THE UNIVERSITY OF QUEENSLAND

AUSTRALIA

Structural presentation of influenza epitopes on modular virus-like particles

Doan Thanh Tam

Master of Biotechnology

A thesis submitted for the degree of Doctor of Philosophy at

The University of Queensland in 2018

Australian Institute for Bioengineering and Nanotechnology

Abstract

Vaccine development has increasingly developed toward subunit vaccine approach which uses highly-purified and well-characterised components, such as antigenic peptides, to elicit an immune response. An antigenic peptide requires a protein carrier or platform e.g. virus-like particle (VLP), to improve its poor immunogenicity. Based on this approach, murine polyomavirus (MuPyV) VLP and its subunit (capsomere) have been employed as vaccine platforms to present heterogeneous epitopes or antigenic elements from targeted pathogens. The design rule for epitope presentation on VLPs still remains unclear. A tandem repeat display strategy, in which antigenic element was repetitively arrayed, led to the induction of a significant antibody titer against its native structure, compared with the use of flanking sequence. Tandem repeat display strategy might retain the epitope conformation that is close to its native structure. This work further investigates the structural presentation of Helix 190 (H190) and Helix A (HA2A), from influenza hemagglutinin, using tandem repeat strategy. Different tools are exploited to characterise the structures of presented epitopes. The major outcomes of this research are: (i) demonstrating the limitation of phage antibody panning as a tool to inform the structural presentation of epitope on modular VLPs; (ii) developing a simple tool for rapid screening of presented epitope conformation using circular dichroism spectroscopy coupled with singular value decomposition algorithms (CD-SVD); (iii) demonstrating the capability of tandem repeat strategy in promoting the helicity of the helical structure; (iv) establishing initial steps in determining atomic structure of presented epitope using X-ray crystallography. This work introduces, to the best of our knowledge, a CD-SVD method, a simple and rapid tool to characterise epitope conformation, particularly helical structure, on modular capsomeres. Due to the simplicity and speed, it enables the CD-SVD method to be employed as a screening tool for the structural presentation of antigenic modules. It allows the rapid identification of potential vaccine constructs in the early stage of vaccine development and therefore, reduces the further cost of pre-clinical and clinical trials. This work also demonstrates the structural presentation of HA2A using tandem repeat display strategy led to induce antibody titer against full-length hemagglutinin in the chicken trial. The finding reveals that a tandem repeat strategy promotes the helicity of inserts derived from native helical regions; this approach deserves further research into its potential as a generic design for the presentation of the helical antigenic elements on modular VLPs.

.**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 policy and procedures of The University of Queensland, the thesis be made available for research and study in accordance with the Copyright Act 1968 unless a period of embargo has been approved by the Dean of the Graduate School.

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.

Publications during candidature

Conference abstract

Tam T. Doan, Nani Wibowo, Linda H.L. Lua, Anton P.J. Middelberg* (2014). Structural presentation of an influenza antigenic element on the virus-like particle. 5th International NanoBio Conference. 6-10 July 2014. Brisbane, Australia.

*Corresponding author

Publications included in this thesis

None.

Contributions by others to the thesis

This PhD project was performed by the candidate under the guidance of supervisory team including Prof. Linda Lua as a primary supervisor and Prof. Anton Middelberg as co-supervisors. This thesis was written by the candidate under the supervision of the supervisory team.

Modular capsomeres simulations in chapter 4 were performed by Dr. Natalie Connors.

Animal handling during *in vivo* immunogenicity study in chapter 4 was performed by the candidate with the assistance of Dr. Nani Wibowo, Jarurin Waneesom, Andrea Schaller and Dr. Natalie Connor. Technical supports in animal handling were provided by Dr. Arun Kumar in Poultry Research Unit, Gatton campus, The University of Queensland. The daily animals care the was performed by Vivek Guruswamy in Poultry Research Unit, Gatton campus, The University of Queensland.

The screening of crystallisation conditions in chapter 5 was performed in collaboration with Dr. Santosh Panjikar from the Beamline team of Macromolecular Crystallography (Protein crystallography), the Australian Synchrotron (Melbourne, Australia).

All of the data and results presented in the figures and tables of this thesis are solely the work of the candidate with the exception of the following figures and table in which the data collected with collaborators:

- Figure 5.18: Protein crystallisation screening was performed by Dr. Santosh Panjikar.
- Figure 5.19: X-ray diffraction was performed by Dr. Santosh Panjikar.
- Table 5.5: Collection of X-ray diffraction was performed by Dr. Santosh Panjikar.

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

None.

Acknowledgements

I would like to thank and express my gratitude to my principal supervisor Prof. Linda Lua and my co-supervisor Prof. Anton Middelberg for their valuable guidance and encouragement during my PhD A special thanks to my former principal supervisor, Dr. Nani Wibowo, for her mentorship and helpful comments and suggestions. Thank you very much Prof. Linda Lua, Prof. Anton Middelberg and Dr. Nani Wibowo for your support during my pregnancy.

I would like to acknowledge The University of Queensland for providing me The University of Queensland International (UQI) Scholarship including tuition fee and living allowance during my study.

Thanks to all the CBE members for their support and friendship. Thanks to all the members of Protein Expression Facility (PEF) who provided me helpful technical support. A special thanks to Prof. Steve Mahler and Dr. Martina Jones for your technical support and useful suggestions during my work with phage antibody panning in Ritchie Lab.

Thanks to all my friends in Australia for your friendship and encouragement.

Finally, a very big thanks to my family. Cám ơn ba mẹ đã vất vả cực nhọc dạy dỗ và nuôi con ăn học cho đến ngày học nay. Những thành tựu mà con đạt được đều nhờ vào công sức quý báu của ba mẹ. Em xin cám ơn anh, người bạn đời của em là chồng em, Lê Lâm Đức Kế, người đã cùng em chia sẽ biết bao nhiêu khó khăn vất vả trong cuộc sống. Xin cám ơn anh trong những đêm khuya hay những sáng sớm giá lạnh mùa đông đã không ngần ngại đưa đón em đến lab. Mọi lúc em cần anh đều bên cạnh em hỗ trợ động viên em.

Without your support and encouragement, I would not be able to fulfill my PhD degree.

Keywords

Virus-like particle, influenza, peptide antigen, structural presentation, vaccine design

Australian and New Zealand Standard Research Classifications (ANZSRC)

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

ANZSRC code: 100403, Medical Molecular Engineering of Nucleic Acids and Proteins, 50%

ANZSRC code: 090403, Chemical Engineering Design, 40%

Fields of Research (FoR) Classification

FoR code: 0904, Chemical Engineering, 60%

FoR code: 1004, Medical Biotechnology, 40%

Table of Contents

List of Figures

.

List of Tables

List of Abbreviations used in the thesis

Å	Angstrom
ARC	Australian Red Cross
CD	Circular dichroism
CDR	Complementary determining regions
Cryo-EM	cryo-electron microscopy
CV	Column volume
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetertraacetic acid
ELISA	Ezyme-linked immunosorbent assay
ER	Endoplasmic reticulum
Fab	Antigen binding fragment
FHV	Flock house virus
GRAVY	Grand average of hydropathy
GSH	Glutathione
GST	Glutathione-S-transferase
H190	Helix 190
HA	Haemagglutinin
HA2A	Helix A
HaPyV	Hamster polyomavirus
HBV	Hepatitis B virus
HCl	Hydrochloric acid
HCV	Hepatitis C virus

HEV	Hepatitis E virus
HIC	Hydrophobic interaction chromatography
HIV	Human immunodeficiency virus
HPV	Human papillomavirus
Hsp	Heat shock protein 70
IPTG	Isopropyl β-D-1-thiogalactopyranoside
kDa	kilo Dalton
kL	kilo Litre
LB	Luria-Bertani
LS	Light scattering
M1	Matrix protein 1
M2	Matrix protein 2
MD	Molecular dynamics
MuPyV	Murine polyomavirus
MRW	Mean residue weight
NA	Neuraminidase
NMR	Nuclear magnetic resonance
NP	Nucleoprotein
NS1	Non-structural protein 1
NS2	Non-structural protein 2
OD600 nm	Optical density at 600nm
PA	Polymerase acidic protein
PB1	Polymerase basic protein 1
PB2	Polymerase basic protein 2
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction

pI	Isoelectric point
RNA	Ribonucleic acid
RNP	Ribonucleoprotein
rpm	revolution per minute
SARS	Severe acute respiratory syndrome
scFv	single chain variable fragment
SD	Standard deviation
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SEC	Size-exclusion chromatography
SVD	Singular value decomposition
TB	Terrific broth
TEVp	Tobacco etch virus protease
UQ	The University of Queensland
UV	Ultraviolet
VLP	Virus-like particle
VP1	Viral protein 1
WHO	World Health Organisation
Wt	Wild-type

Chapter 1

Project Overview

1.1. Influenza

Influenza is an infectious respiratory disease having high morbidity and mortality rates (Szucs, 1999; Thompson *et al.*, 2003). Influenza virus is a negative single-stranded RNA virus from the *Orthomyxoviridae* family. Among the three types of virus, influenza A virus causes severe illness and pandemics in humans (Dawood *et al.*, 2009). Antigenic drift and shift result in the rapid mutation of influenza virus, allowing the virus to evade protective antibodies in the pre-existing immune system. Antigenic shift results in consequences more severe than antigenic drift because this phenomenon can confer in the virus ability to cross host barriers between animals and humans, and causes outbreaks in human populations (Olsen, 2002; Webster *et al.*, 1992). There were a few influenza pandemics in history, for example, the Spanish influenza pandemic (H1N1) in 1918, Hong Kong influenza pandemic (H3N2) in 1968 and the most recent H1N1 pandemic in 2009 (Neumann *et al.*, 2009). The rapid mutation of influenza virus requires the regular re-engineering of vaccine components matching with the circulating strains. Vaccine re-engineering must be quicker than the time, that the virus undergoes antigenic change, to prevent the outbreak of influenza. However, the conventional influenza vaccine based on the egg-culture system responds slowly to influenza pandemics. For example, in the 2009 H1N1 pandemic, the virus spread globally to 43 nations with 86 deaths within only a month, while the egg-manufactured vaccine was released five months later. A new cell-based manufacturing method can potentially produce vaccine in response to a pandemic threat in about two months, which it is still slower than the speed of antigenic drift or shift in the viruses (Barrett, 2010). These considerations reveal a need for a new vaccine technology that can quickly respond to the antigenic change of the virus.

1.2. Influenza vaccines based on the structure-based approach

A recent vaccine development approach has employed highly-purified, well-characterised antigens instead of whole organisms (Thomas and Luxon, 2013). The antigen structure needs to resemble its native structure to induce neutralising antibodies because the antigen recognition of an antibody depends on the structure of antigen (Sela-Culang *et al.*, 2013). Novavax Inc. (Gaithersburg,

Maryland, United States) selected three structural proteins that can form an empty viral shell to develop an influenza vaccine candidate (Bright *et al.*, 2007), leading to the induction of protective immunity (Pushko *et al.*, 2010). Protein Sciences (Meriden, Connecticut, United States) developed recombinant hemagglutinin (rHA) as an influenza vaccine (Flublok®). A study from Protein Sciences reported that the conformational change of rHA, indicated by the formation of non-native disulphide bonds in rHA, led to the loss of potency (Hickey *et al.*, 2014). The studies demonstrated the importance of antigen structure in vaccine efficacy.

The importance of antigen structure on vaccine efficacy led to the idea of "structural vaccinology", in which structural information, obtained from structure determination techniques and computational methods, is used as a basic understanding of vaccine design. Bommakanti *et al*. used computational methods to identify interacting HA1 residues and HA2 hydrophobic residues on the crystal structure of HA. That study retained the interacting HA1 residues and mutated the HA2 hydrophobic residues to stabilise the stalk domain of HA. The rational design of HA2-based immunogens led to mice protection in a lethal challenge (Bommakanti *et al.*, 2010; Bommakanti *et al.*, 2012). Another study also employed the HA crystal structure to design the HA globular head region which encompasses neutralising antibody epitopes, leading to the generation of stable head regions as vaccine candidates (Xuan *et al.*, 2011). The study demonstrated that the rationally designed head part elicited protective immunity in mice (Xuan *et al.*, 2011). This structure-based vaccine design approach also utilises epitopes in inducing immune response because this approach has a rationale that a protective epitope is sufficient to elicit protection against pathogens. Joshi *et al.* employed molecular dynamics (MD) simulation, a computational method, to engineer an epitope from hepatitis B virus for presentation on the surface of human papillomavirus VLP. The study demonstrated the design for epitope presentation led to the binding affinity of antibodies to the epitope similar to the binding to a native virus (Joshi *et al.*, 2011; Joshi *et al.*, 2013).

An epitope requires a protein carrier or platform, e.g. virus-like particle (VLP), to improve its poor immunogenicity (Jennings and Bachmann, 2009). A VLP, an assembly of viral capsid proteins, has particulate nature and highly repetitive and ordered structure, which is the basis of the highly immunogenic property of VLP (Chackerian, 2007). Several studies found that a modular VLP or VLP subunit, presenting an epitope from a distinct pathogen, induced antibody response against the heterogeneous epitope (Chuan *et al.*, 2014a; Dale *et al.*, 2002; Jain *et al.*, 2010; Liu *et al.*, 2000; Tissot *et al.*, 2010; Wibowo *et al.*, 2012). However, very limited studies characterise the conformation of an epitope on VLP (Joshi *et al.*, 2013; Schneeman *et al.*, 2012). The evidence

indicates that there is a knowledge gap in designing epitope presentation in its native structure on the surface of an unrelated VLP.

The microbial vaccine platform at the University of Queensland (UQ), based on murine polyomavirus (MuPyV) major capsid protein VP1, has been exploited to present a heterogeneous epitope or an antigen element (Figure 1.1). The antigenic element is modularised with functional elements (e.g. structure promoting element or spacer element) to create a synthetic module which is then inserted, at the DNA level, into the surface of the VLP, VLP subunit (capsomere), and/or the termini of the capsomere. Studies have demonstrated the immunogenicity of these modular VLPs and capsomeres in inducing potent antibody response (Middelberg *et al.*, 2011; Rivera-Hernandez *et al.*, 2013; Wibowo *et al.*, 2012). A study has demonstrated protective efficacy in mice, measured regarding reduced weight loss during laboratory challenge (Wibowo *et al.*, 2014). Furthermore, the vaccine platform has the potential to enable the production of 320 million doses in 4.7 days for VLP-based vaccines and 2.3 days for VLP subunit-based vaccines using a ten kL bioreactor (Chuan *et al.*, 2014b; Middelberg *et al.*, 2011), indicating the capability of rapid response to influenza pandemics.

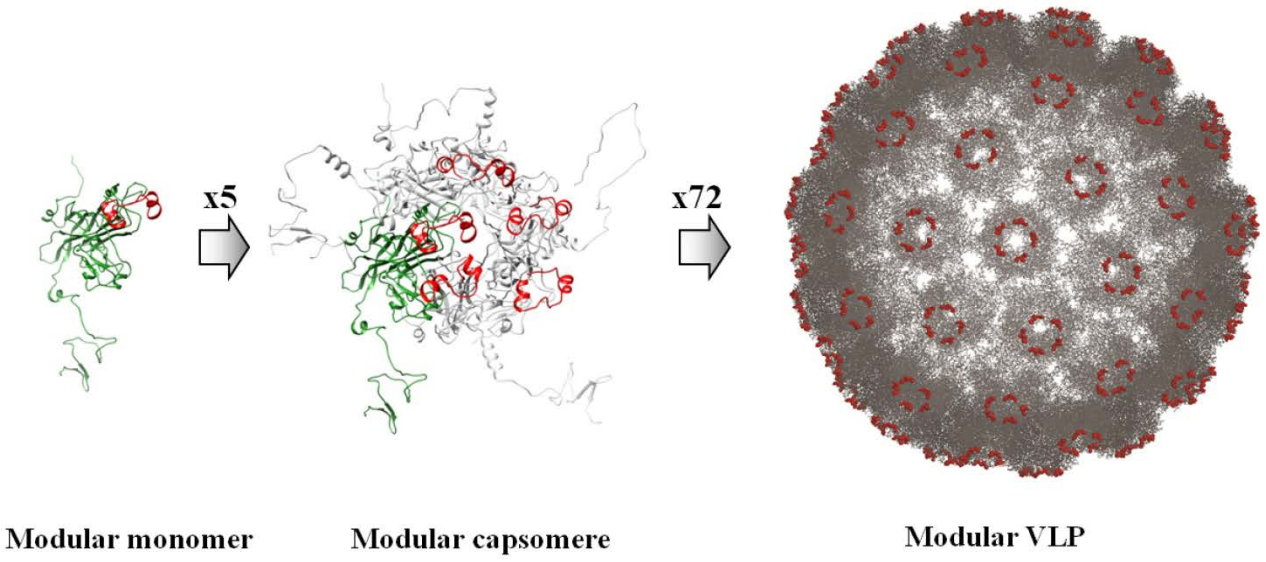


Figure 1.1: Presentation of influenza antigenic module on MuPyV capsomere. The antigenic element (highlighted in red) was presented on MuPyV vaccine platform using tandem repeat display strategy. A modular capsomere consists of five modular monomers (one monomer is highlighted in green), then 72 modular capsomeres self-assemble in modular VLP. The predicted structure was generated using Accelrys Discovery Studio 3.0 performed by Natalie Connors based on 1sid.pdb.

Various strategies, including the usage of structure-promoting element GCN4 and tandem repeat, have been explored to present an antigenic element on MuPyV VLP. GCN4 element, a leucine

zipper of yeast transcription factor, was used to maintain the helical structure of group A *streptococcus* peptide epitope on the surface of modular MuPyV VLP (Rivera-Hernandez *et al.*, 2013). Tandem repeat display, in which the epitope is arrayed repetitively, was employed to promote the native structure of peptide epitope (Fontenot *et al.*, 1995; Fontenot *et al.*, 1993). A study within the UQ group employed these two strategies to display Helix 190 (H190) on MuPyV VLP (Anggraeni *et al.*, 2013). H190 is a highly immunodominant antigenic site within the receptor binding site of influenza hemagglutinin globular sub-domain (HA1) (Caton and Brownlee, 1982; Yu *et al.*, 2008). The study modularised H190 by (i) flanking with GCN4 structure-promoting elements, or (ii) dual H190 copies arrayed as tandem repeats to generate the antigenic modules (Figure 1.2). The study demonstrated that the modularisation of H190 using tandem repeat approach resulted in higher antibody titers against native structure than the flanking with the structure-promoting element. Furthermore, peptide molecular dynamics (MD) simulation predicted that the element H190 within the module H190-GCN4 unfolds whereas a copy of H190 within the module H190-H190 should retain its helical propensity (Anggraeni *et al.*, 2013). The study indicated that the tandem repeat display strategy was more effective to present a helical structure, particularly H190, on a VLP.

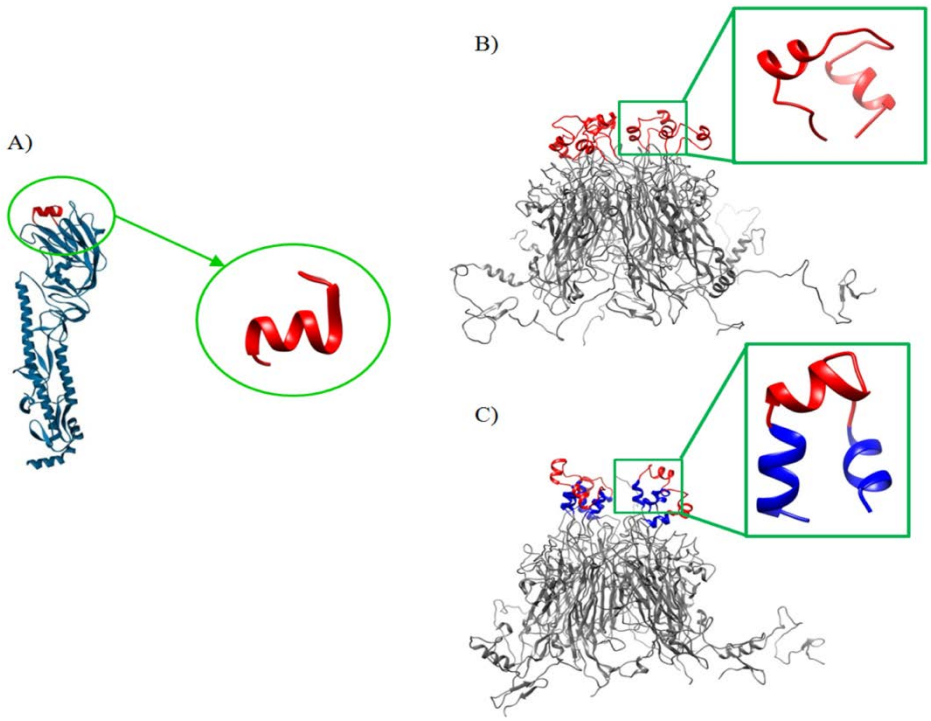


Figure 1.2: Modularisation of an influenza antigenic element. A) The native conformation of an antigenic element from the context of the whole antigen. B) The structure of an antigenic module using tandem repeat display strategy. C) The structure of an antigenic module using the display strategy flanking with a structural promoting element. The antigenic element was highlighted in red, structural promoting element in blue, MuPyV VP1 in grey. The predicted structures of

modular capsomeres were generated using Accelrys Discovery Studio 3.0 performed by Natalie Connors based on 1 SID.pdb.

The observation in the previous study motivates this Ph.D. to investigate the effect of display strategies to a helical structure on modular VLP. The study showed that tandem repeat display strategy would be more effective to present a helical structure (Anggraeni *et al.*, 2013). This research, therefore, further investigated the effect of tandem repeat display on helical epitopes, particularly H190 and another helical epitope, Helix A (HA2A), from influenza hemagglutinin. HA2A is a conserved epitope on the stalk region (HA2) of hemagglutinin (Ekiert *et al.*, 2009). From that, the potential of tandem repeat approach as a generic design for the presentation of a helical structure on VLP can be defined. The design rule of epitope presentation on VLP would accelerate the re-engineering of influenza vaccine components to respond quickly to influenza pandemics, and contribute to the development of modular VLP-based vaccines.

1.3. Research aims and objectives

This Ph.D. seeks to understand better how to present a structural epitope on the modular MuPyV VLP and by extension, on VLPs more generally. Target epitopes from influenza virus are H190 from the globular head region of HA and HA2A from stalk region of HA.

This Ph.D. thesis focuses on two specific objectives:

i. To investigate the structural presentation of antigenic elements, particularly H190 element modularised using two display strategies: (a) flanking with a structure-promoting element or (b) tandem repeat, via the antigen-antibody recognition in phage antibody panning.

ii. To evaluate the effect of tandem repeat display strategy in promoting the helical propensity of antigenic modules, particularly H190 and HA2A, on modular VLP.

The rationale of these research objectives is presented in the following sections.

1.3.1. The structural presentation of the antigenic module using two different strategies

Understanding the structural presentation of an epitope or an antigenic element on the surface of VLP is important to guide the modular designs maintaining the native structure of an antigenic element. The study within the group used two different display strategies: tandem repeat and flanking with structure promoting element to present H190 element on VLP (Anggraeni *et al.*, 2013). Although the study demonstrated that the display strategies varied the quality of immune response, the effect of display strategy to the structure of H190 element is still unknown. The

research question here is whether the display strategies resulted in the structural difference of H190 element presented on modular VLPs. Therefore, it reveals a necessity to investigate the conformation of modularised H190 element presented on modular capsomeres or VLPs.

Phage antibody panning is a simple and rapid technique based on antigen-antibody recognition to select phage antibodies against a specific target antigen (Bradbury and Marks, 2004). Phage antibody panning was utilised to select antibodies against influenza antigens (Kashyap *et al.*, 2008; Sui *et al.*, 2009; Throsby *et al.*, 2008). The antigen-specific antibody has been reported to confirm the native structure of the protein (Lu *et al.*, 2014). This research employed antigen-antibody recognition through phage antibody panning to inform the structural presentation of H190 elements on modular capsomeres. Figure 1.3 shows the concept of how antibody panning can inform the structural presentation of H190 element. If the pool of phages isolated from the panning against modular capsomere contains phages that can bind to native H190 on HA1, it will indicate that the structure of H190 element on modular capsomere is similar to the native conformation on HA1 (Figure 1.3A). In contrast, if the pool of phages isolated from the panning against modular capsomere is unable to bind to HA1 or weakly bind to HA1, it will indicate that the structure of H190 elements on modular capsomere presents in a non-native or weakly binding structure (Figure 1.3B).

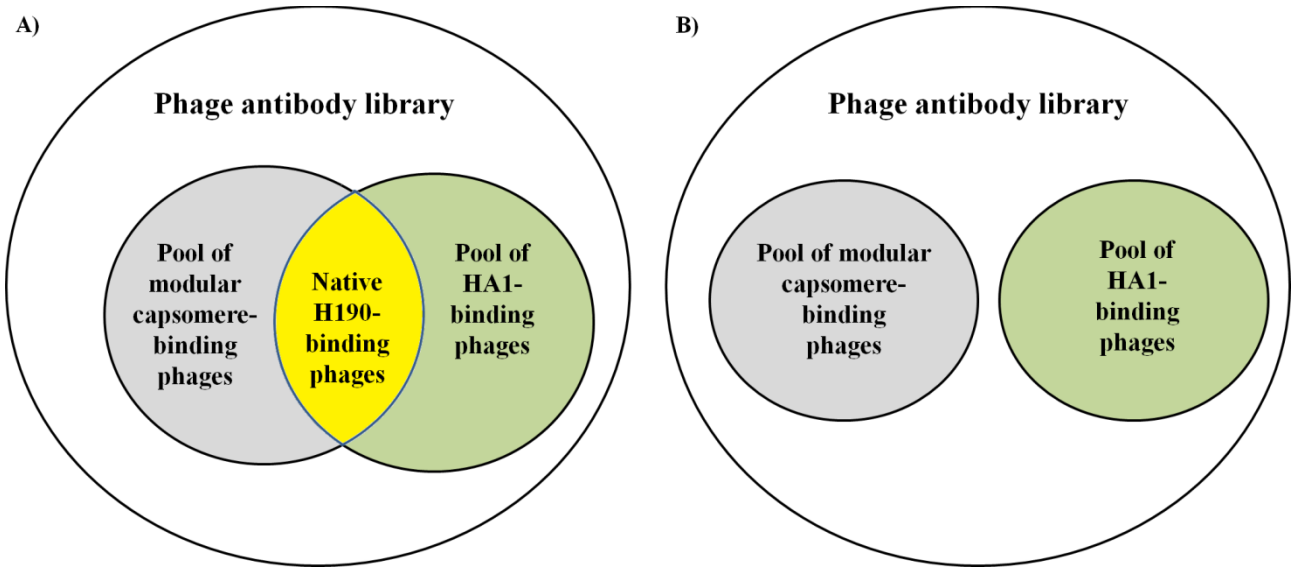


Figure 1.3: The concept how phage antibody panning can inform the structure of modularising H190 on the modular capsomere. (A) If the pool of phages bound to modular capsomere contains phages that can bind to native H190 on HA1, it will indicate that the structure of H190 element on modular capsomere is similar to the native structure on HA1. (B) If the pool of phages bound to modular capsomere is unable to bind to HA1 or weakly bind to HA1, it will indicate that the structure of H190 element on modular capsomere is different from the native structure on HA1.

1.3.2. Effect of tandem repeats to the structure of the antigenic module

The previous study within the group found that the presentation of H190 element using tandem repeat strategy elicited higher quality of immune response compared with the usage of flanking GCN4 elements (Anggraeni *et al.*, 2013). However, the helical structures of antigenic modules, using these display strategies, remain unknown. The effect of tandem repeat in promoting the helical propensity of the antigenic module is still unable to be confirmed. The research questions here are: What is the structure of H190 modules when presented in the two display strategies? Can the increase in the copy number of antigenic elements in tandem repeat lead to improved helicity? To what extent does the tandem repeat strategy affect the helical propensity of HA2A modules? What is the preferred copy number of antigenic element to present HA2A module? Can the modularised HA2A elicit IgGs recognising peptide HA2A and FL-HA, showing the induction of immunogenicity and antibody quality, respectively?

To address the above research questions, this Ph.D. first exploited circular dichroism (CD) spectroscopy, a simple technique for determining the protein secondary structure, to investigate the helical structures of antigenic modules on modular capsomeres. To further examine a high-resolution structure of antigenic module at the atomic level, X-ray crystallography was employed.

The first technique is CD spectroscopy, which is a simple and rapid tool to determine protein secondary structures (Greenfield, 2007; Whitmore and Wallace, 2008), with the support of computation methods, i.e. singular value decomposition algorithms (SVD) (Sreerama and Woody, 2000). From protein CD spectra, SVD algorithms can estimate the ratios of protein secondary structures including α-helix (Sreerama and Woody, 2000). The ratios can then be used to calculate the number of amino acids having the secondary structures. CD spectroscopy coupled with a computation method has been used previously to evaluate the individual secondary structure components of MuPyV capsomere (Yang and Teng, 1998; Yang and Teng, 1999). This Ph.D. investigated CD spectroscopy coupled with SVD algorithms (CD-SVD) to inform the helical structures of antigenic modules, especially H190 and HA2A, on modular capsomeres by analysing the helical amino acid numbers of modular capsomeres in comparison to the unmodified one. The increase in the helical amino acids number of modular capsomeres compared with unmodified capsomere indicated the formation of helical structures of antigenic modules. The conformational information obtained from CD-SVD was then validated using immunogenicity data from *in vivo* immunogenicity testing. Due to the simplicity and speed of CD-SVD, this Ph.D. also examines CD-SVD method as a screening tool for the structural presentation of antigenic modules through informing the helical amino acid numbers of the antigenic module. Though CD-SVD can inform the

helicity of the antigenic module, it was unable to indicate which region of antigenic module forms a helical structure. Therefore, a technique that can determine the atomic structure of the antigenic module is required.

The second technique is X-ray crystallography, which can resolve the atomic structure of proteins larger than 35kDa (Pusey *et al.*, 2005). This technique requires achieving optimum crystals, for instance having sufficient size and high-resolution diffraction, to enable to solve the atomic structure using X-ray crystallography. The diffraction of optimum crystal results in the unique patterns of spots which reflect the electron distributions of the protein. From that, electron density maps are obtained to determine the atomic structure of the protein (Kleywegt and Jones, 1998). Some studies used X-ray crystallography to characterise the three-dimensional structure of VLP (Dong *et al.*, 1998; Wynne *et al.*, 1999). The structure of wild-type MuPyV capsomere was also solved using X-ray crystallography (Stehle and Harrison, 1997). X-ray crystallography, therefore, could be a potential tool to obtain the atomic structure of modular capsomeres, leading to characterise the conformation of antigenic modules on modular capsomere. However, due to time constraints and time consumption of crystallisation screening and optimisation, this work could not be finished within the timescale of this Ph.D.

1.4. Thesis organisation

This Ph.D. thesis consists of six chapters including this introductory chapter and five following chapters listed below.

Chapter 2 provided a literature review of the main topics describing the background of influenza and influenza virus, the UQ microbial vaccine platform, and some available fundamental tools aiding characterisation of protein structure.

Chapter 3 investigated the structural presentation of antigenic elements using different display strategies via phage antibody panning (objective i). Because phage antibody panning requires the regular re-construction of phage antibody library that is relevant to the current emerging strain of the virus, this chapter demonstrated the limitation of this technique in informing the structural presentation of an antigenic element, particularly influenza epitopes.

Chapter 4 developed a simple method, CD spectroscopy coupled with SVD algorithms (CD-SVD), to characterise the helical structures of antigenic modules, leading to disclosing the effect of tandem repeat to the helicity of the antigenic module (objective ii). Due to the poor resolution of structural information obtained from CD-SVD, this chapter revealed a necessity of exploiting another technique to obtain the atomic structures of antigenic modules, i.e. X-ray crystallography.

Chapter 5 implemented some fundamental steps in X-ray crystallography to obtain the atomic structures of antigenic modules, leading to disclose the effect of display strategies to the helical structure of antigenic modules (objective ii). This chapter produced the highly purified modular capsomeres which were then used to screen crystallisation conditions for the growth of the optimum crystals. Optimisation of crystallisation condition is time-consuming. At the time that this Ph.D. thesis was written, the crystallisation optimisation is still ongoing to find conditions for growing optimum crystals.

Chapter 6 summarised the findings of this research and proposed future research directions.

References

Anggraeni, M.R., Connors, N.K., Wu, Y., Chuan, Y.P., Lua, L.H.L., Middelberg, A.P.J., 2013. Sensitivity of immune response quality to influenza helix 190 antigen structure displayed on a modular virus-like particle. Vaccine 31, 4428-4435.

Barrett, P.N., 2010. Development cell culture-derived pandemic vaccines. Curr Opin Mol Ther 12, 21-30.

Bommakanti, G., Citron, M.P., Hepler, R.W., Callahan, C., Heidecker, G.J., Najar, T.A., Lu, X., Joyce, J.G., Shiver, J.W., Casimiro, D.R., Ter Meulen, J., Liang, X., Varadarajan, R., 2010. Design of an HA2-based *Escherichia coli* expressed influenza immunogen that protects mice from pathogenic challenge. Proc Natl Acad Sci USA 107, 13701-13706.

Bommakanti, G., Lu, X., Citron, M.P., Najar, T.A., Heidecker, G.J., ter Meulen, J., Varadarajan, R., Liang, X., 2012. Design of *Escherichia coli*-expressed stalk domain immunogens of H1N1 hemagglutinin that protect mice from lethal challenge. J Virol 86, 13434-13444.

Bradbury, A.R.M., Marks, J.D., 2004. Antibodies from phage antibody libraries. J Immunol Methods 290, 29-49.

Bright, R.A., Carter, D.M., Daniluk, S., Toapanta, F.R., Ahmad, A., Gavrilov, V., Massare, M., Pushko, P., Mytle, N., Rowe, T., Smith, G., Ross, T.M., 2007. Influenza virus-like particles elicit broader immune responses than whole virion inactivated influenza virus or recombinant hemagglutinin. Vaccine 25, 3871-3878.

Caton, A.J., Brownlee, G.G., 1982. The antigenic structure of the influenza virus A/PR/8/34 hemagglutinin (H1 subtype). Cell 31, 417-427.

Chackerian, B., 2007. Virus-like particles: flexible platforms for vaccine development. Expert Rev Vaccines 6, 381-390.

Chuan, Y.P., Wibowo, N., Connors, N.K., Wu, Y., Hughes, F.K., Batzloff, M.R., Lua, L.H.L., Middelberg, A.P.J., 2014a. Microbially synthesized modular virus-like particles and capsomeres

displaying group A *streptococcus* hypervariable antigenic determinants. Biotechnol Bioeng 111, 1062-1070.

Chuan, Y.P., Wibowo, N., Lua, L.H.L., Middelberg, A.P.J., 2014b. The economics of virus-like particle and capsomere vaccines. Biochem Eng J 90, 255-263.

Dale, C.J., Liu, X.S., De Rose, R., Purcell, D.F.J., Anderson, J., Xu, Y., Leggatt, G.R., Frazer, I.H., Kent, S.J., 2002. Chimeric human papilloma virus-simian/human immunodeficiency virus virus-like particle vaccines: Immunogenicity and protective efficacy in macaques. Virology 301, 176-187.

Dawood, F.S., Jain, S., Shaw, M., 2009. Emergence of a novel swine-origin influenza A (H1N1) virus in humans. N Engl J Med 360, 2605-2615.

Dong, X.F., Natarajan, P., Tihova, M., Johnson, J.E., Schneeman, A., 1998. Particles polymorphism caused by deletion of a peptide molecular switch in a quasiequivalent icosahedral virus. J Virol 72, 6024-6033.

Ekiert, D.C., Bhabha, G., Elisliger, M.A., Friesen, R.H.E., Jongeneelen, M., Throsby, M., Goudsmit, J., Wilson, I.A., 2009. Antibody recognition of a highly conserved influenza virus epitope. Science 324, 246-251.

Fontenot, J., Mariappan, S., Catasti, P., Domenech, N., Finn, O.J., Gupta, G., 1995. Structure of a tumor associated antigen containing a tandemly repeated immunodominant epitope. J Biomol Struct Dyn 13, 245-260.

Fontenot, J.D., Tjandra, B., D., Ho, C., Montelaro, R.C., Finn, O.J., 1993. Biophysical characterization of one-, two-, and three-tandem repeats of human mucine (muc-1) protein core. Cancer Res 53, 5386.

Greenfield, N.J., 2007. Using circular dichroism spectra to estimate protein secondary structure. Nat Protocols 1, 2876-2890.

Hickey, J.M., Holtz, K.M., Manikwar, P., Joshi, S.B., McPherson, C.E., Buckland, B., Srivastava, I.K., Middaugh, C.R., Volkin, D.B., 2014. Mechanism of a decrease in potency for the recombinant influenza A virus hemagglutinin H3 antigen during storage. J Pharm Sci 103, 821-827.

Jain, S., Patrick, A.J., Rosenthal, K.L., 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, 7070-7080.

Jennings, G.T., Bachmann, M.F., 2009. Immunodrugs: Therapeutics VLP-based vaccines for chronic diseases. Annu Rev Pharmacol Toxicol 49, 303-326.

Joshi, H., Cheluvaraja, S., Somogyi, E., Brown, D.R., Ortoleva, P., 2011. A molecular dynamics study of loop fluctuation in human papillomavirus type 16 virus-like particles: a possible indicator of immunogenicity. Vaccine 29, 9423-9430.

Joshi, H., Lewis, K., Singharoy, A., Ortoleva, P.J., 2013. Epitope engineering and molecular metrics of immunogenicity: A computational approach to VLP-based vaccine design. Vaccine 31, 4841-4847.

Kashyap, A.K., Steel, J., Oner, A.F., Dillon, M.A., Swale, R.E., Wall, K.M., Perry, K.J., Faynboym, A., Ilhan, M., Horowitz, M., Horowitz, L., Palese, P., Bhatt, R.R., Lerner, R.A., 2008. Combinatorial antibody libraries from survivors of the Turkish H5N1 avian influenza outbreak reveal virus neutralization strategies. Proc Natl Acad Sci USA 105, 5986-5991.

Kleywegt, G.J., Jones, T.A., 1998. Databases in protein crystallography. Acta Cryst D 54, 1119-1131.

Liu, W.J., Liu, X.S., Zhao, K.N., Leggatt, G.R., Frazer, I.H., 2000. Papillomavirus virus-like particles for the delivery of multiple cytotoxic T cell epitopes. Virology 273, 374-382.

Lu, Y., Welsh, J.P., Swartz, J.R., 2014. Production and stabilization of the trimeric influenza hemagglutinin stem domain for potentially broadly protective influenza vaccines. Proc Natl Acad Sci USA 111, 125-130.

Middelberg, A.P.J., Rivera-Hernandez, T., Wibowo, N., Lua, L.H.L., Fan, Y., Magor, G., Chang, C., Chuan, Y.P., Good, M.F., Batzloff, M.R., 2011. A microbial platform for rapid and low-cost virus-like particle and capsomere vaccines. Vaccine 29, 7154-7162.

Neumann, G., Noda, M., Kawaoka, Y., 2009. Emergence and pandemic potential of swine-origin H1N1 influenza virus. Nature 459, 931-939.

Olsen, C.W., 2002. The emergence of novel swine influenza viruses in North America. Virus Res 85, 199-210.

Pusey, M.L., Liu, Z.-J., Tempel, W., Praissman, J., Lin, D., Wang, B.-C., Gavira, J.A., Ng, J.D., 2005. Life in the fast lane for protein crystallization and X-ray crystallography. Prog Biophys Mol Biol 88, 359-386.

Pushko, P., Kort, T., Nathan, M., Pearce, M.B., Smith, G., Tumpey, T.M., 2010. Recombinant H1N1 virus-like particle vaccine elicits protective immunity in ferrets against the 2009 pandemic H1N1 influenza virus. Vaccine 28, 4771-4776.

Rivera-Hernandez, T., Hartas, J., Wu, Y., Chuan, Y.P., Lua, L.H.L., Good, M., Batzloff, M.R., Middelberg, A.P.J., 2013. Self-adjuvanting modular virus-like particles for mucosal vaccination against group A *streptococcus* (GAS). Vaccine 31, 1950-1955.

Schneeman, A., Speir, J.A., Tan, G.S., Khayat, R., Ekiert, D.C., Matsuoka, Y., Wilson, I.A., 2012. A virus-like particle that elicits cross-reactive antibodies to the conserved stem of influenza virus hemagglutinin. J Virol 86, 11686-11697.

Sela-Culang, I., Kunik, V., Ofran, Y., 2013. The structural basis of antibody-antigen recognition. Front Immunol 4, 302.

Sreerama, N., Woody, R.W., 2000. Estimation of protein secondary structure from circular dichroism spectra: Comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. Anal Biochem 287, 252-260.

Stehle, T., Harrison, S.C., 1997. High-resolution structure of a polyomavirus VP1-oligosaccharide complex: implications for assembly and receptor binding. Embo J 16, 5139-5148.

Sui, J., Hwang, W.C., Perez, S., Wei, G., Aird, D., Chen, L., Santelli, E., Stec, B., Cadwell, G., Ali, M., Wan, H., Murakami, A., Yammanuru, A., Han, T., Cox, N.J., Bankston, L.A., Donis, R.O.,

Liddington, R.C., Marasco, W.A., 2009. Structural and functional bases for broad-spectrum neutralization of avian and human influenza A viruses. Nat Struct Mol Biol 16, 265-272.

Szucs, 1999. The socio-economic burden of influenza. J Antimicrob Chemother 44, 11-15.

Thomas, S., Luxon, B.A., 2013. Vaccines based on structure-based design provide protection against infectious diseases. Expert Rev Vaccines 12, 1301-1311.

Thompson, W.W., Shay, D.K., Weintraub, E., et al., 2003. Mortality associated with influenza and respiratory syncytial virus in the united states. J Am Med Assoc 289, 179-186.

Throsby, M., van den Brink, E., Jongeneelen, M., Poon, L.L.M., Alard, P., Cornelissen, L., Bakker, A., Cox, F., van Deventer, E., Guan, Y., Cinalt, J., ter Meulen, J., Lasters, I., Carsetti, R., Peiris, M., de Kruif, J., Goudsmit, J., 2008. Heterosubtypic neutralizing monoclonal antibodies cross-protective against H5N1 and H1N1 recovered from human IgM^{+} memory B cells. PLoS One 3, e3942.

Tissot, A.C., Renhofa, R., Schmitz, N., Cielens, I., Meijerink, E., Ose, V., Jennings, G.T., Saudan, P., Pumpens, P., Bachmann, M.F., 2010. Versatile virus-like particle carrier for epitope based vaccines. PLoS One 5, e9809.

Webster, R.G., Bean, W.J., Gorman, O.T., Chambers, T.M., Kawaoka, Y., 1992. Evolution and ecology of influenza A viruses. Microbiol Rev 56, 152-179.

Whitmore, L., Wallace, B.A., 2008. Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases. Biopolymers 89, 392-400.

Wibowo, N., Chuan, Y.P., Lua, L.H.L., Middelberg, A.P.J., 2012. Modular engineering of a microbially-produced viral capsomere vaccine for influenza. Chem Eng Sci 103, 12-20.

Wibowo, N., Hughes, F.K., Fairmaid, E.J., Lua, L.H.L., Brown, L.E., Middelberg, A.P.J., 2014. Protective efficacy of a bacterially produced modular capsomere presenting M2e from influenza: Extending the potential of broadly cross-protecting epitopes. Vaccine 32, 3651-3655.

Wynne, S.A., Crowther, R.A., Leslie, A.G.W., 1999. The crystal structure of the human hepatitis B virus capsid. Mol Cell 3, 771-780.

Xuan, C., Shi, Y., Qi, J., Zhang, W., Xiao, H., Gao, G.F., 2011. Structural vaccinology: structure-based design of influenza A virus hemagglutinin subtype-specific subunit vaccines. Protein Cell 2, 997-1005.

Yang, Y.-W., Teng, C.-C., 1998. Stability of polyomavirus major capsid protein VP1 under denaturants guanidine hydrochloride and urea. Int J Biol Macromol 22, 81-90.

Yang, Y.-W., Teng, C.-C., 1999. Conformational changes of polyomavirus major capsid protein VP1 in sodium dodecyl sulfate solution. J Pept Res 53, 75-81.

Yu, X., Tsibane, T., McGraw, P.A., House, F.S., Keefer, C.J., Hicar, M.D., Tumpey, T.M., Pappas, C., Perrone, L.A., Martinez, O., Stevens, J., Wilson, I.A., Aguilar, P.V., Altschuler, E.L., Basler, C.F., Jr. Crowe, J.E., 2008. Neutralizing antibodies derived from the B cells of 1918 influenza pandemic survivors. Nature 455, 532-536.

Chapter 2

Literature review

2.1. Influenza

Influenza is a contagious disease causing symptoms such as fever, headache, sore throat and cough. Influenza is a concern because it causes high mortality and economic burdens for the society. In the US, influenza annually leads to 200,000 hospitalisations and 36,000 deaths (Jennings and Bachmann, 2008; Keller *et al.*, 2010). Most of these hospitalisations are caused by an infection in the lower respiratory tract, leading to a decrease in immunity to infection. Young children, the elderly and patients suffering other severe medical conditions are in high-risk groups endangered by influenza (WHO, 2009). Moreover, the economic burden of influenza was estimated to cost $16.3 billion annually (Brown *et al.*, 2002). Influenza causes negative impacts on the development of society and the economy in developing countries which have high population density and poor living conditions (Brooks *et al.*, 2010; Moorthy *et al.*, 2012). Although there have been many efforts in influenza prevention, the control of influenza has still been challenging.

The prevention of influenza has many difficulties due to the highly mutagenic ability of the virus that permits the virus to escape a pre-induced immune system. The influenza virus has a low fidelity RNA polymerase lacking exonuclease proofreading capability, leading to high mutation rates from approximately 1×10^{-3} to 8×10^{-3} substitutions per site per year (Chen and Holmes, 2006). These mutations result in minor antigenic changes known as antigenic drift. Moreover, the nature of its segmented genome enables influenza virus to undergo significant antigenic changes called antigenic shift. The acquisition of mutations during antigenic drift also increases virus pathogenicity, leading to high mortality rates. The co-infection of two strains into one host cell leads to the genetic reassortment among gene segments of both parental viruses, generating new subtypes. The genetic reassortment also gives rise to host switch events of the influenza virus (Garten *et al.*, 2009). Influenza viruses have caused a few influenza pandemics in history. For example, Spanish flu in 1918 resulting from the H1N1 subtype led to 50 million deaths. In 1968, Hong Kong flu with the emergence of the H3N2 subtype killed approximately 34,000 people (Neumann *et al.*, 2009). The most recent pandemic was the H1N1 pandemic in 2009. The pandemic outbreak originated in

Mexico and rapidly spread around the world (Centers for Disease Control and Prevention, 2009; Dawood *et al.*, 2009).

Influenza virus is a negative single-stranded RNA virus in *Orthomyxoviridae* family. Among the three types of influenza virus, influenza A virus causes more serious annual epidemics and more devastating pandemic than other strains. The structure of influenza A virus has a lipid membrane to envelop genetic materials inside (Figure 2.1). The viral envelope consists of three integral membrane proteins including hemagglutinin (HA), neuraminidase (NA) and matrix protein 2 (M2). While HA participates in viral entry via binding to cellular receptors and the fusion of the viral and endosomal membranes, NA plays a major role in packaging and releasing a new virus from the infected cell. M2 protein, an ion channel to transport protons, which is necessary for viral replication (Takeda *et al.*, 2002); M2 also modulates pH change of Golgi during HA generation, leading to the structural integrity of HA (Ito *et al.*, 1991). Under the viral envelope, a shell comprising matrix protein 1 (M1) encapsulates nucleoprotein, polymerase proteins (PB2, PB1, and PA), and non-structural proteins. Nucleoprotein encapsulates viral RNA and constitutes structural ribonucleoproteins. Nucleoprotein was shown to have the ability to induce a cytotoxic T-cell immune response (Webster *et al.*, 1992). Polymerase proteins associated with viral ribonucleoproteins play a role in RNA replication and transcription.

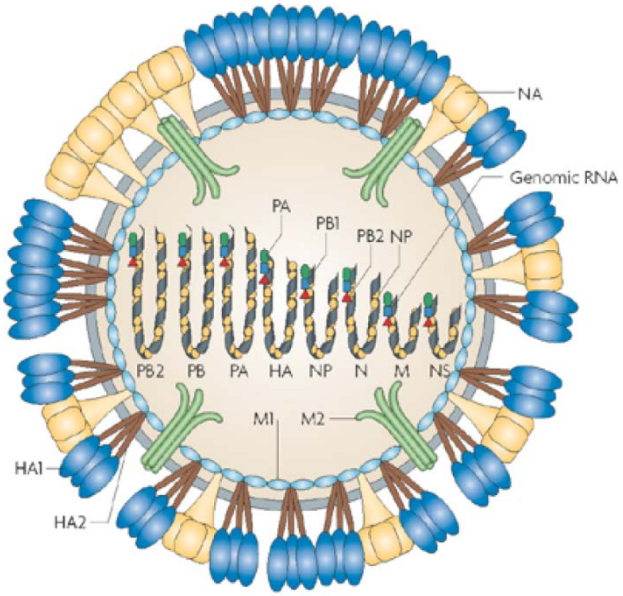


Figure 2.1: Schematic of influenza virus structure

In the outer membrane of influenza virus, there are three viral proteins: hemagglutinin (HA) consisting of HA1 (in blue) and HA2 (in grey) and neuraminidase (NA) (in orange) and M2 (in green) presented on the surface. Underneath the membrane, M1 protein (in light blue) forms a viral shell to package the viral genomes, nucleocapsid protein (NP) and polymerases PA, PB1, and PB2 (Karlsson Hedestam et al., 2008).

Influenza virus infects a target host cell through five fundamental steps: attachment, entry, replication, budding and release (Figure 2.2). When influenza virus initiates viral penetration, HA on the viral membrane firstly binds to the cell surface sialic acid receptor. The virus internalises into endosomes through endocytosis and fusion. Endosomal acid pH then activated M2 channel, allowing proton transport across the viral membrane. It leads to the dismantling of M1 protein, which gives rise to the release of RNP into the cytoplasm. The RNP enters the nucleus to replicate and transcribe the viral genome inside the nucleus. After the replication and transcription, assembly and budding of progeny virions occur at the membrane. In the final step, NA cleaves off sialic acid from glycans on the host cell to release new virions. (Cheng *et al.*, 2012)

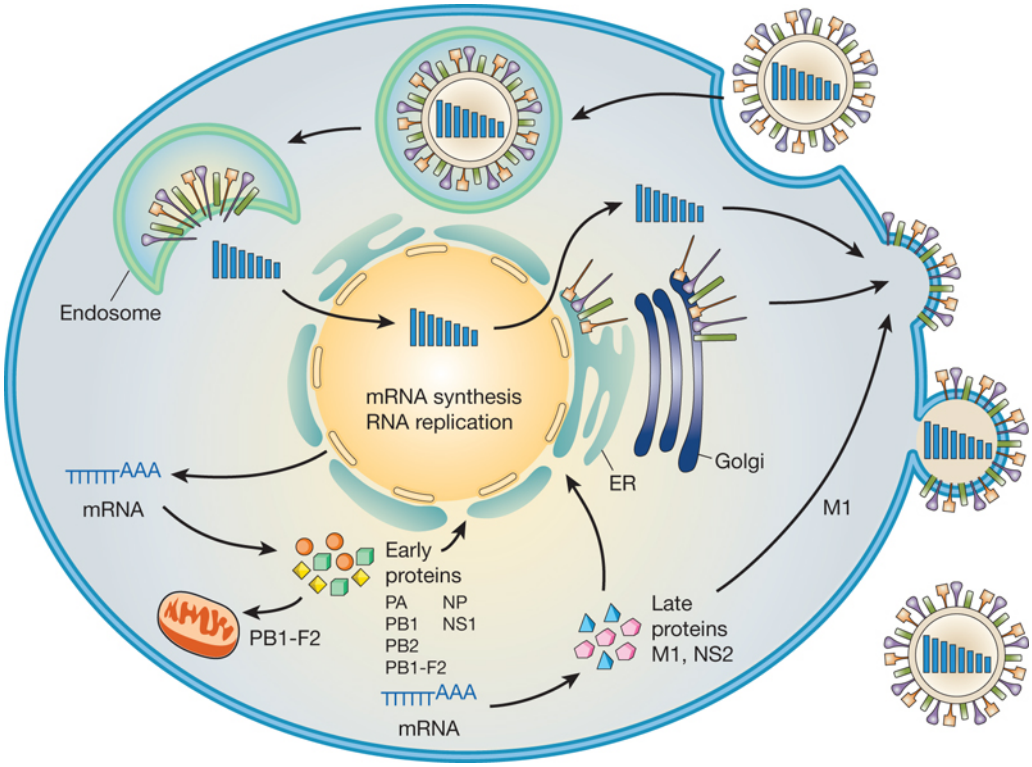


Figure 2.2: The life cycle of influenza virus (Neumann et al., 2009).

Understanding the life cycle of influenza virus facilitates the prevention of influenza virus by interrupting one of those critical steps in the life cycle. Interrupting the viral infection via inhibiting HA's activity is a more favourable strategy as HA is the most abundant protein and is highly accessible for antibody engagement. The structure of HA will be presented in more detail in section 2.2, but primarily, HA protein consists of a globular head region (contains HA1) and a stalk region (includes a small portion of HA1 and entirely HA2). HA1 and HA2 play different functions in the infection of influenza virus. HA1 is involved in the attachment of the virus to the receptor on host cell while HA2 participates in the membrane fusion of virus. HA1- and HA2-specific antibodies, thus, can block these processes, respectively, leading to inhibition of the viral penetration.

2.2. Hemagglutinin (HA)

Hemagglutinin is the most abundant glycoprotein on the surface of the influenza virus and highly accessible for antibody binding. HA, therefore, is a good target for influenza vaccines. HA is initially in an immature form or HA0. The precursor HA0 contains a signal peptide that enables the polypeptide to insert into the endoplasmic reticulum (ER); the signal peptide is then removed from the precursor (Bertram *et al.*, 2010). In the ER and Golgi apparatus, HA undergoes post-translational modification processes: folding, trimerisation, glycosylation and acidification. The most important step for the activation and infectivity of HA is extracellular proteolytic cleavage of the precursor HA0 at the arginine residues into a mature HA consisting of two subunits (HA1 and HA2) tightly linked by a disulphide bond (Figure 2.3).

