate was detected by addition of Turbo-TMB ELISA reagent (Pierce, USA). Plates were incubated at 25°C for 10-30 minutes and the reaction stopped with 2 M H_2SO_4 . The absorbance of the resulting yellow-coloured substrate was read at 450nm.

2.9.2 Direct ELISA assay for Strep-tagged proteins

ELISA plates were coated with antigen, washed, treated with blocking agent, and washed again with PBSt as described above. A 1:5000 dilution of HRP-conjugated Strep-tactin[®] (IBA, Germany) was added to the wells overnight. Bound antibodies were detected with Turbo-TMB ELISA (Pierce, USA) as described above.

2.9.3 Measurement of VHH affinity by competitive ELISA

Competitive ELISA as described by (Friguet *et al.*, 1985) was used to determine the affinity of the VHHs. ELISA plates were coated with various concentrations of antigen (please refer to Section 2.9.1 for concentrations used) in carbonate buffer (pH 9.6) overnight at 4°C. The plates were washed with water. VHHs were diluted in PBS (pH7.5 containing 1% BSA and 2 mM EDTA) to a concentration of 5×10^{-10} M. The samples were mixed with 1×10^{-12} M to 1×10^{-2} M antigen in PBS (pH7.5) and incubated in the wells overnight at 4°C. The plates were washed three times with PBS containing 0.05% (v/v) Tween 20 (PBSt). A 1:5000 dilution of HRP-conjugated Ni-NTA (Qiagen, USA) was then added and incubated overnight at 4°C. Bound antibodies were detected with the Turbo-TMB ELISA reagent (Pierce, USA) as described above.

2.9.4 Measurement of the heat stability of VHHs

Heat stability of VHHs was measured as described by Ladenson *et al.* (2006). Briefly, 1 ng/ml dilution of CLDHS or CBESynHS was prepared in PBS containing 0.05% Tween. The antibody solutions were incubated at temperatures ranging from 0°C, 25°C, 37°C, 40°C, 50°C, 70°C and 90°C for 20 minutes. They were then re-equilibrated to room temperature for 5 minutes before the residual antigen binding activities were measured by direct ELISA.

2.10 Bacterial expression of VHHs

2.10.1 Cloning of GOIs into the expression vector pET-22b

CLDSyn, CBESyn, RN5 and CRN were amplified using the forward primer 5'TATCCATGGCCCAGGTTTCAGTTGGTT-3' (NcoI site underlined) and the reverse primer 5'ACAGTCTCTTCCGGACTCGAGATA-3' (XhoI site underlined). The resulting PCR fragment was digested with NcoI and XhoI and was cloned into the multiple cloning site (MCS) of pET-22b(+) (Novagen, USA) using the corresponding restriction sites. In pET-22b(+), expression of the GOI is driven by a T7 promoter. The expressed protein contains an N-terminal pelB signal sequence which directs it to the periplasmic space and a C-terminal 6xHis tag. The vector was transformed into *E. coli* BL21(DE3) competent cells (Novagen, USA).

2.10.2 Bacterial growth and induction of protein expression

E. coli BL21(DE3) was grown at 37°C in Luria-Bertani (LB) medium containing 100 µg/mL Carbenicillin to an OD₆₀₀ of 0.6-1.0. Protein expression was induced by adding isopropyl-D-thiogalactopyranoside (IPTG) to a final concentration of 0.5 mM and analyzed by SDS-PAGE. 2mL samples were collected before the addition of IPTG and at 1, 2 and 3 hours after protein induction. The cells were harvested by centrifugation at 8,000rpm for 3 minutes. Whole cell extracts were prepared by boiling cells in 100µL Laemmli buffer for 10 minutes.

Chapter 3. Expression of human scFvs and Camel VHHs in *Nicotiana benthamiana*

3.1 Introduction

Transient expression, the expression of a gene of interest (GOI) over a short time span (Janssen and Gardner, 1990; Kapila *et al.*, 1997), is an increasingly important strategy for recombinant protein production in plant cells. Although recently there has been some success in using vectors based on Bean Yellow Dwarf Virus (BeYDV), which has a single stranded DNA genome, in the production of vaccine antigens and monoclonal antibodies (Huang *et al.*, 2009; Huang *et al.*, 2010; Chen *et al.*, 2011), the most efficient transient expression systems, to date, employ vectors based on sequences derived from plant RNA virus genomes.

The transient expression system developed by ICON Genetics uses *Agrobacterium* to deliver DNA-based copies of a 'deconstructed' tobamovirus genome containing a GOI, into plant cells (Marillonnet *et al.*, 2004; Gleba *et al.*, 2007). Tobamoviruses (e.g. tobacco mosaic virus (TMV), Crucifer-infecting TMV (Cr-TMV) and turnip vein clearing virus (TVCV)) are a group of closely related rod-shaped viruses that possess a single-stranded plus-sense RNA genome (for detailed description, see Creager *et al.*, 1999; Scholthof, 2004). The genome comprises three ORFs encoding four proteins: an RNA-dependent RNA polymerase (RdRP), a Methyltransferase/RNA helicase (MT-Hel), a movement protein (MP) and a coat protein (CP). The RdRP and the MT-Hel form the large and small subunit of the replicase complex, respectively. This directs RNA replication using positive strand RNA as the template.

The ICON expression system consists of a pair of separate 5' and 3' pro-vector modules flanked by left and right T-DNA border sequences. The GOI is cloned into the 3' provector which also contains the 3' nontranslated region (3' NTR) from the genome of Cr-TMV. The larger 5' pro-vector module contains approximately two-thirds of the genome of TVCV (as cDNA) including the regions coding for the viral RdRP and the MP gene, which have been placed under the transcriptional control of an *Arabidopsis Actin2* promoter. Different versions of the 5' pro-vector module are

available which encode an N-terminal targeting pre-sequence that directs accumulation of the GOI protein in the apoplast or the chloroplast; when targeting pre-sequences are absent, the encoded protein accumulates in the cytoplasm. When delivered into cells via *Agrobacterium*, the 5' and 3' provector modules are ligated together in the nucleus by site-specific recombination, mediated by the phage C31 integrase encoded on a third provector module. Initial transcription of the recombined (or reconstituted) viral genome by host RNA polymerase II gives rise to RNA transcripts, containing the ORF of the GOI that replicate in the cytoplasm (cytoplasmic RNA replicons). Because these RNA replicons lack the viral coat protein gene, which is essential for long-range systemic movement, and for virus particle formation, the replicons are non-infectious and do not spread efficiently between cells. However, the lack of spread is compensated for by the high efficiency with which *Agrobacterium*, when infiltrated into leaves at a suitable titre, delivers the vector modules to adjacent leaf cells (almost 100% efficiency).

Exceptionally high yields of recombinant proteins can be obtained using the ICON system, typically exceeding 30% tsp depending on the protein in question, in a short amount of time, and it can even be scaled up using an automated vacuum infiltration procedure (Gleba *et al.*, 2007). The ICON expression system has been successfully used to produce candidate vaccine antigens (Santi *et al.*, 2006; Dorokhov *et al.*, 2007; Huang *et al.*, 2008; Webster *et al.*, 2009; Kalthoff *et al.*, 2010), monoclonal antibodies (Lai *et al.*, 2010), various pharmaceutical proteins like human growth hormone (Gils *et al.*, 2005) and even multimeric proteins by the co-infection of two non-competing viral vectors (Giritch *et al.*, 2006). More recently, non-competing vectors have been used to produce full idiotype (Id)-containing IgGs of 20 non-Hodgkin's lymphoma patients at levels between 0.5 and 4.8 mg/g fresh weight in a short amount of time. Moreover, these IgGs were shown to confer protection on vaccinated mice in tumour challenge studies, compared with hybridoma-produced IgGs (Bendandi *et al.*, 2010).

In this chapter, we describe our initial investigations into the influence that gene-intrinsic factors (in particular the addition of different purification tags), and the subcellular compartment in which the recombinant protein accumulates, have on overall expression, yield, functionality, stability, and recovery of antibody fragments in plants. Two types of antibody fragment, scFv and VHH (or nanobody), were expressed in *Nicotiana benthamiana* using the ICON Genetics expression system.

3.2 Results

3.2.1 Cloning and expression of scFv and VHH constructs

Two related human single chain variable fragments (scFv) and two Camel single-domain variable fragments (VHH) were initially chosen for expression via the deconstructed TMV vector system described by Marrillonnet *et al.* (2004) in order to investigate the feasibility of expressing scFv and VHH at high levels in *N. benthamiana* leaves. The human scFvs recognize the antigen β -galactosidase. However, one of them was a mutant selected for efficient expression and functionality in the reducing cytoplasm of *E. coli* (163-R4; renamed ABG) while the other was the original wild-type clone (renamed 163) (Martineau *et al.*, 1998). The two Camel VHHs were cAb-Lys3 (renamed CL3) which recognizes hen egg-white lysozyme (HEWL) and cAb-TT2 (renamed TT2) which recognizes tetanus toxoid from *Clostridium tetani* (Arbabi-Ghahroudi *et al.*, 1997). These particular antibody fragment genes were chosen because the structure, properties and expression levels of the encoded proteins in *E. coli* were well documented.

Sequences encoding a 6xHis tag (HS), Haemagglutinin tag (HA) or Strep tag (ST) were added to the 3' end of the scFv and VHH open reading frame (ORF) by PCR. scFv and VHH transgenes without 3' tags were also made (please refer to Table 3.1 for a list of scFvs and VHHs that were initially expressed). The constructs were then cloned into the 3' provector pICH11599 (Marillonnet *et al.*, 2004) and transformed into *Agrobacterium*. Expression of a gene of interest (GOI) in this system (the ICON expression system) requires two additional *Agrobacterium* strains which are co-infiltrated with the strain containing the 3' provector: one contains a 5' provector module (i.e. either pICH10570 or pICH10570 or pICH10530); the second contains the phiC31 integrase gene.

When expressed *in planta*, the integrase protein joins the 3' provector module containing the GOI to a 5' module (which encodes TMV replication functions), via site-specific recombination (between the provector attB/attP sites) (Figure 3.3b). This results in reassembly of the TMV genome at the DNA level, which following transcription then behaves as a runaway RNA replicon and, in most cases, results in high-level expression of the GOI-encoded protein.

The expression system provides a choice of 3 different 5' pro-vector modules as a partner for the single 3' provector module (which contains the GOI) (Figure 2.1). Each of the 5' provectors encodes TMV replication functions. However, they differ with respect to the presence/absence of cleavable protein targeting signals that will ultimately be fused to the N-terminus of the protein encoded by the GOI (see Table 2.1). Importantly, in the case of proteins targeted to the chloroplast or the apoplast, the mature proteins will have significantly different N-termini as illustrated in Figure 3.1.

3.2.2 Accumulation of scFvs and VHHs in different cellular compartments

For the scFvs, it was found that relatively high amounts of the ABG protein accumulated in the cytoplasm regardless of the presence or absence of a C-terminal tag, while lower amounts were detected when the protein was targeted into chloroplasts. Moreover, the addition of a 6xHis or a Strep tag did not affect protein accumulation or change the ABG expression profile (Table 3.1 and Figure 3.2; ABGHS and ABGST. Data for ABG without a tag or with a HA tag are not shown). In contrast, the pattern of accumulation for 163, the original wild-type version of the scFv, was strikingly different from that of ABG. High-level accumulation of Strep-tagged 163 was observed only when it was targeted to the chloroplast; accumulation in the cytoplasm was much lower (Figure 3.2; 163ST). No accumulation of a 6xHis-tagged version of 163 was observed. Furthermore for both ABG and 163, no protein was detected when they were targeted to the apoplast.

For the VHHs, it was found that the Strep-tagged version of CL3 (CL3ST) accumulated highly when targeted to the apoplast, but no protein was detected by Western blot in the cytoplasm or when it was targeted to the chloroplast (Figure 3.2; CL3ST). A similar pattern of accumulation was observed for HA-tagged CL3 (data not shown). However, when a C-terminal His tag was added to CL3 (CL3HS), no protein was detected in any cell compartment (data not shown). In addition, no expression of TT2 was detected in any cell compartment irrespective of the presence or absence of the various C-terminal tags.

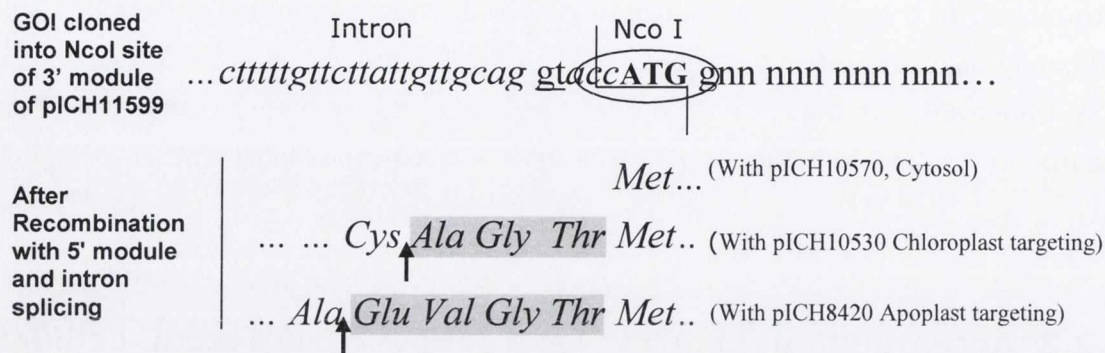


Figure 3.1 Predicted N-terminal amino acid sequences of proteins encoded by a GOI when expressed in the cytoplasm or targeted to different cellular compartments. In this example, the GOI was cloned into the 3' module vector pICH11599. Top line: the nucleotide sequence of the region encompassing the intron, the splice acceptor site (gt), the Nco I cloning site (encircled) and the reading frame of the GOI (in relation to Nco I). The bottom lines show the expected N-terminal amino acid sequence of the protein encoded by the GOI when expressed in the cytoplasm (using the 5' module of the vector pICH10570), or targeted to the chloroplast (with pICH10530), or the apoplast (with pICH8420). Vertical arrows indicate the expected pre-protein cleavage site; the grey boxes highlight the residual amino acids of the signal sequence at the N-terminus of the GOI protein.

Table 3.1 List of the scFvs (163 and ABG) and Camel VHHs (CL3 and TT2) initially investigated, with and without different purification tags.

	Antibodies			
	163	ABG	CL3	TT2
No tag		+	+	-
6x His tag	-	+	-	-
Haemagglutinin (HA) tag		+	+	-
Strep tag	+	+	+	-

+ represents successful attempts at detecting soluble protein by Western Blot.
- represents unsuccessful attempts at obtaining soluble protein.

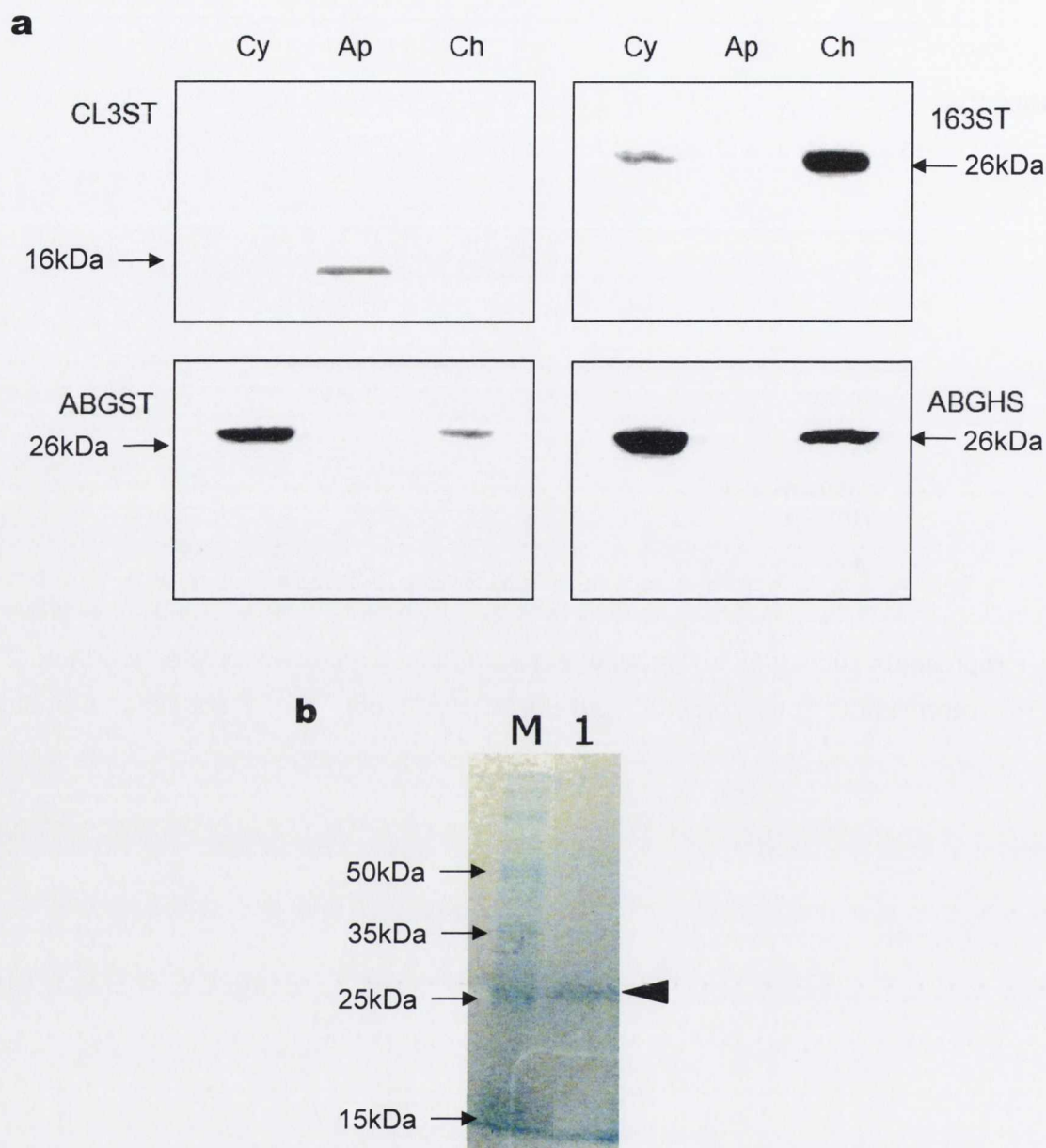


Figure 3.2. (a) The accumulation of CL3ST, 163ST, ABGST and ABGHS in the cytoplasm (Cy), apoplast (Ap) and Chloroplast (Ch) at 7 days post-infiltration (dpi) as detected by Western Blot. Approximately 30 μ g of total soluble protein was loaded in each lane for 163ST, ABGST and CL3ST, while 50 μ g was loaded for ABGHS. (b) PAGE gel showing ABG (no tag; lane 1) purified using ammonium sulphate precipitation. M, Perfect Protein™ Markers. The black arrow head indicates the scFv (approx. 26kDa). PAGE gel and western blots is representative of the three experiments conducted.

3.2.3 Northern blot analyses of mRNA accumulation

Our failure to detect the 6xHis-tagged version of scFv 163 and either the original or tagged versions of the VHH TT2 in any cell compartment by SDS-PAGE and Western blot analysis prompted an investigation of their expression at the mRNA level. Total RNA was extracted from 7 dpi leaves of plants infiltrated with the various constructs targeting the GOI protein to different cellular compartments, blotted and probed with a gene-specific DIG-labelled probe. The Northern blot showed that mRNA for all constructs containing the GOI accumulated to a high level (please refer to Figure 3.3a for TT2ST and 163HS mRNA expression). The upper band of approximately 6 kb corresponds to the mature primary transcript (i.e. after intron splicing) produced by the recombined 5' and 3' provector modules (Figure 3.3b). The lower bands of about 3kb and 1.2kb correspond to the mRNAs produced by RdRP-mediated transcription from the subgenomic promoters located near the 3' end of the RdRP gene and the 3' end of the 30 kDa Movement Protein (MP), respectively (Figure 3.3c). The molecular weight differences between the three mRNAs for a given GOI reflect the presence/absence of additional sequences encoding subcellular targeting signals.

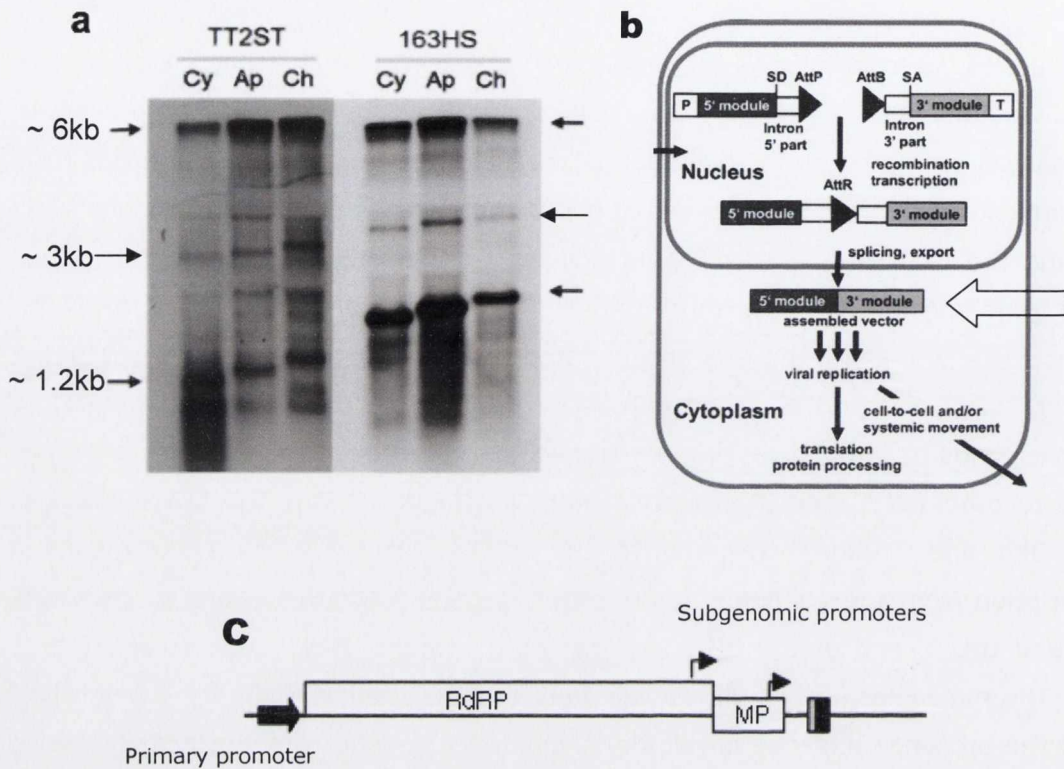


Figure 3.3 (a) Detection of mRNA transcripts in plants infiltrated with the TT2ST or 163HS constructs which encode either a cytoplasmically-expressed antibody fragment (Cy), or variants targeted to the apoplast (Ap) or the chloroplast (Ch). Total RNA was isolated from 7dpi infiltrated plants and analyzed by Northern blot using TT2ST or 163HS gene-specific probes. Approximately 10 μ g of total RNA was loaded onto each lane. Arrows correspond to mRNA containing the primary transcript (top arrow) and RdRP-transcribed mRNA driven by different subgenomic promoters (b) Diagram depicting the ICON expression system showing the assembly and processing of the 5' and 3' provector modules *in planta*. A white block arrow indicates the mature spliced primary transcript. SD, splice donor site; SA, splice acceptor site (Adapted from Marillonnet *et al.*, 2004). (c) Diagram depicting the position of subgenomic promoters (indicated by thin stemmed arrows) in the TMV expression vector (adapted from Lindbo, 2007).

3.2.4 Design of the CL3 variant CLD

When the Nco I site in the expression vector pICH11599 is used as the insertion site for a GOI, and the ATG within Nco I (ccATGg) is used as the translation initiation codon, the identity of the second amino acid is constrained by the fact that the second codon must begin with G. This limits the choice to either Alanine (A), Aspartic acid (D), Glutamic acid (E), Valine (V) or Glycine (G). In the case of CL3, on the basis that the N-terminal first four amino acids of most Camel VHHs correspond to Q (Glutamine), V, Q, L (Leucine) (please refer to Saerens *et al.*, 2005 for amino acid comparison of several Camel VHH), the codon GCC (encoding Alanine) was chosen for the constrained second position, thus giving an N-terminal sequence (after Met): A, Q, V, Q, L.

However, we also designed a CL3 variant, CLD, in which the N-terminal first 4 amino acids (after Met) were D, V, Q, L to better match the N-terminus of the original cAb-Lys3 clone (Arbabi-Ghahroudi *et al.*, 1997). Surprisingly, in contrast to CL3-HS which did not accumulate in any cellular compartment, CLD-HS accumulated at a low level when targeted to the apoplast (please refer to Figure 4.4a for CLDHS).

We therefore used CLD (specifically CLDHS) in subsequent detailed studies to investigate the effects of codon optimization as well as amino acid changes in the framework regions, on the yield of VHHs *in planta* (Chapter 4).

3.2.5 Purification and characterization of scFvs and VHHs

Leaves expressing chloroplast-targeted 163ST or cytoplasmically-expressed ABGST and ABGHS were harvested at 7 dpi and the expressed antibodies affinity-purified using either a Strep-Tactin column (for the Strep-tagged proteins) or a Nickel-NTA column (for the His-tagged proteins) (Figure 3.4). Approximately 0.3 mg protein/gram fresh leaf tissue was obtained for both the Strep-tagged and His-tagged ABG (Table 3.2). However, the yield was much lower for 163ST (about 0.06 mg/g fresh weight).

A yield of 0.03 mg/g fresh weight was obtained for apoplast-targeted CL3ST harvested at 11dpi. An extra band of slightly higher molecular weight was also visible on the PAGE gel (Figure 3.4; CL3ST crude extract) which may correspond to protein that is incorrectly folded. Another possibility is the incomplete cleavage of the apoplast targeting presequence as reported by Santi *et al.* (2006). Similar bands were also observed in the PAGE gels of both CLDHS and CLDST (please refer to Chapter 4).

The specificities of the purified antibodies were determined by direct ELISA using HRP-conjugated Ni-NTA to detect the His-tags or HRP-conjugated Strep-Tactin to detect the Strep tags. It was found that all the antibodies bound with high specificity to their respective antigens (Figure 3.5). Thus ABGHS, ABGST and 163ST bound to β -galactosidase and not to HEWL, while CL3ST bound to HEWL, and not β -galactosidase.



Figure 3.4 PAGE gel showing the crude extract (C) and eluted fraction (E) of 163ST, ABGST, CL3ST and ABGHS. Approximately 3 g infiltrated leaf material was harvested per construct and 30 μ L of each fraction was loaded per lane. M, Perfect ProteinTM Markers. The white, black and grey arrow heads indicate, respectively, the large RUBISCO subunit (54kDa), the scFvs (approx. 26kDa) and the VHH (approx. 16kDa). PAGE gel is representative of the three experiments conducted.

Table 3.2 Table showing the compartment with the highest accumulation and the mean yield after purification ($\pm$ SD; as determined by Bradford Assay) for ABGHS, 163ST, ABGST, CL3ST and CLDST. Data collected from three leaf samples from three independent experiments (total sample size=9).

Antibodies	Compartment with highest accumulation	Yield after purification [†] (mg/g fresh weight)
ABGHS	Cytoplasm	0.28 $\pm$ 0.05
163ST	Chloroplast	0.02 $\pm$ 0.01
ABGST	Cytoplasm	0.20 $\pm$ 0.06
CL3ST	Apoplast	0.03 $\pm$ 0.02
CLDST	Apoplast	0.05 $\pm$ 0.03

[†]Protein concentration determined using Bradford assay

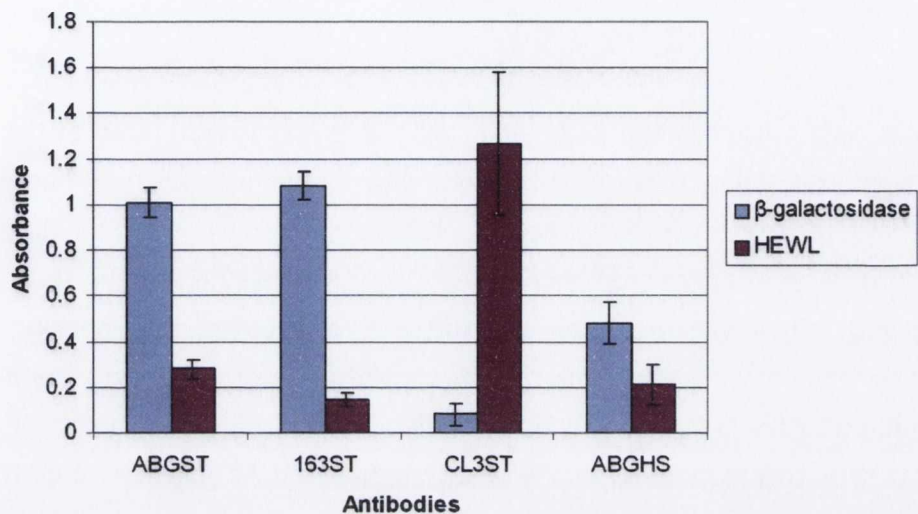


Figure 3.5 Antigen recognition specificity of the scFvs and Camel VHHs as determined by ELISA. Each anti- β -galactosidase scFv (ABGST, 163ST and ABGHS) recognized only β -galactosidase (■) while the anti-HEWL VHH (CLDST) recognized only HEWL (■). Each data point represents the mean of three experiments. Bars indicate 2 standard deviations.

3.2.6 Expression of the human anti- β -galactosidase scFv ABGST over time at the protein and mRNA level

Accumulation of ABGST in the cytoplasm over time was investigated at the protein and mRNA level. Leaves were infiltrated with the construct and harvested at different time points. For investigation of mRNA expression levels over time, leaves from 0, 3, 4, 5, 7, 9 and 11 dpi were used. Total RNA was extracted from the leaf material and mRNA was detected by Northern blot using gene-specific probes. Three main bands were detected on the blot: a band of approximately 6 kb corresponding to the provector primary transcript, and two lower bands of about 3 kb and 1.2 kb corresponding to the subgenomic mRNAs (please refer to Figure 3.3c). Detectable amounts of ABGST mRNA were present at 3 dpi which peaked at 4-5 dpi (Figure 3.6). Then, the mRNA level as well as the quality of the total RNA obtained, decreased with time as the leaf turned progressively more necrotic.

For investigation of protein expression, leaves corresponding to 0, 2, 5, 7, 9, 11, 13 and 15 dpi were used. Total soluble protein was extracted from the leaf material and was fractionated by PAGE. Detectable amounts of the ABGST protein can be seen on Coomassie blue-stained gels as early as 5 dpi (Figure 3.7). It continued to accumulate until 7 dpi where the levels stayed the same (or increased slightly, if the decrease in RUBISCO and other plant proteins is taken into account) up to 15 dpi. Interestingly, the level of ABGST in the plant stayed constant even though the leaves turned more and more necrotic as indicated by the decreasing amounts of RUBISCO large subunit (53 kDa). This suggests that either the plant redirected most of its translational machinery to producing ABGST, or that ABGST can withstand proteolysis better than other plant proteins.



Figure 3.6 The accumulation of mRNA transcripts in plants infiltrated with a construct directing ABGST expression in the cytoplasm at 0, 3, 4, 5, 7, 9, and 11 dpi compared with non-infiltrated leaves (WT). Approximately 10 μ g of total RNA per lane was loaded. The indicated upper band of approximately 6kb corresponds to the provector primary transcript. The two lower bands of about 3 kb and 1.2 kb correspond to the subgenomic mRNAs. Ethidium bromide stained gels of the total RNA were used as loading controls (upper panel).

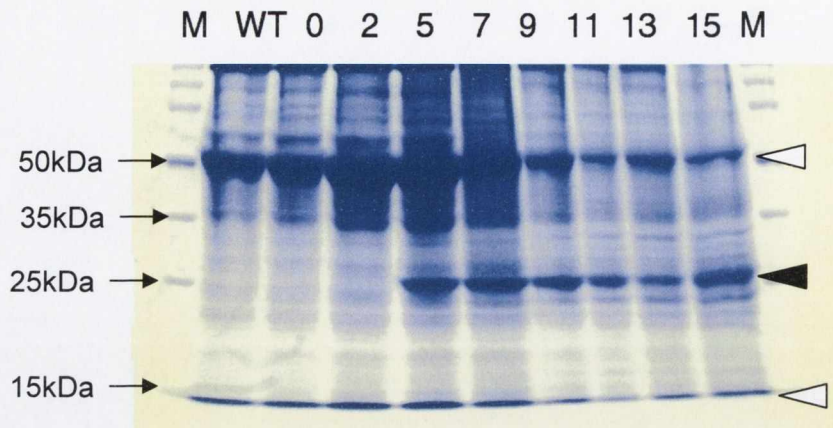


Figure 3.7 Coomassie blue-stained gel showing accumulation of ABGST in the cytoplasm at 0, 2, 5, 7, 9, 11, 13, and 15 dpi compared with uninfiltrated leaves (WT). Approximately 30 μ g of total soluble protein was loaded in each lane along with Perfect ProteinTM Markers (M). The white and black arrow heads indicate, respectively, the large and small RUBISCO subunit (54 kDa and 14kDa, respectively) and ABGST (approx. 26 kDa). PAGE gel is representative of the three experiments conducted.

3.3 Discussion

3.3.1 Comparison of yields obtained by the ICON system to alternative expression systems in plants and other production platforms

In plants, scFvs have been expressed in different organs such as leaves (Ram *et al.*, 2002; Vaquero *et al.*, 2002), tubers (Artsaenko *et al.*, 1998) and seed endosperm (Van Droogenbroeck *et al.*, 2007; Rademacher *et al.*, 2008). In general, the yields of scFvs expressed to date have been disappointingly low. For example, Vaquero *et al.* (2002) reported 5 mg/kg fresh weight for an anti-human carcinoembryonic antigen scFv in *N. tabacum*, while Ram *et al.* (2002) reported 0.031% TSP for an anti-Hepatitis B surface antigen scFv in *N. tabacum*. Remarkably, however, Van Droogenbroeck *et al.* (2007) reported yields of up to 36% TSP for an scFv-Fc fusion protein in *Arabidopsis* seeds, the highest reported to date.

The expression of VHHs in plants is relatively new. So far, a VHH has been expressed to an estimated level of 0.03% TSP in potato chloroplasts (Jobling *et al.*, 2003). Stably-transformed tobacco plants expressing a VHH against cancer-associated mucin (MUC1) has been reported, but with yields ranging from about 20 to 100 µg/g leaf tissue (Ismaili *et al.*, 2007). More recently, Winichayakul *et al.* (2009) expressed anti-tumour necrosis factor α (TNF-α) VHHs either as monomers or as heterodimers in *Arabidopsis thaliana*. Among the various subcellular compartments investigated (cytoplasm, ER, protein storage vacuole and chloroplast), it was found that accumulation was the highest in plants transformed with tandem VHH constructs directing protein accumulation in the ER, but with a yield of 0.1%-0.7% TSP. By using the ICON system, which has previously been used to produce proteins with yields from 0.5 to 5 mg/g fresh weight (Gleba *et al.*, 2007), we obtained up to 0.3 mg/g fresh weight for ABGHS and ABGST, and approximately 0.1 mg/g fresh weight for 163ST. For the VHHs, we initially obtained a yield of up to 0.05 mg/g fresh weight for the Strep-tagged anti-HEWL VHH. Further investigations, described in Chapter 4, revealed that this initial yield could be substantially improved upon.

Apart from plants, scFvs have been produced in other expression platforms such as *E. coli* (Martineau *et al.*, 1998; Müller-Schiffmann *et al.*, 2009), *Saccharomyces*

cerevisiae (Shusta *et al.*, 1998; Miller *et al.*, 2005), *Pichia pastoris* (Shi *et al.*, 2003; Gurkhan *et al.*, 2004), insect cells (Kretzschmar *et al.*, 1996; Johansson *et al.*, 2007), mammalian cells (Cardinale *et al.*, 1998; Braren *et al.*, 2007), and cell free systems (Hoffmann *et al.*, 2004; Galeffi *et al.*, 2006). Miller *et al.* (2005) compared *E. coli*, *S. cerevisiae* and *P. pastoris* for the production of several human scFv and found that *E. coli* was the most efficient system giving the most consistent yields. For example, Yang *et al.* (2009) reported a yield of 120-500 µg/L for an anti-canola stem rot fungus scFv. Martineau *et al.*, (1998) reported production levels of 0.5 g/L for the production of ABG (i.e. the AMEF mutant 163-R4) in *E. coli*. Camel VHHS have also previously been expressed in many platforms including *E. coli* (Arbabi-Ghahroudi *et al.*, 1997; Rahbarizadeh *et al.*, 2005), yeast (Frenken *et al.*, 2000; Rahbarizadeh *et al.*, 2006), and filamentous fungi (Joosten *et al.*, 2005). When comparing the ICON system with other platforms, it was found that the yields were higher than in *E. coli*, the best VHH production platform (Saerens *et al.*, 2005 had obtained yields of 5 mg/L for CLDHS and 4.5 mg/L for CBESynHS).

Comparatively, more time was required to obtain the recombinant protein from the Agro-infiltrated plants (7-11 days for the ICON system) compared with 2-3 days required in the case of *E. coli*. However, plants have the advantage of possessing a eukaryotic post-translational modification system which is similar but not identical to that found in animal cells. However, *N. benthamiana* plants with human-like N-glycosylation patterns have recently been produced using antisense technology to downregulate expression of key enzymes of the glycosylation pathway (Strasser *et al.*, 2008) and this human-like glycosylation pattern can also be obtained using transient expression vectors (Vézina *et al.*, 2009). Moreover, proteins that are highly expressed in *E. coli* often precipitate or aggregate in insoluble inclusion bodies which require refolding (Makrides, 1996; Misawa and Kumagai, 1999). Plant-based platforms have other advantages over the *E. coli* and other production platforms. For example, plants are considerably cheaper and less difficult to maintain than insect cells or mammalian cell platforms because they do not require the use of bioreactors or complex growth media. It is, therefore, possible to rapidly scale up production. In the case of the ICON platform, existing technologies can be used to automate the process (Gleba *et al.*, 2007). There is also a lower risk of contamination with human pathogens since plants are not their native hosts, and also the possibility of delivering a therapeutic antibody directly, without purification, by simply consuming raw plant material expressing the antibody (Twyman *et al.*, 2003; Bock, 2007).

3.3.2 The effect of protein targeting to various cellular compartments on yield

The ICON system gives a choice of three alternative 5' provector modules, with different targeting presequences, that direct the accumulation of expressed proteins in different subcellular compartments. Different targeting presequences have been found to affect protein stability and, subsequently, overall protein accumulation in plants (Schillberg *et al.*, 1999; Ram *et al.*, 2002). This is due to differences in the microenvironments of the different subcellular compartments. Compartments differ not only in respect of their redox status (the cytoplasm and chloroplast are reducing environments, while the endoplasmic reticulum (ER) and apoplast are oxidizing), but also in relation to the spectrum of resident proteolytic activities. In addition, recombinant proteins that have been targeted through the secretory system (en route to the apoplast), or retained in the ER, may undergo glycosylation, which can significantly affect (negatively or positively) protein functionality and stability (Vierstra, 1996; Gomord *et al.*, 2004).

Of the two anti β -galactosidase scFvs investigated, 163 showed the highest level of accumulation in chloroplasts, while ABG accumulated most highly in the cytoplasm. The existing literature indicates that the secretory pathway is particularly suitable for the folding and assembly of scFvs and other antibody formats both in stably- and transiently-transformed plants (Ram *et al.*, 2002; Sainsbury *et al.*, 2008). The particular suitability of the secretory system is largely due to the oxidizing environment of the ER which facilitates inter- and intramolecular disulphide bond formation and hence correct folding; the comparatively lower level of proteolytic activity; and the presence of molecular chaperones which are critical for correct folding, assembly and post-translational processing of antibodies (Ma *et al.*, 2003). However, there are some instances where an scFv has been successfully expressed in chloroplasts (Daniell *et al.*, 2005) or expressed in the cytoplasm (Schouten *et al.*, 1997; Zimmermann *et al.*, 1998; Schouten *et al.*, 2002). It was, therefore, not surprising that ABG, which was selected specifically for its ability to fold correctly in the reducing cytoplasm of *E. coli* (Martineau *et al.*, 1998), accumulated at a high level in the plant cell cytoplasm; whereas 163, the cysteine-containing original from which it was derived, did not.

The apoplast was found to be the best place among the three compartments investigated for high-level accumulation of the anti-HEWL VHH CL3. Previous

reports have shown that when expressed via the ICON system, human growth hormone (Gils *et al.*, 2005) and *Yersinia pestis* antigens (Santi *et al.*, 2006) also accumulate highly in the apoplast compared with other subcellular compartments. A VHH has also been targeted to the chloroplast for the immunomodulation of starch content (Jobling *et al.*, 2003), and expressed in the cytoplasm (Ismaili *et al.*, 2007) and the ER (Winichayakul *et al.*, 2009), although at a very low (barely detectable) level.

3.3.3 The effect of C-terminal purification tags on yield

The various scFvs and VHHs were expressed with different purification tags and it was found that for certain proteins, such as 163 and CL3, the addition of a tag affected protein accumulation while for others, such as ABG, it did not. On the other hand, no soluble protein was obtained when TT2 was expressed either without a tag or with a 6xHis, Hemagglutinin (HA) or Strep tag. This was in contrast to the results of Arbabi-Ghahroudi *et al.* (1997) who obtained soluble TT2 when the protein was targeted to the periplasmic space in *E. coli*. We speculate that the inability to obtain soluble protein in the case of TT2, 163HS and CL3HS might be attributable to post-translational processing and/or degradation since Northern blots showed significant levels of TT2 mRNA in infiltrated leaves.

Furthermore, when we compared the expression levels of the His-tagged and the Strep-tagged versions of the VHHs, the differential influence of the tags was evident, especially between CLDHS and CLDST. Menassa *et al.* (2003) reported a protective effect of the C-terminal His-tag on human interleukin-10 produced in plants. Also, the position of the His-tag can affect the binding properties of an scFv (Goel *et al.*, 2000), and one might speculate that the position of a tag can also affect the yield. This might explain the difference in yield, both between the His-tagged and the Strep-tagged scFvs and VHHs.

