The Francis Crick Institute

Mill Hill Laboratory, London, UK

(Formerly the MRC National Institute for Medical Research)

Studies on low pH-activated HA2 from Influenza haemagglutinin

A thesis submitted by

Roksana Wiktoria Ogrodowicz

In partial fulfilment of the requirements of

University College London

For the degree of Doctor of Philosophy

February 2017

Declaration

I, Roksana Wiktoria Ogrodowicz, declare that the work presented in this thesis was performed in the laboratory of Sir John Skehel & Dr Steven Gamblin in the Division of Molecular Structure and Division of Virology at the Francis Crick Institute, Mill Hill Laboratory, in London, UK (Formerly the MRC National Institute for Medical Research). I confirm that the work presented in this thesis is my own. It has been indicated throughout the text where information has been obtained from other sources.

Roksana W. Ogrodowicz

Date:

Acknowledgements

I would like to thank Sir John Skehel and Dr Steven Gamblin for giving me the opportunity to work in the Molecular Structure/Virology lab at the Francis Crick Institute, Mill Hill Laboratory in London, and my thesis committee: Dr Lesley F. Haire, Dr Katrin Rittinger and Dr Peter Rosenthal for their advice and guidance during the course of my study.

I am especially thankful to Dr Patrick Collins for help with all aspects of this project. For this I will be forever grateful ;) Thank you!

I would like to thank for incredible help I received from my colleagues, particularly Dr Steven Smerdon, Dr Ursula Neu, Dr Xiaoli Xiong, Dr Andrew Purkiss and Dr Jon Wilson for help with X-ray data processing and analysis, Dr Phil Walker for data collection and Baculovirus/insect cell protein expression, Dr Lesley Calder for electron microscopy, Dr Steve Martin for circular dichroism spectroscopy and biolayer interferometry and Dr Yipu Lin & Dr Yan Gu for their assistance with the FI6 antibody neutralization assays.

I would like to say thanks to Dr Lasse Stach, Dr Ian Taylor, Dr David Schwefel, Dr Neil Ball and Dr Simon Pennell for their support, advice and patience throughout my Ph.D.

Abstract

Influenza A haemagglutinin is a surface glycoprotein of Influenza virus, responsible for the initial attachment of the virus to the target cell and, at a later stage, for viral membrane fusion. At the acidic pH of the endosome, the HA molecule undergoes an irreversible structural rearrangement. In consequence, the hydrophobic terminal segments of HA2 are moved to the same end of the refolded molecule, promoting membrane fusion.

16 haemagglutinin subtypes (H1-H16) identified to date can be divided into two groups based on characteristic structural features. The low pH-induced structures of proteolytically prepared and *E.coli*-expressed fragments of influenza A H3 HA2 (group 2 HA) were previously determined by X-ray crystallography.

This study presents structures of proteolytically prepared and recombinantly-expressed fragments of H1 HA2 in a postfusion conformation. Refolded H1 HA2, belonging to group 1 HA, adopts a hairpin-like conformation, similar to that of a rearranged H3 HA2. Structures were compared to the known structures of low pH-activated HA2, to gain a better understanding of the structural differences between the two groups of HA.

The data show the structures of the refolded HA2 to be conserved between the HA groups with minor differences.

These structural data are supplemented with functional studies involving the cross-reactive FI6 antibody. FI6 antibody binds near the conserved fusion subdomain of the HA molecule and thus interferes with the low pH-triggered conformational change of HA. Additional methods employed in this study, such as limited proteolysis, electron microscopy, biolayer interferometry and MDCK1 cell infection, give insight into the mechanism of FI6 antibody-mediated neutralization, and highlight the differences in infectivity of H1N1 and H3N2 viruses neutralized by the FI6 antibody.

Table of contents

Declaration 2

Abstract 4

Table of contents 5

List of figures 8

List of tables 10

List of abbreviations 11

1. Introduction 14
1.1 Influenza Viruses 14
1.1.1 Initial isolation 14
1.1.2 Classification and nomenclature of Influenza Viruses 15
1.1.3 Influenza pandemics and epidemics 16
1.2 Virus morphology and structure 18
1.3 Virus replication cycle 22
1.4 Influenza virus HA and its functions 24
1.4.1 Receptor-binding 26
1.4.2 Membrane fusion 28
1.4.3 Antigenic evolution 30
1.5 Conformational changes in HA leading to membrane fusion 32
1.6 Structural information on the low pH-induced structural refolding of HA obtained to date 35
1.6.1 Crystal structure of H3 HA0 precursor 35
1.6.2 Crystal structure of H3 BHA 35
1.6.3 Crystal structure of postfusion viral H3 TBHA2 38
1.6.4 Crystal structure of *E.coli*-expressed H3 HA2 40
1.6.5 Crystal structure of *E.coli*-expressed Influenza B HA2 44
1.7 Group-specific structural features of HA 46
1.8 Anti-HA antibodies 51
1.8.1 Antibodies blocking receptor-binding 51
1.8.2 Antibodies interfering with the conformational change of HA 52
1.9 Current treatments 54
1.10 Objectives for thesis 55

2 Materials and Methods 56
2.1 Materials 56
2.1.1 Virus strains 56
2.1.2 Antibodies 56
2.1.3 Reagents 56
2.1.4 Cell lines 56
2.2 Protein expression 57
2.2.1 Viral H1 HA2 in a postfusion conformation 57
2.2.1.1 Virus infection and growth in embryonated eggs 57
2.2.1.2 Virus purification 58
2.2.1.3 Release of HA from purified virus with bromelain 59
2.2.1.4 Detergent extraction 60
2.2.2 Recombinant H1 HA2 in a low pH conformation 61
2.3 Protein preparation 63
2.3.1 Quantification of viruses using SDS-PAGE 63
2.3.2 Determination of protein concentration 65
2.3.3 SDS-PAGE 65

2.3.4 Ion-exchange chromatography 65
2.3.5 Size-exclusion chromatography 66
2.3.6 Low pH treatment 67
2.3.7 Limited proteolysis 68
2.3.8 Immobilized metal affinity chromatography (IMAC) 70
2.3.9 Protein crystallisation and crystal freezing 72
2.4 Analytical techniques 75
2.4.1 Haemagglutination assay 75
2.4.2 Dynamic Light Scattering (DLS) 75
2.4.3 Circular dichroism (CD) spectroscopy 77
2.4.4 N-terminal sequencing 78
2.4.5 Electron microscopy (EM) 78
2.4.6 Biolayer interferometry 80
2.4.7 MDCK1 cell infection 81
2.4.7.1 Estimation of the amount of infectious virus in the prepared samples using the plaque assay 81
2.4.7.2 Infection of MDCK1 cells in the presence of FI6 and Hc31 antibodies 85

3 Crystal structure of viral H1 TBHA2 86
3.1 Virus growth and purification 86
3.2 Isolation of H1 HA 86
3.3 Induction of H1 BHA conformational change by incubation at low pH 89
3.4 Removal of HA1 domains by limited proteolysis 95
3.5 Removal of the fusion peptide by limited proteolysis 97
3.6 Crystallisation 101
3.7 Data collection and processing 105
3.8 Crystal structure of B59 H1 TBHA2 113

4 Crystal structure of recombinant H1 HA2 115
4.1 Protein expression and purification 115
4.2 Conformational assessment of recombinantly-expressed H1 HA2 118
4.3 Comparison of recombinantly-expressed postfusion H1 HA2 and H3 HA2 using CD 122
4.3.1 Secondary structure 122
4.3.2 Thermal stability and pH dependence of unfolding transitions 122
4.4 Crystallisation 123
4.5 Data collection and processing 124
4.6 Crystal structure of Bac B59 H1 HA2 131

5 Postfusion H1 HA2 and its comparison to other HA2 structures 133
5.1 Viral and recombinantly-expressed H1 HA2 133
5.2 Postfusion viral H1 and H3 TBHA2 137
5.3 Recombinantly-expressed H1 and H3 HA2 152
5.4 Recombinantly-expressed H1 HA2 and Influenza B HA2 156

6 Characterisation of FI6 antibody binding to the two groups HA 159
6.1 Binding of the FI6 antibody to virus-bound HA 159
6.1.1 Assessment of the ability of the FI6 antibody to prevent the conformational change of virus-bound HA by limited proteolysis 160
6.1.2 Binding of the FI6 antibody to virus-bound HA by electron microscopy 162
6.2 Binding of the FI6 antibody to THA by electron microscopy 164
6.3 Binding of viruses to human and avian receptor analogs in the presence of FI6 by biolayer interferometry 166
6.4 Determination of infectivity of H1N1 and H3N2 viruses neutralized by the FI6 antibody 167

7 Discussion 169

7.1 Purpose of study 169
7.2 Crystal structure of B59 H1 TBHA2 170
7.3 Analysis of H1 TBHA2 fragment 172
7.4 Crystal structure of Bac B59 H1 HA2 173
7.5 Viral and recombinantly-expressed H1 HA2 174
7.6 Comparison of postfusion H3 and H1 HA2 176
7.7 Final conclusions in relation to the low pH- induced refolding of H3 and H1 HA2 183
7.8 Comparison of postfusion H1 HA2 and Influenza B HA2 185
7.9 Implications for viral membrane fusion 185
7.10 Functional studies using the FI6 antibody 187
7.11 Future work 188