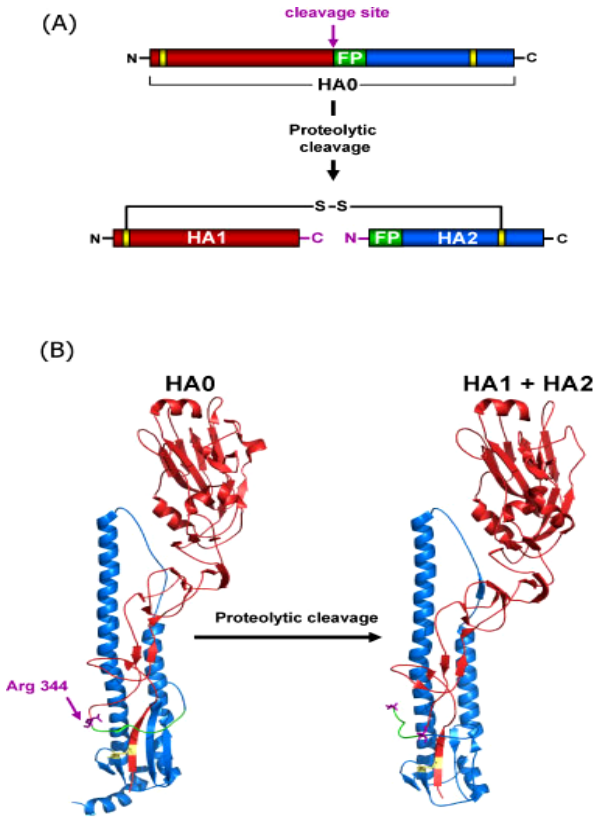


Figure 2.3: Maturation of HA through the proteolytic cleavage of precursor HA0 into mature HA. A) The schematic presentation of the proteolytic cleavage. B) Structure change from HA0 (1rd8.pdb) to mature HA (1ruz.pdb) due to the HA cleavage. HA1 is highlighted in red, HA2 in blue, fusion peptide in green, cleavage site at arginine in purple (Bertram et al., 2010).

Mature HA is an elongated cylinder, which has three homogeneous monomers associating together into a homotrimer to stabilise the structure of HA monomers. Each monomer HA consists of two regions that are a globular head and stalk region (Figure 2.4). A global head region containing entire HA1 presents a receptor binding site (Skehel and Wiley, 2000) and antigenic sites that are highly accessible for antibody binding (Staneková and Varečková, 2010). Due to the high mutagenicity of

amino acids surrounding the receptor-binding site in this region, HA1-targeting antibodies can neutralise a specific strain and may have no cross-reactivity between subtypes. Underneath the globular region, a stalk region consists of whole HA2 and some HA1 residues (Skehel and Wiley, 2000). In contrast, the stem region of HA is more conserved than the globular region; antibodies induced from this region have cross-reactivity with many subtypes (Steel *et al.*, 2010). The next section 2.2.1 and 2.2.2 will present potential antigenic sites on the globular head and stalk regions of HA for influenza vaccine development.

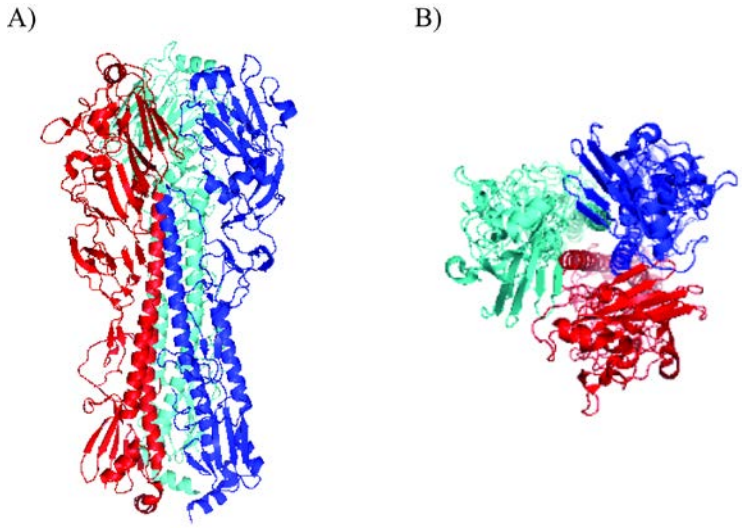


Figure 2.4: 3D structure of A/California/07/2009 hemagglutinin trimer (A) side view (B) top view. Each HA monomer is distinguished by different colours. The picture was generated by UCSF Chimera software version 1.8 based on 3LZG.pdb.

2.2.1. Globular head region or HA1

The globular head region of HA presents a receptor-binding site that attaches to sialic acid receptors when viral infection initiates. The receptor-binding site is a potential target for an influenza vaccine. The receptor-binding site is highly accessible for antibody engagement because it requires exposure to enable binding to the host sialic acid receptors (Vigerust and Shepherd, 2007; Weis *et al.*, 1988). Blocking the receptor binding site can inhibit the viral infection (Xu *et al.*, 2013). About conformation, the receptor binding site is a conserved pocket of amino acids that comprises three structural elements: Helix 190 (residues 188-190); loop 130 (residues 134-138) and loop 220 (residues 221-228) (numbering based on A/Vietnam/1203/2004 strain) (Stevens *et al.*, 2006). The genetics and structure of the receptor-binding site are highly variable among different subtypes (Figure 2.5).

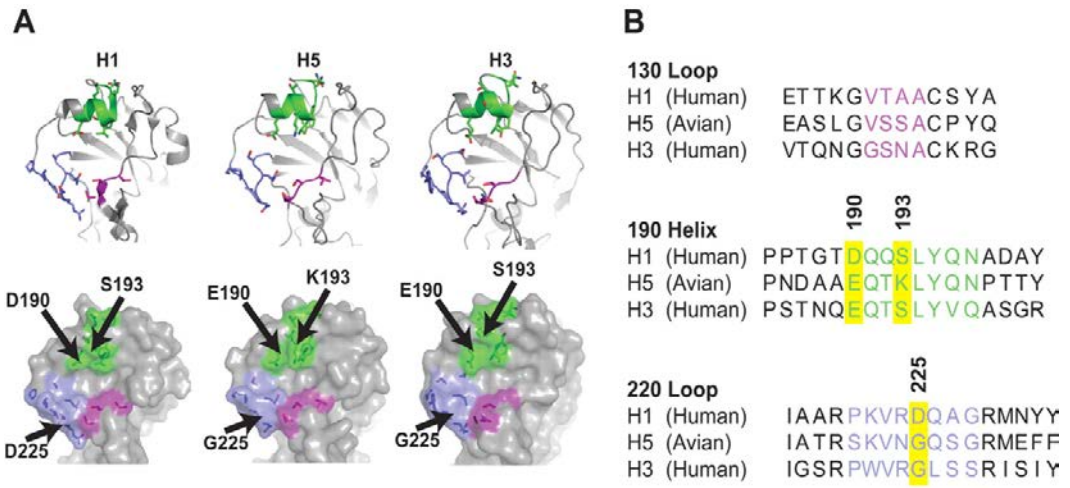


Figure 2.5: Receptor binding site on the globular head region of HA. A) Structural comparison of receptor binding site among H1, H5 and H3 subtypes. B) Sequence alignment of loop 130, helix 190 and loop 220 in receptor binding site from H1, H5 and H5 subtypes (Yang et al., 2007).

The global head region of HA also presents four conformational epitopes including Sa, Sb, Ca and Cb epitopes. The Sb epitope encompasses Helix 190 within the receptor-binding site (Figure 2.6) (Edwards and Dimmock, 2000; Xu *et al.*, 2010). A study found that monoclonal antibody targeting Sb site had protection in hemagglutinin inhibition assay titer in mice (Yu *et al.*, 2008). This evidence motivated this PhD study to exploit Helix 190 as a target antigen for influenza vaccine.

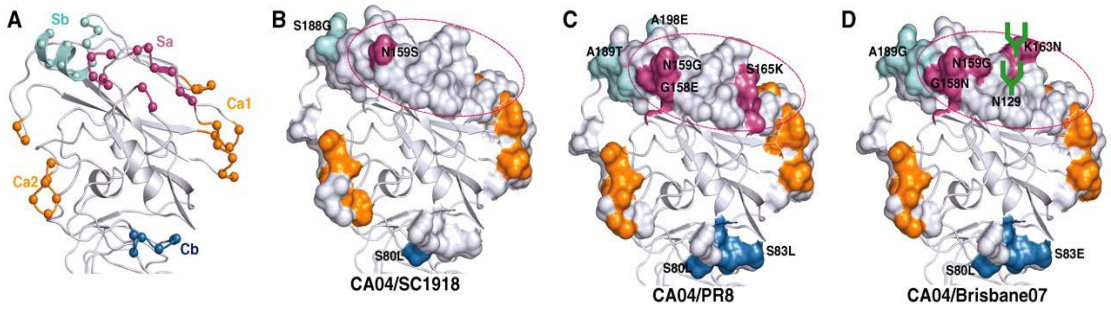


Figure 2.6: Antigenic sites on the globular head region. A) Location of antigenic sites on the globular head region from A/California/04/2009. Structural comparison of antigenic sites between A/California/04/2009 versus B) A/ South-Carolina/1/1918, C) A/Puerto-Rico/8/34, D) A/Brisbane/59/2007. Antigenic sites are highlighted in different colours, especially Sa site in magenta, Sb site in cyan, Ca site in orange, Cb site in blue (Xu et al., 2010).

2.2.2. HA2

Underneath the globular head region, the stalk region mainly consists of the HA2 subunit, which plays an important role in membrane fusion. At low pH in endosomal vesicles, the stalk region undergoes spring-loaded conformational change to facilitate membrane fusion. Some structural elements on the stalk region rearrange their positions or completely change their conformation as shown in Figure 2.7 (Skehel *et al.*, 1982). A study demonstrated that mutations in the stalk region led to fusion defeat (Gruenke *et al.*, 2002). The research found that HA2-specific monoclonal antibodies inhibited the fusion activity of HA (Varečková *et al.*, 2003). It suggested that antigenic sites in the stalk region of HA would be potential targets for influenza vaccine.

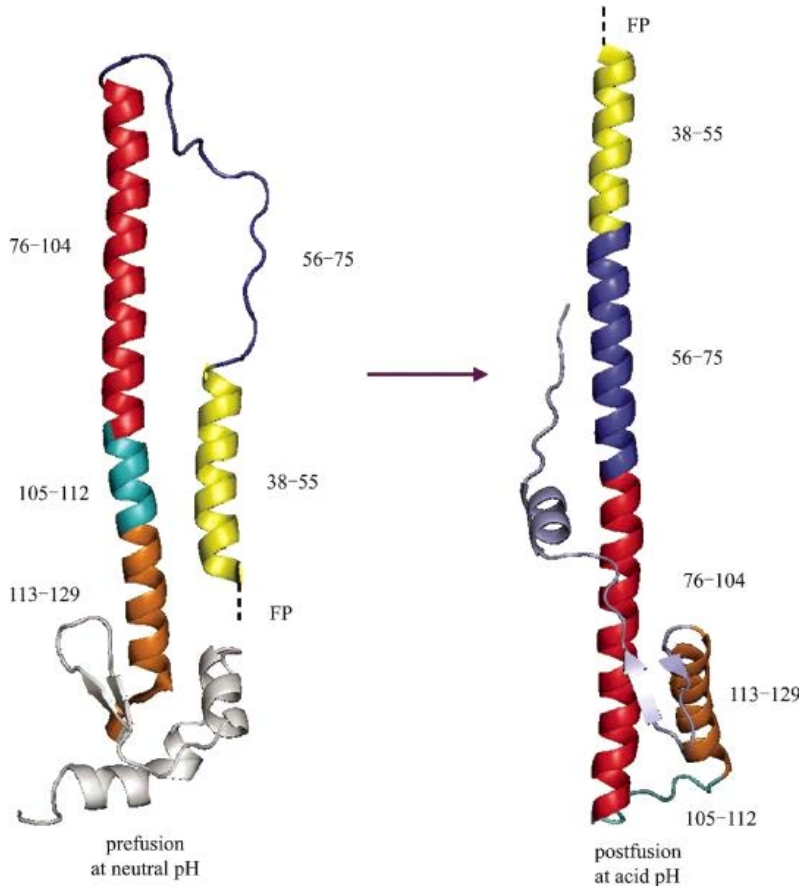


Figure 2.7: Conformational change of HA stalk region at low pH (Jiang et al., 2010).

Helix A is a short helix (residue 38-58) on the HA2 subunit that packs against the central helical bundle of HA (Wiley *et al.*, 1981). Helix A sequence is more highly conserved across different HA subtypes (Figure 2.8). Antibodies against this site can broadly react with different subtypes (Chen *et al.*, 2011). Some studies reveal that Helix A (HA2A) contributes to the interacting surface of neutralising antibodies including CR6261 (Ekiert *et al.*, 2009), F10 (Sui *et al.*, 2009), CR8082 (Ekiert *et al.*, 2011) and FI6v3 (Corti *et al.*, 2011). Particularly, HA2A accounts for 70% of contact area with the neutralising antibodies CR6261 (Ekiert *et al.*, 2009). The study recommended that

Helix A alone is sufficient to induce neutralising antibodies (Ekiert *et al.*, 2009). Although the research found that HA2A showed no protection in mice, there remains a question about the vaccine design of HA2A (Schneeman *et al.*, 2012). There is no clear evidence about the ability of HA2A in eliciting protective antibodies.

```
A/California/07/2009 (H1N1)     DLKSTQNAIDEITNKVNSVIEK
A/Puerto Rico/8/1934(H1N1)      -Q-------NG------T----
A/South Carolina/1/1918(H1N1)   -Q-------DG-----------
A/New Caledonia/20/1999(H1N1)   -Q-------NG-----------
A/Hong Kong/1/1968(H3N2)        ------A---Q-NG-L-R----
A/Viet Nam/1203/2004(H5N1)      -KE---K---GV------I-D-
```

Figure 2.8: Alignment sequence of Helix A derived from A/California/07/2009 against which of different subtypes.

2.3. Influenza vaccine manufacture

A big challenge for influenza vaccine production is the time needed for vaccine manufacturing versus the time of antigenic drift or shift in the virus. To protect humans and animals against the influenza virus, the time of vaccine production must be quicker than the time that the virus undergoes antigenic change (Figure 2.9). Influenza vaccine manufacturing technologies, thus, must consider the time-scale of vaccine production as a priority. The conventional influenza vaccines are based on growing the live viruses in embryonated chicken eggs. This manufacturing method responded slowly to influenza pandemics. For example, in the pandemic H1N1 2009, the virus spread globally to 43 nations with 86 deaths within a month while the egg-based vaccine was released five months afterwards. A new cell-based manufacturing method can reduce the production time to more than two months (Barrett, 2010), which is still considerably slower than the spread of the virus in the case of pandemic H1N1 2009. It suggests that those manufacturing methods would be ineffective to respond to an influenza pandemic.

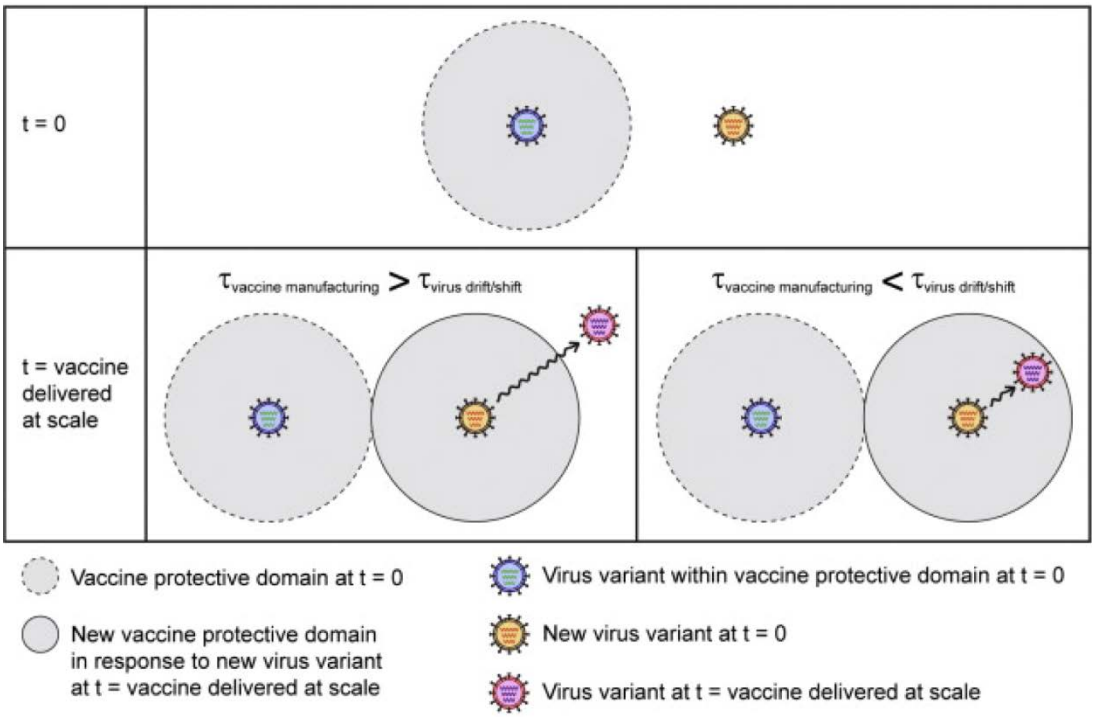


Figure 2.9: Immunoprotein against influenza virus depending on a time-scale of vaccine manufacturing versus time-scale of viral antigenic change (Wibowo et al., 2014).

2.4. Structure-based vaccines

Due to the safety concern and slow response of traditional vaccines, it reveals a necessity for a new vaccine approach to producing a safer vaccine and can rapidly re-engineer vaccine components that match to the circulating strains. A new trend in vaccine development is “structural vaccinology”. Structural vaccinology is an approach employing structural information obtained from protein structure determination techniques and bioinformatics tools to design immunogens to maintain an intact structure and improve vaccine efficacy (Dormitzer *et al.*, 2008). Several methods have been employed to guide the structure-based vaccine designs including molecular dynamics (MD) simulation (Joshi *et al.*, 2011), X-ray crystallography (McLellan *et al.*, 2013; Nuccitelli *et al.*, 2011), bioinformatics (Thomas *et al.*, 2011), and nuclear magnetic resonance (NMR) (Koide *et al.*, 2005).

Some influenza vaccine candidates have been developed based on this structure-based vaccine approach. Kanekiyo *et al.* demonstrated a structure-based nanoparticles vaccine, particularly HA protein incorporated into ferritin nanoparticles, induced broader and more potent immunity than a conventional influenza vaccine (Kanekiyo *et al.*, 2013). A study exploited structure-based vaccine design to engineering the HA head which encompasses neutralising antibody epitopes, leading to the induction of protective immunity in mice (Xuan *et al.*, 2011). Another study rationally designed the HA2 subunit to destabilise the low pH conformation of HA2, leading to mice protection in the

lethal challenge (Bommakanti *et al.*, 2010; Bommakanti *et al.*, 2012). Structure-based vaccine design approach also utilises epitopes in inducing an immune response (Anggraeni *et al.*, 2013; Schneeman *et al.*, 2012; Wibowo *et al.*, 2012) due to that the rationale is that a protective epitope is sufficient to elicit protection against pathogens. However, an antigenic peptide has low immunogenicity, and it requires protein carriers, e.g. virus-like particles (VLP), or adjuvant formulations to improve its immunogenicity (Jennings and Bachmann, 2009).

2.5. Virus-like particles (VLPs) based vaccine

2.5.1. Parental VLPs

VLPs are empty viral shells constituted by the self-assembly of major capsid protein; thus, they mimic the morphology of the pathogen. VLPs can stimulate both humoral and cellular immunity (Liu *et al.*, 2000; Pushko *et al.*, 2005). Initially, VLPs were used to induce an immune response against their parental viruses. Parental VLP vaccines released on the market including Engerix® (Keating and Noble, 2003) and Recombivax® (Bialek *et al.*, 2008) against hepatitis B virus, Cervarix® (Einstein *et al.*, 2009) and Gardasil® (Einstein *et al.*, 2009) against human papillomavirus. Several vaccine candidates based on VLP are in clinical development and produced in various expression systems (Table 2.1).

Table 2.1: VLP-based vaccines in pre-clinical and clinical trials.

Adapted from (Lua et al., 2014; Plummer and Manchester, 2011; Quan et al., 2016)

Virus	***Expression system***	***Recombinant protein***
Chikungunya virus	Baculovirus/insect cells	E1 and E2 glycoprotein subunits
Hepatitis B virus (HBV)	Bacteria, yeast, fungi, transgenic plants, mammalian cells	Surface antigen, core antigen, S protein, M protein, preS1, preS2
Human papillomavirus (HPV)	Bacteria, yeast, baculovirus/insect cells, transgenic plants	L1
Hepatitis E virus (HEV)	Baculovirus/insect cells, transgenic plants	Capsid protein
Influenza A virus	Baculovirus/insect cells	HA, NA, M1, M2, NP
Hepatitis C virus (HCV)	Mammalian cells	Core protein, E1, E2
Poliovirus	Baculovirus/insect cells	VP0, VP1, VP3
Human immunodeficiency virus (HIV)	Yeast, baculovirus/insect cells, mammalian cells	Pr55gag, Pr160gag-pol, gag protein, RT, TN
Ebola virus	Bacteria, baculovirus/insect cells, mammalian cells	VP40 and glycoprotein
Norwalk virus	Baculovirus/insect cells, transgenic plants	Capsid protein
Nipah virus	Cell strain HEK293	Proteins G, F and M
Rotavirus	Baculovirus/insect cells	VP2, VP6, VP7
Severe acute respiratory syndrome (SARS)	Baculovirus/insect cells	SP, EP, MP

2.5.2. Modular VLPs

Another approach using VLP based vaccines is modular VLP. A modular VLP is referred to as a VLP presenting antigenic peptides from an unrelated pathogen. The repetitive array property and particulate nature of VLP lead to multiple copies of antigen being presented on the surface, resulting in eliciting an antibody response against the targeted pathogen (Yin *et al.*, 2011). Table 2.2 describes numerous VLP platforms utilised to present various epitopes from different pathogens.

Table 2.2: Summary of modular VLPs. Adapted from (Lua et al., 2014; Plummer and Manchester, 2011)

VLP platform	***Target pathogen***	***Epitope***
Flock house virus (FHV)	Hepatitis C virus (HCV)	E1
	Bacillus anthrax	ANTXR2 PA-binding domain
	HIV	V3 loop, CCR5
Hepatitis B virus (HBV)	HIV	gp41, p34
	Plasmodium falciparum	CS protein
	Dengue virus	Env protein
	Influenza A	M2 protein
	Human papillomavirus (HPV)	CTL epitopes
Hepatitis E virus (HEV)	HEV	B cell epitope tag
Human papillomavirus (HPV)	HIV	SIV gag, HIV tat, and HIV rev
Bacteriophage Qβ	Allergen	Der p1
	HIV	CCR5
Cowpea mosaic virus (CPMV)	Canine parvovirus	VP2
	Pseudomonas aeruginosa	Outer membrane protein F
	Mink enteritis virus	Mink enteritis virus protein
	HIV	Gp41
Tobacco mosaic virus	Foot and mouth disease virus	VP1
	Murine hepatitis virus	Spike protein
	Pseudomonas aeruginosa	Outer membrane protein F
Cucumber mosaic virus	Hepatitis C virus	R10 epitope
Alfalfa mosaic virus	Respiratory syncytial virus (RSV)	G protein peptide
PapMV	Lymphocytic choriomeningitis virus (LCMV)	Peptide p33
Bacteriophage MS2	HIV	Gp20 or CCR5
Murine polyomavirus	Influenza	M2e, Helix 190, HA1
	Group A *streptococcus*	J8
	Rotavirus	VP8*
Newcastle disease virus	Respiratory syncytial virus	F and G proteins

2.6. The microbial vaccine platform at the University of Queensland

The Centre for Biomolecular Engineering at the University of Queensland (UQ) developed a microbial vaccine platform based on murine polyomavirus VP1 major capsid protein. Five VP1 monomers self-assemble into a capsomere; seventy-two capsomeres self-assemble into a VLP. Both VLP and capsomere have been employed to present heterologous antigens on the surfaces to induce antibodies against the antigens (Middelberg *et al.*, 2011; Rivera-Hernandez *et al.*, 2013; Tekewe *et al.*, 2016; Wibowo *et al.*, 2012) (Figure 2.10). The vaccine platform takes the advantages of a prokaryotic expression system regarding high-yield, speed, and low cost. Leavitt *et al.* (1985) firstly exploited *E.coli* to produce MuPyV VP1 capsomeres; however, the expression level was very poor. The expression level of MuPyV VP1 capsomere was then improved to milligram-per-litre (Chuan *et al.*, 2008) and further to gram-per-litre level (Liew *et al.*, 2010). Process simulation shows that this vaccine technology can produce 320 million doses in 4.7 days for VLP-based vaccines and 2.3 days for capsomere-based vaccines (Chuan *et al.*, 2014b). The cost analysis shows that MuPyV VLP- and capsomere-based vaccines can be produced at 1 cent/dose (Chuan *et al.*, 2014b). This vaccine platform would be promising for influenza vaccine development because it can quickly produce a vaccine to respond to the outbreak of influenza virus.

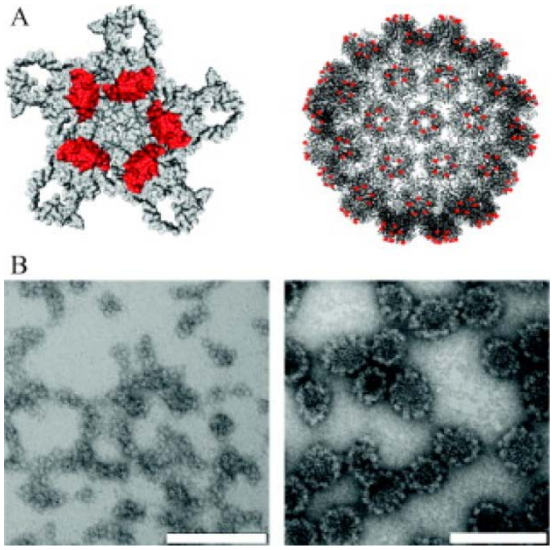


Figure 2.10: Murine polyomavirus modular VP1 displaying a heterologous antigen (red). A) Structural prediction for modular capsomere (left) and VLP (right). B) Transmission electron micrograph showing the self-assembly of modular capsomeres (left) into the modular VLPs (right). Scale bar is 100nm. (Middelberg et al., 2011).

2.6.1. Murine polyomavirus

Murine polyomavirus (MuPyV) belongs to the polyomavirus family. MuPyV mainly infects and causes a variety of tumours in mice (Tegerstedt *et al.*, 2005). The outer shell has an icosahedral lattice structure (T = 7d) in diameter of 50nm, which primarily contains the major capsid protein VP1 (42.5 kDa). VP1 protein constitutes the viral shell by the self-assembly of 72 capsomeres or pentamers (consisting of five VP1 monomers). Although the viral shell also comprises two minor structural proteins: VP2 (34 kDa) and VP3 (23 kDa), these proteins are not essential in the self-assembly of a capsid shell (Salunke *et al.*, 1986; Stehle and Harrison, 1997). It suggests that the major capsid VP1 protein can be exploited as a VLP for vaccine development.

The structure of MuPyV VP1 protein facilitates self-assembly into VLP. The crystal structure of MuPyV VP1 (at 1.9 Å resolution) predominantly comprises a β-sheet structure with "jelly-roll" topology as in Figure 2.11 (Stehle and Harrison, 1997; Stehle *et al.*, 1994). The study identified that both N- and C- termini take part in the assembly and change their structure after assembling (Stehle and Harrison, 1997). The removal of 63 residues at the C-terminus of the VP1 can lead to the loss of ability to assemble into a VLP. In the VLP assembly (Figure 2.12), the C-terminal arm of one capsomere invades a neighbouring capsomere. The N-terminal arm of the adjacent capsomere secures the invading C-terminal arm by forming a clamp. The association between 2 capsomeres completes when they finish exchanging C-terminal arms and create a clamp-clamp interaction across the local dyad (Stehle and Harrison, 1997).

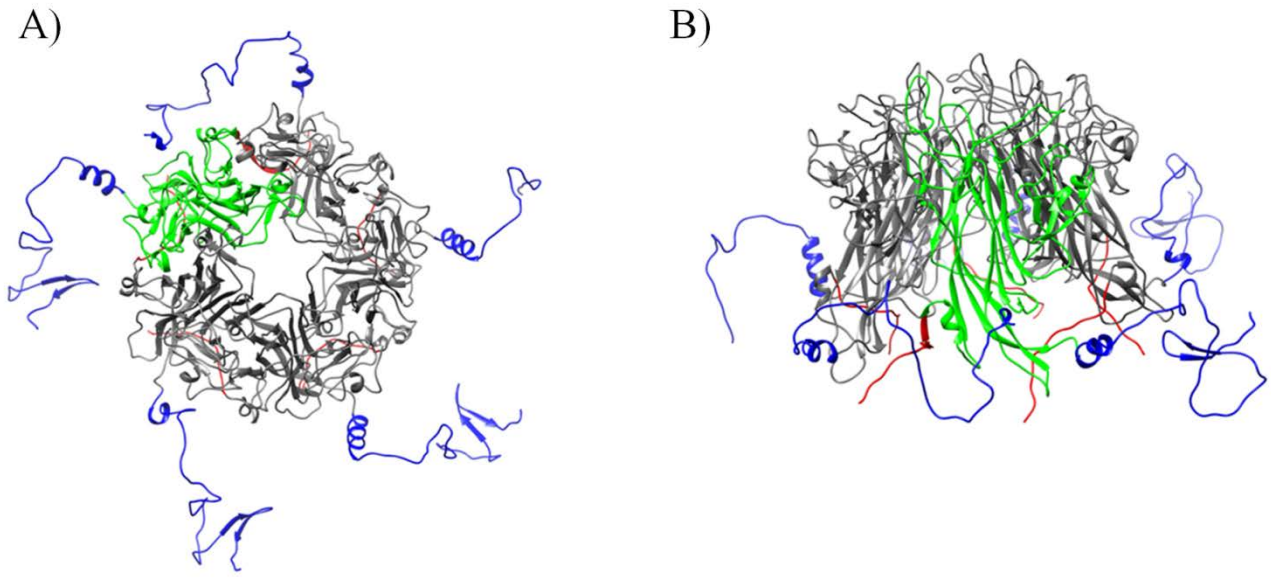


Figure 2.11: Crystal structure of MuPyV VP1 capsomere. A) Top view, B) Side view. One monomer was highlighted in green, N-terminal in red, and C-terminal in blue. The pictures were generated using UCSF Chimera software version 1.8 from 1SID.PDB.

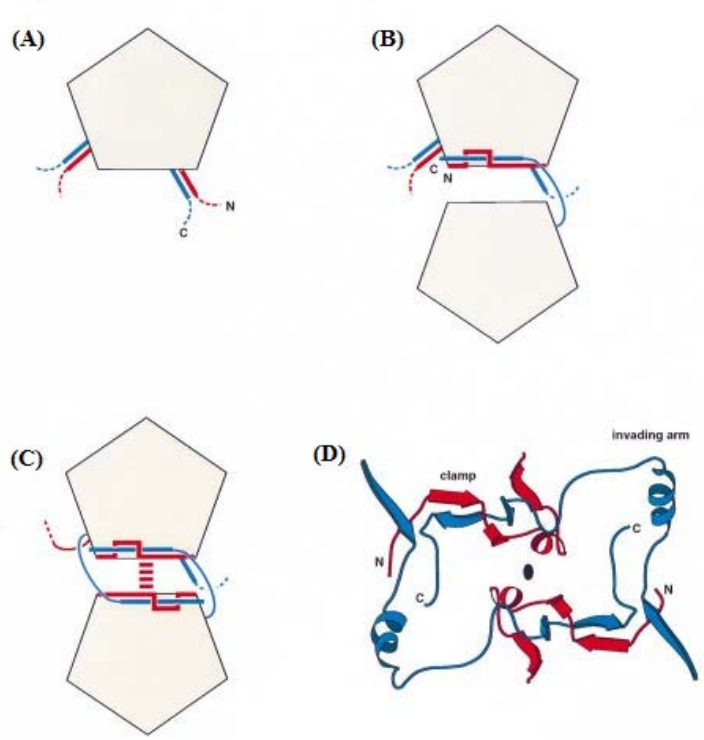


Figure 2.12: Self-assembly process between two MuPyV VP1 capsomeres via the interaction between C-terminal (blue) and N-terminal (red) (A) unassembled capsomere, B) and C) two steps in the self-assembly of two capsomeres. B) the self-assembly initials through the invasion of the C-terminal arm from one capsomere to the other and the formation of the N-terminal clamp to secure the interaction. C) then the invading capsomere exchanges its C-terminal arm with the other, and forms another N-terminal clamp across the dyal of capsomeres. D) ribbon drawing of the interaction occurring in C (Stehle and Harrison, 1997).

2.6.2. Engineering MuPyV VP1

Surface loops on VP1 were engineered with insertion sites to present heterologous antigens on the surface of the assembled VLPs. Stehle and Yan *et al.* (1994) show that MuPyV VP1 protein presents three protruding loops including BC2, DE, and HI loops (Figure 2.13A). Another study found that hamster polyomavirus (HaPV) VP1, which is closely related to MuPyV, has four insertion sites consisting of site 1, 2, 3 and 4 (Gedvilaite *et al.*, 2000). The comparison of VP1 sequence between these polyomaviruses shows the site 1 and site 4 of HaPV is situated at BC2- and HI-loops of MuPyV VP1, respectively. Figures 2.13B and 2.13C show site 1 in loop BC, and site 4 in HI loop expose on the surface of the capsomere. These sites are suitable for the presentation of heterologous antigens on the surface of VLP or capsomere (Stehle *et al.*, 1994). Several studies employed the HI loop to insert different heterologous antigens, particularly a protein domain dihydrofolate reductase (Gleiter *et al.*, 1999) or protein Z (Gleiter and Lilie, 2001) or a polyanionic adapter sequence (Stubenrauch *et al.*, 2001). Gedvilaite *et al.* (2000) ranked that site 4 has the highest potency to tolerate foreign epitopes, compared with the other sites (Gedvilaite *et al.*, 2000).

The insertion at site 4 led to the insolubility of chimeric VP1, reported in a study (Shin and Folk, 2003). Their study suggested that site 4 is a better insertion site compared to site 1. The contradiction between the studies proves a poor understanding of the insertion sites for the presentation of heterologous antigens on VLP or capsomere.

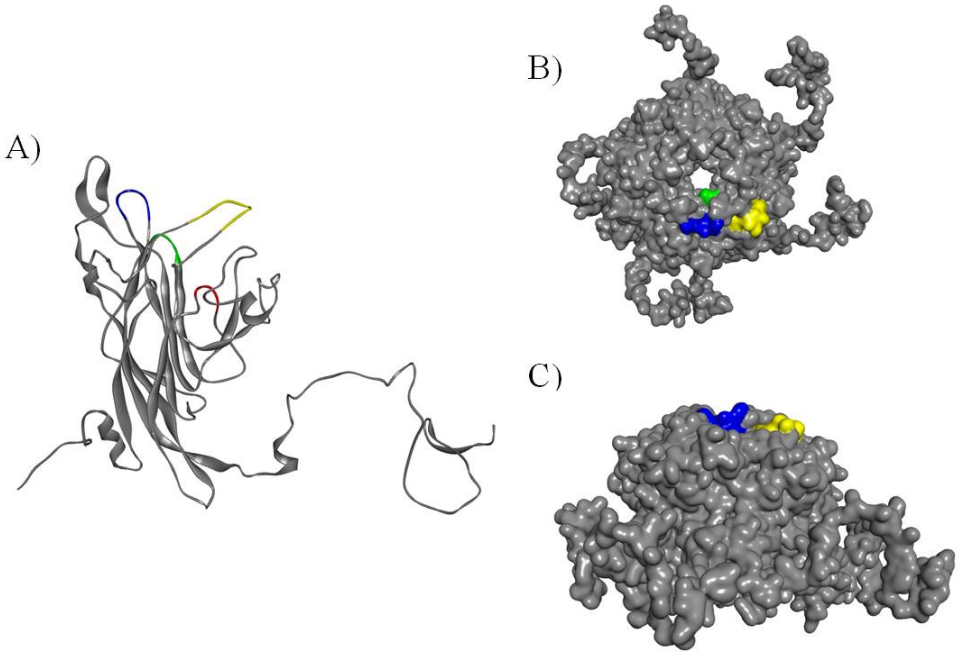


Figure 2.13: MuPyV VP1 insertion sites in comparison with the insertion sites on HaPV VP1. A) MuPyV VP1 monomer, B) MuPyV VP1 capsomere (top view), and C) MuPyV VP1 capsomere (side view). The insertion sites were highlighted in different colours, particularly site 1 in loop BC2 (yellow), site 2 (red), site 3 (green), and site 4 in loop HI (blue). The picture was generated by Natalie Connor using Accelrys Discovery Studio 3.0 based on 1SID.pbd.

Several studies reported the ability of modular VLPs and capsomeres in inducing antibody response against the heterogeneous epitopes or antigenic elements (Chuan *et al.*, 2014a; Dale *et al.*, 2002; Jain *et al.*, 2010; Liu *et al.*, 2000; Tissot *et al.*, 2010; Wibowo *et al.*, 2012). However, very limited studies characterise the conformation of an epitope on VLP (Joshi *et al.*, 2013; Schneeman *et al.*, 2012). The evidence indicates that there is a knowledge gap in designing epitope presentation in its native structure on the surface of an unrelated VLP. Determining the structure of antigenic elements presented is essential to guide the design of epitope presentations on VLPs. The next Section 2.7 will present a potential tool to aid the understanding of structural presentation of antigenic elements on VLP. Antibodies recognise antigens depending on the conformation of antigens (Sela-Culang *et al.*, 2013). An epitope-specific antibody would have the potential to identify the structure of the antigenic element on VLP. Here, isolating epitope-specific antibody using phage display technology is explored. The advent of phage display technology has facilitated the generation of phage antibody libraries. Epitope-specific antibodies can be isolated from a phage antibody library via phage antibody panning technique.

2.7. Phage display technology and phage antibody panning

2.7.1. Phage display technology

Smith (1985) first described phage display technology which involves displaying exogenous peptides, proteins and antibody fragments on the surface of filamentous *E. coli* bacteriophage (Smith, 1985). McCafferty *et al*. (1990) successfully presented a functional fragment of antibody on the bacteriophage surface (Figure 2.14) (McCafferty *et al.*, 1990). The genetic fusion of a single chain variable fragment (scFv) to the gp3 coat protein of a bacteriophage led to the presentation of scFv on the phage coat. Hoogenboom *et al*. (1991) successfully presented a larger and more complex antibody fragment, particularly antigen-binding fragment (Fab), on the phage surface (Hoogenboom *et al.*, 1991). The ability in displaying the antibody fragments on bacteriophage has resulted in the development of a phage antibody library.

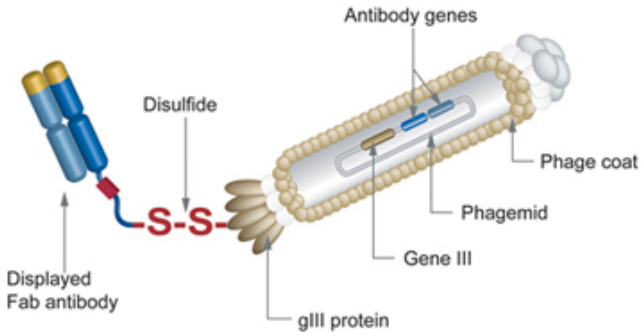


Figure 2.14: Presentation of Fab on the M13 phage surface (http://www.abdserotec.com/hucal-antibody-technology.html).

Phage antibody libraries are categorised into three types: non-immune, immune and synthetic libraries. A non-immune library was constructed from the B lymphocytes of unimmunized donors (Hoogenboom, 2002). This library is suitable for all antigens, even self and non-immunogenic antigens (Hoogenboom, 1997). An immune library originates from the B lymphocytes of immunised animals (Azzazy and Highsmith, 2002). The benefits of this library are the high specificity and affinity maturation of antibodies (Azzazy and Highsmith, 2002). In a synthetic library, oligonucleotide-directed mutagenesis or PCR-based techniques generate the diversity of complementary determining regions (CDRs) antibody (Bostrom and Fuh, 2009).

2.7.2. Phage antibody panning

Phage antibody panning is a technique to screen antigen-specific phage antibodies from a phage antibody library. Phage antibody panning has been widely exploited to screen targeting antibodies mainly for therapeutic purposes (Nielsen *et al.*, 2002; Popkov *et al.*, 2004; Zhang *et al.*, 2012). Phage antibody panning has been utilised to select antibodies against influenza antigens (Kashyap

et al., 2008). The identification of broadly neutralising antibodies against influenza hemagglutinin, e.g. F10 antibody (Sui *et al.*, 2009) and CR6261 antibody (Throsby *et al.*, 2008) via phage antibody panning was reported.

The process of phage antibody panning must be repeated for several panning rounds to enrich antigen-specific phage antibodies. Each round consists of coating, binding, washing, eluting, rescuing and amplifying steps (Figure 2.15). The conditions applied in the panning procedure reflect the stringency of selection. In early rounds, mild selection stringency involving a high antigen concentration and few washes provides an opportunity for the recovery of low-affinity binders. In later rounds, increasing the selection stringency by decreasing the antigen concentration and having more washes allows to the selection of high-affinity binders.

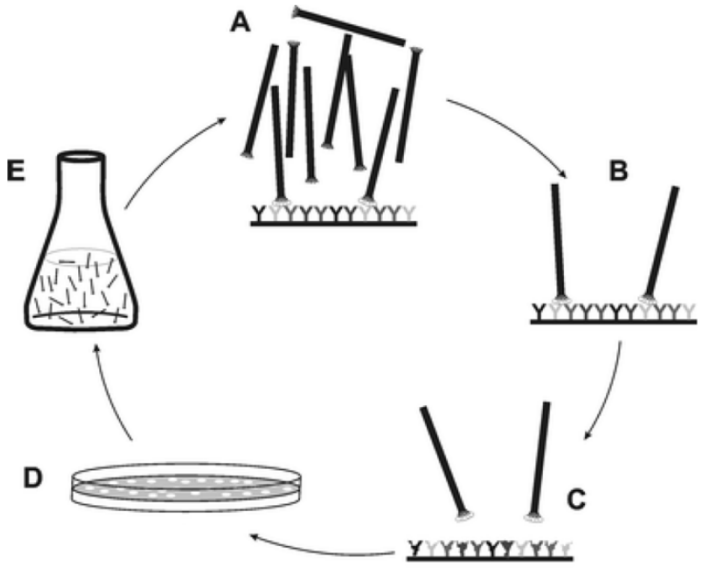


Figure 2.15: A typical phage antibody panning process including several steps: coating (A), binding (B), elution (C), rescue (D) and amplification (E)(Bratkovič, 2010).

Various types of surfaces have been exploited for panning to improve the efficiency of selection. The selection of antigen-specific phage antibodies is performed on immunotubes (Maxisorb) (Marks *et al.*, 1991), ELISA plates (Maxisorb) (Kang *et al.*, 1991), BIA sensorchips (Malmborg *et al.*, 1996), and columns with antigen-activated matrix (McCafferty *et al.*, 1990). Panning has also been performed with a biotinylated antigen (Hawkins *et al.*, 1992) or an antigen presented on the cell surface (De Kruif *et al.*, 1995).

Elution significantly affects affinity selection. A study demonstrated that high-affinity phage antibodies were lost in the elution step of the first round, leading to the failure of selection of the antigen-specific antibody (De Bruin *et al.*, 1999). Schier and Marks (1996) reported that the efficiency of elution to be influenced by the type of eluent (acidic, alkaline and chaotropic solutions) due to differences in phage antibody affinity, a kind of interaction between the antigen and antibody (Schier and Marks, 1996). Different acidic buffers were applied in the elution of panning including 0.1 M HCl pH 1.0 (Schier and Marks, 1996), 0.2 M glycine pH 2.2 (Goletz *et al.*, 2002), 0.76 M citric acid pH 2.8 (Marks *et al.*, 1993) and 0.1 M sodium acetate pH 2.8 containing 0.5 M NaCl; or an alkaline buffer such as 0.1 M triethylamine pH 12 (De Bruin *et al.*, 1999). The interaction between antibody phage and antigen can be disrupted by competition with excess antigen, as reported in (Krishnaswamy *et al.*, 2009). The use of Genenase I in the elution of panning was to cleave at the position between the antibody fragment and coat protein gp3 (Ward *et al.*, 1996).

The next sections will present two protein structure-determining technologies that can be used to identify the structure of antigenic element on capsomere or VLP. The first technique (Section 2.8) is circular dichroism spectroscopy using to determine the secondary structure of protein. The second technique (Section 2.9) is X-ray crystallography using to determine the tertiary structure of protein.

2.8. Circular dichroism spectroscopy

Circular dichroism (CD) spectroscopy is a rapid and convenient method that has been widely used to determine the secondary structure of peptides and proteins. The advent of computation and bioinformatics methods and protein reference data sets has facilitated the analysis of protein CD spectra to obtain the secondary structure of proteins. CD spectroscopy was reported as a tool for examining the conformational changes or stability of MuPyV VP1 protein (Yang and Teng, 1998a; Yang and Teng, 1998b; Yang and Teng, 1999). In these studies, CD spectroscopy associated with the protein secondary structure estimation program proved its ability in providing the ratios of secondary structures, especially α-helix, β-strand, β-turn and random coil. CD spectroscopy would be potentially appropriate to characterise the modularised antigenic element structure presented on

MuPyV modular VP1. This section will provide some basic principles of CD spectroscopy and available tools for analysing protein CD spectra to obtain protein secondary structures.

2.8.1. Definition of circular dichroism (CD) spectroscopy

CD is the unequal absorption of left-handed and right-handed circularly polarised light (Figure 2.16). Due to the relationship with absorbance, CD data are expressed regarding absorbance; the difference in molar absorbance ($\Delta\varepsilon$) is given by:

$$\Delta\varepsilon(\lambda) = \varepsilon_L(\lambda) - \varepsilon_R(\lambda)$$

where ε_L and ε_R are extinction coefficients for the left- and right-circularly polarised components at wavelengthλ. The unit of $\Delta\varepsilon$ is $M^{-1}\,cm^{-1}$; with M is the molar concentration.

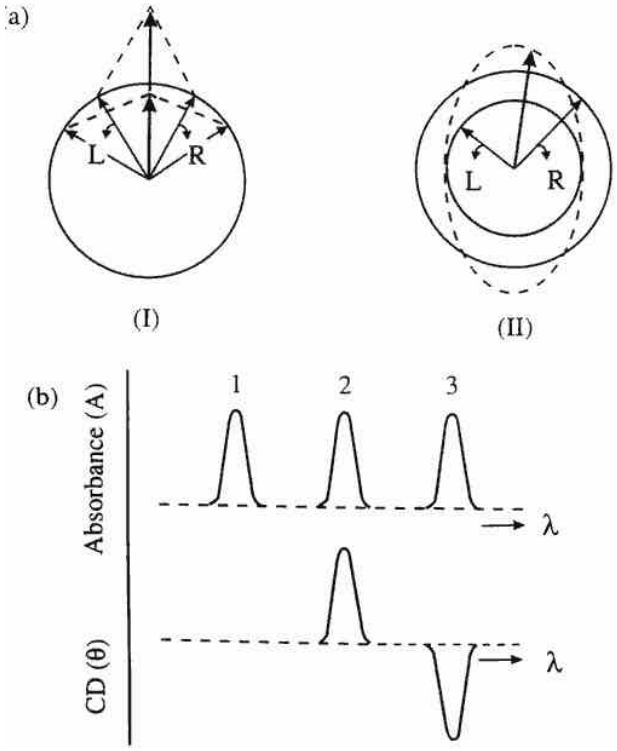


Figure 2.16: Circular dichroism definition. (a) Comparison of optical rotation and circular dichroism. (I) The left (L) and right (R) circularly polarised components have the same amplitude and when combined generate polarised plane radiation. (II) L and R components are of different magnitude, and the resultant (dashed line) is elliptically polarised. (b) The relationship between absorption and CD spectra. Band 1 is not chiral; band 2 has a positive CD spectrum with L absorbed more than R; band 3 has a negative CD spectrum (Sreerama and Woody, 2004).

Alternatively, CD is also reported in mean residue ellipticity [θ], which is defined as the angle whose tangent is the ratio of the minor to the major axis of the ellipse and is defined as:

$$[\theta] = \mathrm{MRW}\theta/10\mathrm{pc}$$

where MRW is mean residue weight, which is obtained by $\mathrm{MRW}=\frac{\text{Molecular weight}}{\text{Total amino acid}-1}$, θ (mdeg) is degree of ellipticity, p (cm) is path length and c (mg mL^{-1}) is the concentration of the sample. The unit of the molar ellipticity is deg cm^2 $dmol^{-1}$ or deg M^{-1} m^{-1}.

Molar ellipticity [θ] and difference in molar absorbance (Δε) are interconvertible by the formula:

$$[\theta] = 3298\Delta\varepsilon$$

A molecule exhibits a CD spectrum when that molecule is a chiral chromophore and placed in an optically asymmetric environment (Pain, 2001). Proteins generate CD spectrum when they meet both these requirements. Chromophores in protein consist of amide, aromatic side chain, disulphides and extrinsic groups. Among three main regions of CD spectra, the far-UV region CD of protein in the wavelength range of 250-150 nm reflects the secondary structure contents of protein due to amides that constitute peptide bonds are the main chromophores in this region (Wallace and Janes, 2009). The different kinds of secondary structure, e.g. α-helix, β-strand and random coil, have characteristic CD spectra. α-Helix shows negative values at two specific wavelengths 222 nm and 208 nm and a positive value at 193 nm (Holzwarth and Doty, 1965). While a β-strand has one negative band at 218 nm and a positive band at 195 nm, a random coil presents the negative band at 195 nm (Greenfield, 2006). CD spectroscopy was initially used to analyse peptide structure, then applied to examine the secondary structures of the protein.

2.8.2. Computational tools for estimation of protein secondary structures

Numerous algorithms have been developed to estimate the secondary conformation of a protein from CD spectra (Sreerama and Woody, 2000; Whitmore and Wallace, 2008; Yang and Teng, 1999). All of these computation methods assume that CD spectrum is a linear combination of spectra of individual secondary structure components, particularly α-helix, β-strand, β-turn and random coil. These methods utilise CD spectra from proteins whose structures are solved by X-ray crystallography as a reference dataset. The ratios of individual secondary structure components were estimated using various algorithms: least-squares analysis (Chang *et al.*, 1978); ridge regression (Provencher and Gloeckner, 1981); singular value decomposition (SVD) (Hennessey and Johnson, 1981); SVD with variable selection (Johnson, 1999; Manavalan and Johnson, 1987); the self-consistent method (Sreerama *et al.*, 2000; Sreerama and Woody, 2004); or neural network

analysis (Unneberg *et al.*, 2001). Table 2.3 compares the advantages and disadvantages of these estimation methods.

Table 2.3: The comparison of available estimation methods (Greenfield, 2006).

Software	***Method***	***Advantages***	***Disadvantages***
SELCON3	Self-consistent method	Good estimation of the globular protein structure	Poor estimation compared to SELCON1, SELCON2, VARSLC and CDNN
CONTIN, and CONTINLL	Ridge regression	Good estimation of α-helix and β-strand of protein and polypeptides	Complicated for quantitative analysis
CDSSTR, VARSLC	Variable selection	High flexibility, and good estimation of the globular protein structure	Not suitable for the analysis of polypeptide and protein fragments
K2D	Neutral networks	Good estimation of α-helix and β-strand of protein and polypeptides	Not estimate turn structures
CDNN	Neutral networks	Not applicable	Not suitable for analysis of polypeptides

CDSSTR software was developed by Manavalan and Johnson (1987) to overcome the difficulties of obtaining a good estimate of protein secondary structures. To have a successful analysis, the ratio value of secondary structure (f_i) firstly must be positive. Total secondary structures (Σf_i) must be equal to 1 or 100% structure. Thus, many studies directly introduced into their estimation methods the constraints that $f_i \geq 0$ and $\Sigma f_i = 1$, which may destroy the validity of the procedure (Manavalan and Johnson, 1985). Thus, CDSSTR employs a different selection method to solve the problems. CDSSTR systematically selected some proteins from a large database of reference proteins to generate a smaller database, thus, leading to the greater flexibility. Additionally, CDSSTR utilises singular value decomposition (SVD), previously described in two studies (Compton and Johnson Jr, 1986; Hennessey and Johnson, 1981) to estimate the secondary structure contents of protein. Due to the usage of the reduced data sets as a reference in structural estimation, SVD provides superior fits of the conformation of globular proteins (Greenfield, 2006). Furthermore, this estimation method requires some criteria to obtain good analyses. The sum of secondary structures should be close to 1.0. The minimal ratio of the individual secondary structure should be above –0.03. The reconstructed CD spectrum should fit the original CD spectrum with only a small error. The ratio of α-helix should be identical to the calculation achieved from the usage of all reference proteins (Johnson, 1999).

CD spectroscopy is a reliable method to obtain the secondary structure of a protein simply and quickly. CD spectroscopy requires smaller sample sizes for the analysis. However, the disadvantage

of CD spectroscopy is the low resolution of structural information. Although far-UV CD can estimate the proportions of α-helix, β-sheet, and β-turns, CD data is unable to indicate which regions of the protein are of which structural type. Therefore, to understand atomic structure of protein, it requires examining another method that can provide higher resolution. X-ray crystallography, a proven technique for determining the atomic structure of protein, will be presented in the next Section 2.9.