Chapter 4. Investigations of codon optimization and framework changes on VHH accumulation in plants

4.1 Introduction

Hamers-Casterman *et al.* (1993) identified a unique class of antibodies in the serum of members of the *Camelidae* family, the heavy chain-only antibodies (HCAbs). These antibodies lacked the light chain as well as the first constant heavy domain (CH1) normally found in conventional antibodies. The minimal antigen binding sites of the HCAb consists of a single domain variable fragment (VHH), also known as a nanobody. Nanobodies have a molecular weight of 14-15 kDa and a structure that is highly similar to the human VH-III domain (Muyldermans *et al.*, 1994; Nguyen *et al.*, 2000). Like VH, they possess four Framework Regions (FRs) that form a β -barrel core or scaffold as well as three Complementarity Determining Regions (CDRs) which are involved in antigen recognition and binding.

Since their discovery, nanobodies have been investigated extensively for use in various biotechnological and immunotherapeutic applications (as elaborated in Section 1.6). It was hypothesized that the range of potential applications might be extended if VHHs directed against a wide variety of antigens contained identical, highly stable scaffold (framework) regions. Such a 'universal' scaffold might, for example, promote their use as probes for biosensors and microarrays because it would facilitate standardization of probe-labelling protocols, immobilization of VHHs on microarray chips and other applications. A 'universal' scaffold might also enable the construction of a synthetic library made purely by randomization of the codons encoding the CDRs, which could then be used to retrieve VHHs against a wide variety of antigens. It might also facilitate construction of more stable VHH derivatives against a given antigen by means of a CDR grafting approach.

For these reasons, attempts have been made to identify a 'universal' framework that can reliably yield stable, high affinity VHHs that are capable of being expressed at a high level. This strategy proved successful with conventional VHs (Gonzales *et al.*, 2005; Hwang *et al.*, 2005), but since the CDRs of VHH do not obey the canonical loop structures, and there is some ambiguity regarding the key elements

that preserve loop architecture (they might, for example, also be located in the framework regions which were not included in the grafting experiments), such an approach has generated mixed results. Saerens *et al.* (2005) grafted the CDRs of several different VHH that bound with high affinity to their antigens onto a candidate universal, stable scaffold. The CDR donor VHHs included anti-HEWL, anti-human lysozyme, anti- β -lactamase, anti-prostate specific antigen and several others. The study found that the framework of cAbBCII10, an anti β -lactamase VHH (Conrath *et al.*, 2001), was potentially useful as a universal acceptor of CDRs grafted from different VHHs. cAbBCII10 is highly stable, has good expression levels in *E. coli* and is able to function in the absence of the disulphide bond found in the framework region of most VHHs which is thought to be a major cause of variability in intracellular expression levels. The grafting of CDR loops from VHHs of the same subclass (VHH2) into cAbBCII10 resulted in functional chimeras with increased thermodynamic stability, while grafting CDR loops from VHHs of a different subclass reduced the antigen-binding specificity and affinity.

In Chapter 3, it was shown that two anti β -galactosidase scFvs, differing only in the presence or absence of intramolecular disulphide bonds, accumulated in different cellular compartments and had strikingly different yields. This prompted us to ask whether changing the framework of a VHH (and ultimately its tertiary structure) might also have a significant impact on overall expression levels and yield in plants. To this end, transgenes encoding 6xHis and Strep-tagged derivatives of the anti-HEWL nanobody cAb-Lys3, in its original unmodified form (Arbabi-Ghahroudi *et al.*, 1997), and a variant in which the CDRs of cAb-Lys3 were grafted onto the 'universal' cAbBCII10 framework (Saerens *et al.*, 2005) were constructed and transiently expressed in *N. benthamiana* using the ICON system.

We also wanted to investigate whether codon optimization might improve VHH expression in plants. Modification of the nucleotide sequence of a GOI so that it more closely matches the codon usage pattern of the host organism, and removal of potentially disruptive *cis*-elements, has been shown to increase the yield of recombinant proteins in plants (Kirby and Kavanagh, 2002; Geyer *et al.*, 2007; Li *et al.*, 2007) and other organisms (Yadava and Ockenhouse, 2003; Burgess-Brown *et al.*, 2008). Therefore, plant-codon optimized versions of the original cAb-Lys3 and its framework-changed derivative were constructed.

4.2 Results

4.2.1 Design of modified anti-HEWL VHH constructs

Three different versions of Camel VHHs recognizing HEWL were constructed to investigate the effects of codon optimization and the influence of different framework regions on VHH production in plants. The three versions were: (i) CLD, a derivative of CL3 (i.e. original camel cAb-Lys3; Arbabi-Ghahroudi *et al.*, 1997) which contains aspartic acid (D) instead of glutamine (Q) at position 2 (ii) CLDSyn, a codon-optimized version of CLD; and (iii) CBESyn, a codon-optimized version of cAbBCII10-L-L-L, a VHH with the complementarity determining regions (CDRs) of CL3, but the framework regions of cAbBCII10, an anti- β -lactamase VHH proposed as a 'universal' framework for VHH engineering (Saerens *et al.*, 2005). Each VHH was expressed with a C-terminal 6x His-tag (HS).

We hypothesized that if the plant codon-optimized versions of the camel VHHs produced higher antibody yields *in planta* than the original versions, then the increase in yield might be due to codon-optimization, which has been shown, in certain cases, to increase recombinant protein production in plants (Geyer *et al.*, 2007; Li *et al.*, 2007). However, if only the version with an altered framework showed a yield increase, this would indicate that the framework regions are more important determinants of VHH yield in plant expression systems.

The codon optimization of CLD and CBESyn for plant expression was performed as described in Section 2.2.1. The aim was to adjust the pattern of codon usage in the original CLD to that more closely matching the codon usage frequencies of *N. benthamiana*, while avoiding undesirable sequences such as rare codons, direct or indirect repeats, potential hairpin structures, polyadenylation signals (Mogen *et al.*, 1992; Li and Hunt, 1995), destabilizing sequences, and 5' intron splice signals (Brown *et al.*, 2002; Simpson *et al.*, 2004). Priority was given to the removal of rare unwanted sequences while maintaining codon diversity since Li *et al.* (2007) had found that constructs with diverse codons usage (but no rare codons) expressed better than constructs with amino acids coded only with the most frequently used codon. Approximately 23% of the nucleotide sequence was altered in CLDSyn compared to the original CLD (Figure 4.1) to give a more even GC distribution and the total GC content was slightly reduced from 56.81% to 48.72% (Figure 4.2). The sequence of CBESyn on the other hand was similar to CLDSyn

except for the 10 amino acid (30 nucleotides) differences in total found in the first 3 framework regions. The 4th framework region was entirely conserved (Figure 4.3).

CLDHS	ATGGACGTGCAGCTGCAGGCGTCTGGAGGAGGCTCGGTGCAGGCTGGAGGGTCT
CLDSynHS	ATGGACGTTTCAGTTGCAAGCATCTGGTGGAGGAAGTGTGCAAGCTGGAGGGAGC
	M D V Q L Q A S G G G S V Q A G G S
CLDHS	CTGAGACTCTCCTGTGCAGCCTCTGGATACACCATCGGTCCCTACTGTATGGGG
CLDSynHS	TTAAGGCTCTCATGCGCTGCATCAGGCTACACAATTGGTCCCTATTGCATGGGC
	L R L S C A A S G Y T I G P Y C M G
CLDHS	TGGTTCGCGCAGGCCCCAGGGAAGGAGCGTGAGGGGGTGCAGCAATTAATATG
CLDSynHS	TGGTTTCGTCAAGCTCCCGGAAGGAACGAGAGGGTGTGCTGCAATCAATATG
	W F R Q A P G K E R E G V A A I N M
CLDHS	GGTGGTGGTATCACCTACTACGCCAACTCCGTGAAGGGCCGATTACCATCTCC
CLDSynHS	GGAGGTGGTATCACGTACTATGCGGATTCTGTCAAAGGGAGATTACCAATTCT
	G G G I T Y Y A D S V K G R F T I S
CLDHS	CAAGACAACGCCAAGAACACGGTGTATCTGCTCATGAACAGCCTAGAACCTGAG
CLDSynHS	CAAGACAATGCCAAGAATACAGTGTATCTACTTATGAACAGTCTGGAACCGGAA
	Q D N A K N T V Y L L M N S L E P E
CLDHS	GACACGGCCATCTATTACTGTGCGGCAGATTCGACCATCTACGCTAGTTATTAT
CLDSynHS	GATACTGCAATATACTACTGCGCAGCTGATAGCACCATATATGCCTCCTACTAT
	D T A I Y Y C A A D S T I Y A S Y Y
CLDHS	GAATGTGGTCACGGTCTTTCCACGGGAGGATATGGGTATGACTCCTGGGGCCAG
CLDSynHS	GAGTGTGGTCATGGCCTTTGACTGGAGGCTATGGGTATGACTCATGGGGACAA
	E C G H G L S T G G Y G Y D S W G Q
CLDHS	GGGACCCAGGTCACCGTCTCCTCAGGCCATCACCATCACCATCACTTAA
CLDSynHS	GGTACTCAGGTAAAGTCTCTTCCGACATCACCATCATCACCACATGATAA
	G T Q V T V S S G H H H H H H *

Figure 4.1 Comparison of the nucleotide sequences of the original His-tagged CLD (CLDHS) and codon-optimized version (CLDSynHS). Nucleotide changes are shown in grey.

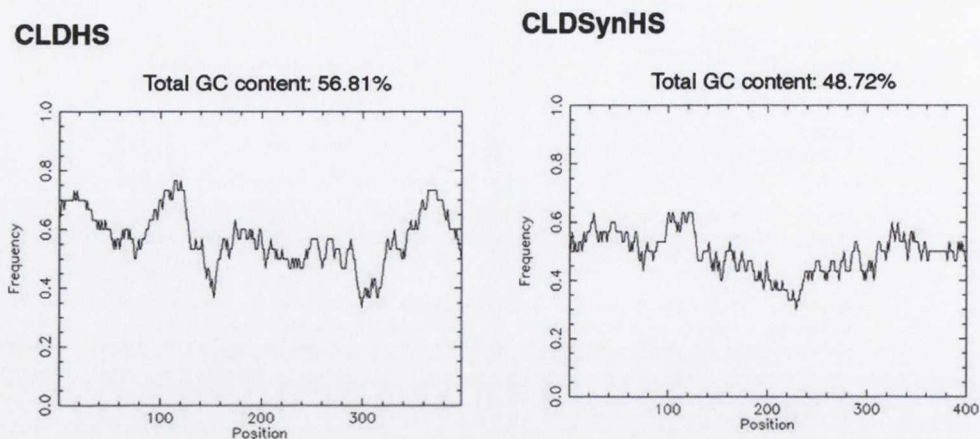


Figure 4.2 A comparison of the GC content of His-tagged CLD (CLDHS) and CLDSyn (CLDSynHS). The graphs show the GC content of each gene expressed as a moving average with a stepping value of 1 and moving window of 30 codons. Total GC content for each gene was displayed on top of the graph.

	-----Framework 1----- ---CDR-H1--- ---Framework 2- ---CDR-H2-
CLD	DVQLQASGGGSVQAGGSLRLSCAAS GYTIGPYCMG WFRQAPGKEREVA AINMGGGITY
CBE	DVQLVESGGGSVQAGGSLRLSCTAS GYTIGPYCMG WFRQAPGQEREVA AINMGGGITY
	----- -----Framework 3----- -----CDR-H3-----
CLD	YADSVKG RFTISQDNAKNTVYLLMNSLEPEDTAIYYCAA DSTIYASYECGHGLSTGGYG
CBE	YADSVKG RFTISRDNKNTVTLMNNLKPEDTAIYYCAA DSTIYASYECGHGLSTGGYG
	--- Framework 4
CLD	YDS WGQGTQVTVSS
CBE	YDS WGQGTQVTVSS

Figure 4.3 Comparison of the amino acid sequences of CLD and CBE (Saerens *et al.*, 2005) (excluding the C-terminal tag). CLD and CBE differ at 10 positions located in Framework regions 1 – 3. Framework region 4 is identical.

4.2.2 Expression and purification of modified anti-HEWL VHHs

The His-tagged (HS) and Strep-tagged (ST) version of the three modified anti-HEWL VHH constructs were cloned into the 3' provector pICH11599 (Santi *et al.*, 2006). The constructs were delivered into *N. benthamiana* along with the other required provectors via *Agrobacterium* infiltration. It was found that both the His-tagged and the Strep-tagged version of each VHH accumulated highly when targeted to the apoplast compared to relatively low accumulation in the cytoplasm and the chloroplast (Figure 4.4a and b).

Furthermore, when comparing the expression levels of His-tagged VHHs against the Strep-tagged versions, His-tagged proteins (Figure 4.4a) generally had higher yields than the Strep-tagged VHHs (Figure 4.4b). It has to be noted that there was some batch-to-batch variation (and in some instances, leaf to leaf variation) which might be attributable to environmental conditions. The data also showed that CBESynHS has a higher yield (up to 0.4 mg/g fresh weight) compared to CLDHS and CLDSynHS (about 0.2 mg/g fresh weight). On the other hand, there was little difference between the original and the codon-optimized CLD but surprisingly, CLDSynHS showed less consistent yields compared to CLDHS (Figure 4.5a). The difference in expression levels was even more evident in the Strep-tagged proteins (Figure 4.4b and 4.5b) where CLDST were barely detectable on a PAGE gel and while CLDSynST has slightly better yield, the expression of CBESynST was the highest among the three (about 0.2 mg/g fresh weight). Furthermore, for both His-tagged and Strep-tagged CLD and CLDSyn, there was a fainter band just below the band with the correct molecular weight in the PAGE gel which may represent a degradation product.

Camel VHHs were recently reported to accumulate highly in the ER compared to the cytoplasm, chloroplast and plant storage vacuole (Winichayakul *et al.*, 2009). To investigate the accumulation of VHHs in the ER, 3 different constructs encoding SEKDEL-tagged CBESyn were made: 1) CBESyn with 3' SEKDEL sequence (CBESynKDEL); 2) CBESyn with a 5' 6xHis tag and 3' SEKDEL sequence (HSCBESynKDEL); 3) CBESyn with a 3' 6xHis tag followed by the SEKDEL sequence (CBESynHSKDEL) (Figure 4.6). The 6xHis tag was placed either at the 5' end of the GOI or in between the GOI and the -SEKDEL tag so as not to interfere with the functionality of the ER retention signal. The constructs were cloned into pICH11599 and delivered via *Agrobacterium* into *N. benthamiana* plants along with the relevant

vectors to add a 5' apoplast targeting presequence to the encoded protein (pICH8420 and pICH10881). Additionally, CBESyn without the 6xHis tag was targeted to the apoplast as a comparison. The apoplast targeting presequence together with the -KDEL retention signal (Gomord *et al.*, 1997) was expected to result in protein accumulation in the ER.

It was found that up to approximately 0.2 mg/g fresh weight of VHH was produced, with the VHHs containing 6xHis tags (HSCBESynKDEL and CBESynHSKDEL) accumulating at much higher levels when compared to the CBESynKDEL, the construct without the purification tag (Figure 4.7). The yield was also slightly lower when compared to CBESynHS targeted to the apoplast (Table 4.1). There were also fainter bands just below the band with the correct molecular weight in the PAGE gel for -SEKDEL tagged VHHs, which may represent degradation products. Furthermore, when the VHH targeted to the apoplast (CBESyn) was compared with that targeted to the ER (CBESynKDEL), it was found that former accumulated more highly than the latter (Figure 4.7). Because targeting of VHHs for accumulation in the ER (i) resulted in bands of lower than expected molecular weight and (ii) did not result in increased VHH accumulation relative to CBESyn, we did not investigate the -SEKDEL-tagged VHHs further.

Instead, leaves expressing CLDHS, CLDSynHS and CBESynHS were harvested at the time when VHH accumulation was the highest (11 dpi for CLDHS and CLDSynHS; 7dpi for CBESynHS), and the corresponding VHHs purified using Nickel-NTA columns (Figure 4.8). Yields of up to about 0.1 mg/g fresh weight were obtained for CLDHS and CLDSynHS (Table 4.1); while for CBESynHS there was a slight increase to about 0.22 mg/g fresh weight.

4.2.3 Expression of anti-HEWL VHH constructs over time

The accumulation of 6xHis- and Strep-tagged version of CLDSyn and CBESyn over time was also investigated by infiltrating leaves with the VHH constructs and harvesting the leaves from 0-15 dpi on alternate days. For both VHHs, detectable amounts of protein could be seen on a Coomassie-stained gel and by Western blot at 3 dpi (Figure 4.9). They continued to accumulate until 7-9 dpi for CBESynHS (Figure 4.9b) and 9-11 dpi for CBESynST (Figure 4.9d), while for CLDHS and CLDST the highest protein accumulation occurred somewhat later at 11-13 dpi (Figure 4.9a and c; CLDSyn showed a similar expression pattern; data not shown). After that, the levels dropped slightly even when the loss of the Rubisco subunits was taken into account. The last time point taken was at 15 dpi and after that time point, the leaves were too necrotic.

From the PAGE and Western blot analyses, the difference in protein accumulation between CLD and CBESyn was even more evident. Furthermore, the previously mentioned lower molecular weight band just below the band of the correct size can also be seen in CLDHS and CLDST (Figure 4.9a and c). Interestingly, the increase in VHH correlates with a decrease in the Rubisco large (54 kDa) and small (14 kDa) subunits which coincides with the progressive necroticity of the infiltrated areas over time. This phenomenon has been observed elsewhere (Gils *et al.*, 2005) and is probably due to the overwhelming amount of mRNA transcripts produced by the viral expression system causing the plant to redirect its translational machinery to produce mostly VHH, which likely results in depletion of essential endogenous proteins (enzymes) and, in consequence, disruption of primary metabolism.

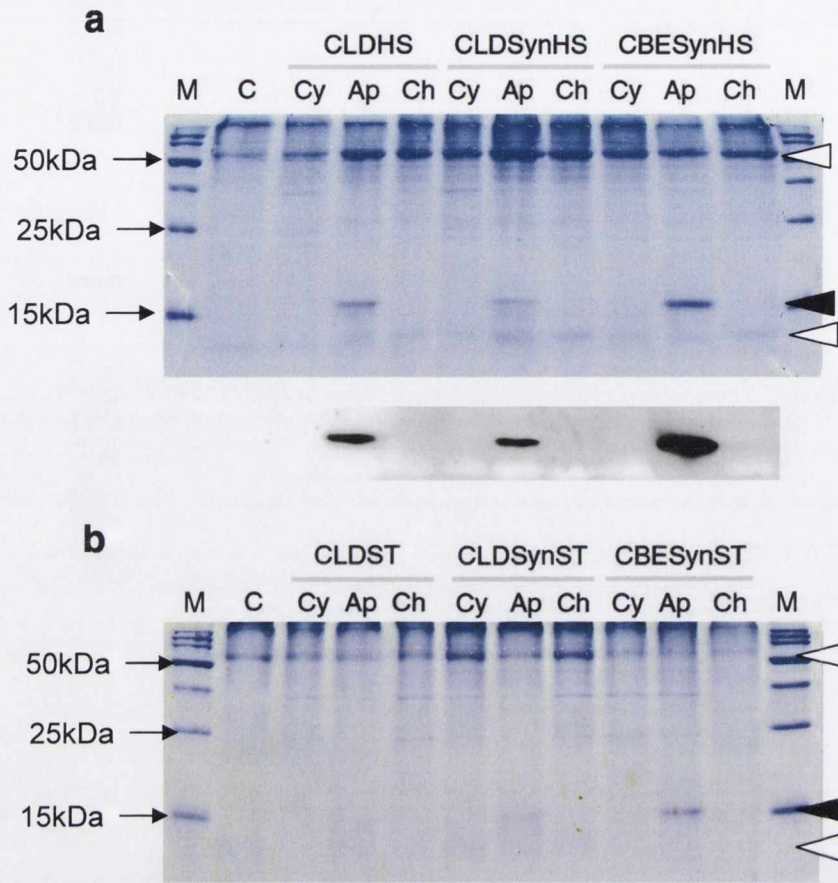


Figure 4.4 Accumulation of anti-HEWL antibodies targeted to different cellular compartments.

(a) PAGE gel (top) and Western blot (bottom) showing accumulation of the three different His-tagged anti-HEWL camel VHHs at 7 dpi in the cytoplasm (Cy), and when targeted to the apoplast (Ap) or chloroplast (Ch), compared with leaves transformed with the empty vector (C). 10 μ g of total soluble protein was loaded in each lane. M, Perfect ProteinTM Markers. White arrow heads indicate the large and small RUBISCO subunits (54 kDa and 14 kDa, respectively). The VHH is indicated by a black arrow head (approx. 16kDa). Gels and blots is representative of the three experiments conducted.

(b) PAGE gel showing accumulation of the three different Strep-tagged anti-HEWL camel VHHs at 9 dpi in the cytoplasm (Cy), and when targeted to the apoplast (Ap) or chloroplast (Ch). Other details as in (a).

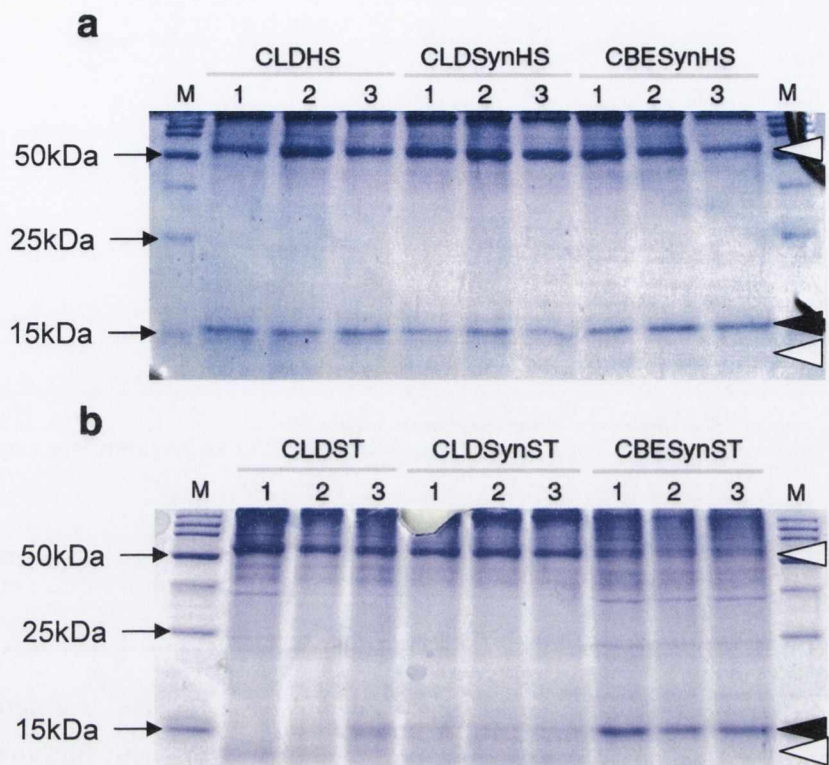


Figure 4.5 PAGE gel showing VHH abundance in 3 randomly sampled leaves infiltrated with the three different His-tagged (HS) (a), and Strep-tagged (ST) (b) anti-HEWL constructs. Leaves were chosen randomly from 3 different plants and harvested at 7 dpi. 10 μ g of total soluble protein was loaded in each lane. M, Perfect ProteinTM Markers (M). White arrow heads indicate the large and small RUBISCO subunits (54 kDa and 14 kDa, respectively). The VHH is indicated by a black arrow head (approx. 16 kDa). PAGE gels is representative of the three experiments conducted.

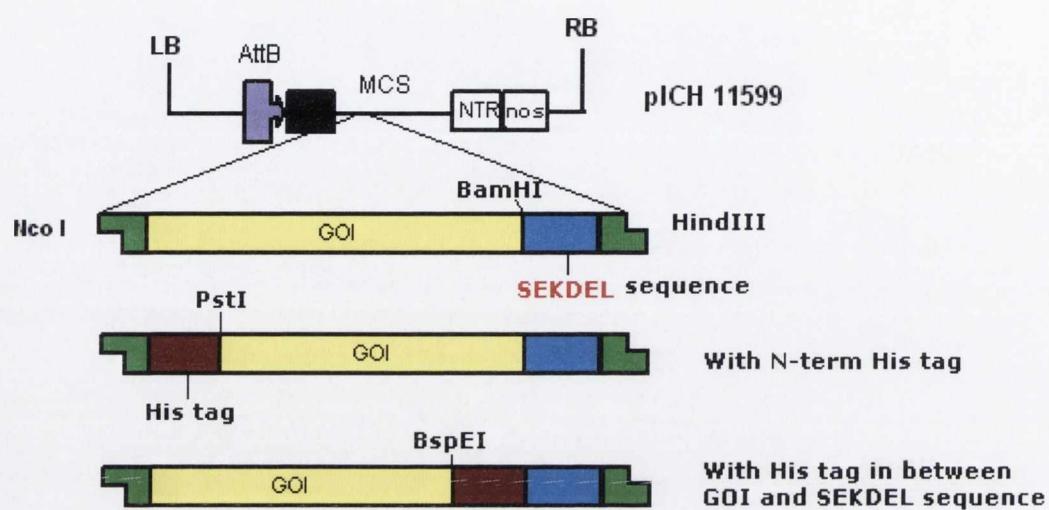


Figure 4.6 Schematic representation of the different SEKDEL-tagged CBESyn constructs. From top to bottom: CBESynKDEL, HSCBESynKDEL and CBESynHSKDEL.

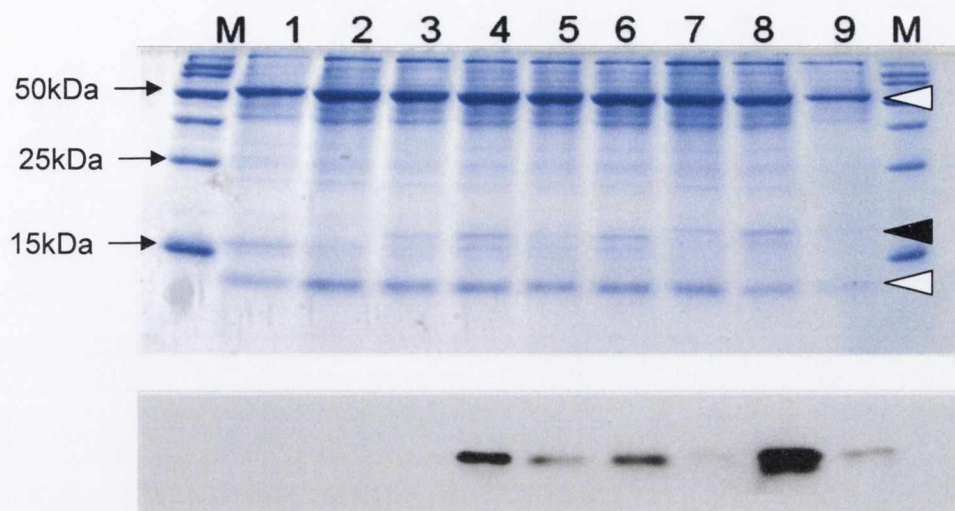


Figure 4.7 Accumulation of anti-HEWL antibodies targeted to the ER. 6xHis purification tags were located either at the 5' end of GOI or between GOI and the -KDEL retention signal.

PAGE gel (top) and Western blot (bottom) showing accumulation of anti-HEWL camel VHHs in 3 different *N. benthamiana* leaves at 7 dpi in endoplasmic reticulum (ER). Leaves were infiltrated with either a CBESynKDEL construct with a 5' His-tag (Lanes 4-6) or a His-tag located between CBESyn and the -KDEL retention signal (Lanes 7-9). Lanes 1, 2 and 3 showed protein expression profile of leaves transformed with CBESyn, CLDSyn and CBESynKDEL respectively. CBESyn and CLDSyn were targeted to the apoplast and did not have any purification tags attached to the GOI. CBESynKDEL possessed the -SEKDEL retention signal but did not possess any purification tags. 10 μ g of total soluble protein was loaded in each lane. M, Perfect Protein™ Markers. White arrow heads indicate the large and small RUBISCO subunits (54 kDa and 14 kDa, respectively). The VHH is indicated by a black arrow head (approx. 16kDa). The Western blot was probed with 1:5,000 mouse anti-His primary antibody followed by 1:10,000 HRP coupled anti-mouse IgG. PAGE gel and western blot is representative of the three experiments conducted.

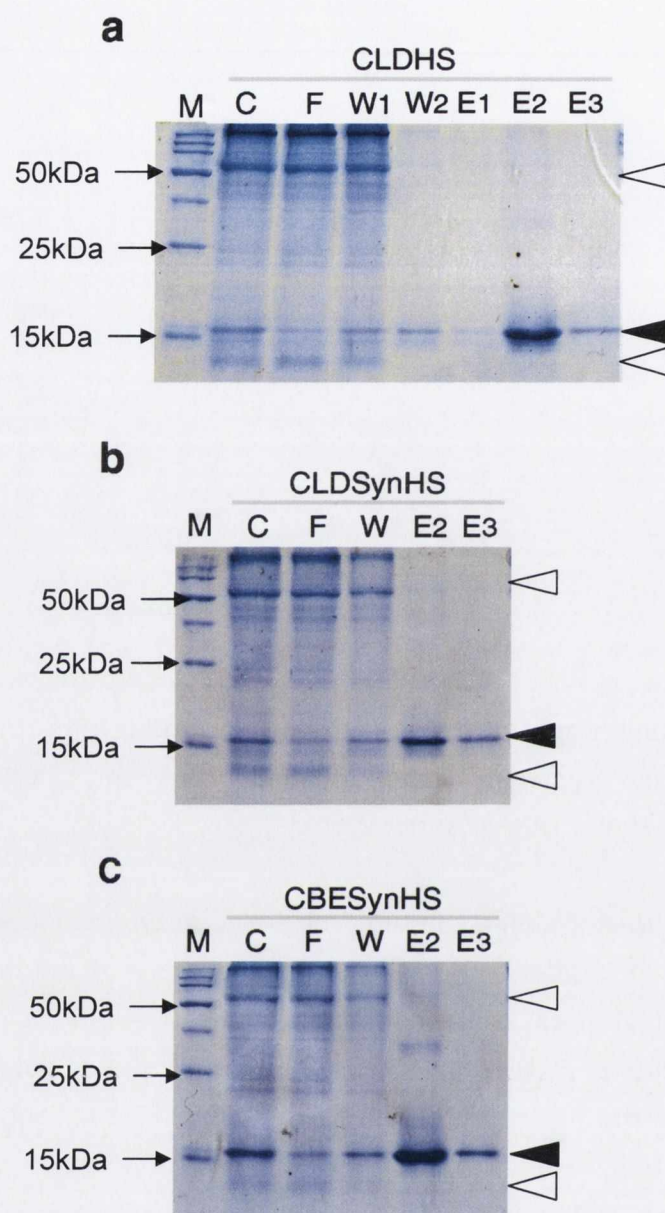


Figure 4.8 Purification and yield of His-tagged CLD, CLDSyn and CBESyn. PAGE gels showing the different fractions during the purification of (a) CLDHS, (b) CLDSynHS and (c) CBESynHS. Approximately 2-2.5 g of infiltrated leaf material was harvested per construct. 10 μ L of each fraction was loaded (except for CBESynHS E2 where 5 μ L was loaded). C, Crude extract; F, Flowthrough; W, Wash; E, Elution. M, Perfect ProteinTM Markers. White arrow heads indicate the large and small RUBISCO subunits (54 kDa and 14 kDa, respectively). The VHH is indicated by a black arrow head (approx. 16 kDa). PAGE gels is representative of the three experiments conducted.

Table 4.1 Table showing for each construct, the compartment with the highest accumulation, the mean estimated yield before purification ($\pm$ SD; obtained using density analysis) and the mean yield after purification ($\pm$ SD; as determined by Bradford Assay). Data collected from three leaf samples from three independent experiments (total sample size=9).

Antibodies	Compartment with highest accumulation	Yield before purification* (mg/g fresh weight)	Yield after purification [†] (mg/g fresh weight)
CLDHS	Apoplast	0.18 $\pm$ 0.04	0.10 $\pm$ 0.03
CLDSynHS	Apoplast	0.20 $\pm$ 0.04	0.13 $\pm$ 0.02
CBEHS	Apoplast	0.35 $\pm$ 0.05	0.22 $\pm$ 0.06

*Protein concentration estimated by comparing protein bands on PAGE gel to Perfect ProteinTM Markers (Novagen, USA) using Multi Guage software (FujiFilm, Japan)

[†]Protein concentration determined using Bradford assay

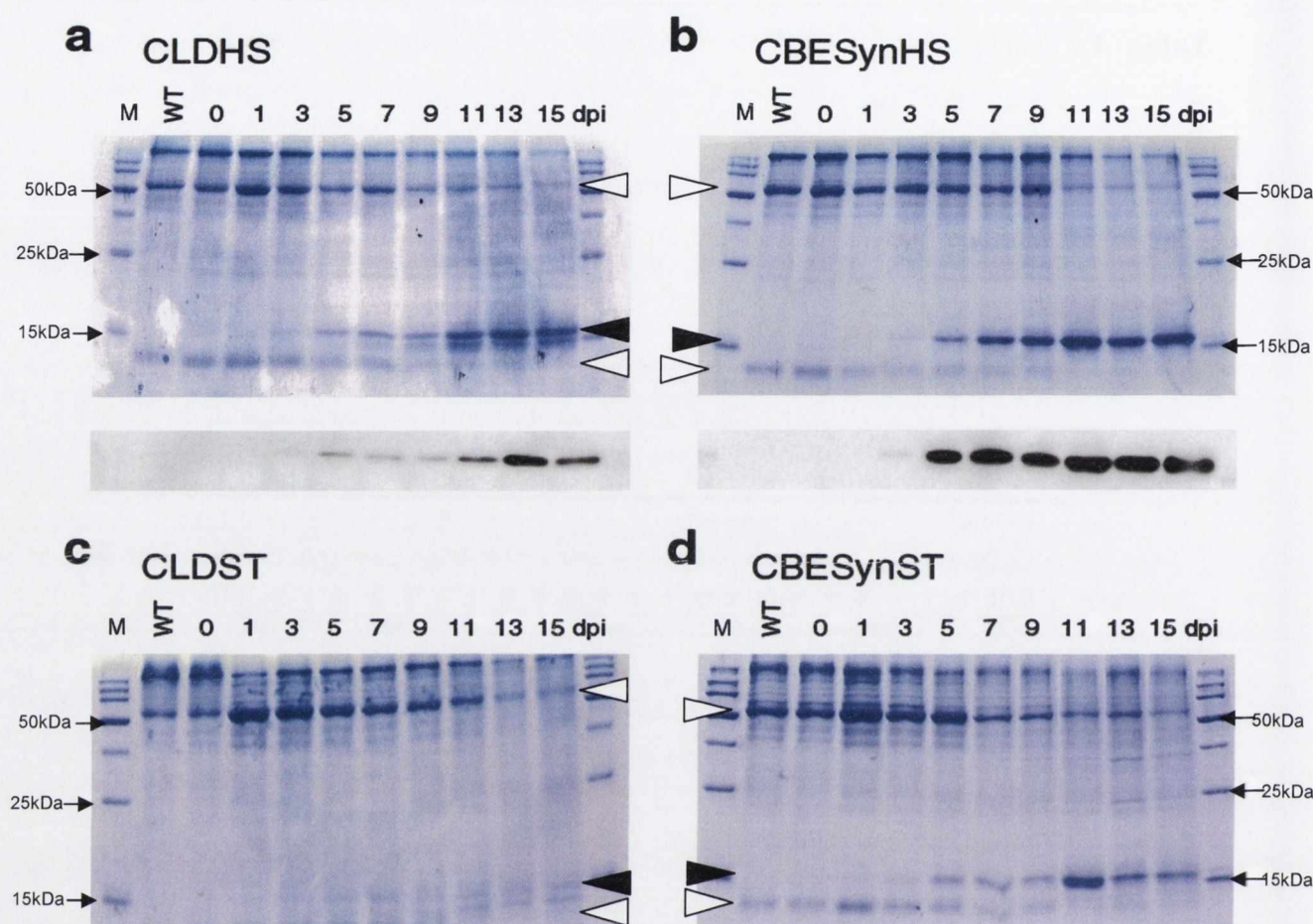


Figure 4.9 Accumulation of apoplast-targeted anti-HEWL antibodies over time.

(a and b) PAGE gels (top) and Western blots (bottom) showing accumulation of apoplast-targeted His-tagged CLD (a) and CBESyn (b) from 0 to 15 dpi at two-day intervals compared to uninfiltrated leaves (WT). 10 μ g of total soluble protein was loaded in each lane. M, Perfect Protein™ Markers. White arrow heads indicate the large and small RUBISCO subunits (54 kDa and 14 kDa, respectively). The VHH is indicated by a black arrow head (approx. 16 kDa). Gels and blots is representative of the three experiments conducted.

(c and d) PAGE gels showing accumulation of apoplast-targeted Strep-tagged CLD (c) and CBESyn (d) from 0 to 15 dpi at two day intervals, compared with uninfiltrated leaves (WT). 10 μ g of total soluble protein was loaded in each lane. Other details as in (a).

4.2.4 Characterisation of anti-HEWL VHHs

The specificities of the purified VHHs were determined by direct ELISA with antigen-bound VHH detected using HRP-conjugated Ni-NTA to detect the His-tags. It was found that all three VHHs bind to their antigen HEWL and not to other unrelated antigens such as the structurally similar α -lactalbumin (Figure 4.10) or β -galactosidase (data not shown).

A competitive ELISA protocol, as described by Friguet *et al.* (1985), was used to determine dissociation constants for CLDSynHS and CBESynHS. Briefly, VHHs diluted to a concentration predetermined by direct ELISA (about 1 to 5×10^{-8} M) were incubated with increasing concentration (1×10^{-12} M to 1×10^{-4} M) of antigen (in this case HEWL). The mixture was then incubated in wells coated with a fixed amount of antigen. Free antibodies not previously bound to the free antigen will bind to the antigen coated on the plate and these antibodies were detected using Ni-NTA probes bound to HRP that will detect the 6xHis-tag on the antibody. The resulting absorbance values (Figure 4.11) were plotted according to Friguet *et al.* (1985). Using this method, the dissociation constants of CLDSynHS and CBESynHS were found 4.73 ± 2.4 nM and 46.3 ± 3.3 nM, respectively (based on 3 repeats). Saerens *et al.* (2005) reported an increase in the dissociation constant (K_D) from 7.2 nM for cAB-Lys3 (CLD) to 52 nM for cBCII10-L-L-L (CBESyn). It has also been shown that there were slight conformational changes in the cAb-Lys3 CDRs (particularly CDR1 and CDR3) when they are placed in the cAbBCII10 framework (Saerens *et al.*, 2005). Therefore, the slight decrease in affinity may be due to the variability in the length of the CDRs and loop contacts with the framework residues.

The thermal stability of CLDHS and CBESynHS was also investigated. Both VHHs were incubated for 20 minutes at different temperatures, re-equilibrated to room temperature for 5 minutes and assayed by direct ELISA. It was found that both antibodies were functional at temperatures above 50°C (Figure 4.12). Camel antibodies have been found to be stable at high temperatures due to their single domain nature and their ability to fold efficiently after thermal or chemical destabilization (Perez *et al.*, 2001; Dumoulin *et al.*, 2002; Ewert *et al.*, 2002) due in part to the increased hydrophilicity caused by their signature mutations in the former VL interface (Davies and Riechmann, 1995; Davies and Riechmann, 1996; Conrath *et al.*, 2005) and to the folding of the extended CDR3 against this region (Bond *et al.*, 2003). However, our investigations revealed that the antigen-binding

activity of CBESynHS decreased far more rapidly than that of CLDHS and that most of its activity is lost after a 20 min incubation at 90°C (Figure 4.12). This suggests that CBESynHS denatures more rapidly than CLDHS although the conformational stability of CBESynHS has been reported to be higher than that of CLDHS (Saerens *et al.*, 2005). Remarkably, CLDHS still retained two-thirds of its activity after incubation at 90°C. It is interesting to note that although the conformational stability of CBESynHS (cBCII10-L-L-L) has been reported to be higher than CLDHS (cAB-Lys3) (Saerens *et al.*, 2005); CLDHS is more stable at higher temperatures compared to CBESynHS.

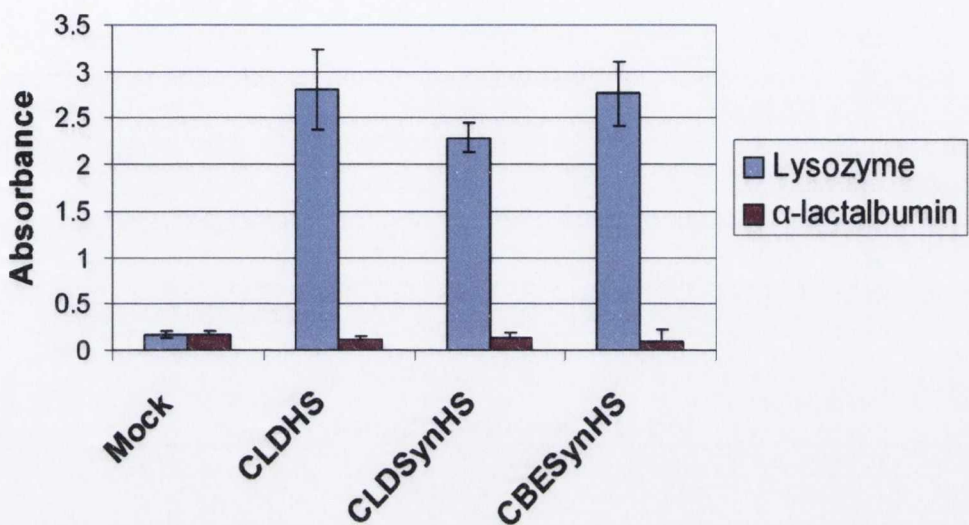


Figure 4.10 Specificity of the anti-HEWL camel antibodies as determined by ELISA. All three versions of the antibody recognized HEWL (■) and did not bind to the structurally similar but unrelated antigen α -lactalbumin (■). Each point represents the mean of three experiments. Bars indicate 2 standard deviations (SDs). 'Mock' represents a control experiment where no nanobody was introduced.