Bibliography 189

List of figures

Figure 1.1 An influenza A virion. 20
Figure 1.2 Electron micrograph of X31 (H3N2) virus. 21
Figure 1.3 Schematic diagram of influenza virus cycle. 23
Figure 1.4 Human and Avian HA receptor-binding sites. 27
Figure 1.5 Hypothetical mechanism of HA-mediated viral membrane fusion. 29
Figure 1.6 Antigenic sites A-E on the surface of HA1 domains of H3 HA. 31
Figure 1.7 The low pH-induced conformational change of HA. 34
Figure 1.8 Crystal structures of the precursor HA0 and cleaved BHA 37
Figure 1.9 Crystal structures of prefusion and postfusion H3 HA 39
Figure 1.10 A trimer of *E.coli*-expressed H3 HA2 42
Figure 1.11 Schematic of H3 HA molecule and the crystallized fragments 43
Figure 1.12 Structural alignment of X31 H3 EHA2 and influenza B HA2 45
Figure 1.13 Sequence alignment of group 1 and group 2 HA in the region of HA2 48
Figure 1.14 Group-specific structures of loop B 49
Figure 1.15 Rotation of HA1 domains relative to the HA2 stem of HA 50
Figure 1.16 The FI6 antibody bound to a trimer of prefusion H3 HA 53
Figure 2.1 Bac B59 H1 HA2 expression construct 62
Figure 2.2 Quantification of viruses using SDS-PAGE. 64
Figure 2.3 Protein solubility phase diagram for a given temperature 73
Figure 2.4 DLS profile for purified B59 H1 TTHA2 76
Figure 3.1 Purification of H1 BHA by sucrose density gradient centrifugation. 87
Figure 3.2 Release of viral H1 HA with BOG 88
Figure 3.3 Induction of the conformational change of influenza virus HA 90
Figure 3.4 Assessment of a conformational state of low pH-treated H1 BHA by limited proteolysis 92
Figure 3.5 Identification of trypsin digestion products of low pH-treated H1 BHA by N-terminal sequencing 94
Figure 3.6 Purification of H1 THA2 aggregates 96
Figure 3.7 Removal of the fusion peptide from H1 rosettes with trypsin 98
Figure 3.8 Removal of the fusion peptide from H1 rosettes with thermolysin 98
Figure 3.9 Removal of the hydrophobic fusion peptide using Proti-Ace proteases 100
Figure 3.10 Removal of the hydrophobic fusion peptide using other enzymes. 100
Figure 3.11 Crystals of PR8 H1 TBHA2 102
Figure 3.12 Crystals of B59 H1 TTHA2 and B59 H1 TBHA2 104
Figure 3.13 Anisotropy analysis of dataset B 106
Figure 3.14 Analysis of I/sigma and Mean (I/sd) versus resolution for dataset B 106
Figure 3.15 Anisotropy analysis of dataset A 107
Figure 3.16 Analysis of I/sigma and Mean (I/sd) versus resolution for dataset A 107
Figure 3.17 Truncations of a search model for molecular replacement 110
Figure 3.18 Ramachandran plot for B59 H1 TBHA2 model 112
Figure 3.19 Crystal structure of B59 H1 TBHA2 114
Figure 4.1 Purification of Bac B59 H1 HA2 116
Figure 4.2 Chromatogram from final size-exclusion chromatography of Bac B59 H1 HA2 117
Figure 4.3 Assessment of the conformational state of Bac B59 H1 HA2 by limited proteolysis 118
Figure 4.4 Far UV CD spectrum of Bac B59 H1 HA2 119
Figure 4.5 Thermal denaturation of Bac B59 H1 HA2 121
Figure 4.6 Crystals of Bac B59 H1 HA2 123
Figure 4.7 Resolution estimate for data collected from a crystal of Bac B59 H1 HA2 125

Figure 4.8 Anisotropy analysis of the 3.4 Å dataset obtained from a crystal of Bac B59 H1 HA2 125
Figure 4.9 Distribution of Rmerge over 1834 frames collected at 0.15°/frame, 0.1s/frame exposure and λ=0.97625 from a crystal of Bac B59 H1 HA2 126
Figure 4.10 Data completeness versus the 1834 frames collected at 0.15°/frame, 0.1s/frame exposure and λ=0.97625 from a crystal of Bac B59 H1 HA2 126
Figure 4.11 Model of postfusion H1 HA2 used as a search model for molecular replacement 128
Figure 4.12 Ramachandran plot for the model of Bac B59 H1 HA2 130
Figure 4.13 Crystal structure of Bac B59 H1 HA2 132
Figure 5.1 A model of postfusion influenza H1 HA2..... 134
Figure 5.2 The low pH-induced refolding of H1 HA2..... 136
Figure 5.3 Structural alignment of prefusion H3 and H1 HA2..... 139
Figure 5.4 Structural alignment of postfusion X31 H3 TBHA2 and B59 H1 TBHA2 140
Figure 5.5 Comparison of turns in X31 H3 TBHA2 and B59 H1 TBHA2 142
Figure 5.6 Loops at residues 130-145 of X31 H3 TBHA2 and B59 H1 TBHA2..... 144
Figure 5.7 Coiled coil prediction profiles for H1 and H3 HA2..... 146
Figure 5.8 Comparison of the H1 and H3 sequences in the region of HA2 residues 38-86 in terms of coiled coil forming potential 148
Figure 5.9 Helical wheel representations of H3 and H1 sequences the region of HA2 residues 38-64 149
Figure 5.10 X31 H3 TBHA2, X31 H3 EBHA2 and B59 H1 TBHA2 proteolytic digestion products 151
Figure 5.11 Recombinantly-expressed X31 H3 EHA2 and Bac B59 H1 HA2 154
Figure 5.12 A symmetry generated trimer of Bac B59 H1 HA2 aligned to a trimer of X31 H3 EHA2 155
Figure 5.13 Comparison of recombinantly-expressed influenza B HA2 and Bac B59 H1 HA2 157
Figure 5.14 Comparison of turns at residues 106-112 of recombinantly-expressed influenza B HA2 and Bac B59 H1 HA2 158
Figure 6.1 Assessment of the ability of the FI6 antibody to prevent a conformational change of virus-bound HA by limited proteolysis 161
Figure 6.2 Electron microscopy of the FI6 antibody binding to the X31 virus..... 163
Figure 6.3 Electron microscopy of X31 H3 THA-FI6 antibody complexes..... 165
Figure 6.4 Binding of B59 H1N1 and X31 H3N2 virus-antibody complexes to sialic acid receptor analogs by biolayer interferometry..... 166
Figure 6.5 Quantification of viral infection in the presence of the Hc31 antibody 168
Figure 6.6 Quantification of viral infection in the presence of the FI6 antibody .. 168
Figure 7.1 Preparation of B59 H1 TBHA2 for crystallographic studies..... 171
Figure 7.2 Composite model of H1 HA2 in a postfusion conformation 175
Figure 7.3 Coiled coil prediction profiles for group 1 HA 178
Figure 7.4 Coiled coil prediction profiles for group 2 HA 179
Figure 7.5 Interhelical loops in prefusion H3 and H1 HA2 structures..... 181
Figure 7.6 Solvent accessible areas in prefusion H3 and H1 HA..... 182
Figure 7.7 Differences in packing of helices E against central helices C in postfusion H3 and H1 HA2..... 184

List of tables

Table 2.1 Summary of a DLS measurement using B59 H1 TTHA2. 76
Table 2.2 Composition of media and reagents used in the plaque assay..... 83
Table 2.3 Immunostaining for influenza NP. 83
Table 3.1 Scores for possible Laue groups for the B59 H1 TBHA2 crystal 108
Table 3.2 Number of B59 H1 TBHA2 monomers in the asymmetric unit of the P321 (150) unit cell and the corresponding % solvent content..... 108
Table 3.3 Data collection and refinement statistics for B59 H1 TBHA2..... 111
Table 4.1 Secondary structure content predictions for Bac B59 H1 HA2 at pH 5.2 by CD spectroscopy at λ=195-260 nm. 120
Table 4.2 Secondary structure content predictions for Bac B59 H1 HA2 at pH 7.2 by CD spectroscopy at λ=195-260 nm. 120
Table 4.3 Melting temperatures for X31 H3 EBHA2, X31 H3 TBHA2 and Bac B59 H1 HA2 by CD spectroscopy at λ= 222 nm..... 122
Table 4.4 Scores for H-3 m Laue group for a crystal of Bac B59 H1 HA2 calculated using Pointless (CCP4). 127
Table 4.5 Number of Bac B59 H1 HA2 monomers in the asymmetric unit of the R 3 2 (155) unit cell and the corresponding % solvent content. 127
Table 4.6 Data collection and refinement statistics for Bac B59 H1 HA2..... 129
Table 5.1 Refolding of H1 HA2 in response to low pH. 135
Table 5.2 Structural comparison of B59 H1 TBHA2 and Bac B59 H1 HA2..... 135
Table 5.3 Refolding of H3 HA2 in response to low pH. 141
Table 5.4 Structural comparison of prefusion H3 HA2 (PDB: 1HGF) and H1 HA2 (PDB: 1RUZ)..... 141
Table 5.5 Structural comparison of X31 H3 TBHA2 (PDB: 1HTM) and B59 H1 TBHA2 141
Table 5.6 Structural comparison of prefusion H3 and H1 HA2 (PDB: 1HGF and 1RUZ), and postfusion X31 H3 EHA2 (PDB: 1QU1) and Bac B59 H1 HA2 .. 153
Table 5.7 Structural comparison of recombinantly-expressed influenza B (PDB: 4NKJ) and Bac B59 H1 HA2 156

List of abbreviations

BHA	Bromelain-released HA
BHA2	Low pH-activated BHA aggregates
Bicine	N, N-Bis (2-hydroxyethyl) glycine
Bis-tris	2,2-Bis (hydroxymethyl)-2, 2′, 2″-nitrilotriethanol
BME	2-Mercaptoethanol
BOG	N-octyl-ß-D-glucoside
BSA	Bovine serum albumin
CD	Circular dichroism
DLS	Dynamic light scattering
DMEM	Dulbecco's modified eagle's medium
E.coli	*Escherichia coli*
EBHA2	*E.coli*-expressed H3 HA2 (residues 38-175)
EDTA	Ethylenediaminetetraacetic acid
EG	Ethylene glycol
EHA2	*E.coli*-expressed H3 HA2 (residues 23-185)
EM	Electron microscopy
F subdomain	Fusion subdomain of HA2
FCS	Fetal calf serum
FP	Fusion peptide (HA2 residues 1-23)
HA	Influenza virus haemagglutinin
HA assay	Haemagglutination assay
HA titer	Haemagglutinin titer
HA0	Uncleaved precursor haemagglutinin
HA1	Receptor-binding domain of HA molecule
HA2	Fusion domain of HA molecule
HAU	HA units
Hepes	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HI assay	Haemagglutination inhibition assay
IGG	Immunoglobulin G
IMAC	Immobilized metal affinity chromatography
kDa	Kilo Dalton
LAIV	Live attenuated influenza vaccine
LDAO	Lauryldimethylamine-oxide
M1	Influenza virus matrix protein
M2	Influenza virus M2 ion channel protein
MDCK cells	Madin-Darby canine kidney cells
MEM	Minimum essential medium
MES	2-(N-morpholino) ethanesulfonic acid
MME	Monomethyl ether
MMS	Microseed matrix screening
MOI	Multiplicity of infection
mRNA	Messenger RNA
Mw	Molecular weight
MWCO	Molecular weight cut-off
NA	Influenza virus neuraminidase
NIMR	National Institute for Medical Research

