



THE UNIVERSITY OF QUEENSLAND
AUSTRALIA

**Engineering of Virus-Like Particles for Alternative Vaccine Candidate
Targeting a Hypervariable Peptide Antigen Element**

Melisa Rike Anggraeni
Master of Engineering Science

*A thesis submitted for the degree of Doctor of Philosophy at
The University of Queensland in 2014
Australian Institute for Bioengineering and Nanotechnology*

Abstract

A promising alternative to replace the current egg- or cell culture-based technology for vaccine production from live viruses is virus-like particle (VLP) technology based on a microbial platform. VLPs are macromolecular assemblies of viral capsid proteins, which have been shown to tolerate insertion of antigen modules via genetic recombinant technology, yielding modular VLPs. Many studies on modular VLPs presume that when a peptide antigen element is taken out from the intact proteins and then modularised on VLPs, it is unable to fold into its native structure. However, until now, presentation of a peptide antigen element on a VLP and the impact of the display strategy to present the antigen element on the quality of the resulting antibodies (i.e. the ability of the antibodies to recognise the intact protein) are not fully understood. This thesis aims to understand the underlying fundamentals regarding modularisation of peptide antigen elements on VLPs for induction of high-quality antibodies. A hypervariable receptor-binding domain, Helix 190 (H190), from the hemagglutinin protein of influenza A virus was used as a model for modularisation on VLPs from murine polyomavirus (MuPyV) VP1 protein. Four major findings are presented. Firstly, two display strategies, i.e. arraying of H190 in tandem repeats and the use of helix promoter elements, were shown to display H190 in its immunogenic form equally. However, modularisation using tandem repeat display induced antibodies of a higher quality than modularisation using helix promoter elements. Secondly, the quality of antibodies induced by the tandem repeat display bearing two copies of H190 was optimum, thus no significant improvement was observed following the use of adjuvant or increasing the copy number of H190. Additionally, the increase in the copy number of H190 was shown to reduce the assembly capability and solubility of modular VP1 in an environment that was optimised for wild-type VP1. Thirdly, this thesis shows the novel finding in the use of flanking ionic elements to stabilise VLP precursors, termed as capsomeres, bearing two copies of H190 containing a hydrophobic stretch, which caused aggregation. Fourthly, the first steps towards obtaining the atomic crystal structure of presented H190 on a modular protein were performed, i.e. a mild and satisfactory laboratory process was developed to achieve high-purity modular VP1 capsomeres, unattainable using previously established expression and purification process of wild-type MuPyV VP1. This thesis shows a step forward towards understanding the presentation of a peptide antigen element on a VLP that enables induction of high-quality antibodies, and towards VLP engineering to manipulate the aggregation and solubility of modular VP1. VLP technology based on a microbial platform presented here is

a potentially safe and effective alternative vaccine candidate that targets a hypervariable peptide antigen element. The speed of the microbial platform allows a rapid response to the hypervariability of the peptide antigen element, which otherwise may be unachievable using the egg- and cell culture-based technologies.

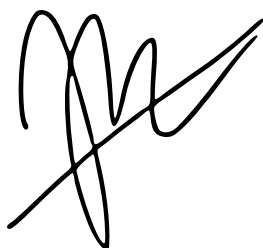
Declaration by author

This thesis is composed of my original work, and contains no material previously published or written by another person except where due reference has been made in the text. I have clearly stated the contribution by others to jointly-authored works that I have included in my thesis.

I have clearly stated the contribution of others to my thesis as a whole, including statistical assistance, survey design, data analysis, significant technical procedures, professional editorial advice, and any other original research work used or reported in my thesis. The content of my thesis is the result of work I have carried out since the commencement of my research higher degree candidature and does not include a substantial part of work that has been submitted to qualify for the award of any other degree or diploma in any university or other tertiary institution. I have clearly stated which parts of my thesis, if any, have been submitted to qualify for another award.

I acknowledge that an electronic copy of my thesis must be lodged with the University Library and, subject to the General Award Rules of The University of Queensland, immediately made available for research and study in accordance with the *Copyright Act 1968*.

I acknowledge that copyright of all material contained in my thesis resides with the copyright holder(s) of that material. Where appropriate I have obtained copyright permission from the copyright holder to reproduce material in this thesis.

A handwritten signature in black ink, consisting of several loops and a long horizontal stroke extending to the right.

Melisa Rike Anggraeni

July 2014

Publications during candidature

Anggraeni, MR, Connors, N, Wu, Y, Chuan, YP, Lua, LHL, Middelberg, APJ. 2013. Sensitivity of immune response quality to influenza helix 190 antigen structure displayed on a modular virus-like particle. Vaccine 31(40):4428-4435.

Anggraeni, MR was responsible for conception and design of experiment, cloning design, data acquisition, analysis, and interpretation of data, as well as drafting and writing. Connors, N was responsible for molecular dynamic simulation, analysis and interpretation of simulation data, drafting and writing. Wu, Y. prepared modular virus-like particles. Chuan, YP contributed in drafting and writing. Lua, LHL contributed to cloning design, as well as drafting and writing. Middelberg, APJ was responsible for conception and design of experiment, analysis and interpretation of data, conception of manuscript, as well as drafting and writing.

Publications included in this thesis

Anggraeni, MR, Connors, N, Wu, Y, Chuan, YP, Lua, LHL, Middelberg, APJ. 2013. Sensitivity of immune response quality to influenza helix 190 antigen structure displayed on a modular virus-like particle. Vaccine 31(40):4428-4435 – incorporated as Chapter 3.

Contributor	Statement of contribution
Anggraeni, MR	Designed experiments (60%), designed cloning (60%), data acquisition, data interpretation and analysis, wrote the paper (70%).
Connors, N	Performed molecular dynamic simulation, analysis and interpretation of simulation data; wrote material and methods for molecular dynamic simulation; provided advice and comments during drafting and writing.
Wu, Y	Prepared modular virus-like particles.
Chuan, YP	Provided advice and comments during drafting and writing.
Lua, LHL	Designed cloning (40%), provided advice and comments during drafting and writing.
Middelberg, APJ	Designed experiments (40%), provided advice and comments during data interpretation and analysis, as well as during drafting and writing.

Contributions by others to the thesis

This thesis was drafted and written by the candidate under the supervision of Professor Anton Middelberg, and Dr. Linda Lua. Dr. Nani Wibowo and Dr. Natalie Connors provided inputs during thesis writing.

Animal handlings during in vivo immunogenicity studies in Chapter 3 (AEC approval number: AIBN/273/10/CBME/AIBN (NF)) and Chapter 4 (AEC approval number: AIBN/058/13/NIRAP/SMART FUTURES) were performed by Dr. Nani Wibowo with assistance from the laboratory animal technicians from the University of Queensland Biological Resources (UQBR).

Animal handling during in vivo immunogenicity study in Chapter 4 (AEC approval number: AIBN/189/12/NIRAP/SMART FUTURE) was performed by the candidate, Dr. Natalie Connors, and Dr. Yap Pang Chuan, with assistance from the laboratory animal technicians from the University of Queensland Biological Resources (UQBR).

Modular VLPs in Chapter 3 was prepared by Yang Wu.

Peptide simulations in Chapter 3 were performed by Dr. Natalie Connors.

Screening of optimum conditions for crystallisation in Chapter 6 was performed in collaboration with Dr. Tom Caradoc-Davis and Dr. Santosh Panjekar from the Beamline team of Macromolecular Crystallography (Protein crystallography), the Australian Synchrotron (Melbourne, Australia).

All of the data and results presented in this thesis are solely the work of the candidate with the exception of the following figures, which were collected by the named collaborators.

Figure 1-3, protein structure pictures were generated by Dr. Natalie Connors.

Figure 3-1D, peptide structure pictures were generated by Dr. Natalie Connors.

Figure 4-5, Transmission electron microscopy (TEM) pictures were taken by Alice Yu.

Statement of parts of the thesis submitted to qualify for the award of another degree

None.

Acknowledgements

I would like to express my greatest appreciation to Professor Anton Middelberg and Dr. Linda Lua for valuable opportunities, guidance, supports, and inspirations over the years. It has been a privilege to be under their supervision.

Thank you Dr. Yap Pang Chuan, Dr. Natalie Connors, and Dr. Nani Wibowo for endless supports and guidance, both intellectually and emotionally.

Thank you for financial support from the University of Queensland in the form of The University of Queensland International Tuition Award (UQI Tuition), University of Queensland International Research Tuition Award, and University of Queensland Research Scholarship. Thank you for financial support from Australian Institute for Bioengineering and Nanotechnology in the form of a top up scholarship.

Thank you Dr. Tania Rivera-Hernandez, Yuan Yuan Fan, Bijun Zheng, Yang Wu, Fiona Hughes, Doan Thanh Tam, Dr. Frank Sainsbury, Lesley Forest, and each member of the Centre for Biomolecular Engineering for shoulders to lean on, ears to listen, hugs to comfort, not to forget for ups and downs, jokes and stories we shared.

Also, thank you Protein Expression Facility for technical supports. Thank you Emily Tan.

This thesis is dedicated to mom, dad, sister, little Jamie, and Luke. Thank you very much for everything.

Thank you God for the life you have given me a chance to live in. I am nothing without You.

“Ad maiorem Dei gloriam inque hominum salutem;

Any work that is not evil, even one that would normally be considered inconsequential to the spiritual life, can be spiritually meritorious if it is performed in order to give glory to God”

Keywords

virus-like particles, influenza, receptor binding, antigen presentation, biomolecular engineering

Australian and New Zealand Standard Research Classifications (ANZSRC)

ANZSRC code: 090403, Chemical Engineering Design, 50%

ANZSRC code: 100403, Medical Molecular Engineering of Nucleic Acids and Proteins, 40%

ANZSRC code: 860801, Human Biological Preventatives (e.g. Vaccines), 10%

Fields of Research (FoR) Classification

FoR code: 0904, Chemical Engineering, 60%

FoR code: 1004, Medical Biotechnology, 40%

Table of Contents

Abstract.....	ii
Declaration by author.....	iv
Publications during candidature.....	v
Publications included in this thesis	v
Contributions by others to the thesis	vi
Statement of parts of the thesis submitted to qualify for the award of another degree.....	vii
Acknowledgements.....	viii
Keywords	ix
Australian and New Zealand Standard Research Classifications (ANZSRC)	ix
Fields of Research (FoR) Classification.....	ix
Table of Contents	x
List of Figures	xv
List of Tables.....	xx
List of Abbreviations used in this thesis	xxi
Chapter 1. Project overview.....	1
1.1 Pandemic influenza A and current influenza vaccine manufacturing	1
1.2 Virus-like particles (VLPs)	2
1.3 Research aims and objectives.....	9
1.4 Thesis organisation	9
Chapter 2. Literature review	11
2.1 Influenza A viruses	11
2.1.1 <i>Haemagglutinin (HA)</i>	14
2.1.2 <i>Antibody-binding site</i>	16
2.1.3 <i>The receptor-binding site</i>	19
2.1.4 <i>Correlation of antibody-binding site and receptor-binding site</i>	21
2.1.5 <i>Glycosylation of HA</i>	23
2.2 Influenza vaccines	25
2.2.1 <i>History of vaccination</i>	25
2.2.2 <i>Structural vaccinology</i>	27
2.2.3 <i>Current technologies for influenza vaccine production</i>	30
2.2.3.1 <i>Egg-based influenza vaccine production</i>	30

2.2.3.2 Cell culture-based influenza vaccine production.....	32
2.3 Virus-like particles (VLPs)	34
2.3.1 VLP production.....	36
2.3.1.1 Baculovirus-insect cell expression vector system (BEVS).....	37
2.3.1.2 <i>Escherichia coli</i> expression system	38
2.3.1.3 Current studies on VLP-based influenza vaccines.....	38
2.4 The UQ Microbial Vaccine Platform (UQ-MVP).....	39
2.4.1 Murine polyomavirus (MuPyV) VP1 protein	39
2.4.2 Modular MuPyV VP1 as an antigen carrier	42
2.4.3 Production of MuPyV VP1 VLPs using <i>E. coli</i> expression system.....	46
2.4.4 Molecular chaperones as contaminants on MuPyV VP1 produced in <i>E. coli</i>	46
2.4.4.1 GroEL.....	47
2.4.4.2 DnaK	48
2.4.4.3 Existing methods for purification of DnaK and GroEL.....	49
2.5 Concluding statement.....	49

Chapter 3. Sensitivity of immune response quality to influenza helix 190 antigen structure displayed on a modular virus-like particle.....	51
Abstract	52
Introduction.....	52
Materials and methods	53
<i>Simulation of peptides GCN4-H190-GCN4 and H190-H190</i>	53
<i>Cloning, purification, and assembly of modular VLPs</i>	53
<i>Characterisation and visualisation of modular VLPs</i>	54
<i>Mice immunisation</i>	54
<i>Dot blot</i>	54
<i>ELISA</i>	54
<i>Statistical analysis</i>	54
Results.....	54
<i>Modularisation of VLP and simulation</i>	54
<i>Formation of modular VLPs</i>	54
<i>H190-specific immunogenicity</i>	54
<i>Specificity of modular VLP antisera to recombinant HA1 protein</i>	54
<i>Glycosylation effects of recombinant H1N1 HA1</i>	54
<i>Specificity of modular VLP antisera</i>	55

Discussion	55
Acknowledgements	58
References	58
Chapter 4. Display strategy improvements for induction of a higher quality of antibodies	60
4.1 Introduction	60
4.2 Materials and methods	62
4.2.1 <i>Generation of modular constructs</i>	62
4.2.2 <i>Expression and purification of modular VP1 proteins</i>	64
4.2.3 <i>Assembly of modular VP1 proteins</i>	65
4.2.4 <i>SDS-PAGE</i>	66
4.2.5 <i>Protein concentration determination</i>	66
4.2.6 <i>Mice immunisation</i>	67
4.2.7 <i>ELISA</i>	67
4.2.8 <i>Statistical analysis</i>	68
4.3 Results and discussion.....	68
4.3.1 <i>Solubility and assembly studies</i>	68
4.3.2 <i>H190-specific immunogenicity</i>	86
4.3.2.1 <i>Effect of AdvaxTM-1 on H190 immunogenicity</i>	87
4.3.2.2 <i>Effect of increasing number of H190 tandem repeats on H190 immunogenicity</i>	89
4.3.3 <i>Specificity of modular VLP antisera to recombinant HA1 protein</i>	92
4.3.3.1 <i>Effect of AdvaxTM-1 on specificity of modular VLP antisera to recombinant HA1 protein</i>	92
4.3.3.2 <i>Effect of increasing number of H190 tandem repeats on specificity of modular VLP antisera to recombinant HA1 protein</i>	94
4.4 Conclusions	97
Chapter 5. Challenging the tandem repeat display strategy using a hydrophobic helix 190 variant	100
5.1 Introduction.....	100
5.2 Materials and methods	108
5.2.1 <i>Generation of modular constructs</i>	108
5.2.2 <i>Protein concentration measurements</i>	110
5.2.3 <i>Expression and purification of modular VP1 proteins</i>	111

5.2.4 SDS-PAGE.....	111
5.3 Results and discussion.....	112
5.3.1 Increased aggregation and heterogeneity of modular VP1 bearing a hydrophobic stretch after GST removal.....	112
5.3.2 The use of charged residues.....	118
5.3.2.1 Effect of flanking ionic elements on the solubility of modular VP1.....	118
5.3.2.2 Effect of flanking ionic elements on the ratio of VP1* to VP1 after GST removal.....	125
5.3.2.3 Effect of flanking ionic elements on the aggregation of modular VP1 and isolation of modular VP1 capsomeres.....	129
5.3.3 Mutation from a thrombin to TEVp cleavage site.....	134
5.3.3.1 Effect of mutation from a thrombin to TEVp cleavage site on the expression and solubility of GST-tagged modular VP1.....	134
5.3.3.2 Effect of mutation from a thrombin to TEVp cleavage site on the ratio of VP1* to VP1 after GST removal.....	137
5.3.3.3 Effect of mutation from a thrombin to TEVp cleavage site on the aggregation of modular VP1 after GST removal.....	143
5.4 Conclusion.....	146

Chapter 6. Production of modular VP1 capsomeres with a high purity to screen the optimum condition for crystallisation.....	148
6.1 Introduction.....	148
6.2 Materials and methods.....	156
6.2.1 Generation of modular constructs.....	156
6.2.2 Expression and purification of GST-tagged modular VP1.....	157
6.2.3 Enzymatic GST tag removal from modular VP1.....	158
6.2.4 Modular VP1 capsomere purification by size-exclusion chromatography (SEC).....	158
6.2.5 Purification using GST-affinity chromatography for SEC-purified modular VP1 capsomeres.....	158
6.2.6 Purification using hydrophobic interaction chromatography (HIC) for GST-purified modular VP1 capsomeres.....	158
6.2.6.1 HIC in batch mode.....	158
6.2.6.2 HIC in bed mode.....	159
6.2.7 Concentrating HIC-purified modular VP1 capsomeres by precipitation.....	159

6.2.8	<i>Protein concentration measurement</i>	160
6.2.9	<i>SDS-PAGE</i>	160
6.2.10	<i>SEC-MALS</i>	160
6.2.11	<i>Western blot</i>	161
6.2.12	<i>Expression and purification of TEVp</i>	161
6.3	Results and Discussion	162
6.3.1	<i>Solubility of GST-tagged modular VP1 bearing thrombin and TEVp cleavage sites</i>	162
6.3.2	<i>GST tag removal efficacy of TEVp and thrombin</i>	166
6.3.3	<i>Contamination of SEC-purified modular VP1 capsomeres with undigested GST-tagged modular VP1, GST tag, and 70-kDa contaminant</i>	171
6.3.4	<i>Removal of GST tag and undigested GST-tagged modular VP1 using tandem GST columns</i>	176
6.3.5	<i>Removal of the 70-kDa contaminant using HIC</i>	177
6.3.5.1	<i>In a batch mode</i>	177
6.3.5.2	<i>In bed mode</i>	182
6.4	Conclusions	192
Chapter 7. Conclusions and future work		194
7.1	Summary of research findings	194
7.1.1	<i>Effects of the display strategy on the quality of the antibodies</i>	196
7.1.2	<i>Improvements in the display strategy to induce higher quality antibodies</i>	198
7.1.3	<i>Challenging the tandem repeat display strategy with a hydrophobic H190 variant</i>	201
7.1.4	<i>Structural determination of the H190 element on modular VLPs</i>	203
7.2	Future work	205
7.3	Concluding thought	208
References		210
Appendix A		244

List of Figures

- Figure 1-1. An illustration of a genetically inserted antigen module in a modular viral capsid protein.
- Figure 1-2. A murine polyomavirus VP1 VLP.
- Figure 1-3. Illustration of generation of a modular MuPyV VP1 VLP presenting a peptide antigen element (helix 190 from HA protein).
- Figure 2-1. A schematic diagram of Influenza A virions.
- Figure 2-2. Schematic diagram of the life cycles of influenza A viruses.
- Figure 2-3. An illustration of HA molecules.
- Figure 2-4. Antibody-binding sites in A/Hong Kong/1/1968 (H3N2) HA1.
- Figure 2-5. The position of the antibody-binding sites on HA from A/PR/8/34 (H1N1) relative to the arrangements on HA from A/Hong Kong/1968 (H3N2).
- Figure 2-6. Comparison of amino acid sequences of antibody-binding sites in various H1 subtypes.
- Figure 2-7. The receptor-binding site within HA molecule.
- Figure 2-8. Comparison of antibody-binding sites and receptor-binding domains in the globular head of HA molecule from A/South Carolina/1/1918 (H1N1) influenza.
- Figure 2-9. Comparison of antibody-binding sites and receptor-binding domains in the globular head of HA molecule from A/Hong Kong/1/1968 (H3N2) influenza.
- Figure 2-10. Manufacturing of seasonal trivalent influenza vaccines using egg-based technology (Treanor 2004).
- Figure 2-11. A schematic flow-chart of vaccine production using cell culture-based technology.
- Figure 2-12. Crystal structures of murine polyomavirus VP1 capsomeres.
- Figure 2-13. Assembly process of murine polyomavirus capsomeres to form VLPs.
- Figure 2-14. Surface-exposed loops in monomeric murine polyomavirus VP1 (1SID.pdb) (Stehle and Harrison 1996).
- Figure 2-15. Amino acid sequence alignment of VP1 from hamster polyomavirus (UniProtKB identifier O73427) and VP1 from murine polyomavirus (UniProtKB identifier P49302, 1SID.pdb).
- Figure 2-16. Insertion sites (spheres) in monomeric murine polyomavirus VP1 protein (1SID.pdb).

- Figure 3-1. (A) Crystal structure of H1N1 HA with H190 highlighted in red. Data for A/California/04/2009 HA (3AL4.pdb), bearing one amino acid difference from A/California/07/2009 HA, is shown here. The picture was created using UCSF Chimera 1.5.2, (B) modular MuPyV VP1 designs showing the heterologous elements in (i) VP1-GCN4-H190-GCN4; and (ii) VP1-H190-H190. Numbers refer to residues of VP1, (C) root-mean-square-deviation (RMSD) of C^α, backbone, side chain, and all heavy atoms of simulated GCN4-H190-GCN4 and H190-H190 peptides with reference to H190 on HA (3MLH.pdb), and (D) final conformation of H190 element in simulated peptides (i) GCN4-H190-GCN4; and (ii) H190-H190, in comparison to (iii) native H190 structure in HA. α -Helical structures are highlighted in blue. The image was created using Accelrys Discovery Studio[®] 3.0. Positions of Ser1 and Tyr16 on the peptides are as indicated.
- Figure 3-2. Characterisation of modular VLPs using AF4 (left panels) and TEM (right panels).
- Figure 3-3. Dot blot, indirect ELISA and competitive ELISA results showing immunogenicity of H190 in modular VLPs.
- Figure 3-4. Dot blot, indirect ELISA and competitive ELISA results showing reactivity of modular VLP antisera against recombinant HA1 protein.
- Figure 3-5. Effects of N-linked glycans on reactivity of modular VLP antisera to HA1.
- Figure 3-6. Dot blot and indirect ELISA results showing reactivity of modular VLP antisera against Fluvax[®].
- Figure 4-1. Schematic diagram of antigen modules for modular constructs reported in Chapter 4.
- Figure 4-2. Schematic diagram of design of DNA fragments containing an antigen module and homologous regions.
- Figure 4-3. SDS-PAGE analysis of modular VP1 bearing three, four, and five copies of H190 after dialysis against PBS.
- Figure 4-4. Principle of separation using asymmetric flow-field flow fractionation (AF4).
- Figure 4-5. Characterisation of modular VP1 bearing three, four, and five copies of H190 after assembly using AF4-MALS and TEM.
- Figure 4-6. The Hofmeister series of anions and cations.
- Figure 4-7. Analysis of modular VP1 bearing four copies of H190 after assembly at various pH.

- Figure 4-8. Analysis of modular VP1 bearing five copies of H190 after assembly at various pH.
- Figure 4-9. Analysis of the effects of $(\text{NH}_4)_2\text{SO}_4$ concentration in Assembly Buffer 1 on the solubility and assembly of modular VP1 bearing four copies of H190.
- Figure 4-10. Analysis of the effects of salt type in Assembly Buffer 1 on the solubility and assembly of modular VP1 bearing four copies of H190.
- Figure 4-11. Schematic diagram of immunisation and tail bleeding for *in vivo* studies conducted in Chapter 4.
- Figure 4-12. Indirect and competitive ELISA results showing the effect of AdvaxTM-1 on the immunogenicity of H190 element.
- Figure 4-13. Indirect and competitive ELISA results showing the immunogenicity of H190 element in modular VLPs bearing one, two, three, and four copies of H190.
- Figure 4-14. Indirect and competitive ELISA results showing the effect of AdvaxTM-1 on reactivity of modular VLP antisera against recombinant HA1 protein.
- Figure 4-15. Indirect and competitive ELISA results showing the effect of increasing number of tandem repeat H190 element on the reactivity of modular VLP antisera against recombinant HA1 protein.
- Figure 5-1. Comparison of crystal structures of A/Human/South Caroline/1/1918 (H1N1) HA1 and A/Duck/Singapore/3/97 (H5N1) HA1.
- Figure 5-2. Amino acid sequence alignment of GST-tagged wt-VP1 bearing thrombin or the TEVp cleavage sites.
- Figure 5-3. Properties of modular VP1 bearing two copies of H190 from A/Victoria/210/2009 (construct VP1-H3-H190-H190).
- Figure 5-4. SDS-PAGE analysis showing the effect of overnight expression temperature on the solubility of GST-tagged wt-VP1 and modular VP1 bearing two copies of H190 from A/Victoria/210/2009 (construct VP1-H3-H190-H190).
- Figure 5-5. SDS-PAGE analysis showing the effects of the addition of glutamic acids on the expression and solubility of GST-tagged VP1.
- Figure 5-6. SDS-PAGE analysis showing the effect of the addition of glutamic acids on the ratio of VP1* to VP1 following GST removal using thrombin for GST-tagged wt-VP1, modular VP1 bearing two copies of H190 from A/Victoria/210/2009 without flanking ionic elements (construct VP1-H3-H190-H190), and modular VP1 bearing two copies of H190 from A/Victoria/210/2009 with flanking ionic elements (construct VP1-H3-H190-H190-4E) at different salt concentrations.

- Figure 5-7. Analysis using size-exclusion chromatography and SDS-PAGE showing the effects of the addition of glutamic acids on the aggregation of modular VP1 bearing H190 from A/Victoria/2010/2009 and feasibility to isolate the modular VP1 as capsomeres at 200 mM and 50 mM NaCl.
- Figure 5-8. SDS-PAGE analysis showing the effect of mutation from a thrombin to TEVp cleavage site on the solubility of GST-tagged VP1 proteins.
- Figure 5-9. SDS-PAGE analysis showing the effect of mutation from a thrombin to TEVp cleavage site on the ratio of VP1* to VP1 following GST removal using thrombin or TEVp.
- Figure 5-10. SDS-PAGE analysis of GST removal for modular VP1 bearing two copies of H190 from A/Victoria/210/2009 without flanking glutamic acids (construct TEVP-VP1-H3-H190-H190), modular VP1 bearing two copies of H190 from A/Victoria/210/2009 with flanking glutamic acids (construct TEVP-VP1-H3-H190-H190-4E), and modular VP1 bearing two copies of H190 from A/California/07/2009 (construct TEVP-VP1-H190-H190) using TEVp at various ratio of TEVp to VP1.
- Figure 5-11. SDS-PAGE analysis of GST removal for wt-VP1 (construct TEVP-wt-VP1), modular VP1 bearing two copies of H190 from A/Victoria/210/2009 without flanking glutamic acids (construct TEVP-VP1-H3-H190-H190), modular VP1 bearing two copies of H190 from A/Victoria/210/2009 with flanking glutamic acids (construct TEVP-VP1-H3-H190-H190-4E), and modular VP1 bearing two copies of H190 from A/California/07/2009 (construct TEVP-VP1-H190-H190) using TEVp at various salt concentrations.
- Figure 5-12. Analysis using size-exclusion chromatography and SDS-PAGE showing the effects of mutation from a thrombin to TEVp cleavage site on the aggregation of modular VP1 bearing H190 from A/Victoria/2010/2009 and feasibility to isolate the modular VP1 as capsomeres.
- Figure 6-1. Crystal structures of wild-type murine polyomavirus VP1 capsomeres.
- Figure 6-2. A schematic diagram of antigen modules in the three modular constructs selected for crystallisation.
- Figure 6-3. A schematic diagram of the engineered cleavage sites in modular VP1 lacking 31 N-terminal residues and 63 C-terminal residues.
- Figure 6-4. SDS-PAGE analysis of the solubility of GST-tagged modular VP1 and wt-VP1 each bearing a thrombin cleavage site.

- Figure 6-5. SDS-PAGE analysis of the solubility of GST-tagged modular VP1 and wt-VP1 each bearing a TEVp cleavage site.
- Figure 6-6. SDS-PAGE analysis showing the impact of OD_{600 nm} values during IPTG induction on the solubility of GST-tagged modular VP1 bearing one copy of H190 and a TEVp cleavage site (construct TEVP-H190).
- Figure 6-7. SDS-PAGE analysis of GST-tag removal from modular VP1 and wt-VP1 (wt-VP1 residues 1-384) using thrombin.
- Figure 6-8. SDS-PAGE analysis showing the impacts of incubation time and TEVp to VP1 mass ratio on GST removal at 30°C.
- Figure 6-9. Size-exclusion chromatograms of modular VP1 following treatment with TEVp.
- Figure 6-10. SDS-PAGE analysis of eluted peaks on SEC for modular VP1.
- Figure 6-11. Identification of the 60-kDa contaminant in the SEC-purified modular VP1 capsomeres using western blot analysis.
- Figure 6-12. SDS-PAGE analysis of flow-through fractions from tandem GST-affinity chromatography columns for modular VP1 bearing two copies of H190 (construct TEVP-H190-H190).
- Figure 6-13. The Hofmeister series of anions and cations.
- Figure 6-14. SDS-PAGE analysis showing the impact of (NH₄)₂SO₄ concentrations on the adsorption of modular VP1 bearing two copies of H190 (modular construct TEVP-H190-H190) to Phenyl Sepharose HP resin.
- Figure 6-15. Removal of 70-kDa contaminant using HiTrap Phenyl HP for modular VP1 bearing two copies of H190 (construct TEVP-H190-H190).
- Figure 6-16. Removal of 70-kDa contaminant using HiTrap Phenyl HP for modular VP1 (VP1 residue 32-320) bearing one copy of H190, flanked by GCN4 helix promoter elements (construct TEVP-GCN4-H190-GCN4).
- Figure 6-17. Removal of 70-kDa contaminant using HiTrap Phenyl HP for modular VP1 bearing one copy of H190 (construct TEVP-H190).
- Figure 6-18. SDS-PAGE analysis of fractions from peaks eluted at 0% (v/v) (NH₄)₂SO₄ at a large scale HIC purification of modular VP1 capsomeres.
- Figure 6-19. Analysis of quaternary structure of HIC-purified modular VP1 capsomeres using SEC-MALS.

List of Tables

- Table 1-1. Influenza A pandemic cases with worldwide impacts.
- Table 2-1. Analysis tools in structural vaccinology, taken from (Dormitzer et al. 2008).
- Table 4-1. List of modular constructs reported in Chapter 4.
- Table 4-2. List of theoretical molecular weight and extinction coefficient of VP1 proteins reported in Chapter 4.
- Table 4-3. Observation of protein solution turbidity after dialysis against PBS.
- Table 5-1. List of predicted secondary thrombin cleavage sites on wt-VP1 and their cleavage products.
- Table 5-2. List of constructs used in Chapter 5.
- Table 5-3. List of theoretical molecular weights and extinction coefficients for proteins used in Chapter 5.
- Table 5-4. Cell density of harvested cells following expression at 12°C and 26°C.
- Table 6-1. List of modular constructs bearing thrombin or TEVp cleavage sites generated in Chapter 6.
- Table 6-2. Theoretical molecular weights and extinction coefficients of proteins used in Chapter 6.
- Table 6-3. Observation of modular VP1 precipitation during GST removal using TEVp at 30°C.

List of Abbreviations used in this thesis

2-ME	2-Mercaptoethanol
AF4-MALS	Asymmetric flow field-flow fractionation – multi-angle light scattering
AGRF	Australian Genome Research Facility
ANOVA	Analysis of variance
DLVO	Derjaguin, Landau, Verwey and Overbeek theory
DNA	Deoxyribonucleic acid
dNTPs	Deoxynucleotide triphosphates
DTT	Dithiothreitol
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	Ethylenediaminetetraacetic acid
ELISA	Enzyme-linked immunosorbent assay
GRAVY	Grand average of hydropathy
GSH	Glutathione
GST	Glutathione-S-transferase
H190	Helix 190
HA	Haemagglutinin
HaPyV	Hamster polyomavirus
HBcAg	Hepatitis B core antigen
HIC	Hydrophobic interaction chromatography
IPTG	Isopropyl β -D-1-thiogalactopyranoside
kDa	Kilo Dalton
LB	Luria-Bertani
LS	Light scattering
M1	Matrix protein 1
M2	Matrix protein 2
MuPyV	Murine polyomavirus
NA	Neuraminidase
NP	Nucleoprotein
NS1	Non-structural protein 1
NS2	Non-structural protein 2
OD _{600 nm}	Optical density at 600 nm
PA	Polymerase acidic protein
PB1	Polymerase basic protein 1

PB2	Polymerase basic protein 2
PBS	Phosphate buffer saline
PCR	Polymerase chain reaction
pI	Isoelectric point
RNA	Ribonucleic acid
rpm	rotation per minute
SD	Standard deviation
SDM	Site directed mutagenesis
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SEC	Size-exclusion chromatography
TB	Terrific Broth
TEM	Transmission electron microscopy
TEMED	Tetramethylethylenediamine
TEVp	Tobacco etch virus protease
UQBR	University of Queensland Biological Resources
UV	Ultraviolet
VLP	Virus-like particles
VP1	Viral protein 1
WHO	World Health Organisation
Wt	Wild-type

Chapter 1. Project overview

1.1 Pandemic influenza A and current influenza vaccine manufacturing

Influenza A viruses belong to the Orthomyxoviridae family (Webster et al. 1992) and have been isolated from various species, including human, birds, pigs, and horses (Horimoto and Kawaoka 2005, Medina and García-Sastre 2011, Olsen et al. 2006). The viruses contain eight single negative-sense RNA strands, which encode eight structural viral proteins and two major internal proteins (Webster et al. 1992). Haemagglutinin (HA) and neuraminidase (NA) are the principal structural proteins of influenza A viruses that determine the antigenic characteristics and subsequent classification of the viruses (Medina and García-Sastre 2011). Antigenic variation in influenza A viruses results from accumulation of mutations within the antibody binding sites in HA and/or NA (i.e. antigenic drift), or from replacement of HA and NA with novel subtypes, which have not circulated in humans for a long period of time (i.e. antigenic shift). Antigenic shift can cause pandemics when a newly emerging influenza A virus can transmit efficiently and sustainably from human to human with little or no existing immune responses against the HA of the virus. There have been four pandemic events with worldwide impacts since 1918, each with antigenically different viruses (Table 1-1).

Table 1-1. Influenza A pandemic cases with worldwide impacts.
(Data were taken from <http://www.flupandemic.gov.au/>)

Year	Influenza virus strain	Mortality rate
1918	A/South Carolina/1/1918 (H1N1)	~ 50 million
1957	H2N2 (A/Singapore/1/1957, A/Japan/305/1957, A/Guiyang/1/57)	~ 2 million
1968	A/Aichi/2/1968 (H3N2)	~ 1 million
2009	A/California/07/2009 (H1N1)	~37,000

One of the most effective and sustainable ways to prevent the spread of influenza viruses is through vaccination (Bagnoli et al. 2011). Vaccination has successfully reduced both occurrence and mortality rates of influenza (Ulmer et al. 2006). Most commercial vaccines in markets are produced using live virus-based technologies, which follow the Pasteur paradigm of “isolate, inactivate, and inject” (Rappuoli 2007). In these technologies, influenza vaccine viruses are passaged in the allantoic fluid of embryonated chicken eggs

(Gerdil 2003, Tree et al. 2001) or cell culture (Doroshenko and Halperin 2009, Tree et al. 2001). The purified viruses are then (i) inactivated (Maassab et al. 1990, Murphy and Coelingh 2002); (ii) treated with diethyl ether or Tween to produce split vaccines (Brandon et al. 1967, Hoyle et al. 1961, Neurath et al. 1971); or (iii) treated with detergent, such as Triton N 101 (Brady and Furminger 1976), ammonium deoxycholate (Laver and Webster 1976), or cetyltrimethylammonium bromide (Bachmayer et al. 1976) to recover the HA and/or NA from the virions.

However, despite the fact that it has played crucial roles in the prevention of influenza disease, live virus-based technologies are slow. Egg-based technology for manufacture of these vaccines requires 6-9 months to produce vaccines (Subbarao and Joseph 2007), while cell culture-based technology requires at least 2 months at scale (Cox and Hollister 2009). The speed limitation of live virus-based vaccine production is exacerbated in a case of an emerging pandemic when a new virus strain appears rather suddenly and unpredictably. Additionally, the process of virus passage results in mutations in the vaccine viruses, leading to a mismatch between the vaccine composition and the dangerous strains it is intended to protect against (Katz and Webster 1992). Due to the slow vaccine manufacturing and mismatches between the vaccine and circulating strains, widespread and timely prophylaxis through rapid manufacturing and deployment of effective vaccine remains a significant unsolved challenge for influenza.

1.2 Virus-like particles (VLPs)

An effective influenza pandemic intervention strategy mandates rapid vaccination at the population scale that is unachievable using live-virus based vaccine technology (Ferguson et al. 2006). One emerging technology that has a potential as the next-generation of influenza vaccines is the virus-like particle (VLP). VLP represents an advanced vaccine technology platform, which is able to combine most of the key features of viruses (1) highly repetitiveness, (2) particulate nature, and (3) effective induction of innate immune response leading to strong T-cell responses (Jennings and Bachmann 2007).

VLPs are defined as macromolecule assemblies, which are composed of self-assembling viral capsid proteins (i.e. capsomeres) (Pattenden et al. 2005). They have been used comprehensively and successfully as a safer alternative to the live virus-based vaccines because they are infection and replication incompetent (Ludwig and Wagner 2007). Approved VLP vaccines are used clinically to protect against the parental viruses from which the VLPs were derived. For example, Engerix-B or Recombivax-B for Hepatitis

B (Averhoff et al. 1998, Keating et al. 2003, Poland and Jacobson 2004), and Gardasil[®] and Cervarix[™] for human papillomavirus (Einstein et al. 2009, Kato et al. 2007, Schiller et al. 2008). Due to their stable structure and tolerance towards modification, VLPs have also been exploited to carry and display antigen modules, which are genetically inserted into the viral capsid protein (Figure 1-1). Antigen modules comprise peptide antigen elements from foreign pathogens and other necessary elements, such as flanking structure promoter or spacer elements. Combination of a VLP and an antigen module yields a modular VLP (Figure 1-1) (Anggraeni et al. 2013, Boisg rault et al. 2002, Garcea and Gissmann 2004, Pattenden et al. 2005, Rold o et al. 2010).

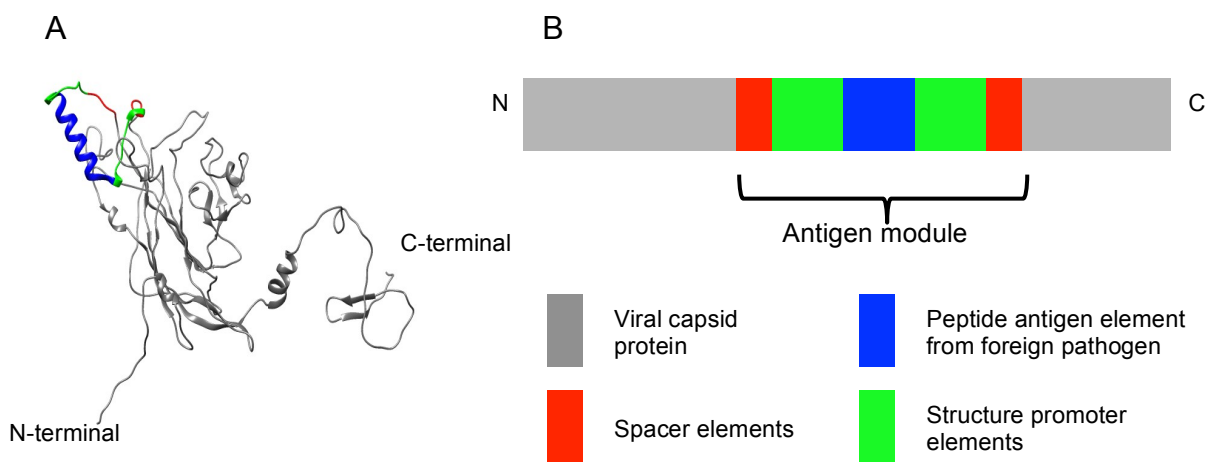


Figure 1-1. An illustration of a genetically inserted antigen module in a modular viral capsid protein. (A) A tertiary structure of a modular monomeric viral capsid protein (grey) displaying an antigen module (red, green, and blue). The figure was generated using UCSF Chimera (Pettersen et al. 2004). The PDB file of the protein was obtained using SWISS MODEL (<http://swissmodel.expasy.org/>) (Arnold et al. 2006, Biasini et al. 2014, Guex et al. 2009, Kiefer et al. 2009); and (B) A schematic diagram of the antigenic module in the modular protein in panel A.

A microbial vaccine platform based on VLPs from murine polyomavirus (MuPyV) VP1 protein has been developed. The use of bacteria allows faster VLP production than other hosts, such as mammalian cells, and thus provides a technology for rapid-response vaccine development (Middelberg et al. 2011). This platform is underpinned by bioprocessing investigations into scale-up techniques (Chuan et al. 2008, Liew et al. 2010, Lipin et al. 2008). The crystal structure of the VP1 protein is available at 1.9   resolution (Stehle and Harrison 1996, Stehle and Harrison 1997, Stehle et al. 1994). Several insertion sites in the VP1 protein have been patented (Lua and Middelberg 2007), allowing for recombinant insertion of antigen modules at a genetic level and projection of the inserted antigen modules to the surface of modular VLPs for immune system recognition

(Gedvilaite et al. 2000, Liddington et al. 1991). The MuPyV VP1 VLPs have been engineered to present antigen modules containing peptide antigen elements from various pathogens (Anggraeni et al. 2013, Chuan et al. 2013, Rivera-Hernandez et al. 2013), indicating the robustness of the VLPs.

The MuPyV VP1 capsid protein is isolated as a basic subunit viral capsid protein (i.e. capsomeres) following bacterial expression. In the presence of Ca^{2+} , purified capsomeres can then self-assemble *in vitro* to form VLPs (Figure 1-2). The assembly mechanism involves (i) invasion of C-terminal arms to the neighbouring capsomeres, which are secured by N-terminal clamps, and (ii) clamp-clamp interactions of N-terminals of neighbouring capsomeres. Compared to *in vivo* VLP assembly, *in vitro* VLP assembly is advantageous due to minimum encapsulated contaminants or cellular materials and can be more controllable yielding highly consistent VLPs (Pattenden et al. 2005).

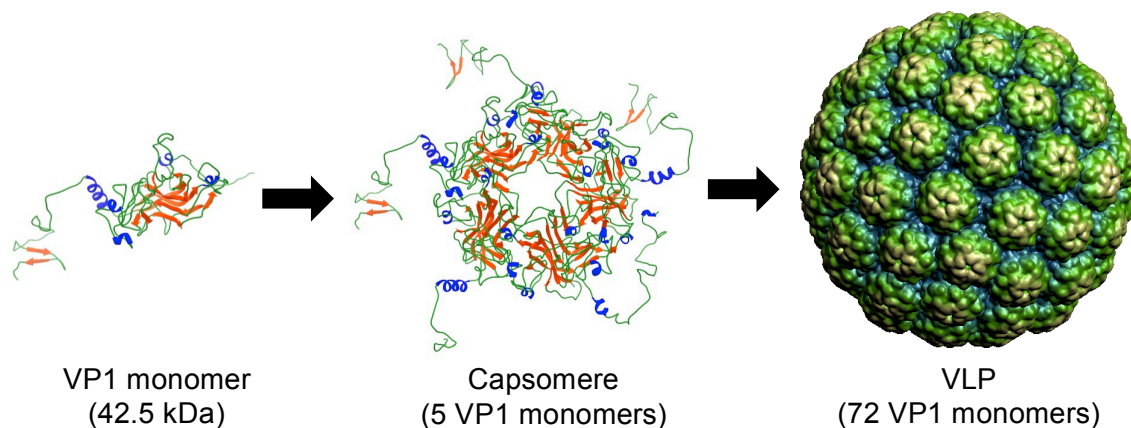


Figure 1-2. A murine polyomavirus VP1 VLP. VP1 capsid proteins are isolated as capsomeres from bacterial expression. Capsomeres self-assemble *in vitro* into VLPs. The VP1 monomer and capsomere structures were obtained from Protein Data Bank, PDB ID 1SID (Stehle and Harrison 1996). Molecular graphic images were created using UCSF Chimera (Pettersen et al. 2004). VLP picture was obtained from Virus Particle Explorer (VIPER) (http://viperd.b.scripps.edu/info_page.php?VDB=1sid) (Reddy et al. 2001).

The manufacturing speed and robustness of MuPyV VP1 VLPs suggest that the VLPs can be used as a potential vaccine platform for pandemic influenza A. An excellent model peptide antigen is from HA. HA is the most important components in influenza vaccines (McCaughey 2010). It is involved in the first step of influenza virus infection, that is binding of the virus onto the sialic acid receptor on host cells (Neumann et al. 2009, Skehel and Wiley 2000). The viral binding occurs in the receptor-binding site, which is located in the membrane distal tip of HA. The receptor binding site is bordered by three hypervariable domains, i.e. loop 220, loop 130, and helix 190 (Weis et al. 1988). Helix 190

(H190) has a defined helicity and contains a B-cell epitope, called site Sb; monoclonal neutralising antibody specific to this site has shown to be protective in mice (Yu et al. 2008). The antigenicity, biological function, and defined structure of H190 suggest that H190 is interesting to be utilised as a model antigen element for insertion into the MuPyV VP1 protein (Figure 1-3).

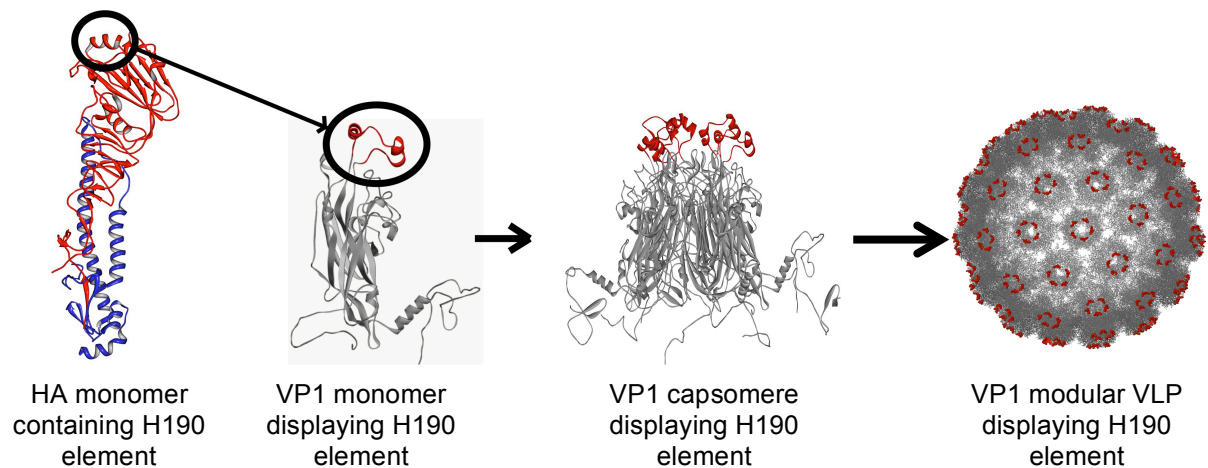


Figure 1-3. Illustration of generation of a modular MuPyV VP1 VLP presenting a peptide antigen element (helix 190 from HA protein). This figure was generated using Accelrys Discovery Studio® 3.0. and RasWin Molecular Graphics version 2.7.5.2. by Dr. Natalie Connors.

The structure of B-cell epitopes can regulate the neutralising mechanisms of antibodies (Dormitzer et al. 2008). This is particularly of interest for immunisation with peptides containing B-cell epitopes. Exogenously administered peptides, removed from the context of the full parent protein(s), may not preserve their conformational integrity, and thus may not necessarily undergo the same pathways of immune processing as the intact proteins. As a result, although B-cell epitope peptides can easily induce antibodies, the quality of antibodies induced against B-cell epitope peptides is different from the quality of those induced against the intact proteins (Purcell et al. 2007, Purcell et al. 2003). The quality of antibodies here is defined as the ability of the antibodies to recognise the native proteins. Immunisation with exogenous epitope peptides very rarely elicits antibodies that recognise the native proteins. Thus, in order to obtain antibodies that cross-react with the native antigen protein, the epitope peptides need to maintain conformational integrity (Purcell et al. 2007, Purcell et al. 2003). The quality of antibodies has been shown to resemble the protection efficacy of the antibodies (Alvarez et al. 2010).

A higher antibody quality indicates a higher potential of protection efficacy, thus a high quality of antibodies is desirable.

Studies in modular VLPs presume that it is nearly impossible for a peptide antigen element to assume its conformational integrity when it is presented on a modular VLP (Jennings and Bachmann 2007, Roldão et al. 2010, Tissot et al. 2010). Until now, presentation of a peptide antigen element on a modular VLP and the impact of display strategy for presentation of the antigen element on the quality of the resulting antibodies are not fully understood. In this thesis, a VLP from MuPyV VP1 is genetically modularised with a H190 element from an influenza A virus using different display strategies, and the display strategies are linked to the quality of the antibodies. Four major problems addressed around the presentation of a H190 element on a modular MuPyV VP1 VLP are:

1. Effect of display strategy on the quality of antibodies

Various strategies have been explored in studies of peptide-based vaccines to drive peptide antigens to fold in their native structures. Examples of these strategies are: (i) the use of flanking helix promoter elements (De Filette et al. 2008, Relf et al. 1996), and (ii) tandem repeat display (Fontenot et al. 1995, Fontenot et al. 1993, Kovacs-Nolan and Mine 2006).

The first strategy explores the use of flanking helix promoter elements derived from GCN4. GCN4 is a well-characterised protein from yeast (Arndt and Fink 1986). Due to its defined helicity, GCN4 has been used in peptide-based vaccine studies as flanking helix promoter elements for a helical peptide antigen from group A *streptococcus* (GAS) bacteria (Hayman et al. 1997, Relf et al. 1996). Incorporation of the GCN4-derived elements to present the GAS antigen in a modular VLP has been reported (Rivera-Hernandez et al. 2013), but the question of whether the incorporated GCN4-derived elements have any impact on the quality of antibodies induced against the modular VLPs was not addressed.

The second strategy utilises tandem repeat display of peptides. Tandem repeat display has been explored to design peptide-based vaccines for various pathogens (Kovacs-Nolan and Mine 2006, Liu et al. 2004, Schuman et al. 2005). In structural studies utilising nuclear magnetic resonance (NMR), tandem repeat display has been shown to promote peptide repeats to fold into their native structures (Fontenot et al. 1995, Fontenot et al. 1993). Furthermore, the display strategy has been utilised solely to increase the immunogenicity of a peptide antigen element (Jain et al. 2010) and has not been linked to the quality of the antibodies generated.

The research question here is whether presentation of the H190 element on a modular VLP using these two display strategies will have any impact on the quality of antibodies induced against the modular VLPs. Comparison of antibody quality raised against H190 element(s) presented on a modular VLP using these two strategies is then required.

2. Display strategy improvements for induction of a higher quality of antibodies

As described above, antibodies with a higher quality are desirable because the quality of antibodies has been correlated to the protection efficacy of the antibodies. After the impact of the two display strategies on the antibody quality has been revealed, further investigation is then required to understand how to induce antibodies with a higher quality.

Two proposed approaches to increase the quality of antibodies are: (i) the use of adjuvant, and (ii) increasing copy number of H190 tandem repeats. The first approach relies on the use of adjuvant to increase the immunogenicity of H190 element on a modular VLP, which is then expected to result in an increased quality of induced antibodies. One of the effective adjuvants for VLPs is AdvaxTM-1. It is a polysaccharide adjuvant derived from delta-inulin (Cooper and Petrovsky 2011). It has been shown to promote the antigen sparring effects of various vaccine candidates, resulting in the increased protective efficacy (Honda-Okubo et al. 2012, Petrovsky et al. 2013, Saade et al. 2013). The second approach focuses on increasing the number of H190 tandem repeats. Previous studies on peptide-based vaccines show that tandem repeat display strategy affects both structure (Fontenot et al. 1995, Fontenot et al. 1993) and immunogenicity of peptide repeats (Jain et al. 2010, Rueda et al. 2004, Zheng et al. 1993).

The research questions here are: Which approach can result in induction of antibodies with a higher quality? What is the preferred display strategy to present the H190 element on a modular VLP?

3. Robustness of display strategy

As described above, the H190 element is a hypervariable region. Following identification of the preferred display strategy, there is a need to investigate the robustness of the display strategy for modularisation of H190 element variants. As an extreme test case is H190 from A/Victoria/210/2009 influenza, which is hydrophobic, and its hydrophobic residues are positioned in the middle of the H190 amino acid sequence. A hydrophobic stretch has been known to cause formation of insoluble aggregates, as well as incorrect folding of proteins. Although the negative effects of a hydrophobic stretch

have been widely recognised (Aleksaitė and Gedvilaitė 2006, Chiti et al. 2003, Esler et al. 1996, Karpenko et al. 2000, Kazaks et al. 2004, Otzen et al. 2000), approaches to minimise its effects on VLP assembly still rely on simultaneous expression of modular and unmodularised viral capsid proteins, which yields mosaic VLPs (Karpenko et al. 2000, Loktev et al. 1996).

Various approaches have been commonly utilised to improve the solubility of proteins. These approaches can be classified into two classes, i.e. (i) sequence-dependent approaches, and (ii) sequence-independent approaches. An example of the first approaches is the use of charged residues. Charged residues can be introduced outside hydrophobic stretches as fusion tags at the N- and C-terminal of proteins (Jung et al. 2011, Kato et al. 2007), or within hydrophobic stretches via point mutations (Miklos et al. 2012, Perchiacca et al. 2012). The second approach includes modifications of conditions for protein expression. One of the important factors during protein expression in *E. coli* is temperature. *E. coli* has a high translation and protein folding rate. Reducing expression temperature can lower translation and folding rate of protein, and subsequently increase the solubility of proteins (Esposito and Chatterjee 2006, Qing et al. 2004).

The questions here are: to what extent does the hydrophobic stretch affect the properties of the modular VLP? Can the addition of charged residues or lowered expression-temperature minimise the effects of the hydrophobic stretch? Between the two approaches, which one is more effective in minimising the effects of a hydrophobic stretch on a modular VLP?

4. Structural determination of H190 element on a modular VLP

The ultimate question that needs to be answered is: what is the structure of the H190 element when presented on a modular VLP using different display strategies? Solving the structure of the H190 element will provide invaluable inputs for future structure-based vaccine designs. A potential method to determine the structure of H190 elements on modular VLPs is X-ray crystallography, which is one proven technology used widely to determine the structure of proteins (Chen et al. 2000, Momany et al. 1996). The crystal structure for assembly-incompetent MuPyV VP1 capsomeres is available at up to 1.9 Å. In addition, the conditions in which the unmodularised VP1 protein can crystallise have been revealed (Stehle and Harrison 1997).

1.3 Research aims and objectives

This thesis aims to understand how to present a structural peptide antigen on a modular VLP in such a way that it can elicit high-quality antibodies. The helical influenza A antigen, H190, is used as a model peptide antigen element to be presented on a modular MuPyV VP1 VLP. Studies in this thesis are directed by the following objectives:

1. To compare effects of two different display strategies: (i) the use of flanking helix promoter elements, and (ii) tandem repeat display on the quality of antibodies.
2. To perform improvements to the tandem repeat display strategy in order to induce antibodies of a higher quality.
3. To investigate the applicability of tandem repeat display strategy for H190 element containing a hydrophobic stretch.
4. To take the first steps toward obtaining the atomic structure, using X-ray crystallography, of a peptide antigen element displayed on a modular VLP.

1.4 Thesis organisation

In this thesis, this introductory chapter is followed by six chapters.

Chapter 2 provides a literature review for central topics in this thesis, including influenza viruses, current influenza vaccine technologies, and VLPs.

Chapter 3 compares the effects of the use of flanking helix promoter elements and tandem repeat on the quality of antibodies raised against the H190 element on a modular VLP [objective (1)].

Chapter 4 investigates what improvements can be done to the display strategies to induce antibodies with a higher quality. Furthermore, this chapter also shows that tandem repeat display comprising two copies of H190 is the preferred display strategy to present H190 element on a modular VLP to induce antibodies with a higher quality [objective (2)].

In Chapter 5, the tandem repeat display strategy is applied for H190 element containing a hydrophobic stretch [objective (3)]. Effects of the hydrophobic stretch on the properties of the modular protein are discussed. Improvements to the modular VP1 protein to minimise the effects of the hydrophobic stretch, including the addition of charge residues and lowered expression temperature in the protein expression, are described.

Chapter 6 demonstrates the development of chromatographic purification processes for production of modular VP1 capsomeres with a high purity [objective (4)].

Chapter 7 contains a summary of findings from accomplished work in this thesis and suggestions for future work.

Chapter 2. Literature review

2.1 Influenza A viruses

Influenza A viruses belong to the Orthomyxoviridae family (Webster et al. 1992). The natural resources of influenza A viruses are aquatic birds (Subbarao and Joseph 2007), but they are also found in various host cells, including humans, pigs, and horses (Horimoto and Kawaoka 2005, Medina and García-Sastre 2011, Olsen et al. 2006). The viruses have a unique characteristic, that is their ability to transmit among wild and domestic birds, as well as mammalian species, (Mubareka and Palese 2011).

Influenza A virions (Figure 2-1) consist of two layers of shells: the outer shell and the inner shell. The outer shells of the virions are a host-derived lipid bilayer envelope. Two surface glycoproteins, haemagglutinin (HA) and neuraminidase (NA), and the M2 ion-channel protein are located within this lipid envelope. The inner shells of the virions are composed of matrix protein (M1). Furthermore, in the centre of the virions, eight single negative-sense RNA strands are encapsulated by nucleoprotein (NP). The RNA strands encode viral polymerase proteins PB1 and PB2, and PA, HA, NP, NA, M1 and M2 proteins and NS1 and NS2 proteins (Horimoto and Kawaoka 2005, Medina and García-Sastre 2011, Webster et al. 1992). The virions must contain these RNA strands to be infectious.

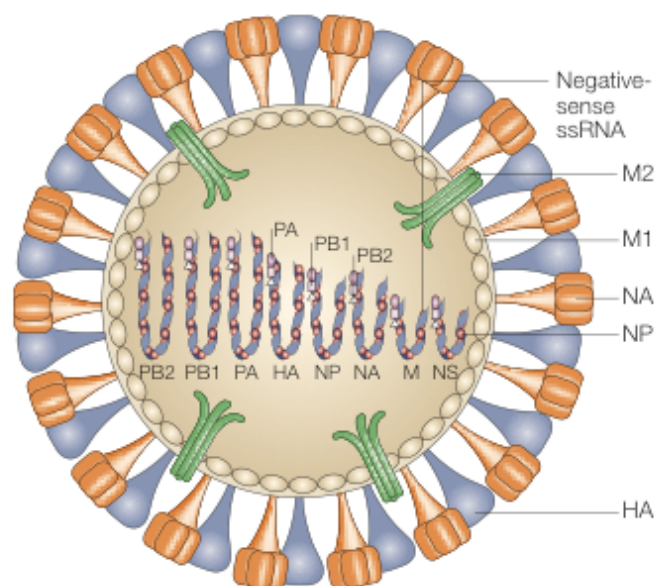


Figure 2-1. A schematic diagram of Influenza A virions. Figure taken from (Horimoto and Kawaoka 2005).

The life cycle of influenza A viruses can be divided into three steps (Figure 2-2). These steps are: (i) viral attachment to the host cells, (ii) viral entry and replication, and (iii) viral assembly budding and release (Mubareka and Palese 2011, Neumann et al. 2009). During the viral attachment process, influenza A viruses bind to host-cellular sialic acid receptors, which are located on the surface of epithelial cells. Human influenza viruses preferentially bind to sialic acid that is linked to penultimate galactose by an α 2,6-linkage (SA α 2,6Gal). This form of sialic acid is the major form of sialosaccharide in human tracheal epithelia, where the influenza viruses replicate. In contrast, avian influenza viruses preferentially bind to SA α 2,3Gal, which is distributed on epithelial cells in the intestinal tract of waterfowl and domestic poultry. The specificity of avian influenza viruses to SA α 2,3Gal reflects the ability of avian influenza viruses to transmit via the faecal-oral route (Skehel and Wiley 2000). Following attachment, the viruses enter the host cells through a process called receptor-mediated endocytosis. During the entry process, the cellular membrane of the host cells invaginates and encloses the virus, leading to proton entry and the formation of an endosome inside the cells. Protons create an acidic environment within the endosome. The pH changes then stimulate structural changes of the HA molecule of the viruses, and subsequently enables the viruses to fuse with the endosomal membrane. Simultaneously, protons are pumped into the viruses to separate the M1 and RNA-nucleoprotein complex (RNP). The RNP then enters the cytoplasm and migrates to the cell nucleus. The entry process is then followed by a replication process, in which the viruses take over the host cells to produce viral RNA and proteins. Viral nucleic acids migrate to the nucleus (i) to produce incoming viral RNA using the viral replicative enzymes, and (ii) to produce the positive strands of the viral RNA, which can function as messenger RNA for subsequent viral protein synthesis in the cytoplasm. Synthesised viral surface proteins consisting of HA, NA, and M2 assemble on the cellular membranes. Subsequently, other viral proteins place themselves under the cellular membranes. Through a budding process, newly synthesised viruses are released when the viral NA enzyme splits the sialic acid receptor, which holds the HA of the new viruses on the host cells.

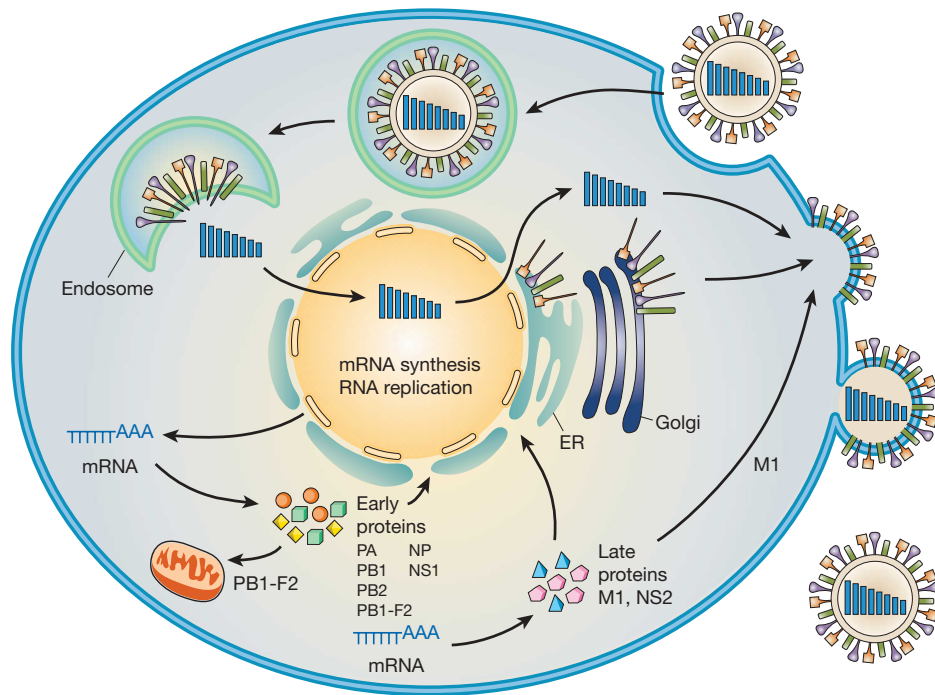


Figure 2-2. Schematic diagram of the life cycles of influenza A viruses. Viruses attach to the host cells via receptor binding. The viruses then enter the host cells through endocytosis and replicate. Following a budding process, newly synthesised viruses are released from the host cells. Figure taken from (Neumann et al. 2009).

Influenza A viruses evolve via a complex pattern of antigenic variation, which primarily occurs in the HA and NA. To date, 16 HA (H1-H16) and 9 NA (N1-N9) have been identified and used for influenza A virus subtyping (Neumann et al. 2009). The evolution of influenza A viruses consists of two mechanisms: (i) antigenic drift, and (ii) antigenic shift. Antigenic drift is the annual antigenic variation in the HA and/or NA due to the accumulation of point mutations, which result from the low fidelity of RNA polymerase. Antigenic drift can also result from immune pressure selection by host cells. Antigenic drift reduces the protection provided by pre-existing antibodies; thus, it has become the major challenge of annual influenza vaccine production. Antigenic shift is the result of viral reassortment, where one virus strain exchanges one or more segments of viral RNA strands with another strain. Due to the ability of influenza A viruses to transmit among birds and mammalian species, antigenic shift can result in the emergence of a novel strain containing HA and/or NA immunologically different from circulating strains, with a possibly distinct host range. The emergence of a novel serotype of HA may lead to a pandemic with a huge impact, such as significant morbidity and mortality globally, and social and economic disturbance (Haaheim et al. 2010, Hay et al. 2001).

There have been three major influenza pandemics in the past. These pandemics occurred in 1918-1919, 1957, and 1968. In 1918-1919, the deadly and highly contagious Spanish influenza virus (H1N1) caused the death of approximately 40 million people worldwide. This pandemic is recognised as the most deadly pandemic in history. In 1957, a less serious pandemic of Asian influenza (H2N2) emerged. HA of the H2N2 subtype only shared 66% identity at the amino acid level with HA of the H1N1 subtype. Similarly, NA of the H2N2 subtypes was 37% similar to NA from the H1N1 subtype. Due to the antigenic variation, pre-existing immune responses in humans could not provide protection against this new H2N2 subtype, resulting in the death of approximately 70,000 people in the United States alone, and many more worldwide. HA of the H2N2 subtypes then further evolved, leading to the emergence of H3N2 subtypes. The pandemic caused by the Hong Kong influenza virus (H3N2) occurred in 1968 and claimed the lives of more than 30,000 people worldwide. In addition to these three pandemics, another occurred in 2009. Within one year, the pandemic resulted in the death of more than 18,000 people and left between 43 and 89 million infected people. The pandemic strain was of the swine-origin H1N1 subtype, which contained genes from avian, swine, and human influenza viruses (Cox and Subbarao 1999, Horimoto and Kawaoka 2005, Medina and García-Sastre 2011, Palese 2004, Potter 2001).

2.1.1 Haemagglutinin (HA)

HA is synthesised as a single polypeptide chain, HA0, and undergoes posttranslational process consisting of proteolytic cleavage, fatty acid acylation, and glycosylation. Signal sequences for transport to the cell membrane are removed from a newly synthesised HA. Carbohydrate side chains are then added in positions that vary between strains. This process is followed by the addition of palmitic acid to cysteine residues, close to the HA carboxyl terminus. The posttranslational process is finalised by cleavage of the HA by a specific host-produced trypsin-like protease. HA is cleaved into subunits HA1 and HA2, which are connected via disulphide bonds (Webster et al. 1992). This cleavage of HA into HA1 and HA2 is required for membrane fusion activity and virus infectivity (Skehel and Wiley 2000).

During maturation, HA molecules form homotrimers (Webster et al. 1992) with a molecular weight of approximately 224 kDa (Figure 2-3). These homotrimers are about 135 Å in length and have a triangular cross section varying in radius from 15 to 40 Å (Wilson et al. 1981). In addition, there are two distinct regions: (i) a globular region composed of an antiparallel β -sheet (Figure 2-3B), and (ii) a stem region, consisting of a

triple-stranded coiled coil of a helix (Figure 2-3C). The coil extends 76 Å from the membrane and consists of HA1 (typically about 328 amino acids) and HA2 (typically about 222 amino acids). The carboxyl terminus of HA2 contains the hydrophobic transmembrane sequence and a terminal cytoplasmic anchor sequence where palmitate is connected (Figure 2-3D). The globular region is solely composed of HA1, and it contains most of the antigenic sites of the molecule as well as the receptor-binding site (Katz and Webster 1992, Wilson et al. 1981).

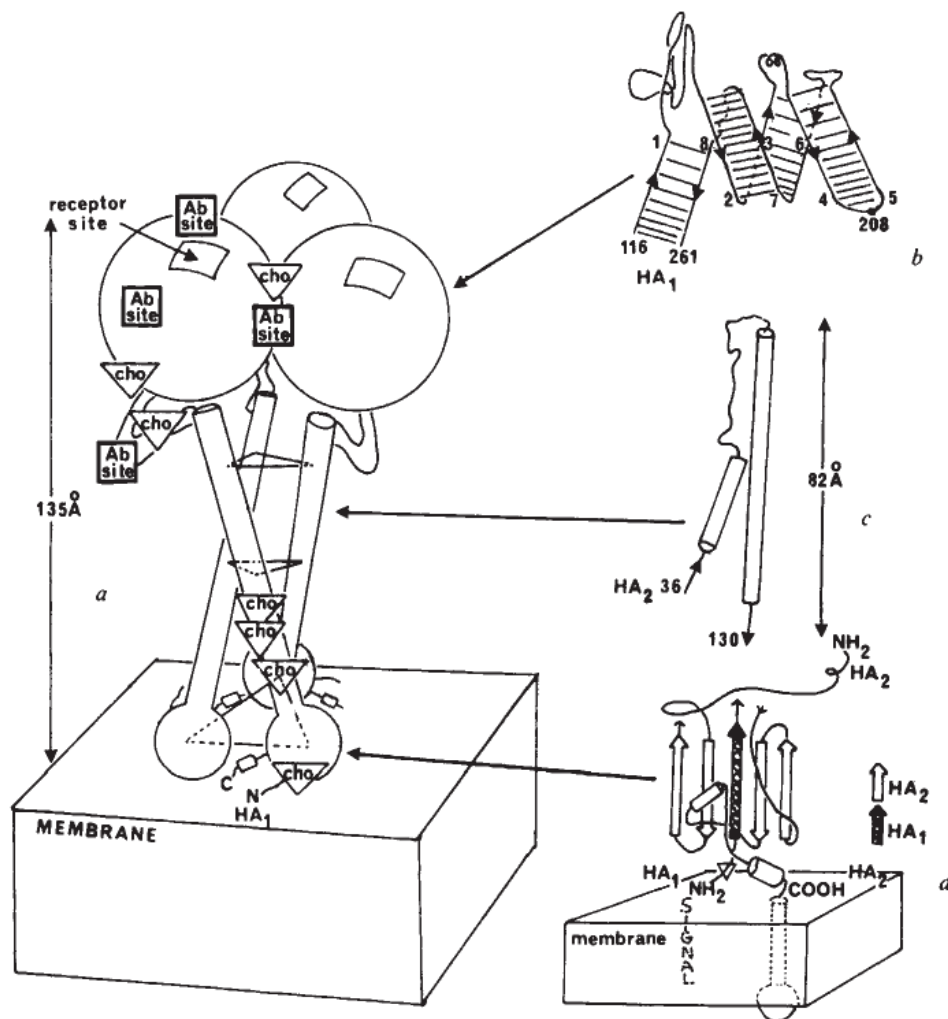


Figure 2-3. An illustration of HA molecules. (A) Schematic diagram of HA trimers, showing its antibody-binding sites (Ab site), the receptor-binding site, carbohydrate attachment sites (CHO) and its position on the membrane; (B) The eight-stranded β -sheet and looped-out region in the globular domain; (C) HA2, showing two α -helices (cylinders). In HA trimers, the long helices from each HA monomer pack together and form a triple-stranded coiled-coil, which stabilises the trimers; and (D) The membrane end of HA contains a five-stranded β -sheet and hydrophobic anchoring peptide (broken line). The central strand (black) is the N-terminal of HA1, and the other strands are the C-terminal of HA2. Figure taken from (Wilson et al. 1981).

2.1.2 Antibody-binding site

HA1 is the primary target for the immune system in a form of neutralising antibodies, which sterically prevent viral attachment and thus neutralise viral infectivity (Das et al. 2010). Antigenic variations in HA due to the evolution of influenza A viruses through antigenic shift and antigenic drift enable the viruses to escape from pre-existing antibodies. To date, the localisation and structure of antibody-binding sites on the three dimensional structure of the HA have been assessed for three influenza A virus subtypes, H3, H5, and H9.

The first structural characterisation and antigenic mapping was conducted for HA from A/Hong Kong/1/1968 (H3N2) (Wiley et al. 1981). Wiley et al. suggested four antibody-binding sites in the HA (Figure 2-4). Amino acid substitutions on these sites were shown to affect the binding of neutralising antibodies. These sites were designated site A, B, C, and D.

1. Site A (●) consists of residues from 140 to 146 and has the structure of a protruding loop. It is 8 Å away from the surface and the most vivid antibody-binding site.
2. Site B (■) is an α -helix, composes of 187 to 196 amino acids and nearby residues along the upper side of receptor-binding sites.
3. The third antigenic site, site C (▲), is located between the disulphide bond formed from Cys52 to Cys277. The site is at about 60 Å from the distal tip of the HA molecule.
4. Site D (◆) is located in the interface region between subunits in the HA trimers. Unlike other antigenic sites, amino acid substitution detected in site D may be recognised because of relative movement of the globular domain HA1.



Figure 2-4. Antibody-binding sites in A/Hong Kong/1/1968 (H3N2) HA1. Red, green, blue, and orange domains refer to the antibody-binding site A, site B, site C, and site D, respectively. The figure is adapted from (Wiley et al. 1981) and generated using UCSF Chimera (Pettersen et al. 2004) using the crystal structure of HA from A/Hong Kong/1/1968 (4FNK.pdb) (Ekiert et al. 2012).

The location of antibody-binding sites in the H3 HA was used to map antigenic sites in HA from A/PR/8/34 (H1N1) (Caton et al. 1982). Caton et al. suggested four antibody-binding sites in the H1N1 HA. These sites were designated Sa, Sb, Ca and Cb (Figure 2-5); the residue numbering of these sites was based on H3 numbering.

1. Sites Sa and Sb reside in the upper part of the globular head of HA1, and are closely linked. These sites cannot accommodate the simultaneous binding of antibodies to each site. Site Sa is in the “front” part of the globular head, and comprises residues 128, 129, 158, 160 and 162 to 167, excluding 164. The site is relatively close to the receptor-binding site of the adjacent monomer in trimeric HA. Site Sb is the “back” part of the globular head, comprises residues that form the upper edge of the receptor-binding site. These residues are 192, 193 and 196 of an α -helix region and residues 198, 156 and 159.
2. Site Ca is composed of two regions, Ca1 and Ca2. In the three-dimensional structure of a HA monomer, these regions are widely separated. Region Ca1 is formed by residues 169, 173, 207, and 240. Residues 182 and 273 are also

possibly involved. Region Ca2 is formed by residues 140, 143, 145, 224 and 225. However, in trimeric HA molecules, regions Ca1 and Ca2 are in close proximity.

3. Site Cb is a region near the bottom of the globular head of HA1, comprised of residues 78 to 83.

Caton et al. highlighted that these four antibody-binding sites correspond to sites in HA from A/Hong Kong/1/1968 (H3N2) identified by Wiley et al. (Wiley et al. 1981). However, no equivalent of site C in H3 HA could be found in the H1 HA.

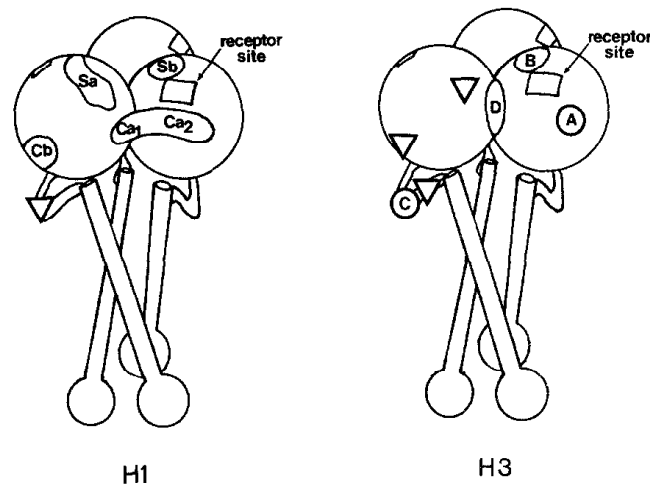


Figure 2-5. The position of the antibody-binding sites on HA from A/PR/8/34 (H1N1) relative to the arrangements on HA from A/Hong Kong/1968 (H3N2). Triangles show the position of the carbohydrate attachment sites and rectangles show the receptor-binding sites. Figure taken from (Caton et al. 1982).

Furthermore, a study by Xu et al. (Xu et al. 2010) successfully identified antibody-binding sites Sa, Sb, Ca, and Cb in various H1 subtypes. In the study, Xu et al. aligned the amino acid sequences of the antibody-binding sites for several different H1 subtypes, as shown in Figure 2-6. The H1 subtypes were A/California/04/2009 and A/South Carolina/1/1918, which are pandemic influenza strains; A/Puerto Rico/8/34, which is a laboratory strain; and A/Brisbane/59/2007, which is a seasonal influenza strain. The alignment shows that, for those H1 subtypes, the antibody-binding sites are discontinued epitopes except for site Sb. In addition, the alignment shows that the amino acid sequences of the four antibody-binding sites vary among the listed H1 subtypes, showing that even within a subtype, the amino acid sequences of the antibody-binding sites are variable.

CLUSTAL O(1.2.1) multiple sequence alignment

```

A/Brisbane/59/2007      MKVKLLVLLCTFTATYADTICIGYHANNSTDVTVDTVLEKNVTVTHSVNLLEN SHNGKLC
A/California/04/2009   -----PGDTLCIGYHANNSTDVTVDTVLEKNVTVTHSVNLLEDKHNGLCK
A/South Carolina/1/1918 -----ADPGDTICIGYHANNSTDVTVDTVLEKNVTVTHSVNLLED SHNGKLC
A/Puerto Rico/8/34     -----ADADDTICIGYHANNSTDVTVDTVLEKNVTVTHSVNLLED SHNGKLC
                        **:*****:

A/Brisbane/59/2007      LKGIAPLQLGNCSVAGWILGNPECELLISKESWSYIVEKPNPENGTCYPGHFADYEELRE
A/California/04/2009   LRGVAPLHLGKCNIAAGWILGNPECESLSTASSWSYIVETPSSDNGTCYPGDFIDYEELRE
A/South Carolina/1/1918 LKGIAPLQLGKCNIAAGWLLGNPECDLLLTASSWSYIVETSNSENGTCYPGDFIDYEELRE
A/Puerto Rico/8/34     LKGIAPLQLGKCNIAAGWLLGNPECDPLLPVRSWSYIVETPNSENGICYPGDFIDYEELRE
                        *:***:*.:.***:*****: *      *****. . :** ****.* *****

A/Brisbane/59/2007      QLSSVSSFERFEIFPKESSWPNHTV-TGVSASC SHNGES SFYRNLLWLTGKNGLYPNLSK
A/California/04/2009   QLSSVSSFERFEIFPKTSSWPNHDSNKGVTAAC PHAGAK SFYKNLIWLVKKGNYPKLSK
A/South Carolina/1/1918 QLSSVSSFKEFEIFPKTSSWPNHETTKGVTAAC SYAGAS SFYRNLLWLT KKGSSYPKLSK
A/Puerto Rico/8/34     QLSSVSSFERFEIFPKESSWPNHNT-NGVTAAC SHEGKS SFYRNLLWLT EKEGSSYPKLN
                        *****:***** ***** .**:* * : * .***:***:* * **:*.:

A/Brisbane/59/2007      SYANNKEKEVLVLWG VHHPP NIGVQKALYHTENAYVSVVSSHYSRKF TPEIAKRPKVRDQ
A/California/04/2009   SYINDKGKEVLVLWG IHHPS TSADQQLYQNADTYV FVGSRRYSKFKFPEIAIRPKVRDQ
A/South Carolina/1/1918 SYVNNKGKEVLVLWG VHHPP TGTDQQLYQNADAYV SVGSSKYNRRF TPEIAARPKVRDQ
A/Puerto Rico/8/34     SYVNNKGKEVLVLWG IHHPP NSKEQQLYQNE NAYVSVVTSNYNRRF TPEIAERPVRDQ
                        ** *. * *****:*** . * : **:. : ** * :*. . :*. ***** *****

A/Brisbane/59/2007      EGRINYWTLL EPGDTIIFEANGNLIAPRYAFALSRGFGSGIINSNAPMDKCDAKCQTPQ
A/California/04/2009   EGRMNYWTLV EPGDKITFEATGNLVVPRYAFAMEERNAGSGIIISDTPVHDCNTTCQTPK
A/South Carolina/1/1918 AGRMNYWTLL EPGDTITFEATGNLIAPWYAFALNRGSGSGIITS DAPVHDCNTKCCQTPH
A/Puerto Rico/8/34     AGRMNYWTLL KPGDTIIFEANGNLIAPMYAFALRRGFGSGIITSN ASMHECNTKCCQTPL
                        **:*****:***.* ***.***:.* ****: * ***** **: :. .:***:****

A/Brisbane/59/2007      GAINSSLPFQNVHPVTIGECPKYVRS AKLRMVTGLRNIPSIQSRGLFGAIAGFIEGGWTG
A/California/04/2009   GAINSSLPFQNIHPITIGKCPKYV KSTKLRLATGLRNIPSIQSR-----
A/South Carolina/1/1918 GAINSSLPFQNIHPVTIGECPKYV RSTKLRMATGLRNIPSIQSR-----
A/Puerto Rico/8/34     GAINSSLPYQNIHPVTIGECPKYV RSAKLRMVTGLRNIPAR-----
                        *****:***:*.:.***:*****:***:*.:.*****:

```

Figure 2-6. Comparison of amino acid sequences of antibody-binding sites in various H1 subtypes. Residues highlighted in blue, pink, orange, and purple refer to antibody-binding site Sb, Sa, Ca, Cb, respectively. The alignment is reproduced from (Xu et al. 2010) using Clustal Omega (<http://www.ebi.ac.uk/Tools/msa/clustalo/>).

Different from the HA molecules from H3 and H1 subtypes, in H5 and H9 HA, only two antibody-binding sites were suggested. In H5 HA, the sites correspond to the antibody-binding sites A and B in H3 HA (Kaverin et al. 2007). Meanwhile, in H9 HA, the first site is composed of residues forming site A (position 135 in H3 numbering) and site B (position 157 and 162 in H3 numbering) of H3 HA. The second site comprises residues at positions 145 (site A in H3), 193 (site B in H3), and 226 (site D in H3). Furthermore, in H9 HA, these two antigenic sites are overlapping (Kaverin et al. 2004).

2.1.3 The receptor-binding site

Topographically, the receptor-binding site is a depression at the membrane distal tip of the HA monomer. The site is composed of conserved residues, and bordered by three secondary structure domains, which are composed of highly variable residues. These

domains are: (i) helix 190, comprising of HA1 residues 190 to 197, (ii) loop 130, consisting of HA1 residues 135 and 138, and (iii) loop 220, consisting of HA1 residues 221 and 228 (numbering based on H3 A/Aichi/1/68; Figure 2-7A). Helix 190 is an α -helical structure. In contrast, loop 130 and loop 220 lack secondary structures. Furthermore, while the secondary structure of helix 190 are similar between subtypes, the conformations adopted by loop 130 and loop 220 are similar within influenza subtypes, but they vary greatly between subtypes (Gamblin et al. 2004). Furthermore, the bottom of the receptor-binding site is composed of the phenolic hydroxyl of Try98 and the aromatic ring of Trp153. The rear of the site is defined by Glu190 and Leu194 from the helix 190 as well as His183 and Thr155. Residues 134 and 138 compose the right side, while residues 224 and 228 compose the left side (Skehel and Wiley 2000, Stevens et al. 2006, Weis et al. 1988, Yang et al. 2007). In the study by Yang et al. (Yang et al. 2007), the amino acid sequences of the three domains in HA from H1, H3, and H5 subtypes are compared. The comparison shows that the three domains in HA molecules from H1 and H5 have greater similarities in structure and genetics to one another than to H3 (Figure 2-7B).

Receptor-binding specificity is important for two reasons. The first is related to the determination of the virus host ranges, which limits virus transmission across species. Studies show that mutation of the conserved residues can alter the receptor-binding specificity, and consequently lead to cross-species transfer. Different influenza virus subtypes require different mutations to alter their receptor-binding specificity. For example, for H2 and H3 subtypes, mutations Gln226Leu and Gly228Ser shift the receptor-binding specificity from SA α 2,3Gal to SA α 2,6Gal (Gamblin et al. 2004, Stevens et al. 2006, Yamada et al. 2006). However, the same mutations in H5 subtypes result in a lack of binding to sialic acid SA α 2,6Gal. In contrast, the changes Ser137Ala and Thr192Ile in H5N1 strains increase binding to SA α 2,6Gal (Yang et al. 2007). The second is related to the selection processes during virus replication *in vitro*, which are relevant for both influenza surveillance and vaccination. It has been shown that during vaccine production, viral propagation in eggs or mammalian cells can result in antigenically different variants, which may not be representative of the circulating viruses (Skehel and Wiley 2000). Yang et al. (Yang et al. 2007) show that if mutations occur on the receptor-binding site, they can change the neutralisation sensitivity of antibodies. In their study, monoclonal antibodies (mAbs) were generated from the vaccination of mice with H5 Glu190Asp, Lys193Ser, Gly225Asp triple mutant HA and wild-type HA. A potent antibody, 9B11, neutralised the triple mutant but not the wild type. Another potent antibody, 9E8, neutralised wild type HA

but showed reduced neutralisation against the triple mutant as well as the Ser137Ala, Thr192Ile mutant.

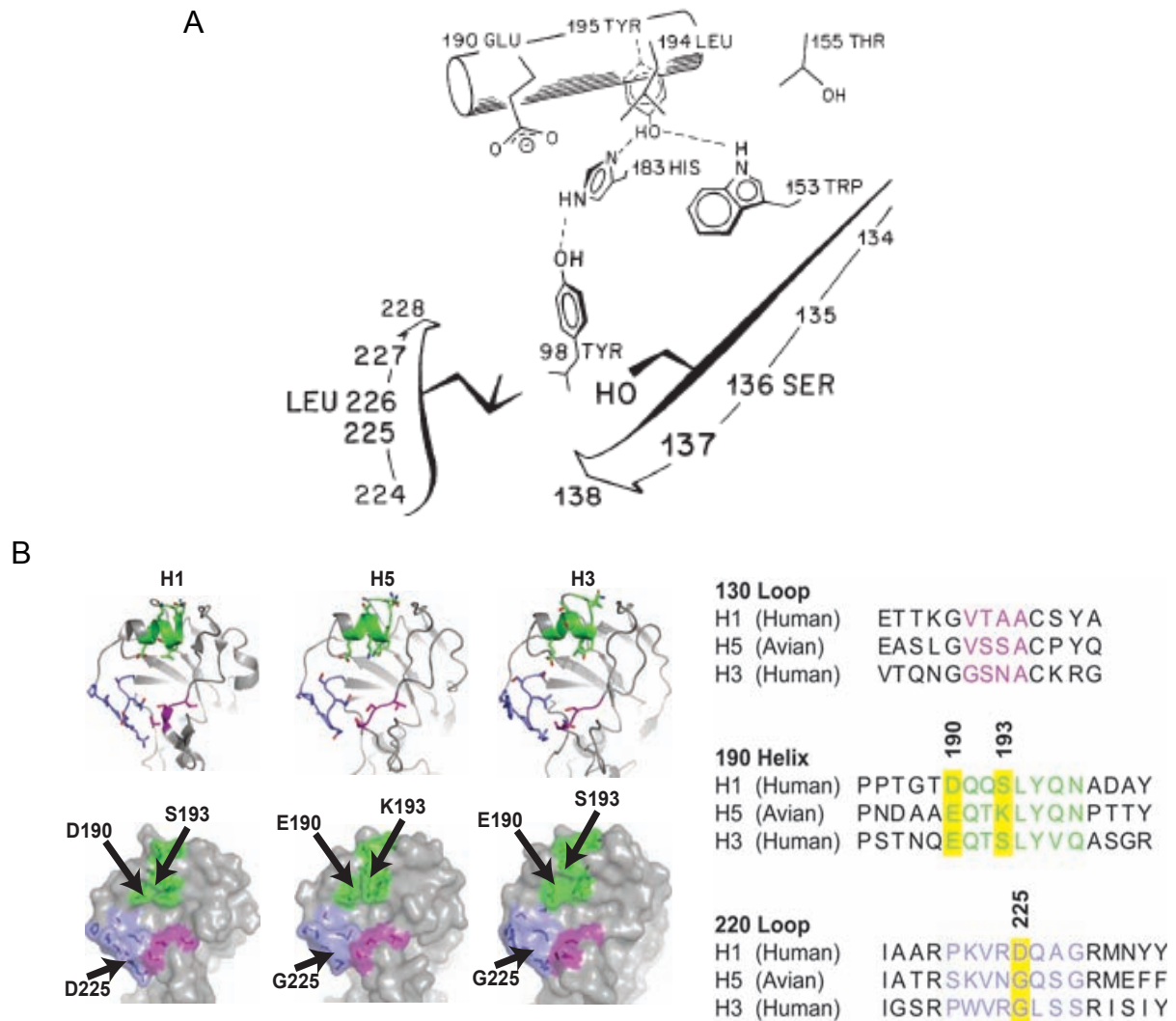


Figure 2-7. The receptor-binding site within HA molecule. (A) A schematic diagram of the receptor-binding site in A/Aichi/2/68 HA. Hydrogen bonds between conserved residues in the receptor-binding site are shown. Figure taken from (Weis et al. 1988); and (B) Comparison of amino acid sequences of loop 130 (pink), helix 190 (green), and loop 220 (purple) in 1918 H1, Vietnam04 H5, and 1968 H3 subtypes. Figure taken from (Yang et al. 2007).

2.1.4 Correlation of antibody-binding site and receptor-binding site

To correlate antibody-binding sites (see Figure 2-6) and receptor-binding domains (see Figure 2-7) in the H1N1 strain, the amino acid sequences of receptor-binding domains and antibody-binding sites in the H1N1 strain are compared (Figure 2-8). The comparison reveals that loop 130 does not overlap with any antibody-binding sites within

the H1N1 HA. In contrast, loop 220 partially overlapped with the antibody-binding site Ca, and helix 190 was contained within antibody-binding site Sb.

A similar comparison was also made for HA from A/Hong Kong/1/1968 (Figure 2-9). The comparison revealed that, similar to the comparison of the H1N1 strain above, helix 190 comprises antibody-binding site B. However, loop 130 and loop 220 do not overlap with any antibody-binding sites.

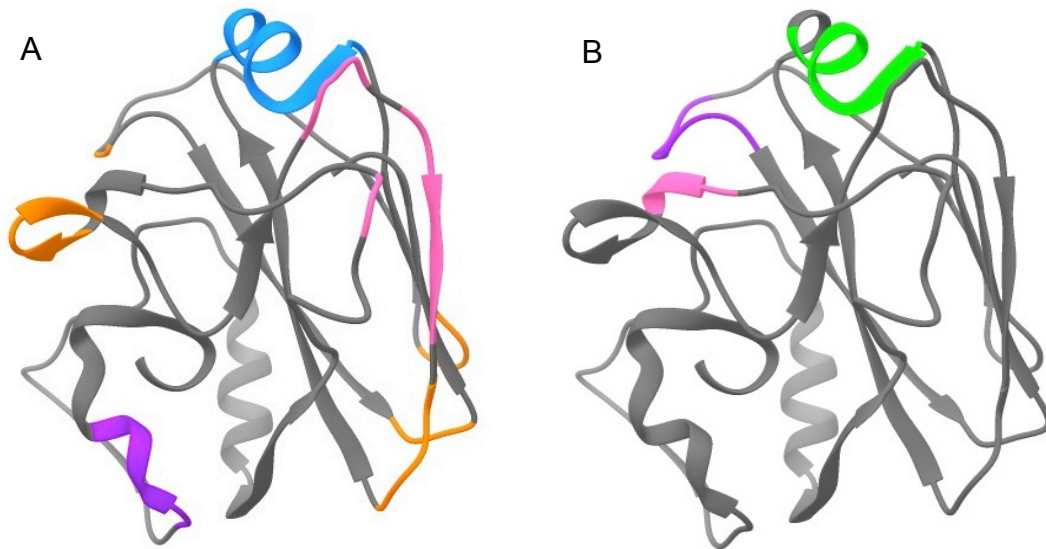


Figure 2-8. Comparison of antibody-binding sites and receptor-binding domains in the globular head of HA molecule from A/South Carolina/1/1918 (H1N1) influenza. In panel A: (blue) antibody-binding site Sb, (pink) antibody-binding site Sa, (orange) antibody-binding site Ca, and (purple) antibody-binding site Cb. In panel B: (pink) loop 130, (green) helix 190, and (purple) loop 220. Figures were generated using UCSF Chimera (Pettersen et al. 2004) with the crystal structure of HA from A/South Carolina/1/1918 (1RUZ.pdb) (Gamblin et al. 2004).

