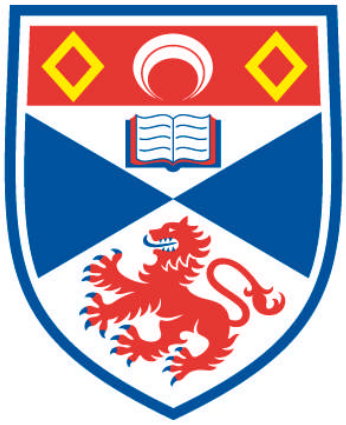
MULTIVALENT SIALIC ACID BINDING PROTEINS AS NOVEL THERAPEUTICS FOR INFLUENZA AND PARAINFLUENZA INFECTION

Nadiawati Alias

A Thesis Submitted for the Degree of PhD
at the
University of St Andrews

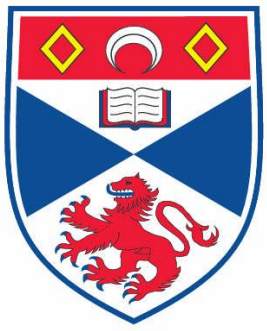


2014

Multivalent sialic acid binding proteins as novel therapeutics for influenza and parainfluenza infection

Nadiawati Alias

A thesis submitted for the degree of PhD
at the
University of St. Andrews



Supervisor: Prof Garry Taylor

2013

Multivalent sialic acid binding proteins as novel therapeutics for influenza and parainfluenza infection

Nadiawati Alias

A thesis submitted for the degree of Doctor of Philosophy

School of Biology
University of St Andrews
December 2013

Declaration

I, Nadiawati Alias, certify that this thesis, which is approximately 44,000 words in length has been written by me, is the record of work carried out by myself and that it has not been submitted in any previous application for a higher degree.

Date: Signature of the candidate:

I was admitted as a research student in October 2009 and as a candidate for the degree of Ph. D in July 2010; the higher study for which this is a record was carried out in the University of St Andrews between 2009 and 2013.

Date: Signature of the candidate:

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of Ph.D in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

Date: Signature of supervisor:

In submitting this thesis to the University of St Andrews we understand that we are giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. We also understand that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that my thesis will be electronically accessible for personal or research use unless exempt by award of an embargo as requested below, and that the library has the right to migrate my thesis into new electronic forms as required to ensure continued access to the thesis. We have obtained any third-party copyright permissions that may be required in order to allow such access and migration, or have requested the appropriate embargo below.

The following is an agreed request by candidate and supervisor regarding the electronic publication of this thesis: Access to Printed copy and electronic publication of thesis through the University of St Andrews.

Date: Signature of the candidate: Signature of supervisor:

Dedicated to
My Husband and Son, Ar-Rayyan

Abstract

In nature, proteins with weak binding affinity often use a multivalency approach to enhance protein affinity via an avidity effect. Interested in this multivalency approach, we have isolated a carbohydrate binding module (CBM) that recognises sialic acid (known as a CBM40 domain) from both *Vibrio cholerae* (Vc) and *Streptococcus pneumoniae* (Sp) *NanA* sialidases, and generated multivalent polypeptides from them using molecular biology. Multivalent CBM40 constructs were designed either using a tandem repeat approach to produce trimeric or tetrameric forms that we call Vc3CBM and Vc4CBM, respectively, or through the addition of a trimerization domain derived from *Pseudomonas aeruginosa* pseudaminidase to produce three trimeric forms of proteins known as Vc-CBMTD (WT), Vc-CBMTD (Mutant) and Sp-CBMTD). Due to the position and flexibility of the linker between the trimerization domain and the CBM40 domain, site directed mutagenesis was employed to introduce a disulphide bond between the monomers at positions S164C and T83C of the CBM40 domain in order to promote a stable orientation of the binding site for easier access of sialic acids.

Data from isothermal titration calorimetry (ITC) reveals that interaction of multivalent CBM40 proteins with α(2,3)-sialyllactose was mainly enthalpy driven with entropy contributing unfavorably to the interaction suggesting that these proteins establish a strong binding affinity to their ligand minimizing dissociation to produce stable multivalent molecules. However, using surface plasmon resonance (SPR), a mixed balance of entropy and enthalpy contributions was found with all constructs as determined by Van't Hoff plots. This proved that binding does not occur through a simple protein-ligand interaction but through disruption of hydrophobic and/or ionic hydration that provide the driving force to the process. Interestingly, the valency of

multiple-linked polypeptides also plays an important part in the protein stabilization. However, little is known about their detailed structure when in multivalent form, as attempts to crystallize the whole protein molecule of Vc-CBMTD (WT) failed due to linker and domain flexibility. Only the trimerization domain (TD) part from *Pseudomonas aeruginosa* pseudaminidase was successfully crystallized and structure was determined to 3.0 Å without its CBM40 domain attached.

In this thesis, we have also reported on the potential anti-influenza and anti-parainfluenza properties of these proteins, which were found to block attachment and inhibit infection of several influenza A and parainfluenza virus strains *in vitro*. As widely mentioned in literature, terminal sialic acids on the cell surface of mammalian host tissue provide a target for various pathogenic organisms to bind. Levels of viral inhibition were greatest against A/Udorn/72 H3N2 virus for Vc4CBM and Vc3CBM constructs with the lowest EC_{50} of 0.59 µM and 0.94 µM respectively, however most of the multivalent proteins tested were also effective against A/WSN/33 H1N1 and A/PR8/34 H1N1 subtypes. For parainfluenza virus, all constructs containing *V. cholerae* sialidase CBM40 domain showed great effect in inhibiting virus infection during cell protection assay. The best EC_{50} values were 0.2 µM from Vc-CBMTD (WT) followed by 1.17 µM from Vc4CBM and 1.78 µM from Vc-CBMTD (Mutant) which was against hPIV2, hPIV3 and hPIV5 infections respectively. Only a construct from *S. pneumoniae* sialidase known as Sp-CBMTD showed negligible effect on cell protection. All constructs were further tested for cytotoxicity in mammalian cell culture as well as undergoing an inhibition study on viral replication proteins. For the *in vivo* study, we also demonstrated the effectiveness of Vc4CBM to protect cotton rats and mice from hPIV3 and *Streptococcus pneumoniae* infections, when given intranasally in advance or on the day of infection. Therefore, these novel multivalent proteins could be promising candidates as broad-spectrum inhibitors or as a prophylactic treatment for both influenza and parainfluenza associated diseases.

Acknowledgements

First and foremost, I would like to express my gratefulness and humbleness to the Almighty for helping me all the way through my studies and final submission of my thesis.

I would like to thank my supervisor Prof. Garry Taylor, for all his guidance, support and encouragement to me during my Ph.D in St Andrews. I also want to express my sincere thanks to Dr. Helen Connaris, for her constant guidance, help, inspiration, and constructive criticisms that encouraged me throughout this research work. It was always a pleasure to discuss scientific and non-scientific issues with her.

I wish to thank Prof. Richard Randall's group for their kind help in the mammalian tissue culture works, Dr. Catherine Botting and Dr. Sally Shirran for their support with the mass spectrometry and Dr. Uli Schwarz-Linek for his help on ITC works. I am also very grateful to the members of GLT group; Dr. Jane Potter who taught me so many lab techniques, Dr Sandra Tauber, Dr Magnus Alphey, Dr Margaret Taylor, Judith Telford, Lei Yang, Valeria Fadda and David Owen for their help and co-operation during my research.

My appreciation and many thanks go to my parents, family members and friends, Dr Mazlina Musa, Dr Mariana Aida and Mrs Juzainita Mohamad for their abundance of love, inspiration and support as I pursued my goal. Special thanks and appreciation also go to my husband and my son for their selfless love, time and encouragement during my study. To them I dedicate this thesis.

Not to forget, my sincerely appreciation to the financial supporter, Ministry of Higher Education Malaysia and Universiti Sultan Zainal Abidin Malaysia for the opportunity and trust given to undergo this PhD journey.

Abbreviations

I. Amino acids

Ala, A	Alanine	Leu, L	Leucine
Arg, R	Arginine	Lys, K	Lysine
Asn, N	Asparagine	Met, M	Methionine
Asp, D	Aspartate	Phe, F	Phenylalanine
Cys, C	Cystine	Pro, P	Proline
Gln, Q	Glutamine	Ser, S	Serine
Glu, E	Glutamate	Thr, T	Threonine
Gly, G	Glycine	Trp, W	Tryptophan
His, H	Histidine	Tyr, Y	Tyrosine
Ile, I	Isoleucine	Val, V	Valine

II. Chemicals

3'SL	α2,3-sialyllactose, Neu5Ac(α2,3)Gal(β1-4)Glc
6'SL	α2,6-sialyllactose Neu5Ac(α2,6)Gal(β1-4)Glc
BSA	Bovine serum albumin
DDM	*n*-dodecyl-β-D-maltoside
DMEM	Dulbecco's Modified Eagle Medium
DMSO	Dimethyl sulfoxide
DNAse	Deoxyribonuclease
DTT	Dithiothreitol
DSLNT	Disialyllacto-N-tetraose
DSL	Disialyllactose
EDTA	Ethylenediaminetetraacetic acid
FCS	Fetal calf serum
Gal	Galactose
GalNAc	N-acetylgalactosamine
GlcNAc	N-acetylglucosamine
GFP	Green fluorescent protein
GM1	Monosialotetrahexosylganglioside
IPTG	Isopropyl thio-β-D-galactoside
LB	Luria Bertani
NAT	N-acetyl trypsin

Neu5Ac	N-acetylneuraminic acid, NANA
PBS	Phosphate buffered saline
PEG	Polyethylene glycol
PSA	Poly α2,8-sialic acid, ploySia, colominic acid
TEV	Tobacco etch virus

III. Sialidases

HN	Haemagglutinin-neuraminidase
IT *trans*-sialidase	Intramolecular *trans*-sialidase
NA	Influenza virus sialidase
NanA	*S. penumoniae* sialidase A
NanB	*S. penumoniae* sialidase B
NanC	*S. penumoniae* sialidase C
NanJ	*C. perfingens* large sialidase
NEU1	Human lysosomal sialidase
NEU2	Human cytosolic sialidase
NEU3	Human plasma membrane-associated sialidase
NEU4	Human mitochondrial sialidase
PaNA	*P. aeruginosa* sialidase
TcTS	*T. cruzi trans*-sialidase

IV. Symbols

a,b,c	Unit cell dimensions	ΔG	Gibbs free energy
Å	Angstrom	ΔS	Entropy
α, β, γ	Angles of unit cell	k_a	Association rate constant
Δ	Macroscopic change	k_D	Dissociation rate constant
δ	Chemical shift	K_D	Equilibrium dissociation constant
σ (*I*)	Error in intensity	$TCID_{50}$	median tissue culture infective dose
ε	Extinction coefficient	EC_{50}	Median effective concentration
λ	Wavelength		
μ	Micro (one millionth)		
T_m	Midpoint temperature		
ΔH	Enthalpy change		

V. Technical & Miscellaneous

3-D	Three dimensions
ATCC	American Type Culture Collection
a.u	Asymmetric unit
AUC	Analytical ultracentrifuge
bp	Base pair
BN-PAGE	Blue-Native polyacrylamide gel electrophoresis
CAZy	Carbohydrate-Active Enzymes
CBM	Carbohydrate binding module
CCP4	Collaborative Computing Project Number 4
Da	Dalton
DLS	Dynamic light scattering
ESRF	European Synchrotron Radiation Facility
HA	Haemagglutinin
I.N	Intranasal
Ig	Immunoglobulin
ITC	Isothermal titration calorimetry
LDL	Low density lipoprotein
LPS	Lipopolysaccharides
M	Molar concentration
MR	Molecular replacement
MS	Mass spectrometry
MW	Molecular weight
OD_{600}	Optical density at 600 nm
PCR	Polymerase chain reactions
PCT	Pre-crystallization test
PDB	Protein Data Bank
Pfu	Plaque-forming unit
p.i	Post infection
r.m.s.d	Root mean square deviation
r.p.m	Round per minute
SD	Standard deviation
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
TD	Trimerization domain
TNF	Tumor necrosis factor

Contents

Declaration i

Abstract ii

Acknowledgements iv

Abbreviations v

Contents viii

List of Figures xiv

List of Tables xxi

Chapter 1: Introduction

1.1 Diversity of sialic acid 1

1.2 Roles of sialic acid 4

1.3 The sialidase superfamily 5

1.4 Relationship between sialic acid and human health 8

1.5 Development of inhibitor against influenza and parainfluenza infections

1.5.1 Antiviral drugs against influenza viruses 14

1.5.2 Antiviral drugs against parainfluenza viruses 19

1.6 Carbohydrate binding module (CBM) 22

1.7 Family 40 CBM domain of bacterial sialidases

1.7.1 *Vibrio cholerae* sialidase 24

1.7.2 *Streptococcus pneumoniae* sialidases 26

1.8 Multivalency of CBMs 29

1.9 Project aims and thesis plan 31

Chapter 2: Development, expression and purification of multivalent CBM40s

2.1 Overview **33**

2.2 General methodology

2.2.1 Production and expression of protein constructs 35

2.2.2 Protein purifications protocol 36

2.3 Multiple tandem repeats of CBM40 Domain **37**

2.4 Production and expression of multivalent CBM40 construct

2.4.1 Expression and purification of Vc3CBM 39

2.4.2 Expression and purification of Vc4CBM 42

2.5 Engineering of multivalent CBM40 using an Oligomerization Domain (TD) from *P. aeruginosa* Pseudaminidase **44**

2.5.1 Development of Vc-CBMTD construct from *V. cholerae* sialidase 47

2.5.2 Site-directed mutagenesis of CBM40 50

2.5.3 Protein expression and purification of Vc-CBMTD (WT) and Vc-CBMTD (Mutant) 55

2.6 Development of Sp-CBMTD-GFP constructs using the CBM40 domain from *Streptococcus pneumoniae NanA* sialidase **61**

2.6.1 Amplification of *Streptococcus pneumoniae* CBM40 62

2.6.2 Expression and purification of Sp-CBMTD-GFP 64

2.7 Discussion **67**

Chapter 3: Biophysical characterization and crystallization of multivalent CBMs

3.1 Overview **71**

3.2 Study of protein conformation and aggregates **72**

3.3 Interaction profile of multivalent CBMs

3.3.1 Thermal Shift Assay: Screening for the best buffer 75

3.3.2 Thermal Shift Assay: Screening for the best ligands/ sialosides 84

3.4 Thermodynamic study of multivalent CBMs

3.4.1 Isothermal Titration Calorimetry (ITC) 98

3.4.2 Surface Plasmon Resonance (SPR) 105

3.5 Structural study of Vc-CBMTD

3.5.1 Protein crystallization and X-ray data collection 113

3.5.2 Crystal optimization 114

3.5.3 X-ray data collection, data processing, molecular replacement and structure refinement 116

3.5.4 Analysis of natively disordered region in full-length Vc-CBMTD 123

3.6 Structural study of Sp-CBMTD

3.6.1 Protein crystallization 125

3.7 Discussion 127

Chapter 4: *In vitro* study of influenza virus infection

4.1 Overview 132

4.2 Effect of multivalent CBMs in inhibiting influenza virus infections

4.2.1 Influenza virus plaque assay protocol 133

4.2.2 Assessment of protein anti-influenza activity 134

4.3 Effect of high concentration multivalent CBMs on cell monolayer

4.3.1 Evaluation of cytotoxic effects of Vc-CBMTD (WT) and Sp-CBMTD peptides 144

4.4 Inhibition study of viral replication protein

4.4.1 Viral inhibition assay 147

4.5 Discussion 151

Chapter 5: *In vitro* study of parainfluenza virus infection

5.1 Overview 154

5.2 Effect of multivalent CBMs in inhibiting parainfluenza virus Infections

5.2.1 Protocol of plaque inhibition assay 156

5.3 Study of different pre-incubation periods on Vc3CBM and Vc4CBM against hPIV3 157

5.4 Inhibitory effects of Vc3CBM and Vc4CBM on A549 cell by hPIV3 163

5.5 Anti-hPIV activity 166

5.6 Effect of high concentration multivalent CBMs on cell monolayer

5.6.1 Evaluation of cytotoxic effects of Vc-CBMTD (WT) and Vc-CBMTD (Mutant) peptides 175

5.6.2 Cell viability assay 177

5.7 Inhibition study of viral replication protein

5.7.1 Viral Inhibition Assay 180

5.8 Discussion 184

Chapter 6: *In vivo* study of Vc4CBM against hPIV3 and *Streptococcus pneumoniae*

6.1 Overview 186

6.2 Inhibition of hPIV3 infection in cotton rats 187

6.3 *In vivo* evaluation of Vc4CBM against acute pneumococcal infection 189

6.4 Discussion 195

Chapter 7: Summary and Future Work 197

Appendix A: Materials and methods

A-1. pEHISTEV vector map 200

A-2. Polymerase chain reaction (PCR) 201

A-3. Blue-Native PAGE staining protocol 201

A-4. Agarose gel electrophoresis 202

A-5. Restriction digestion of PCR products and vector 202

A-6. Ligation 202

A-7. Transformation 203

A-8. Colony PCR 203

A-9. DNA sequencing 204

A-10. Purification of transformed plasmid DNA 204

A-11. SDS-PAGE 204

A-12. Pre crystallisation test 205

Appendix B: The supplementary results **206**

B-1. SPR sensorgrams 206

References **211**

List of Figures

Figure	Title	Page
Figure 1.1	Structure of sialic acid also known as N-acetylneuraminic acid (Neu5Ac).	2
Figure 1.2	The most common linkages of Neu5Ac to other carbohydrates.	3
Figure 1.3	Replication cycle of influenza virus	9
Figure 1.4	Schematic illustration of parainfluenza life cycle	11
Figure 1.5	Chemical structures of adamantane based inhibitors, rimantidine and amantadine	15
Figure 1.6	Chemical structures of neuraminidase inhibitors, zanamivir and oseltamivir	17
Figure 1.7	Molecular model of Fludase (DAS181)	18
Figure 1.8	Chemical structure of haemagglutinin-neuraminidase inhibitor known as BCX2798	21
Figure 1.9	Schematic drawing of *Vibrio cholerae* sialidase.	
Figure 1.10	Schematic views of the three *S. pneumoniae* sialidases.	28
Figure 2.1	Schematic view of the binding of sialic acid to CBM40 domain of *V. cholerae* sialidase.	35
Figure 2.2	Illustration of tandem repeat approach for Vc3CBM and Vc4CBM constructs with their linker sequences.	39
Figure 2.3	Expression and purification of Vc3CBM.	41
Figure 2.4	Expression and purification of Vc4CBM.	43
Figure 2.5	View of the pseudaminidase monomer from *Pseudomonas aeruginosa*.	45
Figure 2.6	Schematic view of pseudaminidase trimer from	46

	Pseudomonas aeruginosa (2W38).	
Figure 2.7	PCR amplification of *P. aeruginosa* trimerization domain (PaTD) from postitive clones.	49
Figure 2.8	PCR amplification of CBM40 domain from *V. cholerae* sialidase.	49
Figure 2.9	Model of Vc-CBMTD (Mutant) shown as trimer.	52
Figure 2.10	Colony PCR amplification of Vc-CBMTD (WT) and Vc-CBMTD (Mutant) clones.	53
Figure 2.11	Amino acid alignments between wild type (WT) and mutant of Vc-CBMTD.	54
Figure 2.12	Expression of Vc-CBMTD (WT) and Vc-CBMTD (Mutant) in LB broth before optimization steps.	56
Figure 2.13	Soluble expression of Vc-CBMTD (WT) and Vc-CBMTD (mutant) after heat shock treatment.	57
Figure 2.14	Expression and purification of Vc-CBMTD (WT).	59
Figure 2.15	Expression and purification of Vc-CBMTD (Mutant).	60
Figure 2.16	PCR amplification of CBM40 and trimerization domain from positive clones.	63
Figure 2.17 (I)	Expression and Ni^{2+}-affinity column purification of Sp-CBMTD-GFP in *E. coli* BL21 (DE3).	65
Figue 2.17 (II)	Gel filtration chromatography of Sp-CBMTD-GFP.	66
Figure 2.18	Mass spectrometry analysis of the Sp-CBMTD clone	66
Figure 3.1	Migration of proteins by BlueNative gel electrophoresis.	74
Figure 3.2	Protein thermal stabilization measured by fluorescence readout in thermal shift assay.	76
Figure 3.3	Dissociation curve of Vc3CBM in 8 different buffers.	79
Figure 3.4	Dissociation curve of Vc4CBM in 8 different buffers.	80

Figure 3.5	Dissociation curve of Vc-CBMTD (WT) in 10 different buffers.	81
Figure 3.6	Dissociation curve of Vc-CBMTD (Mutant) in 10 different buffers.	82
Figure 3.7	Dissociation curve of Sp-CBMTD in 10 different buffers.	83
Figure 3.8	Comparison of T_m when reaction presence with protein alone and protein with ligand	85
Figure 3.9	Chemical structures and abbreviations of ligands used in this study.	86
Figure 3.10	Fluorescent readout of Vc3CBM interaction with 3'SL in two different buffers.	88
Figure 3.11	Fluorescent readout of Vc3CBM interaction with 6'SL in two different buffers.	89
Figure 3.12	Fluorescent readout of Vc4CBM interaction with 3'SL in two different buffers.	90
Figure 3.13	Fluorescent readout of Vc4CBM interaction with 6'SL in two different buffers.	91
Figure 3.14	Fluorescent readout of Vc-CBMTD (WT) interaction with 3'SL in two different buffers.	92
Figure 3.15	Fluorescent readout of Vc-CBMTD (WT) interaction with 6'SL in two different buffers.	93
Figure 3.16	Fluorescent readout of Vc-CBMTD (Mutant) interaction with 3'SL in two different buffers.	94
Figure 3.17	Fluorescent readout of Vc-CBMTD (Mutant) interaction with 6'SL in two different buffers.	95
Figure 3.18	Fluorescent readout of Sp-CBMTD interaction with 3'SL in two different buffers.	96

Figure 3.19	Fluorescent readout of Sp-CBMTD interaction with 6'SL in two different buffers.	97
Figure 3.20	ITC Interaction profile of multivalent constructs.	102
Figure 3.21	Summary of thermodynamic features for six different multivalents constructs with 3'SL.	103
Figure 3.22	Interpretation of binding response curve by SPR	107
Figure 3.23	van't Hoff plot from SPR data at three different temperatures (15, 25 and 37°C).	111
Figure 3.24	Diffraction pattern of Vc-CBMTD (WT) crystals formed before optimization conditions.	115
Figure 3.25	Vc-CBMTD (WT) crystals formed after optimization condition with its x-ray diffraction image.	117
Figure 3.26	Schematic drawing of trimerization domain (TD) of *P. aeruginosa* pseudaminidase from crystallography data.	120
Figure 3.27	Trimerization domain (TD) of *P. aeruginosa* pseudaminidase (PaNA) form as a trimer.	121
Figure 3.28	The electrostatic potential of molecular surface of TD domain calculated by PyMol.	122
Figure 3.29	Prediction of natively disordered region of full-length Vc-CBMTD (WT) by RONN bioinformatics server	124
Figure 3.30	Disorder prediction region of Sp-CBMTD protein predicted by RONN analysis	126
Figure 4.1	Anti-influenza activity of protein tested in plaque inhibition assays.	135
Figure 4.2	Plate representation of plaque inhibition assay on MDCK cells after infection with A/Udorn/72 H3N2.	136

Figure 4.3	Comparison of EC_{50} values of different multivalent proteins constructs against influenza A strains.	138
Figure 4.4	Comparison of different influenza A strains against EC_{50} of multivalent protein constructs.	139
Figure 4.5	Detail of EC_{50} graphs on multivalent protein constructs infected with A/Udorn/72 H3N2.	141
Figure 4.6	Detail of EC_{50} graphs of multivalent protein constructs infected with A/PR8/34 H1N1.	142
Figure 4.7	Detail of EC_{50} graphs of multivalent protein constructs infected with A/WSN/33 H1N1.	143
Figure 4.8	Viability test on Vc-CBMTD (WT).	145
Figure 4.9	Viability test on Sp-CBMTD.	145
Figure 4.10	Percentage of viral replication proteins after treatment with 1 mg/ml of multivalent proteins.	149
Figure 4.11	Effect of anti-influenza activity of Vc3CBM at different concentrations tested on MDCK cells.	150
Figure 4.12	Effect of anti-influenza activity of Vc4CBM at different concentrations tested on MDCK cells.	150
Figure 5.1	Percentage number of plaque compared to mock on Vero cell after hPIV3 infection.	158
Figure 5.2	Plaque assay of Vc3CBM on Vero cells after 3 days of hPIV3 infections.	158
Figure 5.3	Percentage number of plaque compared to mock on Vero cell after hPIV3 infection.	160
Figure 5.4	Plaque assay of Vc4CBM on Vero cells after 3 days of hPIV3 infections.	160
Figure 5.5	Vc3CBM-GFP attached to the surface of Vero cells after	162

	incubation for 1 hour 30 minutes.	
Figure 5.6	Percentage number of plaque compared to mock on A549 cell line after hPIV3 infection.	164
Figure 5.7	Effect of various concentrations of Vc4CBM on A549 cell line infected with hPIV3 after 5 days.	165
Figure 5.8	Detail of EC_{50} graphs on multivalent constructs infected with hPIV2.	169
Figure 5.9	Detail of EC_{50} graphs on multivalent constructs infected with hPIV3.	170
Figure 5.10	Detail of EC_{50} graphs on multivalent constructs infected with hPIV5.	171
Figure 5.11	Comparison of different parainfluenza viruses against EC_{50} of multivalent constructs.	173
Figure 5.12	Comparison of EC_{50} values of different multivalent constructs against hPIVs viruses.	174
Figure 5.13	Effect of high concentration of Vc-CBMTD (WT) on Vero cells after 18 hours post treatment.	176
Figure 5.14	Effect of high concentration of Vc-CBMTD (Mutant) on Vero cells after 18 hours post treatment.	176
Figure 5.15	Viability test on Vc-CBMTD (WT).	178
Figure 5.16	Viability test on Vc-CBMTD (Mutant).	178
Figure 5.17	Percentage of viral replication proteins after treatment with 2mg/ml of different multivalent proteins.	182
Figure 5.18	Effect of anti-hPIVs activity of Vc-CBMTD (WT) at different concentrations on Vero cells.	183
Figure 5.19	Effect of anti-hPIVs activity of Vc4CBM at different concentrations on Vero cells.	183

Figure 6.1	Comparison of $TCID_{50}$ of cotton rats treated with Vc4CBM and with the control (PBS only) group after infection with hPIV3 virus.	188
Figure 6.2	Percentage of mice survival rates after treatment with Vc4CBM proteins.	190
Figure 6.3	Survival times of mice infected with *S. pneumonia*e after treatment with Vc4CBM.	192
Figure 6.4	Bacterial counts from blood sampling at 24, 36 and 48 hours after treatment with Vc4CBM protein.	193
Figure 6.5	Disease sign scores of mice during Vc4CBM treatment at three different time points.	194

List of Tables

Table	Title	Page
Table 1.1	Relation between pathogens and diseases with virulence factors.	13
Table 2.1	List of oligonucleotides primers used to amplify DNA fragments.	38
Table 2.2	Summary of oligonucleotide primers used to amplify Vc-CBM and PaTD.	48
Table 3.1	ITC data on the binding of multivalent proteins to 3'-Sialyllactose.	103
Table 3.2	Biacore kinetic parameters of multivalent CBMs interaction with 3'-sialyllactose.	109
Table 3.3	Thermodynamic parameters for the interaction between multivalent proteins with 3'SL-PAA Biotin.	112
Table 3.4	X-ray data collection and refinement statistics.	118
Table 4.1	Summary of EC_{50} values of influenza A viruses treated with different multivalent CBM40 proteins.	137
Table 5.1	Summary of EC_{50} values (µM) of multivalent CBMs infected with different strains of hPIVs	168
Table 5.2	List of primary antibody used in viral inhibition assay.	181

Chapter 1

Introduction

1.1 Diversity of sialic acid

The term sialic acid is often known as N-acetyl-neuraminic acid (Neu5Ac), which is an O- and N-substituted derivative with a nine-carbon monosaccharide called neuraminic acid (Figure 1.1). This Neu5Ac is linked to galactose residue of carbohydrate chain through α(2,3)- or α(2,6)-linkages or α(2,8) to another sialic acid as shown in Figure 1.2. Sialic acid is a negatively charged molecule and is normally located at a distal end of glycan chains on the cell surface glycoproteins and glycolipids; and usually used for recognition purposes in the immune system (Schauer, 2004; Varki, 1997). It is evident that sialic acids can be recognized by a variety of lectins range from animals, plants and microorganism as well as by certain naturally occurring antibodies (Varki, 1997).

It was reported by Ulloa and Real (2001) that the α(2,3)-linked sialiosides are distributed more abundantly than the α(2,6)-linkages in the apical membrane of epithelial cells. On the other hand, the α(2,8)-linked sialic acid polymers are found in gangliosides and in glycoproteins of neural tissues (Schauer, 1982b) and also on the surface of some bacteria like *E. coli* K1 strain (Kleene and Schachner, 2004).

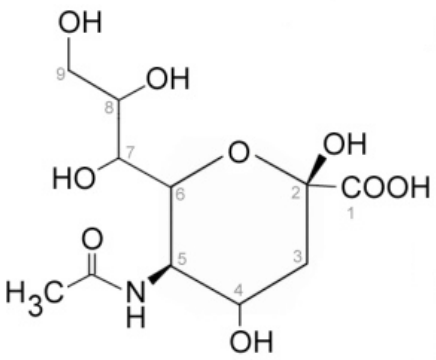


Figure 1.1 Structure of sialic acid also known as N-acetylneuraminic acid (Neu5Ac).

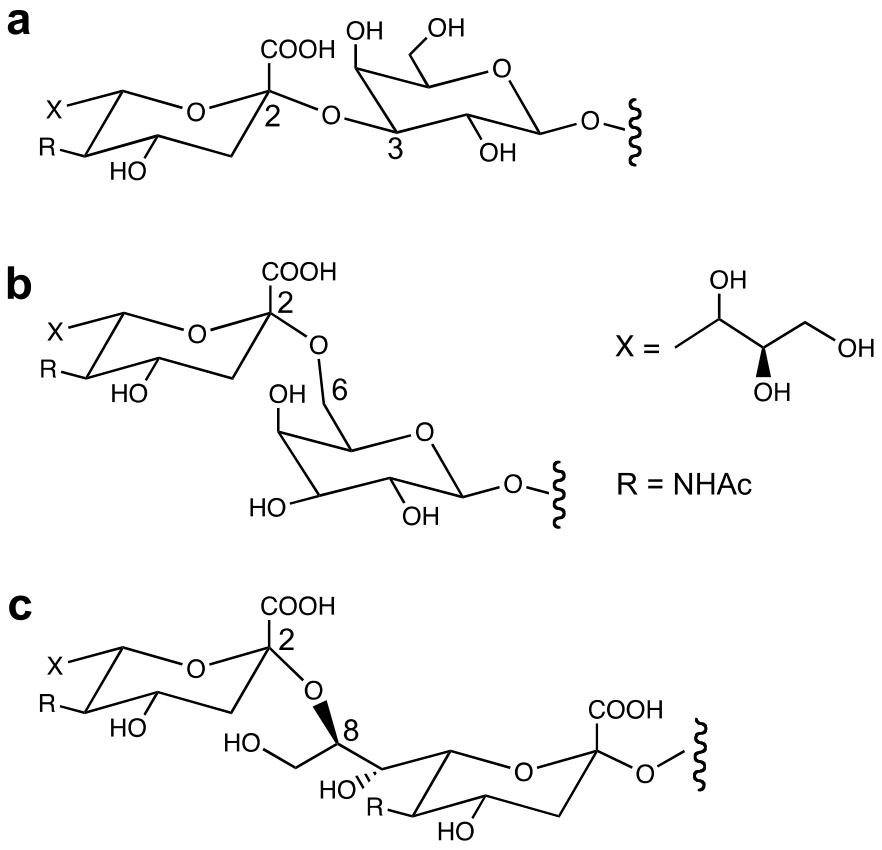


Figure 1.2 The most common linkages of Neu5Ac to other carbohydrates. (a) The α(2,3)-linkages of Neu5Ac to Gal, GalNAc or GlcNAc; **(b)** the α(2,6)-linkages of Neu5Ac to Gal or GalNAc; **(c)** the α(2,8)-linkages of Neu5Ac polymers.

1.2 Roles of sialic acid

Sialic acids are negatively charged substrates with the ability to repel cell-cell interactions but also function as ligands for variety of lectins as mentioned by (Varki, 1997). They are a fundamental component of many biological processes such as cell recognition, adhesion, defence, activation of cellular pathways, sialylation and desialylation events, and cell surface modification. For example, they have been involved in macrophage-lymphocyte interaction (Crocker et al., 1991), virus attachment (Feizi and Childs, 1985) and attachment of bacterial fimbriae to certain mucosal cells (Lindahl et al., 1987). Due to the presence of sialic acids at the termini of glycoconjugates on the cell surface, it can be exploited by various pathogens for recognition sites in order to enter the host cells (Lehmann et al., 2006).

As mentioned before, its negatively charged features help in binding and transporting positive charged residues including pharmaceutical molecules as well as attraction and repulsion of cells such as erythrocytes in the blood stream (Kelm and Schauer, 1997; Schauer, 2004; Schauer and Kamerling, 1997). Besides that, sialic acid also ensures the viscoelastic properties of sialic acid-rich molecules and its intramolecular interactions which help in their molecule shape and form such as mucins (Montreuil, 1982).

While in molecular physiology context, sialic acid can operate to modulate the half-life of some proteins (Ashwell and Harford, 1982; Weigel and Yik, 2002). Under pathological conditions, when sialic acids are missing on the surface of monosaccharides such as on the liver and other organs, it may cause some glycoproteins (biotherapeutic products) to be rapidly cleared away (Raju et al., 2001; Weigel and Yik, 2002).

1.3 The sialidase superfamily

Sialidases also known as neuraminidases are grouped into families of exoglycosidases which catalyze the hydrolysis of terminal sialic acid residues from various glycoconjugates such as glycoproteins and glycolipids (Schauer, 1985). Sialidases have been found widely from various organisms, including viruses, bacteriophage, bacteria, fungi, mycoplasma, protozoa and some higher eukaryotes.

Despite the remarkable diversities in their distributions, amino acids sequences and biological properties, the sialidases still contain significant molecular and structural homologies with each other, therefore are enclosed into the sialidase superfamily with EC 3.2.1.18 (Roggentin et al., 1989). Sialidases can be classified based on their origins or the functional properties as mentioned in this subsection. Moreover, sialidases can also be classified based on enzyme molecular weight (MWs) into two families with the large sialidases family with MWs over 60 kDa and the small sialidases family with MWs of 42 kDa (Roggentin et al., 1993).

One of the most widely studied sialidases is the influenza virus sialidase also known as neuraminidase (NA). This is due to the public health concern on the outbreak of influenza infections which caused, for example, the 1918 influenza pandemic. The role of influenza virus NA is to remove sialic acid from cell surface and progeny virions facilitating virus release from the infected cell. ls (Air and Laver, 1989; Matrosovich et al., 2004b). Related to this, NA is known as one of the virulence factors for influenza viruses whereas, for paramyxoviruses, heamagglutinin-neuraminidase (HN). This HN contains dual functions which recognizes sialic acid for cell attachment and also cleaves sialic acid during infections (Taylor, 1996). Interestingly, these enzymes were proved to be effective inhibitors for *in vitro*

influenza virus infection. It has been observed that when sialic acid was enzymatically removed from cell surface, the cells were less risk to infection by influenza viruses (Gottschalk, 1959).

In addition to the viral sialidases, a variety of sialidases have been identified from some pathogenic bacteria, such as *Vibrio cholerae*, *Clostridium perfringens*, *Streptococcus pneumoniae*, *Salmonella typhimurium* and *Pseudomonas aeruginosa* (Roggentin et al., 1993). In certain situations, this enzyme has a specific role in disease and known to be a virulence factors essential in bacterial colonization and infections (Corfield et al., 1992). There are various non-pathogenic and pathogenic sialidase producing bacteria that also use sialic acid as their carbon and energy source (Corfield et al., 1992). Some pathogenic bacteria are able to project the enzyme directly onto the host tissue and interfere with host immunologic mechanisms and defence system. This approach will unmask some receptors for bacteria adherence and colonization, for example exposure of the monosialotetrahexosylganglioside (GM1) receptor for toxin binding in *V. cholerae* pathogenesis (Galen et al., 1992; Snyder and Walker, 1987).

Another interesting group of sialidases is the mammalian sialidases. Since the first cloning of a mammalian sialidase in 1993, it boosted research on this particular area with 12 different mammalian sialidases successfully cloned and sequenced. Due to its different subcellular distribution, mammalian sialidases are classified into several subtypes such as lysosomal (NEU1), cytosolic (NEU2), plasma membrane-associated (NEU3) and mithocondrial/ ER-associated (NEU4) types (Miyagi and Yamaguchi, 2012; Monti et al., 2010; Monti et al., 2002). All of them are classified as exo-sialidases and share a significant sequence similarity with viral and microbial sialidases (Fanzani et al., 2012). These enzymes provide great importance in various cellular functions, including lysosomal catabolism in which microbial sialidases play roles in nutrition and pathogenesis (Miyagi and Yamaguchi, 2012).

Besides this, there are other unique sialidases known as *trans*-sialidases, which have been identified from parasites such as *Trypanosoma cruzi* (TcTS). This sialidase is different from other hydrolytic sialidases described above due to the parasite unable to synthesize sialic acids on its own, which are vital for its virulence in the human body (Agusti et al., 2007; Cross and Takle, 1993). In order to provide sialic acids, this protozoa expresses an enzyme called *trans*-sialidase (TcTS), which is attached to its outer membrane to crop sialic acids from the host body (Alvarez et al., 2004). Since this parasite does not synthesize its own sialic acid, this specific decoration with host sialic acid by TcTS might protect the protozoan parasite from the host immune system (Haselhorst et al., 2004). Sialidases and *trans*-sialidases and have been found in a few *Trypanosoma* species such as the sialidase in *Trypanosoma rangeli* which although it shares 70% sequence identity to the TcTS, is not a *trans*-sialidase (Buschiazzo et al., 2000).

Different from any other sialidases characterized so far, the sialidase from *Arthrobacter ureafaciens* has the unique property of cleaving the internal sialic acids residues from gangliotetraose (Iwamori et al., 1997). This sialidase was found to cleave internal sialic acid of GM1 and fucosyl GM1 on *V. cholerae* at the highest rate among other sialidases tested. Another unusual type of sialidase is endo-sialidase (endo-N) from bacteriophages. It is a phage specific for *Escherichia coli* strain K1 which encoded de-polymerase that degrades the α(2,8)-linked polysialic acid chains (PSA) of *Escherichia coli* strain K1 which is essential for bacteriophage infection. (Petter and Vimr, 1993). PSA are synthesize usually by certain pathogenic bacteria and more widely distributed on vertebrate neural cell adhesion molecule (NCAM) and sodium channel polypeptide (Rutishauser and Jessell, 1988; Silver and Vimr, 1990). Moreover, sialidase activities also shown present in many parts of human tissues, including salivary glands and lungs (Achyuthan and Achyuthan, 2001).

Despite details above, no two different sialidases have shared identical active site and biochemical properties with each other, even though they exhibit high sequence similarity. This feature of being structurally distinct around the active sites would provide different strategies for drug discovery (Russell et al., 2006).

1.4 Relationship between sialic acids and human health

Sialic acid is found to decorate all cell surfaces of vertebrates and higher invertebrates involved in modulating and mediating a variety of important pathological processes. Sialic acids have been known to serve as binding sites for various pathogens and toxins (Angata and Varki, 2002; Ilver et al., 2003; Schauer, 2000) thus mediating specific roles in human health and disease (Table 1.1).

The binding of influenza viruses to sialic acids on the human airway epithelium is known as the first critical step in the process of infection (Russell and Webster, 2005; Suzuki, 2005). Details of influenza virus replication cycle can be refered to Figure 1.3. Both α(2,3)- and α(2,6)-linked sialic acid can be recognized by influenza viruses as a receptor (Ito, 2000; Schauer, 1982a) but human viruses prefer α(2,6)-linked sialic acid, while avian and equine viruses predominantly prefer α(2,3)-linked sialic acid (Ito, 2000). Human respiratory epithelium basically expresses both forms of sialic asids, with α(2,6)-linked sialic acid being more abundant than α(2,3)-linked sialic acid (Hassid *et al*., 1999; Matrosovich *et al*., 2004). It is highly likely that infections of avian influenza viruses direct to human are possible to occur but it is very inefficient. Such phenomena require the virus haemagglutinin to undergo specific mutations in order to recognize the α(2,6)-linked sialic acid instead of α(2,3)-linked sialic (Stevens et al., 2006; Suzuki, 2005). More commonly, the virus infected wild birds and domestic poultry into other domesticated mammals like pig which is known as a ‘mixing vessel’ as it contains both α(2,3)-linked and α(2,6)-linked sialic acid on the epithelium (Suzuki, 2005).

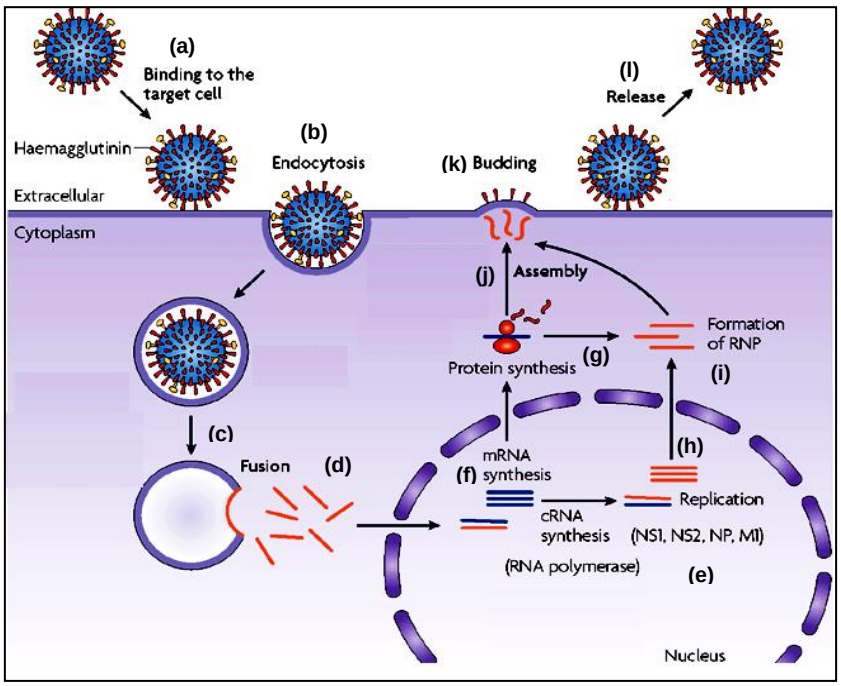


Figure 1.3 Replication cycle of influenza virus. (a) The virus binds to receptors on the surface of the host cell via haemagglutinin, **(b)** virus is internalised into endosomes. **(c)** Fusion and uncoating events occur, which are pH dependent due to low pH in endosomes result in **(d)** the release of the viral genome (in the form of viral ribonucleoproteins; vRNPs) into the cytoplasm. **(e)** The vRNPs are then imported into the nucleus for replication. **(f)** Positive-sense viral messenger RNAs (mRNAs) are exported out of the nucleus into the cytoplasm for **(g)** protein synthesis. **(h)** vRNP assembly, which occur in the nucleus. **(i)** Late in infection, vRNPs form and leave the nucleus, and **(j)** progeny viruses assemble and **(k)** bud from the plasma membrane.**(l)** neuraminidase cleave sialic acid on host cell surface to release and spread virus progeny (von Itzstein, 2007).

There is an example of direct avian to human transfer occurring during the 1918 influenza pandemic which killed tens millions of people (Glaser *et al*., 2005; Stevens *et al*., 2006). This phenomenon happened due to a large dose of virus inhaled by the patients that reached the lower airways contains α(2,3)-linked sialic acids present in humans (Glaser *et al*., 2005). Complications of influenza can include bacterial infections, viral pneumonia, cardiac and other organ system abnormalities. People with chronic medical conditions may be at increased risk for these complications.

Other respiratory pathogens such as viruses belonging to the paramyxovirus families also utilise cell surface carbohydrates to bind and gain entry to a variety of mammalian tissues (Figure 1.4). Parainfluenza viruses (hPIVs) were first discovered in the late 1950s when three different viruses recovered from children with lower respiratory disease proved to be unique but closely related to influenza virus (myxoviruses), which shared a few antigenic sites with the influenza virus. These viruses replicated in the epithelium of the upper respiratory tract and spread to the lower respiratory tract (Moscona, 2005). As stated widely in literature, hPIV is genetically and antigenically divided into types 1 to 4 (Henrickson, 2003).

Each year more than 5 x10^6 of lower respiratory infections occur in younger children under 5 years old as reported in United States (Denny and Clyde, 1986; Glezen et al., 1984). From these cases, hPIV1 and hPIV3 are found in one-third of these infections (Denny and Clyde, 1986; Glezen et al., 1984; Murphy et al., 1980). These hPIVs cause upper respiratory infections in infants, children and adults to a lesser extent while lower respiratory tract infections occur in immunocompromised patients with chronic disease (e.g. heart, lung disease and asthma) and also in elderly people (Falsey, 1991; Falsey et al., 1995; Glezen et al., 1984; Glezen et al., 2000; Lamy and Debacker-Willame, 1973; Muir and Pillay, 1998;

Murphy et al., 1980). Less is known about hPIV4 but still young infants and children are clearly infected by this virus (Henrickson, 2003).

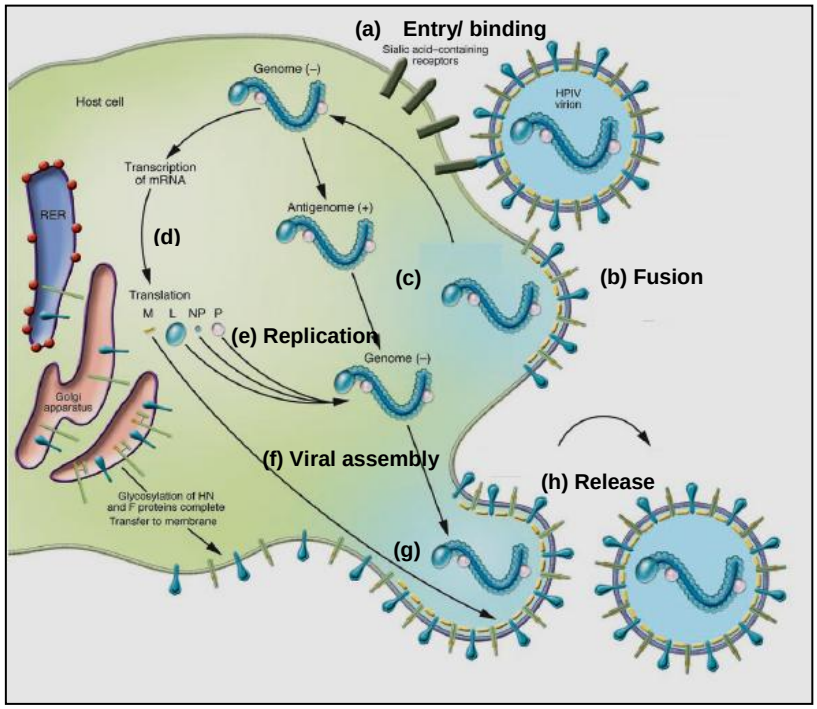


Figure 1.4 Schematic illustration of parainfluenza life cycle. (a) Virus binds to cell surface receptor sialic acid by haemagglutinin-neuraminidase (HN) **(b)** Viral envelope then fuse directly with plasma membrane of the host cell, mediated by fusion protein (F protein), **(c)** thus release viral nucleocapsid into cytoplasm contains genome RNA **(d)** Viral mRNAs and a full-length template are transcribed and translated from the genome RNA **(e)** Replication of the genome RNA occurs in cytoplasm **(f)** New virus progeny is assembled and **(g)** virus bud from plasma membrane **(h)** Neuraminidase cleave sialic acid receptor on the cell surface and allows the release of newly budded virus from the cell to begin a new round of infection (Moscona, 2005).

There are also studies that link the level of sialic acids on lipoprotein to cardiovascular disease risk, apparently by affecting the interactions of lipoproteins in determining uptake of lipids by endothelium (Millar, 2001; Tertov et al., 2001). Research done by Tertov et al. (2001) has identified the presence of *trans*-sialidase activity in human serum which removes sialic acid from low density lipoprotein (LDL) to glycoconjugates of lipoproteins, glycoproteins and sphingolipids in human serum may induce cholesteryl ester accumulation in human aortic intimal smooth muscle cells.

It has been well reported that the relationship between intensity of sialylated glycoconjugates on the cell surface and the possible invasion and metastasis is shown in many types of tumours (Raval et al., 2003; Yamamoto et al., 1997). Furthermore, these sialylated molecules are easily detected in serum and can be used as markers for cancer progression (Varki, 2008). (Miyagi et al., 2004) reported that carcinogenesis, invasion and metastasis of tumor cells are related to abnormal expression of terminal sialic acids and in specific, α(2,3)-linked sialic acids. This association is due to the recognition of malignant cells by selectins which caused interaction and circulating tumor cells with leukocytes, platelets and endothelium and further facilitating metastasis (Varki and Varki, 2002). Interestingly, decrease in migration ability of some cancers such as colon carcinoma cells and gastric cancer are due to the inhibition of α(2,3)-linked sialic acid residues (Ishizuka et al., 2008). This valuable information would provide a helpful clue for early diagnosis and prognosis of certain cancer.

Table 1.1

Relation between pathogens and diseases with virulence factors

Virulence factor	Pathogens	Disease
Neuraminidase (NAs)	**Ortomyxoviruses**	
	- Influenza virus A and B	- Influenza (humans, birds, horses, seals etc)
	Bacteria	
	- *Pseudomonas aeruginosa*	- Cystic fibrosis
	- *Streptococcus*	- Septicaemia, pneumonia, meningitis, periodontal disease
	- *Pneumococcus*	- Septicaemia, haemolyticuraemic syndrome
	- *Bacteroides*	- Gastritis
	- *Helicobacter pylori*	- Periodontal disease
	- *Actinomyces*	- Gas gangrene, peritonitis
	- *Clostridia*	- Septicaemia
	- *Corynebacteria*	- Peritonitis
	- *Enterococcus*	- Peritonitis
	- *Escherichia*	- Cholera
	- *Vibrio cholerae*	
	Parasites	- Colonic parasite (squirrel monkeys)
	- *Trichomonas mobilensis*	- Non-pathogenic in vertebrates
	- *Trypanosoma rengeli*	- African disease of animals
	- *Trypanosoma vivax*	- Meningitis
	Bacteriophage	
	- K1E and K1F	
Haemagglutinin-neuraminidase (HN)	**Paramyxoviruses**	
	- Parainfluenza viruses	- Respiratory disease of humans
	- Sendai virus	- Murine parainfluenza
	- Mumps virus	- Mumps
	- Newcastle disease virus	- Respiratory disease of chickens

		and other birds
Trans-sialidase (TSs)	**Parasites** - *Trypanosome cruzi* - *Trypanosome congolense* - *Trypanosome brucei* - *Pnuemocycstis carinii* - *Eimeria tenella*	 - Chagas disease - African disease of animals - African sleeping sickness - Pneumonia - Diarrhoea disease of chickens

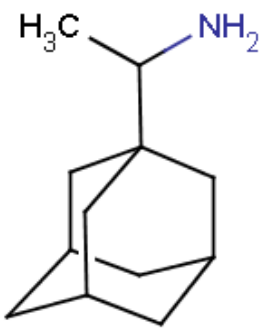
1.5 Development of inhibitor against influenza and parainfluenza infections

1.5.1 Antiviral drugs against influenza viruses

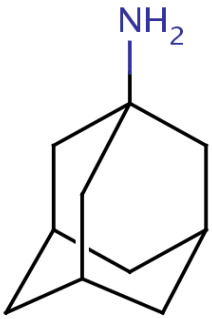
Influenza infection caused by both influenza A (IFV A) and influenza B (IFV B) strains bring enormous effects to the socioeconomic impacts. This virus usually invades cells of the upper and central respiratory tract. Since 1997, the emergence of a new IFV A virus of H5N1 types has been causing epidemics in wild birds and domestic poultry. Alarmingly, in particular in Asia, increased levels of human infections by this virus has made the likelihood of a possible influenza pandemic a serious concern. Research done in the 1960s suggested that pandemics usually arise when strains of avian and human influenza combine. Vaccination is the core means in disease prevention, while antiviral agents are primarily introduced for treatment. At this present time, no antigenically-well matched vaccine is currently available against this future pandemic virus. The concern of human fatality caused by the infection has increased the awareness to this threat and many governments have implemented plans to develop new antiviral drugs as well as stockpiling available anti-influenza drugs for backup.

The first drugs available for the treatment or prophylaxis of influenza were adamantane-based M2 ion channel protein inhibitors, rimantidine (Flumadine®) and amantadine

(Symmetrel®) (Figure 1.5) (Douglas, 1990; Wintermeyer and Nahata, 1995). These drug only been used against influenza A strains because only A strains have the M2 ion channel proteins (Douglas, 1990; Pinto et al., 1992) but have been found to cause CNS side effects (Pinto et al., 1992; Wintermeyer and Nahata, 1995) and also produce rapid emergence of drug-resistant viral strains (Hay, 1992). The rate of adamantine resistance in the United States increased from 2% in 1995-2002 to alarmingly 92% in 2005-2006 influenza seasons (Bright et al., 2006) while 99.6% of seasonal influenza A (H1N1) were found resistant towards oseltamivir as stated by Centers for Disease Control and Prevention (CDC). Rapid emergence of influenza viruses via mechanisms known as antigenic drift and shift (Webster et al., 1992) has cause insensitivity to the available drugs due to mutation or gene reassortment within the current influenza variant (Triana-Baltzer et al., 2009c).



Rimantidine (Flumadine)

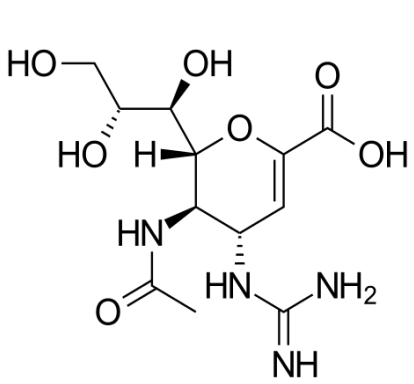


Amantadine (Symmetrel)

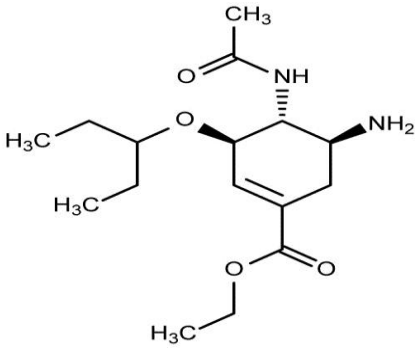
Figure 1.5 Chemical structures of adamantane based inhibitors, rimantidine and amantadine.

Antigenic shift, which is seen only with influenza A viruses, results from the replacement of the haemagglutinin and sometimes the neuraminidase with novel subtypes that have not been present in human viruses for a long time. While, antigenic drift is a subtler process than shift which involves the accumulation of mutations within the genes that code for antibody-binding sites. This results in a new strain of virus particles which cannot be inhibited as effectively by the antibodies that were originally targeted against previous strains (Treanor, 2004). Due to these issues, there has been considerable effort worldwide to discover new therapeutic agents for all types of influenza strains.

Another development of anti-viral drugs are known as neuraminidase inhibitors (NAI). The idea is that neuraminidase inhibitor must fit directly into the enzyme's active site pocket to block the enzyme activity in order to maintain its function (Moscona, 2005). This action will block the function of viral neuraminidase preventing the virus from reproducing and exiting from infected cells (Jones et al., 2006). Such neuraminidase inhibitors show potency against influenza viruses and have been commercialized for prophylaxis and treatment. They are inhaled anti-influenza drug known as zanamivir (Relenza®, GSK) (von Itzstein et al., 1993) and orally bioavailable drug called oseltamivir (Tamiflu®, Roche) (Figure 1.6) (Kim et al., 1997). An inhaled formulation normally will deliver the drugs directly to the specific infection site. These drugs are proven to be effective in reducing the duration of illness and risk of complications. However, because the drugs act at the stage of viral replication, they must be administered as early as possible after the onset of illness. Furthermore, their efficacy diminishes significantly if they are not taken within 36 to 48 hours of the onset of symptoms. Influenza strains can also develop resistance to these drugs, reducing their clinical effectiveness (Hedrick et al., 2000).



Zanamivir (Relenza)



Oseltamivir (Tamiflu)

Figure 1.6 Chemical structures of neuraminidase inhibitors, zanamivir and oseltamivir. The active component of the anti-influenza agent, Relenza and Tamiflu.

Both developments of zanamivir and oseltamivir give rise to new alternatives drugs called Fludase® (DAS181). Fludase® contains a recombinant fusion protein of the catalytic domain of *Actinomyces viscosus* sialidase and the epithelial anchoring domain of human amphiregulin as shown in Figure 1.7 (Triana-Baltzer et al., 2009b). It efficiently binds to respiratory epithelial cells and then removes cell surface sialic acid residues that are used by both avian and human influenza viruses to invade respiratory epithelial cells (Malakov et al., 2006b). Fludase® functions by targeting host cells instead of the virus and may be less likely to induce drug resistance than virus-targeting compounds (e.g. adamantanes and neuraminidase inhibitors) (Belser et al., 2007; Triana-Baltzer et al., 2009c). It therefore represents a first-in-class influenza therapy and potentially an important new weapon in the fight against new strains of influenza virus, including new avian strains.

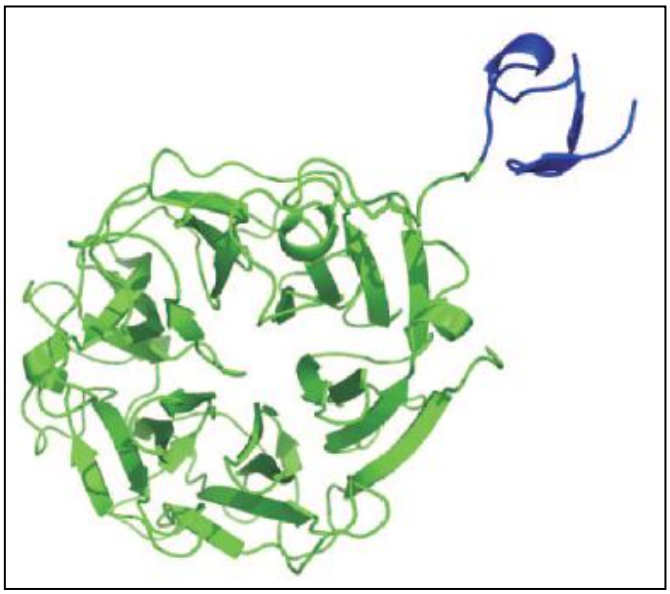


Figure 1.7 Molecular model of Fludase (DAS181). The catalytic domain of *Actinomyces viscosus* sialidase is colored in green and the epithelial anchoring domain of human amphiregulin on C-terminus in blue (Malakov et al., 2006a).

This drug has completed its preclinical development and entered clinical trials development to determine its safety and efficacy in humans. During the preclinical studies, Fludase® displayed potent antiviral activity against clinical influenza strains, including the deadly H5N1 strain (Belser et al., 2007; Triana-Baltzer et al., 2009a). This drug is administered through oral inhalation and the presence of cell surface anchoring-domain aid its attachment to the respiratory epithelium thus increases its retention time and drug potency (Triana-Baltzer et al., 2010). Besides, from the *in vitro* and *in vivo* preclinical test, the drug was found to improve lung function, pathology and the mice survival subjected to viral challenge. While in another study which resembles human influenza infection, administration of Fludase® led to reduced signs of inflammation and illness through significantly inhibiting viral replication or viral shedding (Malakov et al., 2006a). Fludase has now entered Phase II clinical trials to measure the drug effect on influenza viral load, safety and its tolerability. Fludase® is one

from the three antiviral drugs developed by NexBio for the treatment of influenza. The others are known as Inviridin and Viracidin.

The need for novel, effective and safe therapies is acute. In particular, treatments and prophylactics against the entire spectrum of influenza strains, and which are not subject to drug-resistance, are much-needed. Vaccines are only partially effective because of the variety of different circulating strains of influenza at any one time. Current anti-influenza therapy options are also restricted in efficacy by rapid viral evolution and subsequent drug resistance.

1.5.2 Antiviral drugs against parainfluenza viruses.

Parainfluenza viruses (hPIVs) were first discovered in the late 1950s when three different viruses recovered from children with lower respiratory disease were proved to be unique but closely related to influenza virus (myxoviruses)-shared a few antigenic sites with influenza virus. These viruses replicated in the epithelium of the upper respiratory tract and spread from there to the lower respiratory tract (Moscona, 2005).

Respiratory viruses in the paramyxovirus family including hPIV have been left behind by influenza in terms of development of effective antiviral drugs and vaccine. Although effective prophylaxis strategies using for respiratory syncytial virus are available, there are no similar treatments available specific for hPIVs (Moscona et al., 2010). As with influenza, vaccination is also the primary strategy for the prevention, but due to other possibilities like pandemic outbreaks, vaccinations will become inadequate. Vaccines require several months to develop, as well as, phenomena such as antigenic drift in the virus may also render its protective effects. Due to this problem, development of alternative treatment is urgently needed.

Since 1960s, an inactivated trivalent vaccine for hPIV1, hPIV2 and hPIV3 used in infants was found immunogenic and did not offer protection from further infection (Chin et al., 1969; Fulginiti et al., 1969). This has highlighted the challenge to research teams to identify which elements of the immune response provide protection against hPIVs. Current approaches to develop hPIVs vaccines include intranasal administration of live attenuated strains, subunit strategies using the HN and F proteins, recombinant bovine/human viruses and strains engineered using reverse genetics (Henrickson, 2003). From the literature, there is a vaccine developed from recombinant Sendai virus (SV) cocktails to work against hPIV1 and respiratory syncytial virus (RSV) infections in human, but this vaccine is still undergoing clinical trials (Hurwitz, 2008; Hurwitz et al., 1997).

Other vaccines that have been developed, which are intranasally administered are bovine PIV3 (bPIV3) vaccine and cold-adapted PIV3 vaccine (Durbin and Karron, 2003). The bPIV3 vaccine was found effective against both PIV3 and RSV infections in an African green monkey cell line while the cold adapted PIV3 vaccine has been evaluated throughout a range of pediatric age groups but still undergoing efficacy trials (Sato and Wright, 2008). Maeda et al. (2005) have described the development of a live bivalent vaccine against both influenza and parainfluenza viruses by using a reverse genetics approach. Practically, this combined vaccination in which patients are immunized with multiple pathogens at the same time are very desirable in the clinic. However, an important issue that needs to be evaluated is that simultaneous immunization can result in failure if one of the virus affects the replication and causes pathogenicity (Pichichero, 2000). Belshe et al. (2004) had developed another vaccine candidate for hPIV3 known as cp45 vaccine. This vaccine was shown to be safe and immunogenic in young children in Phase II trials but would still needs to be further evaluated in Phase III of efficacy studies.

Antiviral therapy for hPIVs is now being explored due to complexities in vaccine development. This could be a principal weapon against these diseases. Binding, fusion and entry stages of the hPIV's life cycle can be critical for intervention to prevent the infection. One example is zanamivir (Relenza), which does not only inhibit the neuraminidase activity but also the receptor interaction of hPIV3 HN. However, there is a case where serial passaging of the virus in cell culture in the presence of zanamivir resulted in new hPIV3 variant thus reduced sensitivity to the drug (Murrell et al., 2001). This was due to the changes in the catalytic site of hPIV3-HN which contribute to drug resistance (Alymova et al., 2008; Porotto et al., 2004). Another development of an antiviral drug was described by Alymova et al. (2008) of BCX 2798 as a haemagglutinin-neuraminidase (HN) inhibitor against parainfluenza virus infections *in vitro* and *in vivo* (Figure 1.8) (Alymova et al., 2004) and against lethal synergism between hPIV and *S. pneumoniae* in a mouse model (Alymova et al., 2005). This synthetic drug is based on the binding of 2-deoxy-2,3-didehydro-N-acetylneuraminic acid (Neu5Ac2en) compound to receptor on NDV neuraminidase active site and thus function to block the catalytic binding site of the virus HN molecule (Watanabe et al., 2009).

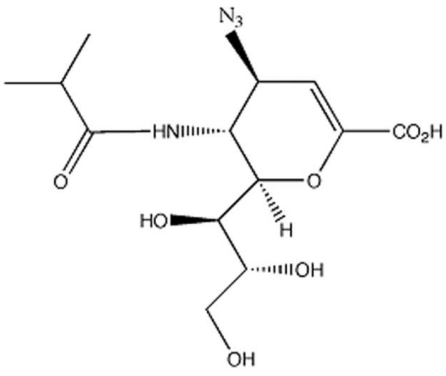


Figure 1.8 Chemical structure of haemagglutinin-neuraminidase inhibitor known as BCX2798 (Alymova et al., 2005).

Recently, a study by Mao et al. (2008) has developed two novel small molecules (C5 and C7) from the screening of a small molecule from ChemBridge library. These compounds exhibited their inhibitory effects on hPIV3 infection in CV-1 and human epithelial A549 cells; moreover, they are also effective in preventing cytopathic effects in infected cells (Mao et al., 2008).

Nexbio (US pharmaceutical company) has developed an alternative treatment for hPIV infections using Fludase (DAS181), which has been used in clinical trials for treatment or prophylaxis of influenza. For this study, Fludase was shown to inhibit a recombinant strain of hPIV3 in a high throughput antiviral screening assay as mentioned by Roth et al. (2009). From the assay, Fludase exhibited the highest potency in EC_{50} among other 23 compounds tested (Roth et al., 2009). Moreover, the anti-PIV activity of Fludase in human airway epithelial cell (HAE), a culture that reflects human airway and in the cotton rat animal model has suggested that this compound to undergo further clinical studies (Moscona et al., 2010). This protein functions by removing sialic acids from the surface of epithelial cells, rendering the virus to bind to the receptors. However, by removing cel-surface sialic acids, other cryptic receptors can be exposed, which may serve as receptors for other pathogens.

1.6 Carbohydrate Binding Modules (CBM)

Carbohydrate binding modules (BMs) were previously classified as cellulose-binding domains (CBDs) based on the initial discovery of several modules that bound cellulose (Gilkes et al., 1988; Tomme et al., 1988) . However, additional modules in carbohydrate-active enzymes were continually being found that bound carbohydrates other than cellulose. This led to the classification of CBMs as carbohydrate-binding modules (CBM).

A carbohydrate-binding module (CBM) is defined as a contiguous amino acid sequence within a carbohydrate-active enzyme with a discrete fold and having carbohydrate-binding activity (Boraston et al., 1998; Boraston et al., 1999)**. T**his type of domain has been found in both hydrolytic and non-hydrolytic proteins. For example, proteins which possess the hydrolytic activity includes cellulases and xylanases encompass a discrete modules (catalytic module and one or more CBMs) joined by unstructured linker sequences. Removing the CBM from the scaffolding will promote decreased enzyme activity (Carrard and Linder, 1999; Coutinho et al., 1993; Goldstein et al., 1993).

As reported, CBMs can be divided into families based on amino acid sequence similarity, binding specificity and structure (Coutinho and Henrissat, 1999; Rodriguez-Sanoja et al., 2005; Volkov et al., 2004). Currently, there are 67 defined families of CBMs based on Carbohydrate-binding Module Family Classification in the CAZY database that display substantial variation in ligand specificity (Cantarel et al., 2009). Some of these modules can be found in glycosyl hydrolases group which are discrete, non-catalytic modules that primarily exist to target the parent enzyme to its substrate for efficient hydrolysis through increasing the concentration of the enzyme at the substrate surface (Boraston et al., 2004b). CBM is able bind specifically to polysaccharide independently when isolated from the parent molecule and can behave in a cooperative manner when isolated in tandem (Boraston et al., 2002b; Crennell et al., 1994).

CBMs usually concentrate enzymes onto the polysaccharide substrate via its sugar-binding activity, maintaining the enzyme in proximity close with the substrate, allowing more efficient degradation of polysaccharides (Bolam et al., 1998a). In fact, dramatically decreased enzyme activity has been evident in cellulose degradation when CBM domain was removed from the scaffolding (Ali et al., 2005; Araki et al., 2004; Bolam et al., 1998b). There have been reports indicating that significant decreases in enzyme activity was due to the

proteolytic excision or genetic truncation of CBMs from the catalytic module (Hall et al., 1995).

Recently, the use of CBMs has been established in different fields of biotechnology and numbers of publications and researches have increased constantly. Its unique properties have made CBM a perfect candidate for various applications. This is due to the fact that CBMs are independently folding units and, therefore, are able to function autonomously in chimeric proteins. Thus, CBMs are excellent model systems to study the mechanism of protein-carbohydrate recognition and their diversity in ligand specificity that underpins the exploration of these protein modules in various biotechnological applications.

1.7 Study of family 40 CBM from bacterial sialidases

1.7.1 *V. cholerae* sialidase

In bacteria, sialidases can contain accessory modules attached to the catalytic core of the protein. For example, the *NanH* sialidase from *Vibrio cholerae* is known to be composed of a canonical six-bladed β-propeller catalytic domain that has two CBMs that flank the catalytic domain as showed in Figure 1.9 (Crennell et al., 1994). Both CBM domains share the same structure topology despite sharing only 23% sequence identity. The N-terminal domain is classified as a family 40 CBM (CBM40) and recognizes single sialic acid moiety as the binding ligand but superimposition of both N- and C- domains showed the C-terminal domain to be shifted in space from the correct position for ligand binding. Thus it is highly unlikely that the C-terminal domain binds to sialic acid (Moustafa et al., 2004). Interestingly, even though other bacteria, for example *Clostridium perfringens* NanJ sialidase also possess the same domain but its displayed lower affinity towards sialic acid as compared to *V. cholerae* CBM40 domain (Boraston et al., 2007).