NMR	Nuclear magnetic resonance
NP	Influenza virus nucleoprotein
NR	Non-reducing
NS-1	Influenza virus non-structural protein 1
NS-2	Influenza virus non-structural protein 2
O/N	Overnight
PB1	Influenza virus polymerase basic protein 1
PB1- F2	Influenza virus polymerase basic protein 1, frame 2
PB2	Influenza virus polymerase basic protein 2
PBS	Phosphate buffered saline
PDB	Protein Data Bank
PDI	Polydispersity index
PEG	Polyethylene glycol
PFU	Plaque forming unit
pI	Isoelectric point
R	Reducing
RBC	Red blood cells
Rmsd	Root-mean-square deviation
RNA	Ribonucleic acid
Rpm	Revolutions per minute
RT	Room temperature
SDS	Sodium dodecyl sulphate
SDS-PAGE	SDS polyacrylamide gel electrophoresis
Sf9 cells	Clonal isolate of *Spodoptera frugiperda* Sf21 cells
Sia	Sialic acid
ssRNA	Single-stranded RNA
SST	Sodium silicotungstate
TBHA2	Bromelain-released and thermolysin/ trypsin-solubilized HA2
TCPK treated	Treated with L-1-tosylamido-2-phenylethyl chloromethyl ketone
THA	Detergent-extracted HA molecule
THA2	Low pH-activated THA aggregates
Tm	Melting temperature
TM anchor	Transmembrane anchor (HA2 residues 185-211)
Tris	Tris [hydroxy methyl] aminomethane
TTHA2	Detergent-extracted and thermolysin/trypsin-solubilized HA2
V_m	Matthews coefficient
vRNA	Viral RNA
vRNP	Viral RNP complex
WHO	World Health Organization
WIC	World Influenza Centre
XDS	X-ray Detector Software

Viruses:

- **X31 -** A/Aichi/2/68 (H3N2)
- **B59 -** A/Brisbane/59/07 (H1N1)
- **PR8 -** A/Puerto Rico/8/34 (H1N1)
- **Tokyo -** A/Tokyo/3/67 (H2N2)

Throughout the text, protein samples were named depending on the source virus strain, and with respect to the expression and purification protocol used.

Viral protein samples:

- **B59 H1 TBHA2** - postfusion viral HA2 released from A/Brisbane/59/07 H1N1 influenza A virus by proteolytic digestion with bromelain, triggered to undergo a conformational change at low pH, and solubilized by digestion with trypsin & thermolysin. The obtained structure contains HA2 residues 65-155.
- **B59 H1 TTHA2** - postfusion viral HA2 released from A/Brisbane/59/07 H1N1 influenza A virus using detergent extraction, triggered to undergo a conformational change at low pH, and solubilized by digestion with trypsin.

Samples obtained from X31, PR8 and Tokyo strains prepared using the two methods, were named accordingly.

Recombinant protein samples:

- **Bac B59 H1 HA2** - a fragment of postfusion H1 HA2 (residues 38-175) from A/Brisbane/59/07 H1N1 influenza A virus, and expressed using Baculovirus/insect cell system for crystallographic studies.
- **X31 H3 EHA2** - a fragment of postfusion H3 HA2 (residues 23-185) from A/Aichi/2/68 H3N2 influenza A virus, expressed in E.coli (Chen, Skehel et al. 1999) (PDB: 1QU1).
- **X31 H3 EBHA2** - a fragment of postfusion H3 HA2 (residues 38-175) from A/Aichi/2/68 H3N2 influenza A virus, expressed in E.coli (Chen, Skehel et al. 1999) (PDB: 1QU1).

1. Introduction

1.1 Influenza Viruses

1.1.1 Initial isolation

Influenza is an acute respiratory disease described by Hippocrates in 412 BC. Its name originated in Italy in the 15th century, from an epidemic, associated with the seasonal movement of the sun and the stars. Scientific research into influenza was undertaken following the 1918 pandemic, which killed ~50 million people worldwide, and the virus was isolated by Wilson Smith, Christopher H Andrewes and Patrick Laidlaw at the MRC National Institute for Medical Research (NIMR), London, in 1933. The virus was extracted from ferrets, infected with filtered nasopharyngeal washings of an infected researcher (Smith 1933). In the research article published in *The Lancet* on the 8th July 1933, the disease-causing agent was identified as a virus, due to its small size, the observed transmissibility between ferrets, and neutralization of its infectivity using sera of recovered patients. In 1936 Burnet successfully grew influenza virus in embryonated hens' eggs, which advanced the development of the first influenza vaccine at NIMR in 1936, and the establishment of the World Influenza Centre (WIC) at NIMR, to study the antigenic drift of influenza strains worldwide. Influenza B and C viruses were isolated in 1940 and 1949, respectively (Francis 1940) (Taylor 1949).

1.1.2 Classification and nomenclature of Influenza Viruses

Influenza viruses are negative sense ssRNA viruses of the *Orthomyxoviridae*. Influenza A, B and C, are three out of six genera of this family of enveloped viruses. Influenza A and B viruses have eight RNA segments coding for viral proteins, while influenza C viruses have seven RNA segments in their genome. Influenza A viruses infect humans, other mammals and birds, and the natural reservoir are aquatic birds. Influenza B viruses infect humans (usually children) but have also been isolated from seals (Osterhaus, Rimmelzwaan et al. 2000). Viruses of this type are less diverse and the mutation rate has been shown to be 2-3 times slower than that of influenza A viruses (Nobusawa and Sato 2006). This low diversity combined with a limited host range lower the risk of a pandemic outbreak. Influenza C viruses infect humans, pigs and dogs, although human infections are rare.

Influenza A viruses have been divided into subtypes since 1980 (WHO 1980), and their classification is based on the antigenic properties of the two glycoproteins present on the virus surface. These antigens, haemagglutinin (HA) and neuraminidase (NA), undergo an independent antigenic evolution, and to date, 16 HA (H1-H16) and 9 NA (N1-N9) antigenic subtypes have been identified (Schild, Newman et al. 1980) (Webster, Bean et al. 1992) (Fouchier, Munster et al. 2005) (Skehel and Wiley 2000) (Klenk, Garten et al. 2011) (Tong, Li et al. 2012) (Tong, Zhu et al. 2013). Classification of viruses into different subtypes relies on their reactivity in double immunodiffusion assays (Schild, Newman et al. 1980) (Hinshaw, Air et al. 1982) (Kawaoka, Yamnikova et al. 1990).

Genetic reassortment between human, swine and avian virus strains can result in a formation of new HA subtypes, characterized by different antigenic properties. Subtypes H1, H2 and H3 have established in humans, subtypes H3 and H7 infect equines, with H1, H2, H3 and H9 specifically infecting swine. Viruses within the subtype are referred to as virus strains.

The identification system currently in place includes virus type, the host of origin (except for human strains), followed by the place of first isolation, lineage number, and finally, the year of isolation. Additionally, HA and NA antigenic subtypes are specified e.g. A/Brisbane/59/2007 (H1N1).

1.1.3 Influenza pandemics and epidemics

Influenza viruses evolve due to genetic mutation and reassortment (Hay, Gregory et al. 2001). Point mutations in RNA segments coding for surface glycoproteins, haemagglutinin (HA) (segment 4) and neuraminidase (NA) (segment 6), are frequent. This continuous genetic evolution through the processes of antigenic drift and shift, allows for new antigenic subtypes to emerge, and survive by escaping host immunity (Zambon 1999).

Gradual accumulation of these point mutations is known as antigenic drift. Apart from the selective immune pressure, the factor responsible for this progressive change is the lack of proofreading ability of viral RNA polymerase (Domingo, Sabo et al. 1978). A genetically changed virus can cause an epidemic in a human population, and antigenic drift is observed in Influenza A and B viruses, giving rise to new virus strains within a subtype.

Genetic reassortment between strains of human, swine and avian origin can result in a significant antigenic change (occurring every 10-40 years). Such antigenic shift has only been observed in influenza A viruses and is promoted by their segmented genome. Viruses from different hosts can exchange genes in pigs, due to the presence of both α- (2,6)- and α- (2,3)- linked sialic acids, allowing for their co-infection with human and avian virus strains at the same time. Genetic reassortment can lead to a dramatic change in the antigenic properties of one, or both glycoproteins (HA and NA) on virus surface, and subsequently result in a pandemic (Itoh, Shinya et al. 2009). A good example of a major antigenic shift is the emergence of H3N2 strain in 1968 in Hong Kong. This virus replaced the H2N2 strain circulating between 1957 and 1967, and caused a pandemic.

The deadliest pandemic recorded to date occurred in 1918 (known as the 'Spanish flu'), and haemagglutinin on the virus surface was of subtype H1 (Gamblin, Haire et al. 2004). The 1918 pandemic was characterized by unusually high mortality rate among young adults (Simonsen, Clarke et al. 1998). The HA gene was reassembled from RNA fragments present in formalin-fixed lung tissues, and crystal structure of the uncleaved precursor of the extinct 1918 pandemic virus HA was obtained at 3 Å resolution, revealing a narrow, avian-like receptor-binding site (Gamblin, Haire et al. 2004). In 1957, H2N2 virus killed up to ~1.5 million people, in a pandemic known as 'Asian flu', and the virus was formed by a genetic reassortment between human and avian viruses, with both glycoproteins (HA and NA) of avian origin (Scholtissek, Rohde et al. 1978) (Kawaoka, Krauss et al. 1989).

The third pandemic occurred in 1968, when H3N2 virus in Hong Kong replaced the H2N2 virus. In this case, only HA glycoprotein was replaced (Dowdle, Coleman et al. 1969). The H3N2 pandemic virus was shown to contain 6 RNA segments of the H2N2 virus, and two segments (coding for HA and PB1) from an avian virus (Scholtissek, Rohde et al. 1978) (Kawaoka, Krauss et al. 1989), thought to be related to an avian A/duck/Ukraine/63 virus (Ha, Stevens et al. 2003).