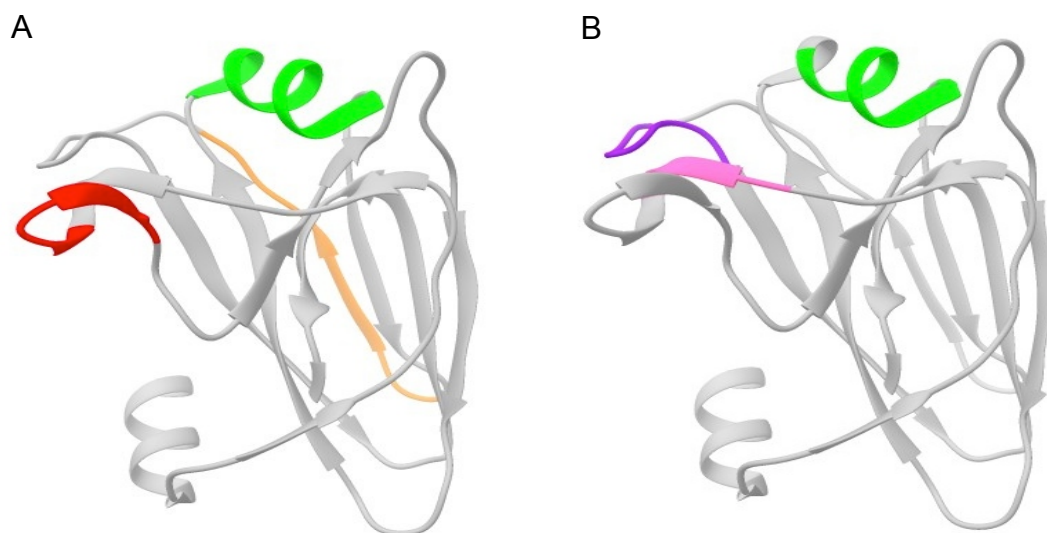


Figure 2-9. Comparison of antibody-binding sites and receptor-binding domains in the globular head of HA molecule from A/Hong Kong/1/1968 (H3N2) influenza. In panel A: (red) antibody-binding site A, (green) antibody-binding site B, and (orange) antibody-binding site D. In panel B: (pink) loop 130, (green) helix 190, and (purple) loop 220. Figures were generated using UCSF Chimera (Pettersen et al. 2004) with the crystal structure of HA from A/Hong Kong/1/1968 (4FNK.pdb) (Ekiert et al. 2012).

2.1.5 Glycosylation of HA

HA undergoes post-translational modification in the form of N-linked glycosylation, which is one of the most common forms of post-translational modification. HA produced in mammalian cells and chicken eggs possesses heterologous complex type N-linked glycans, whereas HA produced in insect cells possesses heterologous high-mannose type N-linked glycans (Chen et al. 2011). N-linked glycan is attached to the amide nitrogen of asparagine in a conserved motif of Asn-X-Ser/Thr (Vigerust and Shepherd 2007).

In HA, N-linked glycans are found in both the stem region and the globular domain of HA molecules (Keil et al. 1985). Glycosylation sites in the stem region are highly conserved and are required to maintain structural integrity and stability during the synthesis process (Daniels et al. 2003, Kawaoka and Webster 1989). These sites are at or near residues 5, 26, 289, 483, and 542 (H3 numbering). Furthermore, these sites are conserved, possibly due to proper interactions between HA and glycan binding ER during HA folding and assembly (Das et al. 2010). Glycosylation in the stem region has been shown to affect virus virulence by changing the optimum pH required for fusion (Ohuchi et al. 1997) and the infectivity of influenza viruses by modulating the cleavage process of HA (Deshpande et al. 1987).

In contrast, glycosylation sites in the globular domain vary from one strain to another. They are added or removed during evolution, and the complexity of the glycans is different. For example, the globular domain of H3 HA gradually gains more glycosylation

sites than H1 HA with a similar circulating time (Xu et al. 2010). There are a maximum of six glycosylation sites in H3 HA, whereas in H1 HA, only three glycosylation sites are encoded, which are centred on residues 92, 129, and 162. In addition, H2 HA maintains one glycosylation site in its globular domain, even though the strain circulated for about 10 years, from 1957 to 1968 (Das et al. 2011, Das et al. 2010). The minimum alteration in glycosylation site in H2 HA indicates that glycosylation greatly affects the biological function of H2 HA (Tsuchiya et al. 2002, Tsuchiya et al. 2002). Furthermore, glycosylation in the globular domain determines the pathogenicity of influenza A virus. It also functions to help the virus evade pre-existing immune responses, and the induction of cross-reactivity polyclonal antibody responses (Abe et al. 2004, Medina et al. 2013, Wanzeck et al. 2011). Analysis of glycosylation sites in pandemic influenza viruses demonstrates that the viruses are highly virulent because of the low number of glycosylation sites in their globular domain. As pandemic strains evolve into seasonal strains, the influenza viruses acquire more glycosylation sites to sterically hinder antibodies binding to the globular domain. It has been shown that N-glycan removal in seasonal influenza viruses increases viral virulence. Additionally, immunisation with seasonal strains lacking N-linked glycans can induce antibodies to cross-react with pandemic strains, although seasonal strains did not share similarities with pandemic strains.

During viral evolution, some glycosylation sites are deleterious, while other glycosylation sites are maintained as they provide fitness and/or escape from immune responses, resulting in antigenic drift (Medina et al. 2013, Reichert et al. 2010, Sun et al. 2011). A study involving the simulation and prediction of glycosylation in HA from H1 strains by Das et al. (Das et al. 2011) demonstrated that glycosylation in one of the antibody-binding sites reduces the variability of residues around the site, and subsequently increases the variability of other sites. Moreover, when variability occurs at residues 190 and 228, the receptor-binding specificity of the influenza virus is consequently altered (Das et al. 2011).

As glycosylation in the globular domain interferes the binding of neutralising antibodies, it has been considered that glycosylation is also responsible for the low efficiency of conventional influenza vaccines based on the use of highly-glycosylated HA. Therefore, glycosylation should be included as one of the most important factors in influenza vaccine design. A study by Wang et al. (Wang et al. 2009) correlated the complexity of N-linked glycan in HA and the protection efficacy of the HA. In the study, the protection efficacy of fully-glycosylated HA (HA_{fg}) was compared to the efficacy of monoglycosylated HA (HA_{mg}). HA_{mg} was produced by trimming down the complexity of N-

linked glycan in HA_{fg} using Endo H enzyme. The enzymatic treatment left a single N-acetylglucosamine (GlcNAc) glycan residue attached to the glycosylation site. Both HA_{mg} and HA_{fg} were shown to be structurally similar, indicating that changes in the complexity of N-linked glycan did not alter the structure of the HA. However, the study showed that HA_{mg} antiserum demonstrated a higher neutralising activity and mice immunised with HA_{mg} showed a higher survival rate.

The single GlcNAc glycan residue of N-linked glycan has been proposed to be important for protein folding and structure stabilisation (Chen et al. 2011, Hanson et al. 2009, Khurana et al. 2010). Therefore, it has been argued that proper folding of HA could not be achieved in the absence of glycosylation. Since the induction of neutralising antibodies necessitates a native conformational epitope, HA lacking of N-linked glycan may not be protective. Khurana et al. (Khurana et al. 2010, Khurana et al. 2011) challenged this assumption by showing that bacterially produced HA1, which was non-glycosylated, was able to induce a higher titre of neutralising antibodies than HA produced in mammalian cells. The study also demonstrated that mice immunised with bacterially produced HA1 were protected from morbidity.

2.2 Influenza vaccines

2.2.1 History of vaccination

Vaccination is one of the most important inventions in human history. It is one of the most cost-effective methods to control and prevent the spread of various infectious diseases (Ehreth 2003, Ulmer et al. 2006). Vaccination was first introduced in China, or India, back in the fifteenth century through the practice of variolation. In variolation, the smallpox virus was artificially given to healthy individuals in order to raise immunity against the virus. There were various methods of variolation, i.e. (i) exposing a piece of cotton filled with pox pus or squama into the nostril of healthy individuals, (ii) putting clothes worn by infected individuals on healthy individuals, or (iii) blowing powdered squama into nostril of healthy individuals using a thin silver tubes. Variolation was widely used, possibly due to its ability to protect the variolated individuals from developing scarring. However, unavoidably, some variolated individuals also died due to illness which developed from the inoculation process itself (Leung 2011).

The practice of variolation was then replaced by Jennerian vaccination. In 1796, Edward Jenner inoculated fluid extracted from a cowpox lesion in the arm of an 8-year old boy. Jenner demonstrated that the boy developed a slight illness from the cowpox and was protected from procuring smallpox following the inoculation. The use of the cowpox

virus to achieve protection against smallpox was a significant development over variolation (Bonanni and Santos 2011, Leung 2011). In addition, Jenner's work was the first demonstration of a safer approach for vaccination using live organism. However, the approach was still empirical, without knowledge of the origin of the virus and an understanding of virus transmission and infection. The term "vaccine inoculation" was used for the first time to describe this process (Letvin 2007, Spohn and Bachmann 2008).

Eighty years later, Louis Pasteur took vaccination to the next level of development in the laboratory, marking the first generation of developments in vaccine technologies. He discovered that exposure to adverse conditions could attenuate the virulence of pathogens. He also showed that inoculation could induce whole-body immunity, rather than just immunity at the point of injection. Pasteur's approach for vaccination was distinct from the practice of variolation and Jennerian vaccination. Variolation and Jennerian vaccination involved inoculation with live pathogens to induce immunity against the pathogen. In contrast, Pasteur's approach elaborated the use of attenuated pathogens. However, similar to variolation and Jennerian vaccination, Pasteur's approach was still empirical, in the absence of knowledge regarding the structure and pathogenicity of the pathogens (Barbosa and Barral-Netto 2013, Bonanni and Santos 2011, Leung 2011). In the last decade of the nineteenth century, the first generation of vaccines for various diseases, such as typhoid fever, cholera, and the plague, were produced and tested (Plotkin and Plotkin 2011). Another example of the first generation vaccines is inactivated whole-virus influenza vaccine that was developed in 1930s using chicken eggs, discussed in Section 2.2.2 below (Thomas and Luxon 2013). In the twentieth century, the first generation of vaccine development reached the golden age through the growth of viruses in cell culture (Enders et al. 1949, Plotkin and Plotkin 2011). Using the cell culture technology, inactivated and live attenuated polio vaccines were produced during 1950s, followed by vaccines for mumps, rubella, rotavirus, and influenza. The first generation of vaccines ended following the development of purified polysaccharide vaccine to combat meningococcus serogroups A, C, Y, and W (Austrian et al. 1976), as well as *Haemophilus influenzae* type b (Anderson et al. 1972). The development in the use of purified-polysaccharide based vaccines then lead to the development of conjugate vaccines, in which the polysaccharides were conjugated to a protein (Avery and Goebel 1929, Robbins and Schneerson 1990).

Although conjugate vaccines were successfully used to combat the two mentioned pathogens, the technology was not applicable for meningococcus serogroup B. The inapplicability of conjugate vaccine drove the development of the second-generation

vaccines, harnessing genomic sequencing. This vaccine technology is often described as reverse vaccinology (Thomas and Luxon 2013). In reverse vaccinology, the predicted antigens are screened for development as vaccine candidates by combining advance genomic sequencing, bioinformatics tools, high-throughput expression and purification systems, and serological screening assays (Dormitzer et al. 2008, Rappuoli 2000, Thomas and Luxon 2013).

2.2.2 Structural vaccinology

Nevertheless, the first and second generation vaccines are not suitable for “refractory pathogens”, such as those with highly variable surface antigens and those which cannot be produced using conventional vaccine production. To overcome these barriers, vaccination is developed towards the third generation, which is structural vaccinology, a.k.a. structure-based antigen design. Structural vaccinology combines genetic engineering and structural information about pathogens obtained using various analysis tools (listed in Table 2-1). It relies on the peptide-antigen-focusing concept, in which the entire protein is not necessary for the induction of protective immune responses. Therefore, in structural vaccinology, vaccination is focused on the selection of specific peptide antigens that elicit protective immune responses based on structural information. It avoids the use of antigens that can induce disease-enhancing immunity and vaccine-mediated diseases (Dormitzer et al. 2008, Thomas and Luxon 2013).

The peptide-antigen-focusing concept led to the emergence of peptide-based vaccines. However, peptide-based vaccines possess various disadvantages. For example, (i) they have poor immunogenicity, (ii) usage is restricted to patients of a given tissue type, and (iii) the antigen has a small size and low copy number. Two approaches to overcome these problems are: (i) the use of the larger domains from the pathogens containing selected peptide antigens but lacking other non-protective and undesirable antigens or domains; and (ii) the presentation of peptide antigens to unrelated carrier proteins for conformational stability (Kulp and Schief 2013). Examples of carrier proteins are: virus-like particles (VLPs) (Middelberg et al. 2011, Schneemann et al. 2012), green fluorescence protein (Abedi et al. 1998), cholera toxin B subunit (Totrov et al. 2010), and self-assembling protein nanoparticles (Wahome et al. 2012).

Table 2-1. Analysis tools in structural vaccinology, taken from (Dormitzer et al. 2008).

Property analysed	Techniques	Utility
Three-dimensional structure of antigen-antibody complexes	X-ray crystallography, NMR, cryo-EM	Allow for rational engineering by defining domain boundaries, epitope structure, and underlying architecture
Antigen structure	ELISA, IP, escape mutant analysis, DXMS, phage display	Define the link between physical structure and the landscapes recognised by antibodies
Post-translational modification	SDS-PAGE, MS, glycosidic linkage analysis, X-ray crystallography, NMR	Assess the authenticity and homogeneity of modifications on recombinantly expressed proteins
Protein folding and stability	CD, ITC, DXMS, NMR, DSC, protease protection, native and SDS-PAGE	Assess antigen conformation and integrity in solution over time for vaccine stability
Non-covalent association and hydrodynamic radius	AUC, DLS, SEC, SPR	Assess antigen valency and aggregation

Abbreviations: AUC, analytical ultracentrifugation; CD, circular dichroism spectroscopy; ELISA, enzyme-linked immunosorbent assay; IP, immunoprecipitation; ITC, isothermal titration calorimetry; EM, electron microscopy; DLS, dynamic light scattering; DXMS, deuterium exchange mass spectrometry; MS, mass spectroscopy; NMR, nuclear magnetic resonance spectroscopy; SEC, size exclusion chromatography; SPR, surface plasmon resonance.

Driven by structural considerations, structural vaccinology allows for the engineering of peptide antigens (i) to improve their feasibility as vaccine candidates, and (ii) to rapidly respond to antigen variability (Correia et al. 2014, Dormitzer et al. 2012). The engineering of peptide antigens is required if B-cell immune responses are targeted. The induction of different immune responses involves different mechanisms. T-cell immune responses are produced through complex intracellular antigen-processing pathways, which are independent of the antigen conformations. Thus, when T-cell immune responses are targeted, engineering of the peptide antigen is not required. In contrast, the neutralising mechanism of antibodies may be dependent on the structure of B-cell epitopes (Dormitzer et al. 2008). Thus, to induce antibodies recognising the parental viruses, B-cell epitopes in their native conformation are necessary.

When peptide antigens are taken out of context from their parental proteins, they may not be able to preserve their native conformation. Similarly, when peptide antigens are presented in unrelated scaffold proteins, they may not be able to assume their native conformation. B-cell epitopes in non-native conformation may not undergo the same pathways of immune processing as the parental proteins. As a result, although B-cell epitopes are highly immunogenic, the induced antibodies may not be able to recognise the parental proteins (Purcell et al. 2007, Purcell et al. 2003). Thus, in order to obtain antibodies that cross-react with the native antigen protein, engineering of peptide antigens is required in order to maintain the native conformation of the epitope (Purcell et al. 2007,

Purcell et al. 2003). Various approaches have been used to induce the native conformation of peptide antigens in peptide-based vaccines. These approaches include: (i) the use of structure promoter (De Filette et al. 2008, Relf et al. 1996), (ii) increasing copy number of B-cell epitope (Fontenot et al. 1995, Fontenot et al. 1993, Kovacs-Nolan and Mine 2006, Schuman et al. 2005), (iii) incorporation of T_H-epitopes and B-cell epitopes (Fitzmaurice et al. 1996), and (iv) cyclisation (Luzzago et al. 1993, Tugyi et al. 2005). However, approaches to induce the native conformation of peptide antigens presented on unrelated carrier proteins are still limited to (i) grafting, and (ii) the use of structural promoters.

In the grafting approach, continuous-structural antigens are transplanted to scaffold proteins, assisted by computational design, whilst maintaining the structure and antigenicity of the peptide antigens (Kwong and Shapiro 2011). Engineering peptide antigens using the grafting approach is comprised of four steps. The first step is finding appropriate scaffold proteins with similar backbone structures to the targeted peptide antigens. The second step is filtering the scaffold proteins by retaining those that can bind to antibodies without any significant clashes. The third step is transplantation of side chains from the targeted peptide antigens at appropriate positions on the selected scaffolds. The fourth step is the introduction of additional mutations into the selected scaffolds to optimise stability, to expose the antigens from the surface, and to minimise undesirable interactions between antibodies and the scaffolds (Ofek et al. 2010). To date, this approach has mostly been used to present helical epitope peptides. Studies of the gp41 epitopes from HIV (Correia et al. 2010, Ofek et al. 2010) have shown that the grafting approach successfully presents the epitope in a conformation that is the same as in the template antibody-antigen complex. Nevertheless, antibodies induced against the grafted epitope are not effective for neutralisation (Ofek et al. 2010, Ofek et al. 2004). The approach has also been used for the Helix A epitope from HA2 of the influenza A virus (Schneemann et al. 2012), and the motavizumab epitope from the Respiratory syncytial virus (McLellan et al. 2011).

In the second approach, structural promoters are used to induce native conformation of peptide antigens present in unrelated VLPs. Studies by Gleiter et al. (Gleiter et al. 1999) and Gleiter and Lilie (Gleiter and Lilie 2001) have shown the use of structural promoters to present the peptide epitope in unrelated VLPs. In these studies, amino acid sequences comprised of glycine and serine residues were used, aiming to provide flexibility for the peptide epitopes on the surface of the VLPs.

2.2.3 Current technologies for influenza vaccine production

2.2.3.1 Egg-based influenza vaccine production

Vaccination remains the main method to prevent seasonal influenza and reduce the rate of morbidity and mortality (Kang et al. 2012). More than 95% of the currently available influenza vaccines in markets are produced in embryonated chicken eggs (Perdue et al. 2011). One million doses of seasonal trivalent influenza vaccines are produced annually using egg-based technology (Lee and Hu 2012). Egg-based technology was licensed in 1945 (Ulmer et al. 2006) to produce inactivated purified whole virus for influenza vaccines (Lee and Hu 2012). In egg-based technology, influenza viruses are grown in the allantoic cavity of embryonated-chicken eggs. Grown viruses are then concentrated using ultracentrifugation and inactivated using formalin or β -propiolactone (Maassab et al. 1990, Murphy and Coelingh 2002). Grown viruses can also be treated using ether or Tween (Brandon et al. 1967, Hoyle et al. 1961, Neurath et al. 1971) to disturb the virion, producing split influenza vaccines, or treated with detergent, such as Triton N 101 (Brady and Furminger 1976), ammonium deoxycholate (Laver and Webster 1976), or cetyltrimethylammonium bromide (Bachmayer et al. 1976) to recover the HA and/or NA from the virions. Vaccines based on split virion and recovered surface antigens had lower adverse reactions than those based on inactivated whole virus. To increase virus yield, high-growth reassortants are used. The reassortant viruses contain six internal genes from the A/PR/8/1934 H1N1 and two external genes (HA and NA) from circulating strains (Lee and Hu 2012, Shaw 2012).

Due to antigenic drift and antigenic shift on the HA and NA of influenza virion, the influenza vaccine must be reformulated every year. The World Health Organisation (WHO) manages an international surveillance system to observe the epidemiology of influenza viruses, and the surveillance data are reviewed by WHO Collaborating Centre investigators in February and September. Based on the review, the WHO then decides which vaccine strains to be included in the preparation of the following season's influenza vaccines. Current influenza vaccines contain two strains of influenza A subtypes (H3N2 and H1N1) and one strain of influenza B. Influenza vaccines are available on the market approximately 6-8 months after the decision is made. Due to the time delay, the vaccine strains may not accurately match the circulating strains. A typical annual vaccine manufacturing process using egg-based technology is shown Figure 2-10.

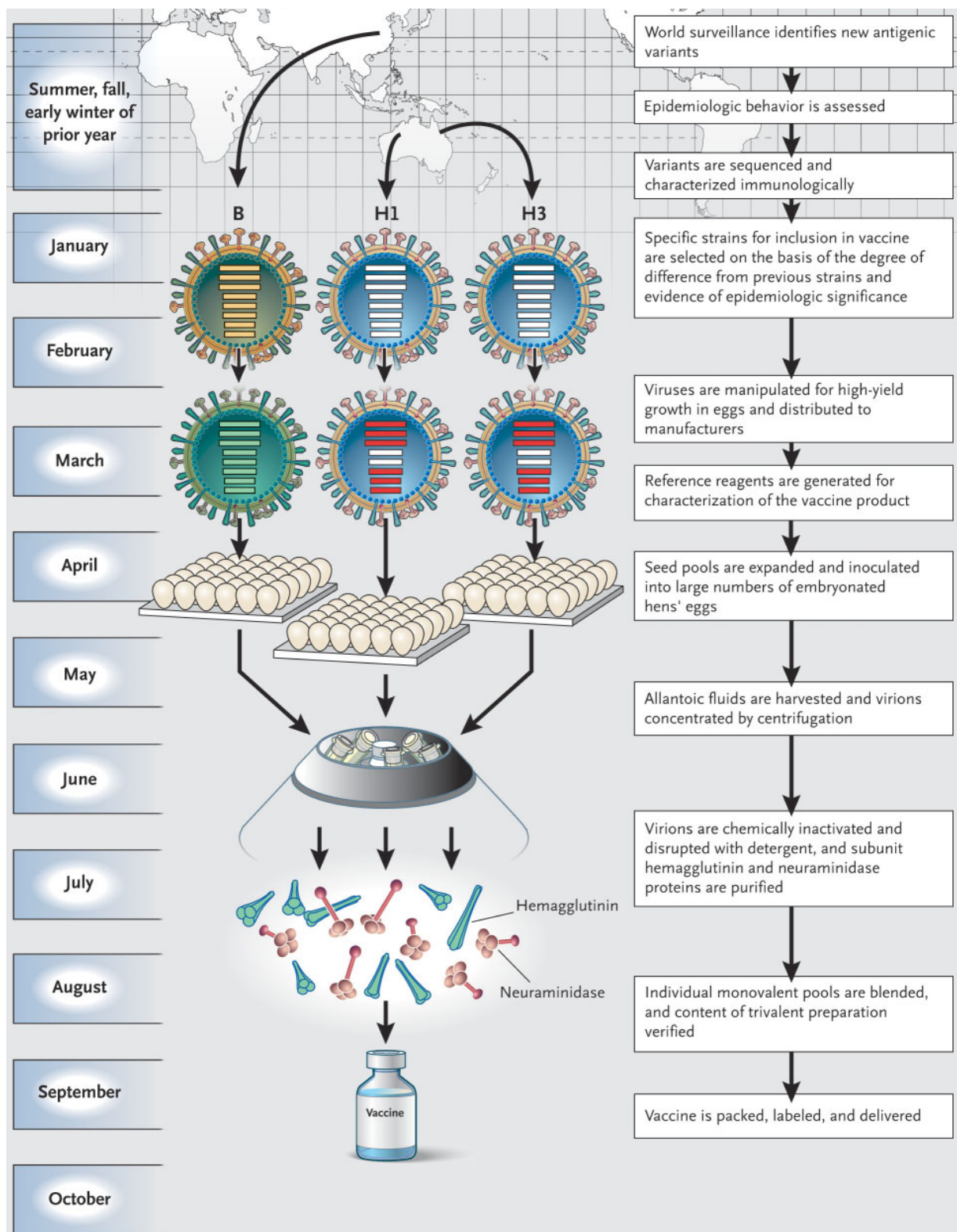


Figure 2-10. Manufacturing of seasonal trivalent influenza vaccines using egg-based technology (Treanor 2004).

Although egg-based vaccine production is widely used, this method has several limitations. First, egg-based vaccine production is time-consuming and labour intensive. In addition, a large number of eggs are required (Kistner et al. 1998, Tree et al. 2001). Therefore, vaccine production using egg-based technology requires an extended period of preparation. Consequently, the vaccine strains may not match the circulating strains when they become available. Second, in egg-based production, eggs are a major part of the production process (Tree et al. 2001). However, the use of eggs is not suitable for avian influenza viruses, such as H5N1. Optimal growth conditions for avian influenza viruses are difficult to achieve because the high pathogenicity of the virus can result in considerable tissue destruction. Although such viruses can still replicate in eggs at high temperatures and shortened incubation times, the yield is low (Stockmann et al. 2009). Third, eggs can also potentially select for variants that grow well in the environment used during the adaptation process (Kistner et al. 1998, Ulmer et al. 2006). Nucleotide sequencing has indicated that the HAs of influenza A strains grown in the egg have a single base substitution and subsequently a single amino change (Katz and Webster 1992, Robertson 1993). Such amino acid substitutions can result in antigenic mismatch between vaccine strains and circulating strains (Tree et al. 2001). Finally, a trace amount of egg protein can become a latent source of allergic reaction for recipients (Ulmer et al. 2006). Furthermore, the handling of pathogenic viruses requires a biosafety level 3 working environment (Kang et al. 2009, Robertson and Engelhardt 2010).

2.2.3.2 Cell culture-based influenza vaccine production

The limitations of egg-based vaccine technology have ignited the emergence of new vaccine production using cell culture. The production of influenza vaccine using cell culture-based technology has been developed since the 1970s. In 1995, the WHO officially recommended cell culture as a new method to produce influenza vaccines (Ghendon et al. 2005). The first licensed cell culture-based influenza vaccine was approved in the Netherlands in 2001. Production of vaccines in cell culture-based technology is similar to egg-based, which is summarised in Figure 2-11. Cell culture-based technology also produces inactivated whole virion influenza vaccines, split influenza vaccines, or surface antigen vaccines. *A priori* differences in the effectiveness of influenza vaccines produced from eggs or cell cultures in commercial production systems have not been observed. Influenza vaccines produced using both technologies have exhibited equal immunological efficacy and safety.

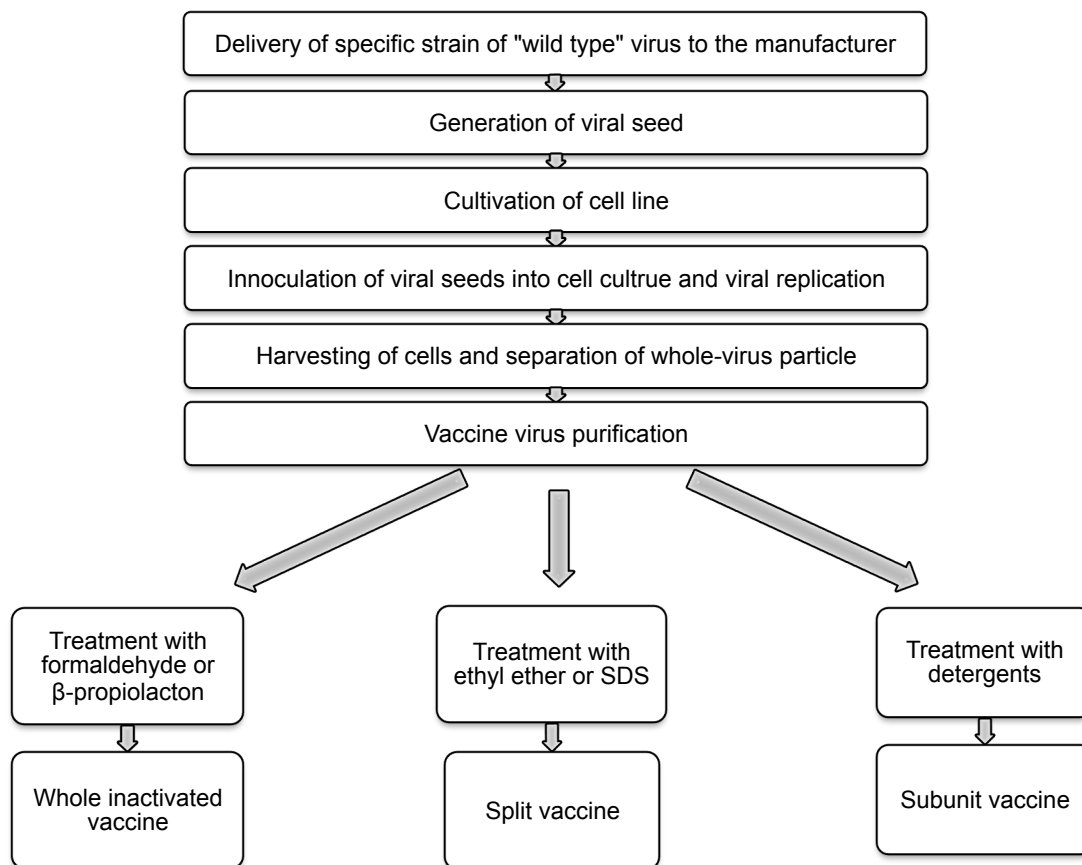


Figure 2-11. A schematic flow-chart of vaccine production using cell culture-based technology. Figure reproduced from (Doroshenko and Halperin 2009).

Various cell lines have been widely used for commercial vaccine development and production. These cell lines are Madin-Darby canine kidney (MDCK) (Doroshenko and Halperin 2009, Hu et al. 2008, Liu et al. 2009), Vero (African green monkey kidney cell) (Kistner et al. 1998), Human diploid cell lines WI-38 and MRC 5 (secondary human lung fibroblasts), chick embryo fibroblast cells, PerC6 (immortalised human cell lines) (Lewis et al. 2005, Pau et al. 2001), and primary monkey kidney (PMK) (Chezzi et al. 1998). Vero and MDCK cells are licensed for the production of human influenza vaccines. The types of vaccines being produced, e.g. inactivated whole virion, subunit or split vaccines, determine the selection of cell lines (Perdue et al. 2011, Youil et al. 2004).

Cell culture-based production offers many advantages over egg-based production. Firstly, cell culture-based production facilitates a faster vaccine production, which is beneficial in the case of pandemics. Vaccine production using egg-based technology requires about 6-8 months, whereas the use of cell culture-based technology vaccine enables production within 2-3 months. The use of various growth media also allows the same system for the production of various different vaccine products. Secondly, cell

culture-based production is controllable and easy to scale-up because continuous cell lines permit the use of fully characterised and standardised cells (Halperin et al. 1998, Murakami et al. 2008, Tree et al. 2001). Lastly, adverse effects, such as allergic reaction to certain egg components, can be minimised in vaccines produced using cell culture-based technology (Murakami et al. 2008). In addition, with the current advanced manufacturing practice, the contamination rate in cell culture-based vaccine production can be reduced to below 1%. In contrast, contamination is inevitable in egg-based technology due to the non-sterile nature of eggs.

However, cell culture-based production also has some limitations. Firstly, to be able to operate at a comparable capacity to egg-based production, cell culture-based technology requires an initial capital cost that is 2-3 times higher. Cell culture-based technology also requires higher maintenance costs. Secondly, cell lines must not contain any adventitious agents. The list of such agents is continually growing, and the testing is laborious, challenging, and time-consuming. Lastly, similar to egg-based production, cell culture can result in cell line-specific gene mutation which can cause further changes in the HA and NA proteins (Shank-Retzlaff et al. 2006) and subsequently lower immunogenicity (Johansson 1999, Wang et al. 2006).

2.3 Virus-like particles (VLPs)

A promising alternative to inactivated whole virion or split virion vaccines is virus-like particles (VLPs). VLPs are produced using genetic recombination technology, without the use of live viruses as the starting material. VLPs are defined as a particulate structure composed of the self-assembling viral capsid proteins. The capsid protein self-assembles to form sub-structures (capsomeres) that are subsequently organised to form particles (capsids). VLPs mimic native viruses, but they lack regulatory proteins and infectious genetic materials. Therefore, VLPs are replication-incompetent and safe for vaccine application (Buonaguro et al. 2001, Noad and Roy 2003, Pattenden et al. 2005, Roldão et al. 2010). The absence of infection genetic materials may also enable the differentiation of immunised and naturally infected animals. Such application is particularly of interest for foot-and-mouth disease vaccines (Parida 2009, Smith et al. 2014). As particulates with repetitive and ordered surface, VLPs are effective for (i) antigen delivery to antigen-presenting cells (APCs), and (ii) stimulating innate and humoral immunity at equal or greater levels than inactivated whole virion and subunit vaccines (Malboeuf et al. 2007, Takamura et al. 2004, Touze and Coursaget 1998). The particulate nature of VLPs can also promote the induction of strong B-cell responses without the need for adjuvants

(Chackerian et al. 2002, Jegerlehner et al. 2002). Furthermore, VLPs have a size range of 20-100 nm, allowing them to reach the draining lymphoid organs following subcutaneous injection and to be taken-up by lymph node-resident cells (Buonaguro et al. 2001).

To date, VLPs have been used to combat parental viruses from which the VLPs are derived (Jennings and Bachmann 2009). Clinically approved VLP-based vaccines for this purpose are Gardasil[®] and Cervarix[™] for human papillomavirus (Jennings and Bachmann 2009, Koutsky et al. 2002, Villa et al. 2005), and Engerix-B or Recombivax-B for Hepatitis B virus (Averhoff et al. 1998, Keating et al. 2003, Poland and Jacobson 2004).

VLPs have also been used to display antigen modules, yielding modular VLPs. Antigen modules commonly contain short antigen elements with expected secondary structures (Crisci et al. 2012, Neiryneck et al. 1999, Xia et al. 2011, Yin et al. 2011) or larger domains with expected tertiary structures (Buonaguro et al. 2001, Kratz et al. 1999, McGinnes et al. 2011, Murawski et al. 2010). Antigen modules may also contain other necessary elements, such as structure promoter elements. The repetitive patterns on the surface of VLPs increase the immunogenicity of the antigen modules significantly. Two commonly used methods for displaying antigen modules are (i) genetic fusion, and (ii) conjugation of the antigen modules on pre-formed VLPs (Zeltins 2013).

In the first method, DNA-encoding antigen modules are cloned into DNA encoding the viral capsid protein. Antigen modules are fused in the region of the viral capsid protein that is exposed on the surface, including the N- and C-termini (Roldão et al. 2010). Genetic fusion of the antigen module combines structural knowledge and DNA recombinant engineering. To assist in the assembly of viral capsid proteins bearing antigen modules to form modular VLPs, modular viral capsid proteins can be co-expressed with wild-type viral capsid proteins. The resultant VLPs contain both wild type and modular capsid proteins at determined proportions. This technique is called mosaic VLPs (Jennings and Bachmann 2007, Karpenko et al. 2000, Loktev et al. 1996). Nevertheless, the production of modular VLPs, in which antigen modules are genetically fused into the viral capsid proteins, is highly unpredictable, depending on (i) the interaction between antigen modules and the viral capsid proteins, (ii) glycosylation efficacy, (iii) the expression system, and (iv) the length of the antigen module. It is thought that it is nearly impossible for the fused antigen module to assume quaternary structures (Roldão et al. 2010). To date, there have been four modular VLP-based vaccines, which are comprised of genetically-fused antigen modules, entered into clinical trials. These vaccines are: (i) RTS.S particle (GSK and PATH Malaria Vaccine Initiative and Bill & Melinda Gates Foundation-Rixensart, Belgium), which contains modular VLPs from the hepatitis B core

antigen (HBcAg) bearing amino acids 207-295 of the circumsporozite (CS) protein in the N-terminus of HBcAg (Nardin et al. 2004); (ii) HIV p17/24:Ty by British Biotech Pharmaceuticals Ltd (Oxford, UK), which contains modular VLPs from the p1 protein encoded by the yeast retrotransposon Ty. The modular VLPs are comprised of a *gag* component of HIV, 33 amino acids of p17 and 177 amino acids of p24 (Jennings and Bachmann 2007, Weber et al. 1995); (iii) Malarivax (ICC-1132), by Apovia, Inc., which contains modular VLPs from HBcAg bearing T-cell and B-cell epitopes from the circumsporozite (CS) protein (Roldão et al. 2010), and (iv) ACAM-FLU-A by Sanofi Pasteur Co., which contains modular HBcAg VLPs bearing a genetically-fused epitope from the external domain of M2 protein (M2e) (Fiers et al. 2009).

In the second method, antigen modules are linked to the surface of preassembled VLPs. VLPs and antigen modules are generated separately. The antigen modules are then conjugated to the VLPs *in vitro* by covalent or non-covalent bonding. Non-covalent bonding exploits the strong interaction between streptavidin and biotin (Chackerian et al. 2001), while covalent bonding utilises chemical cross-linkers. In covalent bonding, two distinct reactive groups are coupled to distinct functional target, one on the VLPs, and the other one on the antigen modules (Roldão et al. 2010). Two modular VLPs with conjugated antigen modules have entered clinical trials. These vaccines are: (i) nicotine Q β VLPs by Cytos Biotechnology AG and Novartis, and (ii) AngII-Q β VLPs by Cytos Biotechnology AG. Nicotine Q β VLPs are generated by chemically conjugating nicotine to the bacteriophage Q β VLPs, which are produced via *E. coli*. The conjugation is performed using a succinate linker with nicotine to VLPs ratio of 580. Nicotine Q β VLPs is the first modular VLPs in clinical trial for non-infectious disease (Maurer et al. 2005). Similar to Nicotine Q β VLPs, AngII-Q β VLPs are generated by chemically conjugated AngII peptide to Q β VLPs using SMPH (Succinimidyl-6-[(β -maleimidopropionamido)hexanoate]) (Ambühl et al. 2007).

2.3.1 VLP production

VLPs can be produced using various expression systems, both eukaryotic (e.g. yeast, insect cells, plants) and prokaryotic (e.g. *Escherichia coli*) host cells. This chapter focuses on baculovirus-insect cells and *E. coli* as the most widely used expression systems for VLP production.

2.3.1.1 Baculovirus-insect cell expression vector system (BEVS)

BEVS is one of the most preferred expression systems for VLP production. The expression system mostly uses Sf9 and High FiveTM (Invitrogen, CA, USA) cell lines. Sf9 cell lines, originating from *Spodoptera frugiperda*, are mostly used for the generation and propagation of recombinant baculovirus (rBVs). High FiveTM cell lines are derived from *Trichoplusia ni* and mostly used to produce recombinant protein (Fernandes et al. 2013, Mena and Kamen 2011). An approved VLP-based vaccine produced in *Trichoplusia ni* cell lines is CervarixTM (Koutsky et al. 2002, Villa et al. 2005).

BEVS offers many advantages as a template for VLP expression system. Firstly, the expression system can reach high cell densities in suspension culture. Thus, high yield recombinant proteins could be achieved (Noad and Roy 2003, Zeltins 2013). Moreover, the proteins undergo eukaryotic post-translational modification, thus the biological activity of the proteins can be retained (Zeltins 2013). Secondly, baculovirus has a very limited host range, including a few species of Lepidoptera. The virus replicates efficiently in insect cells ($>10^8$ pfu/ml). However, the virus is not able to replicate in mammalian cells, although it can infect the cells. In addition, the virus only presents in the nucleus and culture media of insect cell preparations, while VLPs, the product, is in the cytoplasm. For these reasons, VLP produced in this expression system is safe for human use (Noad and Roy 2003). Lastly, BEVS is applicable for large-scale production (Noad and Roy 2003, Zeltins 2013) and allows the production of multiple proteins effectively and simultaneously (Mena and Kamen 2011, Tang et al. 2011).

However, BEVS also faces various challenges for VLP manufacturing. Firstly, the self-assembly of VLPs in insect cells may not be properly performed. BEVS commonly uses two promoters, Polh and p10. The use of both promoters can increase the expression yield. However, the two promoters are activated very late post-infection, when the gradual degradation of insect cells occurs. As a consequence, post-translational modification may not properly performed and thus the self-assembly of VLPs may not be achieved properly (Liu et al. 2013). Secondly, although proteins produced in insect cells can fold more similarly to those produced in mammalian cells, this folding is inconsistent. Consequently, the production of VLPs in insect cells may result in heterologous products (Mena and Kamen 2011). Thirdly, because VLP self-assembly in insect cells occurs *in vivo*, the VLPs formed may contain live baculovirus or host nucleic acid as contaminants. This contamination increases batch-to-batch product variability (Mena and Kamen 2011, Pattenden et al. 2005), and affects the immunology of the VLPs, although it is believed to be at an insignificant level (Haynes 2009). To remove the contaminants, additional

disassemble-reassemble processes are necessary. Lastly, post-translational processes in this expression system are unpredictable, incomplete, and potentially immunogenic (Pattenden et al. 2005).

2.3.1.2 *Escherichia coli* expression system

Various VLPs have been successfully produced using the *E. coli* expression system. For example, VLPs from polyomavirus VP1 protein (Chuan et al. 2008, Leavitt et al. 1985, Salunke et al. 1986), hepatitis B core antigen (Bundy et al. 2008, Zlotnick et al. 1996), human papillomavirus (Schädlich et al. 2009), and hepatitis E (Wu et al. 2012). The last example is the first *E. coli*-produced VLP that is currently undergoing a clinical trial.

The *E. coli* expression system offers many advantages. Firstly, the *E. coli* expression system is well characterised, allowing control and direct access to the optimisation of transcription and translation. These parameters are difficult to regulate in BEVS (Bundy et al. 2008). Secondly, the *E. coli* expression system offers a high yield of proteins, and the expressed proteins are easy to purify and recover (Bundy et al. 2008, Zhang et al. 1998). The well-studied denaturation and refolding mechanism also provides a solution to overcome the solubility problem for some recombinant proteins, especially those which originate from eukaryotic hosts using the *E. coli* expression system, may result in insoluble proteins (Zeltins 2013). Thirdly, in the *E. coli* expression system, the self-assembly of VLPs is performed *in vitro*. Therefore, the optimisation of assembly conditions can be controlled, and consequently more homogenous products can be obtained (Chuan et al. 2008). *In vitro* assembly also minimises the contamination of host nucleic acids and eliminates the disassembly-reassembly process (Pattenden et al. 2005). Lastly, the *E. coli* expression system allows for the production of VLPs within weeks (Lua et al. 2013).

Although the *E. coli* expression system offers various advantages, it still has many challenges, which are yet to be answered. Challenges in VLP production using the *E. coli* expression system include (i) the absence of post-translational modification of proteins expressed in *E. coli*, (ii) the presence of endotoxins in the outer membrane of *E. coli*, and (iii) the presence of heat shock proteins or chaperone proteins, which may affect downstream processes (Zeltins 2013).

2.3.1.3 Current studies on VLP-based influenza vaccines

Current studies on VLP-based influenza A vaccines are focused on the paradigm of broadly cross-protective vaccines, which exploit highly conserved antigens from influenza

A virus, including the HA stalk domain (HA2), M2e, and NA. Antigens from HA2 include (i) a fusion peptide; (ii) a HA cleavage site that is an extended, highly exposed loop structure on the surface, and (iii) the long α -helical polypeptides, which are conserved among H1, H2, H3, H5, and H7 (Kang et al. 2012).

These studies can be classified into two main streams. In the first stream, VLPs are formed from influenza viral capsid proteins and are produced using insect cell expression systems. The first studies in this stream demonstrated the production of VLPs based on the simultaneous expression of four structural proteins, HA, NA, M1 and M2 (Cox 2008, Latham and Galarza 2001). Similarly, many studies have developed VLPs based on HA and/or NA on influenza M1 VLPs both for seasonal (Bright et al. 2007, Quan et al. 2007, Wang et al. 2008) and pandemic (Bright et al. 2008, Haynes 2009, Kang et al. 2009, Mahmood et al. 2008, Prel et al. 2008) influenza viruses. Full-length M2 protein is also presented on influenza M1 VLPs (Song et al. 2011).

In the second stream, studies developed VLP-based influenza vaccines by displaying antigens from influenza viruses on VLPs from unrelated pathogens. Murine leukaemia virus (MLV) gag particles have been used to display genetically fused HA and NA of an influenza virus using mammalian (Szécsi et al. 2006) or insect cell (Haynes et al. 2009) expression systems. Similarly, gag VLPs have been used to display a headless HA protein of influenza (Steel et al. 2010) and Flock House virus (FHV) VLPs have been used to display Helix A from the HA2 influenza virus (Schneemann et al. 2012). In addition, studies have displayed peptide antigens from the M2 protein on VLPs from murine polyomavirus VP1 protein (Wibowo et al. 2013) and HBcAg (De Filette et al. 2006, Neiryneck et al. 1999) via genetic fusion. Antigens from the M2 protein have also been displayed on HBcAg (Fan et al. 2004), human papillomavirus L protein (Ionescu et al. 2006), and phage Q β -derived VLPs (Bessa et al. 2008) via conjugation.

2.4 The UQ Microbial Vaccine Platform (UQ-MVP)

2.4.1 Murine polyomavirus (MuPyV) VP1 protein

Murine polyomavirus (MuPyV) is an endogenous mouse virus from the class of polyomaviridae, and is closely related to simian virus 40 (SV40) (Stehle et al. 1996). It is a small non-enveloped DNA tumour virus, which recognises (α 2,3)-linked α -5-*N*-acetylneuraminic acid (sialic acid) on the surface of host cells. While all strains recognise straight-chain (α 2,3)-sialic acid, some are also able to recognise branched-chain receptor oligosaccharides, which are linked to a second (α 2,6)-sialic acid. The ability of the virus to recognise the branched-chain receptor is linked to a substitution of the residue at position

91 to glycine. The substitution of the residue to glycine has also been shown to reduce tumorigenicity of the virus (Stehle and Harrison 1996, Stehle and Harrison 1997).

The nucleocore in MuPyV consists of 5.3 kb of circular genomic DNA (Bouřa et al. 2005, Gillock et al. 1997). This genome encodes three T antigens (large, middle, and small tumour) and three capsid proteins, i.e. VP1, VP2, and VP3. VP1 is important for virus self-assembly, while VP2 and VP3 have been shown to be insignificant for the assembly process (Stehle and Harrison 1997). The virus particle is about 45 nm in diameter and the shell of the virus consists of 72 pentamers, which are arranged in a $T = 7d$ icosahedral lattice. The pentamers, also called capsomeres, are identical and composed of five copies of the major structural protein, VP1. Each VP1 monomer is about 42.5 kDa (Stehle et al. 1996, Stehle and Harrison 1997).

The crystal structure of VP1 from MuPyV has been resolved at 3.65 Å from purified polyoma strain P16 (Figure 2-12A; 1SID.pdb and UniProtKB P49302). This structure was resolved based on a previously described model of VP1 protein from SV40 (Stehle and Harrison 1996, Stehle et al. 1994). The structure of VP1 from MuPyV has also been resolved at 1.9 Å (Figure 2-12B; 1VPN.pdb) from bacterially expressed VP1 capsomeres lacking 31 amino acids in the N-terminus and 63 amino acids in the C-terminus (Stehle and Harrison 1997). Both crystals show that the structure of the VP1 protein is mainly β -sheets, forming a barrel of “jelly-roll” topology. The only difference between both crystal structures is the arrangement of the N- and C-termini, indicating the role of both termini during capsomere assembly. Similarities between both crystal structures show that in the absence of post-translational modifications, bacterially expressed VP1 protein can fold into its native structure.

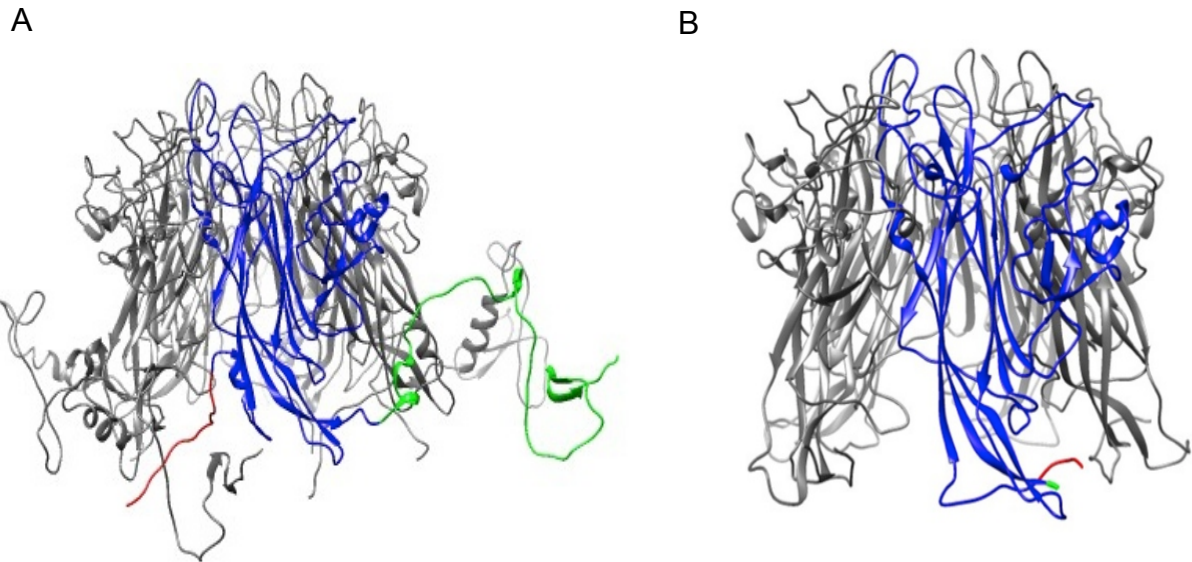


Figure 2-12. Crystal structures of murine polyomavirus VP1 capsomeres. (A) Crystal structure from purified virions (VP1 residues 1-384; 1SID.pdb). Red represents the N-terminal residues (residues 1-31). Green represents the C-terminal residues (residues 321-384); and (B) Crystal structure from recombinant VP1 capsomeres (VP1 residues 32-320; 1VPN.pdb). Red represents residues 32-36. Green represents residues 316-320. Figures were generated using UCSF Chimera (Pettersen et al. 2004).

Similar to SV40, the assembly process of VP1 capsomeres to form VLPs is proposed to involve the interchange of arms to make contact between one capsomere and a neighbouring capsomere. A C-terminal arm from one capsomere invades another (Figure 2-13A). The invading C-terminal affects the structure of the N-terminal on the targeted capsomere. The N-terminal arm forms a clamp to secure the invading C-terminal arm (Figure 2-13B). The invading C-terminal also dislodges another N-terminal arm on the target capsomere to form a clamp with a third capsomere. Afterwards, the exchanging C-terminal arms between two capsomeres are stabilised by additional clamp-clamp interactions between the N-termini of the two capsomeres (Figure 2-13C) (Stehle and Harrison 1997).

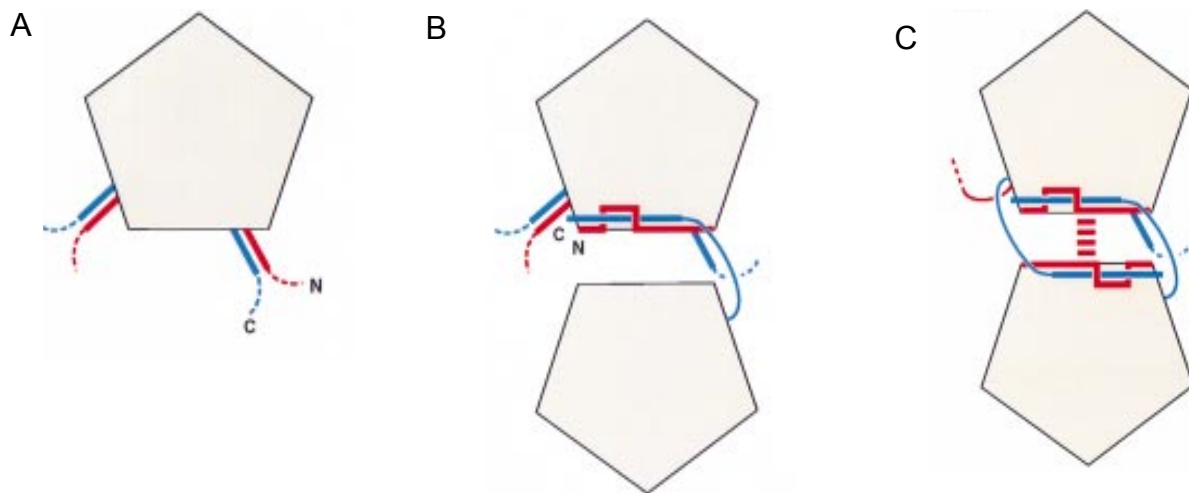


Figure 2-13. Assembly process of murine polyomavirus capsomeres to form VLPs. (A) A free-capsomere, (B) Invading of C-terminal arm (blue) from one capsomere to a neighbouring capsomere. The invading C-terminal arm is secured by the N-terminal (red) clamp, and (C) Formation of clamp-clamp interaction between N-terminal arms of two capsomeres. Figure taken from (Stehle and Harrison 1997).

2.4.2 Modular MuPyV VP1 as an antigen carrier

The structural study of SV40 using X-ray crystallography shows that pentamers are roughly cylindrical, about 80 Å in diameter and 70 Å in length. Each pentamer has a hollow conical interior with 50 Å in diameter at its base and 12 Å at its neck, which is formed by a loop named loop FG (Liddington et al. 1991). The outside of each pentamer is formed by three surface-exposed loops named loops BC2, DE, and HI. These loops interact closely with each other and each loop connects two strands of β -sheets. An additional loop, named loop EF, is on one side of the β -sheet, starting from the inside, facing the end of the β -sheet. Loops BC2, HI and EF are suitable for the insertion of foreign amino acid sequences. However, loop DE is involved in monomer-monomer interactions within the pentamer; thus, it is not suitable for the insertion of a foreign sequence. Due to the structural similarities between SV40 and MuPyV, these four surface-exposed loops are also found in MuPyV VP1 (Figure 2-14). However, in MuPyV VP1, these loops contain 18 additional amino acids (Stehle et al. 1996, Stehle et al. 1994).

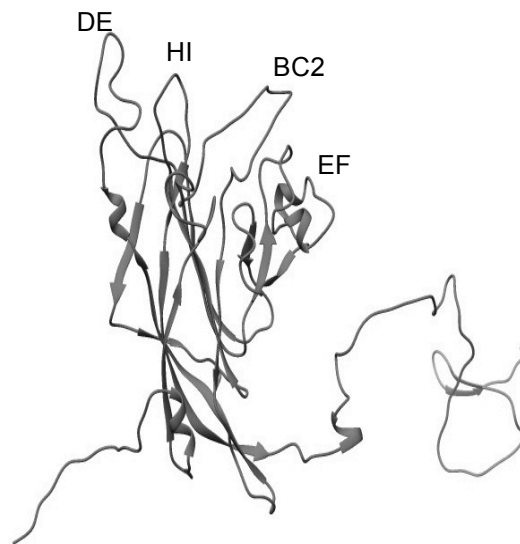


Figure 2-14. Surface-exposed loops in monomeric murine polyomavirus VP1 (1SID.pdb) (Stehle and Harrison 1996). The figure was generated using UCSF Chimera (Pettersen et al. 2004).

Based on the three-dimensional structure of the VP1 major capsid protein of SV40, Gedvilaite et al. (Gedvilaite et al. 2000) conducted a structural prediction in the hamster polyomavirus (HaPyV) VP1 protein (UniProtKB O73427) to determine potential sites for the genetic insertion of foreign antigens. The finding shows that in the HaPyV VP1 protein, four sites in the HaPyV VP1 protein can tolerate the genetic insertion of a foreign sequence. These sites are located in positions 80-89 (site 1), 221-224 (site 2), 243-247 (site 3), and 288-295 (site 4). Solvent accessibility (>30%) calculations indicated that all sites showed relatively high solvent accessibility, except site 2. Furthermore, analysis of atom mobility based on the value of the atomic temperature factor (B-value) given in the X-ray structure analysis showed that the mean B-value of the insertion sites in order were site 4 > site 1 > site 3 >> site 2. These analyses demonstrated that foreign antigens inserted in site 1, 3, and 4 should fold correctly and could therefore induce immune responses. These findings are supported by a study by Gedvilaite et al. (Gedvilaite et al. 2004). The study showed that modular HaPyV VP1 bearing foreign antigens at insertion site 1 and 4 had higher expression levels in yeast than those bearing foreign antigens at insertion sites 2 and 3. Modular HaPyV VP1 bearing foreign antigens in sites 1 and 4 were also able to form VLPs *in vivo*. However, VLPs of modular HapPyV VP1 bearing foreign antigens in sites 2 and 3 could not be obtained.

To be able to locate these insertion sites in MuPyV VP1, the amino acid sequence of MuPyV VP1 and HaPyV VP1 were aligned, as shown in Figure 2-15 (Gedvilaite et al.

2000). The alignment shows that sites 1, 2 and 4 are located in surface-exposed loops BC2, EF and HI, respectively. Site 3 is located in loop FG (Figure 2-16).

CLUSTAL O(1.2.1) multiple sequence alignment

```

O73427      MRKSMAPKRKSGASSRCANPCGKPCPKPANVPKLI MRGGVGVLDLVTGEDSITQIEAYLN
P49302      ----MAPKRKSGV-SKCTKCTKACPRPAPVPKLLIKGGMEVLDLVTGPDSVTEIEAFLN
              ***** . * : * . * * * * : * * * * : : * * : * * * * * * * * : * * * * * * * *

O73427      PRMGQNKPG--TGTDGQYYGFSQSIIKVNSSLTAEV KANQLPYYSMAKIQLPTLNEDLTC
P49302      PRMGQPPTPESLTEGGQYYGWSRGINLATS DTEDSPGNNTLPTWSMAKQLPMLNEDLTC
              *****          ***** : * : . * : : * * * . * * * : * * * * : * * * * * * * *

O73427      DTLQMW EAVSVKTEVVGSGLLNVHGYGSRSET-KDIGISKPV EGTTYHMFVAVGGEPLDL
P49302      DTLQMW EAVSVKTEVVGSGLLDVHGFNKPTDTVNTKGI STPVEG SQYHVFVAVGGEPLDL
              ***** * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *

O73427      QGLVQNYNANYE-AAIVSIKTVTGKAMTSTNQVLDPTAKAKLDK DGRYP I EIWGPDPSPKN
P49302      QGLVTDARTKYKEEGVVTIKTITKKDMVNKDQVLNPI SKAKLDK DGMYPVEIWHDPDPKN
              * * * * : . : * * : . : * * * * : * * * . . . : * * * * : * * * * * * * * * * * * * * * * *

O73427      ENSRYYG NFTGGTGTPPVMQFTNTLTTVLLDENG VGPLCKGDGLYLSAADVMGWYIEYNS
P49302      ENRYFGNYTGGTTTPPVLQFTNTLTTVLLDENG VGPLCKGEGLYLSCVDIMGWRVTRNY
              * * : * * : * * : * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *

O73427      AGWHWRGLPRYFNVTLRKRWVKNPYPVTSLLASLYNNMLPTIEGQ PMEGEAAQVEEVRIY
P49302      DVHHRGLPRYFKITLRKRWVKNPYPMASLISSLFNNMLPQVQGPMEGENTQVEEV RVY
              * * * * * * * * : * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *

O73427      EGTEAVPGDPDVNR FIDKYGQOHTKPPAKPAN
P49302      DGTEVPVPGDPMTRYVDRFGKTKTVFPGN---
              : * * * * * * * * : . * : * : * : * : * * * * * * * * : * * * * * * * *

```

Figure 2-15. Amino acid sequence alignment of VP1 from hamster polyomavirus (UniProtKB identifier O73427) and VP1 from murine polyomavirus (UniProtKB identifier P49302, 1SID.pdb). The alignment is adopted from (Gedvilaite et al. 2000) and performed using Clustal Omega (<http://www.ebi.ac.uk/Tools/msa/clustalo/>).

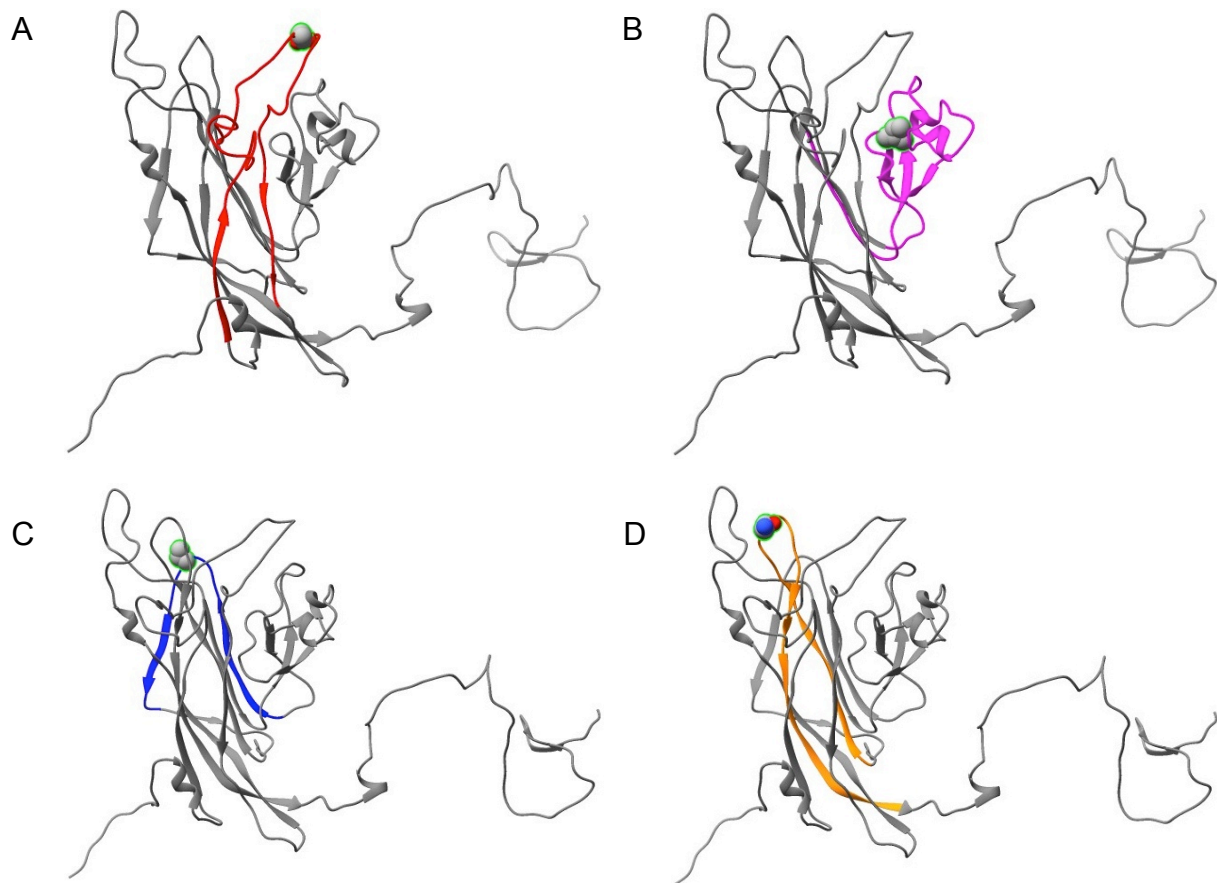


Figure 2-16. Insertion sites (spheres) in monomeric murine polyomavirus VP1 protein (1SID.pdb). The insertion sites are located in surface-exposed loops. (A) Site 1 in loop BC, (B) Site 2 in loop EF, (C) Site 3 in loop FG, and (D) Site 4 in loop HI. The figure was generated using UCSF Chimera (Pettersen et al. 2004).

Studies have exploited the use of the insertion sites in MuPyV VP1 to introduce antigens from various pathogens. Based on the study by Gedvilaite et al. (Gedvilaite et al. 2004) explained in the preceding paragraph, site 1 in loop BC2 and site 4 in loop HI in MuPyV VP1 are mostly exploited (Gleiter et al. 1999). Similarly, the insertion site has also been exploited to display the sequence of the immunoglobulin-binding domain, the protein Z (Gleiter and Lilie 2001). In both studies, the inserted antigens were shown to be able to retain their biological functions. Furthermore, in the study by Rivera-Hernandez et al. (Middelberg et al. 2011, Rivera-Hernandez et al. 2013), MuPyV VP1 was utilised as a carrier for peptide antigen from p145 protein from Group A *streptococcus* (GAS) bacteria. The antigen was inserted in insertion site 4 of the VP1 protein. In the study by Wibowo et al. (Wibowo et al. 2013), the same insertion site was utilised to display peptide antigen from M2e protein of influenza protein.

2.4.3 Production of MuPyV VP1 VLPs using *E. coli* expression system

The UQ MVP (Lua and Middelberg 2007) is developed based on production of MuPyV VP1 VLP using an *E. coli* expression system. In the UQ MPV, MuPyV VP1 is expressed as a Glutathione S-transferase (GST)-tagged protein in *E. coli* (Chuan et al. 2008). Highly soluble GST-tagged protein is then purified using GST-affinity chromatography (Lipin et al. 2008). Following tag removal, capsomeres are separated from aggregates and cleaved GST-tag using size-exclusion chromatography (SEC). Purified capsomeres spontaneously self-assemble *in vitro* under the following conditions: (i) the presence of Ca^{2+} ions, (ii) increased ionic strength, and (iii) the absence of reduction agents (Chuan et al. 2008, Chuan et al. 2010, Salunke et al. 1989). The formed VLPs are then analysed using asymmetric flow-field flow fractionation (AF4) coupled with multi-angle light scattering (MALS), and transmission electron microscopy (TEM) (Chuan et al. 2008).

Unlike VP1 protein expressed in eukaryotic systems, bacterially expressed VP1 does not undergo post-translational modification, such as phosphorylation and acetylation. However, VLPs from bacterially expressed MuPyV VP1 mimic the native virus (Salunke et al. 1986). More importantly, the use of the *E. coli* expression system allows for the rapid production of VLPs with a high yield. The production can be scaled-up to produce up to 1 g/L VP1 (Liew et al. 2010). Nevertheless, bacterially expressed MuPyV VP1 has two major contaminants, GroEL and DnaK (Fan and Middelberg 2010). These contaminants are highly conserved *E. coli* molecular chaperones and classified as heat shock proteins 60 and 70, respectively.

2.4.4 Molecular chaperones as contaminants on MuPyV VP1 produced in *E. coli*

Chaperones recognise and selectively bind to non-native proteins to form stable complexes via exposed hydrophobic residues. The complexes can be dissociated via the binding and hydrolysis of ATP. Protein folding *in vitro* is usually performed at low protein concentrations under non-physiological conditions and long incubation times. However, *in vivo*, cells are mostly grown at homeothermic temperature (e.g. 37°C). This condition results in stronger hydrophobic effects, and consequently the higher propensity of protein aggregation and misfolding. Therefore, chaperones are required for protein *in vivo* folding (i) to prevent the aggregation and misfolding of newly synthesised protein, (ii) to prevent non-productive interactions with other cells, (iii) to assist in the assembly of larger proteins and multiprotein complexes, and (iv) to prevent the unfolding of folded proteins due to exposure to stresses. The major class of chaperones involved in *in vivo* protein folding are 40-kDa heat shock protein (Hsp40, the DnaJ family); 60-kDa heat shock protein (Hsp60),

which include GroEL and the T-complex polypeptide 1 ring-complexes; 70-kDa heat shock protein (Hsp70); and 90-kDa heat shock protein (Hsp90).

2.4.4.1 GroEL

GroEL is a homotetradecamer protein and is composed of two back-to-back rings. Each ring comprises seven monomeric subunits of 57 kDa. The crystal structure of free GroEL shows a cylindrical complex that is 145 Å in height and 135 Å in diameter. The central channel in each terminus is 45 Å in diameter (Braig et al. 1994, Braig et al. 1993). GroEL is found in all biological compartments except the ER of prokaryotes, chloroplasts, and mitochondria. It functions to prevent the aggregation of partially folded intermediates as well as to facilitate the folding and assembly of the intermediates in a favourable environment. It also allows misfolded proteins to unfold and refold (Fink 1999).

The GroEL subunit is composed of three domains: the apical domain, the equatorial domain, and the intermediate domain. The apical domain forms the end part of GroEL cylinder and is composed of an orthogonal β -sheet structure flanked by α -helices. The domain is suggested to have intrinsic flexibility, which is necessary to accommodate the binding of various polypeptides. The equatorial domain is the structural foundation of the cylinder structure of GroEL, and allows for the major contacts between subunits within and between rings. Long and parallel α -helices compose this domain. The apical domain is covalently connected to the equatorial domain via the intermediate domain. Mutations of the intermediate domain have been shown to affect measurable functions of GroEL (Fenton and Horwich 1997, Fink 1999, Xu et al. 1997).

The activity of GroEL is assisted by the co-chaperone GroES (cpn10). Similar to GroEL, GroES is a ring-like protein, composed of seven subunits that are 10 kDa in size. In the presence of nucleotides, GroES binds to one or both ends of the GroEL cylinder, resulting in the upward and outward opening of the GroEL apical domains. Furthermore, the activity of GroEL is also regulated by the binding and hydrolysis of ATP (Chandrasekhar et al. 1986, Chen et al. 1994, Fenton and Horwich 1997, Fink 1999, Langer et al. 1992, Saibil et al. 1991).

The reaction pathway of GroEL-GroES in mediating folding consists of four steps. First, GroEL binds to a non-native polypeptide while GroES caps the other end of the cylinder (*trans* ternary complex). The binding occurs in the presence of physiological levels of ATP/ADP. Second, the *trans* complex then becomes a *cis* ternary complex, in which GroES binds to GroEL on the same ring as polypeptides. The formation of the *cis* ternary complex is achieved (i) by releasing GroES from the *trans* ring, followed by binding in *cis*

ring; or (ii) by binding of the second GroES in the *cis* ring. Third, in the presence of ATP, the bound polypeptide is released from its binding site in GroEL, allowing for folding inside the sequestered space of GroES. Fourth, GroES is released from GroEL, and the polypeptide departs in three possible structures: native conformation, native intermediate, and non-native intermediate. The latter can rebind to GroEL for another folding attempt (Fenton and Horwich 1997).

2.4.4.2 DnaK

DnaK is the major bacterial Hsp70 and expressed abundantly in *E. coli*. Together with the co-chaperone DnaJ and the regulator GrpE, DnaK functions in *de novo* protein folding (Calloni et al. 2012). Similar to GroEL, DnaK prevents the aggregation of misfolded proteins and promotes refolding. Due to its function in repairing denatured proteins, DnaK is viable at high temperatures (>37°C). DnaK only binds to unfolded proteins, and relatively unfolded intermediates (Mayer et al. 2000). It does not bind to folded and strongly native-like intermediates. DnaK binds to unfolded polypeptides via short, hydrophobic regions in the centre of the polypeptides. The hydrophobic region preferably consists of four to five hydrophobic residues, particularly enriched in Leu, Ile, Val, Phe, and Tyr (Rüdiger et al. 1997). The binding is also favoured by a strong ionic interaction between DnaK and negative charged residues in the region (Fink 1999).

DnaK consists of two major functional domains: (i) the ATPase domain of 45 kDa, and (ii) the COOH-terminal domain of 25 kDa. The ATPase domain is highly conserved and binds to ADP and ATP very tightly in the presence of Mg²⁺ and K⁺. Hydrolysis of the ATP also takes place in the domain. Furthermore, the ATPase domain comprises of four sub-domains. These sub-domains form two lobes with a deep gap where MgATP and MgADP bind. Meanwhile, the COOH-domain of DnaK binds to a polypeptide. The domain consists of a β -sandwich subdomain and four α -helical segments. The β -sandwich comprises two groups of four antiparallel β -strands, which are connected by upward-protruding loops. These loops form a channel for the binding of polypeptide in an extended conformation. The complex is then stabilised by one of the α -helical segments without any direct contact. The other three α -helical segments build a hydrophobic helical core, forming a lid-like structure, presumably to maintain the bound polypeptide (Harrison et al. 1997, Zhu et al. 1996).

The reaction cycle of DnaK comprises three major pathways. First, ATP binds to DnaK, inducing a conformational change of the COOH-terminal domain in DnaK to a low-affinity state. The conformational change is hypothesised to involve the elevation of the

hydrophobic helical core. Due to the low affinity of DnaK, the DnaK-ATP complex binds and releases polypeptides rapidly. Second, in the presence of DnaJ, ATP is hydrolysed, resulting in the formation of a stable DnaK-ADP complex. Polypeptides bind to the DnaK-ADP complex, which has a high affinity and lives for a relatively long period. Third, GrpE promotes the dissociation of ADP from DnaK, and ATP consequently binds to DnaK. The cycle returns to (i) the first pathway, leading to the release of polypeptide, and the binding of new polypeptides, or (ii) to the second pathway, where ATP is hydrolysed, yielding the DnaK-ADP complex (Calloni et al. 2012, Fink 1999, Mayer et al. 2000).

2.4.4.3 Existing methods for purification of DnaK and GroEL

Studies have shown that ATP-affinity chromatography effectively purified both chaperones (Khandekar et al. 1993, Peng et al. 1997). Treatment with ATP has also been shown to separate both chaperones from bound proteins (Rial and Ceccarelli 2002, Schön and Schumann 1995). In addition, both chaperones can also be purified based on ion exchange chromatography (IEX) utilising Q-sepharose (Fan and Middelberg 2010, Henot 2010, Henot et al. 2008, Schön and Schumann 1995, Zylicz and Georgopoulos 1984), DEAE (Kamireddi et al. 1997, Zylicz and Georgopoulos 1984), or P-11 cellulose (Zylicz and Georgopoulos 1984) columns. Size-exclusion chromatography (SEC) using a Sephacryl column can also be utilised (Henot 2010, Kamireddi et al. 1997, Schön and Schumann 1995) as a polishing step. The purification of GroEL using hydrophobic interaction chromatography (HIC) (Schoel and Kaufmann 1991) and DnaK using a hydroxyapatite type II column (Henot 2010, Zylicz and Georgopoulos 1984) has also been shown.

2.5 Concluding statement

In this chapter, relevant literature on the influenza A viruses and influenza vaccines are gathered. Vaccination has been shown to be the most cost effective way to prevent the spread of various infectious diseases, including influenza A. However, influenza A viruses undergo continuous antigen variability through antigenic drift and antigenic shift mechanisms. Vaccine production using conventional egg-based and cell culture-based technologies has not provided a solution for the variability of influenza A surface antigens. A promising alternative to the conventional vaccine technologies is virus-like particles (VLPs) based on a microbial platform. The literature gathered in this chapter show extensive studies in VLPs from the murine polyomavirus (MuPyV) VP1 protein. An attractive feature of MuPyV VP1 VLPs is that they can be used as a carrier for foreign

peptide epitopes. Rapid and high-yield production of the VLPs via a bacterial expression system has also been established.

This chapter also showed the importance of the receptor-binding site within the globular domain of HA during the first step of viral infection. The receptor-binding site is bordered by three secondary elements that are hypervariable. One of the elements is named helix 190. It has a helical secondary structure and studies show that it contains an antigenic site in both H1 and H3 subtypes. Therefore, helix 190 is an interesting candidate for peptide antigens to be presented on MuPyV VP1 VLPs.

This chapter has also shown an increasing need to develop vaccines based on the structure of antigens, especially for pathogens undergoing constant changes on the surface antigens, such as influenza virus. Various approaches used to preserve the native conformation of peptide antigens presented in unrelated protein carriers are described. Information from these literatures can be used to present helix 190 in MuPyV VP1 VLPs and to retain the native conformation of helix 190 when it is presented on the VLPs.

Chapter 3. Sensitivity of immune response quality to influenza helix 190 antigen structure displayed on a modular virus-like particle

The entire Chapter 3 consists of the journal article published as:

Anggraeni, MR, Connors, N, Wu, Y, Chuan, YP, Lua, LHL, Middelberg, APJ. 2013. Sensitivity of immune response quality to influenza helix 190 antigen structure displayed on a modular virus-like particle. *Vaccine* 31(40):4428-4435.

The following modifications were made to the article:

1. Page numbers of the original article were crossed out.
2. Page numbers consistent with those on the remainder of the thesis pages were inserted.
3. Published supplementary information was attached as Appendix A. Font and reference format were adjusted.



Sensitivity of immune response quality to influenza helix 190 antigen structure displayed on a modular virus-like particle



Melisa R. Anggraeni^a, Natalie K. Connors^a, Yang Wu^a, Yap P. Chuan^a, Linda H.L. Lua^b, Anton P.J. Middelberg^{a,*}

^a The University of Queensland, Australian Institute for Bioengineering and Nanotechnology, Centre for Biomolecular Engineering, St. Lucia, QLD 4072, Australia

^b The University of Queensland, Protein Expression Facility, St. Lucia, QLD 4072, Australia

ARTICLE INFO

Article history:

Received 24 January 2013
Received in revised form 24 May 2013
Accepted 25 June 2013
Available online 8 July 2013

Keywords:

Virus-like particle
VLP
Influenza
Receptor
Biomolecular engineering
Modular

ABSTRACT

Biomolecular engineering enables synthesis of improved proteins through synergistic fusion of modules from unrelated biomolecules. Modularization of peptide antigen from an unrelated pathogen for presentation on a modular virus-like particle (VLP) represents a new and promising approach to synthesize safe and efficacious vaccines. Addressing a key knowledge gap in modular VLP engineering, this study investigates the underlying fundamentals affecting the ability of induced antibodies to recognize the native pathogen. Specifically, this quality of immune response is correlated to the peptide antigen module structure. We modularized a helical peptide antigen element, helix 190 (H190) from the influenza hemagglutinin (HA) receptor binding region, for presentation on murine polyomavirus VLP, using two strategies aimed to promote H190 helicity on the VLP. In the first strategy, H190 was flanked by GCN4 structure-promoting elements within the antigen module; in the second, dual H190 copies were arrayed as tandem repeats in the module. Molecular dynamics simulation predicted that tandem repeat arraying would minimize secondary structural deviation of modularized H190 from its native conformation. *In vivo* testing supported this finding, showing that although both modularization strategies conferred high H190-specific immunogenicity, tandem repeat arraying of H190 led to a strikingly higher immune response quality, as measured by ability to generate antibodies recognizing a recombinant HA domain and split influenza virion. These findings provide new insights into the rational engineering of VLP vaccines, and could ultimately enable safe and efficacious vaccine design as an alternative to conventional approaches necessitating pathogen cultivation.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

Synthetic biomolecular engineering has enhanced understanding of biological processes and design of novel proteins [1,2]. Fusion of different protein domains to realize a synergic combination of modules [3] has been used extensively to create hybrid enzymes [4] and humanized antibodies [5]. This approach has yet to impact vaccines significantly, but rational vaccine design based on presentation of protective antigen modules [6,7] is emerging.

VLPs are protein assemblies that mimic viral structure [8–14]; modular VLPs combine a VLP with an antigen module from an unrelated pathogen [9,15–17]. Two types of modules are commonly used: (i) short peptides expected to have minimal structure [18–21]; and (ii) larger domains with expected tertiary structures [22–25]. There remains a significant knowledge gap about the

display of peptide antigen with specific secondary structure *e.g.* α -helix, on an unrelated VLP. Here we address this knowledge deficit using influenza as a model pathogen.

Hemagglutinin (HA) is an influenza surface glycoprotein comprising globular HA1 and stalk HA2 domain. HA mediates viral entry in two steps: (i) viral attachment to the sialic acid receptor; and (ii) fusion of the viral and endosomal membrane [26,27]. Antibodies that bind within the receptor binding region on HA1 block viral attachment and are neutralizing [28,29]. The receptor binding site within HA1 determines viral specificity [30] and is bordered by hypervariable regions including helix 190 [31]. Helix 190 (H190; Fig. 1A) possesses a defined helicity and contains the immunodominant antigenic site Sb [32,33]; monoclonal antibodies against this site have been shown to be protective in mice [34]. The immunogenicity, function and defined secondary structure of H190 suggest it is an interesting candidate for structural display on an unrelated VLP.

We chose to explore murine polyomavirus (MuPyV) VLP for modularization with H190. High-efficiency microbial expression

* Corresponding author. Tel.: +61 7 3346 4189; fax: +61 7 3346 4197.
E-mail address: a.middelberg@uq.edu.au (A.P.J. Middelberg).

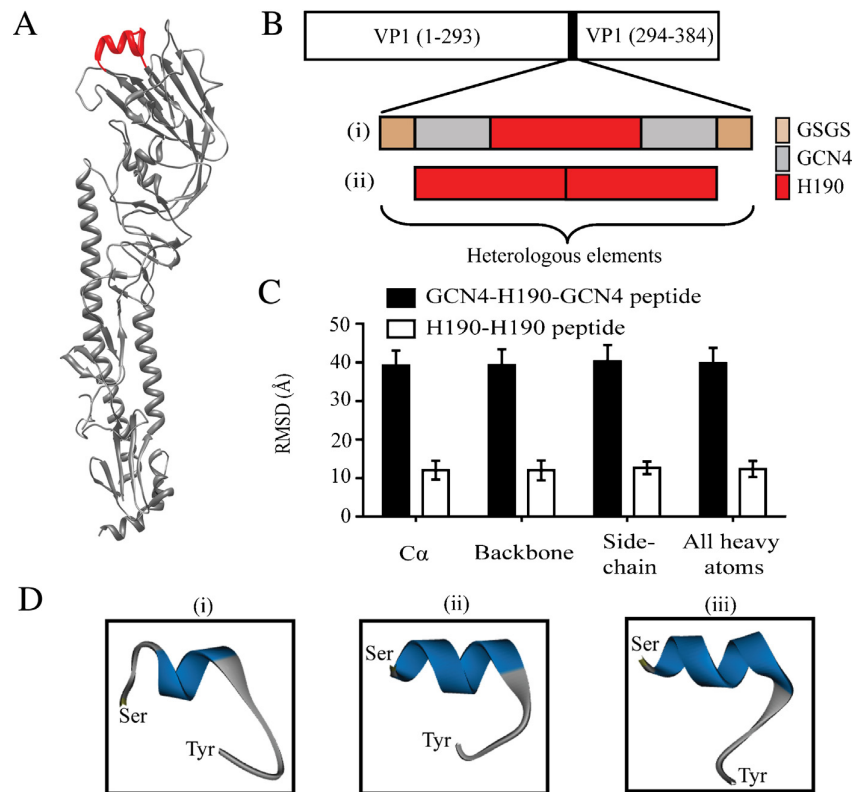


Fig. 1. (A) Crystal structure of H1N1 HA with H190 highlighted in red. Data for A/California/04/2009 HA (3AL4.pdb), bearing one amino acid difference from A/California/07/2009 HA, is shown here. The picture was created using UCSF Chimera 1.5.2 [65]. (B) modular MuPyV VP1 designs showing the heterologous elements in (i) VP1-GCN4-H190-GCN4; and (ii) VP1-H190-H190. Numbers refer to residues of VP1. (C) root-mean-square-deviation (RMSD) of C α , backbone, side chain, and all heavy atoms of simulated GCN4-H190-GCN4 and H190-H190 peptides with reference to H190 on HA (3MLH.pdb), and (D) final conformation of H190 element in simulated peptides (i) GCN4-H190-GCN4; and (ii) H190-H190, in comparison to (iii) native H190 structure in HA. α -Helical structures are highlighted in blue. The image was created using Accelrys Discovery Studio[®] 3.0. Positions of Ser1 and Tyr16 on the peptides are as indicated. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

and purification processes for modular MuPyV VLPs have been reported [9]. We explore whether the helicity of different H190 antigen module designs affect the quality, *i.e.* native protein recognition, of the antibodies raised by vaccination with the modular VLPs.

Two display strategies were compared: (i) display of H190 using flanking helix promoting elements; and (ii) tandem repeat arraying of H190. The first strategy explores GCN4 as flanking structure-promoting elements within the antigen module. GCN4 is well characterized and has been used as an α -helical scaffold [35,36] or as a helix promoter [37,38]. We previously reported the inclusion of GCN4 in an antigenic module designed to display Group A *Streptococcal* antigen on a VLP [9]. However, these studies have not established whether the GCN4 elements have any positive effect on the immune response quality. The second strategy explores the use of tandem repeats, which are ubiquitous in proteins, and form domains with extensive secondary structures including helices [39]. The use of this strategy has been restricted to present conformational proline-rich peptide to effect the necessary structure [40,41], or to simply increase immunogenicity [42–44].

Modularized VLPs were assessed computationally and *in vivo*. Molecular dynamics (MD) simulation predicted that modularization using tandem repeats would provide a closer structural match of modularized H190 and the equivalent native region in HA. *In vivo* testing confirmed this strategy was superior as it induced antibodies able to bind recombinant HA1 protein and split influenza virion. These findings illustrate, for the first time, the importance of VLP

antigen module design on the quality of an immune response, providing new insights into the design rules for synthetic biomolecular engineering of VLP vaccines.

2. Materials and methods

2.1. Simulation of peptides GCN4-H190-GCN4 and H190-H190

Peptides GCN4-H190-GCN4 and H190-H190 were simulated using GROMACS version 3.3.3 in water containing 137 mM NaCl for 20 ns, 3 times for each peptide, as described in Supplementary information.

2.2. Cloning, purification and assembly of modular VLPs

E. coli codon-optimized DNA oligomers corresponding to the antigen modules containing heterologous elements (Fig. 1B) were cloned into pGEX-VP1 vectors as described previously [9]. Modular proteins VP1-GCN4-H190-GCN4 and VP1-H190-H190 were recombinantly produced and prepared as described [45,46]. After endotoxin removal to below 5 EU/ml, purified modular capsomeres were assembled *in vitro* into modular VLPs and dialyzed against PBS [9]. The endotoxin levels of the capsomeres were below the suggested limit for a recombinant protein based vaccine, which is 20 EU/ml [47].

2.3. Characterization and visualization of modular VLPs

VLP size distribution was characterized using AF4-MALS and TEM [46,48].

2.4. Mice immunization

Female Balb/c mice aged 6–8 weeks, 8 mice/group, were purchased from Monash Animal Research Platform (Monash University, Australia). Ethical clearance no. AIBN/273/10/CBME/AIBN (NF) was granted from the University of Queensland Animal Ethics Committee. Mice were injected subcutaneously at the tail base with 50 µg of VP1-GCN4-H190-GCN4 or VP1-H190-H190 VLPs in a volume of 50 µl on days 0, 21, and 42. Negative control groups were immunized with an equal dose of wild-type (wt) VLP or 50 µl PBS. Mice were bled on tail on days 0, 14, 35, and 56. Sera from day 56 were used for immunoassays.

2.5. Dot blot

Antigens (3 µg) were applied onto membrane and were probed with sera (1/200 dilution) for 1 h. After washing, horseradish peroxidase (HRP)-conjugated anti-mouse IgGs (Sigma–Aldrich, St. Louis, USA) was added. After 1-h incubation, HRP activities were detected via luminescence intensity.

2.6. ELISA

For indirect ELISA, antigens were immobilized and probed with sera for 1.5 h. After washing, HRP-conjugated anti-mouse IgGs was added. After 1-h incubation, HRP activities were detected via luminescence intensity. For competitive ELISA, an additional step was included in which sera were pre-mixed with various concentrations of inhibitor prior to incubation on plate.

2.7. Statistical analysis

Statistical analysis was performed on log-transformed data using GraphPad Prism 5.03 (GraphPad, Inc.). Comparison of antibody titer for multigroup was conducted using one-way ANOVA with Tukey's *post hoc* test. Total amount of HA in Fluvax was calculated based on product information for Fluvax season 2011 which specifies that Fluvax contains total HA of 45 µg in 500 µl. The concentration of recombinant HA1 was determined using absorbance at 280 nm with extinction coefficient (ϵ 1% in water) of 1.526 L/g cm.

Additional details are provided in the Supplementary information.

3. Results

3.1. Modularization of VLP and simulation

MuPyV VLP assembled from the major structural protein VP1 [9] was modularized with H190 (STSDAQSLYQNADAY) from H1N1 A/California/07/2009 influenza (Fig. 1A). Two modular MuPyV VP1 designs shown in Fig. 1B were investigated. In VP1-GCN4-H190-GCN4, the module comprised an H190 element flanked immediately by GCN4 helix-promoting elements (VKQLEDKV), with GSGS spacer elements connected to the VP1 backbone (Fig. 1Bi). In VP1-H190-H190, two H190 elements were simply arrayed as tandem repeats (Fig. 1Bii).

Molecular dynamics (MD) was performed to predict the structure of the H190 element in these designed modules. Root-mean-square-deviation (RMSD) was significantly higher in the GCN4-H190-GCN4 module than in the H190-H190 module (Fig. 1C). Final conformation of simulated peptide (Fig. 1D)

showed unfolding of H190 element in GCN4-H190-GCN4 after 20 ns, whereas a copy of H190 in H190-H190 retained its helical propensity. These data suggest that the H190 elements are likely to have different structures in the two antigen modules when presented on a VLP, though actual module structure may be further modified by the VLP junction regions. Nevertheless, the simulation results suggest that a tandem repeat arraying strategy could present the H190 in a structure that might raise an immune response of higher quality. This suggestion was investigated experimentally.

3.2. Formation of modular VLPs

Fig. 2 shows analyses of the assembled VLPs. The AF4 peak at approximately 20 min corresponded to VLPs [48] with averaged root-mean-square (rms) radius of 20 nm. The polydispersity index was 1.00 for all VLPs, indicating that the particles were monodisperse. TEM visualization showed that the modular VLPs were morphologically indistinguishable from the wild-type (wt) VLP.

3.3. H190-specific immunogenicity

To determine the presence of H190 sequence-specific IgGs raised by the modular VLPs, antisera were analyzed against peptide H190 (Peptide2.0, Chantilly, VA, USA) in a dot blot assay (Fig. 3A). Fig. 3B shows that immunization with either modular VLP induced a high IgG titer ($>10^4$) against H190 peptide. The IgG titers were similar for both modular designs. This result confirms that both modular VLPs were comparably effective at inducing a high titer of sequence-specific IgGs against the H190 element.

To determine the immunogenicity of the GCN4 element, a competitive ELISA was performed (Fig. 3C and D). Fig. 3C shows that peptide GCN4-H190-GCN4 (Peptide2.0, Chantilly, VA, USA) at a concentration of 239 µM completely abrogated binding of VP1-GCN4-H190-GCN4 serum IgGs to immobilized GCN4-H190-GCN4 peptide target, while peptide H190 at the same concentration only partially inhibited binding. In contrast, peptides GCN4-H190-GCN4 and H190 were comparably effective at inhibiting binding of VP1-H190-H190 VLP serum IgGs against the immobilized peptide GCN4-H190-GCN4, with complete inhibition achieved at 239 µM of competing peptide. These results confirmed the immunogenicity of the GCN4 element.

3.4. Specificity of modular VLP antisera to recombinant HA1 protein

The modular VLP antisera were tested against recombinant H1N1 A/California/07/2009 HA1. Fig. 4A shows that HA1-specific antibodies were present in the VP1-GCN4-H190-GCN4 VLP and VP1-H190-H190 VLP antisera. The ELISA result (Fig. 4B) showed that immunization with VP1-GCN4-H190-GCN4 or VP1-H190-H190 VLPs induced a high HA1-specific IgG titer against HA1, achieving greater than 10^4 for VP1-H190-H190. Importantly, the result also showed that VP1-H190-H190 induced a HA1-specific IgG titer significantly higher than the VP1-GCN4-H190-GCN4 VLP. Fig. 4C shows that IgGs binding to the immobilized HA1 protein could be completely inhibited by competing peptide H190. The complete inhibition concentration was 0.2 µM and 2 µM for VP1-GCN4-H190-GCN4 and VP1-H190-H190 VLPs, respectively.

3.5. Glycosylation effects of recombinant H1N1 HA1

HA1 used in this study was purchased from the Protein Expression Facility, The University of Queensland (Australia). It was produced in insect cells and was likely glycosylated [49]. SEC-MALS

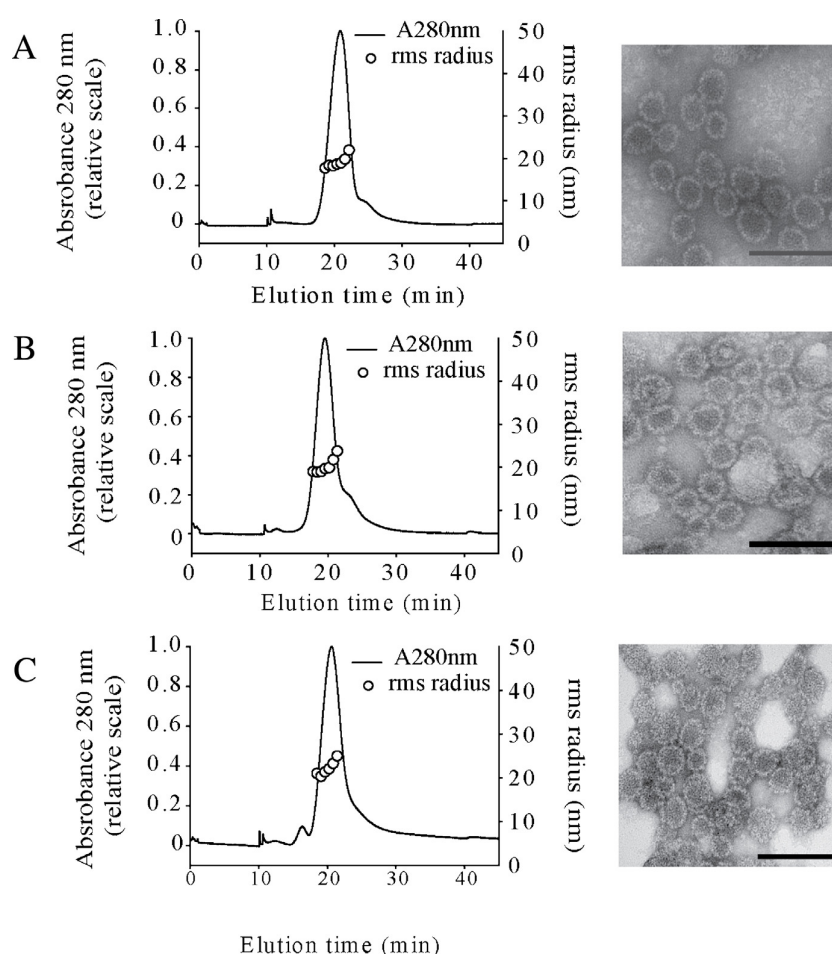


Fig. 2. Characterization of modular VLPs using AF4 (left panels) and TEM (right panels). (A) Wild-type VLP; (B) VP1-GCN4-H190-GCN4 VLP; and (C) VP1-H190-H190 VLP. Scale bars represent 100 nm.

analysis confirmed HA1 was predominantly monomeric (49.3 kDa; Fig. S1A) but larger than expected (35.6 kDa), in agreement with SDS-PAGE (Fig. S1B). Glycoprotein staining confirmed N-linked glycosylation (Fig. S1C lane 1).