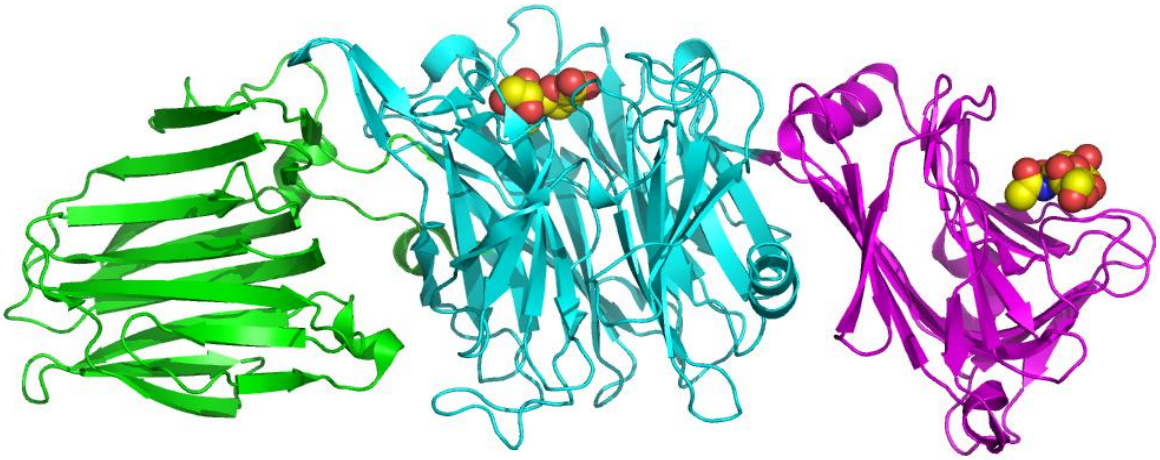


Figure 1.9 Schematic drawing of *Vibrio cholerae* sialidase. The central catalytic domain is flanked by two lectin-like domains. The C-terminal lectin domain is in green, the N-terminal lectin domain, known as CBM40, is in *maroon* (Moustafa et al., 2004).

As mentioned previously, CBMs are not involved in catalytic activity but serve to bring the substrate to the active site of the catalytic domain to help in hydrolysis (Rodriguez-Sanoja et al., 2005). Moreover, it is shown that the sialic acid-specific CBM from *V. cholerae* sialidase can be isolated independently from its parent enzyme and exploited to generate multivalent polypeptides that can bind to sialic acid (Figure 1.1), the monomer having a relatively high affinity K_d ~ 30 µM (Connaris et al, 2009). This was reported as one of the highest affinity for a monosaccharide-protein interaction that recognized sialic acid (Moustafa et al., 2004), when compared to other sialic acid-protein interactions, such as wheat germ agglutinin, demonstrated a binding affinity of K_d ~ 100 µM with sialyllactose, involving also contributions from interactions of the galactose (Kronis and Carver, 1982), the influenza virus haemagglutinin monomer that has a K_d ~ 2.5 mM (Sauter et al., 1989) as does the sialic acid binding domain of rhesus rotavirus (Dormitzer et al., 2002).

Due to its relatively high monovalent affinity for sialic acid, tandem repeat polypeptides have been engineered to achieve higher affinity when interacting with a multivalent surface. Avidity effect is a term usually used to describe the strength of binding of a molecule with multiple binding sites by a larger one, for example the binding of a complex antigen by an antibody. In nature, protein-carbohydrate interactions (eg: CBM) may have evolved to have a weak binding capability towards glycans due to some restrictions such as during elevated temperature. This phenomena occur frequently in thermo- or hyperthermophilic enzymes where binding affinity interactions can be improved by multivalency approach (Boraston et al., 2004a).

1.7.2 *Streptococcus pneumoniae* sialidases (*NanA*, *NanB* and *NanC*)

S. pneumoniae is a Gram-positive, alpha-haemolytic bacterium which belongs to the *Bacillus/Lactobacillus/Streptococcus* group. Besides being a major human pathogen, this bacteria is responsible for diseases like pneumonia, otitis media, septicaemia and meningitis (Tuomanen, 2004). Today, the broad-spectrum antibiotic treatments for these infections are not successful due to the increased emergence of drug-resistant strains. Each year, *S. pneumoniae* has caused more than 1 million of deaths each year worldwide (Williams et al., 2002). Normal clinical infection of *S. pneumoniae* is preceded by bacterial colonization in the upper respiratory tract, predominantly in the nasopharynx. Interestingly, about 40% of healthy people are asymptotic carriers, which bring a significant reason in the increment of pneumococcal infections (Bogaert et al., 2004).

During the successful *S. pneumoniae* genome sequencing, a number of genes encoding proteins have been identified contributing to the bacterial virulent roles and/or as protective antigens (Tettelin et al., 2001). Among the virulence factors mentioned, sialidases are identified to contribute to the pneumococcal colonization and its early infection processes (Jedrzejas, 2001). Sialidases specifically catalyze the removal of sialic acid from cell surface

glycoconjugates, this action would caused damage to the host cells and unmask some important receptors for bacterial adherence and colonization (Mitchell, 2000).

S. pneumoniae expresses three distinct sialidases, known as *NanA* (115 kDa), *NanB* (78 kDa) and *NanC* (82 kDa). From a gene screening study, *NanA* gene is the main sialidase found from 342 clinical isolates; while *nanB* and *nanC* genes were identified in 96% and 51% of these strains respectively (Pettigrew et al., 2006). From gene knockout study in the MF1 mouse models, *NanA* and *NanB* sialidases are found to be important during the *S. pneumoniae* infection in the respiratory tract and sepsis as mentioned by Manco et al. (2006). Furthermore, *NanA* also plays an important role in long-term nasopharynx and middle ear colonization of *S. pneumoniae* (Long et al., 2004; Simell et al., 2006).

NanA and *NanB* share very little similarity in their amino acid sequences (around 24%) while *NanC* is a close homolog to *NanB* with over 50% sequence identity. From sequence analysis, an N-terminal signal peptide is shown to be present in all the sialidases to direct their secretion to the outside cells. Only *NanA* is known to have a surface-anchored protein with the presence of 'LPTEG' sequence motif. Moreover, this protein also contains a C-terminal P/G- and T/S-rich region which is predicted to be naturally disordered. This feature promotes flexibility to the *NanA* protein when fixed to the pneumococci cell wall (Camara et al., 1994).

Moreover, all three *S. pneumoniae* sialidases possess a CBM40 domain which flanks the catalytic domain similar to some other large sialidases. The RIP motif within the catalytic domain is also found conserved in all three sialidases having a catalytic arginine, with four repeating Asp-box motifs at similar positions within the whole sequence (Figure 1.10). This domain functions to recognize and bind sialic acids similar to the *Vibrio cholerae* sialidase

(Connaris et al., 2009). A study on CBM40 domain of *NanA* will be further discussed in Chapter 2 of this thesis.

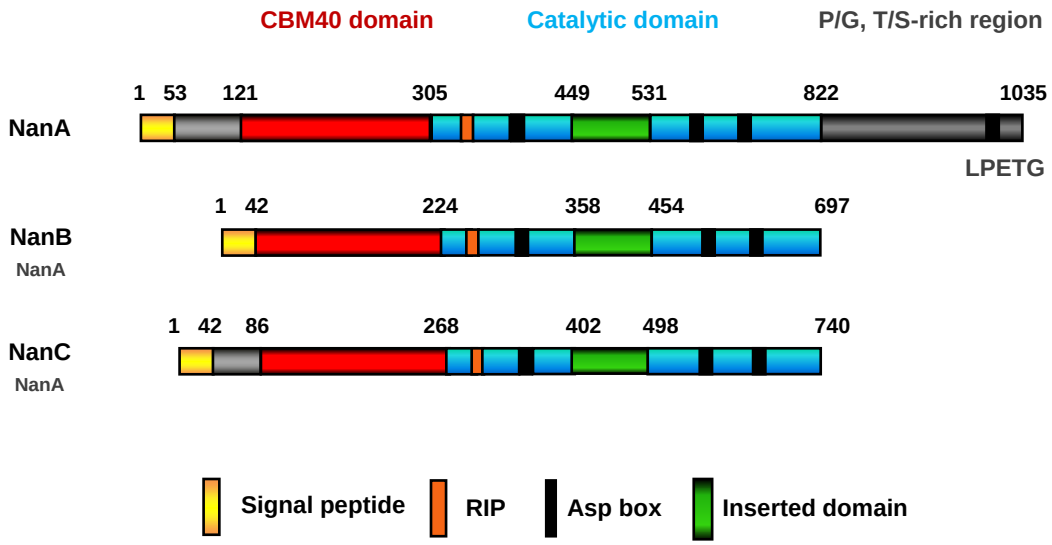


Figure 1.10 Schematic views of the three *S. pneumoniae* sialidases. The sequences of *NanA* (UniProt accession number: P62575), *NanB* (Q54727) and *NanC* (Q97Q99) showing their domain structures.

1.8 Multivalency of CBMs

Recently, many researchers have focused on receptor–ligand interactions including interactions of molecules with its substrate, which are the key element to the functioning of all biological systems. Ligands that show multiple copies of recognition sites and elements are termed as multivalent (Zverlov et al., 2001). Multivalent carbohydrate ligands may be in the form of branched saccharides, clustered cell-surface glycans or polysaccharides as is the most relevant to CBMs (Boraston et al., 2004a).

Some of sialic acid binding lectins such as viral glycoproteins, bacterial toxins as well as the mammalian lectin superfamilies (siglecs and selectins) bind to their receptor with high affinity due to multivalent nature of the molecules (Crocker, 2002). Generally, association constant for the binding of monovalent and divalent sialosides by lectins can reach 10^4 M^{-1} but due to their multivalency, some sialic acid binding lectins can interact with multivalent cell surface glycans and reach affinities up to 10^9 M^{-1}. These enhanced affinities have been shown due to the improved structural packing of proteins promoted by ligand binding, associated with favourable binding energetics (Williams et al., 2004; Williams et al., 2003). Research by Mammen et al. (1998b) reported that one of the best studied multivalent lectin-sialic acid interactions is the influenza virus trimeric heamagglutinin, which can achieve affinities up to 10^8 M^{-1} compared with 4 X 10^2 M^{-1} when one or both of the entities are not in a multivalent state.

A recent report carried out by Connaris et al. (2009) demonstrated the multivalency of engineered polypeptides containing tandem repeats from *V. cholerae* CBM40. The constructs were designed containing two, three and four CBM40 modules with varying linker lengths between the modules. The four-CBM40 module protein has been proved to have 700 to 1500 fold increases in affinity compared with single CBM40 module. This is due to the

stabilization of oligomers and their interaction when the CBM40 exists as tandem-linked polypeptides. Moreover, linker lengths between each CBM40 had shown little effect on the binding with only 1.2-fold increase in affinity when using 5 amino acid linker to 15 amino acids linker (Connaris et al., 2009).

A multivalency approach is proven to be effective to overcome relatively weak binding between protein and ligand. One such example is the two family 2b CBMs from *Cellulomonas fimi* xylanase 11A. The association constant for xylan in tandem, was approximately increased to 10^6 M^{-1} compared to its individual association constants which was approximately 10^4 M^{-1} (Bolam et al., 2001). Interestingly, the appearance of multiple CBMs in glycoside hydrolases appears to occur most frequently in thermo- or hyperthermophilic enzymes (Boraston et al., 2003). This may be due to the need of these proteins to overcome the loss of binding affinity in most molecular interactions at elevated temperatures (Boraston *et al*., 2003).

A previous report of synergy interaction between two discrete CBM domains joined by two cellulose binding domain from *Trichoderma reesei*, demonstrated cooperative binding to cellulose (Linder et al., 1996). This was shown that through the binding of one domain to its substrate helped to increase the affinity of second module to its target due to a proximity effect. A good example of CBMs naturally found to be multi-modular in glycoside hydrolases is the α-amylase from *Lactobacillus amylovorus* which contains five modules of CBM in tandem. The ability of the enzyme to absorb starch granules were found to be increased as the number of domain increased in which they strongly absorb cornstarch (Guillen et al., 2007). Another example is xylanase from *Clostridium stercorarium* which has two repeated domains of family 6 CBM at the C-terminal that bind insoluble cellulose. The affinity of cellulose was enhanced due to avidity effect resulting from the co-operative binding to their polysaccharide (Boraston et al., 2002a).

1.9 Project aims and thesis plan

As nature often uses multivalency to increase affinity through avidity (Lee and Lee, 2000), this led us to question whether we could increase the affinity of CBM40 domain for sialic acid-rich cell surfaces by linking copies of CBM40 unit together. One of the project aims was to isolate the gene encodes the CBM40 from *V. cholerae* sialidase, and generate polypeptides containing up to 4 repeat modules in tandem, linked by certain linker length. These polymers are able to bind α(2,3)-, α(2,6)- and α(2,8) sialiosides. Biacore studies using an immobilized, multivalent biotinylated α(2,3)-sialyllactose, demonstrated that increasing the number of linked modules increases the affinity of these CBM40 polypeptides for sialic acid, with a 7000-10,000-fold increase compared to the monomeric-monovalent interaction from K_d~18μM to 3nM, at 25 °C. These results have been published by (Connaris et al., 2009). Moreover, other CBM40 domain was also explored which was from *NanA* of *S. pneumoniae* sialidase in order to generate another multivalent construct to study its multivalency and affinity effect. Furthermore, this project would explore the use of a simpler and more efficient method to obtain mutlivalency by fusing the CBM40 to a domain that oligomerizes as a trimer. It would involve the use of C-terminal domain of *Pseudomonas aeruginosa* pseudaminidase which is able to self-oligomerize to generate trimeric CBM molecules. This domain is known as oligomerization domain (Xu et al., 2009a). Proteins will then be expressed and tested for binding affinity to carbohydrates by calorimetry and surface plasmon resonance as well as protein crystallography trials. Overall, this project will involve the broad use of molecular techniques, protein expression and purification, biophysical characterization including spectroscopy, calorimetry (isothermal titration calorimetry), surface plasmon resonance and protein crystallography.

Since the project focuses on the development of engineered multivalent CBM40s from *V. cholerae* sialidase, a large amount of experimental data is distributed in Chapter 2 which

include upstream work of this project. This includes the expression and purification works of these multivalent CBM40s from this bacterium. Chapter 3 presents the biochemical characterization of the constructs which includes thermal shift assays, ITC, SPR and crystallization trials.

Due to the fact that mammalian cells surfaces comprise of numerous types of molecules including sialic acid, certain pathogens have exploit the presence of this molecule to bind/adhere and/or gain entry to host cells during pathogenesis. By exploiting the CBM's affinity to the cell surface carbohydrates, the use of these will provide a useful insight for the potential treatment against respiratory viruses such as influenza and parainfluenza virus. The *in-vitro* study of these multivalent constructs against influenza virus and parainfluenza virus together with cytotoxicity test and inhibition study, are all presented in Chapter 4 and Chapter 5. Finally, the additional chapter on *in-vivo* works of one of the best multivalent construct against hPIV3 and *Pneumococca*l infection is discussed in Chapter 6. The experimental procedures, as well as some supplemental data, are included in the Appendix section.

Chapter 2

Development, expression and purification of multivalent CBM40s

2.1 Overview

Bacterial sialidases are normally found as modular enzymes and have accessory domains attached to the catalytic core of the protein (Gaskell et al., 1995). One of these accessory domains has been identified as belonging to a family of carbohydrate-binding modules (CBMs) that are normally found in glycosyl hydrolases (Lis and Sharon, 1998). Currently, there are 67 defined CBM families entered in the CAZy (Carbohydrate-Active enZYmes) database based on their ability to adhere to carbohydrates and on their amino acid sequence similarity (Coutinho and Henrissat, 1999). Studies have also shown that CBMs are capable of binding specifically to either simple saccharides or to polysaccharides independently when isolated from the parent molecule and behave in a cooperative manner when isolated in tandem (Boraston et al., 2002c; Crennell et al., 1994). The role of CBMs is to concentrate the enzymes onto the substrate via sugar-binding activity to increase the catalytic efficienty of the enzyme (Thobhani et al., 2003).

Crennell *et al.* (1994) had reported that the sialidase from *Vibrio cholerae* possesses two CBMs that flank the central catalytic domain. From a structural study of the N-terminal CBM from *V. cholerea* sialidase, it has been proved to recognize a single sialic acid moiety (Figure 2.1) with relatively high affinity, K_D ~ 30 µM which is one of the highest reported for a sialic

acid-binding protein. Due to its identification of binding to sialic acid, the N-terminal CBM was thus designated as a Family 40 CBM (Moustafa et al., 2004).

In nature, protein-carbohydrate interactions often display a weak binding affinity, but nature achieves higher affinities through multivalent interactions, often through protein oligomers that gain affinity through an avidity effect. Examples of nature's multivalent molecules are viral glycoproteins, bacterial toxins (Lehmann et al., 2006), as well as mammalian lectin superfamilies (siglecs and selectins), which bind to their target receptor with high affinity (Crocker, 2002). Thus, this form of multivalency is effectively used by nature to overcome weak binding interactions in order to improve its affinity (Boraston et al., 2004a).

Due to its high affinity and an interest in developing a multivalent protein-carbohydrate approach, tandem repeats of the CBM40 from *V. cholerae* sialidase have been engineered. The constructs were designed containing three or four repeats of the CBM40 domain with a short specific amino acid linker between the modules. It was hoped that the affinity towards sialic acid would increase along with the addition of CBM40 domains by an avidity effect. Other constructs were also engineered consisting of a CBM40 domain from *Vibrio cholerae* sialidase fused with an oligomerisation domain from *Pseudomonas aeruginosa* pseudaminidase (Xu et al., 2011). The incorporation of the oligomerization domain from *P. aeruginosa* pseudaminidase produced a trimeric protein with three binding sites for sialic acids. The linkers between the trimerisation domain and the CBM40 domain would produce an unknown orientation of the CBM40 domains, therefore site directed mutagenesis was carried out to introduce a disulphide bond between the monomers in order to create a more rigid trimer that may have its CBM40 domains pointing in the same direction. As well as the *V. cholerae* CBM40, we were also interested to study multivalent forms using the CBM40 domain from *Streptococcus pneumoniae NanA* sialidase fused with the trimerization domain from *P. aeruginosa* pseudaminidase.

This chapter focuses on the development and engineering of the multivalent constructs mentioned above as well as the expression and purification steps involved in the process.

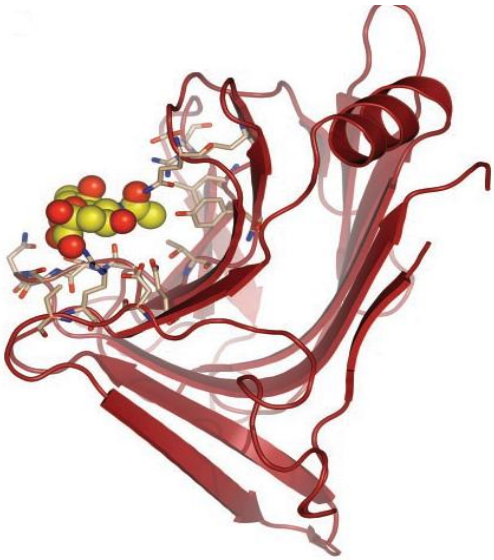


Figure 2.1 Schematic view of the binding of sialic acid to the CBM40 domain of *V. cholerae* sialidase.

2.2 General Methodology

2.2.1 Production and expression of protein constructs

All constructs (Vc3CBM, Vc4CBM, Vc-CBMTD (WT), Vc-CBMTD (Mutant) and Sp-CBMTD) were cloned into pEHISTEV vector (Liu and Naismith, 2009). Briefly, *E. coli* BL21 (DE3) culture harbouring the constructs were grown in Luria-Bertani broth containing 30 µg/ml Kanamycin and incubated at 37°C until the optical density of the culture reached 0.6 at 600 nm. For soluble protein expression in *E. coli*, a heat shock step at 42°C for 30 minutes was introduced before cultures were cooled down to room temperature for about 10 minutes. The cells were then induced with 1 mM IPTG for protein expression and left to grow overnight at 18°C with 200 rpm. A lower temperature (25°C to 18°C) was used in order to

prevent insoluble protein expression, which would have led to accumulation of inclusion bodies. Cells were harvested at 8000 rpm for 15 minutes (4°C) by centrifugation and cell pellets were frozen at - 20°C prior to use.

For lysing, cell pellets were lysed in a Resuspension buffer composed of 0.3 M NaCl, 10 mM Imidazole in Phosphate-buffered saline, supplemented with EDTA-free Protease inhibitor cocktail tablets (Roche Diagnostic, one tablet per 25 ml extract) and 0.2 mg/ml RNA'ase for 30 minutes at room temperature with a gentle shake. The cells were then sonicated on ice at 12 Amp for 30 seconds with 1 minute intervals to further lyse the cells. Cells were subjected to centrifugation at 17000 rpm for 20 minutes and the supernatant was then filtered through a 0.22 µm filter.

2.2.2 Protein purifications protocol

For all constructs, purification steps were performed, according to the following protocols unless stated otherwise.

To obtain high purity of protein, a three-step purification strategy was carried out, which included nickel affinity, gel filtration and anion exchange chromatography. All polypeptides containing His-tag were initially purified by nickel affinity chromatography using a Resuspension Buffer (0.3 M NaCl, 10 mM imidazole in Phosphate-buffered saline) prior to being washed with a gradient concentration of imidazole (10 mM and 20 mM). The proteins were eluted from the column using an imidazole concentration of 250 mM. Samples were analyzed using SDS-PAGE and partially purified polypeptides were dialyzed into Tobacco Etch Virus (TEV) protease cleavage buffer (Phosphate-buffered saline, 0.3 M NaCl, 1 mM dithiothreitol, 0.5 mM EDTA, 20 mM imidazole) and digested overnight with TEV protease (1 mg/ml). The polypeptides were further dialyzed in Resuspension buffer before another purification on a nickel affinity column to remove undigested His-tagged polypeptides. All polypeptides were

concentrated to a certain volume using a 30 kDa cut-off concentrator tubes before final application on a HiPrep 16/60 Sephacryl S-100HR column for gel filtration chromatography purification. As an optional step, an anion exchange chromatography using a HiTrap Q (FF) column (GE Healthcare) was also performed with a linear gradient of increasing NaCl concentration to further purify the protein. Peak fractions containing purified protein were pooled and stored at - 80°C prior to use.

2.3 Multiple tandem repeats of CBM40 domain

The DNA fragment encoding the CBM40 from *V. cholerae* sialidase (residues 25-216) was amplified by PCR from pET30b+ construct containing *NanH* gene (Moustafa et al., 2004) using primer pair 1F and 1R (Table 2.1). The amplified DNA fragment (573 bp) was digested with NcoI and XhoI to be cloned into a pEHISTEV vector containing an N-terminal His-tag with a TEV protease cleavage site upstream of the multiple cloning site (Liu and Naismith, 2009). This construct was then used to transform *Escherichia coli* DH5α.

In order to incorporate different restriction endonuclease sites at the 5'- and 3'- termini, the following sets of primers were used (Table 2.1). This is to allow ligation of individual copies of the DNA fragment to generate three and four copies in tandem. The resulting fragments were then cloned into pEHISTEV vector until the desired number of modules was achieved. These were labeled as Vc3CBM and Vc4CBM, presenting three and four repeating sialic acid-binding domains, respectively.

The tandem repeats of CBM40 *V. cholerae* sialidase were attached to different peptide linker lengths (Figure 2.2). Choices of linkers were designed composed of flexible residues like glycine and serine so that the adjacent protein domains are free to move relative to one another. It has been reported that, length and nature of the linkers (hydrophilic linkers) are

important to ensure two adjacent domains do not sterically interfere with one another and affect its functional activity (Robinson and Sauer, 1998a). This is because the linker sequence composition could have a significant effect on the folding stability of a fusion protein.

All constructs were propagated in *E. coli* DH5α and positive clones were verified by DNA sequencing before transforming expression host *E. coli* BL21 (DE3) strain for protein production. Further information on both constructs can be referred to a paper by (Connaris et al., 2009).

Table 2.1 List of oligonucleotides primers used to amplify DNA fragments

Primer	Oligonucleotide sequence (5'- 3')
1F	CGTCCCATGGCACTTTTTGACTATAACGC (NcoI)
1R	CCGGCTCGAGCTAGTCGCCTTGAATTTCAAAC (XhoI)
3F(5)*	CTGCAAGCTTTGGGAATGGCACTTTTTGAC (HindIII)
3R(5)*	GCACTTCCAAAGCTTGCAGGTCGCCTTGAATTTC (HindIII)
4F(5)*	GGTAGGGAATTCGGGAATGGCACTTTTTGACTATAAC (EcoRI)
4R(5)*	GCACTCCCGAATTCCCTCCGTCGCCTTGAATTTC (EcoRI)

*Note: number of amino acids in linker between each domain fused.

Restriction enzyme sequences for each primer were shown as underlined.

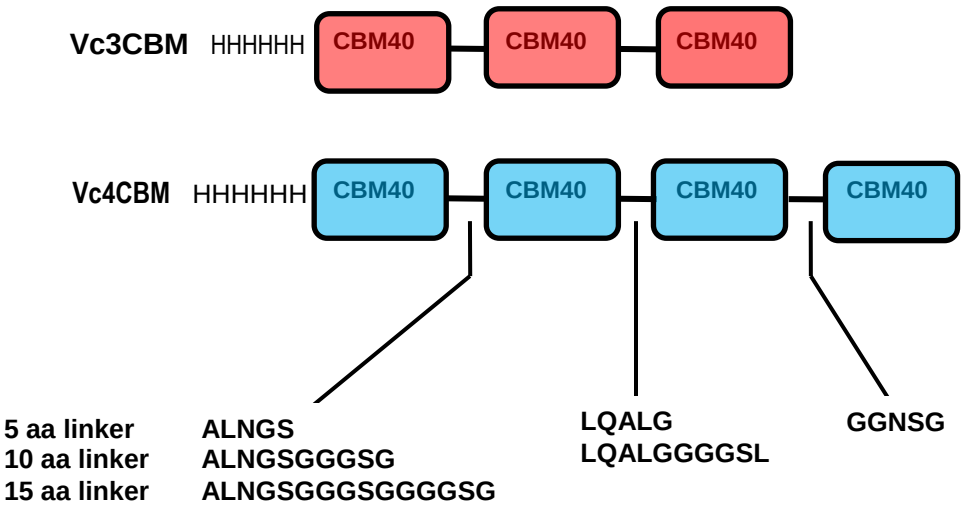


Figure 2.2 Illustration of tandem repeat approach for Vc3CBM and Vc4CBM constructs with their linker sequences.

2.4 Production and expression of multivalent CBM40 construct

2.4.1 Expression and purification of Vc3CBM

Briefly, small scale expression trials of Vc3CBM in *E. coli* BL21 (DE3) strain were performed in order to find out the optimal conditions for protein production. All proteins extracted from the cell lysates and supernatant were kept separately and assessed using sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) using 4-12% NuPAGE gel (Life Technologies). Successful soluble protein expression (Figure 2.3 (A)) was achieved after a heat shock step (42°C for 30 minutes) prior to IPTG induction. This step was introduced to the cultures after they reached an OD_{600} of 0.6 at 37°C. A lower temperature (20°C to 18°C) was used for protein expression in order to prevent insoluble protein and accumulation of

inclusion bodies. A complete procedure of the protein expression was discussed further in Section 2.2.1.

For purification, four-step strategies were carried out as mentioned in Section 2.2.2, comprising two stages of nickel affinity, followed by gel filtration and anion exchange chromatography. During the TEV digestion, the whole process was performed at 4°C with a gentle stirring to avoid protein degradation and to ensure optimum digestion of his-tag protein (Figure 2.3 (B). In certain cases, more than 1 mg/ml of TEV protease was needed to completely digest the protein due to high concentration of protein present. As seen in Figure 2.3 (A) and (C), after two rounds of Ni-column purification followed by a gel filtration chromatography (Figure 2.3 (E)) there was still some lower molecular weight protein contaminants associated with the protein of interest. A chromatogram of a peak with a shoulder was produced represented the quality of the purify protein (Figure 2.3 (D)). In order to remove the contaminants, the protein was therefore concentrated and dialysed against phosphate buffer containing 20 mM NaCl prior for loading into anion-exchange column, HiTrap Q (FF) column (GE Healthcare). Improved protein purity was observed after anion-exchange chromatography as shown in Figure 2.3 (F). Protein identity was confirmed by a mass spectrometry (MS) fingerprint. The yield of pure Vc3CBM was around 35 mg from 2 L culture. An expected protein size of 63 kDa was shown after the N-terminal 6-His tag was removed, which corresponded to Vc3CBM construct.

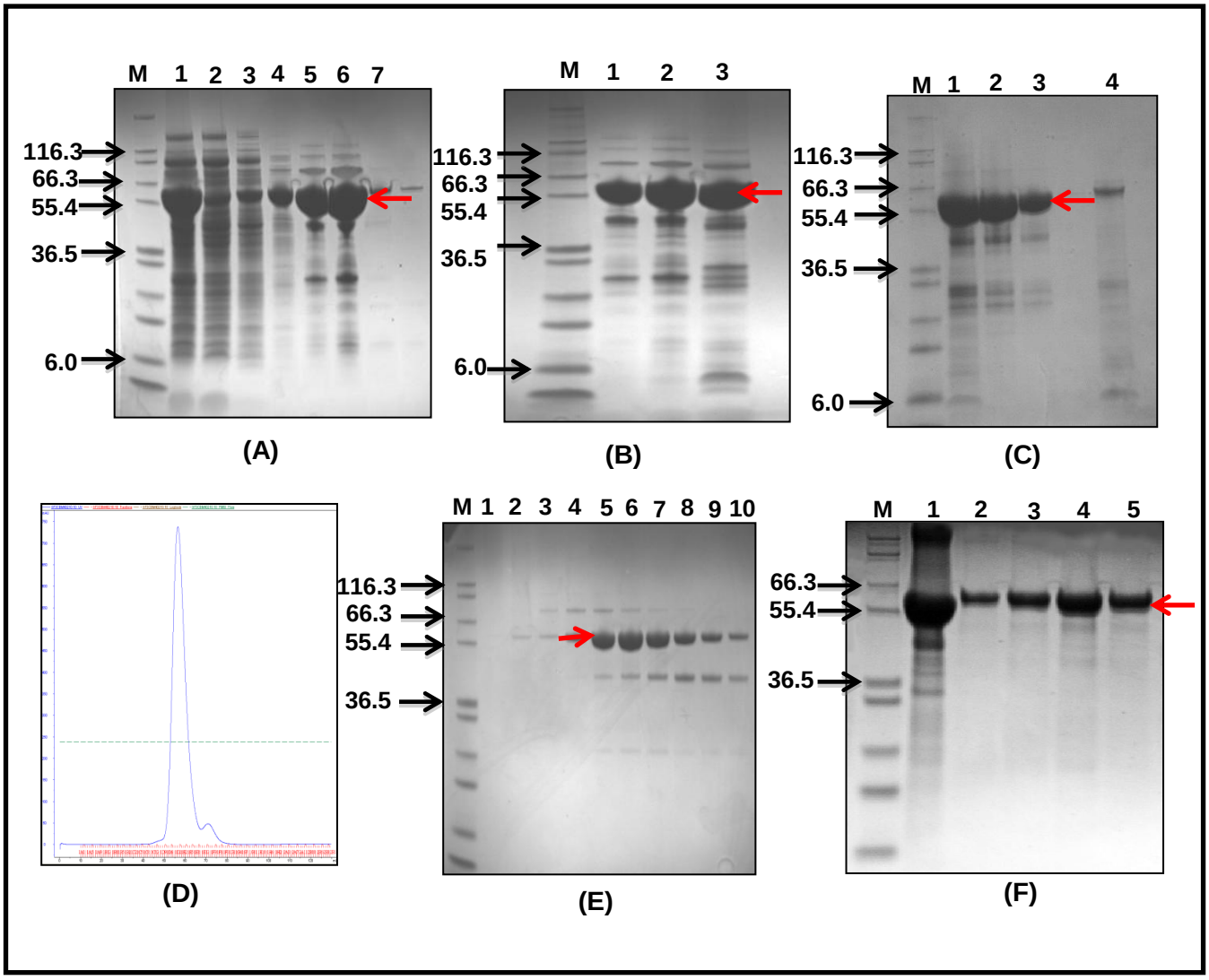


Figure 2.3 Expression and purification of Vc3CBM. (**A**) First Ni^{2+}- affinity column purification. M, protein marker (Mark12 ™, Invitrogen); lane 1, cell lysate after IPTG induction; lane 2, flow-through (unbound protein) after sample loading onto the column; lane 3-4, column wash with the buffer containing 10 mM and 20 mM imidazole; lane 5-7, elution in buffer containing 250 mM imidazole. (**B**) TEV protease digestion. Lane 1-2, undigested sample of Vc3CBM, lane 3, fully digested Vc3CBM at 4 °C, overnight. (**C**) Second Ni^{2+}-affinity column. Lane 1, TEV digested sample, lane 2-3, the digested sample after wash with 10 mM and 20 mM imidazole containing buffer. Lane 4, elution of undigested protein with 250 mM imidazole. (**D**) Gel filtration chromatogram (**E**) Eluted protein after gel filtration chromatography. (**F**) Lane 1, Vc3CBM before anion exchange column, Lane 2-5, Eluted protein after anion-exchange chromatography. Protein samples were run on 4-12% NuPAGE gel (Life Technologies). Red arrow represents protein of interest, Vc3CBM.

2.4.2 Expression and purification of Vc4CBM

For Vc4CBM production, the same expression strategies as Vc3CBM were used (1 mM IPTG induction at 18°C for overnight) with a heat shock step at 42°C for 30 minutes after the culture reached an OD_{600} of 0.6. Expressed Vc4CBM proteins were purified via a four-step purification strategy (nickel affinity, gel filtration and anion exchange chromatography). For anion-exchange purification, a Hi-Trap Q (FF) column (GE Healthcare) was then attached onto an AKTA purification system (GE Healthcare) and the protein was eluted with a linear gradient of increasing NaCl concentration. Most of the contaminants present in Vc4CBM (Figure 2.4 (E)) were successfully removed by anion-exchange chromatograpy. As seen in Figure 2.4 (F), no obvious contaminants could be seen from the SDS-PAGE gels after the anion-exchange purification. In order to make sure we are working with the correct protein, a gel fragment of Vc4CBM was cut out from the SDS-PAGE and sent for mass spectrometry analysis. The expected molecular weight of untagged Vc4CBM was found to be 85 kDa with a yield of pure protein to be approximately 30 - 40 mg from 2 L cultures.

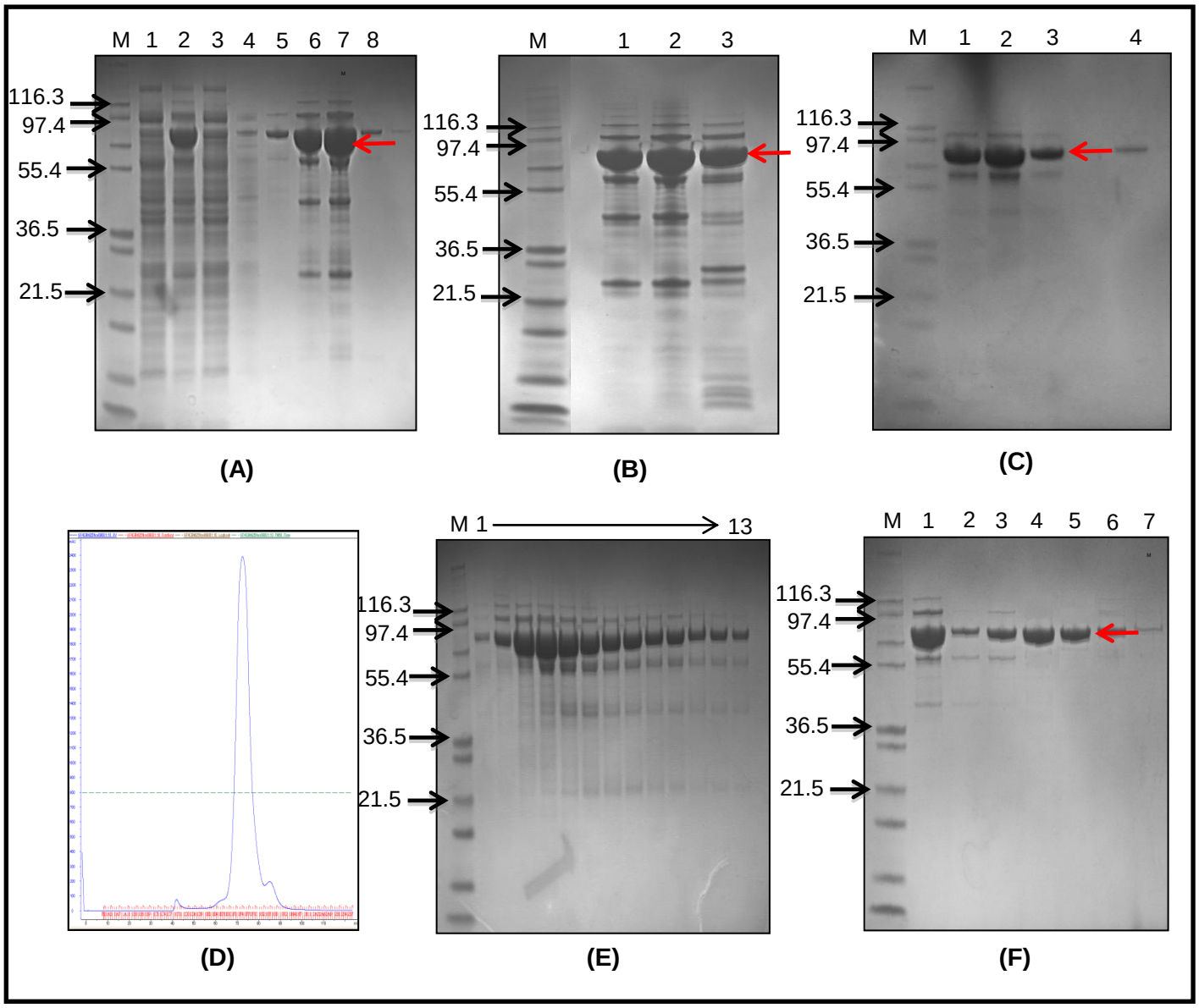


Figure 2.4 Expression and purification of Vc4CBM. (**A**) First Ni^{2+}-affinity column purification. M, protein marker (Mark12 ™, Invitrogen); lane 1, flow-through (unbound protein) after sample loading on to the column; lane 2, cell lysate after IPTG induction; lane 3-4, column wash with the buffer containing 10 mM and 20 mM imidazole; lane 5-7, elution in buffer containing 250 mM imidazole. (**B**) TEV protease digestion. Lane 1-2, undigested sample of Vc4CBM, lane 3, fully digested Vc4CBM at 4°C, overnight. (**C**) Second Ni^{2+}-affinity column. Lane 1, TEV digested sample, lane 2-3, the digested sample after wash with 10 mM and 20 mM imidazole containing buffer. Lane 4, elution of undigested protein with 250 mM imidazole. (**D**) Gel filtration chromatogram. (**E**) Eluted protein after gel filtration chromatography. (**F**) Lane 1, Vc4CBM before anion exchange column, Lane 2-7, Eluted protein after Anion-exchange chromatography. Protein samples were run on 4-12% NuPAGE gel (Life Technologies). Red arrow represents protein of interest, Vc4CBM.

2.5 Engineering of multivalent CBM40 using an oligomerization domain (TD) from *Pseudomonas aeruginosa* pseudaminidase.

Pseudomonas aeruginosa is reported to encode an enzyme mimic sialidase known as pseudaminidase. This enzyme is structurally related to sialic acid but has a different stereochemistry at the 5-, 7-, and 8-positions and normally found in lipopolysaccharide (LPS) and on pili of *Pseudomonas aeruginosa.* The enzyme has a six-bladed β-propeller catalytic domain followed by a C-terminal of a small domain (Figure 2.5) with an immunoglobulin (Ig)-like fold that self-oligomerizes to form a trimer (Figure 2.6). This Ig-like domain is described as a trimerization domain or oligomerization domain by Xu et al. (2009a).

In this study, the trimerization domain was exploited to generate trimeric CBM molecules by subcloning the gene fragment encoding the trimerization domain (TD) and ligating it to the CBM40 domain to study its multivalent effect and affinity towards sialic acid. With this idea, we used the TD domain and fused it in tandem with a CBM40 domain either from the *V. cholerae* sialidase or the *Streptococcus pneumoniae NanA* sialidase to create two constructs designated as Vc-CBMTD and Sp-CBMTD.

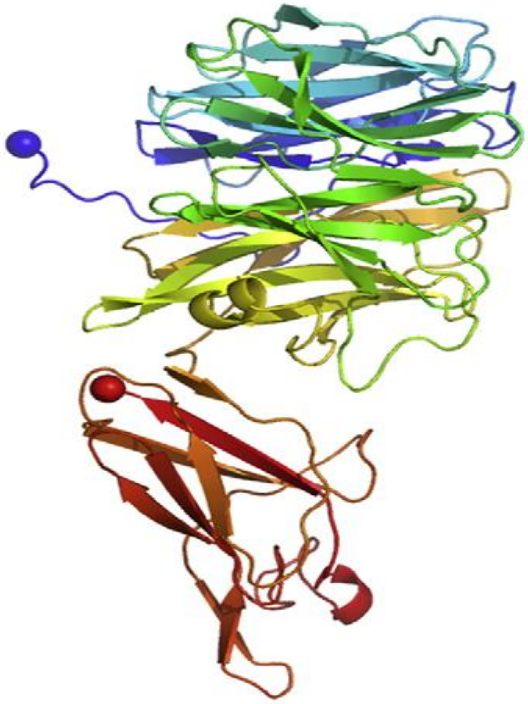


Figure 2.5 View of the pseudaminidase monomer from *Pseudomonas aeruginosa*. The blue and red spheres indicating the N- and C-terminal respectively. The C-terminal domain (red) is known as the trimerization domain (Xu et al., 2009a).

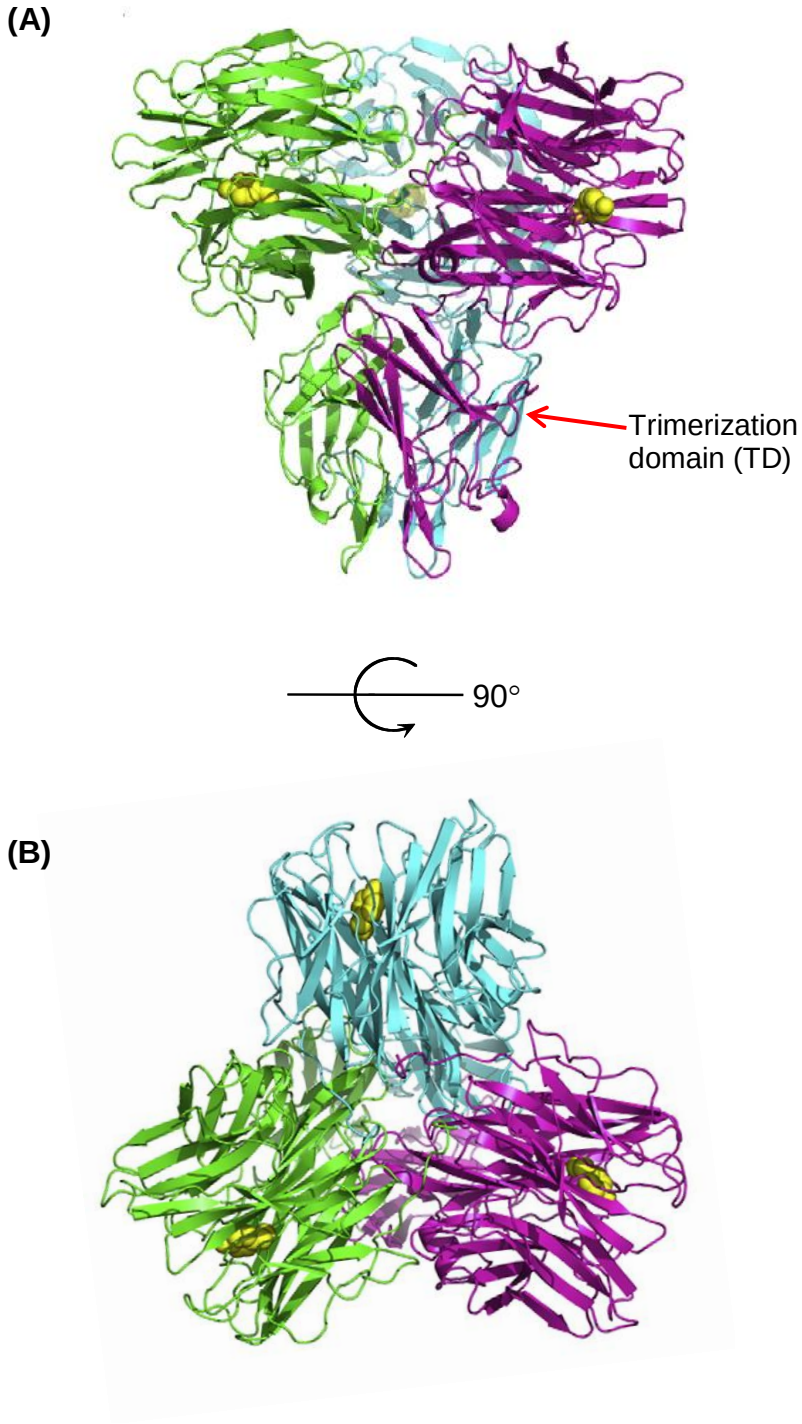


Figure 2.6 Schematic view of pseudaminidase trimer from *Pseudomonas aeruginosa* (2W38). (A) Views from the side of the molecule; **(B)** Views from the three-fold symmetric axis. The protein shown as trimer due to the presence of the trimerization domain (TD).

2.5.1 Development of Vc-CBMTD construct from *Vibrio cholerae* sialidase

2.5.1.1 Amplification of the pseudaminidase trimerization domain from *Pseudomonas aeruginosa* (PaTD)

Amplification of the trimerization domain (TD) from *P. aeruginosa* was undertaken with a set of primers designated as PaTD_BamHI (F): 5'- GGT GGA TCC GTC CCG GAT TTT GAG TCA G -3' and PaTD_XhoI(R): 5'- CCG ACT CGA GCT AAA TCC ATG CTC TGA CCC G -3'. Furthermore, primers used to amplify Vc-CBM were listed as in Table 2.2. Both amplifications of PaTD and Vc-CBM domains used the same amplification protocol as described below.

PCR was carried out in a final volume of 50 µl reaction mixture containing 10 µM dNTP, 1U Pfu DNA Polymerase (Promega), 1X Pfu Buffer, 1 µg template DNA from a recombinant PaNA-pEHISTEV plasmid and 1.0 µM forward and reverse primers. Typically, DNA amplification was started with an initial denaturation at 95°C for 30 seconds followed with 30 cycles of 95°C for 30 seconds, 53°C for 1 minute, 72°C for 30 seconds and a final cycle at 72°C for 7 minutes. The amplified fragments were cleaned with Wizard SV PCR Clean-Up System. The amplified gene was digested with the appropriate restriction enzymes before ligation to the pEHISTEV vector.

Ligation was carried out as followed. The PCR products were directly cloned into the vector after digested with *Xho*1 and *Bam*H1. The vector: insert ratio of 1:3 was used. About 50 ng of pEHISTEV vector was mixed with 1 µg of PCR products, 1X T4 DNA ligase buffer and 1 U of T4 DNA ligase. Sterile distilled H_2O was added up to 20 µl final volume followed with incubation at 22°C for 1 hour. The mixture was then centrifuged briefly and used to

transform *E. coli* DH5α competent cells. Positive clones were identified by isolating DNA and sent for DNA sequencing (University of Dundee, DNA Sequencing Service).

Table 2.2 Summary of oligonucleotide primers used to amplify Vc-CBM and PaTD.

Primer	Oligonucleotide sequence (5'- 3')
Vc-CBM (F)	GGCTCCATGGCACTTTTTGACTATAACGC (NcoI)
Vc-CBM (R)	GCACGGATCCACCACCGTCGCCTTGAATTTC (BamHI)
PaTD (F)	GGTGGATCCGTCCCGGATTTTGAGTCAG (BamHI)
PaTD (R)	CCGACTCGAGCTAAATCCATGCTCTGACCCG (XhoI)

Restriction site sequence shown in each primer is underlined.

Results from the successful clones carrying both PaTD and CBM40 domains were shown in Figure 2.7 and Figure 2.8. PCR amplifications were performed to determine the presence of PaTD and CBM40 gene fragments in clones using appropriate primers outlined in Table 2.2. Figure 2.7 and Figure 2.8 showed an insert size of 321 bp and 588 bp, which corresponded to the TD domain of *P. aeruginosa* pseudaminidase and CBM40 domain from *V. cholerae* sialidase, respectively. For further confirmation of the constructs, both plasmids carrying the genes were sent out for sequencing (University of Dundee, DNA Sequencing Service). For long term storage, glycerol stocks of both clones were prepared in 20% glycerol and stored at - 80°C.

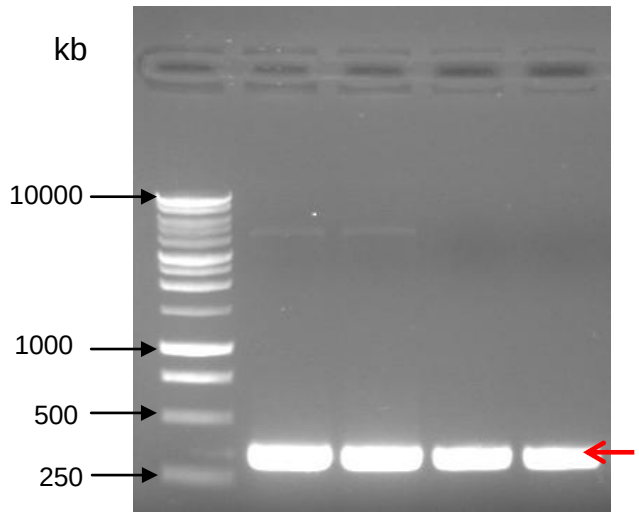


Figure 2.7 PCR amplification of *P. aeruginosa* trimerization domain (PaTD) from postitive clones. Bands size of 321 bp corresponds to the PaTD domain on 0.8% agarose gel, as denoted by a red arrow.

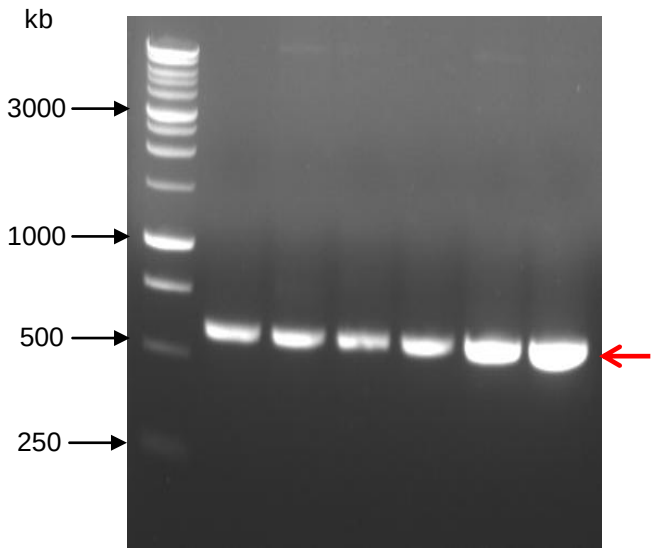


Figure 2.8 PCR amplification of CBM40 domain from *V. cholerae* sialidase. The CBM40 domain showed size of 588 bp on 0.8% agarose gel, as marked with red arrow.

2.5.2 Site-directed mutagenesis of CBM40

Our next approach was to mutate two amino acids in the CBM40 domain of Vc-CBMTD in order to introduce a disulphide bond between amino acids 83 and 164, which required the following mutations, S164C and T83C. The choice of serine (S) and threonine (T) mainly because they were surface exposed and on opposite sides of the CBM40 domain, which we had speculated may form disulphides bridge that would lock the orientation of the CBM40 ddomains. The idea was to create a trimer where all three domains were oriented in the same direction such that they would face the cell surface and might help to increase the affinity of the domain towards sialic acid (Figure 2.9).

Mutagenesis was carried out using the QuickChange Site-Directed Mutagenesis Kit (Stratagene) following the manufacturer's instructions. Constructs containing the carbohydrate-binding module (CBM40) from *Vibrio cholerea* sialidase fused with the trimerization domain from *Pseudomonas aeruginosa* pseudaminidase, labelled as CBMTD, were used for this study. Based on structural analysis of the CBM40 domain, site-specific mutations were introduced at S164C and T83C to allow the formation of a disulfide bond between both amino acids.

The method relies on two complement PCR primers, which incorporates both the mutations at the selection site and the desired single base substitutions at the mutant site. The pair of primers used were CBMS164C (F): 5'-GAA TTG GTA TTC CTT CCT GGA TGT AAC CCA TCC GCT AGC TTT TAC-3' and CBMT83C (F): 5'-CAA TGG ACA TAT TCT CTC TCT TGC AAT CAA CAT GCC CAA GCA TC-3'. The integration of CBM40 and PaTD domain (Figure 2.10) were confirmed by colony PCR using primers as described previously in Table 2.2. From Figure 2.10, the presence of both genes encoding the domains (CBM40 and PaTD domain) in the contructs were confirmed with correct insert sizes of 588 bp and 321 bp

corresponding to each domains size, respectively. Mutants were confirmed by DNA sequencing service at University of Dundee. Sequencing analysis has revealed two amino acid substitutions, which were S164C and T83C as shown in amino acid aligments from both mutant and wild type version of Vc-CBMTD (Figure 2.11).

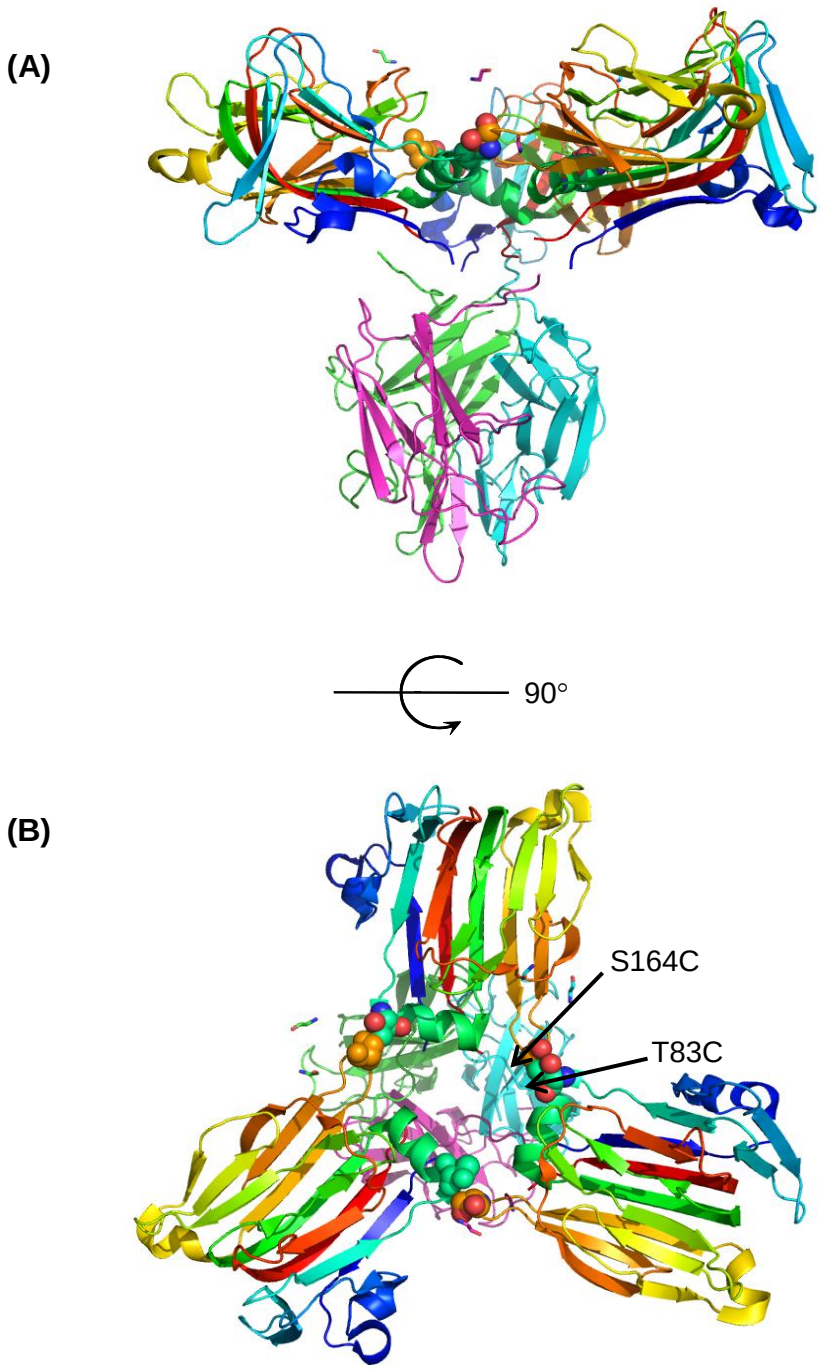


Figure 2.9 Model of Vc-CBMTD (Mutant) shown as trimer. Molecules consist of the three identical CBM40 domains from *V. cholerae* with three identical TD domains *P. aeruginosa*. **(A)** views from the side of the molecule; **(B)** views from the three-fold symmetric

axis. Molecule was constructed and built using PyMol software. The space filling compounds represent amino acids involved in the mutation (S164C and T83C).

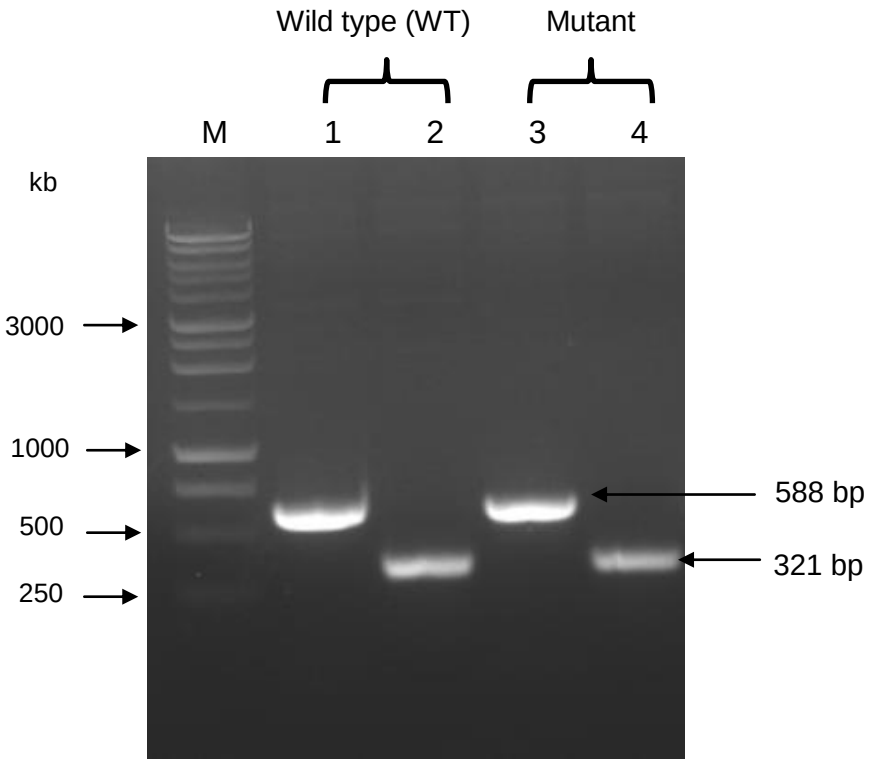


Figure 2.10 Colony PCR amplification of Vc-CBMTD (WT) and Vc-CBMTD (Mutant) clones. Two bands appeared correspond to CBM40 (588 bp) and PaTD (321 bp) domain. Samples were run on 0.8% agarose gel.

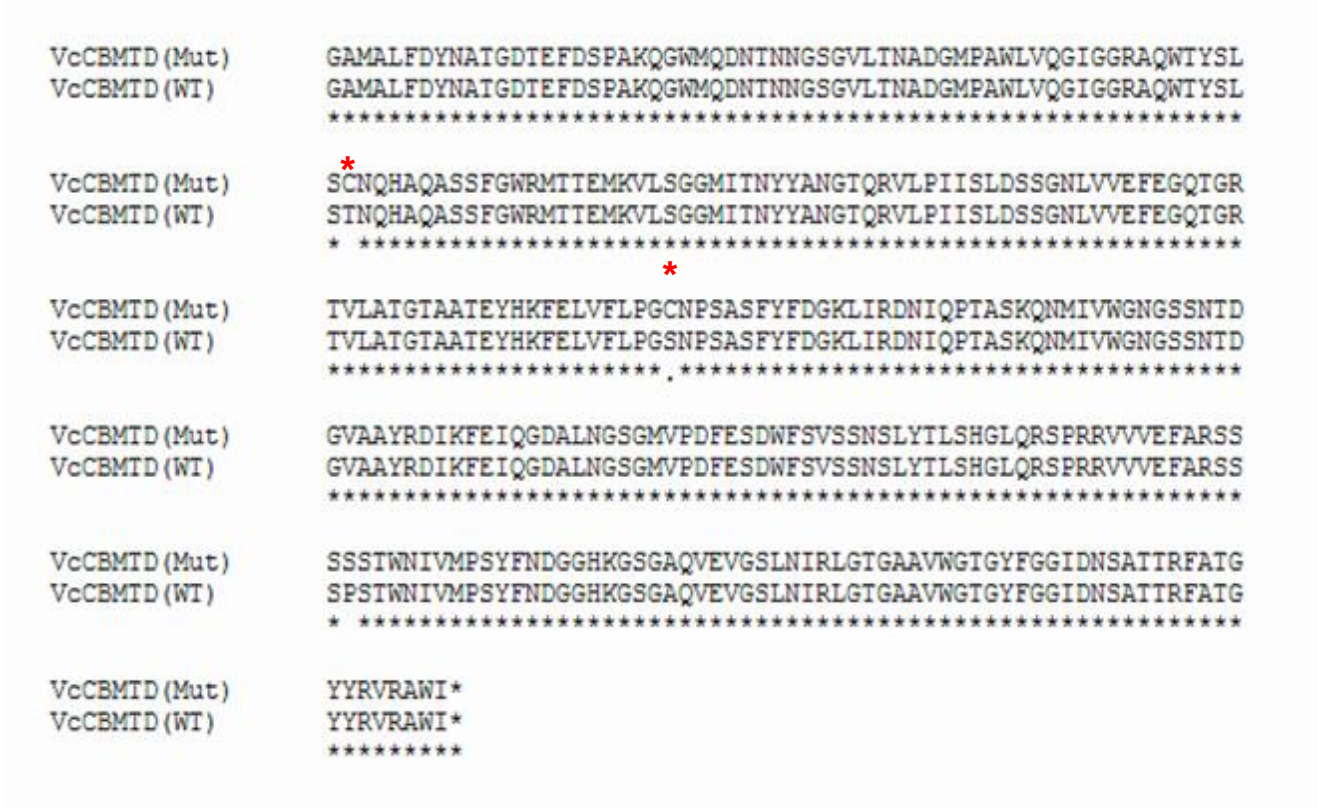


Figure 2.11 Amino acid alignments between wild type (WT) and mutant of Vc-CBMTD. Amino acid replacement at S164C and T83C were shown with red asterisks. Sequence alignment was done with Cluster Omega version 1.1.0 (EMBL-EBI).

2.5.3 Protein expression and purification of Vc-CBMTD (WT) and Vc-CBMTD (Mutant)

Expression and purification protocols of both Vc-CBMTD (WT) and Vc-CBMTD (Mutant) were as described in Section 2.2.1 and Section 2.2.2. For the purpose of studying the expression of these proteins, samples were taken at several time points, which were at 0, 6 and 18 hours after induction with 1 mM IPTG. Both pellets and cell supernatants collected were run through SDS-PAGE to study the expression pattern of the protein.

It was shown that in the absence of heat shock treatment (42°C for 30 minutes), most of the proteins were expressed insolubly despite using 18°C as the induction expression temperature (Figure 2.12). As expected, the size of both Vc-CBMTD was observed as 33 kDa on SDS-PAGE, which is the monomer of Vc-CBMTD. From Figure 2.12 (B), there was some evidence of slightly soluble protein expressed after 6 and 18 hours of incubation. However, compared to the cell pellet fractions, a significant amount of the proteins remained in an insoluble state. It is interesting to find out, after incorporation of the heat shock step (42°C for 30 minutes) most of the proteins were expressed soluble as shown in Figure 2.13. The over-expressed protein band was later cut out for protein identification by mass spectrometry.

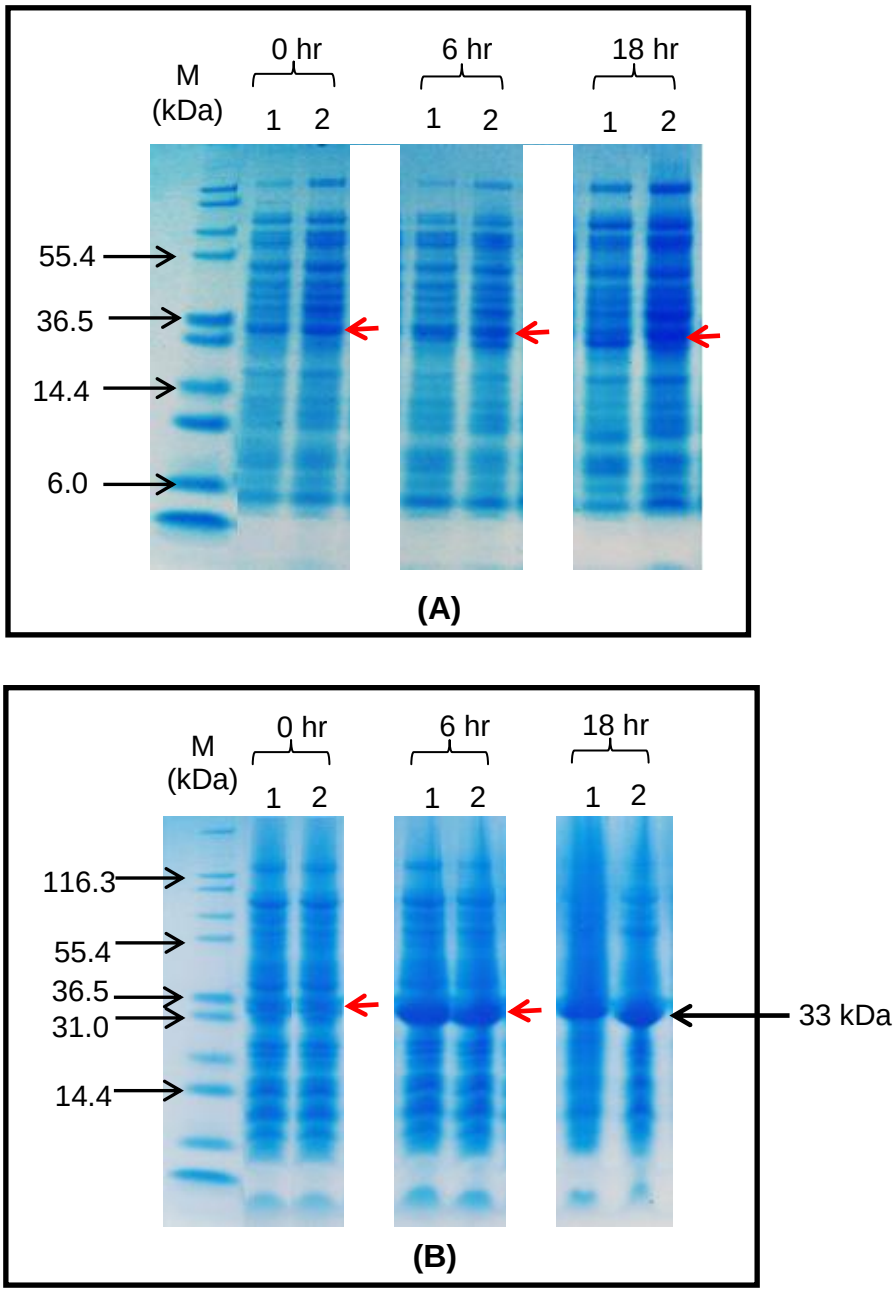


Figure 2.12 Expression of Vc-CBMTD (WT) and Vc-CBMTD (Mutant) in LB broth before optimization steps. Fractions of samples were taken at 0 hr, 6 hr and 18 hr during expression at 18 °C. Lane 1, sample from Vc-CBMTD (WT); Lane 2**,** sample from Vc-CBMTD (Mutant) **(A)** Soluble fractions from supernatant of Vc-CBMTD (WT) and Vc-CBMTD (Mutant). **(B)** Insoluble fraction from cell pellet of Vc-CBMTD (WT) and Vc-CBMTD (Mutant). Protein samples were run on 4-12% NuPAGE gel (Life Technologies). Small red arrow represents protein of interest.

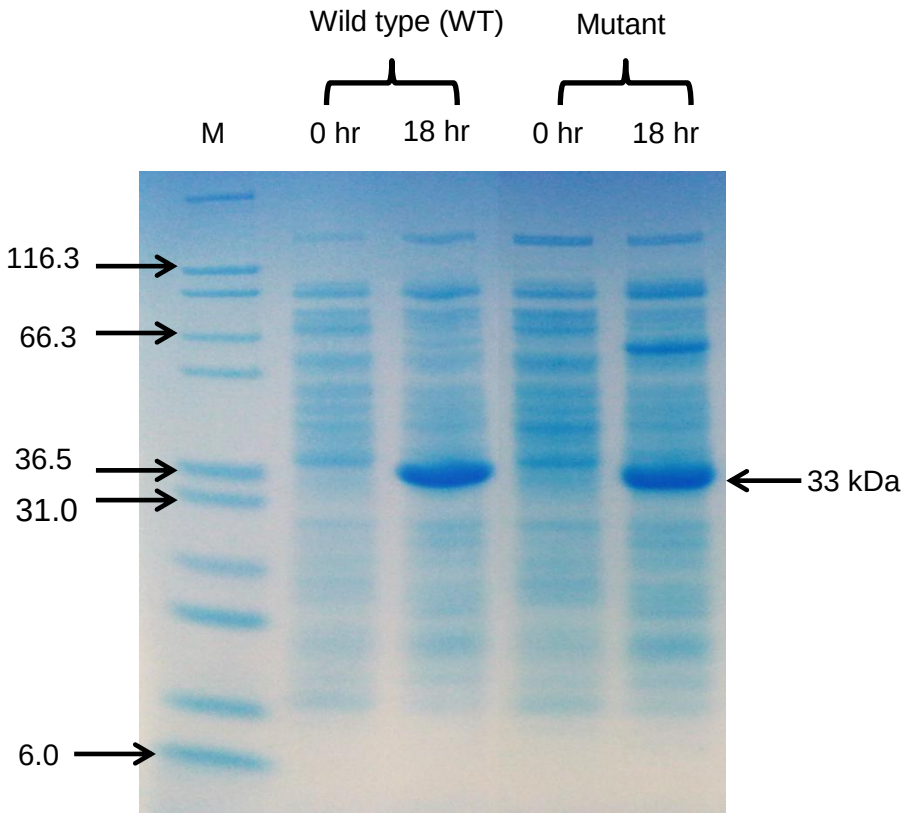


Figure 2.13 Soluble expressions of Vc-CBMTD (WT) and Vc-CBMTD (mutant) after heat shock treatment. Both proteins were expressed at 0 hour and 18 hours in Luria-Bertani broth. Protein samples were run on 4-12% NuPAGE gel (Life Technologies). The expressions of both proteins were denoted by red asterisk.