Human 1918-derived H1N1 virus reappeared in Russia in 1977, and then again in Mexico in 2009, and the virus it thought to have originated in swine (Smith, Vijaykrishna et al. 2009). The 2009 H1N1 virus was a reassortant of a European avian-like virus, and a North American swine H1N2 triple-reassortant (Dunham, Dugan et al. 2009), and the resulting virus was characterised by a higher transmissibility than seasonal flu (Fraser, Donnelly et al. 2009).

Human viruses preferentially bind to α- (2,6) sialic acid receptors, found on cells of the upper respiratory tract, and avian viruses bind to α- (2,3) sialic acid receptors in birds' intestine, and human lower respiratory tract (Rogers and Paulson 1983) (Matrosovich, Zhou et al. 1999) (van Riel, Munster et al. 2006) (Gambaryan, Tuzikov et al. 1997). Avian and swine viruses containing haemagglutinin of subtypes H5, H6, H7, H9 and H10 can cross the species barrier by acquiring the ability to bind to the human receptor (Ito, Couceiro et al. 1998), and the acquired ability of these subtypes to transmit efficiently between humans would pose a serious pandemic threat (Ito 2000) (Shortridge, Gao et al. 2000). In addition to receptor-binding specificity, one factor limiting human influenza virus replication to the respiratory tract is the availability of proteases, which are needed for cleavage and activation of HA molecules on the surface of newly formed viral particles. HA of the highly pathogenic avian strain A/H5N1 however, is efficiently cleaved in the cell, by ubiquitous subtilisin-like proteases. These enzymes can be found in the Golgi apparatus of most cells, allowing replication of avian A/H5N1 virus outside of the human respiratory tract (Guo, Li et al. 2008). Human infection with H5N1 was first reported in Hong Kong in 1997, and its antigenic evolution potentially leading to human-to-human transmissibility, may result in a future pandemic (Smith, Fan et al. 2006) (Xiong, Xiao et al. 2014). This would require a change in its binding preference, from avian to human receptor (Chen, Blixt et al. 2012) (Herfst, Schrauwen et al. 2012) (Imai, Watanabe et al. 2012) (Xiong, Coombs et al. 2013) (Zhang, Shi et al. 2013) (Xiong, Xiao et al. 2014). In contrast, the H7N9 virus was found to be able to transmit via aerosols without further adaptation (Belser, Gustin et al. 2013). In 2013, an avian H10N8 virus manifested in China (To, Tsang et al. 2014), and its characteristic feature was high affinity for human

receptors. The structural conformation of the bound receptor was shown to be similar to that of 1918 H1N1 pandemic virus, and to a human H7 virus (Vachieri, Xiong et al. 2014).

1.2 Virus morphology and structure

Spherical virus particles measure ~100 nm in diameter, and filamentous particles can measure up to 1 µm (Calder, Wasilewski et al. 2010) (Lamb RA 2001). Viral core, containing eight ssRNA segments (McGeoch, Fellner et al. 1976), is enclosed in a lipid envelope, and this layer is derived from host cell membrane and formed during budding (Kates, Allison et al. 1962). The segmented genome exists as a structure called viral RNP (vRNP) complex (Compans, Content et al. 1972). The helical RNA segments code for viral proteins (PB2, PB1, PA, HA, NP, NA, M1, M2, NS1, NS2 and PB1-F2), and are associated with the trimeric polymerase complex, composed of polymerase basic protein 1 (PB1), polymerase basic protein 2 (PB2) and polymerase acidic protein (PA) (Area, Martin-Benito et al. 2004). The core of the polymerase complex is formed by PB1, and each segment is associated with multiple NPs. The ribonucleoprotein core is surrounded by a shell made of matrix protein (M1), and located directly beneath the lipid envelope (Schulze 1970) (Ruigrok, Barge et al. 2000). Viral non-structural protein 1 (NS1) is responsible for inhibition of the host antiviral response, for example by suppressing maturation and translation of interferon mRNA (Kochs, Garcia-Sastre et al. 2007), and therefore promoting viral replication. Non-structural protein 2 (NS2), also known as nuclear export protein (NEP), mediates export of viral ssRNA from the nucleus (O'Neill, Talon et al. 1998). Polymerase basic protein 1 frame 2 (PB1-F2) is responsible for virus-induced cell death by affecting the mitochondrial function of infected cells (Gibbs, Malide et al. 2003), and has been implicated in virus pathogenicity (Zamarin, Ortigoza et al. 2006).

Two types of glycoprotein spikes are anchored in the lipid envelope (Laver and Valentine 1969), and these are rod-shaped haemagglutinin (HA) trimers (~350-400 copies per virion), and mushroom-shaped neuraminidase (NA) tetramers (~50 copies per virion) (Murti and Webster 1986). While NA releases newly formed virus particles from infected cells by cleaving sialic acid residues from the surface of the cell and the newly formed viral particles, HA is involved in the initial binding of the virus to the target cell, and later for fusion of viral and host cell membranes,

eventually leading to the release of the viral genome into the host cytoplasm and its replication (Skehel and Wiley 2000).

M2 is another integral membrane protein (~14-68 copies per virion) (Zebedee and Lamb 1988). It is a homotetrameric, low pH-activated, proton-selective channel, whose transmembrane domain forms a pore (Sugrue and Hay 1991). The main function of M2 is the control of pH across the viral membrane. Opening of the channel, and the resulting protonation of virus interior, are associated with protonation of highly conserved His37 residues of the M2 protein (Pinto, Holsinger et al. 1992). This acidification eventually leads to dissociation of M1 and uncoating of vRNP complexes prior to membrane fusion. M2 is also involved in maintaining a metastable state of newly produced HA molecules, by pumping protons out of the acidic trans-Golgi-network, where the translation takes place, and the residue responsible for forming this protective 'channel gate' is highly conserved Trp41.

A schematic representation of influenza A virion and an electron micrograph of an influenza A H3N2 virus are shown in Figures 1.1 and 1.2, respectively.

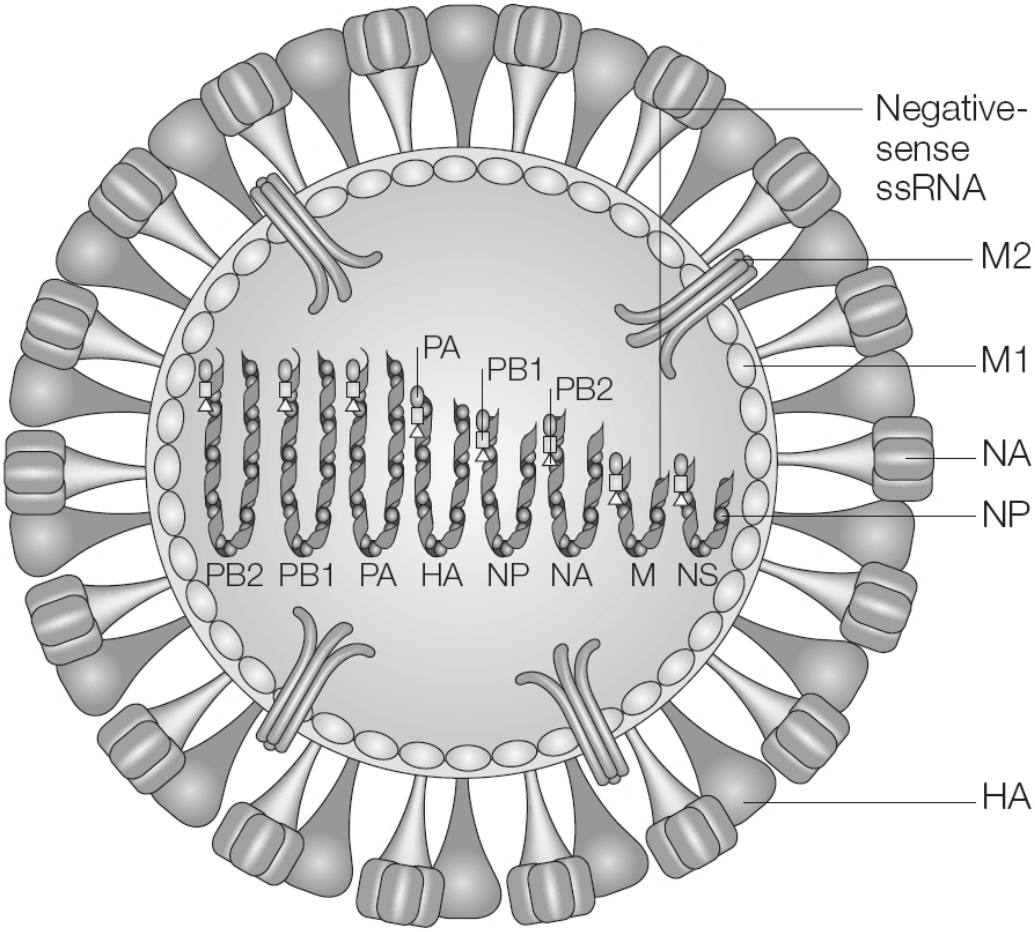


Figure 1.1 An influenza A virion. Haemagglutinin (HA), neuraminidase (NA) and the M2 ion- channel protein anchored in the viral envelope are indicated. M1 is a matrix protein. RNA segments coding for viral proteins are shown. Together with viral nucleoprotein (NP) and polymerase proteins (PA, PB1, and PB2) form a ribonucleoprotein complex. Figure adapted from (Horimoto and Kawaoka 2005).