Dot blot analysis indicated that modular VLP antisera reacted with both PNGase A- and PNGase F-treated HA1 (Fig. 5A). Indirect ELISA performed to test the modular VLP serum IgGs against PNGase F treated-HA1 protein (Fig. 5B) showed that both VP1-GCN4-H190-GCN4 VLP and VP1-H190-H190 VLP antisera contained IgGs that recognized the treated-HA1, at a titer comparable to that for untreated HA1. This result indicates that the HA1 titer observed was independent of glycosylation. Competitive ELISA (Fig. 5C) showed that IgG binding to PNGase F-treated HA1 could be completely inhibited by competing peptide H190. Together, these results indicated that binding of the modular VLP antisera to H190 on HA1 protein was not affected by glycosylation on HA1, even though immunization was performed with non-glycosylated VLPs made using bacteria.

3.6. Specificity of modular VLP antisera

Modular VLP antisera were tested against the trivalent 2011 influenza vaccine Fluvax[®] (containing split H1N1 A/California/07/2009 influenza) by dot blot (Fig. 6A) and indirect ELISA (Fig. 6B). Both analyses showed that only VP1-H190-H190 VLP induced IgGs which recognized the split virion. For

VP1-H190-H190 VLP antisera, the amount of IgGs that bound to the split virion was lower than to the HA1 protein at an equal HA molar concentration. However, full-length HA protein of the split influenza virion has complex and poorly understood self-association characteristics that can lead to the formation of aggregate-like structures that might alter both immunogenicity and the availability of H190 region, reducing the effective concentration of H190.

4. Discussion

This study explores an important question in the synthetic biomolecular engineering of virus-like particle (VLP) vaccines – does antigen module design alter the quality of the resulting immune response, specifically the ability of serum antibodies to recognize the native structure on the target pathogen? It has been shown that many neutralizing antibodies isolated from natural infections are raised against conformational epitopes [50–52]. However, free peptide antigens comprising the same residues as these epitopes in aqueous solutions are generally non-structural [53].

Two modularization strategies (Fig. 1B) were investigated, using: (i) flanking helix promoting elements; and (ii) tandem repeat arraying, to drive helical propensity of the antigen element toward that of the native structure. VP1-GCN4-H190-GCN4 and VP1-H190-H190 modular VLPs were created based on (i) and

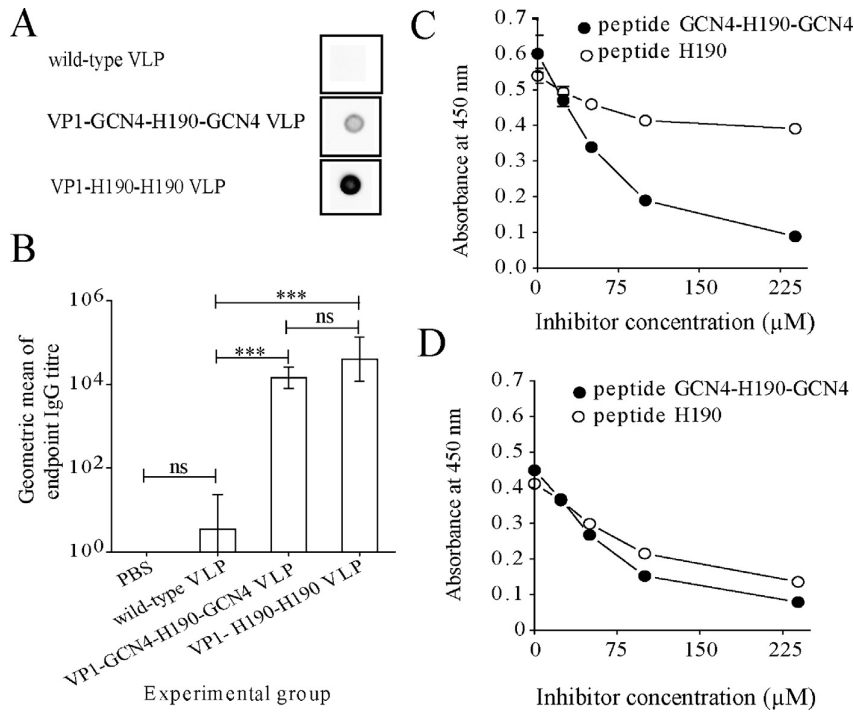


Fig. 3. Dot blot, indirect ELISA and competitive ELISA results showing immunogenicity of H190 in modular VLPs. (A) Dot blot analysis showing serum reactivity to peptide H190, (B) peptide H190 sequence-specific IgG titers induced by modular VLPs, evaluated using indirect ELISA with immobilized peptide H190. Bars represent geometric mean of IgG titer from 8 mice. Error bars represent 95% confidence intervals. ns, not significant ($P > 0.05$); ***, extremely significant ($P < 0.001$). (C and D) competitive inhibition of antiserum (1/1000 dilution) binding to immobilized peptide GCN4-H190-GCN4 by peptide GCN4-H190-GCN4 (●) or peptide H190 (○) in solution. (C) VP1-GCN4-H190-GCN4 VLP antiserum; and (D) VP1-H190-H190 VLP antiserum. Arithmetic means and standard deviations, from triplicates, are presented.

(ii), respectively. Computational prediction provided evidence that helical propensity of a helical antigen module on an unrelated VLP could be manipulated by the antigen module display strategy. More importantly, *in vivo* data showed that the display strategy could significantly affect the quality of the immune response.

The use of peptides from a larger protein domain for structural prediction is a widely used strategy [41]. Particularly for helix structure, it has been shown that helical peptides adopt secondary structure that is preserved when the same sequence is present in a protein. Helix localization occurs due to encoded signals in

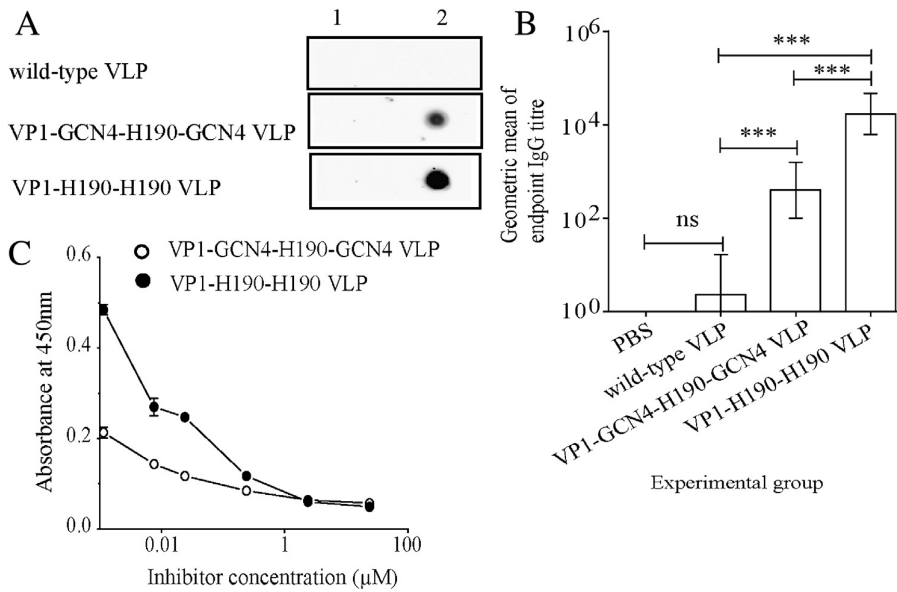


Fig. 4. Dot blot, indirect ELISA and competitive ELISA results showing reactivity of modular VLP antisera against recombinant HA1 protein. (A) Dot blot analysis showing serum reactivity to (1) HA1 storage buffer; and (2) HA1, (B) HA1-specific IgG titer induced by modular VLPs, evaluated using indirect ELISA. Bars represent geometric mean of IgG titer from 8 mice. Error bars represent 95% confidence intervals. ns, not significant ($P > 0.05$); ***, extremely significant ($P < 0.001$) and (C) competitive inhibition of antiserum (1/100 dilution) binding to immobilized HA1 by peptide H190 in solution. Arithmetic means and standard deviations from triplicates are presented.

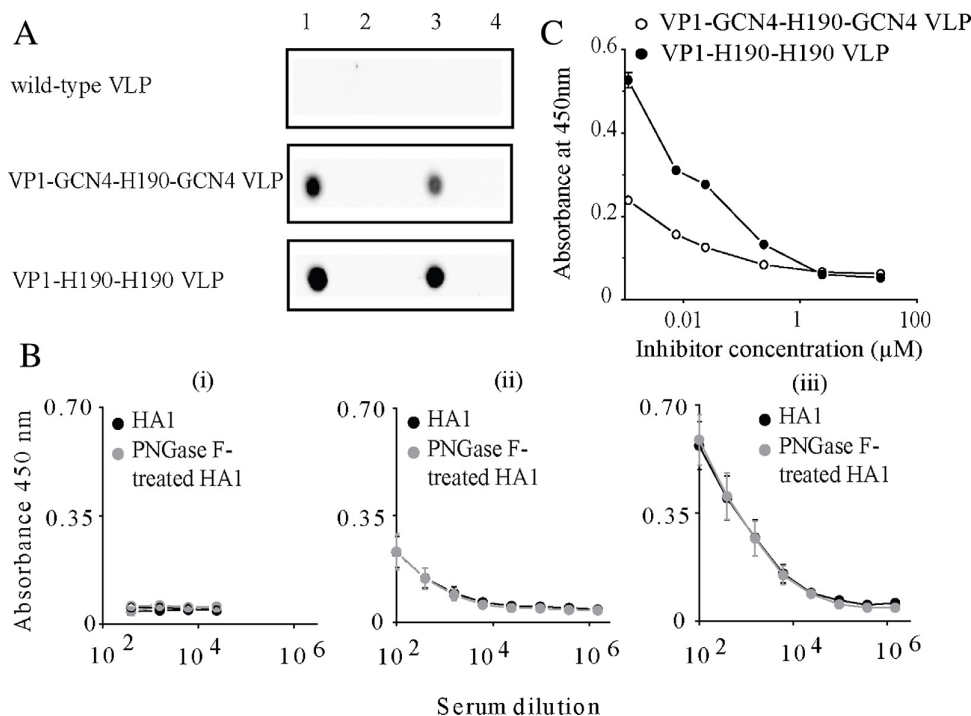


Fig. 5. Effects of N-linked glycans on reactivity of modular VLP antisera to HA1. (A) Dot blot analysis showing serum reactivity to (1) PNGase A-treated HA1; (2) PNGase A enzyme; (3) PNGase F-treated HA1; and (4) PNGase F enzyme, (B) indirect ELISA showing binding of IgGs from (i) wild-type VLP; (ii) VP1-GCN4-H190-GCN4 VLP; and (iii) VP1-H190-H190-VLP antisera, to HA1 and PNGase F-treated HA1 protein at various dilutions. Geometric mean ($n=8$) and standard error of the mean (SEM) are presented, and (C) competitive inhibition of antiserum (1/100 dilution) binding to immobilized PNGase F-treated HA1 by peptide H190 in solution. Arithmetic means and standard deviations from triplicates are presented.

the local sequence, rather than tertiary interaction. These signals are hydrophobic capping at the N- and C-terminal and hydrogen bonding to backbone NH or CO group [54]. We hypothesized that antigen module conformation predicted by molecular dynamics peptide simulation would allow prediction of the superior VLP modularization design. Simulation results suggest that H190 displayed by tandem repeat arraying on the VLP exhibits less structural deviation from its native conformation. In agreement with RMSD values, final conformation of simulated peptide after 20 ns (Fig. 1D) showed that the H190 element within module GCN4-H190-GCN4 unfolded whereas one copy of H190 in module H190-H190 retained

its higher helical propensity. The results suggested that module design would impact immune response quality, and that tandem repeat arraying would be more effective for H190 display.

Synthesis of modular VP1-GCN4-H190-GCN4 and VP1-H190-H190 VLPs was successful using a platform process [9] and robust bioprocessing techniques [45,55,56]. It is important to validate that the quaternary structural integrity of VLPs is maintained after modularization, as structural changes such as VLP aggregation may result in the antigen module being buried or denatured, thus introducing unpredictable experimental artifacts. Analyses by both asymmetrical flow field-flow fractionation [48] and transmission

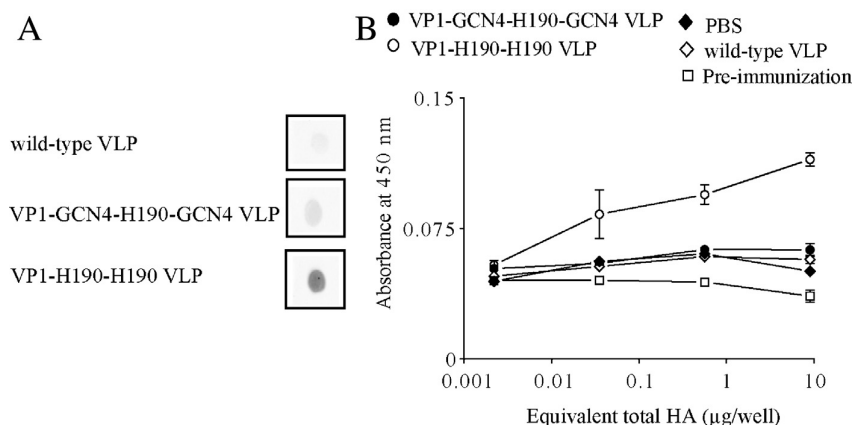


Fig. 6. Dot blot and indirect ELISA results showing reactivity of modular VLP antisera against Fluvax®. (A) Dot blot analysis showing serum reactivity to Fluvax® and (B) indirect ELISA showing binding of IgGs from modular VLP antiserum to Fluvax®. Arithmetic means and standard deviations from duplicates are presented.

electron microscopy (Fig. 2) confirmed that both modular VLP preparations were highly homogenous, and that the VLPs were morphologically indistinguishable from the related virion [57].

In vivo testing was performed to evaluate the immunological properties of the modular VLPs. This testing confirmed that the H190 element in both modular VLPs was highly immunogenic with sequence-specific IgG titers higher than 10^4 (Fig. 3B). The high sequence-specific IgG titer observed in this study is consistent with previous studies showing that presentation by a VLP carrier can promote immunogenicity for peptides which are otherwise non-immunogenic [18]. Interestingly, the H190 element in both modular VLPs led to sequence-specific IgG titers which were not significantly different, indicating that both module designs were comparably effective in presenting the H190 sequence in an immunogenic form.

We next sought to assess the quality of the immune response to each modular VLP. Levels of authentic structure-specific IgGs were evaluated by quantifying serum IgGs able to recognize H190 on recombinant HA1. Fig. 4B shows that VP1-H190-H190 VLP elicited a HA1-specific IgG titer exceeding 10^4 , 10-fold higher than that by immunization with VP1-GCN4-H190-GCN4 VLP. Competitive ELISA showed that IgG binding to HA1 was inhibited by peptide H190 (Fig. 4C), indicating that the HA1-specific IgGs indeed bound to HA1 via the H190 element instead of non-specific sites. These two lines of evidence demonstrated that VP1-H190-H190 VLP was able to induce a significant level of H190 authentic structure-specific IgGs. In contrast, the VP1-GCN4-H190-GCN4 VLP yielded a lower level of H190 authentic structure-specific IgGs despite a high titer of sequence-specific IgGs. These results are in a good agreement with the MD simulation predicting that the structure of the H190 element in VP1-H190-H190 VLP was more closely matched to the native influenza conformation.

The binding of modular VLP antiserum IgGs to split virion within a commercial vaccine preparation (Fluvax[®]) was next evaluated. Only VP1-H190-H190 VLP was able to induce IgGs that exhibited reactivity to Fluvax[®] (Fig. 6). This finding confirmed that the module display strategy for H190 affects native protein recognition of antibody induced by the modular VLPs. Together with the simulation result, the *in vivo* data demonstrated a correlation between helical propensity of the antigen module, as a result of antigen module display strategy, and quality of the immune response induced by the modular VLPs. Importantly, our data show that a design strategy based on tandem repeat arraying of H190 is more effective than using flanking GCN4 helix elements to drive structure within the H190 element. Tandem repeat arraying is also preferable due to its simplicity. The use of an element unrelated to the pathogen and the VLP backbone, such as GCN4, introduces unnecessary complexity, due to the inherent immunogenicity of that element. The GCN4 elements used in this study were indeed immunogenic (Fig. 3C and D), consistent with other reports [38]. However, other variants were shown to be not immunogenic [37]. The response to this foreign element cannot be predicted *a priori*, introducing additional product complexity.

N-glycans have been shown to have important roles in modulating biochemical properties of protein, such as protein folding and bioactivity, and as determinants in molecular recognition events [58]. They have been shown to be important in promoting and maintaining the structure of folded glycoproteins [59–62]. N-glycans affect protein structure locally at residues near the glycans [63] and globally [64]. As such, glycosylation engineering is often heralded as the pre-requisite for correct glycoprotein structure engineering [49]. Here we display vaccine-relevant antigen element (H190) from glycoprotein HA on an unrelated VLP and show that the proper secondary structure needed for native protein recognition of antibodies induced by modular VLPs could be mediated via rational antigen module display strategies, in a

system lacking glycosylation. Antibodies induced by non-glycosylated VLPs could recognize H190 within the glycosylated HA1 (Fig. 4), and the level of antibody binding was not influenced by HA1 glycosylation (Fig. 5). This result indicates that simple structure-based biomolecular engineering enables antigen display in such that the antigen can raise immune responses recognizing a glycosylated native protein.

The results of this study provide new insights into protein engineering of VLP vaccines based on rational display of a heterologous peptide antigen in the correct secondary structure. The rational display strategies here could ultimately enable safe and efficacious vaccine alternatives to historical approaches based on cultivation of an entire virus. This study is particularly relevant to prophylaxis of influenza, by demonstrating the proof-of-concept for an influenza vaccine based on modularization of MuPyV VLPs to display a hyper-variable region from HA, H190. The manufacture responsiveness of MuPyV VLP technology based on microbial platform [9] potentially enables the use of influenza hypervariable antigenic determinants to produce rapidly customized vaccines with genetic sequences precisely tailored for circulating and emerging viral strains.

Acknowledgements

M.A. acknowledges support in the form of an International Research Tuition Award and a Research Scholarship provided by The University of Queensland (UQ). A.P.J.M. acknowledges support from The Queensland Government Smart Futures Fund in the form of the 2010 Premier's Fellowship. The authors acknowledge support for this research through award of Grant 2010000484 from the National and International Research Alliances Program of the Queensland Government. M.A. acknowledges access to facilities at the Australian Microscopy & Microanalysis Research Facility at the Centre for Microscopy and Microanalysis, UQ, as well as training and technical assistance in their proficient use. The authors acknowledge Tania Rivera-Hernandez for generating the expression vector for VP1-H190-H190.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.vaccine.2013.06.087>.

References

- [1] Kaplan J, DeGrado WF. *De novo* design of catalytic proteins. *Proc Natl Acad Sci USA* 2004;101(32):11566–70.
- [2] Perchiaiccia JM, Ladiwala ARA, Bhattacharya M, Tessier PM. Structure-based design of conformation- and sequence-specific antibodies against amyloid β . *Proc Natl Acad Sci USA* 2012;109(1):84–9.
- [3] Ryu DDY, Nam DH. Recent progress in biomolecular engineering. *Biotechnol Prog* 2000;16(1):2–16.
- [4] Ostermeier M. Engineering allosteric protein switches by domain insertion. *Protein Eng Des Sel* 2005;18(8):359–64.
- [5] Boulianne GL, Hozumi N, Shulman MJ. Production of functional chimaeric mouse/human antibody. *Nature* 1984;312(5995):643–6.
- [6] Dormitzer PR, Ulmer JB, Rappuoli R. Structure-based antigen design: a strategy for next generation vaccines. *Trends Biotechnol* 2008;26(12):659–67.
- [7] Wibowo N, Chuan YP, Lua LHL, Middelberg APJ. Modular engineering of a microbially-produced viral capsomere vaccine for influenza. *Chem Eng Sci* 2012. <http://dx.doi.org/10.1016/j.ces.2012.04.001>.
- [8] Palomares LA, Ramirez OT. Challenges for the production of virus-like particles in insect cells: the case of rotavirus-like particles. *Biochem Eng J* 2009;45(3):158–67.
- [9] Middelberg APJ, Rivera-Hernandez T, Wibowo N, Lua LHL, Fan Y, Magor G, et al. A microbial platform for rapid and low-cost virus-like particle and capsomere vaccines. *Vaccine* 2011;29(41):7154–62.
- [10] Pattenden L, Middelberg A, Niebert M, Lipin D. Towards the preparative and large-scale precision manufacture of virus-like particles. *Trends Biotechnol* 2005;23(10):523–9.

- [11] Chuan YP, Lua LHL, Middelberg APJ. Virus-like particle bioprocessing. In: Subramanian G, editor. Biopharmaceutical production technology. Weinheim, Germany: Wiley-VCH Verlag & Co. KGaA; 2012. p. 139.
- [12] Wu T, Li SW, Zhang J, Ng MH, Xia NS, Zhao Q. Hepatitis E vaccine development: a 14-year odyssey. *Hum Vaccin Immunother* 2012;8(6):823–7.
- [13] Zhang J, Li SW, Wu T, Zhao Q, Ng MH, Xia NS. Hepatitis E virus: neutralizing sites, diagnosis, and protective immunity. *Rev Med Virol* 2012;22(5):339–49.
- [14] Bundy BC, Franciszkowicz MJ, Swartz JR. *Escherichia coli*-based cell-free synthesis of virus-like particles. *Biotechnol Bioeng* 2008;100(1):28–37.
- [15] Roldão A, Mellado MCM, Castilho LR, Carrondo MJT, Alves PM. Virus-like particles in vaccine development. *Expert Rev Vaccines* 2010;9(10):1149–76.
- [16] Garcea RL, Gissmann L. Virus-like particles as vaccines and vessels for the delivery of small molecules. *Curr Opin Biotechnol* 2004;15(6):513–7.
- [17] Boisgérault F, Morón G, Leclerc C. Virus-like particles: a new family of delivery systems. *Expert Rev Vaccines* 2002;1(1):101–9.
- [18] Neirynek S, Deroo T, Saelens X, Vanlandschoot P, Jou WM, Fiers W. A universal influenza A vaccine based on the extracellular domain of the M2 protein. *Nat Med* 1999;5(10):1157–63.
- [19] Xia M, Tan M, Wei C, Zhong W, Wang L, McNeal M, et al. A candidate dual vaccine against influenza and noroviruses. *Vaccine* 2011;29(44):7670–7.
- [20] Crisci E, Fraile L, Moreno N, Blanco E, Cabezón R, Costa C, et al. Chimeric calicivirus-like particles elicit specific immune responses in pigs. *Vaccine* 2012;30(14):2427–39.
- [21] Yin Y, Li H, Wu S, Dong D, Zhang J, Fu L, et al. Hepatitis B virus core particles displaying *Mycobacterium tuberculosis* antigen ESAT-6 enhance ESAT-6-specific immune responses. *Vaccine* 2011;29(34):5645–51.
- [22] Kratz PA, Böttcher B, Nassal M. Native display of complete foreign protein domains on the surface of hepatitis B virus capsids. *Proc Natl Acad Sci USA* 1999;96(5):1915–20.
- [23] Murawski MR, McGinnes LW, Finberg RW, Kurt-Jones EA, Massare MJ, Smith G, et al. Newcastle disease virus-like particles containing respiratory syncytial virus G protein induced protection in BALB/c mice: with no evidence of immunopathology. *J Virol* 2010;84(2):1110–23.
- [24] Buonaguro L, Buonaguro F, Tornesello M, Mantas D, Beth-Giraldo E, Wagner R, et al. High efficient production of Pr55^{gag} virus-like particles expressing multiple HIV-1 epitopes: including a gp120 protein derived from an Ugandan HIV-1 isolate of subtype A. *Antiviral Res* 2001;49(1):35–47.
- [25] McGinnes LW, Gravel KA, Finberg RW, Kurt-Jones EA, Massare MJ, Smith G, et al. Assembly and immunological properties of Newcastle disease virus-like particles containing the respiratory syncytial virus F and G proteins. *J Virol* 2011;85(1):366–77.
- [26] Skehel JJ, Wiley DC. Receptor binding and membrane fusion in virus entry: the influenza hemagglutinin. *Annu Rev Biochem* 2000;69:531–69.
- [27] Neumann G, Noda T, Kawaoka Y. Emergence and pandemic potential of swine-origin H1N1 influenza virus. *Nature* 2009;459(7249):931–9.
- [28] Whittle JRR, Zhang R, Khurana S, King LR, Manischewitz J, Golding H, et al. Broadly neutralizing human antibody that recognizes the receptor-binding pocket of influenza virus hemagglutinin. *Proc Natl Acad Sci USA* 2011;108(34):14216–21.
- [29] Knossow M, Gaudier M, Douglas A, Barrere B, Bizebard T, Barbey C, et al. Mechanism of neutralization of influenza virus infectivity by antibodies. *Virology* 2002;302(2):294–8.
- [30] Naeve C, Hinshaw V, Webster R. Mutations in the hemagglutinin receptor-binding site can change the biological properties of an influenza virus. *J Virol* 1984;51(2):567–9.
- [31] Weis W, Brown J, Cusack S, Paulson J, Skehel J, Wiley D. Structure of the influenza virus haemagglutinin complexed with its receptor: sialic acid. *Nature* 1988;333(6172):426–31.
- [32] Caton AJ, Brownlee GG, Yewdell JW, Gerhard W. The antigenic structure of the influenza virus A/PR/8/34 hemagglutinin (H1 subtype). *Cell* 1982;31(2 II):417–27.
- [33] Brownlee G, Fodor E. The predicted antigenicity of the haemagglutinin of the 1918 Spanish influenza pandemic suggests an avian origin. *Philos Trans R Soc B* 2001;356(1416):1871–6.
- [34] Yu X, Tsibane T, McGraw PA, House FS, Keefer CJ, Hicar MD, et al. Neutralizing antibodies derived from the B cells of 1918 influenza pandemic survivors. *Nature* 2008;455(7212):532–6.
- [35] Sia SK, Kim PS. Protein grafting of an HIV-1-inhibiting epitope. *Proc Natl Acad Sci USA* 2003;100(17):9756–61.
- [36] Domingues H, Cregut D, Sebald W, Oschkinat H, Serrano L. Rational design of a GCN4-derived mimetic of interleukin-4. *Nat Struct Mol Biol* 1999;6(7):652–6.
- [37] Relf WA, Cooper J, Brandt ER, Hayman WA, Anders RF, Pruksakorn S, et al. Mapping a conserved conformational epitope from the M protein of group A *Streptococci*. *Peptide Res* 1996;9(1):12–20.
- [38] De Filette M, Martens W, Roose K, Deroo T, Vervalle F, Bentahir M, et al. An influenza A vaccine based on tetrameric ectodomain of matrix protein 2. *J Biol Chem* 2008;283(17):11382–7.
- [39] Kajava AV. Tandem repeats in proteins: from sequence to structure. *J Struct Biol* 2012;179(3):279–88.
- [40] Fontenot J, Mariappan S, Catasti P, Domenech N, Finn OJ, Gupta G. Structure of a tumor associated antigen containing a tandemly repeated immunodominant epitope. *J Biomol Struct Dyn* 1995;13(2):245–60.
- [41] Fontenot JD, Tjandra N, Bu D, Ho C, Montelaro RC, Finn OJ. Biophysical characterization of one-, two-, and three-tandem repeats of human mucin (muc-1) protein core. *Cancer Res* 1993;53(22):5386.
- [42] Rueda P, Morón G, Sarraseca J, Leclerc C, Casal JI. Influence of flanking sequences on presentation efficiency of a CD8+ cytotoxic T-cell epitope delivered by parvovirus-like particles. *J Gen Virol* 2004;85(3):563–72.
- [43] Jain S, Patrick AJ, Rosenthal KL. Multiple tandem copies of conserved gp41 epitopes incorporated in gag virus-like particles elicit systemic and mucosal antibodies in an optimized heterologous vector delivery regimen. *Vaccine* 2010;28(43):7070–80.
- [44] Zheng B, Graham FL, Johnson DC, Hanke T, McDermott MR, Prevec L. Immunogenicity in mice of tandem repeats of an epitope from herpes simplex gD protein when expressed by recombinant adenovirus vectors. *Vaccine* 1993;11(12):1191–8.
- [45] Chuan YP, Lua LHL, Middelberg APJ. High-level expression of soluble viral structural protein in *Escherichia coli*. *J Biotechnol* 2008;134(1/2):64–71.
- [46] Lipin DI, Lua LHL, Middelberg APJ. Quaternary size distribution of soluble aggregates of glutathione-S-transferase-purified viral protein as determined by asymmetrical flow field flow fractionation and dynamic light scattering. *J Chromatogr A* 2008;1190(1/2):204–14.
- [47] Brito LA, Singh M. Acceptable levels of endotoxin in vaccine formulations during preclinical research. *J Pharm Sci* 2011;100(1):34–7.
- [48] Chuan YP, Fan YY, Lua L, Middelberg APJ. Quantitative analysis of virus-like particle size and distribution by field-flow fractionation. *Biotechnol Bioeng* 2008;99(6):1425–33.
- [49] Chen J-R, Ma C, Wong C-H. Vaccine design of hemagglutinin glycoprotein against influenza. *Trends Biotechnol* 2011;29(9):426–34.
- [50] Burston DR. A vaccine for HIV type 1: the antibody perspective. *Proc Natl Acad Sci USA* 1997;94(19):10018–23.
- [51] Lai CY, Tsai WY, Lin SR, Kao CL, Hu HP, King CC, et al. Antibodies to envelope glycoprotein of dengue virus during the natural course of infection are predominantly cross-reactive and recognize epitopes containing highly conserved residues at the fusion loop of domain II. *J Virol* 2008;82(13):6631–43.
- [52] Macagno A, Bernasconi NL, Vanzetta F, Dander E, Sarasini A, Revello MG, et al. Isolation of human monoclonal antibodies that potently neutralize human cytomegalovirus infection by targeting different epitopes on the gH/gL/UL128–131A complex. *J Virol* 2010;84(2):1005–13.
- [53] Sundaram R, Lynch MP, Rawale SV, Sun Y, Kazanji M, Kaumaya PTP. De novo design of peptide immunogens that mimic the coiled coil region of human T-cell leukemia virus type-1 glycoprotein 21 transmembrane subunit for induction of native protein reactive neutralizing antibodies. *J Biol Chem* 2004;279(23):24141–51.
- [54] Baldwin RL, Rose GD. Is protein folding hierarchical? I. Local structure and peptide folding. *Trends Biochem Sci* 1999;24(1):26–33.
- [55] Lipin DI, Raj A, Lua LHL, Middelberg APJ. Affinity purification of viral protein having heterogeneous quaternary structure: modeling the impact of soluble aggregates on chromatographic performance. *J Chromatogr A* 2009;1216(30):5696–708.
- [56] Liew MWO, Rajendran A, Middelberg APJ. Microbial production of virus-like particle vaccine protein at gram-per-litre levels. *J Biotechnol* 2010;150(2):224–31.
- [57] Salunke DM, Caspar DLD, Garcea RL. Self-assembly of purified polyomavirus capsid protein VP1. *Cell* 1986;46(6):895–904.
- [58] Cumming DA. Glycosylation of recombinant protein therapeutics: control and functional implications. *Glycobiology* 1991;1(2):115–30.
- [59] Mitra N, Sinha S, Ramya TNC, Suroli A. N-linked oligosaccharides as out-fitters for glycoprotein folding: form and function. *Trends Biochem Sci* 2006;31(3):156–63.
- [60] O'Connor SE, Imperiali B. A molecular basis for glycosylation-induced conformational switching. *Chem Biol* 1998;5(8):427–37.
- [61] Carlos J, Tschampel SM, Woods RJ, Imperiali B. Effects of glycosylation on peptide conformation: a synergistic experimental and computational study. *J Am Chem Soc* 2004;126(27):8421–5.
- [62] Helenius A, Aebi M. Roles of N-linked glycans in the endoplasmic reticulum. *Annu Rev Biochem* 2004;73(1):1019–49.
- [63] Petrescu AJ, Milac AL, Petrescu SM, Dwek RA, Wormald MR. Statistical analysis of the protein environment of N-glycosylation sites: implications for occupancy, structure, and folding. *Glycobiology* 2004;14(2):103–14.
- [64] Helenius A. How N-linked oligosaccharides affect glycoprotein folding in the endoplasmic reticulum. *Mol Biol Cell* 1994;5(3):253.
- [65] Petterson EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem* 2004;25(13):1605–12.

Chapter 4. Display strategy improvements for induction of a higher quality of antibodies

4.1 Introduction

Structural vaccinology has provided opportunities to combat viruses which otherwise may not be achievable through conventional vaccinology (Grimm and Ackerman 2013). In structural vaccinology, it is believed that whole antigen proteins are not obligatory for induction of immune response against viruses (Thomas and Luxon 2013). Vaccine candidates against viruses are developed based on identification of peptide antigens from the proteins that are critical for immunity. The peptide antigens are then engineered to increase their feasibility as vaccine candidates (Gori et al. 2013).

For induction of strong and long lasting protective antibodies, engineering of peptide antigens can be driven by structural consideration, which exploits: (i) conformation integrity, and (ii) key characteristics of viruses. Information regarding the conformation integrity of peptide antigens can be attained from available structural information of the viruses (Gori et al. 2013, Kulp and Schief 2013). Conformation integrity of peptide antigens is important, as it determines the ability of induced antibodies to recognise the native proteins (i.e. conformational integrity determines the quality of antibodies induced). Key characteristics of viruses are also important for induction of strong and long lasting protective antibodies against viruses. Of particular interest is repetitive peptide antigen display on the surface of viruses, which is efficiently recognised by B-cell receptors, leading to efficient induction of antibodies. This repetitive display of peptide antigens can be efficiently simulated by presenting the peptide antigens on modular virus-like particles (VLPs; reviewed in Section 2.3) (Spohn and Bachmann 2008).

Chapter 3 of this thesis reports the engineering of peptide antigens to achieve both conformational integrity and repetitive pattern display of peptide elements on modular VLPs. A peptide antigen Helix 190 (H190) from influenza virus A/California/07/2009 H1N1 was presented on a modular VLP from murine polyomavirus (MuPyV) VP1 protein. Two display strategies were investigated, i.e. (i) tandem repeat display, and (ii) the use of GSGS spacer elements together with GCN4 helix promoter elements (refer to Fig. 1 in Chapter 3 for details). The chapter concludes that modular VLPs using the tandem repeat H190 display strategy induced higher quality antibodies, resulting in higher antibody level against the recombinant HA1 protein and split influenza virions.

The findings reported in Chapter 3 lead us to the question what improvements can be made to the tandem repeat display strategy to induce antibodies of a higher quality? A higher quality of antibodies is desired because the quality of antibodies has been shown to represent the protection level of the antibodies (Alvarez et al. 2010). In an attempt to increase the quality of antibody, two approaches are explored in this chapter: (i) the use of adjuvants, and (ii) increasing the number of tandem repeat H190. The first approach focuses on the use of adjuvant to induce a higher titre of antigen-specific antibodies. One of the promising adjuvants is AdvaxTM-1, which is a polysaccharide adjuvant derived from delta-inulin (Cooper and Petrovsky 2011). It has been shown to enhance the immunogenicity of various vaccine candidates and consequently their protective efficacy (Honda-Okubo et al. 2012, Petrovsky et al. 2013, Saade et al. 2013). The second approach focuses on increasing the number of H190 tandem repeats from one copy to three, four, and five copies (Figure 4-1). Previous studies on peptide-based vaccines show that tandem repeat display strategy affects both structure (Fontenot et al. 1995, Fontenot et al. 1993) and immunogenicity of peptide repeats (Jain et al. 2010, Rueda et al. 2004, Zheng et al. 1993).

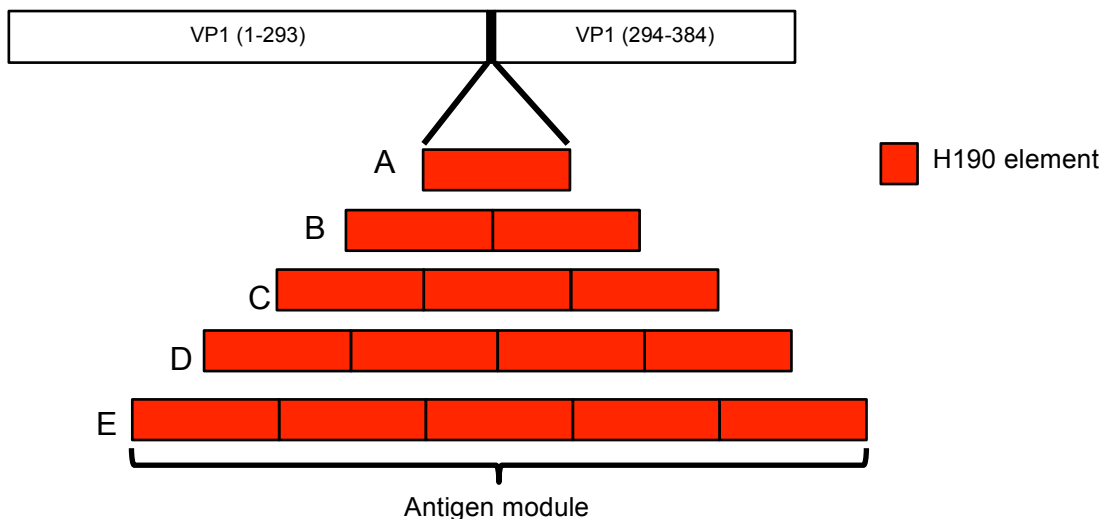


Figure 4-1. Schematic diagram of antigen modules for modular constructs reported in Chapter 4. These antigen modules comprised of various copy number of H190: (A) One copy of H190, (B) Two copies of H190, (C) Three copies of H190, (D) Four copies of H190, and (E) Five copies of H190. Numbers in the white block refer to amino acid residue of wild-type murine polyomavirus VP1 protein. The diagram is not to scale.

4.2 Materials and methods

4.2.1 Generation of modular constructs

The commercial plasmid pGEX-4T-1 (Novagen, Billerica, MA, USA) with gene insert encoding murine polyomavirus VP1 protein (M34958) was generously provided by Professor Robert Garcea (University of Colorado) (Middelberg et al. 2011). This plasmid was designated as pGEX-VP1. *NaeI* and *AfeI* recognition sites were inserted into pGEX-VP1 at positions 86 and 293 of VP1, yielding plasmid pGEX-VP1-S1S4 (Middelberg et al. 2011). Similarly, plasmid pGEX-VP1-S4 was generated by inserting *AfeI* recognition site at position 293 of VP1 (Middelberg et al. 2011).

DNA sequence of antigen module containing H190 from influenza virus A/California/07/2009 H1N1 (STSADQQSLYQNADAY) was codon optimised for *E. coli*. Homologous DNA sequences flanking (21-24 bp) *AfeI* site of plasmid pGEX-VP1-S1S4 or pGEX-VP1-S4, were added to the 5' and 3' end DNA sequence of the antigen module (Figure 4-2).

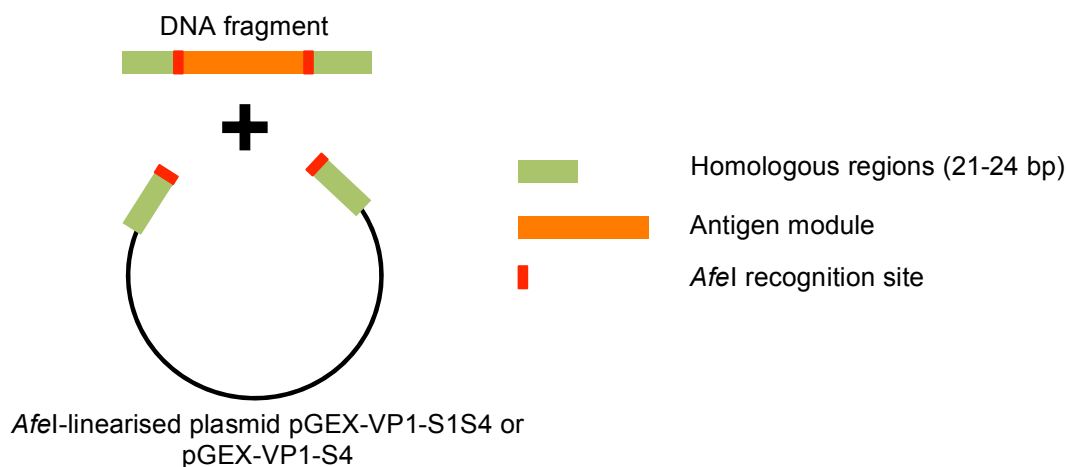


Figure 4-2. Schematic diagram of design of DNA fragments containing an antigen module and homologous regions.

A DNA fragment containing the antigen module and homologous regions was synthesised by (i) annealing and phosphorylation of complementary oligonucleotides (Sambrook and Russell 2001) or, (ii) assembling a set of oligonucleotides (gene assembly), which were designed using DNAWorks (Hoover and Lubkowski 2002). Details of oligonucleotides are given in Table 4-1. Gene assembly was conducted in a polymerase

chain reaction (PCR) mixture (1 U Phusion Hot Start II DNA polymerase^{*}, 1x Phusion HF Buffer, 0.2 μ M mixed oligonucleotides, and 200 μ M dNTPs), and performed at an annealing temperature of 62°C. The synthesised DNA fragment was then amplified using 1 μ l from initial assembly PCR. The PCR reaction for final amplification was the same as for the assembly PCR, except that utmost 5' and 3' oligonucleotides were used as primers at final concentration 25 μ M. The amplified DNA fragment was then purified using PureLinkTM Quick Gel Extraction Kit (Life Technologies, Carlsbad, CA, USA).

The purified DNA fragment was cloned into *AfeI*-linearised vector pGEX-VP1-S1S4 or pGEX-VP1-S4 using *in vivo* homologous recombination (Bubeck et al. 1993, Jones 1994, Oliner et al. 1993, Parrish et al. 2004). Briefly, purified *AfeI*-linearised pGEX-VP1-S1S4 or pGEX-VP1-S4 was mixed with purified DNA fragment at an optimised ratio of 1 to 5 (the ratio referred to the ratio of DNA copy number per volume[†]) to a maximum final volume of 7 μ l. The mixture was then added into 50 μ l of One Shot[®] OmniMAXTM 2 T1^R Chemically Competent *E. coli* strain (InvitrogenTM, Life Technologies, Carlsbad, CA, USA) using the heat shock method (Sambrook and Russell 2001). DNA sequences of designated constructs were confirmed by sequencing conducted by the Australian Genome Research Facility (AGRF, Brisbane, Australia).

* Thermo Scientific, Pittsburgh, PA, USA.

† DNA copy number was calculated based on the assumption that one mole of a base pair weighs 650 g and one mole of a molecule equals to 6.022×10^{23} molecules.

Table 4-1. List of modular constructs reported in Chapter 4.

Construct name	Antigen module	Oligonucleotides to synthesise DNA fragment containing the antigen module and the homologous regions	Clone into
VP1-H190 (Figure 4-1A)	One copy of H190	Complementary oligonucleotides: 5'ggctggagagttacaagaagctctacctctgcgaccagcagtc tctgtaccagaacgcggacgcgtatgcttatgatgtccatcactgg3'	pGEX-VP1-S1S4
VP1-H190-H190 (Figure 4-1B)	Two copies of H190	Reported in Chapter 3	pGEX-VP1-S1S4
VP1-H190-3x (Figure 4-1C)	Three copies of H190	Oligonucleotide set: 5'atggctggagagttacaagaagctctacctct3' 5'cagagactgctgatccgcagaggtagagcttctgtaactc3' 5'gcggatcagcagtcctctgatcaaaacgcgcgacgcgtattc3' 5'agggactgttggtcagcggacgtagaatacgcgtcggcgtt3' 5'cgctgaccaacagtcctctaccagaacgcggacgcataca3' 5'gctttgctggtcggcagacgtgctgatgctcgcgcttct3' 5'ctgccgaccagcaaagcctctatcagaatgcagacgcctac3' 5'tctccagtgatggacatcataagcgtaggcgtctgcattctgata3'	pGEX-VP1-S1S4
VP1-H190-4x (Figure 4-1D)	Four copies of H190	Oligonucleotide set: 5'atggctggagagttacaagaagctctac3' 5'actgctggctgctgctcgtagagcttctgtaactctcca3' 5'gagcgcagaccagcagtcctctaccaaaatgcggatgca3' 5'tggtccgcagaagtgcgtgatgcatccgcattttggtaca3' 5'agcacttctgcgaccacaacagtcctctgatacagaacgcag3' 5'gcgctagtagagtaggcgtctgcttctgatacagagact3' 5'acgcctactctactagcgcggatcagcaatctcttacc3' 5'ggagtacgcacatcggcattctggtagagagattgctgatcc3' 5'gaatgccgatgcgtactccacctctgccgatcaacagagc3' 5'agcgtcagcgtttgatagaggctctgtgatcggcagag3' 5'tctatcaaaacgctgacgcttatgcttatgatgtccatcactg3' 5'tctccagtgatggacatcataagcata3'	pGEX-VP1-S1S4
VP1-H190-5x (Figure 4-1E)	Five copies of H190	Oligonucleotide set: 5'atggctggagagttacaagaagca3' 5'tgctggctggcagaggtgctgcttctgtaactctccagc3' 5'cctctgccgaccagcaatctctaccaaaatgcggatgc3' 5'gatcggcggacgtgctgtaagcatcggcattttggtagag3' 5'agcagtcgccgatcaacagagcctgtatcaaaacgccg3' 5'ccgcgctggtctatatgcgtcggcgttttgatacaggct3' 5'catatagcaccagcgcggatcagcaaaagcctctatcagaa3' 5'cgtggagtacgcattctgatagaggcttctgta3' 5'gcagatgcgtactccacgtctgctgaccagcagtcctgt3' 5'gagtaggcgtctgcttctggtacagggactgctggtcag3' 5'gaacgcagacgcctactctactctgcgaccacaacagct3' 5'aagcgtccgcgttttgatacagagactgttggtccgcaga3' 5'tcaaaacgcggacgcttatgcttatgatgtccatcactgg3' 5'tctccagtgatggacatcataagca3'	pGEX-VP1-S1S4

4.2.2 Expression and purification of modular VP1 proteins

Modular VP1 proteins were expressed as Glutathione S-transferase (GST)-tagged proteins in *E. coli* and purified using affinity chromatography as described previously (Chuan et al. 2008, Lipin et al. 2008, Middelberg et al. 2011). Following GST removal, modular VP1 capsomeres were separated from aggregates and GST through size-

exclusion chromatography (SEC). Briefly, plasmid was transformed into *E. coli* Rosetta (DE3) pLysS cells (EMD Millipore, Merck KGaA, Darmstadt, Germany) using the heat shock method (Sambrook and Russell 2001). A single colony of the transformed cells was inoculated in 5 ml Terrific Broth (TB) medium for overnight incubation at 30°C, 180 rpm using a rotary shaker (Bioline, Edwards Instrument Company, Australia). Then, 400 µl of the seed culture was cultivated in 400 ml of TB medium at 37°C, 180 rpm. When the cell OD_{600 nm} reached 0.5, culture was cooled down to 26°C and subsequently induced with 0.2 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). The culture was then incubated at 26°C, 180 rpm for 16 h, and harvested by centrifugation at 4°C, 6000g for 20 min. All cultures were supplemented with 50 µg ml⁻¹ of ampicillin and 34 µg ml⁻¹ of chloramphenicol.

Cell pellets from 800 ml culture were resuspended in 40 ml of Lysis Buffer (40 mM Tris-base, 200 mM NaCl, 1 mM EDTA, 5% (v/v) glycerol, 5 mM DTT, pH 8.0). The suspension was sonicated for 4 cycles of 45 s using a Branson Sonifier 250 cell disruptor mounted with a microtip (Branson Ultrasonics Corporation, Connecticut, USA), and subsequently clarified by centrifugation at 4°C, 18000g for 20 min. Supernatant was collected and was filtered using 0.45 µm filters (Pall, New York, USA). The clarified supernatant was then loaded into a GST affinity column (GSTrap HP 5 ml, GE Healthcare, UK) at a flow rate of 0.5 ml min⁻¹. Bound GST-tagged VP1 protein was eluted with Elution Buffer (40 mM Tris-base, 200 mM NaCl, 1 mM EDTA, 5% (v/v) glycerol, 5 mM DTT, 10 mM GSH, pH 8.0).

GST tag was removed from VP1 by incubating GST-tagged protein with thrombin (Catalogue# 27-0846-02, GE Healthcare, UK) at a ratio of 40:1 (thrombin unit/ml protein) for 2 h at room temperature. Digested protein was centrifuged and the supernatant was then loaded into size exclusion chromatography (SEC) column Superdex 200 10/300 GL (GE Healthcare, UK) pre-equilibrated with Lysis Buffer, as described previously (Chuan et al. 2008), to separate modular VP1 capsomeres from soluble aggregates and GST. Capsomeres were eluted at a volumetric outflow of approximately 12.5 ml after sample injection.

4.2.3 Assembly of modular VP1 proteins

Endotoxins were removed from the purified modular VP1 capsomeres to below 20 EU ml⁻¹ using Vivapure Q Maxi H spin column (Sartorius Stedim, Goettinge, Germany) as previously described in Chapter 3 (Anggraeni et al. 2013). Modular VP1 capsomeres were assembled *in vitro* into VLPs by dialysis against Assembly Buffer 1 (0.5 M (NH₄)₂SO₄, 20

mM Tris-base, 5% (v/v) glycerol, 1 mM CaCl₂, pH 7.4) for 15 h at room temperature (Gleiter and Lilie 2001) in an endotoxin-free environment. Formed VLPs were then dialysed against phosphate buffer saline (PBS; E404, Amresco, USA) at 4°C for 24 h. VLPs were characterised using asymmetrical flow field-flow fractionation (AF4) coupled with multi-angle light scattering (MALS) and visualised using transmission electron microscopy (TEM) (Chuan et al. 2008).

4.2.4 SDS-PAGE

For analysis by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE), samples were mixed with 5x sample buffer (10% (w/v) sodium dodecyl sulphate, 10 mM DTT, 20% (v/v) glycerol, 0.2 M Tris-HCl, pH 6.8, 0.05% (w/v) Coomassie Brilliant Blue R250) at volume ratio 4:1. Pre-treated samples were then heated at 95°C for 5 min. SDS-PAGE analysis was performed by resolving 12 µl of the samples on polyacrylamide gels (10% (w/v) polyacrylamide, 0.375 M Tris pH 8.8, 0.1% (w/v) ammonium persulfate, 0.1% (w/v) SDS, 0.04% (v/v) TEMED). Electrophoresis was conducted in Tris-Glycine running buffer (25 mM Tris-HCl, 200 mM glycine, 0.1% (w/v) SDS) for 55 min at 150 V. Then, polyacrylamide gels were incubated in staining solution (30% (v/v) methanol, 10% (v/v) acetic acid, 0.1% (w/v) Coomassie Brilliant Blue R250). Subsequently, the gels were incubated in destaining solution (45% (v/v) methanol, 15% (v/v) acetic acid) for visualisation of protein bands.

Each SDS-PAGE gel picture in Figure 4-3 was taken from one gel. Relevant sample lanes were selected and positioned on the right hand side of the ladder. The SDS-PAGE gel picture in Figure 4-8 was taken from two different gels which were aligned to a ladder.

4.2.5 Protein concentration determination

Protein concentration was determined using UV absorbance at 280 nm, based on the Beer-Lambert Law (Aitken and Learmonth 1996),

$$A = \epsilon b c$$

where A is the measured absorbance at 280 nm, ϵ is the extinction coefficient of protein ($M^{-1} cm^{-1}$) at 280 nm measured in water, b is the sample path length (10 mm), and c is the protein concentration (M). Theoretical molecular weight and extinction coefficient of each protein was obtained using the ProtParam tool (Gasteiger et al. 2005), and summarised in Table 4-2.

Table 4-2. List of theoretical molecular weight and extinction coefficient of VP1 proteins reported in Chapter 4.

Protein	Theoretical molecular weight (g mol ⁻¹)	Extinction coefficient, at 280 nm in water (M ⁻¹ cm ⁻¹)
Wild-type (wt) VP1	42763.6	58057
VP1-H190	44507.4	61225
VP1-H190-H190	46251.2	64205
VP1-H190-3x	47995.0	67185
VP1-H190-4x	49738.7	70165
VP1-H190-5x	51482.5	73145

4.2.6 Mice immunisation

In one study, mice were immunised with 50 µg of unmodified VP1 (wt-VP1) VLPs, VP1-GSGS-H190 VLPs, VP1-GCN4-H190 VLPs, VP1-H190 VLPs, VP1-H190-H190 VLPs, VP1-H190-3x VLPs, and VP1-H190-4x VLPs in total volume of 50 µl. Immunisation was conducted as described in Chapter 3, with the following modifications: (i) number of mice was five mice/group; (ii) ethical clearance number AIBN/189/12/NIRAP/SMART FUTURE.

In another study, mice were immunised with 25 µg of VP1-H190-H190 VLPs, with and without adjuvant, in a total volume of 50 µl. The adjuvant was AdvaxTM-1 (Vaxine Pty. Ltd., Adelaide, Australia), which was supplied at 50 mg ml⁻¹. AdvaxTM-1 was then diluted into 1 mg ml⁻¹ in PBS, and gently mixed with an equal volume of VP1-H190-H190 VLPs at the same concentration. Immunisation was conducted as described in Chapter 3, with the following modifications: (i) number of mice was five mice/group; (ii) ethical clearance number AIBN/058/13/NIRAP/SMART FUTURES.

4.2.7 ELISA

Indirect ELISA and competitive ELISA were conducted using sera from day 56 as previously reported in Chapter 3. In this chapter, immobilised antigens for both ELISA assays were: (i) biotinylated peptide H190 (biotin-GGGGSSTSADQQSLYQNADAY-NH₂) from Peptide 2.0 (Chantilly, VA, USA), and (ii) recombinant A/California/07/2009 H1N1 HA1 protein, produced in *Trichoplusia ni* cell line by the Protein Expression Facility (The University of Queensland, Australia). Biotinylated peptide H190 in solution was used as the inhibitor in competitive ELISA. For competitive ELISA, data were presented as percentage relative absorbance [(B/B₀) x 100%]. B was absorbance at 450 nm at a certain concentration of peptide H190, and B₀ was absorbance at 450 nm when no peptide H190 was present.

4.2.8 Statistical analysis

Statistical analysis was performed on log-transformed data using GraphPad Prism 5.03 (GraphPad, Inc.). Comparison of antibody titre for multigroup was conducted using one-way ANOVA with Tukey's post hoc test.

4.3 Results and discussion

4.3.1 Solubility and assembly studies

Three modular constructs: (i) VP1-H190-3x, (ii) VP1-H190-4x, and (iii) VP1-H190-5x (Figure 4-1C-E) bearing three, four, and five copies of tandem repeat H190 elements, were generated respectively as described in Section 4.2.1. The modular VP1 proteins were then expressed and purified as described in Section 4.2.2 to yield modular VP1 capsomeres. These purified modular VP1 capsomeres were assembled to form VLPs in Assembly Buffer 1 (0.5 M (NH₄)₂SO₄, 20 mM Tris-base, 5% (v/v) glycerol, 1 mM CaCl₂, pH 7.4) and the formed VLPs were dialysed against PBS as described in Section 4.2.3.

After dialysis against Assembly Buffer 1 and subsequently PBS, precipitations were observed for all modular constructs as indicated from the turbidity of the protein solutions. Table 4-3 compares the turbidity of protein solution for the three modular VP1 constructs. As can be seen from Table 4-3, protein solutions for VP1-H190-4x and VP1-H190-5x were visually more turbid than for VP1-H190-3x, indicating that both modular VP1 capsomeres precipitated more than VP1-H190-3x. The observation in Table 4-3 was confirmed by SDS-PAGE analysis. Precipitates were removed through centrifugation and protein solution sampled before and after centrifugation was analysed by SDS-PAGE gel. Figure 4-3 shows that after centrifugation, small proportions of VP1-H190-4x and VP1-H190-5x were found in the supernatant, confirming that the majority of both modular VP1 variants had precipitated. In contrast, the majority of VP1-H190-3x was found in the supernatant. Because precipitates were observed after dialysis against Assembly Buffer 1, both Figure 4-3 and Table 4-3 indicated that modular proteins VP1-H190-4x and VP1-H190-5x were less stable in Assembly Buffer 1 than was VP1-H190-3x.

Table 4-3. Observation of protein solution turbidity after dialysis against PBS.

Modular VP1 protein	Number of H190 tandem repeats	Turbidity of protein solution
VP1-H190-3x	3	+
VP1-H190-4x	4	++++
VP1-H190-5x	5	++++

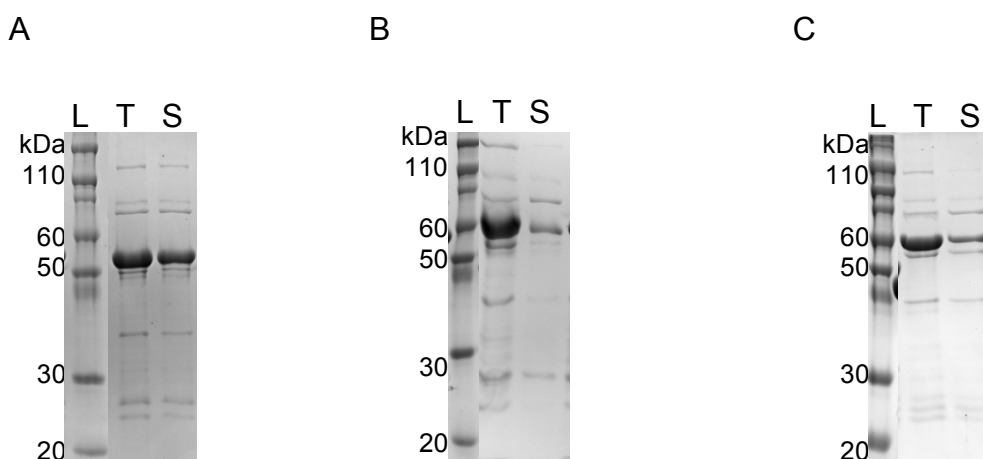


Figure 4-3. SDS-PAGE analysis of modular VP1 bearing three, four, and five copies of H190 after dialysis against PBS. (A) VP1-H190-3x, (B) VP1-H190-4x, and (C) VP1-H190-5x. Lanes: (L) Pre-stained protein marker, (T) Total protein, (S) Soluble protein. The dominant bands correspond to the GST-tagged modular VP1.

After precipitates were removed by centrifugation, supernatants were analysed using asymmetric flow-field flow fractionations (AF4), coupled with multi-angle light scattering (MALS). AF4 is a flow-based separation technique capable of separating particles with size range of 1-1000 nm at high resolution. The separation is performed inside a narrow ribbon-like channel (Figure 4-4A), where carrier liquid flows in a parabolic laminar Newtonian flow profile. The velocity at the wall is zero, and increases to a maximum at the centre of the flow. Additional to the parabolic flow, cross-flow is applied perpendicularly to the flow of carrier liquid. The cross-flow forces particles toward the wall, and driven by concentration difference, particles start to diffuse back toward the centre of the parabolic flow (Figure 4-4B). The diffusion rates of the particles depend on the size of the particles. Smaller particles diffuse faster, and hence they are closer to centre of the laminar flow and travel with a faster velocity (Fraunhofer and Winter 2004, Giddings 2002, Wahlund 2000).

Coupled with a suitable detection technique, such as MALS, AF4 has been shown to be a powerful method for VLP analysis (Chuan et al. 2008). Analysis of VLPs using AF4-MALS comprises of two main steps. These steps are (i) focusing of injected samples at near the injection point, and (ii) elution step. Fractionation of VLPs from the MuPyV VP1 in AF4-MALS typically reveals three peaks, i.e. (i) a void peak at 10 min, (ii) a VLP peak at 20 min, and (iii) a soluble amorphous aggregate peak at 40 min. In the study by Chuan et al., the void peak marked the end of focusing step and the start of the elution step. The

peak was observed due to elution of materials that were too small to be retained at a certain cross-flow value, such as would be expected for unassembled capsomeres.

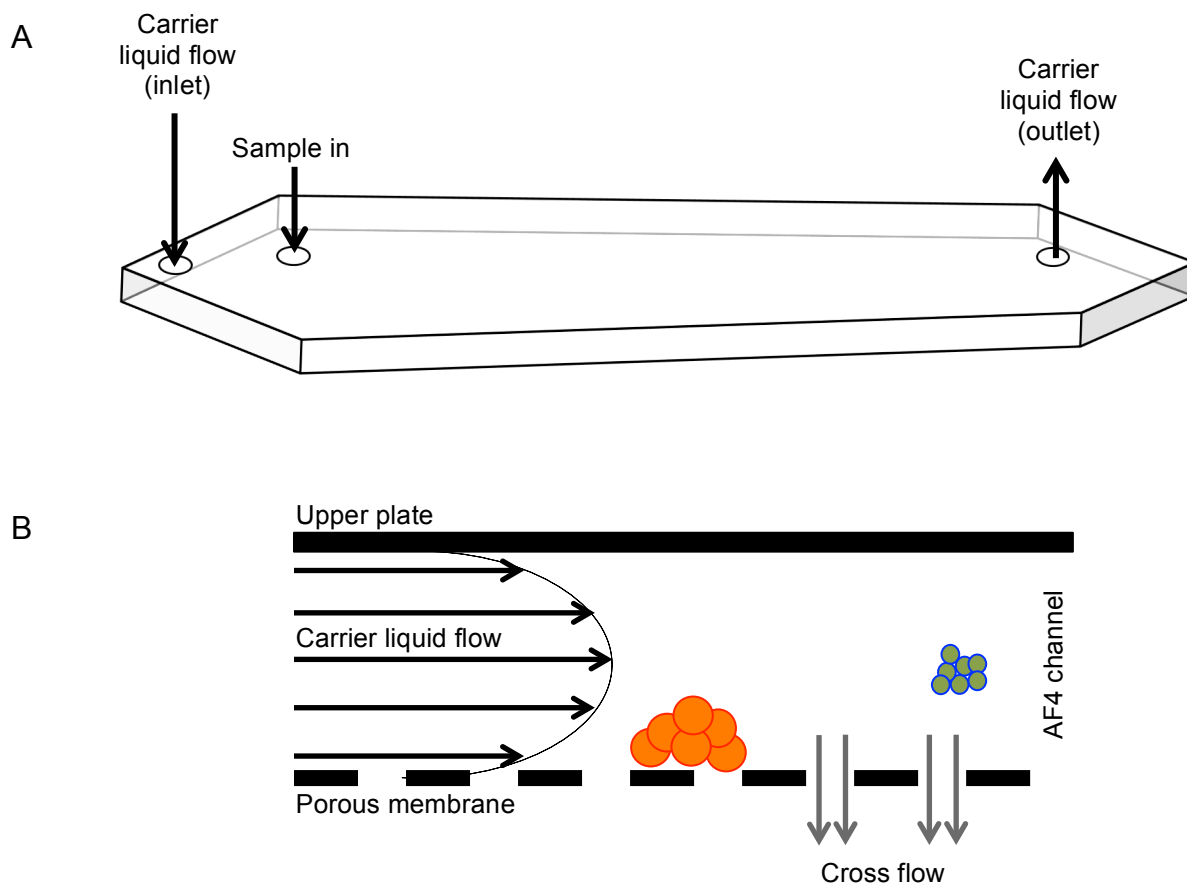


Figure 4-4. Principle of separation using asymmetric flow-field flow fractionation (AF4). (A) Trapezoidal channel used in AF4, and (B) Schematic diagram to illustrate the directions of carrier liquid flow and cross flow inside an AF4 channel. Cross flow forces particles toward the wall of a porous membrane, and driven by concentration different, the particles diffuse back toward the centre of the channel. The figure was reproduced and modified from (Chuan et al. 2008).

AF4-MALS analysis of assembled modular VP1 bearing three, four, and five copies of H190 is shown in Figure 4-5. Figure 4-5A shows construct VP1-H190-3x forming modular VLPs although a small fraction of unassembled capsomeres was observed. In contrast, Figure 4-5B shows construct VP1-H190-4x yielding mostly unassembled capsomeres with a small proportion of VLPs formed. Figure 4-5C shows construct VP1-H190-5x yielded a negligible fraction of VLPs. The proportion of VLPs to unassembled capsomeres in construct VP1-H190-5x was lower than the proportion in construct VP1-H190-4x. This result shows that as the number of H190 tandem repeat increased, the ability of modular VP1 capsomeres to assemble into VLPs decreased.

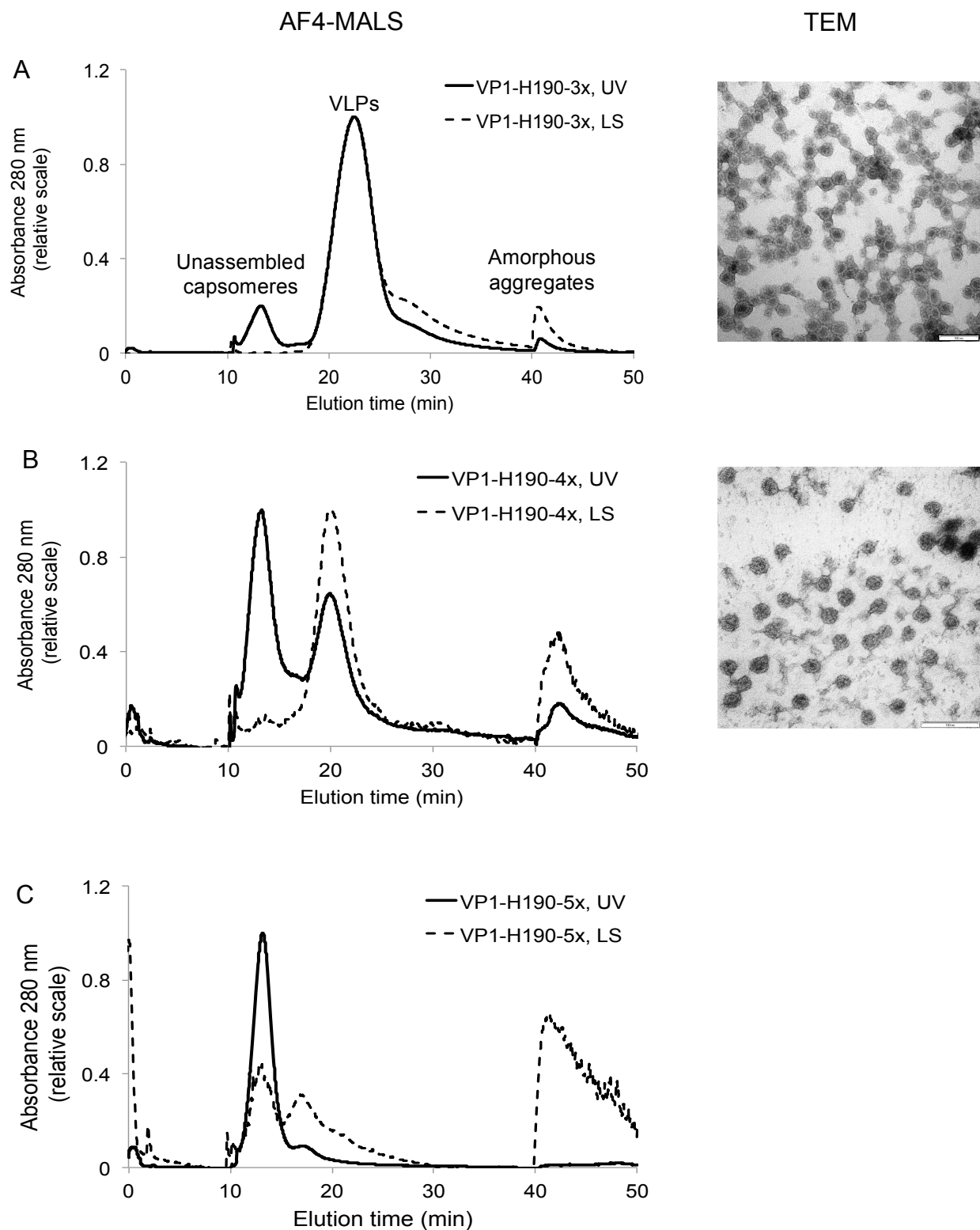


Figure 4-5. Characterisation of modular VP1 bearing three, four, and five copies of H190 after assembly using AF4-MALS and TEM. (A) VP1-H190-3x, (B) VP1-H190-4x, and (C) VP1-H190-5x. Peaks containing unassembled capsomeres, VLPs, and amorphous aggregates eluted at about 10 min, 20 min, and 40 min, respectively. UV absorbance at 280 nm (UV) and light scattering signals (LS) were shown on a relative scale for AF4-MALS analysis. Bars represent 100 nm.

Results in Table 4-3, Figure 4-3, and Figure 4-5 show that there were two difficulties for the assembly of modular VP1 bearing four and five copies of H190 under the chosen buffer conditions that had been previously optimised for wt-VP1. These problems were: (i) reduced solubility, and (ii) reduced capability of assembly.

Like many protein phenomena, protein solubility and formation of aggregates are driven by protein-protein interactions mediated by salt ions and water (Curtis et al. 2002a). Generally, in a dilute aqueous salt solution (< 0.1 M), protein-protein interactions depend on the favourable interaction between charges on the surface of proteins and surrounding ions (Chiew et al. 1995, Curtis et al. 1998). The net surface charge of proteins for a given primary sequence, or the zeta potential of a protein, is determined by pH. As pH increases away from the isoelectric point (pI), where proteins have neutral surface charges and do not migrate in electric field, proteins become net-negatively charged. Conversely, as pH decreases away from pI, proteins become net-positively charged (Yang and Honig 1993). Therefore pH is the strongest factor that influences protein-protein interaction, for a given amino acid sequence, in aqueous salt solution (Curtis et al. 1998).

Protein-protein interactions can be measured using various techniques, such as osmometry, sedimentation, and static-light scattering. These techniques yield the parameter for protein-protein interaction that is a protein-protein osmotic second virial coefficient (B_{22}) (Curtis et al. 1998, Curtis et al. 2002a). The B_{22} is a parameter to assess the thermodynamics of protein solutions. It reflects the direction and magnitude of protein-protein interactions (Prausnitz 2003). Furthermore, the coefficient is the sum of the potentials of mean force (pmf). Negative derivative of pmf with respect to distance (r) reflects the force between two solute molecules at infinite dilution (McMillan Jr and Mayer 2004). Because the function of pmf over r is negative, B_{22} is modelled as a function of W_{22} , that is the free energy of two protein molecules as a function of the centre-to-centre separation (Eq. 1 and 2). The assumption of this equation is that W_{22} is spherically symmetric.

$$B_{22} = N_{AV} B'_{22} / M_2^2 \quad (1)$$

$$B'_{22} = \frac{1}{2} \int_0^{\infty} [1 - e^{\beta W_{22}}] 4\pi r^2 dr \quad (2)$$

where $\beta = (k_b T)^{-1}$, k_b is Boltzmann's constant, and T is absolute temperature (Curtis et al. 1998, Curtis et al. 2002a).

In aqueous salt solution, W_{22} is modelled using DLVO theory, named after the scientists Deryaguin and Landau, and Verwey and Overbeek. In this theory, proteins are

treated as a rigid sphere with uniform surface charge. The sphere is assumed to be in a continuous dielectric medium containing salt ions, which behave as point charges (Deryaguin and Landau 1941, Verwey and Overbeek 1947). Using this approach, W_{22} is the sum of: (i) the hard-sphere potential, (ii) the electric double-layer repulsion, and (iii) the attractive van der Waals forces.

$$W_{22} = W_{hs}(r) + W_{disp}(r) + W_{elec}(r) \quad (3)$$

where W_{hs} is the hard-sphere potential, W_{disp} is the attractive van der Waals force, and W_{elec} is the electric double-layer repulsion (Curtis et al. 1998, Curtis et al. 2002a). Approximation for each force is explained as follows.

The hard-sphere potential determines the excluded volume of a protein. The simplest form of hard-sphere potential is given by

$$W_{hs}(r) = \begin{cases} \infty & \text{for } r \leq d_2 + 2\sigma \\ 0 & \text{for } r > d_2 + 2\sigma \end{cases} \quad (4)$$

Eq. 4 shows two variables that determine the hard-sphere potential of a protein. These variables are d_2 and σ . d_2 is defined as the effective spherical diameter, and can be calculated from the crystal structure dimension of the protein. It equals to the distance of the closest approach between two protein molecules. σ is thickness of the impenetrable layer of water. It is assumed constant and independent of pH and ionic strength (Curtis et al. 1998, Curtis et al. 2002b). Both variables in Eq. 4 are not a function of pH. Consequently, the hard-sphere potential is not affected by pH.

Similarly, variables determining the attractive van der Waals force are pH independent. The attractive van der Waals force is treated using a linear Lifshitz theory (Sabisky and Anderson 1973), which is usually written in a form with Hamaker constant.

$$W_{disp}(r) = -\frac{H}{12} \left\{ \frac{d_2^2}{r^2 - d_2^2} + \frac{d_2^2}{r^2} + 2 \ln \left(1 - \frac{d_2^2}{r^2} \right) \right\} \text{ for } r \geq (d_2 + 2\sigma) \quad (6)$$

H is the effective Hamaker constant, which is affected by the composition and density of the protein, and chemical nature of the solute. Furthermore, Eq. 5 shows that the attractive van der Waals is independent of ionic strength. This independency is because the correlation time required for ions to reach equilibrium between atoms is negligible.

In contrast, pH is one of variables determining the electric double-layer repulsion, which is approximated using non-linear Poisson-Boltzmann theory (Outhwaite and Bhuiyan 1983). If the concentration of electrolyte is very low, less than about 100 mM, and therefore salt ions behave as point charges, the Poisson-Boltzmann equation can be simplified into the Debye-Hückle equation (Verwey and Overbeek 1947). The equation is given by

$$W_{elec}(r) = \frac{z^2 e^2 (1/r) \exp[-\kappa(r-d_2)]}{4\pi\epsilon_0\epsilon_r(1+\kappa d_2/2)^2} \text{ for } r > (d_2 + 2\sigma) \quad (5)$$

where z is the valence of the protein, e is the elementary charge, $4\pi\epsilon_0$ is dielectric permittivity of free space, and ϵ_r is the relative dielectric permittivity of water. Furthermore, κ is the inverse of the Debye length, given by

$$\kappa^2 = (2e^2 N_A I) / (kT \epsilon_0 \epsilon_r) \quad (6)$$

with I is the ionic strength of salt, which is determined by both molar concentration of salt ions and the charge number of salt ions. Additionally, I is also a function of pH, given by the extended Debye-Hückle equation (Eq. 7-8) (Baumgarten 1981).

$$\text{pH} = -\log \gamma_{H^+} [H^+] \quad (7)$$

$$\log \gamma_{H^+} = \frac{-Az_i^2(\sqrt{I})}{1+(B\alpha\sqrt{I})} \quad (8)$$

where γ_{H^+} is the activity coefficient of ion hydrogen, $[H^+]$ is the concentration of ion hydrogen that is 1×10^{-7} , z in Eq. 8 is the charge number of hydrogen ion, I is the ionic strength of salt, A and B are the Debye-Hückle constants, and α is the effective diameter of ion hydrogen. As can be seen from Eq. 5 to 8, changes in pH of solution will affect the ionic strength of the solution at a given type and concentration of salts. Changes in ionic strength will affect the Debye length and consequently the electric double-layer repulsion. Therefore, the importance of the electric double-layer repulsion increases with increasing or decreasing pH away from the pI (Velev et al. 1998).

Different from aqueous salt solution, in concentrated salt solution, such as in biological solutions, protein-protein interactions occur due to unfavourable interaction between salt ions and non-polar surface of proteins, and favourable weak ion binding interaction between salt and charge surface groups (Curtis et al. 2002a). In this solution,

the range of the electric double-layer repulsion is strongly suppressed and reduced, hence the dispersion force becomes dominant (Bostrom et al. 2001). Protein-protein interactions no longer depend on net charge and pH (Curtis et al. 1998); they depend on the specific ion effects.

The specific ion effects are determined by the position of anions and cations in the lyotropic (Hofmeister) series (Curtis et al. 2002a). The effects from anions are more noticeable than from cations. The Hofmeister series were originally developed to explain the different concentrations of neutral salts required to precipitate a protein from solution (Hofmeister 1888). In the series, ions are positioned according to their surface tension, which is the ability of the ions to affect protein stability indirectly by changing the hydrogen-bonding properties of water (Baldwin 1996). Figure 4-6 shows the Hofmeister series for anions and cations in decreasing order of the surface-tension increment. Salts with high surface tension are known as high lyotropics, or kosmotropes, while salts with low surface tension are called low lyotropics, or chaotropes (Curtis et al. 2002a).

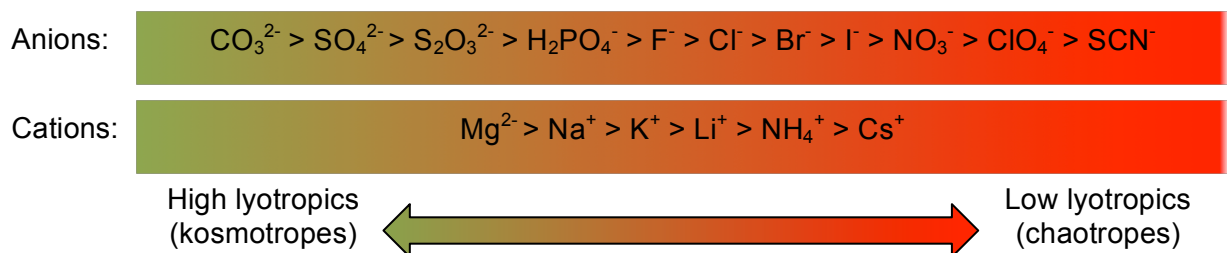


Figure 4-6. The Hofmeister series of anions and cations.

Kosmotropes increase the surface tension between water and proteins by interacting strongly with water. They make water molecules around the salt ions more structured than bulk water, and consequently promote salting-out effects where solubility of proteins is reduced. This behaviour of kosmotropes is in contrast to chaotropes. Chaotropes decrease the surface tension between water and proteins because they interact weakly with water. They break the structure of water molecules and promote salting-in effects where solubility of proteins is increased (Zhang and Cremer 2006).

DLVO theory fails to predict the impact of the specific ion effects because it treats salt ions as point charges. By doing so, DLVO theory could not differentiate the size and

type of ions with the same charges. Therefore, DLVO theory could only be applied at low salt concentration where electrostatics dominates (Bostrom et al. 2001).

Protein-protein interactions have been used to explain the importance of pH and salt on the spontaneous assembly of MuPyV VP1 capsomeres to form VLPs. The study by Salunke et al. (Salunke et al. 1989) showed that reducing the pH of assembly buffer (10 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, 15 mM 2-ME, 5% (v/v) glycerol) from 7.2 to 5 promoted assembly of the VP1 capsomeres to form VLPs. In contrast, increasing pH from 7.2 to 8 did not result in the formation of VLPs. In addition to pH, the assembly is also affected by other factors, i.e. the concentration of calcium and $(\text{NH}_4)_2\text{SO}_4$. In the absence of pH adjustment, addition of 0.5 mM CaCl_2 or 2 M $(\text{NH}_4)_2\text{SO}_4$ promoted the formation of polymorphic aggregates of the VP1 capsomeres.

Chuan et al. (Chuan et al. 2010) investigated how these three parameters, i.e. pH, calcium concentration, and $(\text{NH}_4)_2\text{SO}_4$ concentration, can play role the formation of VLPs by linking the assembly process to protein-protein interactions. In the study, Chuan et al. measured the osmotic second virial coefficient (B_{22}) of MuPyV VP1 capsomeres during assembly process as a function of these three parameters. The study by Chuan et al. showed that reducing the pH of L buffer (40 mM Tris or Bis-tris, 200 mM NaCl, 1 mM EDTA, 5 mM DTT, 5% (v/v) glycerol) from pH 7 to 6 resulted in considerable reduction in B_{22} of MuPyV VP1 capsomeres, which indicates increased attractive interaction between capsomeres to promote assembly of capsomeres to form VLPs. Similar effects on B_{22} were also observed by increasing the concentration of calcium and $(\text{NH}_4)_2\text{SO}_4$. Based on the findings by Chuan et. al., the concentrations of both calcium and $(\text{NH}_4)_2\text{SO}_4$ in Assembly Buffer 1 in this chapter were sufficient to maintain the value of B_{22} at lower than $-20 \times 10^{-4} \text{ mol ml g}^{-2}$. This value is considerably lower than $-8 \times 10^{-4} \text{ mol ml g}^{-2}$, the limit of B_{22} below which protein-protein interactions are classified as a strong attraction (Chuan et al. 2010, Curtis et al. 1998, Curtis et al. 2002a, George and Wilson 1994). The low value of B_{22} suggested that both concentrations of calcium and $(\text{NH}_4)_2\text{SO}_4$ in Assembly Buffer 1 in this chapter had been sufficient to promote strong attractions between capsomeres, and consequently self-assembly of the capsomeres to form modular VLPs. In contrast, at the pH of Assembly Buffer 1 the value of B_{22} was approximately $-2.5 \times 10^{-4} \text{ mol ml g}^{-2}$ (Chuan et al. 2010). This value is higher than the limit of B_{22} for a strong attraction, which suggests that the pH was not sufficient to promote a strong attraction. This analysis indicated that when VLPs are obtained in Assembly Buffer 1, such as observed with wt-VP1, the assembly process is driven by strong attractions resulting from the concentrations of both

calcium and $(\text{NH}_4)_2\text{SO}_4$, rather than from pH. Furthermore, the analysis suggested that a stronger attraction may be achievable via pH optimisation of Assembly Buffer 1.

The Assembly Buffer 1 used in this thesis contained 500 mM $(\text{NH}_4)_2\text{SO}_4$, thus it represented a concentrated salt solution. Because modular VLPs bearing four and five copies could not be obtained when assembly was performed in Assembly Buffer 1, this section investigated the importance of pH and specific ion effects on the assembly and precipitation of VP1-H190-4x and VP1-H190-5x. The importance of pH was investigated by conducting assembly at various pH values, between pH 7.4 and 5.4. Tris-base was replaced with sodium phosphate to maintain the pH of assembly buffers at pH below 7.4. The importance of specific ion effects was investigated by conducting assembly at two conditions, i.e. (i) different concentrations of $(\text{NH}_4)_2\text{SO}_4$ that were 200 mM, 350 mM, and 500 mM; and (ii) different types of salts that were 500 mM NaCl and 500 mM $(\text{NH}_4)_2\text{SO}_4$.

SDS-PAGE and AF4-MALS analysis of VP1-H190-4x and VP1-H190-5x assembled at various pH values are shown in Figure 4-7 and Figure 4-8, respectively. SDS-PAGE analysis in Figure 4-7A showed that at all tested pH values, modular protein VP1-H190-4x precipitated considerably after assembly. There was a slight improvement in the solubility of the modular protein when the pH was reduced from 7.4 to 6. Further pH reduction resulted in decreases in the solubility of the modular protein. The result shows that the pH of Assembly Buffer 1 weakly affected the precipitation of modular VP1 bearing four copies of H190. In contrast, AF4-MALS analysis in Figure 4-7B-F showed that pH of Assembly Buffer 1 influenced considerably the proportion of formed VLPs to capsomeres for VP1-H190-4x. Reduction in the pH of Assembly Buffer 1 from 7.4 to 6 increased the proportion of VLPs to unassembled capsomeres. Further pH reduction to 5.4 resulted in decreased proportion of VLPs to unassembled capsomeres. Thus, for VP1-H190-4x, the optimum proportion of VLPs to capsomeres was obtained at pH 6.

Unlike construct VP1-H190-4x, precipitation of modular protein VP1-H190-5x was not affected by pH of the Assembly Buffer 1. SDS-PAGE analysis in Figure 4-8A showed that, at all tested pH values modular protein VP1-H190-5x precipitated significantly. Similarly, pH of Assembly Buffer 1 did not influence the assembly capability of modular capsomeres VP1-H190-5x. AF4-MALS analysis (Figure 4-8B-F) of assembled protein for VP1-H190-5x showed that at all tested pH values, modular VLPs could not be obtained, using the chosen salt conditions.

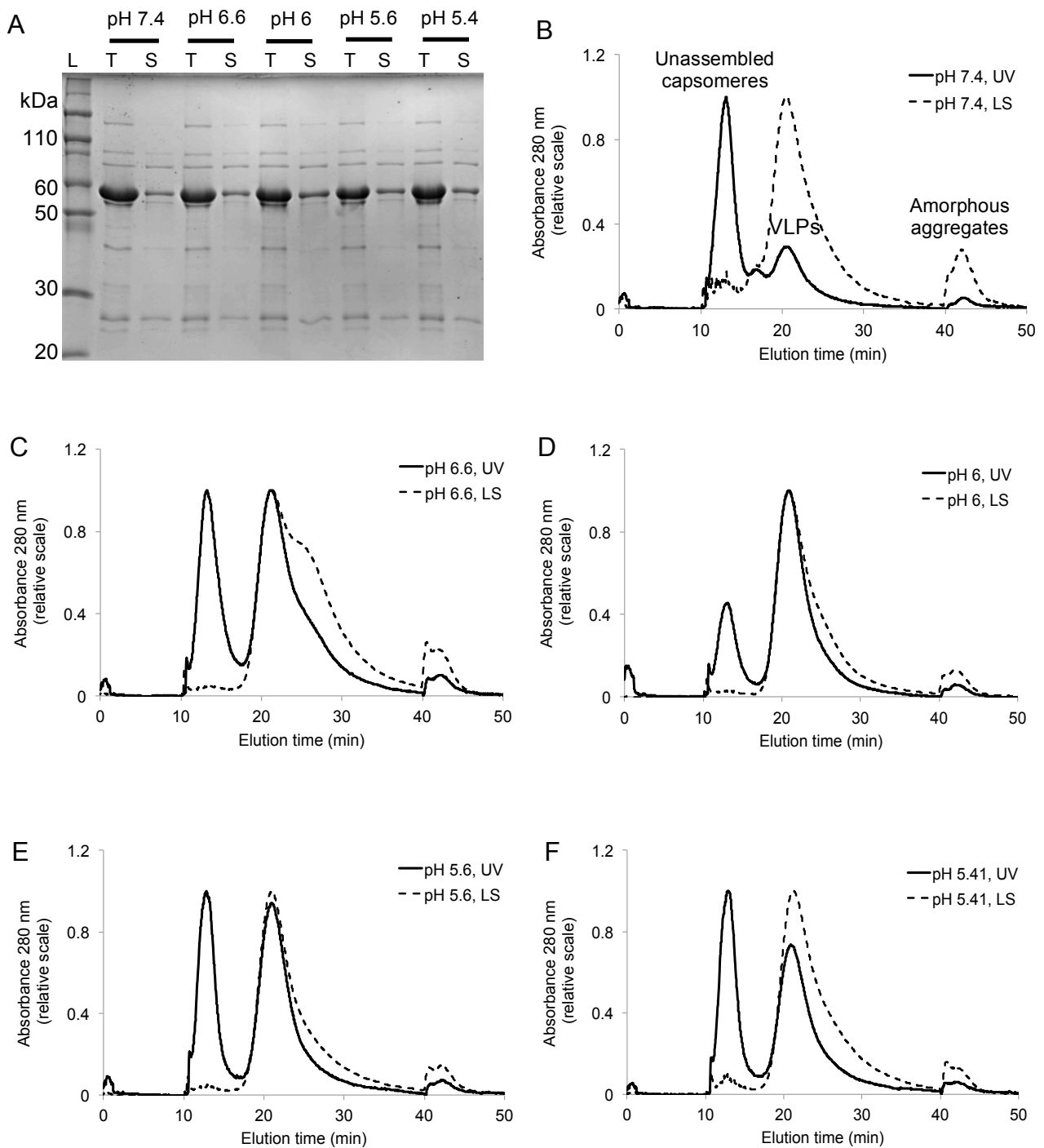


Figure 4-7. Analysis of modular VP1 bearing four copies of H190 after assembly at various pH. (A) Analysis using SDS-PAGE. Lanes: (L) Pre-stained protein marker, (T) Total protein after dialysis against PBS, (S) Soluble protein after dialysis against PBS. The soluble protein after dialysis against PBS was analysed using AF4-MALS. The dominant bands on the gel represent VP1-H190-4x; and (B-F) Analysis using AF4-MALS. UV absorbance at 280 nm (UV) and light scattering signals (LS) were shown on a relative scale for AF4-MALS analysis. Peaks containing unassembled capsomeres, VLPs, and amorphous aggregates eluted at about 10 min, 20 min, and 40 min, respectively.

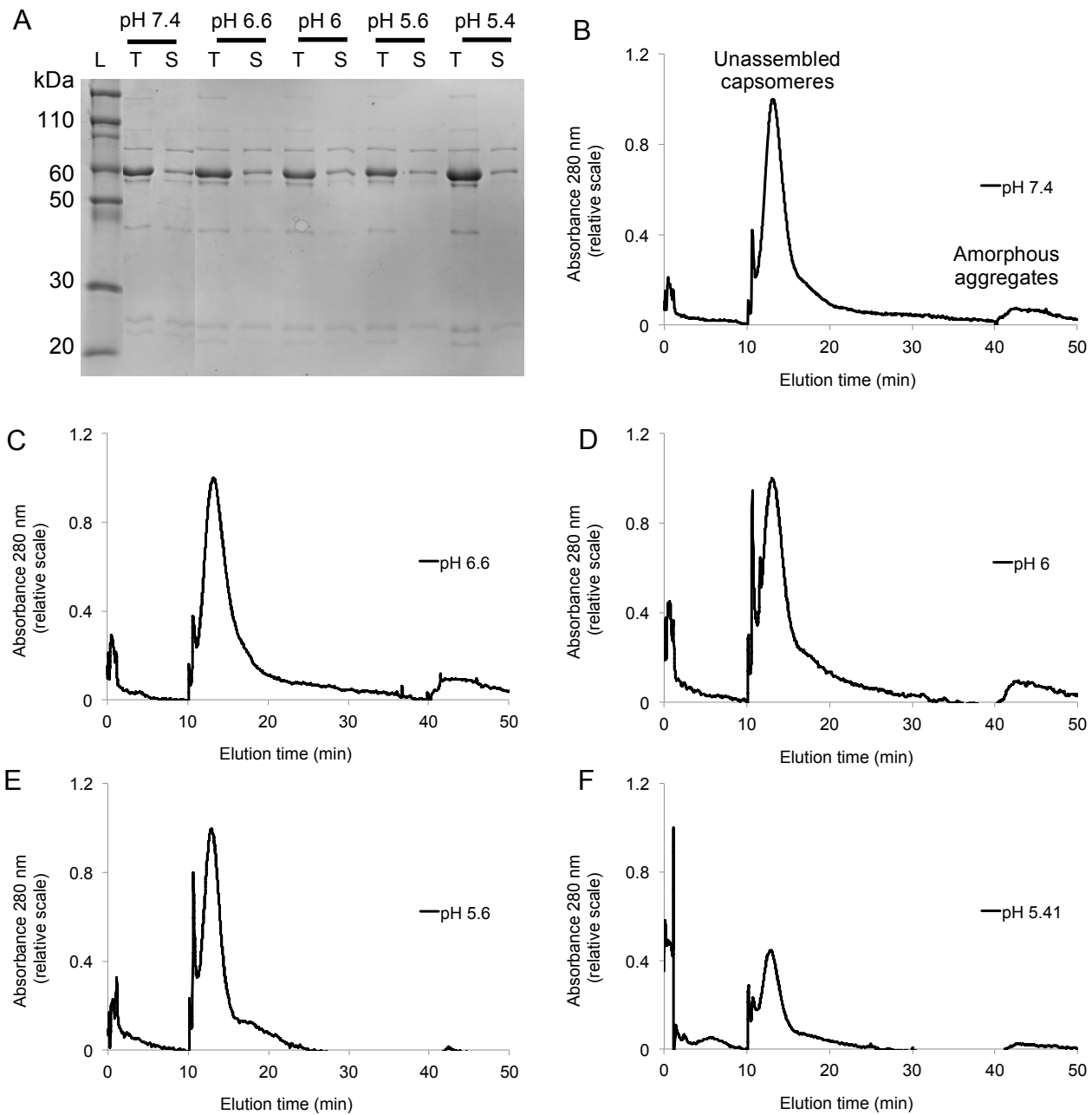


Figure 4-8. Analysis of modular VP1 bearing five copies of H190 after assembly at various pH. (A) Analysis using SDS-PAGE. Lanes: (L) Pre-stained protein marker, (T) Total protein after dialysis against PBS, (S) Soluble protein after dialysis against PBS. The soluble protein after dialysis against PBS was analysed using AF4-MALS. The dominant bands on the gel represent VP1-H190-5x; and (B-F) Analysis using AF4-MALS. UV absorbance at 280 nm (UV) was shown on a relative scale for AF4-MALS analysis. Peaks containing unassembled capsomeres, VLPs, and amorphous aggregates eluted at about 10 min, 20 min, and 40 min, respectively.