2.5.3.1 Purification of Vc-CBMTD (WT) and Vc-CBMTD (Mutant)

For the purification stages, both proteins, Vc-CBMTD (WT) (Figure 2.14) and Vc-CBMTD (Mutant) (Figure 2.15) were purified using the same protocols mentioned in Section 2.2.2. After two nickel affinity chromatography steps, followed by final gel filtration with (HiPrep 16/60 Sephacryl S-100HR) column, both proteins were purified to good quality despite having some aggregated or polymeric form of the protein (Figure 2.14 (D) and Figure 2.15 (D)). These features were clearly shown on Figure 2.15 (A) with the presence of two additional bands due to its trimerization effect on SDS-PAGE gel. The effect was far more obvious for the mutant version compared to Vc-CBMTD (WT) as shown in Figure 2.15 (A) due to the rigid folding of domains with the introduction of a disulphide bond. Purified proteins were analysed by mass spectrometry, confirming the identity of the target proteins as shown in Figure 2.14 (E) and Figure 2.15 (E). Fractions containing purified proteins were concentrated using a 30,000-Da MWCO centrifugal concentrator (Millipore) for crystallization trials later on. Usually, for both proteins, around 20-25 mg of protein can be purified per 2 L cultures.

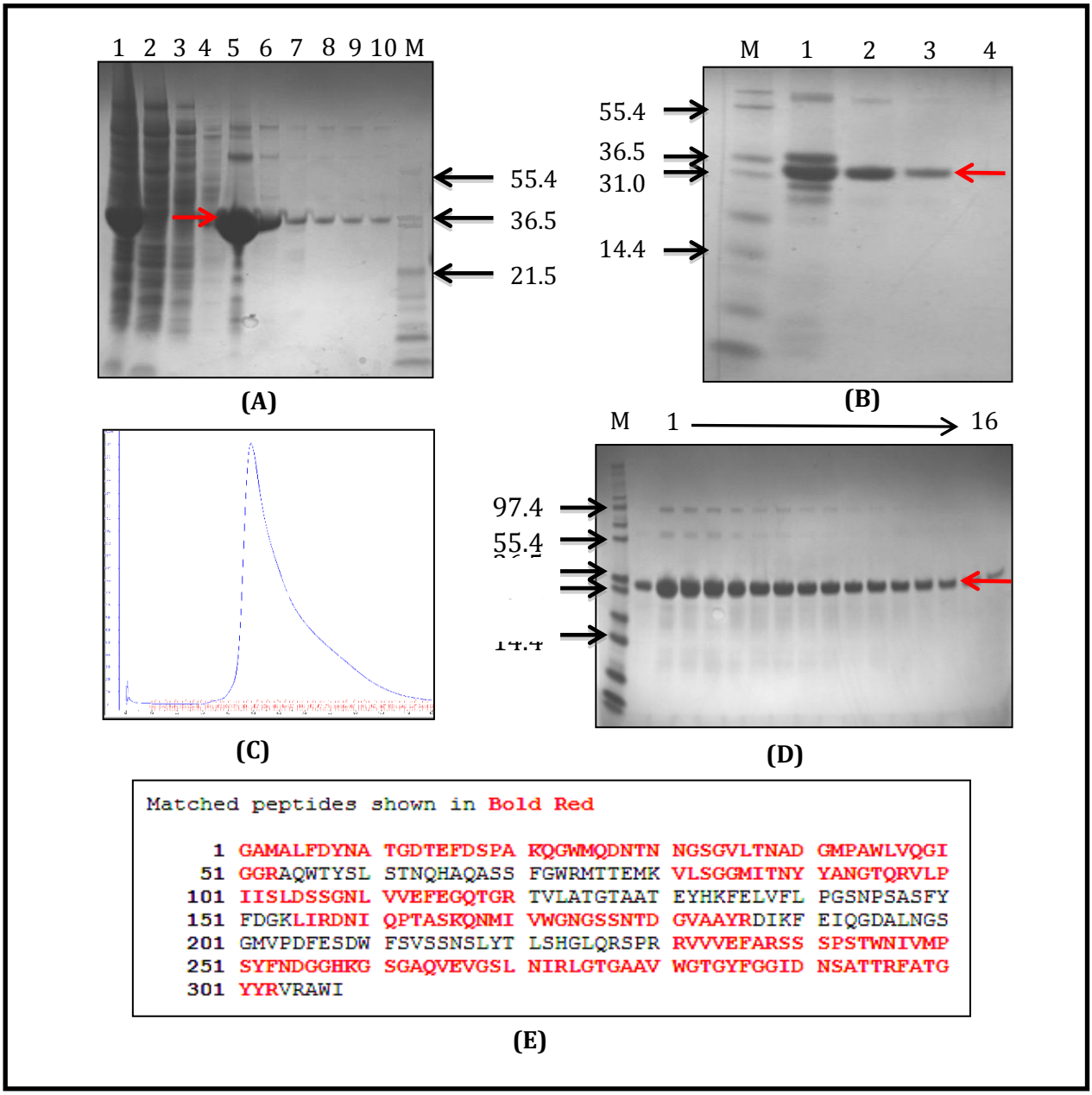


Figure 2.14 Expression and purification of Vc-CBMTD (WT). (**A**) First Ni^{2+}-affinity column purification. lane 1, cell lysate after IPTG induction; lane 2, flow-through after loading sample onto the column; lane 3-4, column wash with the buffer containing 10 mM and 20 mM imidazole; lane 5-10, elution in buffer containing 250 mM imidazole; M, protein marker (Mark12 ™, Invitrogen). (**B**) Second Ni^{2+}-affinity column after completed TEV protease digestion. Lane 1, TEV digested sample, lane 2-3, the sample after wash with 10 mM and 20 mM imidazole containing buffer. Lane 4, elution of undigested protein with 250mM imidazole.(**C**) Gel filtration chromatogram for Vc-CBMTD (WT) protein. (**D**) Eluted protein after gel filtration chromatography. (**E**) Matched peptides of Vc-CBMTD (WT) identified by mass spectrometry. Protein samples were run on 4-12% NuPAGE gel (Life Technologies). Red arrow represents protein of interest, Vc-CBMTD (WT).

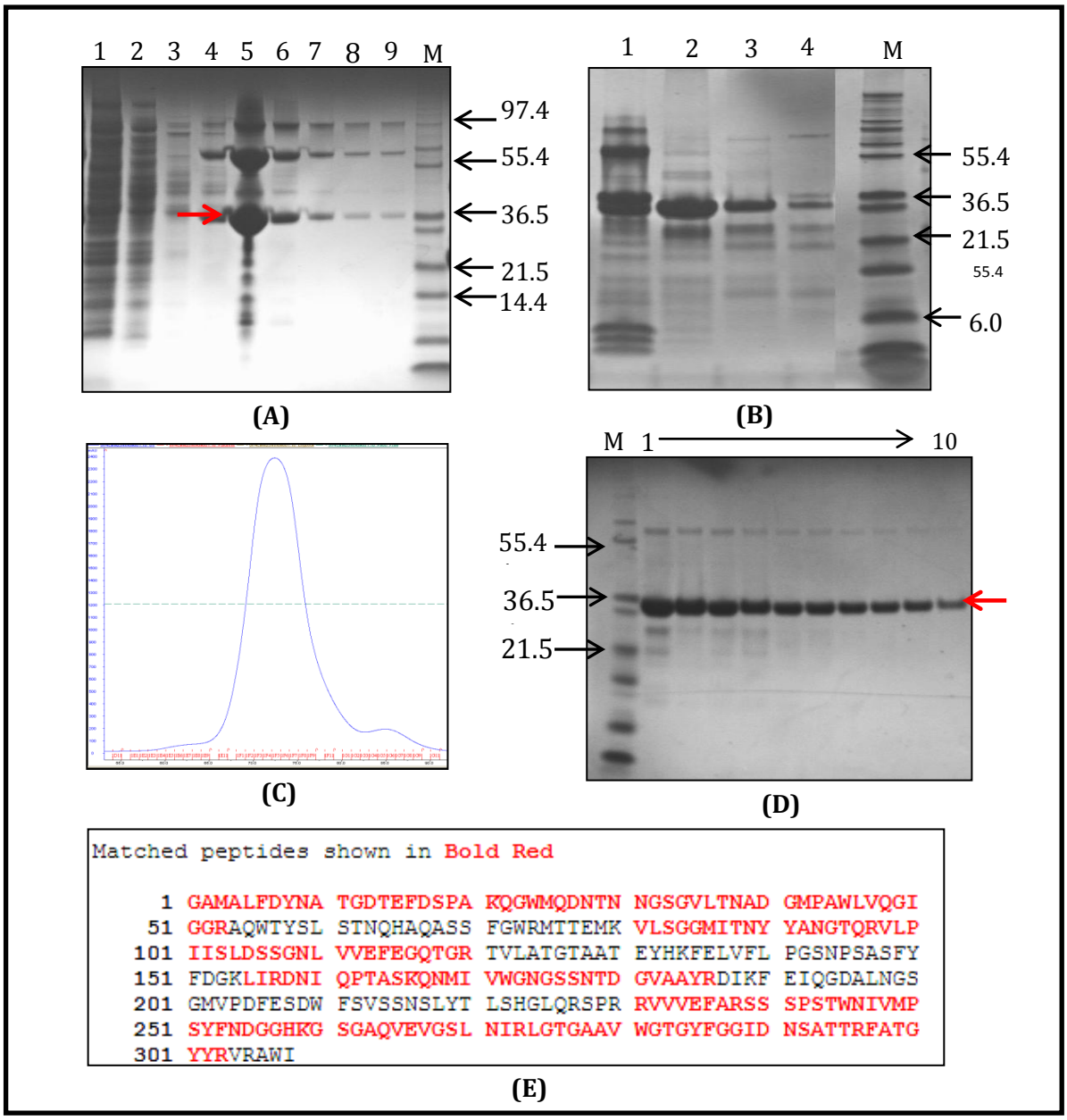


Figure 2.15 Expression and purification of Vc-CBMTD (Mutant). (**A**) First Ni^{2+}- affinity column purification. lane 1, flow-through (unbound protein) after loading sample onto the column; lane 2-3, column wash with the buffer containing 10 mM and 20 mM imidazole; lane 4-9, elution in buffer containing 250 mM imidazole; M, protein marker (Mark12 ™, Invitrogen). (**B**) Second Ni^{2+}-affinity column after completed TEV protease digestion. Lane 1, TEV digested sample, lane 2-3, sample after wash with 10 mM and 20 mM imidazole containing buffer. Lane 4, elution of undigested protein with 250 mM imidazole. (**C**) Gel filtration chromatogram for the protein. (**D**) Eluted protein after gel filtration chromatography. (**E**) Matched peptides of Vc-CBMTD (mutant) identified by mass spectrometry. Protein samples were run on 4-12% NuPAGE gel (Life Technologies). Red arrow represents protein of interest, Vc-CBMTD (Mutant).

2.6 Development of Sp-CBMTD-GFP constructs using the CBM40 domain from *Streptococcus pneumoniae NanA* sialidase

The *Streptococcus pneumoniae* genome encodes three sialidases (*NanA*, *NanB* and *NanC*), which are known to be responsible in removing sialic acid from host cell surface in order to unmask receptors to facilitate bacterial adherence and colonization (Mitchell, 2000). In addition to the catalytic domain, all three sialidases have a lectin-like domain at the N-terminus known as CBM40 where they are predicted to recognize sialic acid moiety. It is also known that, *NanA* and *NanB* are identified as *Pneumococcal* virulence factors and are potential drug targets (Manco et al., 2006; Parker et al., 2009; Pettigrew et al., 2006).

It has been reported by Uchiyama et al. (2009), that *NanA* is anchored to the bacterium's outer membrane and its CBM domain is responsible in enabling the bacterium to cross the blood brain barrier. Despite the fact that their function is not well understood, we believe that they might play cooperative roles in *S. pneumococcal* virulence activity. From this point of view, we have engineered a molecule consisting of CBM40 domain from *S. pneumoniae NanA* sialidase fused with trimerization domain (TD) from *P. aeruginosa* labeled as Sp-CBMTD-GFP to study its effect in multivalent approach. This will be further discussed in next subsection.

2.6.1 Amplification of *Streptococcus pneumoniae* CBM40

The amplification protocol for the *P. aeruginosa* trimerization domain was similar as described in Section 2.6.1, unless stated otherwise and using an annealing temperature of 53 °C. The primer set for the amplification was as followed:

Sp-CBM (F): 5'-GGCTCCATGGTGATAGAAAAAGAAG-3' (NcoI)
Sp-CBM (R): 5'-CCGACTCGAGCTAAATCCATGCTCTG-3' (XhoI)

The DNA fragment encoding Sp-CBM was modified at the 5'- and 3'- termini to incorporate restriction sites to allow ligation of this domain to the trimerization domain of *P. aeruginosa*. The resulting fragments were cloned into an appropriately digested pEHISTEV vector before transforming to *E. coli* DH5α cells. For this construct, we have included a GFP tag as an additional tag to aid in protein solubility as well as to assist in our cell based work.

In order to check for the presence of both domains (CBM40 and TD domains), PCR amplifications were carried out using plasmid pEHISTEV containing Sp-CBMTD gene as template DNA. Suitable primers were used as mentioned in description of Figure 2.16. For the CBM40 domain, a larger size insert (> 1.5 kb) was detected as supposed to 588 bp, due to the use of T7 primers. The size includes the presence of CBM40 domain with GFP-tag in the construct. For the TD domain, the insert size was calculated to be 321 bp, which corresponded to the correct size. Moreover, to identify positive clones, plasmids were isolated and sent for sequencing using the University of Dundee sequencing service (www.dnaseq.co.uk).

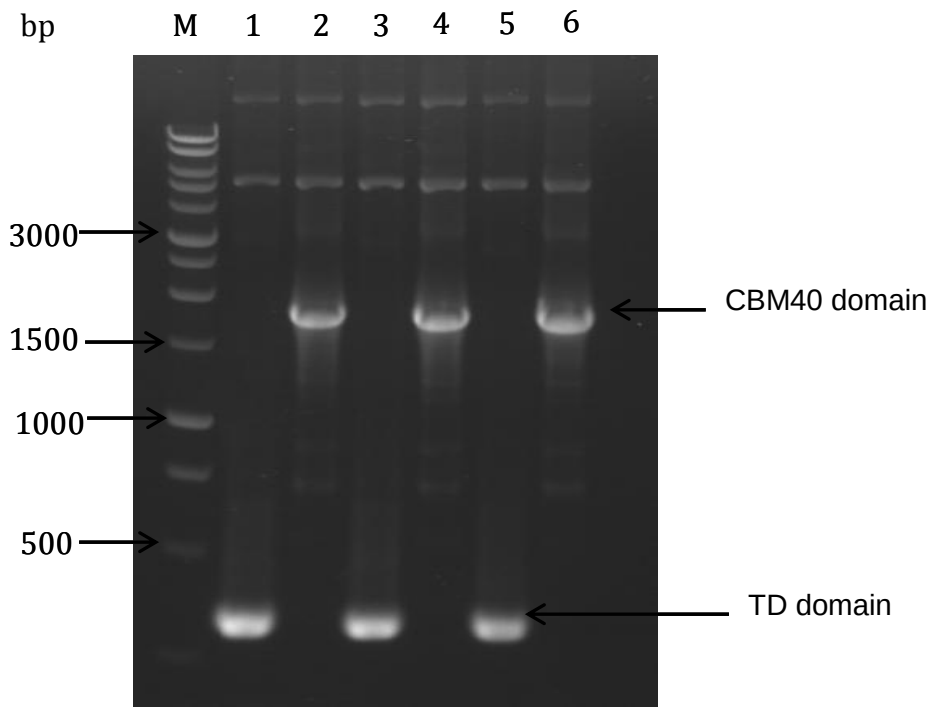


Figure 2.16 PCR amplification of CBM40 and trimerization domain from positive clones. The amplification of CBM40 domain here was using T7 primers, while for TD domain a set of TD primers, PaTD (F) and PaTD (R) were used as described in Table 2.2. Samples were run on 0.8% agarose gel.

2.6.2 Expression and purification of Sp-CBMTD-GFP

In order to express the Sp-CBMTD-GFP construct, successful clones were transformed into protein expression host strain *E. coli* BL21 (DE3). During the expression process, Sp-CBMTD fused with GFP tag was expressed and cultures were observed to have a vibrant green fluorescent colour, due to the presence of GFP protein in the construct. Later, the cultures were then scaled up to 4 L (Figure 2.17 (I)) and induced with 1 mM of IPTG for increase protein yields. Details of the expression and purification methods of the protein were as described in Section 2.2.1 and Section 2.2.2 unless otherwise stated.

After two stages of nickel column purification, histidine tag digestion, followed by gel filtration chromatography, a band size of 33 kDa was obtained on the gel, which corresponded to the the size of Sp-CBMTD-GFP monomer (Figure 2.17 (II)). When using SDS-PAGE, protein that forms trimer will only show one band in an SDS-PAGE, which corresponds to the protein monomer (provided that denaturation is complete). Protein purity was analysed by SDS-PAGE and its identity was confirmed by mass spectrometry (Figure 2.18). The protein samples were concentrated further using a 30,000-Da MWCO centrifugal concentrator (Milipore) prior to storage at -80°C. Usually for 2 L cultures, about 16 mg of pure protein can be obtained.

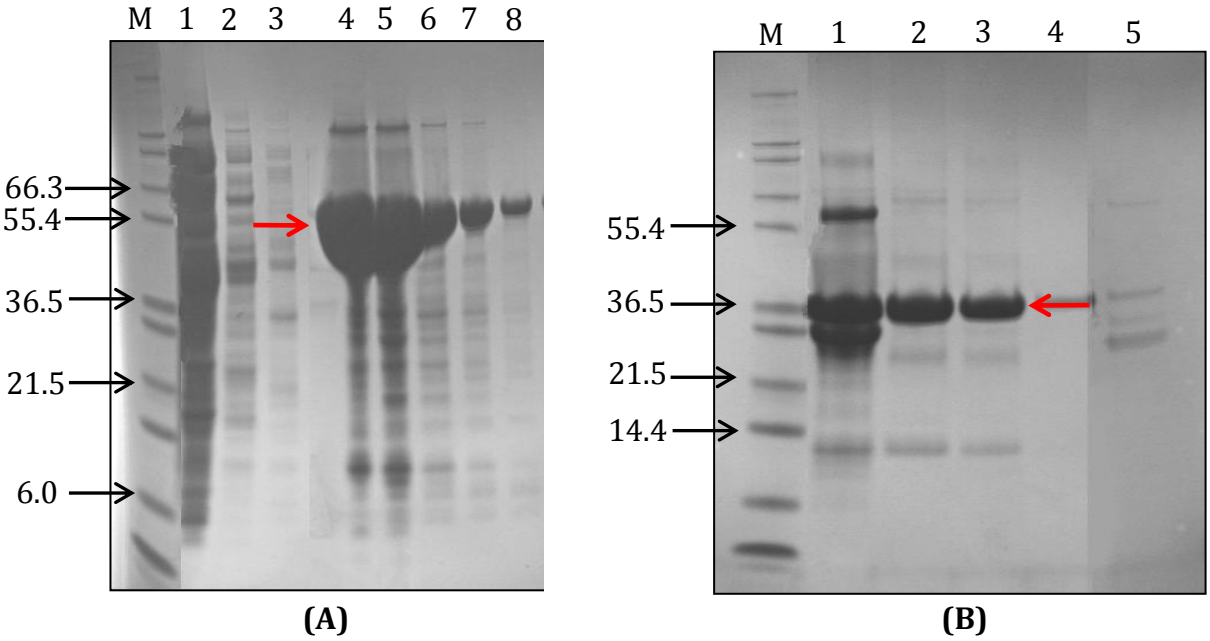
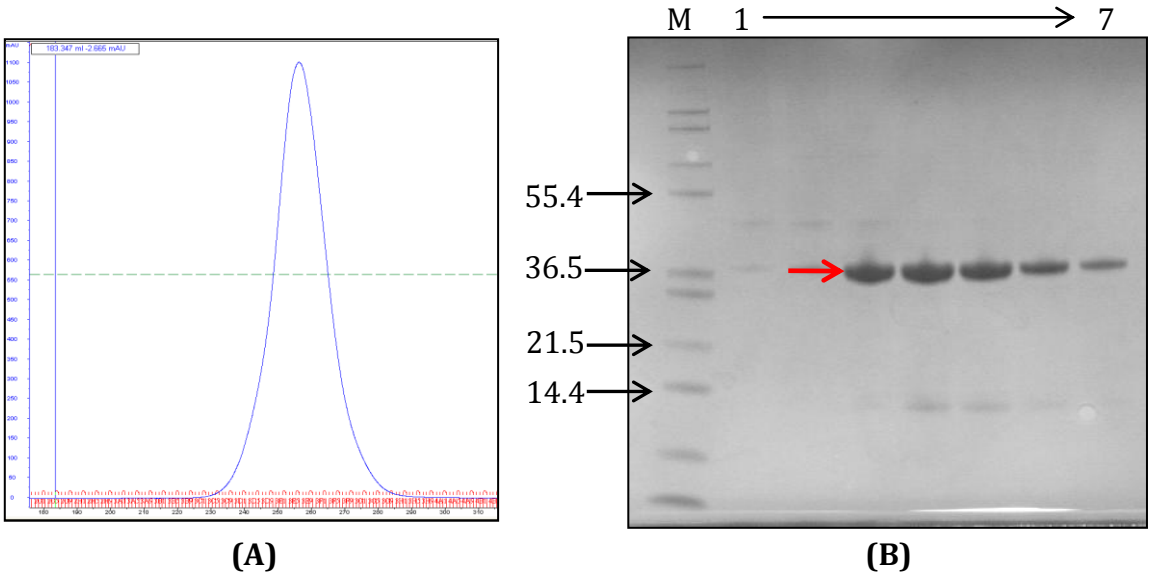


Figure 2.17 (I) Expression and Ni^{2+}-affinity column purification of Sp-CBMTD-GFP in *E. coli* BL21 (DE3). (**A**) First Ni^{2+}- affinity column purification. lane 1, flow-through (unbound Sp-CBMTD-GFP) after loading sample onto the column; lane 2-3, column wash with the buffer containing 10 mM and 20 mM imidazole; lane 4-8, elution in buffer containing 250 mM imidazole; M, protein marker (Mark12 ^{TM}, Invitrogen). (**B**) Second Ni^{2+}-affinity column after completed TEV protease digestion. Lane 1, TEV digested sample, lane 2-4, the digested sample after wash with 10 mM and 20 mM imidazole containing buffer. Lane 5, elution of undigested protein with 250 mM imidazole. Protein samples were run on 4-12% NuPAGE gel (Life Technologies). Red arrow represents protein of interest.



Figue 2.17 (II) Gel filtration chromatography of Sp-CBMTD-GFP. (A) Gel filtration chromatogram for Sp-CBMTD-GFP protein. **(B)** Eluted Sp-CBMTD-GFP protein after gel filtration chromatography. M, protein marker (Mark12 ™, Invitrogen), Lane 1-7, protein elution from the column. Protein samples were run on 4-12% NuPAGE gel (Life Technologies). Red arrow represents protein of interest.

Matched peptides shown in Bold Red

```
     1 MVIEKEDVET NASNGQRVDL SSELDKLKKL ENATVHMEFK PDAKAPAFYN
    51 LFSVSSATKK DEYFTMAVYN NTATLEGRGS DGKQFYNNYN DAPLKVKPGQ
   101 WNSVTFTVEK PTAELPKGRV RLYVNGVLSR TSLRSGNFIK DMPDVTHVQI
   151 GATKRANNTV WGSNLQIRNL TVYNRALTPE EVQKRS
```

Figure 2.18 Mass spectrometry analysis of the Sp-CBMTD-GFP clone. Matched peptides were shown in red and bold.

2.7 Discussion

It has been shown that the CBM40 domains from *Vibrio cholerae* and *Streptococcus pneumoniae* sialidases can be successfully isolated from their parent enzyme and exploited to generate multivalent polypeptides through molecular biology techniques. As widely reported, a multivalent approach can dramatically increase protein affinity and enhanced/altered protein selectivity (Kiessling et al., 2000; Liang et al., 1997). CBMs are not involve in catalytic activity but their role is to bring the substrate to the active site in the catalytic domain to help in hydrolysis (Rodriguez-Sanoja et al., 2005).

Moustafa et al. (2004) had reported that the *V. cholerae* CBM40 recognized sialic acid with high affinity of K_D ~ 30µm. This was reported as one the highest affinity of a single sialic-acid binding protein interaction recognizing a sialic acid moiety (Moustafa et al., 2004). Other sialic acid-protein binding interactions have been reported, such as wheat germ agglutinin has a K_d ~ 100 µm with sialyllactose (Kronis and Carver, 1982), influenza virus haemagglutinin has a K_d ~ 1 µm (Sauter, Bednarski et al. 1989) as does sialic acid binding domain of rhesus rotavirus (Dormitzer et al., 2002). Tandem-repeat polypeptides have been engineered to achieve stronger avidity to their receptor when interacting with a multivalent surface. Using this idea, engineered polypeptides containing multiple repeats of CBM40 from *V. cholerae* sialidase with three and four repeats of the domains (Vc3CBM and Vc4CBM) were successfully developed. Flexible linkers containing glycine and serine residues were introduced to allow the connecting domains to freely twist and rotate through space to recruit their binding partners. These linkers act to promote communication between domains and functional modules of a protein (Gokhale and Khosla, 2000). Several studies were done on linker selection and suggested that flexibility and hydrophilicity of the linkers were important to prevent disturbance of the domain functions, thereby promote stability to the whole domains (Arai et al., 2004; Argos, 1990; Robinson and Sauer, 1998b; Wriggers et al., 2005).

Other multivalent CBM constructs containing an oligomerization domain from *P. aeruginosa* pseudaminidase (PaTD) were also engineered and fused either to a CBM40 domain from *V. cholerae* sialidase or from *Streptococcus pneumoniae* NanA sialidase. Through protein engineering, the TD domain was artificially linked to CBM40 domain in order to form a trimer (Xu et al., 2009a). Examples of a trimer-oligomerizing domain include a C-propeptide of procollagens, a C-terminal portion of Fas ligand, a coiled-coil neck domain of the collectin family of proteins and a bacteriophage T4 fibritin foldon domain (Frank et al., 2001; Holler et al., 2003; Hoppe et al., 1994). Many examples illustrate that oligomerization domains play important role in protein function as well as induced multivalency cause by increasing interactions between domains.

A number of studies on oligomerization systems were observed to have a large increase in thermal stability, for example, a chimera molecule consist of collagen-like peptides attached to the N-terminus of the foldon domain (Engel and Kammerer, 2000). Other effects would involve multivalency and increment in intrinsic concentration by an oligomerization effect (Engel and Kammerer, 2000). This physiological feature is seen in proteins such as collagen, whereby oligomerization of C-type lectin domains are generated by triple helices that help trimerize these domains (Kishore et al., 1997; Weis et al., 1998). Besides using PaTD as an oligomerization domain, α-helical coiled coil is the most widespread subunit oligomerization motif found in proteins. The members of this family are important components found in rigid structures such as hair scales and feathers (keratin), cytoskeleton and also in vesicle/ viral membrane fusion (Engel and Kammerer, 2000; Skehel and Wiley, 1998).

The mutant form of Vc-CBMTD underwent site-specific mutations to introduce a disulphide bridge at S164C and T83C. The idea for this was to create a trimer with its binding domains facing the cell surface and increase its affinity towards sialic acid. This introduction of a disulphide bond, however, depends on its location in the protein structure in order to prevent

or restrict the movement of folded protein domains. It is widely known that disulphide bonds are covalent bonds which are stronger that non-covalent bonds (hydrogen bonds, hydrophobic interaction and Van der Waals interaction), and can thus help to stabilize protein structure through cross-linkage action (Betz, 1993; Matsumura and Matthews, 1991; Wedemeyer et al., 2000).

Purification and expression of these modified multivalent proteins were not as straight-forward as the production of other recombinant proteins that were expressed in *E. coli*. All multivalent proteins were required to go through an additional stage of heat-shock at 42 °C for 30 minutes before induction with IPTG, in order to express soluble protein. By introducing this method, protein solubility was significantly increased whilst induction levels were not affected. During the 42 °C of heat shock stage, *E. coli* heat-shock proteins are generally expressed and triggered to act as molecular chaperones. These molecular chaperones assist protein folding and disassemble protein aggregates (Oganesyan et al., 2007). Moreover, when this stage is performed under carefully controlled conditions, heat shock treatment was found to actually rescue some proteins from inclusion body formation (Chen et al., 2002). For protein expression, high temperature is always unfavourable and can cause protein aggregation, therefore, a temperature of 18 - 20 °C was chosen as the most suitable temperature to express these proteins.

In protein purification, method choice is very important especially for good protein separation and purity. Due to this, two stages of nickel affinity and a gel filtration chromatography with a HiPrep 16/60 Sephacryl S-100HR column were used. This method separates protein molecules according to size as they pass through the column. For our work, the column used in the later stages of purification, exhibited excellent protein separation for all six multivalent proteins with high reproducibility results due to its highly stable matrix. In addition, the hydrophilic nature of the matrix helps to minimize nonspecific binding and maximizes protein recovery. For some of our proteins such as Vc3CBM and Vc4CBM, an

additional step using anion-exchange chromatography was carried out with HiTrap Q (FF) column pre-packed with Q Sepharose matrix to further remove lower molecular weight contaminants, which were present in the protein extracts. This step was done in order to further purify the protein of interest before checking protein identification by mass spectrometry.

Chapter 3

Biophysical characterization and crystallization of multivalent CBMs

3.1 Overview

Protein characterization has never been more important especially for a newly emerged and engineered protein. Characterization profile of a protein is a broad term, which is related to both protein separation and purification also its physical, chemical and biological properties. It is far more challenging to characterize a protein than a small molecule as proteins simply have more properties to investigate, as they are large molecules with complex and varied structures. Most proteins are sensitive to conditions like temperature and pH change as they tend to lose their higher order structure when these conditions are suboptimal.

When it comes to developing a strategy and selecting techniques for protein characterizations, several things need to be considered and this includes protein stability and purity. Therefore, careful sample preparation is very important in protein characterization studies. In this chapter, we will be studying protein conformation and aggregation, thermodynamics as well as structure determination of multivalent CBM40 constructs from *Vibrio cholerae and Streptococcus pneumoniae* sialidases.

A range of biophysical techniques are available to evaluate direct binding between a ligand and a target protein, these include direct fluorescence-based methods, such as thermal shift

assays. In this assay, different sialosides were used in order to look at its selectivity and specificity of binding to multivalent CBM40 proteins. The importance of buffers to formulate proteins to exhibits optimal stability was also studied using the same approach; this includes screening for different buffers and pH conditions. Moreover, protein thermodynamics were also studied with a combination of both isothermal titration chromatography (ITC) and surface plasmon resonance (SPR). Lastly for protein crystallization trials, Vc-CBMTD (WT) and Sp-CBMTD were chosen for the study. For this, protein with high purity and homogeneity are crucial for crystallization to be successful. Hence, it is essential to perform an extensive quality assessment and evaluation of the protein prior to this structural study. This includes a protein pre-crystallization test and its optimization process.

3.2 Study of protein conformation and aggregates

Many proteins are usually stable in solution, undergoing conformational changes due to various stresses during purification, processing and storage. The most common physical instability is protein aggregation, which can be affected by a variety of factors and chemical transformations. There are a few methods available to study protein aggregation such as size-exclusion chromatography (SEC), analytical ultracentrifugation (AUC), dynamic light scattering (DLS). Here in this section, we discussed a method using native gel electrophoresis to study protein aggregation.

We were using NativePAGE^{TM} Bis-Tris Gel system (Invitrogen) that used Coomassie G-250 as a charge-shift molecule without denaturing action. BlueNative-PAGE (BN-PAGE) was carried out using Novex Bis-Tris gel system according to the manufacturer's specifica. Pre-cast NativePAGE^{TM} Novex 4-16% (v/v) Bis-Tris gels were run with near neutral pH at 150 V at room temperature. Protein samples (10 µl) were mixed with the sample buffer provided (2.5 µl) and 5% Coomassie Blue G-250 (0.3 µl). Samples were run without a heating stage.

Gels were stained using Coomassie R-250 staining protocol for NativePAGE™ Gels as described in the protocol (Appendix A-3).

In this technique, sodium dodecyl sulphate (SDS) and reducing agent (DTT) are absent to allow the protein to remain in a nondenatured and associated state. Protein mobility is dependent on its charge, shape and its hydrodynamic size (higher mobility for more compact conformations, lower for larger structures like oligomers). The patterns of Vc-CBMTD (WT) and Vc-CBMTD (mutant) generated by BN-PAGE in the presence and absence of detergents (DDM and Triton-X100) are shown in Figure 3.1. As mentioned before, protein denaturing is not desired for native gel thus solubilisation of the protein is necessary and very important. The used of mild and non-ionic detergents (5% Digitonin, 10% of n-dodecyl-β-D-maltoside (DDM), Triton X-100) are very important to ensure complete solubility.

Proteins which normally have positive net charges will be converted to a negative charge by G-250 dye, so that they can migrate to the anode. As expected in the presence of Coomassie Blue G-250 in the protein sample, all proteins migrated toward the anode. Bovine serum albumin and Vc1CBM were used as positive controls for the experiment and migrate as single bands toward the anode. It was shown that, both proteins are present as monomers on the gel. For Vc-CBMTD (WT), the proteins were shown to migrate as single bands with approximately the same mobility as shown in Lane 3 and Lane 5. Some severe laddering of the protein was also witnessed, which might be an artefact of the protein caused by protein degradation. Furthermore, neither the addition of Coomassie Blue G-250 nor detergents could resolve the laddering issue.

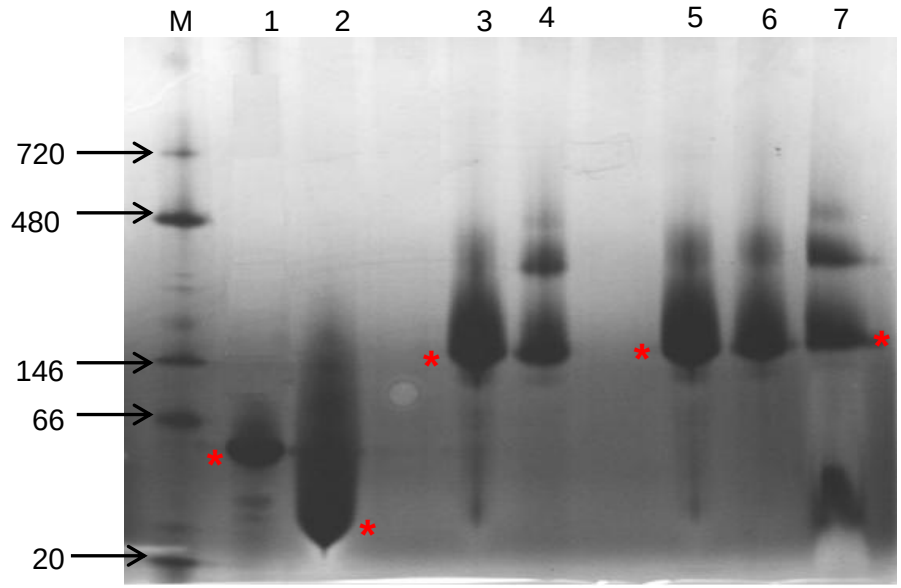


Figure 3.1 Migration of proteins by BlueNative gel electrophoresis. M, NativeMark™ Unstained protein standard, Lane 1, Bovine serum albumin (66 kDa); Lane 2, Vc1CBM (21 kDa); Lane 3, Vc- CBMTD (WT)- without detergent; Lane 4, Vc-CBMTD (mutant)- without detergent; Lane 5, Vc-CBMTD (WT) + 1% (w/v) DDM; Lane 6, Vc-CBMTD (Mutant) + 1% (w/v) DDM; Lane 7, Vc-CBMTD (Mutant) + 0.5% (v/v) Triton X-100. Expected size of Vc-CBMTD (WT) and Vc-CBMTD (Mutant) were 99 kDa. This experiment used pre-cast 4-16% Native-PAGE Bis-tris gel. Proteins size was shown bigger than the expected size due to protein migrate in their native conformation. Asterisk represents protein of interest.

In contrast, the mutant version of Vc-CBMTD with or without detergents, was shown to migrate as multiple bands. Different conformational folded states of protein, plus the presence of disulphide bonds affects the hydrodynamic size and mobility on the gel and lead to the above condition. A different detergent was later used, which was Triton X-100, but the same pattern of protein migration was also observed (Lane 7). One important issue to be aware of when using native gel for protein size estimation is an expected error of ~ 15%, as the size of most proteins will be larger than their expected size due to the protein migration in their native conformation. Both Vc-CBMTD (WT) and Vc-CBMTD (Mutant) showed a larger size than the expected 99 kDa as shown in native gel (Figure 3.1).

As electrophoretic mobility of proteins in the native gel depends on both the electric charge and the hydrodynamic size, it can therefore provide information on protein aggregation. From Figure 3.1, both proteins (Vc-CBMTD (WT) and Vc-CBMTD (Mutant)) presented high molecular weight bands, which correspond to protein aggregates. Usually, aggregation is driven by the presence of a small amount of contaminant such as damaged form of the protein itself, host cell proteins or non-protein materials (silica particles) (Patel et al., 2012). For therapeutic proteins, protein aggregates are known as a potential risk in the generation of immune responses in humans. Minimization of immunogenicity in the product must be accomplished by ensuring stability of the native protein conformation and minimizing formation of high MW species (Rosenberg, 2006).

3.3 Interaction profile of multivalent CBMs

3.3.1 Thermal shift assay: screening for the best buffer

Thermal stability of the protein can be measured through the midpoint or T_m of the protein curve. This is because, proteins usually unfold at certain temperatures and during this

process, temperature at which 50% of protein molecules are unfolded is defined as the melting temperature (T_m). The higher the T_m, the greater the thermal resistance of a protein (Niesen et al., 2007).

During a typical assay, a two-state transition is shown, which are folding and unfolding states of proteins. The fluorescent intensity increases during protein unfolding which exposes the buried hydrophobic core of the protein followed by the binding of the Sypro Orange dye (Figure 3.2) (Lo et al., 2004). However, after reaching the plateau, the fluorescence intensity starts to decrease gradually probably due to aggregation of denatured protein and dye complexes.

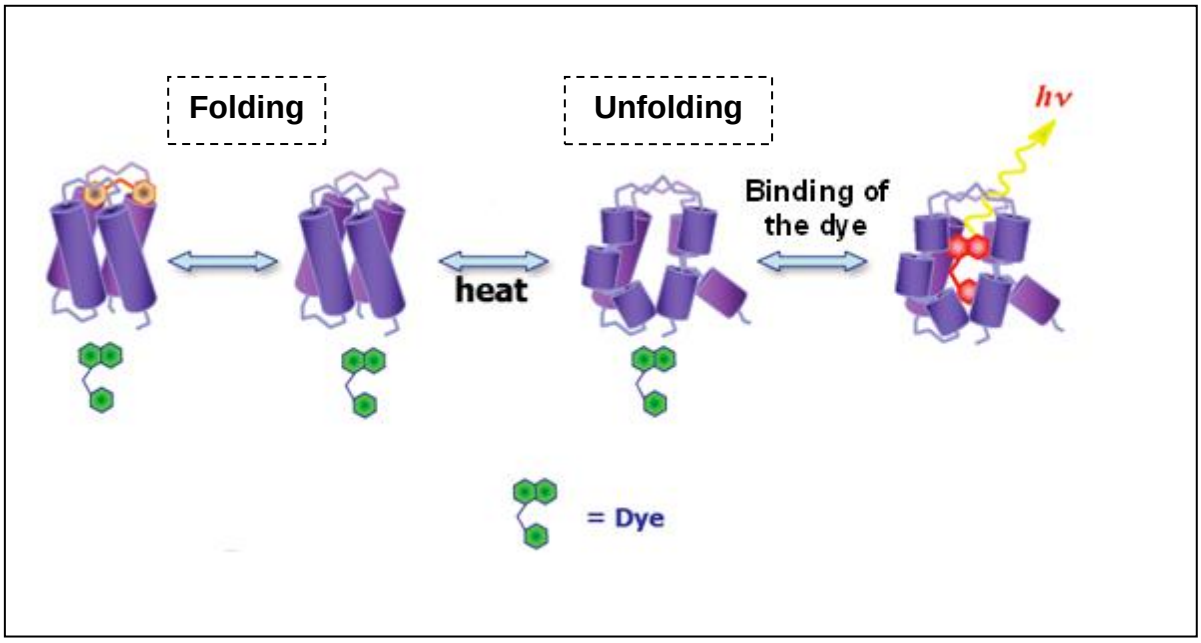


Figure 3.2 Protein thermal stabilization measured by fluorescence readout in thermal shift assay. Figure was adapted from (Pantoliano et al., 2001).

Thermal stability of multivalent CBM proteins was monitored using a thermal shift assay. In this study, we demonstrate the use of this screening tool to find the best buffer conditions for our proteins. The assay was conducted in Applied Biosystems Real-Time PCR Machine using MxPro software. Solutions of 42.5 µl of test buffers with 50 mM NaCl (or 100 mM sodium citrate pH 5.5; 100 mM sodium citrate pH 6.2; 100 mM sodium phosphate pH 6.5; 100 mM sodium phosphate pH 7.5; 100 mM potassium phosphate pH 6.5; 100 mM potassium phosphate pH 7.5; 100 mM HEPES pH 6.5, 100 mM HEPES pH 7.5, 20 mM Tris, pH 6.5 and 20 mM Tris, pH 7.5), 2.5 µl of 50 X SYPRO Orange dye and 5 µl of 2.5 mg/ml proteins were added to the each PCR tubes. The tubes were incubated using the following parameters; 1 cycle at 25 °C for 30 seconds followed with 140 cycles at 25 °C to 100 °C for 1 minute and a final cycle at 25 °C for 1 minute. The fluorescence intensity was measured with Ex/Em: 492/610 nm.

From all buffers used, 100 mM sodium citrate, pH 5.5 + 50 mM NaCl give the highest T_m value for both Vc3CBM and Vc4CBM which were 51°C and 50°C respectively (Figure 3.3 and Figure 3.4). While Vc-CBMTD (Mutant) gave T_m value of 52 °C in 20 mM Tris, pH 6.5 + 50 mM NaCl and Vc-CBMTD (WT) was more stable in 20 mM Tris, pH 7.5 + 50 mM NaCl with the highest T_m value of 69°C, respectively (Figure 3.5 and Figure 3.6). For Sp-CBMTD, the protein showed quite a low melting temperature in all buffers used (33 to 34 °C). This finding was odd as the protein belongs to *Streptococcus pneumoniae*, a common pathogen find in human body. Later, as to confirm, the protein was tested against siallylactose substrates to see how stable the protein when bound to the substrate. Among all buffers, the highest T_m was found at 36.8 °C for Sp-CBMTD in 100 mM sodium phosphate, pH 6.5 + 50 mM NaCl (Figure 3.7). Comparisons of the resulting T_m values can be used to select the best buffer relative to protein stability for the crystallization work later on. Although some of the proteins were found to be more stable in phosphate buffer, this buffer is generally avoided as this would typically form salt crystals (magnesium or calcium phosphate) that may interfere in protein crystal formation.

The presence of two peaks in Figure 3.7 showed two unfolding states that likely represented the two domains of the constructs. This feature was only shown with constructs containing the trimerization domain (Vc-CBMTD (WT), Vc-CBMTD (Mutant) and Sp-CBMTD). The first peak was proposed to be the CBM40 domain while the second was TD domain (Figure 3.6 and Figure 3.7). This has been proved when adding 3'-siallylactose (3'SL) in the assay, only the first sigmoid curve was shifted in temperature showing the binding of 3'SL to the domain, while no temperature shift was observed with the second peak (Figure 3.16).

Moreover, the pH of buffers will affect protein stability, which may alter the electrostatic interactions between charged amino acids (Schein, 1990). Using the screening approach, the consequent changing of T_m of the protein in solution can be used as a measure of the relative stability across the tested pH range. In this assay, range of the pHs tested were 5.5 to 7.5 depending on the buffer used. Addition of salt in the buffer (50 mM NaCl) can also affect protein conformation and stability as the cations and anions of the salt could be potential buffer components.

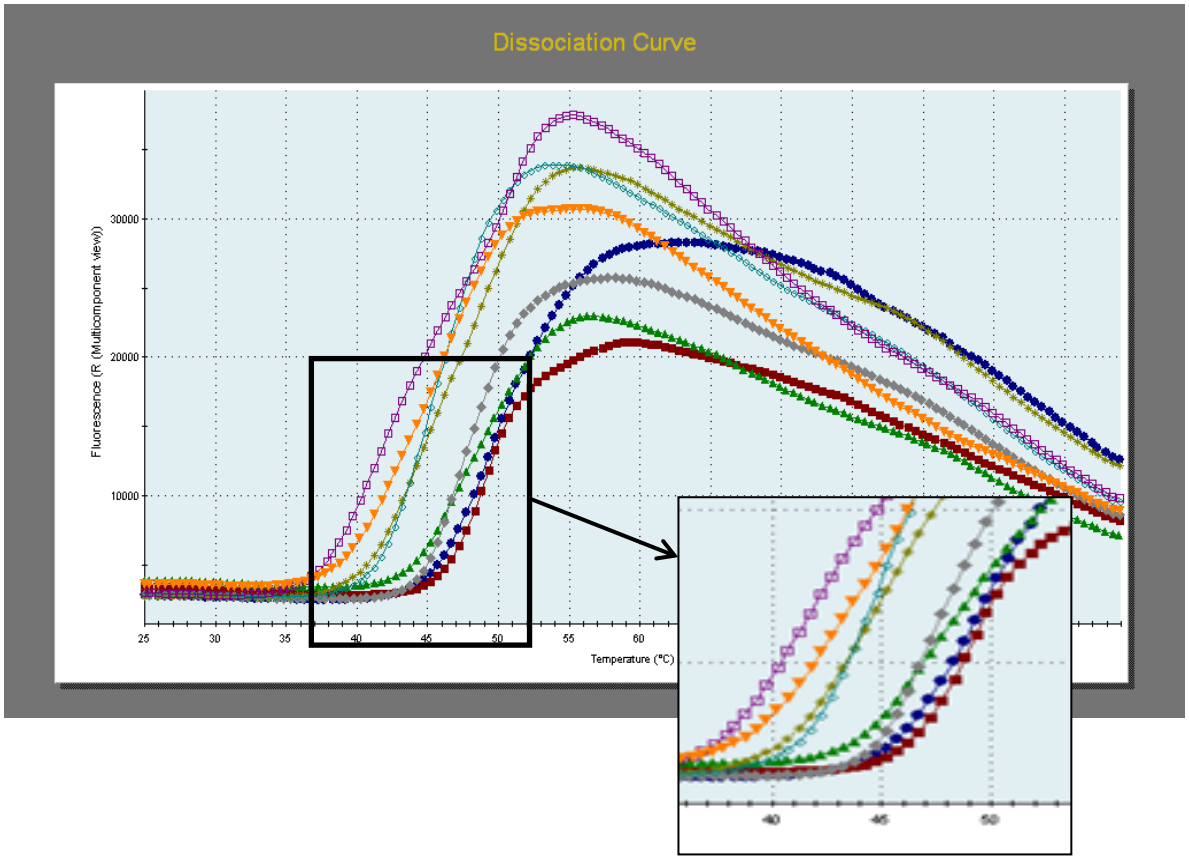


Figure 3.3 Dissociation curve of Vc3CBM in 8 different buffers. (■) 100 mM sodium citrate, pH 6.2 + 50 mM NaCl; (●)100 mM sodium citrate, pH 5.5 + 50 mM NaCl; (◇) 100 mM HEPES, pH 6.5 + 50 mM NaCl; (▲) 100 mM HEPES, pH 7.5 + 50 mM NaCl; (*)100 mM potassium phosphate, pH 6.5 + 50 mM NaCl; (○) 100 mM potassium phosphate, pH 7.5 + 50 mM NaCl; (▲) 100 mM sodium phosphate, pH 6.5 + 50 mM NaCl; (◆)100 mM sodium phosphate, pH 7.5 + 50 mM NaCl.

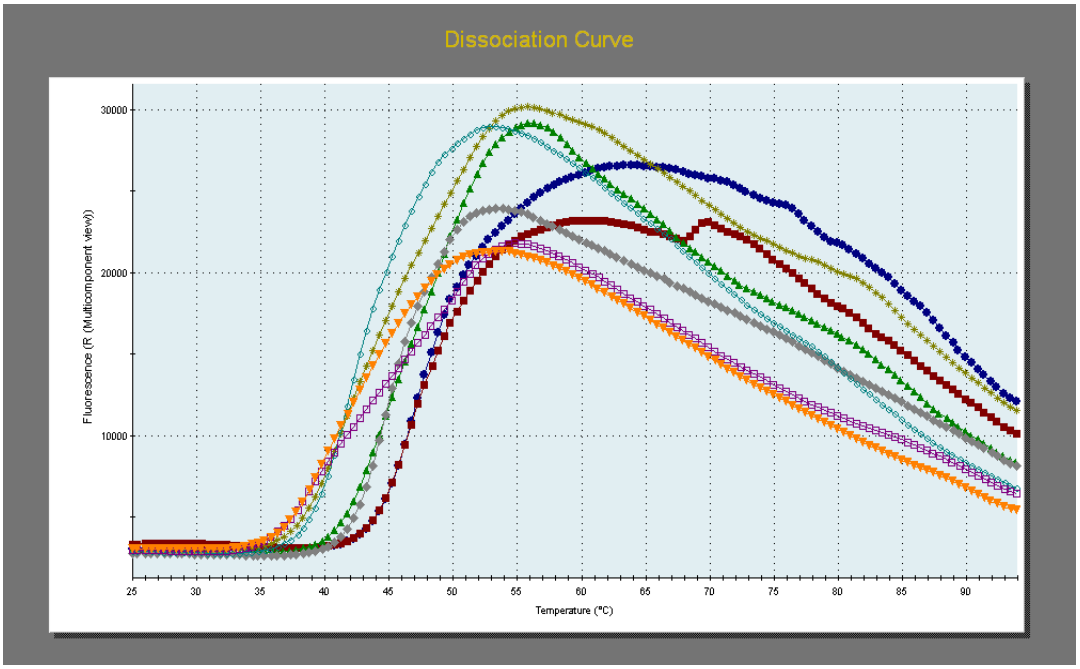


Figure 3.4 Dissociation curve of Vc4CBM in 8 different buffers. (■) 100 mM sodium citrate, pH 6.2 + 50 mM NaCl; (●) 100 mM sodium citrate, pH 5.5 + 50 mM NaCl; (◇) 100 mM HEPES, pH 6.5 + 50 mM NaCl; (▲) 100 mM HEPES, pH 7.5 + 50 mM NaCl; (*)100 mM potassium phosphate, pH 6.5 + 50 mM NaCl; (○) 100 mM potassium phosphate, pH 7.5 + 50 mM NaCl; (▲)100 mM sodium phosphate, pH 6.5 + 50 mM NaCl; (◆)100 mM sodium phosphate, pH 7.5 + 50 mM NaCl.

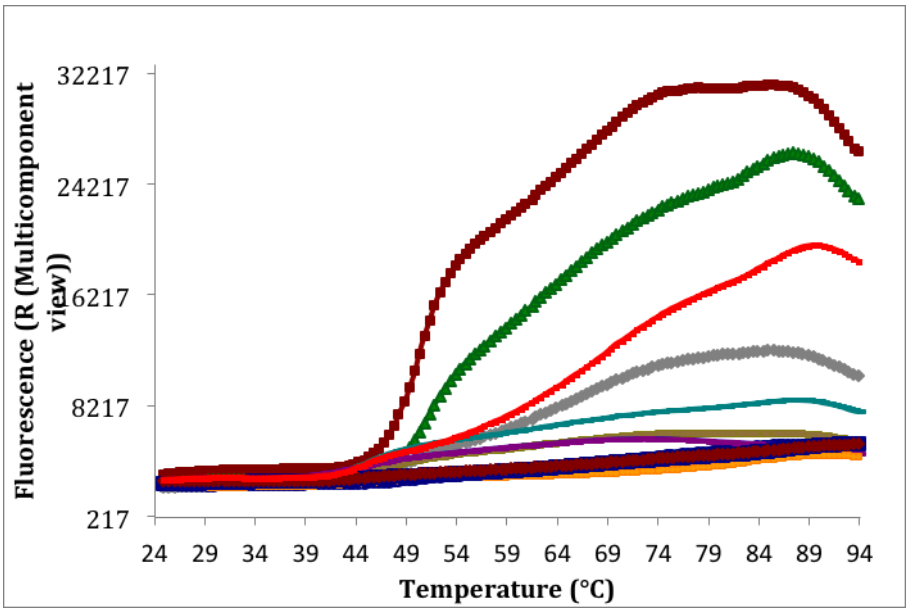


Figure 3.5 Dissociation curve of Vc-CBMTD (WT) in 10 different buffers. **(●)** 100 mM sodium citrate, pH 5.5 + 50 mM NaCl; (■) 100 mM sodium citrate, pH 6.2 + 50 mM NaCl; (▲) 100 mM sodium phosphate, pH 6.5 + 50 mM NaCl; (◆) 100 mM sodium phosphate, pH 7.5 + 50 mM NaCl; (■) 100 mM potassium phosphate, pH 6.5 + 50 mM NaCl; (▬) 100 mM potassium phosphate, pH 7.5 + 50 mM NaCl; (▬) 100 mM HEPES, pH 6.5 + 50 mM NaCl; (■) 100 mM HEPES, pH 7.5 + 50 mM NaCl; (◆) 20 mM Tris, pH 6.5 + 50 mM NaCl; (■) 20 mM Tris, pH 7.5 + 50 mM NaCl.

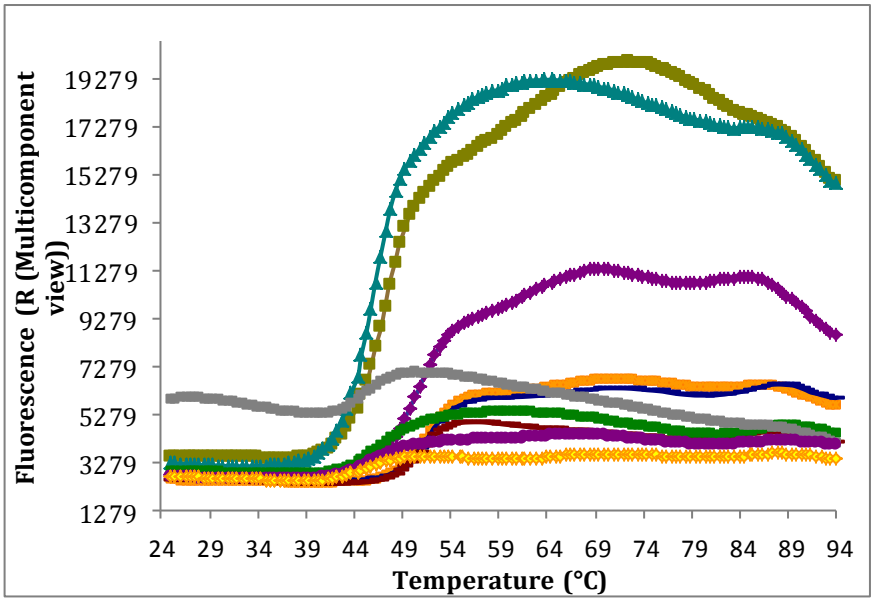


Figure 3.6 Dissociation curve of Vc-CBMTD (Mutant) in 10 different buffers. (■) 100 mM potassium phosphate, pH 6.5 + 50 mM NaCl; (▲) 100 mM potassium phosphate, pH 7.5 + 50 mM NaCl ; (◆) 100 mM sodium citrate, pH 5.5 + 50 mM NaCl ; (■) 100 mM sodium citrate, pH 6.2 + 50 mM NaCl ; (▬) 100 mM sodium phosphate, pH 6.5 + 50 mM NaCl ; (▬) 100 mM sodium phosphate, pH 7.5 + 50 mM NaCl ; (■) 100 mM HEPES, pH 6.5 + 50 mM NaCl : (■) 100 mM HEPES, pH 7.5 + 50 mM NaCl ; (●) 100 mM Tris, pH 6.5 + 50 mM NaCl; (▲) 100 mM Tris, pH 7.5 + 50 mM NaCl.

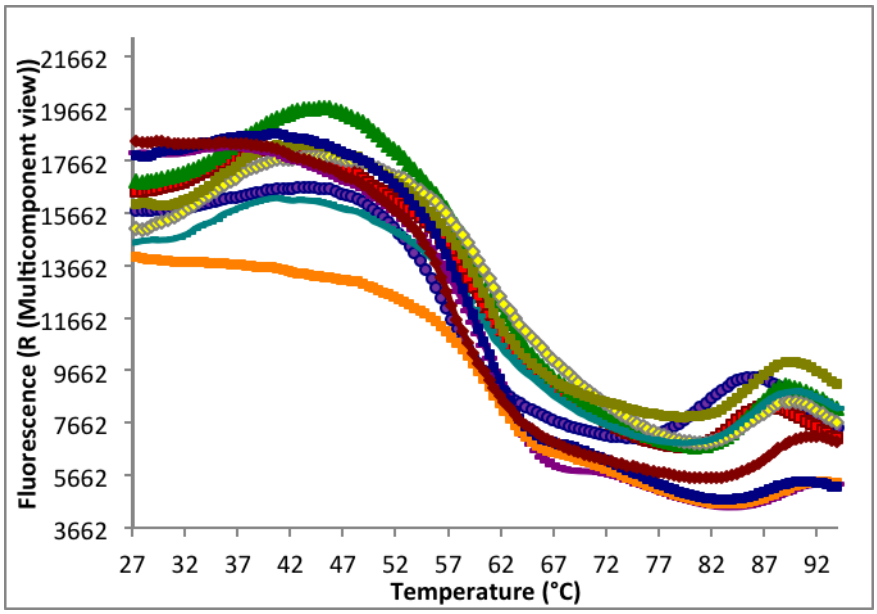


Figure 3.7 Dissociation curve of Sp-CBMTD in 10 different buffers.

(●) 100 mM sodium citrate, pH 5.5 + 50 mM NaCl ;(■) 100 mM sodium citrate, pH 6.2 + 50 mM NaCl, (▲) 100 mM sodium phosphate, pH 6.5 + 50 mM NaCl; (◆) 100 mM sodium phophate, pH 7.5 + 50 mM NaCl; (■) 100 mM pottasium phosphate, pH 6.5 + 50 mM NaCl; (−) 100 mM pottasium phosphate, pH 7.5 + 50 mM NaCl; (−) 100 mM HEPES, pH 6.5 + 50 mM NaCl; (■) 100 mM HEPES, pH 7.5 + 50 mM NaCl; (■) 20 mM Tris, pH 6.5 + 50 mM NaCl; (■) 20 mM Tris, pH 7.5 + 50 mM NaCl.

3.3.2 Thermal shift assay: screening for the best ligands/ sialosides

Thermal-shift assays were also performed in order to identify the best ligand for the multivalent CBM40 constructs. When a ligand binds to a protein native state, it will stabilize the complex, and this can be observed experimentally as an increase in melting temperature (T_m) of the protein in the presence of ligand (Figure 3.8) (Waldron and Murphy, 2003).

The assay was performed using a conventional real-time PCR machine where ligands are added to the solution (protein + dye) and the fluorescence intensity was measured as the temperature was gradually raised. The protocol was as described earlier in Section 3.3.1. As reported by Connaris et al. (2009), in order to characterize sialic acid for substrate specificity, Vc3CBM was fused with GFP tag and tested against a glycan array and screen for 377 glycans. As expected, Vc3CBM-GFP was found to bind glycans with broad linkage specificity and high affinity to α(2,3)- and α(2,6)-linked sialosides, but to a lesser extent to α(2,8)-linked sialosides (Connaris et al., 2009).

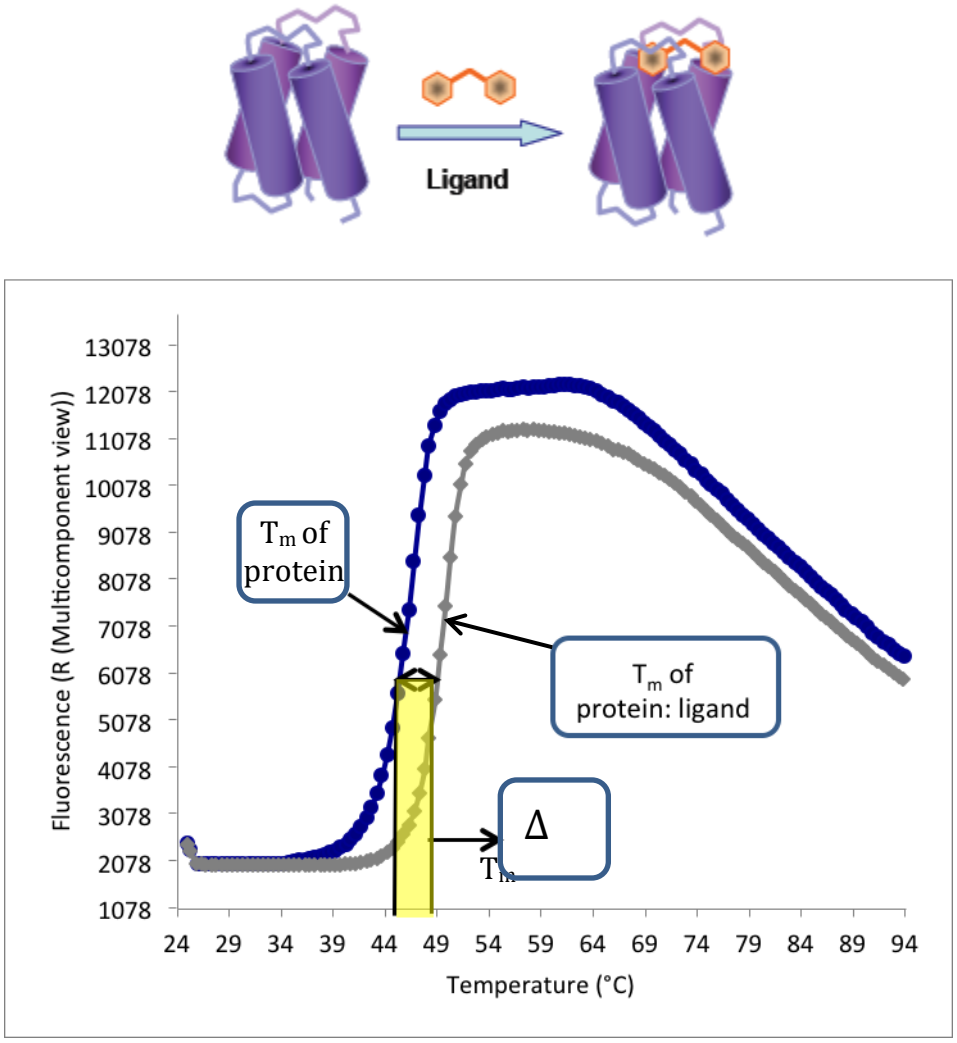


Figure 3.8 Comparison of T_m when reaction presence with protein alone and protein with ligand. A complex of ligand and protein can stabilize the protein structure and increase its melting temperature. This figure was adapted from (Pantoliano et al., 2001).

For all multivalent constructs, five different concentrations of 3'-sialyllactose (3'-SL) and 6'-sialyllactose (6'-SL) (Figure 3.9) were used, which were 0.2, 0.4, 0.6, 0.8 and 1.0 M. In the absence of ligand, the T_m was smaller compared to when the ligand was present (Figure 3.10). In the presence of ligand, five different response curves were measured corresponding to stoichiometric amount of ligand and protein. The increase shift in T_m was identified with an increment of ligand concentrations. From the graph, the best concentrations of ligands were between 0.8 to 1.0 mM for 3'-SL and 6'-SL which was shown to have the highest T_m value from all protein constructs (Figure 3.10- Figure 3.19). While, in terms of buffer used, 100 mM sodium citrate pH 6.2 + 50 mM NaCl and 20 mM Tris, pH 7.5 + 50 mM NaCl were suitable to use without having too much different in T_m values.

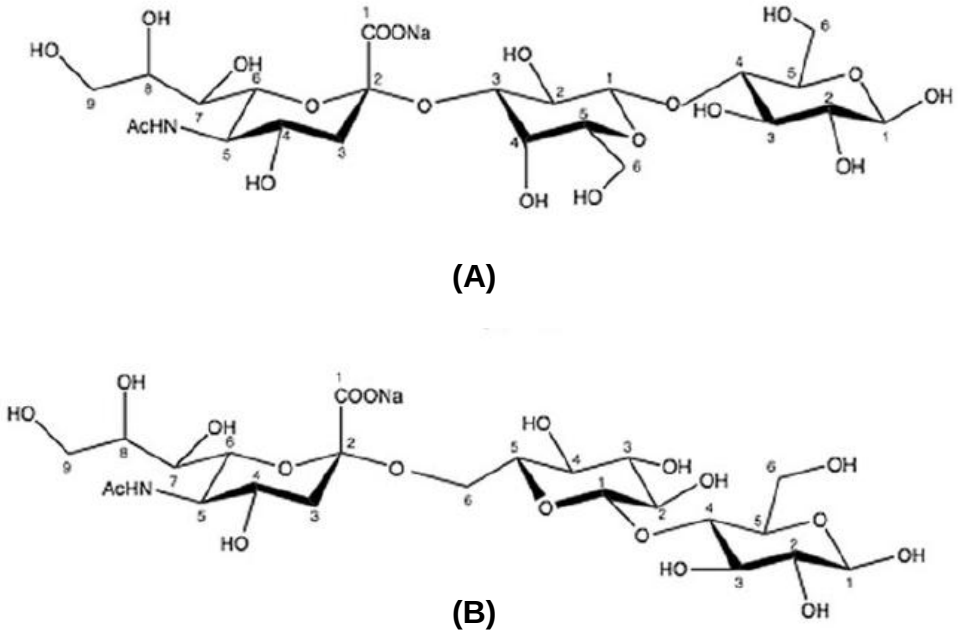


Figure 3.9 Chemical structures and abbreviations of ligands used in this study. (A): 3'-Sialyllactose (3'-SL); **(B)**: 6'-Sialyllactose (6'-SL).

It can be clearly seen from Figure 3.11, that at suboptimal ligand concentrations, little detectable change in T_m was observed compared to the control where ligand was absent. Furthermore, when ligand concentrations were increased, higher melting temperature profile were observed. This pattern was shown on all figures when using both buffers (100 mM sodium citrate pH 6.2 + 50 mM NaCl and 20 mM Tris, pH 7.5 + 50 mM NaCl). As reported by Matulis et al. (2005), T_m value of a protein was expected to increase due to tighter binding between ligands and protein when other variables remain the same (Matulis et al., 2005).

Most of the multivalent constructs showed relevant profile of melting temperatures in different range of buffers and ligands tested. But an interesting finding was shown for Sp-CBMTD which showed quite a low melting temperature (< 37°C) in the tested buffers. This was odd as the protein belongs to *Streptococcus pneumoniae* (a common pathogen in the human body), which should not start to dissociate below 37°C. Later, it was interesting to find out after addition of ligands (3'-SL and 6'-SL), the melting temperature was shifted and stabilized around 37 °C (Figure 3.18 and Figure 3.19). The finding was also in conjunction with data from SPR analysis which showed binding activity occurred starting at lower temperature of 15, 25 and 37 °C. In our case, this assay results will be promising in identifying potential or the best ligand candidates for multivalent constructs. The result derived from this test, will give preliminary data for protein crystallographic trials.

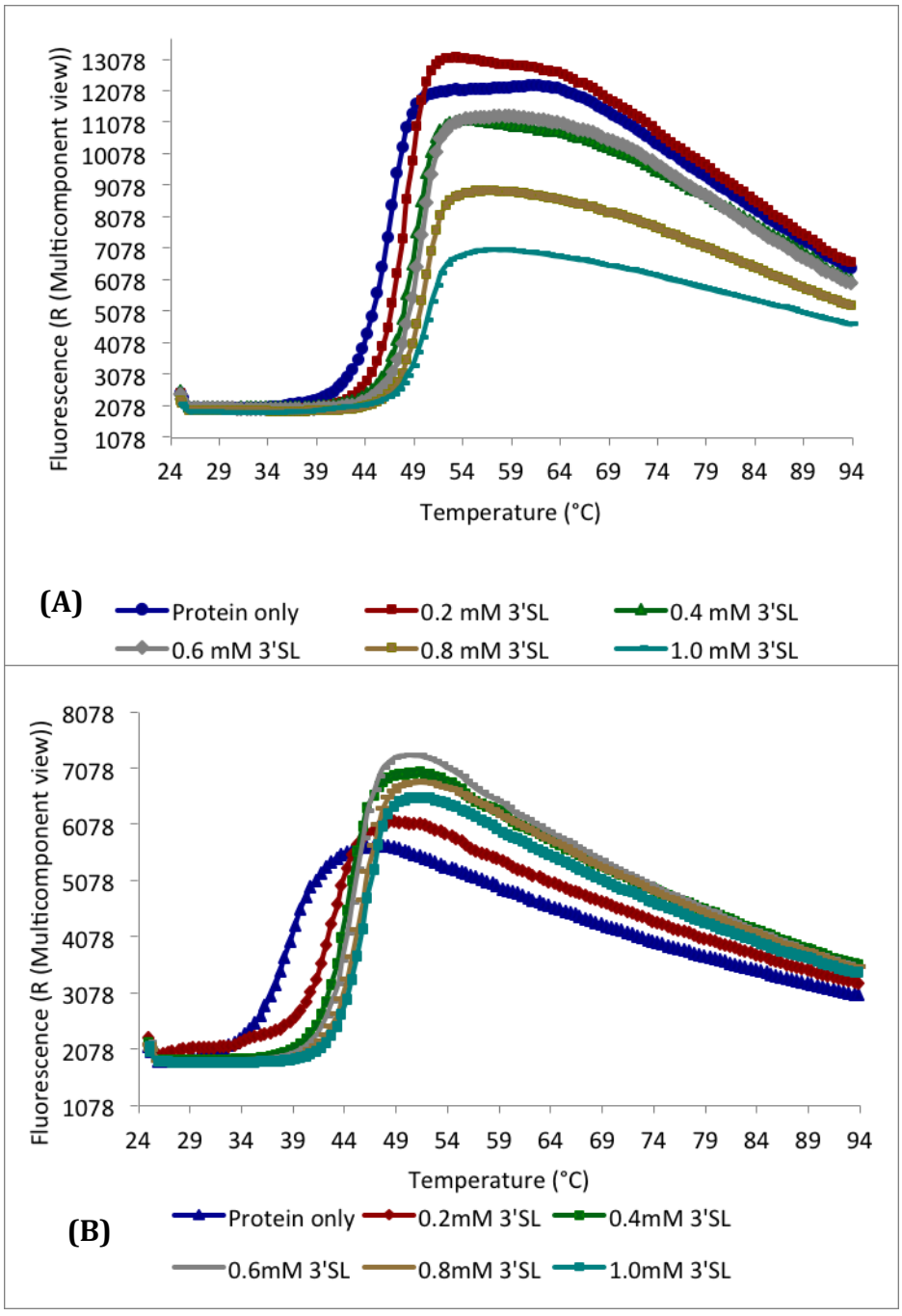


Figure 3.10 Fluorescent readout of Vc3CBM interaction with 3'-Sialyllactose (3'SL) in two different buffers. (A) 100 mM sodium citrate, pH 6.2 + 50 mM NaCl and **(B)** 20 mM Tris, pH 7.5 + 50 mM NaCl with the presence of 3'-SL.