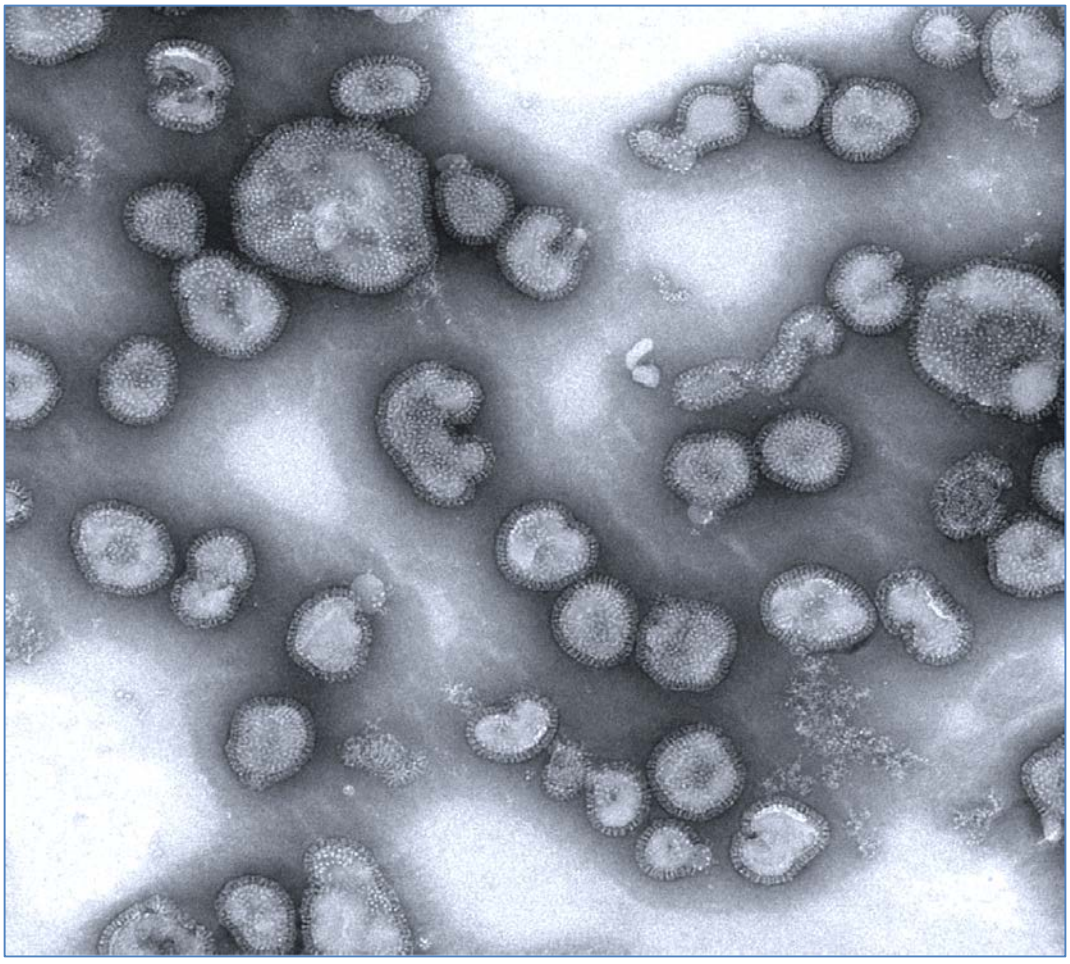


Figure 1.2 Electron micrograph of X31 (H3N2) virus. Negative stain; magnification 30 000 X. Image obtained from Dr Lesley Calder.

1.3 Virus replication cycle

Influenza virus is transmitted via aerosols. An initial step of viral infection is HA-mediated attachment of the virus to cell surface glycan-containing terminal sialic acid (Gottschalk 1959), and host specificity is determined by the type of its linkage to galactose: α- (2, 3) (in birds) or α- (2, 6) (in humans) (Rogers and Paulson 1983). The attached viruses are then taken up into the cells of the respiratory tract by endocytosis (Rossman, Leser et al. 2012), while opening of the M2 proton channel at the pH of endosomes causes acidification of their interior (Sugrue and Hay 1991), and eventually to dissociation of M1 and uncoating of vRNP. Additionally, the acidic pH causes HA on the virus surface to undergo an extensive structural rearrangement, which eventually leads to fusion of viral and endosomal membranes (Skehel, Bayley et al. 1982). The released vRNP are moved to the nucleus, where negative sense viral RNA are initially transcribed to messenger RNA (mRNA), and then to complementary RNA (cRNA), by proteins forming the polymerase complex (PA, PB1 and PB2) (Hay, Lomniczi et al. 1977). While vRNP are formed in the nucleus, proteins associated with viral envelope (HA, NA and M2) are directed to the endoplasmic reticulum (ER) (Compans 1973) (Hay 1974), and inserted into the ER membrane (McCauley, Bye et al. 1979) (Bos, Davis et al. 1984), where they are folded, oligomerized and glycosylated. All viral components are then transported to the budding site on the plasma membrane, where the segmented genome and viral envelope are assembled. The newly formed viruses are then released by the activity of viral NA, which is essential for a successful replication (Palese, Tobita et al. 1974). A schematic diagram of influenza virus replication cycle is shown in Figure 1.3.

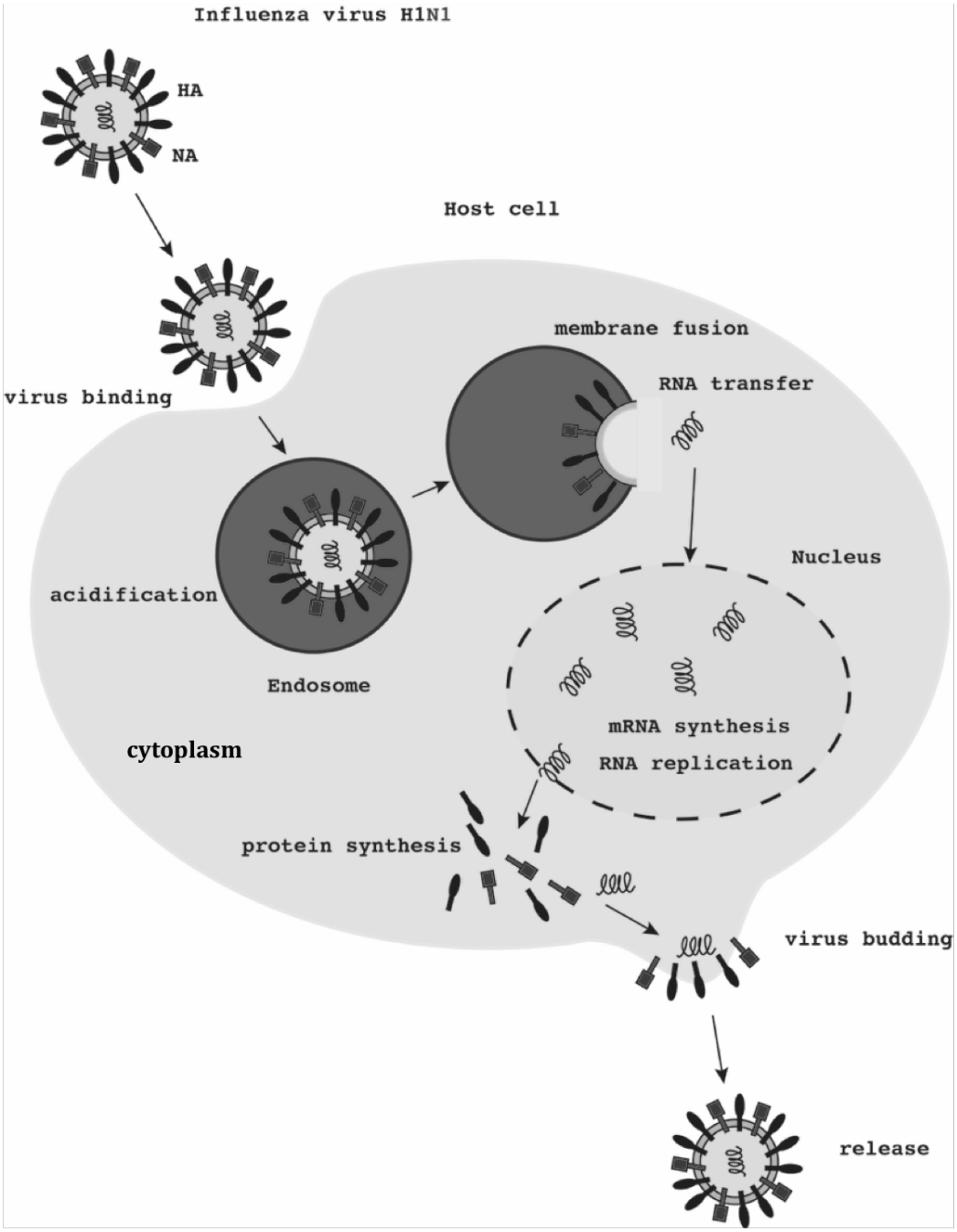


Figure 1.3 Schematic diagram of influenza virus cycle. Viral particles are internalized via receptor-mediated endocytosis. The uncoated vRNP are released into the cytoplasm of infected cell in the process of HA-mediated viral membrane fusion. The released genome is transcribed in the nucleus. Viral proteins are produced in the cytoplasm, and transported to the cell membrane, where virus budding takes place.

1.4 Influenza virus HA and its functions

Influenza A virus haemagglutinin (HA) is encoded by RNA segment 4, and named according to its ability to agglutinate erythrocytes (Hirst 1941). This antigenic integral envelope glycoprotein amounts to ~30% of the total virus protein (Lamb and Choppin 1983), and the number of HA copies on each viral particle is ~400 (Imai, Mizuno et al. 2006). The molecule is involved in initial binding of influenza virus to the target cell (Hirst 1942), and for viral membrane fusion, necessary for entry of the viral genome into the cell (Skehel and Wiley 2000) (Gamblin, Haire et al. 2004). Each monomer of this ~220 kDa homotrimeric molecule (Wiley, Skehel et al. 1977) consists of two domains, HA1 (~56 kDa) and HA2 (~26 kDa), which are disulphide-linked (Skehel and Schild 1971) (Laver 1971). HA1 domains form globular heads of the molecule (Wilson, Skehel et al. 1981), and can be further subdivided into receptor-binding (R) (HA1 residues 115-261) and vestigial esterase (E) (HA1 residues 50-114) subdomains (Skehel and Wiley 2000). Vestigial (inactive) subdomain of influenza A HA is a fragment of a single glycoprotein HEF present in influenza C (Rosenthal, Zhang et al. 1998). Fusion subdomain is composed mainly of HA2 plus HA1 residues 1-49 (F') and 269-309 (F). HA molecule is synthetized as a single precursor polypeptide (HA0) on the rough endoplasmic reticulum (RER) (Compans 1973), to which it is navigated via an N-terminal signalling sequence (McCauley, Bye et al. 1979). Following removal of a 17 amino acid long signal peptide, HA is co-translationally glycosylated and cleaved by trypsin-like proteases present in the respiratory tract (Lazarowitz, Compans et al. 1971) (Klenk, Scholtissek et al. 1972) (Lazarowitz and Choppin 1975). Glycosylation is important for the correct folding and oligomerization of HA trimer (Gallagher, Henneberry et al. 1992) (Chen, Helenius et al. 1995), which is essential for its transition from the rough endoplasmic reticulum to the Golgi network (Gething, McCammon et al. 1986) (Copeland, Zimmer et al. 1988), where further glycosylation takes place. In the process of antigenic evolution, HA can acquire new glycosylation sites, majority of which are located in globular head domains (HA1) (Skehel and Wiley 2000). These carbohydrates mask the antigenic sites from recognition by antibodies (Skehel, Stevens et al. 1984), and can affect binding of HA to the sialic acid receptors, by overlapping with the receptor-binding site (Aytay and Schulze 1991) (Mir-Shekari, Ashford et al. 1997).