Both Figure 4-7 and Figure 4-8 showed how pH changes affected the precipitation and assembly of modular VP1 bearing four. The observation in both figures could be explained using the principal of protein-protein interactions in concentrated salt solution explained above. However, the effect of pH on the assembly of VP1-H190-4x was not in agreement with what would be predicted based on protein interaction literature. The observation showed that pH affected the assembly of the modular VP1 considerably, indicating that electrostatic interaction still played a role during assembly. The optimum pH for the assembly was achieved at pH 6. This pH was close to the theoretical pI of the protein, which is predicted to be 6.27 based on the amino acid sequence of the modular VP1 using ProtParam Tool (Gasteiger et al. 2005). At pH near the pI, modular VP1 bearing four copies of H190 might have neutral net surface charge, which may enhance the capsomere-capsomere interactions and promote formation of VLPs. An alternative explanation is that the role of electrostatic interaction was manifested in how the pH changes affected the C-terminal region of VP1 protein. C-terminal arm of the VP1 protein is involved in the assembly process and sensitive to electrostatic switching. When calcium ions are added to assembly buffer, they bind to the C-terminal arm, reducing electric double-layer repulsion of the arm and promoting assembly (Salunke et al. 1989). In this chapter, modular VLPs bearing four copies of H190 could be obtained by reducing pH of Assembly Buffer 1. The pH reduction leads to the addition of protons, which have a similar charge to calcium. Therefore, it is possible that added protons bound to the C-terminal arm of VP1 in a similar way to calcium ions, shielding electric double-layer repulsion of the arm and subsequently promoting self-assembly of modular VP1 capsomeres (Salunke et al. 1989).

The investigation on the importance of specific ion effects was performed using modular VP1 bearing four copies of H190. Modular capsomeres VP1-H190-4x were assembled in Assembly Buffer containing (i) different concentrations of $(\text{NH}_4)_2\text{SO}_4$, and (ii) different salt type. The effect of different concentration of $(\text{NH}_4)_2\text{SO}_4$ was determined by performing assembly at 200 mM and 350 mM $(\text{NH}_4)_2\text{SO}_4$. The effects of different type of salt were determined by performing assembly at 500 mM $(\text{NH}_4)_2\text{SO}_4$ or NaCl. pH of the assembly buffers was maintained at 6 as results in Figure 4-7B-F show that the optimum ratio of VLPs to unassembled capsomeres could be obtained at this pH.

SDS-PAGE and AF4-MALS analysis of modular capsomeres VP1-H190-4x assembled at different concentrations of $(\text{NH}_4)_2\text{SO}_4$ are shown in Figure 4-9. The SDS-PAGE analysis (Figure 4-9A) showed that assembly of VP1-H190-4x in Assembly Buffer 1 containing 200 mM and 350 mM $(\text{NH}_4)_2\text{SO}_4$ resulted in considerably less precipitation

compared to assembly at 500 mM $(\text{NH}_4)_2\text{SO}_4$. This finding showed that the concentration of $(\text{NH}_4)_2\text{SO}_4$ in Assembly Buffer 1 determined the level of precipitation of VP1-H190-4x. However, analysis of the assembled VP1-H190-4x using AF4-MALS (Figure 4-9B-C) showed that VP1-H190-4x VLPs assembled in Assembly buffer containing 200 mM and 350 mM $(\text{NH}_4)_2\text{SO}_4$ were highly heterogeneous in size as indicated by the broadness of the VLP peaks. The VLP peak for 200 mM $(\text{NH}_4)_2\text{SO}_4$ was broader than the peak for 350 mM $(\text{NH}_4)_2\text{SO}_4$, indicating that the modular VLPs assembled in 200 mM $(\text{NH}_4)_2\text{SO}_4$ were more heterogeneous. Moreover, the VLP peak for 200 mM $(\text{NH}_4)_2\text{SO}_4$ eluted at around 35 min, later than the peak for 350 mM $(\text{NH}_4)_2\text{SO}_4$ which eluted at around 28 min. This result indicates that the VLPs assembled at lower concentrations of $(\text{NH}_4)_2\text{SO}_4$ had bigger size. The AF4-MALS analysis results also showed that at both 200 mM and 350 mM $(\text{NH}_4)_2\text{SO}_4$ the amount of unassembled capsomeres was similar to the amount of VLPs. Furthermore, the proportion of soluble aggregates to VLPs for 200 mM $(\text{NH}_4)_2\text{SO}_4$ was higher than the proportion for 350 mM $(\text{NH}_4)_2\text{SO}_4$. This finding suggested that a lowered concentration of $(\text{NH}_4)_2\text{SO}_4$ promoted formation of soluble aggregates.

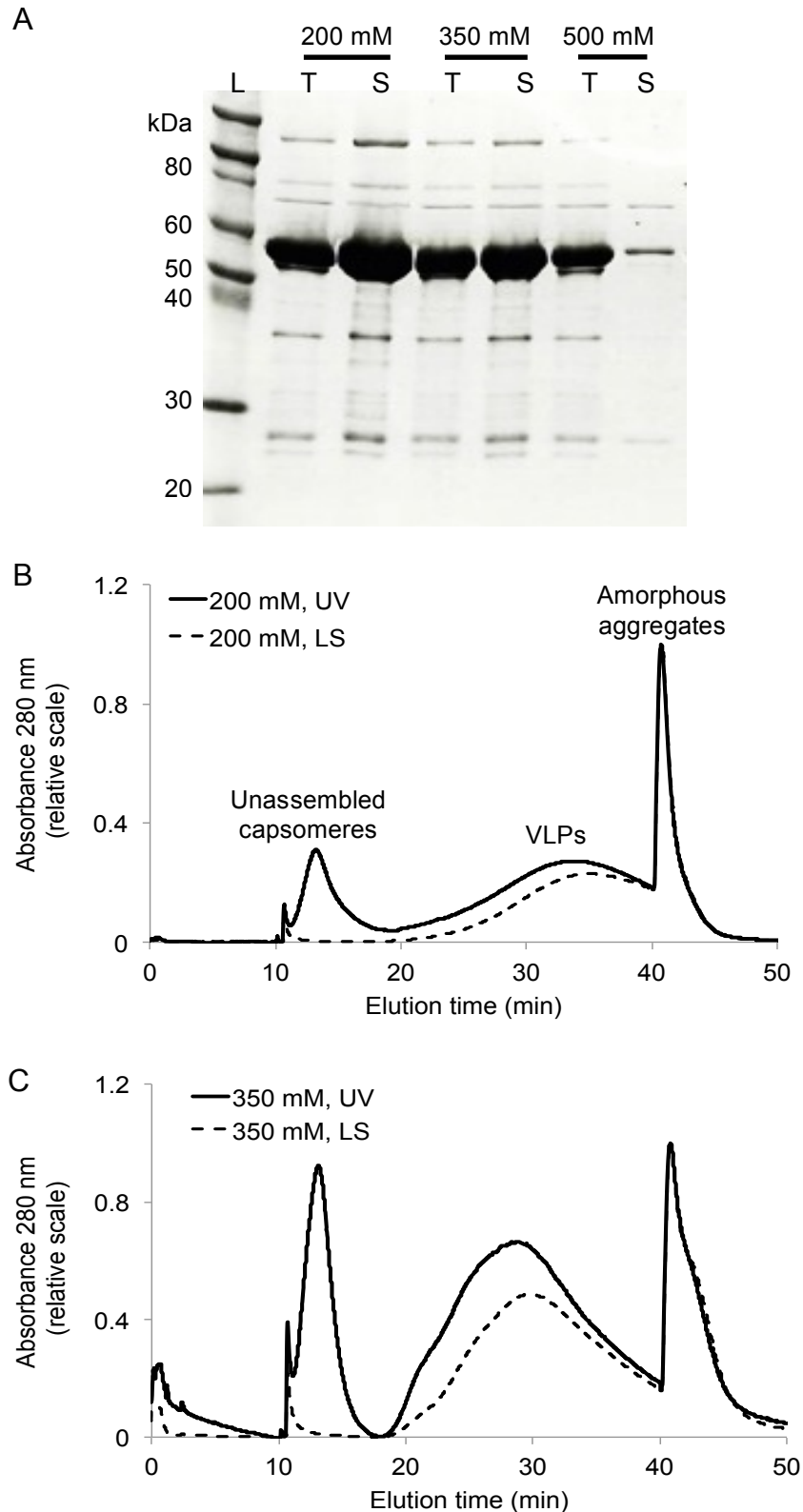


Figure 4-9. Analysis of the effects of $(\text{NH}_4)_2\text{SO}_4$ concentration in Assembly Buffer 1 on the solubility and assembly of modular VP1 bearing four copies of H190. (A) Analysis using SDS-PAGE. Lanes: (L) Pre-stained protein marker, (T) Total protein after dialysis against PBS, (S) Soluble protein after dialysis against PBS. The soluble protein after dialysis against PBS was analysed using AF4-MALS. The dominant bands on the gel represent VP1-H190-4x; and (B-C) Analysis using AF4-MALS. UV absorbance at 280 nm (UV) and light scattering signals (LS) were shown on a relative scale for AF4-MALS analysis. Peaks containing unassembled capsomeres, VLPs, and amorphous aggregates eluted at about 10 min, 20 min, and 40 min, respectively.

Similarly to the impact of $(\text{NH}_4)_2\text{SO}_4$ concentration on the precipitation of VP1-H190-4x, replacement of $(\text{NH}_4)_2\text{SO}_4$ with NaCl also affect the precipitation of the modular VP1 protein. SDS-PAGE analysis in Figure 4-10A showed that when assembly was performed at 500 mM NaCl in place of $(\text{NH}_4)_2\text{SO}_4$, solubility of VP1-H190-4x was considerably higher. Furthermore, AF4-MALS analysis in Figure 4-10B-C showed that VP1-H190-4x VLPs assembled at 500 mM NaCl were more heterogeneous than those assembled at the same concentration of $(\text{NH}_4)_2\text{SO}_4$. Additionally, VP1-H190-4x assembled in NaCl contained less VLPs and higher unassembled capsomeres, which were at similar amount to soluble aggregates. These findings indicated that the use of NaCl rather than $(\text{NH}_4)_2\text{SO}_4$ prevented the formation of VLPs.

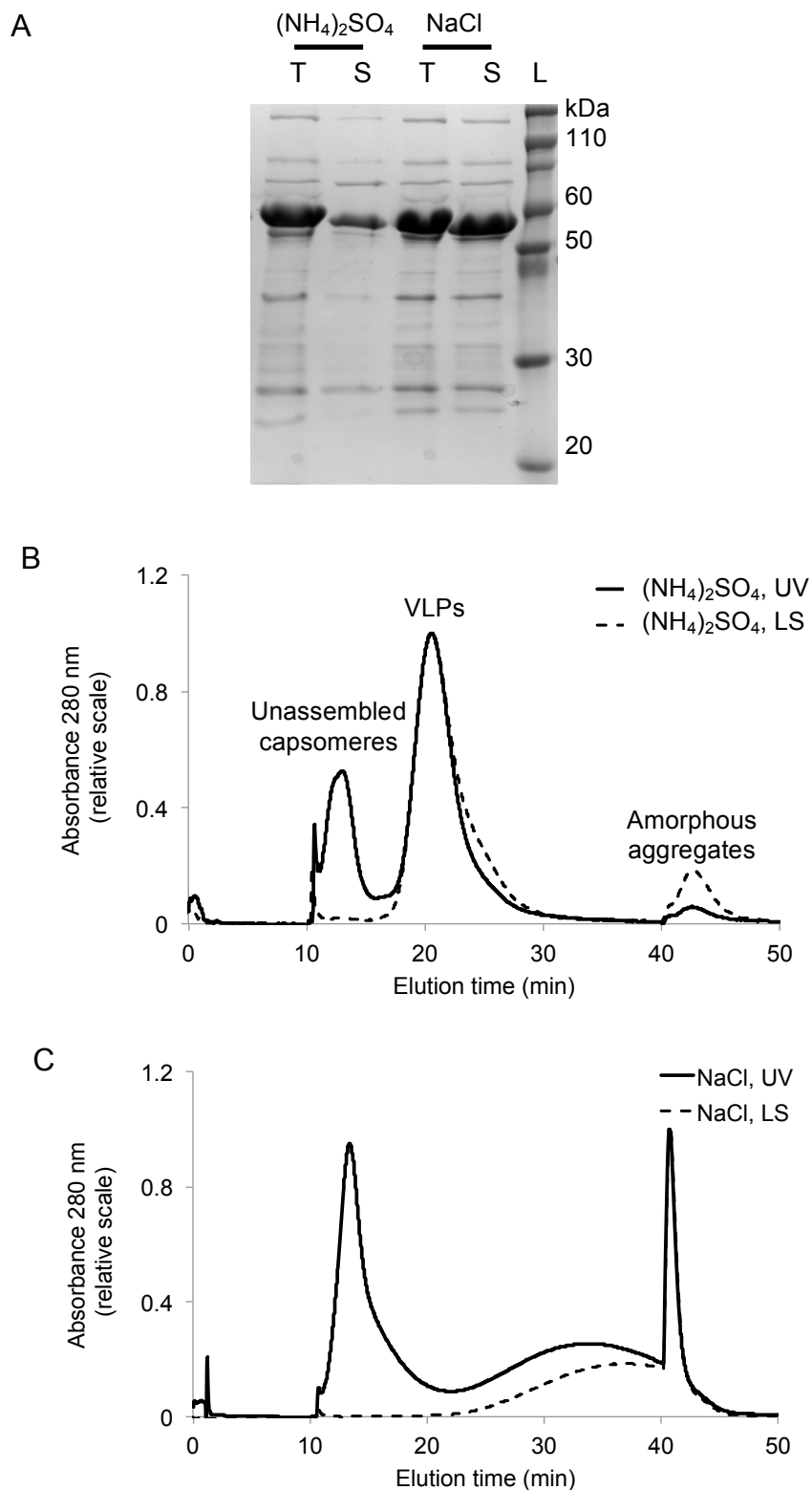


Figure 4-10. Analysis of the effects of salt type in Assembly Buffer 1 on the solubility and assembly of modular VP1 bearing four copies of H190. (A) Analysis using SDS-PAGE. Lanes: (L) Pre-stained protein marker, (T) Total protein after dialysis against PBS, (S) Soluble protein after dialysis against PBS. The soluble protein after dialysis against PBS was analysed using AF4-MALS. The dominant bands on the gel represent VP1-H190-4x; and (B-C) Analysis using AF4-MALS. UV absorbance at 280 nm (UV) and light scattering signals (LS) were shown on a relative scale for AF4-MALS analysis. Peaks containing unassembled capsomeres, VLPs, and amorphous aggregates eluted at about 10 min, 20 min, and 40 min, respectively.

The impacts of salt concentration and type on the solubility and assembly capability of VP1-H190-4x observed in Figure 4-9 and Figure 4-10 were possibly linked to the interactions of protein with ions. Both concentration and type of the salt have been shown to affect the surface tension of water. Increasing the concentration of salts increases the surface tension, and the type of the salts determines the extent of the surface tension increase. Both SO_4^{2-} and Cl^- are categorised as Hofmeister ions (Zhang and Cremer 2006), which affect protein stability indirectly by changing the hydrogen-bonding properties of water. SO_4^{2-} and Cl^- affect the surface tension of waters in different ways. The anion SO_4^{2-} is a kosmotrope, it interacts strongly with water and increases the surface tension between water and proteins. Consequently, water molecules around the salt ions are more structured than bulk water. Such changes in the structure of water molecule promote salting-out effects where solubility of proteins is reduced. In contrast, the anion Cl^- is an example of chaotropes. It interacts weakly with water and decreases the surface tension between water and proteins. Consequently, it breaks the structure of water molecules and promotes salting-in effects where solubility of proteins is increased (Zhang and Cremer 2006). The level of water surface tension have been widely accepted as explanations for changes in protein solubility in salt solution via cavity model (Baldwin 1996).

The results above show that, at an assembly condition previously optimised for wt-VP1, (i) solubility and (ii) assembly capability of the modular VP1 decreased as number of tandem repeat H190 increased. In this chapter, these two difficulties were linked to two important factors of protein-protein interactions, i.e. (i) the net surface charge for a given amino acid sequence (the zeta potential), and (ii) specific ion effects. For modular VP1 bearing four copies of H190, the net surface charge was not the major driver for the solubility of the modular VP1 because pH adjustment of the Assembly Buffer 1 weakly affected the precipitation of the protein (Figure 4-7A). In contrast, manipulation of the net surface charge of modular VP1 bearing four copies of H190 resulted in changes in the proportion of formed VLPs to unassembled capsomeres. The optimum proportion was achieved at pH 6 (Figure 4-7B-F). Furthermore, the impacts of pH on the solubility and assembly capability observed for modular VP1 bearing four copies of H190 were not observed when the number of H190 tandem repeats was increased to five copies. The net surface charge of modular VP1 bearing five copies of H190 did not affect the solubility and assembly capability of the protein (Figure 4-8). In contrast to how pH solution impacted the solubility of modular VP1 bearing four copies of H190, the results in this section show that specific ion effects, in this case concentration and type of salt, regulated the solubility and assembly capability of modular VP1 bearing four copies of H190 considerably. Reducing

salt concentration (Figure 4-9A) and replacing ammonium sulphate to sodium chloride (Figure 4-10A) in Assembly Buffer 1 increased the solubility of the modular VP1. However, increase in the solubility of the protein was compensated by reduced capability of the modular VP1 to assemble into VLPs and increased size heterogeneity of the formed VLPs (Figure 4-9B-C and Figure 4-10B-C).

The importance of net surface charge (zeta potential) and specific ion effects on the solubility and assembly of both modular VP1 bearing four and five copies of H190 agreed with literature. In concentrated salt solution, such as the Assembly Buffer 1, protein-protein interaction is dependent of the specific ion effects rather than the net surface charge of the proteins. In this case, the specific ion effects were the concentration and type of salt. However, the dependency of assembly capability of modular VP1 bearing four copies of H190 on pH and hence the net surface charge suggested that electric double-layer repulsion were still involved in the assembly process.

It is also important to note that the impacts of specific ion effects on protein-protein interactions in a given solution condition are determined by the amino acid sequence of the protein. Increasing number of H190 results in changes in the amino acid sequence and can consequently affect protein-protein interactions. Additionally, changes in the amino acid can result in structural perturbation of the VP1 protein. As the number of H190 tandem repeats increases, the size of antigen modules increases. Bigger antigen modules can perturb the structural integrity of modular VP1. This could result in exposure of buried residues to the surface, and consequently changes in the amino acid sequence exposed to the surface. Unfortunately, results in this chapter were insufficient to prove this possibility regarding structural perturbation.

4.3.2 H190-specific immunogenicity

The *in vivo* studies reported in this chapter were undertaken in two main groups to investigate (i) the use of AdvaxTM-1 as adjuvant, and (ii) the effect of increasing number of H190 tandem repeats. Groups of mice were immunised on days 0, 21, and 42, and bled on tail on days 0, 14, 35, and 56 (Figure 4-11). Sera from day 56 were used to determine the immunogenicity of H190 by performing immunoassays against peptide H190 as described in Section 4.2.7.

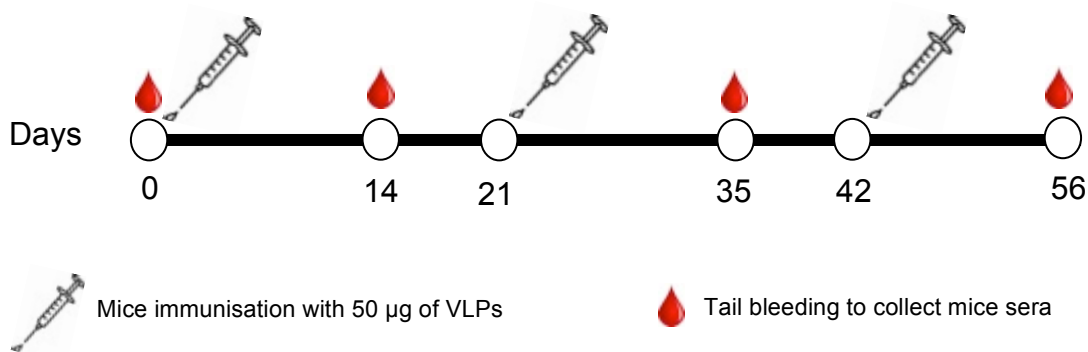


Figure 4-11. Schematic diagram of immunisation and tail bleeding for *in vivo* studies conducted in Chapter 4.

Two types of immunoassays were performed, i.e. (i) indirect ELISA, and (ii) competitive ELISA. In indirect ELISA, mice sera at various dilutions were tested against immobilised peptide H190. Endpoint titres of H190 sequence-specific IgGs were determined as the lowest dilutions of sera with absorbance readings higher than 3 times the standard deviations above the mean absorbance obtained from pre-immunised mouse sera (Relf et al. 1996). Competitive ELISA was performed to confirm the results from indirect ELISA. The binding of IgGs to immobilised peptide H190 was competed with peptide H190 in solution at various concentrations. The competitive ELISA results were presented as the percentage of relative absorbance $[(B/B_0) \times 100\%]$, with B was absorbance 450 nm at a certain concentration of competitor, and B_0 was absorbance 450 nm in the absence of competitor.

4.3.2.1 Effect of AdvaxTM-1 on H190 immunogenicity

The effect of the use of AdvaxTM-1 on immunogenicity of H190 was investigated by performing immunisation using VP1-H190-H190 VLPs with and without AdvaxTM-1. Figure 4-12A shows the endpoint titres of H190-sequence specific IgGs for adjuvanted and non-adjuvanted VLPs. Both adjuvanted and non-adjuvanted VLPs induced H190 sequence-specific IgG titre higher than 10^4 . Furthermore, adjuvanted VLPs induced a slightly higher titre than non-adjuvanted VLPs, although statistical analysis showed that the titre difference was not significantly different ($p= 0.6206$).

Figure 4-12B shows the competitive binding of the IgGs to immobilised peptide H190 by peptide H190 in solution. Similar to result from indirect ELISA in Figure 4-12A, the competitive ELISA result shows that there was a slight difference in H190 sequence-specific IgG titre between adjuvanted and non-adjuvanted VLPs. Peptide H190 at about 240 µM inhibited approximately 20% of IgG binding to HA1 for non-adjuvanted VLPs.

However, the same concentration of peptide H190 only inhibited 40% of IgG binding to immobilised peptide H190 for adjuvanted VLPs. The competitive ELISA result confirms that H190 sequence-specific IgG titre for adjuvanted VLPs was higher than the titre for non-adjuvanted VLPs.

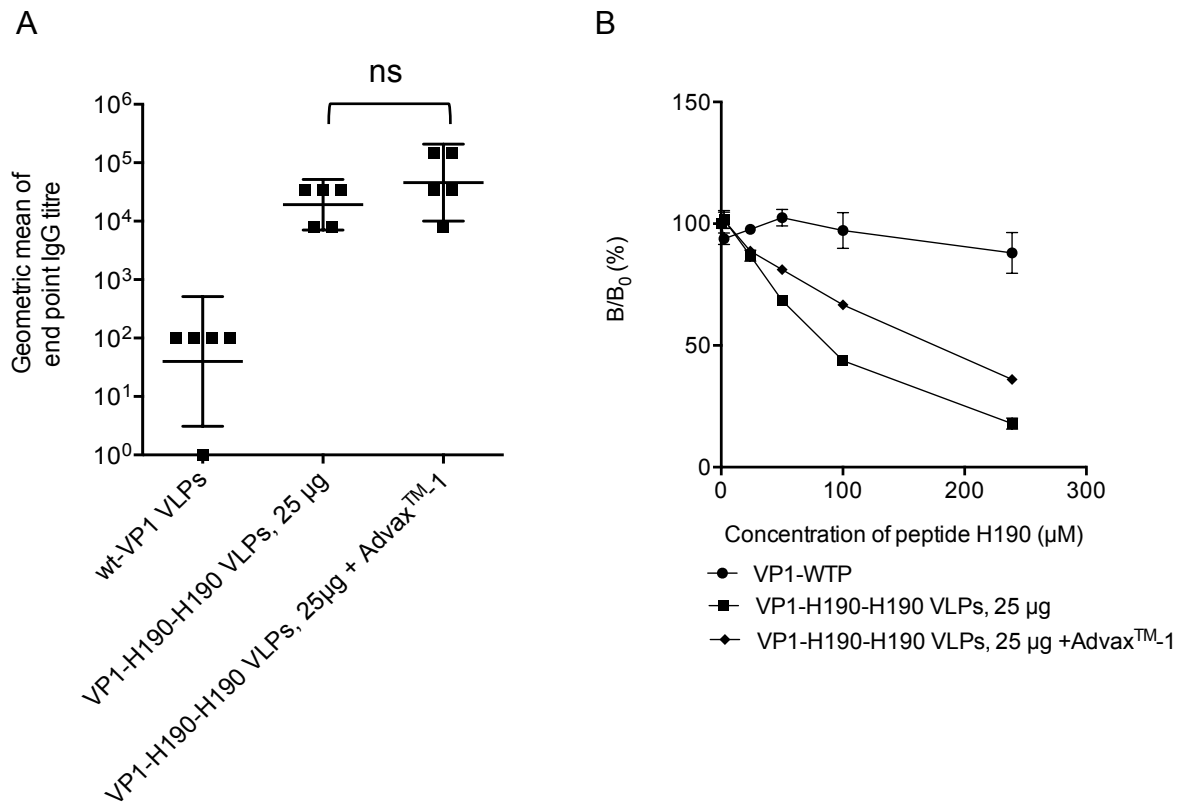


Figure 4-12. Indirect and competitive ELISA results showing the effect of AdvaxTM-1 on the immunogenicity of H190 element. (A) Endpoint titre of H190 sequence specific IgGs induced by modular VLPs. Geometric mean (n=5) and 95% confidence intervals were presented; and (B) Inhibition of antiserum binding (1/1000) to immobilised peptide H190 by peptide H190 in solution. Arithmetic mean (n=5) and standard deviation from duplicate measurements are presented.

Comparison between the result of indirect and competitive ELISA suggested that the IgG titre difference might be deceptively greater in the competitive ELISA. A possible explanation is that the latter assay is more sensitive as has been reported in other studies (Goda et al. 2000, Reddington et al. 1991, Wreghitt et al. 1986). Mice sera used in both assays may contain non-immune IgGs and other serum factors, which can bind non-specifically to the immobilised antigen and consequently cause background interference. Such background can create false-positive reading as well as conceal the low level of IgG binding, reducing the sensitivity of the assay (Saxinger and Gallo 1987).

The novel AdvaxTM-1 is a polysaccharide adjuvant based on particles of β-D-(2→1)poly(fructo-furanosyl)-α-D-glucose (delta inulin) (Saade et al. 2013). It is highly

stable (Cooper and Petrovsky 2011) and has been shown able to have antigen sparing effects through enhancement of antibody titre without skewing the immune response in either Th1 or Th2 direction (Honda-Okubo et al. 2012). The antigen sparing of AdvaxTM-1 is achieved via an increase in total antibody secreting cells in bone marrow and spleen (Saade et al. 2013). In their study, Saade et al. suggested that AdvaxTM-1 assists additional rounds of antigen-specific B-cell proliferation within secondary lymphoid tissues. Alternatively, AdvaxTM-1 confers a survival advantage to antigen-specific B cells (Saade et al. 2013).

AdvaxTM-1 has also been shown to boost immune activity *in vitro* (Cooper and Petrovsky 2011). A study by Honda-Okubo et al. showed that immunisation of influenza vaccine with AdvaxTM-1 enhanced the induction of humoral immune responses IgM, IgG1, and IgG2, as well as CD4 and CD8 T-cell proliferation. Consequently, formulation with AdvaxTM-1 resulted in increased neutralising antibodies and protection level (Honda-Okubo et al. 2012). Similar antigen sparing effect of AdvaxTM-1 was observed for vaccination Japanese encephalitis in mice and horses (Lobigs et al. 2010), HIV in mice (Cristillo et al. 2011), avian (H5N1) influenza in ferrets (Layton et al. 2011), and African Horse Sickness and Glanders in camels (Eckersley et al. 2011).

In good agreement with the aforementioned studies, both results in Figure 4-12 indicate that addition of AdvaxTM-1 increased the immunogenicity of the H190 element in VP1-H190-H190 VLPs, resulting in an increase in H190 sequence-specific IgG titre. Nevertheless, the increase in the IgG titre was not statistically significant. Hence, the findings in this section suggested that the modular VLPs presenting two copies of H190 were self-adjuvanting and able to induce optimum immune responses without the need for adjuvant.

4.3.2.2 Effect of increasing number of H190 tandem repeats on H190 immunogenicity

The effect of increasing number of tandem repeat H190 on immunogenicity of H190 was investigated by comparing four modular constructs containing one to four copies of H190, which were designated as constructs VP1-H190, VP1-H190-H190, VP1-H190-3x, and VP1-H190-4x. Modular construct VP1-H190, as the simplest display strategy that does not incorporate a structural hypothesis, was used as a control. Together with VP1-H190-H190, VP1-H190 was used as a control. Groups of mice were immunised separately with each construct, and antisera were then tested against peptide H190 to determine the presence of H190 sequence-specific IgGs.

Figure 4-13A shows the endpoint titres of H190 sequence-specific IgG for the four modular VLPs. The four modular VLPs were able to induce H190 sequence-specific IgG titre higher than 10^4 . The IgG titre for modular VP1 bearing two to four copies of H190 were similar, and all were slightly higher than the titre for modular VP1 bearing one copy of H190. Statistical analysis showed that IgG titres for VP1-H190 and VP1-H190-H190 were not significantly different ($p= 0.5491$). Furthermore, statistical analysis showed that IgG titre comparisons between VP1-H190-3x VLPs and VP1-H190-H190 VLPs ($p= 0.9919$), and between VP1-H190-4x VLPs and VP1-H190-H190 VLPs ($p= 0.5491$) were not significantly different. Similarly, IgG titre for VP1-H190-3x VLPs and VP1-H190-4x VLPs were not significantly different ($p= 0.9919$).

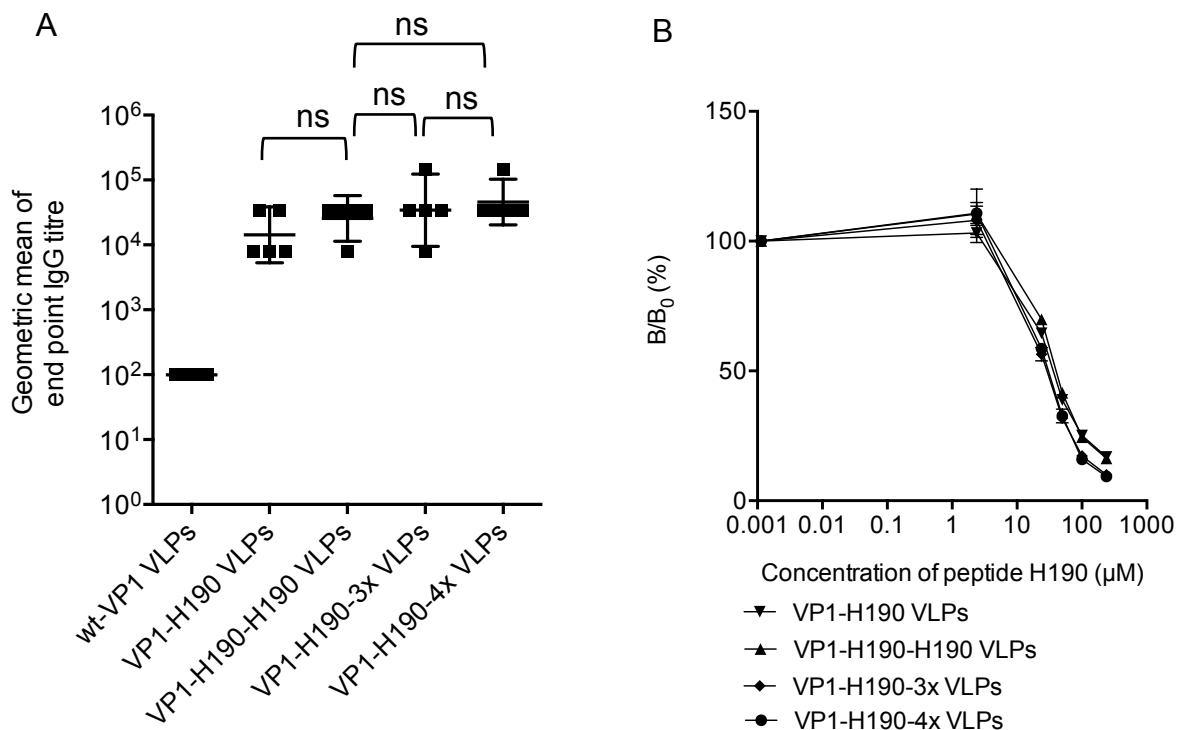


Figure 4-13. Indirect and competitive ELISA results showing the immunogenicity of H190 element in modular VLPs bearing one, two, three, and four copies of H190. (A) Endpoint titre of H190 sequence specific IgGs induced by modular VLPs. Geometric mean ($n=5$) and 95% confidence intervals were presented; and (B) Competition of antiserum binding ($1/1000$) to immobilised peptide H190 by peptide H190 in solution. Arithmetic mean ($n=5$) and standard deviation from triplicate measurements are presented.

Competitive ELISA was then performed to confirm the statistical analysis in Figure 4-13A. Competition profiles of peptide H190 at all tested concentrations, shown in Figure 4-13B, were similar for all modular VLPs. This result is in good agreement with the indirect ELISA result, confirming that differences of H190 sequence-specific IgG titre for all three

modular VLPs were statistically not significant. Furthermore, the difference of IgG titre between modular VP1-H190 and the other three constructs observed in Figure 4-13A was not amplified in the competitive ELISA.

Results in Figure 4-13 confirm that all modular VLPs were equally effective in inducing high titres of H190 sequence-specific IgGs. Comparison of the IgG titre of modular construct VP1-H190 against the other three constructs suggested that modularisations of H190 on a VLP using the simplest display strategy was comparably effective to tandem repeat display strategy in presenting H190 in an immunogenic form. These results are also in good agreement with findings in Chapter 3, confirming that display strategy did not affect the immunogenicity of H190.

Furthermore, the results in Figure 4-13 show that increasing the number of tandem repeat H190 from one copy to four copies did not result in the increase in the H190 immunogenicity. The finding on the immunogenicity of one and two copies of H190 was in good agreement with other studies that exploited the use of the MuPyV VP1 VLPs to present the J8 antigen from Group A *streptococcus* (GAS) (Chuan et al. 2013, Rivera-Hernandez et al. 2013). In those two studies, administration of modular VLPs bearing one and two copies of J8 antigen resulted in induction of a similar level of J8-specific IgG titre. However, the finding on the immunogenicity of three and four copies of H190 could not be compared against other studies, because reports on the studies of modular VLPs bearing more than two copies of peptide antigen within an antigen element have not been found. The most similar report available for comparison is the study on DNA-based HIV vaccine by Jain et al. (Jain et al. 2010). In the study, administration of DNA vectors encoding for VLPs presenting one, three and five copies of MPER or ELD antigen resulted in induction of increased antigen specific IgG titre. Different from the finding in this chapter, Jain et al. showed that the significant increase in antigen specific IgG titre was obtained from one to three copies. The difference between the findings in this chapter and those by Jain et al. could be antigen specific. The difference was also possible because in this chapter mice were immunised with modular VLPs instead of DNA vectors. As DNA vaccines must *in vivo* be turned into protein, there may also be dose-dependent effects manifested in this result. Nevertheless, supported by the two studies in modular VLPs presenting J8 antigen from GAS, the findings in this chapter highlighted the superiority of VLPs from the MuPyV VP1 in presenting the modularised antigens in their immunogenic forms.

4.3.3 Specificity of modular VLP antisera to recombinant HA1 protein

The specificity of modular VLP antisera to recombinant HA1 protein was determined by testing the sera from day 56 against recombinant HA1 A/California/07/2009 (H1N1) produced in insect cells as described in Section 4.2.7. Similar to the previous section, two types of immunoassays were performed, i.e. (i) indirect ELISA, and (ii) competitive ELISA. In indirect ELISA, mice sera at various dilutions were tested against immobilised HA1 protein, while in competitive ELISA, the binding of IgGs to immobilised HA1 protein was competed with peptide H190 in solution at various concentrations.

4.3.3.1 Effect of Advax™-1 on specificity of modular VLP antisera to recombinant HA1 protein

Figure 4-14A shows the endpoint titres of HA1-specific IgGs in modular VLP antisera. Both adjuvanted and non-adjuvanted VLPs induced high titres of HA1-specific IgGs, higher than 10^4 . Importantly, the IgG titre for adjuvanted VLPs (slightly below 10^5) was higher than the titre for non-adjuvanted VLPs (slightly above 10^4). Statistical analysis of HA1-specific IgG titres of both VLPs showed that the pair had an adjusted P value of 0.0552. Using a family-wise significance level of 5% ($p = 0.05$), statistical analysis showed that the titres of both VLPs were not significantly different. However, because the adjusted P value was very close to the limit, the significance of the titre difference might change if the family-wise significance level was changed. When the significance level was relaxed to 10% ($p = 0.1$), the titres of both VLPs became significantly different. This finding indicated that the use of Advax™-1 marginally increased the quality of antibodies induced by adjuvanted VLPs. Whether such increase was statistically significant or not depended on how the parameter of the analysis was determined.

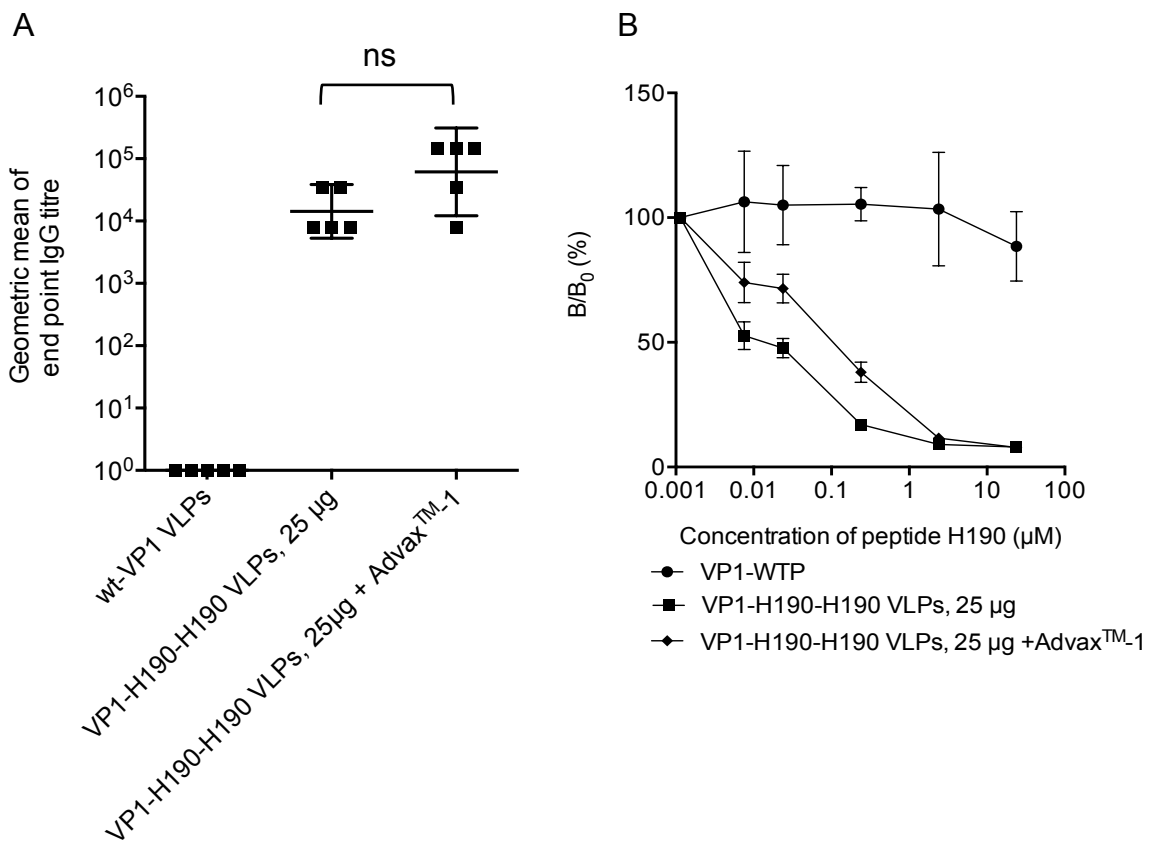


Figure 4-14. Indirect and competitive ELISA results showing the effect of Advax™-1 on reactivity of modular VLP antisera against recombinant HA1 protein. (A) Endpoint titre of HA1-specific IgGs induced by modular VLPs. Geometric mean (n=5) and 95% confidence intervals were presented; and (B) Inhibition of modular VLP antisera binding (1/200) to immobilised HA1 protein by peptide H190 in solution. Arithmetic mean (n=5) and standard deviation from duplicate measurements are presented.

The titre difference observed in the result from indirect ELISA was also observed in the result from competitive ELISA as shown in Figure 4-14B. Competitive binding of IgGs to HA1 with peptide H190 showed that peptide H190 at about 0.2 µM was required to compete the IgG binding down to 20% for both non-adjuvanted VLPs. In contrast, to achieve a similar competition, peptide H190 at a concentration of 10-fold higher was required for adjuvanted VLPs. Therefore, confirming the result from the indirect ELISA; the competition profile indicated that adjuvanted and non-adjuvanted VLPs induced different quality of antibodies.

Both results in Figure 4-14 indicate that the use of Advax™-1 increased the quality of antibodies induced by modular VLPs. However, the increase in the antibody quality was not statistically significant, and might have been deceptively greater in the competitive ELISA due to its high sensitivity as explained in Section 4.3.2.1. Therefore, the biological efficacy of the increase observed in this section may not be significant. The use of Advax™-1 associated with an increase in antibody quality has been reported previously by

Saade et al (Saade et al. 2013). In the study using guinea pigs, Advax™-1 was shown able to induce antibodies of higher quality as indicated from the increase in antibody avidity. However, a higher protective efficacy following the avidity increase was not reported. Furthermore, Saade et al. suggested that the increased antibody avidity was a result of additional rounds of B-cell proliferation, which promoted antibody maturity. This suggestion was in good agreement with the finding that temporal association and direct contact between Advax™-1 with the antigen were not observed, suggesting that Advax™-1 did not promote conformational changes of the antigen (Saade et al. 2013).

Furthermore, as explained in Section 4.3.2.1, enhancement in the protective efficacy of vaccine candidates adjuvanted with Advax™-1 is majorly driven by an increase in antibody titre (Honda-Okubo et al. 2012, Saade et al. 2013), in the absence of activation of innate immune inflammation (Honda-Okubo et al. 2012). Since Section 4.3.2.1 has shown that the use of Advax™-1 increased the immunogenicity of H190 slightly, the slight increase in antibody quality observed in Figure 4-14 was therefore expected.

4.3.3.2 Effect of increasing number of H190 tandem repeats on specificity of modular VLP antisera to recombinant HA1 protein

Figure 4-15A shows HA1-specific IgG titres for modular VLPs bearing one to four copies of H190. VP1-H190 VLPs induced HA1-specific IgG titre slightly less than 10^3 . In contrast, the other modular VLP antisera contained high titres of IgGs specific to HA1, between 10^3 and 10^4 . In addition, the figure shows that the HA1-specific IgG titres increased with the increase in H190 copy number from one to three copies. The highest titre was for modular VLPs bearing three copies of H190. HA1-specific IgG titre for four copies of H190 was slightly lower than the titre for three copies of H190. Furthermore, Figure 4-15A shows that the group of mice immunised with modular VLPs bearing three copies of H190 gave a more uniform responses against HA1, in contrast to other modular VLP groups that gave responses against HA1 with larger ranges.

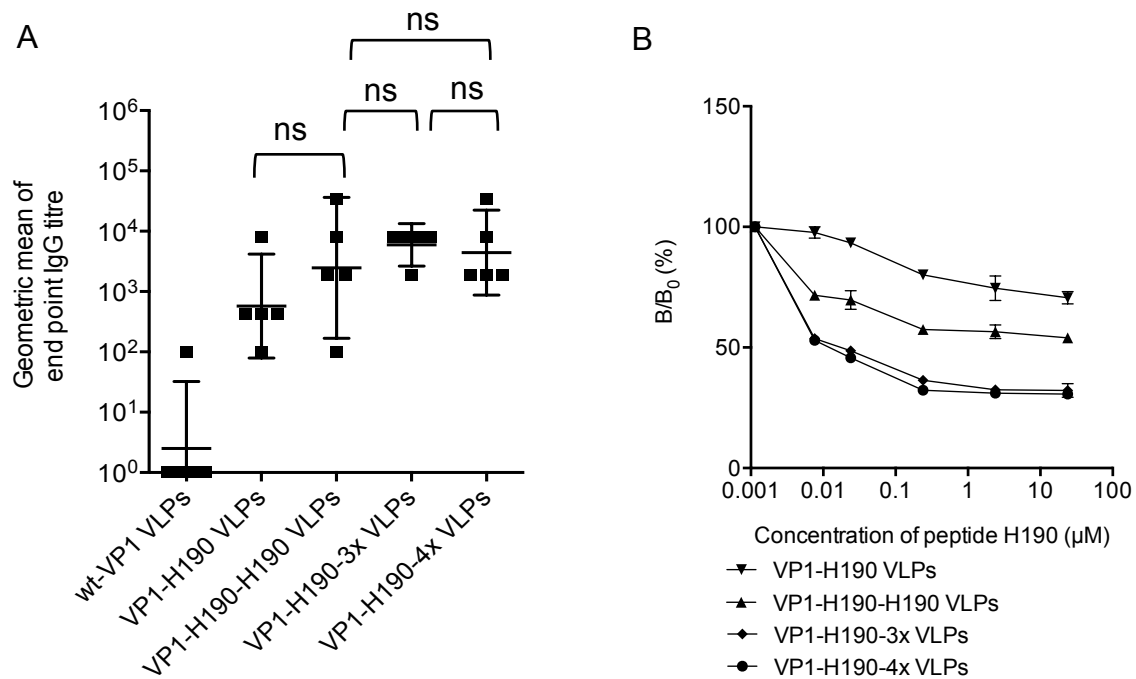


Figure 4-15. Indirect and competitive ELISA results showing the effect of increasing number of tandem repeat H190 element on the reactivity of modular VLP antisera against recombinant HA1 protein. (A) Endpoint titre of HA1-specific IgGs induced by modular VLPs. Geometric mean (n=5) and 95% confidence intervals were presented; and (B) Inhibition of modular VLP antisera binding (1/200) to immobilised HA1 protein by peptide H190 in solution. Arithmetic mean (n=5) and standard deviation from triplicate measurements are presented.

However, statistical analysis showed that HA1-specific IgG titre for VP1-H190 VLPs and VP1-H190-H190 VLPs were not significantly different ($p= 0.8213$). Similarly, statistical analysis showed that IgG titre comparison between VP1-H190-3x VLPs and VP1-H190-H190 VLPs ($p= 0.7565$), and between VP1-H190-4x VLPs and VP1-H190-H190 VLPs ($p= 0.8444$) were not significant. Statistical analysis also showed that HA1-specific IgG titre for both VP1-H190-3x VLPs and VP1-H190-4x VLPs were not statistically different ($p= 0.9784$). Furthermore, comparison of IgG titre for VP1-H190 VLPs and VP1-H190-3x VLPs ($p= 0.1433$) was not statistically significant. Similar result was obtained for comparison between VP1-H190 VLPs and VP1-H190-4x VLPs with a higher adjusted P value of 0.5093.

The increase in HA1-specific IgG titres was also observed in the competitive ELISA as shown in Figure 4-15B. The competitive ELISA result shows that the qualities of antibodies induced by modular VLPs bearing one to four copies of H190 were different. For VP1-H190 VLPs, increasing concentration of peptide H190 resulted in a moderate reduction in IgG binding to HA1. However, this was not the case for VP1-H190-H190 VLPs. IgG binding to HA1 for VP1-H190-H190 VLPs was more sensitive towards the

changes in the concentration of peptide H190. This finding suggested that IgG binding to HA1 for VP1-H190 VLPs was not specific via H190. Similar to the competition profile for VP1-H190-H190 VLPs, the profiles for VP1-H190-3x and VP1-H190-4x VLPs showed that IgG bindings for both modular VLPs were responsive to the changes in the concentration of peptide H190 in solution. Furthermore, the levels of response for both modular VLPs were similar and were higher than for VP1-H190-H190 VLPs. This finding indicated that IgGs raised against VP1-H190-3x VLPs and VP1-H190-4x VLPs had a similar quality, which was higher than the quality of antibodies induced against VP1-H190-H190 VLPs.

Both results from indirect and competitive ELISA in Figure 4-15A show two important findings from comparison between modular VLPs bearing one to four copies of H190. Firstly, both ELISA results show that two copies of H190 tandem repeats (VP1-H190-H190 VLPs) and one copy of H190, which represents the simplest display strategy but without incorporating a structural hypothesis (VP1-H190 VLPs), could induce similar levels of HA1-specific IgG titre. However, for one copy of H190, the HA1-specific IgG bound non-specifically to HA1 (Figure 4-15B). This result indicates that the simplest modularisation of H190 was not sufficient to induce a high quality of antibodies, and therefore display strategy was necessary to obtain a higher quality of antibodies. Two copies of H190 tandem repeats and one copy of H190 were equally effective in presenting H190 in its immunogenic form (Figure 4-13). Therefore, the difference in the quality of antibodies observed in this section may be attributed by a structural difference of the modularised H190 in both modular VLPs, which was similar to the observation reported in Chapter 3 of this thesis. One copy of H190 may not be able to retain its conformation integrity. The finding on the comparison of antibody quality induced by two copies of H190 tandem repeats and the simplest modularisation of H190 was an initial scientific attempt to support the presumption that peptide antigen may not be able to assume its conformation integrity when it is presented on a modular VLP (Jennings and Bachmann 2007, Roldão et al. 2010, Tissot et al. 2010). This presumption has underpinned many studies in modular VLPs, however to the best knowledge of the author, the finding in this thesis is the first reported attempt to support the presumption.

Secondly, the results in Figure 4-15 show that the number of H190 tandem repeats weakly affected the quality of antibodies. Similar to the observation on the increase in antibody quality due to the use of AdvaxTM-1 (Figure 4-14), the increase in antibody quality in Figure 4-15 was marginal and not statistically significant. The increase could be detected in competitive ELISA assay, but not in indirect ELISA, which was possibly because the former assay has a higher sensitivity (Goda et al. 2000, Reddington et al.

1991, Wreghitt et al. 1986). The finding that increasing number of H190 tandem repeats from two to three and four copies of H190 did not result in a statistically significant increase in antibody quality indicated that two copies of H190 was sufficient to induce a high quality of antibodies. This finding is in good agreement with structural studies in peptide-based vaccines (Fontenot et al. 1993). The study showed that the native conformation of antigen peptide could be achieved via incorporation of at least two copies of antigen tandem repeats. Incorporation of more than two copies did not change the structure of the tandem repeat peptides.

The two findings above were drawn mostly from the statistical analysis result. However, it could not be overlooked that both indirect and competitive ELISA results presented in Figure 4-15 show that as the number of H190 tandem repeats increased from one to three copies, the quality of antibodies qualitatively increased. Further increasing from three to four copies of H190 did not result in a further increase in antibody quality. Therefore, without using statistical analysis, it can be concluded that the highest antibody quality was achieved by immunisation with modular VLPs bearing three copies of H190 tandem repeats. This analysis was supported by the fact that mice immunised with this modular VLPs gave the most consistent responses against HA1. The uniform responses indicate that unlike other modular VLPs, modular VLPs bearing three copies of H190 could induce consistent antibody quality. For potential vaccine applications, such consistent responses can be advantageous, although it should be noted that consistency is not quantified by a simple statistical comparison of mean antibody responses.

4.4 Conclusions

In this chapter, Objective 2 stated in Section 1.3 was accomplished by addressing the question concerning what improvements to a tandem repeat display strategy can be made to induce antibodies having a higher quality. Two approaches were explored, i.e. (i) by using AdvaxTM-1 as an adjuvant for modular VLPs bearing two copies of H190, and (ii) by increasing the number of H190 tandem repeats from two copies to five copies. Exploration of these two approaches led to two major findings. The first finding was correlated to the solubility and assembly capability of modular VP1 that was explained in a context of protein-protein interactions in a concentrated salt solution. Meanwhile, the second finding was correlated to *in vivo* studies undertaken in this chapter to analyse antibody quality raised by modular VLPs. Because VLPs could not be obtained for modular VP1 bearing five copies of H190, only modular VLPs bearing one to four copies of H190 were carried through *in vivo* studies.

Results in this chapter show that increasing number of H190 tandem repeats reduced the solubility and assembly capability of the modular VP1 capsomeres in the concentrated salt solution of Assembly Buffer 1, which has been optimised previously for wt-VP1. This chapter showed that in this buffer, modular VP1 bearing four and five copies of H190 (i) had low solubility levels and (ii) could not assemble into VLPs. In this chapter, these two difficulties were investigated in a context of protein-protein interactions. In the investigation, two parameters were varied, i.e. (i) the net surface charge of the proteins (zeta potential), and (ii) specific ion effects. The net surface charge of each modular VP1 was manipulated by pH adjustment, while variation of specific ion effect comprised of concentrations and types of salts. pH adjustment of Assembly Buffer 1 only slightly affected the solubility of modular VP1 bearing four copies of H190. However, pH changes affected the assembly capability of this modular VP1 considerably. In contrast, the effects of pH adjustment were not observed for modular VP1 bearing five copies of H190. Furthermore, results in this chapter show that both concentrations and types of salts in Assembly Buffer 1 affected both solubility and assembly capability of modular VP1 bearing four copies of H190. The results in this chapter are as expected from literature, insofar as that in a concentrated salt solution, protein-protein interactions are regulated by specific ion effects rather than the net surface charges of a given protein sequence. The exception was on how pH changes affected the solubility of modular VP1 bearing four copies of H190. Such exception showed that protein-protein interactions are complex, and could involve electrostatic force and changes in amino acid sequence of the modular VP1 due to addition of modularised H190 and the possibility of structural perturbation.

In addition, using indirect and competitive ELISA for sera analysis, results from *in vivo* studies in this chapter suggest that both the use of AdvaxTM-1 and increasing the number of H190 tandem repeats did not increase the immunogenicity of H190. However, both improvement approaches increased the quality of antibodies, although the increases were marginal. Results from both serological assays in this chapter show that adjuvanting modular VLPs bearing two copies of H190 increased HA1-specific IgG titre for about 10-fold. Statistical significance of the increase depended on the selected significance level. If the family-wise significance level was determined at 5%, the increase was not statistically significant. In contrast, at 10% family-wise significance level, the increase can be concluded as statistically significant. Similarly, results from two serological assays show that increasing number of H190 tandem repeats increased the quality of antibodies. Both assays showed a trend of increases in HA1-specific IgG titres. The presented results suggest that the highest antibody quality could be achieved by increasing number of H190

tandem repeats to three copies. Further increase in the number of H190 did not result in increase in antibody quality. In addition, modular VLPs bearing three copies of H190 was shown able to induce more consistent responses against HA1. Statistical analysis showed that the increases were not statistically significant. Taking into account statistical analysis, the results in this chapter suggest that modular VLPs bearing two copies of H190 was sufficient to induce a high titre of H190-sequence specific IgGs and HA1-specific IgGs. Therefore, improvements to the tandem repeat display strategy to increase the quality of antibodies were not necessary.

Chapter 5. Challenging the tandem repeat display strategy using a hydrophobic helix 190 variant

5.1 Introduction

Chapter 4 of this thesis accomplished Objective 2 stated in Section 1.3, by addressing possible improvements to the tandem repeat display strategy used to induce antibodies of higher quality. Two approaches were explored in Chapter 4: (i) the use of adjuvant, and (ii) increasing copy number of H190 up to five copies. Results in the chapter showed that both approaches did not significantly increase the quality of antibodies compared to those induced by two copies of H190. This finding highlighted that two copies of H190 were sufficient to display the peptide antigen in its immunogenic form, resulting in the induction of high titres of H190-sequence specific IgGs (as also reported in Chapter 3). The induced IgGs were shown to be of a high quality, which was measured as their abilities to recognise recombinant HA1 protein produced in insect cells.

In addition to its ability to induce high titres of antibody of a high quality, the tandem repeat display also has other advantages. Firstly, it is a simple display strategy; an antigen module containing two copies of H190 tandem repeats was modularised without any additional structure-promoter elements or spacer elements. The absence of additional elements is advantageous because such elements can be immunogenic, and the binding of antibodies to these elements may interfere with the antibody binding to the modularised H190 (as reported in Chapter 3). Secondly, modularisation of two copies of H190 was not associated with considerable precipitation of modular VP1 after assembly or reduced assembly capability of the modular VP1 under buffer conditions that were previously optimised for the wild-type (wt) VLP. Section 4.3.1 in Chapter 4 showed that as the copy number of modularised H190 increased, the tendency of modular VP1 to precipitate in standard buffer after assembly also increased, and that the assembly capability of modular VP1 decreased. Modular VP1, bearing three, four, and five copies of H190, precipitated considerably after assembly. Furthermore, under the limited set of excipient-free buffer conditions screened manually, modular VLPs displaying three and four copies of H190 had a lower capability to assemble into VLPs, resulting in a lower yield of modular VLPs, and modular VP1 presenting five copies of H190 could not assemble into VLPs.

The superiority of the tandem repeat display, as reported in Chapters 3 and 4, was built using H190 from influenza virus A/California/07/2009 (H1N1; STSADQQSLYQNADAY) as an antigen model. However, H190 is a hypervariable region

(reviewed in Chapter 2); the amino acid sequence of H190 differs from one strain to another. Of particular interest is H190 (VTDKDQIFLYAQASGR) from influenza virus vaccine strain A/Victoria/210/2009 (H3N2/X-187), which is a high-growth reassortant virus strain. Together with the A/California/07/2009 strain, the A/Victoria/210/2009 strain was included in the trivalent split influenza vaccines, Fluvax® (CSL Limited, Australia) for the 2011 and 2012 seasons. Unlike the H190 element from the A/California/07/2009, the H190 element from the A/Victoria/210/2009 is hydrophobic, as indicated by its GRAVY score (grand average of hydropathicity index) of -0.419. The GRAVY score is a scoring system that can be used as an indicator of protein solubility (Kyte and Doolittle 1982). In this chapter, the GRAVY score of H190 from the A/Victoria/210/2009 was obtained from ProtParam tool (Gasteiger et al. 2005). The hydrophobic non-polar residues within the H190 element are centred in the middle region of the element, forming a hydrophobic stretch (amino acid residues IFLY).

Although amino acid sequence, and consequently the biophysical and biochemical properties of H190 from both strains are different, their helical structures are conserved. Conservation of H190 helicity across various influenza strains has been indicated in several studies. For example, a study by Stevens et al. (Stevens et al. 2006) compared the structure of HA1 from two different influenza strains, i.e. (i) A/Human/South Caroline/1/1918 (H1N1; 1918 H1), and (ii) A/Duck/Singapore/3/97 (H5N1; Sing97). The study showed that Sing97 and 1918 H1 were not closely related in sequence, differing by approximately 40% in the sequence of HA1. The amino acid sequences of H190 from these two strains were also different (Figure 5-1A). However, superimposition of HA1 monomer from these strains showed that the structure of these HA1 were similar (Figure 5-1B). This superimposition also showed that the helicity of H190 from these strains were similar. This comparison was supported by studies of Russell et al. (Russell et al. 2004, Russell et al. 2006). In these studies, Russell et al. compared the structure of HA1 of 1918 H1 and Sing97 against A/X-31 (H3N2), A/Swine/Hong Kong/9/98 (H9N2), and A/Turkey/Italy/02 (H7N3). HA1 from 1918 H1 differed from HA1 from the H3N2, H7N3, and H9N2 by approximately 60%. In comparison, HA1 from the H1N1 strain was more similar to the one from Sing97, differing by about 40%. Nevertheless, confirming the aforementioned study, Russell et al. showed, regardless of their sequence differences, the helicity of H190 from these strains was similar. These studies indicated that, independent of similarities or differences in the sequence of HA1, the secondary structure of H190 in these strains was conserved. Furthermore, these studies also suggested that the helicity

of H190 was conserved among influenza strains, which recognise different types of sialic acid receptors.

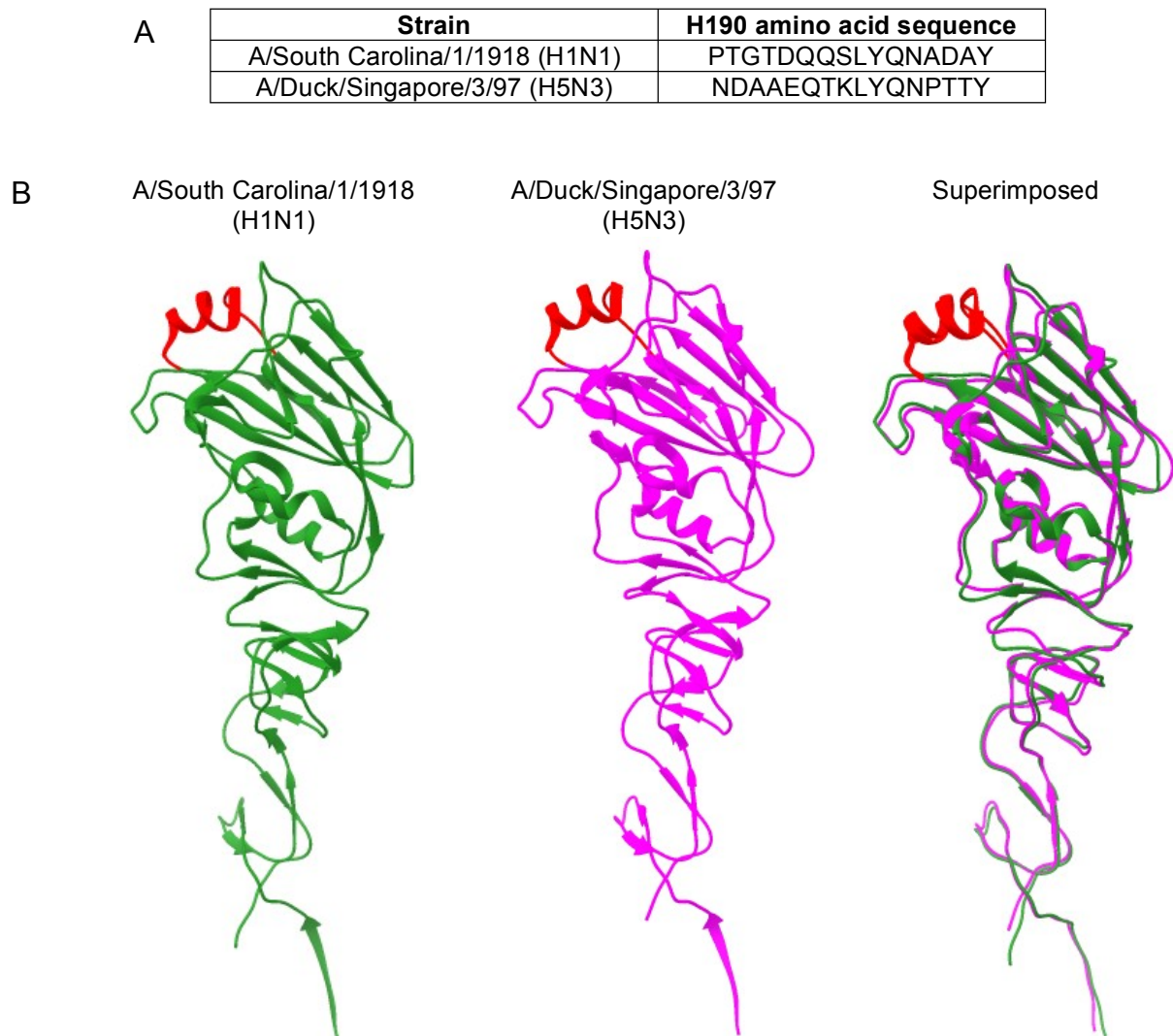


Figure 5-1. Comparison of crystal structures of A/Human/South Carolina/1/1918 (H1N1) HA1 and A/Duck/Singapore/3/97 (H5N1) HA1. (A) Comparison of amino acid sequences of H190 elements from both strains; and (B) Structural comparison and superposition of HA1 monomers from the H1N1 (1RUZ.pdb) and H5N1 (1JSM.pdb). Helix 190 is coloured in red. Figure B was generated using UCSF Chimera 1.8.1 (Pettersen et al. 2004).

A stretch of hydrophobic amino acids within a protein sequence is known to be able to cause the pathological protein aggregation that is involved in a wide variety of diseases, such as Parkinson's disease (Singleton et al. 2003), Alzheimer's disease (Harper et al. 1997, Selkoe 1991), and Prion disease (Lindquist et al. 2001, Scheibel et al. 2004). Such pathological protein aggregation can be lethal for people carrying detrimental mutations. In addition, a hydrophobic stretch within a protein can cause non-pathological protein

aggregation, as observed in antibodies (Wu et al. 2010) and the globular protein Scr Homology 3 (Ventura et al. 2004). Non-pathological protein aggregation can become a nuisance factor in *in vitro* studies, or cause major economic and technical problems in biotechnology and pharmaceutical industries, such as during product storage and delivery (Fink 1999). Both types of protein aggregation are non-native aggregates and distinguishable from native protein aggregates, which can result from high concentrations of protein exceeding the solubility limit, salting out, and isoelectric precipitation. While native protein aggregation retains its native conformation and can be re-dissolved, non-native protein aggregates cannot maintain their native folding and can only be dissolved in buffer containing, for example, a high concentration of detergents (Fink 1999, Kopito 2000).

Modularisation of an antigen module bearing a hydrophobic stretch in a viral capsid protein has been considered to be challenging as it promotes incorrect folding of the capsid protein (Kazaks et al. 2004), and consequently increases the tendency of the capsid proteins to aggregate (Aleksaitė and Gedvilaitė 2006, Shin and Folk 2003). By doing so, the hydrophobic antigen module prevents the proper self-assembly of the modular capsid protein into VLPs. The study using hamster polyomavirus VP1 reported that modular VP1 bearing hydrophobic tumour associated antigens (27 amino acids) formed VLPs with a heterogeneous diameter. Additionally, irregular aggregates were also observed (Aleksaitė and Gedvilaitė 2006). Similarly, the study by Karpenko et al. (Karpenko et al. 2000) showed that when an antigen module with a high hydrophobicity score was modularised into Hepatitis B core protein (HBcAg), the modular protein could not assemble into VLPs. They suggested that the hydrophobic antigen module was likely to fold in such a way that it blocked the contact region between the subunits of HBcAg dimer. Formation of the HBcAg dimer initiates the spontaneous assembly of HBcAg into VLPs. Thus, when the contact region between the dimer-subunit was blocked, HBcAg dimers could not be formed, and consequently the modular HBcAg could not assemble into VLPs.

Incorrect folding of the capsid protein can also affect the quality of modular capsid proteins. In cases where VLPs are assembled *in vitro*, modular capsid proteins are often expressed with fusion tags (Pattenden et al. 2005). The fusion tags are removed from the capsid proteins using proteases prior to assembly processes under controlled buffer conditions. Incorrect folding of the capsid proteins can result in non-specific proteolysis (Li et al. 1983), and may consequently increase the heterogeneity of cleavage products after the tag-removal process. An example of a viral capsid protein that has been expressed

with a fusion tag is MuPyV VP1 protein. The VP1 protein has been expressed as a Glutathione S-transferase (GST)-tagged protein in *E. coli* (Chuan et al. 2008). Removal of GST tag from the wt-VP1 protein using thrombin has been shown to result in the generation of a secondary cleavage product, designated VP1*. Additional purification processes are needed to remove the VP1* in order to improve the quality of the wt-VP1 protein. VP1* also reduced the quality of assembled VLPs. The study by (Connors et al. 2013) showed that, in the presence of VP1*, the assembled VLPs contained a higher amount of misformed VLPs. In contrast, in the absence of VP1*, VLPs with a higher homogeneity were obtained.

Although the influences of hydrophobic antigen modules are known, approaches to minimise these impacts are still limited to the use of mosaic VLPs (Karpenko et al. 2000, Kazaks et al. 2004). Mosaic VLPs are VLPs composed of a modular capsid protein and sterically less challenging capsid protein, e.g. wild-type capsid protein (Vogel et al. 2005). They can be obtained *in vivo* in *E. coli* by performing the co-expression of wild-type and modular capsid protein at a determined ratio using two compatible plasmids (Beterams et al. 2000) or limited translational read-through (Kazaks et al. 2004, Koletzki et al. 1997). Alternatively, they can be obtained from a column-based assembly process, which is comprised of three steps: (i) immobilisation of both wild-type and modular capsid proteins on the surface of a Ni²⁺ matrix via affinity interactions, (ii) assembly step, and (iii) and the elution of mosaic VLPs from the matrix (Vogel et al. 2005). To some extent, this column-based process is rather a complicated process.

The impacts of hydrophobic antigen modules and the limited approach to minimise these impacts motivated this chapter to investigate the application of the preferred display strategy for a hydrophobic H190 variant (VTDKDQIFLYAQASGR), which is from the influenza virus vaccine strain A/Victoria/210/2009 (H3N2/X-187). Modularisation of H190 from the A/Victoria/210/2009, which contains a hydrophobic stretch, on MuPyV VP1 led to two questions: (i) to what extent does the hydrophobic H190 variant affect the properties of the modular VLP, and (ii) how can the impacts of the hydrophobic H190 variant be minimised?

Based on the rationale discussed, this chapter reasons that modularisation of H190 from A/Victoria/210/2009 could potentially result in: (i) increased aggregation, and (ii) increased non-specific thrombin proteolysis. This chapter aims to explore two approaches simultaneously to minimise the effects of hydrophobic stretch on the properties of modular VP1. These approaches are: (i) the addition of charged residues into the antigen module, and (ii) the use of a protease with a higher specificity to minimise non-specific proteolysis.

In this chapter, investigations to answer these two questions are focused on the established buffer conditions that have been previously optimised for wt-VP1.

In the first approach, charged residues were utilised to minimise the tendency of modular VP1 to form soluble aggregates, and consequently obtain modular VP1 capsomeres. Charged residues have been shown to improve protein solubility (Mosavi and Peng 2003). They are commonly used to substitute hydrophobic residues within a hydrophobic stretch (Armstrong et al. 2011, Bolognesi et al. 2013, Xu et al. 2013), but studies on VLPs have not been reported. However, the direct substitution of hydrophobic residues can impact the structure, bioactivity, and stability of proteins, and therefore this approach necessitates extensive information regarding the protein structure (Kato et al. 2007). Alternatively, charged residues can be utilised as a fusion tag at the N- and/or C-termini of proteins (Jung et al. 2011, Kato et al. 2007). The use of charged residues as a fusion tag has not been reported to have any impact on the structure of proteins. Nevertheless, both N- and C-termini of MuPyV VP1 are involved in the assembly of VP1 capsomeres into VLPs (Stehle and Harrison 1997). Thus, the addition of charged residues at the N- and/or C-termini of the VP1 protein is not preferable, as it may interfere with the assembly process. Besides, the addition of charged residues at the N-terminal of a protein has been shown to reduce production of the protein due to interference in the initiation of translation at the N-terminus (Lilie et al. 2013). Based on these considerations, in this chapter, charged residues were added as ionic elements that flank the two copies of H190 tandem repeats.

The flanking ionic elements in this chapter were comprised of glutamic acid residues. Glutamic acid residue was selected because in VLP studies, it has been previously used for various purposes, such as peptide conjugation and purification. An ionic sequence comprising glutamic acids and cysteine (CEEEEEEEE) was explored in a study on MuPyV VP1 protein (Stubenrauch et al. 2000). The ionic sequence was used to aid in the purification of the viral capsid protein using ion-exchange chromatography. In the study, the ionic sequence was genetically fused between Asn294 and Tyr295 (HI loop) of the VP1 protein. The study showed that insertion of the ionic sequence did not have any impact on the VLP assembly process, which was performed *in vitro*. The same ionic sequence was also used in a study using papillomavirus L1 protein (Viscidi and Bossis 2014). However, unlike the study by Stubenrauch et al., the study by Viscidi and Bossis generated VLPs *in vivo*. In the later study, amino acids Cys347-Glu355 (HI loop) of papillomavirus L1 protein were replaced with the ionic sequence CEEEEEEEE. The study showed that the replacement did not interfere with the capability of the L1 protein to

assemble into VLPs *in vivo*. The study also showed that a similar result was obtained when the replacement was performed at amino acids Cys413-Glu421 (H4 loop), and Asn129-Glu137 (DE loop). Additionally, glutamic acid is negatively charged, which is preferable to positively charged residues, such as Lysine and Arginine. Positively charged residues could increase the binding of DNA and RNA, and hence reduce the purity of the VP1 protein. Moreover, the peptide bonds of arginine has been known to be susceptible to proteolysis degradation (Lilie et al. 2013).

In an attempt to reduce non-specific thrombin proteolysis following GST-tag removal, this chapter compares the use of thrombin and tobacco etch virus protease (TEVp). A study by Connors et al. (Connors et al. 2013) conducted bioinformatic analyses to predict thrombin and TEVp cleavage sites in wt-VP1 from murine polyomavirus. Python programming was used to search for (i) amino acid sequences recognised by both proteases experimentally, and (ii) amino acid sequences that follow the cleavage rules outlined by the ExPASy Peptide Cutter program (Gasteiger et al. 2003). The study predicted that thrombin had five of secondary cleavage sites in the wt-VP1; these sites are at Lys4, Arg22, Lys28, Arg58, and Arg305 (Figure 5-2). The size of truncated VP1 fragments resulting from predicted thrombin proteolysis in these sites were determined using the ProtParam tool (Gasteiger et al. 2005) and are shown in Table 5-1. In contrast, TEVp was predicted not to have any secondary cleavage site in the wt-VP1 protein. The only TEVp cleavage site found was on the engineered cleavage site (Figure 5-2). These predictions were then proven experimentally. Experiments performed by Connors et al. proved that the removal of the GST tag from the wt-VP1 using thrombin resulted in the generation of a smaller secondary cleavage product with a theoretical molecular weight of about 37 kDa, designated VP1*. N-terminal protein sequencing identified this fragment as a product of thrombin secondary cleavage at Arg58. In contrast, the experiments proved that wt-VP1 with a higher homogeneity could be obtained following GST removal using TEVp.

```

Thrombin      MSPILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYID 60
TEVp          MSPILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYID 60
*****

Thrombin      GDVKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSR IAYSKDFETLKV 120
TEVp          GDVKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSR IAYSKDFETLKV 120
*****

Thrombin      DFLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFK 180
TEVp          DFLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFK 180
*****

Thrombin      KRIEAIPOIDKYLKSSKYIAWPLQGWQATFGGGDHPPKSD--LVPRGSGGMAPKRKSGVS 238
TEVp          KRIEAIPOIDKYLKSSKYIAWPLQGWQATFGGGDHPPKSDENLYFOGSGGMAPKRKSGVS 240
*****
          Thrombin      Thrombin      Thrombin
          ___|___|___      ___|___
Thrombin      KCETKCTKACPRPAPVPKLLIKGGMEVLDLVTGPDSVTEIEAFLNPRMGQPPTPESLTEG 298
TEVp          KCETKCTKACPRPAPVPKLLIKGGMEVLDLVTGPDSVTEIEAFLNPRMGQPPTPESLTEG 300
*****

Thrombin      GQYYGWSRGINLATSDESPGNNTLPTWSMAKLQLPMLNEDLTCDTLQWWEAVSVKTEV 358
TEVp          GQYYGWSRGINLATSDESPGNNTLPTWSMAKLQLPMLNEDLTCDTLQWWEAVSVKTEV 360
*****

Thrombin      VGSGSLLDVHGFNKPTDVTNKTGISTPVEGSQYHVFAVGGEPLDLQGLVTDARTKYKEEG 418
TEVp          VGSGSLLDVHGFNKPTDVTNKTGISTPVEGSQYHVFAVGGEPLDLQGLVTDARTKYKEEG 420
*****

Thrombin      VVTIKTITKKDMVNKDQVLNPI SKAKLDKDGMYPVEIWHPDPAKNENTRYFGNYTGTTT 478
TEVp          VVTIKTITKKDMVNKDQVLNPI SKAKLDKDGMYPVEIWHPDPAKNENTRYFGNYTGTTT 480
*****
          Thrombin
          ___|___
Thrombin      PPVLQFTNTLTTVLLDENGVGPLCKGEGLYLSCVDIMGWRVTRNYDVHHRGLPRYFKIT 538
TEVp          PPVLQFTNTLTTVLLDENGVGPLCKGEGLYLSCVDIMGWRVTRNYDVHHRGLPRYFKIT 540
*****

Thrombin      LRKRWVKNPYPMASLISSLFNNMLPQVQGQPMEGENTQVEEVRVYDGTEPVP GPDPMTRY 598
TEVp          LRKRWVKNPYPMASLISSLFNNMLPQVQGQPMEGENTQVEEVRVYDGTEPVP GPDPMTRY 600
*****

Thrombin      VDRFGKTKTVFPGN 612
TEVp          VDRFGKTKTVFPGN 614
*****

```

Figure 5-2. Amino acid sequence alignment of GST-tagged wt-VP1 bearing thrombin or the TEVp cleavage sites. The figure reproduced from (Connors et al. 2013). Alignment was performed using ClustalW2 (<http://www.ebi.ac.uk/Tools/msa/clustalw2/>). Residues highlighted in yellow refer to the engineered thrombin cleavage site, while residues highlighted in green refer to the engineered TEVp cleavage site. Five potential thrombin secondary cleavage sites are also shown. Cleavages are predicted to occur after the marked lysine (K) or arginine (R).

Table 5-1. List of predicted secondary thrombin cleavage sites on wt-VP1 and their cleavage products. The table reproduced from (Connors et al. 2013).

Cleavage site	VP1 fragment	Mass (Da)
Lys4	1-4	445
	5-384	42,052
Arg22	1-22	2,408
	23-384	40,089
Lys28	1-28	2,999
	29-384	39,499
Arg58	1-58	6,109
	59-384	36,414
Arg305	1-305	33,412
	306-384	9,111

In addition to its high specificity, TEVp is also more cost effective than thrombin for both small- and large-scale production. It is also not involved in human physiological and biochemical networks (Connors et al. 2013). In contrast, thrombin regulates haemostasis and thrombosis in humans, which are important processes in the formation of blood clots. Excessive thrombosis and haemostasis can cause death, stroke, and a number of other medical complications (Maryanoff 2004). Because the future intended application of the study reported in this chapter is as a human vaccine, a trace amount of thrombin in the products may be undesirable. Therefore, this chapter explores the use of TEVp, which is a proven alternative for thrombin, to reduce the heterogeneity of modular VP1.

5.2 Materials and methods

5.2.1 Generation of modular constructs

Plasmid pGEX-VP1 was generously provided by Professor Robert Garcea (University of Colorado). Plasmid pGEX-VP1-S1S4 was obtained as described in Chapter 4 (Section 4.2.1). The DNA sequence of an antigen module containing H190 was codon optimised for *E. coli*. Homologous DNA sequences (21-24 bp) flanking the *AfeI* site of plasmid pGEX-VP1-S1S4 were added to the 5' and 3' ends of DNA sequence encoding the antigen module (Figure 4-2). A DNA fragment containing the antigen module and the homologous regions were synthesised by assembling a set of oligonucleotides, which were designed using DNAWorks (Hoover and Lubkowski 2002) as outlined below:

1. Construct VP1-H190-H190

The antigen module in construct VP1-H190-H190 consisted of two copies of the H190 element (STSDQQLYQNADAY) from A/California/07/2009 (H1N1). Construct VP1-H190-H190 was generated as previously described in Chapter 3 of this thesis.

2. Construct VP1-H3-H190-H190

The antigen module in construct VP1-H3-H190-H190 consisted of two copies of the H190 element (VTDKDQIFLYAQASGR) from influenza A/Victoria/210/2009 (H3N2/X-187). A DNA fragment containing the antigen module with homologous regions was generated by assembling the following set of oligonucleotides:

5' atgggctggagagttacaagaagcgtgaccgacaagga 3',
5' cctgcgcatcacaggaaaatttgatccttgcggtcacgct 3',
5' aaatttcctgtatgcgagcggcgtctggctgtgttacgga 3',
5' cgtacagaaagatctggctttgtccgtaaacacgaccagacg 3',
5' caaagaccagatctttctgtacgctcaagcgagcggccgt 3', and
5' tctccagtgatggacatcataagcacggccgctcgcttg 3'.

3. Construct VP1-H3-H190-H190-4E

The antigen module in construct VP1-H3-H190-H190-4E consisted of two copies of the H190 element from influenza A/Victoria/210/2009 (H3N2/X-187) flanked by ionic elements, which were composed of four consecutive glutamic acid residues:

EEEEVTDKDQIFLYAQASGRVTDKDQIFLYAQASGREEEE.

A fragment of DNA containing the antigen module with homologous regions was generated by assembling the following set of oligonucleotides:

5'ggctggagagttacaagaagcgaagaagaggaagttac 3',
5'cggaagcctgcgctacagaaaaatctgatccttgcggttaactcctcttcttcgcttc 3',
5'acgcgcaggctccggcgtgtgaccgataaagaccagatcttctgtatgctcaggcgt 3', and
5'ccagtgatggacatcataagcctcttctcctcacgaccagacgcctgagcatacagga 3'.

Each set of oligonucleotides was assembled, amplified, and purified as described in Section 4.2.1. The purified DNA fragment was cloned into *AfeI*-linearised vector pGEX-VP1-S1S4 using the *in vivo* homologous recombination (Bubeck et al. 1993, Jones 1994, Oliner et al. 1993, Parrish et al. 2004), as described in Section 4.2.1.

The engineered thrombin cleavage site in pGEX-VP1 was mutated into a TEVp cleavage site, as described previously (Connors et al. 2013), yielding the construct TEVP-VP1. The mutation was performed by the Protein Expression Facility (The University of Queensland, Australia) using the QuikChange Lightning Site-directed Mutagenesis kit (Agilent Technologies, Inc., Santa Clara, CA, USA), according to the manufacturer's recommendations, with the complementary oligonucleotide 5' tgggtggcgaccatcctccaaaatcggatgaaaactgtactccaaggatccggaggaatggcc 3'. Similarly, the

engineered thrombin cleavage sites in modular constructs VP1-H190-H190, VP1-H3-H190-H190, and VP1-H3-H190-H190-4E were mutated into TEVp cleavage sites, generating the constructs TEVP-VP1-H190-H190, TEVP-VP1-H3-H190-H190, and TEVP-VP1-H3-H190-H190-4E, respectively.

DNA sequences of all generated constructs were verified by the Australian Genome Research Facility (AGRF, Brisbane, Australia). The constructs used this chapter are listed in Table 5-2.

Table 5-2. List of constructs used in Chapter 5.

Constructs	Cleavage site	Antigen module
Wild-type VP1 (wt-VP1)	Thrombin	No antigen module
VP1-H190-H190		Two copies of H190 from A/California/07/2009
VP1-H3-H190-H190		Two copies H190 from A/Victoria/210/2009
VP1-H3-H190-H190-4E		Two copies H190 from A/Victoria/210/2009 flanked with ionic elements consisting of four glutamic acids
TEVP-wt-VP1	TEVp	No antigen module
TEVP-VP1-H190-H190		Two copies H190 from A/California/07/2009
TEVP-VP1-H3-H190-H190		Two copies H190 from A/Victoria/210/2009
TEVP-VP1-H3-H190-H190-4E		Two copies H190 from A/Victoria/210/2009 flanked with ionic elements consisting of four glutamic acids

5.2.2 Protein concentration measurements

Protein concentration was determined using UV absorbance at 280 nm, based on the Beer-Lambert Law (Aitken and Learmonth 1996),

$$A = \epsilon b c$$

where A is the measured absorbance at 280 nm, ϵ is the extinction coefficient of protein ($M^{-1} \text{ cm}^{-1}$) at 280 nm measured in water, b is the sample path length (10 mm), and c is the protein concentration (M). Theoretical molecular weight and extinction coefficient of each protein was obtained using the ProtParam tool (Gasteiger et al. 2005), and summarised in Table 5-3.

Table 5-3. List of theoretical molecular weights and extinction coefficients for proteins used in Chapter 5.

Protein	Theoretical molecular weight of monomeric protein (g mol ⁻¹)	Extinction coefficient of monomeric protein, at 25°C in water (M ⁻¹ cm ⁻¹)
wt-VP1	42763.6	58057
TEVP-VP1		
VP1-H190-H190	46251.1	64205
TEVP-VP1-H190-H190		
VP1-H3-H190-H190	46351.7	61225
TEVP-VP1-H3-H190-H190		
VP1-H3-H190-H190-4E	47384.6	61225
TEVP-VP1-H3-H190-H190-4E		
TEVp	28600	30035

5.2.3 Expression and purification of modular VP1 proteins

Expression and purification of modular VP1 protein for each construct was conducted as described in Section 4.2.2, unless stated otherwise. Briefly, generated plasmids were transformed into *E. coli* Rosetta (DE3) pLysS competent cells (EMD Millipore, Merck KGaA, Darmstadt, Germany). Modular VP1 protein was expressed as a GST-tagged protein at 26°C or 12°C. Following overnight incubation, harvested cells were lysed by sonication (Branson Ultrasonics Corporation, Connecticut, USA) at output 30 for 4 cycles of 40 s in 40 ml of Lysis Buffer (40 mM Tris-base, 200 mM NaCl, 1 mM EDTA, 5% (v/v) glycerol, 5 mM DTT, pH 8.0). The supernatant was separated from cell debris by centrifugation, filtered through 0.45 µm filters (Pall, New York, USA), and then loaded into a GST affinity column (GSTrap HP 5 ml, GE Healthcare, UK) at a flow rate of 0.5 ml min⁻¹. Bound GST-tagged VP1 protein was eluted with Lysis Buffer containing 10 mM GSH (40 mM Tris-base, 200 mM NaCl, 1 mM EDTA, 5% (v/v) glycerol, 5 mM DTT, 10 mM GSH, pH 8.0).

GST tag was removed by incubating GST-tagged modular VP1 (about 473 µg) with 10 U of thrombin or about 48 µg of TEVp (The Protein Expression Facility, The University of Queensland, Australia) at room temperature for 2 or 1 h, respectively, in a final volume of 250 µl. Digested protein was centrifuged and the supernatant was then loaded into a size exclusion chromatography (SEC) column (Superdex 200 10/300 GL, GE Healthcare, UK) pre-equilibrated with Lysis Buffer, as described previously (Chuan et al. 2008), to separate modular VP1 capsomeres from soluble aggregates and GST.

5.2.4 SDS-PAGE

Analysis using sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) was conducted as described in section 4.2.4. Each SDS-PAGE gel picture in

Figures 5-5, Figure 5-6A, Figure 5-6B, Figure 5-7, Figure 5-9A, Figure 5-9B, and Figure 5-12 was taken from one gel. Relevant sample lanes were selected and positioned on the right hand side of the ladder.

5.3 Results and discussion

5.3.1 Increased aggregation and heterogeneity of modular VP1 bearing a hydrophobic stretch after GST removal

The effects of a hydrophobic stretch on the properties of a modular VP1 was examined by investigating the properties of a modular VP1 displaying two copies of H190 elements (VTDKDQIFLYAQASGR) from influenza virus A/Victoria/210/2009 (H3N2/X-187). An antigen module containing two copies of the H190 tandem repeats was genetically fused into MuPyV VP1, yielding the modular construct VP1-H3-H190-H190. The modular protein VP1-H3-H190-H190 was expressed as a GST-tagged protein at 26°C using the protocol in Section 5.2.3, side-by-side with wt-VP1.

The expression and fractional solubility of GST-tagged VP1-H3-H190-H190 (about 72.5 kDa) in comparison with GST-tagged wt-VP1 (about 69 kDa) was analysed using SDS-PAGE, as shown in Figure 5-3A. The analysis showed that the total fraction for VP1-H3-H190-H190 was similar to that of wt-VP1. Furthermore, the majority of GST-tagged wt-VP1 was found in the soluble fraction. However, a smaller proportion of GST-tagged VP1-H3-H190-H190 was found in the soluble fraction, indicating that the modular VP1 protein had lower solubility under the chosen buffer conditions. Thus, the results show that modularisation of two copies of hydrophobic H190 did not affect the expression of the modular VP1, but lowered the fractional solubility of modular VP1, as might be predicted.

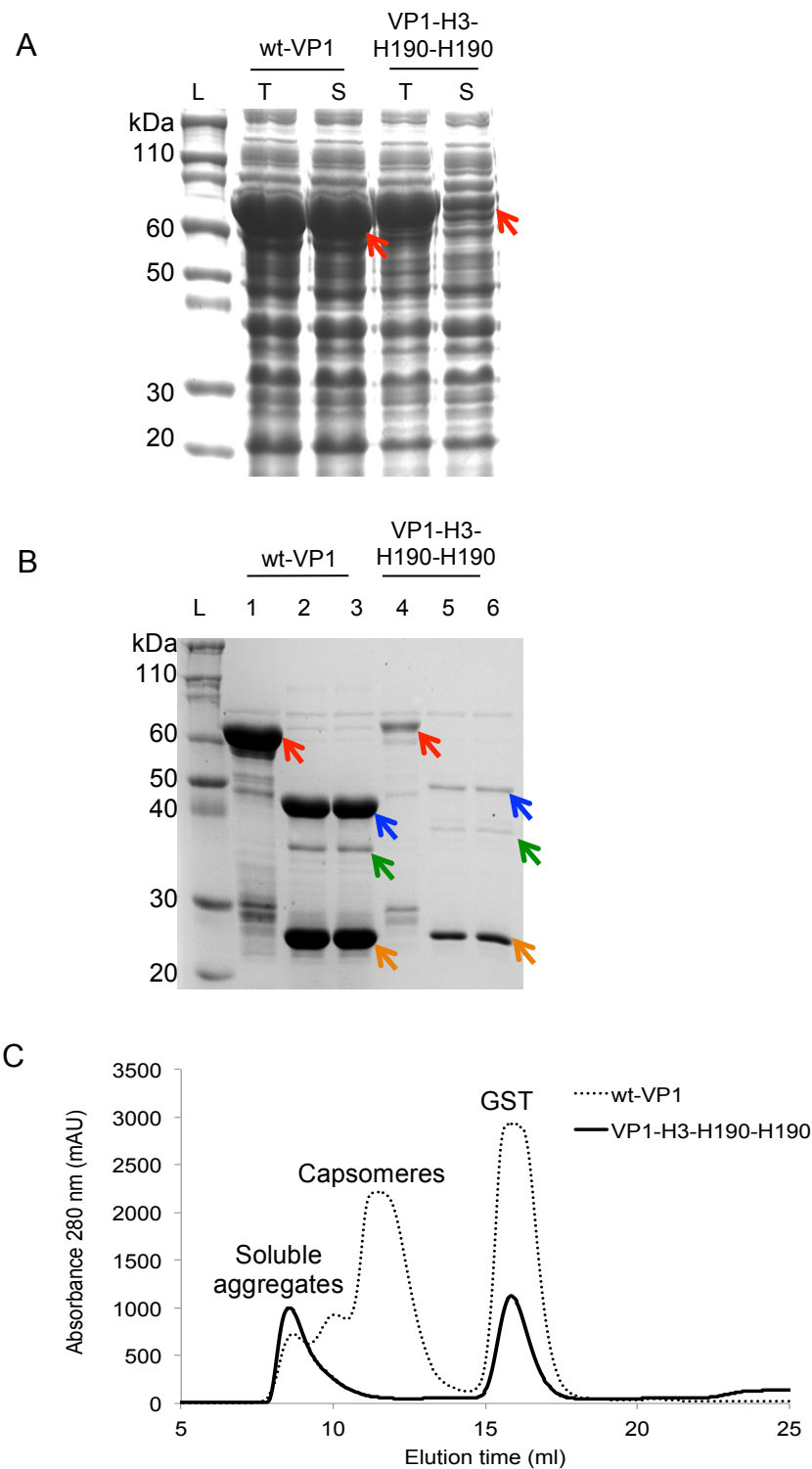


Figure 5-3. Properties of modular VP1 bearing two copies of H190 from A/Victoria/210/2009 (construct VP1-H3-H190-H190). (A) Analysis of expression and solubility of GST-tagged wt-VP1 and VP1-H3-H190-H190 using SDS-PAGE following an overnight expression at 26°C. Red arrows refer to GST-tagged VP1. Lanes: (L) Pre-stained protein marker, (T) Total protein, (S) Soluble protein; (B) SDS-PAGE analysis of GST removal using thrombin. Lanes: (1) GST-tagged wt-VP1, (2) Total protein of thrombin-digested GST-tagged wt-VP1, (3) Soluble protein of thrombin-digested GST-tagged wt-VP1, (4) GST-tagged VP1-H3-H190-H190, (5) Total protein of thrombin-digested GST-tagged VP1-H3-H190-H190, and (6) Soluble protein of thrombin-digested GST-tagged VP1-H3-H190-H190. The soluble protein after thrombin digestion was then loaded into an SEC column. Blue arrows refer to the VP1 protein. Green arrows refer to secondary thrombin cleavage products. Orange arrows refer to the GST tag; and (C) Size-exclusion chromatograms of wt-VP1 and VP1-H3-H190-H190 following treatment with thrombin.