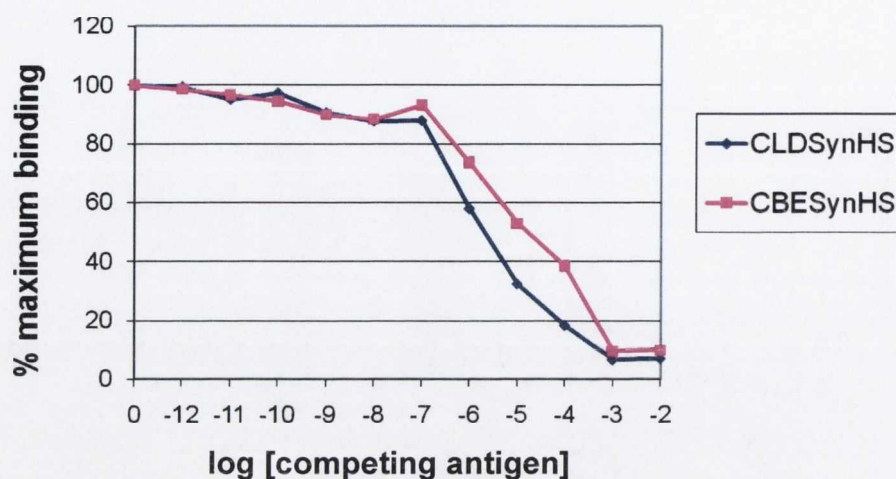


Figure 4.11 Affinity of His-tagged CLDSyn (◆) and CBESyn (■) measured according to Friguet *et al.* (1985). VHHs diluted to a concentration predetermined by direct ELISA (about 1 to $5 \times 10^{-8} \text{M}$) were incubated with increasing concentration ($1 \times 10^{-12} \text{M}$ to $1 \times 10^{-2} \text{M}$) of antigen (in this case HEWL). The mixture was then incubated in wells coated with a fixed amount of antigen. Free antibody not previously bound to the free antigen will bind to the antigen coated on the plate. The 6xHis-tags of these antigen-bound antibodies were detected using an Ni-NTA probe bound to HRP. The resulting absorbance values were plotted according to Friguet *et al.* (1985).

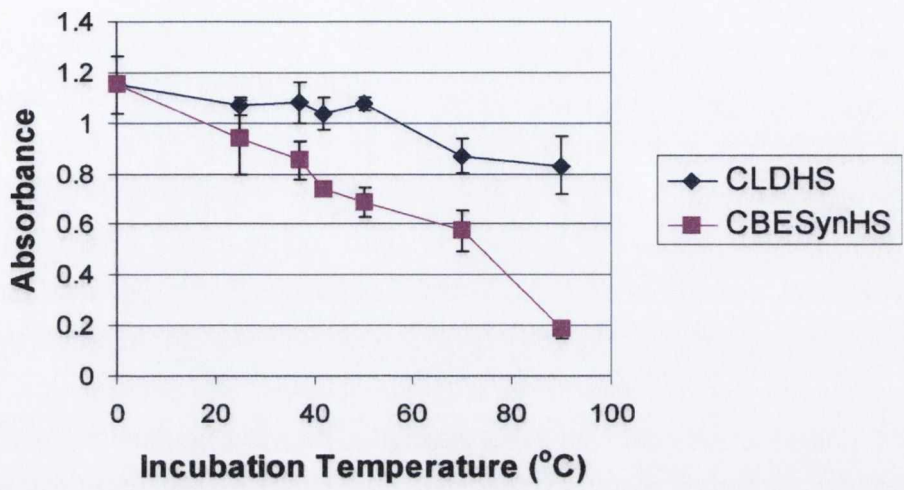


Figure 4.12 Comparison of thermal stability of CLDHS (◆) and CBESynHS (■). Approximately 0.1 ng/mL of each antibody was incubated at increasing temperatures for 20 minutes, reequilibrated to room temperature and assayed using ELISA. Each point represents the mean of three experiments. Bars indicate 2 standard deviations.

4.2.5 Cloning and expression of a CBESyn-GFP fusion protein

A fusion protein construct consisting of CBESyn joined to a Strep-tagged Green Fluorescent Protein (GFP) via a free linker was made by PCR amplification of the individual components and ligating them into pICH11599 (Figure 4.13). This is to investigate whether the fusion of CBESyn with another protein, might alter its accumulation pattern or improve its stability (particularly at the C-terminal region). Following sequence verification, the construct was then infiltrated with appropriate 5' provector strains into *N. benthamiana* leaves in order to express the fusion protein in the cytoplasm, and to target it to the apoplast and chloroplasts. At 7 dpi, strong GFP fluorescence could be seen when leaves were illuminated with a long-wave UV lamp (Figure 4.14, top row). GFP fluorescence was strongest when the CBESyn-GFP fusion protein was expressed in the cytoplasm (a), followed by targeting to the chloroplast (c) and the apoplast (b).

At 200 x magnification (Figure 4.14, bottom row), it was found that GFP fluorescence was present throughout the cytoplasm of cells expressing the cytoplasm-targeted version, but was also found in discrete subcellular foci (as indicated by white arrow) which might correspond to sites of protein accumulation (Figure 4.14a; bottom row). For the apoplast-targeted fusion, GFP fluorescence was distributed throughout the cell but showed specific localization in the apoplastic spaces between the cell walls. Strong fluorescence could also be seen in subcellular bodies that might correspond to the ER (Figure 4.14b; bottom row). On the other hand, fluorescence in the plant infiltrated with the chloroplast-targeted construct could be found in small clusters consistent with the spatial distribution of chloroplasts (Figure 4.14c; bottom row).

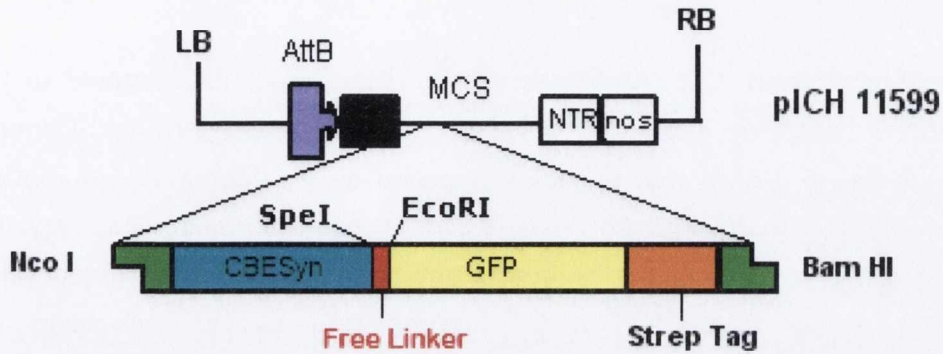


Figure 4.13 Schematic diagram of the CBESyn-GFP fusion constructs. Anti-HEWL VHH CBESyn was joined to Strep-tagged GFP using a free linker (GGAGCTGGAGCT).

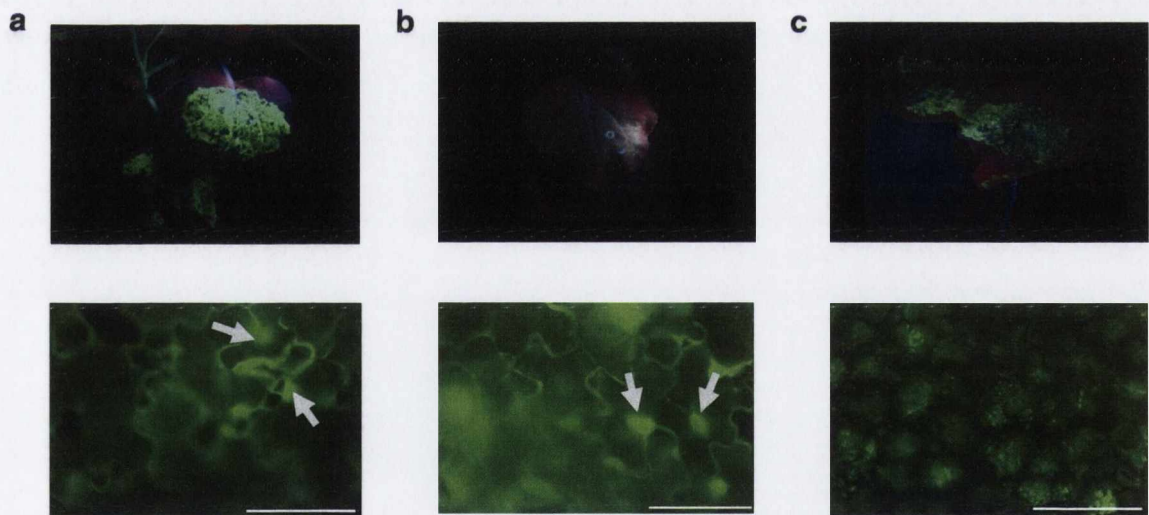


Figure 4.14 Fluorescence of 7 dpi leaves of *N. benthamiana* accumulating CBESyn-GFPST in cytoplasm (a), targeted to the apoplast (b) or chloroplast (c). Bottom row of photos shows cells at 200 x magnification. White arrows indicate areas of high GFP accumulation which may correspond to protein bodies. Scale bar: 10 µm.

4.2.6 Accumulation of CBESyn-GFPST targeted to different cellular compartments

Proteins extracted from 7 dpi leaves expressing CBESyn-GFPST targeted to three cellular compartments were analysed by SDS-PAGE and by Western blot. Compared with the levels of CBESynST observed previously in Coomassie blue-stained polyacrylamide gels (e.g. Figure 4.9d), CBESyn-GFPST accumulated at very much lower irrespective of the cellular compartment to which it was targeted (Figure 4.15a). The levels of CBESyn-GFPST accumulating in the cytoplasm were sufficiently high to be detected on Western blots probed with a 1:5,000 dilution of the HRP-bound Strep Tactin probe. Surprisingly, however, neither the apoplast- nor the chloroplast-targeted fusion proteins were detected by the Strep Tactin probe, despite the strong GFP fluorescence observed in leaves expressing both the corresponding constructs. However, when the blot was reprobed with a polyclonal antibody against GFP, the fusion protein was detected in all three cellular compartments, albeit at very different levels. Protein accumulation was strongest when the fusion protein was expressed in the cytoplasm (Cy), followed by the chloroplast (Ch), consistent with the *in vivo* levels of green fluorescence previously observed. The lack of detection when using the Strep Tactin probe is likely, therefore, due to proteolytic removal of part or all of the 8 amino acid C-terminal Strep tag.

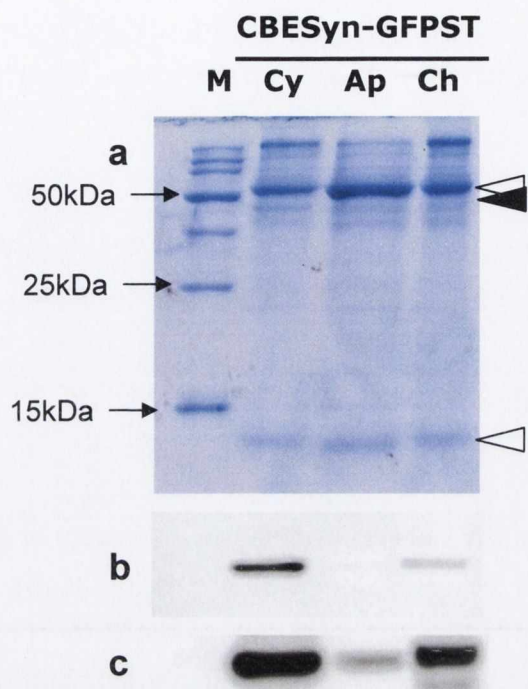


Figure 4.15 PAGE gel (a) and Western blot (b and c) of CBESyn-GFPST accumulating in cytoplasm (Cy), apoplast (Ap) and chloroplast (Ch) at 7 dpi. 10 μ g of total soluble protein was loaded in each lane. The Western blot was probed with 1:5,000 Streptactin-HRP (b) or 1:1,000 polyclonal antibody against GFP (c). White arrow heads indicate the large and small RUBISCO subunits (54 kDa and 14 kDa, respectively). A black arrow head indicates the CBESyn-GFPST fusion protein (approx. 43 kDa). PAGE gel and western blots is representative of the three experiments conducted.

4.3 Discussion

4.3.1 Accumulation of the three versions of anti-HEWL VHH constructs in *N. benthamiana*

In this work we have shown, for the first time, that Camelid VHHs, or nanobodies, can accumulate to very high levels in plants. By 7–9 dpi, apoplast-targeted nanobodies expressed from a deconstructed viral vector and delivered via agroinfiltration accumulated at levels that either equalled or greatly exceeded that of the similarly-sized Rubisco SSU in total leaf protein extracts. Remarkably, between 11 and 15 dpi, the highly expressed His6-tagged VHHs comprised the most prominent protein band visible on Coomassie blue-stained gels, corresponding to an estimated 10–30% total soluble protein, and giving post-purification yields of 0.1–0.22 mg/g fresh weight. These values approximate the highest yields so far achieved for any antibody format expressed in plants (0.2–0.5 mg/g fresh weight; Giritch *et al.*, 2006).

The folding and stability of many different candidate recombinant proteins expressed in plants, especially antibodies, has been shown to be affected by the subcellular compartment in which they accumulate, which in turn may also affect overall yield and functionality (Sharp and Doran, 2001; Doran, 2006). This is likely due to compartment-specific differences in redox potential, pH, and the presence/absence of specific molecular chaperones and foldases, glycosyltransferases and proteolytic enzymes (Doran, 2006; Benchabane *et al.*, 2008). The oxidizing environment of the secretory pathway, in particular that obtaining within the ER, has been shown to promote correct folding and assembly of full-sized and smaller antibody derivatives by facilitating the formation of essential disulphide bonds (Conrad and Fiedler, 1998; Schillberg *et al.*, 1999; Peeters *et al.*, 2001). Winichayakul *et al.* (2009) reported significant accumulation of anti-TNF- α nanobodies only when they were targeted for retention within the ER in stably transformed *Arabidopsis* plants. However, we did not observe high-level accumulation of anti-HEWL VHHs in agroinfiltrated *N. benthamiana* leaves when an ER-retention signal (KDEL) was added to their C-terminus (Section 4.2.2). Instead, with the exception of CBESynST which showed significant accumulation of truncated forms in the cytoplasm and chloroplasts in some instances, all of the nanobody variants we investigated accumulated at a high level only when targeted to the

apoplast. The apoplast has also been found to be the compartment of choice for high-level accumulation of other recombinant proteins including human growth hormone (Gils *et al.*, 2005), *Yersinia pestis* antigens (Santi *et al.*, 2006) and recombinant bovine aprotinin (Badri *et al.*, 2009). However, it is also interesting to note that the addition of a fusion partner (GFP) to VHH, as in the case of CBESyn-GFPST, changed the accumulation pattern in the three compartments: CBESyn-GFPST accumulated highly in the cytoplasm instead of the chloroplast and was not detected in the apoplast.

In addition, the observation by Menassa *et al.* (2003) that human interleukin-10 accumulated in plants only if the protein contained a C-terminal 6xHis tag, led us to investigate whether different purification tags might also influence nanobody accumulation. Remarkably, the 6xHis-tagged VHHs accumulated at levels several-fold higher than the corresponding Strep-tagged versions, suggesting an overall stabilizing effect of this tag on nanobody accumulation. The influence of the Strep tag was more complex. Its presence at the C-terminus of CBESyn (but not CLD or CLDSyn) facilitated accumulation of truncated versions of this VHH in the cytosol and in chloroplasts, while the 6xHis tag did not. This suggests that the amino acid sequence of the VHH itself and possibly its interaction with the tag may play an important role in determining whether a particular tag exerts a stabilizing or destabilizing effect on accumulation in different subcellular compartments. However, other possibilities, such as degradation and/or cleavage by specific proteases cannot be ruled out (Outchkourov *et al.*, 2003; Faye *et al.*, 2005; Doran, 2006; Benchabane *et al.*, 2008a). These factors can have a major impact on protein yield and might explain the slightly lower molecular weight band observed below the correct-sized protein in extracts containing CLDHS, CLDSynHS and CLDSynST.

4.3.2 The effect of codon optimization/amino acid changes on yield

When considering the expression of a GOI in a heterologous organism, intrinsic elements in the gene itself need to be taken into account. It has been well-documented for example, that different organisms have their own particular codon usage bias (Ikemura, 1985) and that codon usage differences between the host organism and the GOI can significantly affect gene expression. In addition, GOIs that are especially AT-rich may contain cryptic polyadenylation signals and cryptic splice donor or acceptor sites, the presence of which may drastically reduce the levels of mRNAs containing the original uninterrupted open reading frame. For these reasons, modification of the nucleotide sequence of a GOI to improve codon usage and remove potentially disruptive *cis*-elements can increase the yield of recombinant protein in plants (Kirby and Kavanagh, 2002; Geyer *et al.*, 2007; Li *et al.*, 2007) and in other organisms (Yadava and Ockenhouse, 2003; Burgess-Brown *et al.*, 2008). In addition, changing specific amino acid residues can significantly affect protein stability (Sawant *et al.*, 2001), how the protein is processed (Hondred *et al.*, 1999), and even influence post-translational targeting of the protein (Benchabane *et al.*, 2008b).

A significant increase in yield occurred when the first glutamine (Q) of CL3HS (as described in Chapter 3) was changed to aspartic acid (D) in CLDHS. Also, the change in framework from CLD to CBESyn (10 amino acid differences) resulted in an approximately two-fold increase in yield. This might have been due to the elimination of protease recognition sites as a consequence of the framework alterations. Although the targeting of proteins, especially antibodies, through the secretory pathway has been shown to promote protein accumulation, the pathway is not protease free. Subtilisin/kexin-like serine and metallo-proteases which process pro-proteins are known to be present in the secretory pathway (Viale *et al.*, 1999; Raharjo *et al.*, 2001) and cysteine proteases are known to degrade misfolded proteins (Otsu *et al.*, 1995) or to mobilize seed storage proteins (Schlereth *et al.*, 2001). Moreover, Delannoy *et al.* (2008) reported the presence of aspartic, cysteine and serine proteases in leaf intercellular (extracellular) fluid. In the past, it has been shown that certain proteins targeted to the plant secretory system can be adversely affected by proteases (Stevens *et al.*, 2000; Sharp and Doran, 2001). For example, Outchkourov *et al.* (2003) reported that sea anemone equistatin, a

protease inhibitor, was sequentially degraded when targeted through the secretory pathway in potato. Identification of the precise cleavage sites by N-terminal sequencing of affinity-purified equistatin degradation products, strongly implicated arginine/lysine-specific and Asn-specific cysteine proteases. This kind of analysis can facilitate the rational modification of key residues in a candidate recombinant protein (via site-directed mutagenesis) in order to minimize proteolytic attack while preserving biological activity, and thereby increase protein accumulation and yield.

4.3.3 The effect of framework changes on yield

Among the three versions of anti-HEWL VHH expressed, the apoplast-targeted VHH with the alternative framework (CBESynHS) had the highest yield. The enhanced accumulation and stability of CBESynHS was therefore likely due to the different framework upon which it is based. This framework, derived from cAbBCII10 (Saerens *et al.*, 2005), differs from that of CLDHS at 10 amino acid positions, and was identified in a search for a 'universal' framework suitable for CDR loop-grafting applications. The cAbBCII10 framework was chosen for its high stability, good expression level in *E. coli*, and its ability to retain functionality even when two conserved disulphide bond-forming cysteines were removed. Indeed, the chimeric nanobody cAbBCII10-L-L-L (corresponding in this study to CBESynHS) was shown by Saerens *et al.* (2005) to have an enhanced thermodynamic stability relative to the CDR loop donor cAbLys3 (in this study, CLDHS), confirming the suitability of the cAbBCII10 scaffold as a CDR loop acceptor, and consistent with its high-level accumulation and improved resistance to proteolysis reported here. Although we have not yet determined whether the conserved cysteines in CBESynHS are necessary for the stability and functionality of the plant-expressed protein, previous experiments with scFvs (Chapter 3) showed that an anti β -galactosidase scFv with its cysteine residues eliminated through mutagenesis (for better expression in the bacterial cytoplasm) accumulated highly in the plant cytoplasm compared to the original version which accumulated highly in the chloroplast. Furthermore, the yield of the mutated scFv was several times higher than the original version.

Another possible explanation of why the VHH with the alternative framework had a higher yield is that its tertiary structure affects overall protein stability in the cell and resistance to protease-mediated degradation, compared to the tertiary structure of the native framework construct. He *et al.* (2011) reviewed 31 successful examples of human recombinant proteins expressed in the mammary

glands of transgenic animals with the resultant proteins accumulating in their milk. They compared the codon usage of these proteins against codon usage of mammary gland-specific genes and genes coding for milk proteins in human, mouse and five other livestock animals but found no correlation between the expression levels of the human recombinant proteins and the similarities of the codon usage patterns. However, when they compared the main protein domains of the recombinant proteins against the protein domains of casein, the main component of milk proteins, they found significant differences in the yield between proteins that share high domain similarities with casein than those that did not, even after accounting for other variables such as the type of promoter used and whether genomic DNA or cDNA was used (introns can have a major positive influence on gene expression levels). They also found that out of the top 13 highly accumulated recombinant proteins, 10 shared a similar alpha-beta based major domain similar to casein while the other 3 shared common motifs with casein. They then went on to propose that the tertiary structure of a recombinant protein adapts to the synthesis and secretion process of milk proteins, thus enabling them to be highly expressed in the animal mammary gland. Proteins which do not have the correct structure may not interact properly with the milk proteins and with the ER environment which the milk proteins are targeted to. These protein may then aggregate together and may be degraded by the ER-associated degradation pathway (ERAD) (Vembar and Brodsky, 2008), ultimately affecting their overall yield.

Chapter 5. Exploring the importance of Framework Regions for VHH accumulation and yield in plants

5.1 Introduction

Like the VH-III antigen-binding domain of conventional human antibodies, the Camelid VHH domain possesses 4 framework regions (FRs) that form a core scaffold and 3 complementarity-determining regions (CDRs) that are involved in antigen recognition and binding (Figure 5.1). Several publications have described the crystallographic structure of VHH (Desmyter *et al.*, 1996; Spinelli *et al.*, 1996; Decanniere *et al.*, 1999; Spinelli *et al.*, 2000). The NMR structure of a camelized human VH has also been reported (Riechmann, 1996). The crystal structures showed that VHH and VH are similar in several respects. Both comprise nine β -strands folded into two sheets which are stabilized by a conserved disulphide bond (Figure 5.2). However, VHH differs from VH at four highly conserved amino acid positions in the FR2 region: F/Y37→V, E44→G, R45→L and G47→W (positions based on Kabat numbering; Kabat *et al.* 1991); substitutions shown as black dots in Figure 5.1). Additionally, W103→R substitutions are found in 10% of VHHs (Desmyter *et al.*, 2001). Also, an L11→S substitution in FR1 was reported in camel VHH but is less common in llama VHH (Vu *et al.*, 1997).

Davies and Riechmann (1994; 1996) grafted the four hallmark hydrophilic residues of VHH onto a human VH fragment and found that the 'camelized' human VH fragment was more stable and less prone to aggregation than the original VH fragment. On the other hand, Conrath *et al.* (2005) replaced the hallmark substitutions to its counterpart in human VH. This process restores the VL binding capabilities of the 'decamelized' VHH and although it still behaves as a monomer, the thermodynamic stability and antigen binding affinity was reduced. This is because those hallmark amino acid substitutions in the FR2 region render the framework region more hydrophilic (Spinelli *et al.*, 1996). Furthermore, the rotation of the adjacent hydrophobic chains to expose the most hydrophilic parts makes the VHH even more hydrophilic (Desmyter *et al.*, 1996).

The largest structural difference between VH and VHH is at the CDR regions. In human and mouse VH, these loops are folded in a limited number of canonical structures (Chothia and Lesk, 1987; Chothia *et al.*, 1992). In the VHH, many deviations from these canonical structures were found and these could not have been predicted from their sequence. For example, although the CDR1 of an anti-carbonic anhydrase VHH forms a type-I loop (Desmyter *et al.*, 2001), none of the other published crystallography structures has a CDR1 adopting the canonical structures (Decanniere *et al.*, 1999; Decanniere *et al.*, 2000; Desmyter *et al.*, 2001). Furthermore, the CDR2 of VHH also appears to be more variable than in human or mouse (Decanniere *et al.*, 2000). This is caused by amino acid insertions or deletions in or around CDR1 and CDR2 which occurs more frequently and this widens the possible antigen-binding loop conformations (Decanniere *et al.*, 2000; Harmsen *et al.*, 2000; Nguyen *et al.*, 2000).

The most interesting feature is the CDR3 loop on the VHH. These are typically more extended than the CDR3 in a conventional VH and are stabilized by a disulphide bond with a cysteine in CDR1 or FR2 (Muyldermans *et al.*, 1994). It was found that VHH CDR3 is normally turned outwards from the β -barrel, folding over the region that would interact with the VL in conventional VH (Desmyter *et al.*, 1996; Decanniere *et al.*, 1999; Desmyter *et al.*, 2001; Desmyter *et al.*, 2002). This increases the potential interaction surface with the antigen (Vu *et al.*, 1997), contributes to the overall domain stability (Bond *et al.*, 2003) and in the process prevents any interaction with a VL domain as the loop covers the hydrophobic patches in the FR2 region. Also, the CDR3 loop folds over CDR1 and CDR2 and provides specificity and high binding affinity to the VHH (Desmyter *et al.*, 2001). Several VHHs have been isolated that act as potent enzyme inhibitors due to their extended CDR3 loops penetrating into the active site cleft. A good example of this is provided by cAb-Lys3. The 24 amino acid long CDR3 which extends outwards from the antigen-binding site inserts into the hen egg white lysozyme (HEWL) active site (Figure 5.3; Desmyter *et al.*, 1996), mimicking the natural substrate of lysozyme and thus inhibiting its enzymatic activity (Transue *et al.*, 1998). Thus the extended CDR3 enables VHHs to reach antigenic sites inaccessible to conventional antibodies such as enzyme active sites (Lauwereys *et al.*, 1998; De Genst *et al.*, 2006) and conserved cryptic epitopes (Stijlemans *et al.*, 2004). However, it should also be noted that not all VHHs work in this way, as non-inhibiting VHH fragments have also been isolated (Lauwereys *et al.*, 1998). Presumably, these VHH would interact with other clefts or surfaces of their antigen.

In Chapter 4, we showed that the use of an alternative 'universal' framework increased the yield *in planta* of an anti-HEWL VHH, in comparison with its original-framework counterpart. In this chapter, we investigate whether the use of this alternative framework might also increase the yield of other VHHs, relative to their original-framework versions. To this end, two VHHs, one recognizing human Prostate Specific Antigen (PSA) and the other bovine RNase A, were investigated. The yields of codon-optimized versions of these VHHs possessing the 'universal' cAbBCII10 framework were compared to versions with the codon-optimized original framework. Furthermore, chimeric anti-HEWL VHHs containing different combinations of FRs from cAbBCII10 and the native VHH were also constructed and expressed in *N. benthamiana* in order to identify which FR or combinations of FRs are important for maintaining high protein accumulation and stability *in planta*. Lastly, the native and alternative framework VHHs were expressed in *E. coli* to (i) investigate whether differences in yield between VHHs with native and alternative frameworks were also evident in a prokaryotic expression system; and (ii) to compare yields between VHHs expressed in *N. benthamiana* and in *E. coli*.

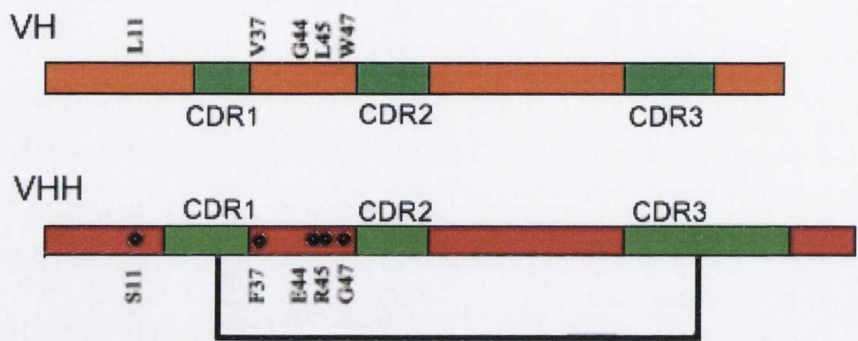


Figure 5.1 Schematic diagram showing the differences between VH and VHH based on sequence comparisons of cDNA clones. Both VH and VHH consists of 4 framework regions (FRs) and 3 CDRs. CDR1 and CDR3 are larger in a VHH than in a VH, and they are often connected by a disulfide bond (thick line). The hallmark amino acid substitutions in FR1 and FR2 are shown as black dots. The numbering refers to the position of the amino acid along the sequence according to the Kabat numbering (Kabat *et al.*, 1991; adapted from Muyldermans, 2001).



Figure 5.2 Ribbon representation of the X-ray structure of the VHH cAb-Lys3 (below) in complex with its antigen hen egg-white lysozyme (HEWL) (top). Lysozyme contains an α -helical domain (blue) and β -sheet domain (purple). For cAb-Lys3, the framework consists of a four-stranded β -sheet region (dark green) and a 5-stranded β -sheet region (site where the VL and VH interact in conventional Ab; light green). There are three CDR regions: CDR1 (red), CDR2 (orange) and CDR3 (yellow). The disulphide bond between CDR1 and CDR3, and the Glu35 side chain of lysozyme are represented in ball-and-stick form (from Desmyter *et al.*, 1996).

5.2 Results

5.2.1 Design of anti-PSA and anti-RNase A VHH constructs

In the previous chapter, an increase in the accumulation and yield of an anti-HEWL VHH was observed when the original framework (as found in CLDSyn) was replaced (in CBESyn) by the alternative 'universal' VHH framework proposed by (Saerens *et al.*, 2005). In order to determine whether the framework effect was specific for the anti-HEWL antibody, or was more general, four additional VHH transgenes were constructed and their expression *in planta* investigated.

The four constructs were: (i) PSAHS, a plant codon-optimized version of cAb-PSA11 (Saerens *et al.*, 2005), encoding a camelid VHH fragment recognizing Human Prostate Specific Antigen (PSA); (ii) CPSHS, a codon-optimized version of cAbBCII10-P-P-P, which encodes the CDRs of cAb-PSA11 grafted into the 'universal' framework of cAbBCII10; (iii) RN5HS, a codon-optimized version of cAb-RN05 (Saerens *et al.*, 2005), encoding a camelid VHH fragment recognizing Bovine RNase A; and (iv) CRNHS, a codon-optimized version of cAbBCII10-R-R-R, encoding the CDRs of cAbRN05 grafted into the cAbBCII10 framework. The codon optimization procedure for each construct is described in Materials and Methods. The aim was to keep the codons for each VHH as close as possible to CBESynHS, the highly expressed anti-HEWL VHH containing the cAbBCII10 framework (Figure 5.3). Codons which differed from CBEHS were replaced with the most frequently used *N. benthaminana* codons obtained from the Codon Usage Database (Nakamura *et al.*, 2000) without, wherever possible, introducing undesirable sequences such as direct or indirect repeats, potential hairpin structures, candidate polyadenylation signals (Mogen *et al.*, 1992; Li and Hunt, 1995), destabilizing sequences, or intron splice site signals (Brown *et al.*, 2002; Simpson *et al.*, 2004).

	-----Framework 1----- ---CDR-H1--- ---Framework 2-
PSAHS	DVQLVESGGG SVQ PGGSLRLS CAAS GY RFS NLAVG WFRQAPGKER EVVA
RN5HS	DVQLVESGGG LVQ AGGSLRLS CAAS GY TYTYI YMG WFRQAPGKER EGVA
CLDSynHS	DVQL QAS GGG SVQ AGGSLRLS CAAS GY TIG PYCMG WFRQAPGKER EGVA
CBESynHS	DVQLVESGGG SVQ AGGSLRLS CTAS GY TIG PYCMG WFRQAPGGER EAVA
	-----CDR-H2----- -----Framework 3-----
PSAHS	VIFVGGIT YYDT *SVKG RFTVSRGNANNTVYLEMNSLKPEDTAM YYCAA
RN5HS	AMDSGGG GLY ADSVKG RFTISRD K GKNTVYLQMDSLKPEDTAT YYCAA
CLDSynHS	AINMGGG ITY ADSVKG RFTIS Q DNAKNTVYLLMNSL EP EDTAI YYCAA
CBESynHS	AINMGGG ITY ADSVKG RFTISRDNAKNTVT LQ MNNLKPEDTAI YYCAA
	-----CDR3----- -----Framework 4-----
PSAHS	NRIPD TDSRN PSAY RY ***** WGQGTQVTVSSGHHHHHHH
RN5HS	GGYELRD RTY G Q ***** WGQGTQVTVSSGHHHHHHH
CLDSynHS	DSTIYASY YEC GHGLSTGGYGYDS WGQGTQVTVSSGHHHHHHH
CBESynHS	DSTIYASY YEC GHGLSTGGYGYDS WGQGTQVTVSSGHHHHHHH

Figure 5.3 Comparison of the amino acid sequences of His-tagged anti-PSA, anti-RNase A and anti-HEWL VHHs. The CDRs of VHHs in the first 3 rows are shown in their native framework context i.e. PSAHS (anti-PSA, 1st row); RN5HS (anti-RNase A, 2nd row); CLDSynHS (anti-HEWL, 3rd row). CBESynHS (4th row) contains the anti-HEWL CDRs of CLDSynHS in the universal framework context. Amino acid changes are shown in grey.

5.2.2 Expression and purification of anti-PSA and anti-RNase A VHHs

The anti-PSA and anti-RNase VHH constructs in their original and 'universal' framework contexts were cloned into the 3' provector pICH11599 (Santi *et al.*, 2006) and delivered into *N. benthamiana* leaves with the other required provectors via *Agrobacterium* infiltration. It was found that the anti-PSA VHHs with the original and the 'universal' frameworks (labelled PSAHS and CPSHS, respectively) accumulated highly in the cytoplasm compared to relatively low accumulation when targeted to the apoplast or the chloroplast (Figure 5.4a). In contrast, the anti-RNase A VHHs with the original and 'universal' frameworks (labelled RN5HS and CRNHS, respectively) accumulated highly when targeted to the apoplast, compared to the cytoplasm or chloroplast (Figure 5.4b).

In terms of yield, it was found that PSAHS had an estimated yield of less than 0.01 mg/g fresh weight while approximately 0.01 mg/g fresh weight was obtained for CPSHS. For the anti-RNase A VHH, RN5HS yielded about 0.02 mg/g fresh weight while CRNHS had a yield of about 0.05 mg/g fresh weight. Thus the framework-changed versions gave higher yields than their original framework counterparts.

7 dpi leaves expressing the anti-PSA and anti-RNase A VHH were harvested and the corresponding VHHs purified using Nickel-NTA columns (Figure 5.5). A yield of up to 0.02 mg/g fresh weight was obtained for RN5HS. The yield increased to 0.05 mg/g fresh weight for CRNHS (Table 5.1). Insufficient protein was obtained for subsequent biochemical analysis in the case of the anti-PSA VHHs (PSAHS and CPSHS) despite efforts to optimize protein purification such as increasing initial leaf material to 10 g and increasing the column volume. However, based on the intensity of the protein bands on the PAGE gel, the yield of PSAHS was estimated to be less than 0.01 mg/g fresh weight while that of CPSHS was estimated to be around 0.01 mg/g fresh weight (Table 5.1).

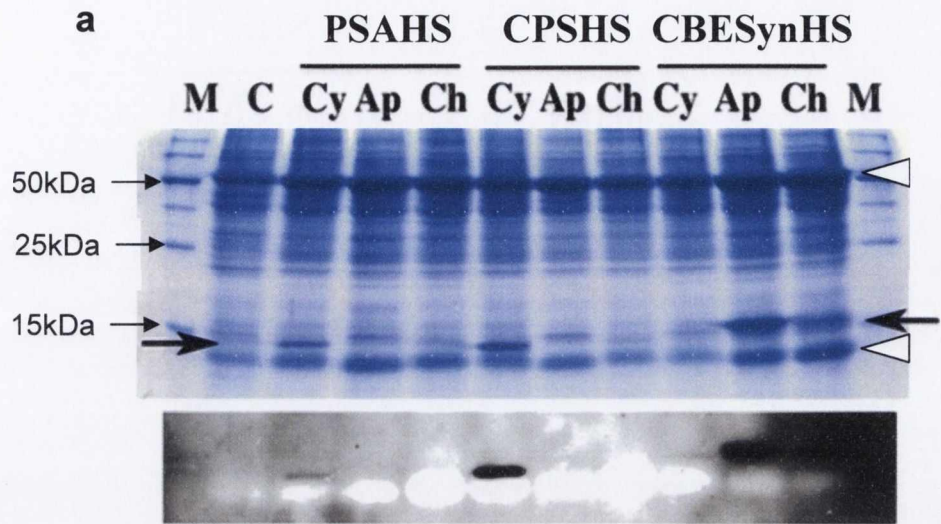


Figure 5.4 Accumulation of anti-PSA and anti-RNase A camel antibodies in various cellular compartments.

(a) PAGE gel (top) and Western blot (bottom) showing accumulation, at 7 dpi, of PSAHS, a His-tagged anti-PSA camel VHH with an original framework, and CPSHS, a derivative containing the 'universal' framework, in the cytoplasm (Cy), and when targeted to the apoplast (Ap) or chloroplasts (Ch). The anti-HEWL antibody CBESynHS was used as a positive control. 20 μ g of total soluble protein was loaded in each lane. C, Extract from leaves transformed with empty vector. M, Perfect ProteinTM Markers. White arrow heads indicate the large and small RUBISCO subunits (54 kDa and 14 kDa, respectively). The VHH is indicated by a black arrow head (approx. 15 - 16 kDa). PAGE gel and western blot is representative of the three experiments conducted.

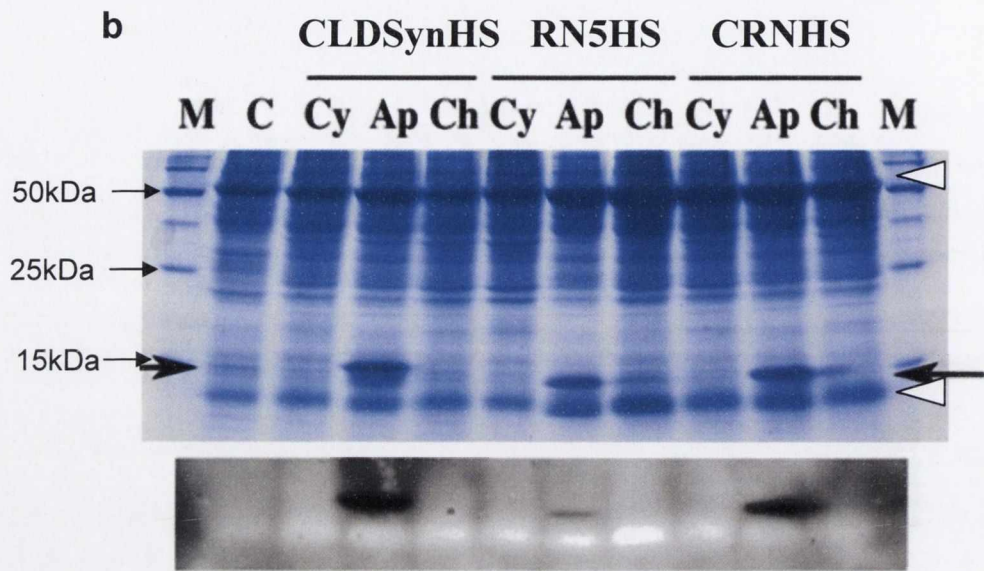


Figure 5.4 (continued) Accumulation of anti-PSA and anti-RNase A camel antibodies in different cellular compartments.

(b) PAGE gel (top) and Western blot (bottom) showing accumulation, at 7 dpi, of RN5HS, a His-tagged anti-RNase A camel VHH with an original framework, and CRNHS, a derivative containing the 'universal' framework, in the cytoplasm (Cy), and when targeted to the apoplast (Ap) or chloroplasts (Ch). The anti-HEWL antibody CLDSynHS was used as positive control. Other details are as in (a). The VHH of interest is approx. 15 – 16 kDa (black arrow). PAGE gel and western blot is representative of the three experiments conducted.

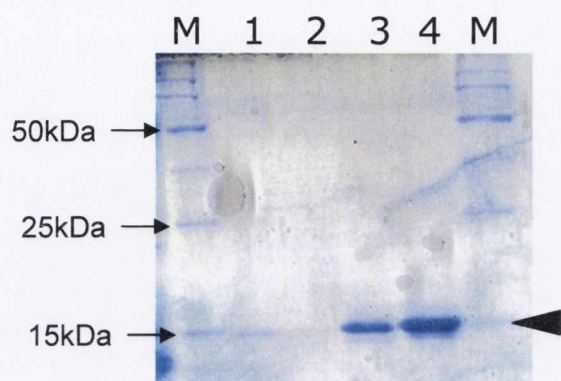


Figure 5.5 Purification and yield of His-tagged anti-PSA and anti-RNase A VHH. PAGE gel showing the purified products of the framework-changed anti-PSA VHH (CPSHS; Lane 1) and its original framework counterpart (PSAHS; Lane 2), as well as the original framework anti-RNase A VHH (RN5HS; Lane 3) and its framework-changed counterpart (CRNHS; Lane 4). Approximately 10 g of infiltrated leaf material was harvested for the purification of anti-PSA VHHs while 2-2.5 g of infiltrated leaf material was used for purification of the anti-RNase A VHHs. 5 μ L of purified protein was loaded on the gel. The black arrow head indicates the VHH of interest (approx. 15kDa). PAGE gel is representative of the two experiments conducted.

Table 5.1 Table showing estimated yields for the anti-PSA and anti-RNase A VHHs, the compartment with the highest accumulation, the mean estimated yield before purification ($\pm$ SD; obtained using density analysis) and the mean yield after purification (obtained using density analysis for anti-PSA VHH and as determined by Bradford Assay for anti-RNase A VHH). Data collected from three leaf samples from three independent experiments (total sample size=9).

Antibodies	Compartment with highest accumulation	Yield before purification* (mg/g fresh weight)	Yield after purification [†] (mg/g fresh weight)
PSAHS	Cytoplasm	<0.01	<0.01 [*]
CPSHS	Cytoplasm	$\approx$ 0.01	$\approx$ 0.01 [*]
RN5HS	Apoplast	0.02 $\pm$ 0.01	$\approx$ 0.02 [†]
CRNHS	Apoplast	0.05 $\pm$ 0.01	$\approx$ 0.05 [†]

*Protein concentration estimated by comparing protein bands on PAGE gel to Perfect ProteinTM Markers (Novagen, USA) using Multi Guage software (FujiFilm, Japan)

[†]Protein concentration determined using Bradford assay

5.2.3 Expression of anti-PSA and anti-RNase VHH constructs over time

The accumulation of the original framework and alternative framework anti-PSA and anti-RNase A VHHs over time was also investigated. Leaves at 0, 3, 5, 7, 9 and 11 dpi were harvested for analysis. It was found that PSAHS accumulated in the cytoplasm to levels detectable by PAGE and Western blot at 5 to 9 dpi while in the case of CPSHS, detectable amounts of protein could be seen at 5 dpi and the levels stayed the same until 11 dpi (Figure 5.6a). On the other hand, RN5HS accumulation when targeted to the apoplast was detected on PAGE gels and by Western blot in 7 dpi leaves but not in leaves harvested before or after (Figure 5.6b). The accumulation of CRNHS however started at 5 dpi, continued to accumulate until 7-9 dpi, before decreasing after 11 dpi.