Activation of HA by proteolytic cleavage takes place on the surface of the cell, or after virus budding (Compans, Klenk et al. 1970, Schulze 1970), and is performed

by trypsin-like proteases (Klenk, Rott et al. 1975) (Lazarowitz and Choppin 1975) or ubiquitous subtilisin-like proteases like furin, which recognize a sequence of polybasic residues (Arg-X-Arg/Lys-Arg) (Garten, Bosch et al. 1981) (Garten, Hallenberger et al. 1994) (Stieneke-Grober, Vey et al. 1992) found at the cleavage site of highly pathogenic H5 and H7 viruses (Webster and Rott 1987) (Perdue, Garcia et al. 1997) (Guo, Li et al. 2008). Proteolytic cleavage primes HA for fusion in the next cycle of infection, and produces a new N-terminus of HA2, known as the fusion peptide (FP), which is highly conserved among influenza strains and rich in hydrophobic residues (Phe, Leu, Ile and Val) (Skehel and Waterfield 1975) (Skehel and Wiley 2000). The fusion peptide (HA2 residues 1-23) contains large hydrophobic residues interspersed with glycine, and is essential for a successful membrane fusion (Gething, White et al. 1978) (Chen, Lee et al. 1998). Deletion of Gly1 (Garten, Bosch et al. 1981) (Cross, Wharton et al. 2001), or its substitution with Val or Glu (Qiao, Armstrong et al. 1999) have been previously shown to block membrane fusion. Cleavage of HA into HA1 and HA2 subunits is essential for virus infectivity (Klenk, Rott et al. 1975), and for human H1, H2 and H3 subtype viruses the cleavage point is a single Arg residue at position 329 of HA2 (Skehel and Wiley 2000). The conserved glycine, adjacent to Arg329, becomes the new N-terminus of HA2 (Skehel, Waterfield et al. 1980) (Steinhauer 1999), and residue 328 of HA0 becomes the new C-terminus of HA1. The C-terminal region of HA2 contains a 27-amino acid transmembrane domain (TM) and a 10-amino acid cytoplasmic tail.

1.4.1 Receptor-binding

One of the main functions of influenza virus HA is binding of the virus to sialic acid receptors on the target cell (Gottschalk 1959). Sialic acids belong to a class of nine-carbon acidic amino monosaccharides, found at the termini of N-linked carbohydrate side chains of cellular glycoproteins and glycolipids of all species (Castells, Ballesta et al. 1990). Differences in the receptor specificity of HA molecules from human and animal viruses were discovered in experiments with cells expressing a particular type of sialic acid and red blood cells (Rogers and Paulson 1983). The avian receptor found on the intestinal epithelium of avian species contains sialic acid in α-(2,3)- linkage to galactose (Ito, Suzuki et al. 2000), while sialic acids of the human upper respiratory tract receptor are linked to galactose via α-(2,6)- linkage (Rogers and Paulson 1983) (Baum and Paulson 1990) (Couceiro, Paulson et al. 1993) (Connor, Kawaoka et al. 1994) (Skehel and Wiley 2000) (Matrosovich, Matrosovich et al. 2004). Equine viruses recognize α-(2,3)-linkage, while swine viruses bind to sialic acid in both, α-(2,3)- and α-(2,6)- linkages to galactose (Rogers and D'Souza 1989), both found in porcine tracheae (Ito, Couceiro et al. 1998).

Receptor-binding site is located at the tip of each HA monomer (Rogers, Paulson et al. 1983), in a cavity surrounded by 190 helix (residues 189-199), 220 loop (residues 224-228) and 130 loop (residues 133-138) of HA1 head domains (Weis, Brown et al. 1988) (Sauter, Bednarski et al. 1989) (Skehel and Wiley 2000), and conformations of these loops vary between subtypes (Gamblin, Haire et al. 2004). The bottom of this shallow cavity is formed by conserved aromatic residues, which include Tyr98, Ser136, Trp153, His183 and Tyr195. Binding of sialic acid occurs via formation of hydrophobic interactions and hydrogen bonds between 130- and 220- loops, and the conserved amino acids found at the base of the receptor-binding site (Gamblin and Skehel 2010). The superposed human and avian H2 receptor-binding sites are shown in Figure 1.4. Several variable residues surrounding receptor-binding site have been implicated in the receptor-binding specificity. For H2 and H3 subtypes, Gln226Leu and Gly228Ser mutations cause the site to open, and the resulting rearrangement of residues within the receptor-binding site can cause a switch in specificity from avian to human receptors (Rogers, Paulson et al. 1983) (Connor, Kawaoka et al. 1994) (Matrosovich, Tuzikov et al. 2000) (Ha, Stevens et al. 2003). In case of H1 HA from the 1918 pandemic virus, this change is correlated with

Glu190Asp and Gly225Asp mutations (Matrosovich, Tuzikov et al. 2000) (Gamblin, Haire et al. 2004) (Stevens, Blixt et al. 2006).

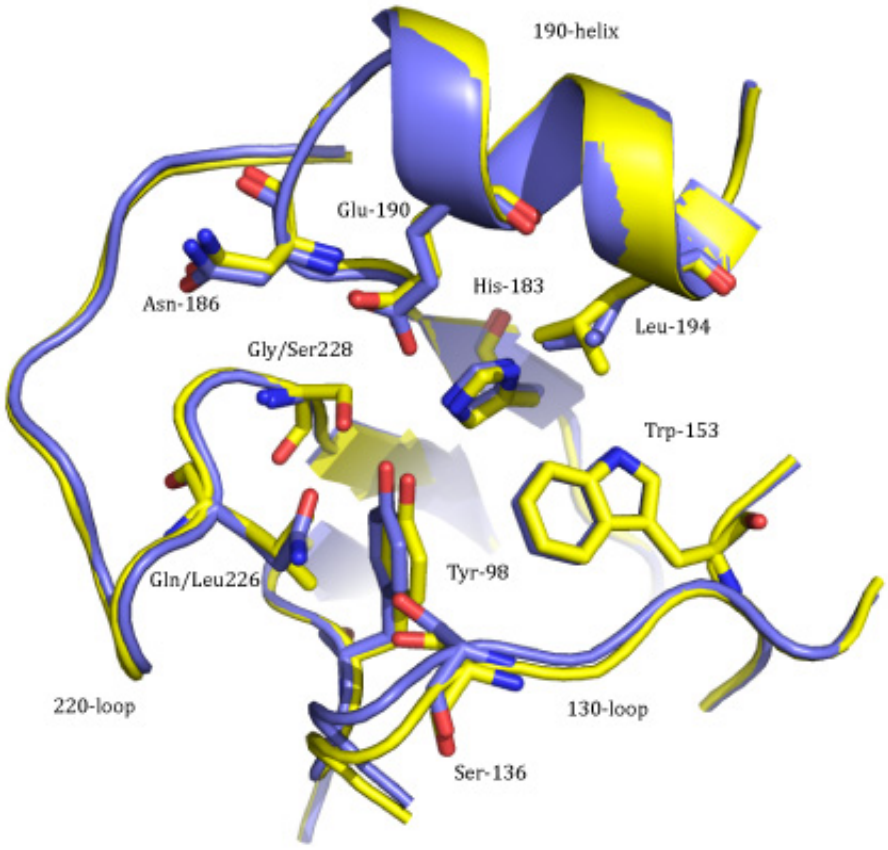


Figure 1.4 Human and Avian HA receptor-binding sites. Human H2 A/Singapore/1/57 (PDB: 2WRC) and Avian H2 A/Duck/Ontario/77 (PDB: 2WR5) receptor binding sites are shown in yellow and purple, respectively. Conserved residues such as Tyr98, Ser136, Trp153 and His183, together with other residues involved in receptor binding specificity are shown as sticks. Rmsd= 0.310 Å.

1.4.2 Membrane fusion

The stem region of HA composed of HA2 is highly conserved across the known HA subtypes. This viral membrane-proximal region of the molecule is responsible for performing fusion of viral and endosomal membranes, in order to release the viral ssRNA genome into the host cell cytoplasm for replication (Skehel and Wiley 2000). In order to complete this stage of viral infection, HA0 precursor must first undergo a proteolytic cleavage into HA1 and HA2, as described in Section 1.4. Receptor-bound viruses enter the cell using endosomes, and receptor binding is described in Section 1.4.1. The acidic pH of late endosomes (pH 5-6) induces an irreversible structural rearrangement of cleaved HA (Skehel, Bayley et al. 1982) (Skehel, Daniels et al. 1986) (Wharton 1987), and the pH value at which the conformational change of HA takes place varies, depending on virus strain (Wharton, Skehel et al. 1986).

It has been proposed, that the extension of the α-helical coiled coil of HA at low pH serves to transport the N-terminal fusion peptide (HA2 residues 1-23), from the hydrophobic cavity where it is located after cleavage of the HA0 precursor, towards the endosomal membrane (Skehel, Bayley et al. 1982) (Doms, Helenius et al. 1985), and this process is thought to be related to protonation of highly conserved His and Arg residues at low pH (Russell, Gamblin et al. 2004). Conformational changes at the viral membrane-proximal end of the molecule result in relocation of the C-terminal transmembrane anchor of HA2 to the top of the rearranged molecule (Bullough, Hughson et al. 1994), and its proximity to the hydrophobic fusion peptide is thought to be crucial in the process of viral membrane fusion (Skehel, Bayley et al. 1982). Although to date not completely understood, viral membrane fusion is thought to proceed through the formation of an extended intermediate, which then collapses, leading to hemifusion (lipid mixing) and the formation of a fusion pore (content mixing), through which the viral genome enters the cytosol of infected cell (Harrison 2008) (Cross, Langley et al. 2009). Fusogenic properties of influenza viruses have been studied utilizing their ability to haemolyse erythrocytes at low pH (Maeda and Ohnishi 1980) (Huang, Rott et al. 1981), and by measuring decrease in FRET between two liposome membrane-bound fluorescent probes upon fusion (Stegmann, Hoekstra et al. 1985) (Wharton, Skehel et al. 1986). The hypothetical mechanism of viral membrane fusion is shown in Figure 1.5.