Following purification using affinity chromatography, GST-tagged wt-VP1 and VP1-H3-H190-H190 were incubated with thrombin to remove GST tag from VP1. Figure 5-3B shows SDS-PAGE analysis of GST removal using thrombin. For wt-VP1 (lane 1-3), two major products were observed after GST removal (Lipin et al. 2008, Middelberg et al. 2011). These products were: (i) VP1 (bands at about 42 kDa; blue arrow), and (ii) GST tag (bands at about 25 kDa; orange arrow). A minor thrombin proteolysis product was observed at about 37 kDa (green arrow). The study by Connors et al. identified a similarly-located 37-kDa band as a degraded VP1 fragment, designated VP1*, which resulted from secondary thrombin proteolysis at Arg58 of VP1 (Connors et al. 2013). A similar digestion profile was observed for VP1-H3-H190-H190 (lane 4-6). Three cleavage products were observed for the modular construct: (i) modular VP1 (bands at about 47 kDa; blue arrow), (ii) a GST tag (bands at about 25 kDa; orange arrow), and (iii) a cleavage product, which was represented as bands at approximately 39 kDa (green arrow). The molecular weight of product (iii) was comparable to the theoretical molecular weight of VP1* in wt-VP1 added to the theoretical molecular weight of two copies of H190 (about 3.6 kDa). Based on this analysis, product (iii) was suggested to result from secondary thrombin proteolysis at Arg58, like VP1*. Furthermore, comparison of VP1* and VP1 in lanes 2 and 3 (Figure 5-3B) showed that, for wt-VP1, the proportion of VP1* was considerably smaller than that of VP1. In contrast, for VP1-H3-H190-H190, lanes 5 and 6 in the figure showed that the proportion of VP1* was almost equal to that of VP1. This comparison showed that VP1-H3-H190-H190 had a higher VP1* to VP1 ratio than wt-VP1. In other words, after thrombin proteolysis, the quality of VP1-H3-H190-H190 was lower than that of wt-VP1.

The higher ratio of VP1* to VP1 for VP1-H3-H190-H190 indicated that, the thrombin secondary cleavage site at Arg58 of VP1 was more accessible in VP1-H3-H190-H190 than in wt-VP1. The different way in which thrombin accessed Arg58 in wt-VP1 and VP1-H3-H190-H190 may indicate structural alterations between these proteins due to modularisation of two copies of hydrophobic H190. A possible structural alteration due to modularisation of an antigen module has also been highlighted in Chapter 4 (Section 4.3.1) of this thesis. The chapter compared the stability of modular VP1 bearing three to five copies of H190 in Assembly Buffer 1, which has been optimised for wt-VP1. The results in the chapter show that modular VP1 bearing four and five copies of H190 were less stable than modular VP1 bearing three copies of H190. Analysis in the chapter interpreted this stability difference as a reflection of protein-protein interactions in a concentrated salt solution, which is dependent on the specific ion effects. Furthermore, the impacts of specific ion effects on a protein are determined by the amino acid sequence of

the protein. Modularisation of an antigen module can change the amino acid sequence of VP1, which can be attributed to (i) the modularised antigen module itself, as well as (ii) structural perturbation of the VP1 protein. Structural perturbation can result in the exposure of buried residues on the surface, and consequently changes in the amino acid sequence exposed on the surface.

After incubation with thrombin, digested protein was centrifuged to remove any precipitate, and the supernatant was loaded onto an SEC column to separate capsomeres from soluble aggregates and GST tags. Figure 5-3C shows size-exclusion chromatograms for wt-VP1 and VP1-H3-H190-H190. The chromatogram of wt-VP1 consisted of three major peaks: (i) an excluded peak corresponding to soluble aggregates, which eluted at about 8.5 ml, (ii) a peak at about 11.5 ml corresponding to capsomeres, and (iii) a peak at about 16 ml corresponding to the released GST tag (Middelberg et al. 2011). Differently, for VP1-H3-H190-H190, only two peaks were observed. These peaks corresponded to soluble aggregates and GST tags. A peak corresponding to capsomeres was not observed, indicating that, following removal of the GST tag, VP1-H3-H190-H190 formed soluble aggregates under this buffer condition.

Results in Figure 5-3 show two major problems associated with modular VP1 bearing two copies of H190 from A/Victoria/210/2009. Firstly, the modular VP1 formed soluble aggregates after GST-tag removal in the standard buffer conditions optimised for wt-VP1; hence, modular VP1 capsomeres could not be obtained. Secondly, the modular VP1 had a lower quality following thrombin proteolysis for removal of the GST-tag from the modular VP1. The lower quality was assessed from the increased amount of secondary digestion product, designated VP1*. Before conducting investigations on these two problems, it is essential to note that the solubility analysis of GST-tagged VP1-H3-H190-H190 in Figure 5-3A shows that the protein was highly expressed but had a relatively low fractional solubility. Structural, functional, and biochemical studies of proteins require proteins with good quality (Gopal and Kumar 2013). Therefore, there was a need to improve the solubility of GST-tagged VP1-H3-H190-H190.

The formation of aggregates *in vivo* is determined by the rate of expression, folding, and aggregation. *E. coli* has a high rate of protein translation and transcription, meaning that protein concentration in the cytoplasm could reach 300-400 mg ml⁻¹. In such an environment, protein folding is a great challenge, especially for high molecular weight proteins, which have slower folding kinetics and often require assistance during folding. Failure to achieve native folding in a timely manner can prompt the proteins to aggregate before folding (Esposito and Chatterjee 2006, Francis and Page 2010).

Various methods have been used to maximise production of soluble recombinant protein in *E. coli*. These methods can be classified into two main streams: (i) strategies where proteins of interest are modified, and (ii) strategies where proteins of interest are not modified. In the first stream, target proteins are genetically engineered to improve their solubility. Such strategies include the use of fusion tags and the generation of soluble variants through genomics and proteomics. In contrast, strategies in the second stream avoid any modification of target proteins. Strategies classified in this stream include reduced expression temperature, the selection of *E. coli* strains, modification of cultivation strategies (Sørensen and Mortensen 2005), the use of chaperones that interact with folding intermediates, and the use of foldases to accelerate rate-limiting steps during folding (Baneyx 1999).

This chapter selected a strategy from the second stream, which is reduced expression temperature, in order to increase the fractional solubility of GST-tagged VP1-H3-H190-H190. A decrease in expression temperature can slow down the growth rate of bacteria, resulting in a decreased amount of cells. It also reduces the transcription and translation rate. Consequently, the concentration of protein in the cytoplasm is decreased, and therefore proper folding and protein solubility can be improved (Baneyx and Mujacic 2004, Esposito and Chatterjee 2006, Francis and Page 2010). Additionally, the use of a low expression temperature can also reduce the strength of hydrophobic interactions, which can cause the incorrect folding of proteins (Baneyx and Mujacic 2004).

In an attempt to improve the solubility of GST-tagged VP1-H3-H190-H190, the expression temperature for GST-tagged modular VP1 was lowered from 26°C to 12°C. The density of harvested cells following expression at both temperatures was measured as absorbance at 600 nm (OD_{600 nm}). The measurements were compared, and the results are shown in Table 5-4. For both wt-VP1 and VP1-H3-H190-H190, OD_{600 nm} after expression at 12°C was lower than OD_{600 nm} at 26°C. OD_{600 nm} at 12°C was about one-sixth the OD_{600 nm} at 26°C. The decrease in OD_{600 nm} following reduction in the expression temperature agreed with results from the aforementioned studies, which stated that the bacterial replication rate of *E. coli* is slower at a lower expression temperature.

Table 5-4. Cell density of harvested cells following expression at 12°C and 26°C.

Constructs	OD_{600 nm}, 12°C	OD_{600 nm}, 26°C
wt-VP1	1.09	6.64
VP1-H3-H190-H190	1.18	6.14

Harvested cell pellets were then resuspended in 40 ml of Lysis Buffer containing 200 mM NaCl for cell lysis using sonication. After sonication, the suspension was clarified using centrifugation. Samples were taken before and after centrifugation, and were analysed using SDS-PAGE. Figure 5-4A and B show the SDS-PAGE analyses of samples at 26°C and 12°C, respectively. Comparison of total fractions for wt-VP1 in Figure 5-4A and B showed that the total fraction of wt-VP1 obtained from expression at 26°C was higher than that obtained at 12°C. Similar observations were also reported for the construct VP1-H3-H190-H190. The lower total fraction obtained from expression at 12°C was highly likely to be a result of a lower cell density. As described in Section 5.2.3, the harvested cells from 26°C and 12°C were resuspended in the same volume of Lysis Buffer. Because expression at 12°C resulted in a lower OD_{600 nm}, the total fractions detected on the gels were also smaller. Furthermore, comparison of total fractions of GST-tagged VP1-H3-H190-H190 and wt-VP1 in Figure 5-4A showed that when expressed at 26°C, the total fraction of the modular construct was similar to that of wt-VP1. Similarly, the comparison in Figure 5-4B showed that when expression was performed at 12°C, both GST-tagged wt-VP1 and VP1-H3-H190-H190 had similar total fractions. This observation confirmed the finding shown in Figure 5-3A that modularisation of two copies of H190 bearing a hydrophobic stretch did not affect the expression of the modular VP1.

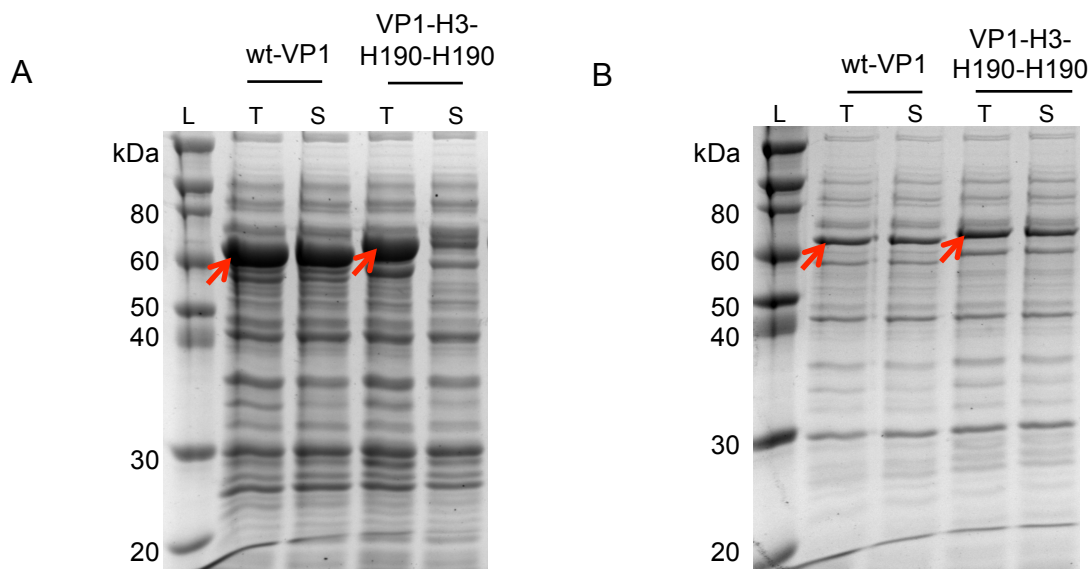


Figure 5-4. SDS-PAGE analysis showing the effect of overnight expression temperature on the solubility of GST-tagged wt-VP1 and modular VP1 bearing two copies of H190 from A/Victoria/2010/2009 (construct VP1-H3-H190-H190). (A) 26°C, and (B) 12°C. Lanes: (L) Pre-stained protein marker, (T) Total protein, and (S) Soluble protein. Red arrows refer to GST-tagged VP1.

Figure 5-4 also shows that GST-tagged VP1-H3-H190-H190 exhibited different solubility following expression at different temperatures. Figure 5-4A shows that a small proportion of GST-tagged VP1-H3-H190-H190 was found in the soluble fraction, indicating that the majority of the protein precipitated. In contrast, the majority of GST-tagged wt-VP1 was found in the soluble fraction. This observation showed that when expression was performed at 26°C, the solubility of VP1-H3-H190-H190 was considerably lower than wt-VP1. In contrast, Figure 5-4B shows that, for both GST-tagged wt-VP1 and VP1-H3-H190-H190, the majority of the proteins were found in soluble fractions. This observation showed that when expression was performed at 12°C, the solubility of GST-tagged VP1-H3-H190-H190 and wt-VP1 was similar. The results in Figure 5-4 show that reduced expression temperature improved the solubility of GST-tagged modular VP1 bearing two copies of H190 from A/Victoria/210/2009, confirming the above rationale that lowering the temperature during protein expression can increase protein solubility.

5.3.2 The use of charged residues

5.3.2.1 Effect of flanking ionic elements on the solubility of modular VP1

Section 4.3.1 of this thesis explained that many protein phenomena, such as protein solubility and the formation of aggregates, are driven by protein-protein interactions (Curtis et al. 2002a). Thus, this section reasons that the formation of soluble aggregates of VP1-H3-H190-H190 following GST-tag removal may also reflect protein-protein interactions.

As explained in Section 4.3.1, in a dilute aqueous salt solution (<0.1 M), protein-protein interactions can be modelled using DLVO theory (Deryaguin and Landau 1941, Verwey and Overbeek 1947), which is comprised of three types of forces: (i) the hard-sphere potential (W_{hs}), (ii) the attractive van der Waals force (W_{disp}), and (iii) the electric double-layer repulsion (W_{elec}). In the theory, protein-protein interactions are expressed in terms of the free energy between two protein molecules (W_{22}) as a function of the centre-to-centre separation (r) (Curtis et al. 1998, Curtis et al. 2002b).

$$W_{22}(r) = W_{hs}(r) + W_{disp}(r) + W_{elec}(r) \quad (1)$$

Among these three forces, only the electric double-layer repulsion is affected by the ionic strength of salt ions (I) in a given solution. The ionic strength of salt ions is determined by both molar concentration of salt ions and the charge number of salt ions. The ionic strength of salt ions and the repulsion force are connected by the Debye length, which is given by:

$$\kappa^2 = (2e^2 N_A I) / (kT \epsilon_0 \epsilon_r) \quad (2)$$

where κ is the inverse of the Debye length, e is the elementary charge, N_A is the Avogadro's number, k is the Boltzmann's constant, T is temperature, ϵ_0 is dielectric permittivity of free space, and ϵ_r is the relative dielectric permittivity of water. The Debye length is approximately 100 Å for 10^{-3} M electrolytes. It decreases to 10 Å for 10^{-1} M electrolytes. In a biological environment, such as in a concentrated salt solution, the Debye length is reduced to <8 Å (Ninham 1999).

If the concentration of electrolyte is very low, less than about 100 mM, and salt ions therefore behave as point charges, the electric double-layer repulsion as a function of the Debye length can be approximated using the Debye-Hückle equation (Verwey and Overbeek 1947). This equation is given by:

$$W_{elec}(r) = \frac{z^2 e^2 (1/r) \exp[-\kappa(r-d_2)]}{4\pi \epsilon_0 \epsilon_r (1+\kappa d_2/2)^2} \text{ for } r > (d_2 + 2\sigma) \quad (3)$$

where z is the valence of the protein.

While the DLVO theory accurately models protein-protein interactions at low salt concentrations (10^{-3} M – 5×10^{-2} M), it fails to predict protein solubility in concentrated salt solutions. This is because in such solution conditions, protein-protein interactions are affected by specific ion effects, which are determined by the position of anions and cations on the lyotropic (Hofmeister) series (Curtis et al. 2002a). The DLVO theory treats salt ions as point charges. Therefore, it cannot differentiate between the size and type of salt ions.

To be applicable for concentrated salt solutions, non-DLVO forces have been defined and added into Eq. 1. These non-DLVO forces are called solvation forces (W_{solv}). Thus, W_{22} is given by:

$$W_{22}(r) = W_{hs}(r) + W_{disp}(r) + W_{elec}(r) + W_{solv}(r) \quad (4)$$

W_{solv} is given by:

$$W_{solv}(r) = \begin{cases} -A(r)(f_a \sigma_a + f_p \sigma_p) & \text{for } d_2 < r < d_2 + r_c \\ 0 & \text{for } r > d_2 + r_c \end{cases} \quad (5)$$

where A is solvent accessible surface area, f_a and f_p are the surface fractions of nonpolar and polar groups, and σ_a and σ_p are the characteristic surface energies of nonpolar and polar groups in contact in water (Curtis et al. 1998, Curtis et al. 2002b).

According to Eq. 5, solvation forces are composed of: (i) hydration effects, and (ii) hydrophobic effects. The hydration effects are the forces between hydrophilic surfaces, which are created by charged or polar residues; they absorb water. Therefore, they create repulsion and increase the distance between two hydrated protein molecules in water. Consequently, the free-energy changes required to remove water molecules around the polar or charged residues are positive (Curtis et al. 2002b). In contrast, hydrophobic effects are created by the attractive interactions between non-polar residues in water (Israelachvili and Pashley 1984). They decrease the distance between two non-polar protein surfaces in water, and consequently increase the free-energy change required to remove water molecules around the non-polar residues (Curtis et al. 2002b).

Although hydration effects are the most widely studied force, the origin of hydration effects is still debatable. Currently, there are two proposed mechanisms. Firstly, the effects possibly result from the anomalous polarisation of water near the interfaces. Such polarisation consequently alters the dielectric responses of water. This mechanism suggests that hydration effects originate from an electrostatic interaction. Although computer simulations can prove such an anomalous dielectric response of water, the simulations indicate that the observed phenomenon is different from the proposed electrostatic theory. Additionally, experimental evidences regarding changes in water structure or dielectric responses of water close to the interfaces are not yet available. Secondly, the hydration effects are suggested to result from the entropic (osmotic) repulsion of thermally excited molecular groups that protrude from the surfaces. However, this concept is invalid for charged systems. Additionally, computer simulations predict that, in neutral systems, protrusions are insignificant (Liang et al. 2007).

While many possible mechanisms have been proposed for hydration effects, currently, the origin and characteristics of hydrophobic effects are still unknown. Generally accepted theories on how the effects have developed are also not yet available. However, many believe that the effects are a result of overlapping solvation zones when two hydrophobic species are in close proximity (Liang et al. 2007).

The results in Section 5.3.1 show that modularisation of H190 containing a hydrophobic stretch on the MuPyV VP1 resulted in the formation of soluble aggregates after GST-tag removal in the standard buffer conditions optimised for wt-VP1. Hence, modular VP1 capsomeres could not be obtained. Based on the rationale explained in

Section 5.1, this section explored the use of glutamic acids as ionic elements within an antigen module. The ionic elements were designed to flank the two copies of H190 tandem repeats, and were comprised of four glutamic acid residues (see Section 5.2.1). The addition of flanking ionic elements is hypothesised to minimise the formation of soluble aggregates VP1-H3-H190-H190 after GST-tag removal by affecting electrostatic double-layer repulsion, as well as solvation forces. The addition of flanking ionic elements may change the surface charge of the protein. Consequently, it may affect protein valence (z) as well as the Debye length (κ). According to Eq. 2 and 3, changes in z and κ will result in changes in the electric double-layer repulsion that are expected to reduce the tendency of the modular VP1 to form soluble aggregates. The addition of the ionic elements into the protein may also tune the surface fraction of nonpolar and polar groups (f_a and f_p), which, according to Eq. 5, will affect the solvation forces. Changes in solvation forces are also expected to reduce the tendency of modular VP1 to form soluble aggregates.

In this chapter, investigation of the effects of glutamic acids was performed using established purification processes, which were developed for wt-VP1. These purification processes used Lysis Buffer (see Section 5.2.3 for details), which contains 200 mM NaCl. Based on its salt concentration, this buffer is classified as a concentrated salt solution. According to Eq. 2 and 3, such buffer conditions can suppress the Debye length and electric double-layer repulsion. Thus, the effects of the addition of glutamic acids may not be observable. Therefore, in order to observe the effects of glutamic acids, the salt concentration in Lysis Buffer was varied using two concentrations: (i) 50 mM NaCl, and (ii) 200 mM NaCl. Lysis Buffer containing 50 mM NaCl was designated as LS Buffer and Lysis Buffer containing 200 mM NaCl was designated HS Buffer. The type of salt ions in the buffer was not changed.

An antigen module comprising two copies of H190 from A/Victoria/210/2009, flanked with four glutamic acids, was genetically inserted into MuPyV VP1, yielding the modular construct VP1-H3-H190-H190-4E (Section 5.2.1). The effects of the addition of glutamic acids were observed by comparing the modular constructs VP1-H3-H190-H190-4E and VP1-H3-H190-H190. Both modular constructs were expressed as GST-tagged VP1 side-by-side, with wt-VP1 as a control (Section 5.2.3). Based on the results in Figure 5-4, the expression of these three constructs was performed at 12°C. Following overnight expression, the cultures were harvested, and cell pellets were then resuspended in 40 ml of LS Buffer or HS Buffer for cell lysis using sonication. After sonication, the suspensions were clarified using centrifugation. Samples were taken before and after centrifugation, and were analysed using SDS-PAGE.

Figure 5-5 shows SDS-PAGE analysis of expression and fractional solubility of GST-tagged wt-VP1, VP1-H3-H190-H190, and VP1-H3-H190-H190-4E at 200 mM and 50 mM NaCl. The comparison of total fractions for these three constructs showed that VP1-H3-H190-H190-4E had a similar total fraction to wt-VP1 and VP1-H3-H190-H190. This result indicates that the expression of VP1-H3-H190-H190-4E was similar to that of wt-VP1 and VP1-H3-H190-H190. This result suggests that the use of glutamic acids did not affect the expression of modular VP1. Alternatively, it was possible that the use of glutamic acid affected the expression of modular VP1 but that the effects were only noticeable when more than four glutamic acids were used. No reports have been found on studies of modular VLPs, in which glutamic acid was used for solubility improvement. Therefore, the correlation between the number of glutamic acid residues present and the expression of the modular VP1 observed in this study cannot be compared with other studies. The most similar report available for comparison is the study by (Jung et al. 2011). In that study, arginine was used to improve the solubility of CalB protein. Arginine was used as a fusion tag in the C-terminal of the CalB protein. In contrast to the findings in this chapter, the study by Jung et al. showed that the number of arginine residues present affected the expression level of the protein. CalB protein bearing zero and two arginine residues had similar total protein levels. However, increasing the number of arginines to six and ten resulted in a lower total protein. A further decrease in the total protein of CalB was observed when the number of arginine residues was increased to eleven. The difference between the findings in this chapter and those reported by Jung et al. could be protein-specific. The difference was also possible because this chapter used different charged residues in a different format.

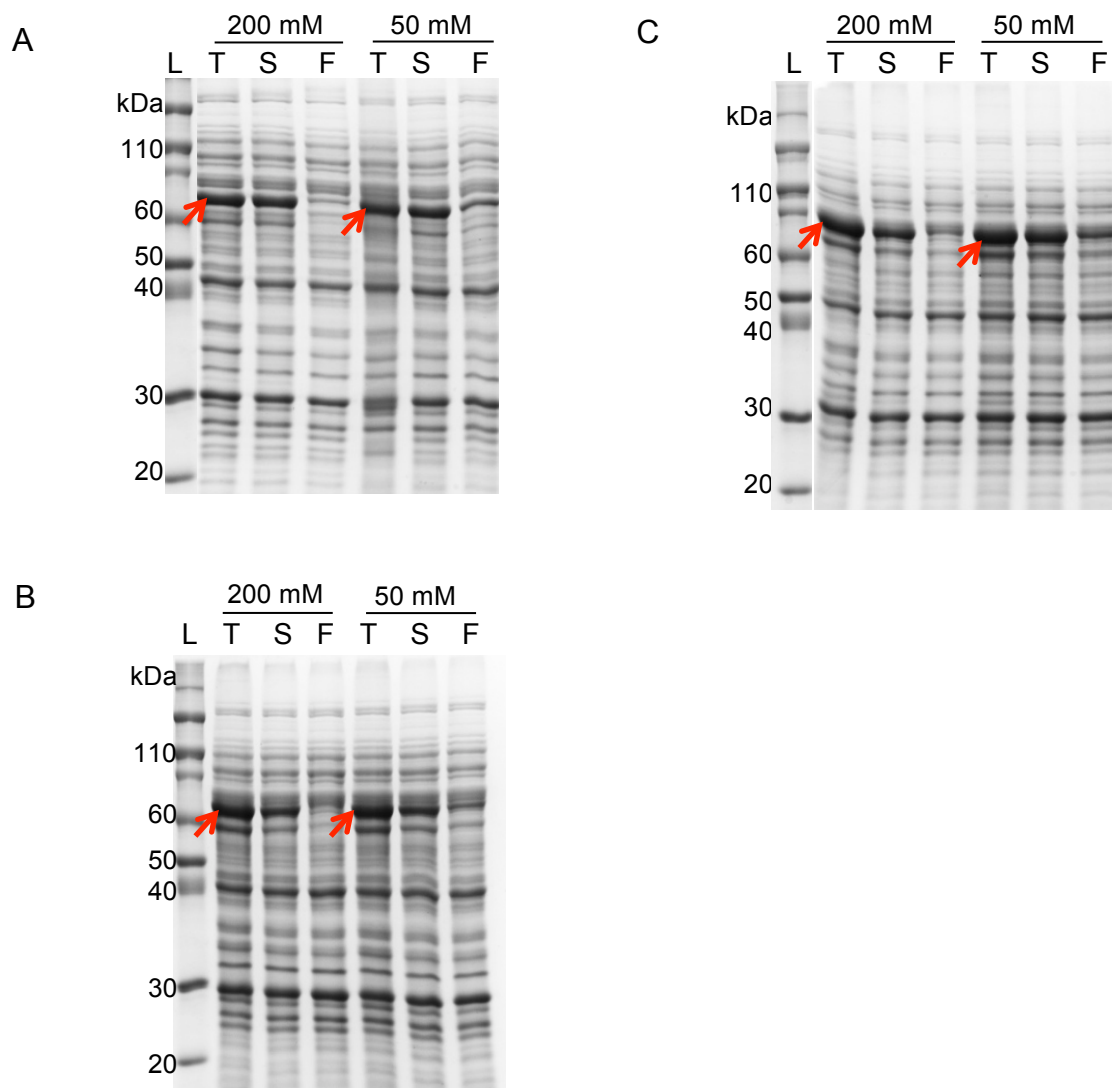


Figure 5-5. SDS-PAGE analysis showing the effects of the addition of glutamic acids on the expression and solubility of GST-tagged VP1. (A) wt-VP1, (B) Modular VP1 bearing two copies of H190 from A/Victoria/210/2009 without flanking glutamic acids (construct VP1-H3-H190-H190), and (C) Modular VP1 bearing two copies of H190 from A/Victoria/210/2009 with flanking glutamic acids (construct VP1-H3-H190-H190-4E). Lanes: (L) Pre-stained protein marker, (T) Total fraction, (S) Soluble fraction, and (F) Flow-through fraction from GST column during purification of GST-tagged VP1. Red arrows refer to the GST-tagged VP1 proteins.

Comparison of soluble fractions for wt-VP1 (Figure 5-5A) showed that the soluble fractions of GST-tagged wt-VP1 at 200 mM and 50 mM NaCl were similar, indicating that the salt concentration did not affect the solubility of GST-tagged wt-VP1. Similarly, Figure 5-5B shows that GST-tagged VP1-H3-H190-H190 had a similar soluble fraction at both salt concentrations. Furthermore, comparison between soluble fractions of wt-VP1 (Figure 5-5A) and VP1-H3-H190-H190 (Figure 5-5B) at both salt concentrations showed that VP1-H3-H190-H190 had a lower solubility than wt-VP1. This result is similar to the findings in

Figure 5-3, confirming that modularisation of H190 bearing hydrophobic stretch resulted in the reduced solubility of GST-tagged modular VP1.

In contrast, Figure 5-5C shows that the soluble fraction for VP1-H3-H190-H190-4E at 200 mM NaCl was slightly smaller than the soluble fraction at 50 mM NaCl. This observation suggested that GST-tagged VP1-H3-H190-H190-4E was less soluble at 200 mM NaCl than at 50 mM NaCl. Because changes in the salt concentration did not affect the solubility of VP1-H3-H190-H190, the increase in the solubility of VP1-H3-H190-H190-4E with a decrease in the salt concentration was suggested as an effect of the addition of the flanking ionic elements. This suggestion was supported by comparing soluble fractions of VP1-H3-H190-H190 and VP1-H3-H190-H190-4E at each salt concentration. The comparison showed that, at 200 mM NaCl, the soluble fractions of modular VP1 with and without flanking were similar. In contrast, at 50 mM NaCl, VP1-H3-H190-H190-4E had a larger soluble fraction than VP1-H3-H190-H190. These observations confirmed that modular VP1 bearing the flanking ionic elements, VP1-H3-H190-H190-4E, had an improved solubility. However, the improvement was more visible at 50 mM NaCl than 200 mM NaCl, implying that the effects of glutamic acids were dependent on the salts concentration.

The higher solubility of VP1-H3-H190-H190-4E at 50 mM NaCl than at 200 mM NaCl may reflect how ionic strength affects the electrostatic double-layer repulsion, as explained in the beginning of this section. As glutamic acid is a charged residue, the addition of this residue into modular VP1 bearing two copies of H190 tandem repeats may affect (i) the Debye length, as well as (ii) the valence of modular VP1 proteins. These two variables determine the strength of electrostatic double-layer repulsion (Eq. 2). At a high concentration of salt, such as 200 mM NaCl, the Debye length was suppressed, and consequently, the electrostatic double-layer interactions were reduced. At such conditions, the effects of the addition of glutamic acids could not be observed. In other words, the solubility of VP1-H3-H190-H190-4E and VP1-H3-H190-H190 appeared to be similar. In contrast, at a lower salt concentration, in this case 50 mM NaCl, the Debye length was increased. Consequently, the electrostatic double-layer interactions were increased, and the effects of the addition of glutamic acids could be observed.

The addition of glutamic acid into modular VP1 bearing two copies of H190 tandem repeats may also affect the proportion of hydrophobic and hydrophilic surface fractions. This variable determines the strength of salvation forces (Eq. 5), which may start to play a role in protein-protein interactions in a concentrated salt solution as explained at the beginning of this section. However, the findings in Figure 5-5 show that the solubility of

GST-tagged VP1-H3-H190-H190 and VP1-H3-H190-H190-4E at a concentration of 200 mM NaCl was similar, indicating that the impact of solvation forces on the solubility of GST-tagged modular VP1 proteins was not observed. This was possibly because the number of added glutamic acids was not sufficient to modify the proportion of the hydrophobic and hydrophilic surface fractions in GST-tagged modular VP1.

GST-tagged wt-VP1, VP1-H3-H190-H190, and VP1-H3-H190-H190-4E were purified using GST-affinity chromatography at respective salt concentrations. Fractions from the flow-through of the GST column were then analysed using SDS-PAGE, as shown in lane F of Figure 5-5. The analysis showed that, for all constructs, flow-through at 50 mM NaCl contained more GST-tagged VP1 than flow-through at 200 mM NaCl. This result indicates that when affinity purification was performed at 50 mM NaCl, GST-tagged VP1 bound more weakly to the ligand. Furthermore, the SDS-PAGE analysis also showed that the addition of glutamic acids did not affect the binding of the GST-tagged VP1-H3-H190-H190-4E to the ligand, as expected.

5.3.2.2 Effect of flanking ionic elements on the ratio of VP1* to VP1 after GST removal

Purified GST-tagged modular VP1 was then incubated with thrombin at respective salt concentrations to remove the GST-tag from the modular VP1. SDS-PAGE analyses of GST removal for wt-VP1, VP1-H3-H190-H190, and VP1-H3-H190-H190-4E at 200 and 50 mM NaCl are shown in Figure 5-6A and B, respectively. The results were analysed for two aspects, i.e. (i) GST-removal efficiency, which is assessed as the proportion of undigested GST-tagged modular VP1, and (ii) the ratio of VP1* to VP1 after GST removal.

Firstly, analysis was conducted to assess GST-removal efficiency. Lanes 2 and 3 of Figure 5-6A show that, after thrombin treatment, a very small amount of GST-tagged wt-VP1 (red arrow; about 69 kDa) was detected in the gel. A similar observation is shown in lanes 2 and 3 of Figure 5-6B. These observations confirmed that GST was almost completely removed from wt-VP1 at both 50 mM NaCl and 200 mM NaCl, indicating that GST-removal efficiency for wt-VP1 was independent of the salt concentration. These observations further suggested that the activity of thrombin at both salt concentrations was similar.

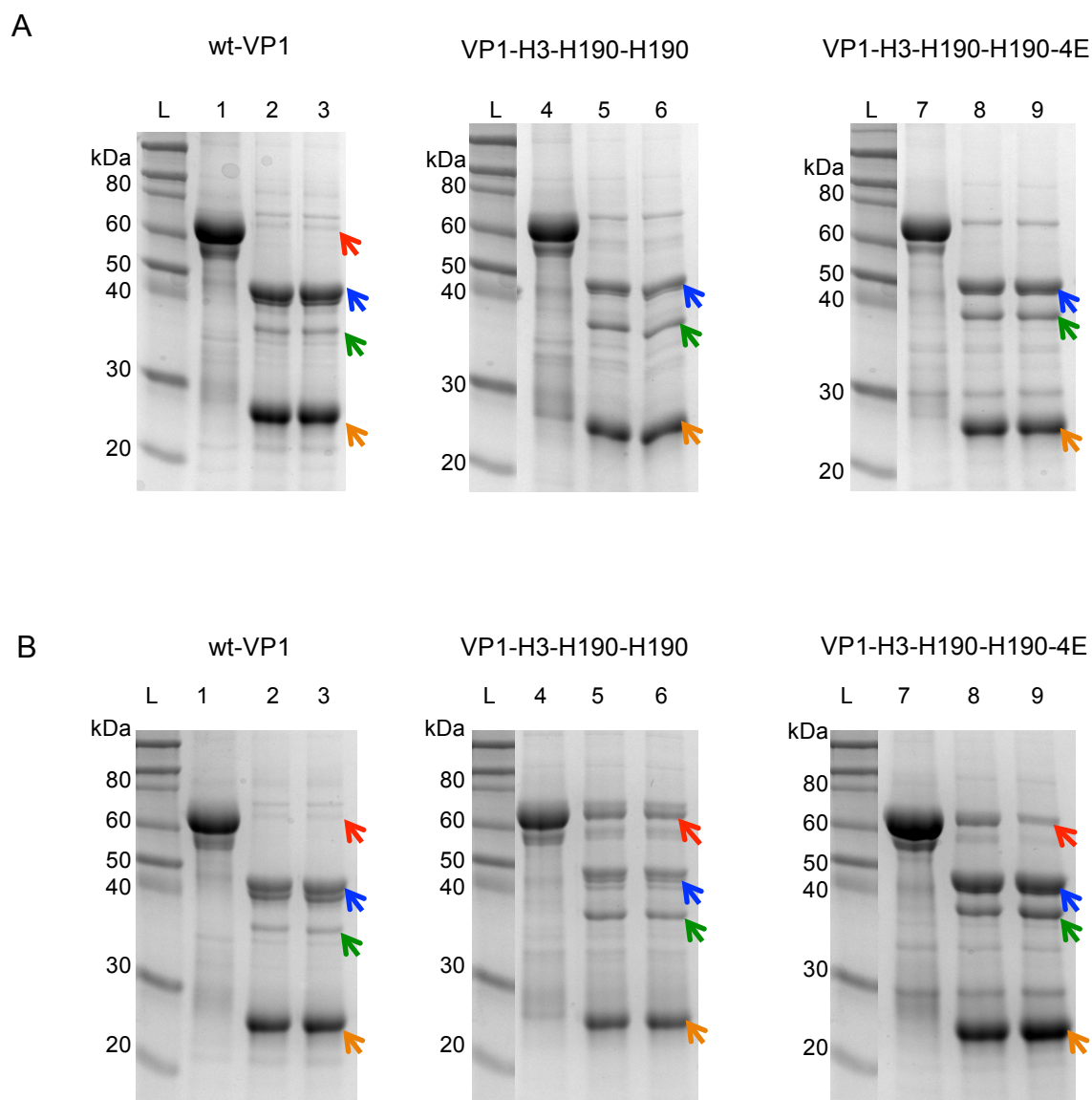


Figure 5-6. SDS-PAGE analysis showing the effect of the addition of glutamic acids on the ratio of VP1* to VP1 following GST removal using thrombin for GST-tagged wt-VP1, modular VP1 bearing two copies of H190 from A/Victoria/210/2009 without flanking ionic elements (construct VP1-H3-H190-H190), and modular VP1 bearing two copies of H190 from A/Victoria/210/2009 with flanking ionic elements (construct VP1-H3-H190-H190-4E) at different salt concentrations. (A) 200 mM NaCl, and (B) 50 mM NaCl. Lanes: (L) Pre-stained protein marker, (1) GST-tagged wt-VP1, (2) Total protein of thrombin-digested GST-tagged wt-VP1, (3) Soluble protein of thrombin-digested GST-tagged wt-VP1, (4) GST-tagged VP1-H3-H190-H190, (5) Total protein of thrombin-digested GST-tagged VP1-H3-H190-H190, (6) Soluble protein of thrombin-digested GST-tagged VP1-H3-H190-H190, (7) GST-tagged VP1-H3-H190-H190-4E, (8) Total protein of thrombin-digested GST-tagged VP1-H3-H190-H190-4E, (9) Soluble protein of thrombin-digested GST-tagged VP1-H3-H190-H190-4E. Red arrows refer to GST-tagged VP1. Blue arrows refer to VP1. Green arrows refer to the secondary thrombin cleavage product. Orange arrows refer to GST.

In contrast to wt-VP1, observations of GST removal for VP1-H3-H190-H190 and VP1-H3-H190-H190-4E suggested that, for these modular constructs, GST removal by thrombin was lower at reduced salt concentrations. Comparisons between lanes 5 and 6 of Figure 5-6A and B show that, for VP1-H3-H190-H190, GST removal was more efficient at 200 mM NaCl than at 50 mM NaCl. At 200 mM NaCl, bands corresponding to GST-tagged VP1-H3-H190-H190 were not detected in the gel, indicating that GST was removed completely. In contrast, at 50 mM NaCl, a small fraction of GST-tagged VP1-H3-H190-H190 (red arrow) was detected in the gel, indicating that GST removal was not completed at such salt concentrations. Likewise, lanes 8 and 9 of Figure 5-6A and B show that GST removal for VP1-H3-H190-H190-4E was more efficient at 200 mM NaCl than at 50 mM NaCl. At 200 mM NaCl, bands corresponding to GST-tagged VP1-H3-H190-H190-4E (about 73.5 kDa) were not detected in the gel, while a small fraction of GST-tagged VP1-H3-H190-H190-4E (red arrow) was detected in the gel for GST removal at 50 mM NaCl. These results indicate that GST removal efficiency for both modular constructs were lower at 50 mM NaCl than at 200 mM NaCl. The preceding paragraph has established that the analysis of GST removal for wt-VP1 suggested that changes in salt concentration did not affect thrombin activity. Therefore, the effect of salt concentration on the GST removal efficiency observed for both modular constructs was suggested to be due to the presence of hydrophobic stretches. It is possible that at different salt concentrations, the presence of hydrophobic stretches within the GST-tagged modular VP1 proteins affected the structures of the proteins, and consequently affected the accessibility of the engineered thrombin cleavage site. Furthermore, the comparison of GST removal for VP1-H3-H190-H190-4E and VP1-H3-H190-H190 at 200 mM NaCl showed that there was no noticeable difference in the efficiency of thrombin for both modular constructs. A similar observation was also noted for both modular constructs at 50 mM NaCl. This observation suggested that the added flanking ionic elements did not have any significant effect on GST-removal efficiency, which is expected because the glutamic acids were inserted at the *Afel* recognition site close to position 293 of VP1 (see Section 4.2.1), while the thrombin cleavage site was in the N-terminal of VP1.

Secondly, analysis was performed to assess the ratio of VP1* to VP1 after GST removal. The digestion profiles of wt-VP1 (lanes 2 and 3) and VP1-H3-H190-H190 (lanes 5 and 6) in Figure 5-6A show that three products were observed after thrombin treatment at 200 mM: (i) VP1 protein (blue arrows), (ii) VP1* (green arrows), and (iii) GST tag (orange arrows). The ratio of VP1* to VP1 for wt-VP1 was lower than the ratio for VP1-H3-H190-H190. This result is similar to the digestion profiles of both constructs reported in

Section 5.3.1 (Figure 5-3B), although wt-VP1 and VP1-H3-H190-H190 in this section were expressed at different temperatures to those in Section 5.3.1. In this section, wt-VP1 and VP1-H3-H190-H190 were expressed at 12°C, and the modular GST-tagged VP1 exhibited a similar solubility to GST-tagged wt-VP1 at 200 mM NaCl. In contrast, in Section 5.3.1, wt-VP1 and VP1-H3-H190-H190 were expressed at 26°C, and the modular GST-tagged VP1 exhibited a considerably lower solubility than GST-tagged wt-VP1 at 200 mM NaCl. The similarities of digestion profiles between VP1-H3-H190-H190 in this section and that reported in Section 5.3.1 therefore suggested that an improvement in the solubility of GST-tagged VP1-H3-H190-H190 did not affect the ratio of VP1* to VP1 for VP1-H3-H190-H190 after GST removal using thrombin. Furthermore, the digestion profiles of wt-VP1 and VP1-H3-H190-H190 at 200 mM shown in Figure 5-6A were similar to those at 50 mM NaCl, as shown in the corresponding lanes of Figure 5-6B. After digestion at the lower salt concentration, the three major products were observed, and the ratio of VP1* to VP1 for wt-VP1 was also lower than the ratio for VP1-H3-H190-H190. This observation suggested that a decrease in the salt concentration does not result in changes in the ratio of VP1* to VP1 for either construct.

Similar to wt-VP1 and VP1-H3-H190-H190, GST removal of GST-tagged VP1-H3-H190-H190-4E at 50 mM and 200 mM NaCl (Figure 5-6, lanes 8 and 9) resulted in the three products: (i) VP1 protein (blue arrows), (ii) VP1* (green arrows), and (iii) GST tag (orange arrows). However, unlike the two aforementioned constructs, the digestion profiles of VP1-H3-H190-H190-4E were affected by salt concentration. The ratio of VP1* to VP1 at 200 mM NaCl was slightly higher than the ratio at 50 mM NaCl. Furthermore, the SDS-PAGE analysis in Figure 5-6A showed that at 200 mM NaCl, the ratio of VP1* to VP1 for VP1-H3-H190-H190-4E (lanes 8 and 9) was lower compared to VP1-H3-H190-H190 (lanes 5 and 6). Similarly, Figure 5-6B showed that VP1-H3-H190-H190-4E had a lower VP1* to VP1 ratio than VP1-H3-H190-H190 at 50 mM NaCl. These findings indicated that the addition of glutamic acid residues could reduce the ratio of VP1* to VP1, and consequently improve the quality of modular VP1 bearing two copies of H190 after GST removal. Although the effects were not considerable, they were more noticeable at the lower salt concentration than at the higher one. This dependency of VP1* to VP1 ratio on salt concentration is in an agreement with the observation reported in Section 5.3.2.1, that the solubility of GST-tagged VP1-H3-H190-H190-4E was higher than the solubility of GST-tagged VP1-H3-H190-H190, and the solubility improvement was more noticeable at 50 mM NaCl than at 200 mM NaCl. In the section, the solubility improvement and its dependency on salt concentration has been linked to how the addition of glutamic acid

may affect the electric double-layer repulsion included in DLVO theory. Therefore, it was appealing to conclude that the lower VP1* to VP1 ratio observed for VP1-H3-H190-H190-4E correlated with improved solubility. Nevertheless, it has been observed that, for VP1-H3-H190-H190, improvements in the solubility of GST-tagged VP1-H3-H190-H190 did not increase the quality of the modular VP1 after GST removal. Therefore, the increase in the quality of VP1-H3-H190-H190-4E after GST removal may not be correlated with the solubility improvement of GST-tagged VP1-H3-H190-H190-4E.

5.3.2.3 Effect of flanking ionic elements on the aggregation of modular VP1 and isolation of modular VP1 capsomeres

After incubation with thrombin, digested proteins were clarified using centrifugation to remove precipitates. The supernatants were then loaded into SEC to separate capsomeres from aggregates and GST tags. Figure 5-7 shows the size-exclusion chromatograms of wt-VP1, VP1-H3-H190-H190, and VP1-H3-H190-H190-4E at 200 and 50 mM NaCl. The figure also shows SDS-PAGE analysis of peaks observed in the chromatograms.

Figure 5-7A shows the SEC chromatograms for wt-VP1 at 200 mM and 50 mM NaCl. At 200 mM NaCl, three typical peaks were observed: (i) an excluded peak corresponding to soluble aggregates (A1), eluted at about 8.5 ml, (ii) a peak at about 11.5 ml corresponding to capsomeres (A11), and (iii) a peak at about 16 ml corresponding to the released GST tag (Middelberg et al. 2011). This result is as reported in Section 5.3.1, although wt-VP1 in this section was expressed at 12°C and wt-VP1 in Section 5.3.1 was expressed 26°C. In contrast, at 50 mM NaCl, only two peaks were observed. These peaks were: (i) peak A1, containing soluble aggregates, and (ii) a peak containing the GST tag. At this lower salt concentration, a distinct peak containing capsomeres was not observed. Furthermore, peak A1 at 50 mM NaCl was bigger than the peak at 200 mM NaCl. Fractions collected from peaks A1 and A11 were then analysed using SDS-PAGE. The analysis showed that peak A11 majorly contained wt-VP1 (blue arrow), as expected. The analysis also showed that peak A1 at 200 mM NaCl was composed of equal amounts of VP1 (blue arrow) and VP1* (green arrow). Meanwhile, the peak at 50 mM NaCl was predominantly composed of VP1. This result suggests that, at 200 mM NaCl, the formation of soluble aggregates of wt-VP1 may be promoted by the ratio of VP1 to VP1*, while at 50 mM NaCl, soluble aggregates of wt-VP1 may be due to instability of the VP1 protein at low salt concentrations.

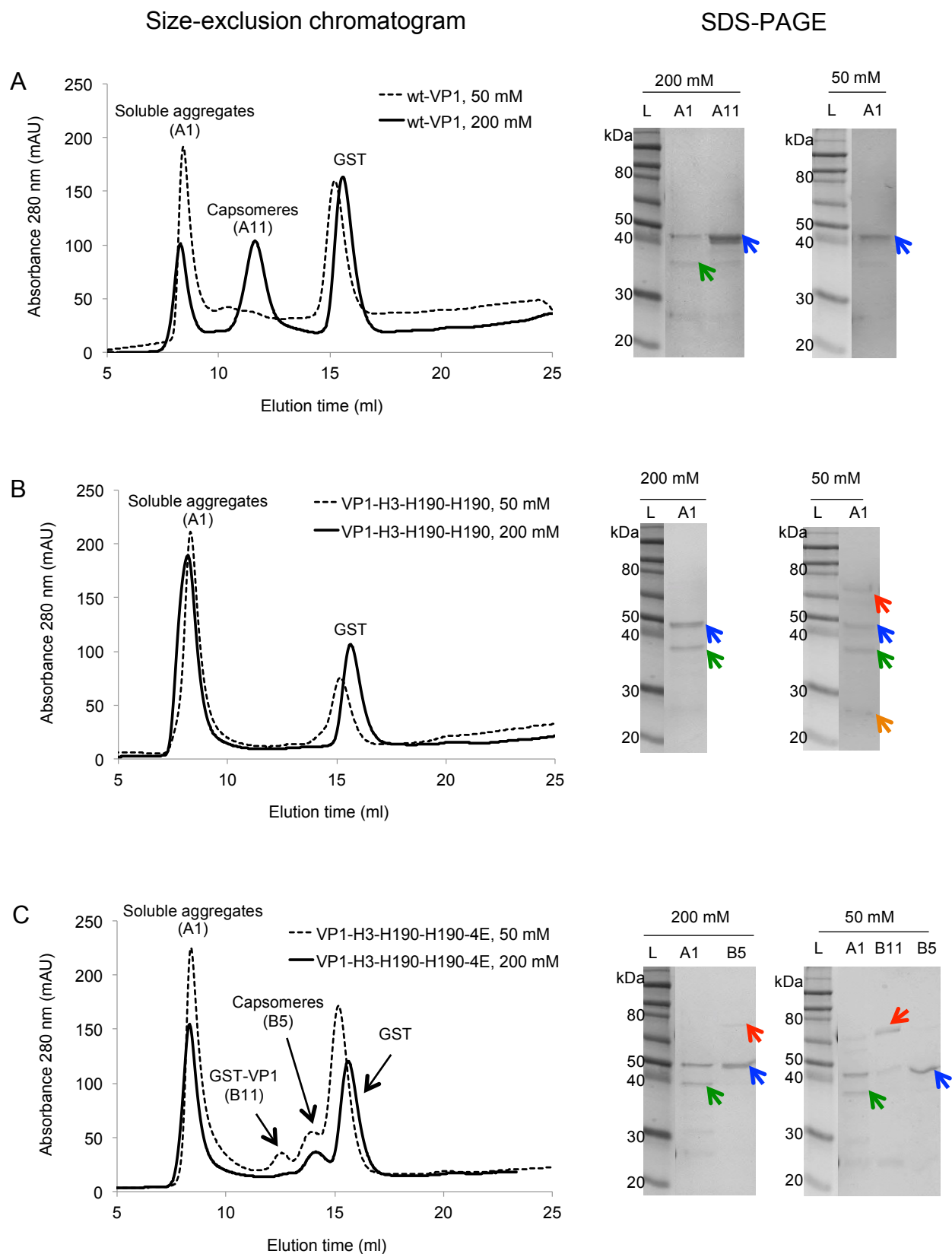


Figure 5-7. Analysis using size-exclusion chromatography and SDS-PAGE showing the effects of the addition of glutamic acids on the aggregation of modular VP1 bearing H190 from A/Victoria/2010/2009 and feasibility to isolate the modular VP1 as capsomeres at 200 mM and 50 mM NaCl. (A) wt-VP1, (B) Modular VP1 bearing two copies of H190 from A/Victoria/210/2009 without flanking ionic elements, and (C) Modular VP1 bearing two copies of H190 from A/Victoria/210/2009 with flanking ionic elements. Red arrows refer to GST-tagged VP1 protein. Blue arrows refer to VP1 protein. Green arrows refer to the secondary thrombin cleavage product. Orange arrows refer to GST.

While three peaks were observed for wt-VP1 at 200 mM, the chromatograms for VP1-H3-H190-H190 in Figure 5-7B shows that, at the same salt concentration, two major peaks were observed. These peaks corresponded to soluble aggregates (A1) and GST. Peaks corresponding to capsomeres were not observed. Analysis of a fraction collected from peak A1 using SDS-PAGE revealed that the soluble aggregates were composed of VP1 (blue arrow) and VP1* (green arrow) at about 1:1 ratio, similar to peak A1 of wt-VP1 at 200 mM NaCl. Furthermore, this chromatogram for VP1-H3-H190-H190 at 200 mM NaCl was similar to that reported in Section 5.3.1. However, in this section, VP1-H3-H190-H190 was expressed at 12°C, and the modular GST-tagged VP1 exhibited a similar solubility to GST-tagged wt-VP1 at 200 mM NaCl. In contrast, in Section 5.3.1, VP1-H3-H190-H190 was expressed at 26°C, and the modular GST-tagged VP1 exhibited a considerably lower solubility than GST-tagged wt-VP1 at 200 mM NaCl. The similarities of the chromatograms for VP1-H3-H190-H190 in this section and that reported in Section 5.3.1 therefore suggested that an improvement in the solubility of GST-tagged VP1-H3-H190-H190 did not reduce the propensity of VP1-H3-H190-H190 to form soluble aggregates. The chromatogram for VP1-H3-H190-H190 at 200 mM NaCl was also similar to that at 50 mM NaCl. The chromatogram for 50 mM showed two main peaks corresponding to soluble aggregates (A1) and GST. SDS-PAGE analysis of the A1 peak showed that the peak was composed of GST-tagged VP1, VP1, VP1*, and GST tag. Furthermore, two negligible peaks were observed at about 12.5 ml and 14 ml. Unfortunately, because these peaks had very low absorbance reading, they could not be analysed in SDS-PAGE, and therefore the content of these peaks remains unknown. Thus, the results in Figure 5-7B indicate that regardless of the variation in salt concentrations, modular protein VP1-H3-H190-H190 was prone to aggregation, which consequently prevented the isolation of modular VP1 capsomeres. This finding further suggested that reducing the salt concentration alone was not sufficient to reduce the aggregation of modular VP1 and increase the feasibility of obtaining modular VP1 capsomeres.

Similar to VP1-H3-H190-H190, the chromatograms for construct VP1-H3-H190-H190-4E at both salt concentrations appear to be different from wt-VP1, as shown in Figure 5-7C. At 200 mM NaCl, three peaks were observed: (i) peak A1, corresponding to soluble aggregates, (ii) peak B5, which was eluted at about 14 ml, and (iii) a peak corresponding to the GST tag. Fractions from these peaks were then analysed using SDS-PAGE. The analysis showed that peak A1 was composed of VP1 (blue arrow) and VP1* (green arrow) at an equal ratio, similar to the observation in VP1-H3-H190-H190.

Furthermore, the analysis showed that peak B5 contained VP1-H3-H190-H190-4E (blue arrow) with a negligible proportion of GST-tagged VP1-H3-H190-H190-4E (red arrow), indicating that the peak corresponded to the elution of capsomeres. Comparisons between the chromatogram of VP1-H3-H190-H190-4E and VP1-H3-H190-H190 at 200 mM showed that the size of peaks corresponding to the GST tag in both chromatograms were similar. However, peak A1 for VP1-H3-H190-H190-4E was smaller than the peak for VP1-H3-H190-H190. This comparison showed that although the amount GST-tagged VP1-H3-H190-H190 and VP1-H3-H190-H190-4E was similar, the amount of soluble aggregates in construct VP1-H3-H190-H190-4E was lower. Furthermore, the chromatogram of VP1-H3-H190-H190-4E shows that peak B5 was eluted considerably later than the wt-VP1 capsomeres (peak A9 in Figure 5-7A), which were eluted at about 11.5 ml. In SEC, proteins are separated based on their hydrodynamic radius. Thus, the shifted elution time may indicate that the hydrodynamic radius of VP1-H3-H190-H190-4E capsomeres was smaller than wt-VP1 capsomeres. In Section 5.3.1, the results in Figure 5-3B showed that the digestion of GST-tagged wt-VP1 and VP1-H3-H190-H190 resulted in a different ratio of VP1* to VP1. This finding suggested that thrombin had different access to its secondary cleavage site in both constructs, possibly due to structural alterations, which resulted from the modularisation of two copies of H190 bearing a hydrophobic stretch. It is possible that such structural alterations not only affected how thrombin accessed its secondary cleavage site, but also modified the tertiary structure, and consequently the hydrodynamic radius of VP1-H3-H190-H190 capsomeres. Alternatively, the later elution time for VP1-H3-H190-H190-4E may have resulted from non-specific interactions between capsomeres and the stationary phase of the SEC column. Such non-specific interactions may also be a result of the aforementioned structural alterations. Unfortunately, data in this chapter were not sufficient to prove the hypothesis regarding structural alterations.

Figure 5-7C also shows that the chromatogram for construct VP1-H3-H190-H190-4E at 200 mM NaCl was similar to that at the lower salt concentration of 50 mM NaCl. While the same amount of GST-tagged VP1-H3-H190-H190-4E was digested at each salt concentration, the three peaks had higher absorbance readings at 50 mM NaCl than those at 200 mM NaCl. In addition, the chromatogram at 50 mM NaCl showed that a fourth peak was observed at about 12.5 ml (B11). SDS-PAGE analysis of a fraction from this peak showed that the negligible peak contained undigested GST-tagged modular protein. This observation probably resulted from the addition of the flanking ionic elements. Repulsion forces from the glutamic acids may become stronger at lower salt concentrations, and the stronger repulsion force may consequently minimise non-specific interactions between

proteins with ligands on the column. The way in which charged residues affect the separation of proteins in SEC has not been previously reported. The most similar study for a comparison in the study by Ejima et al. (Ejima et al. 2005). In their study, Ejima et al. used arginine as a buffer additive, which was different from the use of glutamic acid residues as flanking ionic elements in this chapter. They showed that the addition of arginine to a concentrated salt solution could suppress the hydrophobic binding of proteins in the stationary phase of the column, which was otherwise enhanced at a high salt concentration. As a result, the separation of the proteins could be improved and a larger amount of proteins could be recovered.

Figure 5-7 shows that modular VP1 capsomeres were observed for VP1-H3-H190-H190-4E, but not for VP1-H3-H190-H190 at both salt concentrations. The fact that modular VP1 capsomeres were not observed for VP1-H3-H190-H190 indicated that reducing the expression temperature alone, or changing the salt concentration alone, was not sufficient to minimise the formation of soluble aggregates of the modular VP1 and obtain modular VP1 capsomeres. In contrast, by using glutamic acid as flanking ionic elements, the formation of soluble aggregates was reduced. Furthermore, the flanking ionic elements allowed for the isolation of modular VP1 capsomeres bearing H190 from A/Victoria/210/2009. The modular VP1 capsomeres could be obtained at both 50 mM NaCl and 200 mM NaCl, suggesting that modular VP1 capsomeres could be obtained by using flanking ionic elements, without the need to change the salt concentration of the buffer. These results show that the use of glutamic acid as a flanking ionic element enabled the modification of forces between modular VP1, which otherwise may not be achieved by changing the salt concentration in both buffers, or reducing the expression temperature. As explained in Section 5.3.2.1, the addition of glutamic acids may affect (i) the electrostatic double-layer repulsion, and (ii) the solvation forces. Results in this section may indicate the role of the electrostatic double-layer repulsion from the added glutamic acids. At a lower salt concentration, such as 50 mM NaCl, the electrostatic double-layer repulsion dominates protein-protein interactions. Therefore, it is possible that the repulsion created by glutamic acid residues can overcome the aggregation of modular VP1; thus, modular VP1 capsomeres could be isolated. However, in concentrated salt solutions, such as at 200 mM NaCl, the electrostatic double-layer repulsion is suppressed. Therefore, the fact that modular VP1 capsomeres could still be isolated at 200 mM NaCl may suggest that other forces were involved in the protein-protein interactions. The forces were likely to be solvation forces, as discussed in Section 5.3.2.1. Solvation forces are composed of hydration and hydrophobic effects. The addition of glutamic acid as a flanking ionic

element may modify the proportion of the surface fraction of nonpolar and polar groups (f_a and f_p) in VP1-H3-H190-H190-4E (Eq. 5), and consequently modify the solvation forces.

In summary, the results in Section 5.3.2 indicate three important impacts of the use of flanking ionic elements on the modular VP1 bearing two copies of H190 from A/Victoria/210/2009 in a buffer condition that has been optimised for wt-VP1. These impacts were:

1. The solubility of the GST-tagged modular VP1 bearing the flanking elements was improved, and the improvement was greater at lower salt concentrations.
2. The flanking ionic elements slightly reduced the ratio of VP1* to VP1 after GST removal using thrombin.
3. The modular VP1 bearing the flanking ionic elements had a lowered tendency to aggregate, and consequently the modular VP1 could be obtained as capsomeres.

5.3.3 Mutation from a thrombin to TEVp cleavage site

5.3.3.1 Effect of mutation from a thrombin to TEVp cleavage site on the expression and solubility of GST-tagged modular VP1

Section 5.3.1 above has shown that VP1-H3-H190-H190 had a lower quality than wt-VP1 after enzymatic GST removal using thrombin. The lower quality of the modular VP1 was assigned based on the higher proportion of secondary thrombin digestion product VP1* to VP1. The VP1* fragment of VP1-H3-H190-H190 has a theoretical molecular weight of about 39 kDa. Analysis in Section 5.3.1 suggested that the VP1* of VP1-H3-H190-H190 resulted from the same cause as VP1* of wt-VP1, which was the non-specific thrombin proteolysis at Arg58 of VP1. Connors et al. (Connors et al. 2013) showed that the VP1* fragment could be eliminated by replacing thrombin with tobacco etch virus protease (TEVp). The removal of VP1* was important because it has been shown to promote the formation of mis-formed VLPs. In the absence of this fragment, VLPs with a higher homogeneity could be obtained. Therefore, in order to improve the quality of modular VP1 bearing two copies of hydrophobic H190 following GST removal, this chapter investigated the use of thrombin in comparison to TEVp.

The thrombin cleavage site in constructs wt-VP1, VP1-H3-H190-H190, and VP1-H3-H190-H190-4E was mutated into a TEVp cleavage site (Section 5.2.1), yielding constructs TEVP-wt-VP1, TEVP-VP1-H3-H190-H190, and TEVP-VP1-H3-H190-H190-4E. Construct VP1-H190-H190 (reported in Chapters 3 and 4) was also modified, yielding construct TEVP-VP1-H190-H190, which was used for comparison. The antigen module in construct VP1-H190-H190 bears two copies of H190 from A/California/07/2009 (H1N1).

The modular constructs bearing a TEVp cleavage site were then expressed and purified side-by-side with those bearing a thrombin cleavage site. Based on the results shown in Figure 5-4, the overnight protein expression was conducted at 12°C. Furthermore, based on the findings in Section 5.3.2.3, purification of modular VP1 was performed using LS buffer, which was Lysis buffer containing 200 mM NaCl.

Figure 5-8 compares SDS-PAGE analysis for the expression and solubility of GST-tagged modular VP1 bearing thrombin or TEVp cleavage sites. The expression and solubility of thrombin-modular constructs were consistent with previous results outlined in Section 5.3.2. Figure 5-8A shows that as reported in Section 5.3.2, the solubility of GST-tagged VP1-H3-H190-H190 and VP1-H3-H190-H190-4E in the buffer optimised for wt-VP1 was lower than that of GST-tagged wt-VP1. In contrast, the proportion of soluble to total fraction of VP1-H190-H190 was similar to wt-VP1, suggesting that GST-tagged VP1-H190-H190 had a higher solubility than VP1-H3-H190-H190 and VP1-H3-H190-H190-4E.

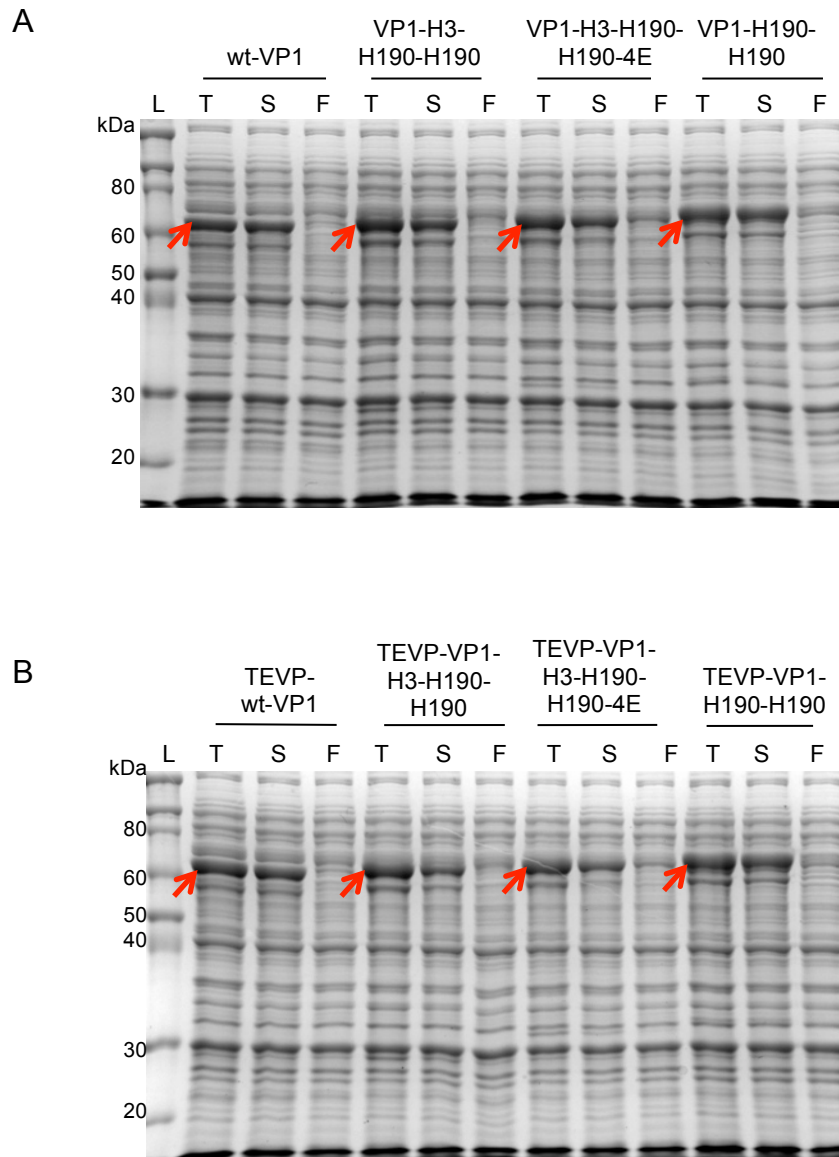


Figure 5-8. SDS-PAGE analysis showing the effect of mutation from a thrombin to TEVp cleavage site on the solubility of GST-tagged VP1 proteins. (A) wt-VP1, modular VP1 bearing two copies of H190 from A/Victoria/210/2009 without flanking glutamic acids (construct VP1-H3-H190-H190), modular VP1 bearing two copies of H190 from A/Victoria/210/2009 with flanking glutamic acids (construct VP1-H3-H190-H190-4E), and modular VP1 bearing two copies of H190 from A/California/07/2009 (construct VP1-H190-H190). Each protein bears a thrombin cleavage site, and (B) Same as (A), but each protein bears a TEVp cleavage site. Lanes: (L) Pre-stained protein marker, (T) Total protein, (S) Soluble protein, and (F) Flow-through fraction from the GST column during purification of GST-tagged VP1. Red arrows represent GST-tagged VP1.

Similarly, Figure 5-8B shows that the proportion of GST-tagged TEVP-VP1-H3-H190-H190 and TEVP-VP1-H3-H190-H190-4E found in soluble fractions was smaller than the proportion of TEVP-wt-VP1, confirming that both GST-tagged modular VP1 had a lower solubility than TEVP-wt-VP1. However, this was not the case for TEVP-VP1-H190-H190, which was the modular VP1 bearing H190 from A/California/07/2009. The

proportion of GST-tagged TEVP-wt-VP1 and TEVP-VP1-H190-H190 found in soluble fractions was similar, indicating that the solubility of GST-tagged TEVP-VP1-H190-H190 was higher than that of TEVP-VP1-H3-H190-H190 and TEVP-VP1-H3-H190-H190-4E. The difference between H190 element from A/Victoria/210/2009 and the one from A/California/07/2009 is the presence of a hydrophobic stretch within the element. The H190 element from A/Victoria/210/2009 contains a hydrophobic stretch, while the H190 element from A/California/07/2009 does not. Therefore, the low solubility of GST-tagged modular VP1 bearing H190 element from A/Victoria/210/2009 was confirmed due to the presence of a hydrophobic stretch within the H190.

Figure 5-8 also shows that the ratio of soluble to total fractions for GST-tagged VP1-H3-H190-H190 was higher than the ratio for GST-tagged TEVP-VP1-H3-H190-H190. A similar observation was also obtained for constructs VP1-H3-H190-H190-4E and TEVP-VP1-H3-H190-H190-4E. The finding indicates that mutation from a thrombin to TEVp cleavage site caused the lower solubility of both GST-tagged TEVP-VP1-H3-H190-H190 and TEVP-VP1-H3-H190-H190-4E. However, the same result was not obtained for constructs wt-VP1 and TEVP-wt-VP1, as well as constructs VP1-H190-H190 and TEVP-VP1-H190-H190, suggesting that the decrease in solubility due to mutation from a thrombin to TEVp cleavage site was rather a construct-dependent phenomenon. This suggestion was also supported by the fact that the thrombin cleavage site is more hydrophobic than the TEVp cleavage site. The thrombin cleavage site (LVPRGS) has a GRAVY score of 0.117. In contrast, the GRAVY score of the TEVp cleavage site (ENLYFQGS) is -0.8. Analysis of both GRAVY scores indicated that the thrombin cleavage site is more hydrophobic than the TEVp cleavage site. Thus, it is expected that mutation from a thrombin to TEVp cleavage site can improve the solubility of GST-tagged VP1.

SDS-PAGE analysis of flow-through fractions from the GST column (lane F, Figure 5-8) showed that GST-tagged wt-VP1 and modular VP1 bearing thrombin and TEVp cleavage sites were effectively captured in the affinity column, and the capture efficiency was similar. This expected finding implied that the mutation did not affect the efficiency of the affinity purification.

5.3.3.2 Effect of mutation from a thrombin to TEVp cleavage site on the ratio of VP1* to VP1 after GST removal

Purified GST-tagged modular VP1 were then incubated with thrombin or TEVp accordingly to remove the GST tag from modular VP1. After incubation, digested proteins

were clarified by centrifugation to remove precipitates. Samples were taken before and after centrifugation, and analysed using SDS-PAGE, as shown in Figure 5-9.

Figure 5-9A shows that the GST removal profiles for constructs bearing a thrombin cleavage site were consistent with the previous results in Section 5.3.2.1. After GST removal using thrombin, three products were observed: (i) VP1 protein (blue arrows), (ii) VP1* fragment (green arrows), and (iii) GST tag (orange arrows). The persistent VP1* bands at about 37 kDa were observed for all constructs. As reported in Section 5.3.2.1, the digestion profiles of wt-VP1 (lanes 3 and 4) were similar to the profiles of both VP1-H3-H190-H190 (lanes 5 and 6) and VP1-H3-H190-H190-4E (lanes 8 and 9). However, a considerably higher VP1* to VP1 ratio was observed for both modular constructs. In addition, in this section, the modular construct VP1-H190-H190 was used as a comparison. Figure 5-9A shows that VP1-H190-H190 (lanes 11 and 12) had a considerably lower VP1* to VP1 ratio in comparison to both VP1-H3-H190-H190 (lanes 5 and 6) and VP1-H3-H190-H190-4E (lanes 8 and 9). This finding indicated that specificity of thrombin proteolysis was lower for both modular VP1 bearing H190 from the A/Victoria/210/2009 than for modular VP1 bearing H190 from the A/California/07/2009. The difference between H190 element from A/Victoria/210/2009 and the one from A/California/07/2009 is the presence of a hydrophobic stretch within the element. Hence, this finding further indicated that the higher VP1* to VP1 ratio for VP1-H3-H190-H190 and VP1-H3-H190-H190-4E was due to the hydrophobic stretch. Results in Section 5.3.1 show that the VP1* to VP1 ratio for wt-VP1 was lower than the ratio for VP1-H3-H190-H190. Based on this result, it was suggested that modularisation of two copies of H190 from the A/Victoria/210/2009 caused structural alterations, and thus resulted in an increase in non-specific thrombin proteolysis. In this section, this suggestion was proven again by comparing the ratio of VP1* to VP1 between VP1-H3-H190-H190, VP1-H3-H190-H190-4E, and VP1-H190-H190. The difference of thrombin specificity for the modular VP1 proteins indicates that there was a modification of the structure of modular VP1 due to the presence of hydrophobic stretch. The modification may expose a secondary cleavage site that is otherwise less accessible by thrombin.

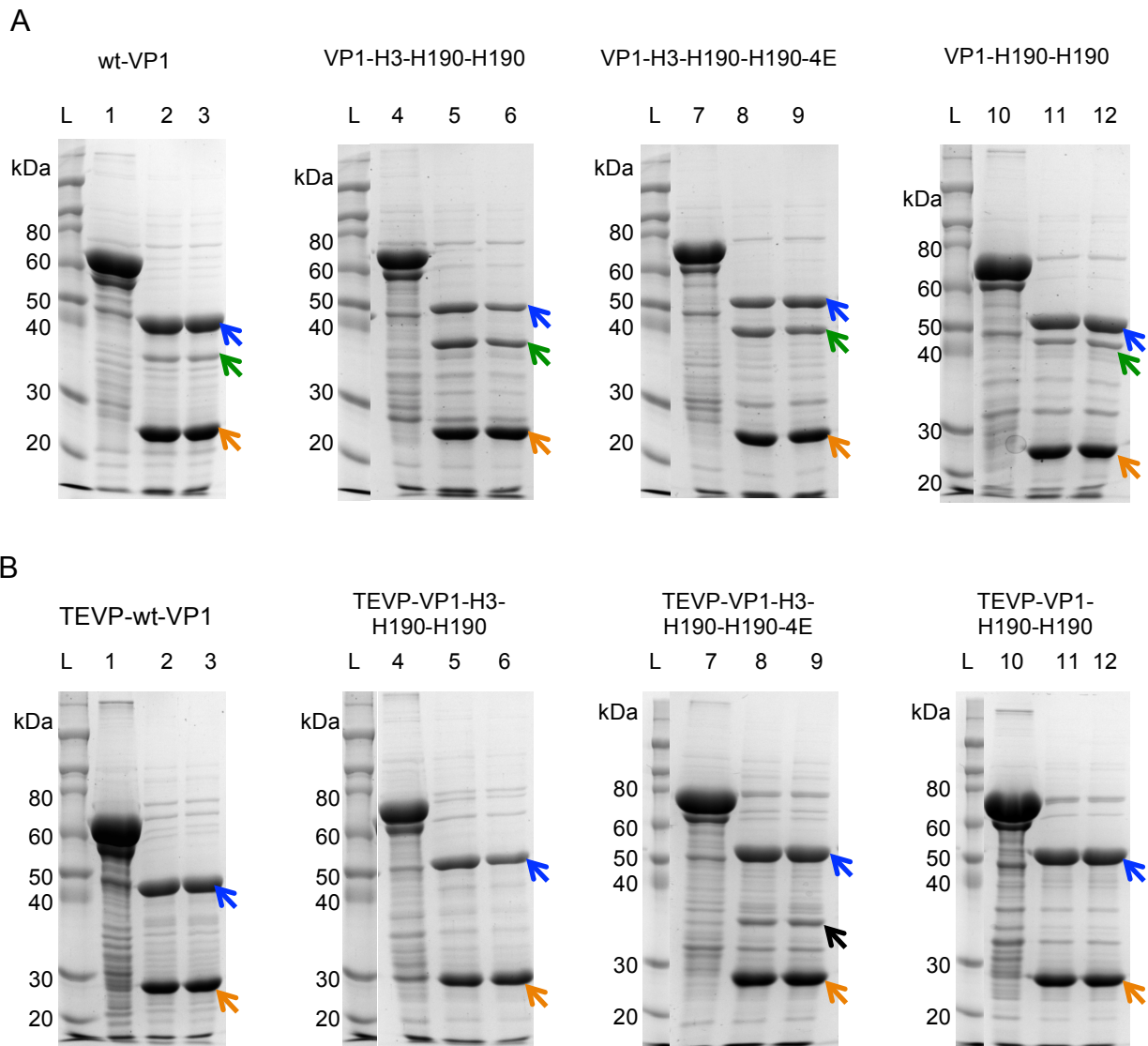


Figure 5-9. SDS-PAGE analysis showing the effect of mutation from a thrombin to TEVp cleavage site on the ratio of VP1* to VP1 following GST removal using thrombin or TEVp. (A) wt-VP1, modular VP1 bearing two copies of H190 from A/Victoria/210/2009 without flanking glutamic acids (construct VP1-H3-H190-H190), modular VP1 bearing two copies of H190 from A/Victoria/210/2009 with flanking glutamic acids (construct VP1-H3-H190-H190-4E), and modular VP1 bearing two copies of H190 from A/Victoria/210/2009 (construct VP1-H190-H190). Each protein bears a thrombin cleavage site, and (B) same as (A), but each protein bears a TEVp cleavage site. Lanes: (L) Pre-stained protein marker, (1) GST-tagged wt-VP1, (2) Total protein of digested GST-tagged wt-VP1, (3) Soluble protein of digested GST-tagged wt-VP1, (4) GST-tagged VP1-H3-H190-H190, (5) Total protein of digested GST-tagged VP1-H3-H190-H190, (6) Soluble protein of digested GST-tagged VP1-H3-H190-H190, (7) GST-tagged VP1-H3-H190-H190-4E, (8) Total protein of digested GST-tagged VP1-H3-H190-H190-4E, (9) Soluble protein of digested GST-tagged VP1-H3-H190-H190-4E, (10) GST-tagged VP1-H190-H190, (11) Total protein of digested GST-tagged VP1-H190-H190, and (12) Soluble protein of digested GST-tagged VP1-H190-H190. Blue arrows refer to the VP1 protein. Green arrows refer to the secondary thrombin cleavage product. Orange arrows refer to GST. Black arrow refers to the secondary TEVp cleavage product.

In contrast to the findings in Figure 5-9A, after GST removal using TEVp, the quality of wt-VP1 and the modular VP1 was considerably improved, as shown in Figure 5-9B. Lanes 2 and 3 in Figure 5-9B show that VP1* was not observed after GST removal from wt-VP1 using TEVp. This finding was in good agreement with the study by Connors et al. (Connors et al. 2013). Similarly, lanes 5 and 6 in Figure 5-9B show that VP1* was not observed for TEVP-VP1-H3-H190-H190 following treatment using TEVp. In addition, the absence of VP1* following TEVp treatment was also observed for TEVP-VP1-H190-H190, as shown in lanes 11 and 12 of the same figure. These results confirm that VP1* observed in Figure 5-6 was a result of non-specific thrombin proteolysis. The use of TEVp, which is more specific than thrombin (Connors et al. 2013), eliminated the generation of VP1*, and therefore increased the quality of modular protein. However, this was not the case for TEVP-VP1-H3-H190-H190-4E. While the VP1* fragment was eliminated following TEVp digestion, a fragment at approximately 35 kDa (black arrow) was observed for TEVP-VP1-H3-H190-H190-4E.

The 35-kDa fragment observed for TEVP-VP1-H3-H190-H190-4E was not observed for the other three constructs, showing that the fragment was not related to *E. coli* host proteins or proteins from the TEVp. This observation suggested that the 35-kDa fragment resulted from non-specific proteolysis of TEVp. This suggestion was also supported by the result shown in Figure 5-10. The figure shows SDS-PAGE analysis of GST-removal using TEVp for the constructs TEVP-VP1-H3-H190-H190, TEVP-VP1-H3-H190-H190-4E, and TEVP-VP1-H190-H190. The digestions were performed at various mass ratios of TEVp to VP1, in Lysis Buffer containing 200 mM NaCl. As described in Section 5.2.3, TEVp digestion was performed by adding TEVp at a certain concentration to protein solutions. At higher ratios of TEVp to VP1, the final volumes of digested proteins were expected to increase, resulting in a dilution factor. Therefore, for the same volume of samples taken from the final solutions, the amount of digested proteins was less due to the dilution. As can be seen in Figure 5-10, for all modular constructs, as the ratio of TEVp to VP1 increased, the intensity of bands corresponding to TEVp (about 27 kDa, purple arrows) increased, and the amount of GST-tagged modular VP1 (red arrows) decreased, indicating that more GST-tagged VP1 was digested. Consequently, as the ratio of TEVp to VP1 increased, the intensity of bands corresponding to VP1 (blue arrows) increased, even though the figure shows the opposite observation due to the aforementioned dilution factor. Based on these observations, it was expected that with increasing TEVp to VP1 ratio, the intensity of bands corresponding to the 35-kDa fragment (black arrow) also subsided. However, the figure shows that the intensity of the 35-kDa fragment

strengthened with increases in the ratio of TEVp to VP1. This result shows that generation of the 35-kDa fragment for the modular construct TEVP-VP1-H3-H190-H190-4E was due to non-specific interactions between TEVp and modular VP1, which was aggravated by the increasing ratio of TEVp to VP1.

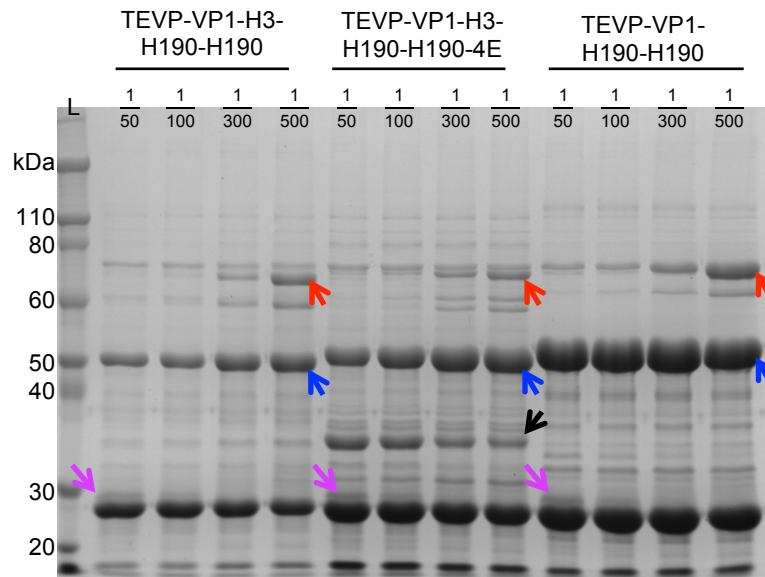


Figure 5-10. SDS-PAGE analysis of GST removal for modular VP1 bearing two copies of H190 from A/Victoria/210/2009 without flanking glutamic acids (construct TEVP-VP1-H3-H190-H190), modular VP1 bearing two copies of H190 from A/Victoria/210/2009 with flanking glutamic acids (construct TEVP-VP1-H3-H190-H190-4E), and modular VP1 bearing two copies of H190 from A/California/07/2009 (construct TEVP-VP1-H190-H190) using TEVp at various ratio of TEVp to VP1. Lanes: (L) Pre-stained protein marker, and (1/50, 1/100, 1/300, and 1/500) Mass ratio of TEVp to VP1. Red arrows refer to GST-tagged VP1 protein. Blue arrows refer to VP1 protein. Black arrow refers to TEVP secondary cleavage product. Purple arrows refer to TEVp.

The molecular size of the fragment suggested that the 35-kDa fragment resulted from non-specific proteolysis of TEVp on the H190 elements flanked by glutamic acid residues. The amino acid sequence of H190 element of A/Victoria/210/2009 (VTDKDQIFLYAQASGR) contains a similar amino acid sequence to the TEVp cleavage site (ENLYFQG). Furthermore, the fact that the fragment was not observed for VP1-H3-H190-H190 indicates that the non-specific interaction was likely attributed to the presence of the flanking ionic elements. TEVp has a neutral net surface charge at pH 8.5 (Kapust and Waugh 1999). During GST removal, which was performed at pH 8, the net charge of TEVp was positive. The charge differences between TEVp and the flanking ionic elements may favour ionic interactions, and subsequently result in non-specific interactions of TEVp on the H190 element.

Interactions between proteins driven by electrostatic interaction have been known to be sensitive to salt concentrations (Golovanov et al. 2004, Zhang et al. 2007). Therefore, to prove the hypothesis that the 35-kDa fragment was the result of ionic interactions between TEVp and glutamic acid, the concentration of NaCl used during GST cleavage was increased. Increasing NaCl concentration can reduce the Debye length and subsequently suppress the electrostatic double-layer interactions between TEVp and the flanking ionic elements. GST removal at various concentrations of NaCl ranging from 200 to 600 mM was tested and analysed using SDS-PAGE as shown in Figure 5-11. The analysis showed that with increasing salt concentration, the amount of the 35-kDa fragment was reduced, although the reduction was only slight. The slight reduction in the 35-kDa fragment when the salt concentration was increased considerably suggested that increasing salt concentration was ineffective at reducing the non-specific ionic interactions. This finding further suggested that the 35-kDa fragment may not have resulted from ionic interactions between TEVp and glutamic acid.

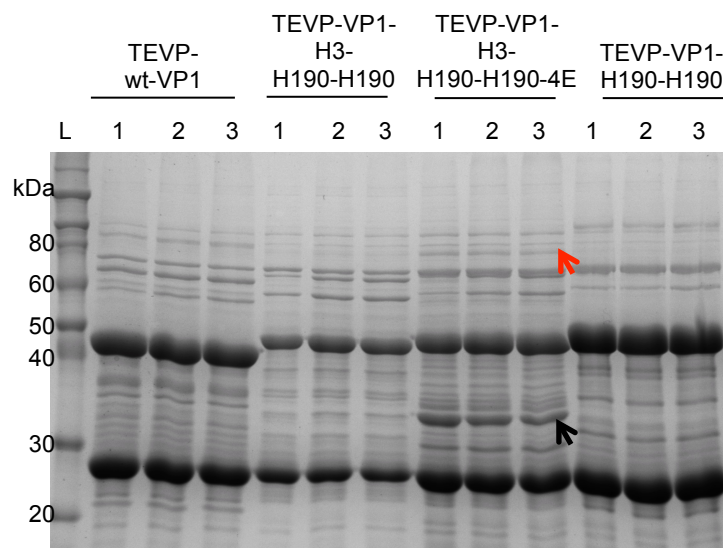


Figure 5-11. SDS-PAGE analysis of GST removal for wt-VP1 (construct TEVP-wt-VP1), modular VP1 bearing two copies of H190 from A/Victoria/210/2009 without flanking glutamic acids (construct TEVP-VP1-H3-H190-H190), modular VP1 bearing two copies of H190 from A/Victoria/210/2009 with flanking glutamic acids (construct TEVP-VP1-H3-H190-H190-4E), and modular VP1 bearing two copies of H190 from A/California/07/2009 (construct TEVP-VP1-H190-H190) using TEVp at various salt concentrations. (1) 200 mM NaCl, (2) 400 mM NaCl, and (3) 600 mM NaCl. Red arrow shows GST-tagged VP1. Black arrow shows TEVp secondary cleavage product.

5.3.3.3 Effect of mutation from a thrombin to TEVp cleavage site on the aggregation of modular VP1 after GST removal

Following GST removal, digested proteins were clarified using centrifugation to remove precipitates. The supernatants were then loaded into SEC. Figure 5-12 shows chromatograms for modular constructs bearing thrombin and TEVp cleavage sites.

In general, the chromatograms for wt-VP1, VP1-H3-H190-H190, and VP1-H3-H190-H190-4E, carrying thrombin cleavage sites (Figure 5-12A-C), were consistent with the results shown in Figure 5-7A-C. Furthermore, these chromatograms were similar to those for the same constructs but bearing the TEVp cleavage site.

Figure 5-12A shows the chromatograms for TEVP-wt-VP1 and wt-VP1. Both chromatograms showed three peaks: (i) peak A1, which corresponded to soluble aggregates, (ii) peak A11, corresponding to capsomeres, and (iii) the GST tag. Peak A1 for TEVP-wt-VP1 appeared to have a similar level of absorbance 280 nm to the peak for wt-VP1. However, the capsomere and GST peaks of TEVP-wt-VP1 had a smaller level of absorbance 280 nm than those of wt-VP1, indicating that TEVP-wt-VP1 had more aggregates. Fractions from these peaks were then analysed using SDS-PAGE. The analysis of peak A1 for wt-VP1 showed that the peak contained VP1 and VP1*. In contrast, peak A1 for TEVP-wt-VP1 contained VP1 (blue arrow) without VP1* (green arrow). This was probably because the VP1* was eliminated following GST removal using TEVp. Furthermore, the SDS-PAGE analysis showed that peak A11 of both wt-VP1 and TEVP-wt-VP1 majorly contained VP1.

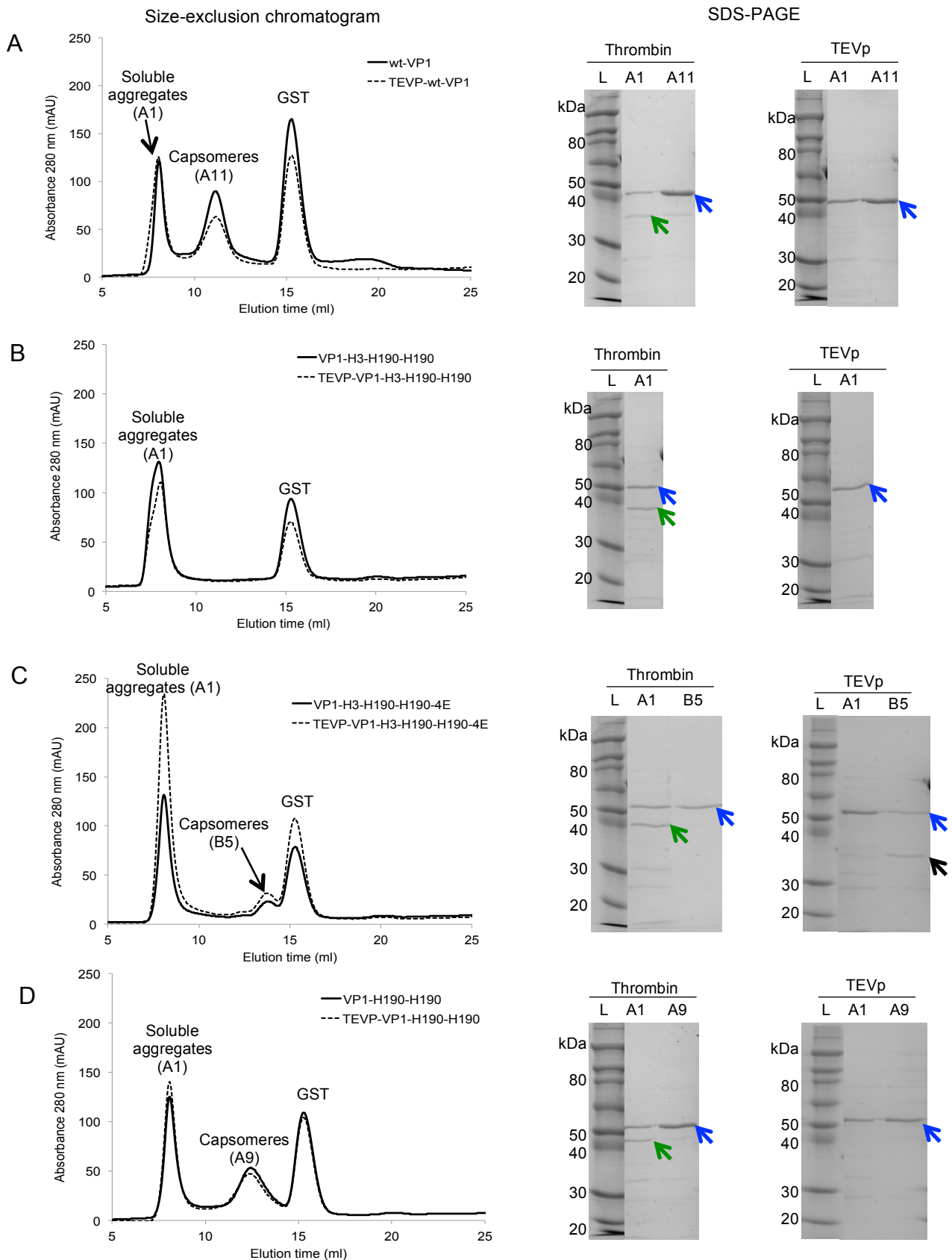


Figure 5-12. Analysis using size-exclusion chromatography and SDS-PAGE showing the effects of mutation from a thrombin to TEVp cleavage site on the aggregation of modular VP1 bearing H190 from A/Victoria/2010/2009 and feasibility to isolate the modular VP1 as capsomeres. (A) wt-VP1 and TEVP-wt-VP1, (B) Modular VP1 bearing two copies of H190 from A/Victoria/210/2009 without flanking ionic elements, (C) Modular VP1 bearing two copies of H190 from A/Victoria/210/2009 with flanking ionic elements, and (D) Modular VP1 bearing two copies of H190 from A/California/07/2009. Blue arrows refer to VP1*. Green arrows refer to VP1*. Black arrow refers to TEVp secondary cleavage product.

Likewise, the chromatograms for VP1-H3-H190-H190 and TEVP-H3-H190-H190 were similar, although a slight difference in absorbance 280 nm readings was detected (Figure 5-12B). Two peaks were observed in both chromatograms, which corresponded to soluble aggregates (peak A1) and the GST tag. Despite the elimination of VP1* by removing GST using TEVp in place of thrombin, peaks corresponding to capsomeres were not observed. SDS-PAGE analysis of a fragment from peak A1 of VP1-H3-H190-H190 showed that the peak contained both VP1 (blue arrow) and VP1* (green arrow). In comparison, the analysis showed that peak A1 from TEVP-H3-H190-H190 only contained VP1. The lack of a peak corresponding to capsomeres following GST removal from TEVP-VP1-H3-H190-H190, supported by the SDS-PAGE analysis, showed that the presence of VP1* was not correlated with aggregation. These findings showed that mutation from a thrombin to TEVp cleavage site did not reduce the tendency of modular VP1 to form soluble aggregates and did not increase the feasibility of isolating modular VP1 capsomeres bearing hydrophobic H190.