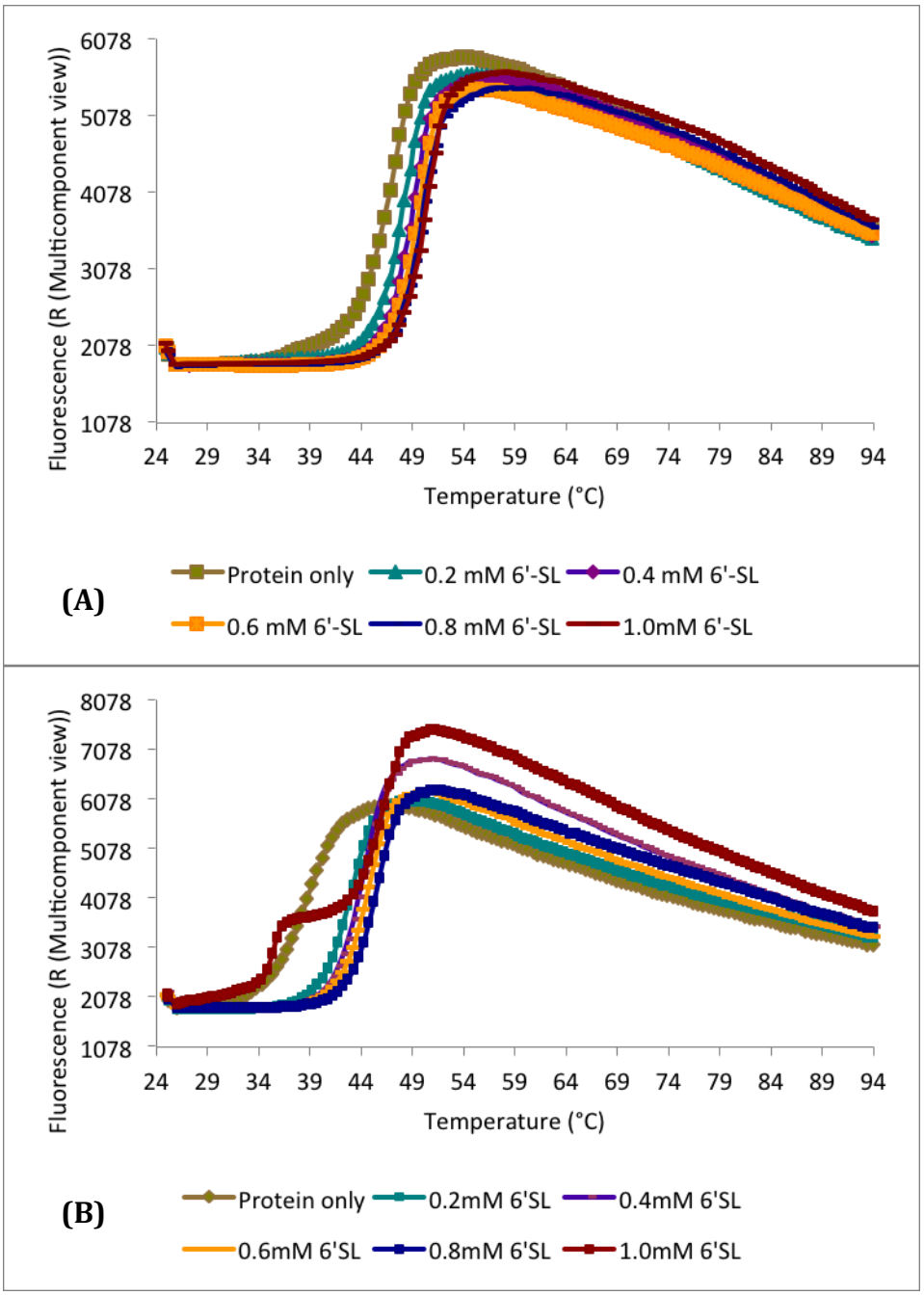


Figure 3.11 Fluorescent readout of Vc3CBM interaction with 6'-Sialyllactose (6'SL) in two different buffers. (A) 100 mM sodium citrate, pH 6.2 + 50 mM NaCl and **(B)** 20 mM Tris-, pH 7.5 + 50 mM NaCl with the presence of 6'SL.

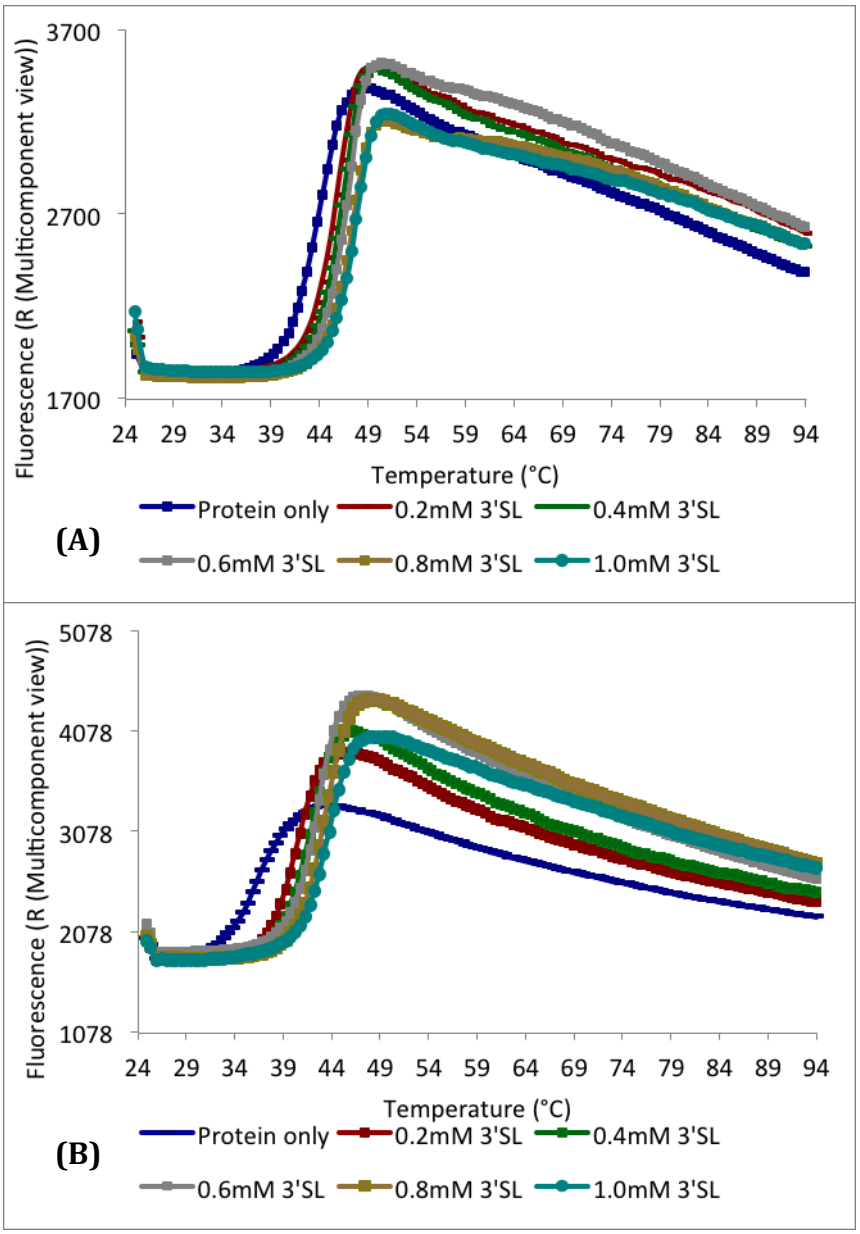


Figure 3.12 Fluorescent readout of Vc4CBM interaction with 3'-Sialyllactose (3'SL) in two different buffers. (A) 100 mM sodium citrate, pH 6.2 + 50 mM NaCl and **(B)** 20 mM Tris, pH 7.5 + 50 mM NaCl with the presence of 3'SL.

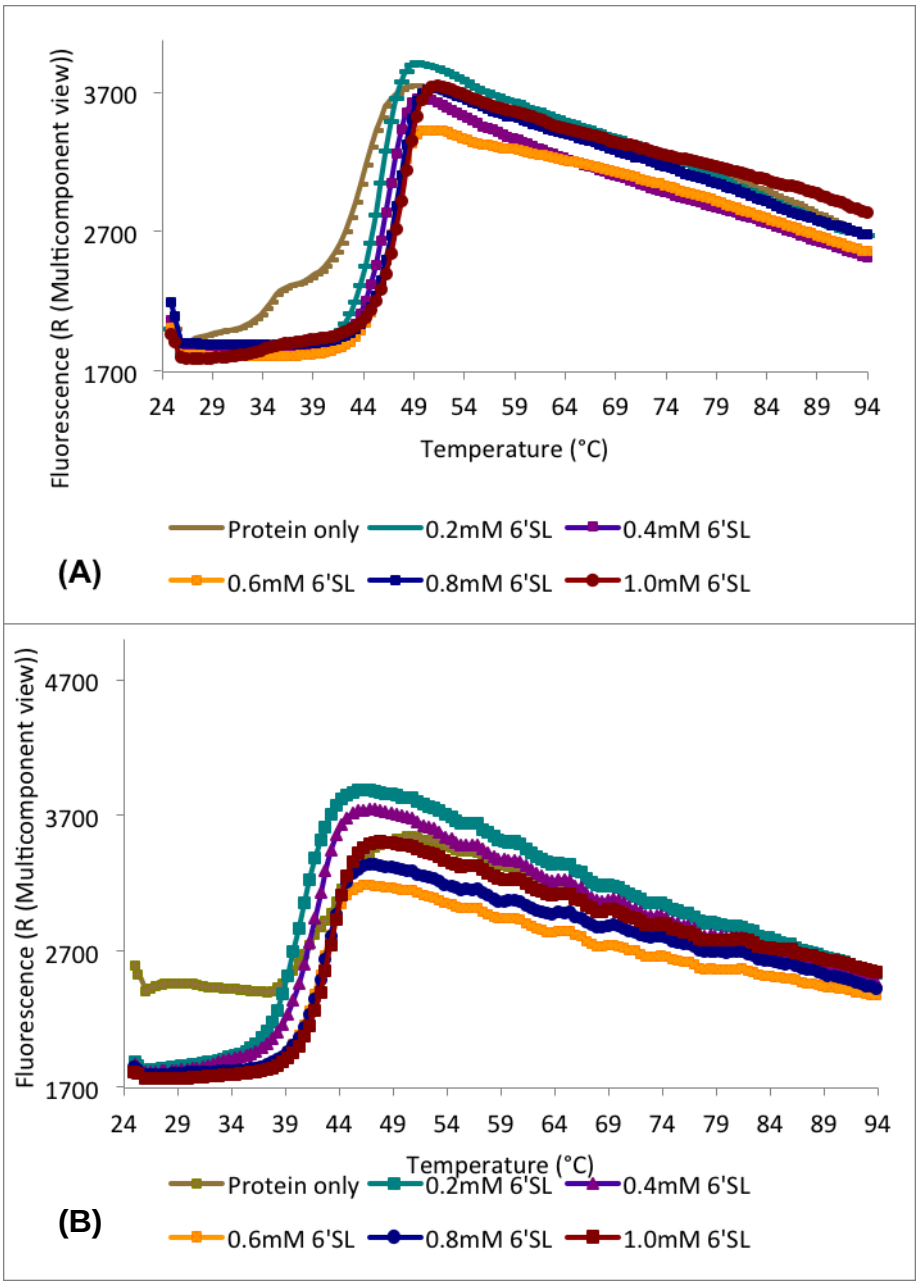


Figure 3.13 Fluorescent readout of Vc4CBM interaction with 6'-Sialyllactose (6'SL) in two different buffers. (A) 100 mM sodium citrate, pH 6.2 + 50 mM NaCl and **(B)** 20 mM Tris, pH 7.5 + 50 mM NaCl with the presence of 6'SL.

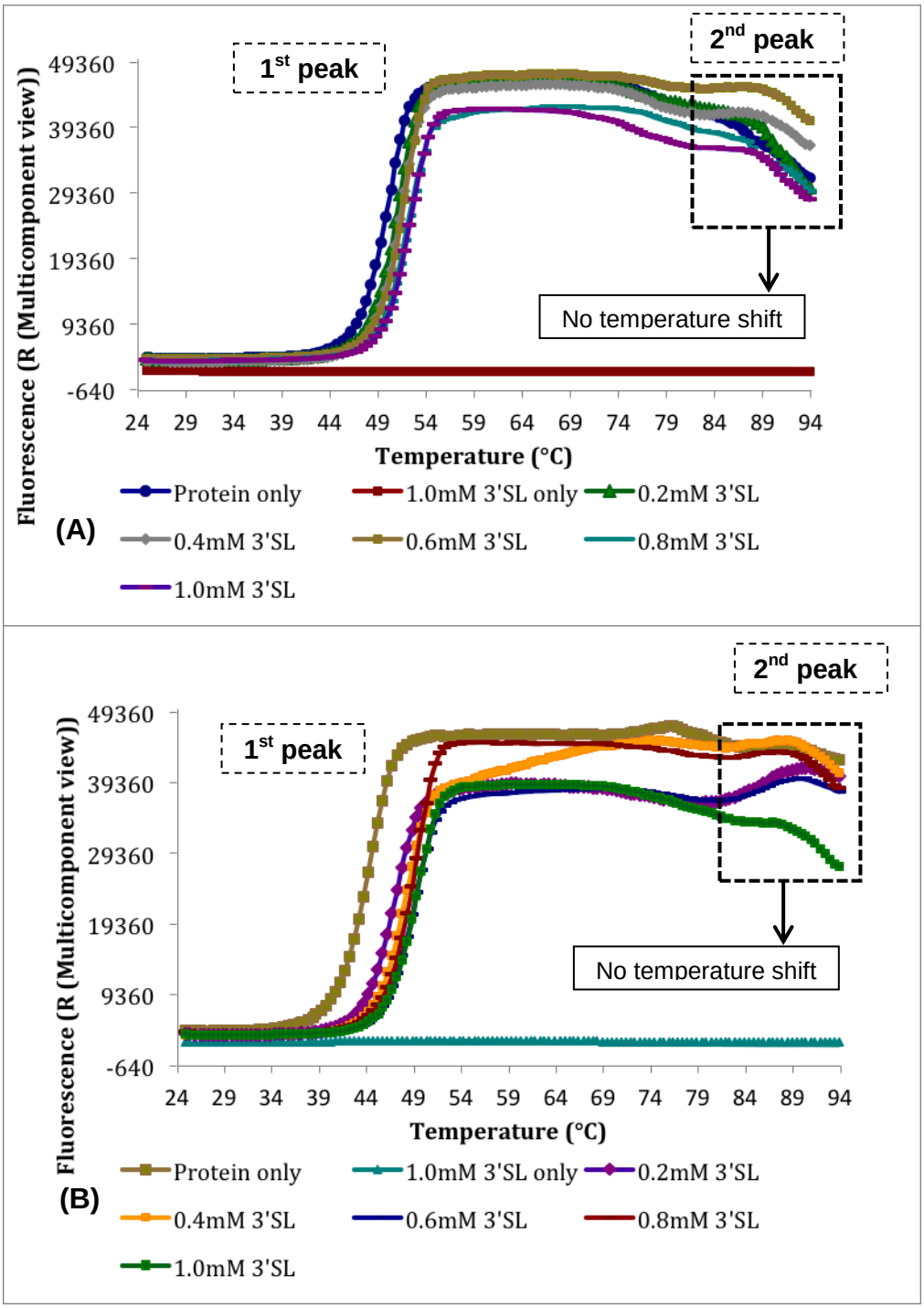


Figure 3.14 Fluorescent readout of Vc-CBMTD (WT) interaction with 3'-Sialyllactose (3'SL) in two different buffers. (A) 100 mM sodium citrate, pH 6.2 + 50 mM NaCl and **(B)** 20 mM Tris, pH 7.5 + 50 mM NaCl with the presence of 3'SL.

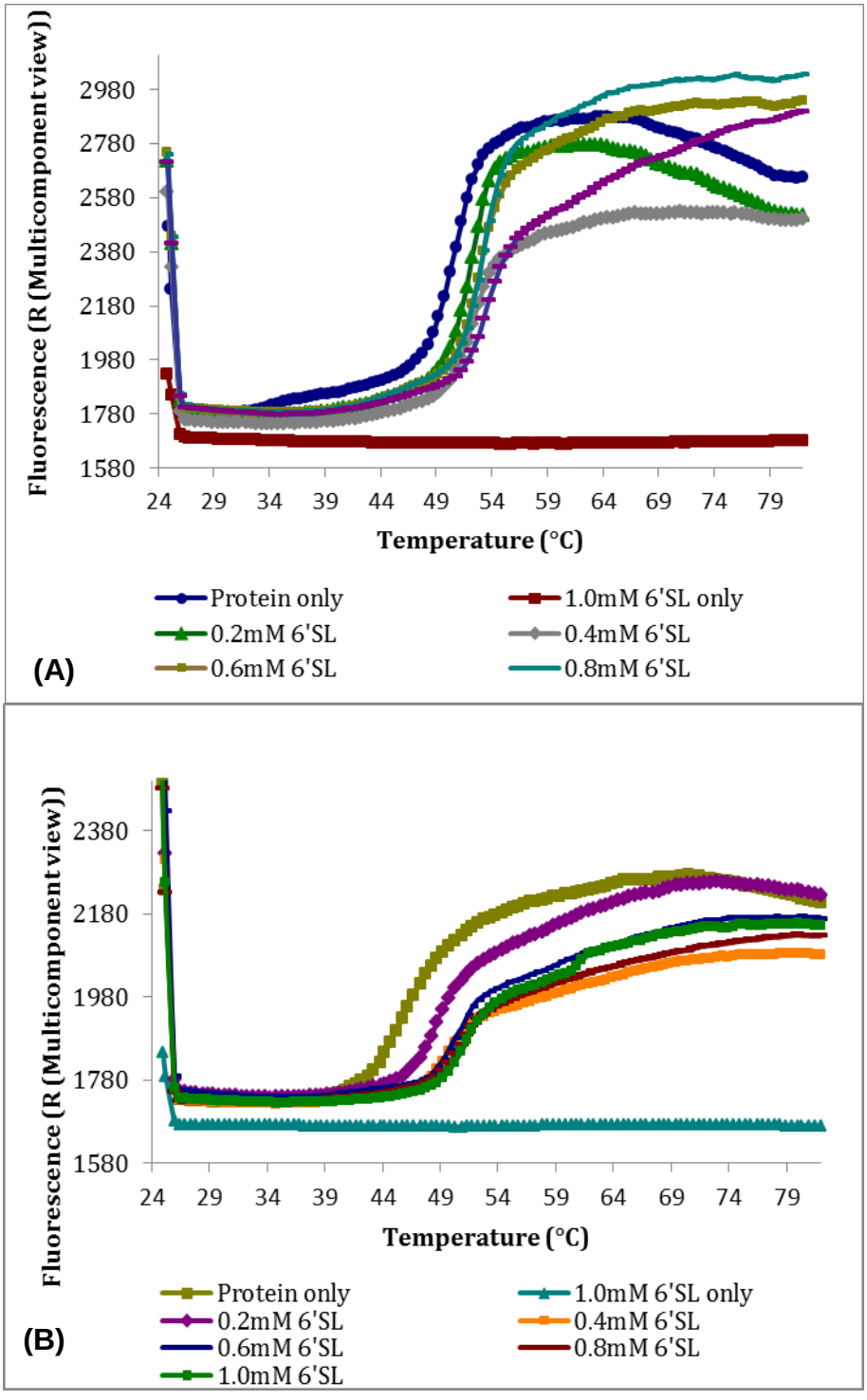


Figure 3.15 Fluorescent readout of Vc-CBMTD (WT) interaction with 6'-Sialyllactose (6'SL) in two different buffers. (A) 100 mM sodium citrate, pH 6.2 + 50 mM NaCl and **(B)** 20 mM Tris, pH 7.5 + 50 mM NaCl with the presence of 6'SL.

.

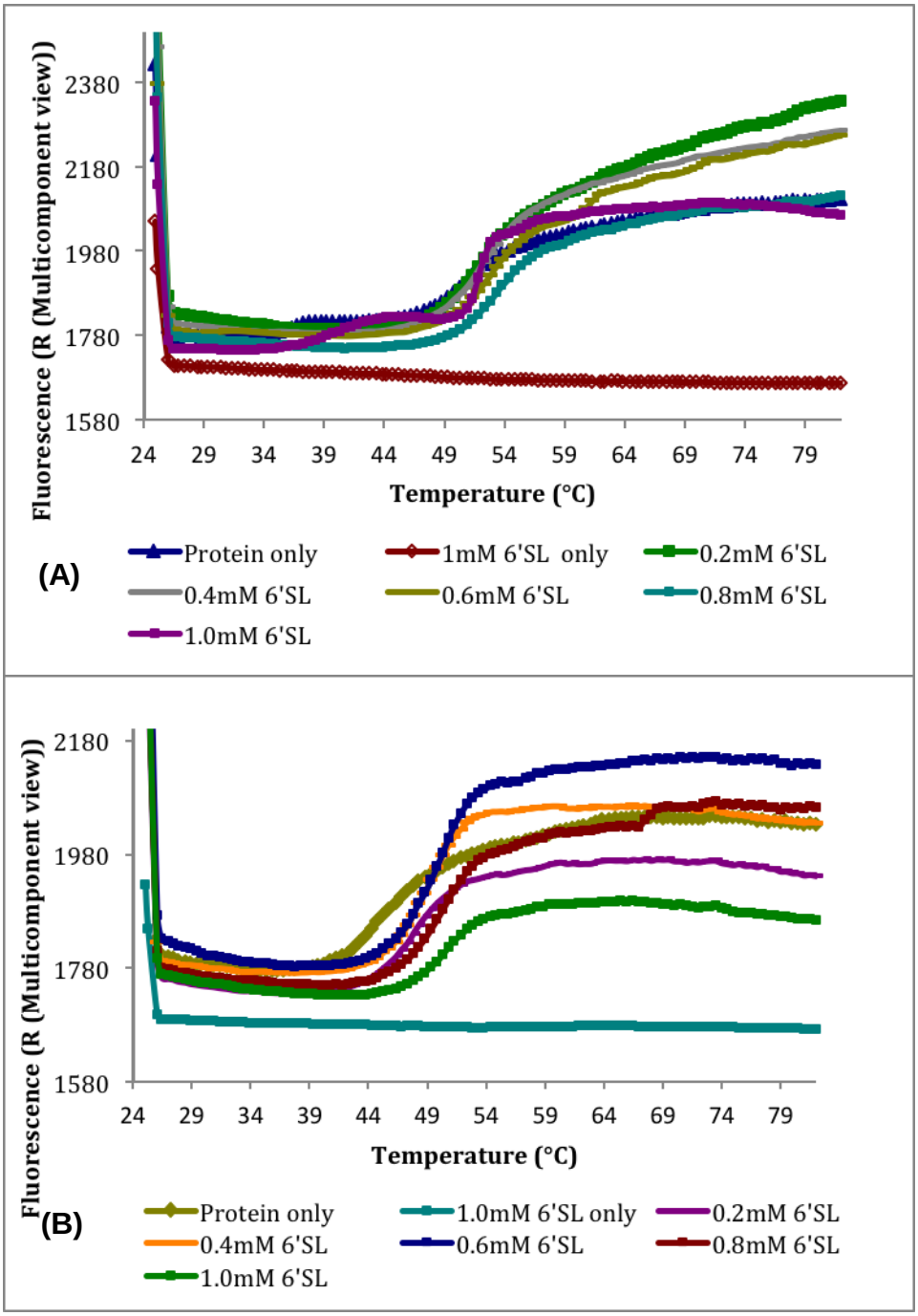


Figure 3.16 Fluorescent readout of Vc-CBMTD (Mutant) interaction with 3'-Sialyllactose (3'SL) in two different buffers. (A) 100 mM sodium citrate, pH 6.2 + 50 mM NaCl and **(B)** 20 mM Tris, pH 7.5 + 50 mM NaCl with the presence of 3'SL.

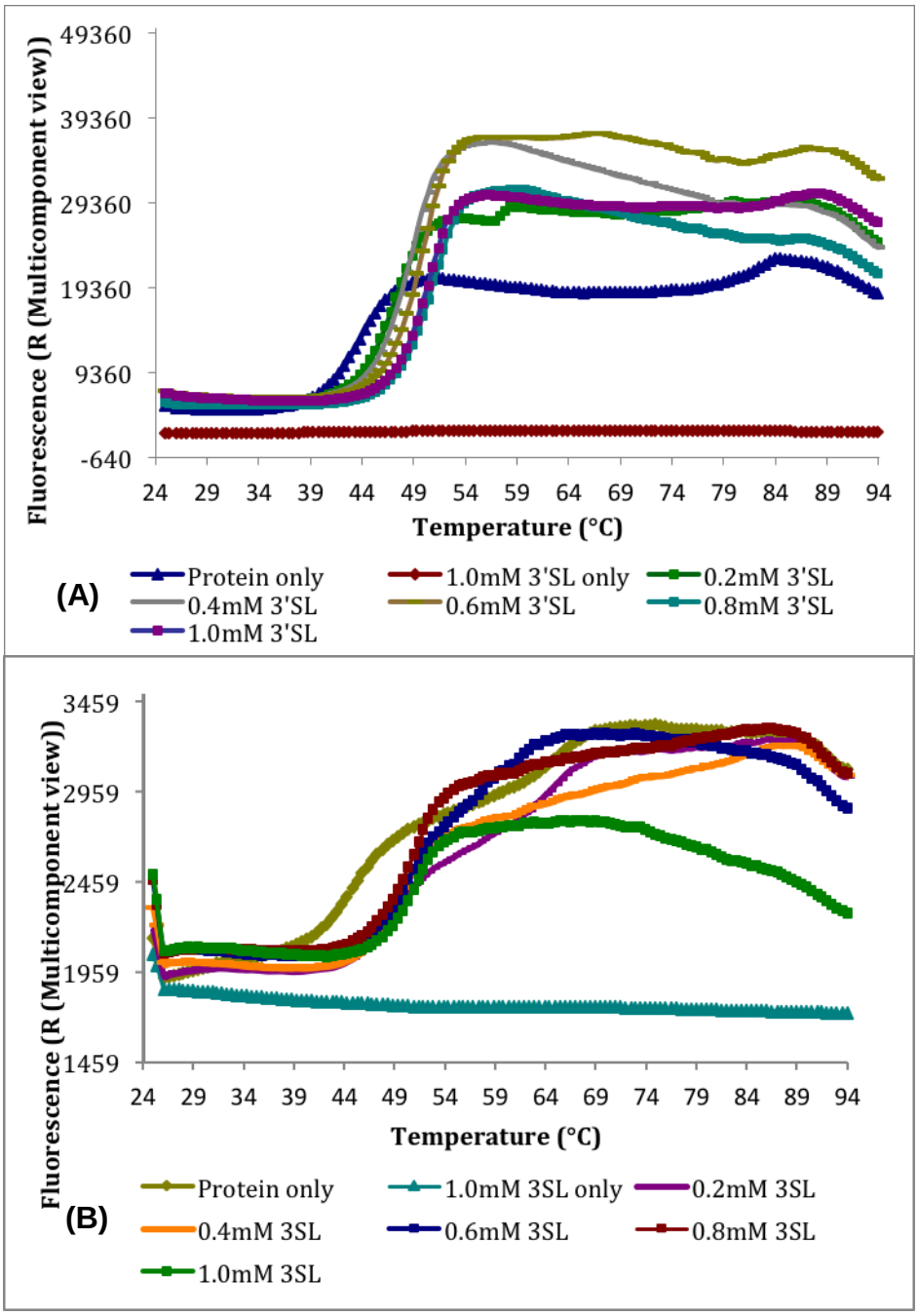


Figure 3.17 Fluorescent readout of Vc-CBMTD (Mutant) interaction with 6'-Sialyllactose (6'SL) in two different buffers. (A) 100 mM sodium citrate, pH 6.2 + 50 mM NaCl and **(B)** 20 mM Tris, pH7.5 + 50 mM NaCl with the presence of 6'SL.

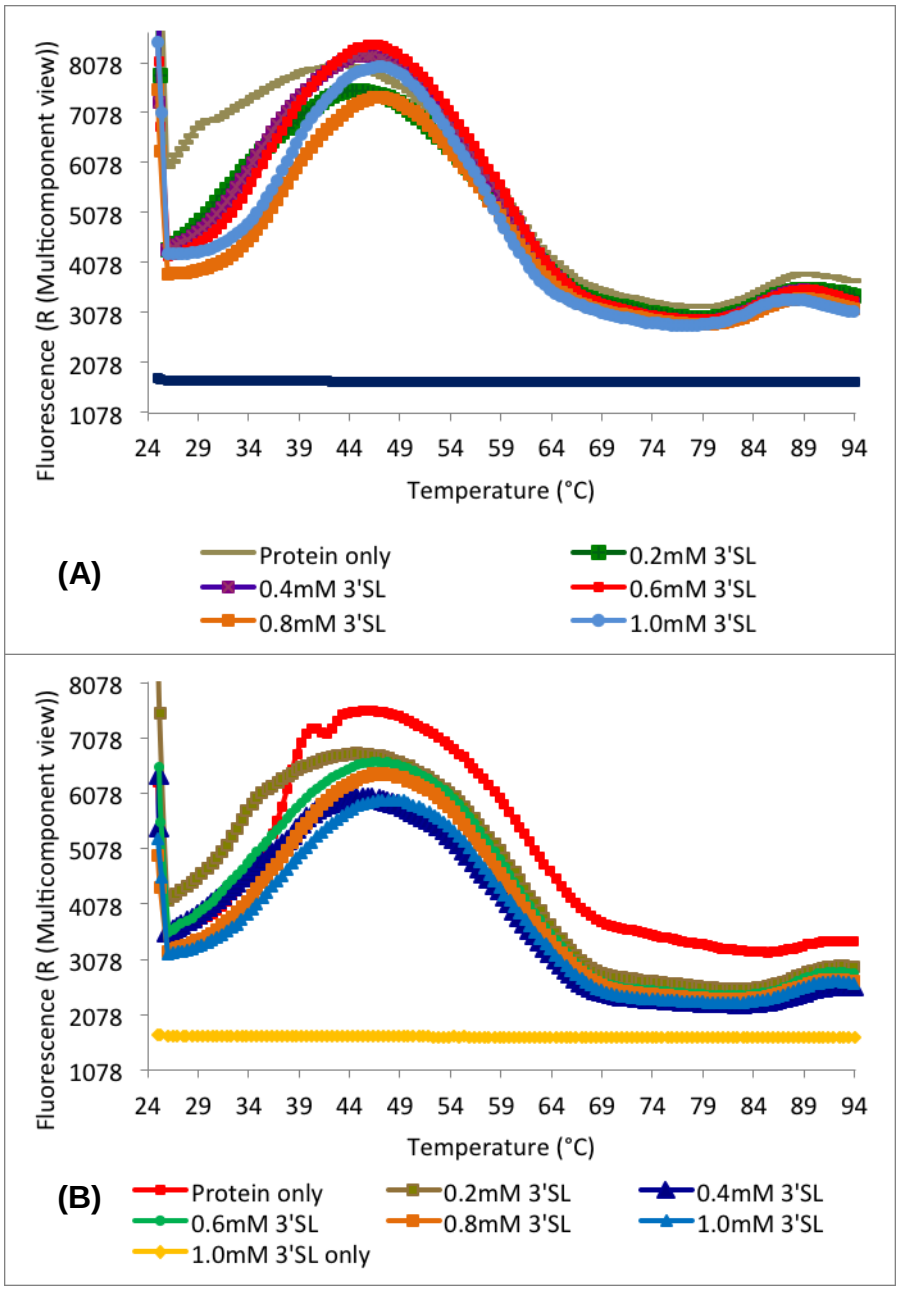


Figure 3.18 Fluorescent readout of Sp-CBMTD interaction with 3'-Sialyllactose (3'SL) in two different buffers. (A) 100 mM sodium citrate, pH 6.2 + 50 mM NaCl and **(B)** 20 mM Tris, pH 7.5 + 50 mM NaCl with the presence of 3'SL.

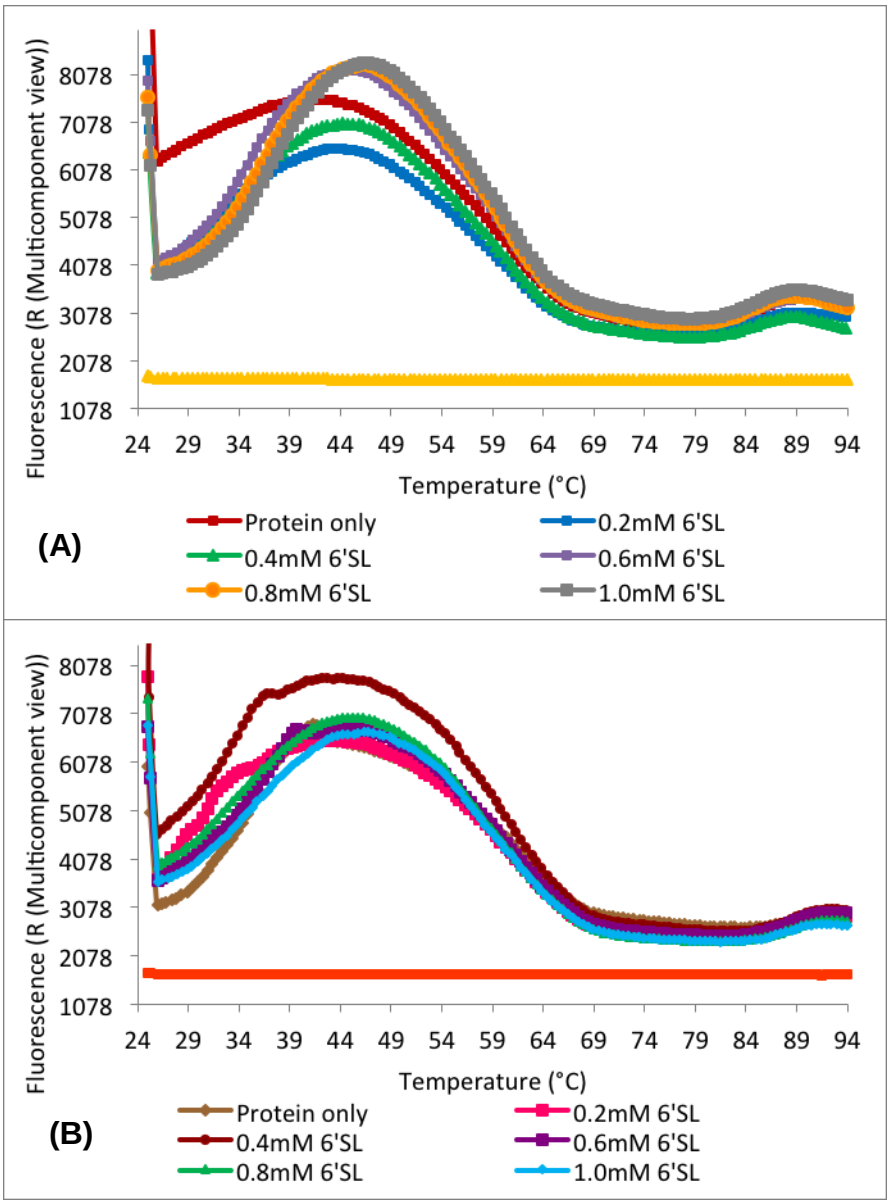


Figure 3.19 Fluorescent readout of Sp-CBMTD interaction with 6'-Sialyllactose (6'SL) in two different buffers. (A) 100 mM sodium citrate, pH 6.2 + 50 mM NaCl and **(B)** 20 mM Tris, pH 7.5 + 50 mM NaCL with the presence of 6'SL.

3.4 Thermodynamic study of multivalent CBMs

3.4.1 Isothermal titration calorimetry

Isothermal titration calorimetry (ITC) is a quantitative technique that can directly measure thermodynamic parameters of protein-protein interaction. When protein binds to ligand, heat is either generated or absorbed during a biomolecular event. Measurement of this heat allows accurate determination of binding constants (K_a), reaction stoichiometry (n), enthalpy (ΔH) and entropy (ΔS) (Freire et al., 1990). With regards to protein thermodynamics, protein stability depends on two opposing forces, which are enthalpy and entropy. Both are temperature dependent.

ITC measurements were done using a VP-ITC microcalorimeter from MicroCal Inc. (Northampton, MA) with a working cell volume of 1.4 mL at 25°C unless stated otherwise. Proteins were dialyzed extensively in 10 mM HEPES, pH 7.4 containing 0.15 M NaCl and ligands were dissolved in the degassed, filtered dialysis buffer. 3'-Sialyllactose (3'SL) and 6'-Sialyllactose (6'SL) were used as ligands and purchased from Dextra Labs (Reading, UK). Protein concentrations were determined from A_{280} using calculated molar extinction coefficients for Vc3CBM (113790 $M^{-1}cm^{-1}$), Vc4CBM (151720 $M^{-1}cm^{-1}$), Vc-CBMTD (WT) (67380 $M^{-1}cm^{-1}$), Vc-CBMTD (Mutant) (67505 $M^{-1}cm^{-1}$) and Sp-CBMTD (50880 $M^{-1}cm^{-1}$) respectively. The working protein concentration used were 0.06 mM for Vc1CBM, 0.02 mM for Vc3CBM, 0.005 mM for Vc4CBM, 0.03 mM for Vc-CBMTD (WT) and Vc-CBMTD (Mutant) and 0.06 mM for Sp-CBMTD constructs.

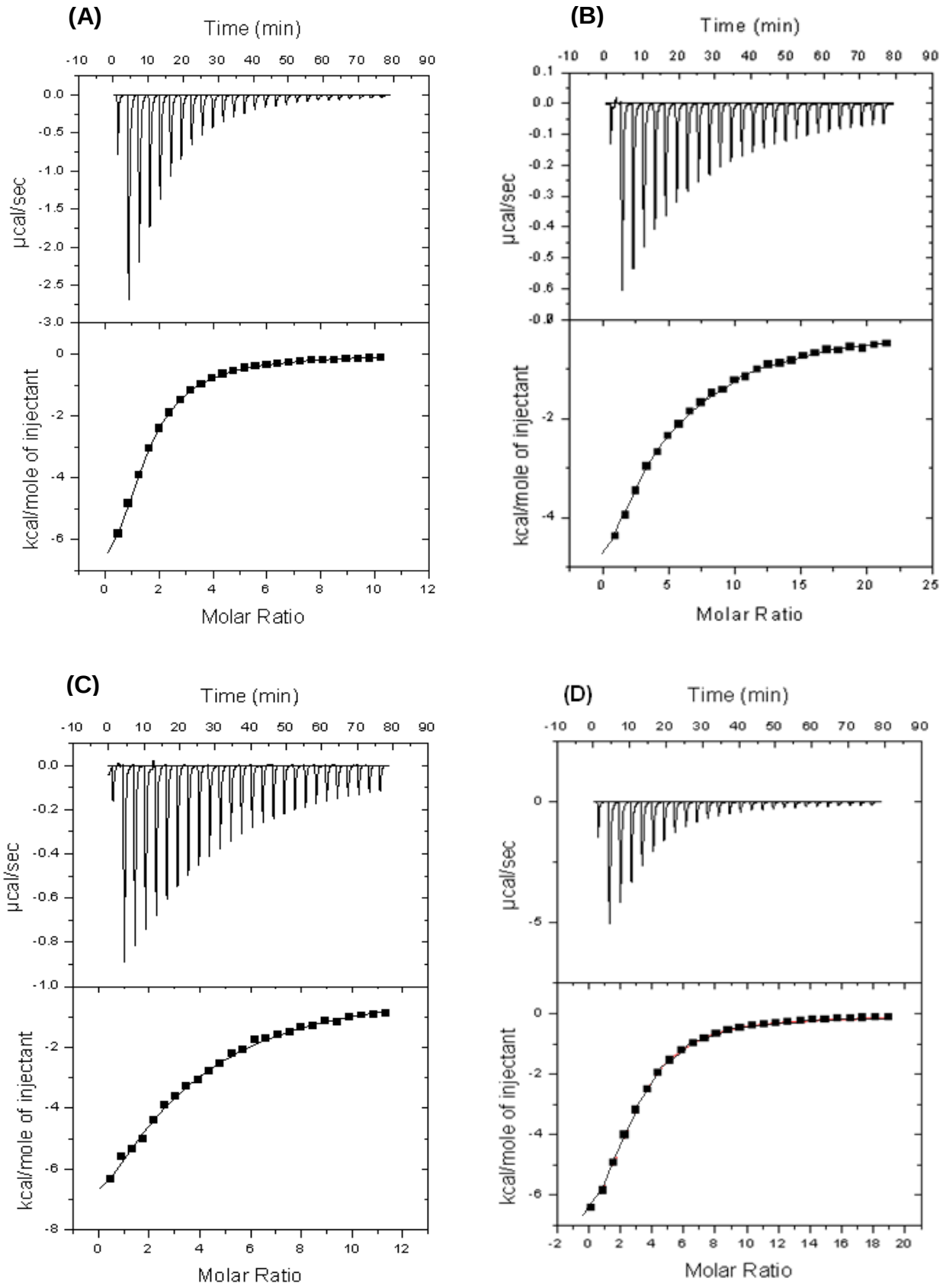
During titration, the protein and ligand mixture was stirred at 307 rpm in 1.4 mL reaction volume. A higher mixing speed was used to ensure rapid mixing of the titrant upon injection. 26 cell injections were used with 10 µL aliquots of ligands at 60 seconds intervals. Raw binding data were corrected for both protein and ligand. The heats of dilution were subtracted from binding isotherm data before data were fitted by nonlinear regression using a single-site binding model from MicroCal Origin software.

Research done by Connaris et al. (2009) to determine Vc1CBM specificity towards monovalent and divalent sialosides such as disialyllacto-*N*-tetraose (DSNT) and disialyllactose (DSL) revealed that Vc1CBM exhibits broad binding specificity towards different linked-α-sialic acids; these being α(2,3)-, α(2,6)- and α(2,8)-linked sialosides. From all sialosides tested, 3'SL and 6'SL promotes similar K_d value of 18 and 19 µM (Connaris et al., 2009). Using this information, ITC was performed using 3'-sialyllactose as a ligand with multivalent polypeptides to study their binding isotherms. Outputs of raw microcalorimetry data, binding isotherms and best curve fitting are as shown in Figure 3.20. From the data, the study demonstrated the exothermic nature of binding between multivalent CBM40 proteins and 3'SL. It was shown that the binding of 3'SL to the multivalent CBM40 polypeptides were enthalpy driven with ΔH values range from -5.4 to -16.2 kcal/mole at 25°C with entropy contributing unfavourably to the interaction (Table 3.1). This reflects the strength of the protein interaction to the ligand due to hydrogen bond formation and van der Waals interaction.

Figure 3.21 shows a summary of different thermodynamic features that were observed for six multivalent proteins with the same target, 3'SL. These results support the previous research from Connaris et al. (2009) which found that all six constructs have favourable ΔH that correspond to formation of direct and water-mediated hydrogen bonds and van der Waals interactions with their ligand. Therefore, a favourable enthalpy is a good indication

that the protein established a net strong interaction to its ligand that avoided any unfavourable entropy (Freire, 2004).

Ligand interaction with protein usually involves changes in the intermolecular or intramolecular interactions and dynamics of the whole system, and this includes the protein, ligand, water molecules and other components that may be present (Salemme et al., 1997). In this study, the ΔG was shown to be negative corresponding to a spontaneous reaction. This was also observed to be similar across all constructs, which related to binding affinity, suggesting that the more negative the value the more tighter the ligand binding to its protein. The *n* values demonstrated the appropriate number of binding sites for each multivalent constructs interacting with 3'SL based on single-side binding model.


(A)
Time (min)
µcal/sec
kcal/mole of injectant
Molar Ratio
(B)
Time (min)
µcal/sec
kcal/mole of injectant
Molar Ratio
(C)
Time (min)
µcal/sec
kcal/mole of injectant
Molar Ratio
(D)
Time (min)
µcal/sec
kcal/mole of injectant
Molar Ratio

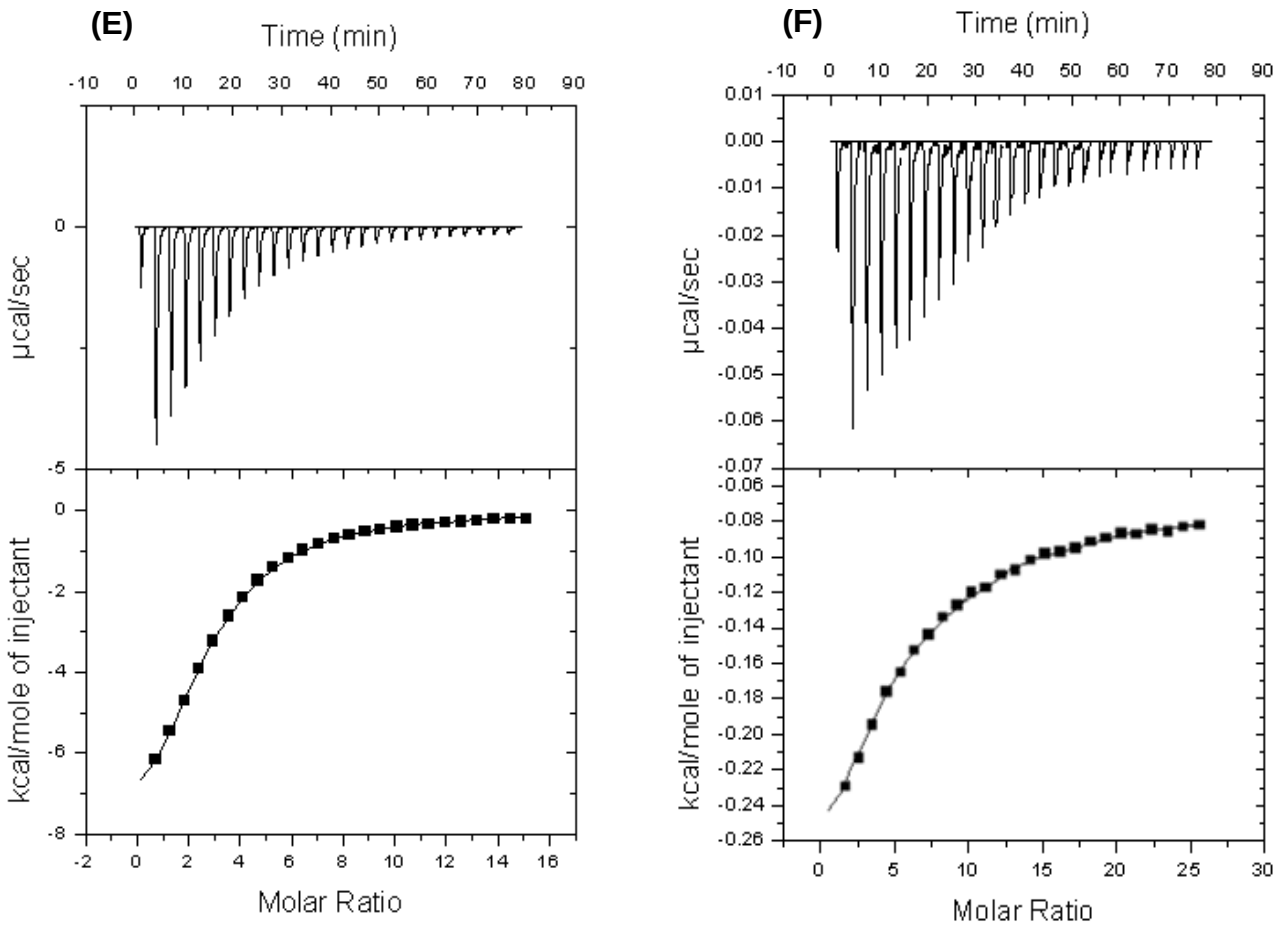


Figure 3.20 ITC Interaction profile of multivalent constructs. (A) Vc1CBM; **(B)** Vc3CBM; **(C)** Vc4CBM; **(D)** Vc-CBMTD (WT); **(E)** Vc-CBMTD (Mutant); **(F)** Sp-CBMTD. Upper panel: shows enthalpy changes upon injection of 3'-sialyllactose into the measurement cell containing polypeptides. Bottom panel: shows integrated power peaks fitted with a single-site binding model.

Table 3.1 ITC data on the binding of multivalent proteins to 3'-Sialyllactose.

Peptide	*n*	ΔH	$T\Delta S$	ΔG	K_a	K_d
		kcal/mol	***kcal/mol***	***kcal/mol***	***10^{-4} M^{-1}***	***µM***
Vc1CBM	1.08 ± 0.008	-12.93 ± 0.11	-6.5	-6.43	5.19 ± 0.06	19.23
Vc3CBM	2.93 ± 0.362	-16.17 ± 2.25	-9.96	-6.21	3.56 ±0.24	28.09
Vc4CBM	4.26 ± 0.22	-13.79 ± 1.30	-7.60	-6.19	3.13 ± 0.5	31.90
Vc-CBMTD (WT)	2.81 ± 0.041	-10.01 ± 0.19	-4.03	-5.99	2.41 ± 0.07	41.50
Vc-CBMTD(Mutant)	2.78 ± 0.024	-9.46 ± 0.104	-3.46	-6.00	2.46 ± 0.044	40.65
Sp-CBMTD	2.93 ± 0.033	- 5.4 ± 1.17	-0.17	-5.57	1.21 ± 0.005	82.60

*Note: Results from at least three repeat experiments.

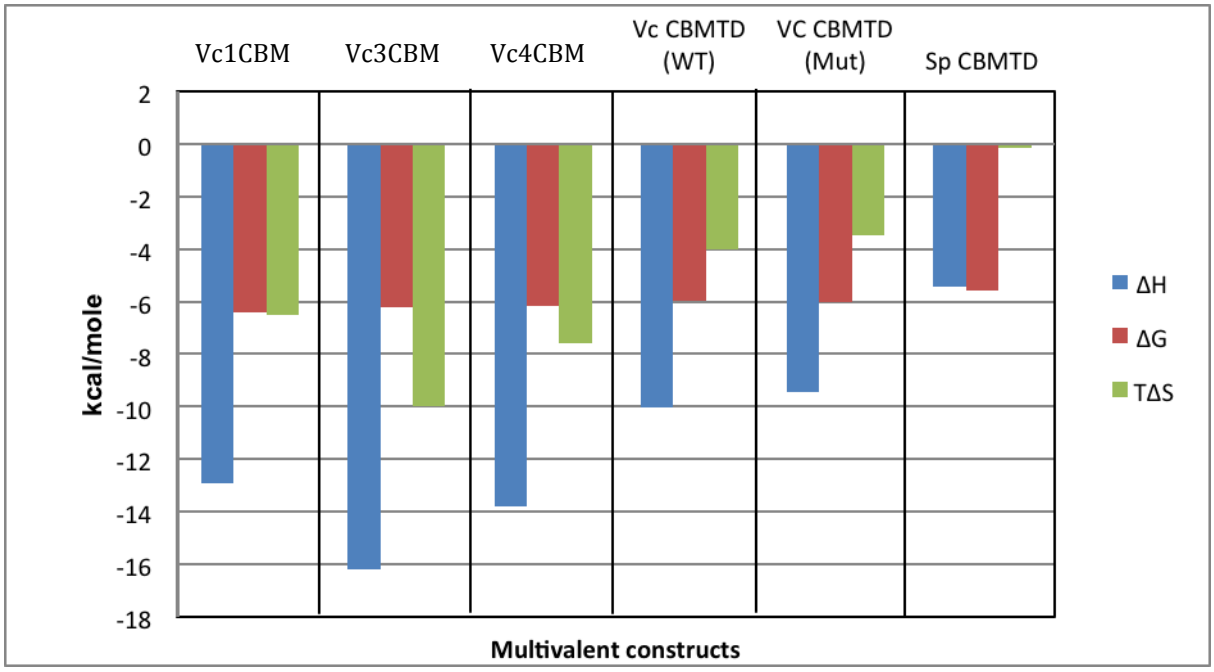


Figure 3.21 Summary of thermodynamic features for six different multivalents constructs with 3'SL. Details of each value were as summarized in Table 3.1.

Binding enthalpy of Vc3CBM compared to Vc1CBM increased by 3.24 kcal/mole, balanced by the increase in entropy, whereas the binding affinity of other multivalent constructs seemed to reduce slightly compared to Vc1CBM, corresponding to the reduction in Gibbs energy of binding. Interestingly, the binding enthalpy of Vc3CBM compared to the constructs containing the trimerization domain (Vc-CBMTD (WT) and Vc-CBMTD (Mutant)) were about 6.16 and 6.71 kcal/mole more favourable than the trimers. Both trimers were also shown to have reduced binding affinity compared to Vc3CBM corresponding to a loss in Gibbs energy. From Figure 3.21, it is shown that binding interaction between Vc3CBM and Vc4CBM with 3'SL has both favourable enthalpy (ΔH) and entropy (ΔS), which indicates that binding was driven by hydrogen bond formation as well as hydrophobic interaction. The fact that there was no obvious increase in affinity seen as the number of linked modules increase, suggests that the interaction between multivalent polypeptides to the sialic acid is similar to a multivalent-monomeric interaction. This indicates a very simple bimolecular association among multivalent CBM40 constructs.

All constructs containing PaTD domain, which is covalently linked to the CBM40 domain (Vc-CBMTD (WT), Vc-CBMTD (Mutant) and Sp-CBMTD, were found to have the smallest k_a. This might be due to the stereochemical restrictions resulting from domain arrangement. Vc-CBMTD (WT) and Vc-CBMTD (Mutant) have quite similar thermodynamic profile with each other, which indicates that mutation in the construct has a small effect on the binding affinity of the protein. Sp-CBMTD containing the CBM40 domain from *Streptococcus pneumoniae NanA* sialidase, were shown to have the lowest K_a of 1.2 x 10^{-4} M^{-1}. This suggests that its binding affinity to 3'SL is slightly weaker than any other constructs having *V. cholerae* CBM40 domain. For this protein, the binding enthalpy contributes -5.4 kcal/mol to the binding energy and the free Gibbs energy was about -5.57 kcal/mol.

From all the constructs, Vc1CBM was found to has the highest binding affinity with 3'SL with K_a of 5.19 X 10^{-4} M^{-1} followed by Vc3CBM (K_a: 3.56 X 10^{-4} M^{-1}) and Vc4CBM (K_a: 3.13 X 10^{-4}

M^{-1}) construct. This is in conjunction with the slowest dissociation rate showing stronger interaction to 3'SL. In overall, binding interaction of all the constructs with 3'SL were proved enthalpically driven like many other CBM-carbohydrate interactions previously reported (Boraston et al., 2004b).

3.4.2 Surface plasmon resonance (SPR)

Surface Plasmon Resonance has been established as a powerful method to monitor label-free macromolecular interactions in liquids. It provides high sensitivity and real-time monitoring of the recognition, binding and separation of two interacting molecules. With the goal to evaluate how well SPR measurements match the ITC data performed in solution, we compared the binding equilibrium constants, thermodynamics and kinetics of all multivalent CBM40 polypeptides. Moreover, in this section, we were also interested to study whether the avidity effect for sialic acid can be achieved through multivalency approach.

Binding kinetics of all multivalent constructs were determined using Biacore T-200 biosensor. In the SPR analysis, measuring binding reactions using SPR biosensors requires one of the binding partners be immobilized onto a surface (streptavidin-coated chip). Prior to use, this chip was pre-conditioned with three consecutive 1 min injections of 1 M NaCl in 50 mM NaOH. Biotinylated 3'SL-Polyacrylamide (PAA) was diluted to 1 mg/ml in PBS buffer containing 0.01% v/v surfactant P20 before injected over the flow cells in the chip. A reference surface was done to substract bulk effects and nonspecific interactions with streptavidin using PBS as running buffer at 30 µl/min flow rate.

For this binding analysis, each multivalent polypeptide interaction with immobilized 3'SL was tested in running buffer (PBS buffer containing 0.01% v/v surfactant P20) at three different temperatures (15, 25 and 37°C). All polypeptides were diluted in running buffer before being

injected over the flow cell surface at 100 µl/min with a contact time for each injection of 120 seconds. Purified protein constructs were diluted in running buffer to give a series of concentrations, for Vc3CBM, which were 1 nM, 5 nM, 20 nM, 62.5 nM and 125 nM. For other proteins (Vc4CBM, Vc-CBMTD (WT), Vc-CBMTD (Mutant) and Sp-CBMTD), proteins were tested at 18.0, 6.0, 2.0, 0.67 and 0.22 nM respectively.

For the dissociation test of analyte, the running buffer was used to flow over the cell surface at the same flow rate for 240 seconds. Surface was regenerated with two consecutive injections of 10 mM glycine-HCl, pH 3 containing 0.05% w/v SDS for 30 seconds at 30 µl/min. The response data for the binding affinity, as described by the equilibrium dissociation constant (K_D) were fit simultaneously to k_d/k_a model with a simple (1:1) Langmuir binding interaction model by using BIAevaluation T200 software (BIAcore). An example of the SPR binding response curves were as illustrated in Figure 3.22. Original SPR sensorgrams showing the binding of the CBM40 modules to immobilized 3'SL receptors are as shown in Appendix B.

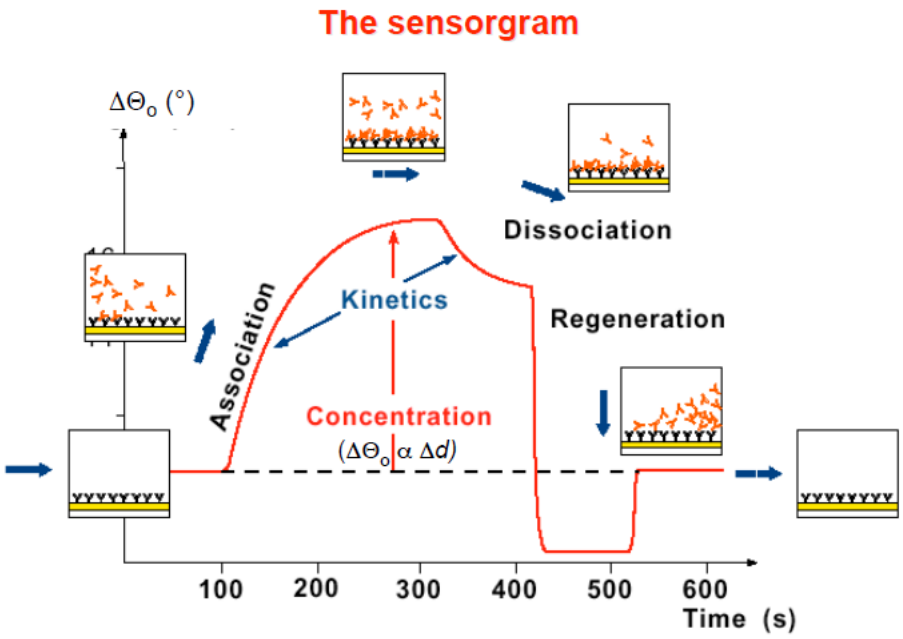


Figure 3.22 Interpretation of binding response curve by SPR. The result analysis was interpreted in three different response stages (association, dissociation and regeneration). This figure was adapted from (Roos et al., 1998).

The mean kinetic rate constants describing Vc3CBM and Vc4CBM were calculated to be k_a: 5.6 x 10^5 $M^{-1}s^{-}1$; 3.56 x 10^6 $M^{-1}s^{-}1$ and k_d: 2.3 x 10^{-3} s^{-1}; 1.6 x 10^{-3} s^{-1} respectively, which yielded equilibrium dissociation affinity of K_D: 4.0 x 10^{-9} and 4.5 x 10^{-10} M correspond to 8.9 fold increase in affinity at 25°C. A work done by Connaris et al. (2009) on Vc1CBM, has revealed the k_a values of 4.3 x 10^3 $M^{-1}s^{-1}$ with the equilibrium dissociation affiinity (K_D) of 1.8 x 10^{-6} M, thus provide an increment of 450 to 4000 fold from single-CBM40 module to three- and four-CBM40 modules linked together at 25°C. The disparity found in data between ITC and SPR were due to the different physical properties when the two assays were done to measure each protein binding.

Furthermore, the mutant version of Vc-CBMTD was found to display higher binding affinity, which was 6.3 X 10^{-11} M compared to 1.71 X 10^{-9} M for its corresponding WT. This finding

might be due to the domain arrangment cause by the mutation at S164C and T83C, which introduce a disulphide bond and thus alter the orientation of the binding sites on the protein for efficient substrate binding. As for trimers, the binding affinity of Sp-CBMTD was found to be better than Vc-CBMTD (WT) but lower than Vc-CBMTD (Mutant). Surprisingly, at 25°C, binding affinity of Sp-CBMTD to 3'SL was also stronger than Vc3CBM and Vc4CBM in SPR analysis which was different to the ITC data of binding. There are various reasons that could contribute to this result. One possibility could be due to differences in mode of interaction between a liquid-liquid state as in ITC to a semi-liquid/solid state as seen in SPR data.

Different temperatures were used (15°C to 37°C) to monitor temperature-dependent binding constants on the biosensor. In the current analysis, proteins and ligand interaction gave acceptable affinity readings throughout the temperatures tested. The highest affinity to 3'SL was found at 15°C reaction, which displayed the highest K_D of 6.4 X 10^{-11} M for Sp-CBMTD followed by Vc4CBM at 9.4 X 10^{-10} M. By using SPR, a 64- to 134-fold increase in affinity can be achieved through a multivalent approach, depending on temperature used during interaction. In order to obtain true thermodynamic parameters describing multivalency, a system must be in equilibrium and not kinetically-trapped in non-equilibrium state.

Table 3.2 Biacore kinetic parameters of multivalent CBMs interaction with 3'-sialyllactose.

Multivalent construct	T (°C)	k_a ($M^{-1}s^{-1}$)	k_d (s^{-1})	K_D (nM)
Vc3CBM	15	(3.0 ± 0.3) x 10^5	(1.1 ± 0.04) x 10^{-3}	3.7
	25	(5.6 ± 0.15) x 10^5	(2.3 ± 0.08) x 10^{-3}	4.0
	37	(5.2 ± 0.35) x 10^5	(4.4 ± 0.2) x 10^{-3}	8.44
Vc4CBM	15	(0.54 ± 0.005) x 10^6	(0.51 ±0.006) x 10^{-3}	0.94
	25	(3.56 ± 0.016) x 10^6	(1.60 ± 0.006) x 10^{-3}	0.45
	37	(2.98 ± 0.024) x 10^6	(5.41 ± 0.037) x 10^{-3}	1.82
Vc-CBMTD (WT)	15	(2.55 ± 0.016) x 10^6	(4.20 ± 0.019) x 10^{-3}	1.65
	25	(7.24 ± 0.059) x 10^6	(12.4 ± 0.096) x 10^{-3}	1.71
	37	(20.7 ± 0.54) x 10^6	(107 ± 2.8) x 10^{-3}	5.15
Vc-CBMTD (Mutant)	15	(2.9 ± 0.006) x 10^6	(0.31 ± 0.001) x 10^{-3}	0.106
	25	(2.6 ± 0.004) x 10^6	(0.17 ± 0.063) x 10^{-3}	0.063
	37	(1.8 ± 0.002) x 10^6	(0.21 ± 0.001) x 10^{-3}	0.114
Sp-CBMTD	15	(16.9 ± 0.01) x 10^6	(1.1 ± 0.004) x 10^{-3}	0.064
	25	(7.8 ± 0.035) x 10^6	(1.2 ± 0.004) x 10^{-3}	0.154
	37	(21.5 ± 0.04) x 10^6	(5.2 ± 0.006) x 10^{-3}	0.243

The equilibrium constants determined from the temperature-dependent analysis were used to determine van't Hoff enthalpies by plotting $\ln K_D$ versus $1/T$ which yielded values for $\Delta H°/R$ from the slope and $-\Delta S°/R$ from the y intercept, where R is the universal gas constant, 1.987 cal mol^{-1} K^{-1} (McNaught, 1997). Detailed kinetic and thermodynamic information will help in characterizing how these multivalent polypeptides interact with their macromolecular targets.

Figure 3.23 shows the van't Hoff plots for all multivalent constructs at three different temperatures measured (15, 25 and 37°C). The plot data were found to be linear and also consistent to the ITC binding parameters that were run at 25°C (Table 3.1). Data derived from van't Hoff plots for each multivalent constructs showed slight increases in ΔG with the addition of CBM40 domain as shown for Vc3CBM (-11.46 kcal/mol) and Vc4CBM (-12.75 kcal/mol). Other constructs containing the trimerization domain (Vc-CBMTD (WT), Vc-CBMTD (Mutant) and Sp-CBMTD) also displayed some increment in Gibbs free energy with values of -11.96, -13.92 and -13.38 kcal/mol respectively. All the data showed a large favourable entropic contribution upon complexation as shown in Table 3.3. These large changes in entropy could be due to the release of structurally ordered water molecules from the hydration shells influenced by tight-protein packing. It is very likely that the valency of the engineered tandem-linked proteins may play an important role in stabilization of protein oligomers and their interaction. Still, the Gibbs free energy was found strongly negative, which suggest spontaneous interaction.

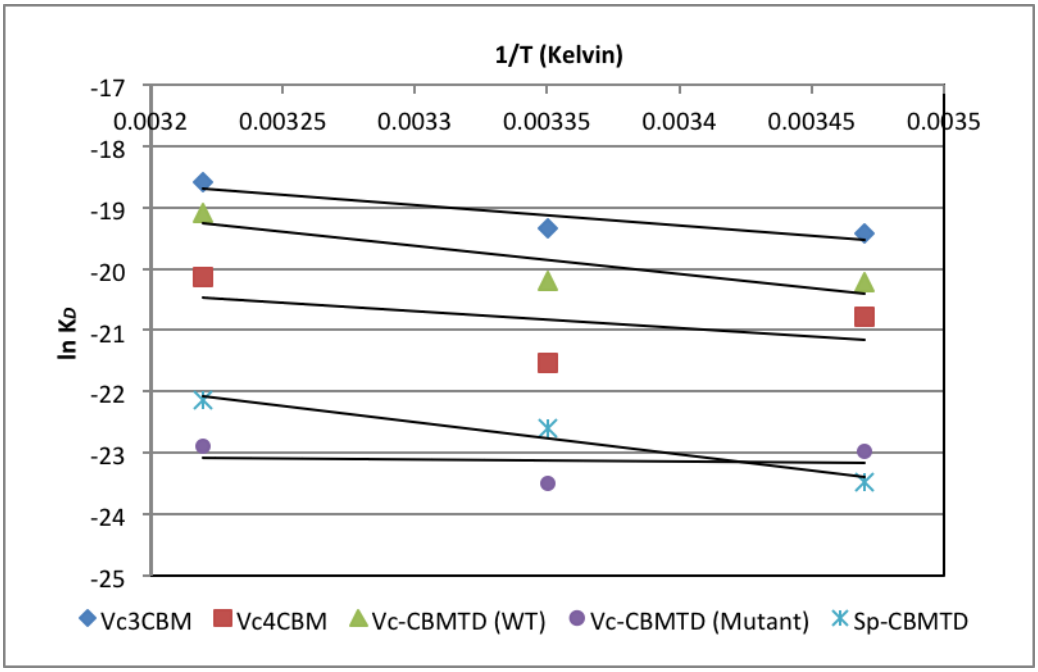


Figure 3.23 van't Hoff plot from SPR data at three different temperatures (15, 25 and 37°C) for all multivalent constructs.

Table 3.3 Thermodynamic parameters of the interaction between multivalent proteins with 3'SL-PAA Biotin.

Multivalent construct	ΔG *(kcal/mol)*	ΔH *(kcal/mol)*	ΔS (*cal $K^{-1}mol^{-1}$*)	TΔS *(kcal/mol)*
Vc3CBM	-11.46	-6.59	15.94	4.87
Vc4CBM	-12.75	-5.55	22.78	7.2
Vc-CBMTD (WT)	-11.96	-9.17	8.71	2.79
Vc-CBMTD (Mutant)	-13.92	-0.75	43.41	13.17
Sp-CBMTD	-13.38	-10.52	9.97	2.86

* Parameters were derived from van't Hoff plot using kinetic rate constant determined by SPR. The Gibbs free energy of activation at 25°C (298 K) was derived from the relationship ΔG = ΔH - TΔS.

Moreover, significant changes in enthalpy and entropy contributions can be seen from all the constructs. Out of all the proteins tested, Sp-CBMTD gave the largest enthalpic change with ΔH value of -10.52 kcal/mol followed by Vc-CBMTD (WT) with -9.17 kcal/mol. Meanwhile, Vc3CBM and Vc4CBM only showed a slight different in enthalpy at -6.59 and -5.55 kcal/mol and entropy contribution of 15.94 and 22.78 kcal/mol respectively. This small decrease in binding enthalpy provided a 10-fold enhancement in affinity probably due to formation of stable aggregated on SPR chip surface (Poon, 2007).

In contrast, Vc-CBMTD (Mutant) showed the lowest favourable enthalpy change of -0.75 kcal/mol but with the highest entropy gains of 43.41 cal $K^{-1}mol^{-1}$. The large differences in equilibrium dissociation energy were also obviously seen between these two constructs (Table 3.2). Through mutation it showed that, even a single amino acid change in the sequence can significantly alter the thermodynamics/ kinetics of the protein. This finding

could be based on various factors such as protein conformational arrangement, ligand accessibility to the binding sites, mode of binding including intra- or intermolecular binding and structural packaging of the protein (Connaris et al., 2009).

3.5 Structural study of Vc-CBMTD

3.5.1 Protein crystallization and X-ray data collection

A pre-crystallization test (PCT) was performed (Appendix A-12), which estimate a starting protein concentration of 10 mg/ml for Vc-CBMTD (WT) construct in 20 mM Tris, pH 7.5 + 50 mM NaCl. Seven commercialized crystal screens kits such as JCSG+ (Molecular Dimension), Wizard I + II (Emerald Bioscience), Crystal screen (Hampton Research), PACT (Molecular Dimension), Index (Hampton Research), PEGs (Qiagen) and Cryo I + II (Emerald Bioscience) were used for initial crystallization trials. Crystal plates were initially setup using Honeybee crystallization robot with 150 nl of buffer and 150 nl of protein. This experiment was carried out using the sitting drop method. All of the plates were stored at room temperature of 20°C. Two different protein concentrations were used in the initial screens; higher concentration was 18 mg/ml and lower concentration was 10 mg/ml.