2.9. X-ray crystallography

The techniques for determination of protein atomic structure are nuclear magnetic resonance (NMR), cryo-electron microscopy (cryo-EM) and X-ray crystallography. X-ray crystallography is the most commonly utilised method to solve the three-dimensional structure of a protein. Protein Data Bank statistics show that 8670 protein structures solved by X-ray crystallography in a total of 8843 protein structures solved (Protein Data Bank, 2017). The stages of determining recombinant proteins structures utilising X-ray crystallography are protein preparation (cloning, expression, and purification), crystallisation, diffraction data collection, electron density analysis and model building. However, the most important and time-consuming step in X-ray crystallography is protein crystallisation to achieve a single high-order three-dimensional crystal.

Protein crystallisation refers to a process in which proteins form crystals. Although protein crystallisation was previously utilised to purify the protein, the most prevalent application of this technique today is to produce of single, high-quality crystals to determine the atomic structure of protein. Figure 2.17 illustrates the difficulty in achieving protein diffracting crystals and protein atomic structure determined by X-ray crystallography. It shows that 3.5% of protein forming diffracting crystals after an extreme crystallisation screening and 1.8% of protein determined atomic structure in a total of proteins desired to obtain a crystal structure. Different proteins have specific crystallisation conditions; thus, there never have any principles for choosing starting screening conditions. The screening of crystallisation conditions is tedious due to it requiring many parameters to be varied including precipitation, pH, and temperature. There is no indication that a condition is close to conditions for crystal formation until obtaining a crystalline and precipitant or the first crystals (Chayen *et al.*, 1996).

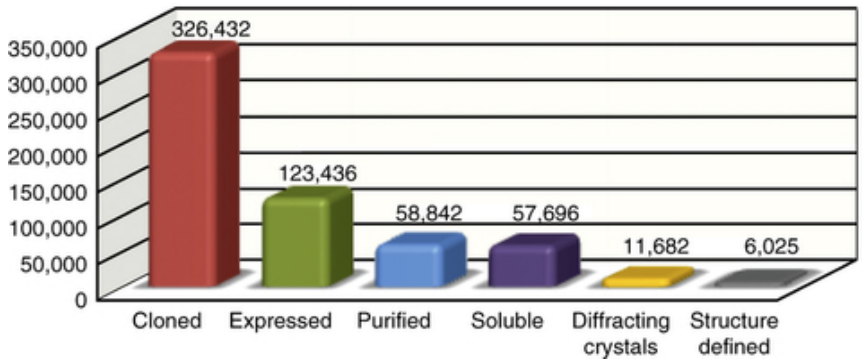


Figure 2.17: The diagram illustrates the difficulty of obtaining diffracting crystal and three-dimensional structure of a protein using X-ray crystallography (Khurshid et al., 2014).

Protein purity, concentration and dispersity can affect the success of protein crystallisation. Impurities can prevent nucleation, which is a requisite of crystal formation. It results in the required purity of proteins to be at least 95% (Pusey *et al.*, 2005). Protein concentration can affect the chance to identify conditions under which proteins achieve crystallisation phases. The concentration of protein needs to be at about 10 mg ml^{-1}, reported by (Blundell *et al.*, 2002). Protein monodispersity is also typically an indicator of good crystallisation propensity (Oberthuer *et al.*, 2012).

The phases of protein crystallisation include supersaturation (or precipitation zone), moderate supersaturation (or labile zone or nucleation zone), lower supersaturation (or metastable zone) and undersaturation zones (Figure 2.18). Proteins can only form a crystal in nucleation and metastable zones. The ordered aggregation of protein molecules form crystalline nuclei in the nucleation zone. The nucleation step can be highly possible to reach when protein concentration is sufficiently high so that it corresponds to nucleation or labile zone. Then, crystalline nuclei grow to a critical size. Crystal growth can take place in both labile zones and metastable zone; but, the latter is the best conditions for the growth of large, well-ordered crystals (Li and Chang, 2009).

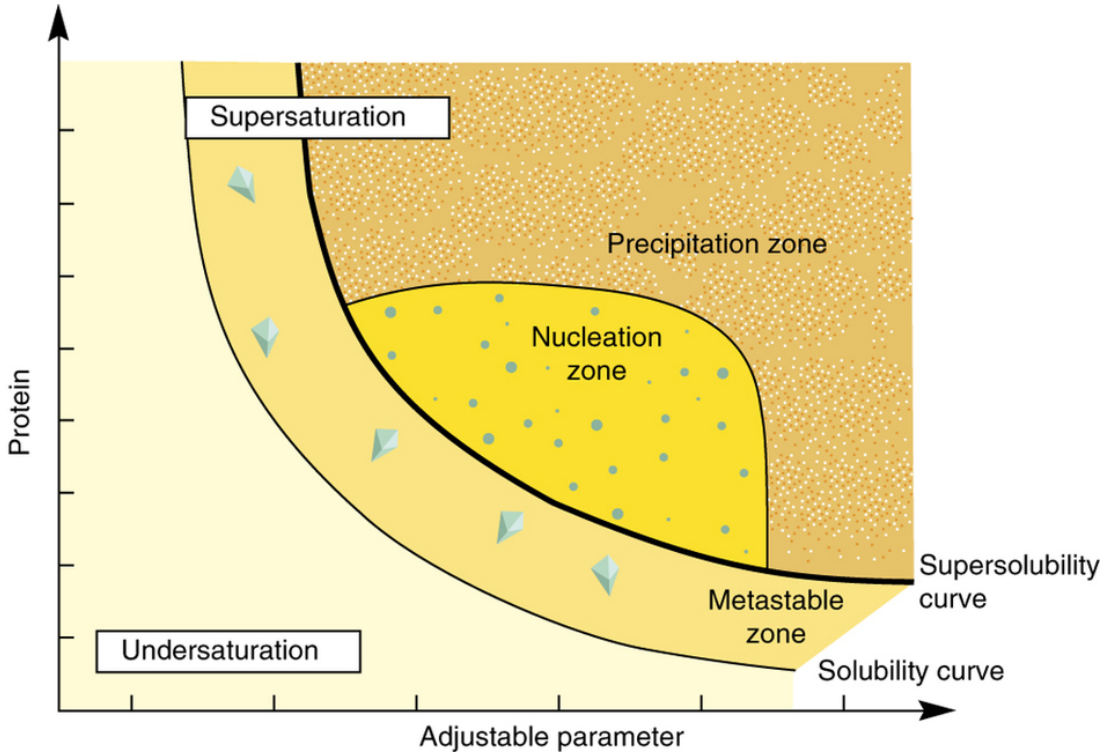


Figure 2.18: The diagram of crystallisation phases: (1) high supersaturation zone which the protein will precipitate. (2) moderate supersaturation zone which nucleation will occur. (3) lower supersaturation zone which crystals may grow, but no further nucleation will take place. (4) an undersaturated area which the protein is fully dissolved and will never crystallise (Chayen and Saridakis, 2008).

The manual screening of crystallisation conditions is laborious and time-consuming. Automated or robotic high-throughput screening has facilitated crystallisation screening (Chayen and Saridakis, 2002). Vapour diffusion, including the hanging-drop and sitting-drop vapour diffusion techniques, is the most commonly used technique (Chayen and Saridakis, 2008). Robotic automation has supported these techniques (Stevens, 2000). High-throughput crystallisation screening has been developed for micro-batch crystallisation (Chayen 1998). Dialysis is a method for crystallisation screening, but it is laborious to set up a screening (Figure 2.19) (Luft and DeTitta, 1997).

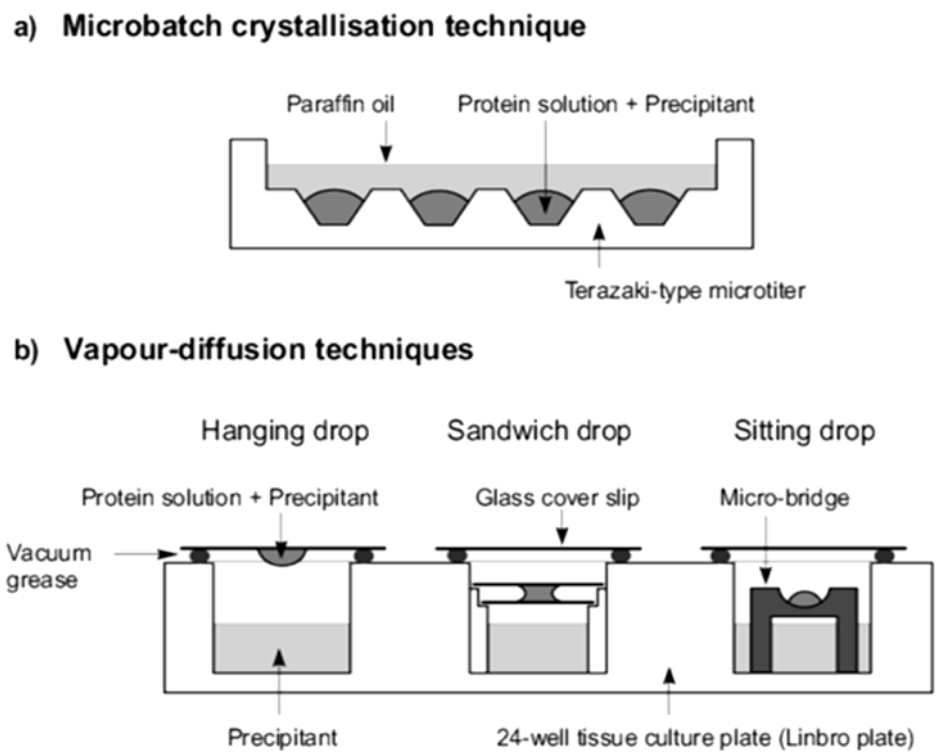


Figure 2.19: Crystallisation techniques.

(http://www.xray.bioc.cam.ac.uk/xray_resources/whitepapers/xtal-in-action/node3.html).

2.10. Concluding Remarks/ Summary

The rapid mutation of influenza viruses requires a vaccine manufacturing technology that can produce vaccine faster than the time that the antigenicity of the virus changes. Modular platform technology based on MuPyV, with advantages such as speed, safety and low cost, is a promising approach to develop an influenza vaccine. Two potential influenza HA epitopes of influenza virus, Helix 190 on HA1 and Helix A on HA2, were reviewed in this chapter. The structural presentation of these epitopes on modular VLP in their native structures is essential to induce neutralising antibodies. It is useful to establish design rules for conformational epitope presentation on a VLP. To define the design rule, tools that inform or characterise the structure of epitope presented on modular VLPs are essential. This chapter reviews some potential techniques for isolating antigen-specific antibodies and determining secondary and tertiary protein structure, including phage antibody panning, CD spectroscopy and X-ray crystallography, which can inform and characterise the structure of epitopes presented on the VLP's surface, respectively.

References

Anggraeni, M.R., Connors, N.K., Wu, Y., Chuan, Y.P., Lua, L.H.L., Middelberg, A.P.J., 2013. Sensitivity of immune response quality to influenza helix 190 antigen structure displayed on a modular virus-like particle. Vaccine 31, 4428-4435.

Azzazy, H.M.E., Highsmith, W.E.J., 2002. Phage display technology: clinical applications and recent innovations. Clin Biochem 35, 425-445.

Barrett, P.N., 2010. Development cell culture-derived pandemic vaccines. Curr Opin Mol Ther 12, 21-30.

Bertram, S., Glowacka, I., Steffen, I., Kühl, A., Pöhlman, S., 2010. Novel insights into proteolytic cleavage of influenza virus hemagglutinin. Rev Med Virol 20, 298-310.

Bialek, S.R., Bower, W.A., Novak, R., Helgenberger, L., Auerbach, S.B., Williams, I.T., Bell, B.P., 2008. Persistence of protection against Hepatitis B virus infection among adolescents vaccinated with recombinant Hepatitis B vaccine beginning at birth: a 15-year follow-up study. Pediatr Infect Dis J 27, 881-885.

Blundell, T.L., Jhoti, H., Abell, C., 2002. High-throughput crystallography for lead discovery in drug design. Nature 1, 45-54.

Bommakanti, G., Citron, M.P., Hepler, R.W., Callahan, C., Heidecker, G.J., Najar, T.A., Lu, X., Joyce, J.G., Shiver, J.W., Casimiro, D.R., Ter Meulen, J., Liang, X., Varadarajan, R., 2010. Design of an HA2-based *Escherichia coli* expressed influenza immunogen that protects mice from pathogenic challenge. Proc Natl Acad Sci USA 107, 13701-13706.

Bommakanti, G., Lu, X., Citron, M.P., Najar, T.A., Heidecker, G.J., ter Meulen, J., Varadarajan, R., Liang, X., 2012. Design of *Escherichia coli*-expressed stalk domain immunogens of H1N1 hemagglutinin that protect mice from lethal challenge. J Virol 86, 13434-13444.

Bostrom, J., Fuh, G., 2009. Design and construction of synthetic phage-displayed Fab libraries. Methods Mol Biol 562, 17-35.

Bratkovič, T., 2010. Progress in phage display: evolution of the technique and its application. Cell Mol Life Sci 67, 749-767.

Brooks, W.A., Goswami, D., Rahman, M., Nahar, K., Fry, A.M., Balish, A., Iftekharuddin, N., Azim, T., Xu, X., Klimov, A., Bresee, J., Bridges, C., Luby, S., 2010. Influenza is a major contributor to childhood pneumonia in a tropical developing country. Pediatr Infect Dis J 29, 216-221.

Brown, W.L., Mastico, R.A., Wu, M., Heal, K.G., Adams, C.J., Murray, J.B., Simpson, J.C., Lord, J.M., Taylor-Robinson, A.W., Stockley, P.G., 2002. RNA bacteriophage capsid-mediated drug delivery and epitope presentation. Intervirology 45, 371-380.

Centers for Disease Control and Prevention, 2009. Outbreak of swine-origin influenza A (H1N1) virus infection - Mexico, March-April 2009, Morb Mortal Wkly Rep, pp. 467-470.

Chang, C.T., Wu, C.S., Yang, J.T., 1978. Circular dichroic analysis of protein conformation: inclusion of the beta-turns. Anal Biochem 91, 13-31.

Chayen, N.E., Boggon, T.J., Cassetta, A., Deacon, A., Gleichmann, T., Habash, J., Harrop, S.J., Helliwell, J.R., Nieh, Y.P., Peterson, M.R., Raftery, J., Snell, E.H., Hadener, A., Niemann, A.C., Siddons, D.P., Stojanoff, V., Thompson, A.W., Ursby, T., Wulff, M., 1996. Trends and challenges in experimental macromolecular crystallography. Q Rev Biophys 29, 227-278.

Chayen, N.E., Saridakis, E., 2002. Protein crystallization for genomics: towards high-throughput optimization techniques. Acta Crystallogr D Biol Crystallogr 58, 921-927.

Chayen, N.E., Saridakis, E., 2008. Protein crystallization: from purified protein to diffraction-quality crystal. Nat Methods 5, 147-153.

Chen, J.-R., Ma, C., Wong, C.-H., 2011. Vaccine design of hemagglutinin glycoprotein against influenza. Trends Biotechnol 29, 426-434.

Chen, R., Holmes, E.C., 2006. Avian influenza virus exhibits rapid evolutionary dynamics. Mol Biol Evol 23, 2336-2341.

Cheng, V.C.C., To, K.K.W., Tse, H., Hung, I.F.N., Yuen, K.-Y., 2012. Two years after pandemic Influenza A/2009/H1N1: What have we learned? Clin Microbiol Rev 25, 223-263.

Chuan, Y.P., Lua, L.H.L., Middelberg, A.P.J., 2008. High-level expression of soluble viral structural protein in *Escherichia coli*. J Biotechnol 134, 64-71.

Chuan, Y.P., Wibowo, N., Connors, N.K., Wu, Y., Hughes, F.K., Batzloff, M.R., Lua, L.H.L., Middelberg, A.P.J., 2014a. Microbially synthesized modular virus-like particles and capsomeres displaying group A *streptococcus* hypervariable antigenic determinants. Biotechnol Bioeng 111, 1062-1070.

Chuan, Y.P., Wibowo, N., Lua, L.H.L., Middelberg, A.P.J., 2014b. The economics of virus-like particle and capsomere vaccines. Biochem Eng J 90, 255-263.

Compton, L.A., Johnson Jr, W.C., 1986. Analysis of protein circular dichroism spectra for secondary structure using a simple matrix multiplication. Anal Biochem 155, 155-167.

Corti, D., Voss, J., Gamblin, S.J., Codoni, G., Macagno, A., Jarrossay, D., Vachieri, S.G., Pinna, D., Minola, A., Vanzetta, F., Silacci, C., Fernandez-Rodriguez, B.M., Agatic, G., Bianchi, S., Giacchetto-Sasselli, I., Calder, L., Sallusto, F., Collins, P., Haire, L.F., Temperton, N., Langedijk, J.P.M., Skehel, J.J., Lanzavecchia, A., 2011. A neutralizing antibody selected from plasma cells that binds to group and group 2 influenza A hemagglutinins. Science 333, 850-856.

Dale, C.J., Liu, X.S., De Rose, R., Purcell, D.F.J., Anderson, J., Xu, Y., Leggatt, G.R., Frazer, I.H., Kent, S.J., 2002. Chimeric human papilloma virus-simian/human immunodeficiency virus virus-like particle vaccines: Immunogenicity and protective efficacy in macaques. Virology 301, 176-187.

Dawood, F.S., Jain, S., Shaw, M., 2009. Emergence of a novel swine-origin influenza A (H1N1) virus in humans. N Engl J Med 360, 2605-2615.

De Bruin, R., Spelt, K., Mol, J., Koes, R., Quattrocchio, F., 1999. Selection of high-affinity phage antibodies from phage display libraries. Nature 17, 397-399.

De Kruif, J., Terstappen, L., Boel, E., Logtenberg, T., 1995. Rapid selection of cell subpopulation-specific human monoclonal antibodies from a synthetic phage antibody library. Proc Natl Acad Sci USA 92, 3938-3942.

Dormitzer, P.R., Ulmer, J.B., Rappuoli, R., 2008. Structure-based antigen design: a strategy for next generation vaccines. Cell 26, 659-667.

Edwards, M.J., Dimmock, N.J., 2000. Two Influenza A virus-specific Fabs neutralize by inhibiting virus attachment to target cells, while neutralization by their IgGs is complex and occurs simultaneously through fusion inhibition and attachment inhibition. Virology 278, 423-435.

Einstein, M.H., Baron, M., Levin, M.J., Chatterjee, A., Edwards, R.P., Zepp, F., Carletti, I., Dessy, F.J., Trofa, A.F., Schuind, A., 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. Hum Vaccin 5, 705-719.

Ekiert, D.C., Bhabha, G., Elisliger, M.A., Friesen, R.H.E., Jongeneelen, M., Throsby, M., Goudsmit, J., Wilson, I.A., 2009. Antibody recognition of a highly conserved influenza virus epitope. Science 324, 246-251.

Ekiert, D.C., Friesen, R.H.E., Bhabha, G., Kwaks, T., Jongeneelen, M., Yu, W., Ophorst, C., Cox, F., Korse, H.J.W.M., Brandenburg, B., Vogels, R., Brakenhoff, J.P.J., Kompier, R., Koldijk, M.H., Cornelissen, L., Poon, L.L.M., Peiris, M., Koudstaal, W., Wilson, I.A., Goudsmit, J., 2011. A highly conserved neutralizing epitope on group 2 influenza A viruses. Science 333, 843-850.

Garten, R.J., Davis, C.T., Russell, C.A., Shu, B., Lindstrom, S., Balish, A., Sessions, W.M., Xu, X., Skepner, E., Deyde, V., Okomo-Adhiambo, M., Gubareva, L., Barnes, J., Smith, C.B., Emery, S.L., Hillman, M.J., Rivailler, P., Smagala, J., de Graaf, M., Burke, D.F., Fouchier, R.A., Pappas, C., Alpuche-Aranda, C.M., Lopez-Gatell, H., Olivera, H., Lopez, I., Myers, C.A., Faix, D., Blair, P.J., Yu, C., Keene, K.M., Dotson, P.D., Jr., Boxrud, D., Sambol, A.R., Abid, S.H., St George, K., Bannerman, T., Moore, A.L., Stringer, D.J., Blevins, P., Demmler-Harrison, G.J., Ginsberg, M., Kriner, P., Waterman, S., Smole, S., Guevara, H.F., Belongia, E.A., Clark, P.A., Beatrice, S.T., Donis, R., Katz, J., Finelli, L., Bridges, C.B., Shaw, M., Jernigan, D.B., Uyeki, T.M., Smith, D.J., Klimov, A.I., Cox, N.J., 2009. Antigenic and genetic characteristics of swine-origin 2009 A(H1N1) influenza viruses circulating in humans. Science 325, 197-201.

Gedvilaite, A., Frommel, C., Sasnauskas, K., Micheel, B., M., O., Bersing, O., Staniulis, J., Scherneck, S., Ulrich, R., 2000. Formation of immunogenic virus-like particles by inserting epitopes into surface-exposed regions of hamster polyomavirus major capsid protein. Virology 273, 21-35.

Gleiter, S., Lilie, H., 2001. Coupling of antibodies via protein Z on modified polyoma virus-like particles. Protein Sci 10, 434-444.

Gleiter, S., Stubenrauch, K.A.Y., Lilie, H., 1999. Changing the surface of a virus shell fusion of an enzyme to polyoma VP1. Protein Sci 8, 2562-2569.

Goletz, S., Christensen, P.A., Kristensen, P., Blohm, D., Tomlinson, I., Winter, G., Karsten, U., 2002. Selection of large diversities of antiidiotypic antibody fragments by phage display. J Mol Biol 315, 1087-1097.

Greenfield, N.J., 2006. Using circular dichroism spectra to estimate protein secondary structure. Nat Protoc 1, 2876-2890.

Gruenke, J.A., Armstrong, R.T., Newcomb, W.W., Brown, J.C., White, J.M., 2002. New insights into the spring-loaded conformational change of influenza virus hemagglutinin. J Virol 76, 4456-4466.

Hawkins, R.E., Russell, S., Winter, G., 1992. Selection of phage antibodies by binding affinity mimicking affinity maturation. J Mol Biol 226, 889-896.

Hennessey, J.P., Johnson, W.C., 1981. Information content in the circular dichroism of proteins. Biochemistry 20, 1085-1094.

Holzwarth, G., Doty, P., 1965. The ultraviolet circular dichroism of polypeptides. J Am Chem Soc 87, 218-228.

Hoogenboom, H.R., 1997. Designing and optimizing library selection strategies for generating high-affinity antibodies. Trends Biotechnol 15, 62-70.

Hoogenboom, H.R., 2002. Overview of antibody phage-display technology and its applications. Methods Mol Biol 178, 1-37.

Hoogenboom, H.R., Griffiths, A.D., Johnson, K.S., Chiswell, D.J., Hudson, P., Winter, G., 1991. Multi-subunit proteins on the surface of filamentous phage: methodologies for displaying antibody (Fab) heavy and light chains. Nucleic Acids Res 19, 4133-4137.

Ito, T., Gorman, O.T., Kawaoka, Y., Bean, W.J., Webster, R.G., 1991. Evolutionary analysis of the influenza A virus M gene with comparison of the M1 and M2 proteins. J Virol 65, 5491-5498.

Jain, S., Patrick, A.J., Rosenthal, K.L., 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, 7070-7080.

Jennings, G.T., Bachmann, M.F., 2008. The coming of age of virus-like particles vaccines. Biol Chem 389, 521-536.

Jennings, G.T., Bachmann, M.F., 2009. Immunodrugs: Therapeutics VLP-based vaccines for chronic diseases. Annu Rev Pharmacol Toxicol 49, 303-326.

Jiang, S., Li, R., Du, L., Liu, S., 2010. Roles of the hemagglutinin of influenza A virus in viral entry and development of antiviral therapeutics and vaccines. Protein Cell 1, 342.

Johnson, W.C., 1999. Analyzing protein circular dichroism spectra for accurate secondary structures. Proteins: Struct, Funct Bioinf 35, 307-312.

Joshi, H., Cheluvaraja, S., Somogyi, E., Brown, D.R., Ortoleva, P., 2011. A molecular dynamics study of loop fluctuation in human papillomavirus type 16 virus-like particles: a possible indicator of immunogenicity. Vaccine 29, 9423-9430.

Joshi, H., Lewis, K., Singharoy, A., Ortoleva, P.J., 2013. Epitope engineering and molecular metrics of immunogenicity: A computational approach to VLP-based vaccine design. Vaccine 31, 4841-4847.

Kanekiyo, M., Wei, C.-J., Yassine, H.M., McTamney, P.M., Boyington, J.C., Whittle, J.R.R., Rao, S.S., Kong, W.-P., Wang, L., Nabel, G.J., 2013. Self-assembling influenza nanoparticle vaccines elicit broadly neutralizing H1N1 antibodies. Nature 499, 102-106.

Kang, A.S., Barbas, C.F., Janda, K.D., Benkovic, S.J., Lerner, R.A., 1991. Linkage of recognition and replication functions by assembling combinatorial antibody Fab libraries along phage surfaces. Proc Natl Acad Sci USA 88, 4363-4366.

Karlsson Hedestam, G.B., Fouchier, R.A.M., Phogat, S., Burton, D.R., Sodroski, J., Wyatt, R.T., 2008. The challenges of eliciting neutralizing antibodies to HIV-1 and to influenza virus. Nat Rev Micro 6, 143-155.

Kashyap, A.K., Steel, J., Oner, A.F., Dillon, M.A., Swale, R.E., Wall, K.M., Perry, K.J., Faynboym, A., Ilhan, M., Horowitz, M., Horowitz, L., Palese, P., Bhatt, R.R., Lerner, R.A., 2008. Combinatorial antibody libraries from survivors of the Turkish H5N1 avian influenza outbreak reveal virus neutralization strategies. Proc Natl Acad Sci USA 105, 5986-5991.

Keating, G.M., Noble, S., 2003. Recombinant hepatitis B vaccine (Engerix-B): a review of its immunogenicity and protective efficacy against hepatitis B. Drugs 63, 1021-1051.

Keller, S.A., Bauer, M., Manolova, V., Muntwiler, S., Saudan, P., Bachmann, M.F., 2010. Cutting edge: Limited specialization of dendritic cell subsets for MHC class II-associated presentation of viral particles. J Immunol 184, 26-29.

Khurshid, S., Saridakis, E., Govada, L., Chayen, N.E., 2014. Porous nucleating agents for protein crystallization. Nat. Protocols 9, 1621-1633.

Koide, S., Yang, X., Huang, X., Dunn, J.J., Luft, B.J., 2005. Structure-based design of a second-generation Lyme disease vaccine based on a C-terminal fragment of Borrelia burgdorferi OspA. J Mol Biol 350, 290-299.

Krishnaswamy, S., Kabir, M.E., Miyamoto, M., Furuichi, Y., Komiyama, T., 2009. Cloning antifungal single chain fragment variable antibodies by phage display and competitive panning elution. Anal Biochem 395, 16-24.

Li, M., Chang, W., 2009. Protein crystallization. Photosynth Res 102, 223-229.

Liew, M.W.O., Rajendran, A., Middelberg, A.P.J., 2010. Microbial production of virus-like particles vaccine protein at gram-per-liter levels. J Biotechnol 150, 224-231.

Liu, W.J., Liu, X.S., Zhao, K.N., Leggatt, G.R., Frazer, I.H., 2000. Papillomavirus virus-like particles for the delivery of multiple cytotoxic T cell epitopes. Virology 273, 374-382.

Lua, L.H.L., Connors, N.K., Sainsbury, F., Chuan, Y.P., Wibowo, N., Middelberg, A.P.J., 2014. Bioengineering virus-like particles as vaccines. Biotechnol Bioeng 111, 425-440.

Luft, J.R., DeTitta, G.T., 1997. Kinetic aspects of macromolecular crystallization. Methods Enzymol 276, 110-131.

Malmborg, A.C., Duenas, M., Ohlin, M., Soderlind, E., Borrebaeck, C.A.K., 1996. Selection of binders from phage displayed antibody libraries using the BIAcore^{TM} biosensor. J Immunol Methods 198, 51-57.

Manavalan, P., Johnson, W.C., 1985. Protein secondary structure from circular dichroism spectra. J Biosci 8, 141-149.

Manavalan, P., Johnson, W.C., Jr., 1987. Variable selection method improves the prediction of protein secondary structure from circular dichroism spectra. Anal Biochem 167, 76-85.

Marks, J.D., Hoogenboom, H.R., Bonnert, T.P., McCafferty, J., Griffith, A.D., Winter, G., 1991. By-passing immunization human antibodies from v-gene libraries displayed on phage. J Mol Biol 222, 581-597.

Marks, J.D., Ouwehand, W.H., Bye, J.M., Finnern, R., Gorick, B.D., Voak, D., Thorpe, S.J., Hughes-Jones, N.C., Winter, G., 1993. Human antibody fragments specific for human blood group antigens from a phage display library. Nature 11, 1145-1149.

McCafferty, J., Griffiths, A.D., Winter, G., Chiswell, D.J., 1990. Phage antibodies: filamentous phage displaying antibody variable domains. Nature 348, 552-554.

McLellan, J.S., Chen, M., Joyce, M.G., Sastry, M., Stewart-Jones, G.B., Yang, Y., Zhang, B., Chen, L., Srivatsan, S., Zheng, A., Zhou, T., Graepel, K.W., Kumar, A., Moin, S., Boyington, J.C., Chuang, G.Y., Soto, C., Baxa, U., Bakker, A.Q., Spits, H., Beaumont, T., Zheng, Z., Xia, N., Ko, S.Y., Todd, J.P., Rao, S., Graham, B.S., Kwong, P.D., 2013. Structure-based design of a fusion glycoprotein vaccine for respiratory syncytial virus. Science 342, 592-598.

Middelberg, A.P.J., Rivera-Hernandez, T., Wibowo, N., Lua, L.H.L., Fan, Y., Magor, G., Chang, C., Chuan, Y.P., Good, M.F., Batzloff, M.R., 2011. A microbial platform for rapid and low-cost virus-like particle and capsomere vaccines. Vaccine 29, 7154-7162.

Moorthy, M., Samuel, P., Peter, J.V., Vijayakumar, S., Sekhar, D., Verghese, V.P., Agarwal, I., Moses, P.D., Ebenezer, K., Abraham, O.C., Thomas, K., Mathews, P., Mishra, A.C., Lal, R., Muliyil, J., Abraham, A.M., 2012. Estimation of the burden of pandemic (H1N1) 2009 in developing countries: Experience from a tertiary care center in South India. PLoS One 7, e41507.

Neumann, G., Noda, T., Kawaoka, Y., 2009. Emergence and pandemic potential of swine-origin H1N1 influenza virus. Nature 459, 931-939.

Nielsen, U.B., Kirpotin, D.B., Pickering, E.M., Hong, K., Park, J.W., Refaat Shalaby, M., Shao, Y., Benz, C.C., Marks, J.D., 2002. Therapeutic efficacy of anti-ErbB2 immunoliposomes targeted by a phage antibody selected for cellular endocytosis. Biochim Biophys Acta - Mol Cell Res 1591, 109-118.

Nuccitelli, A., Cozzi, R., Gourlay, L.J., Donnarumma, D., Necchi, F., Norais, N., Telford, J.L., Rappuoli, R., Bolognesi, M., Maione, D., Grandi, G., Rinaudo, C.D., 2011. Structure-based approach to rationally design a chimeric protein for an effective vaccine against Group B *Streptococcus* infections. Proc Natl Acad Sci USA 108, 10278-10283.

Oberthuer, D., Melero-García, E., Dierks, K., Meyer, A., Betzel, C., Garcia-Caballero, A., Gavira, J.A., 2012. Monitoring and scoring counter-diffusion protein crystallization experiments in capillaries by *in situ* dynamic light scattering. PLoS One 7, e33545.

Pain, R., 2001. Determining the CD spectrum of a protein. Curr Protoc in Protein Sci 38, 7.6.1-7.6.24.

Plummer, E.M., Manchester, M., 2011. Viral nanoparticles and virus-like particles: platform for contemporary vaccine design. Wiley Interdiscip Rev Nanomed Nanobiotechnol 3, 174-196.

Popkov, M., Rader, C., Barbas, C.F., 2004. Isolation of human prostate cancer cell reactive antibodies using phage display technology. J Immunol Methods 291, 137-151.

Protein Data Bank, 2017. PDB statistics.

Provencher, S.W., Gloeckner, J., 1981. Estimation of globular protein secondary structure from circular dichroism. Biochemistry 20, 33-37.

Pusey, M.L., Liu, Z.-J., Tempel, W., Praissman, J., Lin, D., Wang, B.-C., Gavira, J.A., Ng, J.D., 2005. Life in the fast lane for protein crystallization and X-ray crystallography. Prog Biophys Mol Biol 88, 359-386.

Pushko, P., Tumpey, T.M., Bu, F., Knell, J., Robinson, R., Smith, G., 2005. Influenza virus-like particles comprised of the HA, NA and M1 proteins of H9N2 influenza virus induce protective immune responses in BALB/c mice. Vaccine 23, 5751-5759.

Quan, F.-S., Lee, Y.-T., Kim, K.-H., Kim, M.-C., Kang, S.-M., 2016. Progress in developing virus-like particle influenza vaccines. Expert Rev Vaccines 15, 1281-1293.

Rivera-Hernandez, T., Hartas, J., Wu, Y., Chuan, Y.P., Lua, L.H.L., Good, M., Batzloff, M.R., Middelberg, A.P.J., 2013. Self-adjuvanting modular virus-like particles for mucosal vaccination against group A *streptococcus* (GAS). Vaccine 31, 1950-1955.

Salunke, D.M., Caspar, D.L.D., Garcea, R.L., 1986. Self-assembly of purified polyomavirus capsid protein VP1. Cell 46, 895-904.

Schier, R., Marks, J.D., 1996. Efficient *in vitro* affinity maturation of phage antibodies using BIAcore guided selections. Hum Antibodies Hybridomas 7, 97-105.

Schneeman, A., Speir, J.A., Tan, G.S., Khayat, R., Ekiert, D.C., Matsuoka, Y., Wilson, I.A., 2012. A virus-like particle that elicits cross-reactive antibodies to the conserved stem of influenza virus hemagglutinin. J Virol 86, 11686-11697.

Sela-Culang, I., Kunik, V., Ofran, Y., 2013. The structural basis of antibody-antigen recognition. Front Immunol 4, 302.

Shin, Y.C., Folk, W.R., 2003. Formation of polyomavirus-like particles with different VP1 molecules that bind the urokinase plasminogen activator receptor. J Virol 77, 11491-11498.

Skehel, J.J., Bayley, P.M., Brown, E.B., Martin, S.R., Waterfield, M.D., White, J.M., Wilson, I.A., Wiley, D.C., 1982. Changes in the conformation of influenza virus hemagglutinin at the pH optimum of virus-mediated membrane fusion. Proc Natl Acad Sci USA 79, 968-972.

Skehel, J.J., Wiley, D.C., 2000. Receptor binding and membrane fusion in virus entry: the influenza hemagglutinin. Annu Rev Biochem 69, 531-569.

Smith, G.P., 1985. Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. Science 228, 1315-1317.

Sreerama, N., Venyaminov, S.Y., Woody, R.W., 2000. Estimation of protein secondary structure from circular dichroism spectra: inclusion of denatured proteins with native proteins in the analysis. Anal Biochem 287, 243-251.

Sreerama, N., Woody, R.W., 2000. Estimation of protein secondary structure from circular dichroism spectra: Comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. Anal Biochem 287, 252-260.

Sreerama, N., Woody, R.W., 2004. Computation and analysis of protein circular dichroism spectra, in: Ludwig, B., Michael, L.J. (Eds.), Methods Enzymol. Academic Press, pp. 318-351.

Staneková, Z., Varečková, E., 2010. Conserved epitopes of influenza A virus inducing protective immunity and their prospects for universal vaccine development. Virology 7, 351.

Steel, J., Lowen, A.C., Wang, T.T., Yondola, M., Gao, Q., Haye, K., Garcia-Sastre, A., Palese, P., 2010. Influenza virus vaccine based on the conserved hemagglutinin stalk domain. MBio 1, e00018-00010.

Stehle, T., Harrison, S.C., 1997. High-resolution structure of a polyomavirus VP1-oligosaccharide complex: implications for assembly and receptor binding. Embo J 16, 5139-5148.

Stehle, T., Yan, Y., Benjamin, T.L., Harrison, S.C., 1994. Structure of murine polyomavirus complexed with an oligosaccharide receptor fragment. Nature 369, 160-163.

Stevens, J., Blixt, O., Tumpey, T.M., Taubenberger, J.K., Paulson, J.C., Wilson, I.A., 2006. Structure and receptor specificity of the hemagglutinin from an H5N1 influenza virus. Science 312, 404-410.

Stevens, R.C., 2000. High-throughput protein crystallization. Curr Opin Struct Biol 10, 558-563.

Stubenrauch, K.A.Y., Gleiter, S., Brinkmann, U., Rudolph, R., Lilie, H., 2001. Conjugation of an antibody Fv fragment to a virus coat protein: cell-specific targeting of recombinant polyoma-virus-like particles. Biochem J 356, 867-873.

Sui, J., Hwang, W.C., Perez, S., Wei, G., Aird, D., Chen, L., Santelli, E., Stec, B., Cadwell, G., Ali, M., Wan, H., Murakami, A., Yammanuru, A., Han, T., Cox, N.J., Bankston, L.A., Donis, R.O., Liddington, R.C., Marasco, W.A., 2009. Structural and functional bases for broad-spectrum neutralization of avian and human influenza A viruses. Nat Struct Mol Biol 16, 265-272.

Takeda, M., Pekosz, A., Shuck, K., Pinto, L.H., Lamb, R.A., 2002. Influenza A virus M2 ion channel activity is essential for efficient replication in tissue culture. J Virol 76, 1391-1399.

Tegerstedt, K., Franzen, A.V., Andreasson, K., Joneberg, J., Heidari, S., Ramqvist, T., Dalianis, T., 2005. Murine polyomavirus virus-like particles (VLPs) as vectors for gene and immune therapy and vaccines against viral infections and cancer. Anticancer Res 25, 2601-2608.

Tekewe, A., Fan, Y., Tan, E., Middelberg, A.P.J., Lua, L.H.L., 2016. Integrated molecular and bioprocess engineering for bacterially produced immunogenic modular virus-like particle vaccine displaying 18 kDa rotavirus antigen. Biotechnol Bioeng 114, 397-406.

Thomas, S., Thirumalapura, N.R., Crocquet-Valdes, P.A., Luxon, B.A., Walker, D.H., 2011. Structure-based vaccines provide protection in a mouse model of ehrlichiosis. PLoS One 6, e27981.

Throsby, M., van den Brink, E., Jongeneelen, M., Poon, L.L.M., Alard, P., Cornelissen, L., Bakker, A., Cox, F., van Deventer, E., Guan, Y., Cinalt, J., ter Meulen, J., Lasters, I., Carsetti, R., Peiris, M., de Kruif, J., Goudsmit, J., 2008. Heterosubtypic neutralizing monoclonal antibodies cross-

protective against H5N1 and H1N1 recovered from human IgM^+ memory B cells. PLoS One 3, e3942.

Tissot, A.C., Renhofa, R., Schmitz, N., Cielens, I., Meijerink, E., Ose, V., Jennings, G.T., Saudan, P., Pumpens, P., Bachmann, M.F., 2010. Versatile virus-like particle carrier for epitope based vaccines. PLoS One 5, e9809.

Unneberg, P., Merelo, J.J., Chacon, P., Moran, F., 2001. SOMCD: method for evaluating protein secondary structure from UV circular dichroism spectra. Proteins 42, 460-470.

Varečková, E., Mucha, V., Wharton, S.A., Kostolanský, F., 2003. Inhibition of fusion activity of influenza A haemagglutinin mediated by HA2-specific monoclonal antibodies. Arch Virol 148, 469-486.

Vigerust, D.J., Shepherd, V.L., 2007. Virus glycosylation: role in virulence and immune interactions. Trends Microbiol 15, 211-218.

Wallace, B.A., Janes, R.W., 2009. Modern techniques for circular dichroism and synchrotron radiation circular dichroism spectroscopy. IOS Press, Amsterdam, NLD.

Ward, R.L., Clark, M.A., Lees, J., Hawkins, N.J., 1996. Retrieval of human antibodies from phage-display libraries using enzymatic cleavage. J Immunol Methods 189, 73-82.

Webster, R.G., Bean, W.J., Gorman, O.T., Chambers, T.M., Kawaoka, Y., 1992. Evolution and ecology of influenza A viruses. Microbiol Rev 56, 152-179.

Weis, W., Brown, J.H., Cusack, S., Paulson, J.C., Skehel, J.J., Wiley, D.C., 1988. Structure of the influenza virus haemagglutinin complexed with its receptor, sialic acid. Nature 333, 426-431.

Whitmore, L., Wallace, B.A., 2008. Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases. Biopolymers 89, 392-400.

WHO, 2009. Influenza fact sheet.

Wibowo, N., Chuan, Y.P., Lua, L.H.L., Middelberg, A.P.J., 2012. Modular engineering of a microbially-produced viral capsomere vaccine for influenza. Chem Eng Sci 103, 12-20.

Wibowo, N., Hughes, F.K., Fairmaid, E.J., Lua, L.H.L., Brown, L.E., Middelberg, A.P.J., 2014. Protective efficacy of a bacterially produced modular capsomere presenting M2e from influenza: Extending the potential of broadly cross-protecting epitopes. Vaccine 32, 3651-3655.

Wiley, D.C., Wilson, I.A., Skehel, J.J., 1981. Structural identification of the antibody-binding sites of Hong Kong influenza haemagglutinin and their involvement in antigenic variation. Nature 289, 373-378.

Xu, R., Ekiert, D.C., Krause, J., Hai, R., Crowe, J.E., Wilson, I.A., 2010. Structural basis of preexisting immunity to the 2009 H1N1 pandemic influenza virus. Science 328, 357-360.

Xu, R., Krause, J.C., McBride, R., Paulson, J.C., Crowe, 2013. A recurring motif for antibody recognition of the receptor-binding site of influenza hemagglutinin. Nature 20, 363-370.

Xuan, C., Shi, Y., Qi, J., Zhang, W., Xiao, H., Gao, G.F., 2011. Structural vaccinology: structure-based design of influenza A virus hemagglutinin subtype-specific subunit vaccines. Protein Cell 2, 997-1005.

Yang, Y.-W., Teng, C.-C., 1998a. Circular dichroism and fluorescence studies of polyomavirus major capsid protein VP1. J Protein Chem 17, 61-71.

Yang, Y.-W., Teng, C.-C., 1998b. Stability of polyomavirus major capsid protein VP1 under denaturants guanidine hydrochloride and urea. Int J Biol Macromol 22, 81-90.

Yang, Y.-W., Teng, C.-C., 1999. Conformational changes of polyomavirus major capsid protein VP1 in sodium dodecyl sulfate solution. J Pept Res 53, 75-81.

Yang, Z.-Y., Wei, C.-J., Kong, W.-P., Wu, L., Xu, L., Smith, D.F., Nabel, G.J., 2007. Immunization by avian H5 influenza hemagglutinin mutants with altered receptor binding specificity. Science 317, 825-828.

Yin, Y., Li, H., Wu, S., Dong, D., Zhang, J., Fu, L., Xu, J., Chen, W., 2011. Hepatitis B virus core particles displaying *Mycobacterium tuberculosis* antigen ESAT-6-specific immune responses. Vaccine 29, 5645-5651.

Yu, X., Tsibane, T., McGraw, P.A., House, F.S., Keefer, C.J., Hicar, M.D., Tumpey, T.M., Pappas, C., Perrone, L.A., Martinez, O., Stevens, J., Wilson, I.A., Aguilar, P.V., Altschuler, E.L., Basler, C.F., Jr. Crowe, J.E., 2008. Neutralizing antibodies derived from the B cells of 1918 influenza pandemic survivors. Nature 455, 532-536.

Zhang, J., Li, H., Wang, X., Qi, H., Miao, X., Zhang, T., Chen, G., Wang, M., 2012. Phage-derived fully human antibody scFv fragment directed against human vascular endothelial growth factor receptor 2 blocked its interaction with VEGF. Biotechnol Prog 28, 981-989.

Chapter 3

Identification of structural presentation of presented antigenic elements

3.1. Introduction

Virus-like particles (VLPs) are the macromolecular assemblies of viral capsid proteins. Due to their similar morphology to cognate viruses, VLPs can enhance the immunogenicity of antigenic epitopes. An increasing utilised approach in vaccine development is to employ VLPs as a carrier presenting an antigenic epitope to elicit an immune response against a target pathogen. The Centre for Biomolecular Engineering at the University of Queensland (UQ) has developed vaccine platforms based on the VP1 major capsid protein from murine polyomavirus (MuPyV) (Middelberg *et al.*, 2011). The vaccine platform based on the prokaryotic expression system could potentially manufacture influenza vaccine on a large scale and at low cost (Chuan *et al.*, 2008; Liew *et al.*, 2012). VP1 monomers can self-assemble into a capsomere (comprising five monomers), which can then further assemble into a VLP (comprising 72 capsomeres). Both capsomeres and VLPs have been exploited as delivery systems for a target antigenic element from an unrelated virus, such as the influenza virus. Studies have shown that modular capsomeres and VLPs, presenting epitopes or antigenic elements on their surfaces, induced a potent antibody response (Middelberg *et al.*, 2011; Tekewe *et al.*, 2016; Wibowo *et al.*, 2012). A study has reported that modular capsomeres with M2e peptides elicited protective antibodies in mice against influenza (Wibowo *et al.*, 2014). The presentation of epitopes on VLPs in a native structure is essential to induce neutralising antibodies. Very limited studies characterise the conformation of an epitope on VLP (Joshi *et al.*, 2013; Schneeman *et al.*, 2012). The evidence indicates that there is a knowledge gap in designing epitope presentation in its native structure on the surface of an unrelated VLP.

A recent study employed different display strategies to present the A/California/07/2009 influenza epitope, Helix 190 (H190), on the MuPyV VLP (Anggraeni *et al.*, 2013). H190 is a highly immunodominant antigenic epitope within the receptor binding site of the hemagglutinin globular sub-domain (HA1) (Caton and Brownlee, 1982; Yu *et al.*, 2008). The antigenic element was arrayed as a tandem repeat or flanked with functional elements, to generate an antigenic module. The study demonstrated that the modularisation of H190 using a tandem repeat approach resulted in higher

antibody titres against native structure than the flanking with the structure-promoting element (Anggraeni *et al.*, 2013). The display strategies might lead to the structural difference of the antigenic element within a module, which results in a variation in the quality of induced immune response. Determining the structure of the antigenic element is essential to guide the design of epitope presentations on VLPs. This chapter investigates a potential tool to aid the understanding of structural presentation of the H190 element. Antibodies recognise antigens depending on the conformation of antigens (Sela-Culang *et al.*, 2013). An H190-specific antibody would be potential to identify the structure of H190 element on VLP. Here, isolating H190-specific antibody using phage display technology is explored. The advent of phage display technology has facilitated the generation of phage antibody libraries. Antigen-specific antibodies can be isolated from a phage antibody library via phage antibody panning technique.

Phage display technology was first described in 1985 (Smith, 1985). This technology involves displaying exogenous peptides and proteins on the surface of filamentous *E. coli* bacteriophage. Phage display studies exploited all phage coat proteins of bacteriophage M13 to display fused proteins (Hoogenboom *et al.*, 1998). The phage coat consists of a major coat protein, i.e. gp8 (2700 copies/phage), and minor coat proteins, i.e. gp3, gp6, gp7 and gp9 (5 copies/phage) (Figure 3.1A). G3p is the most commonly used coat protein because of its high tolerance capacity (Zani and Moreau, 2010) and its high efficiency for displaying foreign proteins (Kretzschmar and Geiser, 1995). An unrelated peptide was fused into the area between the signal peptide and N-terminal domain of the gp3 protein, resulting in the presentation of the fused protein on all single gp3 copies at the end of the phage particle. Related research displayed a single chain variable fragment (scFv) of an antibody on the bacteriophage surface (McCafferty *et al.*, 1990). Another study succeeded in presenting a larger antigen-binding fragment (Fab) of the antibody with a more complex conformation on the phage surface (Figure 3.1B) (Bradbury and Marks, 2004). By displaying the antibody repertoire on the phage coat, it led to the construction of a phage antibody library which facilitates the screening or isolation of antibodies specific for a target antigen.

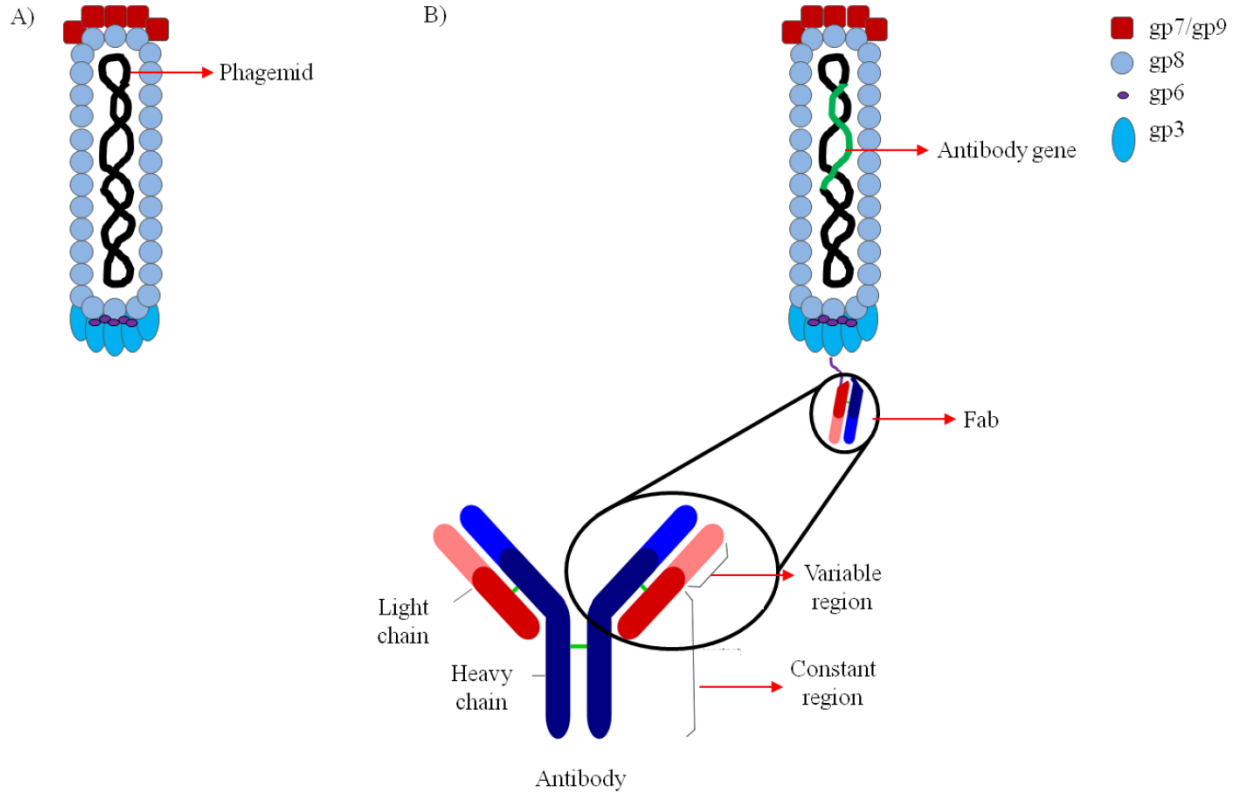


Figure 3.1: Phage display technology. A) The structure of M13 phage coat proteins (Bernal and Willats, 2004), B) Antibody phage presenting Fab on its surface (Brichta et al., 2005).

Phage antibody panning is a technology for the selection of antibody-phages against a specific target antigen from the phage antibody library (Bradbury and Marks, 2004). The panning against the target antigen consists of several crucial stages: coating, binding, washing, elution, rescuing and amplification. The panning rounds are repeated to obtain the enrichment of antigen-specific antibodies. Phage antibody panning has been utilised to select antibody-phages against influenza antigens (Kashyap *et al.*, 2008; Sui *et al.*, 2009; Throsby *et al.*, 2008). Antigen-antibody recognition can confirm the conformation of antigens due to the dependence on the structure of the antigen (Lu *et al.*, 2014; Sela-Culang *et al.*, 2013).

This chapter aims to employ antigen-antibody recognition via phage antibody panning to investigate the conformational presentation of antigenic element H190 on modular capsomeres, using two different display strategies. In the first strategy, the H190 element was flanked with the structure-promoting element GCN4 within a module on modular capsomere H190-GCN4. In the second strategy, dual copies of the H190 element were arrayed as tandem repeats within a module on modular capsomere H190-2 (Figure 3.2). Two sets of panning, A and B, were implemented in parallel to investigate the conformational presentation of the H190 element on H190-GCN4 and H190-2 modular capsomeres, respectively.