In the case of both PSAHS and CPSHS, leaves turned progressively more necrotic which was reflected in the decreased levels of the Rubisco large and small subunits on Coomassie Blue-stained gels. After 13 dpi, the protein extracted from leaves was severely degraded. However, the level of necrosis observed in leaves expressing RN5HS and CRNHS was less severe, although the level of protein accumulation for both these VHHs was higher than for PSAHS and CPSHS. This suggests that the anti-PSA VHHs might be more toxic for plant cells than the anti-RNase A VHHs.

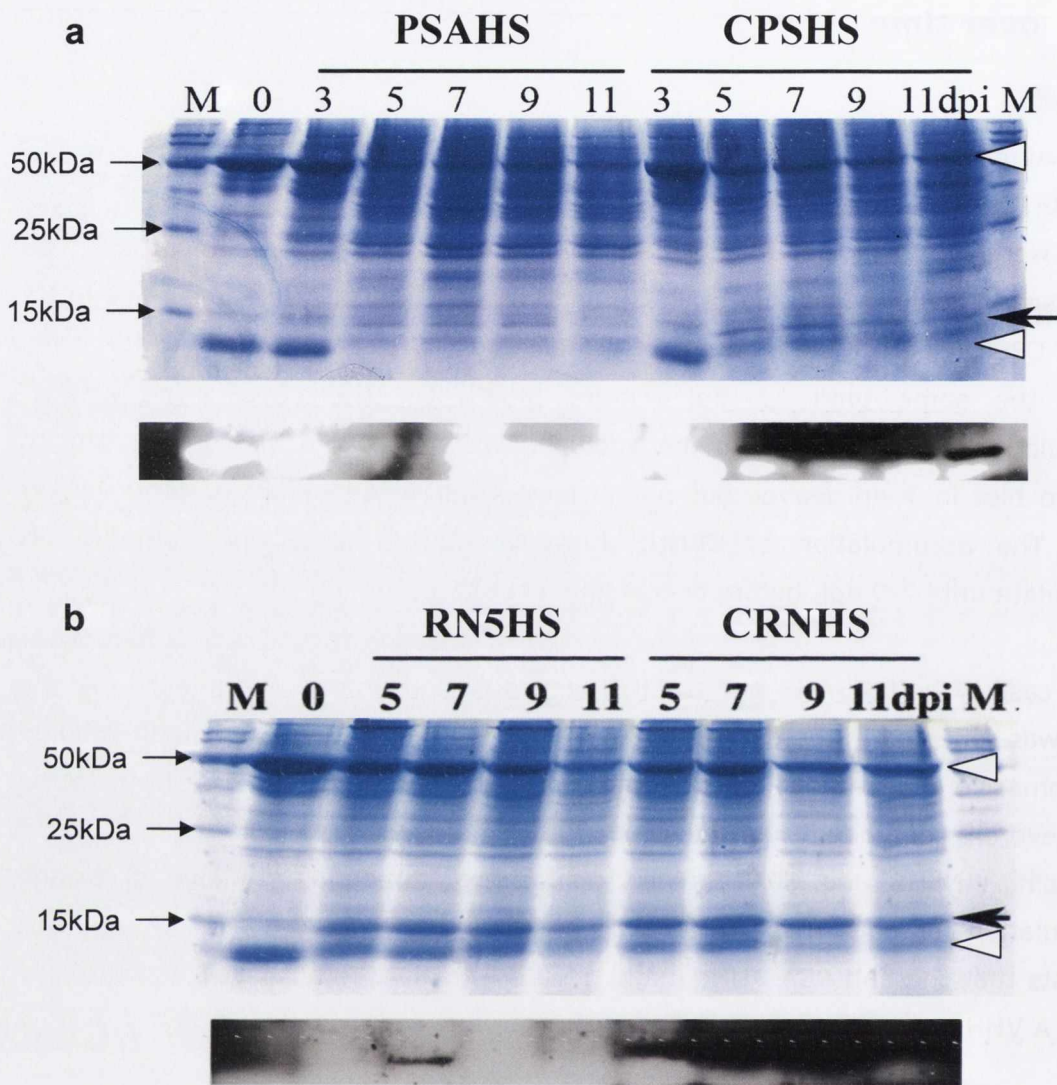


Figure 5.6 Accumulation of anti-PSA VHH and apoplast-targeted anti-RNase A VHH over time. PAGE gels (top) and Western blot (bottom) showing (a) accumulation of the cytoplasmically expressed His-tagged anti-PSA VHHs (PSAHS and CPSHS), and (b) the apoplast-targeted anti-RNase A VHHs (RN5HS and CRNHS) from 0 to 11 dpi at two-day intervals. 20 μ g of total soluble protein was loaded in each lane. M, Perfect ProteinTM Markers (M). White arrow heads indicate the large and small RUBISCO subunits (54 kDa and 14 kDa, respectively). A black arrow head indicates the VHH of interest (approx. 15 to 16 kDa). PAGE gel and western blot is representative of the three experiments conducted.

5.2.4 Characterisation of anti-RNase A VHHs

The specificities of the purified anti-RNase A VHHs were determined by direct ELISA with the antigen-bound VHH detected using HRP-conjugated Ni-NTA which detects the His-tags. It was found that both VHHs, with the original framework and the alternative framework, bound their antigen bovine RNase A, and not other unrelated antigens such as HEWL (Figure 5.7). Moreover, the dissociation constant (K_D) for RN5HS and CRNHS as determined by competitive ELISA (Friguet *et al.*, 1985) was 0.4 ± 0.25 nM and 1.56 ± 0.53 nM respectively (based on 3 repeats; Figure 5.8). Saerens *et al.* (2005) reported an increase in the dissociation constant (K_D) from 0.46 nM for cAb-RN05 (RN5) to 1.64 nM for cBCII10-R-R-R (CRN).

The thermal stability of RN5HS and CRNHS was also investigated. Both VHHs were incubated for 20 minutes at different temperatures, re-equilibrated to room temperature for 5 minutes and residual binding assayed by direct ELISA. It was found that both antibodies retained their functionality at temperatures up to 40°C. However, it was found that the activity of CRNHS decreased by 50% at 50°C while CRNHS still retained almost all of its activity (Figure 5.9). Saerens *et al.* (2005) also reported that the conformational stability of the original framework RN5HS VHH (cAB-RN05) was slightly higher than for CRN (cBCII10-R-R-R). After 70°C, no activity was detected for both RN5HS and CRNHS.

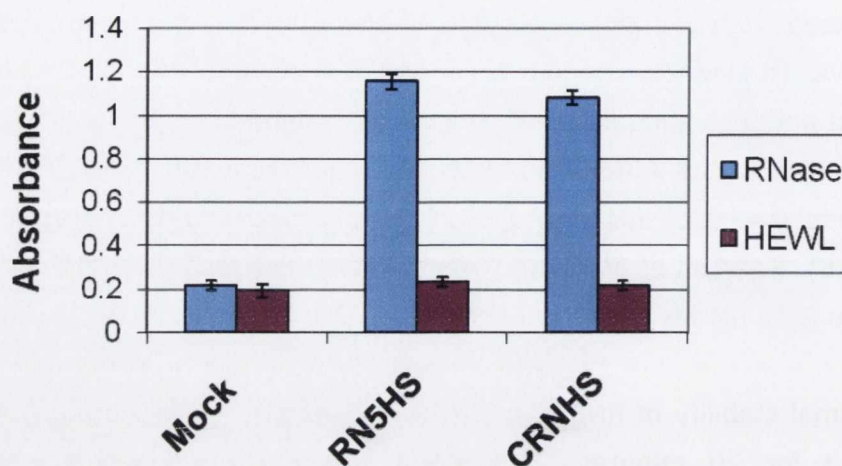


Figure 5.7 Specificity of the anti-RNase A camel antibodies as determined by ELISA. Both the VHH with the original framework (RN5HS) and the alternative framework (CRNHS) recognized RNase A (■) but not the unrelated antigen HEWL (■). Each point represents the mean of three experiments. Bars indicate 2 standard deviations. 'Mock' represents a control experiment where no nanobody was introduced.

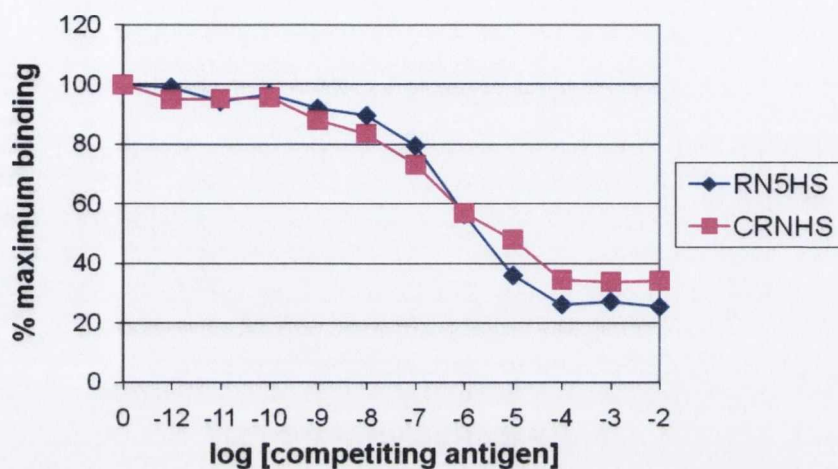


Figure 5.8 Affinity of His-tagged anti-RNase VHH with original framework (RN5HS; $\blacklozenge$) and alternative framework (CRNHS; $\blacksquare$) measured according to Friguet *et al.* (1985). VHHs diluted to a concentration predetermined by direct ELISA (about 1×10^{-8} M) were incubated with increasing concentration (1×10^{-12} M to 1×10^{-2} M) of antigen (in this case RNase A). The mixture was then incubated in wells coated with a fixed amount of antigen. Free antibodies not previously bound to the free antigen will bind to the antigen coated on the plate and these antibodies were detected using Ni-NTA probes bound to HRP which detects the 6xHis-tag on the VHH. The resulting absorbance values were plotted according to Friguet *et al.* (1985).

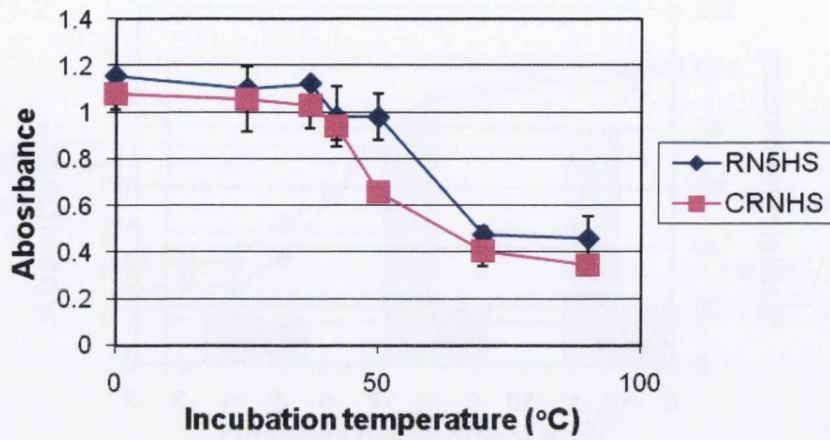


Figure 5.9 Comparison of thermal stability of RN5HS (◆) and CRNHS (■). Approximately 0.1 ng/mL of each antibody was incubated at increasing temperatures for 20 minutes, reequilibrated to room temperature and assayed using ELISA. Each point represents the mean of three experiments. Bars indicate 2 standard deviations (SDs).

5.2.5 Construction of chimeric anti-HEWL VHHs containing combinations of Framework Regions from CLDSyn and CBESyn

Previous sections have shown that VHHs containing the proposed 'universal' framework from cAbBCII10 accumulate to higher levels *in planta*, compared with their original framework counterparts. We therefore set out to determine if specific sequences in the cAbBCII10 framework were responsible for the improvement in VHH accumulation or whether other factors influenced yield. This was investigated by creating chimeric anti-HEWL VHHs in which framework regions were swapped between a VHH with an original framework and the 'universal' cAbBCII10 framework. CLDSyn and CBESyn were chosen for these experiments. They possess identical CDRs recognizing HEWL, but CLDSyn contains an original framework while CBESyn contains the 'universal' cAbBCII10 framework. CLDSyn and CBESyn differ, therefore, at only 10 amino acid positions: 3 in the first framework (FR1), 2 in the second (FR2), and 5 in the third framework (FR3). The fourth framework (FR4) is identical for both VHHs (Figure 5.10).

In order to construct the chimeric VHHs, 4 fragments were generated by PCR using forward primers containing nucleotides specific to FR2 of either CLDSyn or CBESyn (as shown by the region covered by the right-pointing blue or red arrow in Figure 5.10), a common reverse primer located in the 6XHis-tag region, and either CLDSyn or CBESyn as the template. The combinations of forward primers and the VHH templates are detailed in Section 2.3.6. The resulting fragments possess FR2 and FR3 of either CLDSyn, CBESyn or a combination of both (Figure 5.11). The digested fragments were ligated into vectors containing the FR1 of either CLDSyn or CBESyn to create 6 chimeras which contain different permutations of FR1, FR2 and FR3 (Figure 5.12). The constructs were named DBB, DBD, DDB, BDD, BBD and BDB according to the origins of their FR1, FR2 and FR3, respectively, where 'D' denotes FRs derived from CLDSyn, and 'B' denotes FRs derived from CBESyn.

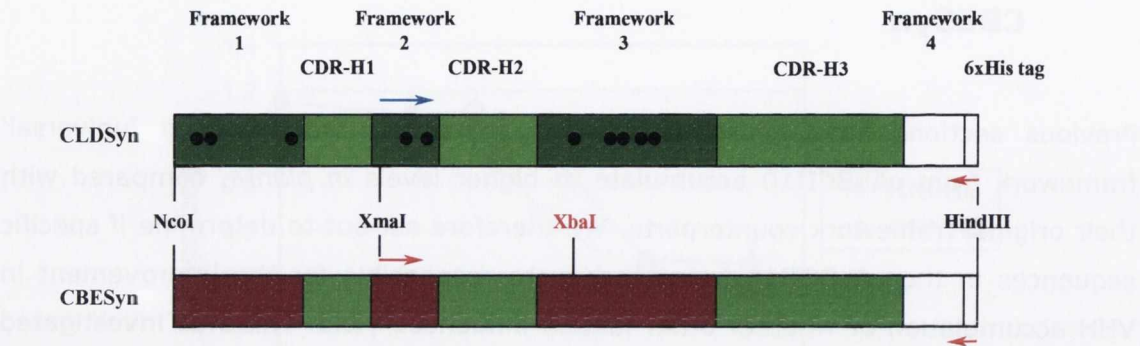


Figure 5.10 Relative positions of the nucleotide differences between the anti-HEWL VHHs CLDSyn and CBESyn. Each VHH contains four framework regions and three identical CDRs (H1-H3) recognizing HEWL. CLDSyn contains the original framework while CBESyn contains the 'universal' cABBCII10-derived framework. Black dots represent the relative positions of codons encoding amino acids that differ between CLDSyn and CBESyn. Arrows represent PCR primers (Blue and red for the forward primer unique to CLDSyn and CBESyn, respectively; the common reverse primer is indicated by a leftward-pointing red arrow) Restriction sites used for construction of the chimeras are also shown.

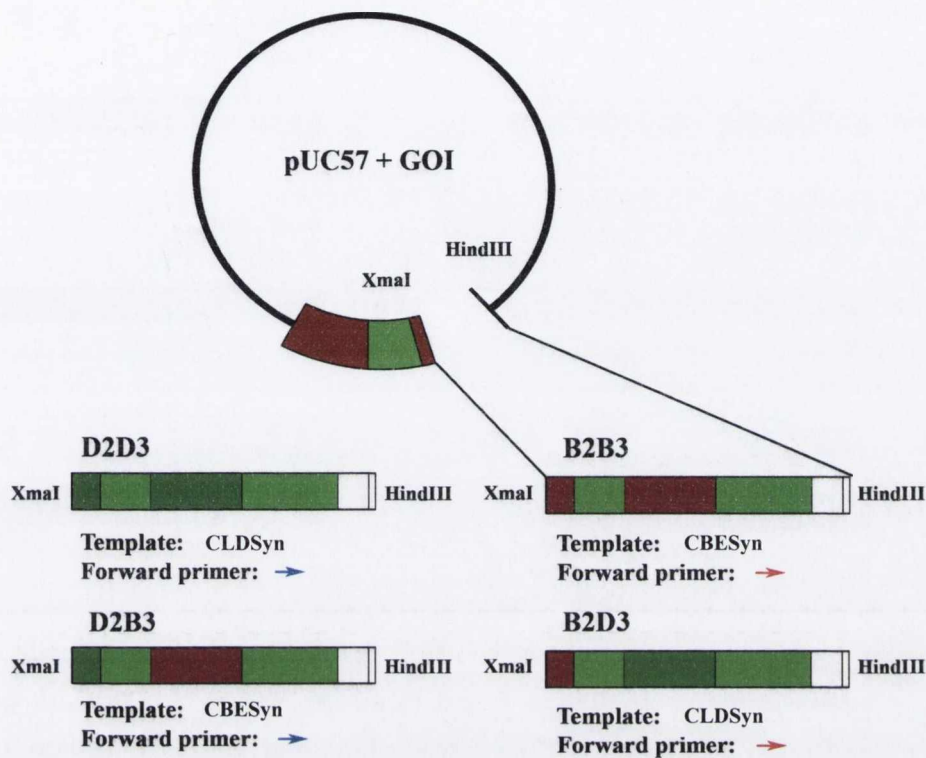


Figure 5.11 Combinations of forward primers and VHH templates used to generate fragments with different combinations of the FR2 and FR3 regions. The resulting fragments were subsequently ligated into vectors containing FR1 of either CLDSyn or CBESyn using the XmaI and HindIII restriction sites. The constructs were named according to the origins of their FR2 and FR3, respectively. 'D' denotes FRs originating from CLDSyn (Dark green); 'B' denotes FRs originating from CBESyn (Red). The numbers denote the specific FR e.g. the fragment D2B3 contains the FR2 region of CLDSyn and FR3 region of CBESyn. All fragments contain the CDR2 and CDR3 regions recognizing HEWL (light green).

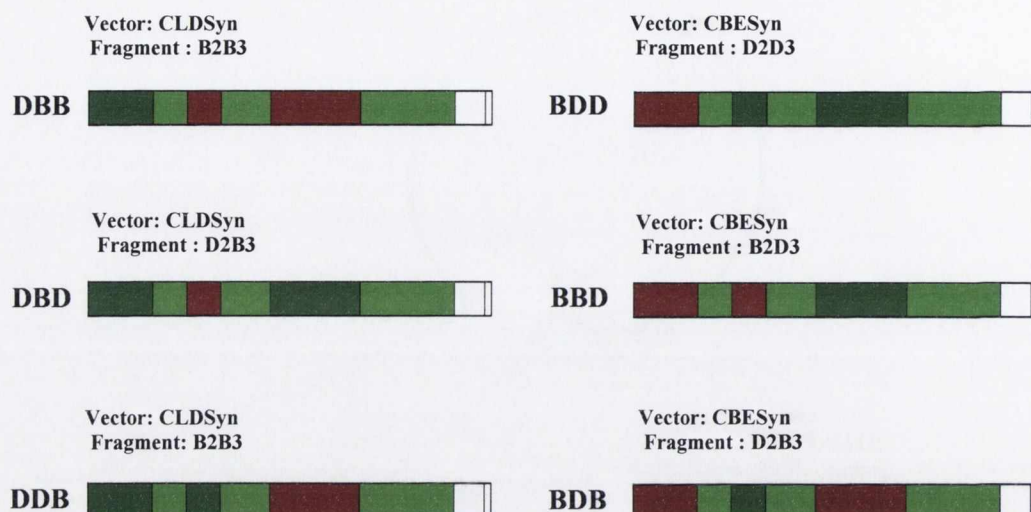


Figure 5.12 Structure of the 6 chimeric VHH genes. The chimeras encode different permutations of FR1, FR2 and FR3 from CLDSyn ('D'; Dark green) or CBESyn ('B'; Red). The vector contained the FR1 region of either CLDSyn or CBESyn (3 constructs each) while the fragments contained the FR2 and FR3 regions of CLDSyn, CBESyn or a combination of both. The constructs were named DBB, DBD, DDB, BDD, BBD and BDB according to the origins of their FR1, FR2 and FR3, respectively. (For example, DBB has FR1 region of CLDSyn as well as the FR2 and FR3 region of CBESyn). All fragments contain the CDR2 and CDR3 regions recognizing HEWL (light green).

5.2.6 Expression and purification of chimeric anti-HEWL VHHs

Constructs encoding the 6xHis-tagged anti-HEWL VHH chimeras were cloned into the 3' provector pICH11599 and were delivered (with the appropriate 5' provector) into *N. benthamiana* via *Agrobacterium* infiltration. It was found that all of the chimeras accumulated highly when targeted to the apoplast, compared with virtually no accumulation in the cytoplasm (Figure 5.13; cytoplasmic accumulation data not shown). Compared to a baseline of 0.24 mg/g fresh weight for CLDSyn and 0.39 mg/g fresh weight for CBESyn, VHHs that contained FR2 of CBESyn and FR3 of CLDSyn (DBD and BBD) had relatively lower yield with BBD having the lowest yield of 0.10 mg/g fresh weight followed by 0.20 mg/g fresh weight for DBD (Table 5.2). Furthermore, additional bands just below the expected band of the correct molecular weight were present in some constructs (e.g. DDB). The identities of these bands were not clear and may represent a degradation product (Figure 5.13).

The chimeras were grouped into two clusters where each cluster was made up of the parent nanobody (CLDSyn or CBESyn) and the chimeras where the FR1, FR2 or FR3 of the parent nanobody was replaced. It was found that when the individual FR of CLDSyn were swapped out, there was no significant drop in yield (Table 5.2). No significant drop in yields were also observed when FR1 was swapped out of CBESyn. However, when FR2 was replaced (in the case of BDB), there was a decrease of 0.15mg/g fresh weight. The most severe drop was when FR3 of CBESyn was replaced, with a decrease of 0.29mg/g fresh weight compared to CBESyn (Table 5.2).

Leaves expressing the 6 chimeras and the parental VHHs (CLDSyn and CBESyn) were harvested at 7 dpi and the VHHs purified using Nickel-NTA resin. VHH yields obtained after purification ranged from an estimated 0.08 mg/g fresh weight to 0.12 mg/g fresh weight (Table 5.2 and Figure 5.14). The relatively low recovery of the VHHs after purification might be due to the use of a batch purification procedure with Nickel-NTA slurry on this occasion, instead of Nickel-NTA columns.

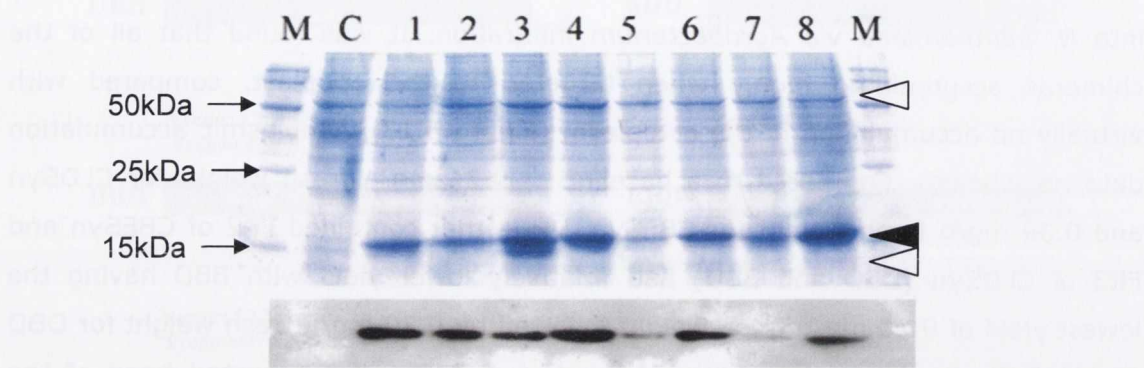


Figure 5.13 PAGE gel (top) and Western blot (bottom) showing accumulation of the six different anti-HEWL camel antibody chimeras DBB, DBD, DDB, BDD, BBD and BDB (Lanes 1-6) when targeted to the apoplast at 7 dpi. C, leaves transformed with empty vector. Lanes 7 – 8: CLDSyn and CBESyn, respectively. 10 μ g of total soluble protein was loaded in each lane. M, Perfect ProteinTM Markers. White arrow heads indicate the large and small RUBISCO subunit (54kDa and 14kDa), respectively; black arrow head indicates the VHH of interest (approx. 16kDa). PAGE gel and western blot is representative of the three experiments conducted.

Table 5.2 Table showing, for the 6 chimeric VHHs as well as the parental controls CLDSyn and CBESyn, the compartment with the highest accumulation, , the mean estimated yield before purification ($\pm$ SD; obtained using density analysis) and the mean yield after purification ($\pm$ SD; as determined by Bradford Assay). Data collected from three leaf samples from three independent experiments (total sample size=9).

Antibodies	Compartment with highest accumulation	Yield before purification* (mg/g fresh weight)	Yield after purification [†] (mg/g fresh weight)
CLDSyn	Apoplast	0.24 $\pm$ 0.07	0.13 $\pm$ 0.007
BDD	Apoplast	0.26 $\pm$ 0.07	0.11 $\pm$ 0.004
DBD	Apoplast	0.22 $\pm$ 0.09	0.11 $\pm$ 0.004
DDB	Apoplast	0.28 $\pm$ 0.02	0.12 $\pm$ 0.005
CBESyn	Apoplast	0.39 $\pm$ 0.08	0.20 $\pm$ 0.004
DBB	Apoplast	0.30 $\pm$ 0.11	0.12 $\pm$ 0.005
BDB	Apoplast	0.24 $\pm$ 0.06	0.10 $\pm$ 0.007
BBD	Apoplast	0.10 $\pm$ 0.04	0.07 $\pm$ 0.007

*Protein concentration estimated by comparing protein bands on PAGE gel to Perfect ProteinTM Markers (Novagen, USA) using Multi Guage software (FujiFilm, Japan)

[†]Protein concentration determined using Bradford assay



Figure 5.14 Purification and yield of 6 anti-HEWL VHH chimeras. PAGE gels showing the purified products of the 6 His-tagged anti-HEWL VHH chimera constructs (Lanes 4-9) compared to the mock infiltration (Lane 1) as well as positive controls CLDSyn and CBESyn (Lanes 3 and 4 respectively). Approximately 2.5-3 g of infiltrated leaf material was harvested for each construct. 5 μ L of purified protein was loaded on the gel. The black arrow head indicates the VHH of interest (approx. 16kDa). PAGE gel is representative of the two experiments conducted.

5.2.7 Characterization of anti-HEWL chimeric VHHs

The specificities of the purified anti-HEWL chimeric VHHs were determined by direct ELISA with antigen-bound VHH detected using HRP-conjugated Ni-NTA to detect the His-tags. It was found that all chimeras bound their antigen HEWL and not an unrelated but structural similar antigen α -lactalbumin (Figure 5.15). The dissociation constant (K_D) was also determined by competitive ELISA (Friguet *et al.*, 1985; Figure 5.16). It was found that the K_D of the 6 chimeras lay between the K_D of CLDSynHS and CBESynHS (Table 5.3), with most of the chimeras having a K_D similar to CLDSynHS (approximately 2 nM) except for DDB, which had a K_D of 20.66 nM.

Furthermore, the thermal stability of all chimeras was investigated. All VHHs were incubated for 20 min at different temperatures, re-equilibrated to room temperature for 5 min and residual binding activity assayed by direct ELISA. It was found that the thermal stability of DBB, DBD, DDB and BBD was similar to CLDSynHS and that they retained most of their antigen-binding activities even up to 90°C (Figure 5.17). The chimera BDD retained almost all its activity until 70°C. At 90°C, half of the antigen-binding capacity of BDD was lost. BDB on the other hand was the least stable of the 6 chimeras. It lost its activity far more rapidly compared to CLDSynHS and all other chimeras. Although BDB activity was slightly higher than that of CBESynHS at 50°C, it had lost almost all its antigen-binding capabilities following exposure to 90°C (Figure 5.17).

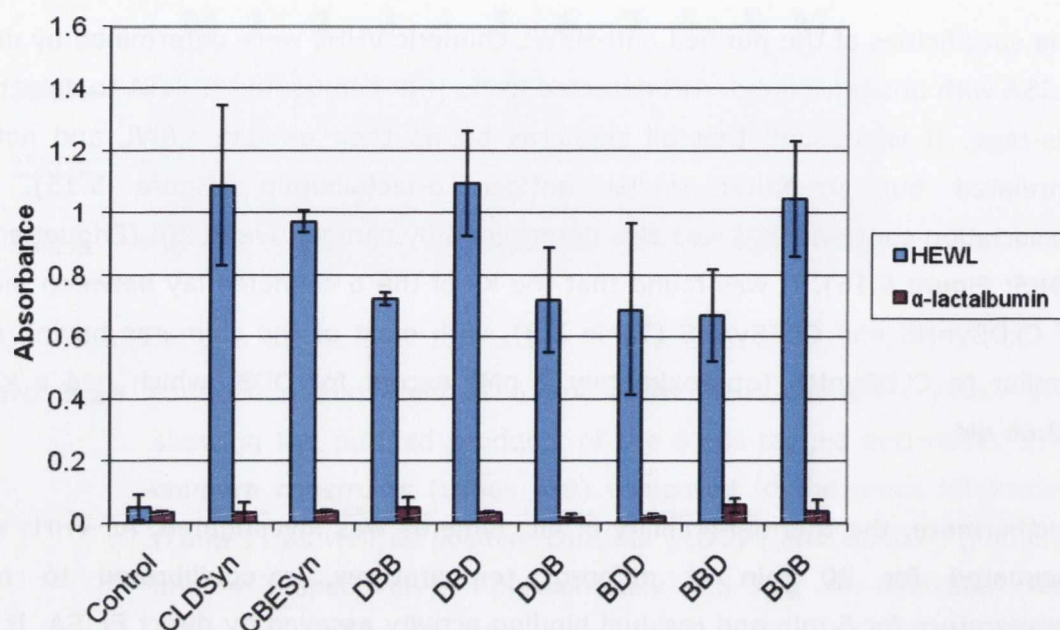


Figure 5.15 Specificity of the anti-HEWL camel antibodies chimeras as determined by ELISA. All of the chimeric VHHs, along with the positive controls CLDSyn and CBESyn recognized HEWL (■), but not the unrelated antigen α -lactalbumin (■). Each point represents the mean of three independent experiments. Bars indicate 2 standard deviations (SDs). 'Control' represents a control experiment where no nanobody was introduced.

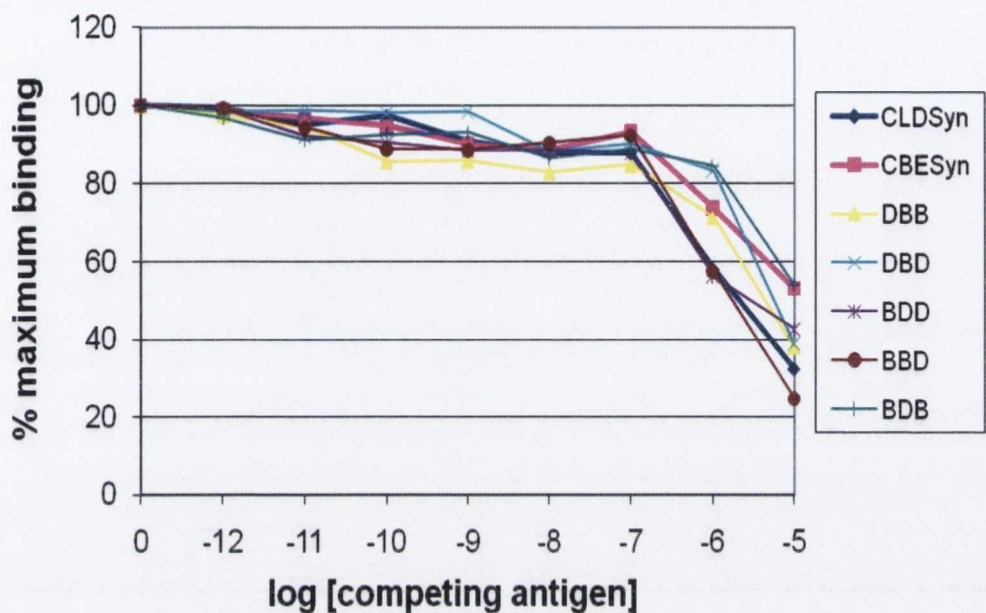


Figure 5.16 Affinity of the anti-HEWL camel antibodies chimeras measured according to Friguet *et al.* (1985). VHHs diluted to a concentration predetermined by direct ELISA (about 5×10^{-8} M) were incubated with increasing concentration (1×10^{-12} M to 1×10^{-5} M) of the antigen HEWL. The mixture was then incubated in wells coated with a fixed amount of antigen. Free antibodies not previously bound to the free antigen bind to the antigen coated on the plate and these antibodies were detected using an Ni-NTA-HRP probe that recognizes the 6xHis tag. The resulting absorbance values were plotted according to Friguet *et al.* (1985).

Table 5.3 Table showing, for the 6 chimeric VHHs as well as the parental controls CLDSyn and CBESyn, their respective dissociation constant (K_D) as determined by competitive ELISA (Friguet *et al.*, 1985)

Antibodies	Dissociation Constant K_D (nM)
CLDSyn	4.73
BDD	5.00
DBD	4.52
DDB	20.66
CBESyn	46.30
DBB	6.17
BDB	1.42
BBD	2.36

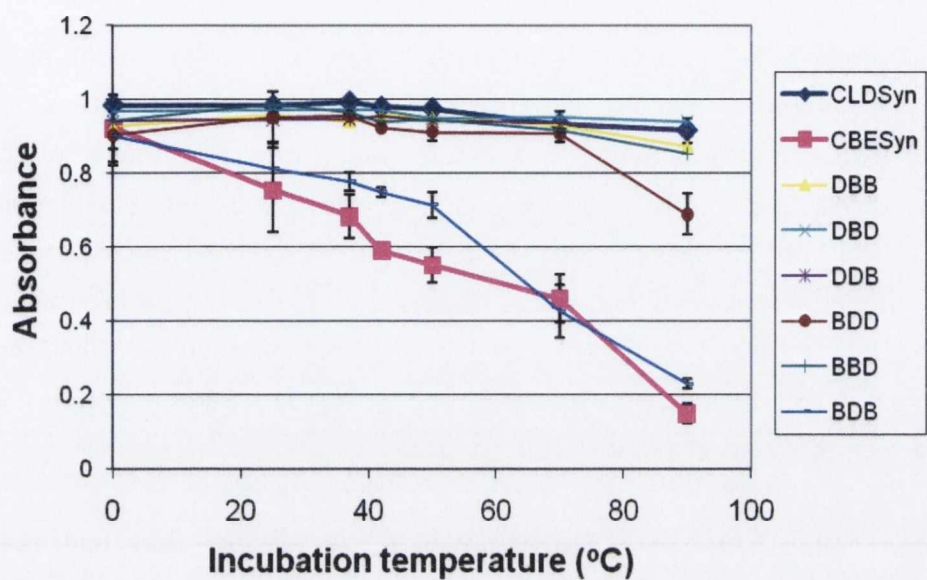


Figure 5.17 Comparison of thermal stability of the anti-HEWL VHH chimeras. Approximately 0.1 ng/mL of each antibody was incubated at increasing temperature for 20 minutes, re-equilibrated to room temperature and assayed using ELISA. Each point represents the mean of three independent experiments. Bars indicate 2 standard deviations.

5.2.8 Expression of an anti-HEWL and anti-RNase A VHH in *Escherichia coli*: a comparison

To compare the VHH production levels and pattern between plant and bacterial expression systems, two types of VHH, anti-HEWL and anti-RNase A, in their original framework (CLDSyn and RN5) as well as in the 'universal' cAbBCII10 framework (CBESyn and CRN), were cloned into the bacterial expression vector pET-22b(+) and introduced into *E. coli*. The pET-22b(+) vector possesses an upstream pelB secretion signal which targets the VHH to the periplasmic space (Lei *et al.*, 1987). Similar vectors encoding the pelB signal, such as pHEN1 and pHEN6 have been used by (Arbabi-Ghahroudi *et al.*, 1997; Saerens *et al.*, 2005) for routine expression of VHHs in *E. coli*.

Transgene expression was induced in bacterial cultures that had reached the exponential growth phase by addition of isopropyl- β -D-thio-galactoside (IPTG). This induces transcription of T7 RNA polymerase which in turn binds to and activates transcription from a T7 promoter located upstream of the VHH. Culture samples were collected before the addition of IPTG (T=0) as well as 1, 2 and 3 hours after induction and whole cell extracts were prepared by boiling cells harvested by centrifugation in SDS-PAGE sample buffer for 10 minutes.

It was found that all four VHHs accumulated at a high level. The estimated yield for the anti-HEWL nanobodies were approximately 5 mg/L for CLDSynHS and 8.5 mg/L for CBESynHS (Table 5.4). However, apart from the expected band with the correct molecular weight, an additional band that was about 1 kDa heavier than the expected band was also detected on Western blots (Figure 5.18a). CLDSyn and CBESyn accumulation was detected at 1 hour after induction, but the levels of CLDSyn were substantially greater than that of CBESyn. CLDSyn accumulation peaked at 2 hours post-induction with levels dropping slightly at 3 hours. In contrast, CBESyn accumulation increased over time, with the highest accumulation being detected at 3 hours post-induction (Figure 5.18a).

Accumulation of the anti-RNase A VHHs (RN5 and CRN) was substantially lower than observed for the anti-HEWL VHHs. The anti-RNase A VHH with the native framework (RN5HS) had a yield of approximately 0.5 mg/L (Table 5.4). On the other hand, the levels of the framework-changed version CRNHS were slightly

higher than that of RN5HS at approximately 1.3 mg/L for CRNHS. For both RN5HS and CRNHS, protein accumulation increased over time with the highest accumulation detected at 3 hours post-induction (Figure 5.18b). Furthermore, the additional band that was about 1 kDa heavier than the expected band as seen in the CLDSynHS and CBESynHS blots were also detected on the RN5HS and CRNHS blots.

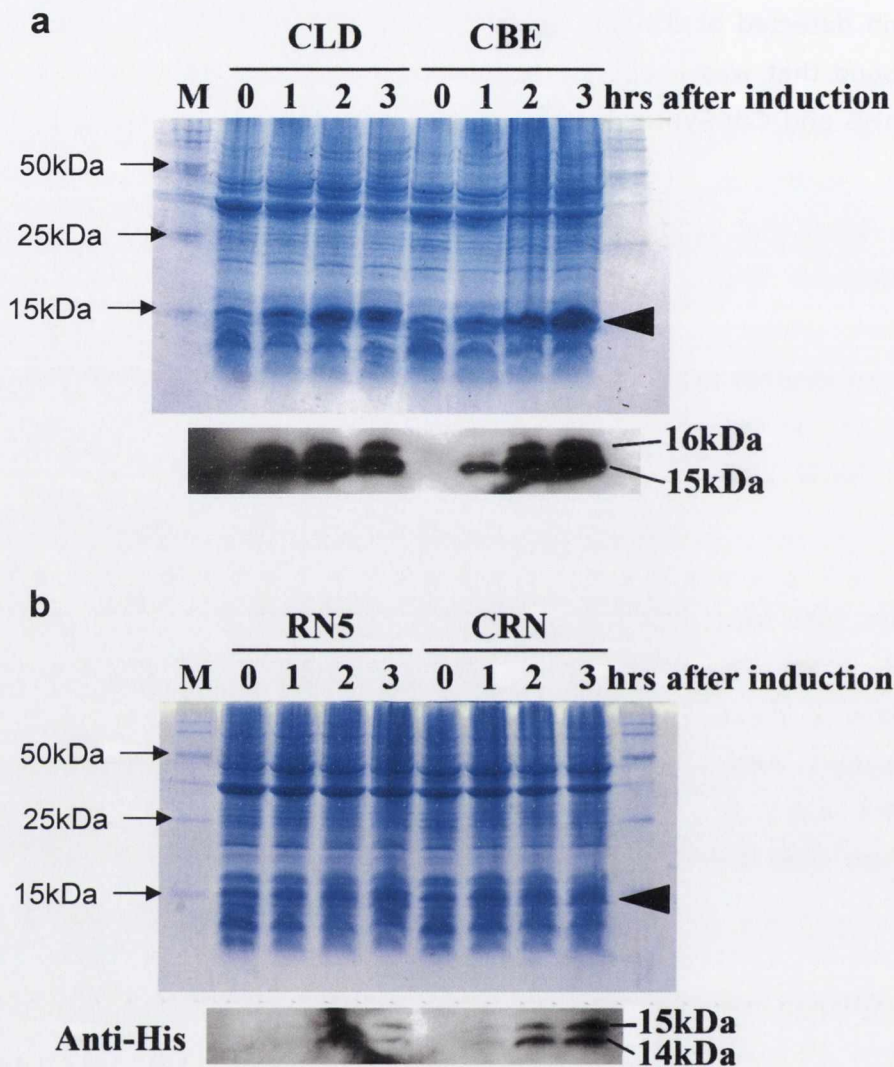


Figure 5.18 Accumulation of anti-HEWL and anti-RNase A VHH in *E. coli* over time. PAGE gels (top) and Western blot (bottom) showing accumulation of (a) His-tagged anti-HEWL VHH and (b) anti-RNase VHH from 0 to 3 hours post induction at hourly intervals. The VHHs were in their original framework (CLD and RN5) or in the alternative 'universal' cAbBCII10 framework. 50 μ g of total soluble protein was loaded in each lane. M, Perfect Protein™ Markers. A black arrow head indicates the VHH of interest (approx. 14 to 15kDa).