Fortunately, after six months, a few beautiful crystals were found in PACT-H5 screening plate when using 18 mg/ml of Vc-CBMTD (WT). The plate condition was 0.2 M sodium nitrate, 20 % PEG3350 and 0.1 M Bis-tris propane, pH 8.5 with equal amounts of protein and reservoir solution (2 µl each). This Vc-CBMTD (WT) crystal that grew in this condition crystallize with a bubble-like shape (Figure 3.24). Wells, with crystalline materials growing in them, were marked for further examination. The initial crystal was found to be very soft and jelly-like when picked out from the drops. In order to distinguish the protein crystals from salt crystals, 0.1 µl Izit Crystal dye (Hampton Research) was mixed with the drops contain tiny

crystals for 30 minutes before observed under a light microscope. Protein crystal should be stained as a dark blue crystal and the corresponding crystallization conditions were selected as the optimization starting points. For further confirmation of proteins crystals, a few crystals were exposed to the X-rays at 100K from an in-house Rigaku/MSC MicroMax-007HF rotating anode equipped with a Saturn 944+ CCD detector at wavelength 1.54178 Å. The crystal-to-detector distance, rotation angle and other parameters were initially set but very poor diffraction data was produced and obtained measurement at 10 Å resolution. For a further comprehensive study, optimization of the protein positive hit condition was performed and discussed in the next subsection.

3.5.2 Crystal optimization

From the initial hit, the crystal condition was optimized by varying the concentration of salt, precipitant and protein. The best quality crystals were obtained in an optimization condition of 0.18 M sodium nitrate, 22% PEG3350 and 0.1 M Bis-tris propane, pH 8.5 with 18 mg/ml of protein used. The crystal displayed a hexagonal shape about the size of 150 µm x 150 µm x 150 µm, which took about 4 to 5 months to form crystals (Figure 3.25). Additive Screen HT (Hampton Research) that contains a library of 96 unique reagents, was also used to improve and manipulate the condition for crystal optimization and alter sample solubility. However this approach did not work at all for the protein. For further data collection, crystal-to-detector distance, rotation angle and other parameters were initially set to measure resolution data based on the diffraction quality of the crystal. The first crystal of Vc-CBMTD (WT) diffracted at 3.2 Å in-house after the optimization stages.

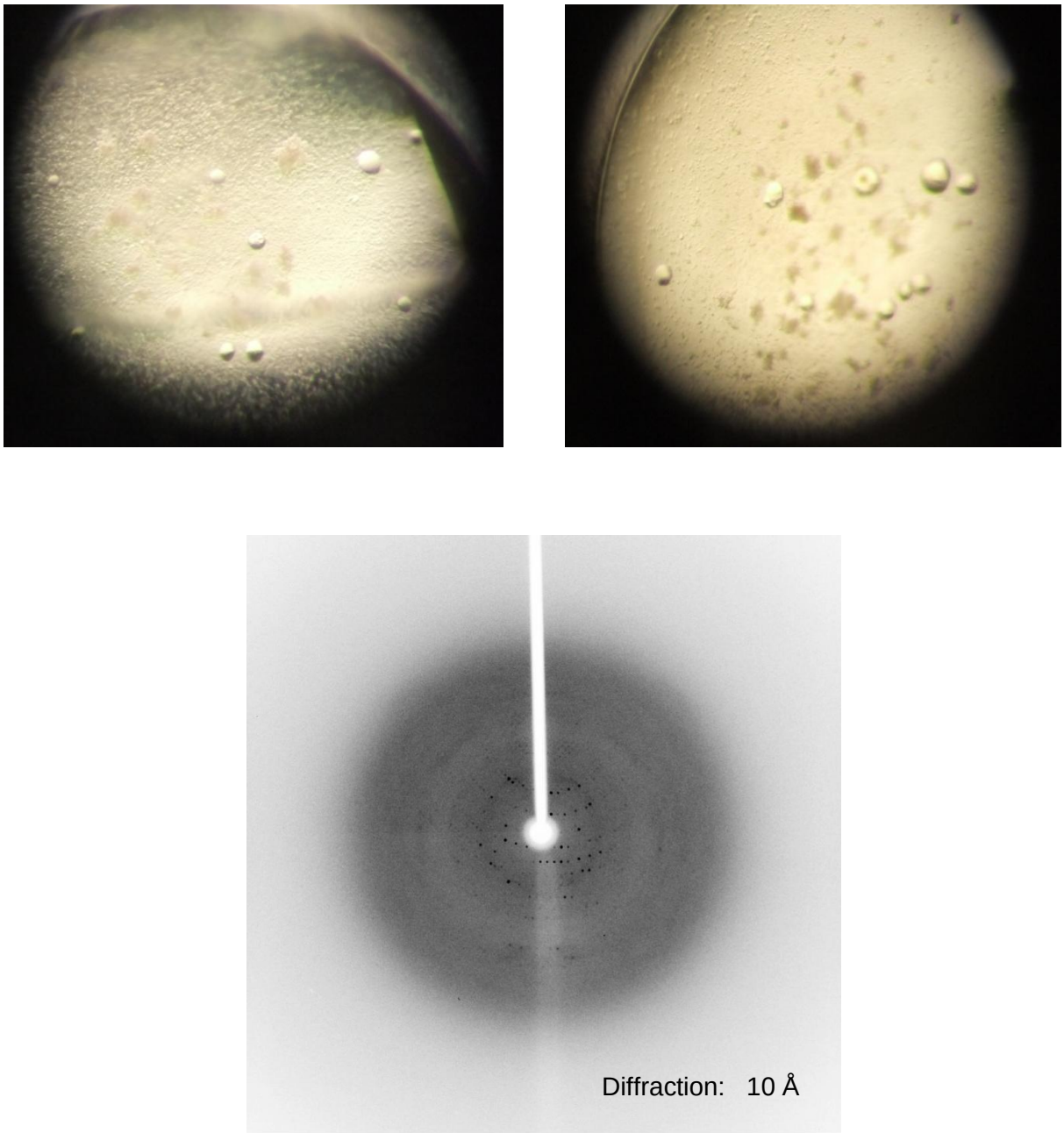


Figure 3.24 Diffraction pattern of Vc-CBMTD (WT) crystals formed before optimization conditions. The initial condition was 0.2 M sodium nitrate, 20% PEG3350 and 0.1 M Bis-tris propane, pH 8.5.

For comprehensive analysis, Vc-CBMTD crystals were sent to the Diamond Light Source facility in Oxford. Prior to that, crystals were picked out from the drops with a crystal loop (1.0-0.5 mm size) and transferred to 5 µl of cryo protectant drop containing 20% glycerol. The crystals were then re-looped and transfer to a storage container containing liquid nitrogen prior to delivery to Diamond Light Source facility. Diffraction data were collected using the I03 macromolecular beamline equipped with Pilatus 6M-F detector**.** The x-rays from this beamline deflect and scatter atoms in equivalent positions in the crystal lattice as concentrated and sharp intense spots known as crystal diffraction pattern.

3.5.3 X-ray data collection, data processing, molecular replacement and structure refinement.

The first dataset was collected at Diamond Light Source facility in Oxford using I03 macromolecular beamline equipped with Pilatus 6M-F detector. The crystallization data were indexed and integrated using the program MOSFLM (Leslie 1992). The integrated data were merged and scaled using the programs called CCP4 suite (CCP4. 1994). From the analysis, the crystal belongs to the group primitive cubic (P $4_1$32) with the unit cell parameters a= 115.17 Å, b= 115.17 Å and c= 115.17 Å. Matthew's coefficient calculation suggested that only one molecule of Vc-CBMTD (WT) was present in the asymmetric unit with solvent content of 36%. The apo form of Vc-CBMTD (WT) dataset was collected with a crystal that diffracted to 3.0 Å at Diamond Light Source facility. Data collection and processing statistics are listed in Table 3.4.

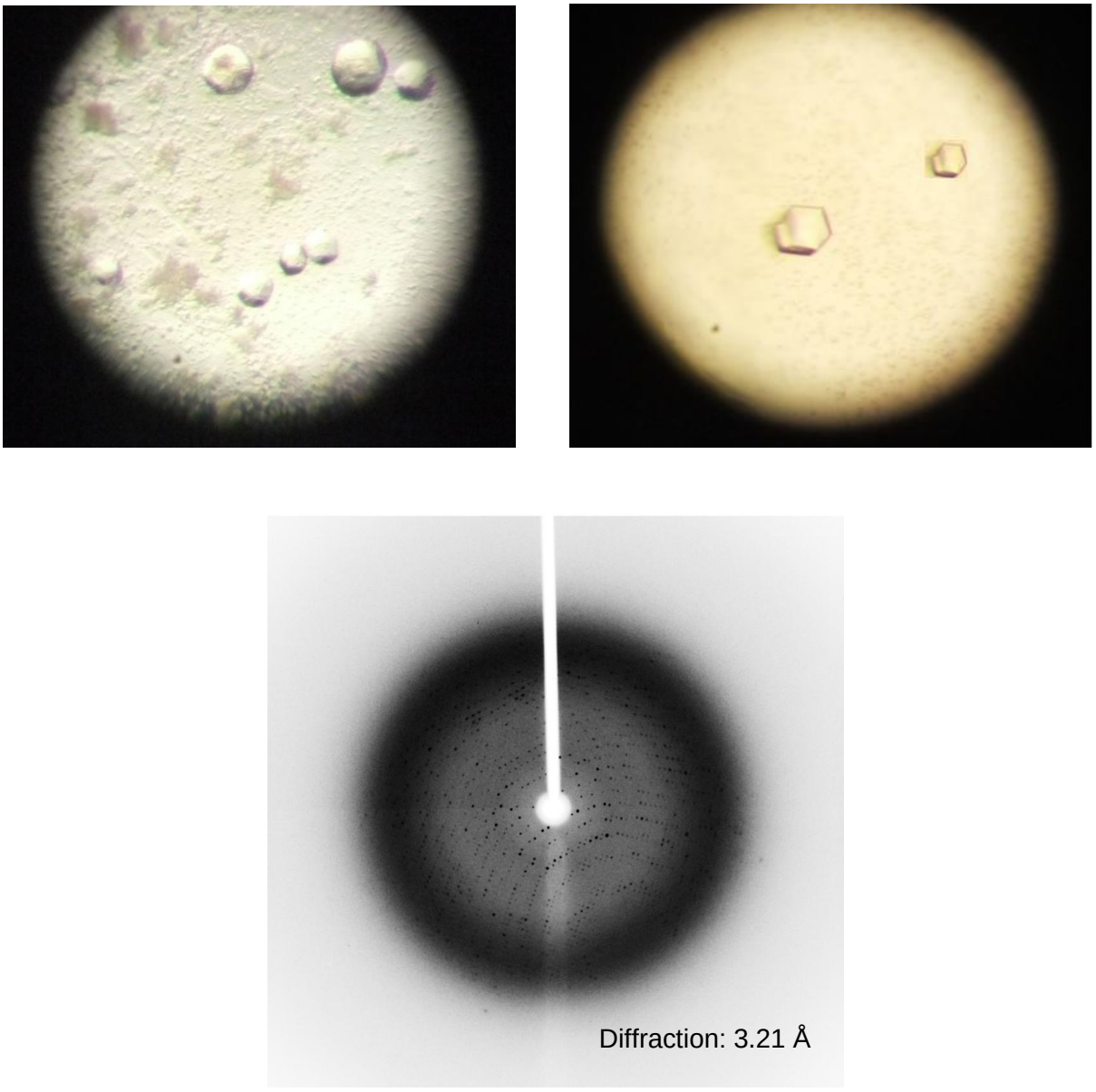


Figure 3.25 Vc-CBMTD (WT) crystals formed after optimization condition with its x-ray diffraction image. The buffer condition was 0.18 M sodium nitrate, 22% PEG3350 and 0.1 M Bis-tris propane, pH 8.5. The crystal size was about 0.15 x 0.15 × 0.15 mm.

Table 3.4 X-ray data collection and refinement statistics.

Data collection		
Space group		P $4_1$32
Unit cell edges	a, b, c (Å)	115.17/ 115.17/ 115.17
	α, β, γ (°)	α = β= γ= 90°
X-ray source		Diamond Light Source
Resolution range (Å)		51.51- 3.0 (3.08-3.00)
No. of unique observations		285223
Completeness (%)		99.7 (98.5)
Redundancy		7.4 (5.4)
R_{merge}		0.144 (0.413)
$I/\delta I$		11.0 (3.5)
Data refinement		
No. of reflections work/ test		37254 / 5618
No. of protein atoms		828
No. of residues		108
No. of chain		1
Average *B*-factors (Å^2)		36.19
R_{cryst} [b]		0.218
R_{free}		0.256
Root mean square deviation bond distance (Å)		0.0170
Root mean square deviation bond angle (°)		1.9088

[a] $R_{merge} = \Sigma_{hkl} \Sigma_i |_{hkl,i} - \langle I_{hkl} \rangle| / \Sigma_{hkl} \langle I_{hkl} \rangle$, [b] R_{cryst} and $R_{free} = (\Sigma||F_o| - |F_c||) / (\Sigma|F_o|)$.

Numbers in parentheses correspond to the highest resolution shell.

Molecular replacement (MR) was selected to solve the structure using the *Vibrio cholerae* sialidase CBM structure (PDB code: 2w68) and the *Pseudomonas aeruginosa* pseudaminidase trimerisation domain (PDB code: 2w38) as the search model. Unfortunately, from automated MR with Phaser program (McCoy, 2007) in Collaborative Computational Project Number 4 (CCP4) program (CCP4, 1994) it was suggested that the molecule was only the trimerization domain (TD) from *Pseudomonas aeruginosa* pseudaminidase, which crystallised without the presence of the CBM40 domain from *Vibrio cholerae* sialidase. It is believed that this CBM40 domain had degraded during the long crystallization periods (4-5 months) and only part of trimerization domain, had successfully crystallized. However, the trimerization domain was built and refined, with an R factor and R free of 0.22 and 0.26 respectively using Coot (Emsley and Cowtan, 2004) and REFMAC5 program (Murshudov et al., 1997). One molecule was present in the asymmetric unit which formed a trimer with two symmetry-related molecules.

Similar to that, described by Xu et al. (2009b), the pseudaminidase from *Pseudomonas aeruginosa* (PaNA) in the present structure consists of an additional domain with an immunoglobulin-like (Ig-like) fold. This domain contains residues from 335-438, which is linked to the C-terminal of the catalytic domain via a 7-residue loop. This Ig-like domain is known as a trimerization domain, which folds into 11 β strands. These strands form two anti-parallel β sheets packing, which form against each other, as shown in Figure 3.26. Such a trimeric association is also similar to the fibronectin type III domain found in certain tumor necrosis factor (TNF)-like cytokines for example the mouse TNF (Baeyens et al., 1999). Moreover, the pseudaminidase oligomerisation domain is also an essential part that trimerizes PaNA monomers to the trimer architecture (Figure 3.27). All figures and protein structures in this section were prepared using Pymol (DeLano, 2007). As shown in Figure 3.28, the surface charge potential of the trimerization domain from *P. aeruginosa* was showed in two orientations of electrostatic distribution.

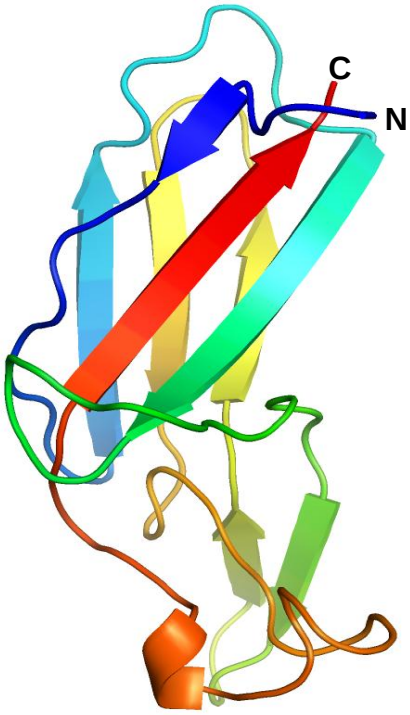


Figure 3.26 Schematic drawing of trimerization domain (TD) of *P. aeruginosa* pseudaminidase from the crystallography data. This domain consists of two anti-parallel β-sheets packed against each other, which stands for an Ig-like β-sandwich fold. The monomer is drawn in rainbow colours from blue at the N-terminus to red at the C-terminus.

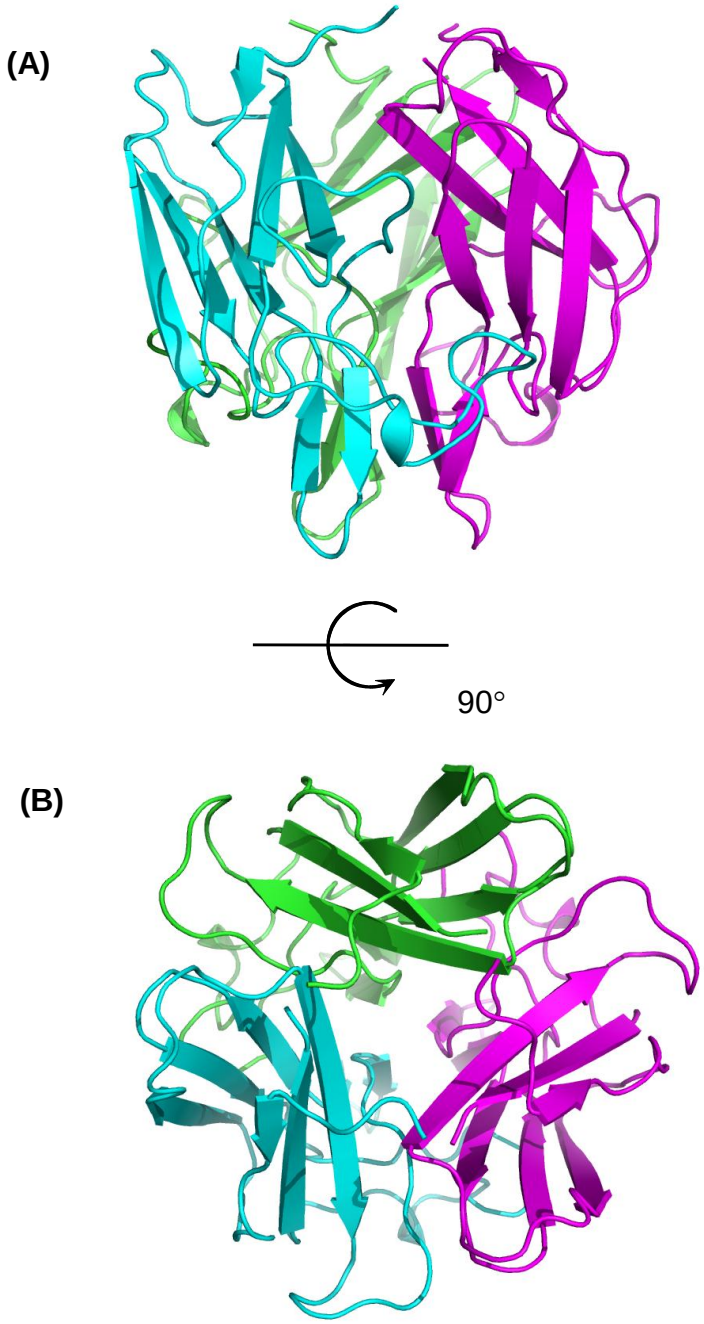


Figure 3.27 Trimerization domain (TD) of *P. aeruginosa* pseudaminidase (PaNA) as a trimer. One molecule (blue) was present in the asymmetric unit and formed a trimer with symmetry-related molecules (green and magenta). **(A)** Views from the side of the TD; **(B)** views from the three-fold symmetric axis.

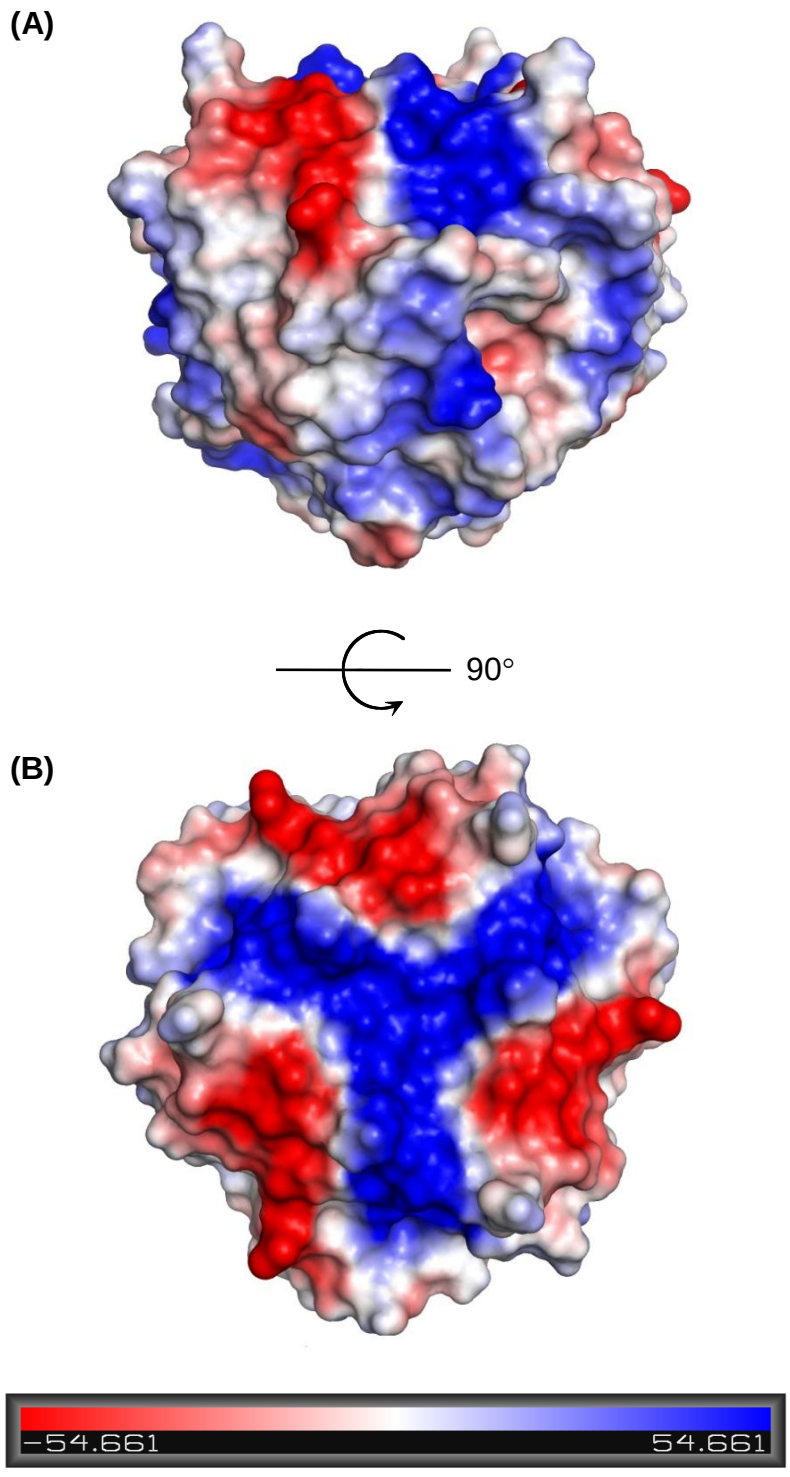


Figure 3.28 The electrostatic potential of molecular surface of TD domain calculated by PyMol. Colours were according to the electrostatic potential.

3.5.4 Analysis of the natively disordered region in full-length Vc-CBMTD

The full-length 308 amino acid sequence of Vc-CBMTD (WT) was submitted to the bioinformatics server, Regional Order Neural Network (RONN) for protein analysis (Yang et al., 2005). As shown in Figure 3.29, 99 residues above the baseline of "Probability of Disorder" (0.5 at y axis) are thought to be disordered. Regions that are involved include residues 1-33, 62-70 and 75-77 at the N-terminus and residues 215-267 located at the C-terminus.

From the amino acid analysis using ProtParam software (Expasy), this protein is very rich in proline/glycine (P/G) and threonine/serine (T/S) residues, which constitute 15.5% and 18.2% in 308 amino acid residues, respectively. The glycine residues would give this domain higher conformational flexibility while proline residues reduce flexibility and forms rigid peptide bonds (Whitford, 2005). Both C-and N-terminal parts of the amino acid from this protein was found to be disordered. This raises the concern that these disordered regions may decrease the possibility of the protein to crystallize. In fact, this full length of Vc-CBMTD (WT) failed to crystallize after screening against more than 670 conditions.

However, despite all that, both domains (CBM40 and PaTD domains) are found to be well ordered on their own. It is in contrast with the RONN analysis where disordered regions were predicted on both domains. This showed some weakness and limitation in the analysis software. Due to this problem, an accurate recognition of proteins disordered regions is very important. But until now, no single disorder prediction techniques are so accurate which can be entirely trusted (Yang et al., 2005). Furthermore, classification of disordered regions in PDB are still incomplete (Berman et al., 2000) due to many protein structures were solved based on truncated polypeptides with some parts (disordered and flexible regions) were

being removed (Linding et al., 2003; Yang et al., 2005). This has clearly led to some limitation in prediction software like the RONN analysis.

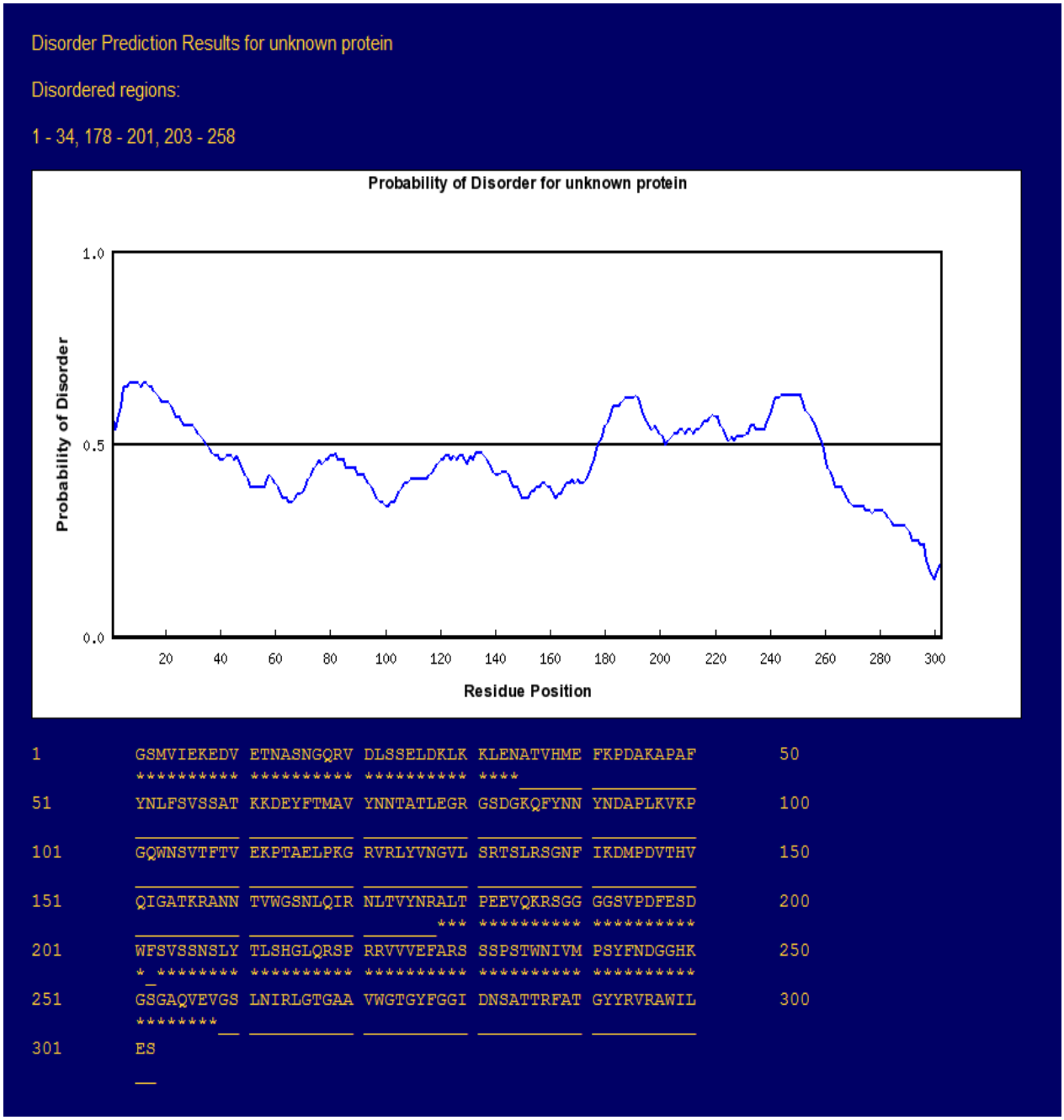


Figure 3.29 Prediction of natively disordered region of full-length Vc-CBMTD (WT) by RONN bioinformatics server. The horizontal lines mark the threshold for disorder prediction and region above the lines correspond to disorder protein sequences.

3.6 Structural study of Sp-CBMTD

3.6.1 Protein crystallization

Analysis by Prot-Param has revealed Sp-CBMTD consist of 302 amino acids with molecular weight of 33252 Da. Its estimated iso-electric point (pI) is found to be 9.32. From a pre-crystallization test, 9 mg/ml of protein concentration was suggested as a starting concentration for crystallization screening of Sp-CBMTD. Seven crystal screens, namely Hampton crystallization screen and Index (Hampton Research), the PEGs (Qiagen), CRYO and Wizard I+II (Emerald Biosciences), PACT and JCSG+ (Molecular Dimensions) were used. All crystal plates were initially setup using Honeybee crystallization robot using a sitting drop method. The lower concentration of 9 mg/ml and the higher concentration of 27 mg/ml of the protein were used for crystal plate set up, but unfortunately no crystals were found to date.

Due to this finding, RONN analysis (Yang et al., 2005) was performed to predict disordered protein regions. From the analysis, there were three disordered regions found, which were amino acid residues 1-34 at the N-terminal and the C-terminal residues of 178-201 and 203-258 of the Sp-CBMTD domain (Figure 3.30). In this Sp-CBMTD construct, the linker region between CBM40 domain and trimerization domain was designed consist of ‘GGGGS’ sequences which were flexible amino acids. This region was also predicted to be disordered by the RONN analysis. Besides that, parts of the N-terminal and C-terminal sequences were also in disordered regions which also promote to the difficulty in crystallization process. Future work plan on this protein is described in Chapter-7.

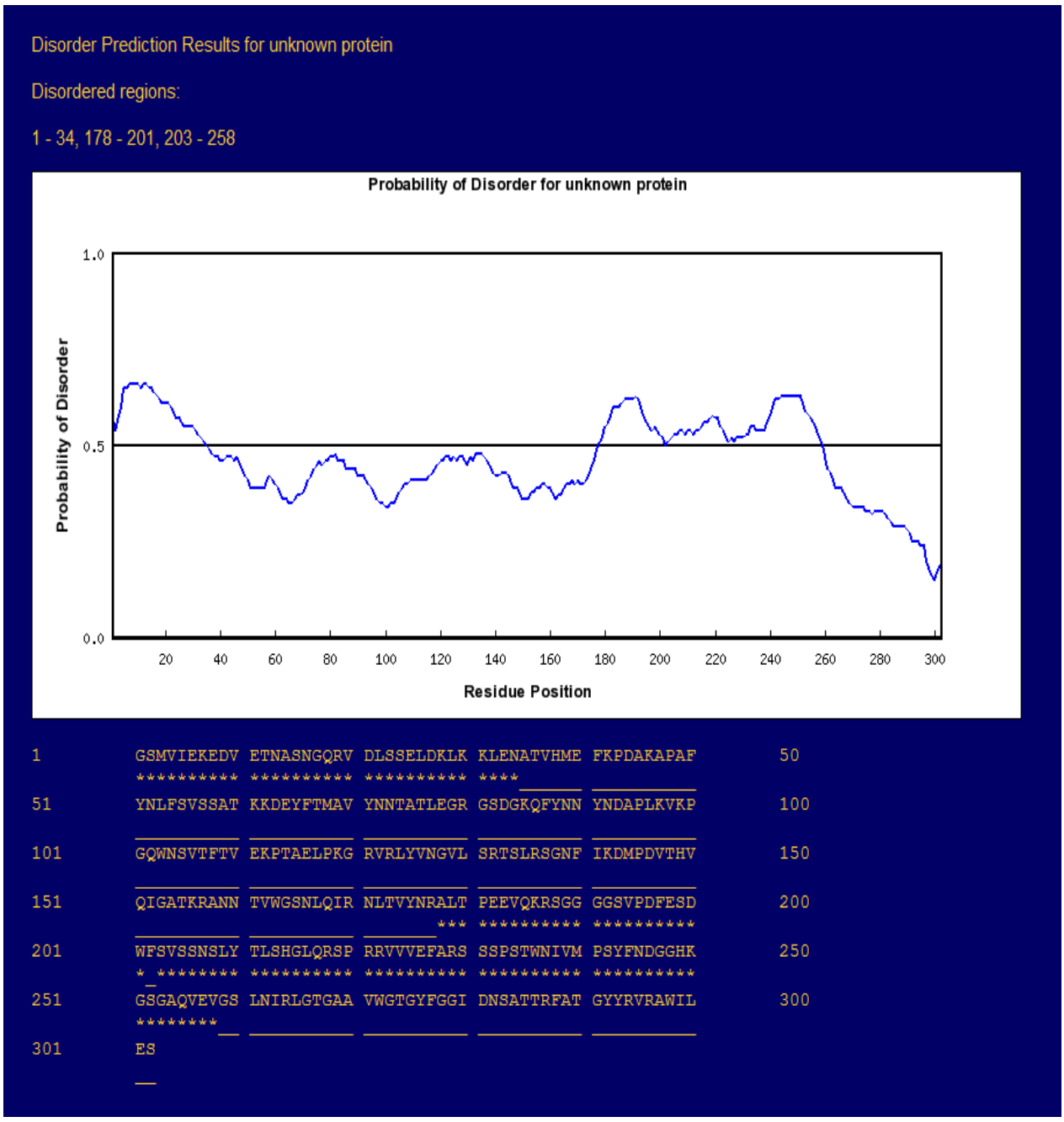


Figure 3.30 Disorder prediction region of Sp-CBMTD protein predicted by RONN analysis. Disordered region of 1-34 were found at the Sp-CBM domain while region of 178-201 and 203-258 were at the TD domain of the construct.

3.7 Discussion

Multiple, simultaneous interactions are often used in biology to enhance protein affinity and binding specificity. This form of approach is known as multivalency. Using native gel electrophoresis, the native state of an engineered multivalent protein can be studied. Protein migration under native conditions is dependent on molecular mass, pI, buffer pH, and type and percentage of gel matrix (Braz and Howard, 2009). The results from the above study have shown that both proteins (Vc-CBMTD (WT) and Vc-CBMTD (Mutant) are presented as high molecular weights species, which correspond to protein aggregates as compared to Vc1CBM. In this assay, other proteins (Vc3CBM, Vc4CBM and Sp-CBMTD) have not been tested because this was a preliminary study to show the presence of high molecular weight protein through multivalency technique that could lead to protein aggregation. It has been reported that structurally altered proteins have a strong tendency to aggregate, which lead to precipitation (Chen et al., 1994).

Systematic identification of ligand and buffer conditions which optimally stabilize protein would significantly improve the success rate of crystallization trials. Due to this, a thermal shift assay was chosen as the technique to identify these conditions. In the assay, ten different buffers were tested to provide suitable condition for protein stability. A good buffer system usually will have maximum buffer capacity at a pH where the protein exhibit optimal stability (Ugwu and Apte, 2004). For the construct containing the *P. aeruginosa* pseudaminiase trimerization domain (Vc-CBMTD (WT), Vc-CBMTD (Mutant) and Sp-CBMTD), two peaks were present, as shown in the Figure 3.17, which correspond to the CBM40 and trimerization domains. Usually this multimeric, or modular proteins often have more than one T_m value due to the presence of additional domains in the protein structure. One such example is the native human factor XIII, which consists of three thermolabile domains and two thermostable domains in each subunit. The intact protein was found to have distinct T_m values which were 69°C and 90°C for both domains respectively (Kurochkin

et al., 1995). Another example is the chimeric protein toxin sCD4 consisting of two domains (HIV binding domains and cytotoxic domain of *Pseudomonas exotoxin*), which is also known to has two transition temperatures from both domains (Davio et al., 1995).

According to Connaris et al. (2009), Vc3CBM showed broad linkage specificity to α(2,3)-, α(2,6)-sialyllactose (3'-SL and 6'-SL) respectively, and di-sialyllactose (α2,8 linked) after undergoing a glycan array screen. The best ligand concentrations of 0.8-1.0 mM were found for all the proteins when using 3'SL and 6'SL as the substrates in the thermal shift assay. Protein interaction with specific ligand would increase the protein thermal stability. This can be measured through increment of melting temperature due to the stabilization of the protein complex. The difference in temperature of the midpoint in the presence and absence of ligands are related to the binding affinity of the substrate molecule. These molecules can be low-molecular-weight compounds, peptides or nucleic acids.

Data from our comparative study revealed that most of the thermodynamic parameters determined from the SPR Biosensor were not equivalent to the parameters determined by solution-based method, isothermal titration calorimetry (ITC). All multivalent CBM40 proteins interaction with 3'SL were found to be exothermic, releasing heat during complex formation. For ITC and SPR assay, the experiments were replicated at least three times at a constant temperature of 25 °C. ITC data measured the interaction of the multivalent polypeptides to the 3'SL shown to be enthalpy driven with entropy contributing unfavourably to the interaction. This favourable enthalpy interaction is a good indication that the protein established a strong binding to its ligand and avoid any unfavourable entropy (Freire, 2004). In SPR study, a mixed balance of entropy and enthalpy contributions were found on multivalent constructs where Vc-CBMTD (Mutant) was predominantly driven by entropy, while Vc-CBMTD (WT) and Sp-CBMTD showed increase in enthalpy. For Vc3CBM and Vc4CBM, their thermodynamic profiles provide equal balance of both enthalpy and entropy contributions. These findings showed that binding does not occur through a simple protein-

protein interaction (ions of opposite charge) but also through the hydrophobic and/or ionic hydration that provide the driving force to the process.

This kind of entropy-driven interaction has been seen in other multivalent protein-carbohydrate interactions due to formation of aggregates by intermolecular binding (Ambrosi et al., 2005; Lundquist and Toone, 2002). Interestingly, increasing the number of linked modules from three to four demonstrated an increase in entropy contribution. The free energy of binding for Vc4CBM was slightly greater than Vc3CBM corresponding to a slight increase in affinity. From this finding, the valency of multiple linked polypeptides was thought to play an important part in the protein stabilization. As described by Poon (2007), when all domains in an oligomer are tethered together (covalent linkage) in a single tandem, oligomerization of the protein becomes a unimolecular folding transition. This conformation would reduce the entropy of dilution to minimize dissociation and produce stable multivalent molecules.

For constructs containing the trimerization domain (Vc-CBMTD (WT), Vc-CBMTD (Mutant) and Sp-CBMTD), results from the study were comparatively different to our findings in ITC. The introduction of a mutation at S164C and T83C in Vc-CBMTD (WT) resulted in an increase in affinity along with an increase in Gibbs free energy, despite having a large increment in entropy contribution. The use of a cysteine mutant will strategically cross-link the adjacent subunits with disulphide bridges and thus promote better orientation of the binding site for easier access of sialic acids. On the other hand, this approach also contributes to a common side effect called irreversible aggregation due to non-specific cross linking which can cause increase in entropy (Poon, 2007).

Even though both results from ITC and SPR were comparable and acceptable, further optimization steps should be done. This includes optimising parameters such as flow rate, types of buffers, number of injections, protein and substrate concentration, temperature and

experimental design (different binding models). In some cases, a number of experimental artefacts (mass transport, nonspecific binding, aggregation, avidity, matrix effects and instrumental drift) can complicate binding responses (Morton and Myszka, 1998; Myszka, 1997). Through this effort, hopefully both data quality and data analysis will be optimized to ensure results are interpreted correctly.

For crystallization trials, two proteins (Vc-CBMTD (WT) and Sp-CBMTD) were initially set up in the optimization buffer (20 mM Tris, pH 7.5 + 50 mM NaCl). These constructs contain an additional domain known as trimerization domain (TD) from *P. aeruginosa* pseudaminidase, which was exploited to generate trimeric CBM40 molecules. Only the TD part was successfully crystallized from Vc-CBMTD (WT) protein in 0.18 M sodium nitrate, 22% PEG3350 and 0.1 M Bis-tris propane, pH 8.5 and took about 5 months to form crystals. Besides that, the presence of glycine at high percentage in the amino acid sequence can promote domain flexibility that might inhibit orderly packing of macromolecules in a crystalline array which complicate the crystallization process (Jin and Babine, 2004; Whitford, 2005).

Multi-domain proteins are hard to crystallize due to the flexibility of the linkers that connect the distinct domain. Due to that, the linker region plays a very important part in any engineered multivalent protein molecule. An ideal linker should be ‘inert’ yet rigid so that the position and orientation of tandem repeats is constrained in its oligomeric state (Mammen et al., 1998a). It is often attempted when synthesizing multivalent ligands to design peptide linkers, which would not interact and perturb the core structure of the oligomer (Kiessling et al., 2000). Due to that, the choice of suitable sequence should consist of hydrophilic residues (glycine and serine) that are not likely to interact with the structural surface charges on the other subunits. In addition, it should avoid any sequence that may contains proteolytic cleavage sites (Poon, 2007).

To date, no crystal of Sp-CBMTD was found on any screening plates. From RONN analysis, three disordered regions were found, which comprised of >100 residues, including regions from both domains that are known to be ordered. This was only a prediction analysis and shows some limitations in the ability of such software to predict the disordered regions. The failure to obtain crystals of full length Vc-CBMTD and Sp-CBMTD, is most likely due to the flexible linker region that was introduced during the engineering of the proteins.

Chapter 4

In vitro study of influenza virus infection

4.1 Overview

Nowadays, influenza continues to be a present danger to human and animal health. The ability of the virus to develop rapid mutation can cause antigenic variants to constantly arise with potential to cause pandemics. The 2009 H1N1 swine flu pandemic, which occurred in Mexico rapidly spread around the world causing major public health costing billions of US dollars to control. The appearance of pandemic influenza is countered by regular seasonal flu epidemics. Although it is less dramatic than pandemics, seasonal influenza has caused more than 200,000 hospitalizations (Thompson et al., 2004) and annual average of >30,000 influenza-associated causing deaths. These estimated also highlight the increased morbidity and mortality in older age groups and variability in disease between seasons (Thompson et al., 2006).

The main strategy to fight influenza is through vaccination. But due to time constraints and other limitations, antiviral drugs are a second choice of defence after vaccination, which may help reduce severity of symptoms and shorten the time course of the disease. It is well documented that there is increasing resistance to current influenza antivirals, such as in the case of adamantine derivatives (Symmetrel® and Flumadine®), which stated by CDC are resistant to most of the influenza A H3N2 (2008/2009) and pandemic strains of H1N1 (2009). While, about 99.6% of seasonal influenza A H1N1 was found resistant towards oseltamivir. This has led to the need of alternative methods to combat influenza.

The ubiquitous distribution and terminal location of sialic acids have become interesting targets for various pathogenic organism and toxins to bind (Ilver *et al*., 2003). The presence of cell surface sialic acids in the airway epithelium, serve as a binding site for influenza viruses and is the first step in the virus infection process (Suzuki, 2005). In this study, we report on the antiviral properties of CBM40 multivalent proteins to block attachment of several influenza A virus strains *in vitro*. The level of viral inhibition is determined by virus plaque inhibition assays and expresses as EC_{50} values. Cytotoxicity tests, in absence of virus, were also performed to see direct effect of proteins at higher concentration on mammalian cell monolayers. For data comparison, quantification of viral replication proteins after treatment with multivalent proteins was also carried out in addition to the plaque inhibition assay mentioned previously.

4.2 Effect of multivalent CBMs protein in inhibiting influenza virus infections

4.2.1 Influenza virus plaque assay protocol

Confluent monolayer cultures of Madin-Darby canine kidney (MDCK) cells were incubated with different concentrations of proteins for 2 hours at 37°C, 5% CO_2 on a rocking platform with gentle shaking. Monolayers were washed once with Dulbecco's modified minimum essential medium (DMEM) before infected with influenza A virus strains at 200 pfu/well per virus for 1 hour with gentle shake.

Later, virus was removed and 13 ml of 2% molten agarose and 13 ml of 2X overlay medium (DMEM/ 90 mM $NaHCO_3$/ 10 mM HEPES, pH 7.4) supplemented with 52 µg of N-acetyl trypsin (NAT) (Sigma) was added to each well. Plates were left at room temperature for 15-20 minutes to solidify. Plates were inverted and incubated at 37°C, 5% CO_2 for 3 days or until plaques formed. For staining, 1 ml of 10% formaldehyde in PBS was added to each

wells and left for 1-2 hours to fix the monolayers. Plaques formed were counted by direct staining using 0.1% crystal violet. The plaques were counted by visual examination and percentage of plaque inhibition was calculated as relative to control plate (contain virus only).

4.2.2 Assessment of protein anti-influenza activity

The ability of various proteins to inhibit influenza virus infection was determined using a plaque inhibition assay on MDCK cells. Figure 4.1 and Figure 4.2 shows a representative assay of Vc4CBM protein. Assays were done with virus in the absence of protein as a mock control and with increasing protein concentrations. From Table 4.1, Vc4CBM is highly efficient at inhibiting the laboratory-adapted A/Udorn/72 H3N2 subtype with an EC_{50} of 0.59 µM. Modification of the Vc3CBM construct with additional CBM40 domain resulted in improved antiviral activity against the same virus from 0.94 µM for Vc3CBM to 0.59 µM for Vc4CBM. Vc4CBM maintained to be the best inhibitor for other influenza A virus strains, A/PR8/34 and A/WSN/33 H1N1 with EC_{50} of 1.8 µM and 1.06 µM as compared to Vc3CBM with 4.7 µM and 1.09 µM respectively.

The CBM40 proteins contained the trimerization domain from *P. aeruginosa* displayed a good inhibition effect with EC_{50} of 1.4 µM for Vc-CBMTD (WT) and 2.1 µM for Vc-CBMTD (Mutant) after infection with A/Udorn/72 H3N2 subtype. Moreover, these proteins also maintained a good level of protection against A/PR8/34 H1N1 virus with EC_{50} values of 3.01 µM for Vc-CBMTD (WT) and 5.01 µM for Vc-CBMTD (Mutant). While for A/WSN/33 H1N1 subtypes, the EC_{50} values are 1.5 µM for Vc-CBMTD (WT) and 2.0 µM for Vc-CBMTD (Mutant) respectively. These two proteins are closely related differing only in the position of mutations in CBM40 domain which are at S164C and T83C. While, for Sp-CBMTD contains a CBM40 domain from *S. pneumoniae* sialidase, are still able to maintain substantial antiviral activity against the influenza A virus strains, resulting in EC_{50} value of 2.2 µM against

A/Udorn/72, 2.03 µM against A/PR8/34 and 3.04 µM against the A/WSN/34 H1N1 subtype. The result, as summarized in Table 4.1, shows proteins containing the CBM40 domain from *V. cholerae* sialidase being the most effective protein in inhibiting influenza A virus replication. Supplementary data and detail graphs related to EC_{50} value of the multivalent protein tested can be referred to Figure 4.5 - Figure 4.7.

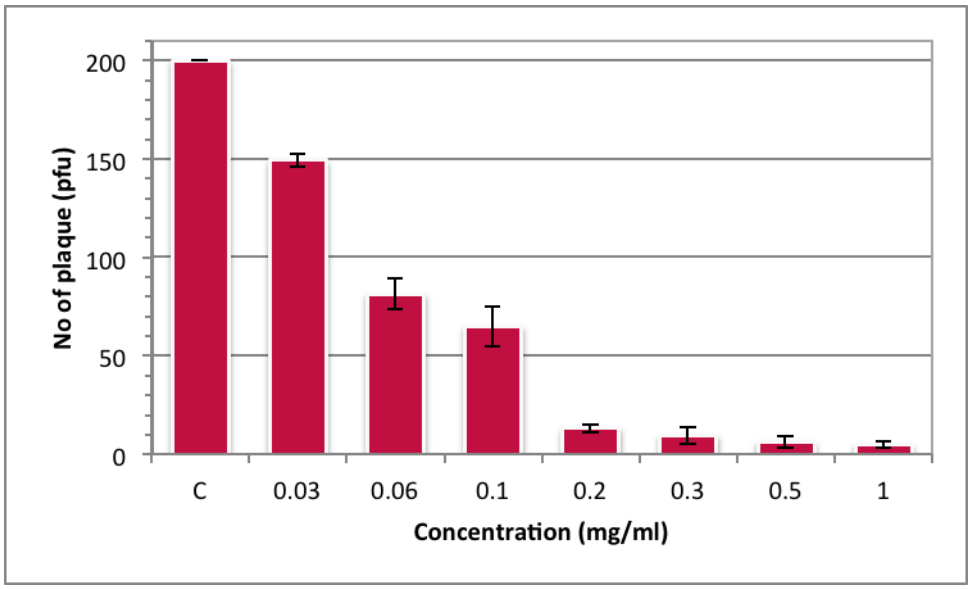


Figure 4.1 Anti-influenza activity of protein tested in plaque inhibition assays. The ability of Vc4CBM to inhibit replication of A/Udorn/72 H3N2 virus (200 pfu) was determined using plaque inhibition assays on MDCK cells. Assays were carried out with the presence of only virus and virus in the presence of increasing concentrations of protein (0.03 mg/ml to 1.0 mg/ml).

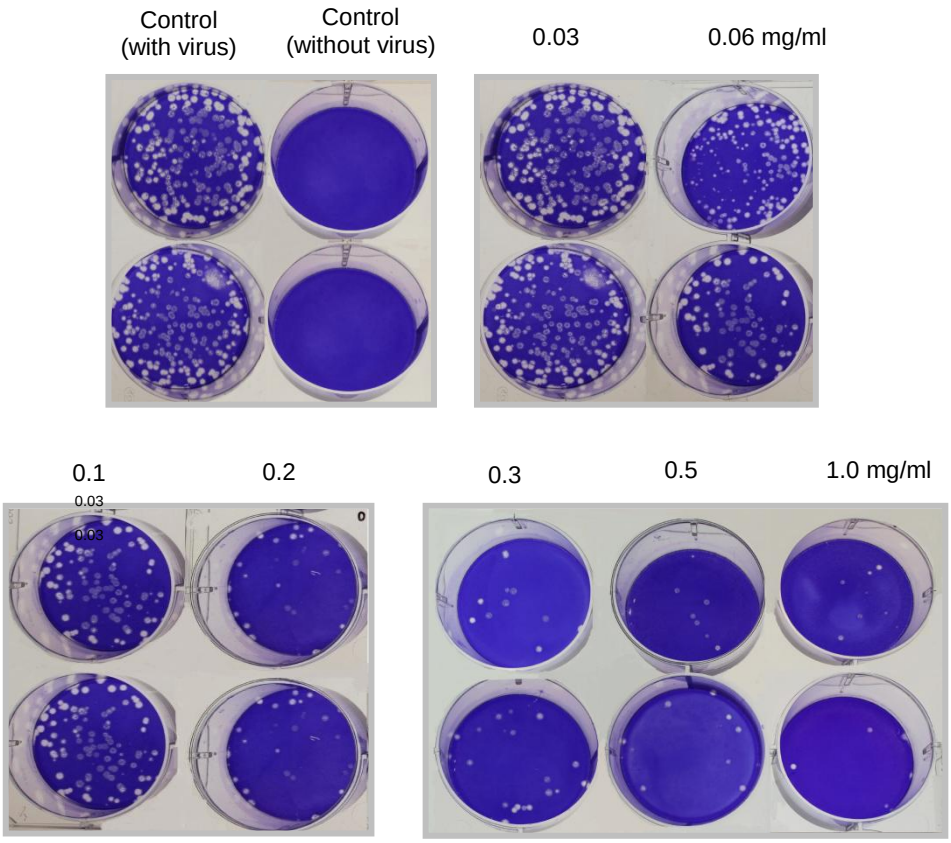


Figure 4.2 Plate representation of plaque inhibition assay on MDCK cells after infection with A/Udorn/72 H3N2 virus. Monolayers were treated with Vc4CBM protein at different concentrations ranging from 0.03 mg/ml to 1.0 mg/ml.

Table 4.1 Summary of EC_{50} values of influenza A viruses treated with different multivalent CBM40 proteins.

Virus strains	Multivalent proteins (μM)*				
	Vc3CBM	Vc4CBM	Vc-CBMTD (WT)	Vc-CBMTD (Mutant)	Sp-CBMTD
A/Udorn/72 H3N2	0.94 ± 0.06	0.59 ± 1.71	1.4 ± 0.86	2.1 ± 1.55	2.2 ± 2.45
A/PR8/34 H1N1	4.7 ± 1.63	1.80 ± 0.40	3.01 ± 1.04	5.01 ± 0.56	2.03 ± 2.41
A/WSN/33 H1N1	1.09 ± 0.29	1.06 ± 0.64	1.5 ± 0.53	2.0 ± 1.07	3.04 ± 2.31

*Note: Results were based on at least three replicate assays for each protein tested.

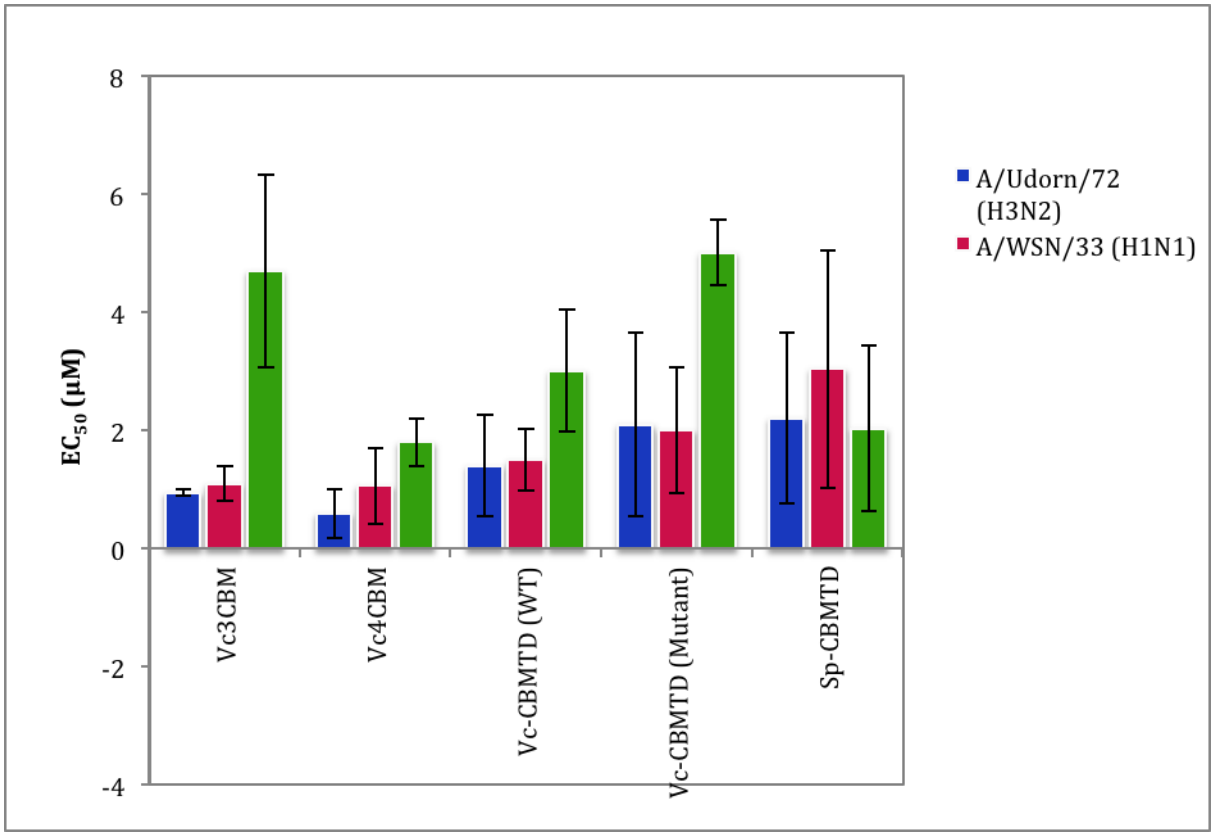


Figure 4.3 Comparison of EC_{50} values of different multivalent proteins constructs against influenza A strains. Details for each EC_{50} with standard error value can be referred to Table 4.1.

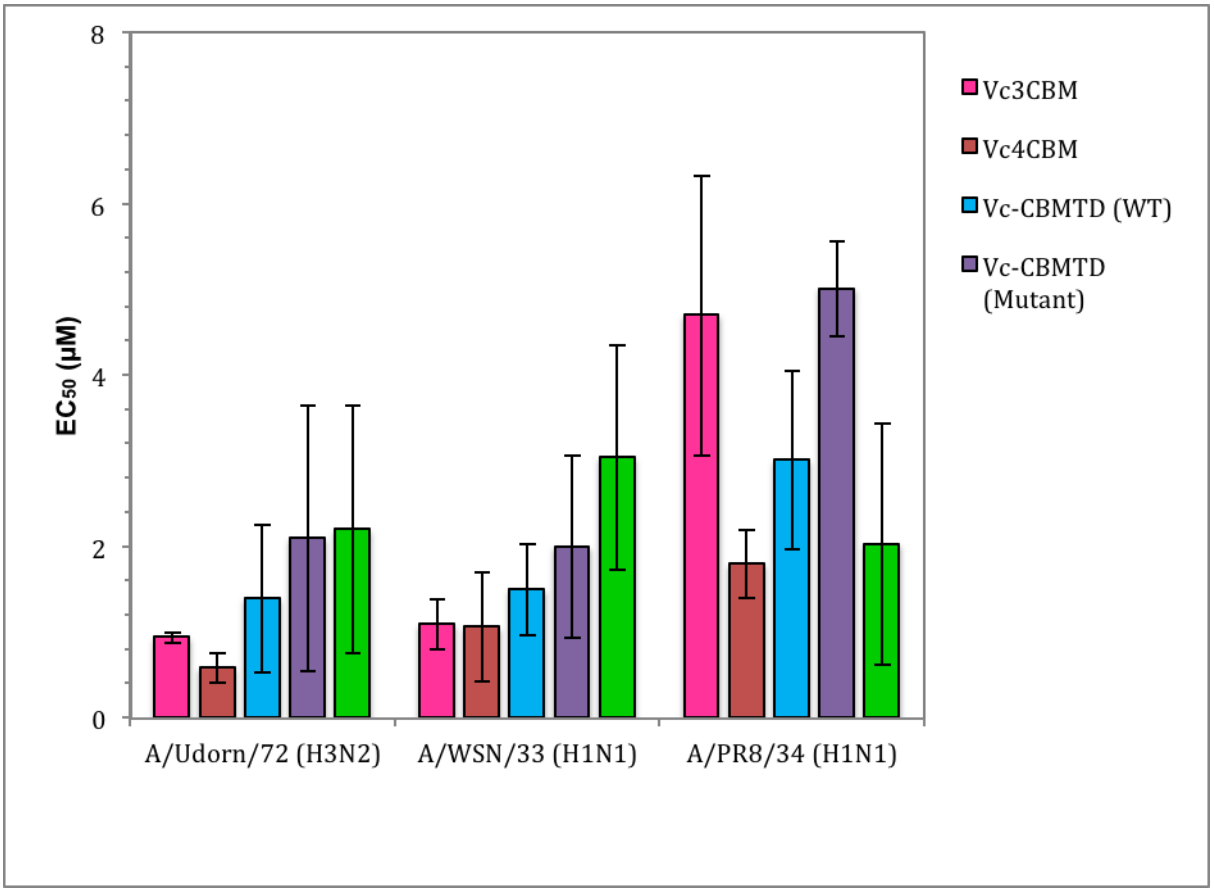


Figure 4.4 Comparison of different influenza A strains against EC_{50} of multivalent protein constructs. Details for each EC_{50} with standard error can be referred to Table 4.1.

Figure 4.3 and Figure 4.4 shows a summary of cell protection by multivalent proteins (Vc3CBM, Vc4CBM, Vc-CBMTD (WT), Vc-CBMTD (Mutant) and Sp-CBMTD) against influenza A strains (A/Udorn/72, A/PR8/34 and A/WSN/33) infections. As shown in the figure, most of the multivalent proteins are found to be least effective against A/PR8/34 H1N1 subtype as this virus is more virulent compared to the other influenza A strains. This subtype is currently endemic in both human and pig populations. H1N1 variant once was responsible for the Spanish flu pandemic which killed around 50-100 million people

worldwide during 1918-1919 (Patterson and Pyle, 1991). Meanwhile, levels of inhibition were greatest against A/Udorn/72 strains H3N2 subtypes. This strain was descended from H2N2 subtype which caused the Hong Kong flu pandemic. The evolution was due to antigenic shift which cause genes from multiple subtypes reassorted to form a new virus (Dumar, 2009).

Inhibition of virus replication were resulting from blocking viral attachment to cell surface sialic acids which will further avoid the re-attachment and infection of neighbouring cells. From the data presented, Vc3CBM, Vc4CBM and Vc-CBMTD (WT) are considered to be the best antiviral agent against all three influenza A strains. Another construct, Sp-CBMTD which is from *S. pneumoniae* CBM40 sialidase was also found to have a strong effect on cell protection with this being most effective against A/Udorn/72 H3N2 virus strain. Overall, these multivalent proteins tested proved to be good candidates for the development of influenza antivirals.

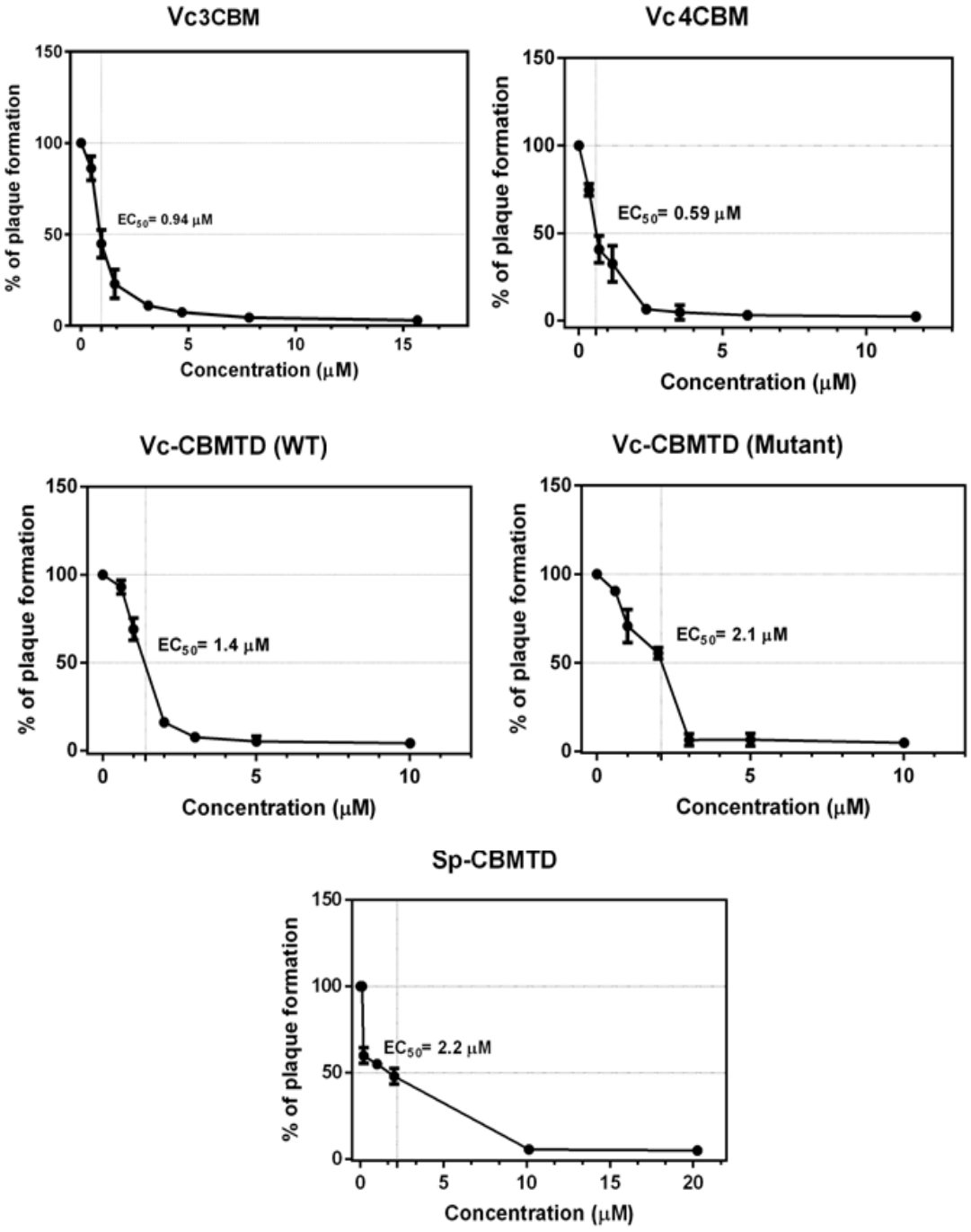


Figure 4.5 Detail of EC_{50} graphs on multivalent protein constructs infected with A/Udorn/72 H3N2.

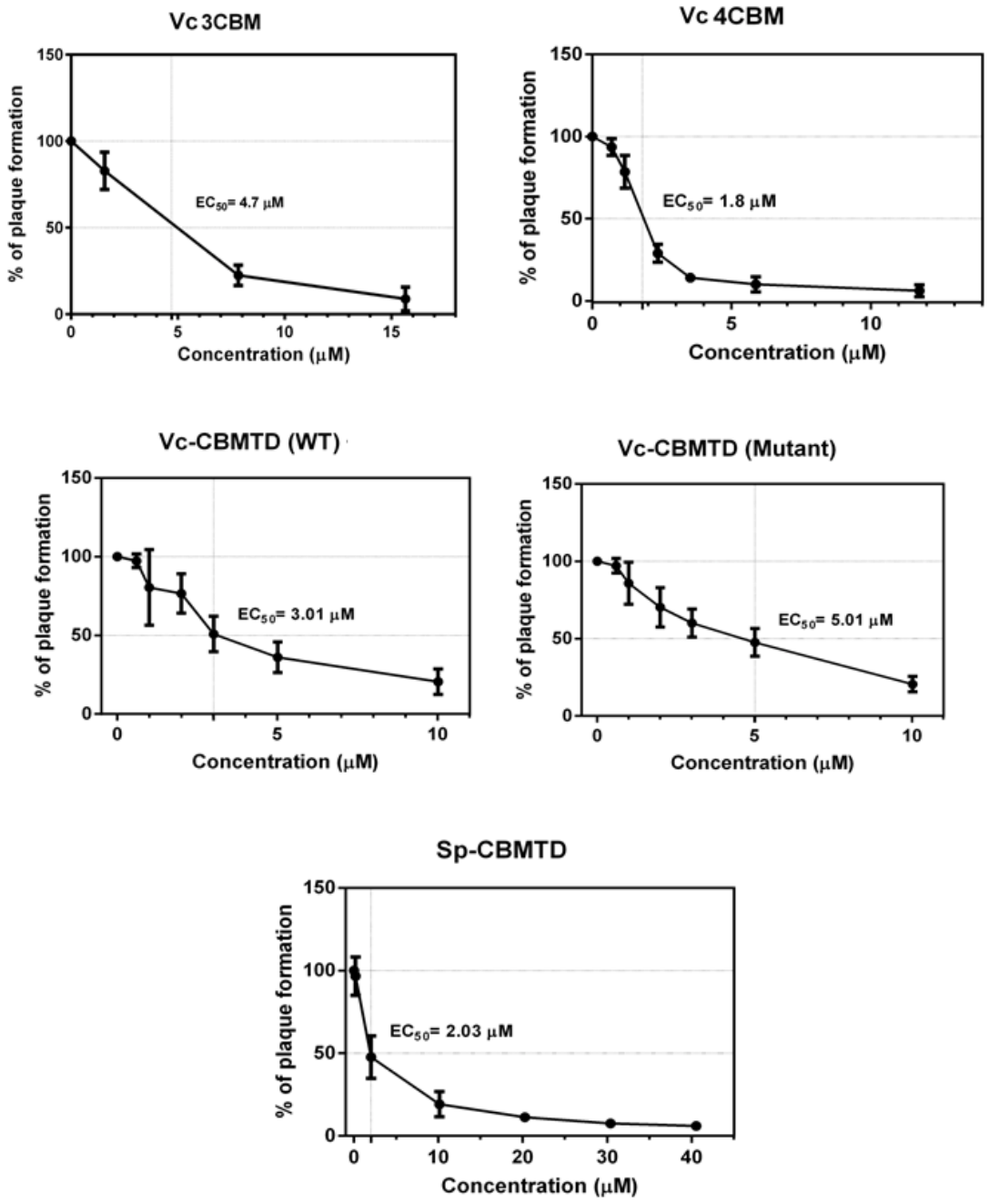


Figure 4.6 Detail of EC_{50} graphs of multivalent protein constructs infected with A/PR8/34 H1N1.

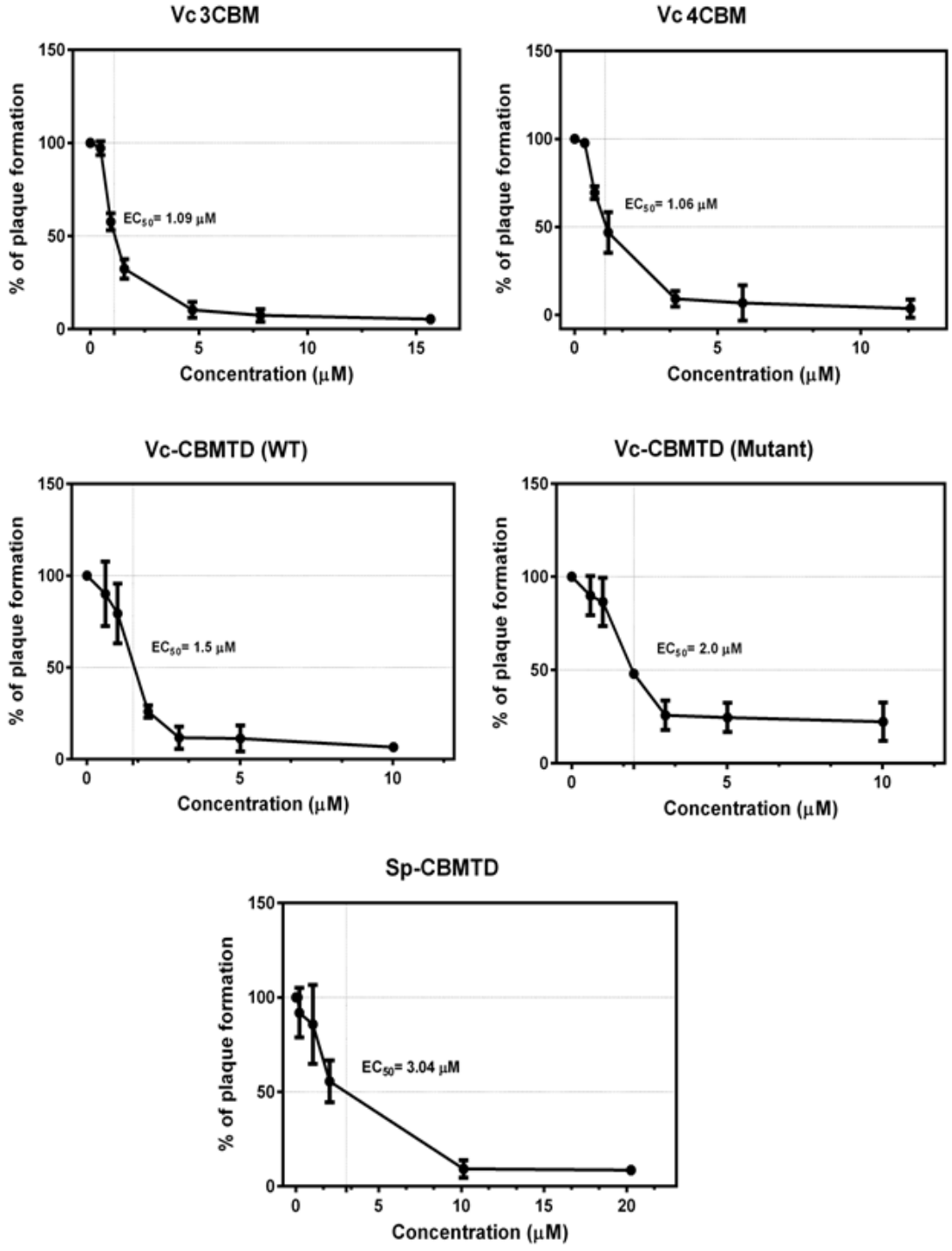


Figure 4.7 Detail of EC_{50} graphs of multivalent protein constructs infected with A/WSN/33 H1N1.