Comparable to the observations in Figure 5-12A and Figure 5-12B, the chromatograms for VP1-H3-H190-H190-4E and TEVP-VP1-H3-H190-H190-4E in Figure 5-12C were similar. Three peaks were observed in both chromatograms, i.e. (i) soluble aggregates (peak A1), (ii) capsomeres (peak B5), and (iii) GST. However, Figure 5-12C shows that the absorbance 280 nm readings of each peak for TEVP-VP1-H3-H190-H190-4E was about twice those for VP1-H3-H190-H190-4E, although the SEC analysis was performed for the same amount of proteins, as explained in Section 5.2.3. Such an observation was also reported in Section 5.3.2.3, where it was suggested that the higher absorbance readings was possibly due to the use of glutamic acids. SDS-PAGE analysis showed that peak A1 from VP1-H3-H190-H190-4E contained VP1 (blue arrow) and VP1* (green arrow) in an equal proportion. Meanwhile, A1 peak from TEVP-VP1-H3-H190-H190-4E contained predominantly VP1. Moreover, the chromatogram of TEVP-VP1-H3-H190-H190-4E shows that, peaks corresponding to capsomeres (B5) were observed for both VP1-H3-H190-H190-4E and TEVP-VP1-H3-H190-H190-4E. SDS-PAGE analysis of these peaks showed that peak B5 from VP1-H3-H190-H190-4E contained majorly VP1 (blue arrow), while peak B5 from TEVP-VP1-H3-H190-H190-4E contained VP1 and the 35-kDa fragment (black arrow). These findings confirmed that the heterogeneity of modular VP1 was not correlated to the aggregation of modular VP1 after GST removal.

Figure 5-12D compares the chromatogram for VP1-H190-H190 and TEVP-VP1-H190-H190. The chromatograms showed that three peaks were observed, similar to the chromatogram for wt-VP1. These peaks were (i) peak A1, corresponding to soluble

aggregates, (ii) peak A9, which corresponded to capsomeres, and (iii) the GST tag. Unlike VP1-H3-H190-H190 and VP1-H3-H190-H190-4E, both chromatograms showed that peaks containing capsomeres were eluted at about 12.5 ml (peak A9). The capsomere elution time was slightly later than the elution time for wt-VP1 capsomeres, which was about 11.5 ml (peak A11). The changed elution time may represent capsomeres VP1-H190-H190 and wt-VP1 having different hydrodynamic radii. SDS-PAGE analysis for peaks from VP1-H190-H190 indicated that peak A1 contained VP1 and VP1* at an equal ratio, while peak A9 contained VP1. In comparison, SDS-PAGE analysis for peaks from VP1-H190-H190 indicated that peak A1 and A9 contained VP1.

In summary, the results in this section show that mutation from a thrombin to TEVp cleavage site eliminated the generation of VP1*, and consequently improved the homogeneity of modular VP1 after GST removal. This finding confirmed that the increased heterogeneity of modular protein after GST removal was due to the non-specific proteolysis of thrombin. Furthermore, results in this section have shown that improved homogeneity of the modular protein was not correlated to the aggregation of modular VP1 and the feasibility of obtaining modular VP1 capsomeres.

5.4 Conclusion

This chapter accomplished Objective 3 stated in Section 1.3, by demonstrating the applicability of tandem repeat display for modularisation of H190 from influenza virus A/Victoria/210/2009 (H3N2/X-187) into murine polyomavirus VP1. The H190 variant is hydrophobic and contains a hydrophobic stretch in the middle of its amino acid sequence.

This chapter shows that modularisation of two copies of the hydrophobic H190 affected the biophysical properties of modular proteins. Firstly, modularisation of the hydrophobic H190 increased the tendency of the modular VP1 to aggregate after GST removal, preventing isolation of the modular VP1 capsomeres. Secondly, modularisation of the hydrophobic H190 enhanced the thrombin non-specific proteolysis, resulting in an increase in the ratio of secondary cleavage product to VP1.

Two approaches were explored simultaneously to minimise these two effects of modularisation of the hydrophobic H190 on the properties of modular VP1. The first approach was the addition of charged residues in the antigen module. This chapter showed that glutamic acids could: (i) improve the solubility of the GST-tagged modular VP1, (ii) reduce the tendency of the modular proteins to aggregate, and increase the feasibility of obtaining modular VP1 capsomeres, and (iii) slightly reduce the ratio of VP1* to VP1. By using the flanking ionic elements, it was possible to obtain modular VP1

capsomeres, which may otherwise not have been achievable by changing the salt concentration of Lysis Buffers or reducing the expression temperature alone.

The second approach was the use of a protease with a higher specificity in order to reduce the heterogeneity of modular VP1 following GST removal. This chapter showed that by replacing thrombin with a more specific protease, in this case TEVp, the heterogeneity of modular VP1 after GST removal was successfully eliminated. However, this was not the case for all constructs.

Furthermore, this chapter showed that the heterogeneity of modular proteins was not correlated with the aggregation of the modular VP1 and the feasibility of obtaining modular VP1 capsomeres. This finding confirmed that the aggregation of modular VP1 capsomeres after GST removal and consequent recalcitrance to isolation resulted from the presence of a hydrophobic stretch within the H190 element rather than being due to the heterogeneity of modular VP1.

Chapter 6. Production of modular VP1 capsomeres with a high purity to screen the optimum condition for crystallisation

6.1 Introduction

Chapters 3 and 4 of this thesis have established that various display strategies used for the modularisation of helix 190 (H190) on murine polyomavirus (MuPyV VP1) have resulted in antibody induction of diverse qualities in mice. In Chapter 3, two display strategies were compared. The first display strategy exploited the use of helix promoter elements, which were derived from the GCN4 protein, to flank one H190 element. Meanwhile, in the second strategy, two copies of H190 were arrayed as a tandem repeat. The results showed that H190 modularised using both strategies were equally immunogenic; however, they induced antibodies of different quality. The two copies of H190 tandem repeats induced a high titre of IgGs recognising recombinant HA1 protein, up to a dilution of 10^4 . In contrast, H190 flanked with helix promoter elements induced a HA1-specific IgG titre less than 10^3 . The induction of different antibody quality by modularised H190 as a result of different display strategies was further verified in Chapter 4. The chapter reports that modular VLPs bearing one to four copies of H190 tandem repeats induced different quality of antibodies, although these modularised H190 were equally immunogenic. A modular VLP bearing one copy of H190 was shown to be unable to induce antibodies that specifically recognised H190 element in recombinant HA1 protein. In contrast, two to four copies of H190 induced a high titre of IgGs recognising the HA1 protein.

These findings in Chapters 3 and 4 may reflect an affinity maturation process during antibody production against H190 displayed on a modular VLP. Antibodies that bind to an antigen with high affinity are developed during an affinity maturation process, which is the key characteristic of antibody production. The affinity maturation is achieved via two mechanisms, both of which are guided by antigens: (i) somatic hypermutation, and (ii) clonal selection. Initially, B cells that encounter an antigen for the first time will bind to the antigen with a low affinity. The antigen-recruited B cells then proliferate in germinal centres and undergo a period of somatic hypermutation. The mutation process results in the production of a small proportion of B cells that bind to the antigen with a higher affinity. Subsequently, both B cells with lower and higher affinity are exposed to the clonal selection process. B cells with a lower affinity are subjected to apoptosis, whereas those with a higher affinity are selectively expanded in an antigen-mediated process, repeating

the two mechanisms until B cells that bind to the antigen with a high affinity are obtained (Batista and Neuberger 1998).

Because the affinity maturation is guided by antigens, the structure of antigens determines the affinity of the induced antibodies. If the affinity maturation is developed against flexible antigens, antibodies with a wide range of affinity may be obtained. This is because antibodies may trap flexible antigens in a conformation that fits with the antigen-binding site. Consequently, the flexible peptide could guide the affinity maturation process down many alternative paths, not only those that optimise binding to the targeted surface in the pathogens of interest. Only a small proportion of these antibodies will be able to recognise the same antigens in the native proteins. In contrast, if the affinity maturation is developed against an antigen that stably presents a specific surface, antibodies that efficiently recognise the surface will be obtained (Dormitzer et al. 2008). Therefore, if the modularised H190 closely resembles the H190 within the native HA protein, the antibodies induced are likely to bind to the HA protein with high affinity. In contrast, if the modularised H190 is structurally different from the one in the HA protein, the antibodies may bind to the HA protein with low affinity. Based on this rationale, the findings in Chapters 3 and 4, which show that various H190 display strategies can result in induction of antibodies of diverse qualities, may suggest the antibodies are raised against modularised H190 of different structure. The structural difference may be a result of different display strategies.

A possible approach to verify this suggestion is by determining the structure of displayed H190 on modular virus-like particles (VLPs) using, for example, X-ray crystallography. X-ray crystallography has long been the most favoured technique to determine the structure of proteins and biological macromolecules (Smyth and Martin 2000). It is an established (Momany et al. 1996) and leading method to resolve structural information in proteins larger than 35 kDa (Pusey et al. 2005). Furthermore, it contributes to the resolution of most of protein structures in the Protein Data Bank archive (Pusey et al. 2005). In X-ray crystallography, a protein of interest is purified and crystallised at a high concentration. The crystal is then subjected to an X-ray beam at high intensity. It diffracts the beam into unique patterns of spots. Intensities and phase of the spots reflect electron distributions in the protein. The electron density maps are then analysed to determine the location of each atom of the protein, yielding an atomic model of the protein. Subsequently, it may be necessary to perform refinement on the crystallographic atomic model. During refinement, a crystallographic atomic model is restrained or constrained using database information of geometrical and stereochemical features, such as bond lengths and angles, planarity and chirality (Kleywegt and Jones 1998). The accuracy of an

atomic model is measured by: (i) resolution, and (ii) R-value. The resolution reflects how much detail can be obtained from a crystallographic atomic model. Meanwhile, the R-value indicates how well a refined atomic model is supported by experimental data on intensities and phase of a diffracted beam.

In order to obtain an atomic model with high accuracy, a high-quality crystal is required. When a crystal with a high quality is used, the atomic model has a higher confidence that it correctly reflects the actual structure of a protein, and *vice versa*. One important factor that determines crystal quality, as well as X-ray beam diffraction, is the purity of the protein of interest. Impurities can affect protein solubility, nucleation, crystal morphology, and irreproducibility in crystallisation experiments. Studies using optical microscopy and interferometry have shown that impurities reduce the growth rate of crystals. Furthermore, they can cause dislocation and cracks, the formation of soluble aggregates, and the degradation of crystal mosaicism (Caylor et al. 1999). Studies on lysozyme using light scattering have shown that impurities could associate with lysozyme, promoting the formation of ill-shaped and twinned crystals (Lorber et al. 1993, Skouri et al. 1995). Because crystal formation is sensitive to many factors, crystallisation of a protein can be the rate-determining step in protein structure determination using X-ray crystallography. It can be difficult and can limit the number of proteins that can be studied using X-ray crystallography (Dessau and Modis 2011, Smyth and Martin 2000).

Of particular relevance to this study, X-ray crystallography has previously been utilised to resolve the molecular structure of various viral capsid proteins (Chen et al. 2000, Hogle et al. 1986, Lattman 1980, Prasad et al. 1999, Speir et al. 1995) including wild-type MuPyV VP1 (wt-VP1) (Stehle and Harrison 1996, Stehle and Harrison 1997). The molecular structure of wt-VP1 was resolved at a resolution of 3.65 Å from purified virions (wt-VP1 residues 1-384; Figure 6-1A) (Stehle and Harrison 1996). Additionally, the molecular structure of wt-VP1 was resolved from capsomeres (wt-VP1 residues 32-320; Figure 6-1B) (Stehle and Harrison 1997). The capsomeres were produced recombinantly using *E. coli* expression and were found to lack: (i) 31 residues from the N-terminus, and (ii) 63 residues from the C-terminus. The 31 N-terminal residues were removed because the flexibility of the residues was expected to generate structural heterogeneity, thus preventing the formation of crystals (Dale et al. 2003). In contrast, the 63 C-terminal residues were removed to prevent the capsomeres from self-assembling into VLPs. Compared to the aforementioned model of wt-VP1, this capsomere-based model has a higher resolution of 1.9 Å even though both resolved structures of wt-VP1 are highly similar. Additionally, the capsomere-based molecular structure provides more structural

details. For example, the capsomere-based model could show the conformation of residues 32-45 in the free capsomeres, which could not be obtained from the virion-based model.

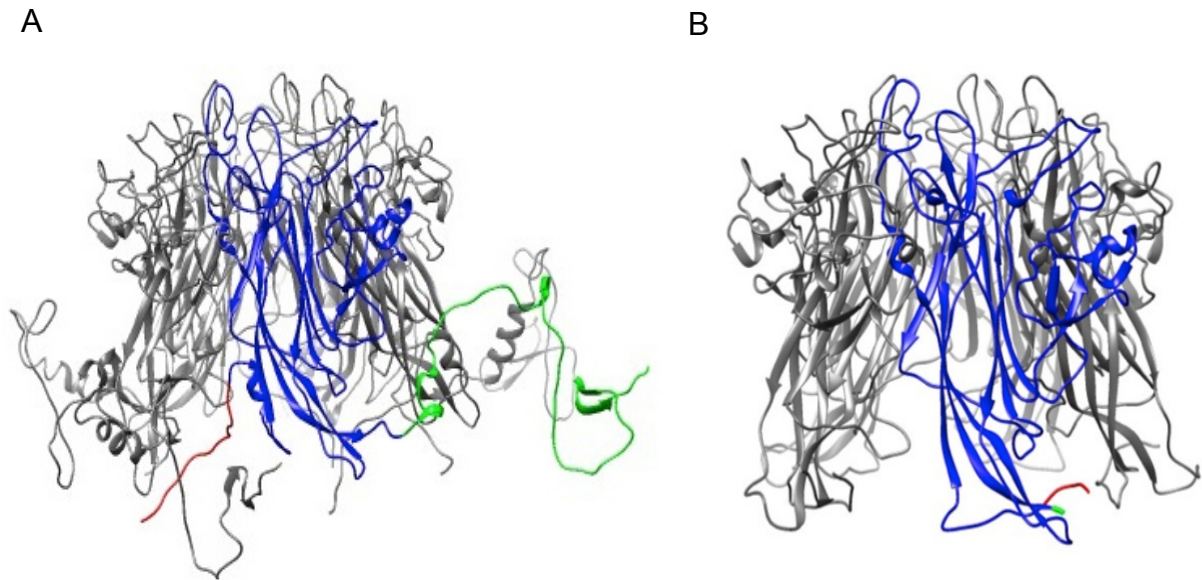


Figure 6-1. Crystal structures of wild-type murine polyomavirus VP1 capsomeres. (A) Crystal structure from purified virions (VP1 residues 1-384; 1SID.pdb). Red represents residues 1-31. Green represents residues 321-384; and (B) Crystal structure from recombinant VP1 capsomeres (VP1 residues 32-320; 1VPN.pdb). Red represents residues 32-36. Green represents residues 316-320. Figures are generated using UCSF Chimera (Pettersen et al. 2004).

This chapter reports the first steps towards obtaining H190 structure displayed on modular VLPs using X-ray crystallography. These first steps are: the production of homogenous and soluble modular VP1, and screening of conditions to obtain optimum crystals. The production of homogeneous modular VP1 capsomeres comprises five main steps: (i) the selection of modular constructs, (ii) design of the selected modular constructs for purification purposes, (iii) expression and purification of the selected modular constructs as Glutathione-S-transferase (GST)-tagged proteins, (iv) enzymatic removal of the GST tag from modular VP1, and (v) the removal of contaminants.

This chapter selects three modular constructs, of which the molecular structures of the modularised H190 in these three constructs are yet to be resolved. These constructs are: (i) VP1-H190-H190 (reported in Chapters 3 and 4), (ii) VP1-GCN4-H190-GCN4 (reported in Chapter 3), and (iii) VP1-H190 (reported in Chapter 4). The antigen module in construct VP1-H190-H190 comprises two copies of H190 arrayed as a tandem repeat (Figure 6-2A). In contrast, in the antigen module of construct VP1-GCN4-H190-GCN4,

H190 is modularised by exploiting the use of GCN4 helix promoter elements and GSGS spacer elements (Figure 6-2B). In the modular construct VP1-H190, one copy of H190 is modularised without incorporating any structural hypothesis (Figure 6-2C). Construct VP1-H190 represent the simplest display strategy to modularise H190 element. The H190 elements in these three selected constructs were from the influenza virus strain A/California/07/2009 (H1N1).

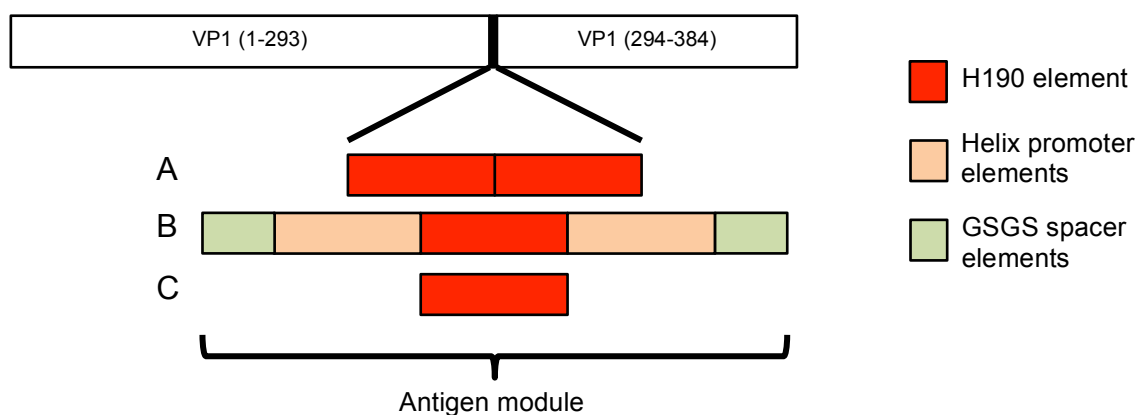


Figure 6-2. A schematic diagram of antigen modules in the three modular constructs selected for crystallisation. (A) VP1-H190-H190, two copies of H190 were arrayed as a tandem repeat, (B) VP1-GCN4-H190-GCN4, one copy of H190 was flanked with GCN4 helix promoter elements and GSGS spacer elements, (C) VP1-H190, the simplest display strategy to modularise H190. Numbers in white boxes refer to amino acid residue of the wt-VP1 protein. The diagram is not to scale.

These three modular constructs were selected for the following reasons. Chapters 3 and 4 of this thesis showed that the H190 element in the modular construct VP1-H190-H190 was displayed in its immunogenic form, and induced a high titre of HA1-specific IgGs. VP1-H190-H190 VLPs were shown to be self-adjuvanting, and thus approaches to improve the quality of antibodies induced by the modular VLPs were shown to be unnecessary. In contrast, the modularised H190 in VP1-GCN4-H190-GCN4 induced a lower titre of IgGs recognising recombinant HA1 protein. Similarly, the H190 in modular construct VP1-H190 raised IgGs that bound to HA1 non-specifically. Comparison between resolved structures of H190 in construct VP1-H190-H190 against VP1-GCN4-H190-GCN4 could provide indications on the effect of display strategy on the structure of modularised H190. Moreover, resolving structures of H190 within VP1-H190 could provide evidence to prove that when one copy of H190 is modularised without any structural hypothesis, H190 could not assume its native conformation.

The three selected modular constructs are then designed in order to increase the possibility of obtaining a high-resolution X-ray structure. The longer-term objective (see

Chapter 7) is an X-ray crystallisation study aiming to resolve the structures of H190 on modular VP1 proteins that are modularised using various display strategies. The resolved molecular structures will be compared to analyse whether the structure of H190 is determined by the display strategy. Finding structural differences of H190 in the three modular VP1 proteins may require atomic models with high resolution. The study by Stehle and Harrison (Stehle and Harrison 1997) showed that a higher resolution of VP1 molecular structure could be obtained from capsomeres (wt-VP1 residues 32-320) instead of purified virions. Therefore, the three selected modular constructs were produced as modular VP1 capsomeres lacking 31 residues from N-terminus and 63 residues from the C-terminus.

Modular VP1 proteins used for crystallisation are expressed as GST-tagged proteins and the GST-tagged modular VP1 are purified using affinity chromatography. These two processes are performed using the previously established methods of expression and purification of GST-tagged wt-VP1 (Chuan et al. 2008, Lipin et al. 2008, Middelberg et al. 2011). Following purification of GST-tagged modular VP1, the GST-tag is enzymatically removed from the modular VP1. The removal of 63 residues from the C-terminal of modular VP1 is not anticipated to have a significant impact during purification. In contrast, the efficiency and rate of enzymatic removal of fusion tag has been shown to be dependent on the length of space between the fusion tag and target protein (Lee et al. 2009). Therefore, the removal of 31 residues from the N-terminus was anticipated to remove a substantial space between GST and modular VP1 (refer to Figure 6-1). Removing 31 residues from the N-terminus may cause a protease to have difficulty accessing its cleavage site, which is located between GST and modular VP1. Consequently, GST removal might be inefficient and undigested GST-tagged modular VP1 might be found in the solution. If the amount of undigested GST-tagged modular VP1 is substantial, the crystal quality and subsequently the resolution of the molecular structure may be decreased.

In order to achieve efficient GST removal for modular constructs in this chapter, two proteases were compared: (i) thrombin and (ii) tobacco etch virus protease (TEVp). Unlike thrombin, TEVp is cost-effective and not involved in human physiological and biochemical networks. It has also been shown to have higher protease cleavage site specificity than thrombin (Connors et al. 2013). The high specificity of TEVp is also reported in Chapter 5 of this thesis, which showed that the use of TEVp eliminated the generation of secondary cleavage product following GST removal. Thus, TEVp is a reasonable comparison for thrombin. Furthermore, the molecular weight of thrombin is rather big, about 37 kDa, compared to the molecular weight of TEVp, which is about 26 kDa. Due to its larger size,

thrombin is anticipated to have difficulty accessing its cleavage site following the removal of 31 residues from the N-terminus. To compare the efficiency of thrombin and TEVp, the selected modular VP1 was designed to carry thrombin or TEVp cleavage sites, in addition to the removal of residues from the N- and C-termini. For modular VP1 bearing a thrombin cleavage site, two glycine residues were added in between GST and VP1 (Figure 6-3A) to provide thrombin with access to its cleavage site. In contrast, for modular VP1 bearing a TEVp cleavage site, no additional residues were added between GST and VP1 (Figure 6-3B).

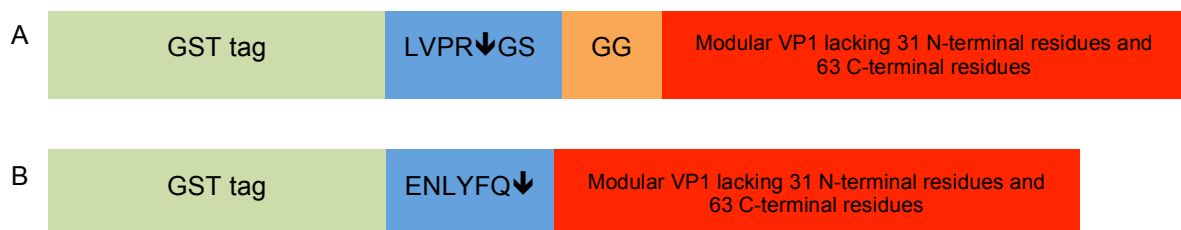


Figure 6-3. A schematic diagram of the engineered cleavage sites in modular VP1 lacking 31 N-terminal residues and 63 C-terminal residues. (A) GST-tagged modular VP1 bearing a thrombin cleavage site, and (B) GST-tagged modular VP1 bearing a TEVp cleavage site. Green represents GST-tag. Blue represents thrombin or TEVp cleavage sites. Red represents modular VP1 protein lacking 31 N-terminal residues and 63 C-terminal residues. The figure is not to scale.

After the GST tag is enzymatically removed, modular VP1 capsomeres are separated from soluble aggregates and GST tag using size-exclusion chromatography (SEC) (Middelberg et al. 2011). A study on wt-VP1 (wt-VP1 residues 1-384) by Fan and Middelberg (Fan and Middelberg 2010) showed that SEC-purified wt-VP1 contained two contaminants, which were detected as bands at about 60 kDa and 70 kDa on SDS-PAGE analysis. Fan and Middelberg identified these bands as GroEL, and DnaK, respectively. Since the GST-tagged modular VP1 are expressed and purified using the same expression system and methods as GST-tagged wt-VP1, it is expected that the SEC-purified modular VP1 also contain these two contaminants.

GroEL and DnaK (reviewed in Chapter 2) are highly conserved *E. coli* proteins, which belong to the family of heat shock protein 60 (Hsp60) and 70 (Hsp70), respectively (Craig et al. 1993, Lindquist and Craig 1988). They act as chaperones, assisting protein folding (Boorstein et al. 1994, Checa and Viale 1997) by shielding hydrophobic surfaces exposed to the surface due to the non-native conformation of proteins (Calloni et al. 2012). They have also been previously shown to promote aggregation and size heterogeneity of

formed VLPs from wt-VP1 capsomeres (wt-VP1 residues 1-384) produced in *E. coli* (Fan and Middelberg 2010).

The study by Fan and Middelberg (Fan and Middelberg 2010) showed two purification methods that could be employed to effectively remove GroEL and DnaK from wt-VP1 capsomeres (wt-VP1 residues 1-384). The first method was ion-exchange chromatography (IEX) using QFF and SP chromatography columns. Purification of the thrombin-treated GST-tagged wt-VP1 using QFF has been shown to reduce the proportion of GroEL in the protein mixture. Similarly, DnaK was decreased after purification of QFF-purified wt-VP1 capsomeres using SP resin. The second method was precipitation using ammonium sulphate, $(\text{NH}_4)_2\text{SO}_4$. The study showed that precipitation using 25% (v/v) $(\text{NH}_4)_2\text{SO}_4$ could reduce the proportion of DnaK in SEC-purified wt-VP1 capsomeres. The precipitation step also reduced the proportion of DnaK following purification using QFF. Furthermore, precipitation using 15% (v/v) $(\text{NH}_4)_2\text{SO}_4$ was able to reduce the proportion of GroEL and DnaK in a mixture of SEC-purified wt-VP1 capsomeres and thrombin-treated GST-tagged wt-VP1.

Although these two methods worked effectively to remove GroEL and DnaK, they were developed for wt-VP1 capsomeres (wt-VP1 residues 1-384). This chapter attempts to remove GroEL and DnaK from modular VP1 proteins lacking 31 residues from the N-terminus and 63 residues from the C-terminus. The modular VP1 proteins bear H190, which is modularised using a different display strategy. In addition, in this chapter, the modular VP1 proteins lack 31 residues from the N-terminus and 63 residues from the C-terminus. These differences may diminish the efficiency of those methods to remove GroEL and DnaK from the modular VP1 proteins. Therefore, this chapter exploits another purification method that is hydrophobic interaction chromatography (HIC) to remove GroEL and DnaK. Unlike SP and Q columns, which separate proteins based on the net surface charge of proteins, HIC separates proteins based on their hydrophobicity.

The quaternary structure of the high purity modular VP1 capsomeres are analysed using SEC coupled with multi-angle light scattering (MALS). In SEC, high purity modular VP1 capsomeres were separated from any remaining impurities based on size. In addition, the use of MALS allows the detection of soluble aggregates. High purity modular VP1 capsomeres were then concentrated and sent to the Australian Synchrotron (Melbourne, Australia) for screening of the optimum conditions for crystallisation. Unfortunately, at the time at which this thesis was written, the conditions required to obtain optimum crystals had not been achieved.

6.2 Materials and methods

6.2.1 Generation of modular constructs

Six modular constructs were generated in this chapter: H190-H190, GCN4-H190-GCN4, H190, TEVP-H190-H190, TEVP-GCN4-H190-GCN4, and TEVP-H190. The first three modular constructs, H190-H190, GCN4-H190-GCN4, and H190, were generated by the Protein Expression Facility (the University of Queensland, Australia).

Plasmid pGEX-VP1 (reported in Section 4.2.1) and plasmid pGEX-TEVP-VP1 (reported in Section 5.2.1) were digested using *Bam*HI and *Xho*I according to the manufacturer's recommendations. Digestion of plasmid pGEX-VP1 yielded linearised DNA fragments pGEX and VP1, while digestion of plasmid pGEX-TEVP-VP1 yielded linearised DNA fragments pGEX-TEVP and VP1. Fragments pGEX and pGEX-TEVP were separated from fragment VP1 using agarose gel electrophoresis, and purified using the PureLink™ Quick Gel Extraction Kit (Life Technologies, Carlsbad, CA, USA).

DNA fragments encoding modular VP1 lacking 31 N-terminal residues and 63 C-terminal residues in plasmids pGEX-VP1-H190-H190, pGEX-GCN4-H190-GCN4, and pGEX-VP1-H190 were amplified using two sets of primers. The first set of primers was: (i) 5' ctggttccgcgctggatccgggggtgggggtatggaggtgctg 3' and (ii) 5' cggaattccgggatcctcagggatagggattttgacctt 3'. Amplification using the first set of primers resulted in DNA fragments carrying additional glycine residues at the 5' end and stop codon at the 3' end. Additionally, the amplified DNA fragments carried 18 bp of DNA sequences homologous to fragment pGEX at both the 5' and 3' ends. The amplifications were performed in a polymerase chain reaction (PCR) mixture (1 U KOD DNA polymerase[‡], 1x KOD Buffer, 25 mM MgSO₄, 1 µl template, 12.5 µM primers, and 2 mM dNTPs). PCR was performed at annealing temperature of 58°C.

The second set of primers was: (i) 5' gatgaaaactgtacttccaagggggtatggaggtgctggac 3' and (ii) 5' cagtcagtcacgatgcggccgctcgagtcagggatagggattttgaccttct 3'. Amplification using the second set of primers resulted in DNA fragments carrying a stop codon at the 5' end and 21 bp of DNA sequences homologous to fragment pGEX-TEVP at both 5' and 3' ends. The amplifications were performed in a PCR mixture (1 U Phusion Hot Start II DNA polymerase[§], 1x Phusion HF Buffer, 1 µl template, 25 µM primers, and 200 µM dNTPs). PCR was performed at annealing temperature of 62°C.

[‡] EMD Millipore, Merck KGaA, Darmstadt, Germany.

[§] Thermo Scientific, Pittsburgh, PA, USA.

The amplified DNA fragments were then purified using PureLink™ Quick Gel Extraction Kit. Purified DNA fragments, which were amplified using the first set of primers, were cloned into purified linearised pGEX, yielding modular constructs H190-H190, GCN4-H190-GCN4, and H190. These constructs carried a thrombin cleavage site. Similarly, purified DNA fragments, which were amplified using the second set of primers, were cloned into purified linearised pGEX-TEVP, yielding modular constructs TEVP-H190-H190, TEVP-GCN4-H190-GCN4, and TEVP-H190. These constructs carried a TEVp cleavage site. Cloning was performed using the *in vivo* homologous recombination as described in Chapter 4 (Section 4.2.1). DNA sequences of the designated constructs were confirmed by sequencing DNA inserts, which was conducted by Australian Genome Research Facility (AGRF, Brisbane, Australia). Modular constructs generated in this chapter were summarised in Table 6-1.

Table 6-1. List of modular constructs bearing thrombin or TEVp cleavage sites generated in Chapter 6.

Construct names	Protease cleavage site	Templates to amplify DNA fragments encoding modular VP1 lacking 31 N-terminal residues and 63 C-terminal residues
H190-H190	Thrombin	Plasmid pGEX-VP1-H190-H190 (reported in Chapter 3)
TEVP-H190-H190	TEVp	
GCN4-H190-GCN4	Thrombin	Plasmid pGEX-VP1-GCN4-H190-GCN4 (reported in Chapter 3)
TEVP-GCN4-H190-GCN4	TEVp	
H190	Thrombin	Plasmid pGEX-VP1-H190 (reported in Chapter 4)
TEVP-H190	TEVp	

6.2.2 Expression and purification of GST-tagged modular VP1

Expression and purification of GST-tagged modular VP1 were conducted as previously described in Section 4.2.1 (Chuan et al. 2008, Lipin et al. 2008, Middelberg et al. 2011), unless stated otherwise. Briefly, plasmids were transformed into *E. coli* Rosetta (DE3) pLysS cells (EMD Millipore, Merck KGaA, Darmstadt, Germany). Modular VP1 protein was expressed as a GST-tagged protein at 26°C after induction at OD_{600 nm} of 0.5 with 0.2 mM isopropyl β -D-1-thiogalactopyranoside (IPTG). Harvested cells were resuspended in 40 ml of Lysis Buffer (40 mM Tris-base, 200 mM NaCl, 1 mM EDTA, 5% (v/v) glycerol, 5 mM DTT, pH 8.0), and lysed by sonication (Branson Ultrasonics Corporation, Connecticut, USA) at output 30 for 4 cycles of 40 s. The supernatant was separated from cell debris by centrifugation, and filtered through 0.45 μ m filters (Pall, New York, USA). GST-tagged modular VP1 was isolated from the supernatant using GST-affinity chromatography. The flow rates of sample loading and elution were reduced to 0.2

ml min⁻¹ to facilitate maximum capture of GST-tagged modular VP1. Bound GST-tagged modular VP1 was eluted using Lysis Buffer containing 10 mM GSH (40 mM Tris-base, 200 mM NaCl, 1 mM EDTA, 5% (v/v) glycerol, 5 mM DTT, 10 mM GSH, pH 8.0).

6.2.3 Enzymatic GST tag removal from modular VP1

GST tag removal using thrombin was conducted as previously described (Chuan et al. 2008). GST-tagged modular VP1 proteins were incubated with thrombin (Catalogue# 27-0846-02, GE Healthcare, UK) at room temperature for 2 h. The ratio of GST-tagged modular VP1 to thrombin was 40:1 (thrombin unit/ml protein).

GST tag removal using TEVp was conducted as previously described (Connors et al. 2013), unless stated otherwise. GST-tagged modular VP1 proteins were incubated with TEVp at a ratio of 1:5 (TEVp mass/protein mass) for 4 h at room temperature or in a ratio of 1:10 for 2 h at 30°C.

6.2.4 Modular VP1 capsomere purification by size-exclusion chromatography (SEC)

Following enzymatic GST tag removal, the protein solutions were centrifuged and the supernatants were subsequently loaded into SEC column Superdex 200 10/300 GL (GE Healthcare, UK). SEC was conducted at a flow rate of 0.5 ml min⁻¹, as previously described in Section 4.2.1 (Chuan et al. 2008, Lipin et al. 2008, Middelberg et al. 2011). Here, X-ray Buffer (50 mM Tris-base, 100 mM NaCl, pH 8.0) was used to perform SEC.

6.2.5 Purification using GST-affinity chromatography for SEC-purified modular VP1 capsomeres

SEC-purified modular VP1 capsomeres in X-ray Buffer were loaded into serially-connected GSTrap HP 1 ml and GSTrap FF 1 ml (GE Healthcare, UK) at a flow rate of 0.2 ml min⁻¹. Modular VP1 capsomeres were collected in the flow-through fractions at 1 ml volume continuously.

6.2.6 Purification using hydrophobic interaction chromatography (HIC) for GST-purified modular VP1 capsomeres

6.2.6.1 HIC in batch mode

Saturated (NH₄)₂SO₄ solution was added into X-ray Buffer containing SEC-purified modular VP1 capsomeres (Section 6.2.4) to final concentrations (v/v) of 0%, 2.5%, 5%, 7.5%, 10%, 12.5%, and 15%. Similarly, X-ray Buffer containing (NH₄)₂SO₄ at these concentrations were also prepared for resin equilibration.

A 500 μl of Phenyl Sepharose HP resin (GE Healthcare, UK) suspension in 20% ethanol was rinsed with 1 ml of filtered and deionised water. The resin was then cleaned with 1 ml of 0.5 M NaOH, rinsed three times with 1 ml of filtered and deionised water, and subsequently equilibrated three times with 1 ml of X-ray Buffers containing $(\text{NH}_4)_2\text{SO}_4$ at various concentrations (v/v) (0%, 2.5%, 5%, 7.5%, 10%, 12.5%, and 15%).

A 50 μl of SEC-purified modular VP1 capsomeres containing $(\text{NH}_4)_2\text{SO}_4$ prepared as described in the beginning of this session was incubated with 50 μl of resin, which had been equilibrated with X-ray Buffer containing $(\text{NH}_4)_2\text{SO}_4$ at the corresponding concentration. The incubation was performed for 20 min at room temperature with gentle agitation. After incubation, the resin and proteins mixture was centrifuged. The supernatant was decanted.

6.2.6.2 HIC in bed mode

A HiTrap Phenyl HP 1 ml column (GE Healthcare, UK) stored in 20% ethanol was cleaned with 5-10 CV of filtered and deionised water, 0.5 M NaOH, and filtered and deionised water, consecutively. The column was pre-equilibrated with 5-10 CV X-ray Buffer containing 10% (v/v) $(\text{NH}_4)_2\text{SO}_4$. Saturated $(\text{NH}_4)_2\text{SO}_4$ solution was added into X-ray Buffer containing GST-purified modular VP1 capsomeres (Section 6.2.5) to final concentrations of 10% (v/v). The mixture was then loaded into the equilibrated column at a flow rate of 1 ml min^{-1} . After sample loading was completed, the column was washed with X-ray Buffer containing 10% (v/v) $(\text{NH}_4)_2\text{SO}_4$ until the UV absorbance reading returned to baseline.

X-ray Buffers with four different concentrations of $(\text{NH}_4)_2\text{SO}_4$ (9%, 7%, 5% and 0% (v/v)) were prepared for elution. The elution started with 6-7 CV of X-ray Buffer containing 9% (v/v) $(\text{NH}_4)_2\text{SO}_4$, and continued step-by-step using the same volume of X-ray Buffer in decreasing concentrations of $(\text{NH}_4)_2\text{SO}_4$. The eluted fractions were collected as 1 ml volumes continuously.

6.2.7 Concentrating HIC-purified modular VP1 capsomeres by precipitation

Saturated $(\text{NH}_4)_2\text{SO}_4$ solution was added to HIC-purified modular VP1 capsomeres (from Section 6.2.6.2) to a final concentration of 50% (v/v). The mixtures were then incubated at 4°C , with gentle rolling. After at least 2 h of incubation, the mixtures were centrifuged and the precipitants were subsequently dissolved in X-ray Buffer.

6.2.8 Protein concentration measurement

Protein concentration was determined using UV absorbance at 280 nm, based on the Beer-Lambert Law (Aitken and Learmonth 1996):

$$A = \epsilon b c$$

A is the measured absorbance at 280 nm, ϵ is the extinction coefficient of protein ($M^{-1} cm^{-1}$) at 280 nm measured in water, b is sample path length (10 mm), and c is protein concentration (M). Theoretical molecular weight and extinction coefficient of each modular protein was obtained using the ProtParam tool (Gasteiger et al. 2005), and summarised in Table 6-2.

Table 6-2. Theoretical molecular weights and extinction coefficients of proteins used in Chapter 6.

Protein	Theoretical molecular weight of monomeric protein ($g mol^{-1}$)	Extinction coefficient of monomeric protein, at 280 nm in water ($M^{-1} cm^{-1}$)
H190-H190	35585.8	60975
TEVP-H190-H190		
GCN4-H190-GCN4	36211.7	57995
TEVP-GCN4-H190-GCN4		
H190	33855.1	57995
TEVP-H190		
wt-VP1 (wt-VP1 residues 1-384)	42763.6	58057
TEVP-wt-VP1 (wt-VP1 residues 1-384)		
TEVp	28600	30035

6.2.9 SDS-PAGE

Protein analysis using sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) was performed as described in Section 4.2.4. Each SDS-PAGE gel picture in Figures 6-4 – 6-6, Figure 6-10, Figure 6-11, and Figure 6-15 was taken from one gel. Relevant sample lanes were selected and positioned on the right hand side of the ladder. The SDS-PAGE gel pictures in Figure 6-8 were taken from different gels which were aligned to a ladder.

6.2.10 SEC-MALS

The quaternary structure of HIC-purified modular VP1 capsomeres was analysed using SEC coupled with multi-angle light scattering (MALS). Samples in X-ray Buffer were loaded isocratically into Superdex 200 10/300 GL (GE Healthcare, UK) at a flow rate of $0.5 ml min^{-1}$.

6.2.11 Western blot

Proteins were resolved in SDS-PAGE gel as described in Section 6.2.9. The resolved protein was then transferred onto Amersham Hybond ECL nitrocellulose membranes (GE Healthcare, UK). Membranes were blocked with blocking buffer (PBS, 0.5% (v/v) Tween 20, 5% (w/v) milk) for 1 h at room temperature. After six 5-min washes with PBST (PBS, 0.5% (v/v) Tween 20), membranes were incubated with mouse GST monoclonal antibody (Catalogue No. 13-6700, Novex®, Life Technologies, Carlsbad, CA, USA) 200-fold diluted in PBSTM (PBS, 0.5% (v/v) Tween 20, 0.5% (w/v) milk) for 1 h at room temperature. The membranes were washed 6 times (5 min each) with PBST, and further incubated with horseradish peroxidase (HRP)-conjugated goat anti-mouse IgG antibodies (Sigma Aldrich, St. Louis, MO, USA) at a 10,000-fold dilution in PBSTM. The incubation was performed for 1 h at room temperature. HRP activity was probed by membrane staining using Novex® Chemiluminescent Substrate (Life Technologies, Carlsbad, CA, USA). After 5-min development in the dark, the stained membrane was placed on the imaging surface of a Molecular Imager® Gel Doc™ XR System (Bio-Rad, Hercules, CA, USA) and digital acquisition of the luminescence signal was recorded.

6.2.12 Expression and purification of TEVp

TEVp was expressed as a His-tagged protein and purified using affinity chromatography as previously described (Cabrita et al. 2007). Plasmids were transformed into *E. coli* Rosetta (DE3) pLysS cells (EMD Millipore, Merck KGaA, Darmstadt, Germany) using the heat shock method (Sambrook and Russell 2001). A single colony of the transformed cells was inoculated in 5 ml Luria Bertani (LB) medium overnight at 30°C, 180 rpm. Then, 400 µl of the seed culture was cultivated in 400 ml of LB medium at 37°C, 180 rpm. When the cell OD_{600 nm} value reached 0.5, culture was cooled down to 26°C and subsequently induced with isopropyl β-D-1-thiogalactopyranoside (IPTG) at a final concentration of 1 mM. The culture was then incubated at 26°C, 180 rpm for 16 h, and harvested by centrifugation at 4°C, 6000g for 20 min. All cultures were supplemented with 100 mg L⁻¹ of ampicillin and 34 mg L⁻¹ of chloramphenicol.

The harvested cell pellet from 800 ml culture was resuspended in 40 ml of TEVp-Lysis Buffer (20 mM sodium phosphate, 500 mM NaCl, 1% (v/v) Triton X-100, 20 mM imidazole, pH 8). One tablet of cOmplete, EDTA-free Protease Inhibitor Cocktail Tablet (Catalogue No. 05056489001; Roche, Mannheim, Germany) was added into the buffer prior to cell resuspension. Then, the suspension was sonicated for 4 cycles of 40 s using a Branson Sonifier 250 cell disruptor mounted with a microtip (Branson Ultrasonics

Corporation, Connecticut, USA), and subsequently clarified by centrifugation at 4°C, 18500g for 30 min. The supernatant was collected and filtered using a 0.45 µm filter (Pall, New York, USA).

A HisTrap FF 1 ml column (GE Healthcare, UK) was equilibrated with 10 CV of TEVp-Lysis Buffer. The clarified supernatant was then loaded into the column at a flow rate of 1 ml min⁻¹. Meanwhile, TEVp-Lysis Buffers containing 39.2 mM and 500 mM imidazole were prepared for elution. After the supernatant was completely loaded, the column was washed with 5 CV of TEVp-Lysis Buffer containing 39.2 mM Imidazole. Afterwards, the column was washed with an increasing concentration of imidazole to 500 mM. The increasing concentration was performed gradually over 15 CV. Finally, the column was washed with 7 CV of TEVp-Lysis Buffer containing 500 mM imidazole.

Selected fractions of purified His-tagged TEVp were applied to a Superdex 75 10/300 GL column (GE Healthcare, UK). The column was pre-equilibrated with Running Buffer (20 mM sodium phosphate, 200 mM NaCl, 10% (v/v) glycerol, 5 mM DTT, pH 8) before sample loading. The purified His-tagged TEVp was loaded at 0.8 ml min⁻¹. TEVp was eluted at a volumetric outflow of approximately 11.5 ml after injection.

6.3 Results and Discussion

6.3.1 Solubility of GST-tagged modular VP1 bearing thrombin and TEVp cleavage sites

The six modular constructs of modular VP1 generated in this chapter (see Section 6.2.1) were expressed as GST-tagged proteins, as described in Section 6.2.2. For the expression of modular VP1 proteins bearing a thrombin cleavage site, wt-VP1 (wt-VP1 residues 1-384) was used as a comparison, while for the expression of modular VP1 proteins bearing a TEVp cleavage site, TEVP-wt-VP1 (wt-VP1 residues 1-384) was used as a comparison. Construct TEVP-wt-VP1 was previously reported in the study by Connors et al. (Connors et al. 2013). As described in Section 6.2.2, following an overnight induction, cell cultures were harvested. The harvested cells were resuspended in Lysis Buffer, and lysed by sonication. After sonication, the suspensions were centrifuged to remove cell debris. Protein solutions were sampled before and after centrifugation, and the samples were analysed using SDS-PAGE.

Figure 6-4 and Figure 6-5 show the SDS-PAGE analysis results for the GST-tagged modular VP1 bearing thrombin and TEVp cleavage sites, respectively. Figure 6-4 shows that after centrifugation, the majority of GST-tagged wt-VP1 was found in the soluble fractions, showing the high solubility of the protein (Chuan et al. 2008, Lipin et al. 2008,

Middelberg et al. 2011). Similar observation was also obtained for GST-tagged GCN4-H190-GCN4. The majority of GST-tagged GCN4-H190-GCN4 was found in the soluble fractions after centrifugation, indicating that the protein had a similar solubility to GST-tagged wt-VP1. In comparison, a smaller proportion of soluble to total protein was observed for GST-tagged H190. These observations indicated that the removal of 31 residues from the N-terminus and 63 residues from the C-terminus affect the solubility of GST-tagged modular VP1, but the effect is a construct-dependent. Furthermore, the figure shows that the smallest proportion of soluble to total protein was observed for GST-tagged H190-H190, indicating that the majority of the protein had precipitated. The lower solubility of GST-tagged H190-H190 compared to the other two modular VP1 proteins was possibly due to a single amino acid mutation that occurred during cloning. The mutation was from Ser81 to Asn81.

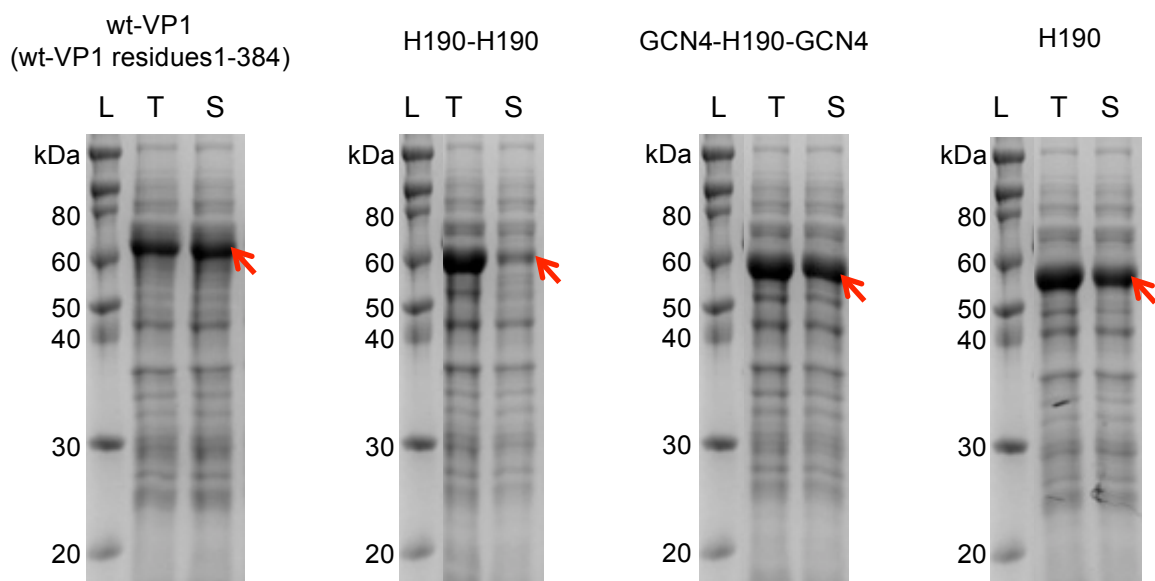


Figure 6-4. SDS-PAGE analysis of the solubility of GST-tagged modular VP1 and wt-VP1 each bearing a thrombin cleavage site. (L) Pre-stained protein marker, (T) Total protein, and (S) Soluble protein. Red arrows represent GST-tagged VP1.

Similar to the observation for GST-tagged wt-VP1 in Figure 6-4, Figure 6-5 shows that after centrifugation, the majority of GST-tagged TEVP-wt-VP1 was found in the soluble fraction. The figure also shows that the majority of GST-tagged TEVP-H190-H190 and TEVP-GCN4-H190-GCN4 was found in the soluble fractions. In contrast, a small

proportion of GST-tagged TEVP-H190 was found in the soluble fraction, showing that the protein had a lower solubility than the other proteins.

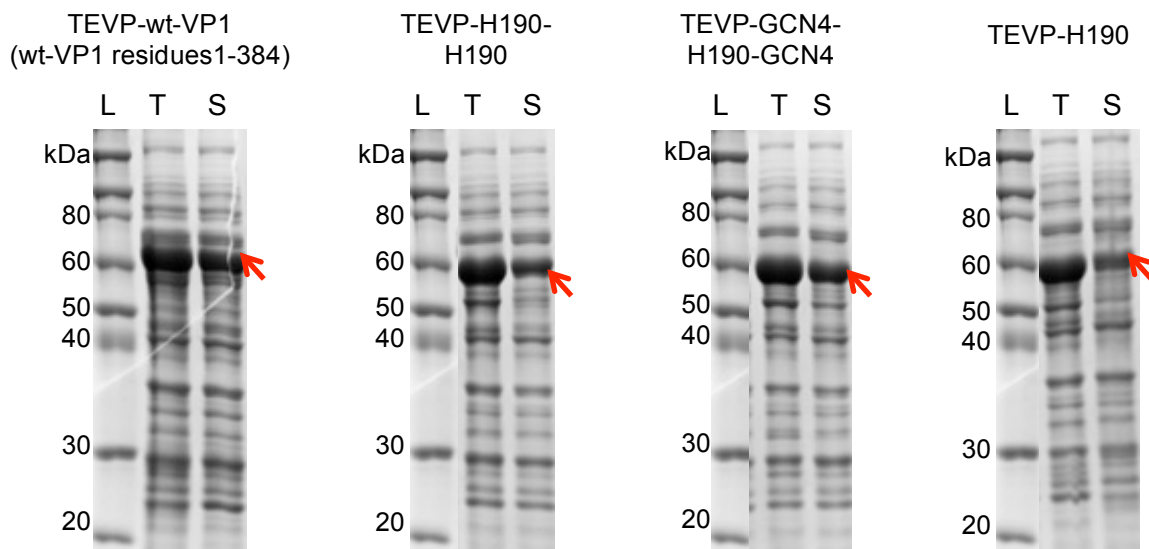


Figure 6-5. SDS-PAGE analysis of the solubility of GST-tagged modular VP1 and wt-VP1 each bearing a TEVp cleavage site. Lanes: (L) Pre-stained protein marker, (T) Total protein, and (S) Soluble protein. The dominant bands represent GST-tagged VP1 proteins.

The solubility analysis result in Figure 6-5 was compared with the result in Figure 6-4. The comparison was applied for all constructs except H190-H190 and TEVP-H190-H190. The solubility result of H190-H190 could not be compared to the solubility result of TEVP-H190-H190 because construct H190-H190 had a single amino acid mutation. Comparison between wt-VP1 and TEVP-wt-VP1 showed that both proteins had a similar solubility. This observation indicated that mutation from a thrombin to TEVp cleavage site did not affect the solubility of GST-tagged wt-VP1. In contrast, the solubility of GST-tagged TEVP-GCN4-H190-GCN4 and TEVP-H190 was lower than the solubility of GST-tagged GCN4-H190-GCN4 and H190. It has been shown that mutation from a thrombin to TEVp cleavage site did not affect the solubility of GST-tagged wt-VP1. Thus, the lower solubility for the modular VP1 bearing TEVp cleavage site may rather be a construct-dependent effect. Such observation has been reported in Chapter 5 of this thesis (Section 5.3.3.1). Results in the section show that the solubility of GST-tagged modular VP1 bearing two copies of H190 from *A/Victoria/210/2009* influenza was lower when a thrombin cleavage site in the protein was mutated into a TEVp cleavage site. In contrast, the solubility of

GST-tagged modular VP1 bearing two copies of H190 from A/California/07/2009 influenza was not affected by the mutation.

Figure 6-5 shows that GST-tagged TEVP-H190 had the lowest solubility compared to GST-tagged TEVP-H190-H190 and TEVP-GCN4-H190-GCN4. Protein solubility has been linked to the ability of the protein to form crystals. Increased solubility has been correlated to a decrease in the crystal-liquid surface tension, resulting in a reduction in the energetic barrier to nucleation and an improved ability of the protein to form crystals (Izaac et al. 2006). Therefore, the solubility of GST-tagged TEVP-H190 required improvement. In this chapter, in an attempt to improve the solubility of GST-tagged TEVP-H190, IPTG induction during protein expression (refer to Section 6.2.2) was performed at two different $OD_{600\text{ nm}}$ values: 0.5 and 0.8. The effect of IPTG induction at different $OD_{600\text{ nm}}$ values on the solubility of the modular VP1 is shown in Figure 6-6. The figure shows that, following induction at $OD_{600\text{ nm}}$ value of 0.5, the majority of GST-tagged TEVP-H190 precipitated. In contrast, following induction at a higher $OD_{600\text{ nm}}$ value of 0.8, the majority of GST-tagged TEVP-H190 was found in the soluble fraction. This result shows that IPTG induction at $OD_{600\text{ nm}}$ of 0.8 resulted in an improved solubility of GST-tagged TEVP-H190. Based on this finding, IPTG induction for the expression of GST-tagged TEVP-H190 was performed at $OD_{600\text{ nm}}$ of 0.8.

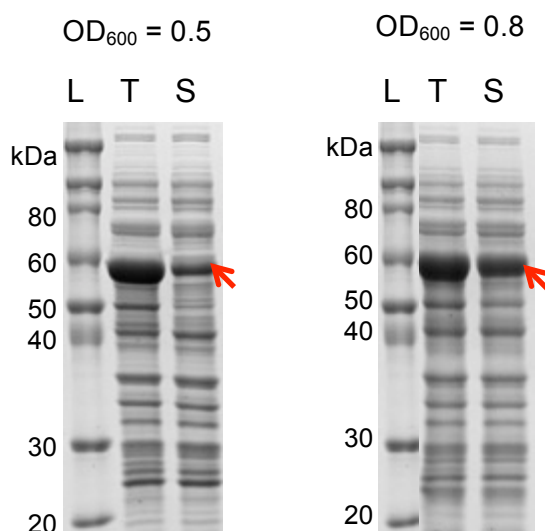


Figure 6-6. SDS-PAGE analysis showing the impact of $OD_{600\text{ nm}}$ values during IPTG induction on the solubility of GST-tagged modular VP1 bearing one copy of H190 and a TEVp cleavage site (construct TEVP-H190). Lanes: (L) Pre-stained protein marker, (T) Total protein, and (S) Soluble protein. Red arrows represent GST-tagged TEVP-H190.

The impact of OD_{600 nm} value during IPTG induction on the solubility of proteins during expression has also been reported by two other studies. The study by Galloway et al. (Galloway et al. 2003) reported that the solubility of recombinant ACF expressed in *E. coli* BL21 cells was higher when IPTG induction was performed at a late-log phase (OD_{600 nm} value of 1.7) than at a mid-log phase (OD_{600 nm} value of 0.6). Galloway et al. reported that they did not know why a higher OD_{600 nm} value could increase the solubility of recombinant ACF, but they suggested two possible explanations. Firstly, at a late-log phase, the bacteria cells have undergone a metabolic and growth shift, thus they become more tolerant towards foreign and potentially toxic proteins. Secondly, at a late-log phase, the cells grow slower, thus the rate of protein synthesis decreases. Hence, less aggregation will occur. In agreement with Galloway et al., the study by Choi et al. (Choi et al. 2000) showed that the solubility of recombinant *E. coli* HB101 harbouring pTrcS1PhoA was higher when IPTG induction was performed at OD_{600 nm} value of 150 than OD_{600 nm} value of 50. Like Galloway et al., Choi et al. reported that the reason of such observation was unclear; however, they suggested that the final concentration of IPTG per cell is possibly lower at a higher OD_{600 nm} value. Hence, induction at a higher cell density gives the same effects as induction with a lower IPTG concentration. Although the recombinant proteins studied by Galloway et al. and Choi et al. were not related to MuPyV VP1, these studies provided possible explanations for the observations regarding the solubility of GST-tagged TEVP-H190.

6.3.2 GST tag removal efficacy of TEVp and thrombin

Following purification of GST-tagged modular VP1 using affinity chromatography, GST was enzymatically removed from modular VP1 using thrombin or TEVp. The purified GST-tagged modular VP1 was incubated with thrombin or TEVp, accordingly, as described in Section 6.2.3.

GST tag removal from modular VP1 using thrombin was performed according to the protocol optimised for wt-VP1 (VP1 residue 1-384) (Chuan et al. 2008). Using this protocol, thrombin digestion was performed for 2 h at room temperature with the ratio of GST-tagged modular VP1 to thrombin of 40:1 (thrombin unit/ml protein). Figure 6-7 shows SDS-PAGE analysis of GST tag removal from modular VP1 using the aforementioned digestion condition. Lanes 1 and 2 of the figure show that after incubating GST-tagged wt-VP1 (red arrow) with thrombin, two major products were observed. These products were (i) wt-VP1, which was shown as a band at about 40 kDa in the gel (blue arrow), and (ii) the GST tag, which was shown as a band at about 25 kDa in the gel (green arrow).

Furthermore, GST-tagged wt-VP1 was not detected in lane 2 of the figure, indicating that thrombin effectively removed GST from the wt-VP1.

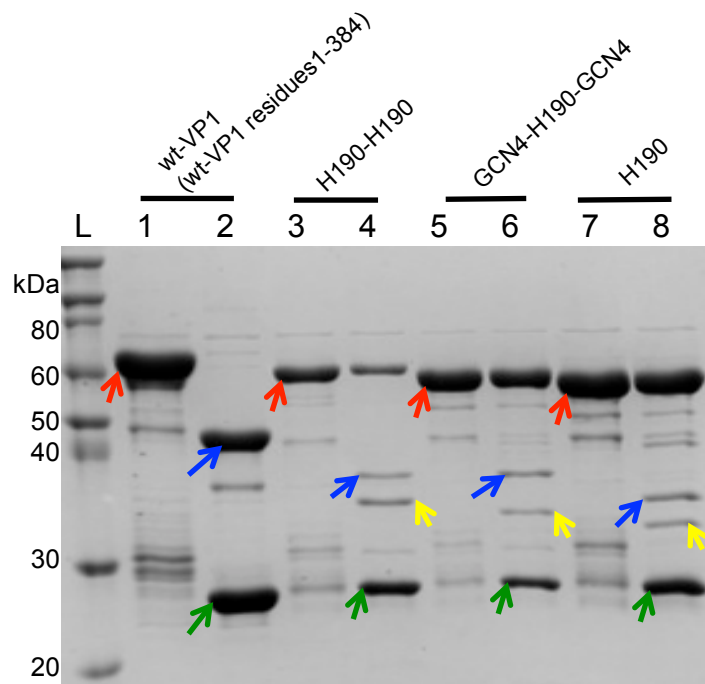


Figure 6-7. SDS-PAGE analysis of GST-tag removal from modular VP1 and wt-VP1 (wt-VP1 residues 1-384) using thrombin. Lanes: (L) Pre-stained protein marker, (1) GST-tagged wt-VP1, (2) Thrombin treated GST-tagged wt-VP1, (3) GST-tagged H190-H190, (4) Thrombin treated GST-tagged H190-H190, (5) GST-tagged GCN4-H190-GCN4, (6) Thrombin treated GST-tagged GCN4-H190-GCN4, (7) GST-tagged H190, (8) Thrombin treated GST-tagged H190. Red, blue, and green arrows refer to GST-tagged VP1, VP1, and GST tag, respectively. Yellow arrows refer to thrombin secondary products.

In contrast, GST tag removal using thrombin for modular constructs H190-H190, GCN4-H190-GCN4, and H190 was less effective. The lower effectiveness of thrombin is shown in lanes 3-8 of Figure 6-7. The lanes show that for each modular construct, a considerable amount of GST-tagged modular VP1 (red arrows) was detected after 2 h treatment with thrombin. Increasing both the ratio of thrombin to modular VP1 and incubation time during the treatment did not increase the effectiveness of thrombin to remove GST (data not shown). The lower thrombin effectiveness to remove GST from the three modular VP1 was likely due to the removal of 31 residues from the N-terminal of VP1. Although two glycine residues were added before Gly31, removal of these 31 residues may cause the elimination of a considerable space between GST and VP1, making it difficult for thrombin to access its cleavage site. In addition, it was possible that the addition of two glycine residues before Gly32 may cause a higher flexibility near the

cleavage site; thus, it was difficult for thrombin to access its cleavage site. Furthermore, lanes 3-8 show that, in addition to GST tag (green arrows), GST removal using thrombin for the three modular constructs resulted in: (i) modular VP1, which was shown by blue arrows, and (ii) secondary digestion product, which was shown by yellow arrows. These results show that GST removal using thrombin for modular constructs H190-H190, GCN4-H190-GCN4, and H190 was not only ineffective, but also resulted in heterogeneous products.

This chapter also explored the use of TEVp as a comparison to thrombin. TEVp has been used to remove GST from wt-VP1 (wt-VP1 residues 1-384) in the study by Connors et al. (Connors et al. 2013). The study shows that TEVp efficiently removed GST from wt-VP1 at 30°C. The digestion temperature is within the range of temperature for an optimum enzymatic activity of TEVp, which is 30-34°C (Connors et al. 2013). Based on this study, GST removal optimisation in this chapter was performed at 30°C. In addition, the study by Connors et al. also showed the ratio of TEVp to VP1 and incubation time required to remove GST from wt-VP1, which was 50:1 (mass ratio) for 2 h. However, in contrast to the study by Connors et al., in this chapter, TEVp was utilised to remove GST from modular VP1 lacking 31 residues from the N-terminus and 63 residues from the C-terminus. Therefore, in this chapter, the TEVp/VP1 ratio and incubation times were optimised for each modular construct.

GST removal for modular constructs TEVP-H190-H190, TEVP-GCN4-H190-GCN4, and TEVP-H190 using TEVp was then analysed using SDS-PAGE, as shown in Figure 6-8A-C. Figure 6-8 shows that for all three modular constructs, two major products were observed following treatment with TEVp. These products were: (i) modular VP1, as shown by the blue arrows, and (ii) the GST tag, as shown by the green arrows. This result shows that TEVp could remove GST from modular VP1 for all three modular constructs, and homogeneous products were obtained. This observation was in contrast to the observation of GST removal using thrombin (Figure 6-7). Thrombin was shown to be ineffective for removing GST from the three modular VP1, and heterogeneous products were obtained.

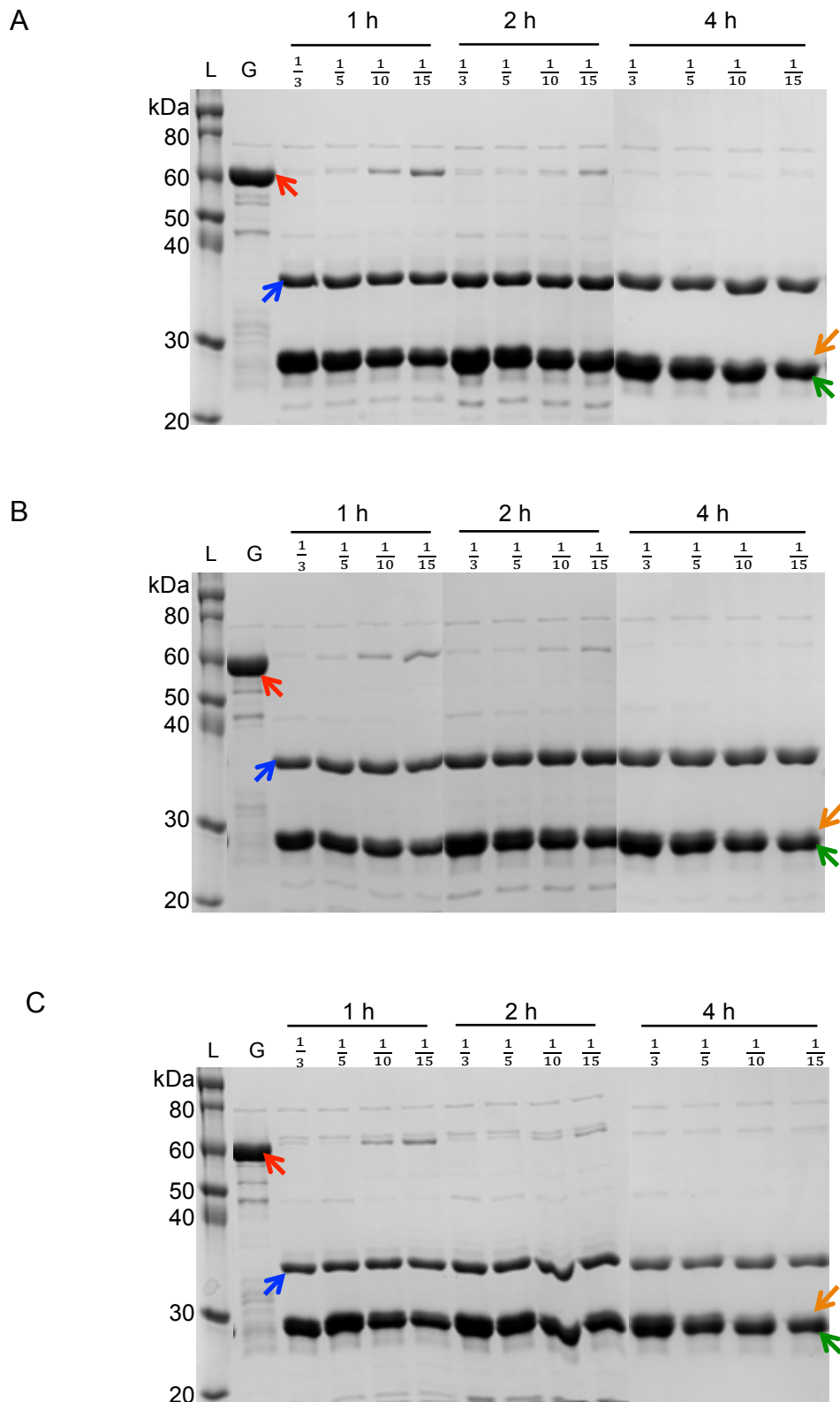


Figure 6-8. SDS-PAGE analysis showing the impacts of incubation time and TEVp to VP1 mass ratio on GST removal at 30°C. (A) TEVP-H190-H190, (B) TEVP-GCN4-H190-GCN4, and (C) TEVP-H190. Lanes: (L) Pre-stained protein marker, (G) GST-tagged modular VP1, (1/3, 1/5, 1/10, 1/15) Mass ratio of TEVp to modular VP1, and (1 h, 2 h, and 4 h) Incubation times. Red, blue, orange, and green arrows refer to GST-tagged modular VP1, modular VP1, TEVp, and GST tag, respectively.

Figure 6-8 also shows that, for all three modular constructs, as the ratio of TEVp to VP1 increased, the amount of undigested GST-tagged modular VP1 (red arrow) detected on the gels decreased. This result indicates that a higher efficiency of GST removal was achieved at a higher TEVp to VP1 ratio. Similarly, a higher efficiency of GST removal was achieved at a longer incubation time. The figure shows that, for each TEVp/VP1 ratio, the amount of undigested GST-tagged modular VP1 for 4 h digestion was less than 2 h and 1 h digestions. Thus, the results in the figure show that an effective GST removal using TEVp could be achieved at a higher TEVp to VP1 ratio and longer incubation time. Furthermore, the results in Figure 6-8 show that, at a digestion temperature of 30°C, the optimum GST removal for all three modular constructs was achieved at TEVp/VP1 ratio of 1:10 with incubation time of 4 h. The optimum TEVp/VP1 ratio and incubation time, however, was considerably higher than the ratio required to remove GST from wt-VP1 (wt-VP1 residues 1-384) efficiently. Such a difference was expected, and most likely due to the removal of 31 residues from the N-terminus.

During optimisation, it was also noticed that, after 1 h treatment at 30°C, precipitations were observed for all modular constructs, as indicated from the turbidity of the protein solutions. Table 6-3 compares the turbidity of protein solutions for the three modular VP1 constructs and TEVP-wt-VP1. As can be seen from the table, protein solutions for the three modular constructs were visually more turbid than the solution of TEVP-wt-VP1, indicating that the modular constructs precipitated more than TEVP-wt-VP1. Furthermore, precipitation of the three modular VP1 suggested that the proteins became unstable at 30°C during extended incubation in the chosen buffer. Unfortunately, the results in this chapter are not sufficient to conclude what caused such instability, but it is likely that the intermolecular factors discussed in Chapters 4 and 5 also operate here. The observed precipitation may be a manifestation of protein-protein interactions under the determined buffer condition, X-ray Buffer. Chapters 4 and 5 have extensively explained that in a concentrated salt solution, such as the X-ray Buffer used during GST removal, protein-protein interactions are determined by the specific ion effects. The impacts of specific ion effects on protein-protein interactions are dependent on the amino acid sequence of the proteins, as outlined in Chapter 4. It is possible that, in the same buffer conditions, the specific ion effects affected protein-protein interactions for TEVP-wt-VP1 and the three modular VP1 proteins differently, and consequently a different degree of precipitation was observed. To minimise precipitations, GST removal of the three modular constructs was alternatively performed at room temperature with a higher TEVp/VP1 ratio of 1:5 for 4 h (data not shown). Because the enzymatic activity of TEVp is not optimum at

room temperature, the higher TEVp to protein ratio was required to achieve effective GST removal.

Table 6-3. Observation of modular VP1 precipitation during GST removal using TEVp at 30°C.

Modular protein	Precipitation
TEVP-H190-H190	++++
TEVP-GCN4-H190-GCN4	++++
TEVP-H190	++++
TEVP-wt-VP1 (VP1 residues 1-384)	+

Comparison between the use of thrombin and TEVp to remove GST for three modular constructs showed that GST removal using thrombin was ineffective, leaving a substantial amount of undigested GST-tagged modular VP1. Additionally, the thrombin treatment also resulted in heterogeneous products. In contrast, GST removal using TEVp was effective, and resulted in homogeneous modular VP1. Therefore, the three modular constructs bearing a thrombin cleavage site, which were H190-H190, GCN4-H190-GCN4, and H190, were not investigated further.

6.3.3 Contamination of SEC-purified modular VP1 capsomeres with undigested GST-tagged modular VP1, GST tag, and 70-kDa contaminant

Following the incubation of GST-tagged modular VP1 with TEVp, digested proteins were clarified using centrifugation to remove precipitates. The supernatants were then loaded into an SEC column to separate modular VP1 capsomeres from soluble aggregates and the GST tag. The size-exclusion chromatograms of TEVP-H190-H190, TEVP-GCN4-H190-GCN4, and TEVP-H190 are shown in Figure 6-9. The figure shows that the chromatograms for all three modular constructs look similar. Furthermore, for all three modular constructs, three major peaks were observed: (i) a peak corresponding to soluble aggregates, which was observed at about 8 ml, (ii) a peak corresponding to capsomeres, which was observed at about 13.5 ml, and (iii) a peak corresponding to the elution of GST tag, which was at about 16 ml. Figure 6-9 shows that for all three modular constructs, capsomeres were eluted at around 13.5 ml. This elution time was later compared to the typical elution time of capsomere wt-VP1 (wt-VP1 residues 1-384), which was at approximately 12.5 ml (Chuan et al. 2010). In SEC, proteins are separated according to their hydrodynamic radius. Proteins with a bigger hydrodynamic radius are eluted earlier than those with a smaller radius. Therefore, the shifted elution time for the three modular constructs suggested that capsomeres TEVP-H190-H190, TEVP-GCN4-

H190, and TEVP-H190 had a smaller hydrodynamic radius than capsomere wt-VP1 (wt-VP1 residues 1-384). The difference in the hydrodynamics radius between the modular VP1 capsomeres and wt-VP1 capsomeres was likely attributed to the removal of residues from both N- and C-termini.

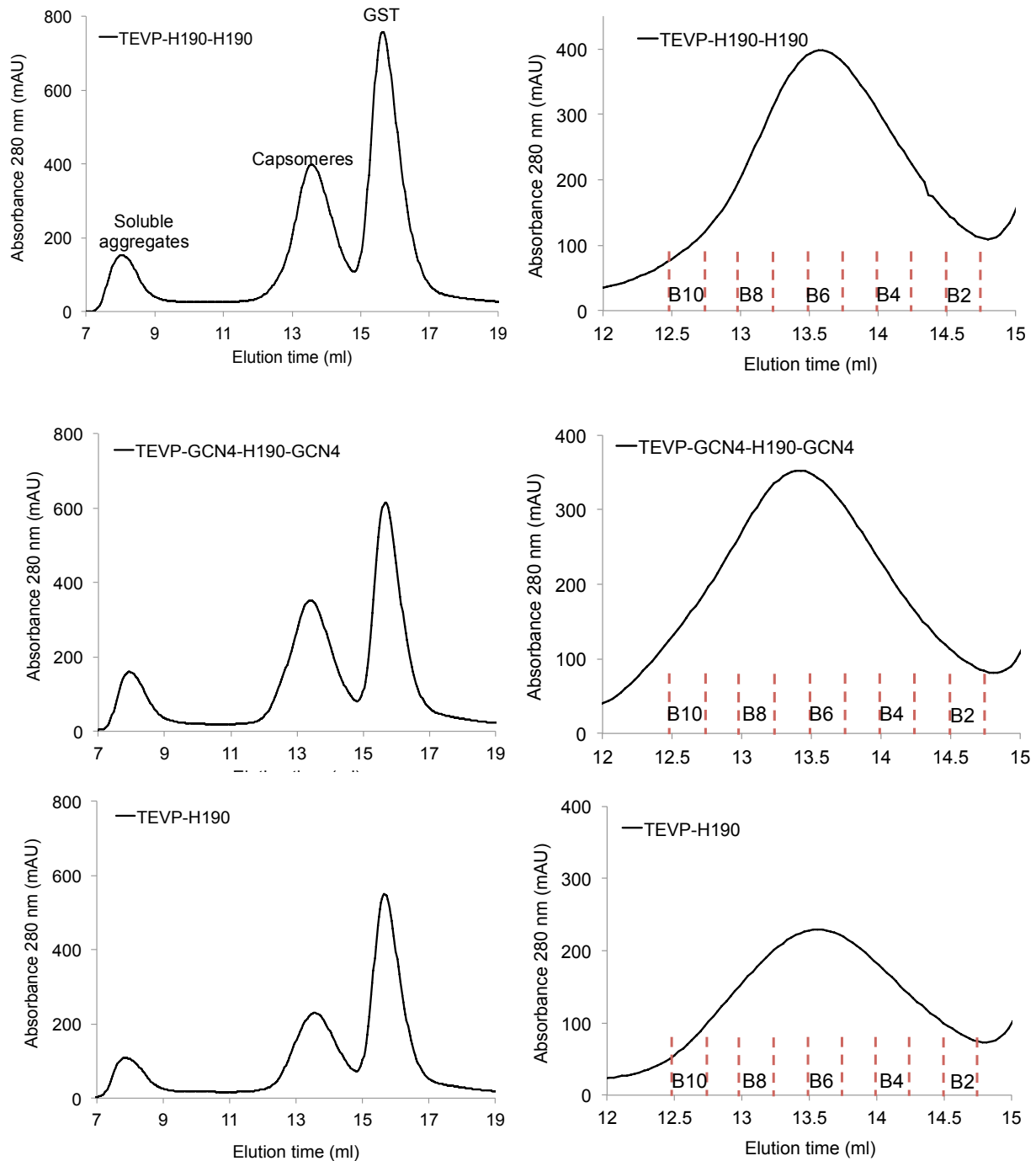


Figure 6-9. Size-exclusion chromatograms of modular VP1 following treatment with TEVP. Left panels show three major peaks on the chromatograms. Right panels show the fractions of capsomere peaks (B10, B8, B6, B4, and B2).

Fractions collected from the soluble aggregate and capsomere peaks were then analysed using SDS-PAGE, as shown in Figure 6-10. The SDS-PAGE analysis results are similar for all three modular constructs. The analysis showed that the soluble aggregates contained both VP1 (blue arrows) and GST (green arrows) in equal proportions. These results are different from those reported in Section 5.3.3.3. In the section, SDS-PAGE analysis result showed that soluble aggregates from SEC purification of modular VP1 bearing two copies of H190 from A/Victoria/2010/2009 influenza were mainly comprised of the modular VP1 protein. Similar results were also obtained for the soluble aggregate peak of modular VP1 bearing two copies of H190 from A/California/07/2009 influenza. The difference between modular VP1 in Section 5.3.3.3 and modular VP1 in this chapter is that the modular VP1 in this chapter lacks of N- and C-terminal residues. Therefore, the difference in the content of the soluble aggregates is likely because of the removal of those N- and C-terminal residues.

The presence of the GST tag was not only observed in fractions from the soluble aggregate peaks, but also in the fractions from the capsomere peaks. Analysis of fractions from capsomere peaks showed that the peaks mainly contained VP1, with a smaller proportion of GST tag detected on the gel. The proportion of GST tag was higher for the later fractions (fractions B6 – B2) than the earlier fractions (fractions B10 – B6). The presence of GST on the fractions of capsomeres indicated that modular VP1 capsomeres were poorly separated from GST during SEC, which was possibly because the capsomere peak was eluted closer to the peak corresponding to the GST tag. This suggestion was supported by the fact that such a high proportion of GST tag in the purified capsomeres was not previously observed in wt-VP1 (wt-VP1 residue 1-384) (Connors et al. 2013). In their study, Connors et al. analysed SEC-purified capsomere wt-VP1 using SDS-PAGE. The analysis showed that the capsomere mainly contained VP1 protein, with a negligible proportion of GST tag. Similarly, a study by Wibowo et al. (Wibowo et al. 2013) showed that following purification in SEC, modular VP1 capsomere lacking 63 residues from the C-terminus contained only a small proportion of GST tag. In the studies by Connors et al. and Wibowo et al., the molecular weight of the monomeric VP1 was bigger than 42 kDa. This molecular weight is bigger than the molecular weight of modular VP1 in this chapter, which was shown to be smaller than 36 kDa. The elution time of wt-VP1 capsomeres in the study by Connors et al. or modular VP1 capsomeres in the studies by Wibowo et al. were farther from the GST compared to modular VP1 capsomeres in this chapter.

In addition, protein separation in SEC is also determined by physical parameters, such as the volume of loaded samples and the buffer flow rate during separation.

Therefore, the poor separation between capsomeres and GST was also possible because optimisation was required for one or both of these physical parameters. Results in Figure 6-9 are obtained by loading about 250 μl of digested proteins to Superdex 200 10/300 GL (GE Healthcare, UK) at a flow rate of 0.5 ml min^{-1} . According to the manufacturer's recommendations, the column can be used to separate 25 – 500 μl samples a flow rate of $0.25 - 0.75 \text{ ml min}^{-1}$. Although the determined conditions are within the recommended values, it may be necessary to further reduce the volume of loaded samples as well as the flow rate during sample loading to increase the resolution of protein separation.

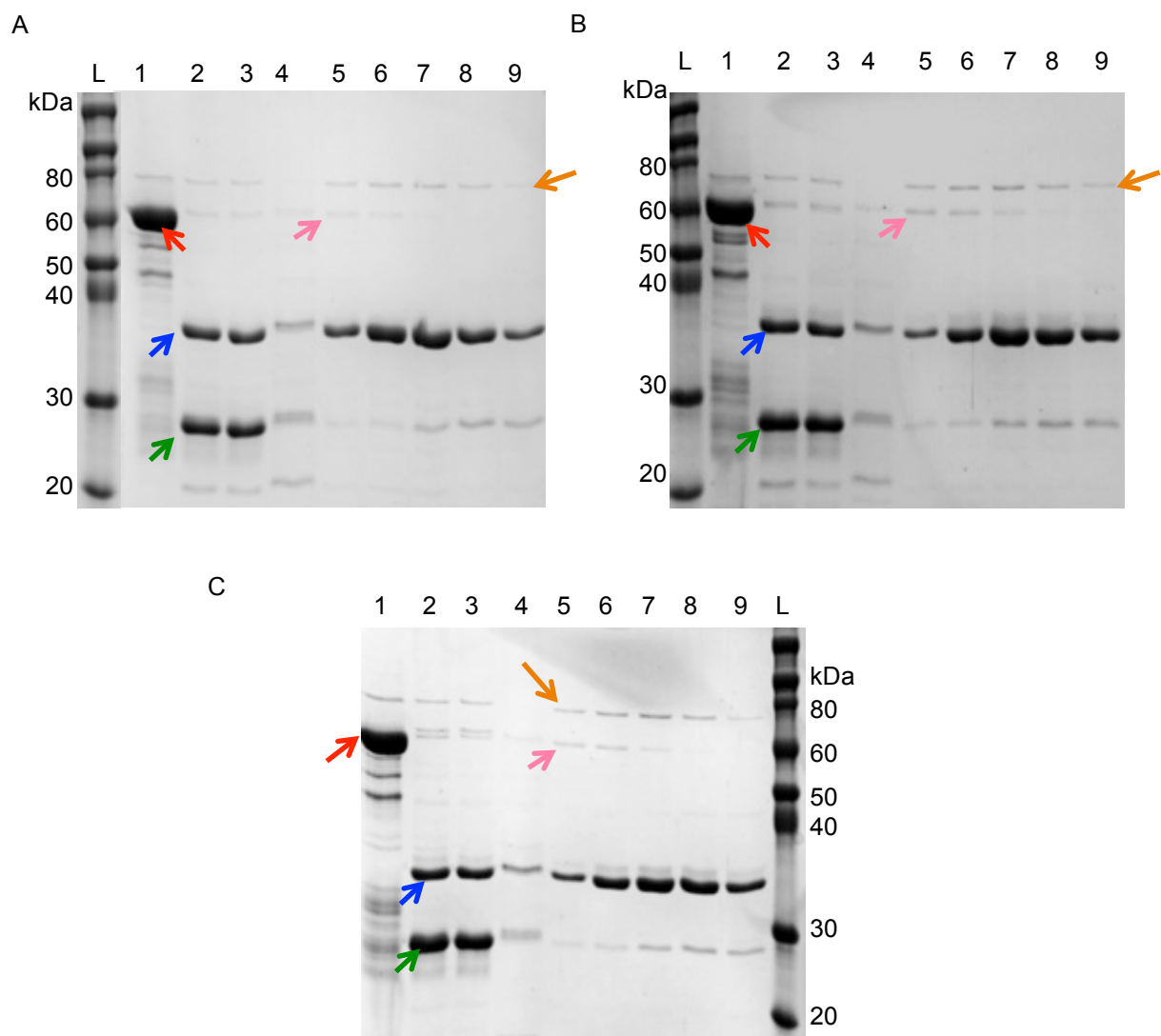


Figure 6-10. SDS-PAGE analysis of eluted peaks on SEC for modular VP1. (A) TEVP-H190-H190, (B) TEVP-GCN4-H190-GCN4, and (C) TEVP-H190. Lanes: (L) Pre-stained protein marker, (1) GST-tagged modular VP1, (2) TEVP-treated GST-tagged modular VP1, (3) Samples analysed in SEC, (4) A fraction from soluble aggregate peak, (5) Fraction B10, (6) Fraction B8, (7) Fraction B6, (8) Fraction B4, and (9) Fraction B2. Red, blue, and green arrows refer to GST-tagged modular VP1, VP1, and GST tag. Orange and pink arrows refer to 70-kDa and 60-kDa contaminants.

Analysis of the fractions from the capsomere peaks also showed that the modular VP1 capsomeres for all three modular constructs contained two other contaminants, which were detected as bands at about 60 kDa (pink arrows) and 70 kDa (orange arrows). These 60-kDa and 70-kDa contaminants were also observed in the study by Fan and Middelberg (Fan and Middelberg 2010). In their study, Fan and Middelberg observed two minor contaminants of SEC-purified wt-VP1 (VP1 residues 1-384) in an analysis using heavily overloaded SDS-PAGE. Peptide mass fingerprinting analysis of these two bands showed that the 60-kDa band matched the GroEL protein from *E. coli*, and the 70-kDa band matched the DnaK protein from *E. coli*. Thus, based on the study by Fan and Middelberg, the 60-kDa and 70-kDa bands observed in Figure 6-10 were suggested to be GroEL and DnaK from *E. coli*, respectively. Alternatively, the 60-kDa contaminant was possibly undigested GST-tagged modular VP1. Analysis using ProtParam Tool (Gasteiger et al. 2005) showed that the theoretical molecular weight of GST-tagged modular VP1 were about 60 kDa, which was similar to the molecular weight of monomeric GroEL. In an attempt to identify the 60-kDa contaminant, western blot analysis using mouse GST monoclonal antibody was performed using SEC-purified modular VP1 capsomeres (Figure 6-11). The result shows that the mouse GST monoclonal antibody recognised the 60-kDa contaminant and GST tag, but not the 70-kDa contaminant. Thus, the analysis suggested that the 60-kDa contaminant was likely to be undigested GST-tagged modular VP1, instead of GroEL protein from *E. coli*.

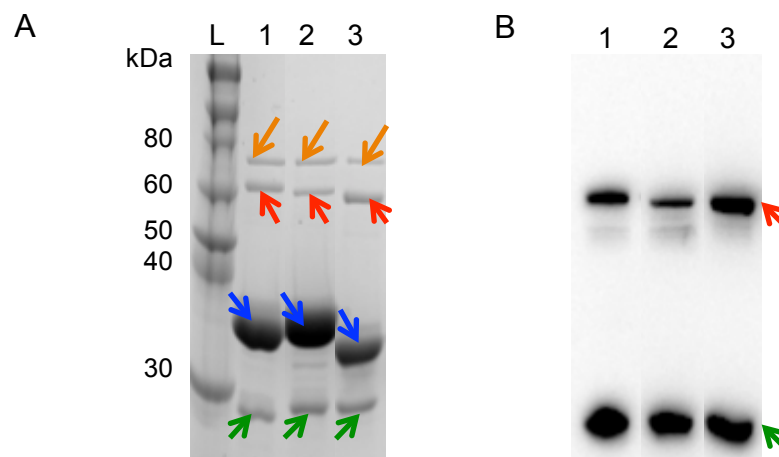


Figure 6-11. Identification of the 60-kDa contaminant in the SEC-purified modular VP1 capsomeres using western blot analysis. (A) SDS-PAGE analysis of SEC-purified modular VP1, and (B) Western blot analysis showing reactivity of mouse GST monoclonal antibody to undigested GST-tagged modular VP1 (red arrows, at about 60 kDa) and GST tag (green arrows, at about 25 kDa). Orange and, blue arrows refer to 70-kDa contaminant and modular VP1. Lane: (L) Pre-stained protein marker, (1) SEC-purified TEVP-H190-H190, (2) SEC-purified TEVP-GCN4-H190-GCN4, and (3) SEC-purified TEVP-H190.

The SDS-PAGE analysis in Figure 6-10 showed that the undigested GST-tagged modular VP1 was detected in earlier fractions of the peaks (fractions B10-B8), but not in later fractions (fractions B6 - B2). The results suggest that the undigested GST-tagged modular VP1 could be removed from purified VP1 capsomeres using SEC, integrated into the purification step where capsomeres were separated from GST tags and soluble aggregates by collecting the later fractions of capsomeres (fractions B6 - B2). However, as mentioned in the preceding paragraph, the later fractions of capsomeres contained a higher proportion of GST tag. Therefore, removal of the undigested GST-tagged modular VP1 using SEC was compensated for by the presence of GST tag. In addition, the SDS-PAGE analysis results show that the 70-kDa contaminant was detected in all fractions for all three modular constructs, indicating that the contaminant could not be removed using SEC.

6.3.4 Removal of GST tag and undigested GST-tagged modular VP1 using tandem GST columns

Section 6.3.3 shows that SEC purified modular VP1 capsomeres contained three contaminants. These contaminants were: (i) undigested GST-tagged modular VP1, which was represented as bands at about 60 kDa, (ii) the GST tag, and (iii) a 70-kDa contaminant. The section also shows that the removal of undigested GST-tagged modular VP1 using SEC was compensated for by contamination of GST tag. Therefore, an alternative purification method was explored to remove both undigested GST-tagged modular VP1 and the GST tag from SEC-purified capsomeres. The alternative purification method was the use of tandem GST-affinity chromatography columns.

Removal of these two contaminants from SEC-purified modular VP1 capsomeres was maximised by increasing the capacity of the columns to capture GST. In this chapter, the increased capture of GST was achieved using two approaches, i.e. (i) the lowered flow rate of sample loading, and (ii) the use of serially-connected GST-affinity chromatography columns. The first approach focuses on the reduction of flow rate during sample loading down to 0.2 ml min^{-1} . The binding of GST to glutathione occurs with slow binding kinetics (Habig et al. 1974). Thus, the reduced flow rate increased the contact time between GST and immobilised glutathione, and consequently allowed more GST to be captured. In the second approach, tandem GST-affinity chromatography columns were used. The first GST-affinity chromatography column was expected to remove the majority of undigested GST-tagged modular VP1 and GST tags from the sample. Exposing the sample containing a residual amount of undigested GST-tagged modular VP1 and GST tag to the second

GST-affinity chromatography column was expected to completely remove both contaminants.

Fractions B10-B2 from the capsomere peak of SEC were pooled and loaded into tandem GST-affinity chromatography columns. Flow-through fractions were collected during sample loading and analysed using SDS-PAGE. Figure 6-12 shows the SDS-PAGE analysis results of fractions of the flow-through from tandem GSTrap columns for modular constructs TEVP-H190-H190, which looked highly similar to the results for TEVP-GCN4-H190-GCN4 and TEVP-H190. The analysis showed that undigested GST-tagged modular VP1 and GST tag was not detected in the flow-through fractions. This result indicated that both contaminants could be removed to below the detection limit from the SEC-purified modular VP1 capsomeres using tandem GSTrap columns. In addition, the analysis showed that the 70-kDa contaminant could not be removed using tandem GST columns, as expected.

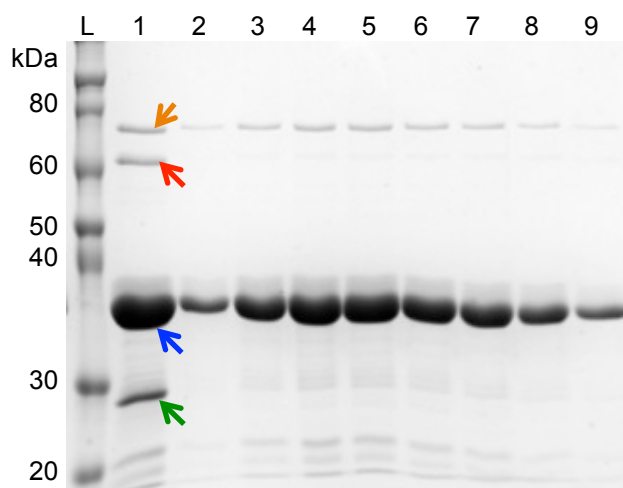


Figure 6-12. SDS-PAGE analysis of flow-through fractions from tandem GST-affinity chromatography columns for modular VP1 bearing two copies of H190 (construct TEVP-H190-H190). Lanes: (L) Pre-stained protein marker, (1) SEC-purified modular VP1 capsomeres, (2-13) Flow-through fractions from tandem GSTrap columns. Orange, red, blue, and green arrows refer to 70-kDa contaminant, undigested GST-tagged modular VP1, modular VP1, and GST tag, respectively.

6.3.5 Removal of the 70-kDa contaminant using HIC

6.3.5.1 In a batch mode

Chapter 4 of this thesis showed that modular VP1 bearing four copies of H190 exhibited a lower solubility and assembly capacity than modular VP1 bearing three copies

of H190 in Assembly Buffer 1, which has been optimised for wt-VP1. The chapter also showed that modification of the type and concentration of salt ions, as well as pH, of Assembly Buffer 1 affected the solubility and the assembly capability of modular VP1 bearing four copies of H190 in the buffer. The results presented in the chapter suggest that these observations reflected protein-protein interactions in a concentrated salt solution, which were dependent on the specific ion effects. The specific ion effects are determined by the position of anions and cations on the lyotropic (Hofmeister) series (Curtis et al. 2002a). The Hofmeister series are shown in Figure 6-13. As explained in Chapter 4, in the series, ions are positioned according to their surface tension, which is their ability to change the hydrogen-binding properties of water (Baldwin 1996). There are two classes of ions in the series, chaotropes and kosmotropes (Curtis et al. 2002a). Chaotropes are salts with low surface tension. They decrease the surface tension between water and proteins, because they interact weakly with water, and break the structure of water molecules. Therefore, chaotropes promote salting-in effects, where the solubility of proteins is increased. In contrast, kosmotropes are salts with high surface tension. They increase the surface tension between water and proteins because they interact strongly with water, forming highly ordered shells. In the presence of kosmotropes, water molecules around the salt ions are more structured than bulk water. Kosmotropes promote salting-out effects, where proteins are forced to merge to minimise the total hydrophobic area of the shells and consequently the solubility of proteins is reduced (Zhang and Cremer 2006).