Table 5.4 Table showing a comparison of the estimated yield before purification (obtained using density analysis) for His-tagged CLDSyn, CBESyn, RN5 and CRN obtained from the *E. coli* expression system and from *N. benthamiana* plant infiltrated with the ICON system

Antibodies	Yield from <i>E. coli</i> expression system* (mg/L)	Yield from ICON expression system* (mg/kg fresh weight)
CLDSynHS	5.0	200
CBESynHS	8.5	350
RN5HS	0.5	20
CRNHS	1.3	50

* Protein concentration estimated by comparing protein bands on PAGE gel to Perfect ProteinTM Markers (Novagen, USA) using Multi Guage software (FujiFilm, Japan)

5.3 Discussion

5.3.1 The use of alternative Framework Regions from cAbBCII10 to increase the yield of Camel VHHs *in planta*

Previous experiments (Chapter 4) showed an increase in VHH accumulation when the CDRs of the anti-HEWL VHH were grafted onto the proposed 'universal' FRs of cAbBCII10 (Saerens *et al.*, 2005) compared to when the VHH was expressed with its native FRs. To determine whether the effect of the change in FRs was specific to anti-HEWLs VHHs or was applicable to plant-expressed VHHs in general, four additional VHH transgenes, two anti-PSA VHHs and two anti-RNase A VHHs, were constructed. In each pair, one codon-optimized VHH possessed the original native FRs while in the other these were replaced with the FRs of cAbBCII10. It was found that, for the two pairs of VHHs, the framework-changed versions gave a higher yield compared to their native framework counterparts. Furthermore, this yield difference between the VHHs with the native and 'universal' framework was also observed when the anti-HEWL and anti-RNase A VHHs were expressed in *E. coli*.

In the previous chapter, it was shown that the enhanced accumulation and stability of the framework-changed VHH was likely due to the different FRs which it contained. In general, the enhanced stability of VHHs compared to conventional VH is largely contributed by 4 hallmark substitutions at positions 37, 44, 45 and 47 (Kabat numbering; Kabat *et al.*, 1991) located in the FR2 region (Davies and Reichmann 1994; 1996; Conrath *et al.* 2005). However, Muyldermans *et al.* (1994) reported that amino acid substitution at position 47 is less well conserved. cAbBCII10 possessed the VHH hallmark substitutions at positions 37, 44 and 45. However, an Alanine (A) residue was present at position 47 instead of the hallmark Glycine (G) residue (Saerens *et al.*, 2005). On the other hand, all four hallmark residues were present in cAbLys3. This difference in amino acid might affect overall stability of the nanobody. Furthermore, the cAbBCII10 FRs were chosen by Saerens *et al.* (2005) as a 'universal' framework suitable for CDR loop-grafting applications. cAbBCII10 was chosen for its good expression level in *E. coli*, its ability to retain functionality even when two conserved disulphide bond-forming cysteines were removed and more importantly, its enhanced thermodynamic stability, which likely also accounts for the increased accumulation and improved resistance to proteolysis shown by VHHs with the cAbBCII10 framework in the present study.

In the present study, when the accumulation of 6 chimeric VHHs with different combinations of 'native' and 'universal' FR1, FR2 and FR3 were compared, it was found that they differed not only in their yield (VHHs with the FR2 of CBESyn and FR3 of CLDSyn had the lowest yield) but also in their susceptibility to proteases (the chimeric construct DDB showed a degradation pattern similar to CLDSyn). Furthermore, when the yields were tabulated to show the effects of the changes of individual framework compared to the parent molecule, it was found that replacing individual framework region of CLDSyn had minimal effect on the yield. On the other hand, swapping out the individual framework region of CBESyn had resulted in significant drop in yields, with the replacement of FR3 being most severe. This suggested that the alternative framework region may be inherently more stable or less susceptible to proteolytic degradation.

Another possible explanation for the improvement in yield is that the tertiary structure of the VHHs based on the 'universal' framework render them more stable in the cell and more resistant to proteases. When He *et al.* (2011) compared the main protein domains of recombinant proteins that were expressed in the mammary glands of transgenic animals with those of milk proteins, they found that proteins which shared high domain similarities with casein, the main component of milk proteins accumulated at a higher level compared to those that did not. Although there are no crystallographic data on the 6 anti-HEWL VHH chimeras constructed here, Saerens *et al.* (2005) have shown that the CDRs of the anti-HEWL VHH adopt slightly different conformations when they are grafted onto the alternative framework. Although the conformation of the CDR-H2 and -H3 loops were identical to those in the native anti-HEWL VHH, the N-terminal part of the CDR-H1 loop was slightly distorted, and the H4 loop in FR3 was presented in a totally different manner. It might, therefore, be possible that differences in tertiary structure, between VHHs with the native and alternative FRs, alter the way in which the VHHs interact with their subcellular environment. This in turn might determine whether they are processed properly or are targeted for degradation by the ER-associated degradation pathway (ERAD) (Vembar and Brodsky, 2008) and/or the ubiquitin-proteasome pathway (Ciechanover, 1998; Hellmann and Estelle, 2002), ultimately affecting their overall yield.

5.3.2 Comparison of Camel VHHs obtained from *E. coli* and transiently-transformed plants (*N. benthamiana*)

Camel VHHs have been expressed via many production platforms including *E. coli* (Arbabi-Ghahroudi *et al.*, 1997; Rahbarizadeh *et al.*, 2005), yeast (Frenken *et al.*, 2000; Rahbarizadeh *et al.*, 2006), and filamentous fungi (Joosten *et al.*, 2005). When the two anti-HEWL nanobodies CLDSynHS and CBESynHS were expressed in *E. coli* using the pET vector system (Novagen, USA), yields of approximately 5 mg/L and 8.5 mg/L were obtained. This was similar to the yields reported by Saerens *et al.* (2005) using pHEN6 vectors (Cortez-Retamozo *et al.*, 2004). However, while Saerens *et al.* (2005) reported a slightly lower yield (4.5 mg/L) for the framework-changed version of CBESynHS (or cAbBCII10-L-L-L in their report) compared to 5 mg/L obtained for the native VHH CLDSynHS (or cAbLys3), we obtained a slightly higher yield for CBESynHS compared to CLDSynHS. This discrepancy might be due to the fact that we were estimating protein concentration from crude extracts while the yields reported in Saerens *et al.* (2005) was the VHH concentration after purification. Another explanation might be that different expression vectors were used. For the anti-RNase A VHH on the other hand, Saerens *et al.* (2005) reported a higher yield for RN5HS and CRNHS (cAbRN05 and cAbBCII10-R-R-R in their report) of 2 mg/L and 10 mg/L respectively. This is a lot higher than what we obtained using the pET vector system (0.5 mg/L and 1.3 mg/L for RN5HS and CRNHS, respectively).

When we compare the *E. coli* and ICON expression system, more time was required to obtain the recombinant protein from Agro-infiltrated plants (7-11 days for the ICON system) compared with 2-3 days required in the case of *E. coli*. However, when we compared the yields obtained from *E. coli* to those obtained from ICON-infiltrated plants, it was found that the yields were much higher using the ICON system compared with the *E. coli* system. For example, in the case of CLDSynHS, 5 mg/L was obtained in both the present study and also by Saerens *et al.* (2005) in the *E. coli* expression system. However, the yield obtained from ICON-infiltrated *N. benthamiana* leaves was approximately 0.2 mg/g (or 200 mg/kg) fresh weight before purification; and was 0.13 mg/g fresh weight after purification. Even taking into account the amount of extraction buffer required for 1kg of leaf material compared to 1L of *E. coli*, 15 times more VHH will be obtained using the ICON system compared to the *E. coli* system. Moreover, it was also reported that proteins that are highly expressed in *E. coli* often precipitate or aggregate in insoluble

inclusion bodies which require refolding (Makrides, 1996; Misawa and Kumagai, 1999). Also, Gleba *et al.* (2007) reported the automation of the ICON expression platform using existing technologies that simplify the infiltration process.

Chapter 6. Investigating the use of different expression vector systems (ICON, pEAQ-HT, pJL-TRBO, pBIN2 and pBIN8) for transient VHH expression

6.1 Introduction

Over the years, numerous vectors for transient expression of recombinant proteins in plants have been developed. The simplest comprise the GOI, flanked by regulatory sequences i.e. a 5' promoter and 3' transcriptional terminator, inserted into the T-DNA of Ti plasmid (for example Popescu *et al.*, 2007). A disadvantage of this system is that the yield is very low and expression drops off after 5 days post-infiltration (dpi). Furthermore, Voinnet *et al.* (2003) have shown that post-transcriptional gene silencing (PTGS) is often induced in the plant after infiltration as a defence mechanism against *Agrobacterium* infection. This could be partially overcome by introducing an RNA silencing suppressor in addition to the GOI. However, this approach necessitates infiltrating relatively high concentrations of *Agrobacterium* cells into leaves, which in some cases can cause a hypersensitivity-induced necrotic reaction in the plant (Wroblewski *et al.*, 2005).

The most efficient transient expression vectors, to date, are based on genomic sequences derived from plant viruses such as the closely-related tobamoviruses Tobacco Mosaic Virus (TMV) and Turnip Vein-Clearing Virus (TVCV), Potato Virus X (PVX), and Cowpea Mosaic Virus (CPMV). These vectors are based either on a single full-length infectious copy of the viral genome (Dawson *et al.*, 1989; Donson *et al.*, 1991; Scholthof *et al.*, 1993; Scholthof *et al.*, 1996; Werner *et al.*, 2006) or on a so-called 'deconstructed' version of the viral genome. In the 'deconstructed virus' strategy, the viral genome is 'stripped down' to the genes required for the expression and amplification of transcripts encoding the GOI. The missing functions such as those initiating infection and DNA integration are provided using non-viral components. *Agrobacterium tumefaciens*, for example, is often used to deliver the vector into plant cells via T-DNA transfer. After incorporation of the T-DNA into the

plant nucleus, the transcribed RNA is able to self-replicate in the cytoplasm due to the presence of the viral replicase complex. These vectors, therefore, not only take advantage of the presence of RNA-dependent RNA replication functions of the plant virus to further amplify the primary RNA transcript initially generated via the host transcription machinery, but also make use of the efficient manner in which *Agrobacterium* infects and transfers T-DNA into the host cells.

The ICON Genetics system (Marillonnet *et al.*, 2004; Gleba *et al.*, 2007) has already been described in some detail (Please refer to Chapter 3). It has been used to express many different recombinant proteins at high level, including antibodies (Gils *et al.*, 2005; Santi *et al.*, 2006; Dorokhov *et al.*, 2007; Gleba *et al.*, 2007; Huang *et al.*, 2008; Webster *et al.*, 2009; Kalthoff *et al.*, 2010). Furthermore, because the ICON vectors are based on tobamovirus sequences, which do not compete with PVX for replication functions, both these systems can be used simultaneously to express multimeric proteins (Giritch *et al.*, 2006).

The TMV RNA-based Overexpression (TRBO) vector (Lindbo, 2007) is a second tobamovirus-based vector reconfigured for transient expression. The vector (pJL-TRBO) comprises an almost full-length cDNA copy of the TMV genome, minus the sequences encoding the coat protein (CP). The construct is transcriptionally driven by a CaMV35S promoter sequence. An advantage that the pJL-TRBO system has over the ICON system is that its use is more straightforward: it comprises a single vector molecule and involves the use of a single *Agrobacterium* strain, rather than the 3 *Agrobacterium* strains required by the ICON system.

pEAQ-HT is an especially interesting virus-based transient expression vector derived from the genome of the Cowpea Mosaic Virus (CPMV) (Sainsbury and Lomonosoff, 2008). The genome of wild-type CPMV consists of two RNAs that are encapsidated in different virus particles: RNA-1, which codes for the replication complex; and RNA-2, which codes for movement and structural proteins and is dependent on RNA-1 for its amplification *in vivo* (Liu *et al.*, 2005; Sainsbury *et al.*, 2010). Sainsbury and Lomonosoff (2008) found that high-level protein expression can be achieved (via transient expression) using sequences derived from RNA-2, without the need for replication functions provided by RNA-1. Amazingly, they found that removing an in-frame start codon (AUG161) upstream of the main translation initiation site on RNA-2 led to a huge increase in the accumulation of the GOI protein. This deletion also eliminated the need to maintain reading frame alignment between the start codon at position 161 and the main start codon at

position 512. The resulting modified sequence functions as a 5' hypertranslatable (HT) leader, which vastly increases the translational efficiency of the GOI (Sainsbury *et al.*, 2008; Sainsbury and Lomonossoff, 2008; Sainsbury *et al.*, 2009). A key feature of the pEAQ-HT system is the presence of a cassette expressing the RNA silencing suppressor protein P19, the presence of which increases GFP expression 7-fold, relative to its absence (Saxena *et al.*, 2010), but apparently without affecting mRNA accumulation (Sainsbury and Lomonossoff, 2008). Thus unlike the TMV-based systems (ICON genetics and pJL-TRBO), high-level protein production by the pEAQ-HT system does not require RdRP-mediated *in vivo* amplification of transcripts. Moreover, pEAQ-HT has been used to express large heteromeric proteins such as Immunoglobulin G (IgGs) (Sainsbury *et al.*, 2008; Sainsbury *et al.*, 2010).

In this chapter, in order to identify the optimal system in which to express VHHs *in planta*, we investigated expression of the anti-HEWL VHH CBESynHS via three different virus-based expression systems: (i) the ICON Genetics system (Marillonnet *et al.*, 2004; Gleba *et al.*, 2007), (ii) the pJL-TRBO system (Lindbo, 2007) and (iii) the CPMV-based pEAQ-HT system (Sainsbury and Lomonossoff, 2008). Moreover, since previous experiments had shown that expression of the anti-Prostate Specific Antigen (PSA) VHHs containing either the native Framework Regions (FRs) or the alternative 'universal' cABBCII10 FRs (named PSAHS and CPSHS respectively) via the ICON system resulted in low yields (Chapter 5; Section 5.2.2), we were especially interested in the possibility that yields of these VHHs might be improved by expression from pJL-TRBO or pEAQ-HT.

6.2 Results

6.2.1 Comparison of GFP expression from three virus-based transient expression vectors

The two tobamovirus-based expression systems i.e. the ICON Genetics system (Marillonnet *et al.*, 2004), and pJL-TRBO (Lindbo, 2007); and the CPMV-based system (pEAQ-HT; (Sainsbury and Lomonossoff, 2008) were investigated in order to determine which was most suitable for transient expression of camel VHHs at high levels in *N. benthamiana*. The gene encoding Green Fluorescent Protein (GFP) was cloned into the 3 vectors, giving pICON-GFP, pEAQ-GFP and pTRBO-GFP. These constructs were used a tool for initial assessment of the expression potential of each system, and the expression timeline. The constructs were infiltrated either on their own (in the case of pEAQ-GFP and pTRBO-GFP) or with additional provectors (in the case of pICON-GFP) into *N. benthamiana* leaves.

At 4-5 dpi, GFP fluorescence could be seen when leaves were illuminated with a long-wave UV lamp. However, the amount of fluorescence was much stronger in leaves infiltrated with pEAQ-GFP and pTRBO-GFP compared to pICON-GFP (Figure 6.1a and b). By 7 dpi, GFP fluorescence in pICON-GFP-infiltrated leaves was equal to or greater than the leaves infiltrated with pEAQ-GFP and pTRBO-GFP (Figure 6.1c and d). When protein extracts from 7 dpi leaves were analysed by PAGE and western blot, it was found that GFP had accumulated to much higher levels in pICON-GFP-infiltrated leaves compared to leaves infiltrated with pEAQ-GFP and pTRBO-GFP (data not shown). Furthermore, when protein extracts from leaves harvested at the peak of protein production (7 dpi for pICON-GFP, 4 dpi for pEAQ-GFP and pTRBO-GFP), were investigated, leaves infiltrated with pICON-GFP showed greater GFP accumulation compared to pTRBO-GFP. On the other hand, no such difference in GFP accumulation was observed when leaves infiltrated with pICON-GFP or pEAQ-GFP were compared (Figure 6.2).

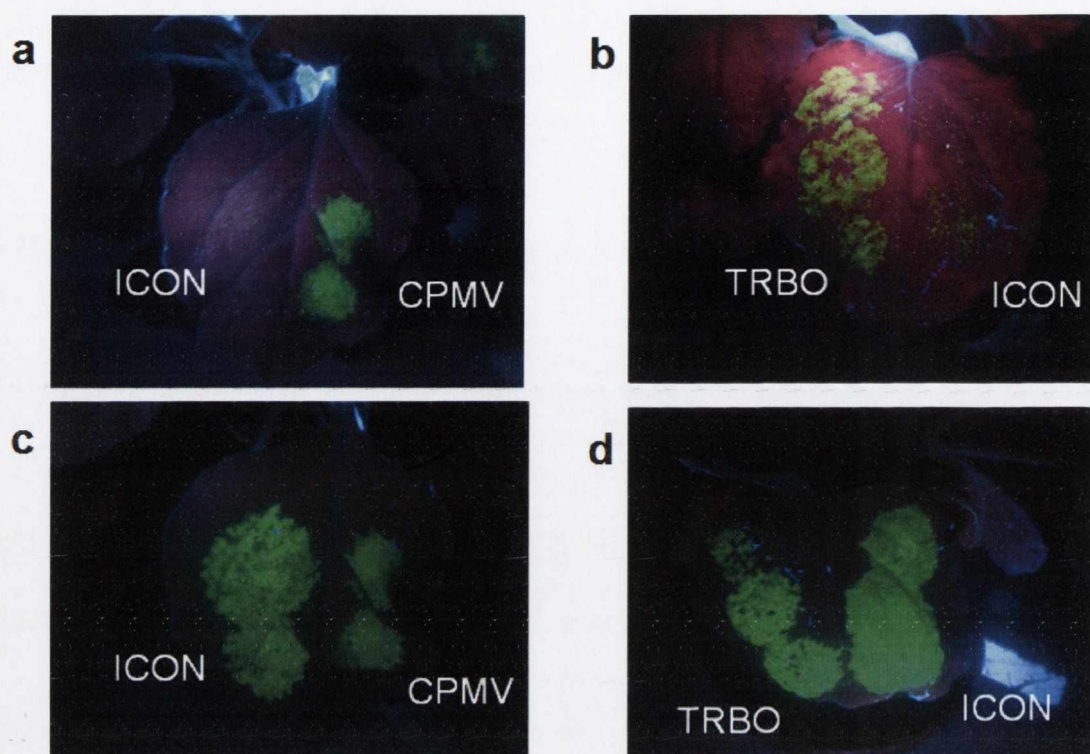


Figure 6.1 Comparison of GFP fluorescence levels in leaves of *N. benthamiana* infiltrated with the pICON-GFP, pEAQ-GFP at 4dpi, pTRBO-GFP at 5 dpi (a and b) and at 7 dpi (c and d).

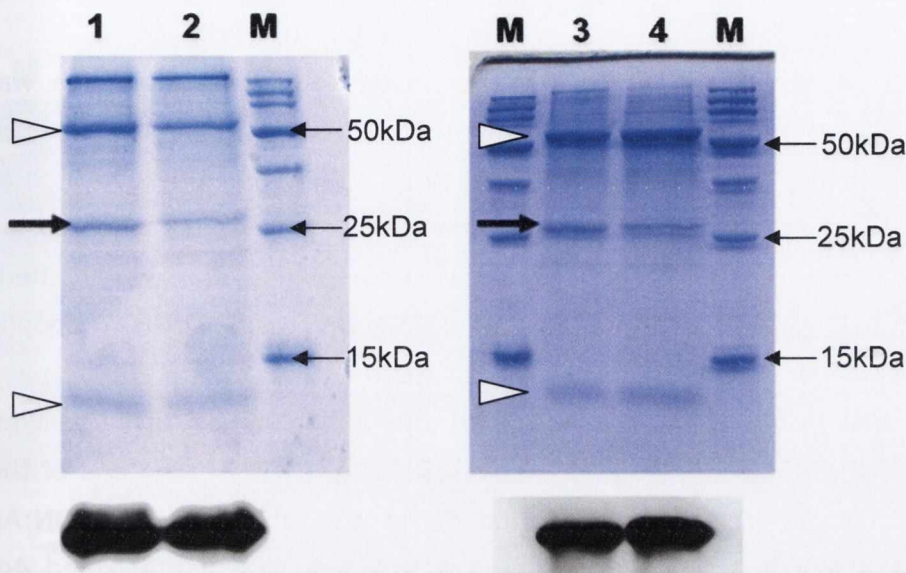


Figure 6.2 PAGE gel (above) and Western blot (below) of GFP accumulating in the cytoplasm of 7 dpi *N. benthamiana* leaves infiltrated with the pICON-GFP (Lanes 1 and 3), 4 dpi leaves infiltrated with pEAQ-GFP (Lane 2) and 5 dpi leaves infiltrated with pTRBO-GFP (Lane 4). 10 μ g of total soluble protein was loaded in each lane. M, Perfect ProteinTM markers. The western blot was probed using a 1:5,000 dilution of a polyclonal antibody against GFP. White arrow heads indicate the large and small RUBISCO subunits (54 kDa and 14 kDa, respectively). A black arrow indicates GFP (approx. 27 kDa). Gels and blots are from one experiment representative of two conducted.

6.2.2 Comparison of apoplast-targeted CBESynHS expression from three virus-based transient expression vectors

The 3 expression systems were also used to investigate which was the most suitable for transient expression of camel VHHs at high levels in *N. benthamiana*. The VHH CBESynHS was chosen due to its already documented high accumulation in *N. benthamiana* using the ICON Genetics system (as documented in Chapter 4). Since CBESynHS showed its highest accumulation level *in planta* when targeted to the apoplast, a derivative of the CBESynHS gene with an apoplast-targeting presequence at its 5' end was constructed and given the name Ap-CBE (as described in Chapter 7). Ap-CBE was then cloned into the vectors pEAQ-HT and pJL-TRBO, giving pEAQ:Ap-CBE and pTRBO:Ap-CBE. In the case of the equivalent ICON Genetics construct (referred to in this chapter as pICON:Ap-CBE), the identical apoplast-targeting signal was provided via a co-infiltrated *Agrobacterium* strain carrying the 5' provector module pICH8420.

In addition, for comparative purposes, Ap-CBE was cloned into two conventional binary vectors, pBIN2 and pBIN8. These are pBINPLUS/ARS-based (Belknap *et al.*, 2008) vectors in which transgene transcription is driven by either the CaMV35S promoter (pBIN2) or the *rbcS* (Rubisco small subunit) promoter (pBIN8); in each case polyadenylation signals were provided by the Nos terminator. The resulting constructs were called pBIN2:Ap-CBE and pBIN8:Ap-CBE.

It was found that CBE accumulation was highest for pICON:Ap-CBE (estimated 0.3 mg/g fresh weight) compared to pEAQ:Ap-CBE and pTRBO:Ap-CBE (estimated 0.06 mg/g and 0.05 mg/g fresh weight, respectively; Figure 6.3). Also, leaves infiltrated with pICON:Ap-CBE were correspondingly more necrotic compared to leaves infiltrated with the other constructs, most probably due to the presence of higher levels of the CBE protein. However, both the pEAQ:Ap-CBE and pTRBO:Ap-CBE systems required less time for maximum protein accumulation (4 dpi and 5 dpi, respectively) compared to the ICON system (7dpi). In contrast, no CBE protein was detectable by PAGE or Western blot when the conventional binary vector constructs pBIN2:Ap-CBE and pBIN8:Ap-CBE were used (Figure 6.3). It was also observed that CBE protein levels decreased significantly after 4 to 5 dpi and could not even be detected on a Western blot after 5 dpi for pEAQ:Ap-CBE or 7 dpi for pTRBO:Ap-CBE.

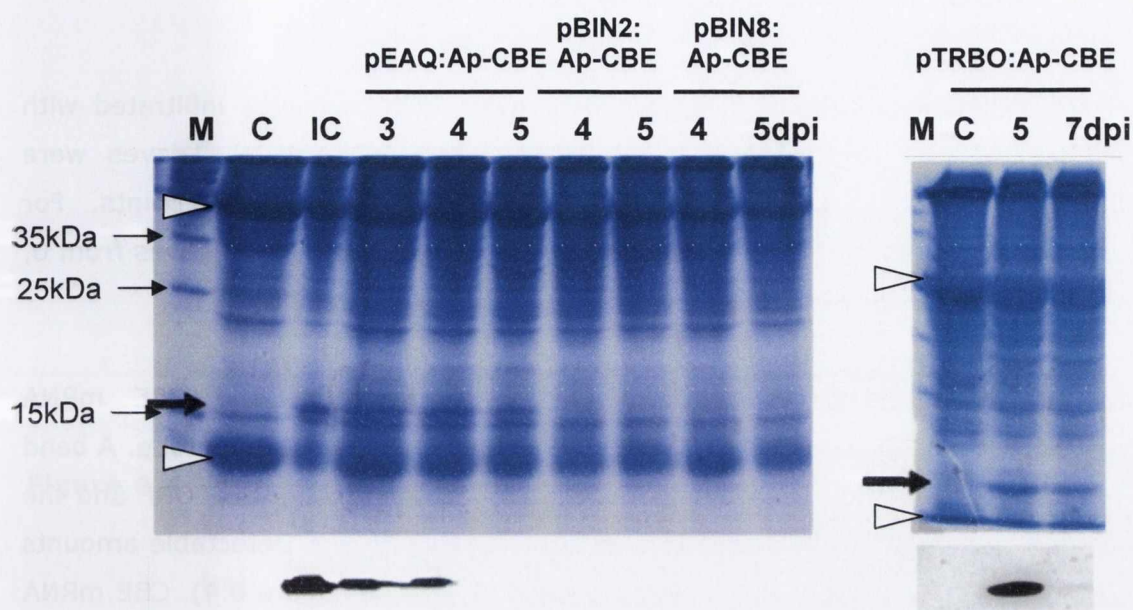


Figure 6.3 PAGE gel (above) and Western blot (below) of CBE accumulation in *N. benthamiana* leaves at various days post-infiltration (dpi) with pEAQ:Ap-CBE, pBIN2:Ap-CBE, pBIN8:Ap-CBE and pTRBO:Ap-CBE, compared to leaves transformed with empty vector (C) or the pICON:Ap-CBE vector (IC; leaves harvested at 7 dpi). The Western blot was probed using a 1:5,000 dilution of anti-6xHis tag monoclonal antibody. 200 μ g of total soluble protein was loaded in each lane, except for protein from leaves infiltrated with pICON:Ap-CBE, where 50 μ g was loaded. M, Perfect Protein™ Markers. White arrow heads indicate the large and small RUBISCO subunit (54 kDa and 14 kDa respectively). The black arrow indicates CBE (approx. 16 kDa). Gels and blots is representative of the three experiments conducted.

6.2.3 Expression, over time, of Ap-CBE from the pEAQ-HT vector at the protein and mRNA level

The accumulation of CBE in the cytoplasm, over time, in leaves infiltrated with pEAQ:Ap-CBE was investigated at the protein and mRNA level. Leaves were infiltrated with the construct and harvested at different time points. For investigation of both mRNA and protein expression levels over time, leaves from 0, 1, 2, 3, 4, 5 and 6 dpi were used.

Total RNA was extracted from transfected leaf material and CBE mRNA accumulation was investigated by Northern blot using a gene-specific probe. A band of approximately 1.4 kb was detected, which corresponds to the CBE ORF and the flanking 5' and 3' UTRs of Cowpea Mosaic Virus (CPMV) RNA-2. Detectable amounts of CBE mRNA were present at 2 dpi and peaked at 4-5 dpi (Figure 6.4). CBE mRNA levels, as well as the quality of the total RNA recovered, started to decrease at 6 dpi.

Total soluble protein was also extracted from the leaf material and was resolved by SDS-PAGE. Detectable amounts of the CBE protein can be seen on Coomassie blue-stained gels as early as 2 dpi (Figure 6.5). It continued to accumulate until 5 dpi where the levels subsequently decreased. Although CBE was clearly visible on Coomassie-stained gels, the yield for leaves infiltrated with pEAQ:Ap-CBE was estimated to be 5 times lower than for leaves infiltrated with the pICON:Ap-CBE vector.

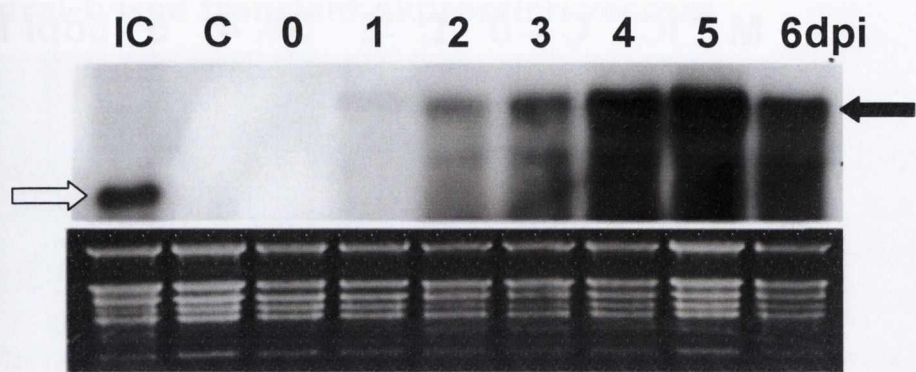


Figure 6.4 Northern blot showing CBE mRNA accumulation at 0 – 6 dpi in leaves of *N. benthamiana* infiltrated with pEAQ:Ap-CBE, compared to leaves transformed with empty vector (C) or pICON:Ap-CBE (IC; leaves harvested at 7dpi). Approximately 10 μ g of total RNA was loaded on the gel and the blot was probed with a DIG-labelled CBE gene-specific probe. The band indicated by a black arrow corresponds to the major transcribed product from the pEAQ:Ap-CBE vector (approximately 1.4kb). The lower band, indicated by the white arrow, corresponds to the CBE mRNA transcribed by the viral RdRP in the ICON system (approximately 1.2 kb). Ethidium bromide stained gel of the total RNA was used as loading controls (bottom panel).

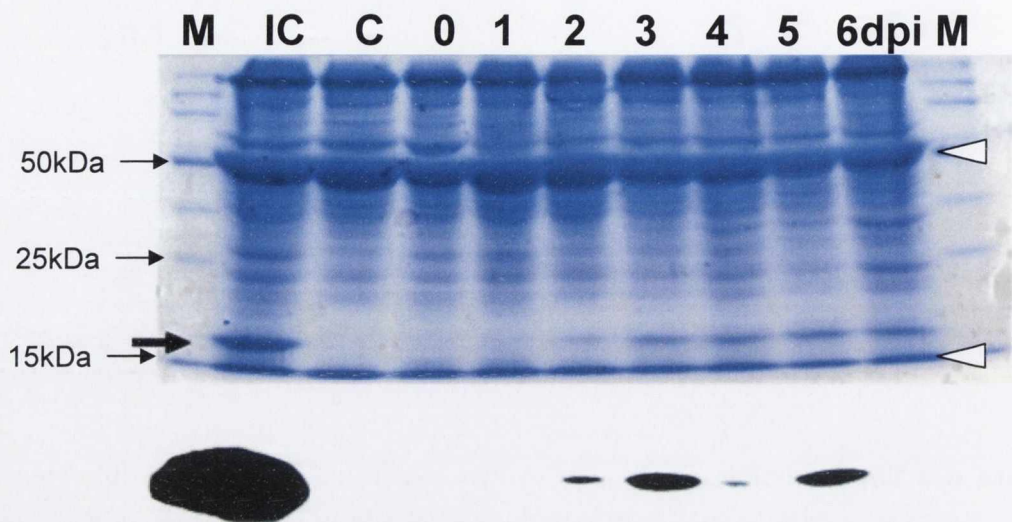


Figure 6.5 PAGE gel and Western blot showing CBE protein accumulation at 0 - 6 dpi in *N. benthamiana* leaves infiltrated with pEAQ:Ap-CBE compared to leaves transformed with empty vector (C) or pICON:Ap-CBE (IC; leaves harvested at 7 dpi). 50 µg of total soluble protein was loaded in each lane. M, Perfect Protein™ Markers. The Western blot was probed using a 1:5,000 dilution of anti-6xHis tag monoclonal antibody. The white arrow heads indicate the large and small RUBISCO subunits (54 kDa and 14 kDa respectively). The black arrow indicates CBE (approx. 16kDa). Gels and blots is representative of the four experiments conducted.

6.2.4 Comparison of PSAHS and CPSHS expression from three viral-based transient expression vectors

When the ICON Genetics vector system was used for transient expression of the anti-prostate-specific antigen (PSA) VHHs, (i.e. PSAHS and its framework changed counterpart CPSHS), low yields were obtained (Section 5.2.2). This prompted an investigation of pEAQ-HT and pJL-TRBO as alternative expression vectors with potential to increase PSAHS and CPSHS protein accumulation in *N. benthamiana*. Among the three compartments investigated, PSAHS and CPSHS were found to accumulate to relatively high levels in the cytoplasm with the ICON system (Figure 5.4a). Therefore, the corresponding GOIs were cloned into pEAQ-HT and pJL-TRBO without added targeting presequences in order to direct their accumulation in the cytoplasm. *N. benthamiana* leaves infiltrated with PSAHS or CPSHS cloned into pEAQ-HT and pJL-TRBO (i.e. constructs pEAQ-PSAHS, pEAQ-CPSHS, pTRBO-PSAHS and pTRBO-CPSHS) were harvested at 0, 3, 5, 7 and 9dpi. For comparison, 0, 3, 5, 7, 9 and 11 dpi leaves infiltrated with PSAHS and CPSHS cloned into the ICON vector pICH11599 (i.e. constructs pICON-PSAHS and pICON-CPSHS) were used.

It was found that among the three expression systems investigated, the ICON system gave the highest yields based on the estimation of the amount of VHH on SDS-PAGE gels and Western blots (Figure 6.6a), compared to pEAQ-HT and pJL-TRBO (Figure 6.6b and c respectively). PSAHS expressed using the ICON system accumulated to levels detectable on SDS-PAGE gels and Western blots at 5 - 9 dpi. For pICON-CPSHS, detectable amounts of protein could be seen at 5 dpi and the levels stayed the same until 11 dpi. For both pICON-PSAHS and -CPSHS, the leaves turned progressively necrotic which was reflected in the drop in the levels Rubisco large and small subunit on the PAGE gel. After 13 dpi, the protein extracted from leaves was severely degraded.

In leaves infiltrated with pEAQ-PSAHS and -CPSHS, no PSAHS protein was detected, while for CPSHS, VHH accumulation was detected at 5 dpi, peaking at 7 dpi with subsequent accumulation levels dropping at 9 dpi (Figure 6.6b). As for leaves infiltrated using pTRBO-PSAHS and -CPSHS, accumulation of PSAHS was observed only at 3 dpi, but not subsequently. On the other hand, CPSHS was detected between 3 - 5 dpi before protein degradation rendered the levels undetectable by SDS-PAGE or Western blot (Figure 6.6c). Leaves infiltrated with the pEAQ- and

pTRBO- constructs also turned progressively necrotic but this was not as severe as seen in leaves infiltrated with ICON provectors. Furthermore, it was also observed that in the case of the pICON- and pEAQ- constructs, the framework-changed VHH (CPSHS) accumulated to higher levels compared to the original version (PSAHS). In the case of pTRBO- constructs, although both proteins accumulated to similar levels, the levels of CPSHS were detected in the leaves for 2 days longer, which suggests that either: (1) the CPSHS gene was transcribed and translated for longer periods in the cell without the interference of post transcriptional gene silencing or (2) the protein itself was less susceptible to degradation.

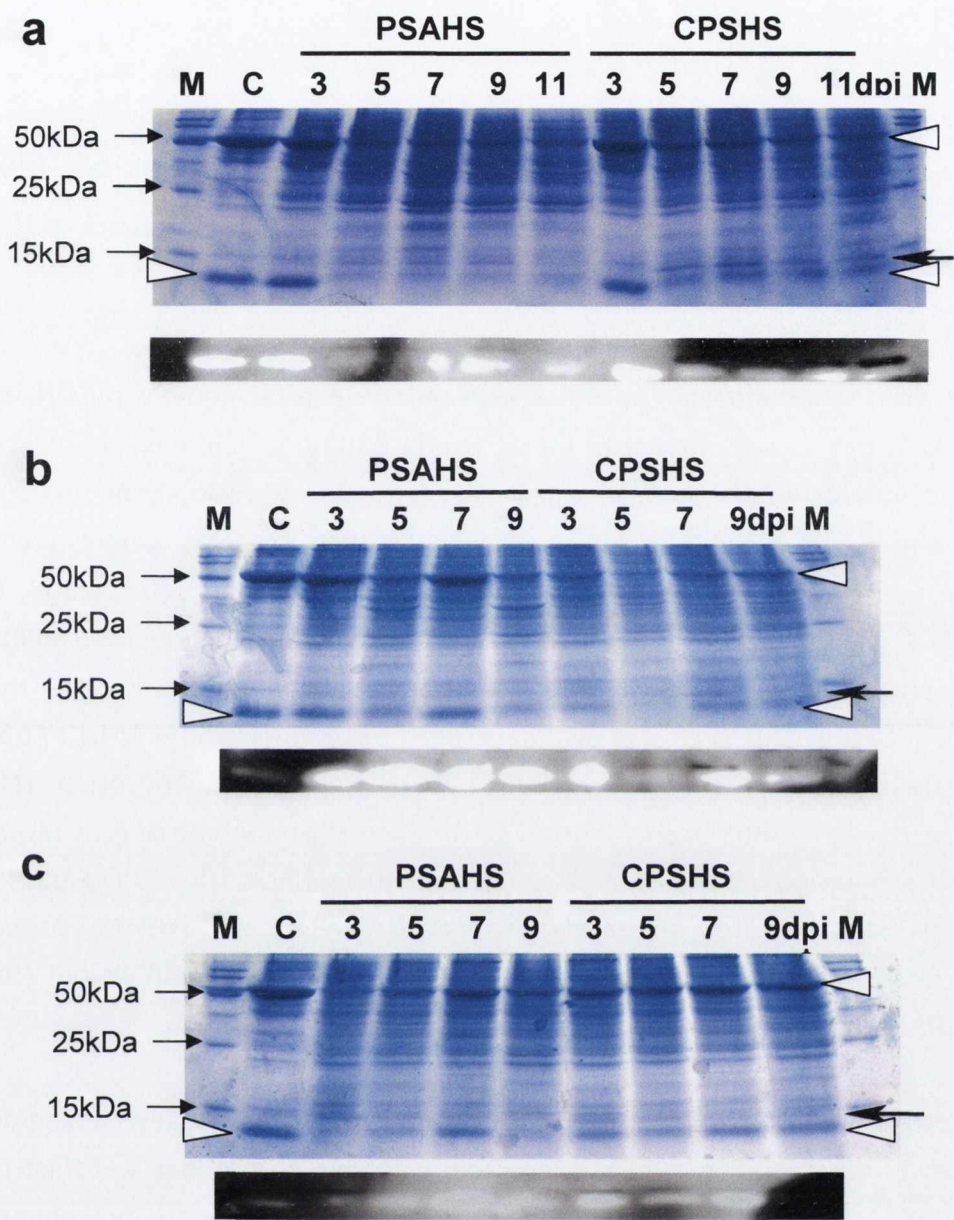


Figure 6.6 PAGE gel (above) and Western blot (below) showing PSAHS or CPSHS protein accumulating at 0-11 days post-infiltration (dpi) in the cytoplasm of *N. benthamiana* leaves infiltrated with the pICON constructs (Panel a), the pEAQ constructs (Panel b) or pTRBO constructs (Panel c) compared to leaves transformed with the corresponding empty vectors (C). 50 µg of total soluble protein was loaded in each lane. M, Perfect Protein™ Markers. White arrow heads indicate the large and small RUBISCO subunits (54 kDa and 14 kDa, respectively). The black arrow indicates the VHH (approx. 15 kDa). Gels and blots is representative of the three experiments conducted.

6.3 Discussion

6.3.1 Comparison of mRNA and protein expression levels of plant transformed with ICON, pEAQ-HT and pJL-TRBO vectors

Three viral genome-based transient expression vectors, namely pICON, pEAQ-HT and pJL-TRBO, were compared in order to determine which was the most suitable for the production of VHHs *in planta*. Since its development in 2004, the tobamovirus-based ICON system has been used to produce various recombinant proteins at a very high level (Marillonnet *et al.*, 2004; Santi *et al.*, 2006; Gleba *et al.*, 2007). More recently it has been used to express individualised idiotype IgG vaccines for non-Hodgkin's lymphoma at levels between 0.5 to 4.8 mg/g leaf (Bendandi *et al.*, 2010). It has also been used to produce Camel VHH to up to 0.45 mg/g fresh weight (Teh and Kavanagh, 2010; Chapter 4). The other TMV-based vector, pJL-TRBO, has been reported to produce recombinant GFP at levels of 3-5 mg/g fresh weight (Lindbo, 2007). On the other hand, the CPMV-based vector pEAQ-HT has been used to express Hepatitis B Core Antigen (HBcAg) particles at 1 mg/g fresh weight and the anti-HIV IgG 2G12 at 0.4 mg/g fresh weight (Sainsbury and Lomonossoff, 2008; Sainsbury *et al.*, 2009).

The anti-HEWL VHH CBE (i.e. CBESynHS; Chapter 4) was initially chosen due to its high-level accumulation when expressed with an N-terminal apoplast-targeting signal via the ICON system. The same apoplast-targeting presequence was therefore added to the 5' end of the CBE ORF (giving Ap-CBE; as described in Chapter 7) when it was expressed from pEAQ-HT and pJL-TRBO. Furthermore, two binary vectors, containing a CamV35S or *rbcS* promoter-driven expression cassette, were used as an additional comparison. It was found that when Ap-CBE was expressed from pEAQ-HT and pJL-TRBO, protein accumulation peaked earlier (4-5 dpi) than when the ICON system was used (7 dpi). This was also evident in the case of GFP expression in the first set of experiments, where GFP fluorescence was clearly observed at 4-5 dpi for plants infiltrated with pEAQ-GFP and pTRBO-GFP, compared to 7 dpi for pICON-GFP infiltrated plants. The observed time difference might be a consequence of the need for the provectors of the ICON system to recombine *in planta* before the mRNA containing the GOI can undergo RdRP-directed amplification in the cytoplasm (Marillonnet *et al.*, 2004; Gleba *et al.*, 2007).

This step is not required for either pEAQ-HT or pJL-TRBO. *In planta* recombination might therefore cause an initial lag in mRNA production and subsequent protein accumulation. This conclusion is supported by comparison of the mRNA expression profiles produced by the ICON and pEAQ-HT systems. In plants infiltrated with ICON vectors, mRNA accumulation was first observed at 2-3 dpi (please refer to Chapter 3) whereas in the case of pEAQ-HT, mRNA was detected in low amounts as early as 1 dpi. However, although the ICON-infiltrated leaves had a slower start, mRNA accumulation for both the ICON and pEAQ-HT systems peaked at 5 dpi.