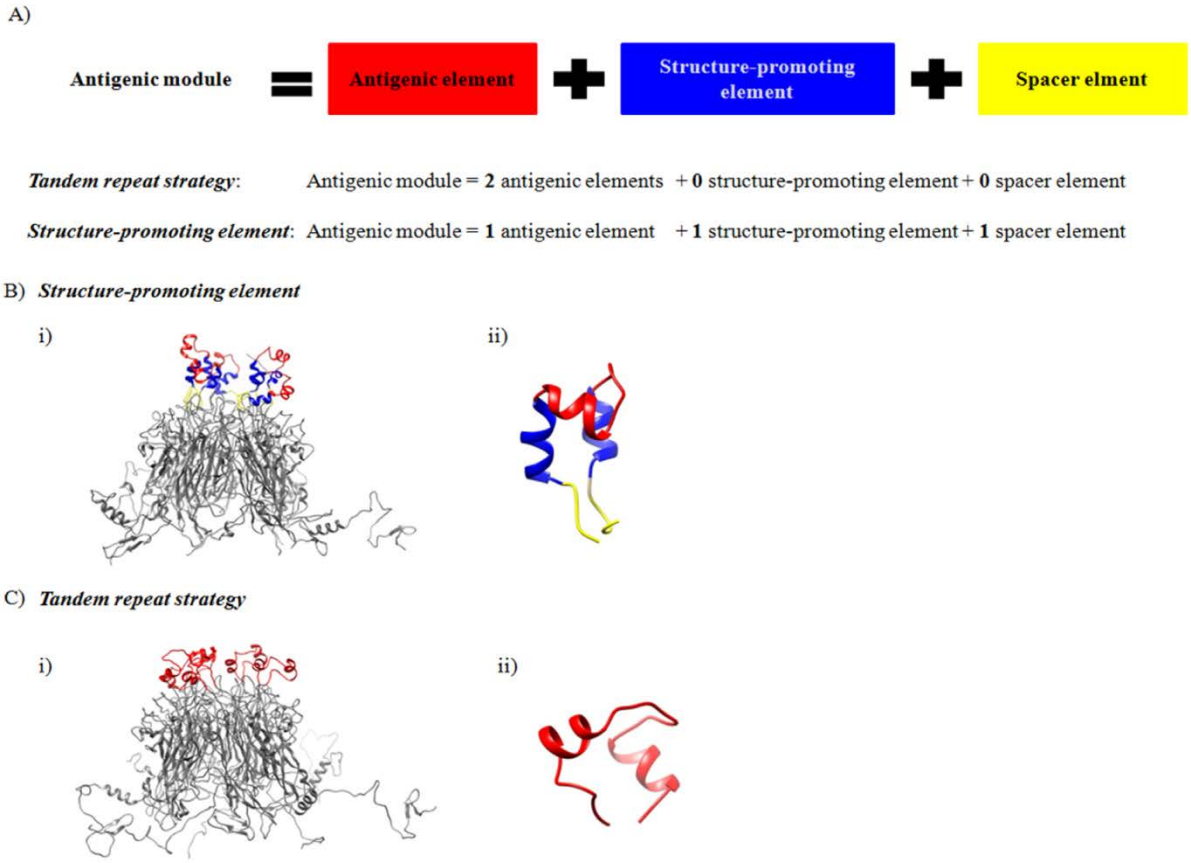


Figure 3.2: Modularisation of an antigenic module H190 from influenza hemagglutinin. A) Diagram illustrating the components within an antigenic module. The presentation of module H190 on MuPyV capsomere using two display strategies: (B, i) modular capsomere H190-GCN4, (B, ii) antigenic module H190-GCN4 consisting of H190 element flanked with GCN4 elements. (C, i) Modular capsomere H190-2 and (C, ii) antigenic module H190-2 comprised of dual copies of H190 element. The module H190 is highlighted in red, structure-promoting element GCN4 in blue and spacer element in yellow. The predicted structure was generated using Accelrys Discovery Studio 3.0 performed by a former UQ research colleague, Natalie Connors, based on 1sid.pdb.

Figure 3.3 illustrates the panning strategy to compare the structure of the H190 element on modular capsomeres and native H190 on HA1. In the first three panning rounds, two subtractive pannings against base capsomeres, bearing no antigenic modules, were used to eliminate VP1-binding phages. Afterwards, selective panning against modular capsomeres was used to isolate the phages bound to the H190 module on the modular capsomere. The isolated pool of phages was panned against HA1 for a further two rounds to select phages bound to both modular capsomere and H1N1 HA1. If the modular capsomere-binding phages can bind to H1N1 HA1 at the same level as the binding to modular capsomere, it would indicate that the structure of the H190 element on the modular capsomere may be identical to the structure of native H190 (Figure 3.3 2A). If modular capsomere-binding phages are unable to bind to H1N1 HA1 or bind weakly to H1N1 HA1, it would indicate that the structure of the H190 element on the modular capsomere may be different from the structure of native H190 (Figure 3.3 2B).

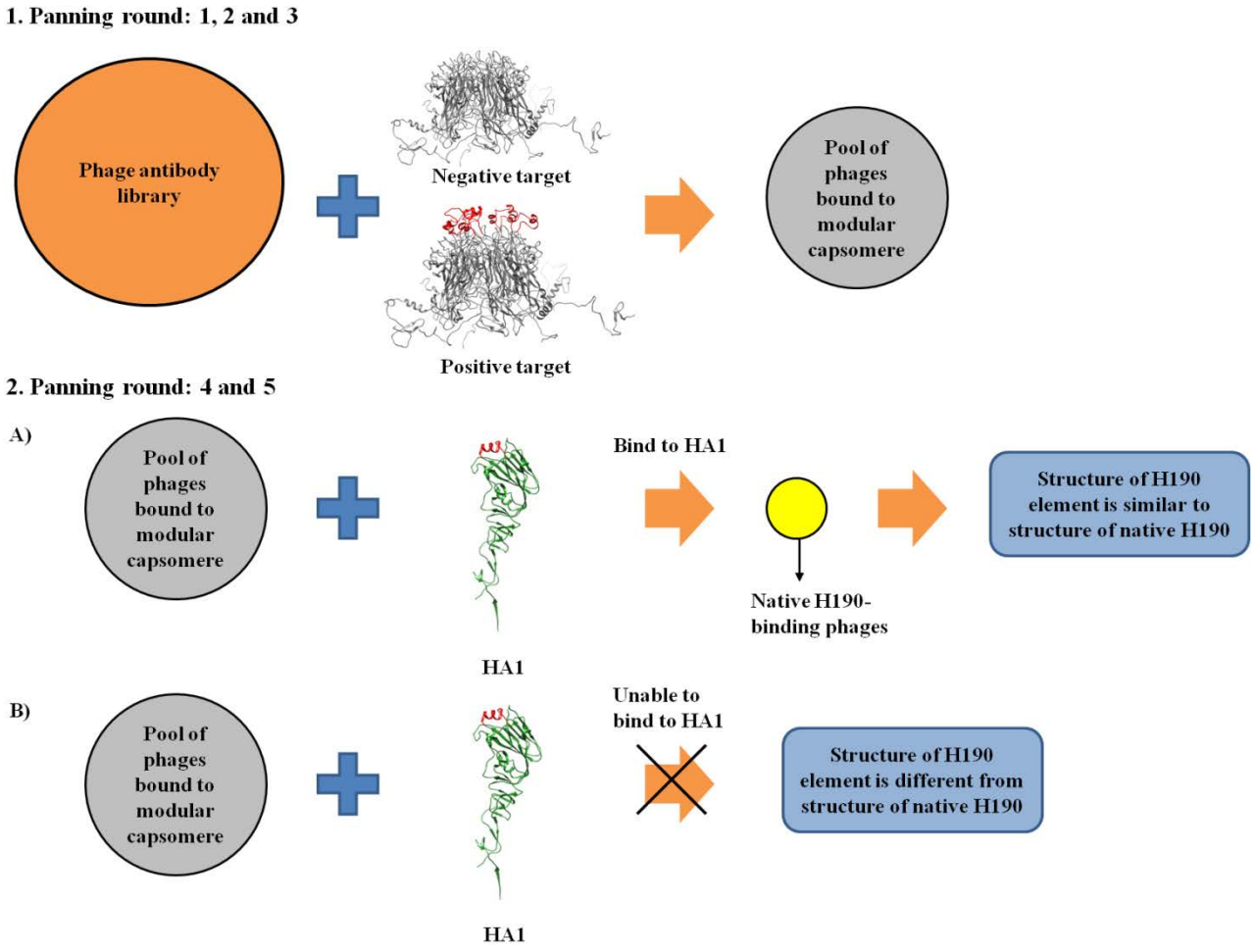


Figure 3.3: Panning strategy to identify the structure of the H190 module on the modular capsomere. 1) The first three panning against base capsomere (negative target) and modular capsomeres (positive target). 2) The fourth and fifth panning against HA1. 2A) If modular capsomere binding-phages bind to HA1 at the same level as to modular capsomere, it would indicate that the structure of the H190 element on modular capsomere may be similar to the native structure on HA1. 2B) If the modular capsomere binding-phages are unable to bind or are weakly bound to HA1, it would indicate that the structure of the H190 element on modular capsomere may be different from the native structure on HA1.

3.2. Materials and Methods

3.2.1. Synthesis of the base and modular capsomeres

3.2.1.1. Molecular cloning

The expression vectors pGEX-VP1-S4-GSGS-GCN4, pGEX-VP1-S4-GSGS-GCN4-H190, pGEX-VP1-S1S4, and pGEX-VP1-S1S4-H190-0020 were generated in past work (Anggraeni *et al.*, 2013). The designated names of these expression vectors are shown in Table 3.1.

Table 3.1: Vector and designated protein names.

Vector	Protein name	Short name
pGEX-VP1-S4-GSGS-GCN4	Capsomere VP1-S4-GSGS-GCN4	Cap-Base A
pGEX-VP1-S4-GSGS-GCN4- H190	Capsomere VP1-S4-GSGS-GCN4-H190	CapH190-GCN4
pGEX-VP1-S1S4	Capsomere VP1-S1S4	Cap-Base B
pGEX-VP1-S1S4- H190- H190	Capsomere VP1-S1S4- H190-H190	CapH190-2

3.2.1.2. Expression

GST tagged-modular and base capsomeres are intracellularly expressed in *E.coli* Rosetta (DE3) pLysS cells (EMD Millipore, Merck KGaA, Darmstadt, Germany) as reported in previous work (Chuan *et al.*, 2008; Lipin *et al.*, 2009). A single colony was inoculated into 5mL of Terrific Broth (TB) containing 1.2% tryptone (w/v), 2.4% yeast extract (w/v), 72 mM K_2HPO_4, 17 mM KH_2PO_4 and 0.4% glycerol (w/v), and incubated at 30°C, 180 rpm for 14-16 hours using a rotary shaker (Bioline, Edwards Instrument Company, Australia). 800µL of seed culture was added to 800mL TB media and then incubated at 37°C, 180 rpm until $OD_{600\ nm}$ reached 0.4. The cultures were cooled to 26°C before being induced with 0.2 mM isopropyl-β-D-thiogalactoside (IPTG). After IPTG induction, the cultures were incubated at 26°C, 180rpm for 16 hours. The cell pellet was harvested by centrifuging at 4000x*g*, 4°C for 20 minutes and kept at -80°C.

3.2.1.3. Glutathione-S-transferase (GST) purification of the base and modular capsomeres

Cell pellets were resuspended in 40 mL L buffer (containing 40 mM Tris-base, 200 mM NaCl, 1 mM EDTA, 5% glycerol, 5 mM DTT, pH 8.0) and sonicated for 4 cycles of 45 seconds each (Branson Sonifer 250). Cell lysates were clarified by centrifugation (Avanti^{TM} J-25 I, Beckman) at 15000 rpm, 4°C for 20 minutes. GST tagged-modular and base capsomeres were purified using a 5mL GSTrap HP column on AKTAxpress^{TM} (GE Healthcare Bioscience). The supernatants were passed through 0.45 µm filters (Pall, New York, USA) before loading to the column. After loading,

unbound proteins were washed with 5 column volumes of L buffer. The GST-modular and base capsomeres were eluted from the column in L buffer containing 10 mM reduced glutathione at pH 8.0.

3.2.1.4. Size exclusion purification of the base and modular capsomeres

GST tag was removed from GST tagged-modular and base capsomeres by enzymatic digestion with thrombin (Catalogue# 27-0846-02, GE Healthcare, UK) at a ratio of 40:1 (thrombin unit/ml protein) for 2 hours at room temperature. Modular and base capsomeres were separated from aggregate, and GST tag using Superdex 200 30/100 GL column (GE Healthcare Bioscience) with a flow rate of 0.5 mL/min in L buffer. Fractions corresponding to capsomeres were collected and stored at -80°C.

3.2.2. Recombinant HA1

Recombinant HA1 (A/California/07/2009 H1N1) produced in insect cells was purchased from Protein Expression Facility, Queensland, Australia.

3.2.3. Determination of protein molar concentration

UV absorbance of protein at 280 nm was measured by UV spectroscopy and converted into protein concentration based on the Beer-Lambert Law (Aitken and Learmonth, 2009):

$$A = \varepsilon \times b \times c$$

A: The absorbance value at 280 nm,

ε: The extinction coefficient of protein ($M^{-1}cm^{-1}$)

b: The path length and c is the protein concentration (mg mL^{-1})

Theoretical molecular weight and extinction coefficient of the proteins were obtained from the ProtParam tool on the Expasy website (Table 3.2).

Table 3.2: Theoretical molecular weight and extinction coefficient of protein.

Protein name	Molecular weight	Extinction coefficient
Cap-Base A	44602.8	58245
CapH190-GCN4	46346.6	61225
Cap-Base B	42562.5	58245
CapH190-2	46050.0	64205
HA1 (A/California/07/2009 H1N1)	35627.1	54360

The concentrations of proteins were converted into molar concentration or molarity (M) based on the formula:

$$Molar\ concentration = \frac{Concentration}{Molecular\ weight}$$

3.2.4. Phage antibody libraries

The Australian Red Cross (ARC) and Sheets' libraries are non-immune libraries, provided by Stephen Mahler and Martina Jones. The ARC library was generated by Stephen Mahler and his colleagues during 2004-2007 and contains over 10^{10} individual clones in Fab form (unpublished). The Sheets' library which was generated by Sheets *et al*. in 1998 contains 6.7 x 10^9 individual clones in scFv form (Sheets *et al.*, 1998).

3.2.5. Phage antibody panning

The Nunc-Immuno^{TM} tubes (Thermo Scientific, Australia) were coated with target antigens overnight and incubated with phages in 2% milk-PBS for 1 hour at room temperature (RT). Unbound phages were washed with PBS-T (0.1% Tween-20), then the bound phages were eluted with 0.1 M triethylamine pH 12 and neutralised with 1 M Tris, pH 7.4. The eluted phages were transfected into *E.coli* XL1-Blue cells (Stratagene, California, US) in log phase and incubated for 30 minutes at 37°C before plating on 2YT agar medium containing 100 µg/ml Ampicillin and 2% glucose (2YT-AmpGlu) plates. The XL1-Blue infected cells were grown overnight and then detached from the plates and grown until the log phase after which they were recovered with 10^{11} cfu ml^{-1} helper phage particles M13K07 (New England Biolabs). The phages were precipitated using PEG-NaCl (containing 20% Polyethene glycol-6000 and 2.5 M NaCl) on ice for 1 hour and harvested for the next panning round.

The stringency of selection was increased by reducing the antigen concentration and increasing washing times as in Table 3.3.

Table 3.3: Antigen concentration and washing times.

Round	Antigen concentration (μM)	Wash (time)
1	1	3
2	0.1	10
3	0.1	10
4	0.1	10
5	0.1	10

3.2.6. Polyclonal phage ELISA

The antigens were coated on 96-well ELISA plates (Nunc, Maxisorp) and incubated with phages in 2% Milk-PBS at RT for 1 hour. After washing, the HRP conjugated anti M13 antibody was added and incubated for 1 hour. TMB (3, 3', 5, 5'- tetramethylbenzidine) solution was added to react with HRP. The signal was measured at A_{450}.

3.2.7. Monoclonal phage ELISA

Eighty-four single colonies of transfecting XL1-Blue cells in the final panning round were randomly picked up and grown in a 96-well round-bottom plate before recovery and harvesting the phages.

Antigens were coated in 96-well ELISA plates (Nunc, Maxisorp) and incubated with phages in 2% Milk-PBS at RT for 1 hour. After washing, HRP conjugated anti M13 antibody was added and incubated for 1 hour. TMB (3, 3', 5, 5'- tetramethylbenzidine) solution was added to react with HRP. The signal was measured at A_{450}.

3.2.8. Competitive phage ELISA against biotinylated-synthetic peptides

Peptide H190-GCN4 (biotin-GGGGSVKQLEDKVSTSADQQSLYQNADAYVKQLEDKV-NH_2) and peptide H190 (biotin-GGGGSSTSADQQSLYQNADAY-NH_2) were purchased from Peptide 2.0 (Chantilly, VA, USA).

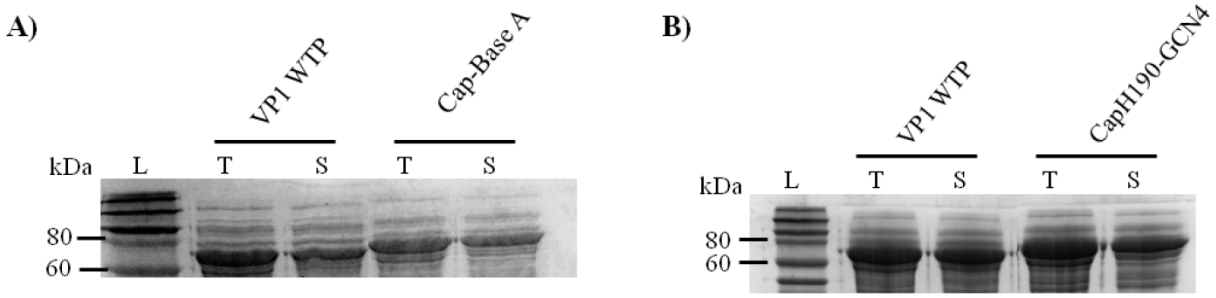
Antigens were coated on Pierce® Streptavidin High Binding Capacity Coated 96-Well Plates (Thermo Scientific, Waltham, MA, USA). Phages were pre-mixed and incubated with various concentration of biotinylated peptides in 2% Milk-PBS at RT for 1 hour before incubation on the plate. After washing, HRP conjugated anti M13 antibody was added and incubated for 1 hour. TMB (3, 3', 5, 5'- tetramethylbenzidine) solution was added to react with HRP. The signal was measured at A_{450}.

3.3. Results and Discussion

3.3.1. Investigation of the conformational presentation of H190 element on CapH190-GCN4

3.3.1.1. Synthesis of Cap-Base A and CapH190-GCN4

The modular VP1 constructs were expressed as GST-tagged proteins (Chuan *et al.*, 2008). After cell disruption using sonication, total cell lysates and soluble fractions were taken for SDS-PAGE to analyse the solubility. Figure 3.4 compares the solubility of wild-type VP1 (VP1 WTP), Cap-Base A and CapH190-GCN4. Cap-Base A and CapH190-GCN4 were highly soluble at about 93-95%, which are similar to the solubility of VP1 WTP. It indicates that the modularisation on the insertion sites of Cap-Base A and CapH190-GCN4 has no impact on the protein solubility.



C)

Constructs	Solubility (%)
VP1 WTP	96
Cap-Base A	93
CapH190-GCN4	95

Figure 3.4: A) SDS-PAGE analysis showing the solubility of Cap-Base A and CapH190-GCN4 in comparison with VP1 WTP. Lanes: (L) molecular weight marker; (T) Total cell lysate; (S) Soluble fraction of cell lysate. B) The solubility of VP1 WTP, Cap-Base A and CapH190-GCN4. The protein solubility was calculated by measuring the relative intensity of the relevant bands in the total fraction in comparison to the soluble fraction using the SDS-PAGE gel image. The analysis was performed using Image Lab software (Bio-rad).

Following GST affinity chromatography purification of Cap-Base A and CapH190-GCN4, GST tags were removed from untagged capsomeres using enzymatic digestion. The Cap-Base A and H190-GCN4 were separated from the aggregate and GST tag using size exclusion chromatography (SEC) (Figure 3.5A). The chromatogram consisted of three major peaks as observed in a previous study (Middelberg *et al.*, 2011). They included: i) a peak corresponding to aggregate eluted at about 8.5 ml; ii) a peak corresponding to capsomere at about 11.5 ml; and iii) a peak corresponding to GST tag at about 16 ml. The SEC analysis shows the Cap-Base A and CapH190-GCN4 eluted at about the same time, suggesting a similar hydrodynamic radius of Cap-Base A and CapH190-GCN4. It demonstrated that the insertion of H190 element on each monomer of CapH190-GCN4 did not affect the hydrodynamic radius of CapH190-GCN4. The Cap-Base A and CapH190-GCN4 were collected and analysed using SDS-PAGE. The SDS-PAGE analysis shows that the molecular weight of CapH190-GCN4 (lane 2) is larger than which of Cap-Base A (lane 1), resulting from the insertion of H190 element on CapH190-GCN4 (Figure 3.5B).

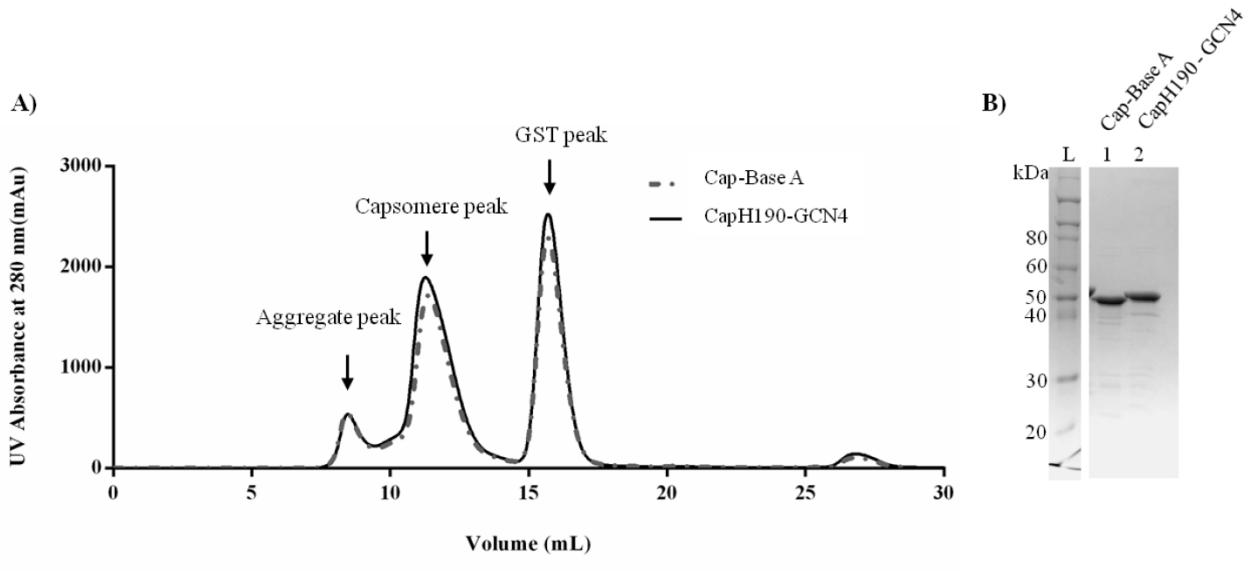


Figure 3.5: A) SEC purification of Cap-Base A (broken line) and CapH190-GCN4 (solid line) after GST removal from untagged capsomeres. B) SDS-PAGE analysis of Cap-Base A and CapH190-GCN4 after SEC purification. Lanes: (L) molecular weight marker; (1) Cap-Base A; (2) CapH190-GCN4.

3.3.1.2. Phage antibody panning

Five panning rounds were performed to isolate modular capsomere-binding phages. Each panning round consisted of 5 fundamental steps: coating, binding, washing, elution and rescue and amplification. The coating step was to immobilise the target antigens on the surface of immunotypes. The binding step allows phages binding to immobilised antigens. In round 1, 2 and 3, subtractive binding against Cap-Base A was to eliminate VP1 capsomere structure-binding phages; then, selective binding against CapH190-GCN4 was to select H190 element-binding phages (Figure 3.6). In round 4 and 5, phages isolated in the previous rounds were panned against H1N1 HA1 in selective binding (Figure 3.7). The washing step was to remove unbound phages. The elution step was to elute bound phages to the target antigen. The rescue and amplification step was to produce and amplify bound phages for the next panning round.

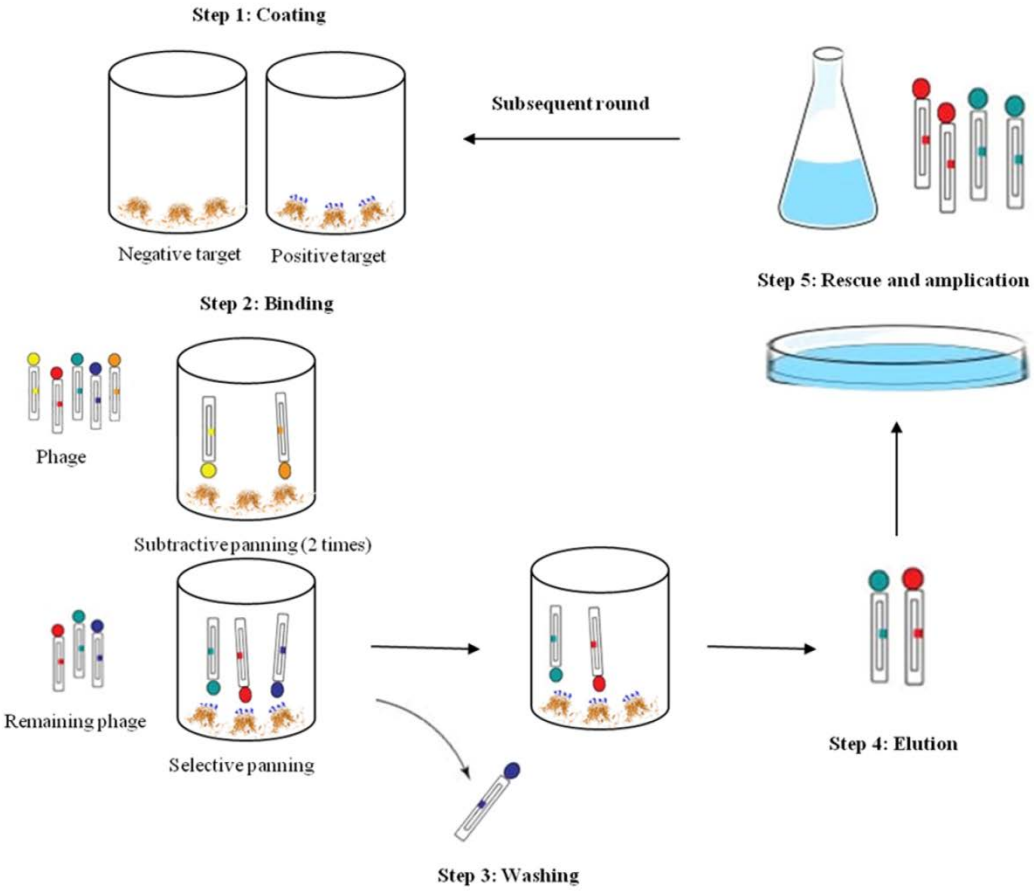


Figure 3.6: Schematic diagram illustrating panning round 1, 2 and 3 against the base and modular capsomeres.

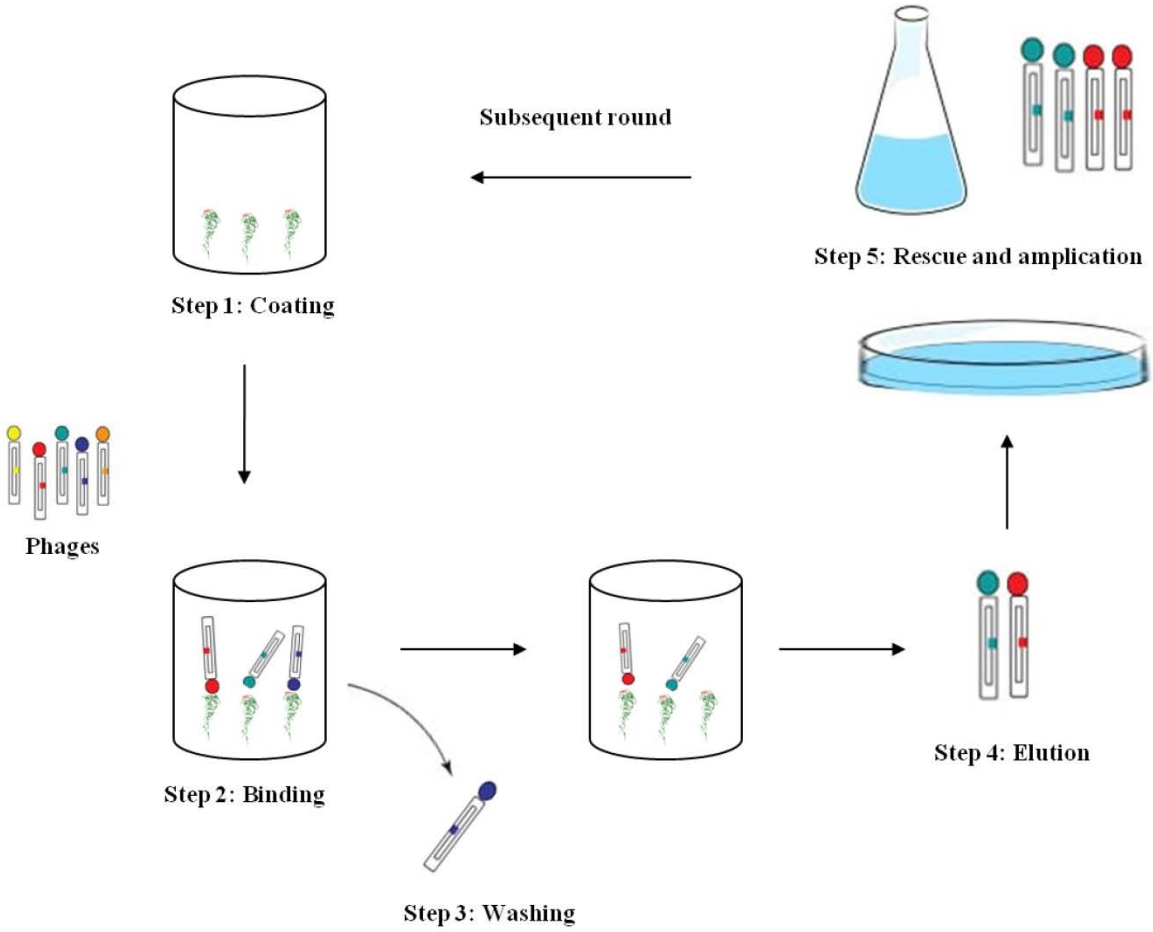


Figure 3.7: Schematic diagram illustrating panning round 4 and 5 against H1N1 HA1.

Among the fundamental steps in the panning process, the elution condition is a critical parameter because it can affect the selection of high- or low-affinity binders. The mild elution condition can result in the recovery of low-affinity antibody-phages and the loss of high-affinity antibody-phages which are still strongly bound to the target antigen. Eluent, particularly 0.2M glycine pH 2.5 (acidic solution) (Kang *et al.*, 1991) or 0.1 M triethylamine pH 12 (basic solution) (Marks *et al.*, 1991), is widely used in phage antibody panning. It is necessary to maximise the amount of eluted phages in the elution step, leading to maximising the chance of recovering high-affinity antibody-phages.

The strategy to increase the amount of eluted phages used two approaches: i) the choice of efficient eluent; and ii) the increment of elution time. In the first strategy, two eluents, 0.2M glycine pH 2.5 and 0.1 M triethylamine pH 12, were investigated to compare the amount of eluted phages at different time points (Figure 3.8). At 10 minutes, 0.1 M triethylamine pH 12 eluted about 8 x 10^5 phages, while 0.2 M glycine pH 2.5 eluted about 2 x 10^5 phages. The better elution capacity of 0.1 M triethylamine pH 12 in comparison with 0.2 M glycine pH 2.5 has also been reported in an earlier study (De Bruin *et al.*, 1999). Therefore, 0.1 M triethylamine pH 12 was chosen to elute bound phages in this study. Furthermore, after 10 minutes, the amount of eluted phages decreased to about 50%, indicating that the harsh elution condition of 0.1 M triethylamine at pH 12 might affect the viability of phages. The result suggests that the tolerance capacity of phages to 0.1 M triethylamine at pH 12 may be 10 minutes, with a longer elution time possibly affecting the survival of phages.

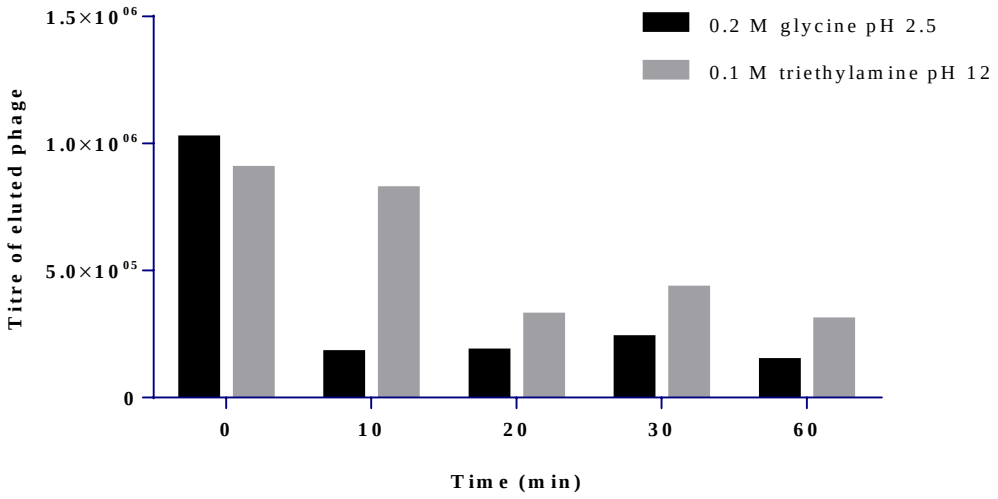


Figure 3.8: Comparison of the effects of 0.2 M glycine pH 2.5 and 0.1 M triethylamine pH 12 on phages elution at different time points.

The second strategy for improving the amount of eluted phages was to increase the elution time. A more extended exposure time of phages to the extreme pH of the eluent can change the structure of the functional antibody fragment on the phage surface (Sahin *et al.*, 2010). This study used step elution to maximise the amount of eluted phages and to reduce the exposure time of phages to the eluent. In step elution, the bound phages were incubated with 0.1 M triethylamine pH 12 for 10 minutes, and then the pH of the eluent was neutralised after the incubation. The eluted phages were harvested before the next step elution. The subsequent elution was repeated every 10 minutes over a period of 40 minutes. The eluted phages from each step elution were combined and used for the next stage of panning. Figure 3.9 shows the amount of phages eluted in each step. The results indicate that about 10^8 phages, 10^7 phages, 10^5 phages, and 10^4 phages eluted in steps 1, 2, 3 and 4, respectively. In step 5, there were no phages eluted, indicating almost all the bound phages were harvested in the step elution process. The results indicate that step elution resulted in more bound phages being harvested compared with standard elution. Therefore, this study applied step elution with 0.1 M triethylamine pH 12 to elute bound phages in the panning process.

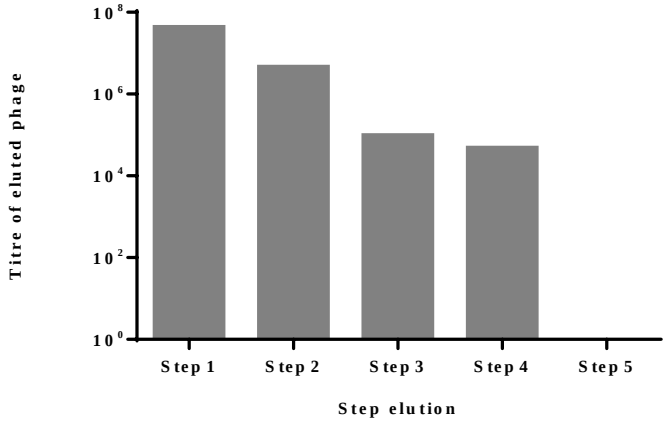


Figure 3.9: Phage titration in each elution step using 0.1 M triethylamine pH 12.

3.3.1.3. Polyclonal phage ELISA

After panning rounds 1, 2 and 3 against Cap-Base A as a subtractive antigen and CapGNC4-H190 as a selective antigen in the binding step (Figure 3.6), the isolated pools after each round were used to examine the enrichment of binding phages using polyclonal phage ELISA. Figure 3.10 indicates the enrichment of binding phages, as reflected by the increase in absorbance value after each round. The isolated pool of binding phages had a weak response to Cap-Base A, indicating that the pool contained a negligible amount of phages bound to Cap-Base A (Figure 3.10A). It demonstrated that the subtractive panning efficiently eliminated Cap-Base A-binding phages. The pool of binding

phages responded highly to CapGNC4-H190, showing that the isolated pool contained a large quantity of phages bound to CapH190-GNC4 (Figure 3.10B). The panning might selectively isolate and enrich phages bound to the H190 module, consisting of GCN4 element and H190 element, on the modular capsomere.

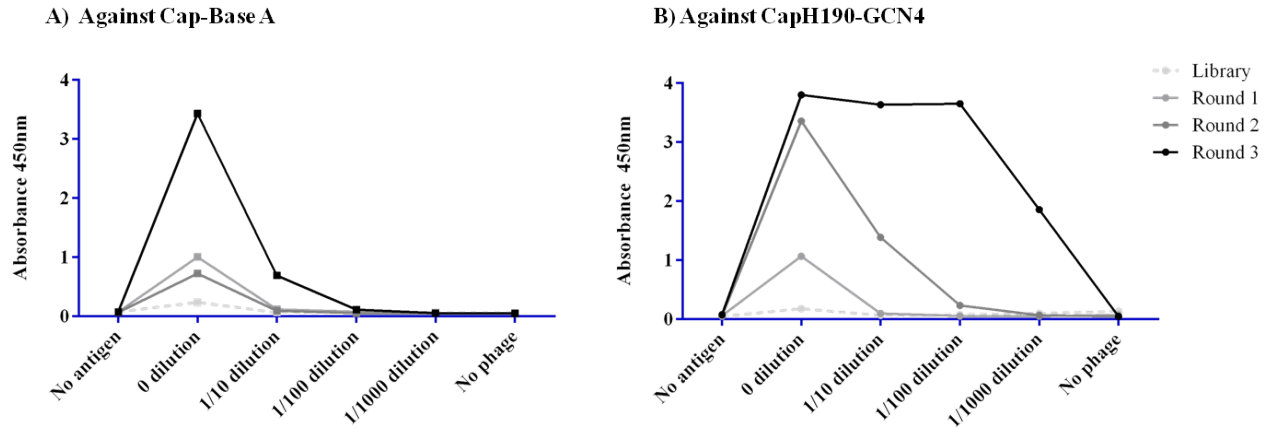


Figure 3.10: Polyclonal phage ELISA of ARC library-phages after panning round 1, 2 and 3 against (A) Cap-Base A, (B) CapH190-GCN4.

The isolated CapH190-GCN4-binding phages after round 3 were panned against H1N1 HA1 in round 4 and 5 (Figure 3.7). The polyclonal phage ELISA result showed that a small quantity of isolated phages in round 4 and 5 had a positive reactivity to H1N1 HA1 (Figure 3.11). The round 4 and 5 phage pools were also tested against CapH190-GCN4 as a positive control in polyclonal phage ELISA. The change of target antigen to H1N1 HA1 yielded no reduction in the binding of round 4 and 5 phage pools to CapH190-GCN4, compared with the binding in round 3 (Figure 3.11). It suggests that the first three rounds of panning isolated phages bound to H190 module on CapH190-GCN4.

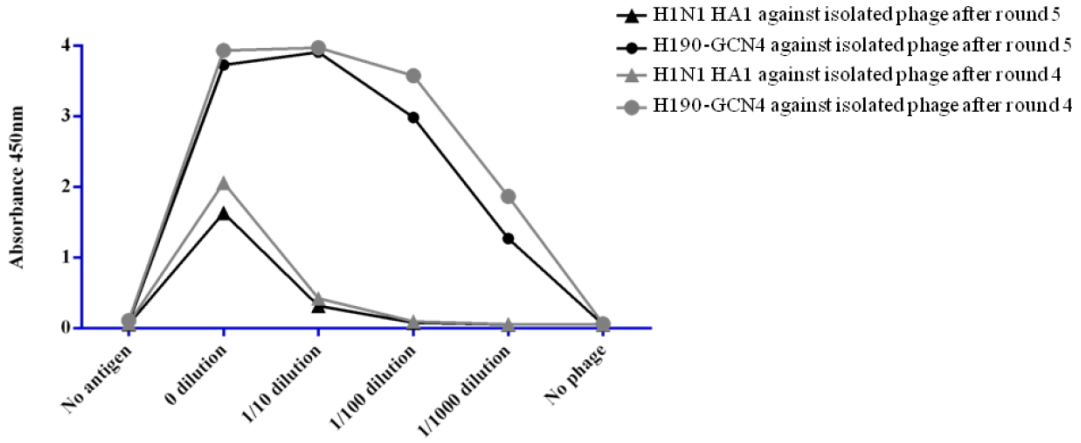


Figure 3.11: Polyclonal phage ELISA of ARC library-phages after the fourth and fifth panning against H1N1 HA1.

3.3.1.4. Monoclonal phage ELISA

The data from polyclonal phage ELISA (Figure 3.11) indicated that the isolated pool of phages after the fifth panning consisted of a few phages bound to both CapH190-GCN4 and H1N1 HA1. The isolated pool then underwent monoclonal phage ELISA to identify individual phages binding to both CapH190-GCN4 and H1N1 HA1. The screening of 84 individual colonies found three positive hits (Figure 3.12). DNA sequencing confirmed that three unique clones (A12, B8, and B12) bound to both CapH190-GCN4 and H1N1 HA1.

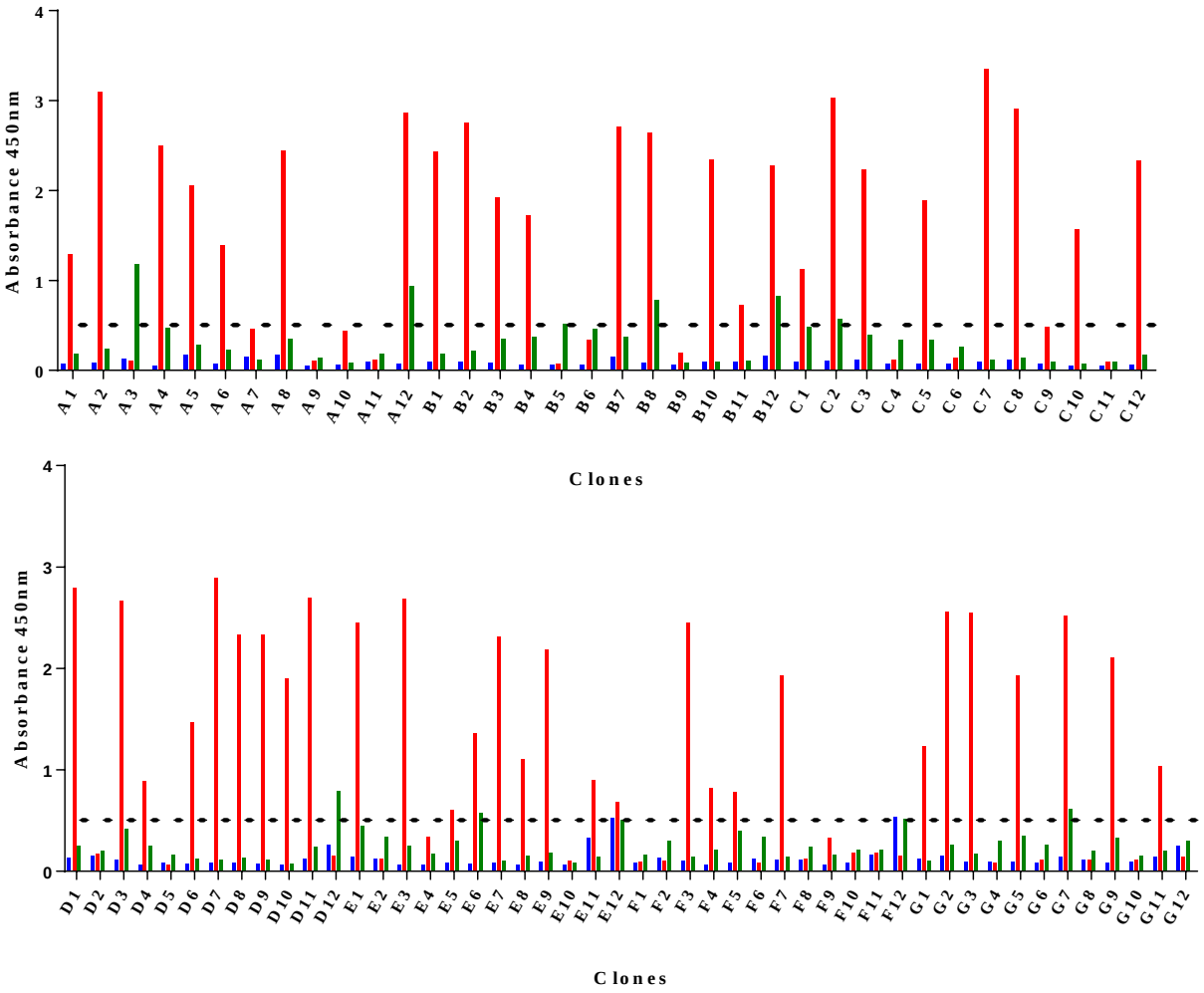


Figure 3.12: Monoclonal phage ELISA of ARC library-phages after the fifth panning against Cap-Base A (blue), CapH190-GCN4 (red) and H1N1 HA1 (green). Dashed line: Mean of systematic error + 5 standard deviations.

Although the identified phages bound to H1N1 HA1, the binding to the HA1 was considerably lower than the binding to CapH190-GCN4, as shown by the level of the absorbance value. It is likely that the H190 module bears the H190 element flanked with the GCN4 element, resulting in the identified phages perhaps bound to the other regions of the H190 module rather than the H190 element. It led to the question: to which part of the H190 module is the identified phages bound. This study used competitive ELISA against immobilised CapH190-GCN4 with the competitors, peptide H190 and peptide H190-GNC4, to examine the binding region of identified phages (A12, B8, and B12). Figure 3.13 shows that the peptide H190-GCN4 at a concentration of 40 µM completely abrogated the binding of phages A12, B8, and B12, to immobilised CapH190-GCN4, while peptide H190 did not inhibit the binding. The unbinding of isolated phages to Cap-Base A, bearing the GCN4 element, is shown in Figure 3.12. It suggests that phages A12, B8, and B12 might bind to the junction between the H190 element and GCN4 element. It explains why the phages showed a high response to CapH190-GCN4 having the H190 and GCN4 elements joined together into a module.

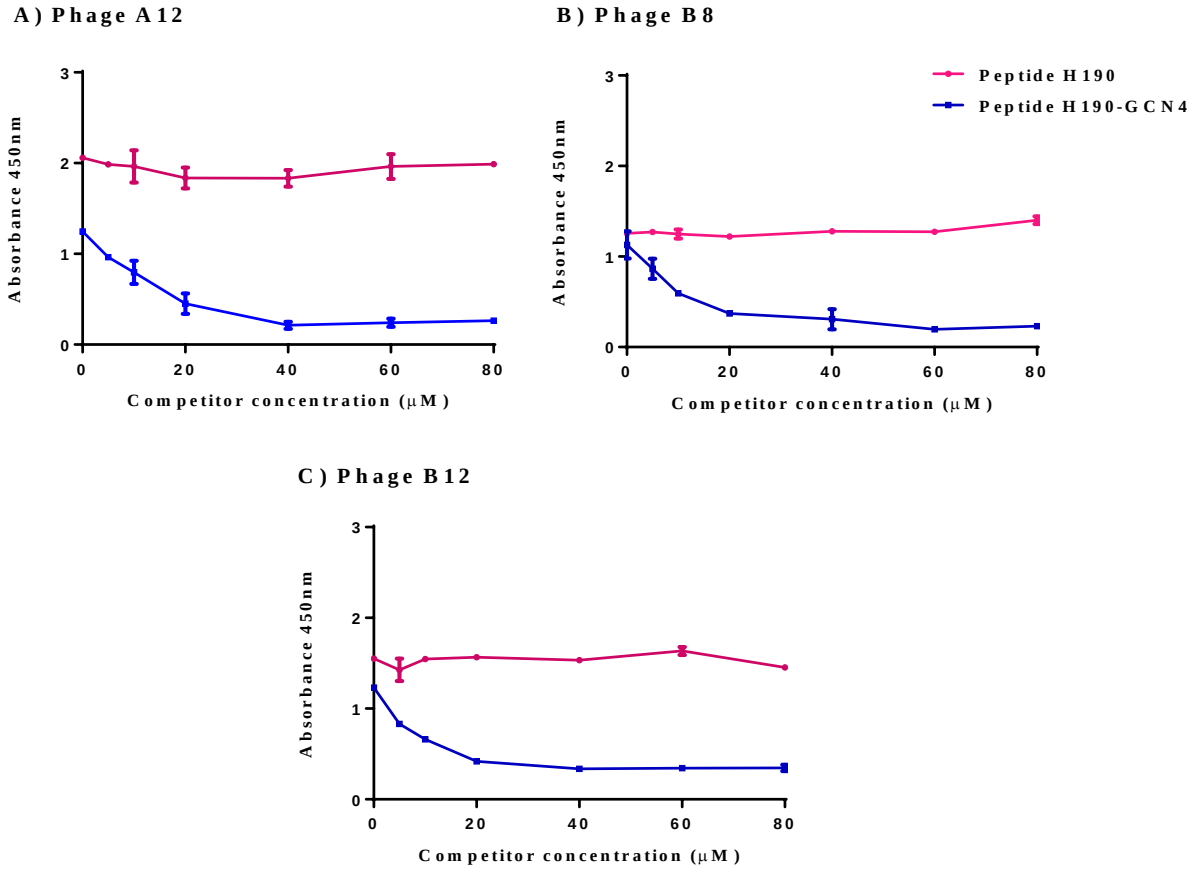


Figure 3.13 Investigation of the binding regions of the phages bound to both CapH190-GCN4 and H1N1 HA1, using competitive ELISA against immobilised modular capsomere H190-GCN4 with two competitors: peptide H190 and peptide H190-GCN4. Phages: (A) A12, (B) B8 and (C) B12.

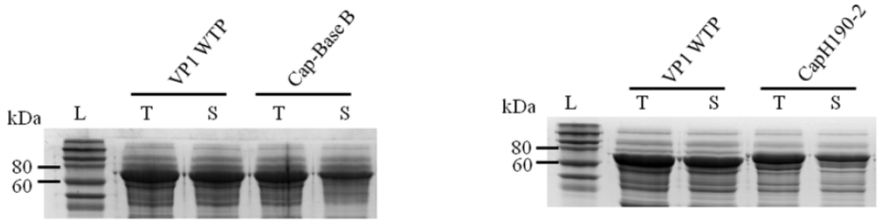
In summary, the experimental works in this section found that the phages bound both CapH190-GCN4 and H1N1 HA1. The further investigation of the binding regions of these phages showed that they might bind to the junction between the H190 and GCN4 elements. It led to the response of CapH190-GCN4 being significantly stronger than the response to H1N1 HA1. The results indicate that these phages can discriminate the conformation differences between the antigenic module and native H190 on H1N1 HA1. If it is possible to select the H190 element-binding phages, the antigen-antibody recognition via phage antibody panning could be a tool to identify the structural difference between H190 element and native H190. However, it is still unclear why H190 element-binding phages were not able to be found in the ARC library.

3.3.2. Investigation of the conformational presentation of H190 element on CapH190-2

3.3.2.1. Synthesis of Cap-Base B and CapH190-2

GST-tagged Cap-Base B and CapH190-2 were expressed as described in (Chuan *et al.*, 2008). The solubility of wild-type VP1 (VP1 WTP), Cap-Base B and CapH190-2 was analysed in SDS-PAGE comparing total cell lysates and soluble fractions (Figure 3.14). Cap-Base B and CapH190-2 were less soluble than VP1 WTP, showing that the modularisation led to a decrease in the solubility of the constructs.

A)



C)

Constructs	Solubility (%)
VP1 WTP	96
Cap-Base B	80
CapH190-2	88

Figure 3.14: A) SDS-PAGE analysis showing the solubility of Cap-Base B and CapH190-2 in comparison with VP1 WTP. Lanes: (L) molecular weight marker; (T) Total cell lysate; (S) Soluble fraction of cell lysate. B) The solubility of VP1 WTP, Cap-Base B and CapH190-2. The protein solubility was calculated by measuring the relative intensity of the relevant bands in the total fraction in comparison to the soluble fraction using the SDS-PAGE gel image. The analysis was performed using Image Lab software (Bio-rad).

GST-tagged Cap-Base B and CapH190-2 were then purified using GST affinity chromatography. GST tags were removed from untagged Cap-Base B and CapH190-2 capsomeres. Cap-Base B and CapH190-2 capsomeres were separated from the aggregate and GST tag using size exclusion chromatography (SEC) (Figure 3.15A). Similar to the chromatogram for Cap-Base A and CapH190-GCN4 in Figure 3.5A, the chromatogram also consisted of three major peaks as observed in a previous study (Middelberg *et al.*, 2011). They were: i) a peak corresponding to the aggregate eluted at about 8.5 ml; ii) a peak corresponding to capsomere at about 11.5 ml; and iii) a peak corresponding to the GST tag at about 16 ml. SEC profile shows that CapH190-2 eluted earlier than Cap-Base B. It indicates that the hydrodynamic radius of CapH190-2 is bigger than for Cap-Base B, resulting from the insertion of H190 modules on each monomer of CapH190-2. The SDS-PAGE analysis (Figure 3.15B) in agreement with SEC result (Figure 3.15A) shows that CapH190-2 (lane 2) has higher molecular weight than Cap-Base B (lane 1).