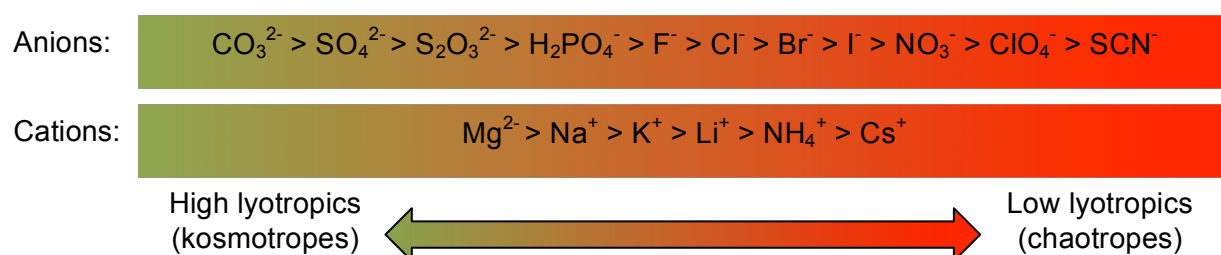


Figure 6-13. The Hofmeister series of anions and cations.

The change in the structure of water molecules is the principal of protein separations in hydrophobic interaction chromatography (HIC). In HIC, interactions occur between hydrophobic regions of proteins and immobilised hydrophobic ligands. The interactions are caused by the salting out effects of salt ions; thus, these interactions are promoted by the addition of kosmotropic salt ions. At high concentrations of kosmotropes, proteins are

adsorbed on the ligand. In contrast, at low concentrations of kosmotropes, proteins are desorbed from the ligand (Xia et al. 2004). However, the increase in kosmotropes concentration is limited up to a point at which proteins precipitate. Furthermore, as flagged by the results in Chapter 4, changes in the structure of water molecules due to the addition of kosmotropes are also dependent on the amino acid sequence of proteins. The addition of a certain kosmotrope at a certain concentration will change the structure of water molecules to a different degree for different proteins. Therefore, at a pre-determined type and concentration of salt ions, interactions between proteins and immobilised hydrophobic ligands will differ from one protein to another, meaning that the proteins can be separated.

Various theories have been developed to explain the adsorption mechanism of proteins on the ligand. The first theory is solvophobic theory by Melander and Horváth (Melander and Horvath 1977) and Melander et al. (Melander et al. 1984). This theory focuses on the association and solvation of salts, proteins, and the ligands. It assumes that the molal surface tension increment of the salts determines protein adsorption (Xia et al. 2004). In this theory, hydrophobic interactions between proteins and ligands are proposed to comprise of two mechanisms: (i) the formation of a cavity in the water above the ligand, and (ii) a protein fills the cavity and adsorbs onto the ligand. The importance of this theory is that it considers the most important parameters on the effects of salt on protein adsorption to be: the salt molality and the molal surface increment of the salt. This theory indicates that, at a high salt concentration, the effect of salt concentrations on the adsorption of proteins should be proportional to the molal surface tension increment of the salt. This indication means that the higher the surface tension increment of the salt, the higher the retention time of a protein (Perkins et al. 1997, Queiroz et al. 2001).

The second theory is based on preferential interaction analysis by Roettger et al. in 1989 (Roettger et al. 1989). This theory correlates protein adsorption to the preferential interactions of salts, proteins, and the ligands. Roettger et al. showed that kosmotropic salts have negative preferential interactions with proteins and stationary phases. Therefore, these salts promote adsorption. This preferential behaviour is in contrast to the behaviour of chaotropic salts. Chaotropic salts have a positive preferential with proteins and the stationary phase, and promote desorption (Queiroz et al. 2001). The advantage of this theory is that, as for the solvophobic theory, this theory is valid for a wide range of salt concentrations. In addition, this theory allows the effects of salts on the observed equilibrium constant to be interpreted in the term of a stoichiometric displacement model (Perkins et al. 1997). This model describes that adsorption of proteins is followed by replacement of a stoichiometric number of water molecules from the interface between

proteins and ligands to the bulk (Geng et al. 1990). Furthermore, preferential interaction analysis has been used to explain various different phenomena, such as protein stabilisation, precipitation, aggregation, ligand binding, and adsorption (Arakawa 1986, Fraaije et al. 1991, Ha et al. 1992).

The third theory is the one by Oscarsson (Oscarsson 1995), which proposes that protein adsorption occurs due to conformational changes. Unlike the other two theories, this theory describes that the effects of salts on protein adsorption is not a simple correlation. According to this theory, protein adsorption occurs because proteins constantly modify their conformations. The conformational changes are enhanced by certain types of ligands and salts, promoting protein interactions with ligands (Queiroz et al. 2001).

The fourth theory was proposed by Lin et al. (Lin et al. 2000), based on the calorimetric studies on the interactions of imidazole (Chen et al. 1996, Wu et al. 1996) and lysozyme (Chen et al. 1997, Lin et al. 1999) in immobilised metal ion affinity chromatography at various pH values and salt concentration ranges. This theory suggests that the adsorption of proteins is comprised of five mechanisms: (i) dehydration or deionisation of proteins, (ii) dehydration or deionisation of the ligand, (iii) hydrophobic interactions between proteins and the ligands, (iv) the structural modification of proteins, and (v) rearrangement of the excluded water molecules in the bulk solution.

In this chapter, HIC was explored to remove the 70-kDa contaminant from GST-purified modular VP1 capsomeres (Section 6.4.3). Hydrophobic ligand Phenyl Sepharose HP (GE Healthcare, UK) was selected, because this resin is based on a 34 μm matrix, which is ideal for laboratory scale separation processes, offering good resolution and capacity trade-off. In addition, a kosmotropic salt, $(\text{NH}_4)_2\text{SO}_4$, was selected. $(\text{NH}_4)_2\text{SO}_4$ is one of the biggest Hofmeister kosmotropes. It has a high solubility and is effective over a wide pH range, between 2 and 10 (Yamniuk et al. 2013). Additionally, $(\text{NH}_4)_2\text{SO}_4$ was selected because it has been known to effectively promote hydrophobic interactions between proteins and ligands due to its bigger surface tension increment (Mahn et al. 2007). At a high concentration, this kosmotropic salt promotes self-association of native hydrophobic surfaces in proteins, and consequently promotes the formation of aggregates (Yamniuk et al. 2013). Furthermore, the sulphate anions have been shown to have a stabilising effect on the native structure of proteins (Arakawa and Timasheff 1982).

A batch mode purification was performed to determine the optimum concentration of $(\text{NH}_4)_2\text{SO}_4$, which allowed modular VP1 to be adsorbed to the resin. The purification utilised SEC-purified capsomere TEVP-H190-H190 as a model. In this simple experiment, saturated $(\text{NH}_4)_2\text{SO}_4$ was added into X-ray Buffer containing capsomere TEVP-H190-H190

to final concentrations of 0%, 2.5%, 5%, 7.5%, 10%, 12.5%, and 15% (v/v). After the addition of $(\text{NH}_4)_2\text{SO}_4$, precipitates were observed for protein solutions containing 12.5% and 15% of $(\text{NH}_4)_2\text{SO}_4$, but not for other concentrations. The protein solutions were then centrifuged to remove precipitates, and the supernatants were incubated with Phenyl Sepharose HP resin. After incubation, the mixtures were centrifuged to separate the resin from the solutions. Samples were taken from the protein solutions before and after incubation with the resin. SDS-PAGE analysis of these samples is shown in Figure 6-14. The analysis shows that the total protein before incubation with the resin at 12.5% and 15% of $(\text{NH}_4)_2\text{SO}_4$ was less than the total proteins at other concentrations. This result confirmed the aforementioned observation that the modular VP1 precipitated at these two concentrations of $(\text{NH}_4)_2\text{SO}_4$. Additionally, the analysis showed that as the concentration of $(\text{NH}_4)_2\text{SO}_4$ increased, the proportion of TEVP-H190-H190 (blue arrow) in the solutions after incubation decreased. This decrease indicated an increase in the proportion of TEVP-H190-H190 adsorbed to the ligand. The effects of $(\text{NH}_4)_2\text{SO}_4$ concentrations on the amount of TEVP-H190-H190 adsorbed to the ligand was in good agreement with literature, which showed that the hydrophobic interactions between proteins and ligands are enhanced at increasing salt concentration (Perkins et al. 1997). The result may also reflect the protein-protein interaction phenomenon discussed at the beginning of this section, as well as in Chapter 4. The X-ray Buffer in this chapter contained 100 mM NaCl. Therefore, the buffer is classified as a concentrated salt solution, in which protein-protein interactions are driven by specific ion effects. Here, the specific ion effects were represented by the variations of the concentration of $(\text{NH}_4)_2\text{SO}_4$ in X-ray Buffer. Furthermore, the SDS-PAGE analysis showed that a complete adsorption of TEVP-H190-H190 to the ligand was achieved at 10% (v/v) of $(\text{NH}_4)_2\text{SO}_4$. The concentration of $(\text{NH}_4)_2\text{SO}_4$ was also optimum, because TEVP-H190-H190 did not precipitate at this concentration.

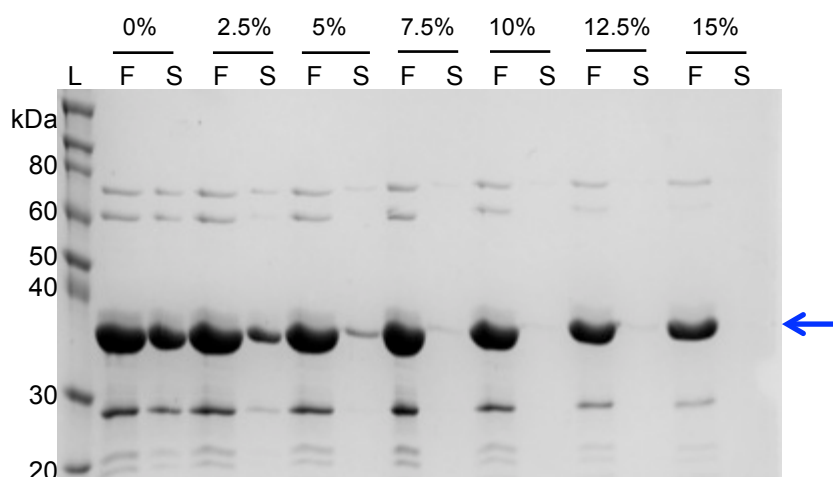


Figure 6-14. SDS-PAGE analysis showing the impact of $(\text{NH}_4)_2\text{SO}_4$ concentrations on the adsorption of modular VP1 bearing two copies of H190 (modular construct TEVP-H190-H190) to Phenyl Sepharose HP resin. Lanes: (L) Pre-stained protein marker, (F) Total protein before incubation with the resin, and (S) Total protein after incubation with the resin. Blue arrow refers to the modular VP1.

6.3.5.2 In bed mode

The results in Section 6.3.5.1 show that it was possible to bind the capsomeres TEVP-H190-H190 in X-ray Buffer to Phenyl Sepharose HP resin, and that complete binding was achieved at 10% (v/v) of saturated $(\text{NH}_4)_2\text{SO}_4$. This finding was then translated into a bed mode purification to remove the 70-kDa contaminant from modular VP1 capsomeres for all three modular constructs: TEVP-H190-H190, TEVP-GCN4-H190-GCN4, and TEVP-H190. Pooled fractions of modular VP1 capsomeres from purification using SEC in X-ray Buffer (as described in Section 6.2.4) were loaded into tandem GST columns (Section 6.2.5). Afterwards, pooled flow-through fractions from the GST columns were mixed with saturated $(\text{NH}_4)_2\text{SO}_4$ to a final concentration of 10% (v/v). The mixtures were then loaded into a HiTrap Phenyl HP column, which was pre-equilibrated with X-ray Buffer containing 10% (v/v) $(\text{NH}_4)_2\text{SO}_4$. Bound proteins were eluted step by step in X-ray Buffer containing 9%, 7%, 5%, and 0% (v/v) $(\text{NH}_4)_2\text{SO}_4$.

Figure 6-15, Figure 6-16, and Figure 6-17 show the chromatograms for modular constructs TEVP-H190-H190, TEVP-GCN4-H190-GCN4, and TEVP-H190, respectively. These figures show that the chromatogram profiles for modular constructs TEVP-H190-H190 and TEVP-H190 were similar, and they were different from the chromatogram profile for TEVP-GCN4-H190-GCN4. This observation indicated that at the determined buffer condition, in this case X-ray Buffer containing 10% (v/v) $(\text{NH}_4)_2\text{SO}_4$, capsomeres for each modular construct interacted differently with the phenyl ligand. As mentioned in Section 6.3.5.1, the adsorption of a protein to a ligand is enhanced by the presence of a

kosmotrope. A kosmotrope interacts strongly with water molecules, forming highly ordered shells. These shells cause the proteins to be thermodynamically unstable, promoting them to adsorb onto the ligand. Therefore, the difference in the capability of modular VP1 capsomeres to adsorb the phenyl ligand under the same buffer composition may reflect the difference in salting out effects of $(\text{NH}_4)_2\text{SO}_4$ on capsomeres from the three modular constructs. The kosmotrope, in this case $(\text{NH}_4)_2\text{SO}_4$, may change the surface tension between water and modular VP1 capsomere molecules differently. Such differences may be caused by possible structural differences of modular VP1 capsomeres for the three constructs. Modularisation of different antigen modules may change the structural integrity of the VP1 protein, exposing residues that were otherwise buried and inaccessible from the surface. However, this possibility is yet to be answered from the resolved crystal structures of capsomeres of the three modular constructs.

Figure 6-15A shows that no peak was observed for TEVP-H190-H190-GCN4 during sample loading, indicating that at the defined buffer condition, the modular VP1 adsorbed strongly onto the Phenyl ligand. Furthermore, during elution, only one peak was observed. The peak was eluted at 0% (v/v) of $(\text{NH}_4)_2\text{SO}_4$, and the peak skewed to the left side (Figure 6-15B). Fractions from this peak were then analysed using SDS-PAGE, as shown in Figure 6-15C. Comparison between lane 1 and 2 in the figure shows that after purification using tandem GST columns, the modular VP1 contained a considerably smaller proportion of undigested GST-tagged TEVP-H190-H190 (red arrow) and GST tag (green arrow). This result was consistent with the findings in the previous section. Furthermore, the SDS-PAGE analysis showed that a small proportion of the 70-kDa contaminant (orange arrow) was detected in the earlier fractions of the eluted peak, which were fractions F1-F3 (lanes 4-6). The proportion of the 70-kDa contaminant was considerably smaller than the proportion of the contaminant in the SEC-purified (lane 1) and GST-purified (lane 2) capsomeres. Furthermore, the contaminant was not detected in the later fractions, which were fractions F4-F7 (lanes 7-10). This result indicates that the contaminant could be removed using a HiTrap Phenyl HP column.

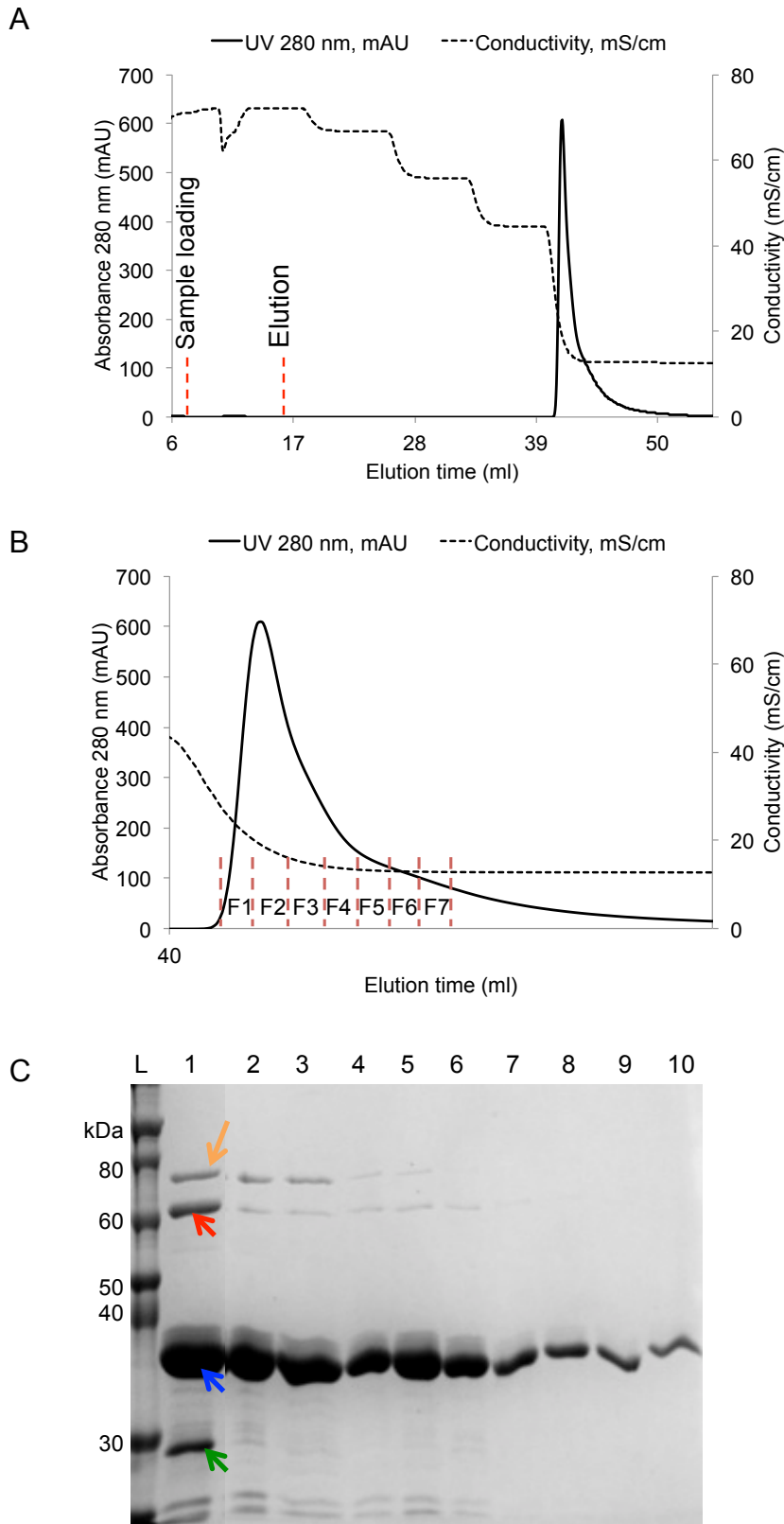


Figure 6-15. Removal of 70-kDa contaminant using HiTrap Phenyl HP for modular VP1 bearing two copies of H190 (construct TEVP-H190-H190). (A) A chromatogram showing sample loading and step elution during purification, (B) A chromatogram showing fractions of the peak eluted at 0% (v/v) $(\text{NH}_4)_2\text{SO}_4$, and (C) SDS-PAGE analysis showing the removal of 70-kDa contaminant. Lanes: (L) Pre-stained protein marker, (1) SEC-purified modular VP1 capsomeres, (2) GST-purified modular VP1 capsomeres, (3) GST-purified modular VP1 capsomeres containing 10% (v/v) $(\text{NH}_4)_2\text{SO}_4$, and (4) Fraction F1, (5) Fraction F2, (6) Fraction F3, (7) Fraction F4, (8) Fraction F5, (9) Fraction F6, and (10) Fraction F7. Orange, red, blue, and green arrows refer to 70-kDa contaminant, undigested GST-tagged modular VP1, modular VP1, and GST tag, respectively.

Different from the chromatogram for TEVP-H190-H190, the chromatogram for TEVP-GCN4-H190-GCN4 in Figure 6-16A shows that two peaks were observed. The first peak was eluted during sample loading until the elution step using X-ray Buffer containing 7% (v/v) of $(\text{NH}_4)_2\text{SO}_4$, indicating that the modular VP1 capsomeres adsorbed weaker to the phenyl ligand. Unfortunately, due to its low absorbance reading, the first peak was not analysed in SDS-PAGE, thus the content of this peak remains unknown. Furthermore, the second peak was eluted at 0% (v/v) of $(\text{NH}_4)_2\text{SO}_4$. Figure 6-16B shows that the eluted peak appears to be a merger of two peaks that were close each other. SDS-PAGE analysis (Figure 6-16C) of fractions from the eluted peak at 0% (v/v) of $(\text{NH}_4)_2\text{SO}_4$ for TEVP-GCN4-H190-GCN4 showed that this peak mainly contained modular VP1, with a negligible amount of the 70-kDa contaminant. Furthermore, the analysis showed that the purity of all fractions from the peak was similar, although Figure 6-16B shows that the peak was comprised of two peaks. This observation was in contrast to the observations in construct TEVP-H190-H190, which showed that the earlier fraction of the peak contained a higher proportion of the 70-kDa contaminant than the later fraction. In addition, the finding shown in Figure 6-16, that the peak eluted at 0% (v/v) of $(\text{NH}_4)_2\text{SO}_4$ contained high purity modular VP1, was in agreement with the findings for TEVP-H190-H190. This suggests that modular VP1 adsorbing strongest to the medium was of high purity.

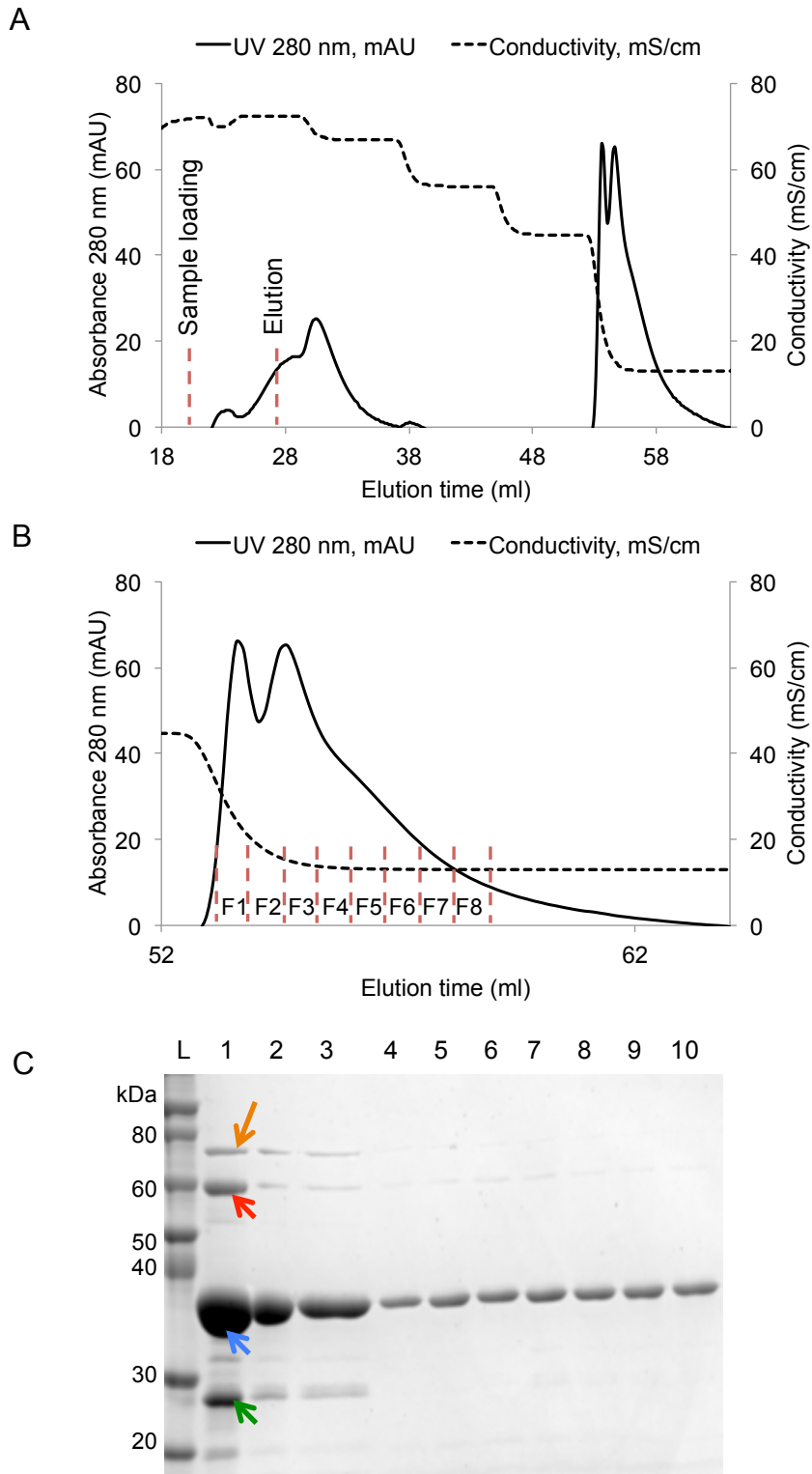


Figure 6-16. Removal of 70-kDa contaminant using HiTrap Phenyl HP for modular VP1 (VP1 residue 32-320) bearing one copy of H190, flanked by GCN4 helix promoter elements (construct TEVP-GCN4-H190-GCN4). (A) A chromatogram showing sample loading and step elution during purification, (B) A chromatogram showing fractions of the peak eluted at 0% (v/v) $(\text{NH}_4)_2\text{SO}_4$, and (C) SDS-PAGE analysis showing the removal of 70-kDa contaminant. Lanes: (L) Pre-stained protein marker, (1) SEC-purified modular VP1 capsomeres, (2) GST-purified modular VP1 capsomeres, (3) GST-purified modular VP1 capsomeres containing 10% (v/v) $(\text{NH}_4)_2\text{SO}_4$, and (4) Fraction F1, (5) Fraction F2, (6) Fraction F3, (7) Fraction F4, (8) Fraction F5, (9) Fraction F6, and (10) Fraction F7. Orange, red, blue, and green arrows refer to 70-kDa contaminant, undigested GST-tagged modular VP1, modular VP1, and GST tag, respectively.

Similar to the chromatogram profile for TEVP-H190-H190, Figure 6-17A shows that for TEVP-H190, the UV absorbance level was maintained at the baseline during sample loading. Like the other two modular constructs, a peak was eluted at 0% (v/v) $(\text{NH}_4)_2\text{SO}_4$ (Figure 6-17B). The chromatogram of TEV-H190 indicated that like capsomere TEVP-H190-H190, capsomere TEVP-H190 bound more strongly to the phenyl ligand than TEVP-GCN4-H190-GCN4. Furthermore, SDS-PAGE analysis (Figure 6-17C) showed that the peak eluted at 0% (v/v) $(\text{NH}_4)_2\text{SO}_4$ mainly contained modular VP1, with a negligible proportion of impurities. Additionally, the analysis showed that the purity of all fractions from the peak was similar. This finding was in good agreement with the analysis result for the other two modular constructs, confirming that modular VP1 bound the strongest to the ligand was of high purity. It is reasonable to suggest that such an observation was linked to the function of a chaperone. In this chapter, the 70-kDa contaminant was not identified, but it was suggested to be the *E. coli* protein, DnaK (Fan and Middelberg 2010). DnaK is the major bacterial Hsp70, which prevents the aggregation of misfolded proteins and promotes refolding. It acts as a chaperone and binds to unfolded polypeptides via hydrophobic regions on the proteins (reviewed in Chapter 2). It was possible that in the presence of 10% (v/v) $(\text{NH}_4)_2\text{SO}_4$, modular VP1 preferred the ligand to DnaK. DnaK binds via hydrophobic regions on the modular VP1; therefore, the lower the proportion of DnaK on the proteins indicates that the more hydrophobic regions are exposed on the proteins. In HIC, protein adsorption to the ligand is modulated via the hydrophobic surfaces on proteins. Thus, the more hydrophobic surfaces available on the proteins, the stronger the protein will be adsorbed to the ligand. Therefore, proteins bound the strongest to the ligands were those with the lowest proportion of DnaK, or in the other words, with the highest purity.

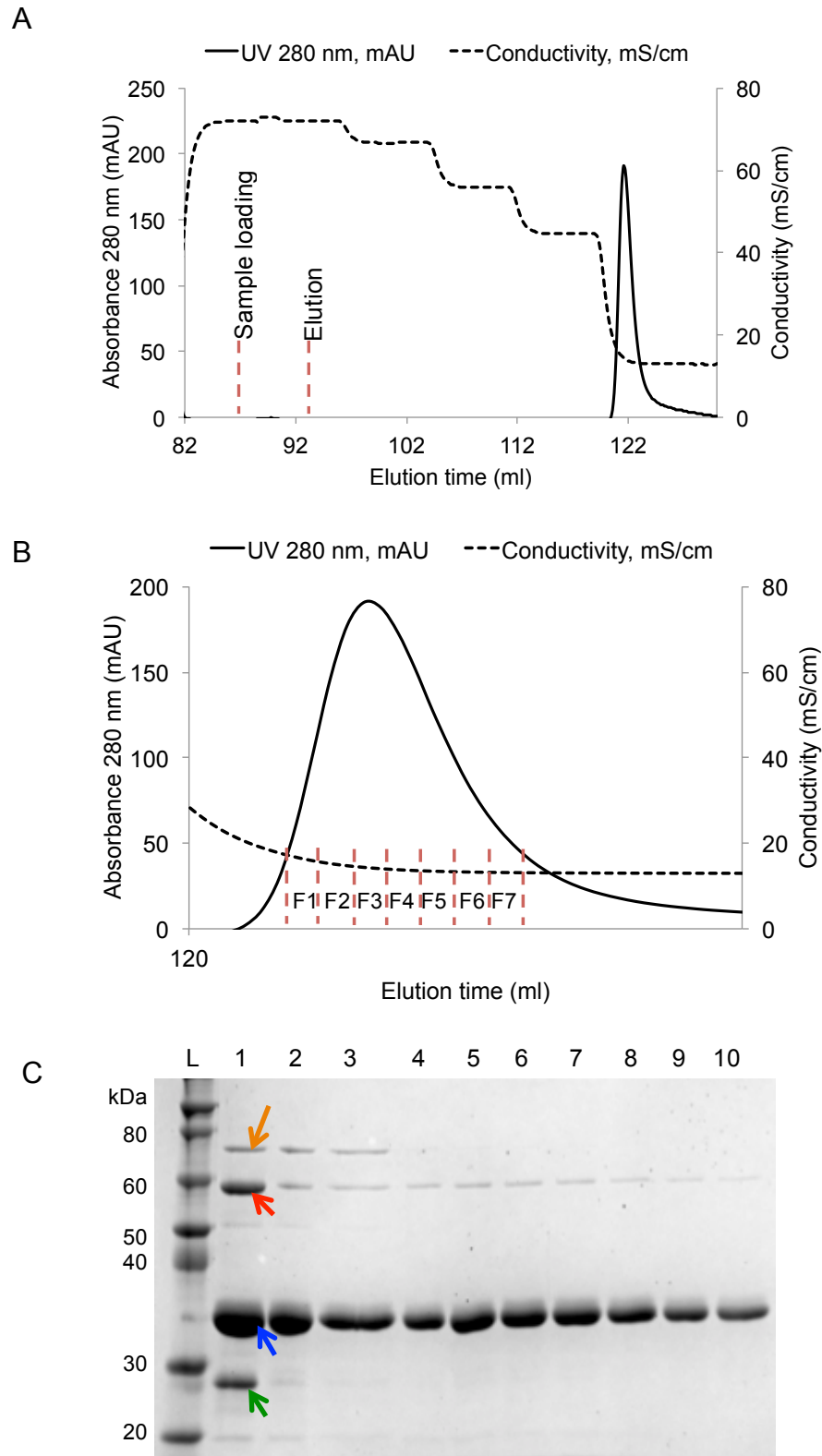


Figure 6-17. Removal of 70-kDa contaminant using HiTrap Phenyl HP for modular VP1 bearing one copy of H190 (construct TEVP-H190). (A) A chromatogram showing sample loading and step elution during purification, (B) A chromatogram showing fractions of the peak eluted at 0% (v/v) $(\text{NH}_4)_2\text{SO}_4$, and (C) SDS-PAGE analysis showing the removal of 70-kDa contaminant. Lanes: (L) Pre-stained protein marker, (1) SEC-purified modular VP1 capsomeres, (2) GST-purified modular VP1 capsomeres, (3) GST-purified modular VP1 capsomeres containing 10% (v/v) $(\text{NH}_4)_2\text{SO}_4$, and (4) Fraction F1, (5) Fraction F2, (6) Fraction F3, (7) Fraction F4, (8) Fraction F5, (9) Fraction F6, and (10) Fraction F7. Orange, red, blue, and green arrows refer to 70-kDa contaminant, undigested GST-tagged modular VP1, modular VP1, and GST tag, respectively.

For all three modular constructs, the majority of the modular VP1 capsomeres was eluted at 0% (v/v) $(\text{NH}_4)_2\text{SO}_4$. For modular constructs TEVP-GCN4-H190-GCN4 and TEVP-H190, SDS-PAGE analysis of fractions from peaks eluted at 0% (v/v) $(\text{NH}_4)_2\text{SO}_4$ showed that the 70-kDa contaminant was not detected. These results indicate that HIC effectively removed the 70-kDa contaminant from capsomeres of both modular constructs, consequently yielding high-purity modular VP1 proteins. However, for TEVP-H190-H190, a considerably smaller proportion of the contaminant was detected. This may indicate that the purification strategy using HIC was less slightly effective for removing the 70-kDa contaminant for the modular VP1 capsomeres. This observation was not in agreement with a previous discussion, which suggested that the chromatogram profiles of TEVP-H190-H190 and TEVP-H190 were similar, but were different to the profile of TEVP-GCN4-H190-GCN4.

HIC purification for modular constructs TEVP-GCN4-H190-GCN4 and TEVP-H190 was then performed on a large scale for crystallisation. Fractions from the peak eluted at 0% (v/v) $(\text{NH}_4)_2\text{SO}_4$ were analysed using SDS-PAGE (Figure 6-18). The analysis showed that, consistent with the results from small-scale purification, for both modular constructs, the eluted peaks contained high purity modular VP1 proteins. The 70-kDa contaminant was not detected in the gel, confirming the effectiveness of HIC at removing the *E. coli* protein from the modular VP1 both on small and large scales.



Figure 6-18. SDS-PAGE analysis of fractions from peaks eluted at 0% (v/v) $(\text{NH}_4)_2\text{SO}_4$ at a large scale HIC purification of modular VP1 capsomeres. (L) Pre-stained protein marker, (1-6) TEVP-H190, and (7-10) TEVP-GCN4-H190-GCN4. Black arrows refer to VP1.

HIC-purified modular VP1 capsomeres were then analysed using SEC coupled with MALS (Figure 6-19). The SEC-MALS analysis results showed that, for both modular constructs, peaks corresponding to soluble aggregates were not observed. The peak should be eluted at approximately 8 ml after sample injection. This observation indicated that following purification using HIC, capsomeres for both modular constructs were stable in solution (20 mM Tris-base, 100 mM NaCl, pH 8.0), and did not form soluble aggregates. Furthermore, Figure 6-19A shows that two major peaks were detected for construct TEVP-GCN4-H190-GCN4. These peaks corresponded to modular VP1 capsomeres and dimeric modular VP1 capsomeres. In contrast, Figure 6-19B shows that one major peak was detected for construct TEVP-H190. The peak corresponded to modular VP1 capsomeres. For TEVP-H190, peaks corresponding to dimeric modular VP1 capsomeres were negligible. The SEC-MAS analysis results showed that under the same buffer conditions, capsomeres for TEVP-GCN4-H190-GCN4 and TEVP-H190 had different quaternary structures. Capsomere TEVP-GCN4-H190-GCN4 had a higher tendency to form dimeric capsomeres than capsomere TEVP-H190. The different quaternary structures between both modular VP1 capsomeres may be because the determined buffer conditions promoted stronger intermolecular interactions for capsomere TEVP-GCN4-H190-GCN4 than for capsomere TEVP-H190.

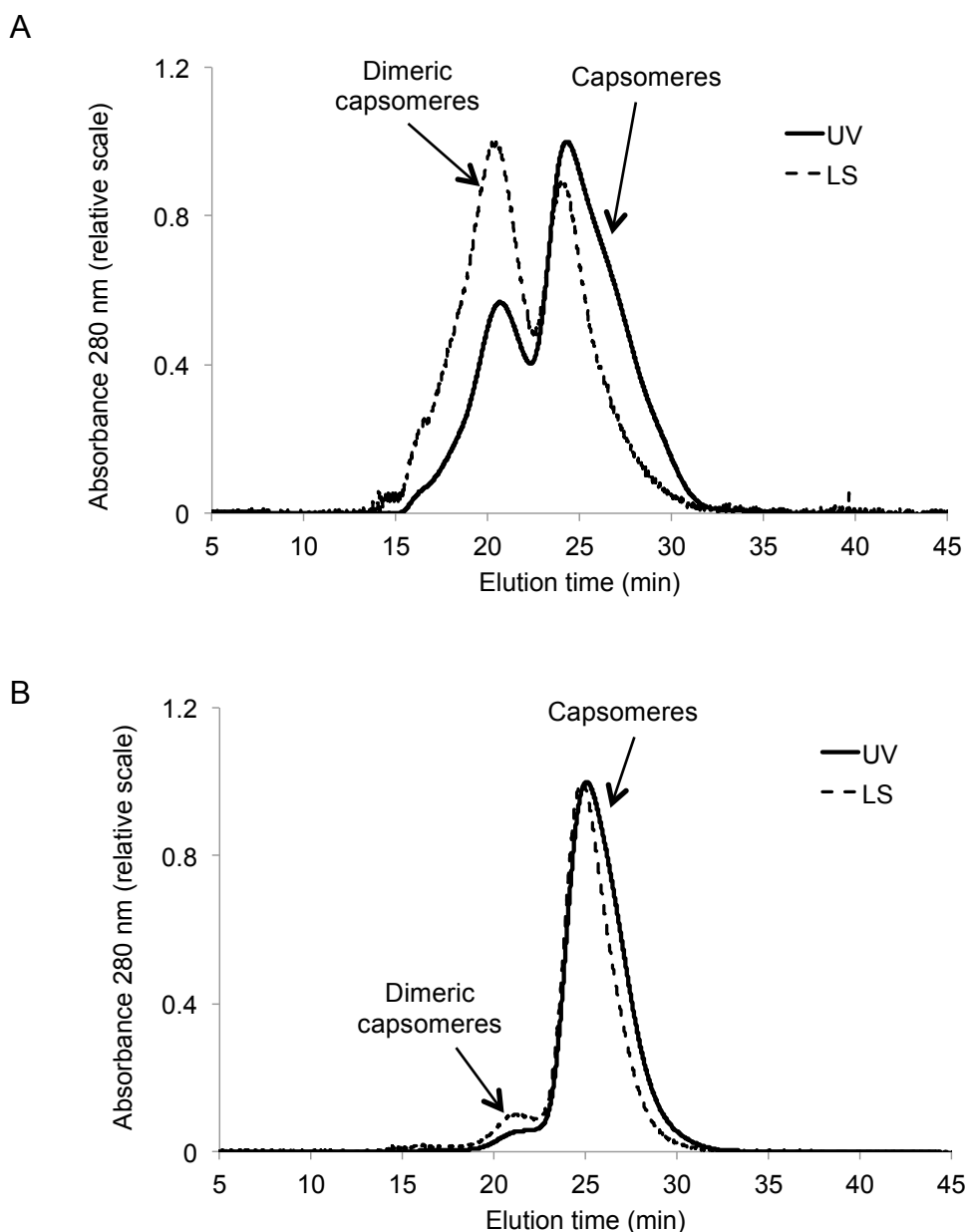


Figure 6-19. Analysis of quaternary structure of HIC-purified modular VP1 capsomeres using SEC-MALS. (A) TEVP-GCN4-H190-GCN4, and (B) TEVP-H190. Continuous lines represent UV absorbance at 280 nm (UV) and dashed lines represent light scattering (LS).

HIC-purified modular VP1 capsomeres were then concentrated using saturated $(\text{NH}_4)_2\text{SO}_4$ at a concentration above 2 mg ml^{-1} to screen for the optimum conditions for crystallisation. These samples were sent to the Australian Synchrotron (Melbourne, Australia) for crystallisation screening to be performed in collaboration with the Beamline team of Macromolecular crystallography (Protein crystallography), Dr Tom Caradoc-Davis and Dr Santosh Panjikar. However, this work could not be completed within the timescale of this thesis; thus, it is considered to be future work and is discussed in Chapter 7.

6.4 Conclusions

This chapter accomplished Objective 4 by (i) demonstrating a mild and satisfactory process for the production of high-purity of modular VP1 capsomeres, and (ii) sending the purified proteins for screening of optimal conditions for crystallisation. Three modular constructs were selected, of which the molecular structures of the modularised H190 in these three constructs are yet to be resolved. VP1 of these modular constructs comprised: (i) one copy of H190, (ii) two copies of H190 tandem repeats, and (iii) one H190 flanked by helix promoter elements. The three selected modular VP1 were also designed to lack the 31 N-terminal residues and 63 C-terminal residues.

The mild and satisfactory process by which to achieve high-purity modular VP1 capsomeres was developed based on the previously reported processes for the expression and purification of wt-VP1 (wt-VP1 residues 1-384) (Chuan et al. 2008, Lipin et al. 2008, Middelberg et al. 2011). Using the previously reported processes, the three modular VP1 were expressed as a GST-tagged protein. The GST tag was then removed using thrombin, and modular VP1 capsomeres were separated from soluble aggregates and GST tag using gel filtration. This chapter shows that these processes were unable to produce high-purity modular VP1 capsomeres, which is mandatory to obtain a high quality crystal. Two obstacles were faced, i.e. (i) ineffective GST removal by thrombin, and (ii) a lower purity of purified modular VP1 capsomeres due to the presence of undigested GST-tagged modular VP1, GST tag, and 70-kDa contaminant.

These obstacles were overcome by performing three key modifications on the previously reported processes. The first modification was the use of a protease of a smaller molecular weight than thrombin to achieve efficient enzymatic removal of the GST tag. In this chapter, tobacco etch virus protease was used to replace thrombin. TEVp was shown to remove GST from modular VP1 more effectively, and resulted in proteins with a higher homogeneity. The second modification was the use of serially connected GST-affinity chromatography columns, comprising GSTrap HP and GSTrap FF columns. This modification was to remove undigested GST-tagged modular VP1 and GST tag from SEC-purified modular VP1 capsomeres using a flow-through purification strategy. The third modification was the use of hydrophobic interaction chromatography (HIC) to remove the 70-kDa contaminant from GST-purified modular VP1 capsomeres. The contaminant was suggested to be DnaK protein from *E. coli*.

The developed purification process for the production of high-purity modular VP1 capsomeres utilised chromatographic systems, of which the media were chemically stable and commercially available. No invasive methods were used during the production

process. This chapter shows that the process could easily be scaled up. Furthermore, the production process was applicable for three modular constructs, indicating that the process could potentially be used to produce other modular VP1 capsomeres of high-purity.

Chapter 7. Conclusions and future work

7.1 Summary of research findings

Virus-like particle (VLP) technology based on a microbial platform is a promising alternative to vaccine production from live viruses in chicken eggs and cell culture. A VLP is a macromolecular assembly of viral capsid proteins. It is tolerant to insertions of peptide antigen elements from foreign pathogens, yielding modular VLPs. Such insertions can be performed via genetic recombination technology. Additionally, the absence of genomic material makes VLPs a safe vaccine for humans. The Centre for Biomolecular Engineering has extensively studied VLPs from murine polyomavirus (MuPyV) VP1 protein, which is produced as a Glutathione S-transferase (GST)-tagged protein using an *E. coli* expression system. VLPs are assembled *in vitro* from purified subunits, yielding pure homogenous VLPs. The microbial platform has been optimised for scaled-up production at a gram-per-litre level. Therefore, it allows for fast vaccine production to respond rapidly in a case of a nascent pandemic.

In this thesis, MuPyV VP1 VLPs are exploited to present antigen modules containing a peptide antigen element from influenza A virus, i.e. helix 190 (H190). H190 is a hypervariable helical element that borders the highly conserved receptor-binding site in the globular domain of the haemagglutinin (HA) protein. It also overlaps an antigenic site that induces B-cell responses. Many studies on VLPs presume that when a peptide antigen is removed from the context of the intact protein and modularised on an unrelated VLP it may not be able to assume its native conformation. However, the native conformation of a peptide antigen is important for the induction of high-quality antibodies. In this thesis, antibody quality is defined as the ability of antibodies to recognise the intact proteins. Presentation of a peptide antigen on a modular VLP is not yet fully understood and the impacts of antigen display strategy approaches on the quality of the induced antibodies are not understood. This thesis aims to understand the presentation of H190 on a modular MuPyV VP1 VLP necessary for the induction of high-quality antibodies.

The studies discussed in this thesis were designed to answer four major unknowns in the presentation of H190 on a modular VLP:

1. *Effect of display strategy on the quality of antibodies*

The experimental work compared two display strategies to modularise peptide antigen H190 from the influenza virus strain A/California/07/2009. These strategies are: (i) the use

of helix promoter elements to flank one H190 element, and (ii) tandem repeat arraying of two copies of H190. The work highlighted that the two display strategies resulted in different antibody qualities even though their immunogenicities were similar. Modularisation of H190 using the tandem repeat display strategy induced higher quality antibodies than modularisation using flanking helix promoter elements (Chapter 3).

2. Improvements to the tandem repeat display strategy for induction of higher quality antibodies

The experimental work investigated what improvements to the tandem repeat display strategy can be made to induce higher quality antibodies. Two approaches were explored, i.e. (i) increasing the number of H190 tandem repeats, and (ii) the use of adjuvant. The work highlighted that the quality of antibodies induced by two copies of H190 tandem repeats was optimised; therefore, further improvements to the display strategy were not required (Chapter 4).

3. Challenging the tandem repeat display strategy with a hydrophobic H190 variant

The applicability of the tandem repeat display strategy for a hydrophobic H190 variant from influenza virus strain A/Victoria/210/2009 was investigated (Chapter 5). The H190 is hydrophobic, and the hydrophobic residues are centred in the middle of its amino acid sequence. The experimental work revealed that following GST tag removal using thrombin, modular VP1 bearing two copies of the hydrophobic H190 tandem repeats had an increased tendency to form soluble aggregates; thus, modular VP1 capsomeres could not be obtained. The modular VP1 also had an increased proportion of a secondary digestion product. In an attempt to minimise the tendency of modular VP1 to form soluble aggregates, charged residues were utilised to flank two copies of the H190. Simultaneously, to minimise the proportion of the secondary digestion product, two proteases, i.e. thrombin and TEVp, were compared. This work opened the way to further studies based around the novel finding that the inclusion of anionic repeat residues within an inserted antigen module could allow the recovery of modular VP1 capsomeres, even for those elements having high hydrophobicity.

4. Structural determination of H190 element on a modular VLP

The first steps toward obtaining the crystal structures of H190 modularised on MuPyV VP1 using three different display strategies were carried out (Chapter 6). The three modular VP1 proteins were designed to lack 31 N-terminal residues and 63 C-terminal residues.

The purification process for wt-VP1 (Chuan et al. 2008, Lipin et al. 2008, Middelberg et al. 2011) was shown to encounter two challenges when it was used to produce highly pure preparations of the three modular VP1 proteins. These challenges were: (i) the ineffectiveness of thrombin in removing the GST tag from the modular VP1, and (ii) the presence of undigested GST-tagged modular VP1, GST tag, and 70-kDa contaminant in the modular VP1 capsomeres after purification using gel filtration. To address these challenges, two modifications were made to the purification process for wt-VP1, i.e. (i) replacement of the thrombin protease with TEVp, (ii) additional purification of modular VP1 capsomeres using tandem GST affinity chromatography columns followed by hydrophobic interaction chromatography. With these modifications, the purification process produced high-purity modular VP1 capsomeres suitable for crystallisation.

The following sections summarise key findings obtained from the experimental work in this thesis:

7.1.1 Effects of the display strategy on the quality of the antibodies

This experimental work (Chapter 3) was conducted to investigate the effects of the display strategy on the quality of the antibodies. Two modular VLP constructs were generated; i.e.: (i) VP1-GCN4-H190-GCN4, and (ii) VP1-H190-H190. These two constructs represented modularisation of H190 from A/California/07/2009 influenza on MuPyV VP1 VLPs using two display strategies. These strategies were: (i) the use of flanking helix promoter elements, and (ii) a tandem repeat display. In the first display strategy (modular construct VP1-GCN4-H190-GCN4), one copy of H190 was flanked by helix promoter elements derived from the GCN4 protein. The flanked H190 was connected to the VP1 protein via GSGS spacer elements comprised of amino acid residues Gly-Ser-Gly-Ser. Glycine and serine have often been used as the major components of linkers used to connect two domains and provide flexibility to peptide epitopes that are genetically inserted into a viral capsid protein (Huston et al. 1988). In the second strategy (modular construct VP1-H190-H190), two copies of H190 tandem repeats were displayed. Tandem repeat arraying of peptide epitopes has been proven to effectively present the peptide epitope folded into its native conformation (Fontenot et al. 1995, Fontenot et al. 1993).

Analysis of the modular VLPs using asymmetric flow-field flow fractionation (AF4) and transmission electron microscopy (TEM) showed that both modular VLPs were morphologically indistinguishable and similar to wild-type (wt)-VP1 VLPs. Additionally, the proportions of soluble aggregates for both modular constructs and wt-VP1 were negligible.

This validation of the quaternary structures for both modular VLPs underlined that any difference in results between the two modular VLPs from the *in vivo* study were due to the difference in the display strategy rather than differences in the morphology of the VLPs or the presence of soluble aggregates.

The immunogenicity of modularised H190 and the quality of the induced antibodies were evaluated *in vivo* using mice. Groups of mice were immunised with wt-VP1, VP1-GCN4-H190-GCN4, and VP1-H190-H190 VLPs. Sera from the mice then were analysed against various antigens in dot blot immunoassay, as well as indirect and competitive ELISA. The immunogenicity of the modularised H190 was analysed by testing the mouse sera against peptide H190. Meanwhile, antibody quality was assessed by testing the mouse sera against recombinant HA1 protein produced in insect cells. The *in vivo* study showed three important results.

First, both modular VLPs were equally able to induce high titre IgGs specific to the H190 peptide ($>10^5$). This result was in agreement with other studies showing that modularisation of a peptide antigen on a VLP could promote the immunogenicity of a peptide antigen which otherwise may not be immunogenic (Neiryneck et al. 1999). Additionally, the similar H190 sequence-specific IgG titres between the two modular VLPs confirmed that the display strategy of H190 did not affect its immunogenicity. Second, both modular VLPs were able to induce IgGs recognising recombinant HA1 protein produced from insect cells. VP1-H190-H190 VLPs induced IgGs recognising the recombinant HA1 protein at a titre of higher than 10^4 . In contrast, VP1-GCN4-H190-GCN4 VLPs induced HA1-specific IgGs at a titre of less than 10^3 . This difference in the titre of the IgGs recognising the HA1 protein suggested that the two modular VLPs induced antibodies of different quality. Furthermore, glycoprotein staining confirmed that the HA1 protein carried N-glycans. The ability of the induced IgGs to bind to the HA1 protein was not abolished when the N-glycans were removed from the protein by treatment with PNGase F enzyme. This result showed that the binding of IgGs induced by modular VLPs was not affected by the glycosylation of the native protein. Third, VP1-H190-H190 VLPs were shown able to induce IgGs recognising the trivalent influenza vaccine, Fluvax[®], containing split influenza virions. In contrast, binding of IgGs induced by VP1-GCN4-H190-GCN4 VLPs to Fluvax[®] was similar to the binding induced by wt-VP1 VLPs. This result confirmed the finding that the two modular VLPs induced antibodies of different quality. Since the display strategy of H190 affected neither the morphology of the modular VLPs nor the immunogenicity of the H190, the *in vivo* study suggested that the antibody quality difference was due to the difference in the display strategy.

These results showed that the key finding of the experimental work in Chapter 3 was that the display strategy used to modularise H190 on the MuPyV VP1 protein did not affect its immunogenicity, but did affect the quality of antibodies induced against the modular VLPs.

The study by Alvarez et al. showed that the quality of the antibodies reflects the protective efficacy of the antibodies (Alvarez et al. 2010). A higher antibody quality indicates a greater potential protection efficacy. Therefore, the key finding in this experimental work leads to the question: What improvements can be made to the tandem repeat display strategy to induce antibodies of a higher quality? The following experimental work (Chapter 4) was performed to address this question.

7.1.2 Improvements in the display strategy to induce higher quality antibodies

Motivated by the findings in Chapter 3, the experimental work in Chapter 4 compared two approaches in an attempt to improve the antibody quality obtained from two copies of H190 tandem repeats (VP1-H190-H190 VLPs). These approaches were: (i) the use of AdvaxTM-1 as an adjuvant for VP1-H190-H190 VLPs, and (ii) increasing the copy number of H190 tandem repeats from one to five copies. The experimental work in this chapter indicated two major findings, which are related to (i) the solubility and assembly capability of modular VP1, and (ii) the antibody quality induced by modular VLPs. Modular VLPs bearing five copies could not be obtained; thus, this modular construct was not investigated in the *in vivo* study.

The results in this chapter showed that increasing the number of H190 tandem repeats from one to five copies (i) reduced the capacity of the modular VP1 to self-assemble to form VLPs, and (ii) increased the tendency of the modular VP1 to precipitate in Assembly Buffer 1, which was optimised for wt-VP1. When assembled in Assembly Buffer 1, modular VP1 bearing four and five copies of H190 precipitated more than modular VP1 bearing three copies of H190. Furthermore, while modular VP1 bearing three copies of H190 assembled into VLPs under the established assembly conditions, only a small proportion of modular VLPs bearing four copies of H190 was obtained and modular VLPs bearing five copies of H190 could not be obtained at all. In order to increase the solubility and assembly capability of modular VP1 bearing four or five copies of H190, two parameters of protein-protein interactions were manipulated, i.e. (i) the net surface charge of the proteins (zeta potential), and (ii) specific ion effects. Manipulation of the zeta potential for each modular VP1 was performed by modifying the pH of Assembly Buffer 1 and hence the ionisation state of residues on the capsomere surface. Meanwhile,

variations in specific ion effects were achieved by changing both the type and concentration of salts in Assembly Buffer 1. The results in Chapter 4 showed that pH modification greatly affected the assembly capability of modular VP1 bearing four copies of H190, but affected the solubility of the modular VP1 only slightly. In contrast, changes in the pH of Assembly Buffer 1 did not affect the assembly capability or solubility of modular VP1 bearing five copies of H190. Furthermore, the results in this chapter showed that both the type and concentration of salt ions in Assembly Buffer 1 affected the assembly capability and solubility of modular VP1 bearing four copies of H190 significantly. These results showed that in Assembly Buffer 1 the influence of specific ion effects on the assembly capability and solubility of modular VP1 was greater than the influence of pH, as expected from literature on protein-protein interactions in a concentrated salt solution. Nevertheless, the result showing how pH changes greatly affected the assembly capability of modular VP1 bearing four copies of H190 was an exception. This indicated that the protein-protein interactions are a complex phenomenon involving electrostatic forces and changes in the amino acid sequence of the modular VP1, which could be due to the amino acid sequence of the modularised H190 and/or a possible structural perturbation.

Mouse immunisation with modular VLPs bearing two copies of H190 (VP1-H190-H190 VLPs) with AdvaxTM-1 induced a slightly higher titre of H190 sequence-specific IgG than non-adjuvanted VP1-H190-H190 VLPs. This increase in immunogenicity was also seen in the result from competitive ELISA assay. In the competitive ELISA, peptide H190 in solution at about 240 μ M inhibited approximately 80% of IgG binding to immobilised peptide H190 for non-adjuvanted VLPs. However, the same concentration of peptide H190 in solution inhibited only about 60% of IgG binding for adjuvanted VLPs. Nevertheless, statistical analysis showed that the difference in IgG titres was not significant ($p=0.602$), suggesting the modular VLPs bearing two copies of H190 were self-adjuvanting and able to induce optimum immune responses without the need for adjuvant. Furthermore, results from the *in vivo* study showed that adjuvanted VLPs induced HA1-specific IgG at a titre slightly less than 10^5 , while the HA1-specific IgG titre from non-adjuvanted VLPs was slightly higher than 10^4 . Statistical analysis showed that the titre difference had an adjusted P value of 0.0552, indicating that the increase was marginal, and the significance of the increase was dependent on the selected significance level. If the family-wise significance level was determined at 5%, the increase was not statistically significant. In contrast, at a 10% family-wise significance level, the increase can be concluded to be statistically significant. These results suggested the use of AdvaxTM-1 increased the

immunogenicity of H190 and consequently the quality of antibodies, although the increase was statistically not significant.

A similar observation was obtained from the *in vivo* study of modular VLPs with increasing copy numbers of H190 tandem repeats. The results in this chapter showed that modular VLPs bearing one to four copies of H190 were equally effective at inducing high titres of H190 sequence-specific IgGs ($> 10^4$). These results suggested that increasing the copy number of the H190 tandem repeats did not increase the antigen element's immunogenicity. In contrast, the *in vivo* study showed that an increase in the copy number of H190 tandem repeats did affect the quality of the induced antibodies. Modular VLPs bearing one copy of H190 induced the lowest HA1-specific IgG titre, which was slightly less than 10^3 , and the IgGs were shown to bind non-specifically to HA1 rather than via the H190 element. To the knowledge of the author, this is the first reported scientific attempt to demonstrate the presumption in modular VLP studies that it is nearly impossible for a peptide element to assume its conformational integrity when it is presented on a modular VLP (Jennings and Bachmann 2007, Roldão et al. 2010, Tissot et al. 2010). The HA1-specific IgG titres increased with increasing H190 copy number from one to three copies, with the highest and most consistent HA1-specific IgG titre achieved by modular VLPs bearing three copies of H190 tandem repeats. Modular VLPs bearing four copies of H190 induced a slightly lower HA1-specific IgG titre than modular VLPs bearing three copies of H190. Nevertheless, statistical analysis showed that the IgG titre differences for modular VLPs bearing one to four copies of H190 were not statistically significant. The presented results suggested that, without taking the statistical analysis into account, the highest antibody quality could be achieved by increasing the number of H190 tandem repeats to three copies. However, when the statistical analysis was taken into account, the results suggested that this increase in the copy number of H190 did not increase the immunogenicity of H190 or the antibody quality.

The key findings from the presented results were that the use of AdvaxTM-1 and increasing the number of H190 tandem repeats did not increase the immunogenicity of H190, and the impacts of these approaches on antibody quality were rather dependent on whether the statistical analysis was taken into account. Without considering the results from the statistical analysis, both approaches increased the quality of the resulting antibodies, although the increases were marginal. Taking into account the statistical analysis, the results in this chapter suggested that modular VLPs bearing two copies of H190 was sufficient to induce a high titre of H190-sequence specific IgGs and HA1-

specific IgGs. Therefore, further improvements to the tandem repeat display strategy to increase the quality of the antibodies were not necessary.

7.1.3 Challenging the tandem repeat display strategy with a hydrophobic H190 variant

The findings in Chapter 4 showing that modular VLPs bearing two copies of H190 were sufficient to induce a high titre of H190 sequence-specific and HA1-specific IgGs using H190 from the influenza strain *A/California/07/2009*. However, H190 is a hypervariable region, although its secondary structure is maintained. Of particular interest as an extreme test case is the H190 from influenza strain *A/Victoria/210/2009*, which is hydrophobic, and its hydrophobic stretches are centred in the middle of the H190 amino acid sequence. Therefore, the experimental work included a challenge of the tandem repeat display using this hydrophobic H190 variant (Chapter 5). The findings in Chapter 5 answered two questions: (i) To what extent does modularisation of H190 containing a hydrophobic stretch affect the properties of the modular VLPs? (ii) How can these effects be minimised?

The work showed that, under the previously established expression and purification conditions for wt-VP1, GST-tagged modular VP1 bearing two copies of H190 containing a hydrophobic stretch (modular construct VP1-H3-H190-H190) had a considerably lower solubility than GST-tagged wt-VP1. Furthermore, following GST tag removal using thrombin, two difficulties were observed. First, the modular VP1 formed soluble aggregates, and consequently modular VP1 capsomeres could not be obtained. Second, the VP1-H3-H190-H190 was of a lower quality than wt-VP1, assessed from the higher proportion of a secondary cleavage product known as VP1*. A study by Connors et al. (Connors et al. 2013) identified VP1* (approximately 37 kDa) in wt-VP1 as a secondary cleavage product by thrombin at Arg58. A molecular weight analysis suggested that the VP1* in the VP1-H3-H190-H190 may also have resulted from thrombin cleavage at Arg58 of VP1-H3-H190-H190.

The results in the chapter showed that the solubility of GST-tagged VP1-H3-H190-H190 was improved by lowering the protein expression temperature from 26°C to 12°C. Following expression at 12°C, GST-tagged VP1-H3-H190-H190 had a similar solubility as GST-tagged wt-VP1 in a buffer that was previously optimised for wt-VP1. This result indicated that the solubility of GST-tagged VP1-H3-H190-H190 was a function of the protein expression temperature.

To minimise the formation of soluble aggregates, an approach utilising glutamic acid was explored. In this approach, two copies of the hydrophobic H190 were flanked by ionic elements comprising four glutamic acid residues (modular construct VP1-H3-H190-H190-4E). The results from the experimental work showed that the solubility of GST-tagged VP1-H3-H190-H190-4E was higher than that of GST-tagged VP1-H3-H190-H190. However, the improvement in solubility was only observed when the salt concentration was lowered from 200 mM to 50 mM NaCl. In addition, the use of the flanking ionic elements slightly reduced the ratio of VP1* to VP1. Unlike the increase in the solubility of the GST-tagged modular VP1, the reduction in the ratio of VP1* to VP1 could be observed at both 200 mM and 50 mM NaCl but the reduction was greater at the lower salt concentration. Furthermore, at both 200 mM and 50 mM NaCl the flanking ionic elements reduced the formation of soluble aggregates, therefore allowing isolation of modular VP1 capsomeres. The key finding of these results was that the use of the glutamic acid as ionic flanking elements was necessary in order to obtain modular capsomeres of VP1 bearing two copies of hydrophobic H190.

In an attempt to increase the quality of the modular VP1 after GST tag removal, two proteases were compared: (i) thrombin, and (ii) tobacco etch virus protease (TEVp), which has a higher specificity than thrombin (Connors et al. 2013). The results in Chapter 5 showed that GST-tagged modular VP1 bearing a thrombin cleavage site (modular constructs VP1-H3-H190-H190 and VP1-H3-H190-H190-4E) had slightly higher solubility than GST-tagged modular VP1 bearing a TEVp cleavage site (modular constructs TEVP-VP1-H3-H190-H190 and TEVP-VP1-H3-H190-H190-4E). This reduced solubility was shown to be a construct-dependent behaviour, since mutation from the thrombin to the TEVp cleavage site did not affect the solubility of GST-tagged wt-VP1 and modular VP1 bearing two copies of H190 from A/California/07/2009 influenza (modular construct VP1-H190-H190, reported in Chapters 3 and 4). Furthermore, following treatment using TEVp, the secondary cleavage product VP1* was not observed for both TEVP-wt-VP1 and TEVP-VP1-H3-H190-H190. However, this was not the case for TEVP-VP1-H3-H190-H190-4E. A secondary cleavage product was observed at approximately 35 kDa for the modular construct. The experimental work suggested that this 35-kDa fragment was a result of non-specific interactions between TEVp and the modular VP1 that were exacerbated by the increase in the ratio of TEVp to VP1 protein. In addition, the salt concentration of the buffer was shown to slightly affect the non-specific interactions. Nevertheless, the presence of the 35-kDa fragment did not affect isolation of the modular VP1 as capsomeres. The results in the chapter showed that capsomeres could be obtained for TEVP-VP1-H3-H190-

H190-4E but not TEVP-VP1-H3-H190-H190. The key finding of these results was that the quality of modular VP1 after GST cleavage could be improved by replacing thrombin with TEVp. Nevertheless, the improved homogeneity of the modular protein was not followed by a reduction in the formation of soluble aggregates or an increase in the likelihood of obtaining modular VP1 capsomeres. This finding highlighted that the aggregation of modular VP1 was not due to heterogeneity of the modular VP1 after GST removal, but rather possibly to the hydrophobic stretch within the H190.

This experimental work showed that a tandem repeat display comprising two copies of H190 was applicable for a hydrophobic H190 from A/Victoria/210/2009 influenza. Nevertheless, the modularisation of two copies of such a hydrophobic H190 into MuPyV VP1 affected the biophysical properties of the VP1 protein and therefore modifications were required to minimise these effects.

7.1.4 Structural determination of the H190 element on modular VLPs

Chapters 3 and 4 of this thesis have shown that modularisation of H190 on MuPyV VP1 using various display strategies resulted in the induction of antibodies of diverse quality in mice. The results in these chapters may suggest that antibodies are raised against modularised H190 elements with different structures. A possible approach to verifying this suggestion is by determining the structure of modularised H190 using X-ray crystallography.

Three modular constructs were selected: (i) VP1-H190-H190, (ii) VP1-GCN4-H190-H190, and (iii) VP1-H190. The H190 elements in these constructs were yet to be resolved using X-ray crystallography. In construct VP1-H190-H190, two copies of H190 are arrayed as a tandem repeat. In contrast, in construct VP1-GCN4-H190-GCN4, one copy of H190 is modularised by exploiting the use of GCN4 helix promoter elements and GSGS spacer elements. In modular construct VP1-H190, one copy of H190 is modularised without incorporating any structural hypothesis. In order to obtain crystal structures with high resolution, the first 31 residues were removed from the N-terminus and 63 residues were removed from the C-terminus.

This chapter (Chapter 6) reports the first steps toward obtaining crystal structure of the three selected modular constructs. These initial steps were (i) production of modular VP1 capsomeres lacking 31 N-terminal residues and 63 C-terminal residues with high purity and (ii) determination of the optimum crystallisation conditions for modular VP1 capsomeres.

The production of highly pure assembly-incompetent modular VP1 relied on previously reported processes to produce high yield wt-VP1 in *E. coli* (Chuan et al. 2008, Lipin et al. 2008, Middelberg et al. 2011). Modular VP1 for the three selected modular constructs were expressed as GST-tagged modular VP1 in *E. coli*. The GST tag was then removed enzymatically from the modular VP1. In this chapter, two proteases were compared: (i) thrombin, and (ii) TEVp. TEVp has a molecular weight of about 26 kDa, which is smaller than the molecular weight of thrombin (about 37 kDa). Its smaller molecular weight is advantageous because the modular VP1 in this chapter lacked 31 residues from the N-terminus, which meant a significant space between the GST tag and VP1 was removed. The results showed that thrombin was ineffective in removing GST from modular VP1, leaving a considerable proportion of undigested GST-tagged modular VP1. In addition, GST tag removal using thrombin resulted in the generation of secondary cleavage products, which are undesirable for crystallisation. In contrast, GST tag removal using TEVp was effective and the digestion products were homogeneous. The optimum GST tag removal using TEVp was achieved at a ratio of 1:5 (TEVp mass/protein mass) at 30°C for 4 h.

Following GST removal, modular VP1 capsomeres were then separated from soluble aggregates and GST tags using size exclusion chromatography (SEC). SDS-PAGE analysis showed that the SEC-purified modular VP1 capsomeres contained three major contaminants. These contaminants were: (i) a 60-kDa contaminant, which was verified to be undigested GST-tagged modular VP1, (ii) GST tags, and (iii) a 70-kDa contaminant, which was suggested to be the *E. coli* protein DnaK according to the study by Fan and Middelberg (Fan and Middelberg 2010). The presence of GST tags as a contaminant was correlated with the removal of 31 residues from the N-terminus and 63 residues from the C-terminus. Removing these residues caused a shift in the elution time of modular VP1 capsomeres in SEC. This shift in elution time consequently resulted in a poor separation between modular VP1 capsomeres and cleaved GST tags, and contamination of modular VP1 capsomeres with GST tags. The results in this chapter showed that both undigested GST-tagged modular VP1 and GST tags could be removed effectively and simultaneously from SEC-purified modular VP1 capsomeres using tandem GST affinity columns. However, the 70-kDa contaminant was still present.

To remove the 70-kDa contaminant, following the purification step using tandem GST-affinity columns, modular VP1 capsomeres were subjected to a further step of purification, which was hydrophobic interaction chromatography (HIC). The result showed that HIC effectively removed the 70-kDa contaminant from modular VP1 capsomeres for

the three selected constructs. Following purification using HIC, the modular VP1 had a high purity and the presence of the 70-kDa contaminant was no longer detected by SDS-PAGE analysis. In this HIC purification, the results showed that the chromatogram profiles for the three modular constructs were different. The profiles for modular VP1 bearing one or two copies of H190 were similar, but distinct from the profile for modular VP1 bearing H190 flanked with helix promoter elements. In addition, although the binding profiles between modular VP1 bearing one and two copies were similar, HIC was less effective in removing the 70-kDa contaminant from modular VP1 bearing two copies than from modular VP1 bearing just one copy of H190.

Using these processes, modular VP1 bearing one copy of H190 or H190 flanked by helix promoter elements were produced in a large scale. These modular VP1 capsomeres with high purity were then sent to a collaborator in Synchrotron Australia for screening of crystallisation conditions.

The experimental work showed the production of high purity modular VP1 capsomeres utilising the developed chromatographic process. The media were chemically stable and commercially available. The chromatography-based process could be easily scaled up and no invasive method was used during the production process. Furthermore, the production process was applicable for two modular constructs, indicating that the process could be applied to produce highly pure modular VP1 capsomeres from other constructs.

7.2 Future work

This thesis has provided the first steps toward understanding how to provide structurally authentic presentation of helical peptide antigens on a VLP, using VP1 from MuPyV displaying H190 from influenza. H190 was presented using display strategies that encouraged the peptide antigen element to assume its native conformation. The effects of the display strategy on the quality of the antibodies induced were evaluated. Additionally, approaches to improving the quality of the antibodies were explored and, subsequently, the preferred display strategy was determined. Furthermore, the applicability of the preferred display strategy for H190 bearing a hydrophobic stretch was also investigated. Finally, the first steps toward obtaining the atomic structure of the presented H190 have been performed.

Nevertheless, further investigations on the following points are required in order to better understand the presentation of a peptide antigen element on MuPyV VP1 VLPs.

1. *In vivo* studies in this thesis demonstrated that the display strategy for presenting H190 on modular VLPs affected the quality of the induced antibodies, and further suggested that the display strategy affected the structure of the modularised H190. A promising technology to determine the atomic structure of modularised H190 is X-ray crystallography. This thesis initiated the first steps toward resolving atomic structures of H190 on modular VP1 capsomeres using X-ray crystallography. However, at the time this thesis was written, crystals of modular VP1 capsomeres had not yet been obtained, and thus the atomic structures of H190 on modular capsomeres have not been resolved. Therefore, further investigations to resolve the atomic structure of modularised H190 are required. In addition, the purification processes reported in Chapter 6 could be further improved and optimised to increase capsomere yield and purity, and hence increase the chances of obtaining high quality crystals. This thesis has shown the production of high purity proteins from for constructs VP1-H190-H190, VP1-GCN4-H190-GCN4 and VP1-H190. Based on the findings in Chapters 3 and 4, it may be necessary to produce high-purity proteins for modular constructs VP1-H190-3x, VP1-H190-4x, and VP1-H190-5x.
2. Chapter 3 in this thesis showed that modular VLPs presenting two copies of H190 (VP1-H190-H190 VLPs) were able to induce high quality antibodies. Furthermore, Chapter 4 in this thesis has shown that the quality of antibodies induced by the modular VLPs could be marginally increased by (i) increasing the number of H190 tandem repeats to three copies, and (ii) using AdvaxTM-1 as an adjuvant for the modular VLPs. However, the protective efficacy of the antibody has not been determined in this study. If the antibodies do not provide protection, then the antibody quality may need further improvement, and thus further examination of the use of other types of adjuvants may be necessary. Various adjuvants have been studied in VLP-based vaccines, such as AAHS (aluminium hydroxylsulfate), and AS04 (combination of AAHS and monophosphoryl lipid or MPL). Determination of whether these adjuvants can be used with VP1-H190-H190 VLPs and whether these adjuvants can increase the quality of the resulting antibodies may be useful and necessary.

In regards to the protective efficacy studies, the studies could be done by detecting the presence of neutralising antibodies in the sera from immunised mice using

various neutralisation assays. Investigation and optimisation studies on the appropriate assays to best assess the protective efficacy of the antibodies generated are also required.

3. Chapter 5 in this thesis showed that flanking ionic elements comprising four glutamic acid residues reduced the formation of soluble aggregates for modular VP1 bearing two copies of hydrophobic H190, and consequently reduced the possibility of obtaining modular VP1 capsomeres. The yield of modular VP1 capsomeres, however, still needs improvement. Such improvements can be made by exploring two types of approaches. The first approach includes optimisations of: (i) the type of ionic residues, such as arginine, histidine, and lysine, (ii) the number of ionic residues, and (iii) the position of the ionic elements. The second approach includes: (i) screening to determine the optimum buffer composition for the modular VP1, or (ii) point mutations within the hydrophobic stretch. Furthermore, once a sufficient yield is achieved, the effects of these optimised approaches on the molecular structure of modularised H190 and the antibody quality will need to be assessed.
4. This thesis has shown that the display strategy comprising two copies of H190 tandem repeats is preferable for modularisation of the helical peptide antigen element on MuPyV VP1 VLPs. This strategy was also shown to be applicable for modularisation of a hydrophobic H190 variant, in which flanking ionic elements were included to allow isolation of modular VP1 capsomeres. Therefore, the display strategy can be explored further for modularisation of a range of helical peptide antigen elements. In addition, exploration of the strategy for (i) antigens with a similar secondary structure, (ii) antigens with a different secondary structure, such as β -sheet antigens, or (iii) non-structural antigens can also be conducted. Further optimisation of the display strategy may be required because it is expected that different antigens and antigens with different secondary structures will behave differently from the H190 modelled in this thesis.
5. This thesis has pursued a rational design approach to the modularisation of a peptide antigen element on MuPyV VP1. The designs were examined *in vivo* and the results from the studies were then used to improve the designs. However, the target antigens from influenza viruses that are able to induce neutralising antibodies

undergo continuous hypervariability. Similar phenomena have also been observed for other pathogens. Therefore, processes that allow for a rational design without the need for *in vivo* studies are increasingly desirable. Such processes can be developed by combining the advantages of resolved atomic structures and molecular modelling. The atomic structures of antigens presented on modularised VLPs can be resolved using various technologies such as NMR and X-ray crystallography. The resolved structures can then be used to verify parameters that are used in molecular modelling to yield a highly accurate molecular model. The verified parameters can then be used to model various antigen designs.

7.3 Concluding thought

Vaccination has significantly reduced the burden of various infectious diseases. One advance in vaccination is the use of selected peptide antigens rather than whole pathogens. The peptide antigens targeted are those that are able to elicit protective immune responses, and these antigens are selected based on available structural information. Furthermore, the omission of other unnecessary proteins from the pathogens has been shown able to increase the safety of the vaccines. Nevertheless, peptide antigen-focusing vaccination has many challenges, including its poor immunogenicity. In addition, protective antigens are usually hypervariable, and change continuously in order to aid pathogens in escaping the immune system. This challenge in the immunogenicity of peptide-based vaccines has been answered by another advance in vaccination, which is the development in VLPs. VLPs are able to promote vigorous immune responses, even without adjuvant (Rivera-Hernandez et al. 2013), and can be formulated to be thermostable (Mohr et al. 2013). Additionally, VLPs have been used as carriers of peptide antigen elements, which can either be genetically inserted into viral capsid proteins or chemically conjugated onto pre-formed VLPs. Various VLP production systems have been studied, including production via bacteria that allows for rapid response vaccines. However, the challenge posed by the hypervariability of protective antigens has not yet been addressed.

Combining peptide antigen-focusing vaccination and VLPs, this thesis has shown the utilisation of VLPs from murine polyomavirus VP1 protein as a carrier for a hypervariable influenza peptide antigen, H190. The use of H190 as a model peptide antigen element shows that an alternative paradigm in vaccinations, targeting the use of hypervariable peptide antigen elements in place of broadly cross-protective peptide antigen elements, is possible. The use of hypervariable peptide antigen elements is empowered by fast and

rapid VLP production using a bacterial platform. The rapid vaccine engineering and production make it possible for new vaccines to be made available at a faster rate than the spread of the viruses and the hypervariability of the peptide antigen elements can undermine (Wibowo et al. 2014).

Furthermore, this thesis also contributed to answering one of the unknowns in the modularisation of a peptide antigen element on a VLP, i.e. the structural engineering of the modularised peptide antigen element. The structures of the modularised peptide antigen elements were shown to be rationally engineered via designing appropriate display strategies. Antigen display strategies have been shown to be important in order for high quality antibodies to be induced. In addition, antigen display strategies have been shown to be an effective tool to control the aggregation and solubility of modular VP1 displaying peptide antigen elements. Thus, it is suggested that the use of a rational design can enable, not only the generation of an effective and protective modular VLP-based vaccine candidate, but also control the quality of the vaccine candidate.

References

- Abe, Y, Takashita, E, Sugawara, K, Matsuzaki, Y, Muraki, Y and Hongo, S. 2004. Effect of the addition of oligosaccharides on the biological activities and antigenicity of influenza A/H3N2 virus hemagglutinin. *Journal of Virology* 78(18):9605-9611.
- Abedi, MR, Caponigro, G and Kamb, A. 1998. Green fluorescent protein as a scaffold for intracellular presentation of peptides. *Nucleic Acids Research* 26(2):623-630.
- Aitken, A and Learmonth, M. 1996. The protein protocols handbook. In: Walker, J M, editor. Protein determination by UV absorption: Springer. p. 3-6.
- Aleksaitė, E and Gedvilaitė, A. 2006. Generation of chimeric hamster polyomavirus VP1 virus-like particles harboring three tumor-associated antigens. *Biologija* (3):83-87.
- Alvarez, MM, López-Pacheco, F, Aguilar-Yañez, JM, Portillo-Lara, R, Mendoza-Ochoa, GI, García-Echauri, S, Freiden, P, Schultz-Cherry, S, Zertuche-Guerra, MI and Bulnes-Abundis, D. 2010. Specific recognition of influenza A/H1N1/2009 antibodies in human serum: a simple virus-free ELISA method. *PLoS ONE* 5(4):e10176.
- Ambühl, PM, Tissot, AC, Fulurija, A, Maurer, P, Nussberger, J, Sabat, R, Nief, V, Schellekens, C, Sladko, K and Roubicek, K. 2007. A vaccine for hypertension based on virus-like particles: preclinical efficacy and phase I safety and immunogenicity. *Journal of Hypertension* 25(1):63-72.
- Anderson, P, Peter, G, Johnston Jr, RB, Wetterlow, LH and Smith, DH. 1972. Immunization of humans with polyribophosphate, the capsular antigen of *Hemophilus influenzae*, type b. *Journal of Clinical Investigation* 51(1):39-44.
- Anggraeni, MR, Connors, NK, Wu, Y, Chuan, YP, Lua, LHL and Middelberg, APJ. 2013. Sensitivity of immune response quality to influenza helix 190 antigen structure displayed on a modular virus-like particle. *Vaccine* 31(40):4428-4435.
- Arakawa, T and Timasheff, SN. 1982. Preferential interactions of proteins with salts in concentrated solutions. *Biochemistry* 21(25):6545-6552.
- Arakawa, T. 1986. Thermodynamic analysis of the effect of concentrated salts on protein interaction with hydrophobic and polysaccharide columns. *Archives of Biochemistry and Biophysics* 248(1):101-105.
- Armstrong, AH, Chen, J, McKoy, AF and Hecht, MH. 2011. Mutations that replace aromatic side chains promote aggregation of the Alzheimer's A β peptide. *Biochemistry* 50(19):4058-4067.
- Arndt, K and Fink, GR. 1986. GCN4 protein, a positive transcription factor in yeast, binds general control promoters at all 5'TGACTC 3' sequences. *Proceedings of the National Academy of Sciences* 83(22):8516-8520.

- Arnold, K, Bordoli, L, Kopp, J and Schwede, T. 2006. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* 22(2):195-201.
- Austrian, R, Douglas, R, Schiffman, G, Coetzee, A, Koornhof, H, Hayden-Smith, S and Reid, R. 1976. Prevention of pneumococcal pneumonia by vaccination. *Transactions of the Association of American Physicians* 89:184-194.
- Averhoff, F, Mahoney, F, Coleman, P, Schatz, G, Hurwitz, E and Margolis, H. 1998. Immunogenicity of hepatitis B vaccines: Implications for persons at occupational risk of hepatitis B virus infection. *American Journal of Preventive Medicine* 15(1):1-8.
- Avery, OT and Goebel, WF. 1929. Chemo-immunological studies on conjugated carbohydrate-proteins: II. Immunological specificity of synthetic sugar-protein antigens. *The Journal of Experimental Medicine* 50(4):533-550.
- Bachmayer, H, Liehl, E and Schmidt, G. 1976. Preparation and properties of a novel influenza subunit vaccine. *Postgraduate Medical Journal* 52(608):360-367.
- Bagnoli, F, Baudner, B, Mishra, RP, Bartolini, E, Fiaschi, L, Mariotti, P, Nardi-Dei, V, Boucher, P and Rappuoli, R. 2011. Designing the next generation of vaccines for global public health. *Omics: A Journal of Integrative Biology* 15(9):545-566.
- Baldwin, RL. 1996. How Hofmeister ion interactions affect protein stability. *Biophysical Journal* 71(4):2056-2063.
- Baneyx, F. 1999. Recombinant protein expression in *Escherichia coli*. *Current Opinion in Biotechnology* 10(5):411-421.
- Baneyx, F and Mujacic, M. 2004. Recombinant protein folding and misfolding in *Escherichia coli*. *Nature Biotechnology* 22(11):1399-1407.
- Barbosa, T and Barral-Netto, M. 2013. Challenges in the research and development of new human vaccines. *Brazilian Journal of Medical and Biological Research* 46(2):103-108.
- Batista, FD and Neuberger, MS. 1998. Affinity dependence of the B cell response to antigen: a threshold, a ceiling, and the importance of off-rate. *Immunity* 8(6):751-759.
- Baumgarten, CM. 1981. A program for calculation of activity coefficients at selected concentrations and temperatures. *Computers in Biology and Medicine* 11(4):189-196.

- Berendsen, HJC, Postma, JPM, van Gunsteren, WF, DiNola, A and Haak, J. 1984. Molecular dynamics with coupling to an external bath. *Journal of Chemical Physics* 81(8):3684-3690.
- Berendsen, HJC, van der Spoel, D and van Drunen, R. 1995. GROMACS: a message-passing parallel molecular dynamics implementation. *Computer Physics Communications* 91(1-3):43-56.
- Bessa, J, Schmitz, N, Hinton, HJ, Schwarz, K, Jegerlehner, A and Bachmann, MF. 2008. Efficient induction of mucosal and systemic immune responses by virus-like particles administered intranasally: implications for vaccine design. *European Journal of Immunology* 38(1):114-126.
- Beterams, G, Böttcher, B and Nassal, M. 2000. Packaging of up to 240 subunits of a 17 kDa nuclease into the interior of recombinant hepatitis B virus capsids. *FEBS Letters* 481(2):169-176.
- Biasini, M, Bienert, S, Waterhouse, A, Arnold, K, Studer, G, Schmidt, T, Kiefer, F, Cassarino, TG, Bertoni, M and Bordoli, L. 2014. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Research* doi: 10.1093/nar/gku1340.
- Boisgérault, F, Morón, G and Leclerc, C. 2002. Virus-like particles: a new family of delivery systems. *Expert Review of Vaccines* 1(1):101-109.
- Bolognesi, B, Cohen, SI, Aran Terol, P, Esbjorner, EK, Giorgetti, S, Mossuto, MF, Natalello, A, Brorsson, AC, Knowles, TP and Dobson, CM. 2013. Single Point Mutations Induce a Switch in the Molecular Mechanism of the aggregation of the Alzheimer's Disease Associated A β_{42} peptide. *ACS Chemical Biology* 9(2):378-382.
- Bommakanti, G, Citron, MP, Hepler, RW, Callahan, C, Heidecker, GJ, Najjar, TA, Lu, XH, Joyce, JG, Shiver, JW, Casimiro, DR, ter Meulen, J, Liang, XP and Varadarajan, R. 2010. Design of an HA2-based *Escherichia coli* expressed influenza immunogen that protects mice from pathogenic challenge. *Proceedings of the National Academy of Sciences of the United States of America* 107(31):13701-13706.
- Bonanni, P and Santos, JI. 2011. Vaccine evolution. *Understanding Modern Vaccines: Perspectives in Vaccinology* 1(1):1-24.
- Boorstein, WR, Ziegelhoffer, T and Craig, EA. 1994. Molecular evolution of the HSP70 multigene family. *Journal of Molecular Evolution* 38(1):1-17.
- Bostrom, M, Williams, D and Ninham, B. 2001. Specific ion effects: why DLVO theory fails for biology and colloid systems. *Physical Review Letters* 87(16):168103-168103.

- Bouřa, E, Liebl, D, Špíšek, R, Frič, J, Marek, M, Štokrová, J, Holáň, V and Forstova, J. 2005. Polyomavirus EGFP-pseudocapsids: analysis of model particles for introduction of proteins and peptides into mammalian cells. *FEBS Letters* 579(29):6549-6558.
- Brady, MI and Furminger, I. 1976. A surface antigen influenza vaccine. 1. Purification of haemagglutinin and neuraminidase proteins. *Journal of Hygiene* 77(02):161-172.
- Braig, K, Simon, M, Furuya, F, Hainfeld, JF and Horwich, AL. 1993. A polypeptide bound by the chaperonin groEL is localized within a central cavity. *Proceedings of the National Academy of Sciences of the United States of America* 90(9):3978-3982.
- Braig, K, Otwinowski, Z, Hegde, R, Boisvert, DC, Joachimiak, A, Horwich, AL and Sigler, PB. 1994. The crystal structure of the bacterial chaperonin GroEL at 2.8 Å. *Nature* 371(6498):578-586.
- Brandon, F, Barrett, C, Hook, A and Lease, G. 1967. Human febrile response to influenza virus or its ether isolated hemagglutinins. *Proceedings of the Society for Experimental Biology and Medicine. Society for Experimental Biology and Medicine (New York, NY)* 125:683-686.
- Bright, RA, Carter, DM, Daniluk, S, Toapanta, FR, Ahmad, A, Gavrillov, V, Massare, M, Pushko, P, Mytle, N and Rowe, T. 2007. Influenza virus-like particles elicit broader immune responses than whole virion inactivated influenza virus or recombinant hemagglutinin. *Vaccine* 25(19):3871-3878.
- Bright, RA, Carter, DM, Crevar, CJ, Toapanta, FR, Steckbeck, JD, Cole, KS, Kumar, NM, Pushko, P, Smith, G, Tumpey, TM and Ross, TM. 2008. Cross-clade protective immune responses to influenza viruses with H5N1 HA and NA elicited by an influenza virus-like particle. *PLoS ONE* 3(1):e1501.
- Bubeck, P, Winkler, M and Bautsch, W. 1993. Rapid cloning by homologous recombination *in vivo*. *Nucleic Acids Research* 21(15):3601.
- Bundy, BC, Franciszkowicz, MJ and Swartz, JR. 2008. *Escherichia coli*-based cell-free synthesis of virus-like particles. *Biotechnology and Bioengineering* 100(1):28-37.
- Buonaguro, L, Buonaguro, F, Tornesello, M, Mantas, D, Beth-Giraldo, E, Wagner, R and al., e. 2001. High efficient production of Pr55^{gag} virus-like particles expressing multiple HIV-1 epitopes, including a gp120 protein derived from an Ugandan HIV-1 isolate of subtype A. *Antiviral Research* 49(1):35-47.
- Cabrita, LD, Gilis, D, Robertson, AL, Dehouck, Y, Rooman, M and Bottomley, SP. 2007. Enhancing the stability and solubility of TEV protease using *in silico* design. *Protein Science* 16(11):2360-2367.

- Calloni, G, Chen, T, Schermann, SM, Chang, H-c, Genevoux, P, Agostini, F, Tartaglia, GG, Hayer-Hartl, M and Hartl, FU. 2012. DnaK functions as a central hub in the *E. coli* chaperone network. *Cell Reports* 1(3):251-264.
- Caton, AJ, Brownlee, GG, Yewdell, JW and Gerhard, W. 1982. The antigenic structure of the influenza virus A/PR/8/34 hemagglutinin (H1 subtype). *Cell* 31(2 II):417-427.
- Caylor, C, Dobrianov, I, Lemay, S, Kimmer, C, Kriminski, S, Finkelstein, K, Zipfel, W, Webb, W, Thomas, B and Chernov, A. 1999. Macromolecular impurities and disorder in protein crystals. *Proteins: Structure, Function, and Bioinformatics* 36(3):270-281.
- Chackerian, B, Lowy, DR and Schiller, JT. 2001. Conjugation of a self-antigen to papillomavirus-like particles allows for efficient induction of protective autoantibodies. *Journal of Clinical Investigation* 108(3):415-423.
- Chackerian, B, Lenz, P, Lowy, DR and Schiller, JT. 2002. Determinants of autoantibody induction by conjugated papillomavirus virus-like particles. *The Journal of Immunology* 169(11):6120-6126.
- Chandrasekhar, GN, Tilly, K, Woolford, C, Hendrix, R and Georgopoulos, C. 1986. Purification and properties of the groES morphogenetic protein of *Escherichia coli*. *Journal of Biological Chemistry* 261(26):12414-12419.
- Checa, SK and Viale, AM. 1997. The 70-kDa heat-shock protein/DnaK chaperone system is required for the productive folding of ribulose-bisphosphate carboxylase subunits in *Escherichia coli*. *European Journal of Biochemistry* 248(3):848-855.
- Chen, J, Ma, C and Wong, C. 2011. Vaccine design of hemagglutinin glycoprotein against influenza. *Trends in Biotechnology* 29(9):426-434.
- Chen, J-R, Ma, C and Wong, C-H. 2011. Vaccine design of hemagglutinin glycoprotein against influenza. *Trends in Biotechnology* 29(9):426-434.
- Chen, S, Roseman, AM, Hunter, AS, Wood, SP, Burston, SG, Ranson, NA, Clarke, AR and Saibil, HR. 1994. Location of a folding protein and shape changes in GroEL-GroES complexes imaged by cryo-electron microscopy. *Nature* 371(6494):261-264.
- Chen, W-Y, Wu, C-F and Liu, C-C. 1996. Interactions of imidazole and proteins with immobilized Cu (II) ions: Effects of structure, salt concentration, and pH in affinity and binding capacity. *Journal of Colloid and Interface Science* 180(1):135-143.
- Chen, W-Y, Lee, J-F, Wu, C-F and Tsao, H-K. 1997. Microcalorimetric studies of the interactions of lysozyme with immobilized Cu (II): Effects of pH value and salt concentration. *Journal of Colloid and Interface Science* 190(1):49-54.

- Chen, XS, Garcea, RL, Goldberg, I, Casini, G and Harrison, SC. 2000. Structure of small virus-like particles assembled from the L1 protein of human papillomavirus 16. *Molecular Cell* 5(3):557-567.
- Chezzi, C, Dommann, C, Blackburn, N, Maselesele, E, McAnerney, J and Schoub, B. 1998. Genetic stability of oral polio vaccine prepared on primary monkey kidney cells or Vero cells—effects of passage in cell culture and the human gastrointestinal tract. *Vaccine* 16(20):2031-2038.
- Chiew, Y, Kuehner, D, Blanch, H and Prausnitz, J. 1995. Molecular thermodynamics for salt - induce protein precipitation. *AIChE Journal* 41(9):2150-2159.
- Chiti, F, Stefani, M, Taddei, N, Ramponi, G and Dobson, CM. 2003. Rationalization of the effects of mutations on peptide and protein aggregation rates. *Nature* 424(6950):805-808.
- Choi, J, Jeong, K, Kim, S and Lee, S. 2000. Efficient secretory production of alkaline phosphatase by high cell density culture of recombinant *Escherichia coli* using the *Bacillus sp.* endoxylanase signal sequence. *Applied Microbiology and Biotechnology* 53(6):640-645.
- Chuan, YP, Fan, YY, Lua, L and Middelberg, APJ. 2008. Quantitative analysis of virus-like particle size and distribution by field-flow fractionation. *Biotechnology and Bioengineering* 99(6):1425-1433.
- Chuan, YP, Lua, LHL and Middelberg, APJ. 2008. High-level expression of soluble viral structural protein in *Escherichia coli*. *Journal of Biotechnology* 134(1–2):64-71.
- Chuan, YP, Fan, YY, Lua, LH and Middelberg, AP. 2010. Virus assembly occurs following a pH- or Ca²⁺-triggered switch in the thermodynamic attraction between structural protein capsomeres. *Journal of The Royal Society Interface* 7(44):409-421.
- Chuan, YP, Rivera-Hernandez, T, Wibowo, N, Connors, NK, Wu, Y, Hughes, FK, Lua, LHL and Middelberg, APJ. 2013. Effects of pre-existing anti-carrier immunity and antigenic element multiplicity on efficacy of a modular virus-like particle vaccine. *Biotechnology and Bioengineering* 110 (9):2343-2351.
- Connors, NK, Wu, Y, Lua, LHL and Middelberg, APJ. 2013. Improved fusion tag cleavage strategies in the downstream processing of self-assembling virus-like particle vaccines. *Food and Bioproducts Processing* 92(2):143–151.
- Cooper, PD and Petrovsky, N. 2011. Delta inulin: a novel, immunologically active, stable packing structure comprising β -d-[2 \rightarrow 1] poly (fructo-furanosyl) α -d-glucose polymers. *Glycobiology* 21(5):595-606.