When we compared recombinant protein accumulation in leaves infiltrated with the three vectors, the initial experiments with GFP showed that leaves infiltrated with pICON-GFP showed higher levels of protein accumulation compared to pTRBO-GFP infiltrated leaves when both proteins were harvested at the peak of production (4-5 dpi for pTRBO-GFP; 7 dpi for pICON-GFP). However, no significant difference in GFP accumulation was observed when compared to pEAQ-GFP infiltrated leaves. In contrast, it was found that for CBE (CBESynHS), pICON:Ap-CBE infiltrated leaves had higher levels compared to leaves infiltrated with pEAQ:Ap-CBE or pTRBO:Ap-CBE when VHH production levels were at their maxima (4 dpi for pEAQ:Ap-CBE, 5 dpi pTRBO:Ap-CBE and 7 dpi for pICON:Ap-CBE). Surprisingly, no protein was detectable by PAGE gel, Western blot or direct ELISA when only the CaMV35S promoter or the *rbcS* promoter was used to drive CBE expression. This suggests that the virus-based expression systems are especially suitable for the transient expression of VHHs and, moreover, that standard non-viral transgene expression cassettes, even those with strong promoters (e.g. CaMV35S and the *rbcS* promoter), are not.

For leaves infiltrated with pEAQ:Ap-CBE and pTRBO:Ap-CBE, VHH accumulation declined after peak production at 4-5 dpi. More interestingly, the VHH could not be detected using probes recognizing the C-terminal 6xHis-tag, although a band of the correct molecular weight was present on SDS-PAGE gels. This is likely the result of protein degradation, most probably directed at the VHH C-terminus, and due to proteases resident in the secretory pathway (Otsu *et al.*, 1995; Viale *et al.*, 1999; Raharjo *et al.*, 2001; Delannoy *et al.*, 2008), which have been shown to affect certain proteins targeted through the plant secretory system (Stevens *et al.*, 2000; Sharp and Doran, 2001; Outchkourov *et al.*, 2003).

6.3.2 Expression of PSAHS and CPSHS using ICON, pEAQ-HT and pJL-TRBO

When expressed in *N. benthamiana* leaves, using the ICON system, the anti-PSA VHHs (PSAHS and CPSHS) accumulated at low levels, with yields of up to 0.01 mg/g (Chapter 5). In order to determine if an alternative expression strategy might improve yields, these two VHHs were also expressed using the pEAQ-HT and pJL-TRBO systems. Since previous experience had shown that PSAHS and CPSHS accumulated most strongly in the cytoplasm, they were expressed without N-terminal targeting presequences. A direct comparison of the three expression systems clearly showed that the ICON system gave the highest yield. It is interesting to note, however, that regardless of which expression vector was used (ICON, pEAQ-HT or pJL-TRBO), CPSHS, the nanobody with the 'universal' framework, had the higher yield and was less susceptible to degradation than its native framework counterpart PSAHS. In pEAQ-HT-infiltrated leaves, although no band was observed for PSAHS, CPSHS was detected from 3-9 dpi. In the case of pJL-TRBO, CPSHS was detectable up to 5 dpi, while PSAHS could only be detected at 3 dpi. The difference in yield was even more evident when the ICON system was used. The yield increase is likely framework-associated and is probably attributable to improved resistance to proteolysis, but might also involve improved transcript stability or translational efficiency (Ciechanover, 1998; Hellmann and Estelle, 2002). Unfortunately, leaves infiltrated with the PSAHS and CPSHS constructs, irrespective of the vector used, became severely necrotic after 7 dpi, most probably due to disruptive effects on plant metabolism by these particular VHH. For this reason, we did not proceed any further with this line of investigation.

Future attempts to improve the yield of PSAHS and CPSHS might include using alternative expression vectors like that based on Bean Yellow Dwarf Virus (BeYDV) (Chen *et al.*, 2011). Unlike TMV and CPMV, BeYDV is a geminivirus with a single-stranded DNA genome that replicates in the nucleus of the plant cell using a virus-encoded replication initiator protein (Rep) and the plant DNA synthesis mechanism. The BeYDV-based vector has been used in the production of vaccine antigens and monoclonal antibodies, achieving yields of up to 0.8 mg/g fresh weight and 0.5 mg/g fresh weight, respectively (Huang *et al.*, 2009; Huang *et al.*, 2010). Another strategy to improve yield might be to anchor the VHH to the cytosolic face of the ER membrane. Barbante *et al.* (2008) successfully fused the HIV protein Nef (Negative

regulatory factor) to the C-terminal domain of mammalian ER cytochrome *b5*, anchoring the Nef protein to the cytosolic face of the ER. The Nef protein is very unstable when targeted to the secretory pathway and accumulated at low levels in transgenic tobacco plants and protoplast (Marusic *et al.*, 2007). However, protein accumulation increased three-fold (from 0.2% to 0.7% TSP) when anchored to the cytosolic face of the ER compared to its non-anchored soluble counterparts accumulating in the cytoplasm (Barbante *et al.*, 2008).

Chapter 7. VHH transgene expression in stably transformed *N. tabacum* and *N. benthamiana* lines

7.1 Introduction

Stable transformation requires the incorporation of a GOI into the plant nuclear or chloroplast genome. Because the GOI is integrated into genomic DNA of the host it can be stably inherited in subsequent generations. The initial transformation event occurs at the single cell (or organelle) level, and therefore requires the use of a selectable marker linked to the GOI, such as the *nptII* gene (which confers kanamycin resistance), the *aadA* gene (which confers streptomycin and spectinomycin resistance), or the *bar* gene (conferring tolerance to the herbicide BASTA), in order to eliminate non-transformed neighbouring cells (or organelles). Recent concerns about the potential human health and ecological impacts of antibiotic resistance- and herbicide tolerance-based markers have been alleviated, to some extent, by the development of post-selection marker removal strategies (Dale and Ow, 1991; Gleave *et al.*, 1999; Zubko *et al.*, 2000; Song *et al.*, 2010), and the development of alternative selection strategies (e.g. Song *et al.*, 2010).

The most commonly used method for nuclear transformation in plants exploits the DNA transfer functions of *Agrobacterium* species (in particular *Agrobacterium tumefaciens* and *Agrobacterium rhizogenes*) (For a review of *A. tumefaciens*-mediated transformation, see Gelvin, 2003). *A. tumefaciens* naturally infects dicotyledonous plants but can also be used to transform some monocotyledonous plants and certain fungal species. Although the simple 'Floral dip' method (Clough and Bent, 1998) can be used for *Arabidopsis thaliana* plants, *Agrobacterium*-mediated transformation of most plant species requires infection of tissue explants, such as leaf discs, or cell cultures, followed by selection and regeneration of transgenic shoots *in vitro* on hormone-containing culture media. A typical stable transformation vector comprises a modified *Agrobacterium* Ti plasmid with a T-DNA region that contains a selectable marker gene linked to the GOI, and a constitutive or inducible promoter to direct expression of the GOI *in planta*. Various alternatives to *Agrobacterium*-mediated transformation exist. In particular, the biolistic method

is widely used in plastid transformation and to transform some monocotyledonous species (Sanford, 1988; Sanford *et al.*, 1993). Other alternatives include polyethyleneglycol (PEG)-mediated DNA transfer into protoplasts (Uchimiya *et al.*, 1986), protoplast and intact cell electroporation (Fromm *et al.*, 1985; Lörz *et al.*, 1985; De la Pena *et al.*, 1987; Arencibia *et al.*, 1995) and microinjection (De la Pena *et al.*, 1987).

The critical advantage of stable, over transient, transformation is that the GOI is heritable. For this reason, stable transformation has been the preferred method for large-scale production of recombinant proteins in plants. Furthermore, the availability of stable transformants facilitates additional genetic manipulation such as crossing different lines to achieve gene stacking (Halpin, 2005). This is advantageous for the modification of multi-step biochemical pathways, or when the aim is to produce multimeric proteins (e.g. secretory antibodies; Ma *et al.*, 1995) on a commercial scale. On the other hand, disadvantages of stable transformation include the length of time required to generate and characterize stable transgenic lines, and unpredictable transgene expression levels due to position effects, gene silencing (De Wilde *et al.*, 2000) and variable transgene copy numbers (Kooter *et al.*, 1999; De Wilde *et al.*, 2000).

Despite the fact that VHHs have been stably expressed in *E. coli* and yeast for over 15 years, there are only four reports that describe the expression of VHHs in stably transformed plants (Jobling *et al.*, 2003; Ismaili *et al.*, 2007; Winichayakul *et al.*, 2009; Conrad *et al.*, 2011). Moreover, in each of these cases the level of protein accumulation was very low, detectable only by ELISA or by immunoblot analyses.

In this chapter, we investigated various strategies for VHH expression in stably transformed *N. tabacum* cv. Samsun and *N. benthamiana* plants. The His-tagged anti-HEWL VHH with the 'universal' cAbBCII10 framework (CBESynHS) was chosen for these experiments. Since it was previously shown that CBESynHS accumulation was highest when targeted to the apoplast (Please refer to Section 4.2.2 for details), sequences encoding the calreticulin apoplast-targeting signal (Ap-) were added to the 5' end of the gene.

Please note that in this chapter, we have abbreviated 'apoplast-targeted CBESynHS' to Ap-CBE and apoplast-targeted CLDSynHS to Ap-CLD.

7.2 Results

7.2.1 Construction of an Ap-CBE transgene by RT-PCR, using RNA from infiltrated leaves as the template

In the ICON expression system, the nucleotide sequences that encode N-terminal targeting signals and the GOI are initially located on separate 5' and 3' provector modules, respectively. Following co-infiltration into leaves with a third *Agrobacterium* strain containing the gene coding for phage C31 integrase, the 5' and 3' modules are joined together in the plant cell nucleus. Expression of the GOI, therefore, is a multi-step process involving: (a) fusion of the 5' and 3' provector modules via site-specific recombination (i.e. processing at the DNA level), followed by transcription of the GOI fusion product, and (b) removal of an intron, generated by the recombination event, which separates the targeting presequence and the GOI coding regions on the resulting mRNA transcript (i.e. processing (splicing) at the RNA level). Thus, in the case of Ap-CBE two types of fusion product may be isolated for subsequent use as a transgene: (i) a cDNA copy of the mature spliced Ap-CBE mRNA transcript and (ii) a genomic DNA copy of Ap-CBE which retains the intron created by site-specific recombination.

In this section we describe the isolation of a cDNA encoding Ap-CBE, while isolation of a genomic copy is described in Section 7.2.4. The Ap-CBE cDNA was obtained by reverse transcription of RNA isolated from *N. benthamiana* leaves that had been infiltrated with the relevant 3' and 5' ICON provector modules, followed by PCR amplification of first-strand cDNA (RT-PCR). The resulting cDNAs were initially cloned into pUC19 (Yanisch-Perron *et al.*, 1985) and their nucleotide sequences determined.

In initial experiments few positive clones were obtained due to poor *E. coli* survival on the ligation plate. Moreover, sequencing of plasmids from these clones revealed mutations in the Ap-CBE cDNA inserts. Also, different *E. coli* clones from the same ligation plate contained inserts that differed from each other. Some of these mutations were base substitutions resulting in silent, missense or nonsense mutations in the Ap-CBE coding region. Remarkably, some Ap-CBE inserts contained incomplete or complete repeats of the Ap pre-sequence (data not shown). Mutations in the CBE region of Ap-CBE were successfully 'repaired' by replacing the

mutant CBE fragment via the flanking restriction sites NcoI and BamHI with the 'correct' fragment obtained from pUC57-CBESynHS (please see Section 2.2.1). However, attempts at repairing the Ap presequence by cutting out the mutated region and ligating in the correct sequence were not successful requiring the use of an alternative strategy.

Inspection of Ap-CBE sequences in pUC19 revealed that the coding region was located downstream of, and in-frame with, the *lacZa* ORF (Figure 7.1). It was speculated that this might result in expression of a LacZa-Ap-CBE fusion protein driven by the upstream *lac* promoter, which in turn might be toxic to *E. coli*. Therefore, a forward primer was designed with a one base deletion between the 5' restriction site (BamHI) and the Ap start codon (ATG) in order to put the Ap-CBE coding region out-of-frame relative to *lacZa* (Figure 7.1; Deleted nucleotide shown in red). Many more *E. coli* colonies were obtained when this primer was used for PCR amplification of Ap-CBE. Furthermore, in the resulting construct the Ap sequence was correct and the CBE coding region had only a single point mutation, which was subsequently fixed by restriction fragment replacement (as described above).

A sequence-correct version of AP-CBE was subsequently cloned into three *Agrobacterium* expression vectors: pBIN2, pBIN8 and pEAQ-HT (please refer to Chapter 6). pBIN2 and pBIN8 are pBINPLUS/ARS-based vectors (Belknap *et al.*, 2008) containing, respectively, a CaMV35S promoter or a Rubisco small subunit (*rbcS*) promoter (RbcS; Amarasinghe *et al.*, 2006) (Figure 7.2a and b). The pEAQ-HT vector (Sainsbury and Lomonossoff, 2008) contains, in addition to the CaMV35S promoter, the CPMV 5' and 3' UTRs flanking Ap-CBE (Figure 7.2c), and a cassette encoding P19, a potent suppressor of RNA silencing. We hypothesized that by comparing Ap-CBE expression in plants stably transformed with the pBIN2 and pBIN8 constructs, we could determine whether the CaMV35S or the *rbcS* promoter was the more suitable for high-level expression of camel VHHs *in planta*. Furthermore, by comparing Ap-CBE expression from the pBIN2 and pEAQ-HT constructs (both of which contain the CaMV35S promoter), we could also determine if the CPMV 5' UTR and 3' UTRs and the P19 suppressor of silencing, shown to dramatically increase protein yield in transient expression experiments (Sainsbury and Lomonossoff, 2008), led to increased VHH expression in stably transformed plants.

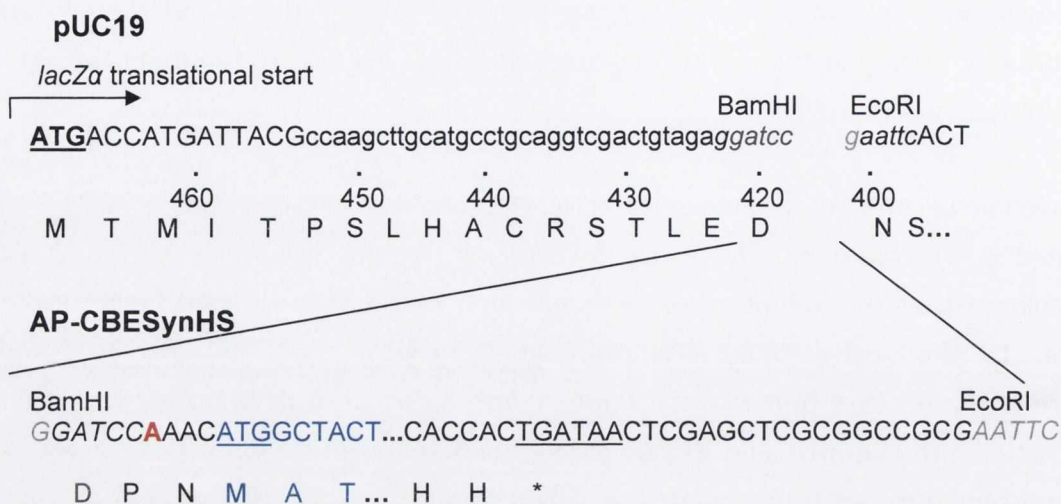


Figure 7.1 Nucleotide and amino acid sequence of an Ap-CBE clone in pUC19 (Yanisch-Perron *et al.*, 1985) starting from the *lacZα* translational start site (underlined and in bold). The start codon of Ap-CBE was found to be in frame with the *lacZα* translational start. Start/Stop codons are underlined. Restriction sites are in italics. Restriction sites in the MCS of pUC19 are shown in lower case. The codons and amino acids of the Ap-targeting sequence are shown in blue. The single base subsequently removed from the forward primer in order to put Ap-CBE out-of-frame relative to the *lacZα* start codon is shown in red.

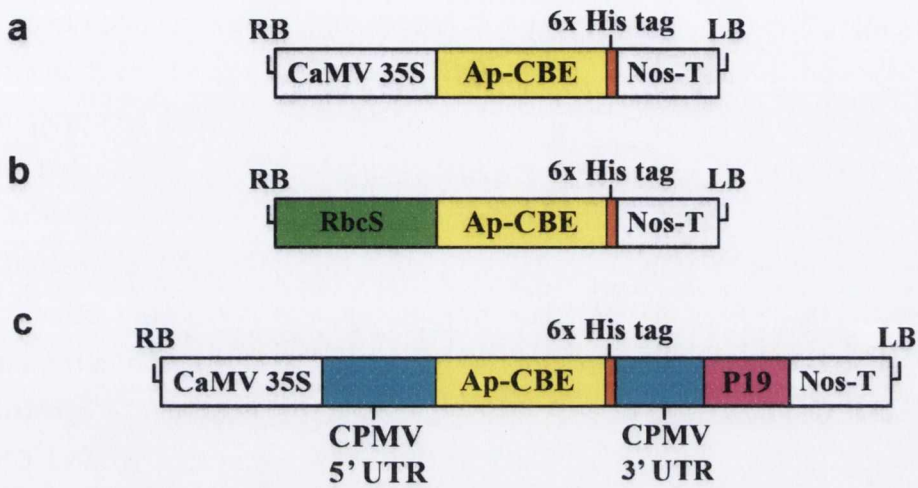


Figure 7.2 Comparison of three constructs for stable transformation derived from cDNA of transiently-transformed *N. benthamiana* leaves. RB-Right border; LB-Left border; CaMV 35S- Cauliflower Mosaic Virus 35S promoter; *RbcS*- Rubisco small subunit promoter; Ap - Tobacco calreticulin apoplast targeting presequence; CBE-anti-HEWL VHH; UTR- untranslated region; P19- Tomato Bushy Stunt Virus p19 suppressor of silencing; Nos-T: Nopaline synthase terminator.

7.2.2 Stable transformation of *N. tabacum* and *N. benthamiana* using the pEAQ-HT vector

The pEAQ-HT vector (Sainsbury and Lomonossoff, 2008) has been used extensively to direct high-level transient expression of recombinant proteins *in planta*. However, so far, there are no reports of pEAQ-HT being used to produce stably transformed plants. Therefore, to investigate the feasibility of pEAQ-HT being used for high-level stable expression of VHH transgenes in *N. tabacum* and *N. benthamiana*, two pEAQ-HT constructs containing the GFP gene were made. One directed expression of GFP in the cytoplasm (no added presequence), while the other contained a calreticulin Ap presequence in order to direct GFP accumulation in the apoplast (Ap-GFP). Two positive controls were included in the investigation: a construct in which GFP expression was driven by the CaMV35S promoter (pBIN2-GFP), and a chloroplast-transformed tobacco line containing a GFP transgene driven by the plastid ribosomal RNA promoter (*Prrn*) (Newell *et al.*, 2003). *Agrobacterium* harbouring the various constructs was used to generate stably transformed lines of *N. tabacum* cv Samsun and *N. benthamiana*. It was found that GFP was expressed in both *N. tabacum* cv Samsun and *N. benthamiana* plants stably transformed with the pEAQ-GFP construct. Furthermore, when the Ap presequence was present, GFP was found to accumulate in the apoplast. In contrast, in the absence of a targeting presequence, GFP accumulated in the cytoplasm (Figure 7.3). Moreover, the level of GFP accumulation in plants transformed with pEAQ-GFP or pEAQ:Ap-GFP was similar to that observed in plants transformed with pBIN2-GFP. However, it was much lower than that observed in a GFP-expressing chloroplast-transformed line (Figure 7.4).

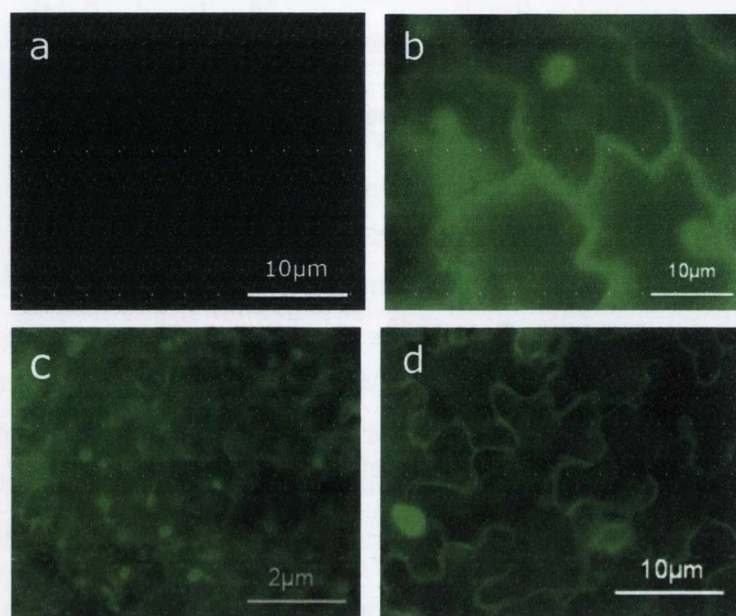


Figure 7.3 Green fluorescence in leaves of *N. tabacum* (b) and *N. benthamiana* (c) plants stably transformed with pEAQ-GFP; and *N. benthamiana* stably transformed with pEAQ:Ap-GFP (d) compared with a wild-type control (a). a, b and d shows cells at 200 x magnification while c shows cells at 50 x magnification. Scale bar 10 µm for 200x magnification and 2 µm for 50 x magnification.

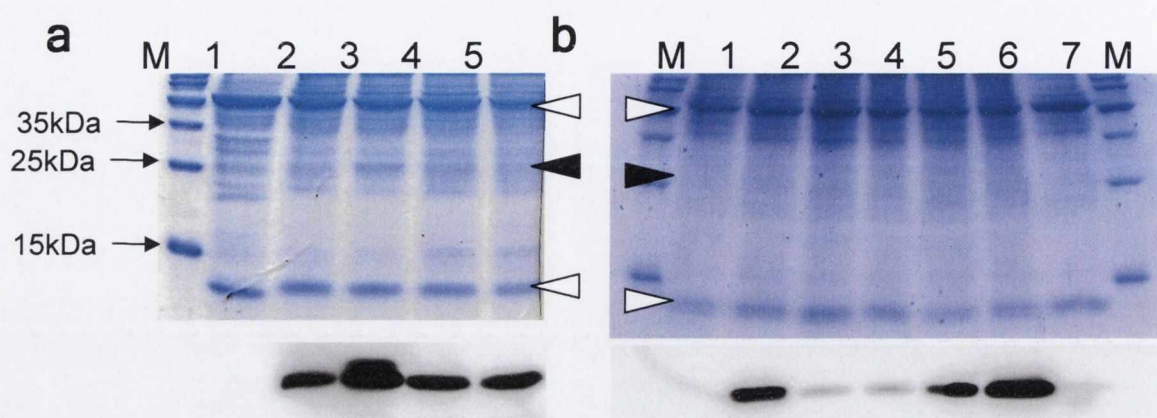


Figure 7.4 GFP accumulation in stably transformed *N. tabacum* cv. Samsun and *N. benthamiana* plants.

(a) PAGE gel (above) and Western blot (below) showing GFP accumulation in stably transformed *N. tabacum* cv Samsun lines, compared to a non-transformed line (Lane 1). GFP expression from pBIN2 (Lane 2) and pEAQ-HT (Lane 4) constructs. Ap-GFP expression from a pEAQ-HT construct (Lane 5). Lane 3 shows GFP expression in a chloroplast-transformed plant. 10 μ g of total soluble protein was loaded in each lane. The Western blot was probed with 1:1,000 rabbit anti-GFP and 1:75,000 goat anti-rabbit IgG-HRP conjugate. White arrow heads indicate the large and small RUBISCO subunits (54 kDa and 14 kDa, respectively). A black arrow head indicates the GFP protein (approx. 27 kDa).

(b) PAGE gel (above) and Western blot (below) showing GFP accumulation in stably transformed *N. benthamiana* lines compared to a non-transformed line (Lane 1). AP-GFP (Lane 2) and GFP (Lane 3-7) expression from pEAQ-HT constructs. 10 μ g of total soluble protein was loaded in each lane. Other details as in (a).

7.2.3 Expression of the Ap-CBE cDNA transgene in *N. tabacum* and *N. benthamiana*

The Ap-CBE cDNA was cloned into plant transformation vectors under the transcriptional control of either the CaMV35S promoter (pBIN2), the Rubisco small subunit (*rbcS*) promoter (pBIN8), or the CaMV35S promoter and the 5' and 3' UTRs from CPMV (pEAQ-HT). However, Western blot analyses of total protein extracts from the resulting transgenic lines failed to detect any AP-CBE protein for any of the three constructs (data not shown), although the Ap-CBE transgene was shown to be present by PCR screening of putative transformants (Figure 7.5a and b). Northern blots were performed to rule out the possibility that the Ap-CBE transgene was not transcribed. Total RNA samples extracted from PCR-positive transformants were electrophoresed in denaturing 1% agarose gels, blotted and probed with a CBE-specific DIG-labelled probe.

Among the *N. tabacum* transformants, only plants transformed with constructs driven by the *rbcS* promoter (pBIN8) showed mRNA accumulation that was detectable by Northern blot (Figure 7.6a). A band of about 1.1 kb corresponds to the transcript containing Ap-CBE. No mRNA transcripts were found in PCR-positive transformants containing the CaMV35S promoter with (pEAQ-HT) or without (pBIN2) the CPMV 5' and 3' UTRs and P19. In contrast, Ap-CBE mRNA accumulation was detected in *N. benthamiana* transformed with all three constructs. Higher levels of mRNA accumulation were detected in transformants in which transgene transcription was driven by the CaMV35S promoter compared to transformants in which transcription was driven by the *rbcS* promoter, regardless of whether the CPMV 5' and 3' UTR and P19 was present or not (Figure 7.6b). The band of about 1.1 kb corresponds to the Ap-CBE transcript produced from pBIN2 and pBIN8, while the band of about 2.1 kb corresponds to the larger Ap-CBE transcript which contains the 5' and 3' UTRs from CPMV.

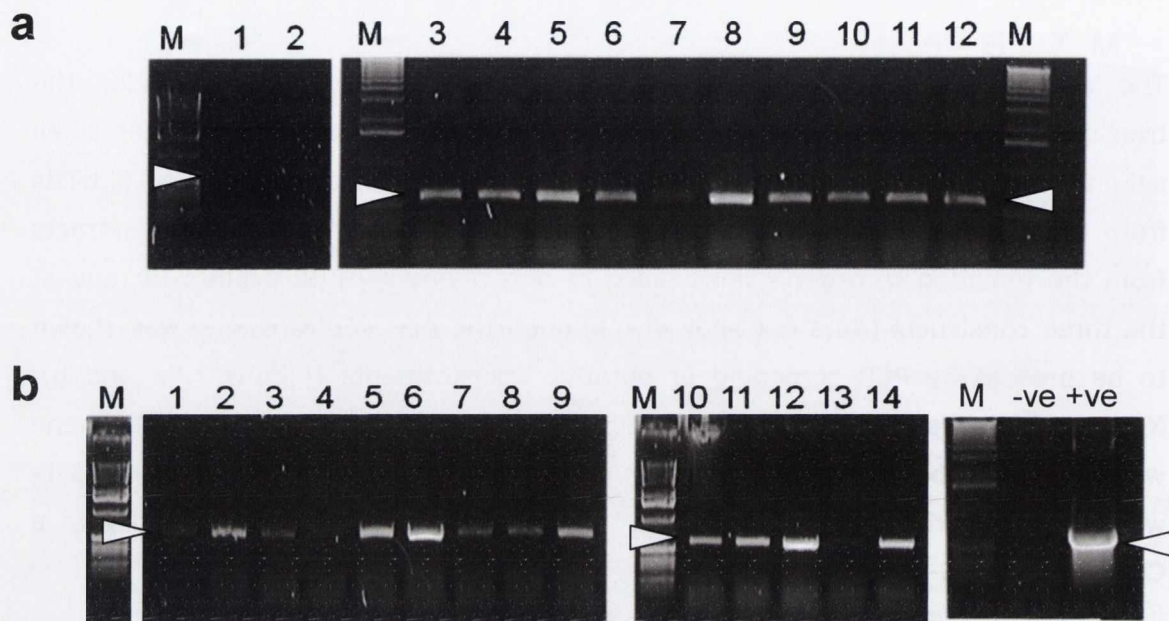


Figure 7.5 PCR screening of putative *N. tabacum* and *N. benthamiana* transformants for the presence of the Ap-CBE transgene.

(a) Agarose gel showing the PCR product (approx. 500bp) amplified from putative *N. tabacum* cv Samsun transformants containing pEAQ:Ap-CBE (Lanes 1 and 2), pBIN2:Ap-CBE (Lanes 3 and 4) and pBIN8:Ap-CBE (Lanes 5 to 12). Untransformed leaves were used as negative control (-ve) and the pBIN2:Ap-CBE vector was used as positive control (+ve).

(b) Agarose gel showing the PCR product (approx. 500bp) amplified from putative *N. benthamiana* transformants containing pBIN2:Ap-CBE (Lanes 1 to 4) and pBIN8:Ap-CBE (Lanes 5 to 15). PCR screening of putative pEAQ:Ap-CBE transformants is not shown. Untransformed leaves were used as a negative control (-ve) and the pBIN2:Ap-CBE vector was used as positive control (+ve).

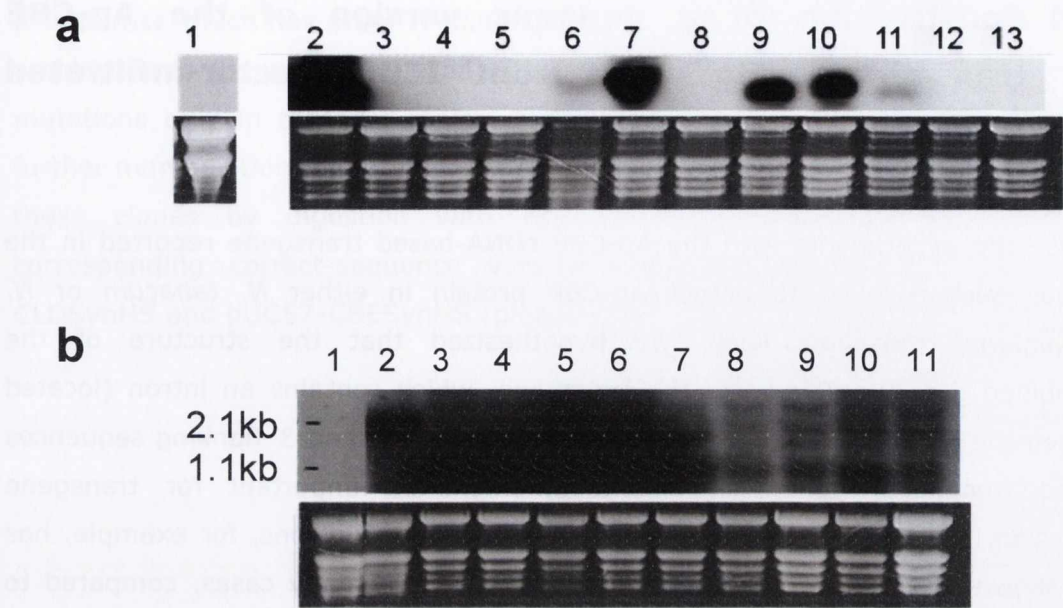


Figure 7.6 mRNA transcripts of *N. tabacum* cv Samsun and *N. benthamiana* transformants containing the Ap-CBE transgene.

(a) Northern blot showing mRNA transcripts in *N. tabacum* cv Samsun plants stably transformed with pBIN2:Ap-CBE (Lanes 4 and 5), pBIN8:Ap-CBE (Lanes 6 to 13) or pEAQ:Ap-CBE (Lane 1) constructs compared to an untransformed plant (Lane 3). *N. benthamiana* leaves transiently expressing apoplast-targeted CBESynHS from the relevant ICON vectors (Marillonnet *et al.*, 2004; Santi *et al.*, 2006) was used as positive control (Lane 2). Total RNA was isolated from leaves of stably transformed plants and analysed by Northern blot using a gene-specific probe. Approximately 10 µg total RNA was loaded into each lane. Ethidium bromide stained gel of the total RNA was used as loading controls (bottom panel).

(b) Northern blot showing mRNA transcripts in *N. benthamiana* plants stably transformed with pEAQ:Ap-CBE (Lane 2), pBIN2:Ap-CBE (Lanes 3 to 6), and pBIN8:Ap-CBE (Lanes 7 to 11) constructs compared to an untransformed plant (Lane 1). Total RNA was isolated from leaves of stably transformed plants and analysed by Northern blot using a gene-specific probe. Approximately 10 µg total RNA was loaded onto each lane. Ethidium bromide stained gel of the total RNA was used as loading controls (bottom panel).

7.2.4 Construction of a genomic version of the Ap-CBE transgene using DNA from ICON vector-infiltrated leaves as the PCR template

Because the experiments with the Ap-CBE cDNA-based transgene reported in the previous section failed to detect Ap-CBE protein in either *N. tabacum* or *N. benthamiana* transgenic lines, we hypothesized that the structure of the recombined genomic version of the transgene, which contains an intron (located between the Ap and CBE coding sequences), and both 5' and 3' flanking sequences derived from the tobamovirus genome, might be important for transgene expression in stably transformed plants. The addition of introns, for example, has been shown to increase transgene expression levels in many cases, compared to identical constructs without introns (for review please see Le Hir *et al.*, 2003).

To obtain the required construct, *N. benthamiana* leaves were infiltrated with *Agrobacterium* strains harbouring the ICON vectors pICH10881, pICH8420 and either pICH11599-CBESynHS or pICH11599-CLDSynHS. DNA was isolated from the infiltrated leaves and used as a template for the amplification of two fragments: the first, named Long Fragment (LF) contains sequences that include the tobamoviral region containing the subgenomic promoter and 5' UTR sequences located near the 3'-end of the Movement Protein (MP) gene, sequences of the apoplast-targeting presequence (Ap), the recombined intron region, the GOI [either CBESynHS (CBE) or CLDSynHS (CLD)], and the tobamoviral 3' non-translated region (3'NTR) (Figure 7.7). The second fragment amplified by PCR, named Short Fragment (SF) contains only sequences encoding Ap, the intron and the His-tagged GOI (Figure 7.7). The amplified fragments were cloned into the vector pBIN2 (in which transgene expression is driven by a CaMV35S promoter). *Agrobacterium* harbouring the LF and SF constructs was then used to stably transform *N. tabacum* cv Samsun and *N. benthamiana* plants.

Sequence analysis of cloned LF and SF PCR fragments revealed several point mutations, particularly in the intron region but also in the GOI (data not shown). Some of the mutations were silent, but most of them were missense mutations. The mutations appeared to be random, as different *E. coli* clones from the same ligation plate contained fragments that were different from each other. The same problem was observed even when Platinum[®] Taq High Fidelity DNA polymerase

(Invitrogen, USA), a mixture of Taq DNA polymerase and *Pyrococcus spp.* GB-D polymerase which has proof-reading capabilities, was used (the fragments failed to amplify when *Pfu* alone was used). Therefore, one LF clone and one SF clone with mutations only in the GOI region and not in the other regions was chosen for further manipulation. The remaining mutation-ridden GOI region was removed from these clones by digestion with NcoI and BamHI and replaced with the corresponding correct-sequence version obtained from the plasmids pUC57-CLDSynHS and pUC57-CBESynHS (please refer to Section 2.2.1).

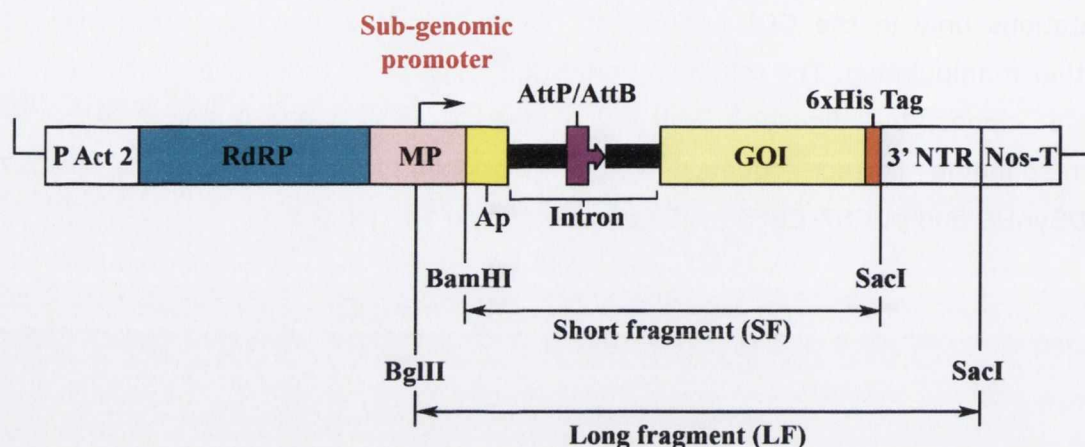


Figure 7.7 Diagram showing the predicted structure of the *in planta* recombined ICON 5' and 3' vector modules, and the regions contained within the PCR-amplified fragments. The Long Fragment (LF) region extends from the MP subgenomic promoter to the tobamoviral 3'NTR. The Short Fragment (SF) region extends from the Ap targeting presequence to the 6xHis-tag at the 3'-end of the GOI. The fragments were amplified using DNA from *N. benthamiana* leaves that had been infiltrated with the relevant ICON vectors. The GOI is either CLD or CBE; however following recombination, which adds the Ap presequence, these are referred to as Ap-CLD and Ap-CBE, respectively. The SF and LF constructs were subsequently cloned into pBIN2 using the restriction sites indicated in the diagram. P Act2 - *Arabidopsis* actin 2 promoter; RdRP - RNA dependent RNA polymerase; MP - Movement Protein; I - Intron; Ap - apoplast-targeting presequence; GOI- Gene of interest; 3'NTR- 3' non-translated region; Nos-T: Nopaline synthase terminator.

7.2.5 Expression of the LF and SF transgenes in *N. tabacum* and *N. benthamiana*

Multiple transgenic lines of both *N. tabacum* and *N. benthamiana*, containing LF and SF transgenes, encoding Ap-CBE and Ap-CLD were obtained and the GOI was shown to be present by PCR screening of putative transformants (Figure 7.8a and b for SF constructs and Figure 7.9a and b for LF constructs). However, no VHH protein was detected by Western blot or ELISA (data not shown) in any line. Similar results were obtained when a polyclonal anti-llama IgG and VHH antibody was used instead of the usual anti-His tag antibody (data not shown).

In order to determine if the GOI was expressed at the mRNA level, Northern blots were performed using total RNA extracted from PCR-positive transformants and a gene-specific DIG-labelled probe. mRNA accumulation was detected in some positive transformant lines for both the SF (Figure 7.10a and b) and LF constructs (Figure 7.11a and b). There was significant between-line variation in mRNA accumulation levels, probably influenced by gene integration position effects. The band of about 1.2 kb corresponded to the expected size of the transcript containing the GOI and the Ap presequence, while the band of about 2.1 kb corresponds to the transcript containing the GOI, the Ap- presequence and the additional upstream and downstream tobamoviral sequences.

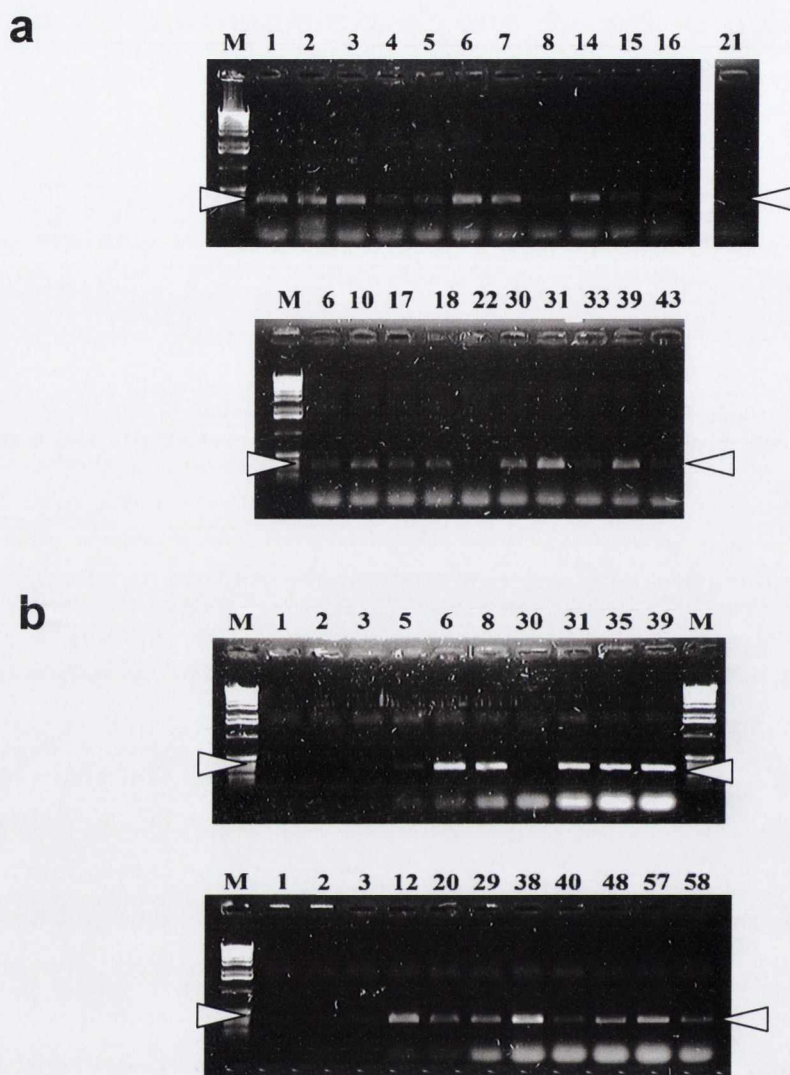


Figure 7.8 PCR screening of putative *N. tabacum* and *N. benthamiana* transformants containing Short Fragment (SF) constructs.

- (a) Agarose gel showing PCR products (approx. 500bp) resulting from the screening of putative *N. tabacum* cv Samsun (upper gel) and *N. benthamiana* (lower gel) Ap-CLD SF transformants, using gene-specific primers. PCR using DNA from a non-transgenic control plant did not yield any product (data not shown).
- (b) Agarose gel showing PCR products (approx. 500bp) resulting from the screening of putative *N. tabacum* cv Samsun (upper gel) and *N. benthamiana* (lower gel) Ap-CBE SF transformants, using gene-specific primers. PCR using DNA from a non-transgenic control plant did not yield any product (data not shown).