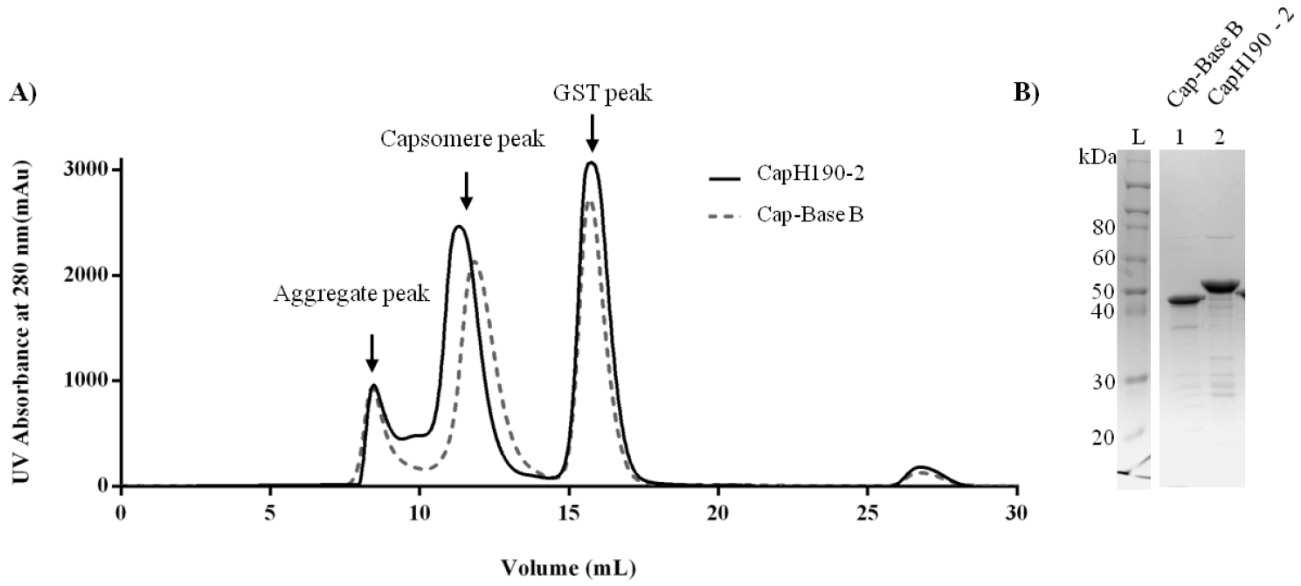


Figure 3.15: A) SEC purification of Cap-Base B (broken line) and CapH190-2 (solid line) constructs after GST removal from untagged capsomeres. B) SDS-PAGE analysis of Cap-Base B and CapH190-2 after SEC purification. Lanes: (L) molecular weight marker; (1) Cap-Base B; (2) CapH190-2.

3.3.2.2. Phage antibody panning

The panning process was done following the procedure described in section 3.3.1.2. Elution conditioner was applied, as was done for the panning of base A and H190-GCN4 in section 3.3.1.2.

3.3.2.3. Polyclonal phage ELISA

After the third panning against Cap-Base B as the subtractive antigen and CapH190-2 as the selective antigen in the binding step (Figure 3.6), the enrichment of binding phages was examined using polyclonal phage ELISA. Figure 3.16 shows the increase in absorbance value after each round, indicating the enrichment of CapH190-2-binding phages. The isolated pool of phages, however, responded similarly to both Cap-Base B and CapH190-2, indicating a lack of selectivity for phages binding to the H190 module, consisting of 2 copies of H190 element, on the modular capsomere.

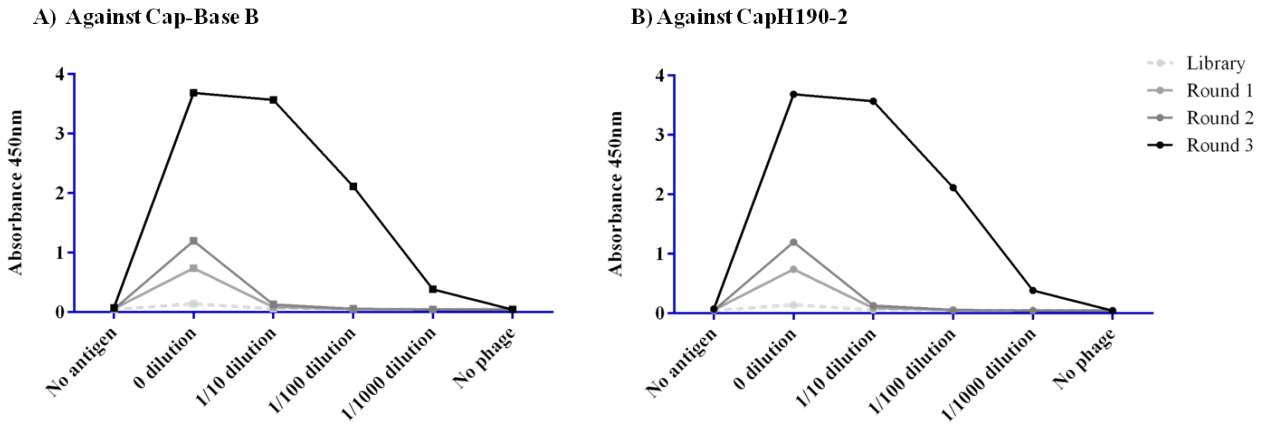


Figure 3.16: Polyclonal phage ELISA of ARC library-phages after panning rounds 1, 2 and 3 against (A) Cap-Base B and (B) CapH190-2.

A lack of selectivity for the CapH190-2-binding phages after the third panning could reflect two possibilities. First, it is possible that the subtractive panning might be inefficient in removing the Cap-Base B-binding phages. The repetition of the subtractive panning against Cap-Base B up to 6 times was to remove Cap-Base B-binding phages. The selectivity of the H190 module binding-phages was still not obtained (data not shown).

Second, the ARC library possibly lacks phages that can bind to the H190 module on the CapH190-2, resulting in the non-selectivity of CapH190-2-binding phages or no binding against the H190

module on CapH190-2. The second possibility was investigated through the panning of the isolated phages after the third pannings against H1N1 HA1 for a further two rounds. Figure 3.17 illustrates that there was a considerable decrease in the quantity of CapH190-2-binding phages after the pannings against H1N1 HA1, demonstrated by the decline of absorbance values in rounds 4 and 5. The finding reveals that the isolated CapH190-2-binding phages, after the third panning, are unable to bind to the H190 module on CapH190-2, resulting in either the unbound to H1N1 HA1 or the elimination of the phages in the pannings against H1N1 HA1. The isolated pool of phages in rounds 4 and 5 also contained some phages bound to H1N1 HA1.

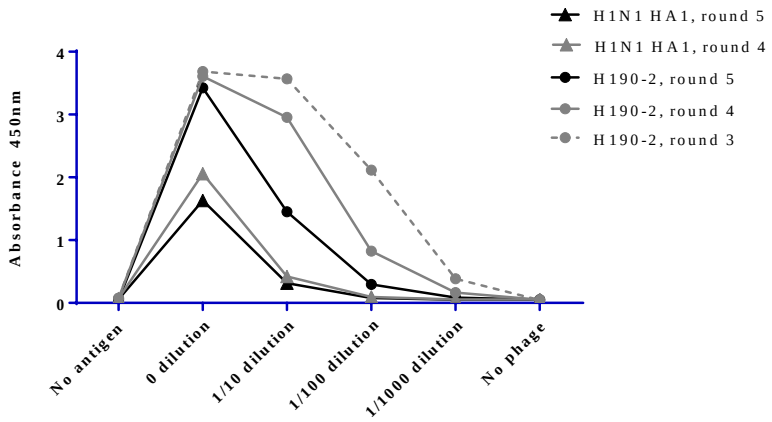


Figure 3.17: Polyclonal phage ELISA of ARC library-phages after the fourth and fifth panning against H1N1 HA1.

3.3.2.4. Monoclonal phage ELISA

The polyclonal phage ELISA (Figure 3.17) indicated that phages bound to both CapH190-2 and H1N1 HA1 might nonpresent in the isolated pool. The screening of 84 clones in monoclonal phage ELISA (Figure 3.18) found two phages highly responded to both Cap-Base B and CapH190-2 and a few phages bound weakly responded to only H1N1 HA1. The results confirmed that the isolated pool of phages, after the panning against H1N1 HA1, contains no phages bound to both CapH190-2 and H1N1 HA1.

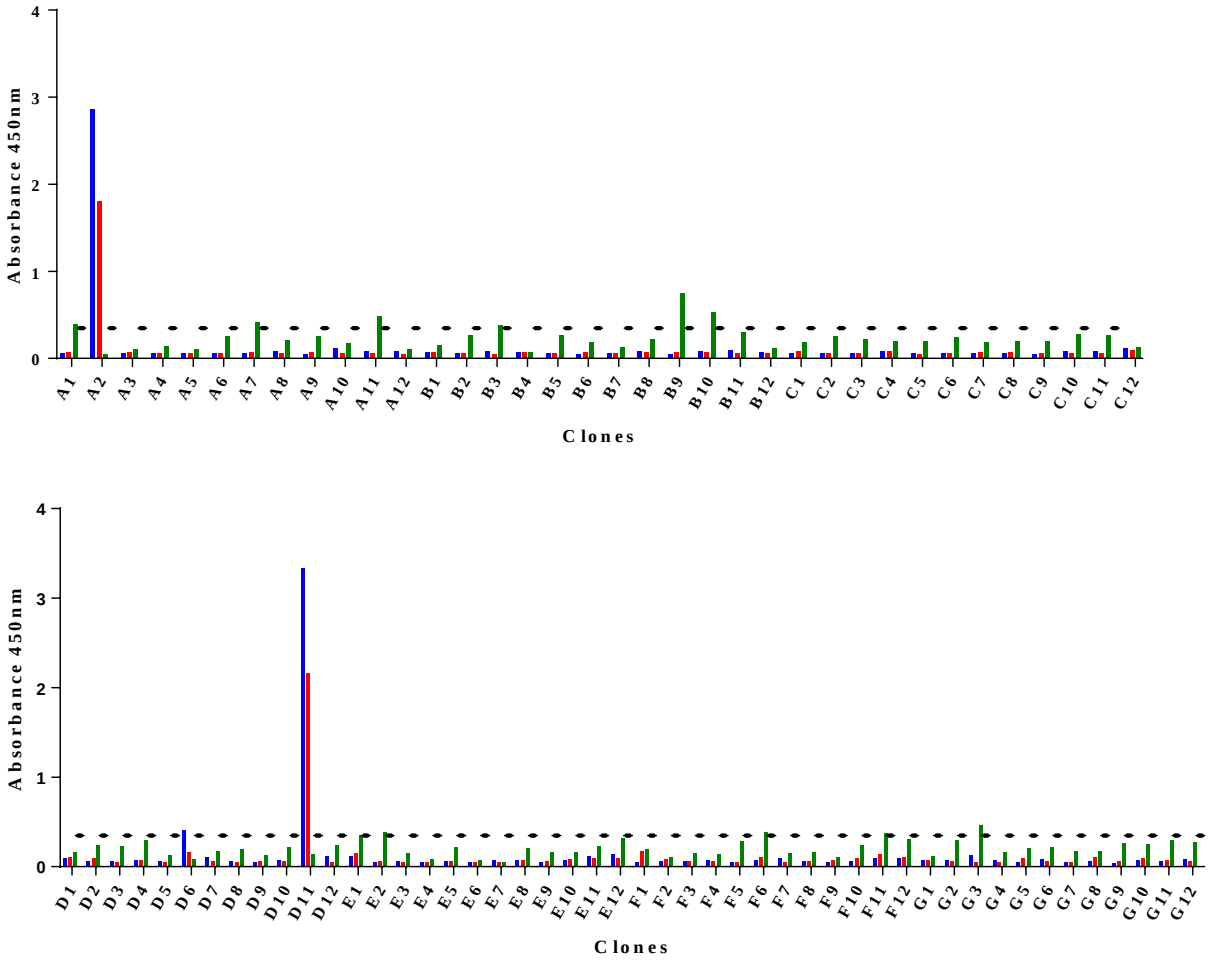


Figure 3.18: Monoclonal phage ELISA of ARC library-phages after the fifth panning against Cap-Base B (blue), CapH190-2 (red) and H1N1 HA1 (green). Dashed line: Mean of systematic error + 5 standard deviations.

In summary, the set of panning A identified phages bound to both CapH190-GCN4 and H1N1 HA1 in ARC library. Competitive ELISA against peptide H190 and peptide H190-GCN4 shows that the phages might bind to the junction between the H190 element and GCN4 element. It indicated that no H190-binding phages were isolated in the panning. The set of panning B identified no phages bound to both CapH190-2 and H1N1 HA1 in ARC library. It is possible that the structures of the H190 element on CapH190-2 may be different from native H190. The native H190 binding-phages, therefore, might fail to recognise the H190 element on CapH190-2. It might lead to a situation of no selectivity for CapH190-2-binding phages (Figure 3.16).

A study illustrated that the presentation of dual H190 elements on CapH190-2 led to a significant induction of HA1-specific IgG titre, compared to the usage of the GCN4 linker (Anggraeni *et al.*, 2013). The result indicated that the structure of the H190 element presented on CapH190-2 more

closely matches to native H190 conformation. It is possible is that H190-specific antibodies might present at low frequency and/or low affinity in the ARC library, leading to the difficulty in isolating CapH190-2-binding phages.

This research also sought CapH190-2-binding phages in another phage antibody library. Sheets' library is a non-immune library generated by Sheets *et al*. in 1998. This library contains 6.7 x 10^9 members in scFv form (Sheets *et al.*, 1998). The pannings, done according to the procedure described in Section 3.3.1.2, led to no enrichment of phages bound to both CapBase-B and CapH190-2 (data not shown). This study also modified the panning strategy by implementing the panning against H1N1 HA1 before the pannings against the CapBase-B and CapH190-2, but no CapH190-2- and H1N1 HA1- binding phages in Sheets' library were isolated. The non-selectivity of H190 specific antibodies is unlikely to be a reflection of library quality, as the other research group at The University of Queensland successfully isolated antibodies specific for target antigens in ARC and Sheet's libraries (Jones *et al.*, 2016; Lebani *et al.*, 2017).

The findings suggest that both ARC and Sheet's libraries might lack the A/California/07/2009 H190-specific phages. The generation of ARC and Sheet's libraries were in 2004 and 1998, respectively. The B lymphocytes, used as the materials for the construction of both libraries, were collected before the emergence of the A/California/07/2009 strain. The donors were unlikely to be infected by the influenza virus strain. Therefore, it is highly possible that antibodies against A/California/07/2009 H190 element may not exist in the immune system of donors, leading to no H190 element-binding phages in ARC and Sheet's libraries.

The possibility of isolating H190-specific antibodies can be increased through panning against immune library generated from A/California/07/2009 strain infected patients (Moreland *et al.*, 2012). The high mutagenic characteristic of influenza virus leads to the rapid change of H190 sequence (Yang *et al.*, 2007). The selection of H190 element-binding phages through antibody panning requires re-generating an immunised library that matches the specific strain of influenza virus. However, the construction of the phage antibody library is costly and time-consuming. It reveals that phage antibody panning is an inappropriate approach to identify the structure of highly mutagenic epitopes, such as influenza epitopes, on modular capsomere for influenza vaccine development.

3.4. Conclusion

In summary, the result suggests that the ARC and Sheet's libraries might lack the A/California/07/2009 H190-binding phages. It requires the generation of an immunised library that is relevant to the A/California/07/2009 strain. The construction of an immunised library is costly, laborious and time-consuming. A necessity of influenza vaccine development, however, is the rapid response to the antigenic change of the virus. The investigation of phage antibody panning in this chapter indicates that this technology is an inappropriate tool to identify the conformational presentation of hypervariable epitopes, such as Helix 190, on VLPs due to the requirement of regular generation of the phage antibody library which matches the target influenza strains.

References

Aitken, A., Learmonth, M., 2009. Protein determination by UV absorption, in: Walker, J. (Ed.), The protein protocols handbook. Humana Press, pp. 3-6.

Anggraeni, M.R., Connors, N.K., Wu, Y., Chuan, Y.P., Lua, L.H.L., Middelberg, A.P.J., 2013. Sensitivity of immune response quality to influenza helix 190 antigen structure displayed on a modular virus-like particle. Vaccine 31, 4428-4435.

Bernal, A.J., Willats, W.G.T., 2004. Plant science in the age of phage. Trends Plant Sci 9, 465-468.

Bradbury, A.R.M., Marks, J.D., 2004. Antibodies from phage antibody libraries. J Immunol Methods 290, 29-49.

Brichta, J., Hnilova, M., Viskovic, T., 2005. Generation of hapten-specific recombinant antibodies: antibody phage display technology: a review. Vet. Med. 50, 231-252.

Caton, A.J., Brownlee, G.G., 1982. The antigenic structure of the influenza virus A/PR/8/34 hemagglutinin (H1 subtype). Cell 31, 417-427.

Chuan, Y.P., Lua, L.H.L., Middelberg, A.P.J., 2008. High-level expression of soluble viral structural protein in *Escherichia coli*. J Biotechnol 134, 64-71.

De Bruin, R., Spelt, K., Mol, J., Koes, R., Quattrocchio, F., 1999. Selection of high-affinity phage antibodies from phage display libraries. Nature 17, 397-399.

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

Jones, M.L., Alfaleh, M.A., Kumble, S., Zhang, S., Osborne, G.W., Yeh, M., Arora, N., Hou, J.J.C., Howard, C.B., Chin, D.Y., Mahler, S.M., 2016. Targeting membrane proteins for antibody discovery using phage display. Sci Rep 6, 26240.

Joshi, H., Lewis, K., Singharoy, A., Ortoleva, P.J., 2013. Epitope engineering and molecular metrics of immunogenicity: A computational approach to VLP-based vaccine design. Vaccine 31, 4841-4847.

Kang, A.S., Barbas, C.F., Janda, K.D., Benkovic, S.J., Lerner, R.A., 1991. Linkage of recognition and replication functions by assembling combinatorial antibody Fab libraries along phage surfaces. Proc Natl Acad Sci USA 88, 4363-4366.

Kashyap, A.K., Steel, J., Oner, A.F., Dillon, M.A., Swale, R.E., Wall, K.M., Perry, K.J., Faynboym, A., Ilhan, M., Horowitz, M., Horowitz, L., Palese, P., Bhatt, R.R., Lerner, R.A., 2008. Combinatorial antibody libraries from survivors of the Turkish H5N1 avian influenza outbreak reveal virus neutralization strategies. Proc Natl Acad Sci USA 105, 5986-5991.

Kretzschmar, T., Geiser, M., 1995. Evaluation of antibodies fused to minor coat protein III and major coat protein VIII of bacteriophage M13. Gene 155, 61-65.

Lebani, K., Jones, M.L., Watterson, D., Ranzoni, A., Traves, R.J., Young, P.R., Mahler, S.M., 2017. Isolation of serotype-specific antibodies against dengue virus non-structural protein 1 using phage display and application in a multiplexed serotyping assay. PLoS One 12, e0180669.

Liew, M.W.O., Chuan, Y.P., Middelberg, A.P.J., 2012. High-yield and scalable cell-free assembly of virus-like particles by dilution. Biochem Eng J 67, 88-96.

Lipin, D.I., Raj, A., Lua, L.H.L., Middelberg, A.P.J., 2009. Affinity purification of viral protein having heterogeneous quaternary structure: Modeling the impact of soluble aggregates on chromatographic performance. J Chromatogr 1216, 5696-5708.

Lu, Y., Welsh, J.P., Swartz, J.R., 2014. Production and stabilization of the trimeric influenza hemagglutinin stem domain for potentially broadly protective influenza vaccines. Proc Natl Acad Sci USA 111, 125-130.

Marks, J.D., Hoogenboom, H.R., Bonnert, T.P., McCafferty, J., Griffith, A.D., Winter, G., 1991. By-passing immunization human antibodies from v-gene libraries displayed on phage. J Mol Biol 222, 581-597.

McCafferty, J., Griffiths, A.D., Winter, G., Chiswell, D.J., 1990. Phage antibodies: filamentous phage displaying antibody variable domains. Nature 348, 552-554.

Middelberg, A.P.J., Rivera-Hernandez, T., Wibowo, N., Lua, L.H.L., Fan, Y., Magor, G., Chang, C., Chuan, Y.P., Good, M.F., Batzloff, M.R., 2011. A microbial platform for rapid and low-cost virus-like particle and capsomere vaccines. Vaccine 29, 7154-7162.

Moreland, N.J., Susanto, P., Lim, E., Tay, M.Y.F., Rajamanonmani, R., Hanson, B.J., Vasudevan, S.G., 2012. Phage display approaches for the isolation of monoclonal antibodies against Dengue virus envelope Domain III from human and mouse derived libraries. Int J Mol Sci 13, 2618-2635.

Sahin, E., Grillo, A.O., Perkins, M.D., Roberts, C.J., 2010. Comparative effects of pH and ionic strength on protein–protein interactions, unfolding, and aggregation for IgG1 antibodies. J Pharm Sci 99, 4830-4848.

Schneeman, A., Speir, J.A., Tan, G.S., Khayat, R., Ekiert, D.C., Matsuoka, Y., Wilson, I.A., 2012. A virus-like particle that elicits cross-reactive antibodies to the conserved stem of influenza virus hemagglutinin. J Virol 86, 11686-11697.

Sela-Culang, I., Kunik, V., Ofran, Y., 2013. The structural basis of antibody-antigen recognition. Front Immunol 4, 302.

Sheets, M.D., Amersdorfer, P., Finnern, R., Sargent, P., Lindqvist, E., Schier, R., Hemingsen, G., Wong, C., Gerhart, J.C., Marks, J.D., 1998. Efficient construction of a large nonimmune phage antibody library: The production of high-affinity human single-chain antibodies to protein antigens. Proc Natl Acad Sci USA 95, 6157-6162.

Smith, G.P., 1985. Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. Science 228, 1315-1317.

Sui, J., Hwang, W.C., Perez, S., Wei, G., Aird, D., Chen, L., Santelli, E., Stec, B., Cadwell, G., Ali, M., Wan, H., Murakami, A., Yammanuru, A., Han, T., Cox, N.J., Bankston, L.A., Donis, R.O., Liddington, R.C., Marasco, W.A., 2009. Structural and functional bases for broad-spectrum neutralization of avian and human influenza A viruses. Nat Struct Mol Biol 16, 265-272.

Tekewe, A., Fan, Y., Tan, E., Middelberg, A.P.J., Lua, L.H.L., 2016. Integrated molecular and bioprocess engineering for bacterially produced immunogenic modular virus-like particle vaccine displaying 18 kDa rotavirus antigen. Biotechnol Bioeng 114, 397-406.

Throsby, M., van den Brink, E., Jongeneelen, M., Poon, L.L.M., Alard, P., Cornelissen, L., Bakker, A., Cox, F., van Deventer, E., Guan, Y., Cinalt, J., ter Meulen, J., Lasters, I., Carsetti, R., Peiris, M., de Kruif, J., Goudsmit, J., 2008. Heterosubtypic neutralizing monoclonal antibodies cross-protective against H5N1 and H1N1 recovered from human IgM^+ memory B cells. PLoS One 3, e3942.

Wibowo, N., Chuan, Y.P., Lua, L.H.L., Middelberg, A.P.J., 2012. Modular engineering of a microbially-produced viral capsomere vaccine for influenza. Chem Eng Sci 103, 12-20.

Wibowo, N., Hughes, F.K., Fairmaid, E.J., Lua, L.H.L., Brown, L.E., Middelberg, A.P.J., 2014. Protective efficacy of a bacterially produced modular capsomere presenting M2e from influenza: Extending the potential of broadly cross-protecting epitopes. Vaccine 32, 3651-3655.

Yang, Z.-Y., Wei, C.-J., Kong, W.-P., Wu, L., Xu, L., Smith, D.F., Nabel, G.J., 2007. Immunization by avian H5 influenza hemagglutinin mutants with altered receptor binding specificity. Science 317, 825-828.

Yu, X., Tsibane, T., McGraw, P.A., House, F.S., Keefer, C.J., Hicar, M.D., Tumpey, T.M., Pappas, C., Perrone, L.A., Martinez, O., Stevens, J., Wilson, I.A., Aguilar, P.V., Altschuler, E.L., Basler, C.F., Jr. Crowe, J.E., 2008. Neutralizing antibodies derived from the B cells of 1918 influenza pandemic survivors. Nature 455, 532-536.

Zani, M.-L., Moreau, T., 2010. Phage display as a powerful tool to engineer protease inhibitors. Biochimie, 1-16.

Chapter 4

Conformational screening of a helical antigenic peptide displayed on a virus-like particle

The entire Chapter 4 consists of the journal article submitted as:

Tam T. Doan[a], Natalie K. Connors[a], Nani Wibowo[a], Anton P.J. Middelberg[a], Linda H.L. Lua[b]*
(2017). Conformational screening of a helical antigenic peptide displayed on a virus-like particle. *Submitted to Protein Science* (Manuscript Number: PRO-17-0142).

[a] The University of Queensland, Australian Institute for Bioengineering and Nanotechnology, St. Lucia, QLD 4072, Australia

[b] The University of Queensland, Protein Expression Facility, St Lucia, QLD 4072, Australia

*Corresponding author: Linda H.L. Lua

The following modifications were made to the article:

- Page numbers of the original article were changed into numbers consistent with those of the overall thesis.
- Figure numbers were changed into numbers consistent with the remainder of figures in the thesis.
- Section and sub-section numbers were changed into numbers consistent with the remainder of sections and sub-sections in all chapters of the thesis.
- Section order was changed into the order consistently used in all chapters of the thesis.
- The reference style of the original article was changed into the style consistently used in all chapters of the thesis.

4.1. Abstract

Modular virus-like particles (VLPs), presenting foreign antigenic elements, have been shown as safe and effective vaccines. The design rule for presenting antigenic elements in native structure is yet to be fundamentally determined. Tools that allow rapid screening of antigenic element structures are useful to enable rapid screening of alternative vaccine designs. Here, we present the use of circular dichroism spectroscopy coupled with singular value decomposition algorithms (CD-SVD) as a potential screening tool to investigate the structural presentation of an antigenic element. Helical influenza hemagglutinin antigenic element, Helix A (HA2A), was modularised and presented on murine polyomavirus (MuPyV) VLP vaccine platform. CD-SVD was used to estimate the ratios of helices within modular and base capsomeres. These ratios were then converted to helical amino acid numbers. The increment in the helical amino acid number of modular capsomeres in comparison to base capsomere indicates the formation of the helical structure of antigenic module on modular capsomeres. The structural analysis of HA2A modular capsomeres was in good agreement with molecular dynamics simulation and *in vivo* immunogenicity testing, indicating that CD-SVD can simply and rapidly identify potential constructs at an early stage of modular VLP-based vaccine development. The study reveals that a tandem repeat strategy, whereby multiple copies of an epitope are inserted, promotes the helicity of inserts derived from native helical regions; this approach deserves further research into its potential as a generic design for the presentation of the helical antigenic elements on modular VLPs.

Keywords: virus-like particle, influenza, peptide antigen, structural presentation, vaccine design

Abbreviations and symbols:

CD: circular dichroism; HA2A: helix A; FL-HA: full-length HA; MD: molecular dynamics; MuPyV: murine polyomavirus; RMSD: root-mean-square deviation; SVD: singular value decomposition algorithms; VLP: virus-like particle.

4.2. Introduction

Vaccines play a major role in fighting against infectious diseases, for example in the eradication of smallpox (Henderson and Fenner, 2001), and in the significant reduction of measles and polio cases (Heinsbroek and Ruitenberg, 2010; Kew *et al.*, 2005). Recent vaccine developments have been directed toward a subunit approach, which uses only parts of the organism in contrast to Louis Pasteur's whole organism approach, for a highly-purified, well-characterised, and safer vaccine (Thomas and Luxon, 2013). However, subunit vaccines have low immunogenicity. To enhance the immunogenicity of a subunit antigen, protein carriers or adjuvant formulations are often used (Jennings and Bachmann, 2008).

A virus-like particle (VLP) is a highly ordered assembly of viral capsid proteins (Chackerian, 2007). The initial use of VLPs as immunogens is to elicit an immune response against their parental virions (Assad and Francis, 1999; McAleer *et al.*, 1984; Schiller *et al.*, 2008). An emerging approach is developing VLPs as immunological carriers to present a heterogeneous peptide antigen or an antigenic element on its surface (Liu *et al.*, 2000; Rivera-Hernandez *et al.*, 2013; Tegerstedt *et al.*, 2005; Yin *et al.*, 2011). These modular VLPs present the target antigenic element in an ordered repetitive array, leading to improving the immunogenicity of the antigenic element (Neirynck *et al.*, 1999).

The presentation of an antigenic element on the surface of a VLP can disrupt the conformational integrity of the antigenic element (Jennings and Bachmann, 2009; Roldao *et al.*, 2010; Tissot *et al.*, 2010). Antigen-antibody recognition depends on antigen structure (Sela-Culang *et al.*, 2013); antigen structure needs to resemble its native structure to induce neutralising antibodies. Antigen design is essential to retain the authentic structure of the antigen (Pantophlet and Burton, 2006). A recent study showed that display strategies of an antigenic element on VLPs influenced the quality of the immune response (Anggraeni *et al.*, 2013). However, the design rules to natively present the conformation of an antigenic element on the surface of VLP remains unclear. The structural design of antigenic element has to date been approached empirically with testing of each construct via *in vivo* study or by solving the protein structure via cryo-electron microscopy or X-ray crystallography. *In vivo* testing is costly; especially to test a large number of constructs required defining design rules. Cryo-electron microscopy (Schneeman *et al.*, 2012) or X-ray crystallography (Lin *et al.*, 1996) is a tedious and time-consuming method which is also unsuitable for testing several constructs. A simpler method that can quantitatively benchmark the conformation of an inserted antigenic element, against its native source structure, is required.

Circular dichroism (CD) spectroscopy is a simple tool able to rapidly determine protein secondary structures (Greenfield, 2007; Whitmore and Wallace, 2008), estimated using computation methods, i.e. singular value decomposition algorithms (SVD) (Sreerama and Woody, 2000). SVD algorithms can analyse protein CD spectra to provide the ratio of protein secondary structures including α-helix (Sreerama and Woody, 2000), which can then be used to calculate the number of amino acids having secondary structure. CD spectroscopy has been used previously to evaluate individual secondary structure components of a VLP subunit, i.e. a pentameric structure called a capsomere (Yang and Teng, 1998; Yang and Teng, 1999). It suggests that the CD spectroscopy coupled with SVD algorithms (CD-SVD) can inform the secondary structure of an antigenic module on a capsomere (modular capsomere) by analysing the helical amino acid number of modular capsomeres in comparison to the base capsomere.

Molecular dynamics (MD) simulation is a computational method that allows *in silico* investigation of a macromolecular structure, e.g. viruses and VLPs. MD simulations have been applied to predict epitope presentation on rhinovirus capsid subunits (Lapelosa *et al.*, 2009) and the structure of modularised antigenic peptides on a VLP (Anggraeni *et al.*, 2013; Joshi *et al.*, 2011; Joshi *et al.*, 2013). This suggests that MD simulation of a modular capsomere could assist CD-SVD to inform the structure of a given antigenic module via predicting the structural deviation between the antigenic module and its native structure.

We hypothesise that CD-SVD, assisted with MD simulation, can characterise the structure of antigenic modules on modular capsomeres. This study employed a vaccine platform based on murine polyomavirus (MuPyV) as a carrier for the presentation of antigenic modules due to the advantages of this platform regarding speed, high yield and low-cost (Middelberg *et al.*, 2011; Waneesorn *et al.*, 2016). A previous study reports that tandem repeat display strategy, which an antigenic element is repetitively arrayed in a module, would be effective to maintain native helical propensity (Anggraeni *et al.*, 2013). The presentation of the dual copies of a helical antigenic element, from influenza hemagglutinin, resulted in the induction of a higher level of IgGs recognising HA1, compared with the use of flanking sequence. This study investigated the effect of the tandem repeat strategy on the structural presentation of the helical structures while antigenic element number was increased within a module. The antigen of interest is Helix A (HA2A) in the stalk region of HA (Figure 4.1A). HA2A is a potentially cross-protective epitope (Ekiert *et al.*, 2009; Sui *et al.*, 2009) and a major determinant for the binding of neutralising antibodies CR6261 and F10 (Ekiert *et al.*, 2009; Sui *et al.*, 2009; Throsby *et al.*, 2008).

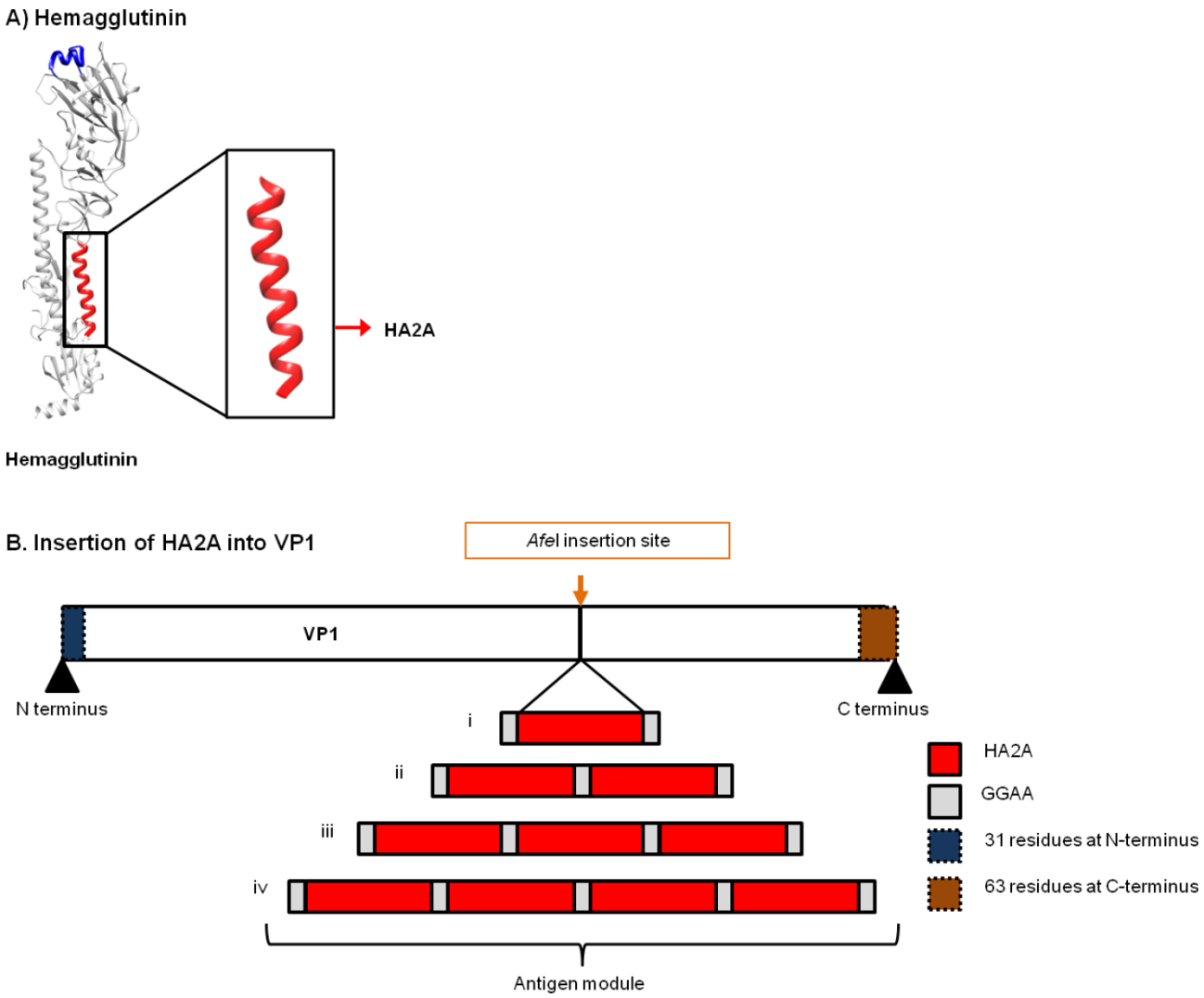


Figure 4.1: Crystal structure of influenza HA and the insertion of HA2A into VP1. A) The location of HA2A on influenza HA. The picture was generated using Chimera UCSF Chimera software version 1.8 from 3LZG.pdb. B) The schematic diagrams show the insertion of HA2A (red) into VP1. The antigenic modules contain i) one antigenic element, ii) two antigenic elements, iii) three antigenic elements or iv) four antigenic elements.

This study utilised CD-SVD assisted with MD simulation to analyse the helical structure of HA2A elements inserted within modular VP1 capsomeres. The data are consistent between structural information from CD-SVD, the conformational prediction from MD simulation and antibody response from *in vivo* immunogenicity testing. Results confirm the ability of a tandem repeat strategy to promote the helicity of HA2A within the modular capsomere. The finding also indicates that CD-SVD allows a rapid identification of potential vaccine constructs in the early stage of vaccine development reducing overall testing cost and time to functional development.

4.3. Materials and Methods

4.3.1. Homology modelling and molecular dynamics (MD) simulation for modular capsomeres

Homology models and simulations were performed as described previously (Lua *et al.*, 2015). A model of each modular HA2A monomer was first constructed using murine polyomavirus PDB templates 1SIE for VP1 and 4JTVB for HA2A antigens, using Accelrys Discovery Studio® 3.0. The single modular monomers were modelled within an unmodified capsomere structure to maintain authentic VP1 structure and steric boundaries for antigen presentation. Modular capsomeres were then modelled using the modular HA2A monomer and 1SIE capsomere as templates.

Molecular dynamics (MD) simulations for modular capsomeres were performed using GROMACS version 4.6 with Gromos96 43a1 force field and simple point charge (SPC) model for water. Each modular capsomere was simulated individually with solvent within a cubic box ($18nm^3$); solvent molecules (Na^+ ions) were randomly placed to charge-neutralise the system. Berendson method was used to control the temperature at 298K, with a time constant of 0.1 ps, and pressure at 1 atm with coupling constant of 1.0 ps. The simulation was run with an integration time step of 2 fs. Particle-mesh Ewald (PME) algorithm was used to account for electrostatic interactions. Neighbouring atom cut-offs, Coulomb and Lennard-Jones (LJ) potential energies were all set to 1.0 nm. Steepest descent energy minimisation was performed for 5000 steps, followed by a 100 ps equilibration with protein heavy atom position restraint. The MD simulation was then performed for 20 ns and was performed in triplicate for each modular capsomere. Simulations were performed on Australia's National Computational Infrastructure High-Performance Cluster (HPC) Raijin (Fujitsu Primergy HPC), scheduled with a PBS Professional job scheduler, using 240 cores (Intel Xeon Sandy Bridge Technology, 2.6GHz) and 4 GB/core memory for approximately 47 h per simulation.

Accelrys Discovery Studio® 3.0 was used to calculate the root-mean-square deviation (RMSD) of the HA2A modules and VP1 capsomere. Final conformations of a single HA2A copy within the modules were compared with the native conformation of HA2A region in the structure on 4JTVB.pdb.

4.3.2. Generation of modular VP1 constructs

Vectors pGEX-VP1, pGEX-VP1-ΔC63 and pGEX-VP1-S1S4 were generated as in previous work (Wibowo *et al.*, 2012). A DNA fragment encoding VP1-ΔN31-ΔC63-S1S4 was generated using designed primers to exclude 31 amino acids at N-terminus and 63 amino acids at C-terminus of VP1-S1S4. The DNA fragment was inserted between *Bam*HI and *Xho*I recognition sites on vector pGEX, yielding vector pGEX-VP1-ΔN31-ΔC63-S1S4.

The DNA sequence of HA2A encoding DLKSTQNAIDEITNKVNSVIEK from A/California/07/2009 H1N1 was flanked with GGAA sequence (Figure 4.1B). The gene inserts of one, two, three or four copies of HA2A were optimised according to codon usage in *E.coli* genes and assembled from oligos calculated from DNAWorks (http://helixweb.nih.gov/dnaworks/). Gene inserts were cloned into *Afe*I recognition sites in the vector pGEX-VP1-S1S4 or pGEX-VP1-ΔN31-ΔC63-S1S4 by *in vivo* cloning using One Shot®OmniMAX™2T1^{R} chemically competent *E.coli* cloning strain (Invitrogen™, Victoria, Australia) to yield the full-length and truncated versions for each construct, respectively (Figure 4.1B). The designated names of these expression vectors and protein names are shown in Table 4.1. All full-length and truncated modular capsomeres were expressed in *E.coli* and purified as described in previous works (Chuan *et al.*, 2008; Lipin *et al.*, 2009).

Table 4.1: The designated names of the expression vectors and proteins

Antigenic module	Expression vectors	Designated protein name
	pGEX-VP1	VP1
	pGEX-VP1-ΔC63	VP1-ΔC
	pGEX-VP1-ΔN31-ΔC63-S1S4	VP1-ΔNΔC or Base capsomere
One copy of HA2A	pGEX-VP1-S1S4-HA2A-1x[1]	VP1-HA2A-1
	pGEX-VP1-ΔN31-ΔC63-S1S4-HA2A-1x[2]	VP1-ΔNΔC-HA2A-1
Two copies of HA2A	pGEX-VP1-S1S4-HA2A-2x[1]	VP1-HA2A-2
	pGEX-VP1-ΔN31-ΔC63-S1S4-HA2A-2x[2]	VP1-ΔNΔC-HA2A-2
Three copies of HA2A	pGEX-VP1-S1S4-HA2A-3x[1]	VP1-HA2A-3
	pGEX-VP1-ΔN31-ΔC63-S1S4-HA2A-3x[2]	VP1-ΔNΔC-HA2A-3
Four copies of HA2A	pGEX-VP1-S1S4-HA2A-4x[1]	VP1-HA2A-4
	pGEX-VP1-ΔN31-ΔC63-S1S4-HA2A-4x[2]	VP1-ΔNΔC-HA2A-4

[1]: full-length version

[2]: Truncated version

4.3.3. Circular dichroism (CD) spectroscopy

Truncated modular capsomeres were diluted to 0.1 mg mL^{-1} concentration in 10mM Tris base, pH 8.0. Quartz cuvette with 0.1 cm path length was utilised to measure CD spectra. Far-UV CD spectra were measured using a Jasco Model J-815 spectropolarimeter (Japan Spectroscopic Co. Ltd, Japan) at 0.1nm data pitch, 2nm bandwidth, 2 second data integration time and 10 nm min^{-1} scanning speed.

4.3.4. Estimation of secondary structure fraction by CDSSTR program

CDSSTR, a program on Dichroweb (Department of Crystallography, Birkbeck College, University of London, London, UK), can estimate protein secondary structure content using a singular value decomposition (SVD) algorithm and equation (Compton and Johnson Jr, 1986; Johnson, 1999; Whitmore and Wallace, 2008). The CD spectra of truncated modular capsomeres and base capsomere were used to estimate the ratio of α-helix using CDSSTR. Preference set 4, containing 43 well-characterised proteins, was used as a protein reference dataset (Wallace and Janes, 2009). The ratios of α-helix were converted into helical amino acid numbers (helical amino acid number = the ratio of helical structure x total amino acids).

4.3.5. Assembly of modular VLPs

Endotoxin was removed from purified full-length modular capsomeres using Vivapure Q Maxi H column (Sartorius stedim, Goettingen, Germany) as previously described (Wibowo *et al.*, 2014). Endotoxin level (below 5 EU/mL) was analysed using LAL-based assay EndosafePTS^{TM}-2005 (Charles River Laboratory, MA, USA). The endotoxin-free full-length modular capsomeres were assembled *in vitro* into modular VLPs and dialysed against PBS.

4.3.6. Animal immunisation

The modular HA2A VLPs were immunised on female *Gallus Gallus* chicken at three weeks of age (8 chickens/group). Chickens were immunised with a dose of 100 μg of modular VLPs on days 0 and 21. Chickens were bled from wing veins on days 0, 21 and 35. Day 35 sera were used for ELISA. The experimental animal work was reviewed and approved by The University of Queensland Animal Ethics Committee (AIBN/235/14/QSF). The animal study was conducted at Poultry Research Unit, Gatton campus, The University of Queensland, Queensland, Australia.

4.3.7. ELISA

Baculovirus-insect cell-expressed recombinant full-length H1N1 (A/California/07/2009) hemagglutinin and antibody CR6261 were purchased from the Sino Biological Inc. (Beijing, China). Biotinylated peptide HA2A (biotin-GGAADLKSTQNAIDEITNKVNSVIEKGGAA-NH_2) was purchased from ChinaPeptides (Shanghai, China).

Antigens were immobilised on Maxisorp 96-well plates (Thermo Scientific, Victoria, Australia) and incubated with anti-modular HA2A VLPs sera at 1:50 dilution in 0.5% skim milk PBS for 1.5 hours. After washing, HRP-conjugated anti-chicken IgG (Sigma-aldrich®, Missouri, USA) was added to each well and incubated for 1.5 hours. UV absorbtion was measured at 450nm.

4.3.8. Statistical analysis

Statistical analysis was performed using GraphPad Prism® Version 6.0 (GraphPad Software Inc., CA, USA). Multiple comparisons of the relative helical amino acid number and anti-modular VLP sera in serological testings were conducted using one-way ANOVA with Tukey's *post hoc* test.

4.4. Results

4.4.1. Synthesis of MuPyV VP1 capsomere variants, truncated modular HA2A capsomeres and modular HA2A VLPs

Synthesis of truncated modular HA2A capsomeres will be discussed in sections 5.3.1, 5.3.2 and 5.3.3 in Chapter 5. Synthesis of MuPyV VP1 capsomere variants and modular HA2A VLPs are described as in Appendices 1 and 2, respectively.

4.4.2. Modularisation of HA2A module using tandem repeat display strategy

A previous study showed that the display strategy could significantly affect the quality of the immune response (Anggraeni *et al.*, 2013). The presentation of an influenza hemagglutinin antigenic element, using tandem repeat display strategy, induced a significant HA1-specific IgG level, in comparison with the use of flanking sequence. It suggested further investigation was needed into the effect(s) of tandem repeat on helical propensity by increasing the antigenic element copy number. This study inserted one, two, three, or four copies of HA2A element(s) into the insertion site of VP1 at DNA level, resulting in a HA2A module on each VP1 monomer bearing one, two, three, or four HA2A element(s), respectively. A VP1 monomer bears one HA2A module, resulting in a modular VP1 capsomere (comprising five modular VP1 monomers) having five HA2A modules (Figure 4.2).

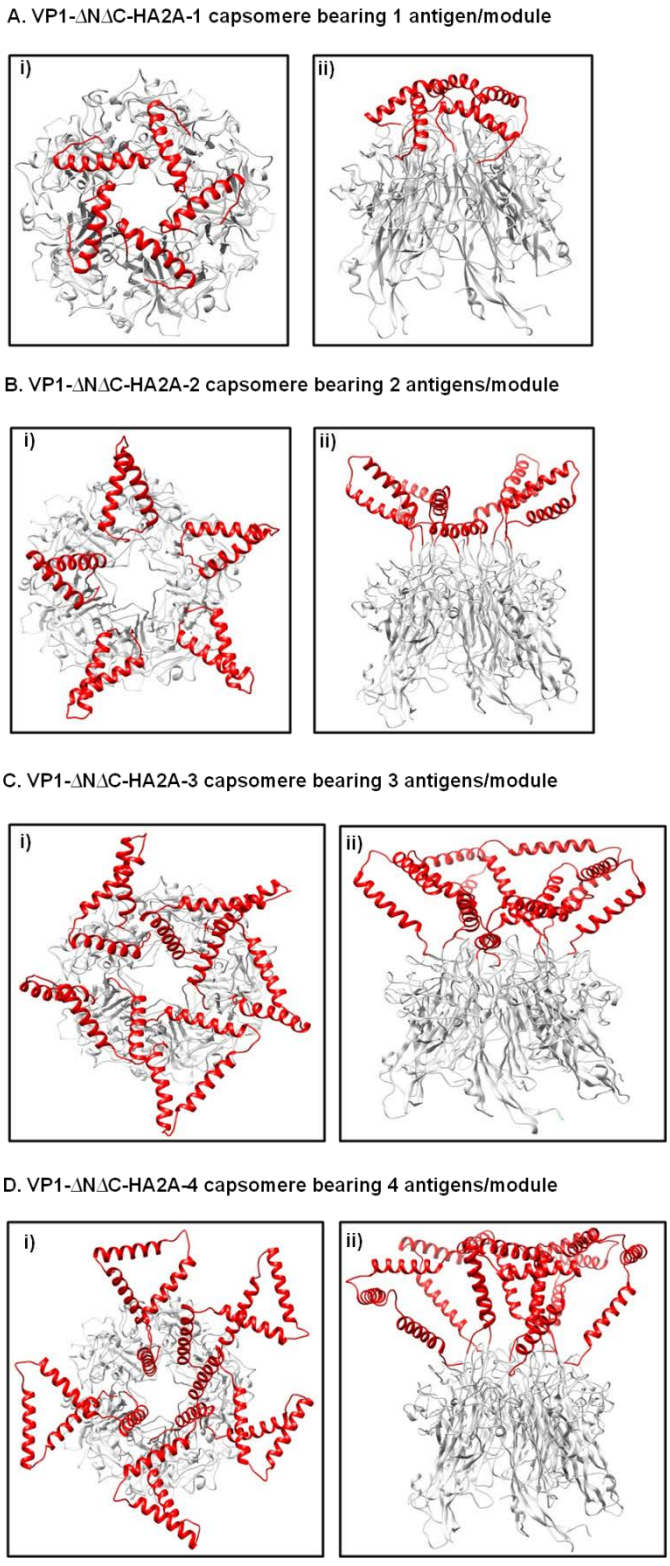


Figure 4.2: The modularisation of HA2A elements on modular HA2A capsomeres using tandem repeat display strategy. A) VP1-ΔNΔC-HA2A-1 capsomere bearing one HA2A element per module; B) VP1-ΔNΔC-HA2A-2 capsomere bearing two HA2A elements per module; C) VP1-ΔNΔC-HA2A-3 capsomere bearing three HA2A elements per module; D) VP1-ΔNΔC-HA2A-4 capsomere bearing four HA2A elements per module; i) Top view; ii) Side view. The HA2A module is highlighted in red and VP1 capsomere in light grey. The homology modelling pictures were generated using Accelrys Discovery Studio 3.0 based on 1SIE.pdb and 4JTVB.pdb by Natalie Connors.

4.4.3. CD-SVD method validation

The capability of CD-SVD to characterise the secondary structure content of capsomeres, especially α-helix, was first validated using MuPyV capsomere variants of known structure, i.e. VP1, VP1-ΔC, and VP1-ΔNΔC. Figure 4.3A presents the structural differences between these constructs. VP1 has an additional helix at C-terminus compared with VP1-ΔC and VP1-ΔNΔC. VP1-ΔC and VP1-ΔNΔC bear the equal quantity of helical amino acids despite the N-terminus truncation of VP1-ΔNΔC. Figure 4.3B compares the far-UV CD spectra of the capsomere variants, which were used to estimate the ratio of α-helix on Dichroweb. These α-helix ratios were converted to the helical amino acid numbers shown in Figure 4.3C. The structural information obtained from CD-SVD was in good agreement with the protein structures in Figure 4.3A, showing that VP1 had more helical amino acids than VP1-ΔC ($P<0.001$) and VP1-ΔNΔC ($0.001<P<0.01$), whereas VP1-ΔC and VP1-ΔNΔC had similar helical amino acid numbers.

4.4.4. Structural analysis of HA2A elements on modular capsomeres

This study employed CD-SVD to analyse the relative helical amino acid numbers of modular and base capsomeres. The increment in the helical amino acid numbers of modular capsomere compared with the base was used to indicate the helical structure of antigenic modules on modular capsomeres. Figure 4.3D reveals that the antigenic module on VP1-ΔNΔC-HA2A-1 may present in a non-helical structure, indicated by non-significant difference with the base capsomere. In contrast, VP1-ΔNΔC-HA2A-2, -3 and -4 shows a significant increase in relative helical amino acid numbers with 9, 11 and 23 amino acids, in comparison with the base capsomere ($P<0.001$), respectively (Appendix 4, Supplementary material). This result indicates the formation of helical structures on antigenic modules of VP1-ΔNΔC-HA2A-2, -3 and -4.

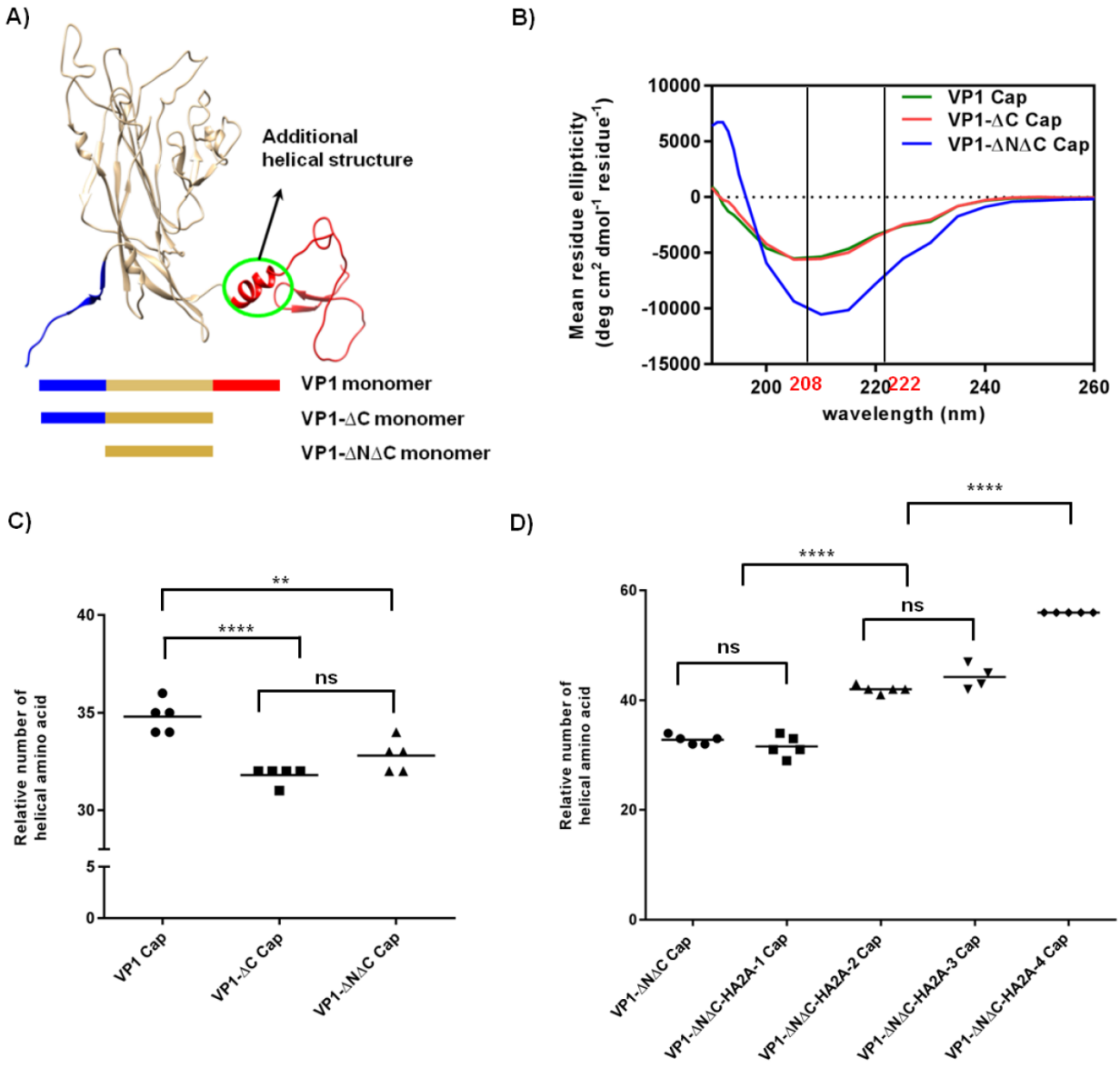


*Figure 4.3: CD-SVD method validation and the structural analysis of modular HA2A capsomeres. The comparison of (A) Conformation (1SID.pdb), (B) Far-UV CD spectra, and (C) Helical amino acids numbers among VP1, VP1-ΔC and VP1-ΔNΔC. D) The comparison of helical amino acid numbers of base capsomere and modular HA2A capsomeres. Statistical analysis was performed using GraphPad Prism Version 6.0. Multiple comparisons of helical amino acid numbers were performed with the Tukey's test. Bars represent geometric mean of helical amino acids from 5 replications. Ns: not significant (P≥0.05);**, very significant (0.001<P<0.01), ****, extremely significant (P<0.0001).*

4.4.5. MD simulation for modular HA2A capsomeres

MD simulation estimated the conformational deviation of the HA2A modules and modular capsomeres from their native structures. The structural simulation suggests that HA2A modules on VP1-ΔNΔC-HA2A-1 may have a high loss of native structure or non-helical structure, indicated by the high RMSD values for modules 1, 3, 4 and 5 (greater than 4Å) (Figure 4.4A). The increase in HA2A element copies per module on VP1-ΔNΔC-HA2A-2, VP1-ΔNΔC-HA2A-3 and VP1-ΔNΔC-HA2A-4 increased the quantity of the HA2A elements retaining relatively native structure (with less than 2Å RMSD) (Figures 4.4B, 4.4C and 4.4D). Figure 4.4E shows that the capsomere structures within VP1-ΔNΔC-HA2A-1, VP1-ΔNΔC-HA2A-2 and VP1-ΔNΔC-HA2A-3 present low RMSD values (about 4Å), indicating the capsomere structures might retain their native structure. MD simulation also suggests that the insertion of 4 HA2A elements per module on VP1-ΔNΔC-HA2A-4 might perturb the capsomere structure, proved by a high RMSD value at 8Å.