- Correia, BE, Ban, Y-EA, Holmes, MA, Xu, H, Ellingson, K, Kraft, Z, Carrico, C, Boni, E, Sather, DN and Zenobia, C. 2010. Computational design of epitope-scaffolds allows induction of antibodies specific for a poorly immunogenic HIV vaccine epitope. *Structure* 18(9):1116-1126.
- Correia, BE, Bates, JT, Loomis, RJ, Baneyx, G, Carrico, C, Jardine, JG, Rupert, P, Correnti, C, Kalyuzhniy, O and Vittal, V. 2014. Proof of principle for epitope-focused vaccine design. *Nature* 507(7491):201-206.
- Cox, M. 2008. Progress on baculovirus-derived influenza vaccines. *Current Opinion in Molecular Therapeutics* 10(1):56-61.
- Cox, MM and Hollister, JR. 2009. FluBlok, a next generation influenza vaccine manufactured in insect cells. *Biologicals* 37(3):182-189.
- Cox, NJ and Subbarao, K. 1999. Influenza. *Lancet* 354(9186):1277-1282.
- Craig, EA, Gambill, BD and Nelson, RJ. 1993. Heat shock proteins: molecular chaperones of protein biogenesis. *Microbiological Reviews* 57(2):402-414.
- Crisci, E, Fraile, L, Moreno, N, Blanco, E, Cabezón, R, Costa, C, Mussá, T, Baratelli, M, Martínez-Orellana, P, Ganges, L, Martínez, J, Bárcena, J and Montoya, M. 2012. Chimeric calicivirus-like particles elicit specific immune responses in pigs. *Vaccine* 30(14):2427-2439.
- Cristillo, AD, Ferrari, MG, Hudacik, L, Lewis, B, Galmin, L, Bowen, B, Thompson, D, Petrovsky, N, Markham, P and Pal, R. 2011. Induction of mucosal and systemic antibody and T-cell responses following prime–boost immunization with novel adjuvanted human immunodeficiency virus-1-vaccine formulations. *Journal of General Virology* 92(1):128-140.
- Curtis, R, Prausnitz, J and Blanch, H. 1998. Protein - protein and protein - salt interactions in aqueous protein solutions containing concentrated electrolytes. *Biotechnology and Bioengineering* 57(1):11-21.
- Curtis, R, Ulrich, J, Montaser, A, Prausnitz, J and Blanch, H. 2002a. Protein–protein interactions in concentrated electrolyte solutions. *Biotechnology and Bioengineering* 79(4):367-380.
- Curtis, R, Steinbrecher, C, Heinemann, M, Blanch, H and Prausnitz, J. 2002b. Hydrophobic forces between protein molecules in aqueous solutions of concentrated electrolyte. *Biophysical Chemistry* 98(3):249-265.
- Dale, GE, Oefner, C and D'Arcy, A. 2003. The protein as a variable in protein crystallization. *Journal of Structural Biology* 142(1):88-97.

- Daniels, R, Kurowski, B, Johnson, AE and Hebert, DN. 2003. N-Linked glycans direct the cotranslational folding pathway of influenza hemagglutinin. *Molecular Cell* 11(1):79-90.
- Das, SR, Puigbò, P, Hensley, SE, Hurt, DE, Bennink, JR and Yewdell, JW. 2010. Glycosylation focuses sequence variation in the influenza A virus H1 hemagglutinin globular domain. *PLoS Pathogens* 6(11):e1001211.
- Das, SR, Hensley, SE, David, A, Schmidt, L, Gibbs, JS, Puigbò, P, Ince, WL, Bennink, JR and Yewdell, JW. 2011. Fitness costs limit influenza A virus hemagglutinin glycosylation as an immune evasion strategy. *Proceedings of the National Academy of Sciences* 108(51):E1417-E1422.
- De Filette, M, Ramne, A, Birkett, A, Lycke, N, Löwenadler, B, Min Jou, W, Saelens, X and Fiers, W. 2006. The universal influenza vaccine M2e-HBc administered intranasally in combination with the adjuvant CTA1-DD provides complete protection. *Vaccine* 24(5):544-551.
- De Filette, M, Martens, W, Roose, K, Deroo, T, Vervalle, F, Bentahir, M, Vandekerckhove, J, Fiers, W and Saelens, X. 2008. An influenza A vaccine based on tetrameric ectodomain of matrix protein 2. *Journal of Biological Chemistry* 283(17):11382-11387.
- Deryaguin, BV and Landau, L. 1941. *Acta Physicochimica. USSR* 14:633.
- Deshpande, KL, Fried, VA, Ando, M and Webster, RG. 1987. Glycosylation affects cleavage of an H5N2 influenza virus hemagglutinin and regulates virulence. *Proceedings of the National Academy of Sciences* 84(1):36-40.
- Dessau, MA and Modis, Y. 2011. Protein crystallization for X-ray crystallography. *Journal of Visualized Experiments: JoVE* (47):art. no. e2285.
- Dormitzer, PR, Ulmer, JB and Rappuoli, R. 2008. Structure-based antigen design: a strategy for next generation vaccines. *Trends in Biotechnology* 26(12):659-667.
- Dormitzer, PR, Grandi, G and Rappuoli, R. 2012. Structural vaccinology starts to deliver. *Nature Reviews Microbiology* 10(12):807-813.
- Doroshenko, A and Halperin, SA. 2009. Trivalent MDCK cell culture-derived influenza vaccine Optaflu® (Novartis Vaccines). *Expert Review of Vaccines* 8(6):679-688.
- Eckersley, A, Petrovsky, N, Kinne, J, Wernery, R and Wernery, U. 2011. Improving the dromedary antibody response: the hunt for the ideal camel adjuvant. *Journal of Camel Practice and Research* 18(1):30-46.
- Ehreth, J. 2003. The global value of vaccination. *Vaccine* 21(7):596-600.

- Einstein, MH, Baron, M, Levin, MJ, Chatterjee, A, Edwards, RP, Zepp, F, Carletti, I, Dessy, FJ, Trofa, AF, Schuind, A and Dubin, G. 2009. Comparison of the immunogenicity and safety of Cervarix™ and Gardasil® human papillomavirus (HPV) cervical cancer vaccines in healthy women aged 18-45 years. *Human Vaccines* 5(10):705-719.
- Ejima, D, Yumioka, R, Arakawa, T and Tsumoto, K. 2005. Arginine as an effective additive in gel permeation chromatography. *Journal of Chromatography A* 1094(1):49-55.
- Ekiert, DC, Kashyap, AK, Steel, J, Rubrum, A, Bhabha, G, Khayat, R, Lee, JH, Dillon, MA, O'Neil, RE and Faynboym, AM. 2012. Cross-neutralization of influenza A viruses mediated by a single antibody loop. *Nature* 489(7417):526-532.
- Enders, JF, Weller, TH and Robbins, FC. 1949. Cultivation of the Lansing strain of poliomyelitis virus in cultures of various human embryonic tissues. *Science* 109(2822):85-87.
- Esler, WP, Stimson, ER, Ghilardi, JR, Lu, Y-A, Felix, AM, Vinters, HV, Mantyh, PW, Lee, JP and Maggio, JE. 1996. Point substitution in the central hydrophobic cluster of a human β -amyloid congener disrupts peptide folding and abolishes plaque competence[†]. *Biochemistry* 35(44):13914-13921.
- Esposito, D and Chatterjee, DK. 2006. Enhancement of soluble protein expression through the use of fusion tags. *Current Opinion in Biotechnology* 17(4):353-358.
- Fan, J, Liang, X, Horton, MS, Perry, HC, Citron, MP, Heidecker, GJ, Fu, T-M, Joyce, J, Przysiecki, CT and Keller, PM. 2004. Preclinical study of influenza virus A M2 peptide conjugate vaccines in mice, ferrets, and rhesus monkeys. *Vaccine* 22(23):2993-3003.
- Fan, YY and Middelberg, AP. Method of preparation of a biological particulate structure US 13/380,300. 2010.
- Fenton, WA and Horwich, AL. 1997. GroEL - mediated protein folding. *Protein Science* 6(4):743-760.
- Ferguson, NM, Cummings, DAT, Fraser, C, Cajka, JC, Cooley, PC and Burke, DS. 2006. Strategies for mitigating an influenza pandemic. *Nature* 442(7101):448-452.
- Fernandes, F, Teixeira, AP, Carinhas, N, Carrondo, MJ and Alves, PM. 2013. Insect cells as a production platform of complex virus-like particles. *Expert Review of Vaccines* 12(2):225-236.
- Fiers, W, De Filette, M, Bakkouri, KE, Schepens, B, Roose, K, Schotsaert, M, Birkett, A and Saelens, X. 2009. M2e-based universal influenza A vaccine. *Vaccine* 27(45):6280-6283.

- Fink, AL. 1999. Chaperone-mediated protein folding. *Physiological Reviews* 79(2):425-449.
- Fitzmaurice, CJ, Brown, LE, McInerney, TL and Jackson, DC. 1996. The assembly and immunological properties of non-linear synthetic immunogens containing T-cell and B-cell determinants. *Vaccine* 14(6):553-560.
- Fontenot, J, Mariappan, S, Catasti, P, Domenech, N, Finn, OJ and Gupta, G. 1995. Structure of a tumor associated antigen containing a tandemly repeated immunodominant epitope. *Journal of Biomolecular Structure and Dynamics* 13(2):245-260.
- Fontenot, JD, Tjandra, N, Bu, D, Ho, C, Montelaro, RC and Finn, OJ. 1993. Biophysical characterization of one-, two-, and three-tandem repeats of human mucin (muc-1) protein core. *Cancer Research* 53(22):5386.
- Fraaije, J, Murriss, R, Norde, W and Lyklema, J. 1991. Interfacial thermodynamics of protein adsorption, ion co-adsorption and ion binding in solution: I. Phenomenological linkage relations for ion exchange in lysozyme chromatography and titration in solution. *Biophysical Chemistry* 40(3):303-315.
- Francis, DM and Page, R. 2010. Strategies to optimize protein expression in *E. coli*. *Current Protocols in Protein Science (SUPPL. 61)*:5.24.21-25.24.29.
- Fraunhofer, W and Winter, G. 2004. The use of asymmetrical flow field-flow fractionation in pharmaceuticals and biopharmaceuticals. *European Journal of Pharmaceutics and Biopharmaceutics* 58(2):369-383.
- Galloway, CA, Sowden, MP and Smith, HC. 2003. Increasing the yield of soluble recombinant protein expressed in *E. coli* by induction during late log phase. *Biotechniques* 34(3):524-526, 528, 530.
- Gamblin, S, Haire, L, Russell, R, Stevens, D, Xiao, B, Ha, Y, Vasisht, N, Steinhauer, D, Daniels, R and Elliot, A. 2004. The structure and receptor binding properties of the 1918 influenza hemagglutinin. *Science* 303(5665):1838-1842.
- Garcea, RL and Gissmann, L. 2004. Virus-like particles as vaccines and vessels for the delivery of small molecules. *Current Opinion in Biotechnology* 15(6):513-517.
- Gasteiger, E, Gattiker, A, Hoogland, C, Ivanyi, I, Appel, RD and Bairoch, A. 2003. ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Research* 31(13):3784-3788.
- Gasteiger, E, Hoogland, C, Gattiker, A, Wilkins, MR, Appel, RD and Bairoch, A. 2005. The proteomics protocols handbook. In: Walker, J M, editor. *Protein identification and analysis tools on the ExPASy server*: Springer. p. 571-607.

- Gedvilaite, A, Frömmel, C, Sasnauskas, K, Micheel, B, Özel, M, Behrsing, O, Staniulis, J, Jandrig, B, Scherneck, S and Ulrich, R. 2000. Formation of immunogenic virus-like particles by inserting epitopes into surface-exposed regions of hamster polyomavirus major capsid protein. *Virology* 273(1):21-35.
- Gedvilaite, A, Zvirbliene, A, Staniulis, J, Sasnauskas, K, Krüger, DH and Ulrich, R. 2004. Segments of puumala hantavirus nucleocapsid protein inserted into chimeric polyomavirus-derived virus-like particles induce a strong immune response in mice. *Viral Immunology* 17(1):51-68.
- Geng, X, Guo, LA and Chang, J. 1990. Study of the retention mechanism of proteins in hydrophobic interaction chromatography. *Journal of Chromatography A* 507:1-23.
- George, A and Wilson, WW. 1994. Predicting protein crystallization from a dilute solution property. *Acta Crystallographica Section D: Biological Crystallography* 50(4):361-365.
- Gerdil, C. 2003. The annual production cycle for influenza vaccine. *Vaccine* 21(16):1776-1779.
- Ghendon, Y, Markushin, S, Akopova, I, Koptiaeva, I, Nechaeva, E, Mazurkova, L, Radaeva, I and Kolokoltseva, T. 2005. Development of cell culture (MDCK) live cold-adapted (CA) attenuated influenza vaccine. *Vaccine* 23(38):4678-4684.
- Giddings, JC. 2002. The Field-Flow Fractionation Family: Underlying Principles. In: Schimpf, M, Caldwell, K and Giddings, J C, editor. *Field flow fractionation handbook*: Wiley-Interscience, Inc. p. 3-30.
- Gillock, E, Rottinghaus, S, Chang, D, Cai, X, Smiley, S, An, K and Consigli, R. 1997. Polyomavirus major capsid protein VP1 is capable of packaging cellular DNA when expressed in the baculovirus system. *Journal of Virology* 71(4):2857-2865.
- Gleiter, S, Stubenrauch, K and Lilie, H. 1999. Changing the surface of a virus shell fusion of an enzyme to polyoma VP1. *Protein Science* 8(12):2562-2569.
- Gleiter, S and Lilie, H. 2001. Coupling of antibodies via protein Z on modified polyoma virus-like particles. *Protein Science* 10(2):434-444.
- Goda, Y, Kobayashi, A, Fukuda, K, Fujimoto, S, Ike, M and Fujita, M. 2000. Development of the ELISAs for detection of hormone-disrupting chemicals. *Water Science and Technology* 42(7-8):81-88.
- Golovanov, AP, Hautbergue, GM, Wilson, SA and Lian, L-Y. 2004. A simple method for improving protein solubility and long-term stability. *Journal of the American Chemical Society* 126(29):8933-8939.

- Gopal, GJ and Kumar, A. 2013. Strategies for the production of recombinant protein in *Escherichia coli*. *The Protein Journal* 32(6):419-425.
- Gori, A, Longhi, R, Peri, C and Colombo, G. 2013. Peptides for immunological purposes: Design, strategies and applications. *Amino Acids* 45(2):257-268.
- Grimm, SK and Ackerman, ME. 2013. Vaccine design: Emerging concepts and renewed optimism. *Current Opinion in Biotechnology* 24(6):1078-1088.
- Guex, N, Peitsch, MC and Schwede, T. 2009. Automated comparative protein structure modeling with SWISS - MODEL and Swiss - PdbViewer: A historical perspective. *Electrophoresis* 30(S1):S162-S173.
- Ha, J-H, Capp, MW, Hohenwalter, MD, Baskerville, M and Record Jr, MT. 1992. Thermodynamic stoichiometries of participation of water, cations and anions in specific and non-specific binding of *lac* repressor to DNA: Possible thermodynamic origins of the "glutamate effect" on protein-DNA interactions. *Journal of Molecular Biology* 228(1):252-264.
- Haaheim, L, Van-Tam, J and Sellwood, C. 2010. Introduction to Pandemic Influenza. In: Van-Tam, J and Sellwood, C, editor. *Basic influenza virology and immunology*: CAB International. p. 14-27.
- Habig, WH, Pabst, MJ and Jakoby, WB. 1974. Glutathione S-transferases: The first enzymatic step in mercapturic acid formation. *Journal of Biological Chemistry* 249(22):7130-7139.
- Halperin, SA, Nestruck, AC and Eastwood, BJ. 1998. Safety and immunogenicity of a new influenza vaccine grown in mammalian cell culture. *Vaccine* 16(13):1331-1335.
- Hanson, SR, Culyba, EK, Hsu, T-L, Wong, C-H, Kelly, JW and Powers, ET. 2009. The core trisaccharide of an N-linked glycoprotein intrinsically accelerates folding and enhances stability. *Proceedings of the National Academy of Sciences* 106(9):3131-3136.
- Harper, JD, Lieber, CM and Lansbury Jr, PT. 1997. Atomic force microscopic imaging of seeded fibril formation and fibril branching by the Alzheimer's disease amyloid- β protein. *Chemistry and Biology* 4(12):951-959.
- Harrison, CJ, Hayer-Hartl, M, Di Liberto, M, Hartl, F-U and Kuriyan, J. 1997. Crystal structure of the nucleotide exchange factor GrpE bound to the ATPase domain of the molecular chaperone DnaK. *Science* 276(5311):431-435.
- Hay, AJ, Gregory, V, Douglas, AR and Lin, YP. 2001. The evolution of human influenza viruses. *Philosophical transactions of the Royal Society of London. Series B, Biological Sciences* 356(1416):1861-1870.

- Hayman, WA, Brandt, ER, Relf, WA, Cooper, J, Saul, A and Good, MF. 1997. Mapping the minimal murine T cell and B cell epitopes within a peptide vaccine candidate from the conserved region of the M protein of group A streptococcus. *International Immunology* 9(11):1723-1733.
- Haynes, JR. 2009. Influenza virus-like particle vaccines. *Expert Review of Vaccines* 8(4):435-445.
- Haynes, JR, Dokken, L, Wiley, JA, Cawthon, AG, Bigger, J, Harmsen, AG and Richardson, C. 2009. Influenza-pseudotyped Gag virus-like particle vaccines provide broad protection against highly pathogenic avian influenza challenge. *Vaccine* 27(4):530-541.
- Henot, F, Legon, T, Piroton, S and Placier, G. Method for the purification of heat shock protein. US 12/311,755. 2008.
- Henot, F. Method for the purification of heat shock proteins. US 12/311,755. 2010.
- Hess, B, Kutzner, C, van der Spoel, D and Lindahl, E. 2008. Gromacs 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *Journal of Chemical Theory and Computation* 4(3):435-447.
- Hofmeister, F. 1888. Zur lehre von der wirkung der salze. *Archiv für experimentelle Pathologie und Pharmakologie* 25(1):1-30.
- Hogle, J, Maeda, A and Harrison, S. 1986. Structure and assembly of turnip crinkle virus: I. X-ray crystallographic structure analysis at 3.2 Å resolution. *Journal of Molecular Biology* 191(4):625-638.
- Honda-Okubo, Y, Saade, F and Petrovsky, N. 2012. Advax™, a polysaccharide adjuvant derived from delta inulin, provides improved influenza vaccine protection through broad-based enhancement of adaptive immune responses. *Vaccine* 30(36):5373-5381.
- Hoover, DM and Lubkowski, J. 2002. DNABWorks: an automated method for designing oligonucleotides for PCR-based gene synthesis. *Nucleic Acids Research* 30(10):e43.
- Horimoto, T and Kawaoka, Y. 2005. Influenza: lessons from past pandemics, warnings from current incidents. *Nature Reviews Microbiology* 3(8):591-600.
- Hoyle, L, Horne, R and Waterson, A. 1961. The structure and composition of the myxoviruses: II. Components released from the influenza virus particle by ether. *Virology* 13(4):448-459.

- Hu, AY-C, Weng, T-C, Tseng, Y-F, Chen, Y-S, Wu, C-H, Hsiao, S, Chou, A-H, Chao, H-J, Gu, A and Wu, S-C. 2008. Microcarrier-based MDCK cell culture system for the production of influenza H5N1 vaccines. *Vaccine* 26(45):5736-5740.
- Huston, JS, Levinson, D, Mudgett-Hunter, M, Tai, M-S, Novotný, J, Margolies, MN, Ridge, RJ, Brucoleri, RE, Haber, E and Crea, R. 1988. Protein engineering of antibody binding sites: recovery of specific activity in an anti-digoxin single-chain Fv analogue produced in *Escherichia coli*. *Proceedings of the National Academy of Sciences* 85(16):5879-5883.
- Ionescu, RM, Przysiecki, CT, Liang, X, Garsky, VM, Fan, J, Wang, B, Troutman, R, Rippeon, Y, Flanagan, E and Shiver, J. 2006. Pharmaceutical and immunological evaluation of human papillomavirus viruslike particle as an antigen carrier. *Journal of Pharmaceutical Sciences* 95(1):70-79.
- Israelachvili, JN and Pashley, RM. 1984. Measurement of the hydrophobic interaction between two hydrophobic surfaces in aqueous electrolyte solutions. *Journal of Colloid And Interface Science* 98(2):500-514.
- Izaac, A, Schall, CA and Mueser, TC. 2006. Assessment of a preliminary solubility screen to improve crystallization trials: uncoupling crystal condition searches. *Acta Crystallographica Section D: Biological Crystallography* 62(7):833-842.
- Jain, S, Patrick, AJ and Rosenthal, KL. 2010. Multiple tandem copies of conserved gp41 epitopes incorporated in gag virus-like particles elicit systemic and mucosal antibodies in an optimized heterologous vector delivery regimen. *Vaccine* 28(43):7070-7080.
- Jegerlehner, A, Storni, T, Lipowsky, G, Schmid, M, Pumpens, P and Bachmann, MF. 2002. Regulation of IgG antibody responses by epitope density and CD21 - mediated costimulation. *European Journal of Immunology* 32(11):3305-3314.
- Jennings, GT and Bachmann, MF. 2007. Designing recombinant vaccines with viral properties: a rational approach to more effective vaccines. *Current Molecular Medicine* 7(2):143-155.
- Jennings, GT and Bachmann, MF. 2009. Immunodrugs: therapeutic VLP-based vaccines for chronic diseases. *Annual Review of Pharmacology and Toxicology* 49:303-326.
- Johansson, BE. 1999. Immunization with influenza A virus hemagglutinin and neuraminidase produced in recombinant baculovirus results in a balanced and broadened immune response superior to conventional vaccine. *Vaccine* 17(15):2073-2080.

- Jones, D. 1994. PCR mutagenesis and recombination *in vivo*. *Genome Research* 3(6):S141-S148.
- Jung, H-J, Kim, S-K, Min, W-K, Lee, S-S, Park, K, Park, Y-C and Seo, J-H. 2011. Polycationic amino acid tags enhance soluble expression of *Candida antarctica* lipase B in recombinant *Escherichia coli*. *Bioprocess and Biosystems Engineering* 34(7):833-839.
- Kamireddi, M, Eisenstein, E and Reddy, P. 1997. Stable Expression and Rapid Purification of *Escherichia coli* GroEL and GroES Chaperonins. *Protein Expression and Purification* 11(1):47-52.
- Kang, S-M, Song, J-M, Quan, F-S and Compans, RW. 2009. Influenza vaccines based on virus-like particles. *Virus Research* 143(2):140-146.
- Kang, S-M, Kim, M-C and Compans, RW. 2012. Virus-like particles as universal influenza vaccines. *Expert Review of Vaccines* 11(8):995-1007.
- Kapust, RB and Waugh, DS. 1999. *Escherichia coli* maltose - binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein Science* 8(8):1668-1674.
- Karpenko, L, Ivanisenko, V, Pika, I, Chikaev, N, Eroshkin, A, Veremeiko, T and Ilyichev, A. 2000. Insertion of foreign epitopes in HBcAg: how to make the chimeric particle assemble. *Amino Acids* 18(4):329-337.
- Kato, A, Maki, K, Ebina, T, Kuwajima, K, Soda, K and Kuroda, Y. 2007. Mutational analysis of protein solubility enhancement using short peptide tags. *Biopolymers* 85(1):12-18.
- Katz, JM and Webster, RG. 1992. Amino acid sequence identity between the HA1 of influenza A (H3N2) viruses grown in mammalian and primary chick kidney cells. *The Journal of General Virology* 73(5):1159-1165.
- Kaverin, NV, Rudneva, IA, Ilyushina, NA, Lipatov, AS, Krauss, S and Webster, RG. 2004. Structural differences among hemagglutinins of influenza A virus subtypes are reflected in their antigenic architecture: analysis of H9 escape mutants. *Journal of Virology* 78(1):240-249.
- Kaverin, NV, Rudneva, IA, Govorkova, EA, Timofeeva, TA, Shilov, AA, Kochergin-Nikitsky, KS, Krylov, PS and Webster, RG. 2007. Epitope mapping of the hemagglutinin molecule of a highly pathogenic H5N1 influenza virus by using monoclonal antibodies. *Journal of Virology* 81(23):12911-12917.

- Kawaoka, Y and Webster, RG. 1989. Interplay between carbohydrate in the stalk and the length of the connecting peptide determines the cleavability of influenza virus hemagglutinin. *Journal of Virology* 63(8):3296-3300.
- Kazaks, A, Borisova, G, Cvetkova, S, Kovalevska, L, Ose, V, Sominskaya, I, Pumpens, P, Skrastina, D and Dislers, A. 2004. Mosaic hepatitis B virus core particles presenting the complete preS sequence of the viral envelope on their surface. *Journal of General Virology* 85(9):2665-2670.
- Keating, GM, Noble, S, Averhoff, FM, Belloni, C, Duval, B, Goldwater, PN, Hall, AJ, Honorati, MC, Kallinowski, B, Leroux-Roels, G and Poovorawan, Y. 2003. Recombinant hepatitis B vaccine (Engerix-B®): A review of its immunogenicity and protective efficacy against hepatitis B. *Drugs* 63(10):1021-1051.
- Keil, W, Geyer, R, Dabrowski, J, Dabrowski, U, Niemann, Hv, Stirm, S and Klenk, H. 1985. Carbohydrates of influenza virus. Structural elucidation of the individual glycans of the FPV hemagglutinin by two-dimensional ¹H n.m.r and methylation analysis. *The EMBO Journal* 4(10):2711.
- Khandekar, SS, Bettencourt, BM, Kelley, KC and Recny, MA. 1993. A simple and rapid method for the purification of GroEL, an *Escherichia coli* homolog of the heat shock protein 60 family of molecular chaperonins. *Protein Expression and Purification* 4(6):580-584.
- Khurana, S, Verma, S, Verma, N, Crevar, CJ, Carter, DM, Manischewitz, J, King, LR, Ross, TM and Golding, H. 2010. Properly folded bacterially expressed H1N1 hemagglutinin globular head and ectodomain vaccines protect ferrets against H1N1 pandemic influenza virus. *PLoS ONE* 5(7):e11548.
- Khurana, S, Verma, S, Verma, N, Crevar, CJ, Carter, DM, Manischewitz, J, King, LR, Ross, TM and Golding, H. 2011. Bacterial HA1 vaccine against pandemic H5N1 influenza virus: evidence of oligomerization, hemagglutination, and cross-protective immunity in ferrets. *Journal of Virology* 85(3):1246-1256.
- Kiefer, F, Arnold, K, Künzli, M, Bordoli, L and Schwede, T. 2009. The SWISS-MODEL Repository and associated resources. *Nucleic Acids Research* 37(suppl 1):D387-D392.
- Kistner, O, Barrett, P, Mundt, W, Reiter, M, Schober-Bendixen, S and Dorner, F. 1998. Development of a mammalian cell (Vero) derived candidate influenza virus vaccine. *Vaccine* 16(9):960-968.
- Kleywegt, GJ and Jones, TA. 1998. Databases in protein crystallography. *Acta Crystallographica Section D: Biological Crystallography* 54(6):1119-1131.

- Koletzki, D, Zankl, A, Gelderblom, HR, Meisel, H, Dislers, A, Borisova, G, Pumpens, P, Kr, D and Ulrich, R. 1997. Mosaic hepatitis B virus core particles allow insertion of extended foreign protein segments. *Journal of General Virology* 78(8):2049-2053.
- Kopito, RR. 2000. Aggresomes, inclusion bodies and protein aggregation. *Trends in Cell Biology* 10(12):524-530.
- Koutsky, LA, Ault, KA, Wheeler, CM, Brown, DR, Barr, E, Alvarez, FB, Chiacchierini, LM and Jansen, KU. 2002. A controlled trial of a human papillomavirus type 16 vaccine. *New England Journal of Medicine* 347(21):1645-1651.
- Kovacs-Nolan, J and Mine, Y. 2006. Tandem copies of a human rotavirus VP8 epitope can induce specific neutralizing antibodies in BALB/c mice. *Biochimica et Biophysica Acta (BBA)-General Subjects* 1760(12):1884-1893.
- Kratz, PA, Böttcher, B and Nassal, M. 1999. Native display of complete foreign protein domains on the surface of hepatitis B virus capsids. *Proceedings of the National Academy of Sciences of the United States of America* 96(5):1915-1920.
- Kulp, DW and Schief, WR. 2013. Advances in structure-based vaccine design. *Current Opinion in Virology* 3(3):322-331.
- Kwong, PD and Shapiro, L. 2011. Vaccine design reaches the atomic level. *Science Translational Medicine* 3(91):art. no. 91ps29.
- Kyte, J and Doolittle, RF. 1982. A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology* 157(1):105-132.
- Langer, T, Pfeifer, G, Martin, J, Baumeister, W and Hartl, F. 1992. Chaperonin-mediated protein folding: GroES binds to one end of the GroEL cylinder, which accommodates the protein substrate within its central cavity. *The EMBO Journal* 11(13):4757.
- Latham, T and Galarza, JM. 2001. Formation of wild-type and chimeric influenza virus-like particles following simultaneous expression of only four structural proteins. *Journal of Virology* 75(13):6154-6165.
- Lattman, EE. 1980. Simian virus 40 crystals. *Science* 208(4447):1048-1050.
- Laver, W and Webster, R. 1976. Preparation and immunogenicity of a purified influenza virus haemagglutinin and neuraminidase subunit vaccine. *Postgraduate Medical Journal* 52(608):373-378.
- Layton, RC, Petrovsky, N, Gigliotti, AP, Pollock, Z, Knight, J, Donart, N, Pyles, J, Harrod, KS, Gao, P and Koster, F. 2011. Delta inulin polysaccharide adjuvant enhances the ability of split-virion H5N1 vaccine to protect against lethal challenge in ferrets. *Vaccine* 29(37):6242-6251.

- Leavitt, A, Roberts, T and Garcea, R. 1985. Polyoma virus major capsid protein, VP1. Purification after high level expression in *Escherichia coli*. *Journal of Biological Chemistry* 260(23):12803-12809.
- Lee, EG, Baek, J-E, Lee, S-H, Kim, T-W, Choi, JH, Rho, M-C, Ahn, J-O, Lee, H-W and Jung, J-K. 2009. Efficient proteolytic cleavage by insertion of oligopeptide linkers and its application to production of recombinant human interleukin-6 in *Escherichia coli*. *Enzyme and Microbial Technology* 44(5):254-262.
- Lee, M-S and Hu, AY-C. 2012. A cell-based backup to speed up pandemic influenza vaccine production. *Trends in Microbiology* 20(3):103-105.
- Letvin, NL. 2007. Correlates of immune protection and the development of a human immunodeficiency virus vaccine. *Immunity* 27(3):366-369.
- Leung, AKC. 2011. "Variolation" and vaccination in late imperial China. In: Plotkin, S A, editor. *History of Vaccine Development*: Springer. p. 5-12.
- Lewis, J, Brown, E and Duncan, P. 2005. Approaches to the release of a master cell bank of PER.C6 cells; a novel cell substrate for the manufacture of human vaccines. *Developments in Biologicals* 123:165-176.
- Li, P, Bellett, AJ and Parish, CR. 1983. A comparison of the terminal protein and hexon polypeptides of avian and human adenoviruses. *Journal of General Virology* 64(6):1375-1379.
- Liang, Y, Hilal, N, Langston, P and Starov, V. 2007. Interaction forces between colloidal particles in liquid: Theory and experiment. *Advances in Colloid and Interface Science* 134:151-166.
- Liddington, R, Yan, Y, Moulai, J, Sahli, R, Benjamin, T and Harrison, S. 1991. Structure of simian virus 40 at 3.8-Å resolution. *Nature* 354(6351):278-284.
- Liew, MWO, Rajendran, A and Middelberg, APJ. 2010. Microbial production of virus-like particle vaccine protein at gram-per-litre levels. *Journal of Biotechnology* 150(2):224-231.
- Lilie, H, Richter, S, Bergelt, S, Frost, S and Gehle, F. 2013. Polyionic and cysteine-containing fusion peptides as versatile protein tags. *Biological Chemistry* 394(8):995-1004.
- Lin, F-Y, Chen, W-Y and Sang, L-C. 1999. Microcalorimetric studies of the interactions of lysozyme with immobilized metal ions: effects of ion, pH value, and salt concentration. *Journal of Colloid and Interface Science* 214(2):373-379.
- Lin, F-Y, Chen, W-Y, Ruaan, R-C and Huang, H-M. 2000. Microcalorimetric studies of interactions between proteins and hydrophobic ligands in hydrophobic interaction

- chromatography: effects of ligand chain length, density and the amount of bound protein. *Journal of Chromatography A* 872(1):37-47.
- Lindahl, E, Hess, B and van der Spoel, D. 2001. GROMACS 3.0: A package for molecular simulation and trajectory analysis. *Journal of Molecular Modeling* 7(8):306-317.
- Lindquist, S and Craig, E. 1988. The heat-shock proteins. *Annual Review of Genetics* 22(1):631-677.
- Lindquist, S, Krobitsch, S, Li, L and Sondheimer, N. 2001. Investigating protein conformation-based inheritance and disease in yeast. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 356(1406):169-176.
- Lipin, DI, Lua, LHL and Middelberg, APJ. 2008. Quaternary size distribution of soluble aggregates of glutathione-S-transferase-purified viral protein as determined by asymmetrical flow field flow fractionation and dynamic light scattering. *Journal of Chromatography A* 1190(1-2):204-214.
- Liu, F, Wu, X, Li, L, Liu, Z and Wang, Z. 2013. Use of baculovirus expression system for generation of virus-like particles: successes and challenges. *Protein Expression and Purification* 90(2):104-116.
- Liu, J, Shi, X, Schwartz, R and Kemble, G. 2009. Use of MDCK cells for production of live attenuated influenza vaccine. *Vaccine* 27(46):6460-6463.
- Liu, W, Peng, Z, Liu, Z, Lu, Y, Ding, J and Chen, Y-H. 2004. High epitope density in a single recombinant protein molecule of the extracellular domain of influenza A virus M2 protein significantly enhances protective immunity. *Vaccine* 23(3):366-371.
- Lobigs, M, Pavy, M, Hall, RA, Lobigs, P, Cooper, P, Komiya, T, Toriniwa, H and Petrovsky, N. 2010. An inactivated Vero cell-grown Japanese encephalitis vaccine formulated with Advax, a novel inulin-based adjuvant, induces protective neutralizing antibody against homologous and heterologous flaviviruses. *Journal of General Virology* 91(6):1407-1417.
- Loktev, V, Ilyichev, A, Eroshkin, A, Karpenko, L, Pokrovsky, A, Pereboev, A, Svyatchenko, V, Ignat'ev, G, Smolina, M and Melamed, N. 1996. Design of immunogens as components of a new generation of molecular vaccines. *Journal of Biotechnology* 44(1):129-137.
- Lorber, B, Skouri, M, Munch, J-P and Giegé, R. 1993. The influence of impurities on protein crystallization; the case of lysozyme. *Journal of Crystal Growth* 128(1):1203-1211.

- Lua, LH, Connors, NK, Sainsbury, F, Chuan, YP, Wibowo, N and Middelberg, AP. 2013. Bioengineering virus - like particles as vaccines. *Biotechnology and Bioengineering* 111(3):425-440.
- Lua, LH-L and Middelberg, APJ. VLP based vaccine delivery system. US 12/444,092. 2007.
- Ludwig, C and Wagner, R. 2007. Virus-like particles—universal molecular toolboxes. *Current Opinion in Biotechnology* 18(6):537-545.
- Luzzago, A, Felici, F, Tramontano, A, Pessi, A and Cortese, R. 1993. Mimicking of discontinuous epitopes by phage-displayed peptides, I. Epitope mapping of human H ferritin using a phage library of constrained peptides. *Gene* 128(1):51-57.
- Maassab, HF, Heilman, CA and Herlocher, ML. 1990. Cold-adapted influenza viruses for use as live vaccines for man. *Advances in Biotechnological Processes* 14:203-242.
- Mahmood, K, Bright, RA, Mytle, N, Carter, DM, Crevar, CJ, Achenbach, JE, Heaton, PM, Tumpey, TM and Ross, TM. 2008. H5N1 VLP vaccine induced protection in ferrets against lethal challenge with highly pathogenic H5N1 influenza viruses. *Vaccine* 26(42):5393-5399.
- Mahn, A, Lienqueo, ME and Asenjo, JA. 2007. A simple method for the estimation of protein retention in hydrophobic interaction chromatography under different operation conditions. *Open Biotechnology Journal* 1:9-13.
- Malboeuf, CM, Simon, DA, Lee, Y-EE, Lankes, HA, Dewhurst, S, Frelinger, JG and Rose, RC. 2007. Human papillomavirus-like particles mediate functional delivery of plasmid DNA to antigen presenting cells *in vivo*. *Vaccine* 25(17):3270-3276.
- Maryanoff, BE. 2004. Inhibitors of serine proteases as potential therapeutic agents: the road from thrombin to tryptase to cathepsin G. *Journal of Medicinal Chemistry* 47(4):769-787.
- Maurer, P, Jennings, GT, Willers, J, Rohner, F, Lindman, Y, Roubicek, K, Renner, WA, Müller, P and Bachmann, MF. 2005. A therapeutic vaccine for nicotine dependence: preclinical efficacy, and Phase I safety and immunogenicity. *European Journal of Immunology* 35(7):2031-2040.
- Mayer, MP, Rüdiger, S and Bukau, B. 2000. Molecular basis for interactions of the DnaK chaperone with substrates. *Biological Chemistry* 381(9-10):877-885.
- McCaughey, C. 2010. Influenza: a virus of our times. *The Ulster Medical Journal* 79(2):46-51.
- McGinnes, LW, Gravel, KA, Finberg, RW, Kurt-Jones, EA, Massare, MJ, Smith, G, Schmidt, MR and Morrison, TG. 2011. Assembly and immunological properties of

- Newcastle disease virus-like particles containing the respiratory syncytial virus F and G proteins. *Journal of Virology* 85(1):366-377.
- McLellan, JS, Correia, BE, Chen, M, Yang, Y, Graham, BS, Schief, WR and Kwong, PD. 2011. Design and characterization of epitope-scaffold immunogens that present the motavizumab epitope from respiratory syncytial virus. *Journal of Molecular Biology* 409(5):853-866.
- McMillan Jr, WG and Mayer, JE. 2004. The statistical thermodynamics of multicomponent systems. *The Journal of Chemical Physics* 13(7):276-305.
- Medina, RA and García-Sastre, A. 2011. Influenza A viruses: new research developments. *Nature Reviews Microbiology* 9(8):590-603.
- Medina, RA, Stertz, S, Manicassamy, B, Zimmermann, P, Sun, X, Albrecht, RA, Uusi-Kerttula, H, Zagordi, O, Belshe, RB, Frey, SE, Tumpey, TM and García-Sastre, A. 2013. Glycosylations in the globular head of the hemagglutinin protein modulate the virulence and antigenic properties of the H1N1 influenza viruses. *Science Translational Medicine* 5(187):art. no. 187ra170.
- Melander, W and Horvath, C. 1977. Salt effects on hydrophobic interactions in precipitation and chromatography of proteins: an interpretation of the lyotropic series. *Archives of Biochemistry and Biophysics* 183(1):200-215.
- Melander, W, Corradini, D and Horváth, C. 1984. Salt-mediated retention of proteins in hydrophobic-interaction chromatography: application of solvophobic theory. *Journal of Chromatography A* 317:67-85.
- Mena, JA and Kamen, AA. 2011. Insect cell technology is a versatile and robust vaccine manufacturing platform. *Expert Review of Vaccines* 10(7):1063-1081.
- Middelberg, AP, Rivera-Hernandez, T, Wibowo, N, Lua, LH, Fan, Y, Magor, G, Chang, C, Chuan, YP, Good, MF and Batzloff, MR. 2011. A microbial platform for rapid and low-cost virus-like particle and capsomere vaccines. *Vaccine* 29(41):7154-7162.
- Middelberg, APJ, Rivera-Hernandez, T, Wibowo, N, Lua, LHL, Fan, Y, Magor, G, Chang, C, Chuan, YP, Good, MF and Batzloff, MR. 2011. A microbial platform for rapid and low-cost virus-like particle and capsomere vaccines. *Vaccine* 29(41):7154-7162.
- Miklos, AE, Kluwe, C, Der, BS, Pai, S, Sircar, A, Hughes, RA, Berrondo, M, Xu, J, Codrea, V and Buckley, PE. 2012. Structure-based design of supercharged, highly thermoresistant antibodies. *Chemistry and Biology* 19(4):449-455.
- Mohr, J, Chuan, YP, Wu, Y, Lua, LH and Middelberg, AP. 2013. Virus-like particle formulation optimization by miniaturized high-throughput screening. *Methods* 60(3):248-256.

- Momany, C, Kovari, LC, Prongay, AJ, Keller, W, Gitti, RK, Lee, BM, Gorbalenya, AE, Tong, L, McClure, J and Ehrlich, LS. 1996. Crystal structure of dimeric HIV-1 capsid protein. *Nature Structural and Molecular Biology* 3(9):763-770.
- Mosavi, LK and Peng, Zy. 2003. Structure - based substitutions for increased solubility of a designed protein. *Protein Engineering* 16(10):739-745.
- Mubareka, S and Palese, P. 2011. Influenza Vaccines for the Future. In: Rappuoli, R and Del Giudice, G, editor. *Influenza virus: The biology of a changing virus*: Springer. p. 3-26.
- Murakami, S, Horimoto, T, Nidom, CA, Chen, H, Muramoto, Y, Yamada, S, Iwasa, A, Iwatsuki-Horimoto, K, Shimojima, M and Iwata, A. 2008. Growth determinants for H5N1 influenza vaccine seed viruses in MDCK cells. *Journal of Virology* 82(21):10502-10509.
- Murawski, MR, McGinnes, LW, Finberg, RW, Kurt-Jones, EA, Massare, MJ, Smith, G, Heaton, PM, Fraire, AE and Morrison, TG. 2010. Newcastle disease virus-like particles containing respiratory syncytial virus G protein induced protection in BALB/c mice, with no evidence of immunopathology. *Journal of Virology* 84(2):1110-1123.
- Murphy, BR and Coelingh, K. 2002. Principles underlying the development and use of live attenuated cold-adapted influenza A and B virus vaccines. *Viral Immunology* 15(2):295-323.
- Nardin, EH, Oliveira, GA, Calvo-Calle, JM, Wetzel, K, Maier, C, Birkett, AJ, Sarpotdar, P, Corado, ML, Thornton, GB and Schmidt, A. 2004. Phase I testing of a malaria vaccine composed of hepatitis B virus core particles expressing *Plasmodium falciparum* circumsporozoite epitopes. *Infection and Immunity* 72(11):6519-6527.
- Neiryneck, S, Deroo, T, Saelens, X, Vanlandschoot, P, Jou, WM and Fiers, W. 1999. A universal influenza A vaccine based on the extracellular domain of the M2 protein. *Nature Medicine* 5(10):1157-1163.
- Neumann, G, Noda, T and Kawaoka, Y. 2009. Emergence and pandemic potential of swine-origin H1N1 influenza virus. *Nature* 459(7249):931-939.
- Neurath, A, Rubin, B, Sillaman, J and Tint, H. 1971. The effect of nonaqueous solvents on the quaternary structure of viruses: a procedure for the simultaneous concentration, purification and disruption of influenza viruses. *Microbios* 4(14):145-150.
- Ninham, B. 1999. On progress in forces since the DLVO theory. *Advances in Colloid and Interface Science* 83(1):1-17.

- Noad, R and Roy, P. 2003. Virus-like particles as immunogens. *Trends in Microbiology* 11(9):438-444.
- Ofek, G, Tang, M, Sambor, A, Katinger, H, Mascola, JR, Wyatt, R and Kwong, PD. 2004. Structure and mechanistic analysis of the anti-human immunodeficiency virus type 1 antibody 2F5 in complex with its gp41 epitope. *Journal of Virology* 78(19):10724-10737.
- Ofek, G, Guenaga, FJ, Schief, WR, Skinner, J, Baker, D, Wyatt, R and Kwong, PD. 2010. Elicitation of structure-specific antibodies by epitope scaffolds. *Proceedings of the National Academy of Sciences* 107(42):17880-17887.
- Ohuchi, R, Ohuchi, M, Garten, W and Klenk, H-D. 1997. Oligosaccharides in the stem region maintain the influenza virus hemagglutinin in the metastable form required for fusion activity. *Journal of Virology* 71(5):3719-3725.
- Oliner, JD, Kinzler, KW and Vogelstein, B. 1993. *In vivo* cloning of PCR products in *E. coli*. *Nucleic Acids Research* 21(22):5192-5197.
- Olsen, B, Munster, VJ, Wallensten, A, Waldenström, J, Osterhaus, ADME and Fouchier, RAM. 2006. Global patterns of influenza A virus in wild birds. *Science* 312(5772):384-388.
- Oscarsson, S. 1995. Influence of salts on protein interactions at interfaces of amphiphilic polymers and adsorbents. *Journal of Chromatography B: Biomedical Sciences and Applications* 666(1):21-31.
- Otzen, DE, Kristensen, O and Oliveberg, M. 2000. Designed protein tetramer zipped together with a hydrophobic Alzheimer homology: a structural clue to amyloid assembly. *Proceedings of the National Academy of Sciences* 97(18):9907-9912.
- Outhwaite, CW and Bhuiyan, LB. 1983. An improved modified Poisson–Boltzmann equation in electric-double-layer theory. *Journal of the Chemical Society, Faraday Transactions 2: Molecular and Chemical Physics* 79(5):707-718.
- Palese, P. 2004. Influenza: old and new threats. *Nature Medicine* 10(12 SUPPL.):S82-S87.
- Parida, S. 2009. Vaccination against foot-and-mouth disease virus: strategies and effectiveness. *Expert Review of Vaccines* 8(3):347-365.
- Parrish, JR, Limjindaporn, T, Hines, JA, Liu, J, Liu, G and Finley, RL. 2004. High-throughput cloning of *Campylobacter jejuni* ORFs by *in vivo* recombination in *Escherichia coli*. *Journal of Proteome Research* 3(3):582-586.

- Pattenden, L, Middelberg, A, Niebert, M and Lipin, D. 2005. Towards the preparative and large-scale precision manufacture of virus-like particles. *Trends in Biotechnology* 23(10):523 - 529.
- Pau, M, Ophorst, C, Koldijk, MH, Schouten, G, Mehtali, M and Uytdehaag, F. 2001. The human cell line PER. C6 provides a new manufacturing system for the production of influenza vaccines. *Vaccine* 19(17):2716-2721.
- Peng, P, Menoret, A and Srivastava, PK. 1997. Purification of immunogenic heat shock protein 70-peptide complexes by ADP-affinity chromatography. *Journal of Immunological Methods* 204(1):13-21.
- Perchiacca, JM, Ladiwala, ARA, Bhattacharya, M and Tessier, PM. 2012. Aggregation-resistant domain antibodies engineered with charged mutations near the edges of the complementarity-determining regions. *Protein Engineering Design and Selection* 25(10):591-602.
- Perdue, ML, Arnold, F, Li, S, Donabedian, A, Cioce, V, Warf, T and Huebner, R. 2011. The future of cell culture-based influenza vaccine production. *Expert Review of Vaccines* 10(8):1183-1194.
- Perkins, TW, Mak, DS, Root, TW and Lightfoot, EN. 1997. Protein retention in hydrophobic interaction chromatography: modeling variation with buffer ionic strength and column hydrophobicity. *Journal of Chromatography A* 766(1):1-14.
- Petrovsky, N, Larena, M, Siddharthan, V, Prow, NA, Hall, RA, Lobigs, M and Morrey, J. 2013. An inactivated cell culture Japanese encephalitis vaccine (JE-ADVAX) formulated with delta inulin adjuvant provides robust heterologous protection against west nile encephalitis via cross-protective memory B cells and neutralizing antibody. *Journal of Virology* 87(18):10324-10333.
- Petterson, EF, Goddard, TD, Huang, CC, Couch, GS, Greenblatt, DM, Meng, EC and Ferrin, TE. 2004. UCSF Chimera—A visualization system for exploratory research and analysis. *Journal of Computational Chemistry* 25(13):1605-1612.
- Plotkin, SA and Plotkin, SL. 2011. The development of vaccines: how the past led to the future. *Nature Reviews Microbiology* 9(12):889-893.
- Poland, GA and Jacobson, RM. 2004. Prevention of hepatitis B with the hepatitis B vaccine. *New England Journal of Medicine* 351(27):2832-2838.
- Potter, CW. 2001. A history of influenza. *Journal of Applied Microbiology* 91(4):572-579.
- Prasad, BV, Hardy, ME, Dokland, T, Bella, J, Rossmann, MG and Estes, MK. 1999. X-ray crystallographic structure of the Norwalk virus capsid. *Science* 286(5438):287-290.

- Prausnitz, JM. 2003. Molecular thermodynamics for some applications in biotechnology. *The Journal of Chemical Thermodynamics* 35(1):21-39.
- Prel, A, Le Gall-Recule, G and Jestin, V. 2008. Achievement of avian influenza virus-like particles that could be used as a subunit vaccine against low-pathogenic avian influenza strains in ducks. *Avian Pathology* 37(5):513-520.
- Purcell, AW, Zeng, W, Mifsud, NA, Ely, LK, Macdonald, WA and Jackson, DC. 2003. Dissecting the role of peptides in the immune response: theory, practice and the application to vaccine design. *Journal of Peptide Science* 9(5):255-281.
- Purcell, AW, McCluskey, J and Rossjohn, J. 2007. More than one reason to rethink the use of peptides in vaccine design. *Nature Reviews Drug Discovery* 6(5):404-414.
- Pusey, ML, Liu, ZJ, Tempel, W, Praissman, J, Lin, D, Wang, BC, Gavira, JA and Ng, JD. 2005. Life in the fast lane for protein crystallization and X-ray crystallography. *Progress in Biophysics and Molecular Biology* 88(3):359-386.
- Qing, G, Ma, L-C, Khorchid, A, Swapna, G, Mal, TK, Takayama, MM, Xia, B, Phadtare, S, Ke, H and Acton, T. 2004. Cold-shock induced high-yield protein production in *Escherichia coli*. *Nature Biotechnology* 22(7):877-882.
- Quan, F-S, Huang, C, Compans, RW and Kang, S-M. 2007. Virus-like particle vaccine induces protective immunity against homologous and heterologous strains of influenza virus. *Journal of Virology* 81(7):3514-3524.
- Queiroz, J, Tomaz, C and Cabral, J. 2001. Hydrophobic interaction chromatography of proteins. *Journal of Biotechnology* 87(2):143-159.
- Rappuoli, R. 2000. Reverse vaccinology. *Current Opinion in Microbiology* 3(5):445-450.
- Rappuoli, R. 2007. Bridging the knowledge gaps in vaccine design. *Nature Biotechnology* 25(12):1361-1366.
- Reddington, JJ, Reddington, GM and MacLachlan, NJ. 1991. A competitive ELISA for detection of antibodies to the group antigen of bluetongue virus. *Journal of Veterinary Diagnostic Investigation* 3(2):144-147.
- Reddy, VS, Natarajan, P, Okerberg, B, Li, K, Damodaran, K, Morton, RT, Brooks, CL and Johnson, JE. 2001. Virus Particle Explorer (VIPER), a website for virus capsid structures and their computational analyses. *Journal of Virology* 75(24):11943-11947.
- Reichert, T, Chowell, G, Nishiura, H, Christensen, RA and McCullers, JA. 2010. Does glycosylation as a modifier of original antigenic sin explain the case age distribution and unusual toxicity in pandemic novel H1N1 influenza? *BMC Infectious Diseases* 10(1):5.

- Relf, WA, Cooper, J, Brandt, ER, Hayman, WA, Anders, RF, Pruksakorn, S, Currie, B, Saul, A and Good, MF. 1996. Mapping a conserved conformational epitope from the M protein of group A *Streptococci*. *Peptide Research* 9(1):12-20.
- Rial, DV and Ceccarelli, EA. 2002. Removal of DnaK contamination during fusion protein purifications. *Protein Expression and Purification* 25(3):503-507.
- Rivera-Hernandez, T, Hartas, J, Wu, Y, Chuan, YP, Lua, LHL, Good, M, Batzloff, MR and Middelberg, APJ. 2013. Self-adjuvanting modular virus-like particles for mucosal vaccination against group A *streptococcus* (GAS). *Vaccine* 31(15):1950-1955.
- Robbins, JB and Schneerson, R. 1990. Polysaccharide-protein conjugates: a new generation of vaccines. *The Journal of Infectious Diseases* 821-832.
- Robertson, J. 1993. Clinical influenza virus and the embryonated hen's egg. *Reviews in Medical Virology* 3(2):97-106.
- Robertson, JS and Engelhardt, OG. 2010. Developing vaccines to combat pandemic influenza. *Viruses* 2(2):532-546.
- Roettger, BF, Myers, JA, Ladisch, MR and Regnier, FE. 1989. Adsorption phenomena in hydrophobic interaction chromatography. *Biotechnology Progress* 5(3):79-88.
- Roldão, A, Mellado, MCM, Castilho, LR, Carrondo, MJT and Alves, PM. 2010. Virus-like particles in vaccine development. *Expert Review of Vaccines* 9(10):1149-1176.
- Rüdiger, S, Germeroth, L, Schneider - Mergener, J and Bukau, B. 1997. Substrate specificity of the DnaK chaperone determined by screening cellulose - bound peptide libraries. *The EMBO journal* 16(7):1501-1507.
- Rueda, P, Morón, G, Sarraseca, J, Leclerc, C and Casal, JI. 2004. Influence of flanking sequences on presentation efficiency of a CD8⁺ cytotoxic T-cell epitope delivered by parvovirus-like particles. *Journal of General Virology* 85(3):563-572.
- Russell, R, Gamblin, S, Haire, L, Stevens, D, Xiao, B, Ha, Y and Skehel, J. 2004. H1 and H7 influenza haemagglutinin structures extend a structural classification of haemagglutinin subtypes. *Virology* 325(2):287-296.
- Russell, RJ, Stevens, DJ, Haire, LF, Gamblin, SJ and Skehel, JJ. 2006. Avian and human receptor binding by hemagglutinins of influenza A viruses. *Glycoconjugate Journal* 23(1-2):85-92.
- Saade, F, Honda-Okubo, Y, Trec, S and Petrovsky, N. 2013. A novel hepatitis B vaccine containing Advax™, a polysaccharide adjuvant derived from delta inulin, induces robust humoral and cellular immunity with minimal reactogenicity in preclinical testing. *Vaccine* 31(15):1999-2007.

- Sabisky, E and Anderson, C. 1973. Verification of the Lifshitz theory of the van der Waals potential using liquid-helium films. *Physical Review A* 7(2):790-806.
- Saibil, H, Dong, Z and Wood, S. 1991. Binding of chaperonins. *Nature* 353(6339):25-26.
- Salunke, D, Caspar, D and Garcea, R. 1989. Polymorphism in the assembly of polyomavirus capsid protein VP1. *Biophysical Journal* 56(5):887-900.
- Salunke, DM, Caspar, DL and Garcea, RL. 1986. Self-assembly of purified polyomavirus capsid protein VP1. *Cell* 46(6):895-904.
- Sambrook, J and Russell, DW. 2001. Plasmids and their usefulness in molecular cloning. In: Janssen, K and Argentine, J, editor. *Molecular cloning: A laboratory manual*, 3rd edition: Cold Spring Harbor Laboratory Press. p. 1.116-1.118.
- Saxinger, WC and Gallo, RC. Competitive ELISA for the detection of HTLV-III antibodies. US 06/737,458. 1987.
- Schädlich, L, Senger, T, Kirschning, CJ, Müller, M and Gissmann, L. 2009. Refining HPV 16 L1 purification from *E. coli*: Reducing endotoxin contaminations and their impact on immunogenicity. *Vaccine* 27(10):1511-1522.
- Scheibel, T, Bloom, J and Lindquist, SL. 2004. The elongation of yeast prion fibers involves separable steps of association and conversion. *Proceedings of the National Academy of Sciences of the United States of America* 101(8):2287-2292.
- Schiller, JT, Castellsagué, X, Villa, LL and Hildesheim, A. 2008. An update of prophylactic human papillomavirus L1 virus-like particle vaccine clinical trial results. *Vaccine* 26(SUPPL. 10):K53-K61.
- Schneemann, A, Speir, JA, Tan, GS, Khayat, R, Ekiert, DC, Matsuoka, Y and Wilson, IA. 2012. A virus-like particle that elicits cross-reactive antibodies to the conserved stem of influenza virus hemagglutinin. *Journal of Virology* 86(21):11686-11697.
- Schoel, B and Kaufmann, SH. 1991. Hydrophobic interaction chromatography for the purification of a mycobacterial heat shock protein of relative molecular mass 60 000. *Journal of Chromatography A* 587(1):19-23.
- Schön, U and Schumann, W. 1995. Overproduction, purification and characterization of GroES and GroEL from thermophilic *Bacillus stearothermophilus*. *FEMS Microbiology Letters* 134(2 - 3):183-188.
- Schuman, JT, Grinstead, JS, Apostolopoulos, V and Campbell, AP. 2005. Structural and dynamic consequences of increasing repeats in a MUC1 peptide tumor antigen. *Biopolymers* 77(2):107-120.
- Selkoe, DJ. 1991. The molecular pathology of Alzheimer's disease. *Neuron* 6(4):487-498.

- Shank-Retzlaff, ML, Zhao, Q, Anderson, C, Hamm, M, High, K, Nguyen, M, Wang, F, Wang, N, Wang, B and Wang, Y. 2006. Evaluation of the thermal stability of Gardasil®. *Human Vaccines* 2(4):147-154.
- Shaw, A. 2012. New technologies for new influenza vaccines. *Vaccine* 30(33):4927-4933.
- Shin, YC and Folk, WR. 2003. Formation of polyomavirus-like particles with different VP1 molecules that bind the urokinase plasminogen activator receptor. *Journal of Virology* 77(21):11491-11498.
- Singleton, A, Farrer, M, Johnson, J, Singleton, A, Hague, S, Kachergus, J, Hulihan, M, Peuralinna, T, Dutra, A and Nussbaum, R. 2003. α -Synuclein locus triplication causes Parkinson's disease. *Science* 302(5646):841-841.
- Skehel, JJ and Wiley, DC. 2000. Receptor binding and membrane fusion in virus entry: the influenza hemagglutinin. *Annual Review of Biochemistry* 69:531-569.
- Skouri, M, Lorber, B, Giegé, R, Munch, J-P and Candau, J. 1995. Effect of macromolecular impurities on lysozyme solubility and crystallizability: dynamic light scattering, phase diagram, and crystal growth studies. *Journal of Crystal Growth* 152(3):209-220.
- Smith, MT, Bennett, AM, Grubman, MJ and Bundy, BC. 2014. Foot-and-mouth disease: Technical and political challenges to eradication. *Vaccine* 32(31):3902-3908.
- Smyth, M and Martin, J. 2000. X-Ray crystallography. *Molecular Pathology* 53(1):8-14.
- Song, J-M, Wang, B-Z, Park, K-M, Van Rooijen, N, Quan, F-S, Kim, M-C, Jin, H-T, Pekosz, A, Compans, RW and Kang, S-M. 2011. Influenza virus-like particles containing M2 induce broadly cross protective immunity. *PLoS ONE* 6(1):e14538.
- Sørensen, HP and Mortensen, KK. 2005. Soluble expression of recombinant proteins in the cytoplasm of *Escherichia coli*. *Microbial Cell Factories* 4(1):1-8.
- Speir, JA, Munshi, S, Wang, G, Baker, TS and Johnson, JE. 1995. Structures of the native and swollen forms of cowpea chlorotic mottle virus determined by X-ray crystallography and cryo-electron microscopy. *Structure* 3(1):63-78.
- Spohn, G and Bachmann, MF. 2008. Exploiting viral properties for the rational design of modern vaccines. *Expert Review of Vaccines* 7(1):43-54.
- Steel, J, Lowen, AC, Wang, TT, Yondola, M, Gao, Q, Haye, K, García-Sastre, A and Palese, P. 2010. Influenza virus vaccine based on the conserved hemagglutinin stalk domain. *mBio* 1(1):e00018-00010.
- Stehle, T, Yan, Y, Benjamin, TL and Harrison, SC. 1994. Structure of murine polyomavirus complexed with an oligosaccharide receptor fragment. *Nature* 369(6476):160 - 163.

- Stehle, T, Gamblin, SJ, Yan, Y and Harrison, SC. 1996. The structure of simian virus 40 refined at 3.1 Å resolution. *Structure* 4(2):165-182.
- Stehle, T and Harrison, SC. 1996. Crystal structures of murine polyomavirus in complex with straight-chain and branched-chain sialyloligosaccharide receptor fragments. *Structure* 4(2):183-194.
- Stehle, T and Harrison, SC. 1997. High-resolution structure of a polyomavirus VP1-oligosaccharide complex: implications for assembly and receptor binding. *The EMBO journal* 16(16):5139-5148.
- Stevens, J, Blixt, O, Tumpey, TM, Taubenberger, JK, Paulson, JC and Wilson, IA. 2006. Structure and receptor specificity of the hemagglutinin from an H5N1 influenza virus. *Science* 312(5772):404-410.
- Stockmann, C, Scheidle, M, Dittrich, B, Merckelbach, A, Hehmann, G, Melmer, G, Klee, D, Buchs, J, Kang, H and Gellissen, G. 2009. Process development in *Hansenula polymorpha* and *Arxula adenivorans*, a re-assessment. *Microbial Cell Factories* 8:art. no. 22.
- Stubenrauch, K, Bachmann, A, Rudolph, R and Lilie, H. 2000. Purification of a viral coat protein by an engineered polyionic sequence. *Journal of Chromatography B: Biomedical Sciences and Applications* 737(1):77-84.
- Subbarao, K and Joseph, T. 2007. Scientific barriers to developing vaccines against avian influenza viruses. *Nature Reviews Immunology* 7(4):267-278.
- Sun, S, Wang, Q, Zhao, F, Chen, W and Li, Z. 2011. Glycosylation site alteration in the evolution of influenza A (H1N1) viruses. *PLoS ONE* 6(7):e22844.
- Szécsi, J, Boson, B, Johnsson, P, Dupeyrot-Lacas, P, Matrosovich, M, Klenk, H-D, Klatzmann, D, Volchkov, V and Cosset, F-L. 2006. Induction of neutralising antibodies by virus-like particles harbouring surface proteins from highly pathogenic H5N1 and H7N1 influenza viruses. *Virology Journal* 3(70):1-7.
- Takamura, S, Niikura, M, Li, TC, Takeda, N, Kusagawa, S, Takebe, Y, Miyamura, T and Yasutomi, Y. 2004. DNA vaccine-encapsulated virus-like particles derived from an orally transmissible virus stimulate mucosal and systemic immune responses by oral administration. *Gene Therapy* 11(7):628-635.
- Tang, X-C, Lu, H-R and Ross, TM. 2011. Baculovirus-produced influenza virus-like particles in mammalian cells protect mice from lethal influenza challenge. *Viral Immunology* 24(4):311-319.

- Thomas, S and Luxon, BA. 2013. Vaccines based on structure-based design provide protection against infectious diseases. *Expert Review of Vaccines* 12(11):1301-1311.
- Tissot, AC, Renhofa, R, Schmitz, N, Cielens, I, Meijerink, E, Ose, V, Jennings, GT, Saudan, P, Pumpens, P and Bachmann, MF. 2010. Versatile virus-like particle carrier for epitope based vaccines. *PLoS ONE* 5(3):e9809.
- Totrov, M, Jiang, X, Kong, X-P, Cohen, S, Krachmarov, C, Salomon, A, Williams, C, Seaman, MS, Cardozo, T and Gorny, MK. 2010. Structure-guided design and immunological characterization of immunogens presenting the HIV-1 gp120 V3 loop on a CTB scaffold. *Virology* 405(2):513-523.
- Touze, A and Coursaget, P. 1998. *In vitro* gene transfer using human papillomavirus-like particles. *Nucleic Acids Research* 26(5):1317-1323.
- Treanor, J. 2004. Weathering the influenza vaccine crisis. *New England Journal of Medicine* 351(20):2037-2040.
- Tree, JA, Richardson, C, Fooks, AR, Clegg, JC and Looby, D. 2001. Comparison of large-scale mammalian cell culture systems with egg culture for the production of influenza virus A vaccine strains. *Vaccine* 19(25-26):3444-3450.
- Tsuchiya, E, Sugawara, K, Hongo, S, Matsuzaki, Y, Muraki, Y, Li, Z-N and Nakamura, K. 2002. Effect of addition of new oligosaccharide chains to the globular head of influenza A/H2N2 virus haemagglutinin on the intracellular transport and biological activities of the molecule. *Journal of General Virology* 83(5):1137-1146.
- Tsuchiya, E, Sugawara, K, Hongo, S, Matsuzaki, Y, Muraki, Y and Nakamura, K. 2002. Role of overlapping glycosylation sequons in antigenic properties, intracellular transport and biological activities of influenza A/H2N2 virus haemagglutinin. *Journal of General Virology* 83(12):3067-3074.
- Tugyi, R, Mezö, G, Fellinger, E, Andreu, D and Hudecz, F. 2005. The effect of cyclization on the enzymatic degradation of herpes simplex virus glycoprotein D derived epitope peptide. *Journal of Peptide Science* 11(10):642-649.
- Ulmer, JB, Valley, U and Rappuoli, R. 2006. Vaccine manufacturing: challenges and solutions. *Nature Biotechnology* 24(11):1377-1383.
- Van Der Spoel, D, Lindahl, E, Hess, B, Groenhof, G, Mark, AE and Berendsen, HJC. 2005. GROMACS: fast, flexible, and free. *Journal of Computational Chemistry* 26(16):1701-1718.

- Velev, O, Kaler, E and Lenhoff, A. 1998. Protein interactions in solution characterized by light and neutron scattering: comparison of lysozyme and chymotrypsinogen. *Biophysical Journal* 75(6):2682-2697.
- Ventura, S, Zurdo, J, Narayanan, S, Parreño, M, Mangues, R, Reif, B, Chiti, F, Giannoni, E, Dobson, CM and Aviles, FX. 2004. Short amino acid stretches can mediate amyloid formation in globular proteins: the Src homology 3 (SH3) case. *Proceedings of the National Academy of Sciences of the United States of America* 101(19):7258-7263.
- Verwey, EJW and Overbeek, JTG. 1947. Theory of the stability of lyophobic colloids. *The Journal of Physical Chemistry B* 51(3):631-636.
- Vigerust, DJ and Shepherd, VL. 2007. Virus glycosylation: role in virulence and immune interactions. *Trends in Microbiology* 15(5):211-218.
- Villa, LL, Costa, RL, Petta, CA, Andrade, RP, Ault, KA, Giuliano, AR, Wheeler, CM, Koutsky, LA, Malm, C and Lehtinen, M. 2005. Prophylactic quadrivalent human papillomavirus (types 6, 11, 16, and 18) L1 virus-like particle vaccine in young women: a randomised double-blind placebo-controlled multicentre phase II efficacy trial. *The Lancet Oncology* 6(5):271-278.
- Viscidi, RP and Bossis, I. Polyionic papilloma virus-like particle (VLP) vaccines. US 13/821,579. 2014.
- Vogel, M, Diez, M, Eisfeld, J and Nassal, M. 2005. In vitro assembly of mosaic hepatitis B virus capsid-like particles (CLPs): rescue into CLPs of assembly-deficient core protein fusions and FRET-suited CLPs. *FEBS Letters* 579(23):5211-5216.
- Wahlund, K-G. 2000. Asymmetrical Flow Field-Flow Fractionation. In: Schimpf, M, Caldwell, K and Giddings, J C, editor. *Field flow fractionation Handbook*: Wiley-Interscience, Inc. p. 279-294.
- Wahome, N, Pfeiffer, T, Ambiel, I, Yang, Y, Keppler, OT, Bosch, V and Burkhard, P. 2012. Conformation-specific Display of 4E10 and 2F5 Epitopes on Self-assembling Protein Nanoparticles as a Potential HIV Vaccine. *Chemical Biology and Drug Design* 80(3):349-357.
- Wang, B-Z, Quan, F-S, Kang, S-M, Bozja, J, Skountzou, I and Compans, RW. 2008. Incorporation of membrane-anchored flagellin into influenza virus-like particles enhances the breadth of immune responses. *Journal of Virology* 82(23):11813-11823.
- Wang, C-C, Chen, J-R, Tseng, Y-C, Hsu, C-H, Hung, Y-F, Chen, S-W, Chen, C-M, Khoo, K-H, Cheng, T-J, Cheng, Y-SE, Jan, J-T, Wu, C-Y, Ma, C and Wong, C-H. 2009.

- Glycans on influenza hemagglutinin affect receptor binding and immune response. *Proceedings of the National Academy of Sciences of the United States of America* 106(43):18137-18142.
- Wang, K, Holtz, KM, Anderson, K, Chubet, R, Mahmoud, W and Cox, MMJ. 2006. Expression and purification of an influenza hemagglutinin—one step closer to a recombinant protein-based influenza vaccine. *Vaccine* 24(12):2176-2185.
- Wanzeck, K, Boyd, KL and McCullers, JA. 2011. Glycan shielding of the influenza virus hemagglutinin contributes to immunopathology in mice. *American Journal of Respiratory and Critical Care Medicine* 183(6):767-773.
- Weber, J, Cheinsong-Popov, R, Callow, D, Adams, S, Patou, G, Hodgkin, K, Martin, S, Gotch, F and Kingsman, A. 1995. Immunogenicity of the yeast recombinant p17p24: Ty virus-like particles (p24-VLP) in healthy volunteers. *Vaccine* 13(9):831-834.
- Webster, RG, Bean, WJ, Gorman, OT, Chambers, TM and Kawaoka, Y. 1992. Evolution and ecology of influenza A viruses. *Microbiological Reviews* 56(1):152-179.
- Weis, W, Brown, J, Cusack, S, Paulson, J, Skehel, J and Wiley, D. 1988. Structure of the influenza virus haemagglutinin complexed with its receptor, sialic acid. *Nature* 333(6172):426-431.
- Wibowo, N, Chuan, YP, Lua, LHL and Middelberg, APJ. 2013. Modular engineering of a microbially-produced viral capsomere vaccine for influenza. *Chemical Engineering Science* 103:12-20.
- Wibowo, N, Hughes, FK, Fairmaid, EJ, Lua, LH, Brown, LE and Middelberg, AP. 2014. Protective efficacy of a bacterially produced modular capsomere presenting M2e from influenza: Extending the potential of broadly cross-protecting epitopes. *Vaccine* 32(29):3651-3655.
- Wiley, D, Wilson, I and Skehel, J. 1981. Structural identification of the antibody-binding sites of Hong Kong influenza haemagglutinin and their involvement in antigenic variation. *Nature* 289(5796):373-378.
- Wilson, I, Skehel, J and Wiley, D. 1981. Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 Å resolution. *Nature* 289(5796):366-373.
- Wreghitt, TG, Hicks, J, Gray, JJ and O'Connor, C. 1986. Development of a competitive enzyme-linked immunosorbent assay for detecting cytomegalovirus antibody. *Journal of Medical Virology* 18(2):119-129.

- Wu, C-F, Chen, W-Y and Lee, J-F. 1996. Microcalorimetric studies of the interactions of imidazole with immobilized Cu (II): effects of pH value and salt concentration. *Journal of Colloid and Interface Science* 183(1):236-242.
- Wu, S-J, Luo, J, O'Neil, KT, Kang, J, Lacy, ER, Canziani, G, Baker, A, Huang, M, Tang, QM and Raju, TS. 2010. Structure-based engineering of a monoclonal antibody for improved solubility. *Protein Engineering Design and Selection* 23(8):643-651.
- Wu, T, Li, S-W, Zhang, J, Ng, M-H, Xia, N-S and Zhao, Q. 2012. Hepatitis E vaccine development: a 14 year odyssey. *Human Vaccines and Immunotherapeutics* 8(6):823-827.
- Xia, F, Nagrath, D, Garde, S and Cramer, SM. 2004. Evaluation of selectivity changes in HIC systems using a preferential interaction based analysis. *Biotechnology and Bioengineering* 87(3):354-363.
- Xia, M, Tan, M, Wei, C, Zhong, W, Wang, L, McNeal, M and Jiang, X. 2011. A candidate dual vaccine against influenza and noroviruses. *Vaccine* 29(44):7670-7677.
- Xu, L, Shan, S and Wang, X. 2013. Single point mutation alters the microstate dynamics of amyloid β -protein A β_{42} as revealed by dihedral dynamics analyses. *The Journal of Physical Chemistry B* 117(20):6206-6216.
- Xu, R, Ekiert, DC, Krause, JC, Hai, R, Crowe, JE and Wilson, IA. 2010. Structural basis of preexisting immunity to the 2009 H1N1 pandemic influenza virus. *Science* 328(5976):357-360.
- Xu, Z, Horwich, AL and Sigler, PB. 1997. The crystal structure of the asymmetric GroEL–GroES–(ADP) 7 chaperonin complex. *Nature* 388(6644):741-750.
- Yamada, S, Suzuki, Y, Suzuki, T, Le, MQ, Nidom, CA, Sakai-Tagawa, Y, Muramoto, Y, Ito, M, Kiso, M and Horimoto, T. 2006. Haemagglutinin mutations responsible for the binding of H5N1 influenza A viruses to human-type receptors. *Nature* 444(7117):378-382.
- Yamniuk, AP, Ditto, N, Patel, M, Dai, J, Sejwal, P, Stetsko, P and Doyle, ML. 2013. Application of a kosmotrope - based solubility assay to multiple protein therapeutic classes indicates broad use as a high - throughput screen for protein therapeutic aggregation propensity. *Journal of Pharmaceutical Sciences* 102(8):2424-2439.
- Yang, A-S and Honig, B. 1993. On the pH dependence of protein stability. *Journal of Molecular Biology* 231(2):459-474.
- Yang, Z-Y, Wei, C-J, Kong, W-P, Wu, L, Xu, L, Smith, DF and Nabel, GJ. 2007. Immunization by avian H5 influenza hemagglutinin mutants with altered receptor binding specificity. *Science* 317(5839):825-828.

- Yin, Y, Li, H, Wu, S, Dong, D, Zhang, J, Fu, L, Xu, J and Chen, W. 2011. Hepatitis B virus core particles displaying *Mycobacterium tuberculosis* antigen ESAT-6 enhance ESAT-6-specific immune responses. *Vaccine* 29(34):5645-5651.
- Youil, R, Su, Q, Toner, T, Szymkowiak, C, Kwan, W, Rubin, B, Petrukhin, L, Kiseleva, I, Shaw, A and DiStefano, D. 2004. Comparative study of influenza virus replication in Vero and MDCK cell lines. *Journal of Virological Methods* 120(1):23-31.
- Yu, X, Tsibane, T, McGraw, PA, House, FS, Keefer, CJ, Hicar, MD, Tumpey, TM, Pappas, C, Perrone, LA, Martinez, O, Stevens, J, Wilson, IA, Aguilar, PV, Altschuler, EL, Basler, CF and Jr, JEC. 2008. Neutralizing antibodies derived from the B cells of 1918 influenza pandemic survivors. *Nature* 455(7212):532-536.
- Zeltins, A. 2013. Construction and characterization of virus-like particles: a review. *Molecular Biotechnology* 53(1):92-107.
- Zhang, F, Skoda, MW, Jacobs, RM, Martin, RA, Martin, CM and Schreiber, F. 2007. Protein interactions studied by SAXS: effect of ionic strength and protein concentration for BSA in aqueous solutions. *The Journal of Physical Chemistry B* 111(1):251-259.
- Zhang, W, Carmichael, J, Ferguson, J, Inglis, S, Ashrafian, H and Stanley, M. 1998. Expression of human papillomavirus type 16 L1 protein in *Escherichia coli*: Denaturation, renaturation, and self-assembly of virus-like particles *in vitro*. *Virology* 243(2):423-431.
- Zhang, Y and Cremer, PS. 2006. Interactions between macromolecules and ions: the Hofmeister series. *Current Opinion in Chemical Biology* 10(6):658-663.
- Zheng, B, Graham, FL, Johnson, DC, Hanke, T, McDermott, MR and Prevec, L. 1993. Immunogenicity in mice of tandem repeats of an epitope from herpes simplex gD protein when expressed by recombinant adenovirus vectors. *Vaccine* 11(12):1191-1198.
- Zhu, X, Zhao, X, Burkholder, WF, Gragerov, A, Ogata, CM, Gottesman, ME and Hendrickson, WA. 1996. Structural analysis of substrate binding by the molecular chaperone DnaK. *Science* 272(5268):1606-1614.
- Zlotnick, A, Cheng, N, Conway, J, Booy, F, Steven, A, Stahl, S and Wingfield, P. 1996. Dimorphism of hepatitis B virus capsids is strongly influenced by the C-terminus of the capsid protein. *Biochemistry* 35(23):7412-7421.
- Zylicz, M and Georgopoulos, C. 1984. Purification and properties of the *Escherichia coli* dnaK replication protein. *Journal of Biological Chemistry* 259(14):8820-8825.

Appendix A

Supporting information

SI Material and Methods

Molecular dynamics simulation for structural prediction of H190 element on modular VLPs

Two peptides used in this simulation were:

- (i) GCN4-H190-GCN4 peptide, Ac-VKQLEDKVSTSADQQSLYQNADAYVKQLEDKV-NH₂;
- (ii) H190-H190 peptide, Ac-STSADQQSLYQNADAYSTSADQQSLYQNADAY-NH₂.

Protein data bank ID codes 2WPZ.pdb and 3MLH.pdb were used as structural templates for GCN4 and H190 elements, respectively. Homology models of the peptides were created using Accelrys Discovery Studio® 3.0 (Accelrys, Inc., San Diego, USA) based on the templates. Molecular dynamics (MD) simulations were performed using GROMACS version 3.3.3 (Berendsen et al. 1995, Hess et al. 2008, Lindahl et al. 2001, Van Der Spoel et al. 2005) with Gromos96 43a1 force field for peptides and the simple point charge (SPC) model for water. Each molecular system was first solvated in a cubic box (5nm³ for H190-H190 peptide; 8nm³ for GCN4-H190-GCN4 peptide), with solvent molecules added randomly around the peptides. The system was neutralized by adding Na⁺ and Cl⁻ as counter ions. Simulations were run in a solution containing 137 mM NaCl and SPC model water atoms. Berendson method (Berendsen et al. 1984) was used to control temperature at 298 K with time constant of 0.1 ps and pressure at 1 atm with coupling constant of 1.0 ps. MD algorithm was used with an integration time step of 2 fs. Particle-mesh Ewald (PME) algorithm was used to account for electrostatic interactions. The cutoffs of neighbour atom list, Coulomb potential, and Lennard-Jones (LJ) potential energies were all set to 1.0 nm. The initial velocities of particles were generated according to a Maxwell distribution at 298.15 K. Then, 5000 steps of steepest descent energy minimization was performed, followed by 100 ps equilibration with position restraints on the protein heavy atoms. The MD simulations were then performed for 20 ns, three times for each peptide.

Accelrys Discovery Studio® 3.0 was used to calculate the root-mean-square deviation (RMSD) of H190 element structures. Final conformations of H190 element of the simulated peptides were compared with the conformation of H190 region in the reference homology template structure 3MLH.pdb. RMSD was calculated for C α atoms, main-chain atoms (N, C α , C, and O), side-chain atoms and all heavy atoms (non-hydrogen), and was reported in Angstroms.

Cloning

Plasmid pGEX-VP1 (Middelberg et al. 2011) containing gene encoding murine polyomavirus VP1 protein (M34958) in the commercial pGEX-4T-1 (Novagen, Billerica, MA, USA) vector was generously provided by Professor Robert Garcea (University of Colorado). An Afel site was introduced at position 293 of VP1 by site directed mutagenesis as described elsewhere (Middelberg et al. 2011), giving plasmid pGEX-VP1-S4. Plasmid pGEX-VP1-S1S4 was constructed by introducing NaeI site at position 86 of VP1 of plasmid pGEX-VP1-S4. Plasmid pGEX-VP1-S4-GSGS was constructed by inserting DNA sequence encoding spacers GSGS at the Afel restriction site of plasmid pGEX-VP1-S4. DNA sequences encoding for H190 element (STSADQQSLYQNADAY) from influenza A(H1N1)pdm09 (A/California/07/2009) or the flanking GCN4 element (VKQLEDKV) were codon-optimized for *E. coli*. Complementary oligos for H190 element flanked with GCN4 elements

(5'GTGAAACAGCTGGAAGATAAAGTGAGCACCCAGCGCGGATCAACAGAGCCTGTATCAGAACGCGGATGCGTATGTGAAACAGCTGGAAGATAAAGTG 3'),

and tandem repeats of H190 element

(5'TCTACCTCTGCGGATCAGCAATCTCTGTACCAGAACGCGGACGCCTATAGCACCTCCGCCGACCAGCAGAGCCTGTATCAAAATGCAGACGCGTAC3')

were annealed and phosphorylated using standard protocol (Sambrook and Russell 2001). These oligos were ligated into Afel-linearized pGEX-VP1-S4-GSGS and pGEX-VP1-S1S4, respectively, using a standard conventional ligation method (Sambrook and Russell 2001), yielding plasmids pGEX-VP1-GCN4-H190-GCN4 and pGEX-VP1-H190-H190. DNA sequences of designated constructs VP1-GCN4-H190-GCN4 and VP1-H190-H190 were verified by Australian Genome Research Facility (AGRF, Brisbane, Australia).

Expression, purification and assembly of modular proteins VP1-GCN4-H190-GCN4 and VP1-H190-H190

Plasmids pGEX-VP1, pGEX-VP1-GCN4-H190-GCN4 and pGEX-VP1-H190-H190 were transformed into *E. coli* Rosetta (DE3) pLysS cells (EMD Millipore, Billerica, MA, USA). GST-tagged VP1 protein was expressed, followed by affinity and size-exclusion chromatographic purification yielding VP1 capsomeres (Chuan et al. 2008). Endotoxins were removed from the purified capsomeres using Vivapure Q Mini M spin column (Sartorius Stedim, Goettingen, Germany) as described (Middelberg et al. 2011). In an endotoxin-free condition, capsomeres were assembled into VLPs in Assembly Buffer 1 and the formed VLPs were dialyzed against PBS as described (Middelberg et al. 2011).

Enzymatic removal of N-Glycans from recombinant HA1

The recombinant H1N1 A/California/07/2009 HA1 protein (residue 1-313) was obtained from the Protein Expression Facility, The University of Queensland (Australia). The protein was expressed in *Trichoplusia ni* insect cell. HA1 was deglycosylated with PNGase F (New England Biolabs) or Glycopeptidase A from almonds (PNGase A) (Sigma Aldrich, St. Louis, USA) following the suppliers' instructions. Briefly, 3 µg protein was incubated with 500 U of PNGase F for 2 h at 37°C. For digestion with PNGase A, 3 µg of protein was incubated with 0.25 mU of PNGase A for 16 h at 37°C. The deglycosylated samples were analysed using SDS-PAGE gel electrophoresis, followed by total protein staining using Coomassie Blue (GE Healthcare, Piscataway Township, NJ, USA) and glycoprotein staining using Pierce® Glycoprotein Staining Kit (Thermo Scientific, Waltham, MA, USA) according to manufacturer's recommendation.

Dot blot immunoassay for detecting antigen specific IgG

Proteins or peptides (3 µg) were applied by pipetting onto Amersham Hybond ECL nitrocellulose membranes (GE Healthcare, Piscataway Township, NJ, USA). Membranes were air-dried and then blocked with blocking buffer (PBS, 0.5% v/v Tween 20, 5% w/v milk) for 1 h at 37°C. Sera from 8 mice were pooled. After six 5-min washes with PBST (PBS, 0.5% v/v Tween 20), membranes were incubated with serum 200-fold diluted in PBSTM (PBS, 0.5% v/v Tween 20, 0.5% w/v milk) for 1 h at 37°C. The membranes were washed 6 times (5 min each) with PBST and further incubated with horseradish peroxidase (HRP)-conjugated goat anti-mouse IgG antibodies (Sigma Aldrich, St. Louis, USA) at 10,000-fold dilution in PBSTM for 1 h at 37°C. HRP activity was probed by membrane staining using Novex® Chemiluminescent Substrate (Life Technologies, Carlsbad, CA, USA). After 5-min development in the dark, the stained membrane was placed on the imaging surface of a Molecular Imager® Gel Doc™ XR System (Bio-Rad, Hercules, CA, USA) and digital acquisition of luminescence signal was recorded.

Peptide-based ELISA to determine antibody titers

Biotinylated-synthetic peptides used in ELISA were purchased from Peptide 2.0 (Chantilly, VA, USA). Sequences of the peptides were as follows:

(i) biotin-GGGGSVKQLEDKVSTSDQQLYQNADAYVKQLEDKV-NH₂ (peptide GCN4-H190-GCN4); and (ii) biotin-GGGGSSTSDQQLYQNADAY-NH₂ (peptide H190).

Pierce® Streptavidin High Binding Capacity Coated 96-Well Plates (Thermo Scientific, Waltham, MA, USA) were washed 3 times with wash buffer (25 mM Tris-HCl, 150 mM

NaCl, pH 7.2, 0.1% w/v BSA, 0.05% v/v Tween-20). Biotinylated-synthetic peptide H190 in wash buffer was incubated on plates (1 µg/well) for 2 h at room temperature to allow coupling of the biotinylated-synthetic peptides to the immobilized streptavidin. Plates were then washed 3 times with wash buffer. Sera were serially diluted, each time at 4-fold dilution, in PBSTM and added to the plates (100 µl/well). Plates were incubated at 37°C for 1.5 h. After washing 4 times with PBST, 100 µl of HRP-conjugated goat anti-mouse IgG antibodies (10,000-fold diluted in in PBSTM) was added to each well and plates were incubated at 37°C for 1.5 h. Plates were washed 4 times with PBST and HRP activity was detected using 0.4 mg/ml o-phenylenediamine dihydrochloride (OPD) substrate (Sigma Aldrich, St. Louis, USA) in 50 mM phosphate citrate buffer containing 0.03% sodium perborate (Sigma Aldrich, St. Louis, USA). Plates were incubated in the dark for 20 min. The absorbance of each well was measured at 450 nm with an Infinite® 200 plate reader (Tecan, Durham, NC, USA). Endpoint titer was defined as the lowest dilution that gave an absorbance higher than 3 times the standard deviation above the mean absorbance obtained from wells containing pre-immunized mouse sera (Relf et al. 1996).

HA1-based ELISA to determine antibody titers

HA1-based ELISA was performed as described (Bommakanti et al. 2010) with minor changes. Briefly, HA1 in 20 mM sodium phosphate, 20 mM NaCl, pH 7 was adsorbed onto Nunc MaxiSorp ELISA plates (Thermo Scientific, Waltham, MA, USA) overnight at 4°C (300 ng/well). Plates were blocked with blocking buffer for 1.5 h at 37°C and were then washed twice with PBST. Plate incubation with sera and HRP-conjugated goat anti-mouse IgG antibodies and HRP activity detection were done as described in peptide-based ELISA procedures.

Split virion-based ELISA

Fluvax® vaccine 2011 (CSL Limited, Victoria, Australia) was diluted in PBS giving equivalent total HA ranging from 0 to 9 µg. Fluvax® was adsorbed onto MaxiSorp ELISA plates overnight at 4°C (100 µl/well). Plates were blocked with blocking buffer for 1.5 h at 37°C and then washed twice with PBST. Plate incubation with sera and HRP-conjugated goat anti-mouse IgG antibodies and HRP activity detection were done as described in peptide-based ELISA protocol.

Competitive binding of sera to peptides or HA1

Pierce® Streptavidin High Binding Capacity Coated 96-Well Plates were coated with biotinylated-synthetic peptide GCN4-H190-GCN4 as described in peptide-based ELISA procedures. Sera from 8 mice were pooled and diluted in PBSTM. The pooled sera (50 µl) was mixed with an equal volume of peptide GCN4-H190-GCN4 or peptide H190 as the inhibitor at concentrations of 0, 2.39, 23.9, 50, 100, and 239 µM and incubated for 1.5 h at 37°C. To each well, mixtures of serum-peptide were then added (100 µl/well), and the plates were incubated for 1.5 h at 37°C. Plates were washed 4 times with PBST. Plate incubation with HRP-conjugated goat anti-mouse IgG antibodies and HRP activity detection was done as described in peptide-based ELISA protocol. Competitive binding to HA1 was conducted similarly, except plates were coated with HA1 as described and the pooled sera was mixed only with peptide H190 as the inhibitor at concentrations of 0, 0.008, 0.02, 0.24, 2.4, and 24 µM.

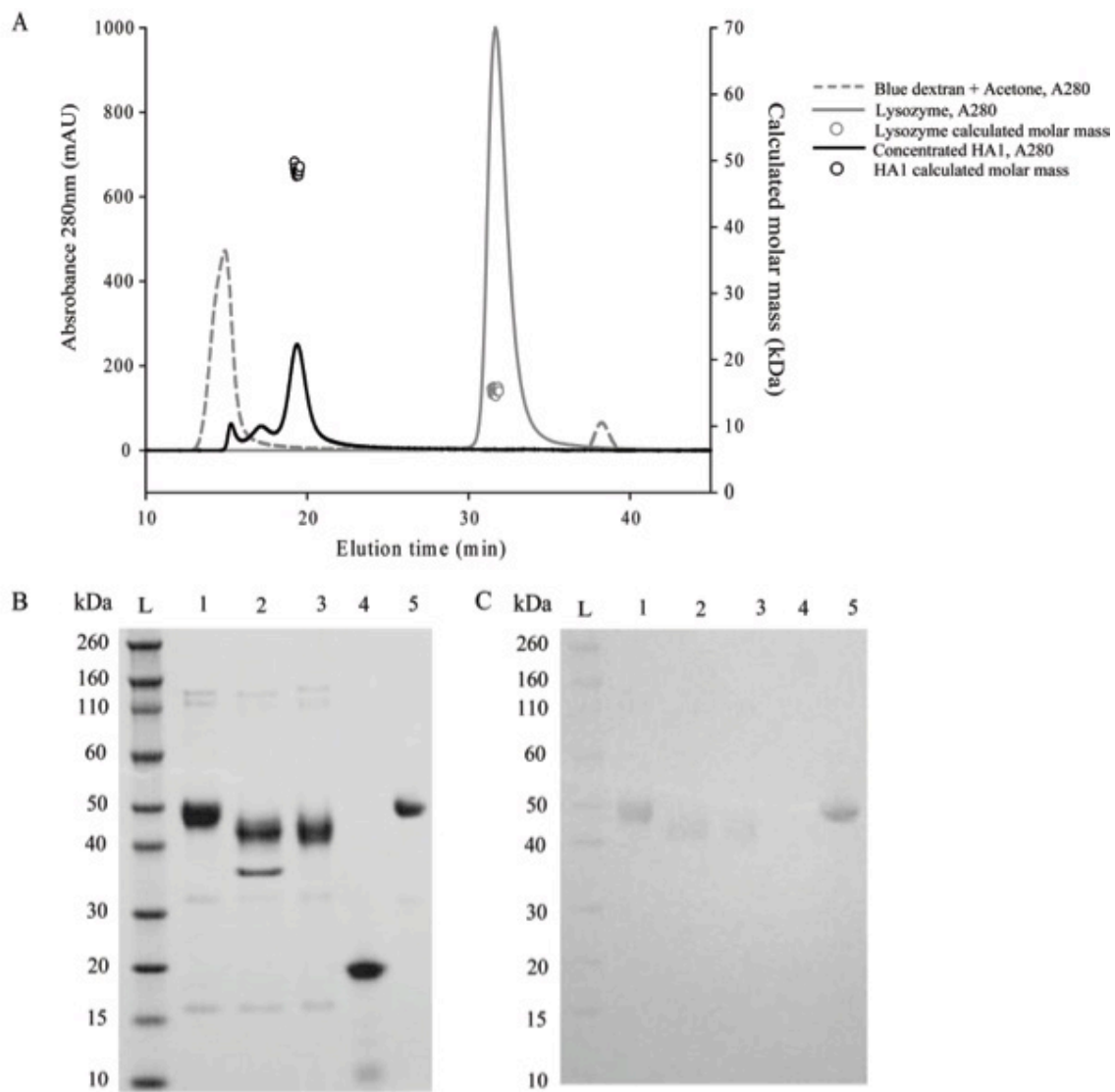


Fig. S1. Characterization of HA1. (A) Quaternary structure analysis using size-exclusion chromatography coupled with multi-angle light scattering in Superdex® 75 100/300 GL (GE Healthcare). Lysozyme, blue dextran and acetone were used as calibration standards. (B) Detection of total proteins using Coomassie blue; and (C) Detection of glycoproteins using Pierce® Glycoprotein Staining Kit following electrophoresis in reduced SDS-PAGE gel. Soybean trypsin inhibitor and horseradish peroxidase (HRP) were included as negative and positive controls, respectively, for glycoprotein staining. Lanes: (L) protein standards; (1) HA1; (2) HA1 treated with PNGase F; (3) HA1 treated with PNGase A; (4) Soybean trypsin inhibitor; and (5) HRP. Each lane contains about 1 µg of HA1 or controls.