4.3 Effect of high concentration multivalent CBMs on cell monolayer

4.3.1 Evaluation of cytotoxic effects of Vc-CBMTD (WT) and Sp-CBMTD proteins

A cytotoxicity assay is used to predict toxicity level of drugs or protein on cultured mammalian cells. This investigation often required testing several different concentrations and drug exposure times using cells in culture. Therefore, we are using cell viability test to evaluate the effect of multivalent proteins on an established mammalian cell line. This protocol provides a simple and fast methodology to analyse viability of MDCK cells against different concentrations of the tested proteins and to test putative cytotoxicity effects associated with exposure to the substances.

MDCK cells were grown in 96-well plates at 7 x 10^4 cells/ well for overnight. Spent media was aspirated out and monolayer was washed with phosphate buffer saline (PBS). CBMs proteins were incubated at different concentrations on the monolayers for 6 hours and 24 hours at 37°C, 5% CO_2 in PBS supplemented with 10% fetal calf serum (FCS). Cell survival in samples was measured using 10 ml of PrestoBlue™ cell viability reagent (Invitrogen, USA) according to the manufacturer's protocol and incubated for 1 hour at 37 °C. Data values were measured at OD readings of 570 nm with reference wavelength at 620 nm using an ELISA reader. Cell viability was calculated using the following formula: 100 x [(OD_{570} of treated sample) / (OD_{570} of untreated sample)].

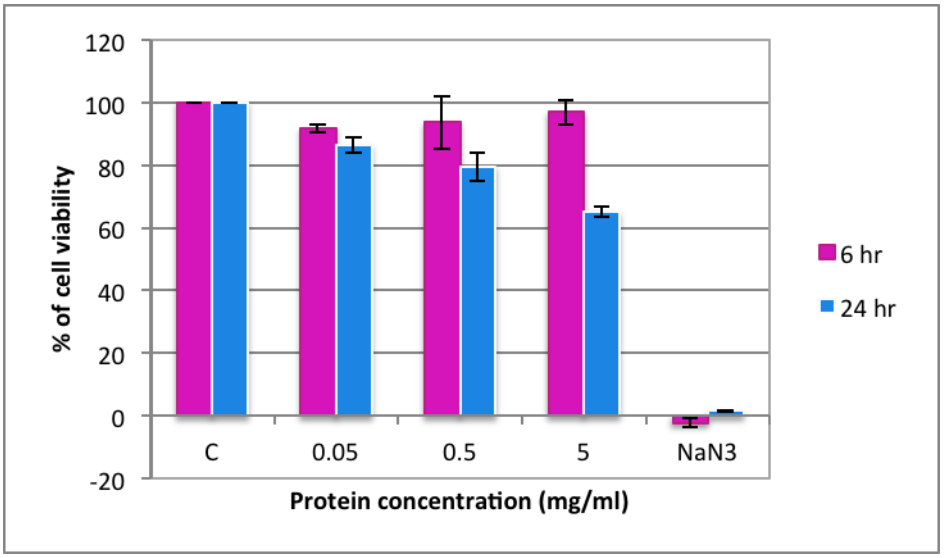


Figure 4.8 Viability test on Vc-CBMTD (WT). Two incubation times were tested which were (■) 6 hours and (■) 24 hours.

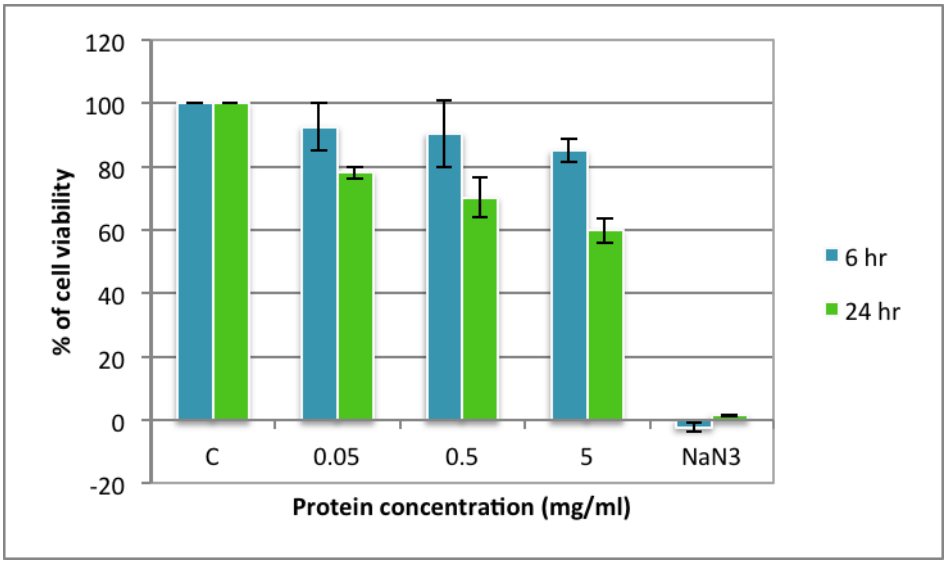


Figure 4.9 Viability test on Sp-CBMTD. Two incubation times were tested which were (■) 6 hours and (■) 24 hours.

The experimental data in Figure 4.8 and Figure 4.9 indicates dose- and time-dependent responses from MDCK cell monolayer treated with Vc-CBMTD (WT) and Sp-CBMTD. Other multivalent proteins (Vc3CBM, Vc4CBM, Vc-CBMTD (Mutant)) were less toxic to the cell after the same treatments were done. For the assay control, cell monolayer without any protein treatment was used besides using a negative control of cell monolayer treated with 20% sodium azide. Two post-treatment periods were used, which were 6 hours and 24 hours.

For both proteins tested (Vc-CBMTD (WT) and Sp-CBMTD), there are obvious differences in cell viability seen after 6 hours of post-treatment as compared to 24 hours. There are small reductions on cell viability percentage from 92.6 % to 85% for Sp-CBMTD and around 92% to 97% of the cell monolayers for Vc-CBMTD (WT) were still intact after 6 hours of treatment. After 24 hours of post treatment, the percentage of cell viability of Sp-CBMTD was reduced from 78.2% at 0.05 to 60 % at 5.0 mg/ml. While for Vc-CBMTD (WT), the monolayer showed reduction from 86.3 % to 65.2% of cell viability. Meanwhile, the same proteins gave about 24.4% increase in cell death for Vc-CBMTD (WT) and about 23.3 % of cell death for Sp-CBMTD at the highest protein concentration tested after 24 hours. When the protein concentration decreased to 0.5 mg/ml, cell death figures reduced to 7.8% for Vc-CBMTD (WT) and 10.2% for Sp-CBMTD, which show less toxic effect at the lower concentration tested. At 0.05 mg/ml, the cell death was seen to increase after 24 hours treatment as compared to 6 hours data at about 6% and 15.5% increment for Vc-CBMTD (WT) and Sp-CBMTD protein respectively.

4.4 Inhibition study of viral replication protein

4.4.1 Viral inhibition assay

This assay was done in order to measure the rate of inhibition of viral replication and viral protein synthesis by different CBMs constructs. The results can be used as comparison to data obtained from plaque inhibition assays. The assay was initiated with replicates (n=3) of MDCK monolayers in 96-well plates were treated with various dilutions of multivalent proteins (1.0, 0.5 and 0.1 mg/ml) in serum-free DMEM for 2 hours at 37°C, 5% CO_2. Both treated and untreated control cells (only virus in plain DMEM) were infected with virus dilution of MOI 0.01. After one hour, the monolayers were washed twice with PBS prior to being incubated with serum-free DMEM supplemented with 2.5 µg/ml N-acetylated Trypsin (NAT, Sigma Alderich) for 8 or 16 hours, 37°C, 5% CO_2.

The monolayers were then fixed with 4% formaldehyde in PBS before adding a permeabilisation solution (0.5 % Triton-X, 20 mM Glycine in PBS) with plates gently shaken for 10-15 minutes (ie if using a non-H3N2 virus). At this point, cell monolayers were blocked with 1% BSA + 0.02 % sodium azide in PBS (Blocking buffer) for one hour on the rocker. Cells were washed twice with PBS before adding 50 µl of X31 as primary antibody (generously given by Prof Richard Elliot's group members) at 1:500 dilution in blocking buffer for 1 hour at 37°C. Excess primary antibody was removed with three washes of PBS. All wells were then incubated with 50 µl of 1:1000 dilution of anti-IgG HRP anti-donkey (Santa-Cruz, USA) diluted in blocking buffer for one hour on a rocker. Excess secondary antibody was removed by washing three times with PBS. Plates were developed by incubating each well with 50 µl of TMB substrate for 30 minutes and stirred slowly. For stop reaction, 50 µl of 1 M H_2SO_4 was added to each well and absorbance was measured at 450 nm (reference

wavelength 620 nm) using spectrometer. Wells containing uninfected cells were used as the background control for the assay.

Having established that a variety of proteins were active against a few influenza A subtypes, we next examined their inhibition potential through replication of viral proteins. We analysed the activity against the same influenza A virus strains using an antibody, X31. This antibody recognized viral proteins (M1, NP and HA) replication after treatment with multivalent peptides. Figure 4.10 showed percentage of viral protein replications after treatment with 1 mg/ml of multivalent proteins for 1 hour. From the data, Vc3CBM and Vc4CBM gave the best inhibition profile for all influenza A strains tested (A/Udorn/72, A/WSN/33 and A/PR8/34). Vc3CBM against A/Udorn/72 H3N2 gave the highest percentage of inhibition of 25.1% as compared to 18% for A/PR8/34 H1N1 and 15% for A/WSN/33 H1N1. Moreover, effect of additional CBM40 domain in Vc4CBM construct, revealed more potent inhibition activity against all the strains with 29.9% inhibition for A/Udorn/72 H3N2, 24.4% inhibition for A/PR8/34 H1N1 and 22.4% for A/WSN/33 H1N1. For Vc-CBMTD (WT), A/Udorn/72 H3N2 showed the highest value in viral replication of 77.6 % followed by A/PR8/34 H1N1 of 92.9% and 85.9% for A/WSN/33 H1N1. However, for Vc-CBMTD (Mutant) and Sp-CBMTD their effect in inhibiting viral replication were acceptable with inhibition percentage of 18.2% (A/Udorn/72 H3N2), 15.1% (A/PR8/34 H1N1) and 3.2% (A/WSN/33 H1N1) for Vc-CBMTD (Mutant) and 15.7% (A/Udorn/72 H3N2), 16.3% (A/PR8/34 H1N1) and 10% (A/WSN/33 H1N1) for Sp-CBMTD respectively. This data were further confirmed with the EC_{50} data from cell protection assay as showed in Table 4.1, with the lowest EC_{50} contributed by Vc4CBM and Vc3CBM at 0.59 µM and 0.94 µM for A/Udorn/72 H3N2 followed by 1.06 µM and 1.09 µM for A/WSN/33 respectively.

Figure 4.11 and Figure 4.12 display a summary of all CBM40 proteins tested, demonstrating two proteins (Vc3CBM and Vc4CBM), being the most effective in preventing influenza A virus replication. A big reduction in viral replication of A/Udorn/72 H3N2 was observed, from

97.5% at 0.1 mg/ml to 74.9 % at 1.0 mg/ml after treatment with Vc3CBM. Whereas for other viruses (A/WSN/33 H1N1 and A/PR8/34 H1N1), small differences in inhibition effect were observed with 94.7 % to 82 % and 96.9 % to 85 % respectively at 0.1 mg/ml to 1.0 mg/ml of protein concentration. Figure 4.12 showed that Vc4CBM is among the best multivalent protein to fight against influenza A infection especially with A/Udorn/72 H3N2 virus strain. The data percentage reduced from 90.7 % to 70.1 % from 0.1 to 1.0 mg/ml, followed by A/WSN/33 H1N1 (92.8 % to 77.8 %) and A/PR8/34 H1N1 (94.2 % to 75.6 %) respectively.

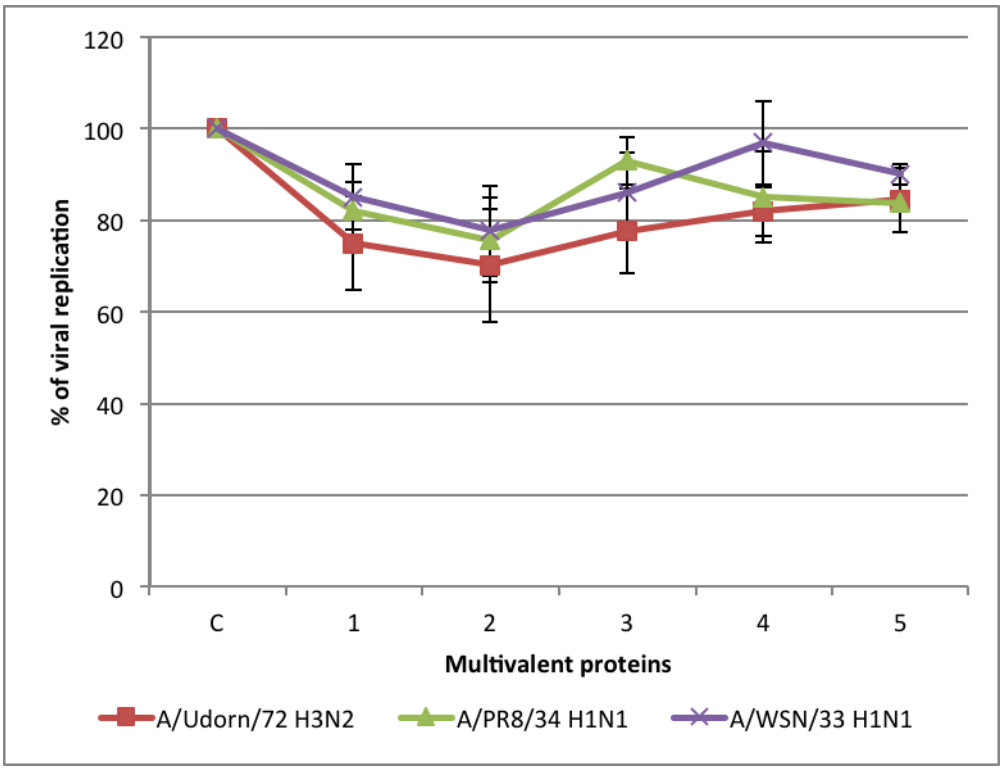


Figure 4.10 Percentage of viral replication proteins after treatment with 1 mg/ml of different multivalent proteins. **(C)** Control; **(1)** Vc3CBM; **(2)** Vc4CBM; **(3)** Vc-CBMTD (WT); **(4)** Vc-CBMTD (Mutant); **(5)** Sp-CBMTD.

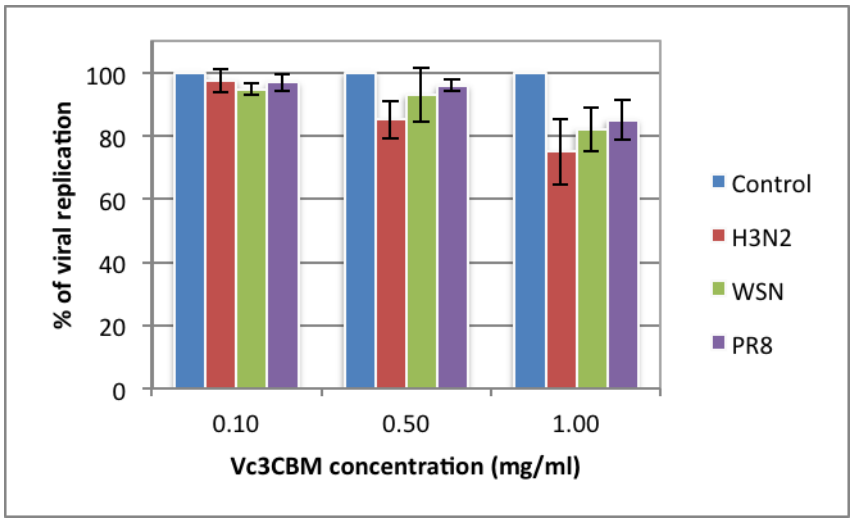


Figure 4.11 Effect of anti-influenza activity of Vc3CBM at different concentrations tested on MDCK cells. Percentage of viral replication is based on the replication of viral proteins (M1, NP and HA). Control represent virus in plain DMEM.

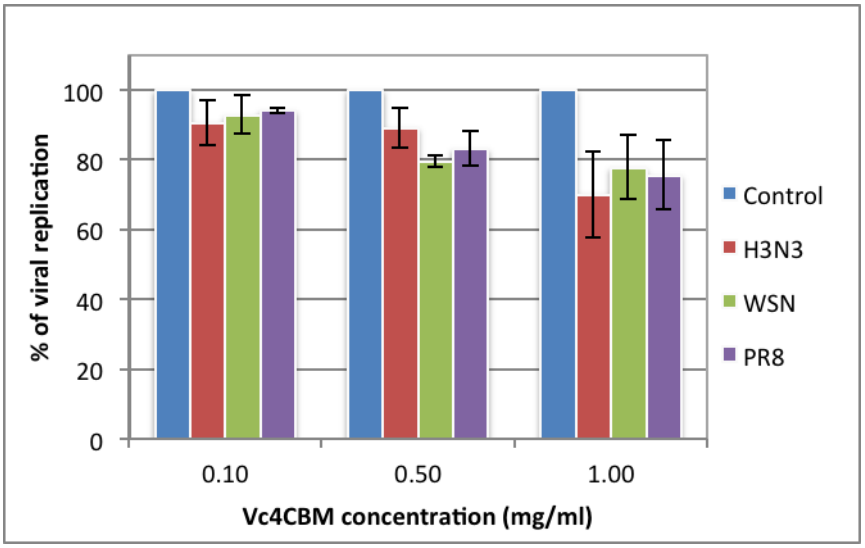


Figure 4.12 Effect of anti-influenza activity of Vc4CBM at different concentrations tested on MDCK cells. Percentage of viral replication is based on the replication of viral proteins (M1, NP and HA). Control represent virus in plain DMEM.

4.5 Discussion

Levels of inhibition were greatest against A/Udorn/72 H3N2 virus, but most of the multivalent proteins (Vc3CBM, Vc4CBM, Vc-CBMTD (WT), Vc-CBMTD (Mutant) and Sp-CBMTD) tested were also effective against A/WSN/33 H1N1 and A/PR8/34 H1N1 subtypes. Peptide inhibition of virus replication resulted from masking protein interactions with sialic acid receptor on the cell surface, which block viral attachment to cells thus inhibit the infections and re-attachment of virus to neighbouring cells.

The initial stages of cell infection by influenza viruses involve the haemagglutinin (HA) binding to sialylglycoconjugates followed by endocytosis, and then protease cleavage of HA in order to provide fusion of the endosomal and viral membranes to allow delivery of the viral contents into the cytoplasm. Therefore, the mechanisms whereby our proteins mediate their antiviral activity are through blocking the binding of virus HA to sialic acid receptors, acting as an entry blocker. The ability of these multivalent CBM40 constructs to bind to cell surface sialic acids has been studied by (Connaris et al., 2009) who demonstrated through FACS that a GFP-fused Vc3CBM, when incubated with human leukocytes can bind to cell surfaces. Result from the flow cytometry revealed that the probe (Vc3CBM-GFP) bound to sialic acids on the surfaces of granulocytes, monocytes and lymphocytes at concentration lower than 20 µg/ml, thus proving its high affinity towards its receptor.

This finding is interesting as it can have an important role in the prevention of influenza infection. Influenza A viruses constantly evolve by mechanisms known as antigenic drift and shift (Webster et al., 1992). The importance to predict the emergence of new virus strains are very important for the development of new antiviral drugs and these cannot be overestimated. In particular, treatments and prophylactics against a spectrum of influenza strains, which are not subject to drug resistance, are much needed. History has shown

pandemic influenza viruses have emerged three times which were in 1918 (Spanish Influenza, H1N1), in 1957 (Asian influenza, H2N2) and 1968 (Hong Kong influenza, H3N2) (Cox and Subbarao, 2005; Webby and Webster, 2003). Moreover, from 1997-2004 there was reported a circulation of highly pathogenic avian H5N1 viruses in Asia, which have caused a small number of human deaths (Claas et al., 1998; Peiris et al., 2004; Subbarao et al., 1998).

Results of *in vitro* studies here have demonstrated the ability of multivalent proteins, especially Vc3CBM and Vc4CBM, to significantly inhibit the replication of influenza A viruses, prevent or significantly reduce its infections. By targeting the host cells rather than the virus, CBM constructs demonstrated distinct anti-influenza properties from the virus targeting sialic acid. These findings were further supported by the lowest EC_{50} values from both proteins, 0.59 µM for Vc4CBM and 0.94 µM for Vc3CBM against A/Udorn/72 H3N2. For other viruses, A/WSN/33 and A/PR8/34, Vc4CBM was also found to be the best antiviral candidate among other multivalent proteins tested, with the EC_{50} values of 1.06 µM and 1.8 µM respectively. Other multivalent proteins are also acceptable in inhibiting influenza A infection to certain extent. The sharp increase in antiviral activity of Vc3CBM to Vc4CBM (EC_{50} 4.7 µM to 1.8 µM) against A/PR8/34 H1N1 suggests that the additional CBM40 domain in the protein construct help to improve its efficiency by providing an additional binding site to cell surface sialic acids. However, there are issues of binding energetics, affinity and conformational steric hindrances that need to be considered if working with multivalent proteins.

Increment in EC_{50} value between Vc-CBMTD (WT) and Vc-CBMTD (Mutant) suggest that there may be cooperative binding between CBM40 domain and trimerization domain (TD) in the constructs to sialic acids that account for differing EC_{50} values among the viruses tested in this study. Also, the initial binding between the protein and host sialic acid could lead to a conformational changes thus decreasing viral affinity towards sialic acids on the cell surface.

Another possibility is that, proteins interact at or near the receptor binding pocket, in this case multivalent proteins may block or cause steric hindrance during receptor docking.

Cytotoxicity data is very important as a predictor of acute systemic toxicity and its application in the *in vitro* study provides useful insight for CBM40s as potential drugs. Through this test, we are able to distinguish that at high concentration, 5 mg/ml can cause morphological changes to the cell monolayers due to the long term incubation with the protein alone. This study was also demonstrated that those proteins did not induce significant cytotoxicity in MDCK cells with the minimum concentration of 2 mg/ml. These has demonstrated that multivalent proteins tested are able to inhibit influenza virus attachment leading to decreased in viral replication as shown in viral inhibition assay.

In future, these multivalent CBM40 constructs could be developed as a successful inhibitor against variety of influenza virus strains. This drug is not a substitute for vaccine, but can be used during the period of vaccine development when novel strains have been identified, when resistance to commercially available antivirals occur, or as an alternative option, as vaccines are only effective with certain populations of society.

Chapter 5

In vitro study of parainfluenza virus infection

5.1 Overview

hPIVs are dangerous respiratory pathogens with the ability to spread infection past ethnic, socioeconomic, gender, age and geographic boundaries. They are able to cause upper respiratory infections in infants, children and adults while lower respiratory tract infections in elderly people and immunocompromised patients with chronic diseases (heart, lung disease and asthma). These viruses can cause annual epidemics and since re-infection occurs throughout life, there is a need for the development of novel and effective antiviral agents.

Antiviral therapy for parainfluenza viruses could be an optional treatment due to the complexities involved in vaccination. This could be a principal weapon against those diseases. hPIVs enter target cells through binding to a host cell receptor molecule and fusing their viral envelope with the cell membrane to gain entry to the host cytoplasm. Binding, fusion and entry are the main stages of the virus life cycle at which could be interfered with to prevent disease.

The surface of most mammalian cells is covered with a complex array of glycoconjugates, which are mostly attached to proteins and lipids. Moreover, the location of sialic acid at the termini of various carbohydrate complexes is often exploited by microbial pathogens to bind/ adhere and/or gain entry to the host cell (Ilver et al., 2003; Varki, 2007). Many of these

pathogens have also acquired sialidases that recognize sialic acids that aid in their pathogenesis.

Our group has discovered that by engineering multivalent CBM40s with high affinity towards its substrate, this will provide a useful method for the treatment of diseases caused by respiratory pathogens that target sialic acid. The main function of CBMs is to target or direct glycosyl hydrolase enzymes (such as sialidases or neuraminidases) to their substrates for efficient hydrolysis. Due to its affinity towards cell surface carbohydrates (e.g. sialic acid, galactose, blood group antigen, N-acetylglucosamine), this high affinity CBM may be used to block the binding of pathogens that recognize those carbohydrates. This action may prevent pathogens from entering or infecting the cell.

Here, we provide data to support the above hypothesis using a rapid and reliable assay (plaque inhibition assay) for determining the susceptibility of parainfluenza viruses to the multivalent CBMs constructs and to compare anti-parainfluenza (anti-hPIVs) activity of these compounds against a few virus isolates. For data comparison, a study on inhibition of viral replication protein was also done in conjunction with the plaque inhibition assays. Lastly, a cytotoxicity assay was performed *in vitro*, to evaluate the toxicity level of the tested CBMs in mammalian cells. The uses of these *in vitro* assays are very important in order to extrapolate the *in vivo* effect in humans.

5.2 Effect of multivalent CBMs in inhibiting parainfluenza virus Infections

5.2.1 Protocol of plaque inhibition assay

This assay was done in order to evaluate efficiency of multivalent proteins against viral infections. Confluent monolayers of African green monkey kidney epithelial (Vero) cells were grown in a Dulbecco's modified Eagle's medium (DMEM) containing 10% of Fetal Calf Serum (FCS) in 6-well tissue culture plates with approximately 4 x 10^5 cells per well. On the following day, different concentrations of proteins (Vc3CBM, Vc4CBM, Vc-CBMTD (WT), Vc-CBMTD (Mutant) and Sp-CBMTD) were prepared in serum-free DMEM and kept on ice. Each well of cells were inoculated with 1 ml of tested protein and incubated for 2 hours and 30 minutes at 37°C in a 5% CO_2 on a rocking platform.

Next, virus dilution was prepared with approximately 200 pfu/ml of virus (hPIV2, hPIV3 and hPIV5), which was previously determined by virus titration. The virus dilution was made in serum-free DMEM and kept on ice prior to use. Previous media were removed from the 6-wells plates and cells were washed once with PBS before adding 1 ml per well of each virus dilution. The virus was allowed to adsorb for 1 hour 30 minutes at 37°C, 5% CO_2 on a rocking platform. Later, 10 ml of overlay medium (DMEM/ 2% FCS and 0.6 % Methocel) was added to each well. The plates were incubated for 5 to 10 days or until plaques formed at 37°C, 5% CO_2 depending on virus used. The monolayers were fixed with 5% formaldehyde in PBS for at least 30 minutes before stained with 0.1% crystal violet. The plaques were counted by visual examination and percentage of plaque inhibition was calculated as relative to control plate (contain virus only).

5.3 Study of different pre-incubation periods on Vc3CBM and Vc4CBM against hPIV3

Blocking viral binding to host cell receptors is an interesting antiviral approach. It is well established both for influenza virus (IFV) and human parainfluenza virus (hPIV) which utilize cell surface sialic acid for binding and entry to initiate viral infection (Maisner et al., 1994; Schauer, 1985). By targeting the host cell rather than the virus, this approach has become an alternative route to combat virus infection and halts the spread of further infection.

For initial screening of inhibitory effects by multivalent CBMs, protein extracts were tested on African green monkey kidney cell line (Vero cells) by plaque assay. As shown in Figure 5.1, Vc3CBM and Vc4CBM proteins were shown to be highly effective in preventing viral attachment, as tested in Vero cells against hPIV3. Vero cells have been identified primarily to display α(2,3)-linked sialic acid on their cell surface (Govorkova et al., 1996) which corresponds to inner human respiratory epithelium, which expresses both forms of α(2,6)-, and α(2,3)-linked sialic acids (Hassid et al., 1999; Matrosovich et al., 2004a).

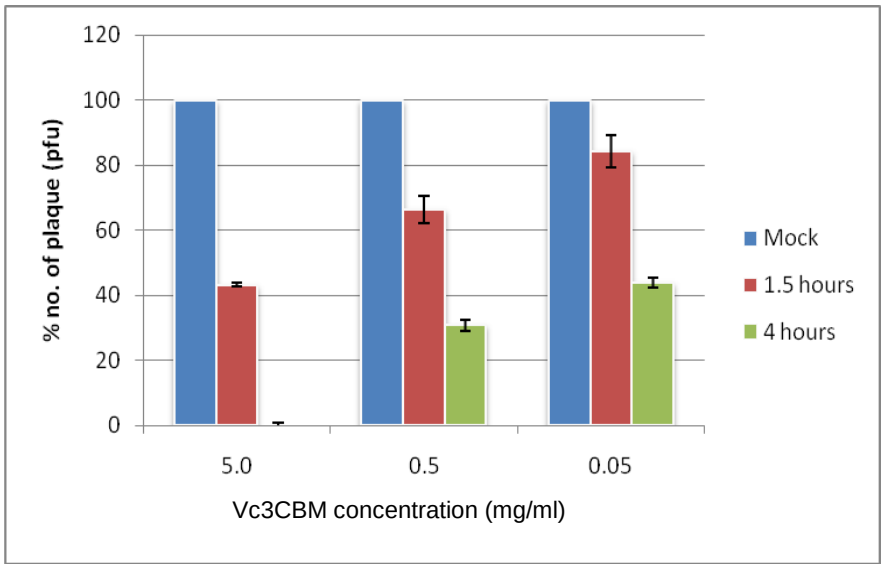


Figure 5.1 Percentage number of plaque compared to mock on Vero cell after hPIV3 infection. Cells were pre-incubated with Vc3CBM for 1.5 hours and 4 hours prior to virus infection.

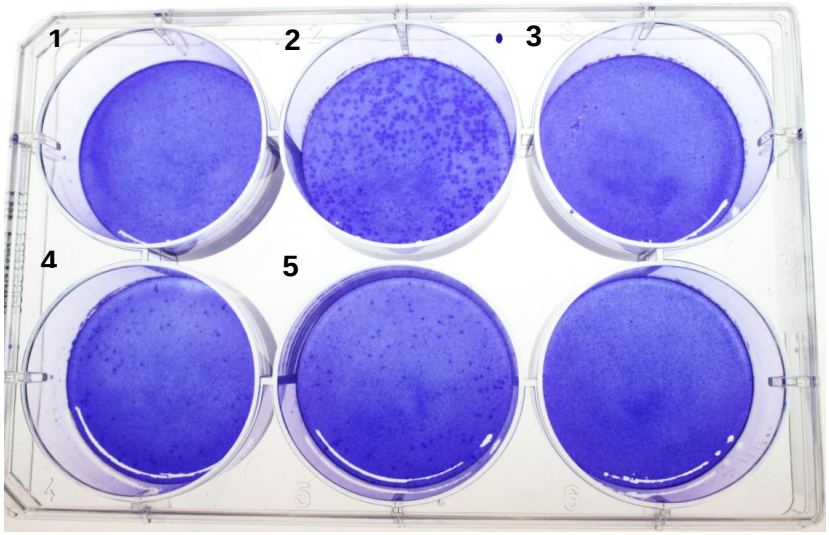


Figure 5.2 Plaque assay of Vc3CBM on Vero cells after 3 days of hPIV3 infections. Cells were pre-incubated with Vc3CBM for 4 hours prior to infection with virus. **(1)** Control (no hPIV3 infection); **(2)** Mock with hPIV3 alone ($1X10^{-6}$ dilution); **(3)** with 5 mg/ml of Vc3CBM; **(4)** with 0.5 mg/ml of Vc3CBM; **(5)** with 0.05 mg/ml of Vc3CBM.

To investigate viral inhibition effect of Vc3CBM and Vc4CBM, Vero cells were treated with different concentrations of each protein prior to hPIV3 infection. Briefly, Vero cells were mock inoculated (virus alone) or inoculated with CBMs at two different incubation times (1 hour 30 minutes, and 4 hours). The virus was then allowed to bind to the cells for 1 hour 30 minutes at 37°C, 5% CO2 to synchronize infection.

In the absence of CBMs, viral infections of cell occurred rapidly in just 3 days as seen in plaque assays (Figure 5.2). As shown in Figure 5.1, pre-treatment with Vc3CBM at 5 mg/ml was highly effective for cell protection from virus, producing 100% viral inhibition. While at 0.5 mg/ml and 0.05 mg/ml, viral inhibitory effects were seen at 69% and 56% respectively, compared to mock cells (Figure 5.1). As for Vc4CBM, 98% inhibitory effect was detected at 2 mg/ml (Figure 5.3 and Figure 5.4). Even at 0.2 mg/ml and 0.02 mg/ml, the percentage of viral inhibition was 72% and 61% respectively (Figure 5.3 and Figure 5.4). Prolongation of the pre-incubation time to 4 hours for both Vc3CBM and Vc4CBM completely protected the cells from hPIV3 infection. This must be due to longer time required for the protein to properly attach to the cell surface for effective inhibition.

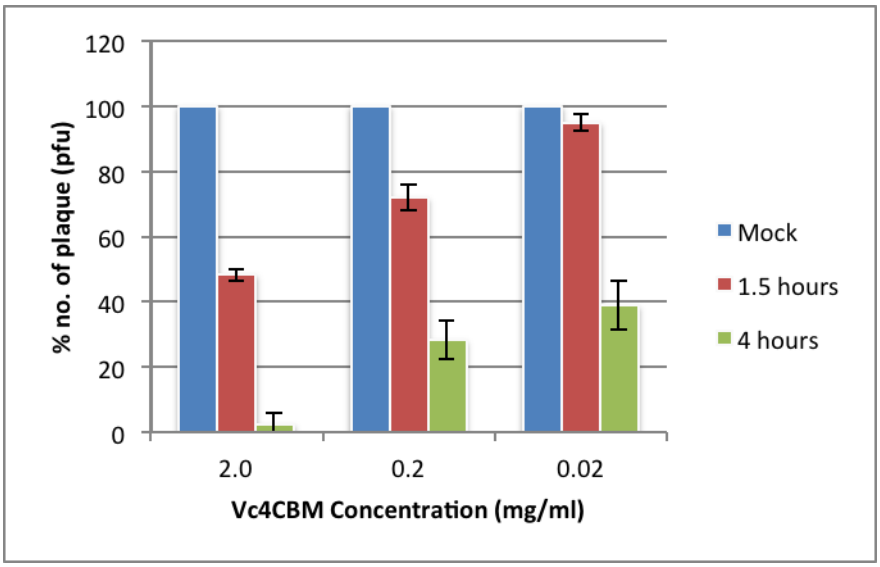


Figure 5.3 Percentage number of plaque compared to mock on Vero cell after hPIV3 infection. Cells were pre-incubated with Vc4CBM for 1.5 hours and 4 hours prior to virus infection.

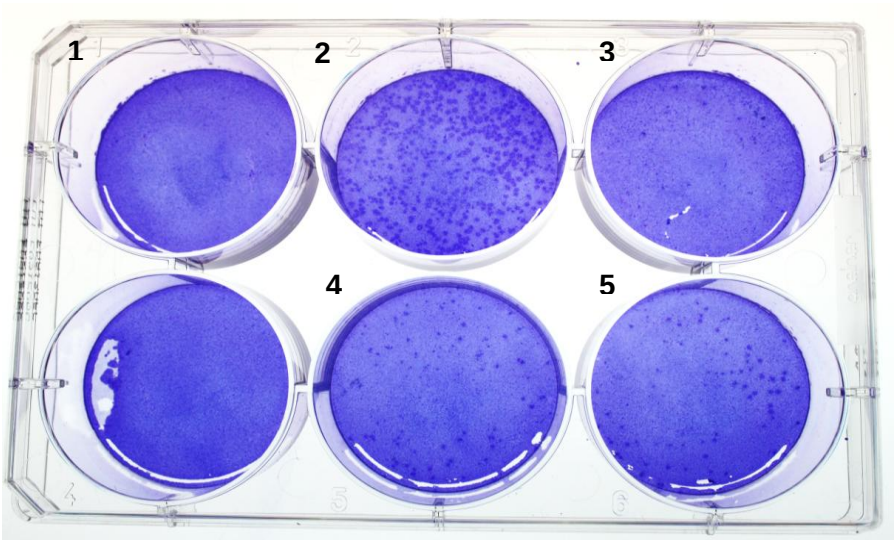


Figure 5.4 Plaque assay of Vc4CBM on Vero cells after 3 days of hPIV3 infections. Cells were pre-incubated with Vc4CBM for 4 hours prior to infection with virus. **(1)** Control (no hPIV3 infection); **(2)** Mock with hPIV3 alone ($1X10^{-6}$ dilution); **(3)** with 2.0 mg/ml of Vc4CBM; **(4)** with 0.2 mg/ml of Vc4CBM; **(5)** with 0.02 mg/ml of Vc4CBM.

From the cell binding study of GFP-Vc3CBM, the protein was shown to attach to the surface of Vero cells evenly at concentration of 0.05 mg/ml and 0.5 mg/ml after incubation with the GFP fused CBM for 1 hour 30 minutes. This has suggested that the CBM binds to the cell surface even at lower concentrations (Figure 5.5). But at this pre-incubation time, it did not sufficiently inhibit hPIV3 infection completely, as seen in plaque assays (Figure 5.2). This initial study determined that virus plaque reduction depended on the duration of protein incubation on the monolayers. It proved that with 4 hours pre-incubation time, for both proteins, number of plaques have reduced to ≥ 50% as compared to the data at 1 hour 30 minutes.

The strong inhibitory effect on Vero cells of both Vc3CBM and Vc4CBM observed throughout the growth phase raises the possibility that the antiviral effect is exerted not only on the initially infecting viruses, but also on subsequent steps in the infectious cycle as the experiments have been prolonged for 5 days after viral infections. It may be that the CBMs are still attached to sialic acid, preventing further viral attachment from newly emerged viral particles.

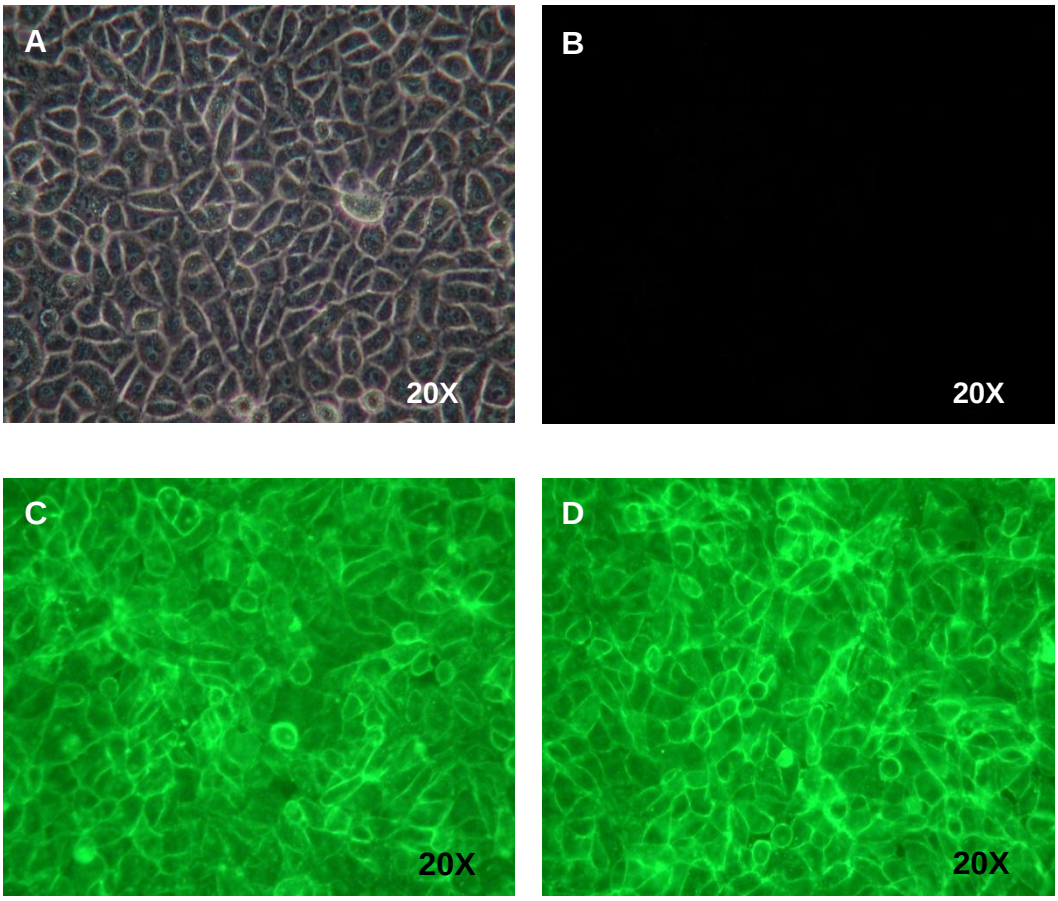


Figure 5.5 Vc3CBM-GFP attached to the surface of Vero cells after incubation for 1 hour 30 minutes. All pictures were taken at 20X magnification. **(A)** Vero cell monolayer; **(B)** Control without Vc3CBM-GFP; **(C)** with 0.5 mg/ml of Vc3CBM-GFP; **(D)** with 0.05 mg/ml of Vc3CBM-GFP.

5.4 Inhibitory effects of Vc3CBM and Vc4CBM on A549 cells by hPIV3

Since hPIV3 infects the epithelial cells of lung as the primary target, we used A549 cells which is human lung carcinoma as a model of airway epithelial (Bose et al., 2001; Gao et al., 2001). To test the antiviral activity of Vc3CBM and Vc4CBM, A549 monolayers were infected with hPIV3 (200 pfu) and both Vc3CBM and Vc4CBM were pre-incubated on the cell for 4 hours prior to viral infection. As mentioned before, both constructs are highly efficient in recognizing sialic acid on the cell surface as described by Connaris *et al*., (2009). Initially, for the plaque assay, three different concentrations of Vc4CBM were used (2.0 mg/ml, 0.2 mg/ml and 0.02 mg/ml). As seen in Figure 5.6, 2 mg/ml of Vc4CBM was found to inhibit 78% of hPIV infection, while at 0.2 mg/ml and 0.02 mg/ml the percentage were 70% and 46% respectively. The experiment was repeated again but with higher concentration of Vc4CBM in order to knock out the infection. Surprisingly, even at 5.0 mg/ml of Vc4CBM did not totally inhibit the infection as seen on Figure 5.7 compared with Vero cells.

This finding had been supported by Bose and Banerjee (2002), who found that A549 cell line interaction with hPIV3 only accounted for ~50% of the infection due to the present of another alternative cell surface moiety, heparan sulfate (HS). But in this case, the presence of Vc4CBM on the cell surface does give good protection against the infection but not sufficient to knock out the infection completely, possibly due to the presence of the alternative receptor. It was demonstrated that hPIV3 binds to HS and the virus interaction with the cell surface HS is important for its efficient entry into the cells (Bose and Banerjee, 2002). Cell surface HS is known to be present as proteoglycan and widely utilize by some viruses to gain entry into the host cells including respiratory syncytial virus (RSV) (Feldman *et al*., 2000) herpes simplex virus (HSV) (WuDunn and Spear, 1989), foot and mouth disease virus (Jackson *et al*., 1996), HIV-1 (Roderiquez *et al*., 1995) and adenovirus type 2 and 5 (Dechecchi *et al*., 2000). It may be possible that 50% inhibition would be sufficient to reduce

disease symptoms. Furthermore, combination therapy can possibly be used through simultaneous use of CBM and a specific inhibitor of HS to eliminate or inhibit viral infection on this cell line.

In another study where A549 cells were treated with neuraminidase, it was revealed that despite the absence of cell surface sialic acids, hPIV3 was still capable of infecting cells with reduced efficiency to 55% inhibition. This suggests that hPIV3 need both cell surface HS and sialic acid molecule for optimal and efficient cellular entry (Bose and Banerjee, 2002). Because of A549 cell line do not support efficient hPIV3 infection, we plan to use another alternative cell lines, which is Vero cell line, to compare the result.

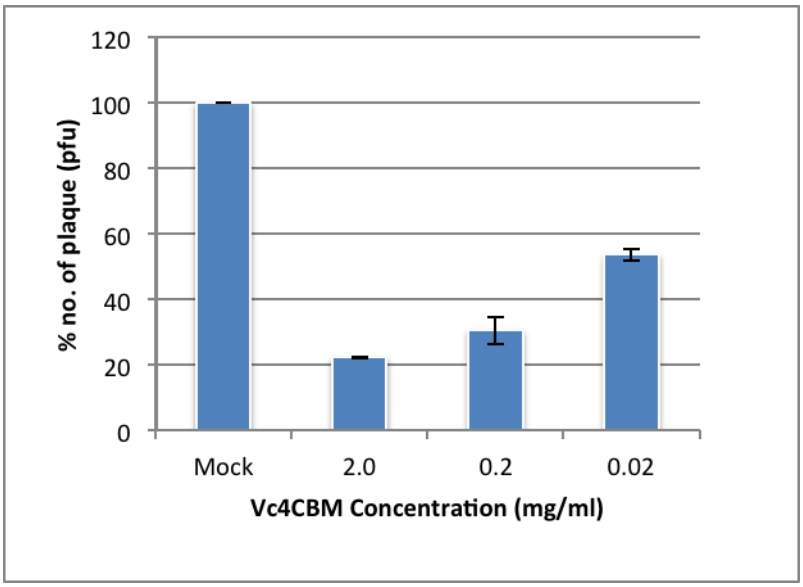


Figure 5.6 Percentage number of plaque compared to mock on A549 cell line after hPIV3 infection. Cells were pre-incubated with Vc4CBM for 4 hours prior to virus infection.

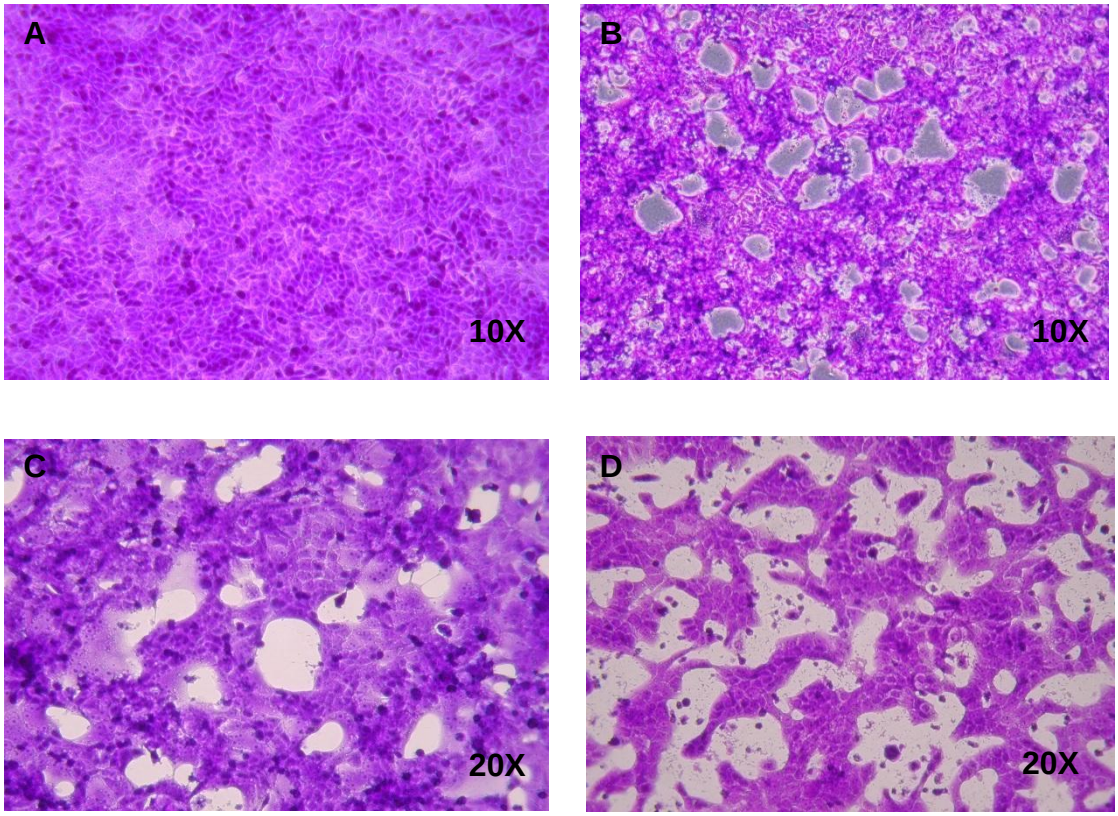


Figure 5.7 Effect of various concentrations of Vc4CBM on A549 cell line infected with hPIV3 after 5 days. The monolayers were stained with 0.05% crystal violet. **(A)** Control without hPIV3 infection. **(B)** 5.0 mg/ml of Vc4CBM; **(C)** 0.5 mg/ml of Vc4CBM; **(D)** 0.05 mg/ml of Vc4CBM.

5.5 Anti-hPIV activity

The relative antiviral activities of multivalent CBMs against contemporary strains of parainfluenza viruses were determined in African green monkey kidney epithelial (Vero) cell monolayers with a plaque inhibition assay. The concentration required to inhibit virus replication to 50% of the level of control (without the compound) (EC_{50}) were determined and calculated by regression analysis of the dose-response curves generated by data. This assay proved to be a reliable and rapid method in determining 50% inhibitory concentrations which correlated well with clinically achievable drug levels and clinical trials results.

For the screening of antiviral activity, each protein compounds were tested by plaque inhibition assay at a different range of concentrations. Firstly, confluent monolayer cultures of Vero cells were treated with multivalent CBM constructs (Vc3CBM, Vc4CBM, Vc-CBMTD (WT), Vc-CBMTD (Mutant) and Sp-CBMTD) for 2 hours and 30 minutes according to the protocol in Section 5.2.1. At this stage, plates must be gently shaken as to avoid monolayers from drying out and to allow even distribution of the tested protein on monolayers. Later, monolayers were infected with parainfluenza virus strains at 200 pfu (hPIV2, hPIV3 and hPIV5) for 1 hour and 30 minutes before adding an overlay medium as described previously (Section 5.2.1). Addition of 2% fetal calf serum was not needed in the overlay medium for hPIV3 and hPIV2 unlike hPIV5 to aid infection. Plates were incubated at 37°C, 5% CO_2 for approximately 3-7 days depending on virus used. The plaques formed were calculated by visual examination and compare with the control plate.

Data represented by a dose-response curve of viruses (hPIV2, hPIV3 and hPIV5) inhibited by multivalent proteins were shown in Figure 5.8, Figure 5.9 and Figure 5.10 respectively. These data represented as the mean of at least three independent experiments, as summarized in Table 1. From the EC_{50} value, all multivalent constructs were able to block

parainfluenza virus infections but at different levels of efficiency. The lowest EC_{50} was found at 1.17 µM with Vc4CBM after hPIV3 infection followed by hPIV2 treatment with EC_{50} of 1.94 µM. Vc4CBM gave the lowest EC_{50} value compared to Vc3CBM after infection with both hPIV2 and hPIV3 viruses. From the data, it appears that an increase in number of CBM40 modules could be important in effectively blocking viral binding to sialic acid on the cell surface. These results indicate that Vc4CBM is more potent than Vc3CBM and their activity is manifested in a dose-dependent manner.

In contrast, for hPIV5, the lowest EC_{50} was found with Vc3CBM at 2.43 µM instead of Vc4CBM which has EC_{50} value at 3.50 µM. There is clearly an increment in potency of CBM inhibition against hPIV2 and hPIV3 strains as the number of linked CBM40 modules increase from 3 to 4 of CBM40. Also, incubation of Vero cells with oligomers from Vc-CBMTD (WT) showed effective inhibition and gave relatively low EC_{50} values as compared to Vc-CBMTD (Mutant) and Sp-CBMTD. The lowest EC_{50} was given by Vc-CBMTD (WT) in protection against hPIV2 (0.2 µM) followed by hPIV5 (1.8 µM) and hPIV3 at 17.5 µM. Other constructs, such as the Vc-CBMTD (Mutant) effectively inhibited hPIV5 infection with EC_{50} at 1.78 µM followed by 2.31 µM for hPIV2 and 80.18 µM for hPIV3 infections. Meanwhile, Sp-CBMTD showed a weak effect in cell protection from all viruses tested with EC_{50} of 37.59 µM followed by 90 µM and 91.47 µM, which correspond to hPIV5, hPIV2 and hPIV3 infections, respectively.

Table 5.1 Summary of EC_{50} values (μM) of multivalent CBMs infected with different strains of hPIVs

Virus strains	Multivalent CBMs (μM)				
	Vc3CBM	**Vc4CBM**	**Vc-CBMTD (WT)**	**Vc-CBMTD (Mutant)**	**Sp-CBMTD**
hPIV2	2.77 ± 0.42	1.94 ± 0.19	0.2 ± 0.85	2.31 ± 1.51	90 ± 6.78
hPIV3	2.6 ± 0.37	1.17 ± 0.88	17.5 ± 3.16	80.18 ± 8.65	91.47 ± 10.01
hPIV5	2.43 ± 0.66	3.50 ± 1.86	1.8 ± 0.45	1.78 ± 0.74	37.59 ± 6.85

*Note: Results were based on at least three replicate assays for each protein tested.

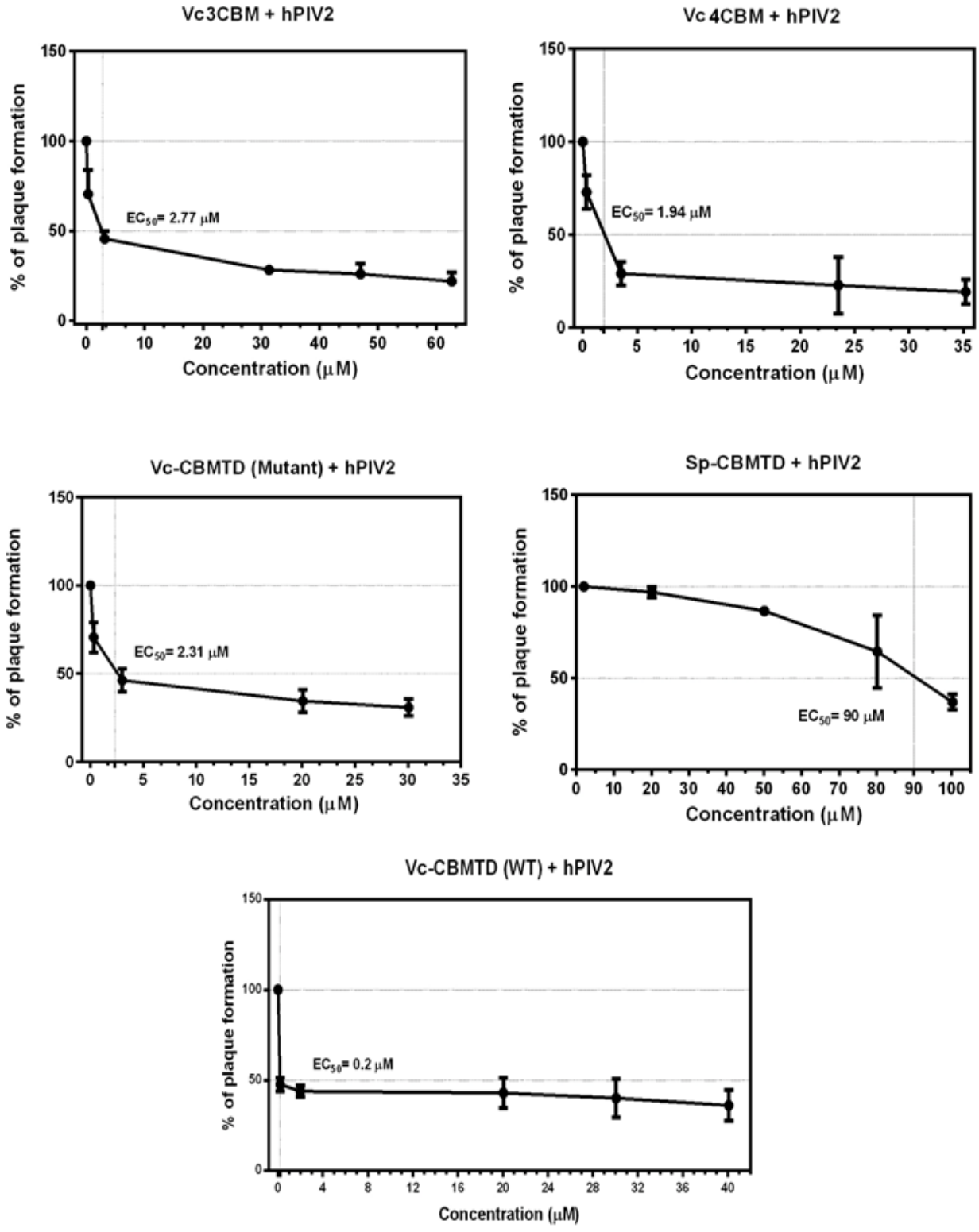


Figure 5.8 Detail of EC_{50} graphs on multivalent constructs infected with hPIV2.

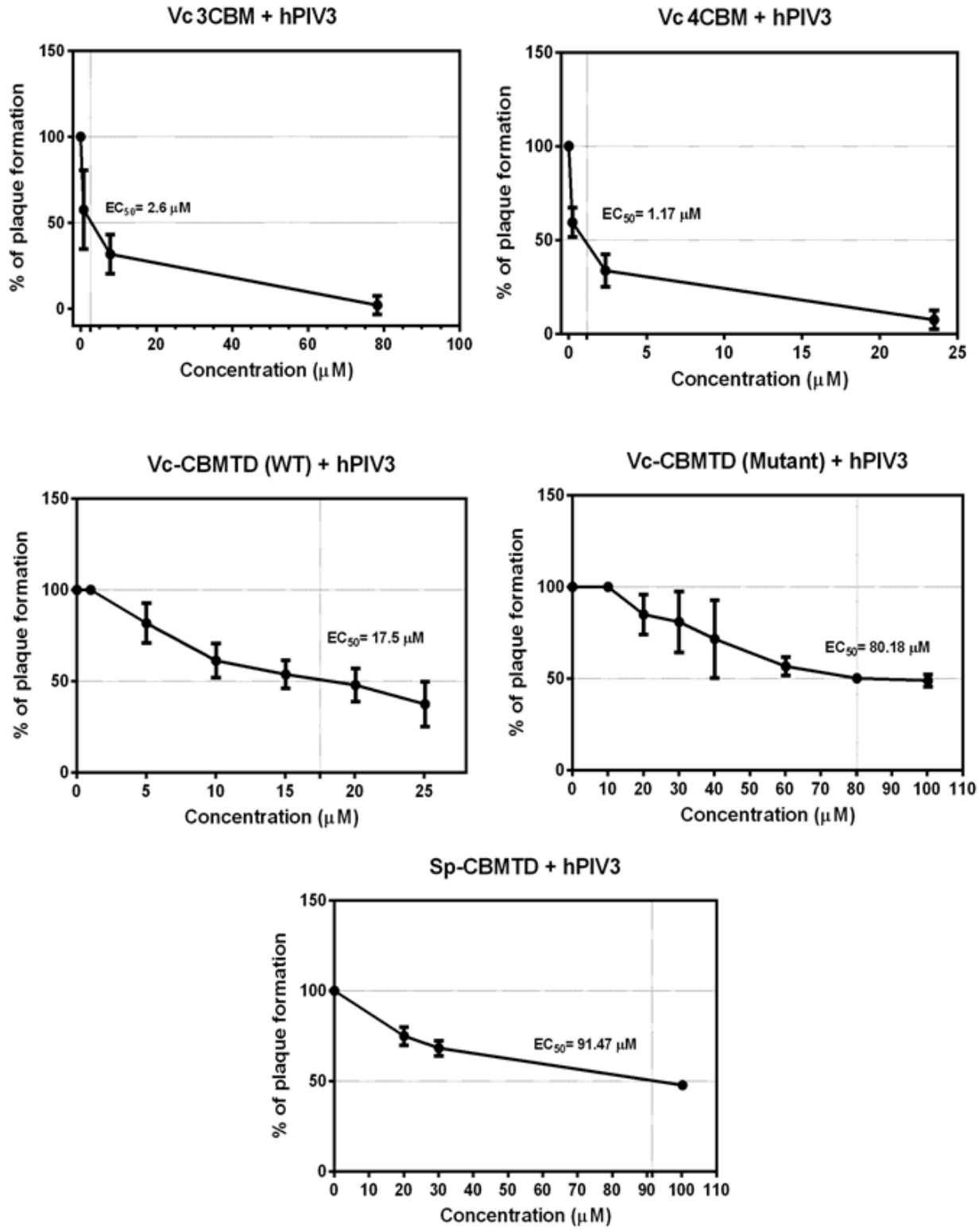


Figure 5.9 Detail of EC_{50} graphs on multivalent constructs infected with hPIV3.

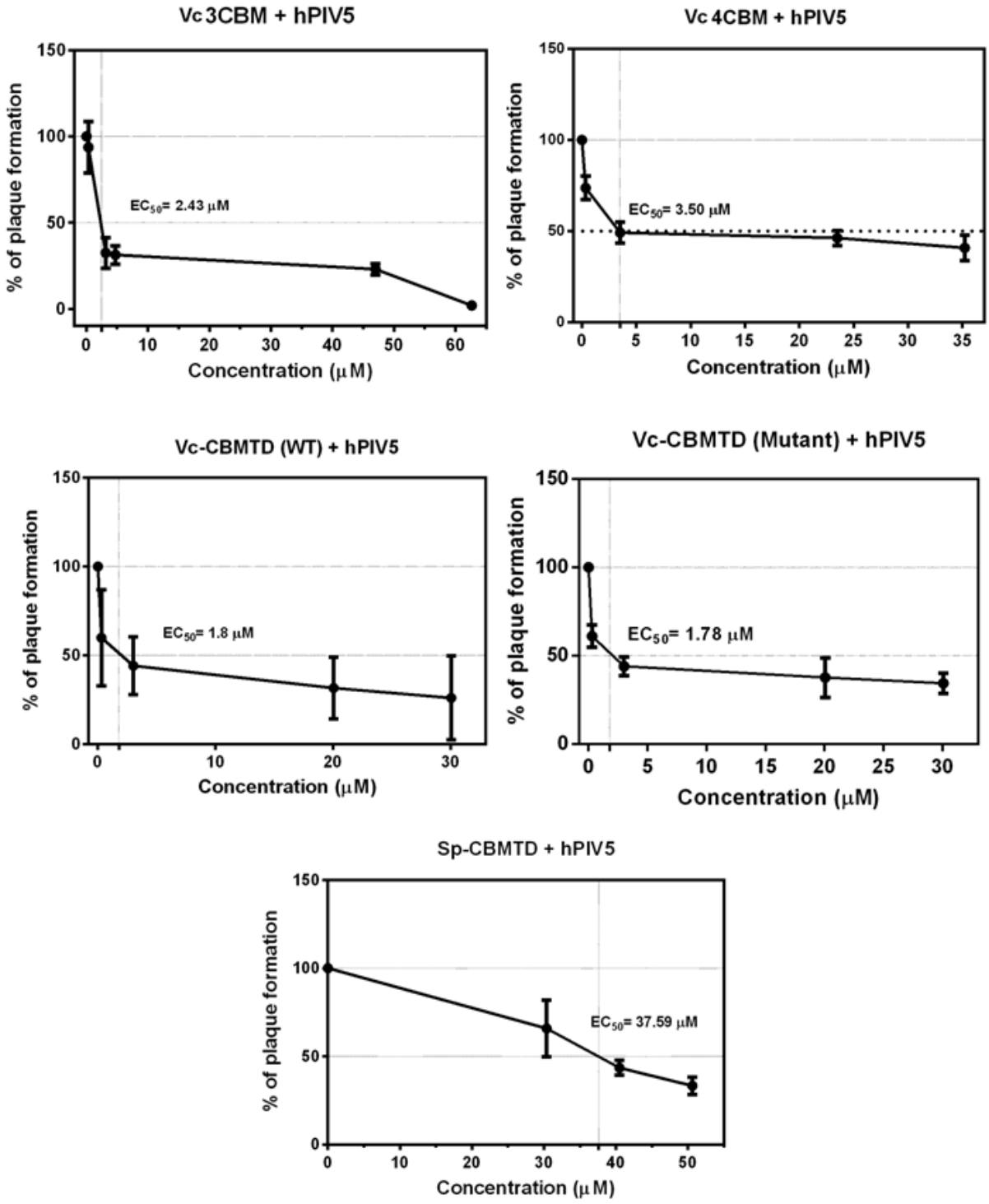


Figure 5.10 Detail of EC_{50} graphs on multivalent constructs infected with hPIV5.

The data presented in Figure 5.11 and Figure 5.12 clearly showed a summary of cell protection of all multivalent polypeptides (Vc3CBM, Vc4CBM, Vc-CBMTD (WT), Vc-CBMTD (Mutant) and Sp-CBMTD) against parainfluenza viruses (hPIV2, hPIV3 and hPIV5) infection. All constructs containing *V. cholerae* sialidase CBM40 domain (Vc3CBM, Vc4CBM, Vc-CBMTD (WT) and Vc-CBMTD (Mutant) showed to have a great effect in inhibiting viruses infection and this includes hPIV2, hPIV3 and hPIV5. Only a construct Sp-CBMTD, which is from *S. pneumoniae* sialidase CBM40, was found to have a little effect on cell protection. Vc3CBM and Vc4CBM are the best multivalent proteins which have the strongest effect in inhibiting parainfluenza virus infection and its spreading. These data demonstrate that the *V. cholerae* sialidase CBM40, whether in tandem-linked 3 or 4 domain repeats or as a self-assembled oligomer containing trimerization domain (PaTD), can effectively protect Vero cells from parainfluenza virus infections. Interestingly, a construct from *S. pneumoniae* sialidase CBM40 with PaTD domain demonstrated a weak protection against the virus infections.

Thus, by joining and conjugating CBM40 monomers together, it is possible to create large polymeric molecules, which exhibit greater affinity for a particular number of substrates through an avidity effect. The decrease in number of cells infected by the pathogen relative to the control assay has indicated that CBM polymers are potentially useful in the treatment of disease and condition caused by the pathogen. Thus, differential binding capabilities of the various CBM polymers can be exploited for the screening, identifying, detecting and/ or labelling carbohydrates.

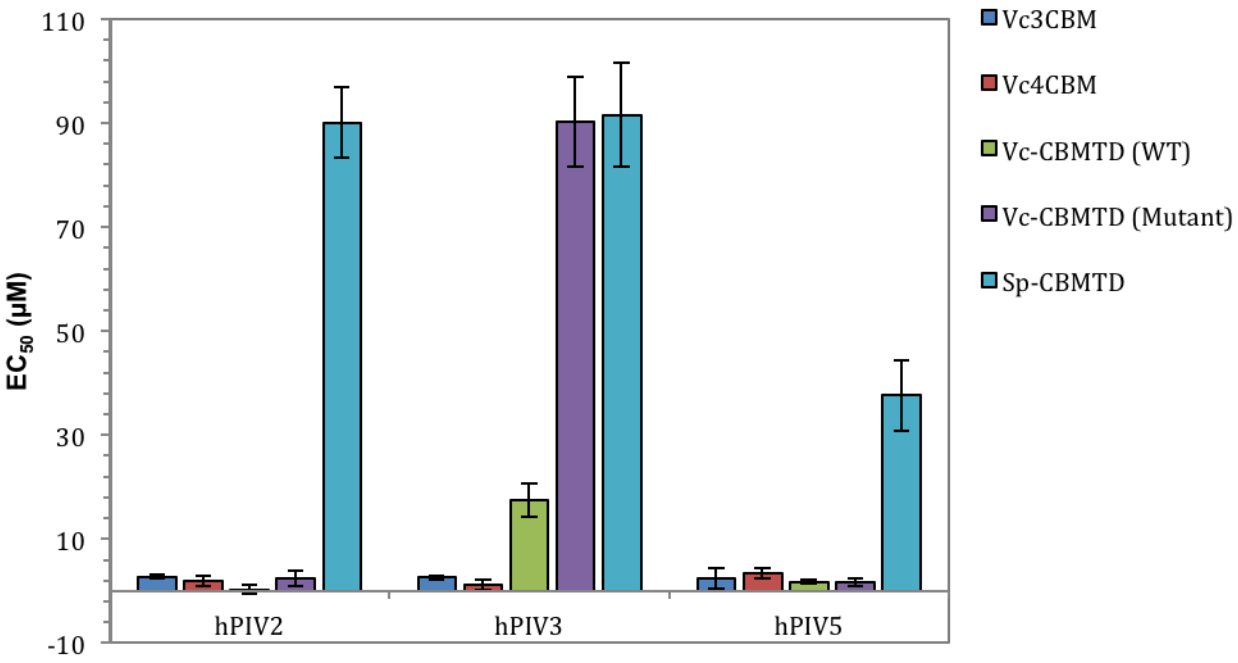


Figure 5.11 Comparison of different parainfluenza viruses against EC_{50} of multivalent constructs. Details of EC_{50} with standard error value can be referred to Table 5.1.