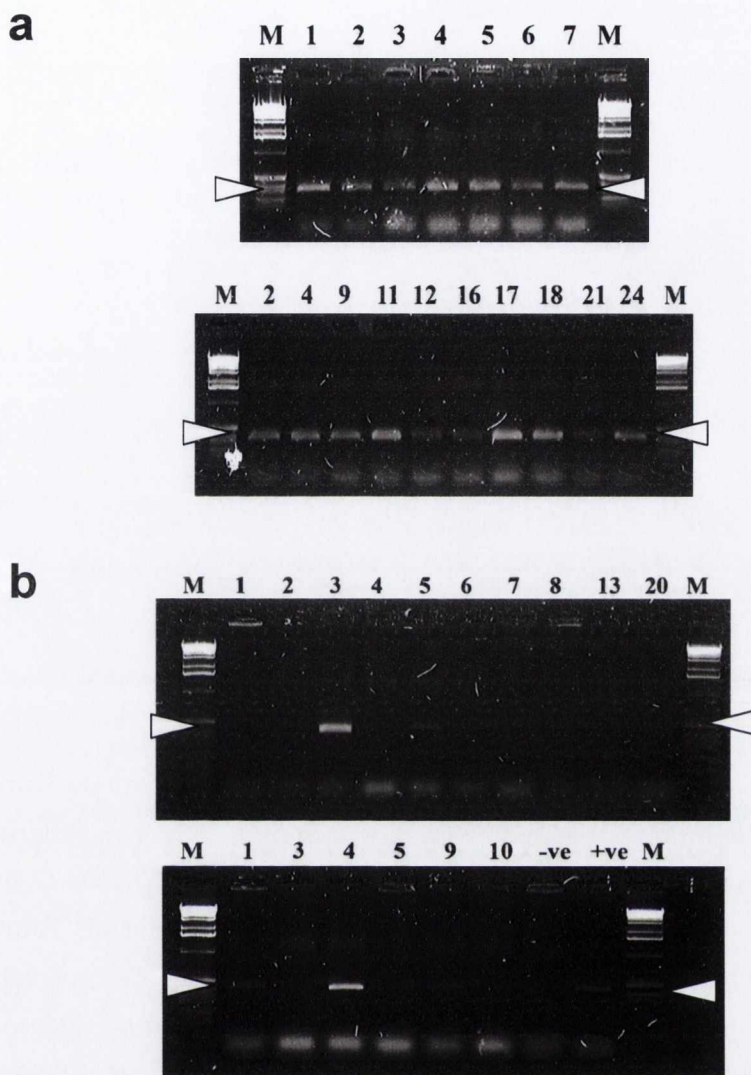


Figure 7.9 PCR screening of putative *N. tabacum* and *N. benthamiana* transformants containing Long Fragment (LF) constructs.

- (a) Agarose gel showing PCR products (approx. 500bp) resulting from the screening of putative *N. tabacum* cv Samsun (upper gel) and *N. benthamiana* (lower gel) Ap-CLD LF transformants using gene specific primers. PCR using DNA from a non-transgenic control plant did not yield any product (data not shown).
- (b) Agarose gel showing PCR products (approx. 500bp) resulting from the screening of putative *N. tabacum* cv Samsun (upper gel) and *N. benthamiana* (lower gel) Ap-CBE LF transformants using gene specific primers. -ve, PCR of DNA from a non-transgenic control plant.

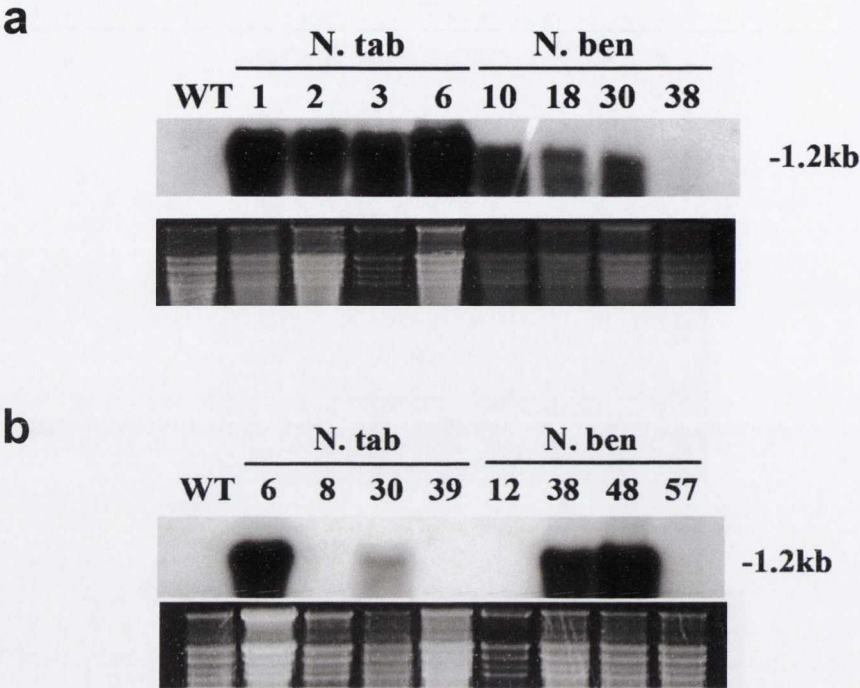


Figure 7.10 Northern blot analysis of total RNA from *N. tabacum* cv Samsun and *N. benthamiana* transformants containing Short Fragment (SF) constructs. Plants were stably transformed with (a) Ap-CLD SF and (b) Ap-CBE SF constructs. WT, non-transformed control. Total RNA was isolated from leaves and analysed by probing the blot with a gene-specific probe. Approximately 10 µg of total RNA was loaded onto each lane. Ethidium bromide stained gels of the total RNA were used as loading controls (bottom panel).

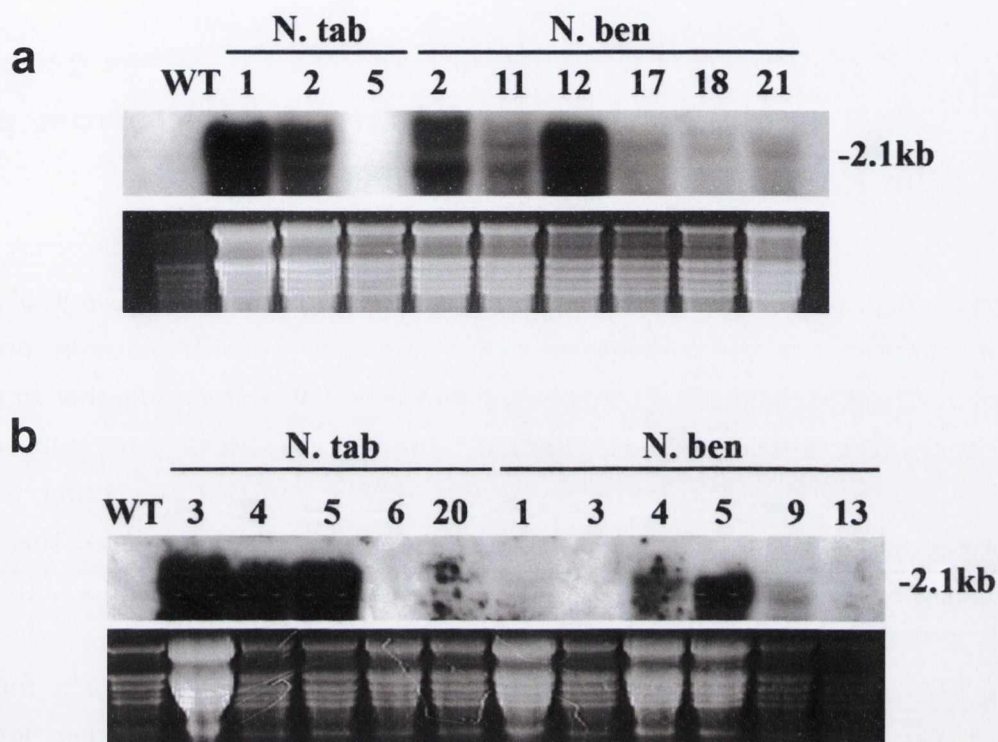


Figure 7.11 Northern blot analysis of total RNA from *N. tabacum* cv Samsun and *N. benthamiana* transformants containing Long Fragment (LF) constructs. Plants were stably transformed with (a) Ap-CLD LF and (b) Ap-CBE LF constructs. WT, non-transformed control. Total RNA was isolated from leaves, blotted and analysed by probing the blot with a gene-specific probe. Approximately 10 µg of total RNA was loaded onto each lane. Ethidium bromide stained gels of the total RNA were used as loading controls (bottom panel).

7.3 Discussion

7.3.1 Difficulties in the construction of transgenes generated from amplified cDNA or recombined DNA from plants infiltrated with ICON provector modules

Infiltration of leaves with *Agrobacterium* strains carrying the relevant 5' and 3' ICON provector modules (and the C31 integrase gene) results in provector recombination in the plant cell nucleus, joining two half-introns together to give an Ap-intron-GOI transgene configuration. Intron sequences are subsequently removed from the primary mRNA transcripts of Ap-intron-GOI by splicing, to give a shorter transcript containing a continuous Ap-GOI open reading frame. This mature mRNA is then massively amplified in the cytoplasm by the tobamovirus RdRP.

In order to speed up the process of making VHH transgene constructs for stable plant transformation experiments, the traditional route of sub-cloning individual vector components separately into plasmids, followed by restriction site-mediated assembly of the relevant sub-fragments was not used. Instead, cDNA copies of Ap-GOI were generated by RT-PCR amplification of RNA isolated from infiltrated leaves, and Ap-intron-GOI fragments were obtained by amplification of DNA from the same source.

However, nucleotide sequencing of the cloned PCR fragments revealed multiple point and frameshift mutations. It was subsequently discovered that the start site of the AP-GOI fragment was in frame with the translational start site of *lacZa* in pUC19 (Yanisch-Perron *et al.*, 1985), resulting in potential expression of the AP-GOI fragment which may have been detrimental to *E. coli*. This might explain the low *E. coli* survival rate and the high frequency of mutations in the plasmids: perhaps only rare clones that contained a mutated version of the AP-GOI construct could survive. This problem was resolved when the AP-GOI construct was placed out-of-frame relative to the *lacZa* start site. However, some point mutations were still present in the GOI as well as the Ap- presequence. This phenomenon was also observed with the SF and LF fragments amplified from *in planta* recombined DNA templates. Each clone possessed about 1-8 point substitutions or deletions, with the fragments amplified from DNA having fewer mutations (1-3 mutations) compared to the

fragments amplified from cDNA (2-8 mutations). In the case of Ap-CBE, while these errors might have occurred occur during the RT-PCR and cloning processes, another possible explanation is that the tobamoviral RdRP which amplifies the GOI mRNA in the cytoplasm lacks a proof reading function and hence is more prone to errors, and that this leads to frequent nucleotide changes (Ward *et al.*, 1988; Steinhauer *et al.*, 1992).

7.3.2 Expression of VHH transgenes in stably transformed *N. tabacum* and *N. benthamiana* plants

An Ap-CBE cDNA transgene, placed under the transcriptional control of the CaMV35S promoter or the *rbcs* promoter, was initially chosen for production of stable *N. tabacum* cv Samsun and *N. benthamiana* transformants because of the high-level accumulation of the encoded VHH in leaves following transient expression. Later, when it was found that transgenic lines with this cDNA-based transgene had no detectable CBE protein, two genomic versions of Ap-CBE (and Ap-CLD), containing the intron generated *in planta* by recombination between ICON 5' and 3' vector modules, were also investigated. However, for all of the constructs assessed in stably transformed lines, although significant transgene expression was observed at the mRNA level, accumulation of the encoded VHH was not observed in any line.

To investigate whether the 5' and 3' UTRs of CPMV and TMV might affect the accumulation of VHHs, three constructs were made: an Ap-CBE cDNA fragment flanked by CPMV 5' HT leader and 3' UTR (Sainsbury and Lomonossoff, 2008) as well as AP-Intron-GOI (GOI being CLD or CBE) flanked by tobamovirus 5' UTR and 3' NTR sequences. Because the construct containing the CPMV 5' HT leader and 3' UTR (pEAQ-HT; Sainsbury and Lomonossoff, 2008) had not previously been used for stable transformation, we first used it to investigate GFP expression in stably transformed *N. tabacum* and *N. benthamiana* plants. While GFP was expressed in these plants and accumulated in the apoplast in the case of Ap-GFP, the presence of the CPMV 5' and 3' UTR did not result in increased accumulation of GFP, relative to a standard CaMV35S-GFP expression cassette which lacked them. This is in stark contrast to the very high levels of GFP observed in plant cells transiently transformed with the pEAQ-GFP construct (Sainsbury and Lomonossoff, 2008),

suggesting that transgene silencing, even in the presence of the P19 suppressor of RNA silencing, is likely responsible for the observed difference.

To our surprise, the constructs containing CPMV and tobamoviral 5' UTR and 3' UTR (or 3' NTR) enhancers of translation did not yield transgenic lines with detectable levels of VHH protein. This is in contrast to Qin *et al.* (2000) for example, who reported that in stably- and transiently- transformed tobacco plants the 5' UTR of TMV (the so-called Ω fragment) not only enhances translation when used alone, but can act with the TMV 3' NTR to enhance the expression of GUS. Furthermore, the tobamovirus 3' NTR and the 3' NTR of some other viruses form pseudoknots that have been demonstrated to enhance translation efficiencies (Gallie *et al.*, 1991; Gallie, 1996). More recently, Sainsbury and Lomonosoff (2008) reported that a modified version of the CPMV 5' UTR, in which an in-frame upstream ATG was removed, resulted in a more than 10-fold increase in the accumulation of GFP in transiently infected leaves (20 – 25% tsp; Sainsbury and Lomonosoff, 2008), relative to the wild-type 5' UTR sequence. Moreover, this occurred without a change in the relative levels of GFP mRNA, clearly indicating that a change in translational efficiency was responsible for the effect.

In the stably transformed *N. tabacum* cv Samsun and *N. benthamiana* lines obtained for all of the VHH constructs investigated, the insert was detected in many plants by PCR screening. However, the level of mRNA produced in these plants varied greatly from high levels to no expression at all. This was most likely due to insert position effects and transgene-induced gene silencing (Matzke and Matzke, 1998; Matzke and Matzke, 1998; Vaucheret *et al.*, 1998). Furthermore, except for the Ap-CLD SF construct where mRNA expression in *N. benthamiana* was slightly lower than that observed in *N. tabacum* lines, there was generally no significant difference in mRNA levels between transgenic *N. tabacum* and *N. benthamiana* plants, whether the P19 suppressor of silencing was present or not (pEAQ constructs vs pBIN2, pBIN8 constructs, respectively). Conley *et al.* (2010) also found that the type of *Nicotiana* species and variety used for stable, as distinct from transient, transformation had little impact on the yield of recombinant proteins.

Our failure to detect any VHH accumulation, using either Western blotting or ELISA, in stably transformed lines was very disappointing and the reasons for this are not entirely clear. One possibility is that the immunological detection methods used were not sufficiently sensitive to detect very low level VHH expression in stably transformed plants. Other possibilities include VHH degradation *in planta* by

unknown proteases (Outchkourov *et al.*, 2003; Faye *et al.*, 2005; Doran, 2006; Benchabane *et al.*, 2008) or misfolding and aggregation due to incompatibilities with the subcellular environment and subsequent degradation by the ER-associated degradation pathway (ERAD) (Vembar and Brodsky, 2008; He *et al.*, 2011). The ER has been shown to be the optimal sub-cellular compartment for facilitating proper protein folding and assembly (Fischer and Emans, 2000), and when an anti-TNF α VHH was expressed as a single monomeric protein, was found to be the only compartment in which VHH accumulation was detected (Winichayakul *et al.*, 2009). The oxidizing environment of the ER is particularly suited for protein assembly because it (a) facilitates inter- and intramolecular disulphide bond formation and hence correct folding, (b) has a comparatively lower level of proteolytic activity and (c) contains molecular chaperones that are critical for correct folding, assembly and post-translational processing of antibodies (Ma *et al.*, 2003).

A possible route to increasing VHH yield might be to create protein fusions. Obregon *et al.* (2006) expressed a fusion protein consisting of the HIV p24 core antigen fused to a human IgA fragment containing the heavy chain domains and found that the expression of the fusion protein was 13-fold higher than when HIV p24 was expressed alone. In the case of VHH, monomers can be fused to generate tandem heterodimers. Winichayakul *et al.* (2009) found that when VHHs were expressed in tandem, they accumulated to a higher level (an estimated of 0.1–0.7% TSP) compared to when the VHHs were expressed as monomers, and not only did they accumulate in the ER but they were also found in the cytoplasm, the chloroplast and protein storage vacuoles in *Arabidopsis*. Another possible fusion partner is the elastin-like peptide (ELP). Several ELP fusion proteins have been successfully expressed in plants (Scheller *et al.*, 2004; Patel *et al.*, 2007; Floss *et al.*, 2008; Floss *et al.*, 2009; Conrad *et al.*, 2011) and have been found to accumulate to up to 20% TSP in seeds (Scheller *et al.*, 2006). Moreover, they can be purified inexpensively and have been found not to compromise the folding and activity of the fusion partner (Floss *et al.*, 2008). Conrad *et al.* (2011) recently reported that fusing ELP to an anti-tumour necrosis factor α (anti-TNF- α) VHH significantly increased protein accumulation (from 0.003 to 1.7% TSP), and importantly, the VHH fusion was shown to be effective in the prevention of death caused by septic shock in mice. The ELP fusion protein was also found to have a longer serum half-life compared to non-ELPylated VHHs.

Chapter 8. Conclusion and Future Work

The primary aim of this thesis was to investigate factors influencing the expression and accumulation of antibody fragments, including single-chain variable fragments (scFvs) and Camel heavy-chain only variable fragments (VHHs), in *N. benthamiana* and *N. tabacum*, using both transient and stable expression strategies. I began with a project to investigate the possibility of expressing two human scFvs and two Camel VHHs (or nanobodies) in *N. benthamiana* plants using the 'deconstructed virus'-based transient expression system developed by ICON Genetics. The two human scFvs recognized the same epitope of the antigen β -galactosidase. However, one of them was a mutant selected for efficient expression, folding and functionality in the reducing cytoplasm of *E. coli*, while the other was the original wild-type clone. Up to 0.3 mg/g fresh weight was obtained for the mutant anti β -galactosidase scFv, while the yield was slightly lower at 0.1 mg/g fresh weight for the wild-type. It was also found that maximum accumulation of the two anti β -galactosidase scFvs occurred in different subcellular compartments: the chloroplast for the wild-type scFv, but the cytoplasm for the mutant scFv. The two scFvs differ in that folding of the wild-type version requires correct disulphide bond formation between conserved cysteine residues, whereas folding of the mutant scFv does not. The wild-type version was presumably, therefore, more stable in the relatively oxidizing environment of the chloroplast (compared with the relatively more reducing environment of the cytoplasm). Thus, the accumulation profiles of both anti β -galactosidase scFvs differed from previously published reports suggesting that the secretory pathway is the most suitable for folding and functionality of scFvs. It is well known that the subcellular compartment in which a protein accumulates can affect protein stability and, in consequence, its overall accumulation in plants. This is likely due to the fact that different subcellular compartments vary not only in their redox status, but also in the spectrum of proteolytic activities they contain. Moreover, when the scFvs were expressed with different purification tags, it was found that the presence of a tag affected accumulation of the wild-type but not the mutant scFv.

On the other hand, the two Camel VHHs (or nanobodies) initially investigated bound different antigens i.e. hen egg-white lysozyme (HEWL) and tetanus toxoid from *Clostridium tetani*. We found that for the anti-tetanus toxoid VHH, protein accumulation was not high enough to be detected by Western blot or ELISA. The

inability to obtain soluble protein for this particular VHH might be attributable to post-translational processing and/or degradation, since Northern blots showed significant levels of mRNA transcripts present in the infiltrated leaves. In contrast, for the anti-HEWL VHH, the yield initially obtained was up to 0.05 mg/g fresh weight for a Strep-tagged version targeted to the apoplast, while no protein was detected for a His-tagged version. We then designed a variant with a modified N-terminus (CLD) to better match that of the original clone cAb-Lys3, which surprisingly showed moderate levels of accumulation in the apoplast. This His-tagged cAb-Lys3 variant was, therefore, used in subsequent detailed studies to investigate the effects of codon optimization and amino acid changes in the framework regions, on VHH yield *in planta*.

Based on previous experience with a recombinant bacterial sialidase that could not be expressed in plants without codon-optimization, we also compared the expression of the native CLD gene and the codon-optimized derivative, CLDSyn. Disappointingly, the redesigned gene showed no clear advantage over the original unmodified gene. However, spurred on by the results of preliminary experiments with the scFvs suggesting that changes in tertiary structure, due to the presence or absence of intramolecular disulphide bonds, affected antibody fragment accumulation in different subcellular compartments, we investigated whether the use of an alternative 'universal' VHH framework, known to alter the tertiary structure of the anti-HEWL VHH, might influence VHH yield *in planta*. The resulting nanobodies were expressed in the cytoplasm or were targeted to the apoplast, chloroplast or ER. The effect of different purification tags on overall protein accumulation was also investigated. Of all the combinations investigated, the apoplast-targeted, His-tagged anti-HEWL VHH with the 'universal' framework (CBESynHS) showed the highest level of accumulation, reaching an estimated 30% total soluble leaf protein at approximately 7 dpi, several-fold higher than the His-tagged native VHH, CLDHS. Moreover, while small quantities of a lower molecular weight form of CLDHS was evident on protein gels, this was not observed with CBESynHS, which suggests enhanced resistance to proteolytic attack. Another possible explanation for this, as well as for the differences in accumulation between CLDHS and CBESynHS, might be differences in tertiary structure which might, in turn, influence VHH folding and interactions with other proteins in the subcellular environment. Also interesting to note was the change in the protein accumulation profile when CBESyn was fused to GFP. Whereas CBESyn on its own accumulated showed maximum accumulation in the apoplast, the fusion protein accumulated most highly in the cytoplasm, followed by the chloroplast.

After we established that the use of the 'universal' framework increased the accumulation of the anti-HEWL nanobody *in planta*, we proceeded to determine whether the framework effect was specific for that VHH or might be applied to nanobodies in general. Two pairs of nanobodies, one an anti-prostate specific antigen (PSA) VHH and the other an anti-RNase A VHH, either in their codon-optimized native framework or the alternative cAbBCII10 framework were constructed and their expression *in planta* investigated. It was found that the anti-PSA nanobodies accumulated highly in the cytoplasm compared with other subcellular compartments. The anti-RNase A VHHs on the other hand accumulated highly only when targeted to the apoplast. In both cases, the framework-changed versions had higher yields than their counterparts containing the original framework. We postulate two possibilities for this phenomenon. The cAbBCII10 framework region was originally proposed as a 'universal' framework because of its enhanced thermodynamic stability, its good expression levels in *E. coli* and its ability to retain functionality even in the absence of a conserved disulphide bond. These properties might account for the high yield and improved resistance to proteolysis of the *in planta*-expressed VHHs. Another possibility is that subtle alterations in tertiary structure induced by the 'universal' framework render them more resistant to proteases compared to their native counterpart, and hence enhanced stability. Crystallographic data has shown that anti-HEWL nanobodies based on the original, compared with the 'universal' framework adopt slightly different conformations. To further investigate framework effects on VHH accumulation, we constructed 6 anti-HEWL chimeras containing different combinations of FRs from cAbBCII10 and the original framework. It was found that the 6 chimeras showed different levels of accumulation, and that some were more susceptible to proteolytic degradation than others. It would be interesting to obtain crystallographic data for these 6 chimeras in order to determine if tertiary structure correlates with the overall yield of the different VHHs.

We then went on to compare VHH accumulation not only between the deconstructed-virus based transient expression systems ICON, pEAQ-HT and pJL-TRBO, but also between the viral-based systems and conventional binary vectors containing the CaMV35S or RUBISCO small subunit (*rbcS*) promoter. We found that among the three transient expression systems, the highest yields were obtained with the ICON system compared with pEAQ-HT and pJL-TRBO, especially when it was used to express anti-HEWL and anti-PSA VHHs. However, VHHs expressed using the ICON system took longer (several more days) to reach maximum protein

accumulation compared with pEAQ-HT and pJL-TRBO, presumably due to the time required for recombination of the provectors *in planta*. On the other hand, no protein accumulation was detected when nanobodies were expressed using the CaMV35S or *rbcS* promoters. Furthermore, we also compared the expression of anti-HEWL and anti-RNase A VHs between two different expression platforms, *E. coli* and plants. The nanobodies were targeted to the apoplast in plants while in *E. coli*, they were targeted to the periplasmic space. Overall, although expression in *E. coli* took less time, the yield was several times higher in plants. Furthermore, an additional band was observed on Western blots in addition to the band of interest. Further steps to purify the nanobodies from the periplasmic space of *E. coli* could be undertaken to investigate whether this additional band accumulates in the periplasmic space or elsewhere in *E. coli*. However, it is interesting to note that no matter which expression vector or expression platform was used, nanobodies with the alternative cAbBCII10 framework showed higher protein accumulation compared with their native framework counterparts.

Finally, we also investigated VHH expression in stably transformed *N. tabacum* cv. Samsun and *N. benthamiana* plants. The anti-HEWL VHH CBESynHS was chosen because of its high yield in transient expression experiments, especially when targeted to the apoplast. Several strategies were investigated, including the use of different promoters (CaMV35S and *rbcS* promoters), the inclusion of an intron in the transgene, as well as the use of 5' and 3' untranslated regions (UTRs) from Tobacco Mosaic Virus (TMV) and Cowpea Mosaic Virus (CPMV). The required components of the different constructs were amplified either from cDNA or from DNA isolated from *N. benthamiana* leaves transiently transformed using the ICON system. An in-frame alignment of the GOI ORF with the lacZ α translational start site in the cloning vector pUC19 caused some initial difficulties in the early stages of cloning. This was resolved by removing one base from the forward primer to put the ORF out of frame, relative to lacZ α . However, point mutations and deletions were still recovered in the resulting cDNA clones, which may have been caused by the lack of proof-reading functions in the RNA-dependant RNA polymerase (RdRP) encoded by the ICON vectors. The VHH constructs were subsequently 'repaired' by swapping mutated fragments with those having the correct sequence using internal restriction sites. When the constructs were stably transformed into *N. tabacum* cv. Samson and *N. benthamiana*, VHH protein could not be detected in the crude leaf extracts, despite the fact that Northern analysis showed the GOI was transcribed at significant levels in some transgenic lines. This led us to conclude that the encoded VHH protein may have been undetectable because of the combined effects of

inefficient translation and proteolytic degradation. Possible solutions to this problem might include expressing the VHHs in tandem or as fusion proteins, in addition to targeting the VHHs to the ER. An alternative strategy might be to express VHH fusion proteins from transgenes located in the chloroplast genome.

References

- Achour I, Cavelier P, Tichit M, Bouchier C, Lafaye P, Rougeon F. 2008. Tetrameric and homodimeric Camelid IgGs originate from the same IgH locus. *Journal of Immunology* 181:2001-09.
- Adams GP, Weiner LM. 2005. Monoclonal antibody therapy of cancer. *Nature Biotechnology* 23:1147-57.
- Amarasinghe B, Faivre-Nitschke E, Wu Y, Udall JA, Dennis ES, Constable G, Llewellyn DJ. 2006. Genomic approaches to the discovery of promoters for sustained expression in cotton (*Gossypium hirsutum* L.) under field conditions: expression analysis in transgenic cotton and *Arabidopsis* of a Rubisco small subunit promoter identified using EST sequence analysis and cDNA microarrays. *Plant Biotechnology* 23:437-50.
- Andersen DC, Reilly DE. 2004. Production technologies for monoclonal antibodies and their fragments. *Current Opinion in Biotechnology* 15:456-462
- Andrianov V, Borisjuk N, Pogrebnyak N, Brinker A, Dixon J, Spitsin S, Flynn J, Matyszczyk P, Andryszak K, Laurelli M. 2010. Tobacco as a production platform for biofuel: overexpression of *Arabidopsis* DGAT and LEC2 genes increases accumulation and shifts the composition of lipids in green biomass. *Plant Biotechnology Journal* 8:277-87.
- Arbabi-Ghahroudi M, Desmyter A, Wyns L, Hamers R, Muyldermans S. 1997. Selection and identification of single domain antibody fragments from camel heavy-chain antibodies. *FEBS Letters* 414:521-6.
- Arbabi-Ghahroudi M, Tanha J, MacKenzie R. 2005. Prokaryotic expression of antibodies. *Cancer and Metastasis Reviews* 24:501-519.
- Arencibia A, Molina PR, de la Riva G, Selman-Housein G. 1995. Production of transgenic sugarcane (*Saccharum officinarum* L.) plants by intact cell electroporation. *Plant Cell Reports* 14:305-09.
- Artsaenko O, Kettig B, Fiedler U, Conrad U, Düring K. 1998. Potato tubers as a biofactory for recombinant antibodies. *Molecular Breeding* 4:313-19.
- Avesani L, Marconi G, Morandini F, Albertini E, Bruschetta M, Bortesi L, Pezzotti M, Porceddu A. 2007. Stability of Potato virus X expression vectors is related to insert size: implications for replication models and risk assessment. *Transgenic research* 16:587-97.
- Azhakanandam K, Weissinger SM, Nicholson JS, Qu R, Weissinger AK. 2007. Amplicon-plus targeting technology (APTT) for rapid production of a highly unstable vaccine protein in tobacco plants. *Plant Molecular Biology* 63:393-404.
- Badri MA, Rivard D, Coenen K, Michaud D. 2009. Unintended molecular interactions in transgenic plants expressing clinically useful proteins: The case of bovine aprotinin traveling the potato leaf cell secretory pathway. *Proteomics* 9:746-56.

- Bang S, Nagata S, Onda M, Kreitman RJ, Pastan I. 2005. HA22 (R490A) is a recombinant immunotoxin with increased antitumor activity without an increase in animal toxicity. *Clinical Cancer Research* 11:1545-50.
- Baral TN, Magez S, Stijlemans B, Conrath K, Vanhollebeke B, Pays E, Muyldermans S, De Baetselier P. 2006. Experimental therapy of African trypanosomiasis with a nanobody-conjugated human trypanolytic factor. *Nature Medicine* 12:580-84.
- Barbante A, Irons S, Hawes C, Frigerio L, Vitale A, Pedrazzini E. 2008. Anchorage to the cytosolic face of the endoplasmic reticulum membrane: a new strategy to stabilize a cytosolic recombinant antigen in plants. *Plant Biotechnology Journal* 6:560-75.
- Barbi T, Drake PMW, Drever M, van Dolleweerd CJ, Porter AR, Ma JKC. 2010. Generation of transgenic plants expressing plasma membrane-bound antibodies to the environmental pollutant microcystin-LR. *Transgenic Research ePub* (ahead of print).
- Baulcombe DC, Saunders GR, Bevan MW, Mayo MA, Harrison BD. 1986. Expression of biologically active viral satellite RNA from the nuclear genome of transformed plants. *Nature* 321:446-49.
- Bazi MR, Rasaee MJ, Foruzandeh M, Rahimpour A, Kiani J, Rahbarizadeh F, Alirezapour B, Mohammadi M. 2007. Production of chimeric recombinant single domain antibody-green fluorescent fusion protein in Chinese hamster ovary cells. *Hybridoma* 26:1-9.
- Belknap WR, Rockhold DR, McCue KF. 2008. pBINPLUS/ARS: an improved plant transformation vector based on pBINPLUS. *BioTechniques* 44:753-58.
- Benchabane M, Goulet C, Rivard D, Faye L, Gomord V, Michaud D. 2008. Preventing unintended proteolysis in plant protein biofactories. *Plant Biotechnol Journal* 6:633-48.
- Bendandi M, Marillonnet S, Kandzia R, Thieme F, Nickstadt A, Herz S, Fröde R, Inogés S, López-Díaz de Cerio A, Soria E. 2010. Rapid, high-yield production in plants of individualized idiotypic vaccines for non-Hodgkin's lymphoma. *Annals of Oncology* 21:2420-27.
- Bock R. 2007. Plastid biotechnology: prospects for herbicide and insect resistance, metabolic engineering and molecular farming. *Current Opinion in Biotechnology* 18:100-6.
- Bond CJ, Marsters Jr JC, Sidhu SS. 2003. Contributions of CDR3 to VHH domain stability and the design of monobody scaffolds for naive antibody libraries. *Journal of Molecular Biology* 332:643-55.
- Braren I, Blank S, Seismann H, Deckers S, Ollert M, Grunwald T, Spillner E. 2007. Generation of human monoclonal allergen-specific IgE and IgG antibodies from synthetic antibody libraries. *Clinical Chemistry* 53:837-44.
- Brown JW, Simpson CG, Thow G, Clark GP, Jennings SN, Medina-Escobar N, Haupt S, Chapman SC, Oparka KJ. 2002. Splicing signals and factors in plant intron removal. *Biochemical Society Transactions* 30:146-9.

- Burgess-Brown NA, Sharma S, Sobott F, Loenarz C, Oppermann U, Gileadi O. 2008. Codon optimization can improve expression of human genes in *Escherichia coli*: A multi-gene study. *Protein Expression and Purification* 59:94-102.
- Cardinale A, Lener M, Messina S, Cattaneo A, Biocca S. 1998. The mode of action of Y13-259 scFv fragment intracellularly expressed in mammalian cells. *FEBS Letters* 439:197-202.
- Chan PH, Pardon E, Menzer L, De Genst E, Kumita JR, Christodoulou J, Saerens D, Brans A, Bouillenne F, Archer DB, Robinson CV, Muyldermans S, Matagne A, Redfield C, Wyns L, Dobson CM, Dumoulin M. 2008. Engineering a Camelid antibody fragment that binds to the active site of human lysozyme and inhibits its conversion into amyloid fibrils. *Biochemistry* 47:11041-54
- Chapman S, Kavanagh T, Baulcombe D. 1992. Potato virus X as a vector for gene expression in plants. *Plant Journal* 2:549 - 57.
- Chen Q, He J, Phoolcharoen W, Mason HS. 2011. Geminiviral vectors based on bean yellow dwarf virus for production of vaccine antigens and monoclonal antibodies in plants. *Human Vaccine* 7:331-38.
- Chothia C, Lesk AM. 1987. Canonical structures for the hypervariable regions of immunoglobulins. *Journal of Molecular Biology* 196:901-17.
- Chothia C, Lesk AM, Gherardi E, Tomlinson IM, Walter G, Marks JD, Llewelyn MB, Winter G. 1992. Structural repertoire of the human VH segments. *Journal of Molecular Biology* 227:799-817.
- Ciechanover A. 1998. The ubiquitin-proteasome pathway: on protein death and cell life. *EMBO Journal* 17:7151-60.
- Clough SJ, Bent AF. 1998. Floral dip: a simplified method for *Agrobacterium* mediated transformation of *Arabidopsis thaliana*. *Plant Journal* 16:735-43.
- Conley AJ, Zhu H, Le LC, Jevnikar AM, Lee BH, Brandle JE, Menassa R. 2010. Recombinant protein production in a variety of *Nicotiana* hosts: a comparative analysis. *Plant Biotechnology Journal* ePub (ahead of print).
- Conrad U, Fiedler U. 1998. Compartment-specific accumulation of recombinant immunoglobulins in plant cells: an essential tool for antibody production and immunomodulation of physiological functions and pathogen activity. *Plant Molecular Biology* 38:101-09.
- Conrad U, Floss DM. 2010. Expression of Antibody Fragments in Transgenic Plants. Berlin: Springer-Verlag. 377-86.
- Conrad U, Plagmann I, Malchow S, Sack M, Floss DM, Kruglov AA, Nedospasov SA, Rose-John S, Scheller J. 2011. ELPylated anti-human TNF therapeutic single-domain antibodies for prevention of lethal septic shock. *Plant Biotechnology Journal* 9:22-31.
- Conrath K, Vincke C, Stijlemans B, Schymkowitz J, Decanniere K, Wyns L, Muyldermans S, Loris R. 2005. Antigen binding and solubility effects upon the veneering of a camel VHH in framework-2 to mimic a VH. *Journal of Molecular Biology* 350:112-25.

- Conrath KE, Lauwereys M, Galleni M, Matagne A, Frere J-M, Kinne J, Wyns L, Muyldermans S. 2001. beta-Lactamase Inhibitors Derived from Single-Domain Antibody Fragments Elicited in the *Camelidae*. *Antimicrobial Agents and Chemotherapy* 45:2807-12.
- Coppieters K, Dreier T, Silence K, de Haard H, Lauwereys M, Casteels P, Beirnaert E, Jonckheere H, Van de Wiele C, Staelens L, Hostens J, Revets H, Remaut E, Elewaut D, Rottiers P. 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 and Rheumatism* 54:1856-66.
- Cortez-Retamozo V, Backmann N, Senter PD, Wernery U, De Baetselier P, Muyldermans S, Revets H. 2004. Efficient cancer therapy with a nanobody-based conjugate. *Cancer Research* 64:2853-57.
- Creager ANH, Scholthof KBG, Citovsky V, Scholthof HB. 1999. Tobacco mosaic virus: pioneering research for a century. *The Plant Cell* 11:301-08.
- Dale EC, Ow DW. 1991. Gene transfer with subsequent removal of the selection gene from the host genome. *Proceedings of the National Academy of Sciences USA* 88:10558-62.
- Daniell H, Kumar S, Dufourmantel N. 2005. Breakthrough in chloroplast genetic engineering of agronomically important crops. *Trends in Biotechnology* 23:238-45.
- Datta K, Baisakh N, Oliva N, Torrizo L, Abrigo E, Tan J, Rai M, Rehana S, Al Babili S, Beyer P. 2003. Bioengineered 'golden' indica rice cultivars with carotene metabolism in the endosperm with hygromycin and mannose selection systems. *Plant Biotechnology Journal* 1:81-90.
- Davies J, Riechmann L. 1994. 'Camelising' human antibody fragments: NMR studies on VH domains. *FEBS Letters* 339:285-90.
- Davies J, Riechmann L. 1995. Antibody VH domains as small recognition units. *Nature Biotechnology* 13:475-79.
- Davies J, Riechmann L. 1996. Single antibody domains as small recognition units: design and in vitro antigen selection of camelized, human VH domains with improved protein stability. *Protein Engineering* 9:531-37.
- Dawson WO, Lewandowski DJ, Hilf ME, Bubrick P, Raffo AJ, Shaw JJ, Grantham GL, Desjardins PR. 1989. A tobacco mosaic virus-hybrid expresses and loses an added gene. *Virology* 172:285-92.
- De Genst E, Silence K, Decanniere K, Conrath K, Loris R, Kinne J, Muyldermans S, Wyns L. 2006. Molecular basis for the preferential cleft recognition by dromedary heavy-chain antibodies. *Proceedings of the National Academy of Sciences USA* 103:4586-91.
- De la Pena A, Lörz H, Schell J. 1987. Transgenic rye plants obtained by injecting DNA into young floral tillers. *Nature* 325:274-76.

- De Wilde C, Peeters K, Jacobs A, Peck I, Depicker A. 2002. Expression of antibodies and Fab fragments in transgenic potato plants: a case study for bulk production in crop plants. *Molecular Breeding* 9:271-82.
- De Wilde C, Van Houdt H, De Buck S, Angenon G, De Jaeger G, A. D. 2000. Plants as bioreactors for protein production: avoiding the problem of transgene silencing. *Plant Molecular Biology* 43:347-59.
- Decanniere K, Desmyter A, Lauwereys M, Ghahroudi MA, Muyldermans S, Wyns L. 1999. A single-domain antibody fragment in complex with RNase A: non-canonical loop structures and nanomolar affinity using two CDR loops. *Structure* 7:361-70.
- Decanniere K, Muyldermans S, Wyns L. 2000. Canonical antigen-binding loop structures in immunoglobulins: more structures, more canonical classes? *Journal of Molecular Biology* 300:83-91.
- Delannoy M, Alves G, Vertommen D, Ma J, Boutry M, Navarre C. 2008. Identification of peptidases in *Nicotiana tabacum* leaf intercellular fluid. *Proteomics* 8:2285-98.
- Desmyter A, Decanniere K, Muyldermans S, Wyns L. 2001. Antigen Specificity and High Affinity Binding Provided by One Single Loop of a Camel Single-domain Antibody. *Journal of Biological Chemistry* 276:26285-90.
- Desmyter A, Spinelli S, Payan F, Lauwereys M, Wyns L, Muyldermans S, Cambillau C. 2002. Three Camelid VHH Domains in Complex with Porcine Pancreatic alpha-Amylase. Inhibition and versatility of binding topology *Journal of Biological Chemistry* 277:23645-50.
- Desmyter A, Transue TR, Ghahroudi MA, Thi MH, Poortmans F, Hamers R, Muyldermans S, Wyns L. 1996. Crystal structure of a camel single-domain VH antibody fragment in complex with lysozyme. *Nature Structural Biology* 3:803-11.
- Dolk E, van der Vaart M, Lutje Hulsik D, Vriend G, de Haard H, Spinelli S, Cambillau C, Frenken L, Verrips T. 2005. Isolation of llama antibody fragments for prevention of dandruff by phage display in shampoo. *Applied Environmental Microbiology* 71:442-50.
- Donson J, Kearney CM, Hilf ME, Dawson WO. 1991. Systemic expression of a bacterial gene by a tobacco mosaic virus-based vector. *Proceedings of the National Academy of Sciences USA* 88:7204-08.
- Doran PM. 2006. Foreign protein degradation and instability in plants and plant tissue cultures. *Trends in Biotechnology* 24:426-32.
- Dorokhov YL, Sheveleva AA, Frolova OY, Komarova TV, Zvereva AS, Ivanov PA, Atabekov JG. 2007. Superexpression of tuberculosis antigens in plant leaves. *Tuberculosis* 87:218-44.
- Drake PMW, Barbi T, Drever MR, van Dolleweerd CJ, Porter AJR, Ma JKC. 2010. Generation of transgenic plants expressing antibodies to the environmental pollutant microcystin-LR. *The FASEB Journal* 24:882-90.