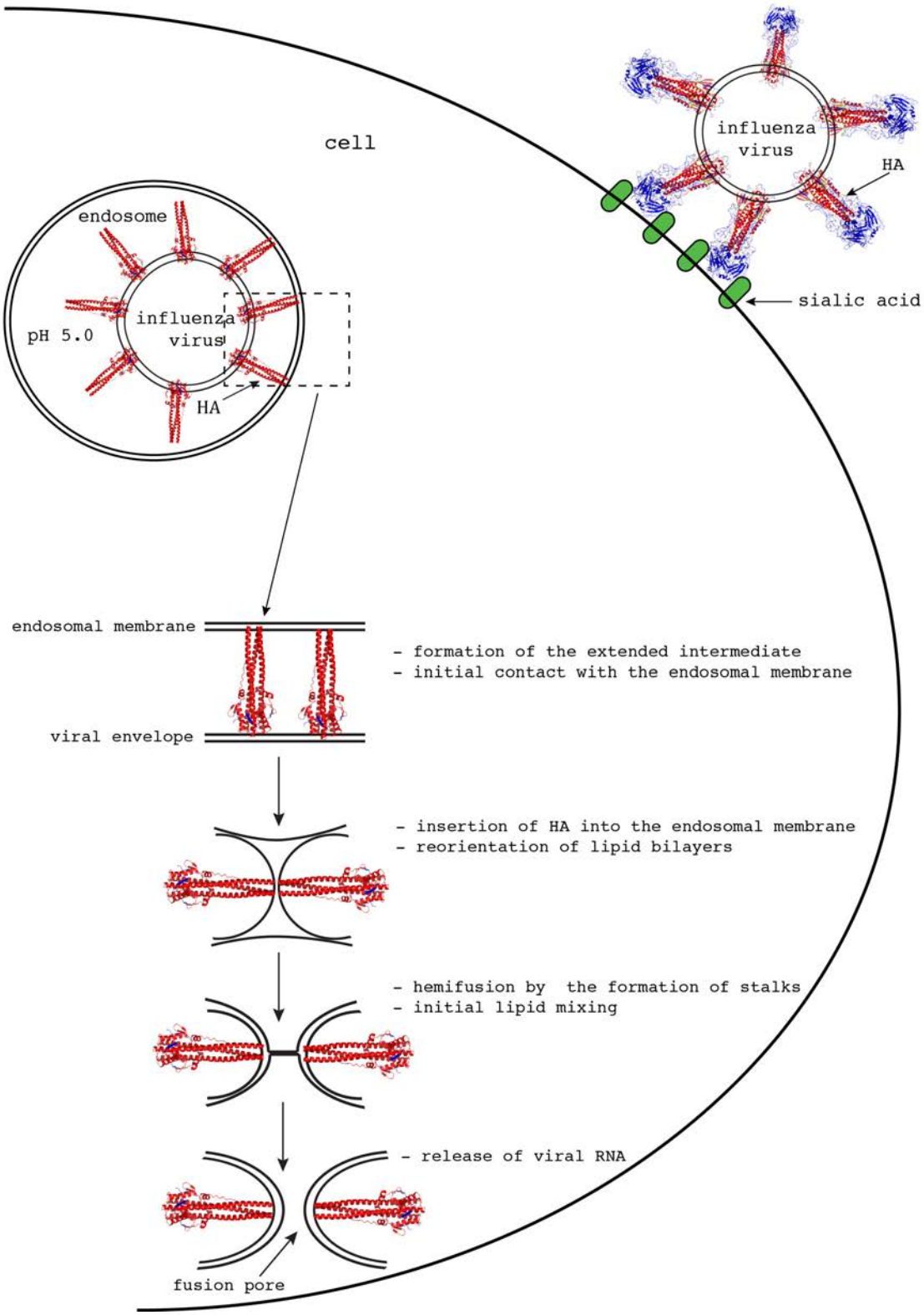


Figure 1.5 Hypothetical mechanism of HA-mediated viral membrane fusion.

1.4.3 Antigenic evolution

In addition to the initial attachment and subsequent entry of the viral genome into the cell through the process of viral membrane fusion (Skehel and Wiley 2000) (Harrison 2008), HA serves as an antigen for antibody neutralization (Laver and Kilbourne 1966). Antigenically novel virus strains are produced in the processes of antigenic drift and shift (Skehel, Daniels et al. 1983), described in Section 1.1.3. These viruses can evade natural immunity gained from previous infections, and thus have a potential to cause epidemics. Mutations in HA arising in the process of an antigenic drift do not affect the overall structure of the molecule. The variable regions of HA include the receptor-binding sites, antigenic sites recognized by antibodies located within HA1 domains (Wiley, Wilson et al. 1981) (Wiley and Skehel 1987) (Skehel and Wiley 2000), and glycosylation sites found on the protein's surface. Due to the proximity of the antigenic sites to the receptor-binding sites however, these point mutations may affect both antibody- and receptor-binding functions of the molecule (Gambaryan, Marinina et al. 1998). New glycosylation sites (Asn-X-Ser/Thr, where X is any amino acid except Pro), acquired in the process of antigenic drift, can inhibit recognition of HA by antibodies (Skehel, Stevens et al. 1984), and thus allow for infection and spread of the virus (Schulze 1997), with large N-glycans sterically interfering with virus-receptor binding.

The principal method for identification of viruses, antigenic analysis of different strains, and for quantification of antibodies (both infection- or immunization-induced) is haemagglutination-inhibition (HI) assay, proposed in 1949 by the Committee on Standard Serological Procedures in Influenza Studies. The test is preceded by a removal of naturally occurring heat resistant inhibitors of haemagglutination from cultures of *vibrio cholera* using NA (Burnet and Stone 1947). Employing monoclonal antibodies against HA facilitated studies of antigenic mutants, which neutralize majority of viruses grown at limited dilutions, and allowed for selection of antibody resistant mutants (Yewdell, Webster et al. 1979). This method, combined with peptide mapping, initially resulted in identification of four antigenic sites on influenza A H1 HA, and three on influenza A H3 HA (Wiley, Wilson et al. 1981) (Breschkin, Ahern et al. 1981). These results, together with sequence analysis of naturally occurring antigenic variants, X-ray crystallographic studies of antibody-selected mutants (such as G146D, G135R and Q189K) and HA-Fab complexes, supported by electron microscopy of HA in complex with

monoclonal antibodies, eventually allowed to map five antigenic sites on H3 HA, (Wilson and Cox 1990) (Wiley and Skehel 1987) (Wiley, Wilson et al. 1981) (Bizebard, Gigant et al. 1995) (Fleury, Barrere et al. 1999) (Skehel and Wiley 2000). These antigenic regions named A-E are located within HA1 domains, and surround the highly conserved receptor binding site (Wiley, Wilson et al. 1981) (Weis, Brown et al. 1988) (Figure 1.6).

As shown in experiments involving passaging human viruses in embryonated chicken eggs, followed by selection of mutant variants with changed receptor-binding specificity, the selective pressure for amino acid substitutions in HA1 affects the antigenic regions of HA and the efficiency of binding to the host cell, and is driven by the presence of antibodies that bind near the binding site. Given the short life span of birds, avian influenza viruses display a high conservation of antigenic sites on HA. In pigs, a short life span and therefore a low immune pressure, result in slow antigenic evolution of influenza virus HA. In humans, selection of new antigenic variants is likely to occur during reinfection of partially immune individuals, or as a result of a prolonged infection in immunocompromised children. Correct identification of antigenic changes in HA of currently circulating viruses is necessary for recommending the composition of the annual influenza vaccine.

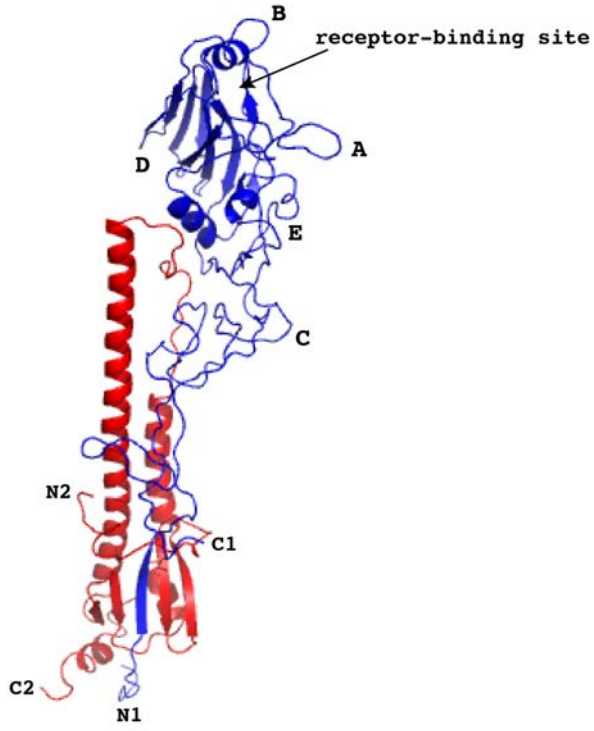


Figure 1.6 Antigenic sites A-E on the surface of HA1 domains of H3 HA. Receptor-binding site is indicated. The N- and C-termini of HA1 and HA2 are labelled N1, C1 and N2, C2, respectively. Figure adapted from (Popova, Smith et al. 2012).