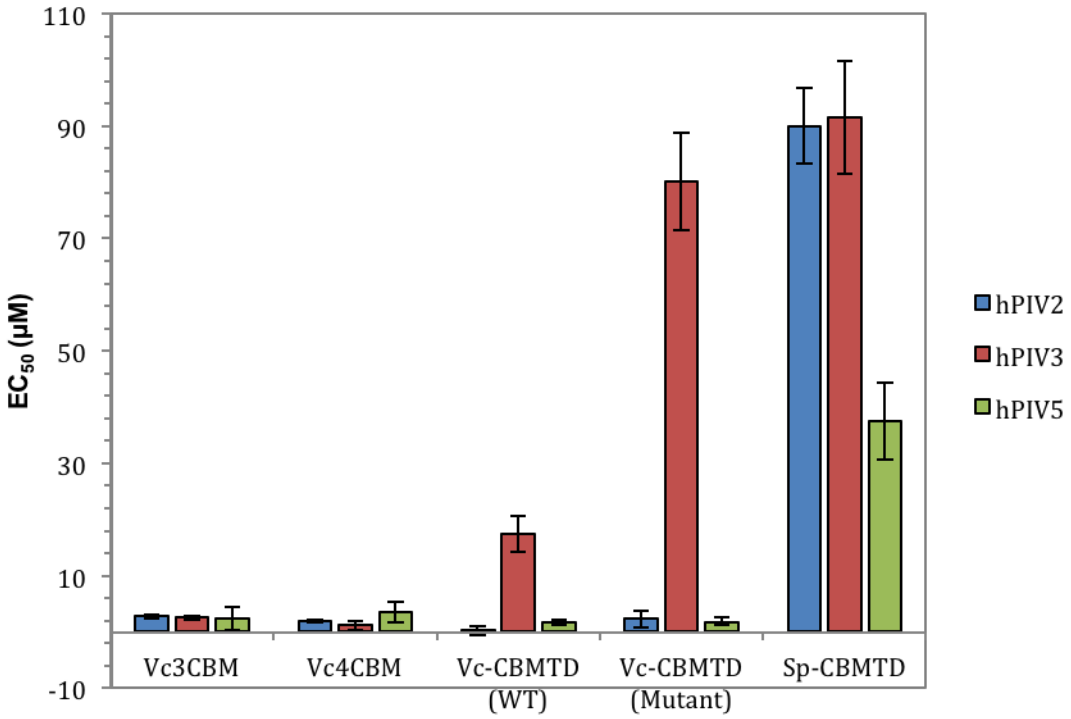


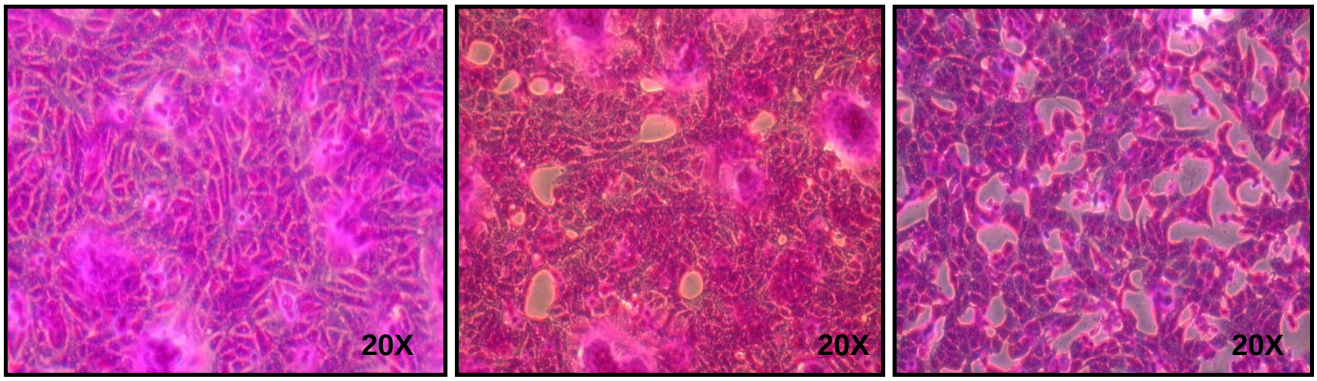
Figure 5.12 Comparison of EC_{50} values of different multivalent constructs against hPIVs viruses. Details of EC_{50} value can be referred to Table 5.1.

5.6 Effect of high concentration multivalent CBMs on cell monolayer

5.6.1 Evaluation of cytotoxic effects of Vc-CBMTD (WT) and Vc-CBMTD (Mutant) peptides

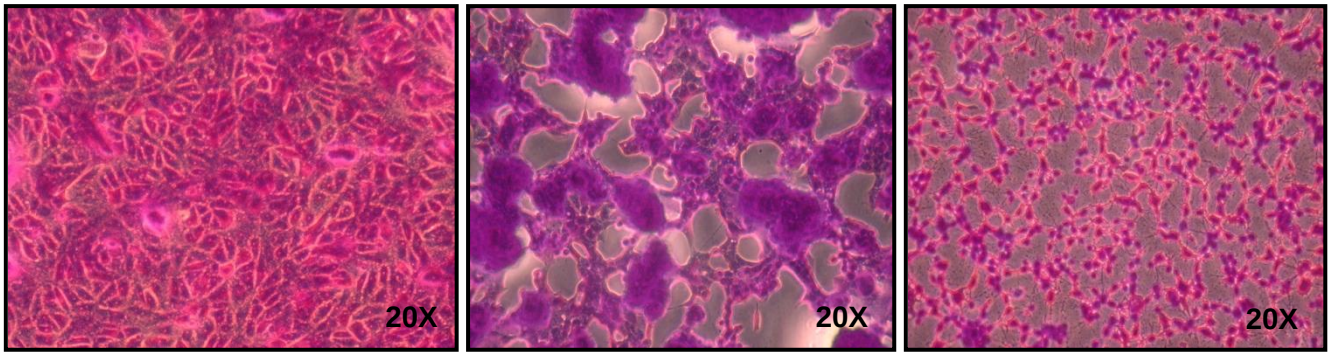
The results from *in vitro* cytotoxicity revealed that the Vc-CBMTD (WT) and Vc-CBMTD (Mutant) at high concentration caused noticeable cytotoxic effects (cell death) and morphological changed on Vero cell lines compared with the non-treated Vero cell culture (Figure 5.13 and Figure 5.14). The morphology of the cells was found to be different, where some cells appeared to have shrunk from their normal cell size. Longer incubation periods of 18 hours with both proteins induced stronger cytotoxic effects (cell death and detachment) than shorter incubation (6 hours), when treated with the similar concentration (0.5 mg/ml) (Figure 5.13 (B) and Figure 5.14 (B)). The cytotoxicity of both proteins was assessed in a parallel blue staining assay with no significant difference to the non-treated control.

The micrographs indicated cellular detachment of the confluent layer when the cell cultures were incubated with series of protein concentration of 0.5 mg/ml respectively of Vc-CBMTD (WT) and Vc-CBMTD (Mutant) for 18 hours. In addition, when the Vero cell culture was incubated with 10 times higher concentration (5.0 mg/ml) of both proteins for a fix of 18 hours, cellular detachment and cytotoxicity were more significant in wells treated with the higher concentrations. In contrast, no apparent cytotoxic effect was observed when the cell line was treated and incubated with concentration less than 0.1 mg/ml of Vc-CBMTD (WT) and Vc-CBMTD (Mutant) for 6 or 18 hours of post treatment (data not shown). No detailed studies of the compounds toxicity on Vero cells were undertaken as this was a preliminary observation during the experiment course. As compared to Vc-CBMTD (WT), more obvious cytotoxicity (cellular detachment and cell death) was observed on cell monolayers treated with Vc-CBMTD (Mutant) after a long incubation of 18 hours (Figure 5.13 (B) and Figure 5.14 (B)).



(A) **(B)** **(C)**

Figure 5.13 Effect of high concentration of Vc-CBMTD (WT) on Vero cells after 18 hours post treatment. Pictures were taken at 20X magnification. Cells were stained with 0.1% crystal violet. **(A)** Control without protein; **(B)** protein at 0.5 mg/ml; **(C)** protein at 2.0 mg/ml.



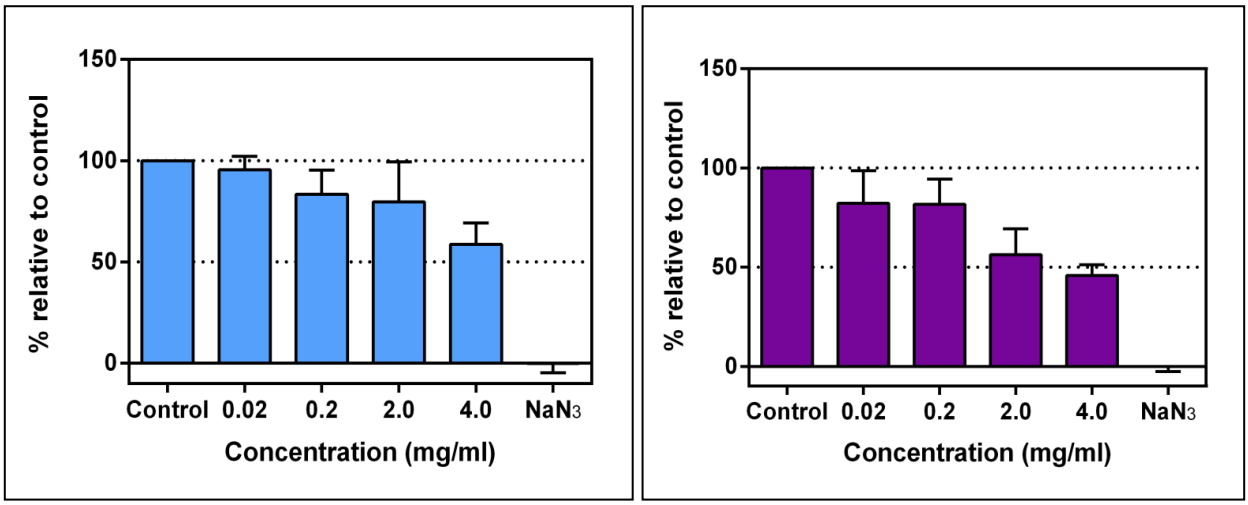
(A) **(B)** **(C)**

Figure 5.14 Effect of high concentration of Vc-CBMTD (Mutant) on Vero cells after 18 hours post treatment. Pictures were taken at 20X magnification. Cells were stained with 0.1% crystal violet. **(A)** Control without protein; **(B)** protein at 0.5 mg/ml; **(C)** protein at 5.0 mg/ml.

5.6.2 Cell viability assay

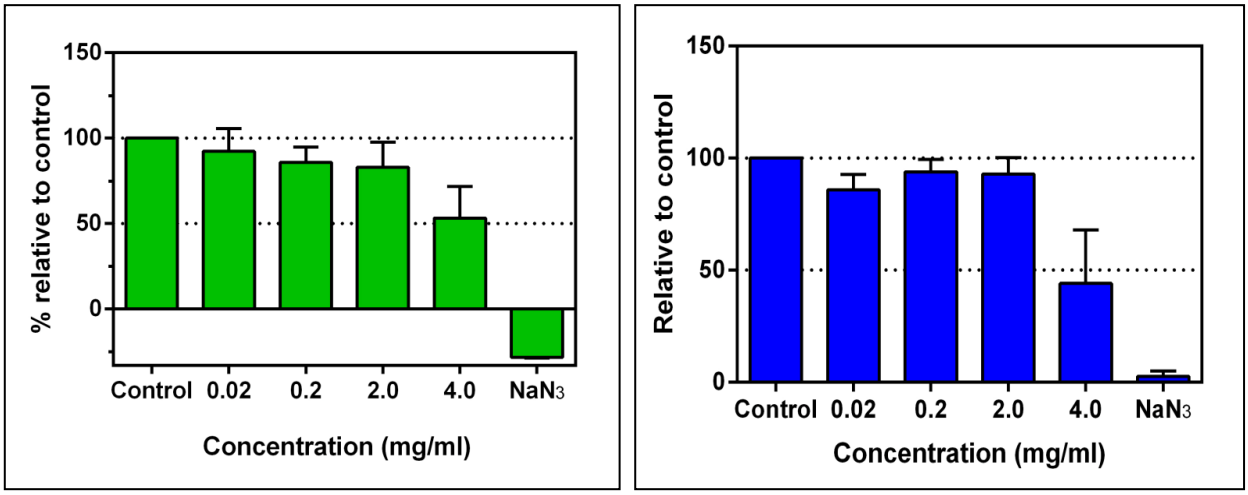
The measurement of cell viability plays a fundamental role in cell culture especially for toxicity assays. This assay also can be used to correlate cell behaviour to cell number, providing more accurate picture of anabolic cell activity (Stoddart, 2011). In this subsection, to further investigated the effect of cell toxicity, a cell viability test was performed with multivalent peptides. MDCK cells were grown in 96-well plates at 7 x 10^4 cells/ well for overnight. Previous media was aspirated out and monolayer was washed with phosphate buffer saline (PBS). CBM proteins were incubated at different concentrations on the monolayers for 6 hours and 18 hours at 37 °C, 5% CO_2 in PBS supplemented with 10% fetal calf serum (FCS). Cell survival in samples was measured using 10 ml of PrestoBlue™ cell viability reagent (Invitrogen, USA) according to the manufacturer's protocol and incubated for 1 hour at 37 °C. This dye provides a shorter incubation step compared to other resazurin-based type assay. Data values were measured at OD readings of 570 nm with reference wavelength at 620 nm using an ELISA reader. Cell viability was calculated using the following formula: 100 x [(OD_{570} of treated sample) / (OD_{570} of untreated sample)].

The assay used PrestoBlue® reagent that function as a cell viability indicator which used the reducing environment on the living cells to quantitatively measure cells proliferation. Basically, the dye will turn from blue to red in a reducing environment and become highly fluorescent, thus providing easier detection using fluorescence/ absorbance measurements. This dye is proven to obtain accurate results with great consistency to monitor living cells over time.



(A) **(B)**

Figure 5.15 Viability test on Vc-CBMTD (WT). The protein was tested against two incubation times; **(A)** 6 hours; **(B)** 18 hours.



(A) **(B)**

Figure 5.16 Viability test on Vc-CBMTD (Mutant). The protein was tested against two incubation times; **(A)** 6 hours; **(B)** 18 hours.

The experimental data presented in Figure 5.15 and Figure 5.16 indicates dose- and time-dependent effects of the Vc-CBMTD (WT) and Vc-CBMTD (Mutant) and the absence of cytotoxic effect after treatment with the negative control (Vero monolayers incubated with 20% sodium azide). This acutely toxic chemical would promote cell death and inhibit cell proliferation, which acts as a negative control of the assay. For the assay, two different post treatment times were used which were 6 hours and 18 hours to complement the previous assay (on-plate assay). Comparison between both proteins (Vc-CBMTD (WT) and Vc-CBMTD (Mutant) showed no obvious changes/ differences in the first two concentrations tested (0.02 and 0.2 mg/ml). An increment of Vc-CBMTD (WT) to 2.0 mg/ml had caused a drop in viable cell count to 29.4% when incubation was prolonged to 18 hours. Meanwhile for Vc-CBMTD (Mutant), the reading was dropped to 56.3 % from 83.0 %, which gave around 32.2 % increment in cell death. Interestingly, when the concentration of Vc-CBMTD (WT) was increased to 4.0 mg/ml, cell viability percentage was dropped drastically to 42.5 % from 79.7 % after 18 hours of incubation. As compared to Vc-CBMTD (Mutant), increment of cell death was indicated to be 44.8 % at 4.0 mg/ml of protein concentration.

Basal cytotoxicity test is considered as a starting point in assessment of potential *in vivo* toxicity of medicated compounds. Due to that, *in vitro* cytotoxicity tests are useful and necessary to define basal cytotoxicity of a compound. Intrinsic ability of a compound to cause cell death can be shown through a consequence damage of basic cellular functions (Eisenbrand et al., 2002).

5.7 Inhibition study of viral replication protein

5.7.1 Viral inhibition assay

Three replicates of Vero monolayers in 96-well plates were treated with various dilutions of multivalent proteins in serum-free DMEM for 2 hours at 37°C, 5% CO_2. Both treated and untreated control cells (only virus in plain DMEM) were infected with virus dilution of 10^{-3} MOI. After 30 minutes, the monolayers were washed two times with PBS and incubated with serum-free DMEM supplemented with 2.5 µg/ml N-acetylated Trypsin (NAT, Sigma Aldrich) for 8 or 16 hours, 37°C, 5% CO_2. The monolayers were then fixed with 4% formaldehyde in PBS before being blocked with 3% BSA in PBS for 30 minutes. Cells were washed once with PBST (PBS supplemented with 0.02% Tween 20). Each well was incubated with 50 µl of primary antibody as shown in Table 5.2 (generously given by Prof Richard Elliot's group members) at 1:1000 dilution for 2 hours at 37°C. Excess primary antibody was removed with four washes of PBST. All wells were then incubated with 50 µl of 1:5000 dilution of IgG HRP anti-mouse (Santa-Cruz, USA). Excess secondary antibody was removed by washing five times with PBST. Plates were developed by incubating each well with 50 µl of TMB substrate for 30 minutes and stirred slowly. For stop reaction, 50 µl of 1 M H_2SO_4 was added to each well and absorbance was measured at 450 nm using spectrometer. Wells containing uninfected cells were used as the background control for the assay.

Table 5.2 List of primary antibody used in viral inhibition assay

Type of virus	Type of Primary Antibody (Anti-mouse antibody)*	Recognition of Viral Protein
hPIV2	49	Nucleoprotein (NP)
hPIV3	4481, 4812, 4721	Nucleoprotein (NP)
hPIV5	214	Nucleoprotein (NP)

The *in vitro* study was performed to measure rate of inhibition of viral reproduction and viral protein synthesis by different types of CBMs. Also, this test can evaluate anti-hPIV activity of multivalent proteins and compared them with the result from plaque inhibition assays. A few of the proteins demonstrated potent efficacy against a panel of laboratory strains of parainfluenza viruses (hPIV2, hPIV3 and hPIV5). Figure 5.17 showed percentage of viral protein replications after treatment with 2 mg/ml of multivalent proteins for 2 hours. From the data, proteins Vc-CBMTD (WT), Vc3CBM and Vc4CBM strongly inhibited hPIV2 and hPIV5 infection. Vc-CBMTD (WT) gave a percentage inhibition of 61.8% and 67.7%, whereas values of 65.9% and 72 % were obtained for Vc4CBM. Moreover, the best inhibition profile for hPIV3 was given by Vc3CBM at 78% and Vc-CBMTD (WT) at 72.5% as compared to other proteins. Vc-CBMTD (Mutant) and Sp-CBMTD both gave low effect in cell protection as shown in the same figure (Figure 5.17). This data was further confirmed with cell protection EC_{50} data as seen in Table 5.1, which showed the lowest EC_{50} attributed by Vc-CBMTD (WT) at 0.2 for hPIV2 followed by Vc4CBM at 1.17 for hPIV3.

From Figure 5.18, about 16% of reduction in viral replication of hPIV3 was observed, from 88.5% at 0.5 mg/ml to 72.5% at 2.0 mg/ml after treatment with Vc-CBMTD (WT), whereas with other viruses, hPIV2 and hPIV5, a small difference in anti-hPIVs effect was observed.

From Figure 5.19, Vc4CBM was showed to be among the best multivalent protein to promote anti-hPIV activity especially with hPIV3 infections. The data percentage reduced from 89% to 75% from 0.5 to 2.0 mg/ml, followed by hPIV5 (85% to 73%) and hPIV2 (68% to 61%) respectively.

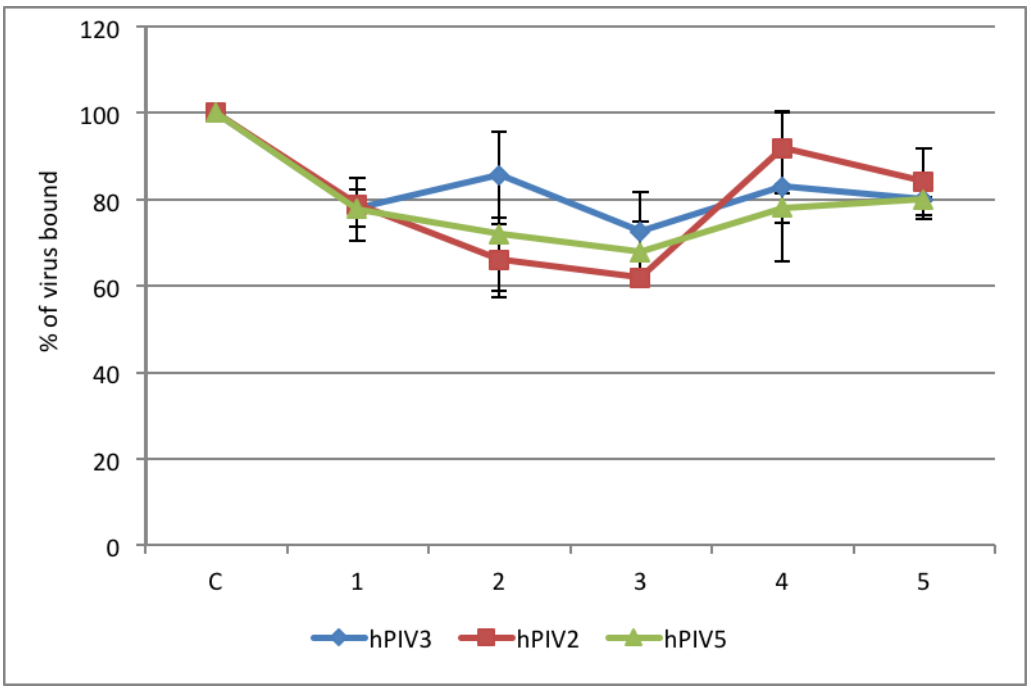


Figure 5.17 Percentage of viral replication proteins after treatment with 2 mg/ml of different multivalent proteins. **(C)** Control; **(1)** Vc3CBM; **(2)** Vc4CBM; **(3)** Vc-CBMTD (WT); **(4)** Vc-CBMTD (Mutant); **(5)** Sp-CBMTD.

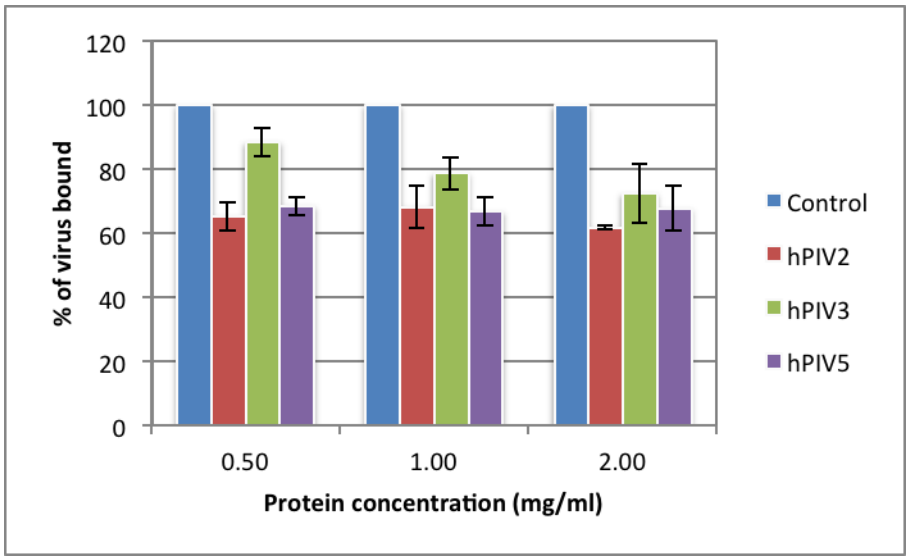


Figure 5.18 Effect of anti-hPIVs activity of Vc-CBMTD (WT) at different concentrations on Vero cells. Percentage of virus bound is based on the replication of viral protein (NP). Control represents virus in plain DMEM.

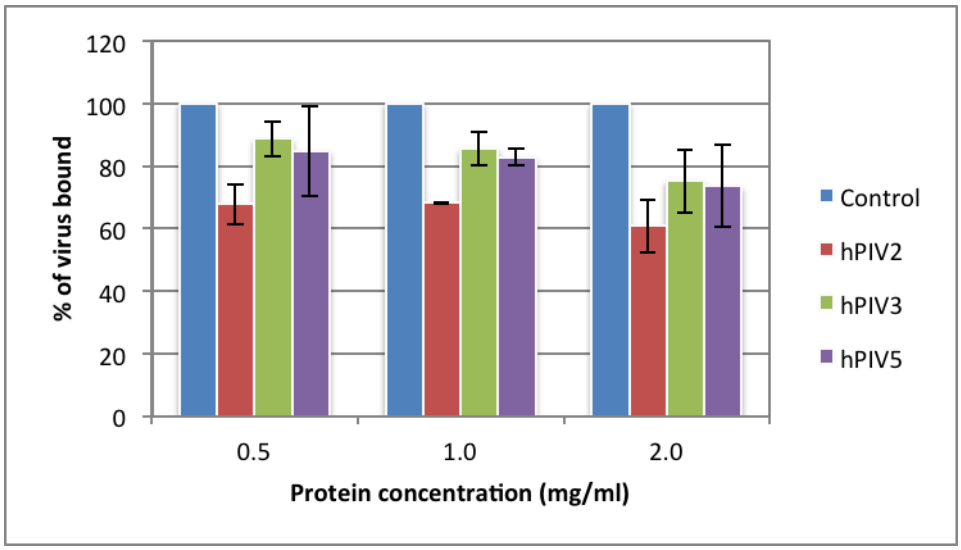


Figure 5.19 Effect of anti-hPIVs activity of Vc4CBM at different concentrations on Vero cells. Percentage of viruses bound is based on the replication of viral proteins presence. Control represent virus in plain DMEM.

5.8 Discussion

hPIVs are known for their threat to human respiratory disease to infants, children and elderly people. Unfortunately, effective vaccines or antivirals to combat hPIVs are not yet unavailable. Therefore, the development and use of antiviral drugs capable of preventing viral infection becomes a strong alternative approach. Recently, antiviral discovery programs have been initiated to target many pathogenic viruses using various mechanism-based screening approaches (Hwang et al., 2003; Mason et al., 2004). Many chemicals and natural product mixtures are being tested for their ability to knock out various viral infections.

In vitro systems are used for screening purposes and for generating more comprehensive toxicological profiles. Also, *in vitro* assays provide very important tools to enhance the extrapolation of data from *in vitro* to *in vivo* in humans. Results of *in vitro* studies have demonstrated the ability of multivalent proteins (Vc3CBM, Vc4CBM and Vc-CBMTD (WT)) to significantly inhibit the replication of hPIVs, prevent or significantly reduce its infections. By targeting the host cells rather than the virus, the CBM constructs demonstrated distinct anti-hPIV properties from the virus targeting sialic acid. These findings were not limited to hPIVs virus but also inhibited influenza viruses as reported by Connaris *et al*., (data not shown), indicated that these polypeptides could be a promising candidate as a broad-spectrum inhibitor for influenza and parainfluenza as a prophylaxis treatment.

Although cell surface sialic acid is known as the primary entry of hPIV3, additional cell surface molecules, such as heparan sulphate, may also act as secondary receptors for the virus as speculated by Moscona and Peluso (1996). However, hPIV3-HN does not contain a consensus heparin sulphate (HS) binding motif like the hPIV3-Fusion (F) protein (Bose and Banerjee, 2002). At this time, whether hPIV3 binds to HS through F or HN or both is not yet recognized. This would require further studies to investigate the interactions of hPIV3

envelope proteins with HS. It is also important to note that apart from HN, the other envelope protein, Fusion protein (F) also plays an important role in cell surface entry by forming a complex with HN during the attachment stage (Yao et al., 1997). Due to that, cell surface specificity for the paramyxovirus envelope proteins may vary depending on its specific function during the virus life cycle.

The use of cytotoxicity data as a predictor of acute systemic toxicity has been considered over the last two decades, which uses basal cytotoxicity data to predict the acute effects of compound *in vivo* (Clemedson et al., 2000). If the compound was found acutely toxic, it is anticipated the same reaction also occurs in humans. The most frequently used outcome in cellular toxicity are changes in cell morphology (Borenfreund and Puerner, 1985), breakdown of cellular permeability barrier, reduced mitochondrial function (Borenfreund and Puerner, 1985; Werner et al., 1999) and changes in cell replication (North-Root et al., 1982). Careful consideration to the choice of an *in vitro* model is important to study the cellular response to toxic chemicals. This includes information about the background levels and the metabolic competence of the models as compared to normal cells *in vivo* that need to be ascertained to interpret the results with confidence. Besides, different responses will be obtained, depending on model, whether a static cell system or proliferation cells system is used.

At the moment there is no specific treatment for parainfluenza infection other than soothing the symptoms until the illness faded away. Antibacteria only are used if secondary bacterial infection has developed. Although CBMs construct are derived from human pathogens that bind to terminal carbohydrate residues (sialic acid), they might not elicit an immune response, but still, immunogenicity remains an important issue to be evaluated by animal studies and clinical trials. After all, based on the evidence found, these compounds could be promising candidates for antiviral therapy as well as useful research tools to study hPIVs infection.

Chapter 6

In vivo study of Vc4CBM against hPIV3 and *Streptococcus pneumoniae*

6.1 Overview

Animal testing and clinical trials are two forms of *in vivo* research tools. The use of animal models to determine compound or protein toxicity of potential therapeutic agents is one of the important parts to validate product safety in humans. The efficacy of data will provide important information about the adverse effect before any clinical trials are conducted.

The rapid evolution of viruses and pathogens and their efficiency against available antivirals and vaccines remains an ongoing public health concern. In the case of respiratory disease, pathogens such as influenza virus, would usually initiate infection through binding to terminal sialic acid receptors on glycoconjugates on the surface of host cells. Therefore, any therapeutic agents or drug with the ability to block the pathogens from binding to those receptors on the human airway epithelium has the potential to prevent infections. One of the multivalent proteins, Vc4CBM has found to have the ability to exhibit potent antiviral activity against both influenza A and parainfluenza virus *in vitro* and has help reduced virus replication. However, the *in vivo* activity of the drug remains unknown and needs to be explored. Here, we demonstrate the effectiveness of Vc4CBM to protect cotton rats and mice from hPIV3 and *Streptococcus pneumoniae* infections, respectively. In this study, for

the hPIV3 experiments, *in vivo* data was obtained through collaboration with Professor Julia Hurwitz from St Jude's Children Research Hospital in Memphis, United State while for the *S. pneumoniae* study, the work was a collaboration with Professor Peter Andrew in University of Leicester, UK. The data contained in this chapter were obtained by these collaborators, and are included here in order to provide a complete story for the thesis.

6.2 Inhibition of hPIV3 infection in cotton rats

The experiment was conducted according to the protocol approved by the Animal Care and Use Committee at the St Jude's Children Research Hospital in Memphis, US. A group of 4 cotton rats were given 1 mg of Vc4CBM intranasally (I.N) in a volume of 100 µl diluted in PBS. After 1 hour, ~1.5 X 10^5 $TCID_{50}$ hPIV3 was given I.N. in 100 µl of volume. Lungs were harvested after 3 days and homogenized in 3.0 ml of PBS. Serial dilutions of 10^{-1} to 10^{-6} were made from the homogenates and about 200 µl/well (6 wells) was placed on MK-2 monolayers and incubated at 33°C for 4 days. Later, 50 µl supernatant was removed and mixed with 50 µl Turkey RBC (1:200 dil.) and incubated at 4°C for 45 minutes. Wells were scored for positive reaction. The tissue culture infectious doses ($TCID_{50}$) in the presence or absence of Vc4CBM protein was calculated that was induced by hPIV3 virus on MK-2 cells. The activity values, $TCID_{50}$ shown in Figure 6.1 indicate the differences between the $TCID_{50}$ of the test protein-treated and the untreated groups.

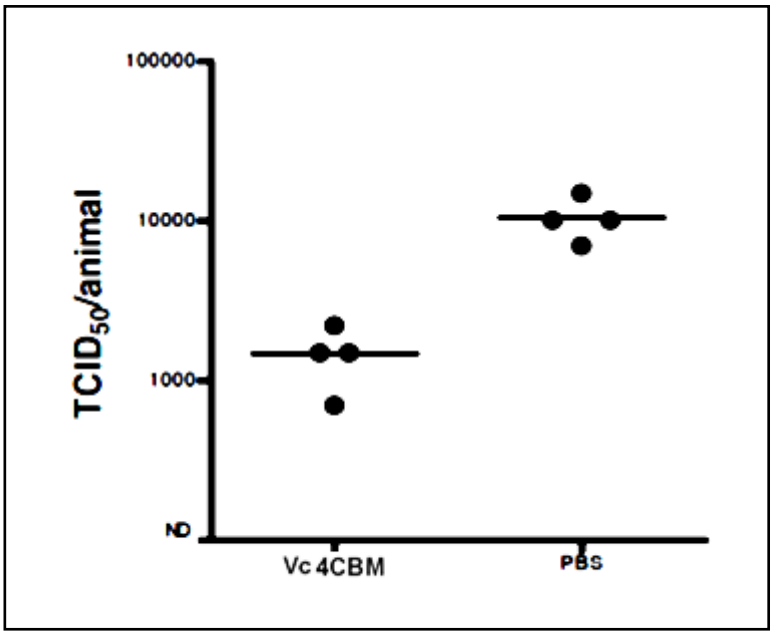


Figure 6.1 Comparison of $TCID_{50}$ of cotton rats treated with Vc4CBM and with the control (PBS only) group after infection with hPIV3 virus.

This model of hPIV3 infection in cotton rats has revealed that the respiratory tract is not only the site of infection but also the site of defence by the host. The Figure 6.1 showed $TCID_{50}$ value of each group tested, which represents median effective dose that will promote pathological change in 50% of the cell cultures inoculated. Interestingly, in the presence of Vc4CBM protein, the magnitude of $TCID_{50}$ reduced more than 9 times as compared to the control cotton rats with PBS treated. The lowest level of detection, as represented by the dotted line, was, in this case, 18 $TCID_{50}$/animal. Noted that the 'ND' value shown in Y-axis appears to have at least 1 log reduction in virus load.

6.3 *In vivo* evaluation of Vc4CBM against acute pneumococcal infection

To evaluate the therapeutic use of Vc4CBM against pneumococcal colonization, BALB/c mice (n=5) were lightly anesthetized by halothane inhalation and intranasally administered with a single dose of 500 µg of Vc4CBM in 50 µl sterile PBS at three time points: a day before (-1), at the day (0), or a day after (+1) infection. The control group received PBS only. For infection, a lethal dose of 1x 10^6 pneumococci were administered intranasally in 50 µl PBS. Animals were scored for clinical signs of disease over 7 days of post infection (168 hours). After the treatment, bacteraemia was determined by sampling the blood at various time points; 24, 36 and 48 hours of post-infection.

We used a mice model infected with *Streptococcus pneumoniae D39* to assess the antiviral activity of Vc4CBM *in vivo*. Figure 6.2 demonstrated efficacy of the protein on the survival of the infected animals when given a 500 µg dose of Vc4CBM. To evaluate the prophylactic and therapeutic potential of Vc4CBM as a treatment for pneumococcal infection, the protein was applied at different time points relative to infection in the mouse model study. Treatments at all time points (-1, 0 and +1) significantly delayed mean time to death up to 48 hours after the infections compared with the mock-untreated group.

It is clearly shown, in Figure 6.2, that treatment given a day before (-1) pneumococcal infection has proven to increase the survival rates to greater > 60 hours as compared to other mice. Interestingly, 80% of the mice were still alive until culled at 168 hours. Treatment on the day (0) of infection showed 80% mice survive until 80 hours of life but percentage of mice decreased drastically to 40% of mice surviving. Whereas, tr eatment given a day after (+1) infection showed 80% of mice surviving after 50 hours of life but the survival rate

dropped to 60% when reached 60 hours followed by 40% that survived after 70 hours of treatment.

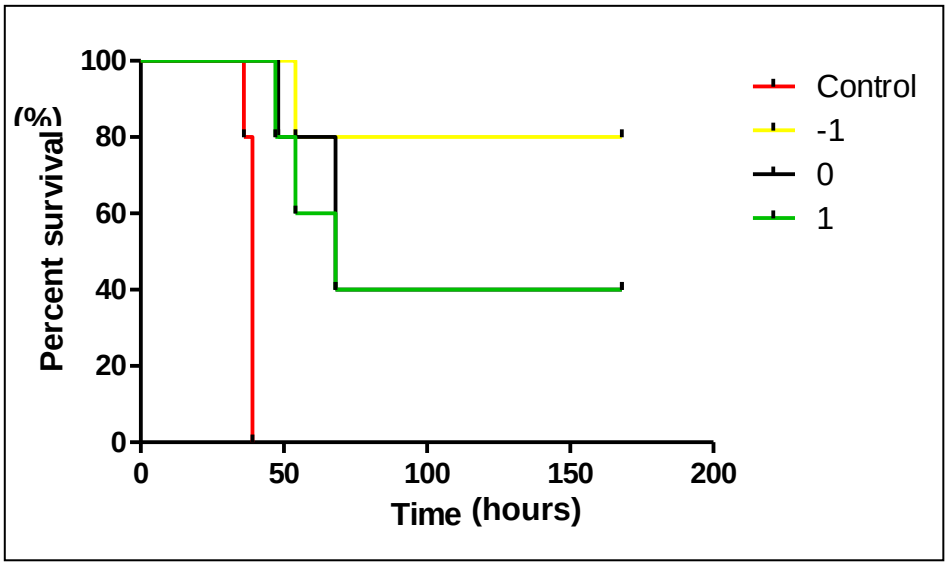


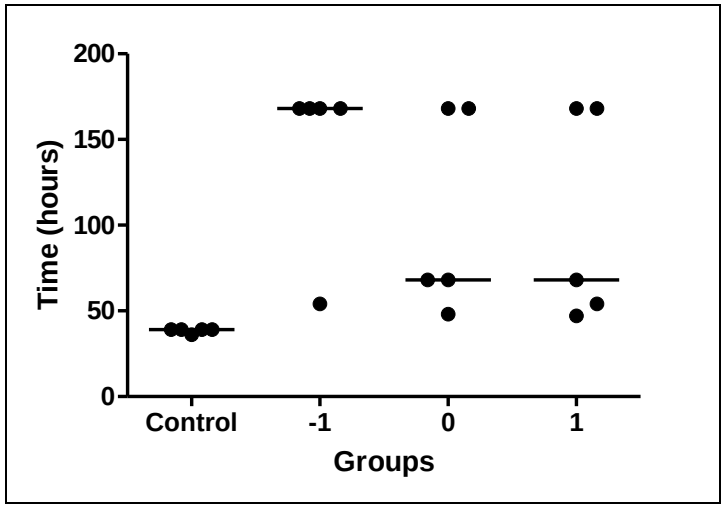
Figure 6.2 Percentage of mice survival rates after treatment with Vc4CBM proteins. Treatments were done at three different time points; (-1) a day before; (0) on the day and (+1) a day after bacterial infections.

Some respiratory bacteria including *Heamophilus influenzae*, *Streptococcus pneumoniae* and *Pseudomonas aeruginosa* have been reported to mediate cellular adhesion by binding to sialic acid receptors on the host cells (Baker et al., 1990; Barthelson et al., 1998; Fakih et al., 1997). By targeting the host rather than the virus, Vc4CBM showed a distinctive anti-pneumococcal infection through masking cellular receptors, which also can be applied to secondary pneumococcal infection due to desialylation of sialic acid receptors on the cell surface after primary infection by influenza viruses (Hedlund et al., 2010). This is due to the damage to the lining of the respiratory epithelium airways and upper respiratory system caused by influenza viruses which may facilitate pneumococcal entry and infections.

Figure 6.3 showed 4 of 5 mice were able to survive until 168 hours when treatment were given a day before (-1 day) bacterial infection, whereas only 2 of the animals had survived after treatment at (0) day and (+1) day of infection. In the control groups, all mice were only able to survive until 39 hours after the infections. Remarkably, 80% of the mice survived the lethal challenge when the treatment was given a day before the bacterial infections. Delaying the treatment given at (0) day and (+1) day of infection reduced overall survival to 40% respectively.

We next determined the effect of Vc4CBM treatment on bacterial titres in mouse blood at 24, 36 and 48 hours post-infection. Figure 6.4 demonstrates that all groups receiving a single dose treatment with Vc4CBM showed decreased bacterial counts and improved clinical signs, compared with PBS-treated control mice, which did not survive after 48 hours. Treatment one day before (-1 day) provided the best way to protect the mice from infection as shown in Figure 6.4. At 24 hours and 36 hours of incubation, bacterial counts were found to be very low, around 5 log lower after 24 hours and up to 7 log lower after 36 hours of treatment compared to mock-untreated mice.

(A)



(B) Median survival times (hour):

Groups	Control	-1	0	1
Median survival time (h)	39	168	68	68
SD	1.3	51	59	62
P value*	-	<0.01	<0.01	<0.01

*relative to the control.

Figure 6.3 Survival times of mice infected with *Streptococcus pneumonia*e after treatment with Vc4CBM. The treatment was done for 7 days. **(A)** Horizontal lines represent median survival time. Each animal is represented with a round dot. **(B)** Summary of the median survival times of each mice groups.

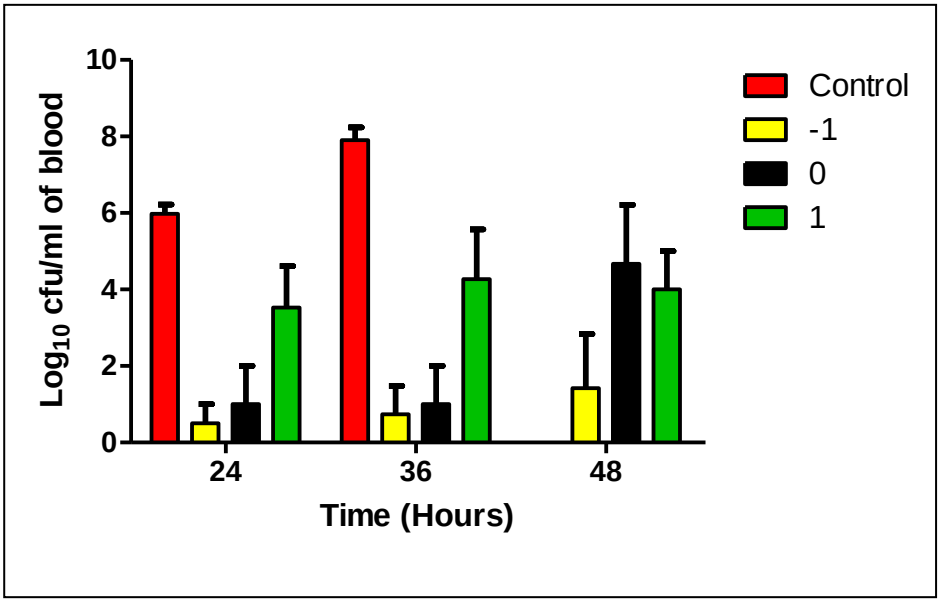


Figure 6.4 Bacterial counts from blood sampling at 24, 36 and 48 hours after treatment with Vc4CBM protein. Treatments were compared between three different time points; (-1) a day before; (0) on the day and (+1) a day after bacterial infections.

Increament of bacterial count was further observed at 48 hours post infection (p.i) when the treatment was given either on the day (0) or a day after (+1) infections at about $\log_{10}$ 5 and $\log_{10}$ 4 cfu/ml respectively. As expected, delaying treatment until after bacterial challenge resulted in substantially more bacterial replication in mice blood, but, nevertheless, titres were lower than those in the control, untreated groups. These results suggests that with Vc4CBM treatment, even when initiated early after infection, was able to control bacterial replication, providing time for adaptive immune responses to come into play.

Figure 6.5 has showed that Vc4CBM protein is found effective at inhibiting *Streptococcus pneumoniae* when given before or after bacterial challenge. The disease sign score was based on number of observation on mice which include the sign of ruffled fur, irritation,

lethargy and etc (Morton, 1985). Mice receiving the protein displayed fewer signs of clinical diseases whereas groups administered with the PBS alone exhibited the highest in clinical signs of illness up to a score level of 6. The test was done only until 48 hours as all the control mice did not survive. Importantly, when Vc4CBM was given a day before (-1) bacterial challenge, mice did not show any disease signs at 24 and 36 hours, retaining a clinical score of 0, whereas, treated mice at 48 hours p.i, were observed a very low disease sign score. At day 0 of challenge, the disease sign scores were shown to slowly decline throughout the remainder of the time courses (24 to 48 hours), with several fold decreases from 4.5 to 3 point. While, at Day +1 of administration, mice showed quite a low significant in clinical signs with the highest point of 2.3 observed after 36 hours of treatment.

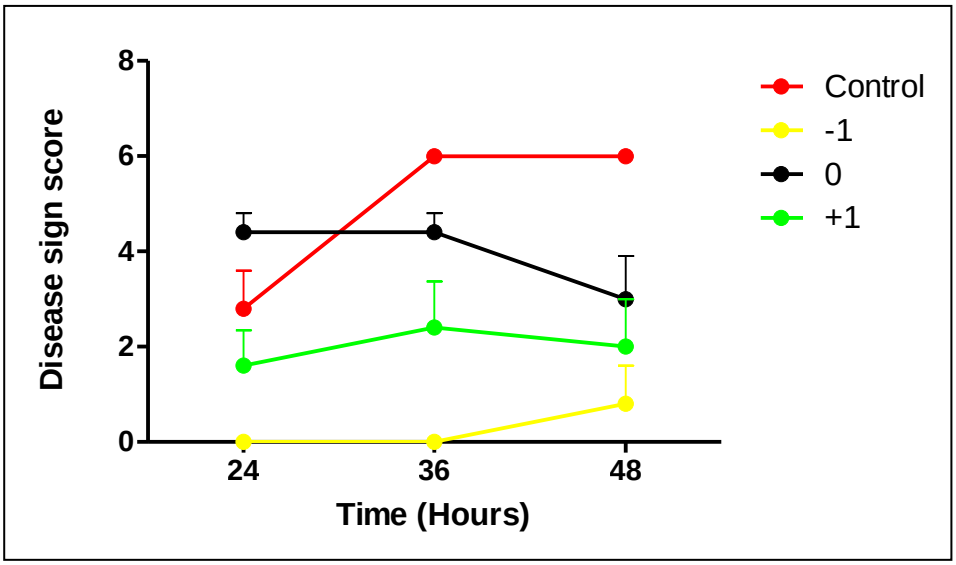


Figure 6.5 Disease sign scores of mice during Vc4CBM treatment at three different time points; (-1) a day before; (0) on the day and (+1) after the day of infections. In this figure, 0 score represented healthy mice while 6 score represented severe mice.

6.4 Discussion

Vc4CBM represents a novel and potentially broad-spectrum anti-influenza, anti-parainfluenza and anti-pneumococcal agent. Although the multivalent protein are designed to be non-immunogenic to humans in that they just bind to cell surface sialic acids, immunogenicity remains an important issue and needs to be evaluated by further animal studies and clinical trials. Since the proteins targets cell receptors (sialic acids) rather than a viral gene product, they can minimize the chance of virus evolution and subsequent drug resistant strain development. Nevertheless, it is still important to carefully evaluate the potential of this construct and other sialic acid binding CBMs on different variants. Moreover, since these proteins acts as entry blocker by binding to extracellular receptors, the protein molecules can reach the target easier than intracellular viral proteins.

Cellular adhesion/ infection by some of the most dangerous respiratory bacteria, such as *Haemophilus influenzae* (Fakih et al., 1997; Kawakami et al., 1998), *Streptococcus pneumoniae* (Barthelson et al., 1998) and *Pseudomonas aeruginosa* (Baker et al., 1990; Hazlett et al., 1986) have been reported to be mediated by recognition of sialic acid receptors on the host cells. It has been reported by Peltola et al. (2005) that mice infected with recombinant influenza virus were associated with higher incidence of secondary pneumonia infection by *Streptococcus pneumonia*e. Desialylation of host cell sialic acid by viral neuraminidase may cause increased in *Streptococcus pneumonia*e cell binding as stated by a study done by (McCullers and Bartmess, 2003).

In this chapter, we have demonstrated that novel Vc4CBM protein has good antiviral activity against a highly virulent hPIV3 and *Streptococcus pneumoniae*. When delivered before challenge at 500 µg, Vc4CBM not only protected the mice from the disease but also reduced the symptoms of infection that increase their survival rate up to 80% for 168 hours. Vc4CBM

was also effective when administered after bacterial challenge. Most importantly, after 36 hours of treatment, there was no significant increase in bacterial titres. This supported the protective action of Vc4CBM by observing decreased bacterial titers in the bloods of infected mice. Although, therapeutic treatment with the protein was not as effective as prophylaxis, it should be noted that the protein was given only one dose per day in these studies.

However, even though both hPIV3 and *S. pneumoniae* bind to similar receptors (sialic acid) on the host cells in order to mediate infections, their replication are totally different from each other. For example, hPIV3 usually bind to target host cell surface receptors via its haemagglutinin-neuraminidase (HN). With the help of fusion protein (F protein), the hPIV viral envelope fuses with the host plasma membrane to release viral nucleocapsid into cytoplasm where they replicate (Lamb, 1993; Plemper et al., 2003). Next, virions are released into the respiratory tract to begin a new round of infection (Huberman et al., 1995; Porotto et al., 2001). In contrast, bacteria such as *S. pneumoniae* use sialic acids present in the nasopharynx as a foci for adhesion, which lead to nasopharyngeal colonisation and further spread to the lungs. These bacteria eventually released into the blood stream where they further replicate (Trappetti et al., 2009; Xu et al., 2008).

In order to obtain full treatment effects, the multivalent protein should be administered intranasally to cover sufficient amount of the upper and lower respiratory tract surfaces of the animal models. Additional studies are under way to further refine the dose and treatment schedules required for the optimum protection. The efficient *in vivo* effect against hPIV3 and *Streptococcus pneumoniae* highlights the potential of Vc4CBM as an effective antiviral agent.

Chapter 7

Summary and future works

One of the efficient methods to prevent viral infections is to block host cell receptors such as sialic acids that are used by viruses to gain entry. In the current study, we have used a multivalent approach to successfully developed five novel sialic-acid binding multivalent proteins containing the CBM40 domain from *V. cholerae* sialidase (Vc3CBM, Vc4CBM, CBMTD-(WT) and CBMTD-(Mutant)) and from *S. pneumoniae* sialidase (*NanA*), Sp-CBMTD. Vc3CBM and Vc4CBM were developed using a tandem repeat approach while Vc-CBMTD (WT), Vc-CBMTD (Mutant) and Sp-CBMTD contain a fusion of CBM40 domain and trimerization domain from *P. aeruginosa* pseudaminidase to generate a trimeric form of protein. It was well understood that by increasing the valency of multiple linked polypeptides would contribute to protein stabilization. This high avidity binding approach, using multivalency, provides simultaneous binding of multi-subunit molecules to more copies of a target molecule present on the cell surface.

Biophysical characterization of the constructs was successfully performed using a variety of techniques such as ITC, SPR and thermal shift analysis to determine CBM40 interaction with different sialosides in order to understand binding energetics of protein-ligand in solution, affinity and kinetics of CBM40 binding to immobilized multivalent surfaces, and thermal stability profile for each protein. From the ITC study, it was shown that the binding of 3'SL to the multivalent CBM40 polypeptides were enthalpy driven with ΔH values range from -9.5 to -16.2 kcal/mole with entropy contribute unfavourably to the interaction. While, through SPR

analysis about 64- to 134-fold increase in affinity can be achieved through multivalent approach which also depending on temperature used during interaction.

Moreover, attempts to crystallize Vc-CBMTD (WT) and Sp-CBMTD were not completely successful. Only the trimerization domain was found to be crystallized but without its CBM40 domain due to flexibility issue with the linker. To date, the crystallographic study of both proteins is still ongoing, and therefore solving the crystal structures in complex with its glycan can aid in the development of selective reagents for use in the field of glycomics. The integration of this information will be one of the main focuses in the near future to provide more insight into the context of protein-carbohydrate interaction.

Interestingly, these multivalent proteins have been identified with the ability to prevent influenza A and parainfluenza infections as tested *in vitro*. Using a host targeted approach, Vc3CBM and Vc4CBM have shown to protect mammalian cells against influenza and parainfluenza strains significantly reducing infection symptoms such as viral titres as discussed in Chapter 4 and 5. Also, for constructs containing the additional TD domain, these were also efficient in preventing influenza A and parainfluenza infection to a certain extent. Preliminary *in vivo* studies using Vc4CBM resulted in significant protection against hPIV3 and *Pneumococcal* infection in cotton rats and mice respectively. These results confirm that the multivalent CBM40 constructs act as entry blockers to inhibit viral attachment to sialic acid on the surface of host cell.

The *in vitro* data obtained using mammalian cell culture should be interpreted with caution because the commonly used cell lines (Vero and MDCK) may not accurately represent the real human respiratory tract in terms of cell types and expression pattern of various sialic acids. In this case, HAE cell line could be used which consist of human airway tracheobronchial epithelial cells. Though, it is essential that these potential anti-infective drug candidates need to be further tested for toxicity and efficiency especially since Vc4CBM has

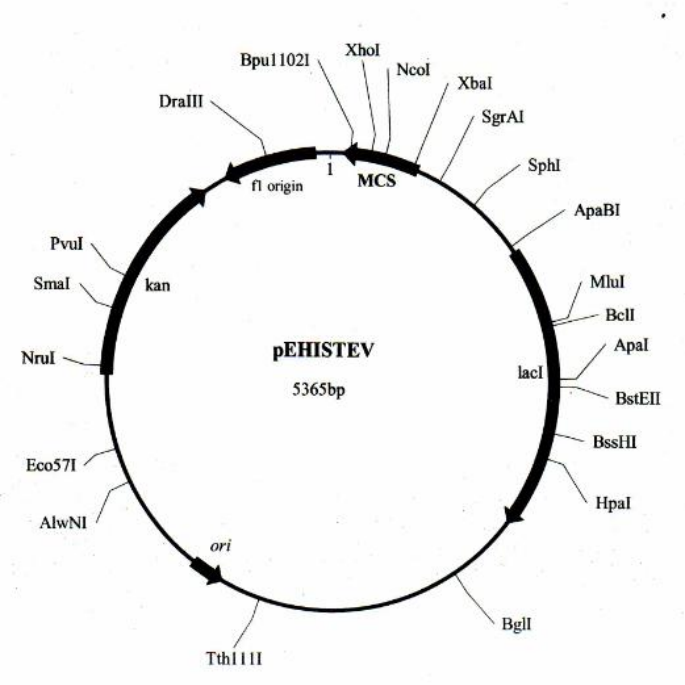
already exhibited positive effects in animal trials. As novel multivalent recombinant fusion proteins, these may contribute to a new generation of potent prophylactic and therapeutic agent for parainfluenza and influenza induced diseases.

The technology used to generate multivalent proteins here can also be applied to other areas where high avidity binding is desired, ie. for glycan screening and profiling. Additional multivalent receptor blockers may be developed to inhibit binding for not only viruses, but also various bioactive ligands such as growth factors, hormones, cytokines and etc. On the other hand, making multivalent receptor agonist can be yet another exciting application of the technology, as it may yield highly desirable therapeutic effects, such as killing cancer cells by multivalent cross-linking and activation of certain receptors that promote apoptosis.

Appendix A

Materials and methods

A-1 pEHISTEV vector map (Liu and Naismith, 2009)



Multiple Cloning Sites of pEHISTEV/pEHISGFPTEV

```
                                  T7 Pro seq primer ▶                                                    XbaI
ATCGATCTCGATCCCGCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTAGAAATAAT
                                  T7 promoter                lac operator

                    rbs          NdeI                                  6xHis
TTTGTTTAACTTTAAGAAGGAGATATACAT ATG TCG TAC TAC CAT CAC CAT CAC CAT CAC GAT TAC GAT
                                  M   S   Y   Y   H   H   H   H   H   H   D   Y   D

                                                                  TEV protease site
pEHISTEV                    GAT TAC GAC ATC CCA ACG ACC GAA AAC CTG TAT TTT CAG GGC
                              D   Y   D   I   P   T   T   E   N   L   Y   F   Q ▲ G

                                                                  TEV protease site
pEHISGFPTEV [eGFP]          GAT TAC GAC ATC CCA ACG ACC GAA AAC CTG TAT TTT CAG GGC
                              D   Y   D   I   P   T   T   E   N   L   Y   F   Q ▲ G

   NcoI       EcoRV    BamHI    EcoRI     SacI      SalI  HindIII     NotI       XhoI
GCC ATG GCT GAT ATC GGA TCC GAA TTC GAG CTC CGT CGA CAA GCT TGC GGC CGC ACT CGA GCA
 A   M   A   D   I   G   S   E   F   E   L   R   R   Q   A   C   G   R   T   R   A

CCA CCA CCA CCA CCA CTG AGA TCC GGC TGC TAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCT
 P   P   P   P   P   L   R   S   G   C end                                        ◀
                                                                                T7 Ter
Bpu1102I
GAGCAATAACTAGCATAACCCCTTGGGGCC TCTAAACGGGTCTTGAGGGGTTTTTTGCTGAA
         ———————  T7 terminator
seq primer
```

A-2 Polymerase chain reaction (PCR)

The target genes were amplified by polymerase chain reaction (PCR) from genomic DNA of *V. cholerae* sialidase. Forward and reverse primers were designed corresponding to the 5′′ and complementary 3′′ ends of target genes with specific recognition sites (as described in Chapter 2). PCRs were performed in a volume of 50 µl containing 2.5 µl of forward primer (20 µM), 2.5 µl of reverse primer (20 µM), 5 µl PCR buffer (10X), 1.5 µl (300 ng) of DNA, 1 µl dNTPs (10 mM), 1 µl pfu polymerase and water up to 50 µl. The amplification conditions were: pre-incubation at 95 °C for 30 seconds, followed by 30 cycles of 95 °C for 30 seconds, 53 °C (depending on the primer pair) for 1 minute, 72 °C for 30 seconds and final extension at 72 °C for 7 minutes before storing samples at 4 °C samples.

A-3 Blue-Native PAGE staining protocol

The NativePAGE™ NovexR Bis-Tris Gels are compatible with any of the standard Coomassie staining procedures. The NovexR Colloidal Blue Staining Kit is recommended for staining NativePAGE™ Gels. The protocol is as described below. Gel was place in 100 ml of Fix solution (40% methanol, 10% acetic acid) and microwave for 45 seconds at 950-1100 watts. Later, the gel was shake on an orbital shaker for 15-30 minutes before decanted the Fix solution. Repeat the above steps once if using NativePAGE™ NovexR 4-16% Bis-Tris gels. Another 100 ml of Coomassie R-250 stain solution (0.02% Coomassie R-250 in 30% methanol and 10% acetic acid) was added onto the gel and microwave at high (950-1100 watts) for 45 seconds. The gel was shake on an orbital shaker at room temperature until the desired background is obtained. For a more sensitive staining method, try using the Colloidal Blue Staining kit which can available at Life Technologies, Invitrogen.

A-4 Agarose gel electrophoresis

For a 0.8% w/v Agarose gel, 800 mg of agarose was dissolved in 100 ml 1X Tris-Acetate-EDTA (TAE) buffer. Solutions were cooled slightly before addition of 5 µl of 10 mg/ml ethidium bromide and then poured into a mould. DNA samples were prepared with 10 µl of gel loading solution being added to 50 µl DNA and loaded into wells. 1 kb DNA ladder (Promega), 100 bp or 1 kb was used depending on the product size. Gels were run at 80 V for 45-55 minutes.

A-5 Restriction digestion of PCR products and vector

The product band of target genes were excised from the gel and purified by Gel Extraction Kit (Promega). DNA was eluted in a volume of 50µl. To open the specific recognition sites in PCR products and vector, varying amount of PCR products and vector DNA (with varying concentration) were digested with 1-2 µl of each of the respective Fast Digest enzyme (Fermentas), 2-4 µl of 10x fast digest buffer (depending on the DNA) and total volume of 20-40 µl was incubated at 37 °C for 15 minutes followed by denaturation of the enzyme at 80 °C for 5 minutes.

A-6 Ligation

For ligation of target genes into vectors, with the ratio of 3:1 (digested PCR product: digested vector), ligation reactions were performed in 20 µl volume of ligation mixture containing 1 µl of T4 DNA ligase, 4 µl T4 ligase buffer (10X) and varying amount of DNA (depends on its concentration) and water. Ligations were incubated overnight at 16 °C or at 22 °C for 1 hour for fast ligation.

A-7 Transformation

1 µl of ligated DNA was transformed into 20-50 µl of competent host cells such as *E. coli* BL21(DE3), by cold shock (in ice 30 minutes), followed by a heat shock (in water bath 42°C for 45 seconds) and cooling on ice for 2 minutes. 450 µl of pre-warmed Luria-Bertani (LB) Media was added to the transformation reaction, and then incubated at 37 °C for 60 minutes. After incubation, the cells were spun down at 4000 rpm for 4 minutes and following removal of excess LB; cells were re-suspended with the remaining supernatant. These transformed cells were plated on LB agar containing appropriate antibiotic and incubated at 37 °C overnight. Clone positive colonies were identified either by blue-white screening (that contains the *lacZ*α gene) or colony PCR.

A-8 Colony PCR

Colonies containing recombinant vector DNA were identified using colony PCR screening. Several colonies were randomly picked by sterile toothpick and each resuspended in 10 µl water, then the wet toothpick was streaked on LB agar containing appropriate antibiotic and allowed to grow at 37 °C overnight. The remainder of each suspension was heated at 99 °C for 5 minutes before cell debris was removed by centrifugation. The cell free lysate was used in a PCR reaction containing 5 µl lysate, and 9 µl master mix which contained 0.18 µl 10 mM dNTPs, 1.4 µl 10X *pfu* DNA polymerase buffer, 0.18 µl *pfu* DNA polymerase, 0.18 µl of 5 µM vector specific forward primer and 0.18 µl of 5 µM vector specific reverse primer and 6.88 µl of water. PCR reactions were carried out at 94 °C for 10 minutes, followed by 30 cycles of 94 °C for 30 seconds, 48 °C for 30 seconds, 72 °C for 1.5 minutes. Before storage at 4 °C, samples were incubated at 72 °C for 10 minutes. Colony PCR was analysed by agarose gel electrophoresis.

A-9 DNA sequencing

The correct sequence of cloned constructs was confirmed by sequencing using T7 forward/reverse or suitable primers. For sequencing, 500 ng of DNA in 30 µl water was sent to the sequencing service at the School of Life Sciences, University of Dundee, Scotland (www.dnaseq.co.uk).

A-10 Purification of transformed plasmid DNA

A transformed colony containing the recombinant vector was inoculated in 10 ml LB containing appropriate antibiotic followed by shaking incubation at 37 °C overnight. After incubation, cells were spun down at 3000 rpm for 10 minutes at 4 °C. The pellet was collected and plasmid DNA was purified by QIA quick spin protocol (QIAGEN), which elutes plasmid DNA in 50 µl of water.

A-11 SDS-PAGE

Pre-cast 4-12% Nupage gels were set up according to the manufacturers' instruction (Invitrogen). Rigs were filled with 800 ml 1X MES buffer and 1ml of antioxidant was added to the upper chamber. Loading buffer, 4X LDS was added to the protein samples to dilute followed by denaturation at 99 °C for 4 minutes and samples were loaded into wells alongside an appropriate unstained marker such as Mark12^{TM} standard (Invitrogen). Electrophoresis was carried out at 200 V and a current of 115-70 mA for 35 minutes. On completion, gels were stained with Coomassie Blue to allow protein detection.

A-12 Pre crystallisation test

The PCT™ (Pre-Crystallization Test) is used to determine the appropriate protein concentration for crystallization screening. The protein samples were concentrated to about 5-15 mg/ml (depending on the protein) by using appropriate MWCO centrifugal concentrator (Millipore) which diluted in a sample buffer which promotes sample stability and homogeneity. Small aliquots of 50 µl of the sample were made and stored at -80 °C. Procedures as described below are according to the manufacture's protocol.

About 1.0 ml of PCT reagent A1 and PCT reagent A2 were pipetted into reservoir A1 and reservoir A2 on the VDX plate with sealant. Add about 0.05 to 1.0 µl of protein sample onto the center of a single glass cover slide. Equal volume of Reagent A1 was then pipetted onto the protein sample on the siliconized cover slide. Do not mix the drop. The cover slide with the drop was inverted and sealed. The same steps were repeated for the reagent and reservoir A2. Leave for about 30 minutes before view the drops under a light microscope with magnification between 20 and 100X. The drops were compared to each other.

The ideal drop should have a microcrystalline or light granular precipitate throughout the drop. Further reference can be found in the manufacturer's manual. Samples too concentrated can result in amorphous precipitate, while too dilute samples can result in clear drops. Precipitate and clear drops are typical crystallization screen results for reagent conditions, which do not promote crystallization and are part of every crystallization screen. However, by optimizing protein concentration for the screen, the number of clear and precipitated results can often be reduced, which at the same time enhancing the chances for crystallization.

Appendix B

The supplementary results

(B) SPR sensorgrams

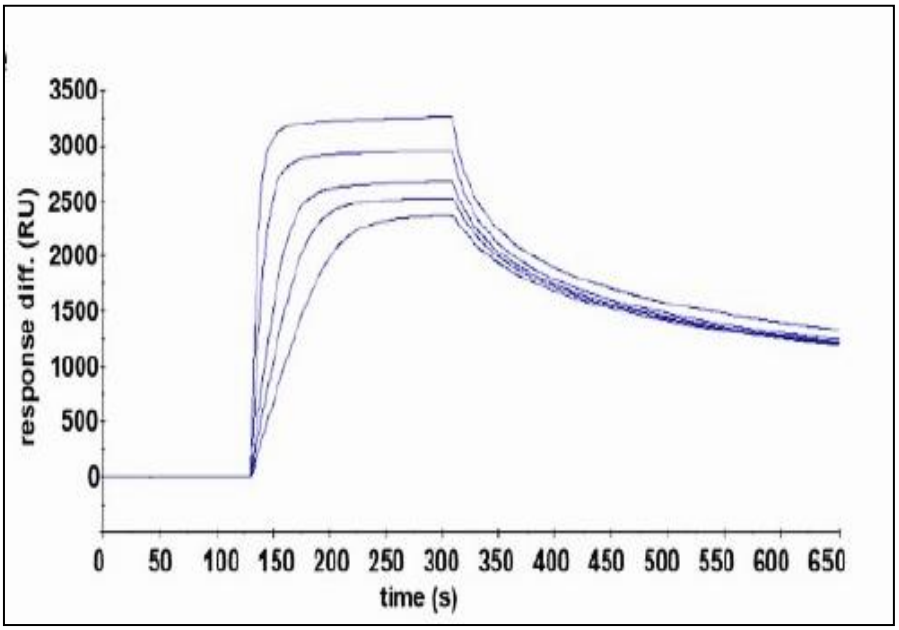


Figure B-1 Sensorgram of Vc3CBM binding profile with 3'SL at 25°C. Protein concentrations used were 1 nM, 5 nM, 20 nM, 62.5 nM and 125 nM.

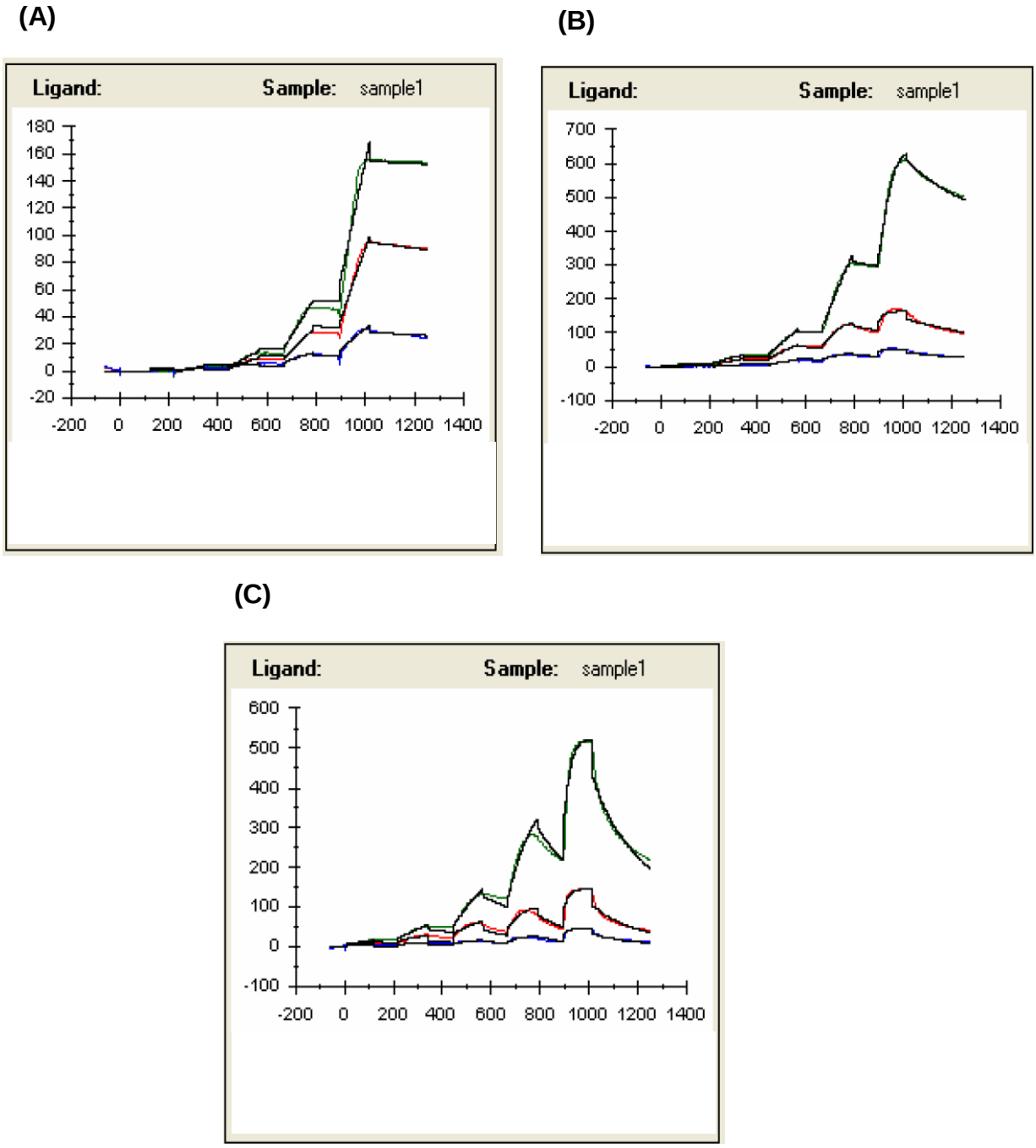


Figure B-2 Vc4CBM sensorgrams at three different temperature tested (A) 15°C; (B) 25; (C) 37°C. The analyte (protein) concentrations used were 18.0, 6.0, 2.0, 0.67 and 0.22 nM. Biotinylated ligand 3'sialyllactose-PAA-biotin (1µg/ml) was immobilised on flow cells at three different levels; green curves: high level, red curves: medium level and blue curves: low level. The black curves were the fitted data. A flow cell containing buffer only was used as the reference surface.