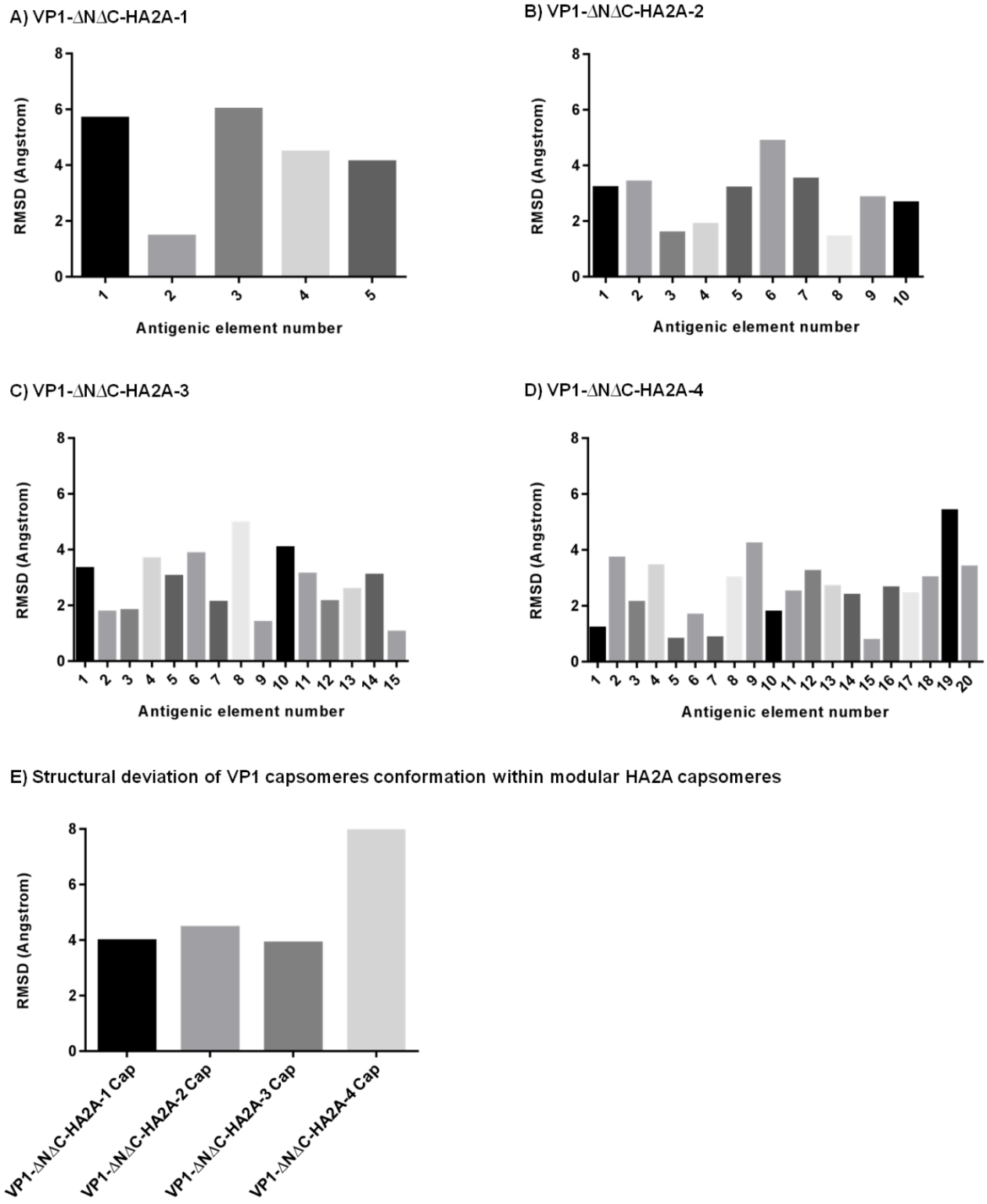


Figure 4.4: The structural prediction of modular HA2A capsomeres using MD simulation. The root-mean-square-deviation (RMSD) of C^{α} of HA2A antigenic elements on A) VP1-ΔNΔC-HA2A-1 capsomere; B) VP1-ΔNΔC-HA2A-2 capsomere; C) VP1-ΔNΔC-HA2A-3 capsomere and D) VP1-ΔNΔC-HA2A-4 capsomere, against the native HA2A conformation (4JTVB.pdb). E) Structural deviation of VP1 capsomeres conformation within modular HA2A capsomeres compared with unmodified VP1 capsomere (1SID.pdb).

4.4.6. *In vivo* immunogenicity testing

This study examined whether the presentation of HA2A modules, using tandem repeat display strategy, on VLP can lead to the induction of IgGs recognising HA2A peptide and the full-length HA (FL-HA). Figure 4.5A shows the immunogenicity of modular HA2A VLPs against peptide HA2A. While VP1-HA2A-1 and -2 were unable to induce antibodies able to bind peptide HA2A, VP1-HA2A-3 ($P<0.0001$) and VP1-HA2A-4 ($P<0.001$) induced a significant immune response recognising HA2A peptide, compared with wild-type VLP (negative control). Figure 4.5B shows the induced antibody quality of modular HA2A VLPs, demonstrated by the binding against the FL-HA. The result shows that VP1-HA2A-1, -2 and -4 were incapable of eliciting antibodies against FL-HA, indicated by non-significant difference with the negative control. In contrast, VP1-HA2A-3 induced IgGs recognising FL-HA significantly higher than the negative control ($0.01<P<0.05$).

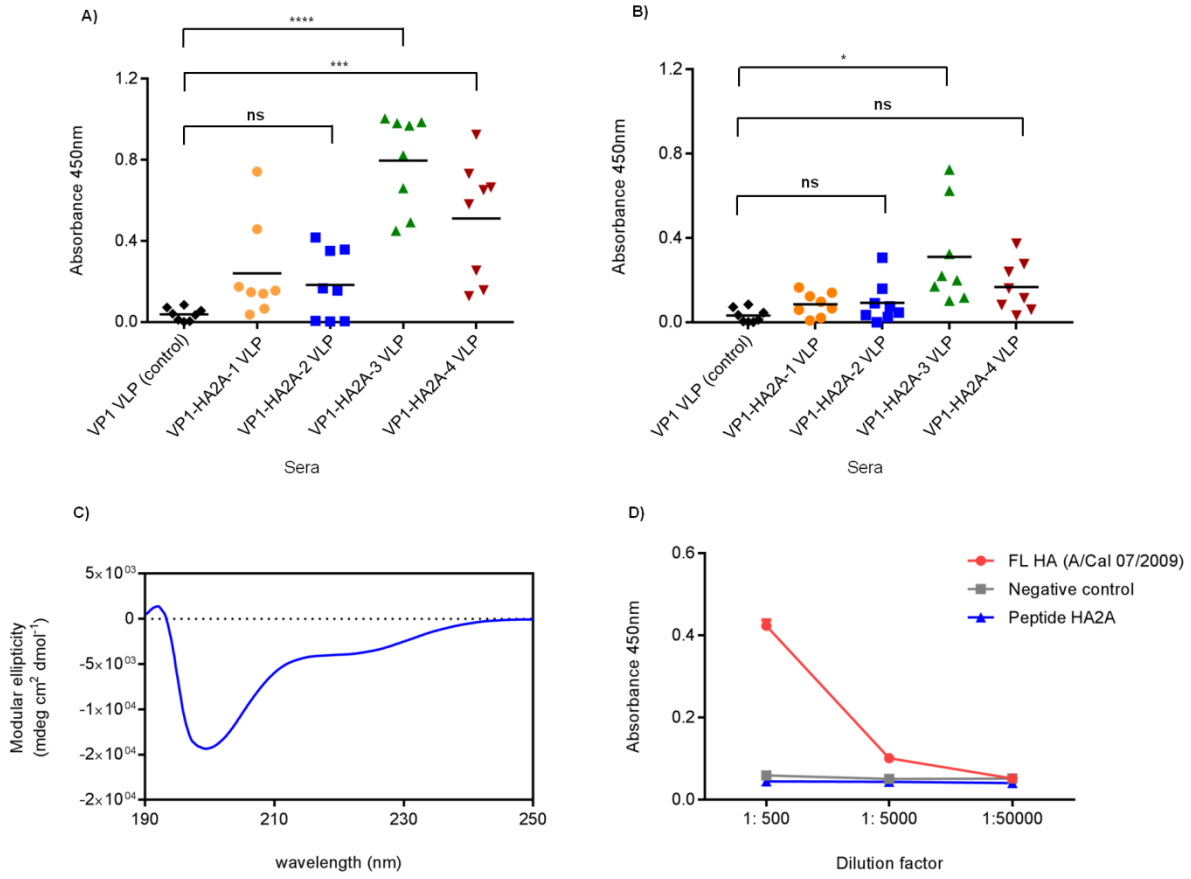


*Figure 4.5: The in vivo immunogenicity testing of modular HA2A VLPs. Indirect ELISA of A) anti-modular HA2A VLP sera against peptide HA2A and B) anti-modular HA2A VLP sera against full-length H1N1 HA, in which bars represent geometric mean of absorbance values from 8 chickens. Statistical analysis was performed using GraphPad Prism Version 6.0. Multiple comparisons of IgG titers and absorbance values were performed with the Tukey's test. Ns: not significant ($P\geq0.05$); *, significant ($0.01<P<0.05$);***, extremely significant ($P<0.001$);****, extremely significant ($P<0.0001$). C) Far-UV CD spectrum of peptide HA2A in the indirect ELISA condition. D) Indirect ELISA of full-length H1N1 HA against CR6261 antibody.*

4.5. Discussion

This study seeks a rapid and simple tool to reduce the number of vaccine candidates for testing in animals and to identify potential vaccine candidates at an early stage of vaccine development. CD spectroscopy alone has been used to characterise antigen presentation on modular MuPyV capsomere in a previous study (Gleiter and Lilie, 2001). The previous study only qualitatively compared the CD spectrum of the modular capsomere and the sum of CD spectra of base capsomeres and antigen proteins, to conclude the structure of the inserted antigen on modular capsomere. In contrast, this study quantitatively analyses the CD spectra of modular and base capsomeres using SVD algorithm to estimate the ratios of secondary structures within capsomeres. The ratios were then converted to amino acid numbers to compare the number of helical amino acids in modular and base capsomeres.

The ability of the CD-SVD method was firstly validated through the structural analysis of MuPyV VP1 capsomere variants. It detected that VP1 capsomere has three more helical amino acids than VP1-ΔC and VP1-ΔNΔC capsomeres (Figure 4.3C), revealing that CD-SVD can identify the secondary structure, particularly helical structure, within the structure of VP1 capsomere. CD-SVD is exploited here to characterise antigenic module structure on modular capsomere. The VP1 monomer structure (1SID.pdb) illustrates that each VP1 monomer comprises 11 helical amino acids at the C-terminus (Figure 4.3A), which is many more helical amino acids than the detection of CD-SVD (Figure 4.3C). It proves that the method undervalues the actual helical amino acid number, suggesting that CD-SVD method can only provide a relative helical amino acid number.

CD-SVD and MD simulation are in good agreement in informing the structure of HA2A antigenic module. CD-SVD (Figure 4.3D) detected that HA2A modules on VP1-ΔNΔC-HA2A-1 formed non-helical structures. Meanwhile, MD simulation (Figure 4.4A) predicted the HA2A modules conformation are different from its native structure, indicated by high RMSD values at above 4Å for 4 out of 5 modules. It is possible that the rigidity of the helical insert, with minimal flexible linkers at either end of the module, forces the helix to unfold for successful insertion into the VP1 surface loop. This result suggests that the presentation of a single HA2A element on each module is an ineffective vaccine design. Increasing element numbers to dual HA2A elements per module, on VP1-ΔNΔC-HA2A-2, led to the increment of 9 helical amino acids, obtained by CD-SVD (Figure 4.3D), and RMSD values of under 4Å for all but one element, obtained by MD simulation (Figure 4.4B). This may be due to the extra length of the inserts, enabling some flexibility of the modules to successfully insert into the VP1 surface loop, while retaining helicity. Similar results were obtained for VP1-ΔNΔC-HA2A-3 and VP1-ΔNΔC-HA2A-4, whereby both CD-SVD and MD simulation

suggest the increase in copy numbers of HA2A element within each module results in improved helical propensity. CD-SVD found the increment in relative helical amino acid numbers on HA2A modules on VP1-ΔNΔC-HA2A-3 and VP1-ΔNΔC-HA2A-4 (Figure 4.3D). MD simulation also reported that the increasing number of elements, and thus the size and flexibility of the insert as a whole, enabled more elements to retain relatively native structure (Figure 4.4C and 4.4D).

The increase in the copy number of HA2A elements, however, might result in increasing the risk of capsomere structure perturbation and hence loss of VLP integrity. The structural prediction (Figure 4.4E) shows the capsomere structure of VP1-ΔNΔC-HA2A-1, VP1-ΔNΔC-HA2A-2 and VP1-ΔNΔC-HA2A-3 maintains low RMSD values (about 4Å). It suggests the modular inserts do not interrupt the capsomere structures significantly, enabling native capsomere structures and likely native VLP assembly. The presentation of 4 HA2A elements per module on VP1-ΔNΔC-HA2A-4 led to an increase in the size of the HA2A module to 11KDa. MD simulation predicted that the increasing size of the module insertions affected the capsomere structure, as seen in a high RMSD value at 8Å (Figure 4.4E). An insertion of 18KDa protein domain on the modular VP1 capsomere has been seen to perturb VP1 capsomere structure (Lua *et al.*, 2015) except in the case of specific linker insertion aimed to counter this effect (Tekewe *et al.*, 2016a; Tekewe *et al.*, 2016b).

This study experimentally identified a variation of induced immunogenicity and antibody quality that might result from the helicity of HA2A antigenic modules. The presentation of a single HA2A element per module results in the non-helical structure of HA2A module, revealed by CD -SVD and MD analysis (Figure 4.3D and 4.4B). Single-element insertion led to a failure to induce antibodies that recognise peptide HA2A and FL-HA (Figure 4.5A and 4.5B). Although CD-SVD indicates that the module presenting dual HA2A elements might form a helical structure (Figure 4.3D), immunological testing showed no antibody response against peptide HA2A and FL-HA (Figure 4.5A and 4.5B). This result could be because the presentation of dual HA2A elements might result in the formation of helical structure but not sufficient antigenicity to evoke an immune response. The increment to triple and quadruple HA2A elements per module resulted in eliciting IgGs recognising peptide HA2A (Figure 4.5A). This study found that the HA2A peptide formed helical structure (Figure 4.5C) without any support obtained by inclusion of a helix stabiliser, e.g. trifluoroethanol, which is required to drive helical propensity of peptides reported in previous studies (Myers *et al.*, 1998; Relf *et al.*, 1996). The binding of anti-HA2A-3 and -4 sera against the helix of peptide HA2A demonstrates the formation of helical structures of antigenic modules on VP1-HA2A-3 and -4. CD-SVD indicates that the modules bearing 4 HA2A elements presented more helical amino acids than on the modules bearing 3 HA2A elements (Figure 4.3D). However,

the serological testing shows that VP1-HA2A-4 induced a lower level of IgGs recognising HA2A peptide compared with VP1-HA2A-3; and, failed to elicit IgGs recognising FL-HA (Figure 4.5A and 4.5B). The insertion of 4 antigenic elements per module, possibly, might perturb VP1 capsomere structure, as suggested by MD simulation. VP1-HA2A-3 successfully induced IgGs recognising FL-HA. The accessibility of HA2A epitope on the FL-HA was also confirmed by the recognition of a conformational antibody, CR6261 antibody (Figure 4.5D). However, VP1-HA2A-3 was not recognised by the CR6261 antibody (Appendix 5, Supplementary material), which may be due to CR6261 requiring some HA1 residues for binding (Ekiert *et al.*, 2009; Schneeman *et al.*, 2012). The result suggests that the helical structure of the HA2A modules on VP1-HA2A-3 may resemble the HA2A native structure on the FL-HA. It reveals that the presentation of triple HA2A antigenic elements is the optimised design for this particular helical structure.

This study found that CD-SVD, assisted with MD simulation, can identify the non-helical structure of HA2A modules on the modular capsomere, resulting in a reduction in the number of vaccine candidates progressed into biological testing. The technique can lead to reduced cost and, more importantly, the number of animals used in accordance with the Three Rs (Replacement, Reduction, Refinement) principle in animal welfare. The findings reveal that CD-SVD can be employed as a rapid screening tool of antigenic module structure, leading to the identification of potential constructs at the early stage of modular VLP based vaccine development.

Supplementary material

The supplementary material provides the α-helix structure ratios and helical amino acid numbers of VP1 capsomere variants and modular HA2A capsomeres in Appendix 3 and 4, respectively.
Appendix 5 shows the indirect ELISA result of VP1-HA2A-3 VLP against CR6261 antibody.

References

Anggraeni, M.R., Connors, N.K., Wu, Y., Chuan, Y.P., Lua, L.H.L., Middelberg, A.P.J., 2013. Sensitivity of immune response quality to influenza helix 190 antigen structure displayed on a modular virus-like particle. Vaccine 31, 4428-4435.

Assad, S., Francis, A., 1999. Over a decade of experience with a yeast recombinant hepatitis B vaccine. Vaccine 18, 57-67.

Chackerian, B., 2007. Virus-like particles: flexible platforms for vaccine development. Expert Rev Vaccines 6, 381-390.

Chuan, Y.P., Lua, L.H.L., Middelberg, A.P.J., 2008. High-level expression of soluble viral structural protein in *Escherichia coli*. J Biotechnol 134, 64-71.

Compton, L.A., Johnson Jr, W.C., 1986. Analysis of protein circular dichroism spectra for secondary structure using a simple matrix multiplication. Anal Biochem 155, 155-167.

Ekiert, D.C., Bhabha, G., Elisliger, M.A., Friesen, R.H.E., Jongeneelen, M., Throsby, M., Goudsmit, J., Wilson, I.A., 2009. Antibody recognition of a highly conserved influenza virus epitope. Science 324, 246-251.

Gleiter, S., Lilie, H., 2001. Coupling of antibodies via protein Z on modified polyoma virus-like particles. Protein Sci 10, 434-444.

Greenfield, N.J., 2007. Using circular dichroism spectra to estimate protein secondary structure. Nat Protocols 1, 2876-2890.

Heinsbroek, E., Ruitenberg, E.J., 2010. The global introduction of inactivated polio vaccine can circumvent the oral polio vaccine paradox. Vaccine 28, 3778-3783.

Henderson, A.D., Fenner, F., 2001. Recent events and observations pertaining to smallpox virus destruction in 2002. Clin Infect Dis 22, 1057-1059.

Jennings, G.T., Bachmann, M.F., 2008. The coming of age of virus-like particles vaccines. Biol Chem 389, 521-536.

Jennings, G.T., Bachmann, M.F., 2009. Immunodrugs: Therapeutics VLP-based vaccines for chronic diseases. Annu Rev Pharmacol Toxicol 49, 303-326.

Johnson, W.C., 1999. Analyzing protein circular dichroism spectra for accurate secondary structures. Proteins: Struct, Funct Bioinf 35, 307-312.

Joshi, H., Cheluvaraja, S., Somogyi, E., Brown, D.R., Ortoleva, P., 2011. A molecular dynamics study of loop fluctuation in human papillomavirus type 16 virus-like particles: a possible indicator of immunogenicity. Vaccine 29, 9423-9430.

Joshi, H., Lewis, K., Singharoy, A., Ortoleva, P.J., 2013. Epitope engineering and molecular metrics of immunogenicity: A computational approach to VLP-based vaccine design. Vaccine 31, 4841-4847.

Kew, O.M., Sutter, R.W., de Gourville, E.M., Dowdle, W.R., Pallansch, M.A., 2005. Vaccine-derived polioviruses and the endgame strategy for global polio eradication. Annu Rev Microbiol 59, 587-635.

Lapelosa, M., Gallicchio, E., Arnold, G.F., Arnold, E., Levy, R.M., 2009. In silico vaccine design based on molecular simulations of rhinovirus chimeras presenting HIV-1 gp41 epitopes. J Mol Biol 16, 675-691.

Lin, T., Porta, C., Lomonossoff, G., Johnson, J.E., 1996. Structure-based design of peptide presentation on a viral surface: the crystal structure of a plant/animal virus chimera at $2.8A^0$ resolution. Fold Des 1, 179-187.

Lipin, D.I., Raj, A., Lua, L.H.L., Middelberg, A.P.J., 2009. Affinity purification of viral protein having heterogeneous quaternary structure: Modeling the impact of soluble aggregates on chromatographic performance. J Chromatogr 1216, 5696-5708.

Liu, W.J., Liu, X.S., Zhao, K.N., Leggatt, G.R., Frazer, I.H., 2000. Papillomavirus virus-like particles for the delivery of multiple cytotoxic T cell epitopes. Virology 273, 374-382.

Lua, L.H.L., Fan, Y., Chang, C., Connors, N.K., Middelberg, A.P.J., 2015. Synthetic biology design to display an 18 kDa rotavirus large antigen on a modular virus-like particle. Vaccine 33, 5937-5944.

McAleer, W.J., Buynak, E.B., Maigetter, R.Z., Wampler, D.E., Miller, W.J., Hilleman, M.R., 1984. Human hepatitis B vaccine from recombinant yeast. Nature 307, 178-180.

Middelberg, A.P.J., Rivera-Hernandez, T., Wibowo, N., Lua, L.H.L., Fan, Y., Magor, G., Chang, C., Chuan, Y.P., Good, M.F., Batzloff, M.R., 2011. A microbial platform for rapid and low-cost virus-like particle and capsomere vaccines. Vaccine 29, 7154-7162.

Myers, J.K., Pace, C.N., Scholtz, J.M., 1998. Trifluoroethanol effects on helix propensity and electrostatic interactions in the helical peptide from ribonuclease T1. Protein Sci 7, 383-388.

Neirynck, S., Deroo, T., Saelens, X., Vanlandschoot, P., Jou, W.M., Fiers, W., 1999. A universal influenza A vaccine based on the extracellular domain of the M2 protein. Nat Med 5, 1157-1163.

Pantophlet, R., Burton, D.R., 2006. GP120: Target for neutralizing HIV-1 antibodies. Annu Rev Immunol 24, 739-769.

Relf, W.A., Cooper, J., Brandt, E.R., Hayman, W.A., Anders, R.F., Pruksakorn, S., Currie, B., Saul, A., Good, M.F., 1996. Mapping a conserved conformational epitope from the M protein of group A *streptococci*. Pept Res 9, 12-20.

Rivera-Hernandez, T., Hartas, J., Wu, Y., Chuan, Y.P., Lua, L.H.L., Good, M., Batzloff, M.R., Middelberg, A.P.J., 2013. Self-adjuvanting modular virus-like particles for mucosal vaccination against group A *streptococcus* (GAS). Vaccine 31, 1950-1955.

Roldao, A., Mellado, M.C.M., Castilho, L.R., Carrondo, M.J.T., Alves, P.M., 2010. Virus-like particles in vaccine development. Expert Rev Vaccines 9, 1149-1176.

Schiller, J.T., Castellsagué, X., Villa, L.L., Hildesheim, A., 2008. An update of prophylactic human papillomavirus L1 virus-like particle vaccine clinical trial results. Vaccine 26, Supplement 10, K53-K61.

Schneeman, A., Speir, J.A., Tan, G.S., Khayat, R., Ekiert, D.C., Matsuoka, Y., Wilson, I.A., 2012. A virus-like particle that elicits cross-reactive antibodies to the conserved stem of influenza virus hemagglutinin. J Virol 86, 11686-11697.

Sela-Culang, I., Kunik, V., Ofran, Y., 2013. The structural basis of antibody-antigen recognition. Front Immunol 4, 302.

Sreerama, N., Woody, R.W., 2000. Estimation of protein secondary structure from circular dichroism spectra: Comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. Anal Biochem 287, 252-260.

Sui, J., Hwang, W.C., Perez, S., Wei, G., Aird, D., Chen, L., Santelli, E., Stec, B., Cadwell, G., Ali, M., Wan, H., Murakami, A., Yammanuru, A., Han, T., Cox, N.J., Bankston, L.A., Donis, R.O., Liddington, R.C., Marasco, W.A., 2009. Structural and functional bases for broad-spectrum neutralization of avian and human influenza A viruses. Nat Struct Mol Biol 16, 265-272.

Tegerstedt, K., Lindencrona, J.A., Curcio, C., Andreasson, K., Tullus, C., Forni, G., Dalianis, T., Kiessling, R., Ramqvist, T., 2005. A single vaccination with polyomavirus VP1/VP2Her2 virus-like particles prevents outgrowth of HER-2/neu-expressing tumors. Cancer Res 65, 5953-5957.

Tekewe, A., Connors, N.K., Middelberg, A.P.J., Lua, L.H.L., 2016a. Design strategies to address the effect of hydrophobic epitope on stability and in vitro assembly of modular virus-like particle. Protein Sci 25, 1507-1516.

Tekewe, A., Fan, Y., Tan, E., Middelberg, A.P.J., Lua, L.H.L., 2016b. Integrated molecular and bioprocess engineering for bacterially produced immunogenic modular virus-like particle vaccine displaying 18 kDa rotavirus antigen. Biotechnol Bioeng 114, 397-406.

Thomas, S., Luxon, B.A., 2013. Vaccines based on structure-based design provide protection against infectious diseases. Expert Rev Vaccines 12, 1301-1311.

Throsby, M., van den Brink, E., Jongeneelen, M., Poon, L.L.M., Alard, P., Cornelissen, L., Bakker, A., Cox, F., van Deventer, E., Guan, Y., Cinalt, J., ter Meulen, J., Lasters, I., Carsetti, R., Peiris, M., de Kruif, J., Goudsmit, J., 2008. Heterosubtypic neutralizing monoclonal antibodies cross-

protective against H5N1 and H1N1 recovered from human IgM^+ memory B cells. PLoS One 3, e3942.

Tissot, A.C., Renhofa, R., Schmitz, N., Cielens, I., Meijerink, E., Ose, V., Jennings, G.T., Saudan, P., Pumpens, P., Bachmann, M.F., 2010. Versatile virus-like particle carrier for epitope based vaccines. PLoS One 5, e9809.

Wallace, B.A., Janes, R.W., 2009. Modern techniques for circular dichroism and synchrotron radiation circular dichroism spectroscopy. IOS Press, Amsterdam, NLD.

Waneesorn, J., Wibowo, N., Bingham, J., Middelberg, A.P.J., Lua, L.H.L., 2016. Structural-based designed modular capsomere comprising HA1 for low-cost poultry influenza vaccination. Vaccine. DOI: 10.1016/j.vaccine.2016.11.058.

Whitmore, L., Wallace, B.A., 2008. Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases. Biopolymers 89, 392-400.

Wibowo, N., Chuan, Y.P., Lua, L.H.L., Middelberg, A.P.J., 2012. Modular engineering of a microbially-produced viral capsomere vaccine for influenza. Chem Eng Sci 103, 12-20.

Wibowo, N., Hughes, F.K., Fairmaid, E.J., Lua, L.H.L., Brown, L.E., Middelberg, A.P.J., 2014. Protective efficacy of a bacterially produced modular capsomere presenting M2e from influenza: Extending the potential of broadly cross-protecting epitopes. Vaccine 32, 3651-3655.

Yang, Y.-W., Teng, C.-C., 1998. Stability of polyomavirus major capsid protein VP1 under denaturants guanidine hydrochloride and urea. Int J Biol Macromol 22, 81-90.

Yang, Y.-W., Teng, C.-C., 1999. Conformational changes of polyomavirus major capsid protein VP1 in sodium dodecyl sulfate solution. J Pept Res 53, 75-81.

Yin, Y., Li, H., Wu, S., Dong, D., Zhang, J., Fu, L., Xu, J., Chen, W., 2011. Hepatitis B virus core particles displaying *Mycobacterium tuberculosis* antigen ESAT-6-specific immune responses. Vaccine 29, 5645-5651.

Chapter 5

Determining the atomic structures of presented antigenic elements

5.1. Introduction

The use of modular virus-like particles (VLP) is a promising approach in vaccine development. The presentation of foreign antigenic peptides or antigenic elements on VLP in their correct conformation is essential to induce neutralising antibodies. Different display strategies were used to promote the conformational presentation of antigenic elements on VLP. An antigenic element, helix 190 (H190) from influenza hemagglutinin (HA), was arrayed in tandem repeat or flanked by GCN4 structure-promoting elements within the antigenic module. The use of the tandem repeat strategy resulted in a high level of IgGs recognising HA1, in comparison with the use of flanking sequence (Anggraeni *et al.*, 2013). The findings in Chapter 4 show that tandem repeat display of one, two, three and four copies of the helix A (HA2A) from influenza HA2 led to increased levels of IgGs recognising peptide HA2A and full-length HA. The tandem repeat display strategy possibly promoted the propensity of the antigenic modules to adopt a helical conformation. Obtaining the structure of the antigenic modules will confirm the effect of the tandem repeat display strategy in promoting helical structure in antigenic modules. The CD-SVD method (in Chapter 4) can identify the probability of helical structure formation of antigenic modules; however, the method is unable to determine the actual structure of antigenic modules. A method that can determine the structure of presented antigenic modules or modular capsomeres will confirm if the desired secondary structure is obtained from the tandem repeat display strategy.

X-ray crystallography is employed to obtain the structure of modular capsomere, for the determination of the conformation of the presented epitope. X-ray crystallography is the most commonly utilised technique for solving the three-dimensional structure of proteins. Approximately 88% of the protein structures in the Protein Data Bank were determined using X-ray crystallography (Li and Chang, 2009). The determination of recombinant protein structures using X-ray crystallography requires some fundamental steps including protein preparation (cloning, expression, and purification), crystallisation, diffraction data collection, electron density analysis and model building. Advances in protein crystallisation such as the availability of commercial screening kits and the automation of high-throughput screening for crystal optimisation have facilitated the determination of protein structures using X-ray crystallography (Blundell *et al.*, 2002; Chayen and

Saridakis, 2002). X-ray crystallography has been used to determine the structure of viral capsid protein (Dong *et al.*, 1998; Wynne *et al.*, 1999). Stehle and Harrison utilised X-ray crystallography to resolve the structure of wild-type (wt) MuPyV VP1 capsomere (Stehle and Harrison, 1996, 1997).

The presence of the C- and N- termini was reported to prevent crystal formation of wt MuPyV VP1 protein in a previous study (Stehle and Harrison, 1997). Thus, 31 residues at the N-terminus and 63 residues at the C-terminus of the modular VP1 constructs were removed. Seven modular constructs were selected for X-ray crystallography to address the effect of (1) flanking sequence and (2) copy number of an antigenic element, on the helical module structure (Figure 5.1). First, to address the effect of flanking sequence, two constructs (i) VP1-ΔNΔC-H190-GCN4 and (ii) VP1-ΔNΔC-H190-2 were selected to compare the conformation of modularised H190 in two different designs. Second, the effect of antigenic element copy number was examined using two helical epitopes, H190 and HA2A. Constructs (ii) VP1-ΔNΔC-H190-2 and (iii) VP1-ΔNΔC-H190-3 were used to compare the helical propensity of H190 elements when the copy number of the H190 element was increased. Constructs (iv) VP1-ΔNΔC-HA-1, (v) VP1-ΔNΔC-HA-2, (vi) VP1-ΔNΔC-HA-3 and (vii) VP1-ΔNΔC-HA-4 were used to examine the effect of an increase in HA2A element copy number on the helical propensity of the HA2A element.

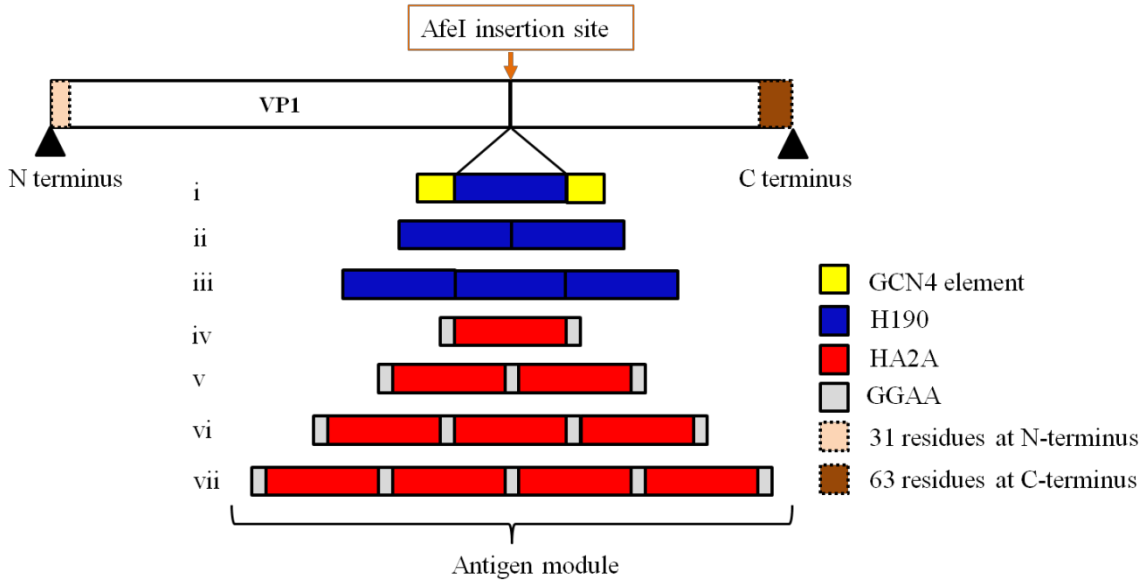


Figure 5.1: Schematic diagram of the antigenic modules in seven modular constructs selected for crystallisation. (i) VP1-ΔNΔC-H190-GCN4, in which one copy of H190 was flanked with GCN4 elements. (ii) VP1-ΔNΔC-H190-2 and (iii) VP1-ΔNΔC-H190-3, in which two and three copies of H190 were arrayed as tandem repeats, respectively. (iv) VP1-ΔNΔC-HA-1, (v) VP1-ΔNΔC-HA-2, (vi) VP1-ΔNΔC-HA-3 and (vii) VP1-ΔNΔC-HA-4, in which one, two, three and four copies of HA2A were arrayed as tandem repeats, respectively.

5.2. Materials and Methods

5.2.1. Molecular cloning

The plasmid pGEX-VP1, consisting of pGEX-4T-1 (Novagen, Billerica, MA, USA) with the murine polyomavirus VP1 sequence (M34958) inserted between the *Bam*HI and *Xho*I recognition sites, was gifted by Professor Robert Garcea (University of Colorado, Colorado, USA). *Nae*I and *Afe*I recognition sites were inserted into pGEX-VP1 at positions 86 and 293 of VP1, yielding the plasmid pGEX-VP1-S1S4. The gene inserts VP1-ΔN31-ΔC63-S1S4 was synthesised using designed primers to exclude 31 amino acids at the N-terminus and 63 amino acids at the C-terminus of VP1-S1S4. The inserts were cloned between the *Bam*HI and *Xho*I recognition sites in the vector pGEX, yielding VP1-ΔN31-ΔC63-S1S4 construct. The performed cloning method was *in vivo* homologous recombination (Bubeck *et al.*, 1993; Parrish *et al.*, 2004) using the One Shot®OmniMAX™2T1^{R} chemically competent *E. coli* cloning strain (Invitrogen™, Victoria, Australia). Table 5.1 lists seven modular constructs produced for X-ray crystallography.

DNA sequences encoding H190 (STSADQQSLYQNADAY) were codon optimised and generated using gene assembly from oligos calculated using DNAWorks (http://helixweb.nih.gov/dnaworks/). Two expression vectors: pGEX-VP1-ΔN31-ΔC63-S4-GSGS-GCN4-H190 and pGEX-VP1-ΔN31-ΔC63-S1S4-H190-2x were generated by Melissa R. Anggraeni. The gene inserts VP1-ΔN31-ΔC63-S4-GSGS-GCN4-H190, VP1-ΔN31-ΔC63-S1S4-H190-2x, and VP1-ΔN31-ΔC63-S1S4-H190-3x were synthesised using designed primers to exclude 31 amino acids at the N-terminus and 63 amino acids at the C-terminus of VP1-S4-GSGS-GCN4-H190, VP1-S1S4-H190-2x, and VP1-S1S4-H190-3x, respectively. The inserts were cloned between the *Bam*HI and *Xho*I recognition sites in the vector pGEX, yielding modular H190 constructs.

DNA sequences encoding HA2A (DLKSTQNAIDEITNKVNSVIEK) from A/California/07/2009 H1N1 were codon optimised and generated using gene assembly from oligos calculated using DNAWorks (http://helixweb.nih.gov/dnaworks/). The HA2A sequence was flanked by GGAA and separated with the same sequence in tandem constructs (Figure 5.1). The gene inserts VP1-ΔN31-ΔC63-S1S4-HA2A, VP1-ΔN31-ΔC63-S1S4-HA2A-2x, VP1-ΔN31-ΔC63-S1S4-HA2A-3x and VP1-ΔN31-ΔC63-S1S4-HA2A-4x were cloned into the *Afe*I recognition sites in the vector pGEX-VP1-ΔN31-ΔC63-S1S4, yielding modular HA2A constructs.

Table 5.1: Vector names and designated protein names.

Expression vectors	**Antigen module**	**Designated name**
pGEX-VP1-ΔN31-ΔC63-S4-GSGS-GCN4-H190	H190 x 1	H190-GCN4
pGEX-VP1-ΔN31-ΔC63-S1S4-H190-2x	H190 x 2	H190-2
pGEX-VP1-ΔN31-ΔC63-S1S4-H190-3x	H190 x 3	H190-3
pGEX-VP1-ΔN31-ΔC63-S1S4-HA2A	HA2A x 1	HA2A-1
pGEX-VP1-ΔN31-ΔC63-S1S4-HA2A-2x	HA2A x 2	HA2A-2
pGEX-VP1-ΔN31-ΔC63-S1S4-HA2A-3x	HA2A x 3	HA2A-3
pGEX-VP1-ΔN31-ΔC63-S1S4-HA2A-4x	HA2A x 4	HA2A-4

5.2.2. Expression

GST-tagged modular VP1 capsomeres were expressed intracellularly in *E. coli* Rosetta(DE3)pLysS cells (Novagen, Virginia, USA) as described in (Chuan *et al.*, 2008; Lipin *et al.*, 2009). A single colony was inoculated into 10 mL TB media (1% tryptone (w/v), 0.5% yeast extract (w/v), 1% NaCl (w/v)) and incubated at 30°C, 180 rpm for 16 hours. The starting culture was inoculated into 800 mL TB media to obtain a starting OD_{600} of 0.001. The cultures were cooled before inducing with 0.2 mM isopropyl-β-D-thiogalactoside (IPTG) and incubating overnight.

5.2.3. Purification

5.2.3.1. First glutathione-S-transferase (GST) purification

After cell lysis, GST-tagged modular VP1 capsomeres were captured with a 5 mL GSTrap HP column on an AKTAxpress^{TM} (GE Healthcare Bioscience, UK) in L buffer (containing 40 mM Tris-base, 200 mM NaCl, 1 mM EDTA, 5% glycerol, 5 mM DTT, pH 8.0) and eluted in L buffer containing 10 mM Glutathione (GSH).

5.2.3.2. Size exclusion chromatography (SEC)

GST-tagged modular VP1 capsomeres were incubated with tobacco etch virus protease (TEVp) at the ratio of 1 TEVp: 5 proteins (w/w) for 4 hours at room temperature to cleave GST tag and yield modular VP1 capsomeres. The sample was loaded onto a Superdex 200 30/100 GL column (GE Healthcare Bioscience, UK) to separate modular capsomeres from the aggregate and GST tag in X-ray buffer containing 20 mM Tris-base, 100 mM NaCl, pH 8.0.

5.2.3.3. Second glutathione-S-transferase (GST) purification

After SEC, some trace amounts of GST tag and undigested materials remained in the samples. The samples were passed through a 1ml GSTrap^{TM} HP column (GE Healthcare Bioscience, UK) at the flow rate of 0.2 ml min^{-1} in X-ray buffer containing 20 mM Tris-base, 100mM NaCl, pH 8.0. The contaminants were captured by the column, and the flow-through was collected.

5.2.3.4. Hydrophobic interaction chromatography (HIC)

Co-purified chaperone 70 (Hsp70) contaminants were removed by HIC using a 1 ml HiTrap Phenyl HP column (GE Healthcare Bioscience, UK). Buffer A contained 20 mM Tris-base, 100 mM NaCl, 0.4 M $(NH_4)_2SO_4$ or 10% (v/v) $(NH_4)_2SO_4$, pH 8.0. Buffer B contained 20 mM Tris-base, 100 mM NaCl, pH 8.0. Saturated $(NH_4)_2SO_4$ was added to the samples to a final concentration of 10% (v/v). The column was equilibrated with buffer A until UV absorbance and conductivity were stable. The samples were passed onto the column with a flow rate of 1 ml min^{-1}, then unbound protein was washed with buffer A until the UV baseline was stable. In elution, the concentration of $(NH_4)_2SO_4$ was decreased using step elution with 10%, 30%, 50%, 75% and 100% of buffer B and prolonged 6–7 column volume (CV) for each elution step.

5.2.4. Protein concentration

The purified modular VP1 capsomeres were concentrated using Amicon Ultra-4 centrifugal concentrators (molecular weight cut off, 10 kDa).

5.2.5. Protein concentration measurement

The UV absorbance of proteins at 280 nm was measured using UV spectroscopy and converted to protein concentration using the Beer-Lambert Law (Aitken and Learmonth, 2009):

$$A = \varepsilon \times b \times c$$

where A is absorbance value measured at 280 nm, ε is the extinction coefficient of protein ($M^{-1}cm^{-1}$), b is the path length and c is the protein concentration (mg mL^{-1}). The theoretical molecular weights (MW) and extinction coefficients of proteins were obtained from the ProtParam tool on the Expasy website and are shown in Table 5.2.

Table 5.2: Theoretical molecular weights and extinction coefficient of the wt-VP1 and modular VP1 constructs

Protein name	**MW of proteins (Da)**	**Extinction coefficient (M^{-1} cm^{-1})**
Wild-type VP1 (Wt-VP1)	42763.6	58057
H190-GCN4	36211.7	57995
H190-2	35586.8	60975
H190-3	37342.6	63955
HA2A-1	35065.6	55015
HA2A-2	37763.6	55015
HA2A-3	40461.6	55015
HA2A-4	43159.5	55015

5.2.6. Crystallisation screening

The crystallisation screening was carried out by Dr. Santosh Panjikar at the Australian Synchrotron (Melbourne, Australia). The hanging drop and sitting drop techniques were used to screen crystallisation conditions for the modular VP1 capsomeres (Figure 5.2). Crystallisation screening was performed by varying conditions including precipitant concentration, pH, and temperature. The screening was performed with a drop ratio of 1:1 (1µl protein: 1µl precipitant) and 0.5 ml precipitant per well in 24-well plates (VDX plate, Hampton Research, California, the US).

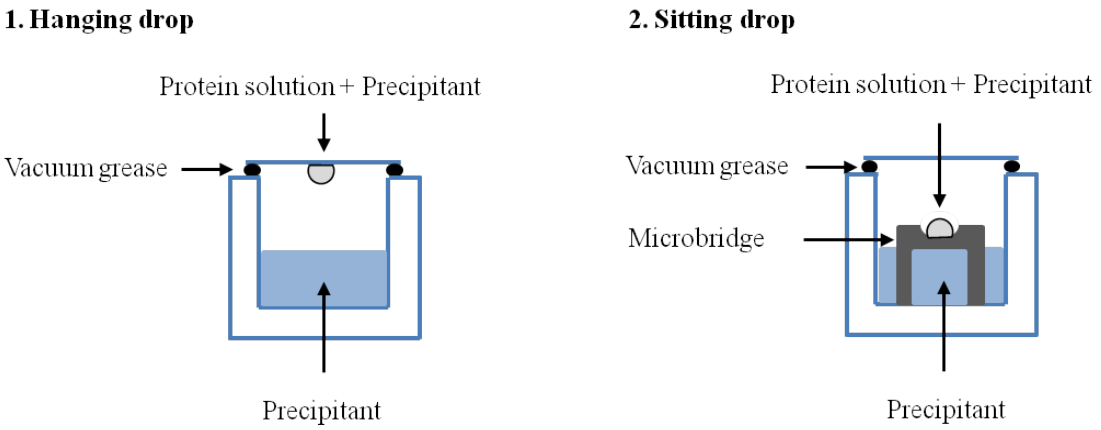


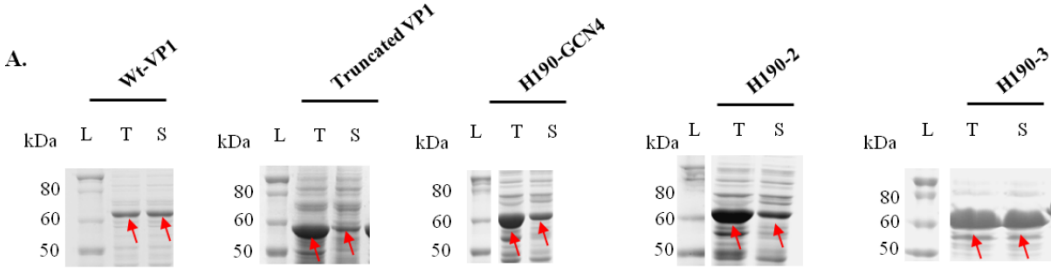
Figure 5.2: Schematic diagram of crystallisation screening techniques: hanging drop and sitting drop.

5.3. Results and Discussion

5.3.1. Expression of modular VP1 constructs

The modular VP1 constructs were expressed as GST-tagged proteins at 26°C as described in previous work (Chuan *et al.*, 2008). After cell disruption using sonication, total cell lysates and soluble fractions were taken for SDS-PAGE to analyse the solubility of the modular VP1 constructs. Figure 5.3 compares the solubility of wt-VP1 (1–384 residues), truncated VP1 (32–321 residues), and the H190 constructs. The solubility of truncated VP1 and H190 constructs was lower than that of wt-VP1, indicating that the truncation at the C- and N- termini of VP1 led to the decreased solubility. The N-terminal truncation might affect the ribosome-binding site which locates on the N-terminus of VP1 (Hartmann and Armengaud, 2014). It could interfere the ribosome binding to mRNA in translation, leading to reduce protein translation or protein yield. The N-terminus of protein was also reported to change the folding mechanism of protein (Korepanova *et al.*, 2001). The deletion at the N-termini, therefore, might perturb the folding of VP1, resulting in the decreased solubility. The H190 constructs have higher protein solubility than truncated VP1, suggesting that the insertion of H190 modules might result in improved solubility of the H190 constructs.

A.



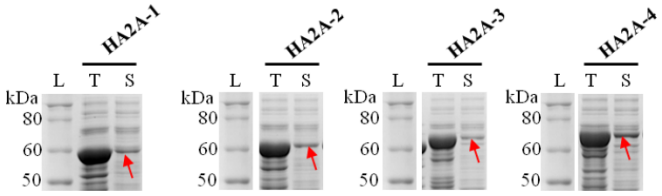
B.

Constructs	Final OD	Solubility (%)
Wt-VP1	6.1	95.9
Truncated VP1	4.1	24.5
H190-GCN4	4.6	43.9
H190-2	4.2	38.5
H190-3	4.8	76.3

Figure 5.3: A) SDS-PAGE analysis showing the solubility of H190 constructs expressed at 26℃ and induced with 0.2mM IPTG. Lanes: (L) molecular weight marker; (T) Total cell lysate; (S) Soluble fraction of cell lysate. B) Final OD and the solubility of wt-VP1, truncated VP1 and H190 constructs. The protein solubility was calculated by measuring the relative intensity of the relevant bands in the total fraction in comparison to the soluble fraction using the SDS-PAGE gel image. The analysis was performed using Image Lab software (Bio-rad).

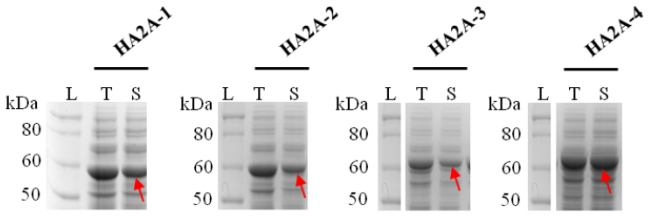
Figure 5.4A shows the solubility of the HA2A constructs. The solubility of the HA2A constructs was lower than that of truncated VP1 (Figure 5.3), suggesting that the insertion of HA2A module led to a decrease in the solubility of the constructs. The insertion of another antigenic element into VP1 was reported to cause reduced solubility of modular VP1 constructs in a previous study (Wibowo *et al.*, 2012). A strategy was therefore applied to improve the solubility of these constructs by decreasing the expression temperature and IPTG concentration, which might influence the solubility of the target protein in *E. coli* expression, as reported in other studies (Baneyx and Mujacic, 2004; Francis and Page, 2010; Tolia and Joshua-Tor, 2006). Figure 5.4B shows that expression at 15°C and induction with 0.1mM IPTG resulted in the improved solubility of all HA2A constructs from below 22% to above 49%. *E. coli* has a high growth rate, which can cause incorrect folding of proteins and aggregation. A decrease in expression temperature and IPTG concentration can reduce the growth rate of *E. coli*, in turn slowing the transcription and translation rate, which can result in improved protein folding and better protein solubility.

A. Expression condition: 26°C, 0.2 mM IPTG



Constructs	Final OD	Solubility (%)
HA2A-1	4.7	16.0
HA2A-2	4.6	15.5
HA2A-3	4.5	12.2
HA2A-4	4.7	21.8

B. Expression condition: 15°C, 0.1 mM IPTG



Constructs	Final OD	Solubility (%)
HA2A-1	1.7	64.9
HA2A-2	1.7	50.5
HA2A-3	1.3	49.3
HA2A-4	1.8	80.8

Figure 5.4: SDS-PAGE analysis of HA2A constructs showing the solubility of the proteins expressed under two different conditions: A) at 26°C and 0.2mM IPTG and B) at 15°C and 0.1mM IPTG. Lanes (L) Molecular weight marker; (T) Total cell lysate; (S) Soluble fraction of cell lysate. The tables show the final OD and the solubility of the HA2A constructs. The protein solubility was calculated by measuring the relative intensity of the relevant bands in the total fraction in comparison to the soluble fraction using the SDS-PAGE gel image. The analysis was performed using Image Lab software (Bio-rad).

5.3.2. Purification of modular VP1 constructs

After cleaving off GST tag in fusion protein using TEVp digestion, size exclusion chromatography (SEC) was used to separate the modular capsomeres from soluble aggregates and GST tag. For all constructs, the SEC profiles show three dominant peaks representing soluble aggregates, modular capsomeres and GST. The modular capsomeres were eluted later than soluble aggregates and earlier than GST. Figure 5.5 shows the comparison of the elution time of wt-VP1, truncated VP1 and HA2A-1 as a representative for the modular H190 and HA2A constructs. Both the truncated VP1 and the HA2A-1 capsomeres eluted later than wt-VP1 capsomere, as reported in a previous study (Wibowo *et al.*, 2012). Meanwhile, HA2A-1 capsomere eluted about the same time with the truncated VP1 capsomere. It suggests that the truncated VP1 and HA2A-1 capsomeres have smaller hydrodynamic radii compared with wt-VP1 capsomere. The removal of 31 residues at N- and 63 residues at C- termini of VP1 might cause the decrease in hydrodynamic radius of HA2A-1 capsomeres and the other modular H190 and HA2A capsomeres, compared with wt-VP1.