- Dumoulin M, Conrath K, Van Meirhaeghe A, Meersman F, Heremans K, Frenken LG, Muyldermans S, Wyns L, Matagne A. 2002. Single-domain antibody fragments with high conformational stability. *Protein Science* 11:500-15.
- Dumoulin M, Last AM, Desmyter A, Decanniere K, Canet D, Larsson G, Spencer A, Archer DB, Sasse J, Muyldermans S, Wyns L, Redfield C, Matagne A, Robinson CV, Dobson CM. 2003. A camelid antibody fragment inhibits the formation of amyloid fibrils by human lysozyme. *Nature* 424:783-8.
- Ewert S, Cambillau C, Conrath K, Pluckthun A. 2002. Biophysical properties of camelid V(HH) domains compared to those of human V(H)3 domains. *Biochemistry* 41:3628-36.
- Faye L, Boulaflous A, Benchabane M, Gomord V, Michaud D. 2005. Protein modifications in the plant secretory pathway: current status and practical implications in molecular pharming. *Vaccine* 23:1770-78.
- Fernandez-San M, A., Mingo-Castel A, Miller M, Daniell H. 2003. A chloroplast transgenic approach to hyper-express and purify Human Serum Albumin, a protein highly susceptible to proteolytic degradation. *Plant Biotechnology Journal* 1:71-79.
- Fiedler U, Phillips J, Artsaenko O, Conrad U. 1997. Optimization of scFv antibody production in transgenic plants. *Immunotechnology* 3:205-16.
- Fischer R, Drossard J, Commandeur U, Schillberg S, Emans N. 1999. Towards molecular farming in the future: moving from diagnostic protein and antibody production in microbes to plants. *Biotechnology and Applied Biochemistry* 30 (Pt 2):101-8.
- Fischer R, Emans N. 2000. Molecular farming of pharmaceutical proteins. *Transgenic Research* 9:279-99.
- Fischer R, Schumann D, Zimmermann S, Drossard J, Sack M, Schillberg S. 1999. Expression and characterization of bispecific single-chain Fv fragments produced in transgenic plants. *European Journal of Biochemistry* 262:810-16.
- Floss DM, Sack M, Arcalis E, Stadlmann J, Quendler H, Rademacher T, Stöger E, Scheller J, Fischer R, Conrad U. 2009. Influence of elastin like peptide fusions on the quantity and quality of a tobacco derived human immunodeficiency virus neutralizing antibody. *Plant Biotechnology Journal* 7:899-913.
- Floss DM, Sack M, Stadlmann J, Rademacher T, Scheller J, Stöger E, Fischer R, Conrad U. 2008. Biochemical and functional characterization of anti-HIV antibody-ELP fusion proteins from transgenic plants. *Plant Biotechnology Journal* 6:379-91.
- Forsman A, Beirnaert E, Aasa-Chapman MMI, Hoorelbeke B, Hijazi K, Koh W, Tack V, Szynol A, Kelly C, McKnight A, Verrips T, Haard Hd, Weiss RA. 2008. Llama antibody fragments with cross-subtype Human Immunodeficiency Virus Type 1 (HIV-1)-neutralizing properties and high affinity for HIV-1 gp120. *Journal of Virology* 82:12069-81.

- Francesca P, Lorenza D. 2008. Investigating recombinant protein exudation from roots of transgenic tobacco. *Environmental Biosafety Research* 7:219-26.
- Frenken LG, van der Linden RH, Hermans PW, Bos JW, Ruuls RC, de Geus B, Verrips CT. 2000. Isolation of antigen specific llama VHH antibody fragments and their high level secretion by *Saccharomyces cerevisiae*. *Journal of Biotechnology* 78:11-21.
- Friguet B, Chaffotte AF, Djavadi-Ohanian L, Goldberg ME. 1985. Measurements of the true affinity constant in solution of antigen-antibody complexes by enzyme-linked immunosorbent assay. *Journal of Immunological Methods* 77:305-19.
- Fromm M, Taylor LP, Walbot V. 1985. Expression of genes transferred into monocot and dicot plant cells by electroporation. *Proceedings of the National Academy of Sciences USA* 82:5824-28.
- Fujiyama K, Furukawa A, Katsura A, Misaki R, Omasa T, Seki T. 2007. Production of mouse monoclonal antibody with galactose-extended sugar chain by suspension cultured tobacco BY2 cells expressing human beta(1,4)-galactosyltransferase. *Biochemical and Biophysical Research Communications* 358:85-91.
- Galeffi P, Lombardi A, Pietraforte I, Novelli F, Di Donato M, Sperandei M, Tornambe A, Fraioli R, Martayan A, Natali P, Benevolo M, Mottotese M, Ylera F, Cantale C, Giacomini P. 2006. Functional expression of a single-chain antibody to ErbB-2 in plants and cell-free systems. *Journal of Translational Medicine* 4:39-51.
- Gallie DR. 1996. Translational control of cellular and viral mRNAs. *Plant Molecular Biology* 32:145-58.
- Gallie DR, Feder JN, Schimke RT, Walbot V. 1991. Functional analysis of the tobacco mosaic virus tRNA-like structure in cytoplasmic gene regulation. *Nucleic Acids Research* 19:5031.
- Gamborg OL, Miller RA, Ojima K. 1968. Nutrient requirements of suspension cultures of soybean root cells. *Experimental Cell Research* 50:151-58.
- Gelvin SB. 2003. *Agrobacterium*-mediated plant transformation: The biology behind the "gene-jockeying" tool. *Microbiology and Molecular Biology Reviews* 67:16-37.
- Geyer BC, Fletcher SP, Griffin TA, Lopker MJ, Soreq H, Mor TS. 2007. Translational control of recombinant human acetylcholinesterase accumulation in plants. *BMC Biotechnology* 7:27-40.
- Geyer BC, Kannan L, Garnaoud PE, Broomfield CA, Cadieux CL, Cherni I, Hodgins SM, Kasten SA, Kelley K, Kilbourne J. 2010. Plant-derived human butyrylcholinesterase, but not an organophosphorous-compound hydrolyzing variant thereof, protects rodents against nerve agents. *Proceedings of the National Academy of Sciences USA* 107:20251-56.
- Gils M, Kandzia R, Marillonnet S, Klimyuk V, Gleba Y. 2005. High-yield production of authentic human growth hormone using a plant virus-based expression system. *Plant Biotechnology Journal* 8:613-20.

- Giritch A, Marillonnet S, Engler C, van Eldik G, Botterman J, Klimyuk V, Gleba Y. 2006. Rapid high-yield expression of full-size IgG antibodies in plants coinfecting with noncompeting viral vectors. *Proceedings of the National Academy of Sciences USA* 103:14701-06.
- Gleave AP, Mitra DS, Mudge SR, Morris BAM. 1999. Selectable marker-free transgenic plants without sexual crossing: transient expression of cre recombinase and use of a conditional lethal dominant gene. *Plant Molecular Biology* 40:223-35.
- Gleba Y, Klimyuk V, Marillonnet S. 2007. Viral vectors for the expression of proteins in plants. *Current Opinion in Biotechnology* 18:134-41.
- Goel A, Colcher D, Koo J-S, Booth BJM, Pavlinkova G, Batra SK. 2000. Relative position of the hexahistidine tag effects binding properties of a tumor-associated single-chain Fv construct. *Biochimica et Biophysica Acta (BBA) - General Subjects* 1523:13-20.
- Goldman ER, Anderson GP, Liu JL, Delehanty JB, Sherwood LJ, Osborn LE, Cummins LB, Hayhurst A. 2006. Facile generation of heat-stable antiviral and antitoxin single domain antibodies from a semisynthetic llama library. *Analytical chemistry*(Washington, DC) 78:8245-55.
- Gomord V, Denmat L-A, Fitchette-Lainé A-C, Satiat-Jeunemaitre B, Hawes C, Faye L. 1997. The C-terminal HDEL sequence is sufficient for retention of secretory proteins in the endoplasmic reticulum (ER) but promotes vacuolar targeting of proteins that escape the ER. *The Plant Journal* 11:313-25.
- Gomord V, Sourrouille C, Fitchette AC, Bardor M, Pagny S, Lerouge P, Faye L. 2004. Production and glycosylation of plant-made pharmaceuticals: the antibodies as a challenge. *Plant Biotechnology Journal* 2:83-100.
- Gonzales NR, De Pascalis R, Schlom J, Kashmiri SVS. 2005. Minimizing the immunogenicity of antibodies for clinical application. *Tumor Biology* 26:31-43.
- Gottlin EB, Xiangrong G, Pegram C, Cannedy A, Campa MJ, Patz EF, Jr. 2009. Isolation of Novel EGFR-Specific VHH Domains. *Journal of Biomolecular Screening* 14:77-85.
- Gurkhan C, Symeonides SN, Ellar DJ. 2004. High-level production in *Pichia pastoris* of an anti-p185HER-2 single-chain antibody fragment using an alternative secretion expression vector. *Biotechnology and Applied Biochemistry* 39:115-22.
- Halin C, Niesner U, Villani ME, Zardi L, Neri D. 2002. Tumor-targeting properties of antibody-vascular endothelial growth factor fusion proteins. *International Journal of Cancer* 102:109-16.
- Halpin C. 2005. Gene stacking in transgenic plants—the challenge for 21st century plant biotechnology. *Plant Biotechnology Journal* 3:141-55.
- Halpin C, Boerjan W. 2003. Stacking transgenes in forest trees. *Trends in Plant Science* 8:363-65.

- Hamers-Casterman C, Atarhouch T, Muyldermans S, Robinson G, Hamers C, Songa EB, Bendahman N, Hamers R. 1993. Naturally occurring antibodies devoid of light chains. *Nature* 363:446-8.
- Harmsen M, van Solt C, van Zijderveld-van Bemmelen A, Niewold T, van Zijderveld F. 2006. Selection and optimization of proteolytically stable llama single-domain antibody fragments for oral immunotherapy. *Applied Microbiology and Biotechnology* 72:544-51.
- Harmsen MM, De Haard HJ. 2007. Properties, production, and applications of camelid single-domain antibody fragments. *Applied Microbiology and Biotechnology* 77:13-22.
- Harmsen MM, Ruuls RC, Nijman IJ, Niewold TA, Frenken LG, de Geus B. 2000. Llama heavy-chain V regions consist of at least four distinct subfamilies revealing novel sequence features. *Molecular Immunology* 37:579-90.
- Harmsen MM, Van Solt CB, Fijten HPD, Van Setten MC. 2005. Prolonged in vivo residence times of llama single-domain antibody fragments in pigs by binding to porcine immunoglobulins. *Vaccine* 23:4926-34.
- He Z, Zhao Y, Mei G, Li N, Chen Y. 2011. Could protein tertiary structure influence mammary transgene expression more than tissue specific codon usage? *Transgenic Research* 19:519-33.
- Hellmann H, Estelle M. 2002. Plant development: regulation by protein degradation. *Science* 297:793.
- Hellwig S, Drossard J, Twyman RM, Fischer R. 2004. Plant cell cultures for the production of recombinant proteins. *Nature Biotechnology* 22:1415-22.
- Hiatt A, Cafferkey R, Bowdish K. 1989. Production of antibodies in transgenic plants. *Nature* 342:76-8.
- Hmila I, Abdallah R, Ben AB, Saerens D, Benlasfar Z, Conrath K, Ayeb ME, Muyldermans S, Bouhaouala-Zahar B. 2008. VHH, bivalent domains and chimeric Heavy chain-only antibodies with high neutralizing efficacy for scorpion toxin AahI. *Molecular immunology* 45:3847-56.
- Hoffmann M, Nemetz C, Madin K, Buchberger B. 2004. Rapid translation system: a novel cell-free way from gene to protein. *Biotechnology Annual Review* 10:1 - 30.
- Hoffmann P, Hofmeister R, Brischwein K, Brandl C, Crommer S, Bargou R, Itin C, Prang N, Baeuerle PA. 2005. Serial killing of tumor cells by cytotoxic T cells redirected with a CD19-/CD3-bispecific single-chain antibody construct. *International Journal of Cancer* 115:98-104.
- Höhl U, Upmeyer B, Barz W. 1988. Growth and nicotinate biotransformation in batch cultured and airlift fermenter grown soybean cell suspension cultures. *Applied Microbiology and Biotechnology* 28:319-23.
- Holliger P, Hudson PJ. 2005. Engineered antibody fragments and the rise of single domains. *Nature Biotechnology* 23:1126-36.

- Hondred D, Walker JM, Mathews DE, Vierstra RD. 1999. Use of ubiquitin fusions to augment protein expression in transgenic plants. *Plant Physiology* 119:713-24.
- Horsch RB, Fry JE, Hoffmann NL, Eichholtz D, Rogers SG, Fraley RT. 1985. A simple and general method for transferring genes into plants. *Science* 227:1229-31.
- Huang J, Sutliff T, Wu L, Nandi S, Bengé K, Terashima M, Ralston A, Drohan W, Huang N, Rodriguez R. 2001. Expression and purification of functional human alpha-1-antitrypsin from cultured plant cells. *Biotechnology Progress* 17:126-33.
- Huang X, von Rad U, Durner J. 2002. Nitric oxide induces transcriptional activation of the nitric oxide-tolerant alternative oxidase in *Arabidopsis* suspension cells. *Planta* 215:914-23.
- Huang Z, Chen Q, Hjelm B, Arntzen C, Mason H. 2009. A DNA replicon system for rapid high-level production of virus-like particles in plants. *Biotechnology and Bioengineering* 103:706-14.
- Huang Z, LePore K, Elkin G, Thanavala Y, Mason HS. 2008. High-yield rapid production of Hepatitis B surface antigen in plant leaf by a viral expression system. *Plant Biotechnology Journal* 6:202-09.
- Huang Z, Phoolcharoen W, Lai H, Piensook K, Cardineau G, Zeitlin L, Whaley KJ, Arntzen CJ, Mason HS, Chen Q. 2010. High level rapid production of full size monoclonal antibodies in plants by a single vector DNA replicon system. *Biotechnology and Bioengineering* 106:9-17.
- Hwang WYK, Almagro JC, Buss TN, Tan P, Foote J. 2005. Use of human germline genes in a CDR homology-based approach to antibody humanization. *Methods* 36:35-42.
- Ikemura T. 1985. Codon usage and tRNA content in unicellular and multicellular organisms. *Molecular Biology and Evolution* 2:13-34.
- Ismaili A, Jalali-Javaran M, Rasaee MJ, Rahbarizadeh F, Forouzandeh-Moghadam M, Memari HR. 2007. Production and characterization of anti-(mucin MUC1) single-domain antibody in tobacco (*Nicotiana tabacum* cultivar Xanthi). *Biotechnology and Applied Biochemistry* 47:11-9.
- Janeway CA, Travers P, Walport M, Shlomchik M. 2001. Immunobiology. New York: Garland Science.
- Janssen BJ, Gardner RC. 1990. Localized transient expression of GUS in leaf discs following cocultivation with *Agrobacterium*. *Plant Molecular Biology* 14:61-72.
- Jobling SA, Jarman C, Teh M-M, Holmberg N, Blake C, Verhoeyen ME. 2003. Immunomodulation of enzyme function in plants by single-domain antibody fragments. *Nature Biotechnology* 21:77-80.

- Johansson DX, Drakenberg K, Hopmann KH, Schmidt A, Yari F, Hinkula J, Persson MAA. 2007. Efficient expression of recombinant human monoclonal antibodies in *Drosophila* S2 cells. *Journal of Immunological Methods* 318:37-46.
- Joosten V, Gouka RJ, van den Hondel CA, Verrips CT, Lokman BC. 2005. Expression and production of llama variable heavy-chain antibody fragments (V(HH)s) by *Aspergillus awamori*. *Applied Microbiology and Biotechnology* 66:384-92.
- Kabat E, Wu T, Perry H, Gottesman K, Foeller C. 1991. Sequences of proteins of immunological interest. In: NIH Publication no. 91-3242 (5th edition). Bethesda, MD National Institutes of Health.
- Kalthoff D, Giritch A, Geisler K, Bettmann U, Klimyuk V, Hehnen HR, Gleba Y, Beer M. 2010. Immunization with plant-expressed hemagglutinin protects chickens from lethal HPAIV H5N1 challenge infection. *Journal of Virology* ePub (ahead of print).
- Kamarck ME. 2006. Building biomanufacturing capacity—the chapter and verse. *Nature Biotechnology* 24:503-05.
- Kapila J, De Rycke R, Van Montagu M, Angenon G. 1997. An *Agrobacterium*-mediated transient gene expression system for intact leaves. *Plant Science* 122:101-08.
- Kirby J, Kavanagh TA. 2002. NAN fusions: a synthetic sialidase reporter gene as a sensitive and versatile partner for GUS. *The Plant Journal* 32:391-400.
- Ko K, Koprowski H. 2005. Plant biopharming of monoclonal antibodies. *Virus Research* 111:93-100.
- Koehler U, Wagner T. 2007. New software from Eurofins MWG Operon optimises sequences for gene synthesis. In: <http://www.eurofinsdna.com/products-services/gene-synthesis/geneius.html>.
- Kooter JM, Matzke MA, Meyer P. 1999. Listening to the silent genes: transgene silencing, gene regulation and pathogen control. *Trends in Plant Science* 4:340-47.
- Kretzschmar T, Aoustin L, Zingel O, Marangi M, Vonach B, Towbin H, Geiser M. 1996. High-level expression in insect cells and purification of secreted monomeric single-chain Fv antibodies. *Journal of Immunological Methods* 195:93-101.
- Krüger C, Hultberg A, Marcotte H, Hermans P, Bezemer S, Frenken L, Hammarström L. 2006. Therapeutic effect of llama derived VHH fragments against *Streptococcus mutans* on the development of dental caries. *Applied Microbiology and Biotechnology* 72:732-37.
- Kurata H, Takemura T, Furusaki S, Kado CI. 1998. Light-controlled expression of a foreign gene using the chalcone synthase promoter in tobacco BY-2 cells. *Journal of Fermentation and Bioengineering* 86:317-23.

- Kwon T-H, Kim Y-S, Lee J-H, Yang M-S. 2003. Production and secretion of biologically active human granulocyte-macrophage colony stimulating factor in transgenic tomato suspension cultures. *Biotechnology Letters* 25:1571-74.
- Kwon T, Seo J, Kim J, Lee J, Jang Y, Yang M. 2003. Expression and secretion of the heterodimeric protein interleukin-12 in plant cell suspension culture. *Biotechnology and Bioengineering* 81:870-75.
- Ladenson RC, Crimmins DL, Landt Y, Ladenson JH. 2006. Isolation and characterization of a thermally stable recombinant anti-caffeine heavy-chain antibody fragment. *Analytical Chemistry* 78:4501-8.
- Lai H, Engle M, Fuchs A, Keller T, Johnson S, Gorlatov S, Diamond MS, Chen Q. 2010. Monoclonal antibody produced in plants efficiently treats West Nile virus infection in mice. *Proceedings of the National Academy of Sciences USA* 107:2419-24.
- Lauwereys M, Arbabi-Ghahroudi M, Desmyter A, Kinne J, Hölzer W, De Genst E, Wyns L, Muyldermans S. 1998. Potent enzyme inhibitors derived from dromedary heavy-chain antibodies. *The EMBO Journal* 17:3512-20.
- Le Hir H, Nott A, Moore MJ. 2003. How introns influence and enhance eukaryotic gene expression. *Trends in Biochemical Sciences* 28:215-20.
- Lee S-Y, Kim Y-H, Roh Y-S, Myoung H-J, Lee K-Y, Kim D-I. 2004. Bioreactor operation for transgenic *Nicotiana tabacum* cell cultures and continuous production of recombinant human granulocyte-macrophage colony-stimulating factor by perfusion culture. *Enzyme and Microbial Technology* 35:663-71.
- Lei SP, Lin HC, Wang SS, Callaway J, Wilcox G. 1987. Characterization of the *Erwinia carotovora* pelB gene and its product pectate lyase. *Journal of Bacteriology* 169:4379.
- Li Q, Hunt AG. 1995. A near-upstream element in a plant polyadenylation signal consists of more than six nucleotides. *Plant Molecular Biology* 28:927-34.
- Li XQ, Wei JZ, Tan A, Aroian RV. 2007. Resistance to root-knot nematode in tomato roots expressing a nematicidal *Bacillus thuringiensis* crystal protein. *Plant Biotechnology Journal* 5:455-64.
- Lindbo JA. 2007. TRBO: A high-efficiency Tobacco Mosaic Virus RNA-based overexpression vector. *Plant Physiology* 145:1232-40.
- Liu L, Ca izaes MC, Monger W, Perrin Y, Tsakiris E, Porta C, Shariat N, Nicholson L, Lomonossoff GP. 2005. Cowpea mosaic virus-based systems for the production of antigens and antibodies in plants. *Vaccine* 23:1788-92.
- Lörz H, Baker B, Schell J. 1985. Gene transfer to cereal cells mediated by protoplast transformation. *Molecular and General Genetics* 199:178-82.
- Ma JKC, Drake PMW, Christou P. 2003. The production of recombinant pharmaceutical proteins in plants. *Nature Reviews Genetics* 4:794-805.

- Ma JKC, Hiatt A, Hein M, Vine ND, Wang F, Stabila P, van Dolleweerd C, Mostov K, Lehner T. 1995. Generation and assembly of secretory antibodies in plants. *Science* 268:716-19.
- Ma JKC, Lehner T, Stabila P, Fux CI, Hiatt A. 1994. Assembly of monoclonal antibodies with IgG1 and IgA heavy chain domains in transgenic tobacco plants. *European Journal of Immunology* 24:131-8.
- Magnuson NS, Linzmaier PM, Reeves R, An G, HayGlass K, Lee JM. 1998. Secretion of biologically active human Interleukin-2 and Interleukin-4 from genetically modified tobacco cells in suspension culture. *Protein Expression and Purification* 13:45-52.
- Makrides SC. 1996. Strategies for achieving high-level expression of genes in *Escherichia coli*. *Microbiological Reviews* 60:512-38.
- Marillonnet S, Giritch A, Gils M, Kandzia R, Klimyuk V, Gleba Y. 2004. In planta engineering of viral RNA replicons: Efficient assembly by recombination of DNA modules delivered by *Agrobacterium*. *Proceedings of the National Academy of Sciences USA* 101:6852-57.
- Martineau P, Jones P, Winter G. 1998. Expression of an antibody fragment at high levels in the bacterial cytoplasm. *Journal of Molecular Biology* 280:117-27.
- Marusic C, Nuttall J, Buriani G, Lico C, Lombardi R, Baschieri S, Benvenuto E, Frigerio L. 2007. Expression, intracellular targeting and purification of HIV Nef variants in tobacco cells. *BMC Biotechnology* 7:12-23.
- Mason HS, Warzecha H, Mor T, Arntzen CJ. 2002. Edible plant vaccines: applications for prophylactic and therapeutic molecular medicine. *Trends in Molecular Medicine* 8:324-29.
- Matzke AJM, Matzke MA. 1998. Position effects and epigenetic silencing of plant transgenes. *Current Opinion in Plant Biology* 1:142-48.
- Matzke MA, Matzke AJM. 1998. Epigenetic silencing of plant transgenes as a consequence of diverse cellular defence responses. *Cellular and Molecular Life Sciences* 54:94-103.
- McCabe M, Klaas M, Gonzalez-Rabade N, Poage M, Badillo-Corona J, Zhou F, Karcher D, Bock R, Gray J, Dix P. 2008. Plastid transformation of high-biomass tobacco variety Maryland Mammoth for production of human immunodeficiency virus type 1 (HIV-1) p24 antigen. *Plant Biotechnology Journal* 6:914-29.
- Menassaa R, Kennettea W, Nguyena V, Rymersona R, Jevnikarb A, Brandle J. 2003. Subcellular targeting of human interleukin-10 in plants. *Journal of Biotechnology* 108:179-83.
- Merle C, Perret S, Lacour T, Jonval V, Hudaverdian S, Garrone R, Ruggiero F, Theisen M. 2002. Hydroxylated human homotrimeric collagen I in *Agrobacterium tumefaciens*-mediated transient expression and in transgenic tobacco plant. *FEBS Letters* 515:114-18.

- Michalet X, Pinaud FF, Bentolia LA, Tsay JM, Doose S, Li JJ, Sundaresan G, Wu AM, Gambhir SS, Weiss S. 2005. Quantum dots for live cells, *in vivo* imaging, and diagnostics Science 307:538-44.
- Miller KD, Weaver-Feldhaus J, Gray SA, Siegel RW, Feldhaus MJ. 2005. Production, purification, and characterization of human scFv antibodies expressed in *Saccharomyces cerevisiae*, *Pichia pastoris*, and *Escherichia coli*. Protein Expression Purification 42:255-67.
- Misawa S, Kumagai I. 1999. Refolding of therapeutic proteins produced in *Escherichia coli* as inclusion bodies. Peptide Science 51:297-307.
- Mogen BD, MacDonald MH, Leggewie G, Hunt AG. 1992. Several distinct types of sequence elements are required for efficient mRNA 3' end formation in a pea *rbcS* gene. Molecular Cell Biology 12:5406-14.
- Müller-Schiffmann A, Petsch B, Leliveld SR, Muylers J, Salwierz A, Mangels C, Schwarzingen S, Riesner D, Stitz L, Korth C. 2009. Complementarity determining regions of an anti-prion protein scFv fragment orchestrate conformation specificity and antiprion activity. Molecular Immunology 46:532-40.
- Murashige T, Skoog F. 1962. A revised medium for rapid growth and bioassays with tobacco culture. Physiologia Plantarum 15:473-97
- Muruganandam A, Tanha J, Narang S, Stanimirovic D. 2002. Selection of phage-displayed llama single-domain antibodies that transmigrate across human blood-brain barrier endothelium. The FASEB Journal 16:240-42.
- Muyldermans S. 2001. Single domain camel antibodies: current status. Journal of Biotechnology 74:277-302.
- Muyldermans S, Atarhouch T, Saldanha J, Barbosa JA, Hamers R. 1994. Sequence and structure of VH domain from naturally occurring camel heavy chain immunoglobulins lacking light chains. Protein Engineering 7:1129-35.
- Nakamura Y, Gojobori T, Ikemura T. 2000. Codon usage tabulated from the international DNA sequence databases: status for the year 2000. Nucleic Acids Research 28:292.
- Newell CA, Birch-Machin I, Hibberd JM, Gray JC. 2003. Expression of green fluorescent protein from bacterial and plastid promoters in tobacco chloroplasts. Transgenic Research 12:631-34.
- Nguyen VK, Hamers R, Wyns L, Muyldermans S. 1999. Loss of splice consensus signal is responsible for the removal of the entire CH1 domain of the functional camel IGG2A heavy-chain antibodies. Molecular Immunology 36:515-24.
- Nguyen VK, Hamers R, Wyns L, Muyldermans S. 2000. Camel heavy-chain antibodies: diverse germline VHH and specific mechanisms enlarge the antigen-binding repertoire. The EMBO Journal 19:921-30.
- Nguyen VK, Muyldermans S, Hamers R. 1998. The specific variable domain of camel heavy-chain antibodies is encoded in the germline. Journal of Molecular Biology 275:413-18.

- Obregon P, Chargelegue D, Drake PMW, Prada A, Nuttall J, Frigerio L, Ma JKC. 2006. HIV-1 p24-immunoglobulin fusion molecule: a new strategy for plant-based protein production. *Plant Biotechnology Journal* 4:195-207.
- Omidfar K, Rasaee MJ, Modjtahedi H, Forouzandeh M, Taghikhani M, Golmakani N. 2004. Production of a novel camel single-domain antibody specific for the Type III mutant EGFR. *Tumor Biology* 25:296-305.
- Otsu M, Urade R, Kito M, Omura F, Kikuchi M. 1995. A possible role of ER-60 protease in the degradation of misfolded proteins in the endoplasmic reticulum. *Journal of Biological Chemistry* 270:14528-961.
- Outchkourov NS, Rogelj B, Strukelj B, Jongsma MA. 2003. Expression of Sea Anemone Equistatin in Potato. Effects of Plant Proteases on Heterologous Protein Production. *Plant Physiology* 133:379-90.
- Pant N, Hultberg A, Zhao Y, Svensson L, Pan-Hammarström Q, Johansen K, Pouwels Peter H, Ruggeri Franco M, Hermans P, Frenken L, Borén T, Marcotte H, Hammarström L. 2006. *Lactobacilli* expressing variable domain of llama heavy-chain antibody fragments (lactobodies) confer protection against Rotavirus-induced diarrhea. *The Journal of Infectious Diseases* 194:1580-88.
- Patel J, Zhu H, Menassa R, Gyenis L, Richman A, Brandle J. 2007. Elastin-like polypeptide fusions enhance the accumulation of recombinant proteins in tobacco leaves. *Transgenic Research* 16:239-49.
- Peeters K, De Wilde C, Depicker A. 2001. Highly efficient targeting and accumulation of a Fab fragment within the secretory pathway and apoplast of *Arabidopsis thaliana*. *European Journal of Biochemistry* 268:4251-60.
- Perez JM, Renisio JG, Prompers JJ, van Platerink CJ, Cambillau C, Darbon H, Frenken LG. 2001. Thermal unfolding of a llama antibody fragment: a two-state reversible process. *Biochemistry* 40:74-83.
- Peters J, Stoger E. 2011. Transgenic crops for the production of recombinant vaccines and anti-microbial antibodies. *Human Vaccine* 7:ePub (ahead of print).
- Pleschberger M, Saerens D, Weigert S, Sleytr UB, Muyldermans S, Sara M, Egelseer EM. 2004. An S-layer heavy chain camel antibody fusion protein for generation of a nanopatterned sensing layer to detect the prostate-specific antigen by surface plasmon resonance technology. *Bioconjugate Chemistry* 15:664-71.
- Pogrebnyak N, Golovkin M, Andrianov V, Spitsin S, Smirnov Y, Egolf R, Koprowski H. 2005. Severe acute respiratory syndrome (SARS) S protein production in plants: development of recombinant vaccine. *Proceedings of the National Academy of Sciences USA* 102:9062-67.
- Popescu SC, Popescu GV, Bachan S, Zhang Z, Seay M, Gerstein M, Snyder M, Dinesh-Kumar SP. 2007. Differential binding of calmodulin-related proteins to their targets revealed through high-density *Arabidopsis* protein microarrays. *Proceedings of the National Academy of Sciences USA* 104:4730-35.

- Qin H, Guo H, Jia Y, Li L, Tian Y. 2000. Effect of UTRs from TMV-RNA on the expression of foreign gene in transgenic plants. *Chinese Science Bulletin* 45:1681-85.
- Rademacher T, Sack M, Arcalis E, Stadlmann J, Balzer S, Altmann F, Quendler H, Stiegler G, Kunert R, Fischer R, Stoger E. 2008. Recombinant antibody 2G12 produced in maize endosperm efficiently neutralizes HIV-1 and contains predominantly single-GlcNAc *N*-glycans. *Plant Biotechnology Journal* 6:189-201.
- Raharjo SB, Emoto N, Ikeda K, Sato R, Yokoyama M, Matsuo M. 2001. Alternative splicing regulates the Endoplasmic Reticulum localization or secretion of soluble secreted endopeptidase. *Journal of Biological Chemistry* 276:25612-20.
- Rahbarizadeh F, Rasaee MJ, Forouzandeh-Moghadam M, Allameh AA. 2005. High expression and purification of the recombinant camelid anti-MUC1 single domain antibodies in *Escherichia coli*. *Protein Expression and Purification* 44:32-8.
- Rahbarizadeh F, Rasaee MJ, Forouzandeh M, Allameh AA. 2006. Over expression of anti-MUC1 single-domain antibody fragments in the yeast *Pichia pastoris*. *Molecular Immunology* 43:426-35.
- Ram N, Ayala M, Lorenzo D, Palenzuela D, Herrera L, Doreste V, Pérez M, Gavilondo J, Oramas P. 2002. Expression of a single-chain Fv antibody fragment specific for the Hepatitis B surface antigen in transgenic tobacco plants. *Transgenic Research* 11:61-64.
- Ramírez N, Lorenzo D, Palenzuela D, Herrera L, Ayala M, Fuentes A, Pérez M, Gavilondo J, Oramas P. 2000. Single-chain antibody fragments specific to the hepatitis B surface antigen, produced in recombinant tobacco cell cultures. *Biotechnology Letters* 22:1233-36.
- Razdan M. 2003. *Introduction to Plant Tissue Culture*. Andover, Hampshire, UK: Intercept.
- Reichert JM, Valge-Archer VE. 2007. Development trends for monoclonal antibody cancer therapeutics. *Nature Reviews Drug Discovery* 6:349-56.
- Riechmann L. 1996. Rearrangement of the former VL interface in the solution structure of a camelised, single antibody VH domain. *Journal of Molecular Biology* 259:957-69.
- Roovers RC, Laeremans T, Huang L, De Taeye S, Verkleij AJ, Revets H, de Haard HJ, van Bergen en Henegouwen PM. 2007. Efficient inhibition of EGFR signaling and of tumour growth by antagonistic anti-EFGR Nanobodies. *Cancer Immunology and Immunotherapy* 56:303-17.
- Rothbauer U, Zolghadr K, Muyldermans S, Schepers A, Cardoso MC, Leonhardt H. 2008. A Versatile Nanotrap for Biochemical and Functional Studies with Fluorescent Fusion Proteins. *Molecular and Cell Proteomics* 7:282-89.
- Saerens D, Ghassabeh GH, Muyldermans S. 2008. Antibody technology in proteomics. *Briefings in Functional Genomics and Proteomics* 7:275-82.

- Saerens D, Ghassabeh GH, Muyldermans S. 2008. Single-domain antibodies as building blocks for novel therapeutics. *Current Opinion in Pharmacology* 8:600-08.
- Saerens D, Pellis M, Loris R, Pardon E, Dumoulin M, Matagne A, Wyns L, Muyldermans S, Conrath K. 2005. Identification of a universal VHH framework to graft non-canonical antigen-binding loops of camel single-domain antibodies. *Journal of Molecular Biology* 352:597-607.
- Sainsbury F, Cañizares MC, Lomonossoff GP. 2010. Cowpea mosaic virus: the plant virus-based biotechnology workhorse. *Annual Reviews of Phytopathology* 48:437-55.
- Sainsbury F, Lavoie P-O, D'Aoust M-A, Vézina L-P, Lomonossoff G. 2008. Expression of multiple proteins using full-length and deleted versions of cowpea mosaic virus RNA-2. *Plant Biotechnology Journal* 6:82-92.
- Sainsbury F, Lomonossoff G. 2008. Extremely high-level and rapid transient protein production in plants without the use of viral replication. *Plant Physiology* 148:1212-18.
- Sainsbury F, Sack M, Stadlmann J, Quendler H, Fischer R, Lomonossoff GP, Stoddart CA. 2010. Rapid transient production in plants by replicating and non-replicating vectors yields high quality functional anti-HIV antibody. *PLoS One* 5:263-70.
- Sainsbury F, Thuenemann EC, Lomonossoff GP. 2009. pEAQ: Versatile expression vectors for easy and quick transient expression of heterologous proteins in plants. *Plant Biotechnology Journal* 7:682-93.
- Sambrook J, Fritsch EF, Maniatis T. 1989. *Molecular cloning: a laboratory manual*. . Cold Spring Harbour: Cold Spring Harbour Laboratory Press.
- Sanford JC. 1988. The biolistic process. *Trends in Biotechnology* 6:299-302.
- Sanford JC, Smith FD, Russell JA. 1993. Optimizing the biolistic process for different biological applications. *Methods in Enzymology* 217:483-509.
- Santi L, Giritch A, Roy CJ, Marillonnet S, Klimyuk V, Gleba Y, Webb R, Arntzenc CJ, Mason HS. 2006. Protection conferred by recombinant *Yersinia pestis* antigens produced by a rapid and highly scalable plant expression system. *Proceedings of the National Academy of Sciences USA* 103 861-66.
- Sawant SV, Kiran K, Singh PK, Tuli R. 2001. Sequence architecture downstream of the initiator codon enhances gene expression and protein stability in plants. *Plant Physiology* 126:1630-36.
- Scheller J, Henggeler D, Viviani A, Conrad U. 2004. Purification of spider silk-elastin from transgenic plants and application for human chondrocyte proliferation. *Transgenic Research* 13:51-57.
- Scheller J, Leps M, Conrad U. 2006. Forcing single-chain variable fragment production in tobacco seeds by fusion to elastin-like polypeptides. *Plant Biotechnology Journal* 4:243-49.

- Schillberg S, Fischer R, Emans N. 2003. Molecular farming of recombinant antibodies in plants. *Cellular and Molecular Life Science* 60:433-45.
- Schillberg S, Zimmermann S, Findlay K, Fischer R. 2000. Plasma membrane display of anti-viral single chain Fv fragments confers resistance to tobacco mosaic virus. *Molecular Breeding* 6:317-26.
- Schillberg S, Zimmermann S, Voss A, Fischer R. 1999. Apoplastic and cytosolic expression of full-size antibodies and antibody fragments in *Nicotiana tabacum*. *Transgenic Research* 8:255-63.
- Schlereth A, Standhardt D, Mock H-P, Müntz K. 2001. Stored cysteine proteinases start globulin mobilization in protein bodies of embryonic axes and cotyledons during vetch (*Vicia sativa* L.) seed germination. *Planta* 212:718-27.
- Schlereth B, Fichtner I, Lorenczewski G, Kleindienst P, Brischwein K, da Silva A, Kufer P, Lutterbuese R, Junghahn I, Kasimir-Bauer S, Wimberger P, Kimmig R, Baeuerle PA. 2005. Eradication of tumors from a human colon cancer cell line and from ovarian cancer metastases in immunodeficient mice by a single-chain Ep-CAM-/CD3-bispecific antibody construct. *Cancer Research* 65:2882-9.
- Scholthof HB, Morris TJ, Jackson AO. 1993. The capsid protein gene of tomato bushy stunt virus is dispensable for systemic movement and can be replaced for localized expression of foreign genes. *Molecular plant-microbe interactions: MPMI (USA)* 6:309-32.
- Scholthof HB, Scholthof KBG, Jackson AO. 1996. Plant virus gene vectors for transient expression of foreign proteins in plants. *Annual Reviews of Phytopathology* 34:299-323.
- Scholthof KBG. 2004. Tobacco mosaic virus: a model system for plant biology. *Phytopathology* 42:13.
- Schornack S, Fuchs R, Huitema E, Rothbauer U, Lipka V, Kamoun S. 2009. Protein mislocalization in plant cells using a GFP-binding chromobody. *The Plant Journal* 60:744-54.
- Schouten A, Roosien J, Bakker J, Schots A. 2002. Formation of disulfide bridges by a single-chain Fv antibody in the reducing ectopic environment of the plant cytosol. *Journal of Biological Chemistry* 277:19339-45.
- Schouten A, Roosien J, de Boer JM, Wilmink A, Rosso MN, Bosch D, Stiekema WJ, Gommers FJ, Bakker J, Schots A. 1997. Improving scFv antibody expression levels in the plant cytosol. *FEBS Letters* 415:235-41.
- Semenyuk EG, Stremovskii OA, Orlova IV, Balandin TG, Nosov AM, Buryanov YI, Deyev SM, Petrov RV. 2003. Biosynthesis of the scFv Antibody to Human Ferritin in Plant and Bacterial Producers. *Molecular Biology* 37:780-86.
- Serruys B, Van Houtte F, Verbrugghe P, Leroux Roels G, Vanlandschoot P. 2009. Llama derived single domain intrabodies inhibit secretion of Hepatitis B virions in mice. *Hepatology* 49:39-49.

- Sharma SK, Bagshawe KD, Begent RH. 2005. Advances in antibody-directed enzyme prodrug therapy. *Current Opinion in Investigational Drugs* 6:611-5.
- Sharp JM, Doran PM. 2001. Strategies for enhancing monoclonal antibody accumulation in plant cell and organ cultures. *Biotechnology Progress* 17:979-92.
- Shi X, Karkut T, Chamankhah M, Alting-Mees M, Hemmingsen SM, Hegedus D. 2003. Optimal conditions for the expression of a single-chain antibody (scFv) gene in *Pichia pastoris*. *Protein Expression and Purification* 28:321-30.
- Shusta EV, Raines RT, Pluckthun A, Wittrup KD. 1998. Increasing the secretory capacity of *Saccharomyces cerevisiae* for production of single-chain antibody fragments. *Nature Biotechnology* 16:773-7.
- Sijmons PC, Dekker BMM, Schrammeijer B, Verwoerd TC, van den Elzen PJM, Hoekema A. 1990. Production of correctly processed human serum albumin in transgenic plants. *Nature Biotechnology* 8:217-21.
- Simpson CG, Jennings SN, Clark GP, Thow G, Brown JW. 2004. Dual functionality of a plant U-rich intronic sequence element. *Plant Journal* 37:82-91.
- Somerville CR, Dario Bonetta D. 2001. Plants as Factories for Technical Materials. *Plant Physiology* 125:168-71.
- Song G, Sink KC, Ma Y, Herlache T, Hancock JF, Loescher WH. 2010. A novel mannose-based selection system for plant transformation using celery mannose-6-phosphate reductase gene. *Plant Cell Reports* 29:163-72.
- Spinelli S, Frenken L, Bourgeois D, Ron Ld, Bos W, Verrips T, Anguille C, Cambillau C, Tegoni M. 1996. The crystal structure of a llama heavy chain variable domain. *Nature Structural and Molecular Biology* 3:752-57.
- Spinelli S, Frenken LGJ, Hermans P, Verrips T, Brown K, Tegoni M, Cambillau C. 2000. Camelid heavy-chain variable domains provide efficient combining sites to haptens. *Biochemistry* 39:1217-22.
- Staub JM, Garcia B, Graves J, Hajdukiewicz PTJ, Hunter P, Nehra N, Paradkar V, Schlittler M, Carroll JA, Spatola L, Ward D, Ye G, Russell DA. 2000. High-yield production of a human therapeutic protein in tobacco chloroplasts. *Nature Biotechnology* 18:333-38.
- Steinhauer DA, Domingo E, Holland JJ. 1992. Lack of evidence for proofreading mechanisms associated with an RNA virus polymerase. *Gene* 122:281-88.
- Stevens LH, Stoopen GM, Elbers IJW, Molthoff JW, Bakker HAC, Lommen A, Bosch D, Jordi W. 2000. Effect of Climate Conditions and Plant Developmental Stage on the Stability of Antibodies Expressed in Transgenic Tobacco. *Plant Physiology* 124:173-82.
- Stijlemans B, Conrath K, Cortez-Retamozo V, Van Xong H, Wyns L, Senter P, Revets H, De Baetselier P, Muyldermans S, Magez S. 2004. Efficient Targeting of Conserved Cryptic Epitopes of Infectious Agents by Single Domain Antibodies: African trypanosomes as paradigm *Journal of Biological Chemistry* 279:1256-61.