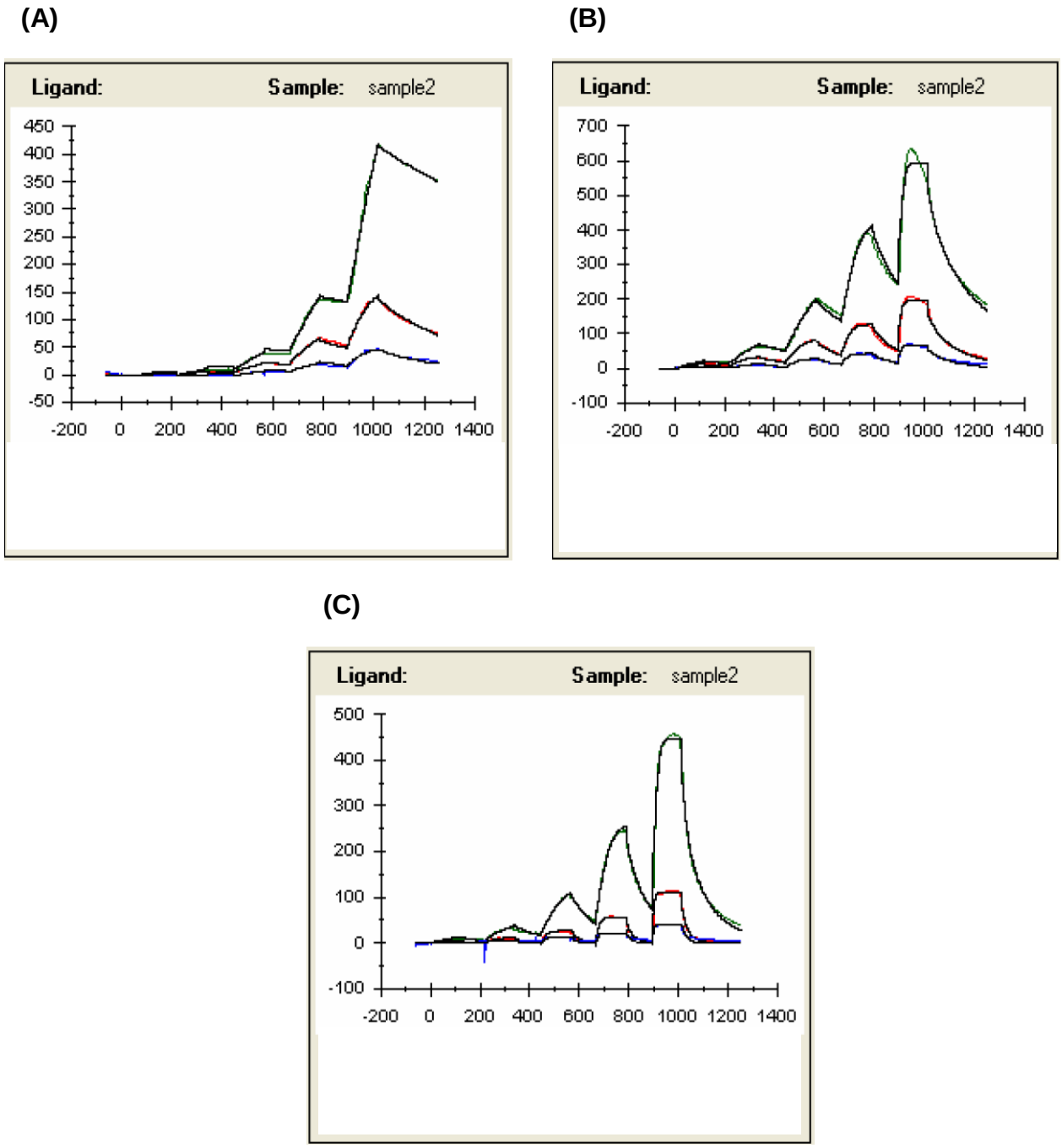


Figure B-3 Vc-CBMTD (WT) sensorgrams at three different temperature tested (A) 15°C; (B) 25°C; (C) 37°C. The analyte (protein) concentrations used were 18.0, 6.0, 2.0, 0.67 and 0.22 nM. Biotinylated ligand 3'sialyllactose-PAA-biotin (1µg/ml) was immobilised on flow cells at three different levels; green curves: high level, red curves: medium level and blue curves: low level. The black curves were the fitted data. A flow cell containing buffer only was used as the reference surface.

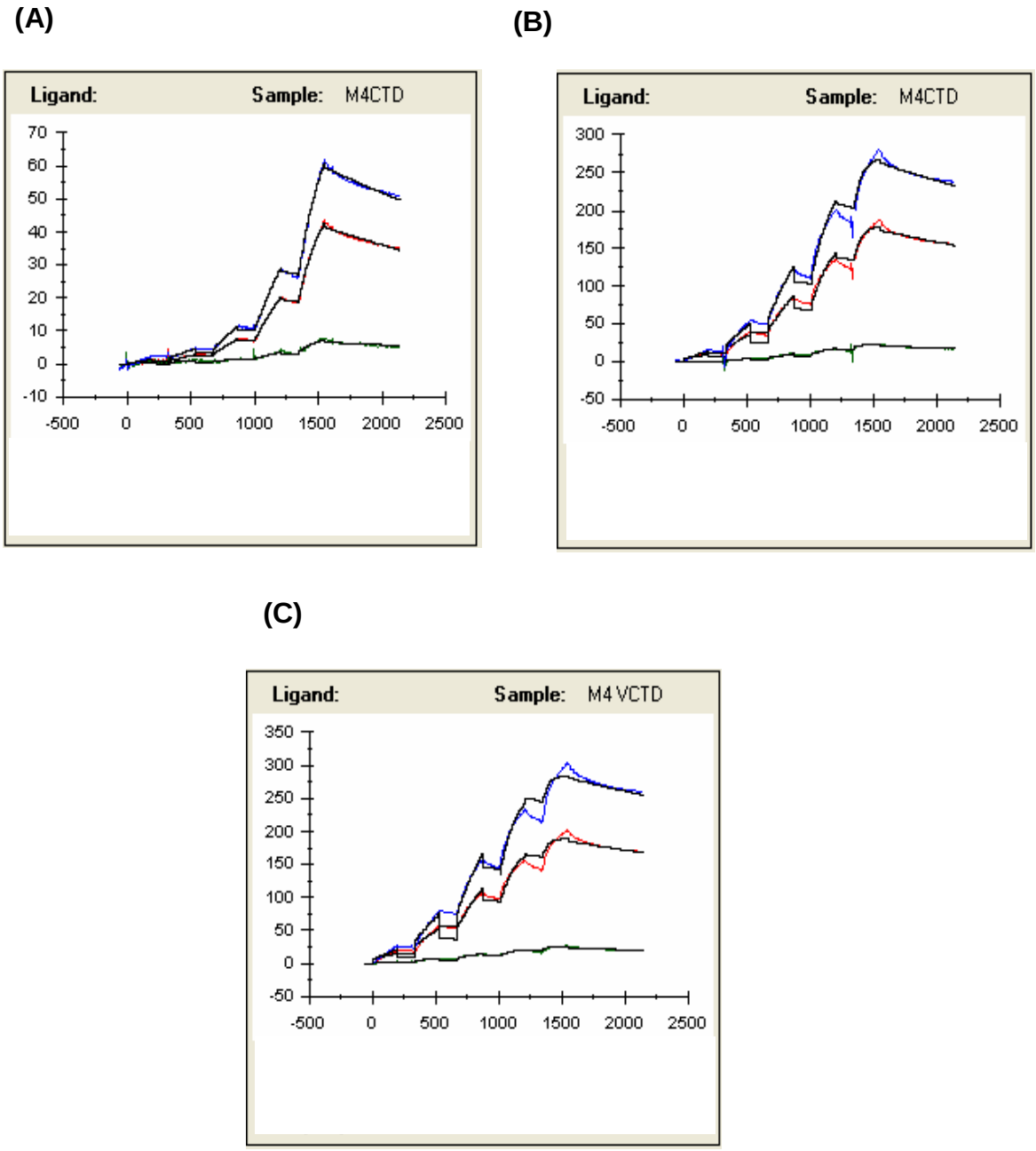


Figure B-4 Vc-CBMTD (Mutant) sensorgrams at three different temperature tested (A) 15°C; (B) 25°C; (C) 37°C. The analyte (protein) concentrations used were 18.0, 6.0, 2.0, 0.67 and 0.22 nM. Biotinylated ligand 3'sialyllactose-PAA-biotin (1µg/ml) was immobilised on flow cells at three different levels; green curves: high level, red curves: medium level and blue curves: low level. The black curves were the fitted data. A flow cell containing buffer only was used as the reference surface.

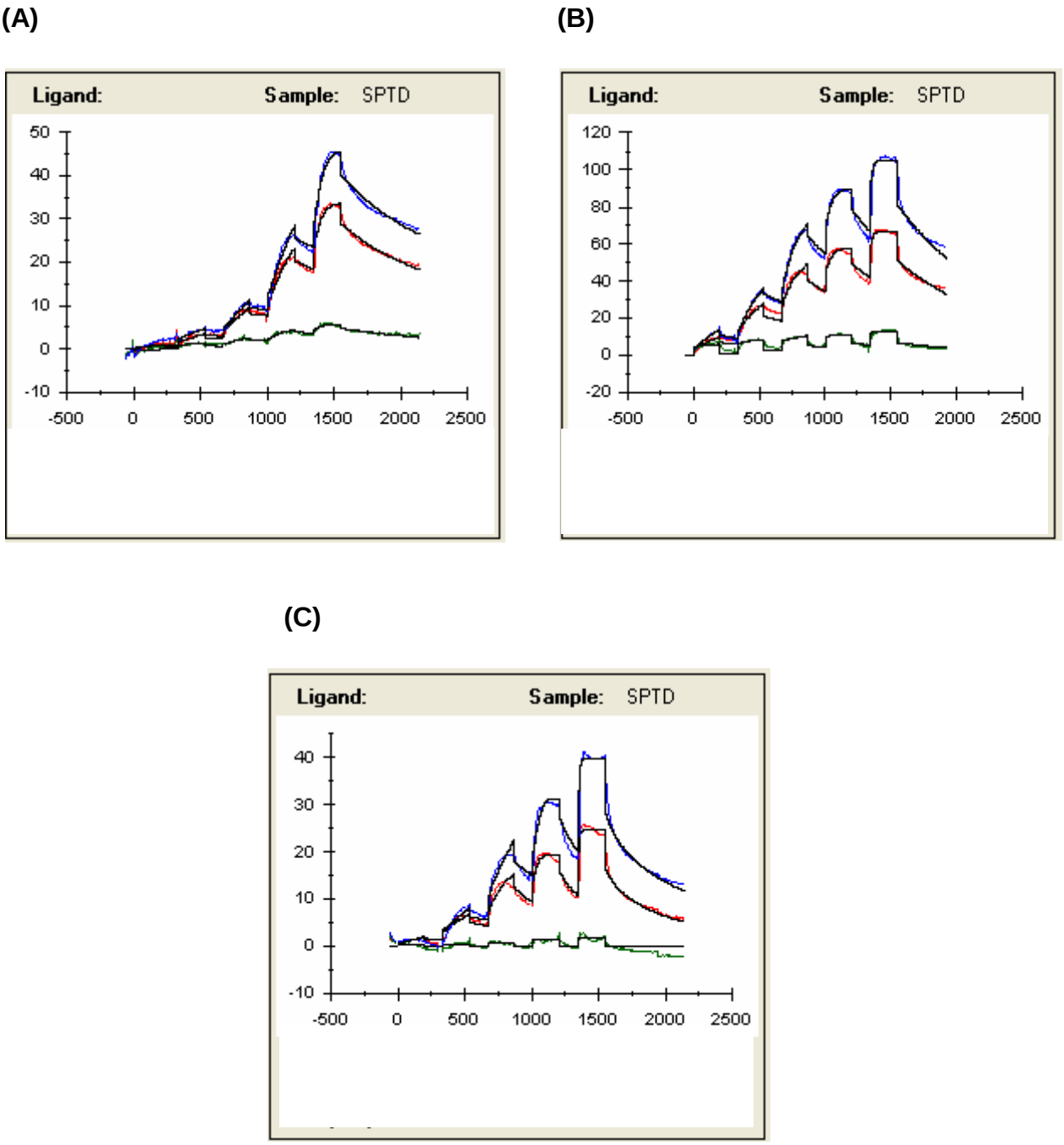


Figure B-5 Sp-CBMTD sensorgrams at three different temperature tested (A) 15°C; (B) 25°C; (C) 37°C. The analyte (protein) concentrations used were 18.0, 6.0, 2.0, 0.67 and 0.22 nM. Biotinylated ligand 3'sialyllactose-PAA-biotin (1µg/ml) was immobilised on flow cells at three different levels; green curves: high level, red curves: medium level and blue curves: low level. The black curves were the fitted data. A flow cell containing buffer only was used as the reference surface.

Reference

Achyuthan, K.E., and Achyuthan, A.M. (2001). Comparative enzymology, biochemistry and pathophysiology of human exo-alpha-sialidases (neuraminidases). Comp Biochem Physiol B Biochem Mol Biol *129*, 29-64.

Agusti, R., Giorgi, M.E., and de Lederkremer, R.M. (2007). The trans-sialidase from Trypanosoma cruzi efficiently transfer alpha-(2,3)-linked N-glycolylneuraminic acid to terminal beta-galactosyl units. Carbohydr Res *342*, 2465-2469.

Air, G.M., and Laver, W.G. (1989). The neuraminidase of influenza virus. Proteins *6*, 341-356.

Ali, E., Zhao, G., Sakka, M., Kimura, T., Ohmiya, K., and Sakka, K. (2005). Functions of family-22 carbohydrate-binding module in Clostridium thermocellum Xyn10C. Bioscience, biotechnology, and biochemistry *69*, 160-165.

Alvarez, P., Buscaglia, C.A., and Campetella, O. (2004). Improving protein pharmacokinetics by genetic fusion to simple amino acid sequences. The Journal of biological chemistry *279*, 3375-3381.

Alymova, I.V., Portner, A., Takimoto, T., Boyd, K.L., Sudhakara Babu, Y., and McCullers, J.A. (2005). The Novel Parainfluenza Virus Hemagglutinin-Neuraminidase Inhibitor BCX 2798 Prevents Lethal Synergism between a Paramyxovirus and Streptococcus pneumoniae. Antimicrobial Agents and Chemotherapy *49*, 398-405.

Alymova, I.V., Taylor, G., Mishin, V.P., Watanabe, M., Gopal Murti, K., Boyd, K., Chand, P., Babu, S., and Portner, A. (2008). Loss of the N-Linked Glycan at Residue 173 of Human Parainfluenza Virus Type 1 Hemagglutinin-Neuraminidase Exposes a Second Receptor-Binding Site. Journal of virology *82*, 8400-8410.

Alymova, I.V., Taylor, G., Takimoto, T., Lin, T.H., Chand, P., Babu, Y.S., Li, C., Xiong, X., and Portner, A. (2004). Efficacy of novel hemagglutinin-neuraminidase inhibitors BCX 2798 and BCX 2855 against human parainfluenza viruses in vitro and in vivo. Antimicrob Agents Chemother *48*, 1495-1502.

Ambrosi, M., Cameron, N.R., Davis, B.G., and Stolnik, S. (2005). Investigation of the interaction between peanut agglutinin and synthetic glycopolymeric multivalent ligands. Org Biomol Chem *3*, 1476-1480.

Angata, T., and Varki, A. (2002). Chemical diversity in the sialic acid and related alpha-keto acids: an evolutionary perspective. Chemical reviews *102*, 439-469.

Arai, R., Wriggers, W., Nishikawa, Y., Nagamune, T., and Fujisawa, T. (2004). Conformations of variably linked chimeric proteins evaluated by synchrotron X-ray small-angle scattering. Protein: Structure, Function and Bioinformatics *57*, 829-838.

Araki, R., Ali, M.K., Sakka, M., Kimura, T., Sakka, K., and Ohmiya, K. (2004). Essential role of the family-22 carbohydrate-binding modules for beta-1,3-1,4-glucanase activity of Clostridium stercorarium Xyn10B. FEBS letters *561*, 155-158.

Argos, P. (1990). An investigation of oligopeptides linking domains in protein tertiary structures and possible candidates for general gene fusion. J Mol Biol *211*, 943-958.
Ashwell, G., and Harford, J. (1982). Carbohydrate-specific receptors of the liver. Annual Review of Biochemistry *51*, 531-554.
Baeyens, K.J., De Bondt, H.L., Raeymaekers, A., Fiers, W., and De Ranter, C.J. (1999). The structure of mouse tumour-necrosis factor at 1.4 A resolution: towards modulation of its selectivity and trimerization. Acta Crystallogr D Biol Crystallogr 55 *55*, 772-778.
Baker, N., Hansson, G.C., Leffler, H., Riise, G., and Svanborg-Eden, C. (1990). Glycosphingolipid receptors for Pseudomonas aeruginosa. Infection and immunity *58*, 2361-2366.
Barthelson, R., Mobasseri, A., Zopf, D., and Simon, P. (1998). Adherence of Streptococcus pneumoniae to respiratory epithelial cells is inhibited by sialylated oligosaccharides. Infection and immunity *66*, 1439-1444.
Belser, J.A., Lu, X., Szretter, K.J., Jin, X., Aschenbrenner, L.M., Lee, A., Hawley, S., Kim do, H., Malakhov, M.P., Yu, M.*, et al.* (2007). DAS181, a novel sialidase fusion protein, protects mice from lethal avian influenza H5N1 virus infection. The Journal of infectious diseases *196*, 1493-1499.
Belshe, R.B., Newman, F.K., Tsai, T.F., Karron, R.A., Reisinger, K., Roberton, D., Marshall, H., Schwartz, R., King, J., Henderson, F.W.*, et al.* (2004). Phase 2 evaluation of parainfluenza type 3 cold passage mutant 45 live attenuated vaccine in healthy children 6-18 months old. The Journal of infectious diseases *189*, 462-470.
Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., and Bourne, P.E. (2000). The Protein Data Bank. Nucleic Acids Research *28*, 235-242.
Betz, S.F. (1993). Disulfide bonds and the stability of globular proteins. Protein Science *2*, 1551-1558.
Bogaert, D., van Belkum, A., Sluijter, M., Luijendijk, A., de Groot, R., Rumke, H.C., Verbrugh, H.A., and Hermans, P.W. (2004). Colonisation by Streptococcus pneumoniae and Staphylococcus aureus in healthy children. Lancet *363*, 1871-1872.
Bolam, D.N., Ciruela, A., McQueen-Mason, S., Simpson, P., Williamson, M.P., Rixon, J.E., Boraston, A., Hazlewood, G.P., and Gilbert, H.J. (1998a). Pseudomonas cellulose-binding domains mediate their effects by increasing enzyme substrate proximity. The Biochemical journal *331 (Pt 3)*, 775-781.
Bolam, D.N., Ciruela, A., McQueen-Mason, S., Simpson, P., Williamson, M.P., Rixon, J.E., Boraston, A.B., Hazlewood, G.P., and Gilbert, H.J. (1998b). Pseudomonas cellulose-binding domains mediate their effects by increasing enzyme substrate proximity. The Biochemical journal *331*, 775-781.
Bolam, D.N., Xie, H., White, P., Simpson, P.J., Hancock, S.M., Williamson, M.P., and Gilbert, H.J. (2001). Evidence for synergy between family 2b carbohydrate binding modules in cellulomonas fimi xylanase 11A. Biochemistry *40*, 2468-2477.
Boraston, A.B., Bolam, D.N., Gilbert, H.J., and Davies, G.J. (2004a). Carbohydrate-binding modules: fine-tuning polysaccharide recognition. The Biochemical journal *382*, 769-781.

Boraston, A.B., Bolam, D.N., Gilbert, H.J., and Davies, G.J. (2004b). Carbohydrate-binding modules: fine-tuning polysaccharide recognition. The Biochemical journal *382*, 769-781.
Boraston, A.B., Bray, M., Burn, E., Creagh, A.L., Gilkes, N., Guarna, M., Jervis, E., Johnson, P., Kormos, J., McIntosh, L.*, et al.* (1998). The structure and function of cellulose binding domains. In Carbohydrate from Trichodema reesei and other microorganisms, M. Claeyssen, W. Nenrinkx, and K. Piens, eds. (Cambridge, United Kingdom: The Royal Society of Chemistry), pp. 139-146.
Boraston, A.B., Ficko-Blean, E., and Healey, M. (2007). Carbohydrate recognition by a large sialidase toxin from Clostridium perfringens. Biochemistry *46*, 11352-11360.
Boraston, A.B., Kwan, E., Chiu, P., Warren, R.A.J., and Kilburn, D.G. (2003). Recognition and hydrolysis of noncrystalline cellulose. Journal of Biology Chemistry *278*, 6120-6127.
Boraston, A.B., McLean, B.W., Chen, G., Li, A., Warren, R.A., and Kilburn, D.G. (2002a). Co-operative binding of triplicate carbohydrate-binding modules from a thermophilic xylanase. Molecular microbiology *43*, 187-194.
Boraston, A.B., McLean, B.W., Kormos, J.M., Alam, M.A., Gilkes, N.R., Haynes, C.A., Tomme, P., Kilburn, D.G., and Warren, A.J. (1999). Carbohydrate-binding modules: diversity of structure and function (Cambridge, United Kingdom: The Royal Society of Chemistry).
Boraston, A.B., Nurizzo, D., Notenboom, V., Ducros, V., Rose, D.R., Kilburn, D.G., and Davies, G.J. (2002b). Differential oligosaccharide recognition by evolutionarily-related beta-1,4 and beta-1,3 glucan-binding modules. Journal of molecular biology *319*, 1143-1156.
Boraston, A.B., Nurizzo, D., Notenboom, V., Ducros, V., Rose, D.R., Kilburn, D.G., and Davies, G.J. (2002c). Differential oligosaccharide recognition by evolutionarily-related beta-1,4 and beta-1,3 glucan-binding modules. Journal of molecular biology *319*, 1143-1155.
Borenfreund, E., and Puerner, J.A. (1985). Toxicity determined in vitro by morphological alterations and neutral red absorption. Toxicology letters *24*, 119-124.
Bose, S., and Banerjee, A.K. (2002). Role of heparan sulfate in human parainfluenza virus type 3 infection. Virology *298*, 73-83.
Bose, S., Malur, A., and Banerjee, A.K. (2001). Polarity of human parainfluenza virus type 3 infection in polarized human lung epithelial A549 cells: role of microfilament and microtubule. Journal of virology *75*, 1984-1989.
Braz, V.A., and Howard, K.J. (2009). Separation of protein oligomers by blue native gel electrophoresis. Analytical biochemistry *388*, 170-172.
Bright, R.A., Shay, D.K., Shu, B., Cox, N.J., and Klimov, A.I. (2006). Adamantane resistance among influenza A viruses isolated early during the 2005-2006 influenza season in the United States. JAMA *295*, 891-894.
Buschiazzo, A., Tavares, G.A., Campetella, O., Spinelli, S., Cremona, M.L., Paris, G., Amaya, M.F., Frasch, A.C., and Alzari, P.M. (2000). Structural basis of sialyltransferase activity in trypanosomal sialidases. EMBO J *19*, 16-24.
Camara, M., Boulnois, G.J., Andrew, P.W., and Mitchell, T.J. (1994). A neuraminidase from Streptococcus pneumoniae has the features of a surface protein. Infect Immun *62*, 3688-3695.

Cantarel, B.L., Coutinho, P.M., Rancurel, C., Bernard, T., Lombard, V., and Henrissat, B. (2009). The Carbohydrate-Active enZymes database (CAZy): an expert resource for Glycogenomics. Nucleic Acids Research *37*, D233-238.
Carrard, G., and Linder, M. (1999). Widely different off rates of two closely related cellulose-binding domains from Trichoderma reesei. European journal of biochemistry / FEBS *262*, 637-643.
CCP4 (1994). The CCP4 suite: programs for protein crystallography. Acta Crystallogr D Biol Crystallogr *50*, 760-763.
Chen, B.L., Arakawa, T., Hsu, E., Narhi, L.O., Tressel, T.J., and Chien, S.L. (1994). Strategies to suppress aggregation of recombinant keratinocyte growth factor during liquid formulation development. Journal of pharmaceutical sciences *83*, 1657-1661.
Chen, J., Acton, T.B., Basu, S.K., Montelione, G.T., and Inouye, M. (2002). Enhancement of the Solubility of Proteins Overexpressed in Escherichia coli by Heat Shock. J Mol Microbiol Biotechnol *4*, 519-524.
Chin, J., Magoffin, R., Shaearer, L., Schieble, J., and Lennette, E. (1969). Field evaluation of a respiratory syncytial virus vaccine and a trivalent parainfluenza virus vaccine in a pediatric population. American journal of epidemiology *89*, 449-463.
Claas, E.C.J., Osterhaus, A.D.M.E.O., van Beek, R., De Jong, J.C., Rimmelzwaan, G.F., Senne, D.A., Krauss, S., Shortridge, K.F., and Webster, R.G. (1998). Human influenza A H5N1 virus related to a highly pathogenic avian influenza virus. Lancet, *351*, 472-477.
Clemedson, C., Barile, F.A., Chesne, C., Cottin, M., Curren, R., Eckwall, B., Ferro, M., Gomez-Lechon, M.J., Imai, K., Janus, J.*, et al.* (2000). MEIC evaluation of acute systemic toxicity. Part VII. Prediction of human toxicity by results from testing of the first 30 reference chemicals with 27 further in vitro assays. ATLA *28*, 159-200.
Connaris, H., Crocker, P.R., and Taylor, G.L. (2009). Enhancing the receptor affinity of the sialic acid-binding domain of Vibrio cholerae sialidase through multivalency. The Journal of biological chemistry *284*, 7339-7351.
Corfield, A.P., Wagner, S.A., Clamp, J.R., Kriasis, M.S., and Hopkins, L.C. (1992). Mucin degradation in the human colon: production of sialidase, sialate O-acetylesterase, N-acetylneuraminatelyase, arylesterase and glycosulfatase activities by strains of fecal bacteria. Infection and immunity *60*, 3971-3978.
Coutinho, J.B., Gilkes, N.R., Kilburn, D.G., Warren, R.A.J., and Miller, R.C. (1993). The nature of the cellulose-binding domain affects the activities of a bacterial endoglucanase on different forms of cellulose. FEMS microbiology letters *113*, 211-218.
Coutinho, P.M., and Henrissat, B. (1999). Carbohydrate-active Enzyme: An Integrated Database Approach (Royal Society of Chemistry, Cambridge).
Cox, N.J., and Subbarao, K. (2005). Global epidemiology of influenza: Past and present. Annu Rev Med *51*, 407-421.
Crennell, S., Garman, E., Laver, G., Vimr, E., and Taylor, G. (1994). Crystal structure of Vibrio cholerae neuraminidase reveals dual lectin-like domains in addition to the catalytic domain. Structure *2*, 535-544.
Crocker, P.R. (2002). Siglecs: sialic-acid-binding immunoglobulin-like lectins in cell-cell interactions and signalling. Current opinion in structural biology *12*, 609-615.

Crocker, P.R., Kelm, S., Dubois, C., Martin, B., McWilliam, A.S., Shotton, D.M., Paulson, J.C., and Gordon, S. (1991). Purification and properties of sialoadhesin, a sialic acid binding receptor of murine tissue macrophages. Journal of EMBO *7*, 1661-1669.
Cross, G.A.M., and Takle, G.B. (1993). The surface trans-sialidase family of *Trypanosoma cruzi*. Annu Rev Microbiol *47*, 385-411.
Davio, S.R., Kienle, K.M., and Collins, B.E. (1995). Interdomain interactions in the chimeric protein toxin sCD4(178)-PE40: a differential scanning calorimetry (DSC) study. Pharmaceutical research *12*, 642-648.
DeLano, W. (2007). The Pymol Molecular Graphics System (DeLano Scientific, Palo Alto, CA).
Denny, F.W., and Clyde, W.A., Jr (1986). Acute lower respiratory infections in nonhospitalized children. Journal of Pediatric *108*, 635-646.
Dormitzer, P.R., Sun, Z.Y., Blixt, O., Paulson, J.C., Wagner, G., and Harrison, S.C. (2002). Specificity and affinity of sialic acid binding by the rhesus rotavirus VP8* core. Journal of virology *76*, 10512-10517.
Douglas, R.G., Jr. (1990). Prophylaxis and treatment of influenza. The New England journal of medicine *322*, 443-450.
Dumar, A.M. (2009). Swine Flu: What you need to know (Wildside Press LLC: Brownstone Books).
Durbin, A.P., and Karron, R.A. (2003). Progress in the Development of Respiratory Syncytial Virus and Parainfluenza Virus Vaccines. Clinical Infectious Disease *37*, 1668-1677.
Eisenbrand, G., Pool-Zobel, B., Baker, V., Balls, M., Blaauboer, B.J., Boobis, A., Carere, A., Kevekordes, S., Lhuguenot, J.C., Pieters, R.*, et al.* (2002). Methods of in vitro toxicology. Food and chemical toxicology : an international journal published for the British Industrial Biological Research Association *40*, 193-236.
Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr *60*, 2126-2132.
Engel, J., and Kammerer, R.A. (2000). What are oligomerization domains good for? Matrix Biology *19*, 283-288.
Fakih, M.G., Murphy, T.F., Pattoli, M.A., and Berenson, C.S. (1997). Specific binding of Haemophilus influenzae to minor gangliosides of human respiratory epithelial cells. Infection and immunity *65*, 1695-1700.
Falsey, A.R. (1991). Noninfluenza Respiratory Virus Infection in Long-Term Care Facilities. Infection Control and Hospital Epidemiology *12*, 602-608.
Falsey, A.R., Cunningham, C.K., Barker, W.H., Kouides, R.W., Yuen, J.B., Menegus, M., Weiner, L.B., Bonville, C.A., and Betts, R.F. (1995). Respiratory Syncytial Virus and Influenza A Infections in the Hospitalized Elderly. The Journal of Infectious Disease *172*.
Fanzani, A., Zanola, A., Faggi, F., Papini, N., Venerando, B., Tettamanti, G., Sampaolesi, M., and Monti, E. (2012). Implications for the mammalian sialidases in the physiopathology of skeletal muscle. Skeletal muscle *2*, 23.
Feizi, T., and Childs, R.A. (1985). Carbohydrate and structures of glycoproteins and glycolipids as differentiation antigens, tumour associated antigens and components of receptor systems. Trends in Biochemical Science *10*, 24-29.
Frank, S., Kammerer, R.A., Mechling, D., Schulthness, T., Landwehr, R., Bann, J., Guo, Y., Lustig, A., Bachinger, H.P., and Engel, J. (2001).

Stabilization of Short Collagen-like Triple Helices by Protein Engineering. J Mol Biol *308*, 1081-1089.
Freire, E. (2004). Isothermal titratiom calorimetry controlling binding forces in lead optimization. Drug Discovery Today: Technologies *1*, 5.
Freire, E., Mayorga, O.L., and Straume, M. (1990). Isothermal Titration Calorimetry. Analytical Chemistry *62*, 950-959.
Fulginiti, V.A., Eller, J.J., Sieber, O.F., Joyner, J.W., Minamitani, M., and Meiklejohn, G. (1969). Respiratory virus immunization. I. A field trial of two inactivated respiratory virus vaccines; an aqueous trivalent parainfluenza virus vaccine and an alum-precipitated respiratory syncytial virus vaccine. American journal of epidemiology *89*, 435-448.
Galen, J.E., Ketley, J.M., Fasano, A., Richardson, S.H., Wasserman, S.S., and Kaper, J.B. (1992). Role of Vibrio cholerae neuraminidase in the function of cholera toxin. Infect Immun *60*, 406-415.
Gao, J., De, B.P., Han, Y., Choudhary, S., Ransohoff, R., and Banerjee, A.K. (2001). Human Parainfluenza Virus Type 3 Inhibits Gamma Interferon-Induced Major Histocompatibility Complex Class II Expression Directly and by Inducing Alpha/Beta Interferon. Journal of virology *75*, 1124-1131.
Gaskell, A., Crennell, S., and Taylor, G. (1995). The three domains of a bacterial sialidase: a beta-propeller, an immunoglobulin module and a galactose-binding jelly-roll. Structure *3*, 1197-1205.
Gilkes, N.R., Warren, R.A., Miller, R.C., Jr., and Kilburn, D.G. (1988). Precise excision of the cellulose binding domains from two Cellulomonas fimi cellulases by a homologous protease and the effect on catalysis. The Journal of biological chemistry *263*, 10401-10407.
Glezen, W.P., Frank, A.L., Taber, L.H., and Kasel, J.A. (1984). Parainfluenza virus type 3: seasonality and risk of infection and reinfection in young children The Journal of Infectious Disease *150*, 851-857.
Glezen, W.P., Greenberg, S.B., Atmar, R.L., Piedra, P.A., and Couch, R.B. (2000). Impact of respiratory virus infections on persons with chronic underlying conditions. Journal of American Medical Association (JAMA) *283*, 499-505.
Gokhale, R.S., and Khosla, C. (2000). Roles of linkers in communication between protein modules. Current Opinion in Chemical Biology *4*, 22-27.
Goldstein, M.A., Takagi, M., Hashida, S., Shoseyov, O., Doi, R.H., and Segel, I.H. (1993). Characterization of the cellulose-binding domain of the Clostridium cellulovorans cellulose-binding protein. A J Bacteriol *175*, 5762-5768.
Govorkova, E.A., Murti, G., Meignier, B., de Taisne, C., and Webster, R.G. (1996). African green monkey kidney (Vero) cells provide an alternative host cell system for influenza A and B viruses. Journal of virology *70*, 5519-5524.
Guillen, D., Santiago, M., Linares, L., Perez, R., Morlon, J., Ruiz, B., Sanchez, S., and Rodriguez-Sanoja, R. (2007). Alpha-amylase starch binding domains: cooperative effects of binding to starch granules of multiple tandemly arranged domains. Applied and environmental microbiology *73*, 3833-3837.
Hall, J., Black, G.W., Ferreira, L.M.A., Millward-Sadler, S.J., Ali, B.R.S., Hazlewood, G.P., and Gilbert, H.J. (1995). The non-catalytic cellulose-binding domain of a novel cellulase from Pseudomonas fluorescens subsp. cellulosa is important for the efficient hydrolysis of Avicel. The Biochemical journal *309*, 749-756.

Haselhorst, T., Wilson, J.C., Liakatos, A., Kiefel, M.J., Dyason, J.C., and von Itzstein, M. (2004). NMR spectroscopic and molecular modeling investigations of the trans-sialidase from Trypanosoma cruzi. Glycobiology *14*, 895-907.
Hassid, S., Choufani, G., Nagy, N., Kaltner, H., Danguy, A., Gabius, H.J., and Kiss, R. (1999). Quantitative glycohistochemical characterization of normalnasal mucosa, and of single as opposed to massive nasal polyps. Annals of Otology, Rhinology and Laryngology *108*, 797-805.
Hay, A.J. (1992). The action of adamantanamines against influenza A viruses: inhibition of the M2 ion channel protein. Seminar in Virology *3*, 21-30.
Hazlett, L.D., Moon, M., and Berk, R.S. (1986). In vivo identification of sialic acid as the ocular receptor for Pseudomonas aeruginosa. Infection and immunity *51*, 687-689.
Hedlund, M., Aschenbrenner, L.M., Jensen, K., Larson, J.L., and Fang, F. (2010). Sialidase-based anti-influenza virus therapy protects against secondary pneumococcal infection. The Journal of infectious diseases *201*, 1007-1015.
Hedrick, J.A., Barzilai, A., Behre, U., Henderson, F.W., Hammond, J., Reilly, L., and Keene, O.M. (2000). Zanamivir for treatment of symptomatic influenza A and B infection in children five to twelve years of age: a randomized controlled trial. The Pediatric infectious disease journal *19*, 410-417.
Henrickson, K.J. (2003). Parainfluenza Viruses. Clinical Microbiology Reviews *16*, 242-264.
Holler, N., Tardivel, A., Bankowski, M.K., Hertig, S., Gaide, O., Martinon, F., Tinel, A., Deperthes, D., Calderara, S., Schulthess, T.*, et al.* (2003). Two Adjacent Trimeric Fas Ligands Are Requires for Fas Signaling and Formation of a Death-Inducing Signaling Comples. Molecular and Cellular Biology *23*, 1428-1440.
Hoppe, H.J., Barlow, P.N., and Reid, K.B.M. (1994). A parallel three stranded a-helical bundle at the nucleation site of collagen triple-helix formation. FEBS letters *344*, 191-195.
Huberman, K., Peluso, R., and Moscona, A. (1995). The hemagglutinin-neuraminidase of human parainfluenza virus type 3: role of the neuraminidase in the viral life cycle. Virology *214*, 294-300.
Hurwitz, J.L. (2008). Development of Recombinant Sendai Virus Vaccines for Prevention of Human Parainfluenza and Respiratory Syncytial Virus Infections. The Pediatric infectious disease journal *27*, 126-128.
Hurwitz, J.L., Soike, K.F., Sangster, M.Y., Portner, A., Sealy, R.E., Dawson, D.H., and Coleclough, C. (1997). Intranasal Sendai virus vaccine protects African green monkeys from infection with human parainfluenza virus-type one. Vaccine *15*, 533-540.
Hwang, S., Tamilarasu, N., Kibler, K., Cao, H., Ali, A., Ping, Y.H., Jeang, K.T., and Rana, T.M. (2003). Discovery of a small molecule Tat-trans-activation-responsive RNA antagonist that potently inhibits human immunodeficiency virus-1 replication. The Journal of biological chemistry *278*, 39092-39103.
Ilver, D., Johansson, P., Miller-Podraza, H., Nyholm, P.G., Teneberg, S., and Karlsson, K.A. (2003). Bacterium-host protein-carbohydrate interactions. Methods Enzymol *363*, 134-157.
Ishizuka, A., Hashimoto, Y., Naka, R., Kinoshita, M., Kakehi, K., Seino, J., Funakoshi, Y., Suzuki, T., Kameyama, A., and Naritmasu, H. (2008).

Accumulation of free complex-type N-glycans in MKN7 and MKN45 stomach cancer cells. The Biochemical journal *413*, 227-237.
Ito, T. (2000). Interspecies transmission and receptor recognition of influenza A viruses. Microbiology and immunology *44*, 423-430.
Iwamori, M., Ohta, Y., Uchida, Y., and Tsukada, Y. (1997). Arthrobacter ureafaciens sialidase isoenzymes, L, M1 and M2, cleave fucosyl GM1. Glycoconjugate journal *14*, 67-73.
Jedrzejas, M.J. (2001). Pneumococcal virulence factors: structure and function. Microbiol Mol Biol Rev *65*, 187-207 ; first page, table of contents.
Jin, L., and Babine, R.E. (2004). Removing protein heterogeneity by truncation. In Protein Crystallography in Drug Discovery, R. Mannhold, H. Kubinyi, and G. Folkers, eds. (British Library Cataloguing: John Wiley & Sons), pp. 210-211.
Jones, J.C., Turpin, E.A., Bultmann, H., Brandt, C.R., and Schultz-Cherry, S. (2006). Inhibition of Influenza Virus Infection by a Novel Antiviral Peptide That Targets Viral Attachment to Cells. Journal of virology *80*, 11960-11967.
Kawakami, K., Ahmed, K., Utsunomiya, Y., Rikitomi, N., Hori, A., Oishi, K., and Nagatake, T. (1998). Attachment of nontypable Haemophilus influenzae to human pharyngeal epithelial cells mediated by a ganglioside receptor. Microbiology and immunology *42*, 697-702.
Kelm, S., and Schauer, R. (1997). Sialic acids in molecular and cellular recognition. International review of cytology *175*, 137-240.
Kiessling, L.L., Gestwicki, J.E., and Strong, L.E. (2000). Synthetic multivalent ligands in the exploration of cell-surface interactions. Curr Opin Chem Biol *4*, 696-703.
Kim, C.U., Lew, W., Williams, M.A., Liu, H., Zhang, L., Swaminathan, S., Bischofberger, N., Chen, M.S., Mendel, D.B., Tai, C.Y.*, et al.* (1997). Infleunza neuraminidase inhibitors possessing a novel hydrophobic interaction in the enzyme active site: Design synthesis and structural analysis of carbocyclic sialic acid analogues with potent anti-influenza activity. Journal of American Chemical Society *119*, 681-690.
Kishore, U., Eggleton, P., and Reid, K.B.M. (1997). Modular organization of carbohydrate recognition domains in animal lectins. Matrix Biology *15*, 583-592.
Kleene, R., and Schachner, M. (2004). Glycans and neural cell interactions. Nat Rev Neurosci *5*, 195-208.
Kronis, K.A., and Carver, J.P. (1982). Specificity of isolectins of wheat germ agglutinin for sialyloligosaccharides: a 360-MHz proton nuclear magnetic resonance binding study. Biochemistry *21*, 3050-3057.
Kurochkin, I.V., Procyk, R., Bishop, P.D., Yee, V.C., Teller, D.C., Ingham, K.C., and Medved, L.V. (1995). Domain structure, stability and domain-domain interactions in recombinant factor XIII. Journal of molecular biology *248*, 414-430.
Lamb, R. (1993). Paramyxovirus fusion: a hypothesis for changes. Virology *197*, 1-11.
Lamy, M.E.P.-S., F, and Debacker-Willame, E. (1973). Respiratory Viral Infections in Hospital Patients with Chronic Bronchitis. Chest *63*, 336-341.
Lee, R.T., and Lee, Y.C. (2000). Affinity enhancement by multivalent lectin-carbohydrate interaction. Glycoconjugate journal *17*, 543-551.

Lehmann, F., Tiralongo, E., and Tiralongo, J. (2006). Sialic acid-specific lectins: occurrence, specificity and function. Cellular and Molecular Life Sciences *63*, 1331-1354.
Liang, R., Loebach, J., Horan, N., Thompson, C., Yan, L., and Kahne, D. (1997). Polyvalent binding to carbohydrates immobilized on an insoluble resin. Proceedings of the National Academy of Sciences of the United States of America *94*, 10554-10559.
Lindahl, M., Brossmer, R., and Wadstrom, T. (1987). Carbohydrate receptor specificity of K99 fimbriae of enterotoxigenic Escherichia coli. Journal of Glycoconjugate *4*, 51-58.
Linder, M., Salovouri, I., Ruohonen, L., and Teeri, T.T. (1996). Characterization of a Double Cellulose-binding Domain. jbc *271*, 5.
Linding, R., Jensen, L.J., Diella, F., Bork, P., Gibson, T.J., and Russell, R.B. (2003). Protein Disorder Prediction: Implications for Structural Proteomics. Structure *11*, 1453-1459.
Lis, H., and Sharon, N. (1998). Lectins: Carbohydrate-Specific Proteins That Mediate Cellular Recognition. Chemical reviews *98*, 637-674.
Liu, H., and Naismith, J.H. (2009). A simple and efficient expression and purification system using two newly constructed vectors. Protein expression and purification *63*, 102-112.
Lo, M.C., Aulabaugh, A., Jin, G., Cowling, R., Bard, J., Malamas, M., and Ellestad, G. (2004). Evaluation of fluorescence-based thermal shift assays for hit identification in drug discovery. Analytical biochemistry *332*, 153-159.
Long, J.P., Tong, H.H., and DeMaria, T.F. (2004). Immunization with native or recombinant Streptococcus pneumoniae neuraminidase affords protection in the chinchilla otitis media model. Infect Immun *72*, 4309-4313.
Lundquist, J.J., and Toone, E.J. (2002). The Cluster Glycoside Effecr. Chem Rev *102*, 555-578.
Maeda, Y., Hatta, M., Takada, A., Watanabe, T., Goto, H., Neumann, G., and Kawaoka, Y. (2005). Live Bivalent Vaccine for Parainfluenza and Influenza Virus Infections. Journal of virology *79*, 6674-6679.
Maisner, A., Schneider-Schaulies, J., Liszewski, M.K., Atkinson, J.P., and Herrler, G. (1994). Binding of measles virus to membrane cofactor protein (CD46): importance of disulfide bonds and N-glycans for the receptor function. Journal of virology *68*, 6299-6304.
Malakov, M.P., Aschenbrenner, L.M., Smee, D.F., Wandersee, M.K., Sidwell, R.W., Gubareva, L.V., Mishin, V.P., Hayden, F.G., Kim, D.H., Ing, A.*, et al.* (2006a). Sialidase Fusion Protein as a Novel Broad-Spectrum Inhibitor of Influenza Virus Infection. Antimicrobial Agents and Chemotherapy *50*, 1470-1479.
Malakov, M.P., Aschenbrenner, L.M., Smee, D.F., Wandersee, M.K., Sidwell, R.W., Larisa, V.G., Mishin, V.P., Hayden, F.G., Kim, D.H., Ing, A.*, et al.* (2006b). Sialidase fusion protein as a novel broad-spectrum inhibitor of influenza virus infection. Antimicrob Agents Chemother *50*, 1470-1479.
Mammen, M., Choi, S.K., and Whitesides, G.M. (1998a). Polyvalent Interactions in Biological Systems: Implications for Design and Use of Multivalent Ligands and Inhibitors. Angewandte Chemie International Edition *37*, 2755-2794.

Mammen, M., Choi, S.K., and Whitesides, G.M. (1998b). Polyvalent Interactions in Biological Systems: Implications for Design and Use of Multivalent Ligands and Inhibitors. Angew Chem Int Ed *37*, 2754-2794.
Manco, S., Hernon, F., Yesilkaya, H., Paton, J.C., Andrew, P.W., and Kadioglu, A. (2006). Pneumococcal neuraminidases A and B both have essential roles during infection of the respiratory tract and sepsis. Infection and immunity *74*, 4014-4020.
Mao, H., Thakur, C.S., Chattopadhyay, S., Silverman, R.H., Gudkov, A., and Banerjee, A.K. (2008). Inhibition of human parainfluenza virus type 3 infection by novel small molecules. Antiviral research *77*, 83-94.
Mason, S.W., Lawetz, C., Gaudette, Y., Do, F., Scouten, E., Lagace, L., Simoneau, B., and Liuzzi, M. (2004). Polyadenylation-dependent screening assay for respiratory syncytial virus RNA transcriptase activity and identificatiion of an inhibitor. Nucleic Acids Research *32*, 4758-4767.
Matrosovich, M.N., Matrosovich, T.Y., Gray, T., Roberts, N.A., and Klenk, H.D. (2004a). Human and avian influenza viruses target different cell types in cultures of human airway epithelium. Proceedings of the National Academy of Sciences of the United States of America *101*, 4620-4624.
Matrosovich, M.N., Matrosovich, T.Y., Gray, T., Roberts, N.A., and Klenk, H.D. (2004b). Neuraminidase is important for the initiation of influenza virus infection in human airway epithelium. J Virol *78*, 12665-12667.
Matsumura, M., and Matthews, B.W. (1991). Stabilization of Functional Proteins by Introduction of Multiple Disulfide Bonds. Methods in Enzymology *202*, 336-356.
Matulis, D., Kranz, J.K., Salemme, F.R., and Todd, M.J. (2005). Thermodynamic stability of carbonic anhydrase: measurements of binding affinity and stoichiometry using ThermoFluor. Biochemistry *44*, 5258-5266.
McCoy, A.J. (2007). Solving structures of protein complexes by molecular replacement with Phaser. Acta Crystallogr D Biol Crystallogr *63*, 32-41.
McCullers, J.A., and Bartmess, K.C. (2003). Role of neuraminidase in lethal synergism between influenza virus and Streptococcus pneumoniae. The Journal of infectious diseases *187*, 1000-1009.
McNaught, A.D.W., A (1997). IUPAC: Compendium of Chemical Terminology, 2nd edn (Oxford: Blackwell Scientific Publications).
Millar, J.S. (2001). The sialylation of plasma lipoproteins. Atherosclerosis *154*, 1-13.
Mitchell, T.J. (2000). Virulence factors and the pathogenesis of disease caused by Streptococcus pneumoniae. Research in microbiology *151*, 413-419.
Miyagi, T., Wada, T., Yamaguchi, K., and Hata, K. (2004). Sialidase and malignancy: a mini review. Glycoconjugate journal *20*, 189-198.
Miyagi, T., and Yamaguchi, K. (2012). Mammalian sialidases: physiological and pathological roles in cellular functions. Glycobiology *22*, 880-896.
Monti, E., Bonten, E., D'Azzo, A., Bresciani, R., Venerando, B., Borsani, G., Schauer, R., and Tettamanti, G. (2010). Sialidases in vertebrates: a family of enzymes tailored for several cell functions. Adv Carbohydr Chem Biochem *64*, 403-479.
Monti, E., Preti, A., Venerando, B., and Borsani, G. (2002). Recent development in mammalian sialidase molecular biology. Neurochemical research *27*, 649-663.

Montreuil, J. (1982). Glycoproteins (Amsterdam: Elsevier).
Morton, D.B. (1985). Pain and Laboratory Animals. Nature *317*, 106.
Morton, T.A., and Myszka, D.G. (1998). Kinetic Analysis of Macromolecular Interactions Using Surface Plasmon Resonance Biosensors. Methods in Enzymology *295*, 268-294.
Moscona, A. (2005). Entry of parainfluenza virus into cells as a target for interrupting childhood respiratory disease. The Journal of Clinical Investigation *115*, 1688-1698.
Moscona, A., and Peluso, R. (1996). Analysis of human parainfluenza virus type 3 receptor binding variants: Evidence for the use of a specific sialic-acid containing receptor. Microbial pathogenesis *20*, 179-184.
Moscona, A., Porotto, M., Palmer, S., Tai, C., Aschenbrenner, L., Triana-Baltzer, G., Li, Q.X., Wurtman, D., Niewiesk, S., and Fang, F. (2010). A recombinant sialidase fusion protein effectively inhibits human parainfluenza viral infection in vitro and in vivo. The Journal of infectious diseases *202*, 234-241.
Moustafa, I., Connaris, H., Taylor, M., Zaitsev, V., Wilson, J.C., Kiefel, M.J., von Itzstein, M., and Taylor, G. (2004). Sialic acid recognition by Vibrio cholerae neuraminidase. The Journal of biological chemistry *279*, 40819-40826.
Muir, D., and Pillay, D. (1998). Respiratory virus infections in immunocompromised patients. J Med Microbiol *47*, 561-562.
Murphy, B., Phelan, P.D.J., I, and Uren, E. (1980). Seasonal pattern in childhood viral lower respiratory tract infections in Melbourne. Medical Journal of Australia *1*, 22-24.
Murrell, M.T., Porotto, M., Greengard, O., Poltoratskaia, N., and Moscona, A. (2001). A single amino acid alteration in the human parainfluenza virus type 3 hemagglutinin-neuraminidase glycoprotein confers resistance to the inhibitory effects of zanamivir on receptor binding and neuraminidase activity. J Virol *75*, 6310-6320.
Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr D Biol Crystallogr *53*, 240-255.
Myszka, D.G. (1997). Kinetic analysis of macromolecular interactions using surface plasmon resonance biosensors
Current Opinion in Biotechnology *8*, 50-57.
Niesen, F.H., Berglund, H., and Vedadi, M. (2007). The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. Nature Protocols *2*, 2212 – 2221.
North-Root, H., Yackovitch, F., Demetrulias, J., and Gacula, M., Heinz, J.E. (1982). Evaluation of an in vitro cell toxicity using rabbit corneal cells to predict the eye-irritation potential of surfactants. Toxicology letters *14*, 107-214.
Oganesyan, N., Ankoudinova, I., Kim, S.H., and Kim, R. (2007). Effect of Osmotic Stress and Heat Shock in Recombinant Protein Overexpression and Crystallization. Protein expression and purification *52*, 280-285.
Pantoliano, M.W., Petrella, E.C., Kwasnoski, J.D., Lobanov, V.S., Myslik, J., Graf, E., Carver, T., Asel, E., Springer, B.A., Lane, P.*, et al.* (2001). High-Density Miniaturized Thermal Shift Assays as a General Strategy for Drug Discovery. Journal of biomolecular screening *6*, 429-440.

Parker, D., Soong, G., Planet, P., Brower, J., Ratner, A.J., and Prince, A. (2009). The NanA neuraminidase of Streptococcus pneumoniae is involved in biofilm formation. Infection and immunity *77*, 3722-3730.
Patel, C.R., Niranjani, D.H., Desai, K.B., Pradhan, P.K., and Upadhyay, U.M. (2012). A review: Aggregation analysis of therapeutic proteins. . World Journal of Pharmaceutical Research *1*, 432-455.
Patterson, K.D., and Pyle, G.F. (1991). The geography and mortality of the 1918 influenza pandemic. Bulletin of the history of medicine *65*, 4-21.
Peiris, J.S., Yu, W.C., Leung, C.W., Cheung, C.Y., Ng, W.F., Nicholls, J.M., Ng, T.K., Chan, K.H., Lai, S.T., Lim, W.L.*, et al.* (2004). Re-emergence of fatal human influenza A subtype H5N1 disease. Lancet *363*, 617-619.
Peltola, V.T., Murti, K.G., and McCullers, J.A. (2005). Influenza virus neuraminidase contributes to secondary bacterial pneumonia. The Journal of infectious diseases *192*, 249-257.
Petter, J.G., and Vimr, E.R. (1993). Complete Nucleotide Sequence of the Bacteriophage K1F Tail Gene Encoding Endo-N-Acylneuraminidase (Endo-N) and Comparison to an Endo-N Homolog in Bacteriophage PK1E. Journal of bacteriology *175*, 4354-4363.
Pettigrew, M.M., Fennie, K.P., York, M.P., Daniels, J., and Ghaffar, F. (2006). Variation in the presence of neuraminidase genes among Streptococcus pneumoniae isolates with identical sequence types. Infection and immunity *74*, 3360-3365.
Pichichero, M.E. (2000). New combination vaccines. Pediatr Clin North Am *47*, 407-426.
Pinto, L.H., Holsinger, L.J., and Lamb, R.A. (1992). Influenza Virus M2 Protein has Ion Channel Activity. Cell *69*, 517-528.
Plemper, R.K., Lakdawala, A.S., Gernert, K.M., Snyder, J.P., and Compans, R.W. (2003). Structural features of paramyxovirus F protein required for fusion initiation. Biochemistry *42*, 6645-6655.
Poon, G.M. (2007). Enhancement of oligomeric stability by covalent linkage and its application to the human p53tet domain: thermodynamics and biological implications. Biochem Soc Trans *35*, 1574-1578.
Porotto, M., Greengard, O., Poltoratskaia, N., Horga, M.A., and Moscona, A. (2001). Human parainfluenza virus type 3 HN-receptor interaction: the effect of 4-GU-DANA on a neuraminidase-deficient variant. J Virology *76*, 7481-7488.
Porotto, M., Murrell, M., Greengard, O., Lawrence, M.C., McKimm-Breschkin, J.L., and Moscona, A. (2004). Inhibition of parainfluenza virus type 3 and Newcastle disease virus hemagglutinin-neuraminidase receptor binding: effect of receptor avidity and steric hindrance at the inhibitor binding sites. Journal of virology *78*, 13911-13919.
Raju, T.S., Briggs, J.B., Chamow, S.M., Winkler, M.E., and Jones, A.J. (2001). Glycoengineering of therapeutic glycoproteins: in vitro galactosylation and sialylation of glycoproteins with terminal N-acetylglucosamine and galactose residues. The Biochemical journal *40*, 8868-8876.
Raval, G.N., Patel, D.D., Parekh, L.J., Patel, J.B., Shah, M.H., and Patel, P.S. (2003). Evaluation of serum sialic acid, sialyltransferase and sialoprotein in oral cavity cancer. Journal of Oral Disease *9*, 119-128.
Robinson, C.R., and Sauer, R.T. (1998a). Optimizing the stability of single-chain proteins by linker length and composition mutagenesis. Proceedings of

the National Academy of Sciences of the United States of America *95*, 5929-5934.
Robinson, C.R., and Sauer, R.T. (1998b). Optimizing the stability of single chain proteins by linker length and composition mutagenesis. Proceedings of the National Academy of Sciences of the United States of America *95*, 5929-5934.
Rodriguez-Sanoja, R., Oviedo, N., and Sanchez, S. (2005). Microbial starch-binding domain. Current opinion in microbiology *8*, 260-267.
Roggentin, P., Rothe, B., Kaper, J.B., Galen, J., Lawrisuk, L., Vimr, E.R., and Schauer, R. (1989). Conserved sequences in bacterial and viral sialidases. Glycoconjugate journal *6*, 349-353.
Roggentin, P., Schauer, R., Hoyer, L.L., and Vimr, E.R. (1993). The sialidase superfamily and its spread by horizontal gene transfer. Mol Microbiol *9*, 915-921.
Roos, H., Karlsson, R., Nilshans, H., and Perrsson, A. (1998). Thermodynamic analysis of protein interactions with biosensor technology. Journal of Molecular Recognition *11*, 204-210.
Rosenberg, A.S. (2006). Effects of protein aggregates: an immunologic perspective. The AAPS journal *8*, E501-507.
Roth, J.P., Li, J.K., Smee, D.F., Morrey, J.D., and Barnard, D.L. (2009). A recombinant, infectious human parainfluenza virus type 3 expressing the enhanced green fluorescent protein for use in high-throughput antiviral assays. Antiviral research *82*, 12-21.
Russell, C.J., and Webster, R.G. (2005). The genesis of a pandemic influenza virus. Cell *123*, 368-371.
Russell, R.J., Haire, L.F., Stevens, D.J., Collins, P.J., Lin, Y.P., Blackburn, G.M., Hay, A.J., Gamblin, S.J., and Skehel, J.J. (2006). The structure of H5N1 avian influenza neuraminidase suggests new opportunities for drug design. Nature *443*, 45-49.
Rutishauser, U., and Jessell, T.M. (1988). Cell adhesion molecules in vertebrate neural development. Physiol Rev *68*, 819-857.
Salemme, F.R., Spurlino, J., and Bone, R. (1997). Serendipity meets precision: the integration of structure-based drug design and combinatorial chemistry for efficient drug discovery. Structure *5*, 319-324.
Sato, M., and Wright, P.F. (2008). Current Status of Vaccines for Parainfluenza Virus Infections. The Pediatric infectious disease journal *27*, 123-125.
Sauter, N.K., Bednarski, M.D., Wurzburg, B.A., Hanson, J.E., Whitesides, G.M., Skehel, J.J., and Wiley, D.C. (1989). Hemagglutinins from two influenza virus variants bind to sialic acid derivatives with millimolar dissociation constants: a 500-MHz proton nuclear magnetic resonance study. Biochemistry *28*, 8388-8396.
Schauer, R. (1982a). Chemistry, metabolism and biological functions of sialic acids. Adv Carbohydr Chem Biochem *40*, 131-234.
Schauer, R. (1982b). Sialic acids: chemistry, metabolism and function (Wien, New York: Springer-Verlag).
Schauer, R. (1985). Sialic acids and their role as biological masks. Trends in Biochemical Science *10*, 357-360.
Schauer, R. (2000). Achievements and challenges of sialic acid research. Glycoconjugate journal *17*, 485-499.

Schauer, R. (2004). Sialic acids: fascinating sugars in higher animals and man. zoology *107*, 49-64.
Schauer, R., and Kamerling, J.P. (1997). Chemistry, biochemistry and biology of sialic acids (Amsterdam: Elsevier Science A.V).
Schein, C.H. (1990). Solubility as a function of protein structure and solvent components. Bio/Tech *8*, 308–315.
Silver, R.P., and Vimr, E.R. (1990). Polysialic acid capsule of Escherichia coli K1. In The Bacteria: Molecular Basis of Bacterial Pathogenesis, B. Iglewski, and V. Miller, eds. (New York: Academic Press, Inc), pp. 39-60.
Simell, B., Jaakkola, T., Lahdenkari, M., Briles, D., Hollingshead, S., Kilpi, T.M., and Kayhty, H. (2006). Serum antibodies to pneumococcal neuraminidase NanA in relation to pneumococcal carriage and acute otitis media. Clin Vaccine Immunol *13*, 1177-1179.
Skehel, J.J., and Wiley, D.C. (1998). Coiled Coils in Both Intracellular Vesicle and Viral Membrane Fusion. Cell *95*, 871-874.
Snyder, J.D., and Walker, W.A. (1987). Structure and function of intestinal mucin: developmental aspects. Int Arch Allergy Appl Immunol *82*, 351-356.
Stevens, J., Blixt, O., Glaser, L., Taubenberger, J.K., Palese, P., Paulson, J.C., and Wilson, I.A. (2006). glycan microarray analysis of the heamagglutinins from modern and pandemic influenza viruses reveals different receptor specificities. Journal of molecular biology *355*, 1143-1155.
Stoddart, M.J. (2011). Cell Viability Assays: Introduction, Vol 740 (Humana Press).
Subbarao, K., Klimov, A., Katz, J., Regnery, H., Lim, W., Hall, H., Perdue, M., Swayne, D., Bender, C., Huang, J.*, et al.* (1998). Characterization of an avian influenza A (H5N1) virus isolated from a child with a fatal respiratory illness. Science *279*, 393-396.
Suzuki, Y. (2005). Sialobiology of influenza: molecular mechanism of host range variation of influenza viruses. Biological & pharmaceutical bulletin *28*, 399-408.
Taylor, G. (1996). Sialidases: structures, biological significance and therapeutic potential. Current opinion in structural biology *6*, 830-837.
Tertov, V.V., Kaplun, V.V., Sobenin, I.A., Boytsova, E.Y., Bovin, N.V., and Orekhov, A.N. (2001). Human plasma trans-sialidase causes atherogenic modification of low density lipoprotein. Atherosclerosis *159*, 103-115.
Tettelin, H., Nelson, K.E., Paulsen, I.T., Eisen, J.A., Read, T.D., Peterson, S., Heidelberg, J., DeBoy, R.T., Haft, D.H., Dodson, R.J.*, et al.* (2001). Complete genome sequence of a virulent isolate of Streptococcus pneumoniae. Science *293*, 498-506.
Thobhani, S., Ember, B., Siriwardena, A., and Boons, G.J. (2003). Multivalency and the Mode of Action of Bacterial Sialidases. J Am Chem Soc *125*, 7154-7155.
Thompson, W.W., Comanor, L., and Shay, D.K. (2006). Epidemiology of seasonal influenza: use of surveillance data and statistical models to estimate the burden of disease. The Journal of infectious diseases *194 Suppl 2*, S82-91.
Thompson, W.W., Shay, D.K., Weintraub, E., Lynnette, B., Carolyn, B.B., Nancy, J.C., and Fukuda, K. (2004). Influenza-Associated Hospitalizations in the United States. JAMA *292*, 1333-1340.

Tomme, P., Van Tilbeurgh, H., Pettersson, G., Van Damme, J., Vandekerckhove, J., Knowles, J., Teeri, T., and Claeyssens, M. (1988). Studies of the cellulolytic system of Trichoderma reesei QM 9414. Analysis of domain function in two cellobiohydrolases by limited proteolysis. European journal of biochemistry / FEBS *170*, 575-581.

Trappetti, C., Kadioglu, A., Carter, M., Hayre, J., Iannelli, F., Pozzi, G., Andrew, P.W., and Oggioni, M.R. (2009). Sialic acid: a preventable signal for pneumococcal biofilm formation, colonization, and invasion of the host. The Journal of infectious diseases *199*, 1497-1505.

Treanor, J. (2004). Influenza vaccine--outmaneuvering antigenic shift and drift. The New England journal of medicine *350*, 218-220.

Triana-Baltzer, G.B., Babizki, M., Chan, M.C., Wong, A.C., Aschenbrenner, L.M., Campbell, E.R., Li, Q.X., Chan, R.W., Peiris, J.S., Nicholls, J.M.*, et al.* (2010). DAS181, a sialidase fusion protein, protects human airway epithelium against influenza virus infection: an in vitro pharmacodynamic analysis. The Journal of antimicrobial chemotherapy *65*, 275-284.

Triana-Baltzer, G.B., Gubareva, L.V., Klimov, A.I., Wurtman, D.F., Moss, R.B., Hedlund, M., Larson, J.L., Belshe, R.B., and Fang, F. (2009a). Inhibition of neuraminidase inhibitor-resistant influenza virus by DAS181, a novel sialidase fusion protein. PloS One *4*, e7838.

Triana-Baltzer, G.B., Gubareva, L.V., Klimov, A.I., Wurtman, D.F., Moss, R.B., Hedlund, M., Larson, J.L., Belshe, R.B., and Fang, F. (2009b). Inhibition of neuraminidase inhibitor-resistant influenza virus by DAS181, a novel sialidase fusion protein. PloS ONE *4*, 1-9.

Triana-Baltzer, G.B., Gubareva, L.V., Nicholls, J.M., Pearce, M.B., Mishin, V.P., Belser, J.A., Chen, L.M., Chan, R.W., Chan, M.C., Hedlund, M.*, et al.* (2009c). Novel pandemic influenza A(H1N1) viruses are potently inhibited by DAS181, a sialidase fusion protein. PloS One *4*, e7788.

Tuomanen, E.I. (2004). The pneumococcus (Washington, D.C.: ASM Press).

Uchiyama, S., Carlin, A.F., Khosravi, A., Weiman, S., Banerjee, A., Quach, D., Hightower, G., Mitchell, T.J., Doran, K.S., and Nizet, V. (2009). The surface-anchored NanA protein promotes pneumococcal brain endothelial cell invasion. The Journal of experimental medicine *206*, 1845-1852.

Ugwu, S.O., and Apte, S.P. (2004). The Effect of Buffers on Protein Conformational Stability. Pharmaceutical Technology, 86-113.

Ulloa, F., and Real, F.X. (2001). Differential distribution of sialic acid in alpha2,3 and alpha2,6 linkages in the apical membrane of cultured epithelial cells and tissues. J Histochem Cytochem *49*, 501-510.

Varki, A. (1997). Sialic acids as ligands in recognition phenomena. The FASEB Journal *11*, 248-255.

Varki, A. (2007). Glycan-based interactions involving vertebrate sialic-acid-recognizing proteins. Nature *446*, 1023-1029.

Varki, A. (2008). Sialic acid in human health and disease. Trends in molecular medicine *14*, 351-360.

Varki, N.M., and Varki, A. (2002). Heparin inhibition of selectin-mediated interactions during the hematogenous phase of carcinoma metastasis: rationale for clinical studies in humans. Seminars in thrombosis and hemostasis *28*, 53-66.

Volkov, I., Lunina, N.A., and Velikodvorskaia, G.A. (2004). Prospects for practical application of substrate-binding modules of glycosyl hydrolases (A review). Prikladnaia biokhimiia i mikrobiologiia *40*, 499-504.
von Itzstein, M. (2007). The war against influenza: discovery and development of sialidase inhibitors. Nature Reviews Drug Discovery *6*, 967-974.
von Itzstein, M., Wu, W.Y., Kok, G.B., Pegg, M.S., Dyason, J.C., Jin, B., Phan, T.V., Smythe, M.L., White, H.F., Oliver, S.W.*, et al.* (1993). Rational design of potent sialidase-based inhibitors of influenza virus replication. Nature *363*, 418-423.
Waldron, T.T., and Murphy, K.P. (2003). Stabilization of Proteins by Ligand Binding: Application to Drug Screening and Determination of Unfolding Energetics. Biochemistry *42*, 7.
Watanabe, M., Mishin, V.P., Brown, S.A., Russell, C.J., Boyd, K., Sudhakara Babu, Y., Taylor, G., Xiong, X., Yan, X., Portner, A.*, et al.* (2009). Effect of hemagglutinin neuraminidase inhibitors BCX2798 and BCX2855 on growth and pathogenicity of sendai/human parainfluenza type 3 chimera virus in mice. Antimicrobial Agents and Chemotherapy *53*, 3942-3951.
Webby, R.J., and Webster, R.G. (2003). Are we ready for pandemic influenza? Science *302*, 1519-1522.
Webster, R.G., Bean, W.J., Gorman, O.T., Chambers, T.M., and Kawaoka, Y. (1992). Evolution and ecology of influenza A viruses. Microbiological reviews *56*, 152-179.
Wedemeyer, W.J., Welker, E., Narayan, M., and Scheraga, H.A. (2000). Disulfide Bonds and Protein Folding. Biochemistry *39*, 4207-4216.
Weigel, P.H., and Yik, J.H. (2002). Glycans as endocytosis signals: the cases of the asialoglycoprotein and hyaluronan/ chondroitin sulfate receptors. Biochimica et biophysica acta *1572*, 341-363.
Weis, W.I., Taylor, M.E., and Drickamer, K. (1998). The C-type lectin superfamily in the immune system. Immunological Reviews *163*, 19-34.
Werner, L., Legeais, J.M., Nagel, M.D., and Renard, G. (1999). Neutral red assay of the cytotoxicity of fluorocarbon coated polymethylmethacrylate intraocular lenses in vitro. Journal of Biological Material Research *148*, 814-819.
Whitford, D. (2005). Proteins : structure and function (Hoboken, NJ: J. Wiley & Sons).
Williams, B.G., Gouws, E., Boschi-Pinto, C., Bryce, J., and Dye, C. (2002). Estimates of world-wide distribution of child deaths from acute respiratory infections. Lancet Infect Dis *2*, 25-32.
Williams, D.H., Stephens, E., O'Brien, D.P., and Zhou, M. (2004). Understanding noncovalent interactions: ligand binding energy and catalytic efficiency from ligand-induced reductions in motion within receptors and enzymes. Angew Chem Int Ed Engl *43*, 6596-6616.
Williams, D.H., Stephens, E., and Zhou, M. (2003). Ligand binding energy and catalytic efficiency from improved packing within receptors and enzymes. Journal of molecular biology *329*, 389-399.
Wintermeyer, S.M., and Nahata, M.C. (1995). Rimantadine: a clinical perspective. The Annals of pharmacotherapy *29*, 299-310.
Wriggers, W., Chakravaty, S., and Jennings, P.A. (2005). Control of Protein Functional Dynamics by Peptide Linkers. Biopolymers (Peptide Science) *80*, 736-746.

Xu, G., Kiefel, M.J., Wilson, J.C., Andrew, P.W., Oggioni, M.R., and Taylor, G.L. (2011). Three Streptococcus pneumoniae sialidases: three different products. Journal of the American Chemical Society *133*, 1718-1721.
Xu, G., Potter, J.A., Russell, R.J., Oggioni, M.R., Andrew, P.W., and Taylor, G.L. (2008). Crystal structure of the NanB sialidase from Streptococcus pneumoniae. Journal of molecular biology *384*, 436-449.
Xu, G., Ryan, C., Kiefel, M.J., Wilson, J.C., and Taylor, G.L. (2009a). Structural studies on the Pseudomonas aeruginosa sialidase-like enzyme PA2794 suggest substrate and mechanistic variations. Journal of molecular biology *386*, 828-840.
Xu, G., Ryan, C., Kiefel, M.J., Wilson, J.C., and Taylor, G.L. (2009b). Structural studies on the Pseudomonas aeruginosa sialidase-like enzyme PA2794 suggest substrate and mechanistic variations. Journal of molecular biology *386*, 828-840.
Yamamoto, H., Saito, T., Kaneko, Y., Kersey, D., Yong, V.W., Bremer, E.G., Mkrdichian, E., Cerullo, L., Leestma, J., and Moskal, J.R. (1997). alpha-2,3-sialyltransferase mRNA and alpha-2,3-linked glycoprotein sialylation are increased in malignant gliomas. Brain Research *755*, 175-179.
Yang, Z.R., Thomson, R., McNeil, P., and Esnouf, R.M. (2005). RONN: the bio-basis function neural network technique applied to the detection of natively disordered regions in proteins. Bioinformatics *21*, 3369-3376.
Yao, Q., Hu, X., and Compans, R.W. (1997). Association of the parainfluenza virus fusion and hemagglutinin-neuraminidase glycoproteins on cell surfaces. Journal of virology *71*, 650-656.
Zverlov, V.V., Volkov, I.Y., Velikodvorskaya, G.A., and Schwarz, W.H. (2001). The binding pattern of two carbohydrate-binding modules of laminarinase Lam16A from Thermotoga neapolitana: differences in beta-glucan binding within family CBM4. Microbiology *147*, 621-629.