The SEC results for the HA2A constructs (Figure 5.6) confirms the solubility test result, showing that the expression at 15°C with 0.1mM IPTG resulted in increased protein solubility and decreased aggregation. Figure 5.6 also compares the SEC profiles of HA2A constructs expressed at two different conditions. A relative improvement in HA2A capsomere yields was observed in the expression at 15°C with 0.1mM IPTG. The results show that the yield of HA2A capsomeres decreased with increasing of antigen copy number, as observed in previous work within the group (Rivera-Hernandez, 2012).

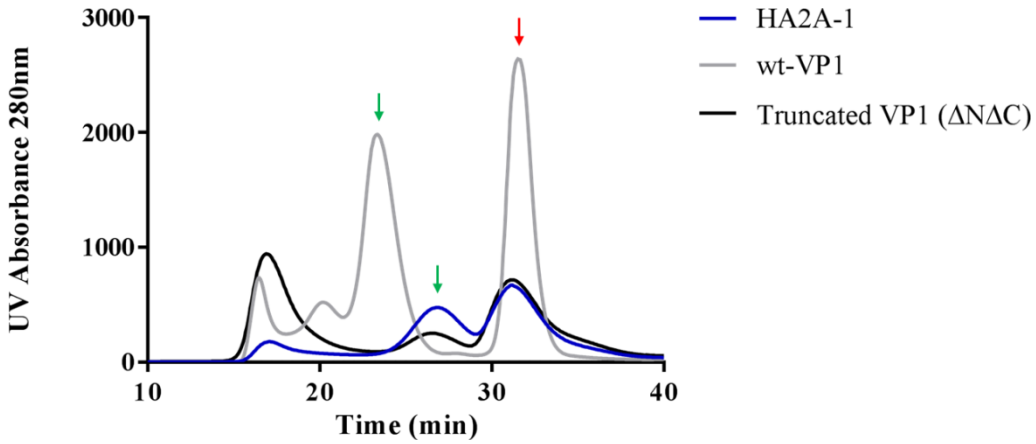


Figure 5.5: SEC chromatograms of wt-VP1 (1–384 residues), truncated VP1 (32–321 residues) and a representative of modular H190 and HA2A construct (HA2A-1). The SEC was run using Superdex™ 200 10/300 GL (GE Healthcare Bioscience, UK). Green arrow: capsomere, red arrow: GST peak.

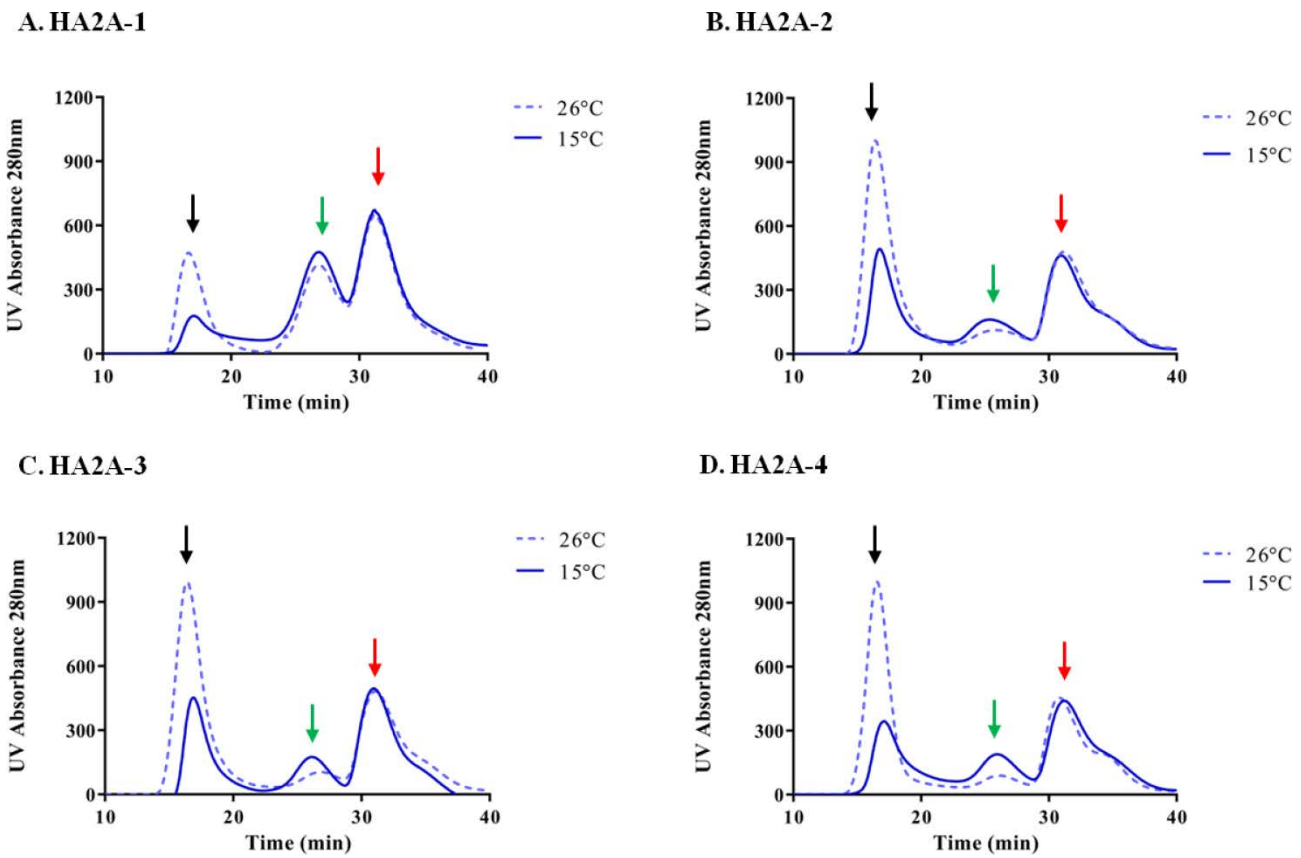


Figure 5.6: Size exclusion chromatograms of HA2A constructs expressed under two different conditions: 26°C and 0.2mM IPTG (blue broken line) and 15°C and 0.1mM IPTG (solid blue line). The modular capsomeres were separated using Superdex™ 200 10/300 GL (GE Healthcare Bioscience, UK). Black arrow: aggregate, green arrow: HA2A capsomere, red arrow: GST peak.

After using SEC to isolate the untagged modular capsomeres, the purity of the capsomeres was analysed using SDS-PAGE. Figure 5.7 shows the SDS-PAGE analysis of the H190 and HA2A constructs, comparing the total protein of TEVp digested GST-tagged proteins or SEC feed (lane 1) and the pool of capsomere fractions (lane 2). Lane 1 shows that the digestion with TEVp was effective to remove GST from untagged modular capsomeres, with only a negligible amount of undigested material remaining after digestion (orange arrows). The late elution of the modular capsomeres resulted in some GST (red arrow) co-purifying with the modular capsomeres (lane 2). After SEC, the modular VP1 capsomeres (lane 2) contained three contaminants: (i) undigested GST-tagged modular VP1, (ii) the GST tag, and (iii) a 70-kDa contaminant.

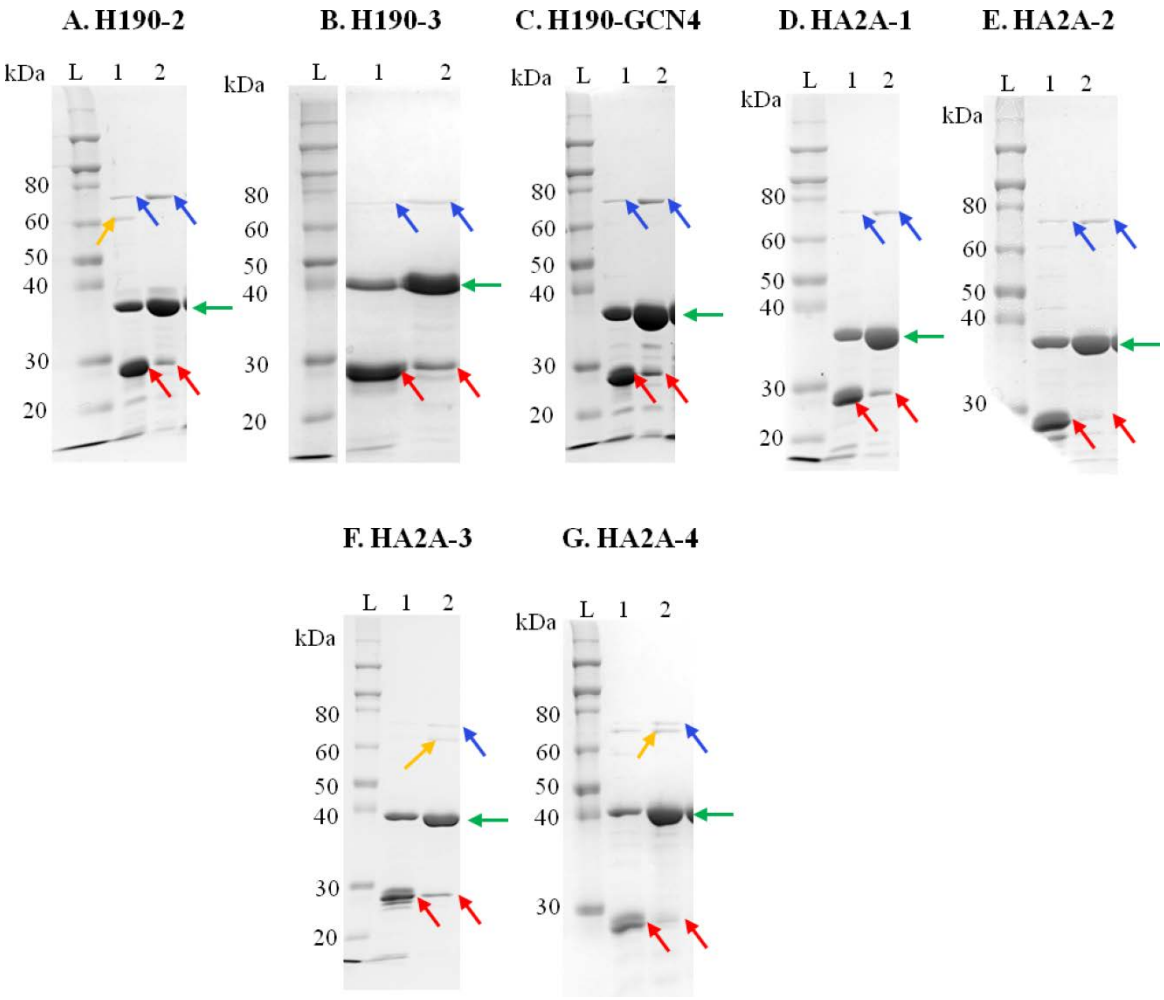


Figure 5.7: SDS-PAGE analysis showing the purity of the modular VP1 constructs after size exclusion chromatography (SEC): A) H190-2, B) H190-3, C) H190-GCN4, D) HA2A-1, E) HA2A-2, F) HA2A-3 and G) HA2A-4. Lanes (L) molecular weight marker; (1) SEC feed or total protein of TEVp digested GST-tagged proteins; (2) the pool of capsomere fractions. Modular VP1: green arrows, undigested GST-tagged modular VP1: orange arrows, GST tag: red arrows and 70-kDa contaminant: blue arrows.

5.3.3. Improving the purity of the modular VP1 capsomeres

A purification strategy involving affinity and hydrophobic interaction was developed to obtain high purity modular capsomeres. First, GST affinity chromatography was used to remove the remaining undigested GST-tagged modular VP1 (orange arrows) and the GST tag (red arrows) from untagged proteins. To maximise the binding capacity of the GST column, the flow rate for sample loading was reduced to 0.2 ml min^{-1}. The binding of GST to glutathione occurs with slow binding kinetics (Habig *et al.*, 1974). A decrease in flow rate enables longer contact time between GST and immobilised glutathione, leading to improving GST binding. Second, hydrophobic interaction chromatography (HIC) was used to eliminate DnaK (a major bacterial Hsp 70), corresponding to the 70-kDa band, from the modular H190 and HA2A capsomeres. DnaK (a major bacterial Hsp 70) was identified as being co-purified with wt-VP1 expressed in *E. coli* in a previous study (Fan and Middelberg, 2011). DnaK acts as a chaperone to prevent aggregation and misfolding of proteins produced in *E. coli*, binding to hydrophobic regions on unfolded polypeptides (Rüdiger *et al.*, 1997). The DnaK contaminant can be eliminated using hydrophobic interaction chromatography (HIC), reported in (Fan and Middelberg, 2011).

Figure 5.8 compares the purity of the capsomeres in purification steps. The GST-tagged proteins (orange arrows) and the GST tag (red arrows) were effectively removed after the second GST purification (lane 3−the pool of GST fractions), in comparison with the total protein of TEVp digested GST-tagged proteins or SEC feed (lane 1) and the pool of capsomere fractions or GST feed (lane 2). Figure 5.8 shows that a DnaK band on the gel was always detected in the presence of the modular VP1 band. It suggests that the interaction between modular capsomeres and DnaK via hydrophobic regions may be very strong and cannot be separated by HIC. There might be two pools of modular VP1, some DnaK-associated and some DnaK-free modular VP1. During the protein expression, only some of the modular VP1 capsomeres may require the support of DnaK in protein folding. HIC separated free DnaK-modular VP1 capsomeres from DnaK-associated modular VP1 capsomeres. The HIC step thus sacrificed some DnaK-associated modular VP1 capsomeres to improve the purity. Figure 5.8 (lane 4-13) indicated that the HIC step enhanced the purity of modular VP1 capsomeres. Figure 5.9 shows the typical HIC chromagraphy of modular VP1 capsomeres and HIC fraction collection. For constructs H190-2, H190-3, GCN4-H190, HA-1 and HA-3, there were no DnaK bands detected in the later fractions of HIC, indicating that HIC completely removed DnaK-associated modular VP1 capsomeres. For constructs HA2A-2 and HA2A-4, there a negligible amount of DnaK remained in HIC fractions in comparison with the amount of DnaK in the pool of GST fractions or HIC feed (lane 3). After HIC, the purity of

modular H190 and HA2A capsomeres was more than 90%, thus meeting the requirement of crystallisation screening.

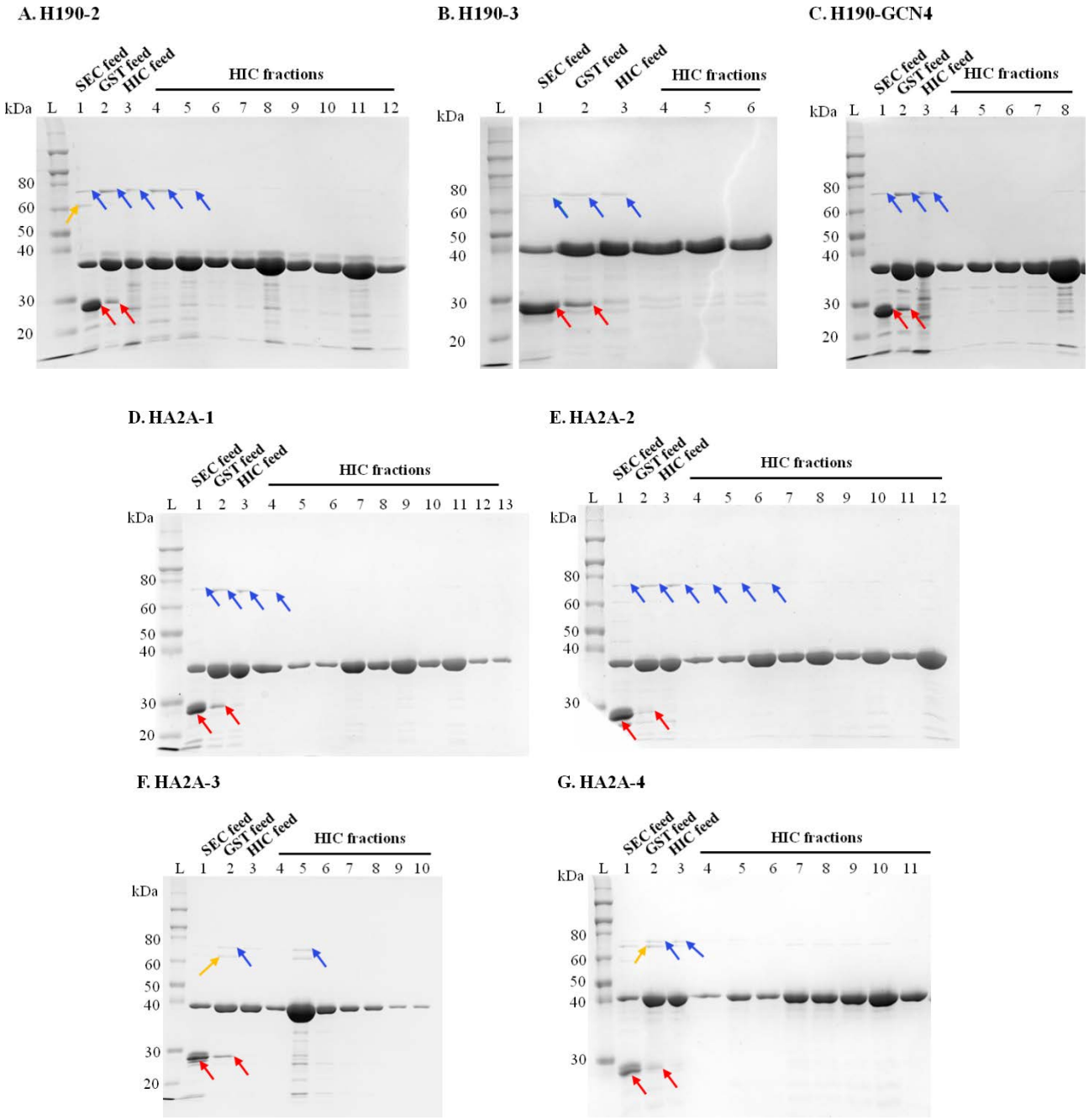


Figure 5.8: SDS-PAGE analysis showing the purity of the modular VP1 constructs: (A) H190-2, (B) H190-3, (C) H190-GCN4, (D) HA-1, (E) HA-2, (F) HA-3 and (G) HA-4. Lanes: (1) SEC feed or total protein of TEVp digested GST-tagged proteins; (2): GST feed or the pool of capsomere fractions; (3): HIC feed or the pool of GST fractions; (4-13): HIC fractions. The most dominant bands, between 30–40 kDa, corresponding to modular VP1. Contaminants are indicated by arrows in different colours, undigested GST-tagged modular VP1 in orange arrows, GST tag in red arrows, DnaK in blue arrows.

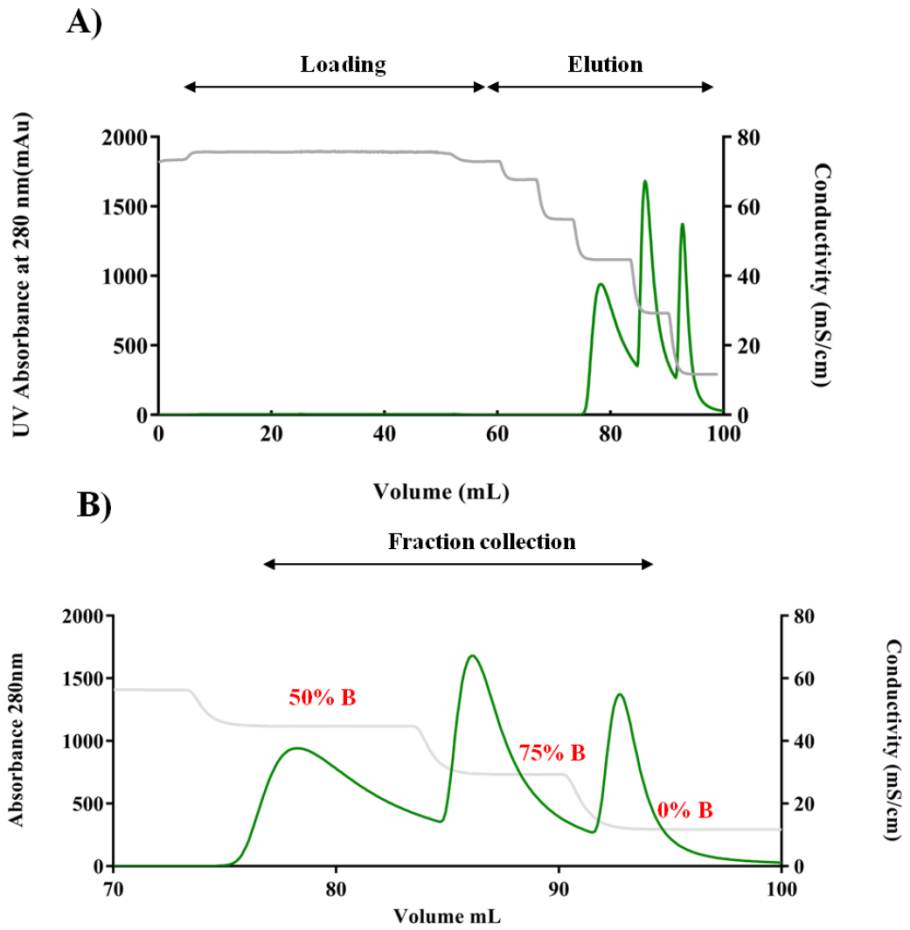


Figure 5.9: Typical HIC chromatogram of modular VP1.

5.3.4. Size exclusion chromatography with multi-angle light scattering (SEC-MALS) analysis

The HIC-purified H190 and HA2A capsomeres were concentrated to a concentration above 8 mg ml^{-1} (Table 5.3) for the crystallisation screening. After concentration with Amicon Ultra-4 centrifugal concentrators (molecular weight cut off, 10KDa), SEC-MALS was used to analyse the quaternary structure and stability of the H190 and HA2A capsomeres (Figure 5.10). For construct H190-2, the SEC-MALS analysis detected a single peak corresponding to modular VP1 capsomeres. For constructs H190-3, H190-GCN4, HA2A-1 and HA2A-2, SEC-MALS profiles of these constructs show a major peak corresponding to modular VP1 capsomeres and a minor peak corresponding to dimeric modular VP1 capsomeres. The peaks corresponding to dimeric modular VP1 capsomeres, however, were negligible. The result suggests that the modular constructs exhibited monodispersity and were stable at capsomere form after concentration. Meanwhile, non-concentrated HA2A-3 and HA2A-4 exhibit polydispersity, indicated by their multiple peaks in SEC-MALS profiles (Figure 5.11). After concentration, constructs HA2A-3 and HA2A-4 were

unstable and tended to form aggregates, shown by the multiple peaks eluted earlier compared with the non-concentrated samples. The polydispersity and aggregation tendency of HA2A-3 and HA2A-4 might result in a lower chance of obtaining protein crystals in crystallisation screening, reported in (Bartlett, 2000).

Table 5.3: The concentrations of concentrated modular VP1 capsomeres.

Protein name	**Protein concentration (mg mL^{-1})**
H190-GCN4	12.6
H190-2	11.5
H190-3	12.8
HA2A-1	10.8
HA2A-2	8.8
HA2A-3	8.3
HA2A-4	12.3

The dispersity, such as monodispersity or polydispersity, of the modular H190 and HA2A capsomeres could be affected by the size of the presented antigenic module. Constructs H190-GCN4, H190-2, H190-3, HA2A-1 and HA2A-2, presenting the antigenic modules having theoretical molecular weights in the range of 3.0–5.7 kDa and exist in monodisperse form. The increase in antigen copy number of constructs HA2A-3 and HA2A-4 resulted in an increase of the theoretical molecular weight of the antigenic modules to 8.3 kDa and 11 kDa, respectively. This increased size possibly causes the strong tendency for intermolecular interactions of the antigenic modules on HA2A-3 and HA2A-4 capsomeres, leading to the polydispersity of these constructs.

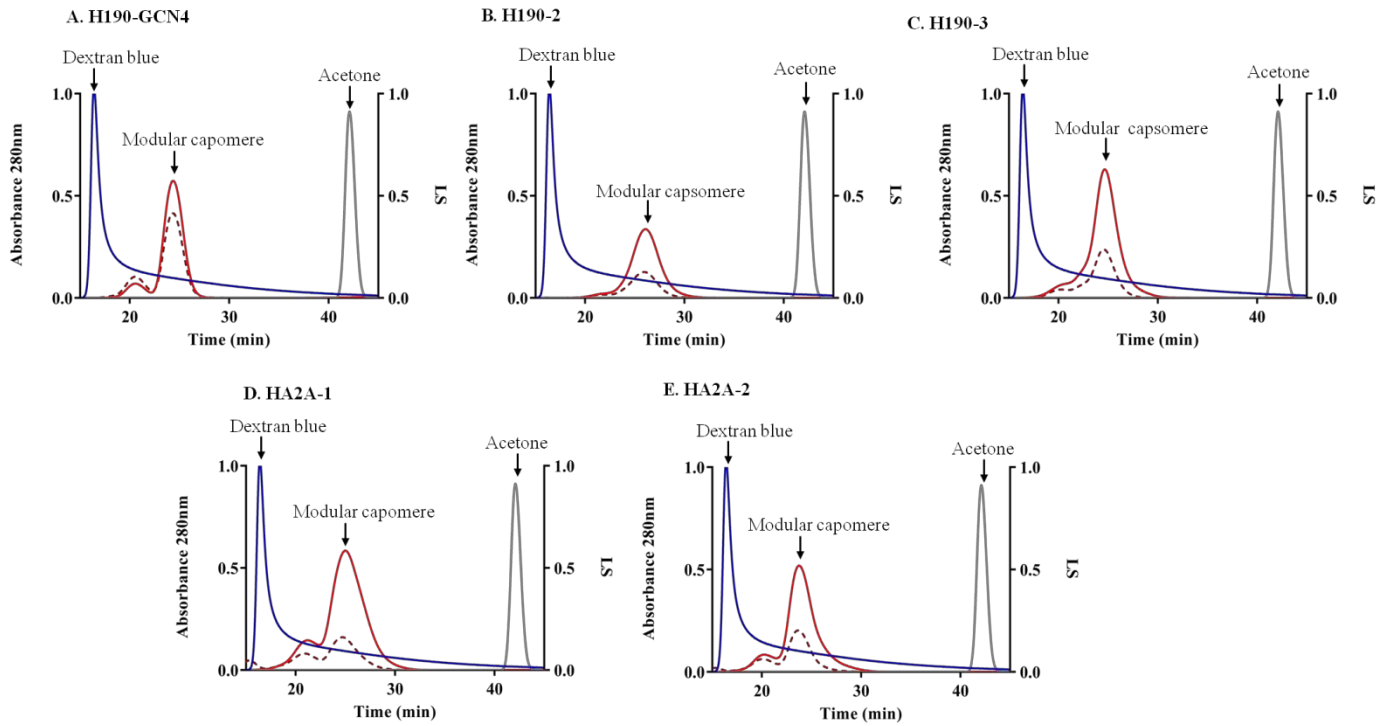


Figure 5.10: SEC-MALS analysis of modular VP1 constructs after concentration (A) H190-GCN4; (B) H190-2; (C) H190-3;(D) HA2A-1;(E) HA2A-2. Dextran blue is represented in blue, acetone in grey and the modular constructs in red. Solid lines represent UV absorbance at 280 nm, and broken lines represent light scattering (LS). The SEC-MALS analysis was run using Superdex™ 200 10/300 GL (GE Healthcare Bioscience, UK).

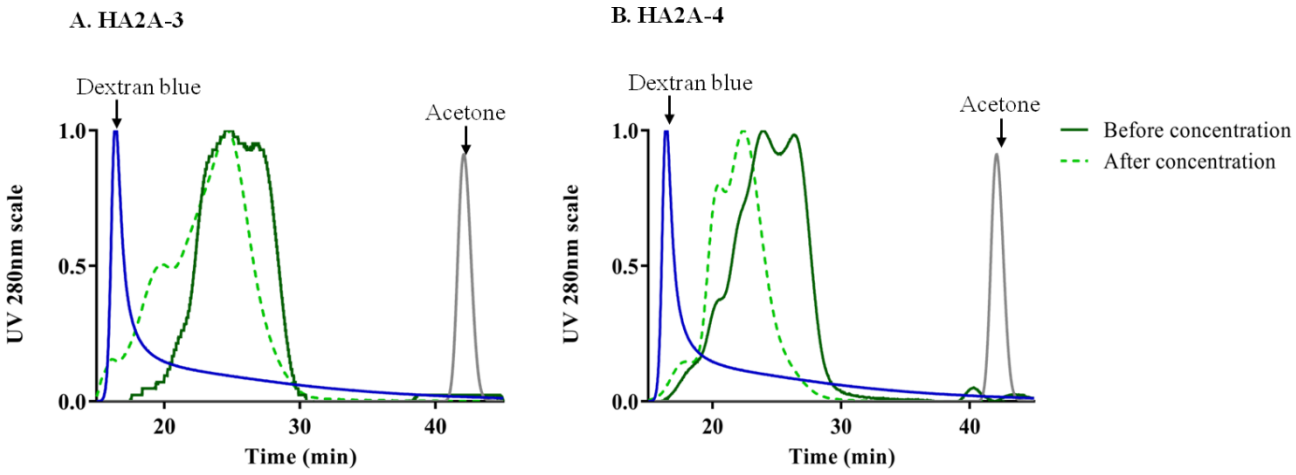


Figure 5.11: SEC-MALS analysis of modular VP1 constructs before and after concentration (A) HA2A-3; (B) HA2A-4. Dextran blue is represented in blue, acetone in grey, modular constructs before concentration in dark green and modular constructs after concentration in light green. The SEC-MALS analyses were run using Superdex™ 200 10/300 GL (GE Healthcare Bioscience, UK).

This study further investigated the cause of the polydispersity of HA2A-3 and HA2A-4. Table 5.4 compares the grand average of hydropathicity (GRAVY) values, measuring the hydrophobicity or hydrophilicity of proteins, and the theoretical isoelectric point of the modular HA2A capsomeres. The data indicate that the increase in antigen copy numbers did not change the GRAVY values of the modular capsomeres, but led to a slight decrease in the theoretical pI. Therefore, ionic interaction is hypothesised to cause the intermolecular interaction of capsomeres HA2A-3 and HA2A-4. The ionic interactions among the modular capsomeres could be weakened or strengthened through adjusting the buffer components such as pH and salt concentration. To test this hypothesis, the pH and salt concentration of the digestion buffer was varied to examine their effect on the polydispersity of HA2A-3 and HA2A-4 after TEVp treatment for GST tag release. The pH used were pH 7.5, 8.0 and 8.5 and salt concentrations of 100 mM, 200 mM and 500 mM NaCl. The total digested product of GST-tagged HA2A-3 and HA2A-4 were then analysed using SEC-MALS and SDS-PAGE. Figure 5.12 shows that the pH and salt concentration of the digestion buffer did not affect the polydispersity of construct HA2A-3, indicated by the similarity in SEC-MALS and SDS-PAGE data for all conditions. It proves that ionic interactions did not affect the polydispersity of construct HA2A-3.

Table 5.4: Comparison of hydrophobic (GRAVY) and charge (theoretical pI) properties among modular capsomeres: HA2A-1, HA2A-2, HA2A-3 and HA2A-4.

Modular constructs	**GRAVY**	**Theoretical pI**
HA2A-1	-0.416	5.36
HA2A-2	-0.420	5.29
HA2A-3	-0.424	5.23
HA2A-4	-0.428	5.18

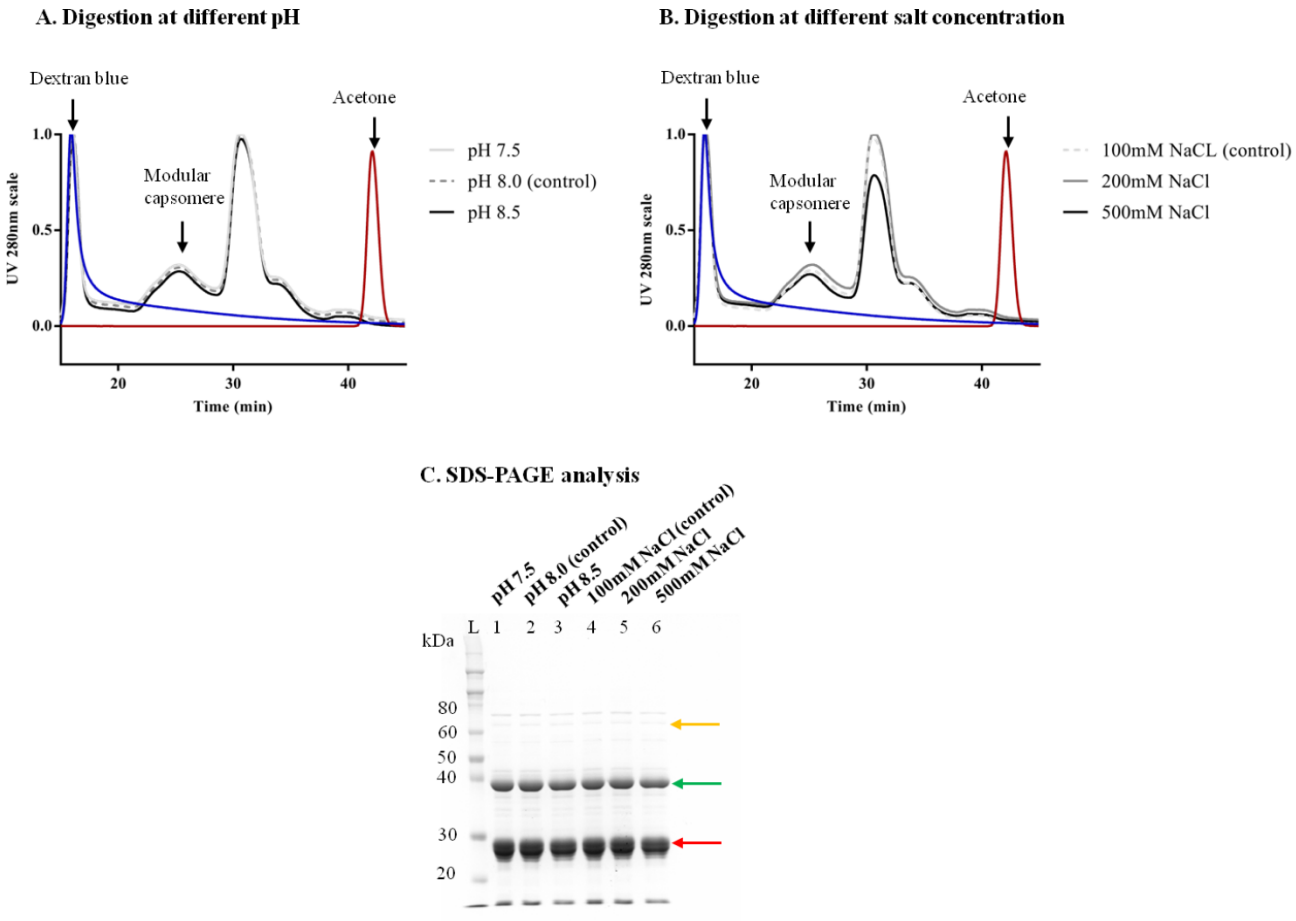


Figure 5.12: SEC-MALS analysis of constructs HA2A-3 after GST tag removal from untagged modular capsomeres in different (A) pH; (B) salt concentration. C) SDS-PAGE analysis of constructs HA2A-3 after the GST tag cleavage in various pH and salt concentrations. Undigested GST-tagged modular VP1: orange arrow, modular VP1: green arrow and GST tag: red arrow. The SEC-MALS analyses were run using Superdex^{TM} 200 10/300 GL (GE Healthcare Bioscience, UK).

For construct HA2A-4, the variation of pH of digestion buffer also did not affect the polydispersity of HA2A-4, shown by the similarity of the SEC-MALS profiles (Figure 5.13A) and SDS-PAGE (lanes 1, 2 and 3 in Figure 5.13C). Furthermore, the increase in NaCl concentration in digestion buffer led to the incomplete TEVp digestion of GST-tagged HA2A-4 to release GST tag. Comparison of lanes 4, 5 and 6 in Figure 5.13C found an increase in the intensity of the band corresponding to undigested GST-tagged HA2A-4 when the NaCl concentration increased from 100 mM to 500 mM. The incomplete digestion in 200 mM and 500 mM NaCl led to the observation of a peak, corresponding to the undigested material, eluted earlier than the modular capsomere peak (Figure 5.13B). Also, the SEC chromatogram for the digestion at 500 mM NaCl shows an abnormal peak after the GST peak (dimer GST) (Figure 5.14A). SDS-PAGE analysis (Figure 5.14B) indicates that the protein band corresponding to the abnormal peak was the same size as GST. This result suggested that the peak could be monomer GST. Some studies have found that GST can exist in both dimeric and monomeric forms (Andújar-Sánchez *et al.*, 2004; Fabrini *et al.*, 2009; Huang *et al.*, 2008). GST exists as a dimer in its active state, and as a monomer in an inactive unfolded state (Hornby *et al.*, 2002). It is possible that the GST removal from untagged HA2A-4 capsomere in 500 mM NaCl could cause disassociation of GST dimers to GST monomers. Two GST monomers interact each other via hydrogen bonds to form a GST dimer, as shown in a previous study (Rufer *et al.*, 2005). The high salt concentration might break down the intramolecular interactions of GST dimers, resulting in a release of GST monomers.

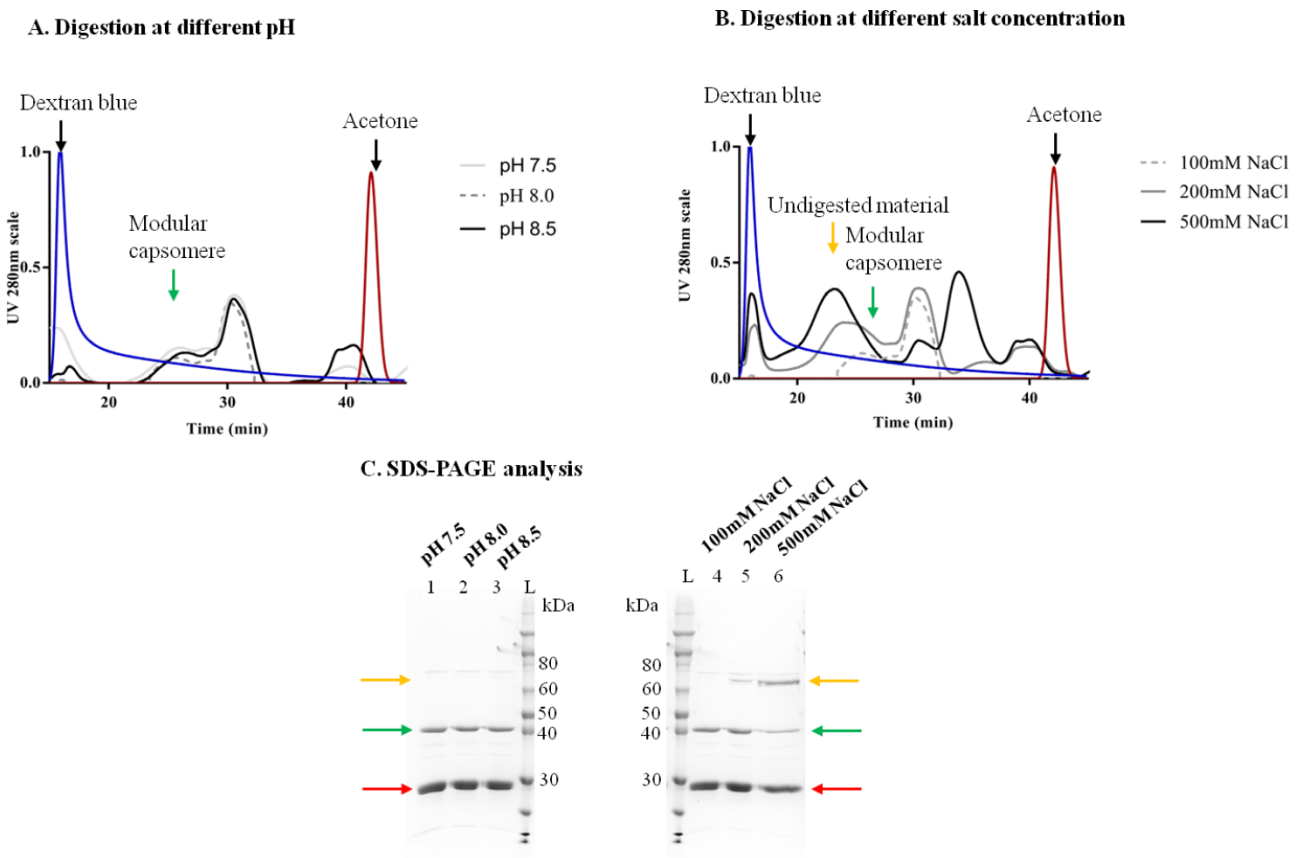


Figure 5.13: SEC-MALS analysis of construct HA2A-4 after GST tag removal from untagged modular capsomeres in different (A) pH; (B) salt concentration. C) SDS-PAGE analysis of the GST tag cleavage of construct HA2A-4 in different pH and salt concentration. Undigested GST-tagged modular VP1: orange arrow, modular VP1: green arrow and GST tag: red arrow. The SEC-MALS analyses were run using Superdex^{TM} 200 10/300 GL (GE Healthcare Bioscience, UK).

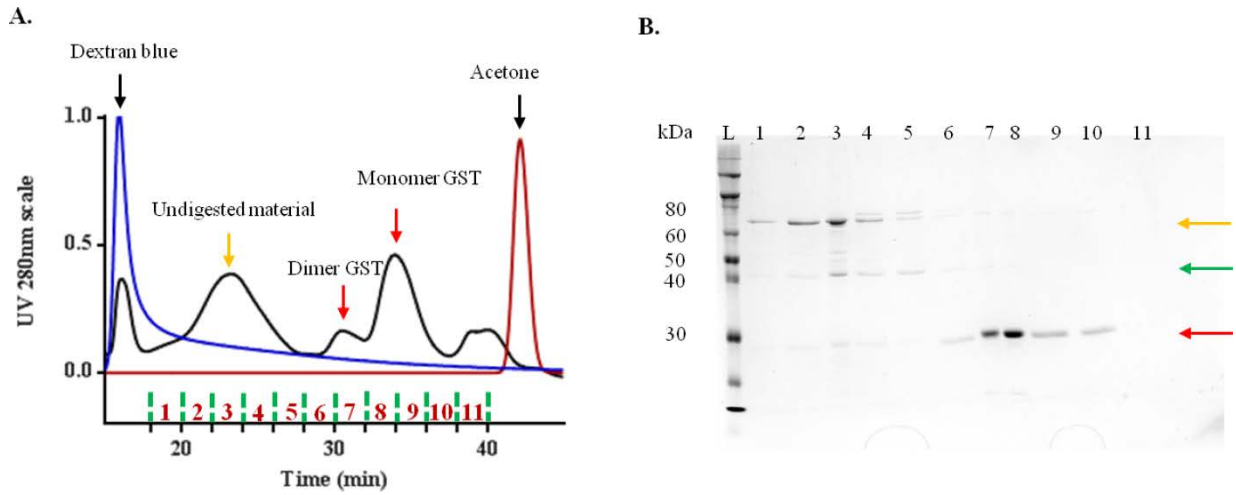


Figure 5.14: A) SEC-MALS analysis of construct HA2A-4 after GST tag removal from untagged modular capsomeres in 500 mM NaCl. B) SDS-PAGE analysis of collected SEC fractions. The lane numbers correspond to the fraction numbers in SEC-MALS analysis. Undigested GST-tagged modular VP1: orange arrow, modular VP1: green arrow and GST tag: red arrow. The SEC-MALS analyses were run using Superdex^{TM} 200 10/300 GL (GE Healthcare Bioscience, UK).

In summary, due to their polydispersity, constructs HA2A-3 and HA2A-4 were not used in crystallisation screening. For construct H190-GCN4, H190-2, H190-3, HA2A-1 and HA2A-2, 3 mg of each purified protein at a concentration above 8 mg ml^{-1} were used for crystallisation screening at the Australian Synchrotron (Melbourne, Australia).

5.3.5. Crystallisation screening

The crystal conditions of 2.0M ammonium phosphate pH 8.0 and 5% ethanol from the study of Stehle and Harrison (1997) were used as a starting point for screening the crystallisation conditions of all modular VP1 capsomeres. However, no protein crystals were formed under these conditions. Therefore, crystallisation screenings were performed by varying parameters such as the precipitants used, the concentration of the precipitant, the pH and the temperature. The screenings led to the formation of microcrystals (about 10 microns in size) for constructs H190-GCN4, H190-2, HA2A-1 and HA2A-2 (Figure 5.15). Despite the formation of these protein crystals, their size was too small for X-ray diffraction experiments. The crystallisation conditions for these constructs, therefore, are currently being optimised to increase the size of the crystals.

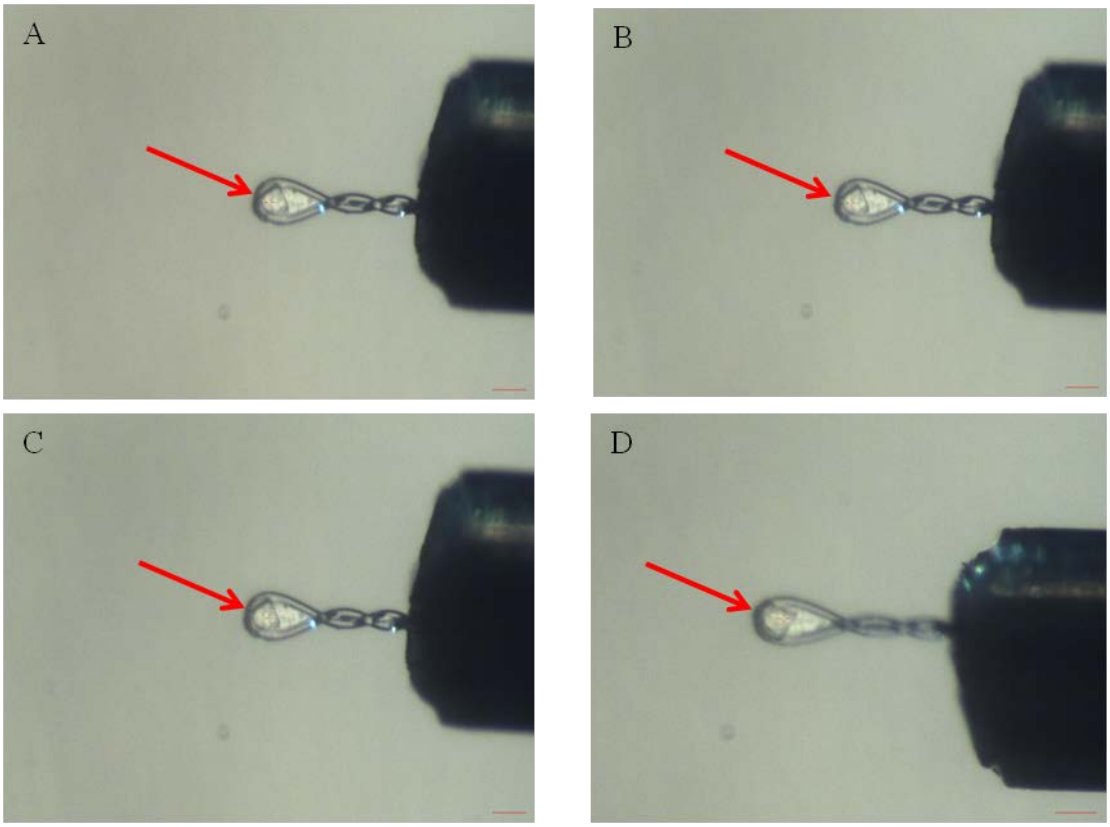


Figure 5.15: The microcrystals of modular constructs: (A) H190-GCN4; (B) H190-2; (C) HA2A-1; (D) HA2A-2. The red arrows indicate the position of microcrystals.

For construct H190-3, the screening resulted in the formation of a 100-micron protein crystal in 0.3M $(NH_4)_2SO_4$. The crystal belongs to a primitive orthorhombic space group. The crystal was mounted and exposed to X-ray beam light. The X-ray beam diffracted the crystal, and diffraction spots were collected on a film (Figure 5.16). Table 5.5 shows the statistics of the diffraction collection data. The results indicate that the diffraction resolution of the crystal was about 8.4 Angstrom (Å). This diffraction data is sufficient to determine unit cell parameters and space groups (Table 5.5). However, the diffraction data had a too low resolution to be able to solve the atomic structure of the protein. The diffraction resolution required for determining protein structure is at least 4 Angstrom (Å). Due to the poor resolution of crystal diffraction, further optimisation of the crystallisation conditions is required to obtain a crystal that will allow the atomic structure to be solved.

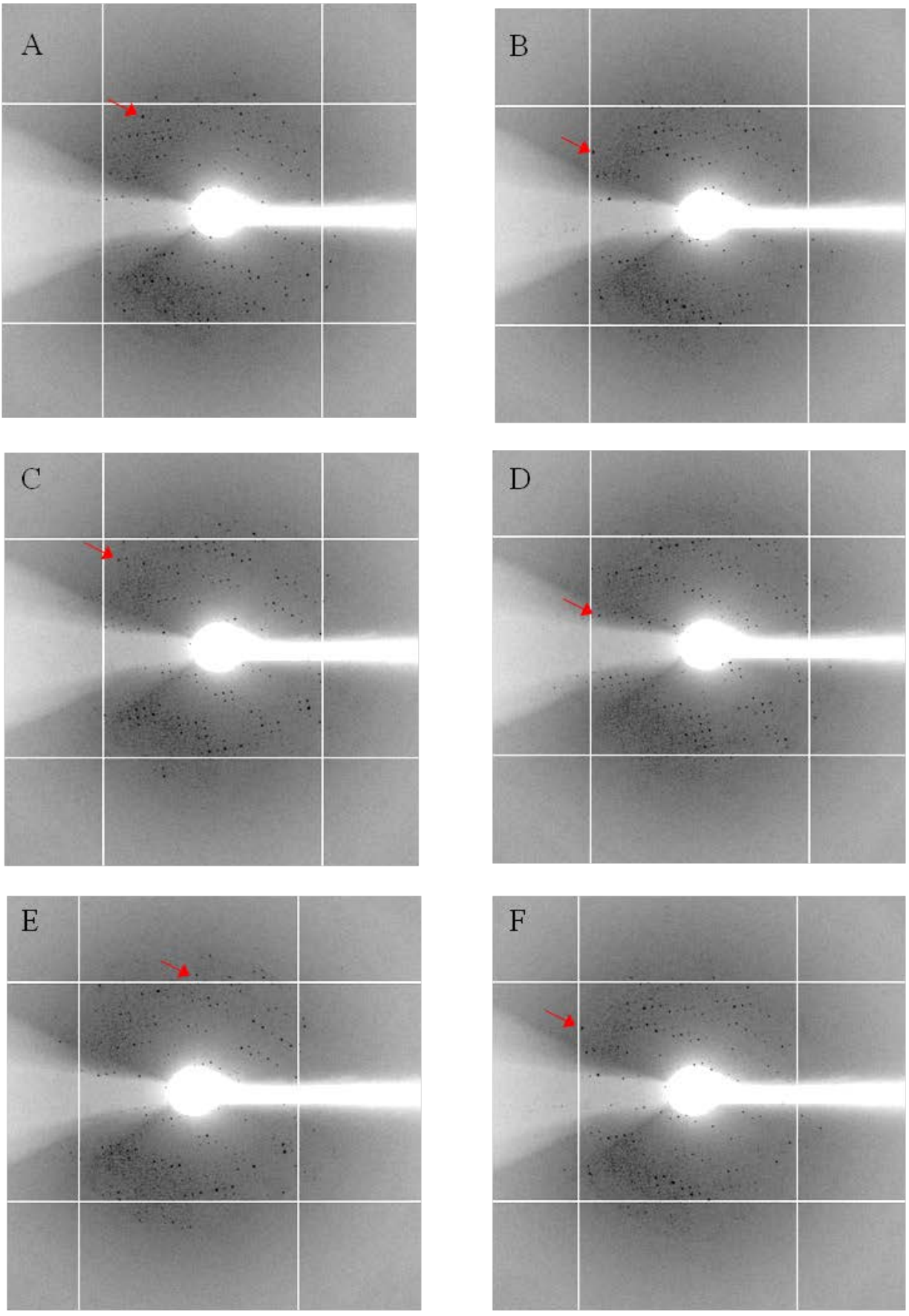


Figure 5.16: The X-ray diffraction pattern of crystallised H190-3. Black spots (red arrows) were diffraction spots of the crystal.

Table 5.5: The diffraction data collection of the H190-3 crystal.

	Dataset
Cell parameters	148.66 153.56 162.63 90.000 90.000 90.000
Resolution	49.24 – 8.36 (8.85 – 8.36)
Completeness	81.3% (80.7%)
I/sigma (I)	4.33 (1.28)
Rmeas	40.4% (147.1%)
Rsym	33.53% (122.7%)
Multiplicity	2.9
Compared	8574 (1283)
Measured	9009
Unique	3118
Rejected	Misfits 164
Wilson scaling (B/Corr)	381.0 (0.94)

In summary, due to time constraints, further crystallisation optimisation for constructs H190-GCN4, H190-2, HA2A-1, HA2A-2 and H190-3 could not be completed within the timescale of this thesis. The crystallisation optimisation was still ongoing at the time this thesis was submitted.

5.4. Conclusions

This chapter reports a strategy to improve the purity of the modular VP1 capsomeres to above 90% purity, which is mandatory to obtain a high-quality crystal for X-ray crystallisation. It is the introduction of two extra purification steps including GST affinity and hydrophobic interaction chromatography. First, the second GST affinity chromatography step is to remove undigested GST-tagged modular VP1 and GST tag. Second, hydrophobic interaction chromatography (HIC) is to eliminate co-purified DnaK contaminants. The application of HIC in modular VP1 capsomere purification, however, is limited to modular VP1 capsomeres presenting hydrophobic antigenic modules. In HIC, the interaction with hydrophobic ligands on media occurs via exposed hydrophobic residues on proteins. Hydrophobic antigenic modules bear hydrophobic residues that are hidden by DnaK. This results in an insufficiency of exposed hydrophobic residues available for interaction with ligands. HIC, therefore, is inappropriate to purify modular VP1 capsomeres presenting hydrophobic antigenic modules.

The findings in this chapter indicated that the increase in the antigenic element copy number could cause the polydispersity of modular capsomere, which can reduce the chance of obtaining crystals in crystallisation screening. The increased antigenic element copy number results in the increased antigenic module size, which might interfere the capsomere structure. Molecular dynamic (MD) simulations predict that the perturbation of the VP1 capsomere structure resulted from the insertion of four HA2A antigenic element copies (chapter 4). The insertion of large antigen was also reported to cause the structural disruption of modular VP1 capsomere (Lua *et al.*, 2015). The result suggests that the capsomere structure perturbation could lead to the polydispersity of the HA2A-4 capsomeres. Polydispersity, therefore, can be used as an early indication of intolerance of the VP1 modular capsomere to the insertion. Although the limit of the antigenic module size has not been determined, the data indicate that modular VP1 capsomeres presenting antigenic modules larger than 8 kDa may be inappropriate for determining atomic structure using X-ray crystallography.

A protein crystal of the modular VP1 capsomere was obtained. Unfortunately, the diffraction resolution was insufficient to determine the atomic structure. Further screening of crystallisation conditions is needed to grow optimum crystals for structure determination. This activity is time-consuming without the certainty of obtaining quality diffraction data (Chayen and Saridakis, 2008; Li and Chang, 2009). At the time that this PhD thesis was written, the optimisation of the crystallisation conditions was still ongoing.

References

Aitken, A., Learmonth, M., 2009. Protein determination by UV absorption, in: Walker, J. (Ed.), The protein protocols handbook. Humana Press, pp. 3-6.

Andújar-Sánchez, M., María Clemente-Jimenez, J., Rodriguez-Vico, F., Javier Las Heras-Vazquez, F., Jara-Pérez, V., Cámara-Artigas, A., 2004. A monomer form of the glutathione S-transferase Y7F mutant from *Schistosoma japonicum* at acidic pH. Biochem Biophys Res Commun 314, 6-10.

Anggraeni, M.R., Connors, N.K., Wu, Y., Chuan, Y.P., Lua, L.H.L., Middelberg, A.P.J., 2013. Sensitivity of immune response quality to influenza helix 190 antigen structure displayed on a modular virus-like particle. Vaccine 31, 4428-4435.

Baneyx, F., Mujacic, M., 2004. Recombinant protein folding and misfolding in *Escherichia coli*. Nat Biotechnol 22, 1399-1408.

Bartlett, P., 2000. The effect of polydispersity on colloidal phase transitions, in: Buckin, V. (Ed.), Trends in colloid and interface science XIV. Springer Berlin Heidelberg, pp. 137-140.

Blundell, T.L., Jhoti, H., Abell, C., 2002. High-throughput crystallography for lead discovery in drug design. Nature 1, 45-54.

Bubeck, P., Winkler, M., Bautsch, W., 1993. Rapid cloning by homologous recombination *in vivo*. Nucleic Acids Res 21, 3601-3602.

Chayen, N.E., Saridakis, E., 2002. Protein crystallization for genomics: towards high-throughput optimization techniques. Acta Crystallogr D Biol Crystallogr 58, 921-927.

Chayen, N.E., Saridakis, E., 2008. Protein crystallization: from purified protein to diffraction-quality crystal. Nat Methods 5, 147-153.

Chuan, Y.P., Lua, L.H.L., Middelberg, A.P.J., 2008. High-level expression of soluble viral structural protein in *Escherichia coli*. J Biotechnol 134, 64-71.

Dong, X.F., Natarajan, P., Tihova, M., Johnson, J.E., Schneeman, A., 1998. Particles polymorphism caused by deletion of a peptide molecular switch in a quasiequivalent icosahedral virus. J Virol 72, 6024-6033.

Fabrini, R., De Luca, A., Stella, L., Mei, G., Orioni, B., Ciccone, S., Federici, G., Lo Bello, M., Ricci, G., 2009. Monomer-dimer equilibrium in glutathione transferases: a critical re-examination. Biochemistry 48, 10473-10482.

Fan, Y.Y., Middelberg, A.P.J., 2011. A method of preparation of a biological particulate structure. University of Queensland, Australia.

Francis, D.M., Page, R., 2010. Strategies to optimize protein expression in *E. coli*. Curr Protoc Protein Sci Chapter 5, Unit 5.24.21-29.

Habig, W.H., Pabst, M.J., Jakoby, W.B., 1974. Glutathione S-transferases. The first enzymatic step in mercapturic acid formation. J Biol Chem 249, 7130-7139.

Hartmann, E.M., Armengaud, J., 2014. N-terminomics and proteogenomics, getting off to a good start. Proteomics 14, 2637-2646.

Hornby, J.A.T., Codreanu, S.G., Armstrong, R.N., Dirr, H.W., 2002. Molecular recognition at the dimer interface of a class Mu glutathione transferase: Role of a hydrophobic interaction motif in dimer stability and protein function. Biochemistry 41, 14238-14247.

Huang, Y.-c., Misquitta, S., Blond, S.Y., Adams, E., Colman, R.F., 2008. Catalytically active monomer of glutathione S-transferase π and key residues involved in the electrostatic interaction between subunits. J Biol Chem 283, 32880-32888.

Korepanova, A., Douglas, C., Leyngold, I., Logan, T.M., 2001. N-terminal extension changes the folding mechanism of the FK506-binding protein. Protein Sci 10, 1905-1910.

Li, M., Chang, W., 2009. Protein crystallization. Photosynth Res 102, 223-229.

Lipin, D.I., Raj, A., Lua, L.H.L., Middelberg, A.P.J., 2009. Affinity purification of viral protein having heterogeneous quaternary structure: Modeling the impact of soluble aggregates on chromatographic performance. J Chromatogr 1216, 5696-5708.

Lua, L.H.L., Fan, Y., Chang, C., Connors, N.K., Middelberg, A.P.J., 2015. Synthetic biology design to display an 18 kDa rotavirus large antigen on a modular virus-like particle. Vaccine 33, 5937-5944.

Parrish, J.R., Limjindaporn, T., Hines, J.A., Liu, J., Liu, G., Finley, R.L., Jr., 2004. High-throughput cloning of Campylobacter jejuni ORfs by in vivo recombination in *Escherichia coli*. J Proteome Res 3, 582-586.

Rivera-Hernandez, T., 2012. Bioengineering virus-like particles for vaccine development, Australian Institute for Bioengineering and Nanotechnology. The University of Queensland.

Rüdiger, S., Germeroth, L., Schneider-Mergener, J., Bukau, B., 1997. Substrate specificity of the DnaK chaperone determined by screening cellulose-bound peptide libraries. Embo J 16, 1501-1507.

Rufer, A.C., Thiebach, L., Baer, K., Klein, H.W., Hennig, M., 2005. X-ray structure of glutathione S-transferase from *Schistosoma japonicum* in a new crystal form reveals flexibility of the substrate-binding site. Acta Crystallogr Sect F: Struct Biol Cryst Commun 61, 263-265.

Stehle, T., Harrison, S.C., 1996. Crystal structures of murine polyomavirus in complex with straight-chain and branched-chain sialyloligosaccharide receptor fragments. Structure 4, 183-194.

Stehle, T., Harrison, S.C., 1997. High-resolution structure of a polyomavirus VP1-oligosaccharide complex: implications for assembly and receptor binding. Embo J 16, 5139-5148.

Tolia, N.H., Joshua-Tor, L., 2006. Strategies for protein coexpression in *Escherichia coli*. Nat Methods 3, 55-64.

Wibowo, N., Chuan, Y.P., Lua, L.H.L., Middelberg, A.P.J., 2012. Modular engineering of a microbially-produced viral capsomere vaccine for influenza. Chem Eng Sci 103, 12-20.

Wynne, S.A., Crowther, R.A., Leslie, A.G.W., 1999. The crystal structure of the human hepatitis B virus capsid. Mol Cell 3, 771-780.

Chapter 6

Conclusions and future work

6.1. Summary of research findings

Influenza is a contagious respiratory illness that causes high mortality and economic burdens on society. The prevention of influenza has many challenges due to the highly mutagenic nature of the virus. Current vaccine manufacturing methods are unable to rapidly respond to pandemics, suggesting a necessity for a new vaccine technology. The MuPyV vaccine platform at UQ is a promising approach that can rapidly respond to influenza pandemics because it exploits the advantages of a prokaryotic expression system — high yield, high speed, and low cost. A research group at UQ has developed this vaccine platform to present heterogeneous epitopes or antigen elements to induce an immune response against the antigenic elements. The presentation of an epitope in its native structure on the surface of VLP is essential to induce neutralising antibodies; however, there is a knowledge gap in this field.

A previous study within the UQ group observed that modularisation of an antigenic element using different display strategies resulted in variation in the quality of immune response. The modularisation of an influenza epitope, Helix 190 (H190), using a tandem repeat display strategy, resulted in higher antibody titres against HA1 than the employment of flanking sequence (Anggraeni *et al.*, 2013). However, the structures of the antigenic modules have not been determined. Determining the structure of the antigenic element will better inform and guide the design of epitope presentations on VLPs. A generic design for epitope presentation on VLP would accelerate the re-engineering of influenza vaccine components to respond quickly to influenza pandemics, and contribute to the development of modular VLP-based vaccines.

The experimental works in this thesis were designed to address two key unknowns in the presentation of structural epitopes on modular MuPyV VLP:

1. Identification of the structure of presented antigenic modules

This preliminary study evaluated the ability of phage antibody panning to identify the structural presentation of H190 modules, displayed using two strategies: flanking with the GCN4 structure promoting element and using tandem repeats (Chapter 3). The need for an alternative method to identify the antigenic module conformation was highlighted.

2. The effect of tandem repeat display strategy in promoting the helical propensity of antigenic modules

This preliminary study evaluated the capacity of a tandem repeat display strategy to promote the helical propensity of antigenic modules. Two protein structure determination techniques were explored to investigate the structure of antigenic modules on modular capsomeres: (i) circular dichroism (CD) spectroscopy (Chapter 4) and (ii) X-ray crystallography (Chapter 5).

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

6.1.1. Identification of the structure of presented antigenic modules

Phage antibody panning was investigated to identify the conformational presentation of an H190 element, using two different display strategies. In the first strategy, a single H190 element is flanked with the GCN4 structure promoting element. In the second strategy, two copies of the H190 element are arrayed in a tandem repeat. The pannings against modular capsomeres and H1H1 HA1 were performed to seek modular capsomere-binding phages that can bind to HA1. The phages were then used to inform the conformation of the H190 element compared with native H190 on HA1. If modular capsomere-binding phages can bind to H1N1 HA1 at the same level that they bind to modular capsomere, it indicates that the structure of the H190 element on the modular capsomere may be identical to the structure of native H190. If modular capsomere-binding phages are unable to bind to H1N1 HA1 or bind weakly to H1N1 HA1, it indicates that the structure of the H190 element on the modular capsomere may be different from the structure of native H190.

Two sets of panning, A and B, were implemented in parallel to investigate the conformational presentation of the H190 element on H190-GCN4 and H190-2 capsomeres, respectively. After three rounds of panning against the modular capsomere and two against H1N1 HA1, polyclonal phage ELISA was performed to examine the enrichment of bound phages. Monoclonal phage ELISA was then examined the binding region of isolated phages.

The polyclonal ELISA against base capsomere A (Cap-Base A) as a subtractive antigen and modular capsomere H190-GNC4 (CapH190-GNC4) as a selective antigen illustrated that the panning might selectively isolate and enrich H190 antigenic module-binding phages. The polyclonal ELISA against H1N1 HA1 suggested that a small quantity of isolated phages were bound to H1N1 HA1. The monoclonal phage ELISA showed that four clones were bound to both CapH190-GNC4 and H1N1 HA1. Further investigation of the binding regions of these phages, using competitive ELISA with peptide H190 and peptide H190-GNC4 as the competitors, demonstrated that the positive clones were bound to the junction between the H190 and GCN4 elements. This explained why the positive clones responded strongly to CapH190-GNC4 that has an H190 element and a GCN4 element joined together into a module, rather than to HA1. It indicates that these phages can discriminate the conformational difference between the H190 antigenic module and native H190 on H1N1 HA1.

The polyclonal ELISA against unmodified capsomere B (Cap-Base B) as a subtractive antigen and modular capsomere H190-2 (CapH190-2) as a selective antigen indicated the lack of selectivity for phages binding to H190 module. This lack of selectivity suggested that the isolated pool of phages might lack phages bound to both CapH190-2 and H1N1 HA1. The panning of the phages isolated after the third pannings against H1N1 HA1 for a further two rounds led to a considerable decrease in the quantity of CapH190-2-binding phages. This observation shows that the isolated CapH190-2-binding phages, after the third panning, are unable to bind to the H190 module on CapH190-2, resulting either in them unbound to H1N1 HA1, or the elimination of the phages in the pannings against H1N1 HA1. The monoclonal phage ELISA result is in agreement with the polyclonal ELISA result, confirming that the isolated pool of phages in round five did not contain phages that were bound to both CapH190-2 and H1N1 HA1.

The findings in this chapter (Chapter 3) reveal that the phage libraries may lack phages bound to the H190 element derived from the A/California/07/2009 strain. The generation of an immunised phage antibody library that is relevant to A/California/07/2009 strain is required. However, the construction of immunised phage antibody library is costly, laborious and time-consuming, which can slow down the response of the vaccine platform to an influenza pandemic. The presented results indicate that phage antibody panning is an inappropriate tool to identify antigenic module structure for influenza VLP-based vaccine development. This finding also highlighted a necessity for other tools to understand the structure of antigenic modules on modular capsomeres.

6.1.2. The effect of the tandem repeat display strategy in promoting the helical propensity of antigenic modules

The structure of the antigenic element is essential to determine the capacity of the tandem repeat strategy to promote the helicity of the antigenic module. Motivated by the findings in chapter 3, the following chapters examined two protein structure determination techniques — circular dichroism (CD) spectroscopy (Chapter 4) and X-ray crystallography (Chapter 5) — to determine the conformation of the antigenic elements. The experimental works performed in Chapter 4 and 5 aim to address the research questions: What is the structure of the H190 modules using the two display strategies? Can the increase in the copy number of antigenic elements in the tandem repeat lead to improved helicity? To what extent does the tandem repeat strategy affect the helical propensity of HA2A modules? What is the preferred copy number of the antigenic element for the presentation of the HA2A module? Can the modularised HA2A elicit IgGs recognising peptide HA2A and FL-HA, showing the induction of immunogenicity and antibody quality, respectively?

6.1.2.1. Circular dichroism spectroscopy

The study in Chapter 4 investigates the combination of simpler techniques, CD spectroscopy coupled with SVD algorithms (CD-SVD), to determine the helical structure of antigenic modules on modular capsomeres. From protein CD spectra, SVD algorithms were applied to estimate the ratios of types of protein secondary structures including α-helix (Sreerama and Woody, 2000). The ratios were then used to calculate the number of amino acids participating in the secondary structures. An increase in the number of helical amino acids in modular capsomeres compared with unmodified capsomeres indicated the formation of helical structures of antigenic modules.

The capability of CD-SVD to detect helical structures within VP1 capsomeres was first validated by analysing the helical amino acid number of VP1 capsomere variants VP1, VP1-ΔC and VP1-ΔNΔC. VP1 has an additional helical structure at the C-terminus compared with VP1-ΔC and VP1-ΔNΔC while VP1-ΔNΔC lacks a flexible structure at N-terminus compared with VP1-ΔNΔC. The validation of the CD-SVD method shows that CD-SVD detected the helical structure at the C-terminal of VP1. This was demonstrated by the detection of a higher helical amino acid number for VP1 and a similar number of helical amino acids for VP1-ΔC and VP1-ΔNΔC. However, the CD-SVD method can only provide a relative number of helical amino acids. While the detection of CD-SVD suggested the C-terminus of VP1 has three helical amino acids, the crystal structure of VP1 monomer (1SID.pdb) illustrates that VP1 monomer comprises 11 helical amino acids.

CD-SVD was then used to analyse the helical amino acid number of modular HA2A capsomeres. Each module of VP1-ΔNΔC-HA-1, VP1-ΔNΔC-HA-2, VP1-ΔNΔC-HA-3 and VP1-ΔNΔC-HA-4 presents one, two, three or four copies of the HA2A element arrayed in tandem repeats, respectively. The VP1-ΔNΔC-HA-1 capsomere, having one antigen/module, has no increase in helical amino acid numbers compared with the base capsomere. The number of antigenic elements increases by two, three or four copies, leading to a significant increase in helical amino acid numbers. VP1-ΔNΔC-HA-2, -3 and -4 showed significant increases in relative helical amino acid numbers compared to the base capsomere ($P<0.001$).

MD simulation was used to assist CD-SVD in the structural analysis of modular HA2A capsomeres. CD-SVD and MD simulations were in good agreement in informing the structure of HA2A antigenic module. Both CD-SVD and MD simulation results indicated that HA2A modules on VP1-ΔNΔC-HA-1 formed non-helical structures. Structural analysis obtained by CD-SVD showed that increasing the element numbers to two, three and four HA2A elements per module, on VP1-ΔNΔC-HA-2, VP1-ΔNΔC-HA-3 and VP1-ΔNΔC-HA-4, respectively, led to an increment in relative helical amino acid number. MD simulation also predicted that increasing the number of elements, on VP1-ΔNΔC-HA-2, VP1-ΔNΔC-HA-3 and VP1-ΔNΔC-HA-4, enabled more elements to retain a relatively native structure.

MD simulation also predicted that the capsomere structure of VP1-ΔNΔC-HA-1, VP1-ΔNΔC-HA-2 and VP1-ΔNΔC-HA-3 maintain low RMSD values (about 4Å), suggesting that the capsomere structures can tolerate the modular inserts. The increasing size of the modular insertions on VP1-ΔNΔC-HA-4, presenting four HA2A elements on each module, might interrupt the capsomere structures significantly. The steric barriers associated with such a large insert have been seen previously in (Lua *et al.*, 2015).

The immunogenicity of modularised HA2A and the quality of the induced antibodies were evaluated using serological testing of anti-modular HA2A VLP sera. The presentation of a single HA2A element per module did not lead to the induction of peptide HA2A- and FL-HA-recognising antibodies. This is in agreement with the MD simulation and the CD-SVD results, showing that the antigenic module on VP1-ΔNΔC-HA-1 may present in a non-helical structure. The presentation of dual HA2A elements per module also did not result in an antibody response against peptide HA2A and FL-HA. CD-SVD showed that the modules, bearing two HA2A elements, formed a helical structure, while MD simulation predicted that three out of ten modules retain relatively native structures, indicated by the low RMSD values (<2 Å). Despite the formation of a helical structure, the antigenicity of the modules was insufficient to evoke an immune response, led to the lack of immune response. The presentation of triple and quadruple HA2A elements per module results in

the elicitation of IgGs recognising peptide HA2A. The helical structure of peptide HA2A was illustrated in this study, and the results show that IgGs recognised the helical structure formed by HA2A modules bearing three or four HA2A elements. It was in agreement with the CD-SVD and MD simulation results, which suggested helical structures on the modules bearing three or four HA2A elements. The module bearing four HA2A elements, however, did not elicit an antibody response against FL-HA. This may be because the insertion of four antigenic elements per module led to the perturbation of the VP1 capsomere structure, as suggested by MD simulation. The module bearing three HA2A elements led to the induction of antibodies recognising FL-HA. The accessibility of the HA2A epitope on the FL-HA was also confirmed by the recognition of a conformational antibody, CR6261. This result suggests that the helical structure of modules bearing three HA2A elements may resemble the HA2A native structure of the FL-HA. It suggests that the presentation of triple HA2A antigenic elements is the optimum design for this particular helical structure.

This finding demonstrates that the CD-SVD method can reduce the number of vaccine candidates in preclinical and clinical testing, reducing the cost and, more importantly, the number of animals used in accordance with the Three Rs (Replacement, Reduction, and Refinement) principles in animal welfare. Due to the simplicity and speed of CD-SVD, the method can be used as a screening tool for the structural presentation of antigenic modules.

6.1.2.2. Using X-ray crystallography

The study in Chapter 5 employed X-ray crystallography, a proven atomic protein structure determination technique, to obtain the structure of antigenic modules on modular capsomeres. Seven modular VP1 capsomeres were selected to address the effect of (1) flanking sequence and tandem repeat display strategies; and, (2) antigenic element copy number on the helical module structure. First, two constructs: (i) H190-GCN4, bearing an H190 element flanked with GCN4 elements within each module, and (ii) H190-2, bearing two copies of the H190 element within each module, were selected to compare the conformation of modularised H190 in two different designs. Second, the effect of the antigenic element copy number was examined using two helical epitopes, H190 and HA2A. Constructs (ii) H190-2 and (iii) H190-3, having two and three copies of H190 element per module, respectively, were used to compare the helical propensity of the H190 element when its copy number was increased. Constructs (iv) HA2A-1, (v) HA2A-2, (vi) HA2A-3 and (vii) HA2A-4, bearing one, two, three and four copies of HA2A elements per module, respectively, were used to examine the effect of increasing the HA2A element copy number on the helical propensity of the HA2A element. The presence of the C- and N- termini was previously reported to prevent the

crystal formation of wt MuPyV VP1 protein (Stehle and Harrison, 1997). Thus, 31 residues at the N-terminus and 63 residues at the C-terminus of the modular VP1 constructs were removed.

Chapter 5 presents initial steps to obtain the atomic structure of the seven selected modular constructs. These steps are (i) production of highly purified and concentrated modular VP1 capsomeres including protein expression, purification, and SEC-MALS analysis of the concentrated protein; (ii) screening and optimisation of protein crystallisation conditions; and (iii) protein crystal diffraction collection.

The modular VP1 constructs were expressed as described in previous work. The solubility levels of the truncated VP1 (31–321 residues) and modular H190 constructs were lower than that of wt-VP1 (1–384 residues), suggesting that the truncation of VP1 at the C- and N- termini led to the decreased solubility. The modular H190 constructs had higher solubility levels than truncated VP1, suggesting that the insertion of H190 modules might result in improved solubility of the H190 constructs. The modular HA2A constructs, however, had lower solubility levels compared with truncated VP1. It indicates that the insertion of the HA2A module led to the decrease in the solubility of the truncated modular HA2A constructs. The effect of antigenic element insertion into VP1 to reduce the solubility of modular VP1 constructs was observed in a previous study (Wibowo *et al.*, 2012). A strategy, therefore, was applied to improve the solubility of these constructs by a decreasing the expression temperature and IPTG concentration. The results demonstrated that low expression temperature at 15°C and reduced IPTG concentration at 0.1mM IPTG resulted in improved solubility of all modular HA2A constructs. The effects of temperature and IPTG concentration on the solubility of the expressed protein were reported in previous studies (Baneyx and Mujacic, 2004; Francis and Page, 2010; Tolia and Joshua-Tor, 2006).

Size exclusion chromatography (SEC) was used to separate the modular capsomeres from soluble aggregates and GST tags. The modular H190 and HA2A capsomeres eluted at the same time with truncated VP1 capsomere (32–321 residues), but later than wt-VP1 capsomere (1–384 residues), as reported in a previous study (Wibowo *et al.*, 2012). It suggests that the removal of 31 residues at N- and 63 residues at C- termini might cause the decrease in hydrodynamic radius of the modular H190 and HA2A capsomeres, leading to the late elution. After SEC, the modular H190 and HA2A capsomeres contained three contaminants: (i) undigested GST-tagged modular VP1, (ii) the GST tag, and (iii) a 70-kDa contaminant.

The purity of the modular H190 and HA2A constructs was improved to at least 90% using two extra polishing steps. First, the GST affinity chromatography, with a decreased flow rate of 0.2 ml min^{-1}, was used to remove undigested GST-tagged modular VP1 and the GST tag. Second, hydrophobic interaction chromatography (HIC) was used to eliminate a 70-kDa contaminant, which was shown to be co-purified DnaK from *E. coli* in a previous study (Fan and Middelberg, 2011). The results indicated that DnaK strongly bound to modular capsomeres via hydrophobic regions. The binding of DnaK to hydrophobic regions of unfolded polypeptides was also reported in a previous study (Rüdiger *et al.*, 1997). Additionally, there may be some DnaK-associated modular VP1 and some free DnaK-modular VP1capsomeres. Because the interaction between DnaK and modular capsomeres in DnaK-modular capsomere complexes could not be broken, HIC separated free DnaK-modular VP1 capsomeres from DnaK-associated modular capsomeres. Therefore, to improve the purity of the modular capsomeres, some of the modular VP1 capsomeres were sacrificed during HIC.

SEC-MALS analysis showed that H190-GCN4, H190-2, H190-3, HA2A-1, and HA2A-2 exhibited monodispersity. In contrast, HA2A-3 and HA2A-4 exhibited polydispersity, which can reduce the chance of obtaining a crystal in crystallisation screening (Bartlett, 2000). The size of the antigenic module presented could affect the dispersity (monodispersity or polydispersity) of modular VP1 capsomeres. The theoretical molecular weights of the antigenic modules in constructs H190-GCN4, H190-2, H190-3, HA2A-1 and HA2A-2, are in the range 3.0–5.7 kDa, which could result in the monodispersity of these modular constructs. The increased antigen copy number of constructs HA2A-3 and HA2A-4 resulted in increased theoretical molecular weights of the antigenic modules of 8.3 kDa and 11 kDa. This increased size is a possible cause of the strong tendency for intermolecular interactions of the antigenic modules on capsomere HA2A-3 and HA2A-4, which leads to the polydispersity of these constructs.

In summary, 3 mg of concentrated modular VP1 capsomere constructs including H190-GCN4, H190-2, H190-3, HA2A-1, and HA2A-2 were sent to the Australian Synchrotron (Melbourne, Australia) for crystallisation screening. Due to their polydispersity, the constructs HA2A-3 and HA2A-4 were not sent for the screening. For the modular constructs H190-GCN4, H190-2, HA2A-1, and HA2A-2, the screening found crystallisation conditions for these modular constructs that formed microcrystals of about 10 microns, which is too small for crystal diffraction. For modular construct H190-3, a protein crystal with a size of 100 microns was obtained and diffracted at 8 Angstrom resolution. Unfortunately, this resolution was insufficient to solve the crystal structure, which requires further optimisation of crystallisation conditions to obtain an optimum crystal that

will allow resolution of the structure. At the time that this PhD thesis was written, the crystallisation optimisation is still ongoing.

6.2. Future work

Determination of the atomic structure of antigenic modules on modular capsomeres is essential to confirm the capacity of design strategy in maintaining and promoting the native structure of epitopes. It could assist in defining design rules for the presentation of structural epitopes on modular capsomeres or VLPs.

Further investigations of the following points are required to better understand how to present structural epitopes on modular MuPyV VLP and, by extension, on VLPs more generally.

1. The experiments in Chapter 4 investigated the capability of the tandem repeat display strategy to promoting helical structure. The results show that VP1-HA-3 VLPs were able to induce antibody titres against the helical structure of peptide HA2A and the full-length HA. It indicates that the helical structure of the HA2A antigenic element on VP1-HA-3 might be identical to the conformation of HA2A on full-length HA, which shows the ability of the tandem repeat display strategy to maintain or promote the helical structure of the HA2A antigenic module. A further investigation into the tandem repeat display strategy in the presentation of β-sheet structures could be necessary. This could enable the exploitation of the tandem repeat strategy as a generic design for the structural presentation of epitopes on modular capsomeres or VLPs.
2. In chapter 4, MD simulation was used to assist CD-SVD in characterising the structure of antigenic modules on modular capsomeres. Both MD simulation and CD-SVD identified the non-helical structure of antigenic modules and the probability of helical structure formation of antigenic modules. MD simulation can also predict the capsomere structure might be perturbed by the insertion of four HA2A elements. It suggests that MD simulation can be employed as a guiding tool for vaccine design. A further investigation into the capability of MD simulation in informing module structure is required.
3. Some neutralising epitopes, which have potential for presentation on modular VLP, have a non-helical structure, i.e. β-sheet structure (Inoue *et al.*, 2012; Pantophlet and Burton, 2006). Further investigation into the capacity of the CD-SVD method for detecting β-sheet structures on modular capsomeres may be necessary to widen the applicability of the CD-SVD method in characterising the structure of antigenic

modules on modular capsomeres or VLPs. This would further facilitate the development of modular VLP-based vaccines.

4. The findings in Chapter 4 suggest that the structure of modularised HA2A on modular capsomeres might be identical to the HA2A epitope on full-length HA. This leads naturally to the next research question — whether the structural presentation of HA2A antigenic elements arrayed in tandem repeats can confer on modular VLPs the ability to elicit protective antibodies. The capacity of HA2A epitopes to elicit protective antibodies remains unclear. Some studies have shown that the HA2A contributes to the epitope conformation of many protective antibodies (Corti *et al.*, 2011; Ekiert *et al.*, 2011; Sui *et al.*, 2009). Another study has suggested that HA2A alone is sufficient to induce neutralising antibodies (Ekiert *et al.*, 2009). However, Schneeman *et al.* presented HA2A on Flock House virus (FHV) VLP and found that modular FHV VLP of HA2A did not induce neutralising antibodies against influenza virus (Schneeman *et al.*, 2012). A possible reason for this is that the HA2A presentation in this study was inadequate to induce neutralising antibodies, suggesting a necessity for re-design of HA2A element. Further investigation into the capacity of the HA2A epitope to elicit protective antibodies is required.

5. Chapter 5 in this PhD thesis reported the initial steps in X-ray crystallography to determine the atomic structure of antigenic modules on modular capsomeres. The findings in this chapter reveal the challenge in obtaining optimum crystals for solving antigenic module structures. To increase the chances of achieving high-quality crystals, the purification process could be further improved and optimised to increase capsomere yield and purity. The use of HIC to remove DnaK, reported in chapter 5, sacrificed about 40% of the modular capsomeres. Furthermore, after HIC the modular capsomeres were still contaminated with trace amounts of DnaK. An alternative method for eliminating DnaK from modular capsomeres, therefore, requires to be further investigated, e.g. the use of ATP to dissociate the DnaK-modular capsomere complex (Amersham Pharmacia Biotech, 2001), or the use of MgATP plus denatured *E. coli* proteins to wash out DnaK contamination (Rial and Ceccarelli, 2002).

   Besides the yield and purity, the polydispersity of modular capsomeres also results in reduced chances of obtaining a crystal in crystallisation screening. Polydispersity can result from intermolecular interactions, i.e. the charge or hydrophobicity of modular capsomeres causes them to form dimer-capsomeres or trimer-capsomeres. Buffer components (pH, salt concentration) can help to prevent these intermolecular

interactions among modular capsomeres. Future work can investigate a high throughput method, e.g. dynamic light scattering (DLS), to screen buffer components that will maintain the monodispersity of modular capsomeres.

6. An alternative method to determine the structure of the antigenic module is cryo-electron microscopy (cryo-EM), an electron imaging technique for visualising three-dimensional (3D) structural details. Many studies have employed cryo-EM to obtain the 3D structure of VLPs (Joshi *et al.*, 2011; Peyret *et al.*, 2015) or modular VLP (Schneeman *et al.*, 2012). Cryo-EM requires a smaller amount of protein as material compared with X-ray crystallography. This suggests that cryo-EM would be a potentially viable technique for obtaining the structure of antigenic modules on modular VLPs.

6.3. Concluding thought

Recent approaches to vaccine development have moved towards highly-purified, well-characterised antigens rather than whole organisms (Thomas and Luxon, 2013). The use of modular VLP, presenting antigenic epitopes from unrelated pathogens on their surface, is a promising approach in vaccine development. Modular VLPs have also been utilised to induce an immune response against presented epitopes (Chuan *et al.*, 2014; Dale *et al.*, 2002; Jain *et al.*, 2010; Liu *et al.*, 2000; Tissot *et al.*, 2010; Wibowo *et al.*, 2012). The presentation of an epitope in its native structure on the surface of VLP is a challenge. There is a necessity to develop tools for the characterisation of antigenic module structures on modular capsomeres or VLPs. The study in this PhD thesis has developed a CD-SVD method, a simple and rapid tool for characterising antigenic module conformations, particularly helical structure, on modular capsomeres. The findings show that the CD-SVD method can allow the rapid identification of potential vaccine candidates in the early stages of vaccine development and therefore, reduce the costs of pre-clinical and clinical trials.

References

Amersham Pharmacia Biotech, 2001. The recombinant protein handbook 18, 1142-1175.

Anggraeni, M.R., Connors, N.K., Wu, Y., Chuan, Y.P., Lua, L.H.L., Middelberg, A.P.J., 2013. Sensitivity of immune response quality to influenza helix 190 antigen structure displayed on a modular virus-like particle. Vaccine 31, 4428-4435.

Baneyx, F., Mujacic, M., 2004. Recombinant protein folding and misfolding in *Escherichia coli*. Nat Biotechnol 22, 1399-1408.

Bartlett, P., 2000. The effect of polydispersity on colloidal phase transitions, in: Buckin, V. (Ed.), Trends in colloid and interface science XIV. Springer Berlin Heidelberg, pp. 137-140.

Chuan, Y.P., Wibowo, N., Connors, N.K., Wu, Y., Hughes, F.K., Batzloff, M.R., Lua, L.H.L., Middelberg, A.P.J., 2014. Microbially synthesized modular virus-like particles and capsomeres displaying group A *streptococcus* hypervariable antigenic determinants. Biotechnol Bioeng 111, 1062-1070.

Corti, D., Voss, J., Gamblin, S.J., Codoni, G., Macagno, A., Jarrossay, D., Vachieri, S.G., Pinna, D., Minola, A., Vanzetta, F., Silacci, C., Fernandez-Rodriguez, B.M., Agatic, G., Bianchi, S., Giacchetto-Sasselli, I., Calder, L., Sallusto, F., Collins, P., Haire, L.F., Temperton, N., Langedijk, J.P.M., Skehel, J.J., Lanzavecchia, A., 2011. A neutralizing antibody selected from plasma cells that binds to group and group 2 influenza A hemagglutinins. Science 333, 850-856.

Dale, C.J., Liu, X.S., De Rose, R., Purcell, D.F.J., Anderson, J., Xu, Y., Leggatt, G.R., Frazer, I.H., Kent, S.J., 2002. Chimeric human papilloma virus-simian/human immunodeficiency virus virus-like particle vaccines: Immunogenicity and protective efficacy in macaques. Virology 301, 176-187.

Ekiert, D.C., Bhabha, G., Elisliger, M.A., Friesen, R.H.E., Jongeneelen, M., Throsby, M., Goudsmit, J., Wilson, I.A., 2009. Antibody recognition of a highly conserved influenza virus epitope. Science 324, 246-251.

Ekiert, D.C., Friesen, R.H.E., Bhabha, G., Kwaks, T., Jongeneelen, M., Yu, W., Ophorst, C., Cox, F., Korse, H.J.W.M., Brandenburg, B., Vogels, R., Brakenhoff, J.P.J., Kompier, R., Koldijk, M.H., Cornelissen, L., Poon, L.L.M., Peiris, M., Koudstaal, W., Wilson, I.A., Goudsmit, J., 2011. A highly conserved neutralizing epitope on group 2 influenza A viruses. Science 333, 843-850.

Fan, Y.Y., Middelberg, A.P.J., 2011. A method of preparation of a biological particulate structure. University of Queensland, Australia.

Francis, D.M., Page, R., 2010. Strategies to optimize protein expression in *E. coli*. Curr Protoc Protein Sci Chapter 5, Unit 5.24.1-5.24.29. DOI: 10.1002/0471140864.ps0524s61.

Inoue, Y., Kubota-Koketsu, R., Yamashita, A., Nishimura, M., Ideno, S., Ono, K.-i., Okuno, Y., Ikuta, K., 2012. Induction of anti-influenza immunity by modified GFP carrying hemagglutinin-derived epitope structure. J Biol Chem 288, 4981-4990.

Jain, S., Patrick, A.J., Rosenthal, K.L., 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, 7070-7080.

Joshi, H., Cheluvaraja, S., Somogyi, E., Brown, D.R., Ortoleva, P., 2011. A molecular dynamics study of loop fluctuation in human papillomavirus type 16 virus-like particles: a possible indicator of immunogenicity. Vaccine 29, 9423-9430.

Liu, W.J., Liu, X.S., Zhao, K.N., Leggatt, G.R., Frazer, I.H., 2000. Papillomavirus virus-like particles for the delivery of multiple cytotoxic T cell epitopes. Virology 273, 374-382.

Lua, L.H.L., Fan, Y., Chang, C., Connors, N.K., Middelberg, A.P.J., 2015. Synthetic biology design to display an 18 kDa rotavirus large antigen on a modular virus-like particle. Vaccine 33, 5937-5944.

Pantophlet, R., Burton, D.R., 2006. GP120: Target for neutralizing HIV-1 antibodies. Annu Rev Immunol 24, 739-769.

Peyret, H., Gehin, A., Thuenemann, E.C., Blond, D., El Turabi, A., Beales, L., Clarke, D., Gilbert, R.J.C., Fry, E.E., Stuart, D.I., Holmes, K., Stonehouse, N.J., Whelan, M., Rosenberg, W., Lomonossoff, G.P., Rowlands, D.J., 2015. Tandem fusion of Hepatitis B core antigen allows assembly of virus-like particles in bacteria and plants with enhanced capacity to accommodate foreign proteins. PLoS One 10, e0120751.

Rial, D.V., Ceccarelli, E.A., 2002. Removal of DnaK contamination during fusion protein purifications. Protein Expr Purif 25, 503-507.

Rüdiger, S., Germeroth, L., Schneider-Mergener, J., Bukau, B., 1997. Substrate specificity of the DnaK chaperone determined by screening cellulose-bound peptide libraries. Embo J 16, 1501-1507.

Schneeman, A., Speir, J.A., Tan, G.S., Khayat, R., Ekiert, D.C., Matsuoka, Y., Wilson, I.A., 2012. A virus-like particle that elicits cross-reactive antibodies to the conserved stem of influenza virus hemagglutinin. J Virol 86, 11686-11697.

Sreerama, N., Woody, R.W., 2000. Estimation of protein secondary structure from circular dichroism spectra: Comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. Anal Biochem 287, 252-260.

Stehle, T., Harrison, S.C., 1997. High-resolution structure of a polyomavirus VP1-oligosaccharide complex: implications for assembly and receptor binding. Embo J 16, 5139-5148.

Sui, J., Hwang, W.C., Perez, S., Wei, G., Aird, D., Chen, L., Santelli, E., Stec, B., Cadwell, G., Ali, M., Wan, H., Murakami, A., Yammanuru, A., Han, T., Cox, N.J., Bankston, L.A., Donis, R.O., Liddington, R.C., Marasco, W.A., 2009. Structural and functional bases for broad-spectrum neutralization of avian and human influenza A viruses. Nat Struct Mol Biol 16, 265-272.

Thomas, S., Luxon, B.A., 2013. Vaccines based on structure-based design provide protection against infectious diseases. Expert Rev Vaccines 12, 1301-1311.

Tissot, A.C., Renhofa, R., Schmitz, N., Cielens, I., Meijerink, E., Ose, V., Jennings, G.T., Saudan, P., Pumpens, P., Bachmann, M.F., 2010. Versatile virus-like particle carrier for epitope based vaccines. PLoS One 5, e9809.

Tolia, N.H., Joshua-Tor, L., 2006. Strategies for protein coexpression in *Escherichia coli*. Nat Methods 3, 55-64.

Wibowo, N., Chuan, Y.P., Lua, L.H.L., Middelberg, A.P.J., 2012. Modular engineering of a microbially-produced viral capsomere vaccine for influenza. Chem Eng Sci 103, 12-20.

Appendices

Appendix 1: Synthesis of MuPyV VP1 capsomere variants.

The VP1-ΔC capsomere was synthesised and provided by Yang Wu. VP1 (named as Wt-VP1) and VP1-ΔNΔC (Truncated VP1) were presented in Section 5.3.1 in Chapter 5.

Appendix 2: Synthesis of modular HA2A VLPs.

Modular HA2A constructs were expressed in *E.coli* as GST fusion protein. Figure 1 indicated that the insertion of HA2A module did not adversely influence expression and solubility of VP1. GST fused modular HA2A capsomeres were purified using GST-affinity chromatography and polished using size exclusion chromatography (SEC).

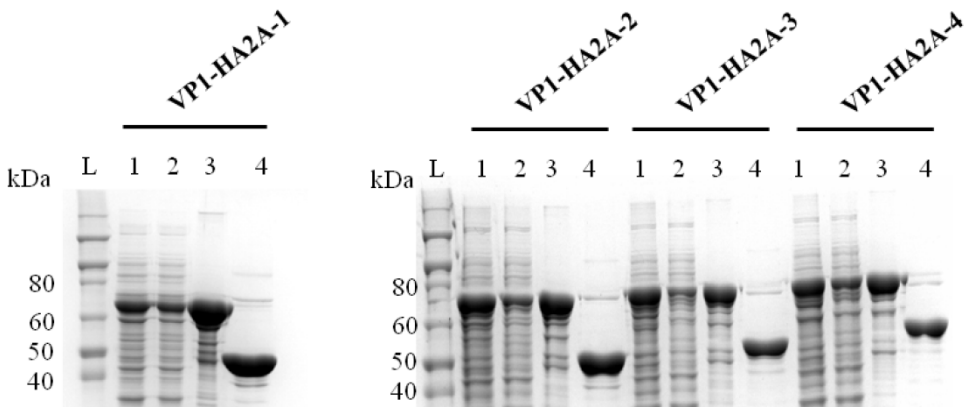


Figure 1: A) SDS-PAGE analysis showing the solubility of Cap-Base B and CapH190-2 in comparison with VP1 WTP. Lanes: (L) molecular weight marker; (1) Total cell lysate; (2) Soluble fraction of cell lysate; (3) GST fraction and (4) S200 fraction.

The modular HA2A capsomeres were assembled *in vitro* into modular VLPs. Modular HA2A VLPs were characterised using asymmetrical flow field-flow fractionation (AF4) and TEM analyses (Figure 2). The AF4 peak at approximately 22 min corresponded to VLPs with averaged root-mean-square (rms) radius of 20 nm. TEM visualisation showed that the modular HA2A VLPs have a similar size distribution and morphology with the Wt-VP1 VLPs reported previously (Chuan *et al.*, 2008; Middelberg *et al.*, 2011).

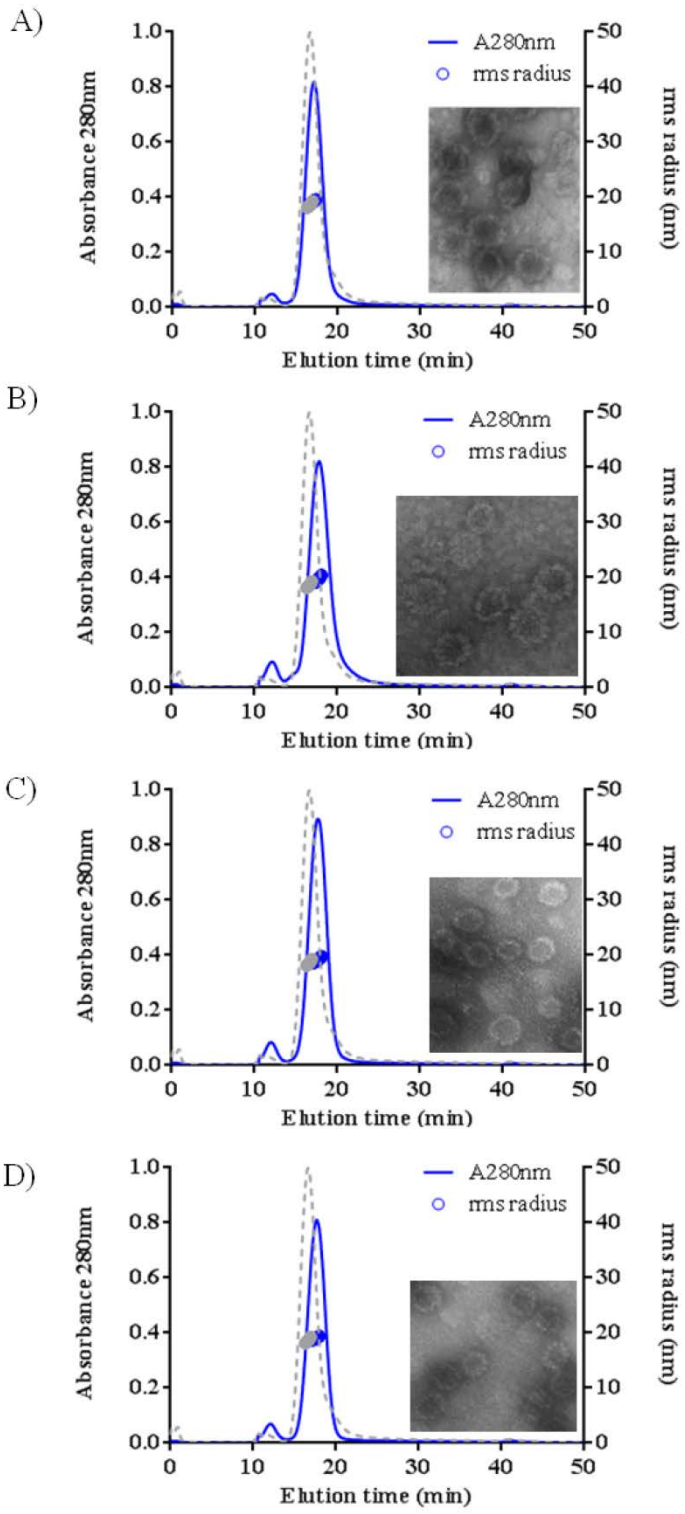


*Figure 2: Characterization of modular VLPs using AF4 (left pannels) and TEM (right panels)**.** A) VP1-HA2A-1, B) VP1-HA2A-2, VP1-HA2A-3 and VP1-HA2A-4.*

Appendix 3: The ratios of α-helix structure estimated from CDSSTR and translated helical amino acid (AA) numbers of VP1, VP1-ΔC and VP1-ΔNΔC.

Samples	CD reading	CDSSTR analysis	Ratio of α-helix	Total AAs	Number of helical AAs	Averaged helical AAs for each CD reading
VP1	I	1	0.09	384	35	35
		2	0.10		38	
		3	0.09		35	
		4	0.09		35	
		5	0.09		35	
	II	1	0.09	384	35	34
		2	0.09		35	
		3	0.09		35	
		4	0.09		35	
		5	0.08		31	
	III	1	0.09	384	35	34
		2	0.09		35	
		3	0.09		35	
		4	0.09		35	
		5	0.08		31	
	IV	1	0.08	384	31	36
		2	0.09		35	
		3	0.13		50	
		4	0.09		35	

		5	0.08		31	
	V	1 2 3 4 5	0.09 0.09 0.09 0.09 0.09	384	35 35 35 35 35	35
VP1-ΔC	I	1 2 3 4 5	0.10 0.10 0.10 0.10 0.10	321	32 32 32 32 32	32
	II	1 2 3 4 5	0.10 0.10 0.10 0.10 0.10	321	32 32 32 32 32	32
	III	1 2 3 4 5	0.10 0.10 0.10 0.10 0.10	321	32 32 32 32 32	32
	IV	1 2 3	0.09 0.10 0.10	321	29 32 32	31

		4	0.10		32	
		5	0.10		32	
	V	1	0.10	321	32	32
		2	0.10		32	
		3	0.10		32	
		4	0.10		32	
		5	0.10		32	
VP1-ΔNΔC	I	1	0.11	291	32	33
		2	0.12		35	
		3	0.11		32	
		4	0.11		32	
		5	0.11		32	
	II	1	0.11	291	32	32
		2	0.11		32	
		3	0.11		32	
		4	0.11		32	
		5	0.11		32	
	III	1	0.12	291	35	34
		2	0.12		35	
		3	0.11		32	
		4	0.12		35	
		5	0.11		32	
	IV	1	0.11	291	32	32
		2	0.11		32	

		3	0.11		32	
		4	0.11		32	
		5	0.11		32	
	V	1	0.11	291	32	33
		2	0.12		35	
		3	0.11		32	
		4	0.11		32	
		5	0.11		32	

The ratio of the helical structure was obtained from CDSSTR program on Dichroweb. The translation is implemented by the formula:

Number of the helical amino acid = ratio of helical structure x Total amino acid

Appendix 4: The ratios of α-helix structure estimated from CDSSTR and translated helical amino acid numbers of base capsomere and modular HA2A capsomeres.

Samples	CD reading	CDSSTR analysis	Ratio of α-helix	Total AAs	Number of helical AAs	Averaged helical AAs for each CD reading
Unmodified	I	1	0.11	291	32	33
		2	0.12		35	
		3	0.11		32	
		4	0.11		32	
		5	0.11		32	
	II	1	0.11	291	32	32
		2	0.11		32	
		3	0.11		32	
		4	0.11		32	
		5	0.11		32	
	III	1	0.12	291	35	34
		2	0.12		35	
		3	0.11		32	
		4	0.12		35	
		5	0.11		32	
	IV	1	0.11	291	32	32
		2	0.11		32	
		3	0.11		32	

		4 5	0.11 0.11		32 32	
	V	1 2 3 4 5	0.11 0.12 0.11 0.11 0.11	291	32 35 32 32 32	33
VP1-ΔNΔC-HA2A-1	I	1 2 3 4 5	0.09 0.10 0.10 0.11 0.09	321	29 32 32 35 29	31
	II	1 2 3 4 5	0.09 0.10 0.11 0.11 0.10	321	29 32 35 35 32	33
	III	1 2 3 4 5	0.09 0.10 0.09 0.10 0.10	321	29 32 29 32 32	31
	IV	1 2	0.09 0.09	321	29 29	29

		3 4 5	0.09 0.09 0.09		29 29 29	
	V	1 2 3 4 5	0.10 0.11 0.11 0.10 0.11	321	32 35 35 32 35	34
VP1-ΔNΔC-HA2A -2	I	1 2 3 4 5	0.12 0.12 0.12 0.12 0.11	347	42 42 42 42 38	41
	II	1 2 3 4 5	0.12 0.12 0.13 0.13 0.12	347	42 42 45 45 42	43
	III	1 2 3 4 5	0.12 0.12 0.12 0.12 0.12	347	42 42 42 42 42	42
	IV	1	0.12	347	42	42

		2	0.12		42	
		3	0.12		42	
		4	0.12		42	
		5	0.12		42	
	V	1	0.12	347	42	42
		2	0.12		42	
		3	0.13		45	
		4	0.12		42	
		5	0.11		38	
VP1-ΔNΔC-HA2A -3	I	1	0.11	373	41	43
		2	0.11		41	
		3	0.12		45	
		4	0.12		45	
		5	0.12		45	
	II	1	0.13	373	48	47
		2	0.12		45	
		3	0.13		48	
		4	0.12		45	
		5	0.13		48	
	III	1	0.11	373	41	42
		2	0.11		41	
		3	0.12		45	
		4	0.11		41	
		5	0.11		41	

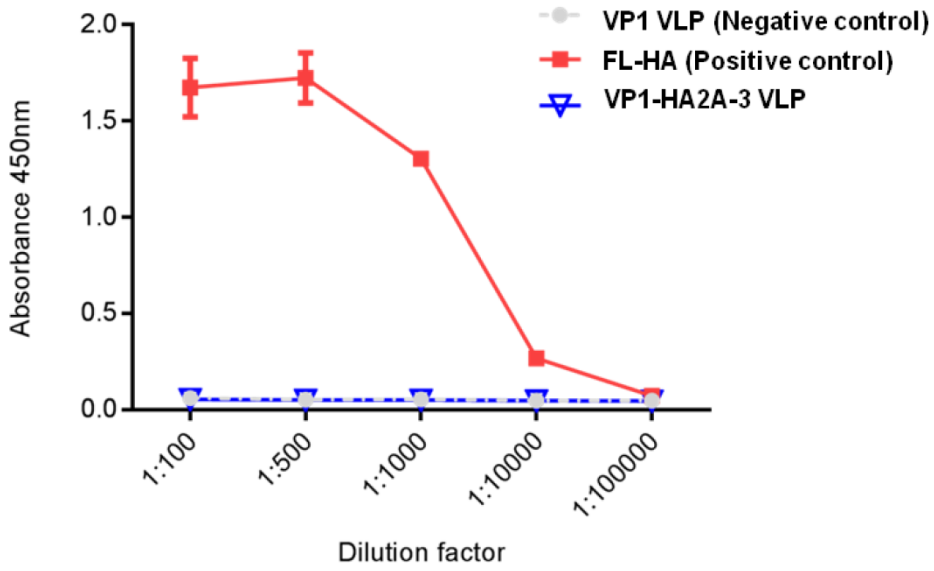
	IV	1 2 3 4 5	0.11 0.11 0.12 0.11 0.12	373	41 41 45 41 45	43
	V	1 2 3 4 5	0.12 0.12 0.12 0.12 0.12	373	45 45 45 45 45	45
VP1-ΔNΔC-HA2A -4	I	1 2 3 4 5	0.14 0.14 0.14 0.14 0.14	399	56 56 56 56 56	56
	II	1 2 3 4 5	0.14 0.13 0.15 0.14 0.14	399	56 52 60 56 56	56
	III	1 2 3 4	0.13 0.14 0.14 0.14	399	52 56 56 56	56

		5	0.15		60	
	IV	1	0.14	399	56	56
		2	0.14		56	
		3	0.14		56	
		4	0.14		56	
		5	0.14		56	
	V	1	0.14	399	56	56
		2	0.13		52	
		3	0.15		60	
		4	0.14		56	
		5	0.14		56	

The ratio of the helical structure was obtained from CDSSTR program on Dichroweb. The translation is implemented by the formula:

Number of the helical amino acid = ratio of helical structure x Total amino acid

Appendix 5: Indirect ELISA of VP1-HA2A-3 VLP against CR6261 antibody.



References

Chuan, Y.P., Fan, Y.Y., Lua, L.H.L., Middelberg, A.P.J., 2008. Quantitative analysis of virus-like particles size and distribution by field-flow fraction. Biotechnol Bioeng 99, 1425-1433.

Middelberg, A.P.J., Rivera-Hernandez, T., Wibowo, N., Lua, L.H.L., Fan, Y., Magor, G., Chang, C., Chuan, Y.P., Good, M.F., Batzloff, M.R., 2011. A microbial platform for rapid and low-cost virus-like particle and capsomere vaccines. Vaccine 29, 7154-7162.