1.5 Conformational changes in HA leading to membrane fusion

At the acidic pH of endosomes (pH 5-6), influenza virus haemagglutinin undergoes an irreversible structural rearrangement (Skehel, Bayley et al. 1982). Triggered by low pH, the molecule changes its conformation, releasing the hydrophobic N-terminal fusion peptide of HA2 (residues 1-23) (Skehel, Daniels et al. 1986), which is moved to the top of postfusion HA2, presumably towards the endosomal membrane (Bullough, Hughson et al. 1994). As previously demonstrated using circular dichroism (CD), electron microscopy and sedimentation in sucrose density gradients, at the pH of membrane fusion, bromelain-released HA (BHA) forms rosettes of ~8 molecules, as a result of a fusion peptide-mediated aggregation (Skehel, Bayley et al. 1982) (Ruigrok, Wrigley et al. 1986). Low pH-treated BHA was also shown to bind lipid vesicles and non-ionic detergents, whose addition prevented the aggregation of the low pH-treated BHA (pH 5.0) (Skehel, Bayley et al. 1982). Previous CD analysis of the low pH-treated BHA (near-UV region), recorded changes in the CD spectrum, and increased signal intensity in the region of 270-280 nm, corresponding to the conformational change of the low pH-activated molecule (Skehel, Bayley et al. 1982).

It is known that the HA1 domains move away from the trimeric stem of HA, allowing the central α-helices to elongate (Skehel, Bayley et al. 1982), and this dissociation is thought to be an initial step in the low pH-induced refolding of influenza haemagglutinin (Godley, Pfeifer et al. 1992). It has been previously shown that cross-linking of HA1 domains prevents the conformational change of HA (Godley, Pfeifer et al. 1992) (Barbey-Martin, Gigant et al. 2002). While HA0 does not respond to acidic pH (Chen, Lee et al. 1998), it has been proposed, that displacement of HA1 domains in BHA is caused by protonation of its solvent-exposed regions, which acquire a positive net charge at acidic pH, and dissociation is thought to be driven by a repulsive force (Huang, Opitz et al. 2002). The observed elongation of haemagglutinin at low pH (Skehel, Bayley et al. 1982) (Bullough, Hughson et al. 1994) (Chen, Skehel et al. 1999) exposes the previously inaccessible trypsin cleavage sites at residues 27 and 224 of BHA1 (H3 subtype) (Skehel, Bayley et al. 1982). The HA1 domains can therefore be secondarily digested using proteases post-purification (Skehel, Bayley et al. 1982) (Doms, Helenius et al. 1985). The low pH-induced refolding of HA exposes the interchain disulphide bond linking HA1 and HA2, and HA1 domains can then be removed using dithiothreitol (DTT) (Graves, Schulman et al. 1983).

This rearrangement of HA1 domains, and the formation of a central cavity, allow for the surrounding solvent to access HA2, its protonation, and subsequent formation of the extended coiled coil (Huang, Opitz et al. 2002). The extended structure is formed by a recruitment of loop B (HA2 residues 55-75) in the prefusion structure (Wilson, Skehel et al. 1981) to the central postfusion coiled coil. This loop (B) adopts an α-helical conformation in the postfusion structure of H3 HA2, and serves to connect the long helix C (HA2 residues 76-105) and the short helix A (HA2 residues 38-55), observed in the crystal structures of postfusion H3 HA2 (Bullough, Hughson et al. 1994) (Chen, Skehel et al. 1999).

At the other end of HA molecule, a part of the prefusion central stem (HA2 residues 106-112) undergoes a helix-to-turn transition, inverting the rest of the polypeptide chain (C-terminal to HA2 residue 112) by 180° (Bullough, Hughson et al. 1994) (Chen, Skehel et al. 1999). This rearrangement results in a relocation of the C-terminal TM anchor (HA2 residues 185-211), which is moved near the N-terminal fusion peptide (HA2 residues 1-23), and the proximity of the two subdomains of HA2 at the same end of the postfusion molecule is thought to promote viral membrane fusion (Bullough, Hughson et al. 1994).

Digestion of bromelain-released and low pH-treated BHA with trypsin, in order to remove disordered HA1 domains (Skehel, Bayley et al. 1982), followed by a digestion with thermolysin to remove the released hydrophobic fusion peptide (HA2 residues 1-23) (Ruigrok, Aitken et al. 1988), result in a fragment of viral HA2 ectodomain in a postfusion conformation, generally referred to as TBHA2 (Thermolysin/Bromelain-released HA2). The enzymatically prepared fragment is suitable for crystallographic studies (Bullough, Hughson et al. 1994), and structural refolding of H3 HA2 at pH ~5.0 is shown in Figure 1.7.

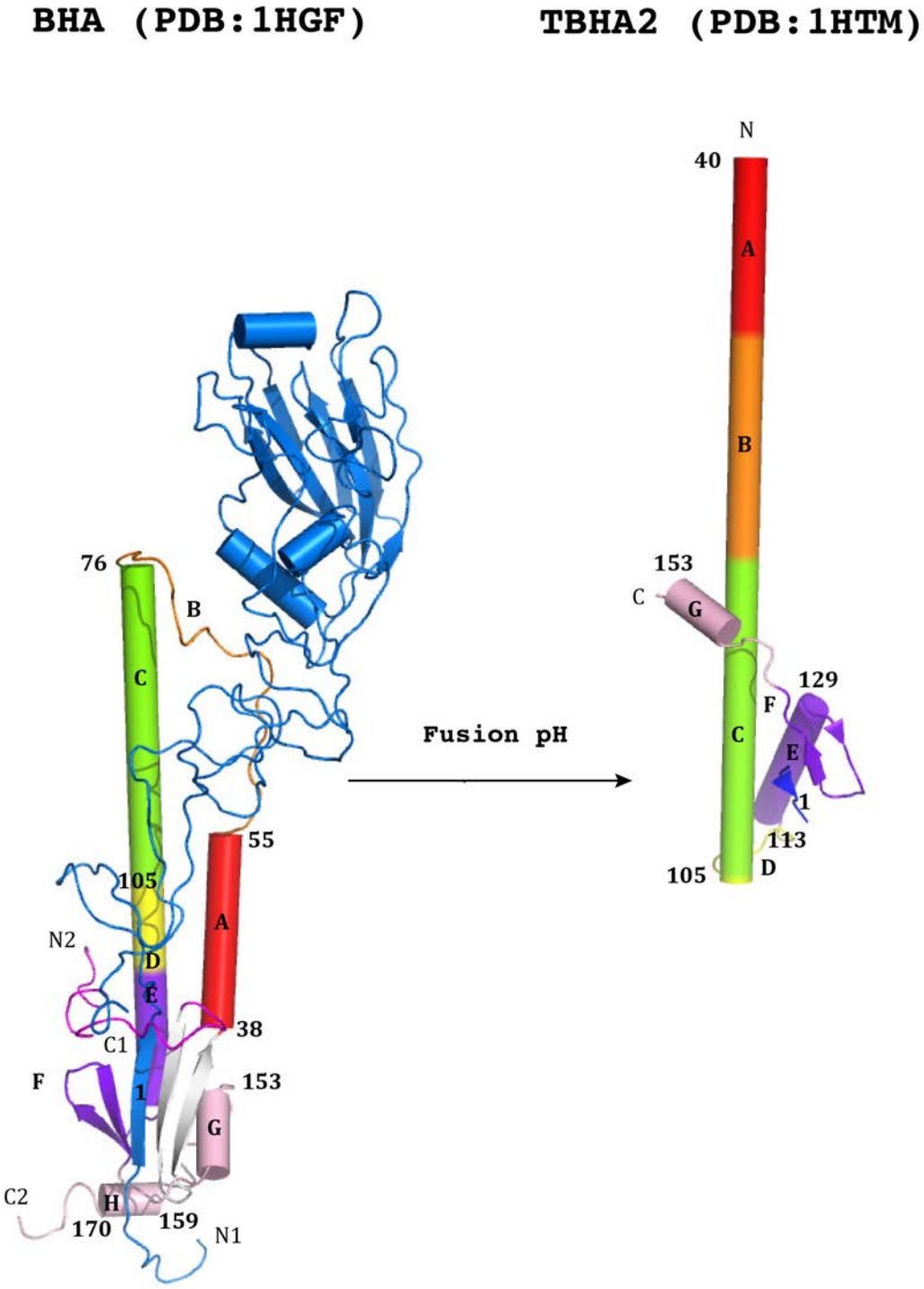


Figure 1.7 The low pH-induced conformational change of HA. Monomeric cartoon representation of bromelain cleaved HA on the left (PDB: 1HGF) (Wilson, Skehel et al. 1981) and thermolysin solubilized, postfusion H3 TBHA2 on the right (PDB: 1HTM) (Bullough, Hughson et al. 1994). The HA1 domain of bromelain-released HA is shown in blue. HA2 molecules are coloured by segments, which undergo refolding at low pH. 4 α-helical segments: A, C, E and G, and 2 β-strands 1 and F are conserved in both structures. Loop B between helices A and C adopts a helical conformation in the refolded molecule. Segment D (BHA residues 106-112) refolds into a loop between helices C and E in TBHA2. Figure adapted from (Bullough, Hughson et al. 1994).

1.6 Structural information on the low pH-induced structural refolding of HA obtained to date

1.6.1 Crystal structure of H3 HA0 precursor

The crystal structure of an uncleaved, single-chain HA0 precursor from human A/Aichi/2/68 H3N2 virus (PDB: 1HA0) was determined in 1998 (Chen, Lee et al. 1998), and an uncleaved trimer is shown in Figure 1.8. The haemagglutinin of this virus is identical to that of the strain that caused the 'Hong Kong' pandemic in 1968. Arg329Gln substitution was introduced to prevent cleavage into HA1 and HA2. Only 19 residues near the cleavage site (HA1 residues 323-328, Gln329 and HA2 residues 1-12) have a different structural conformation when compared to the crystal structure of cleaved H3 HA (BHA) (PDB:1HGF), described in Section 1.6.2. These residues are organized into a loop on the glycoprotein's surface, exposing HA1 residues 327- HA2 residue 5 to proteases (Chen, Lee et al. 1998).

1.6.2 Crystal structure of H3 BHA

The structure of a metastable, cleaved, bromelain-released HA (BHA) of A/Hong Kong/1968 H3N2 virus was previously determined to 3 Å resolution using X-ray crystallography (Wilson, Skehel et al. 1981). This was the first crystal structure of a membrane fusion protein, and was published by Wilson, Skehel and Wiley in *Nature* in 1981. Bromelain solubilized the trimeric fragment of HA ectodomain, by cleaving the hydrophobic transmembrane anchor (TM) at the C-terminus of HA2 (residues 185-211), and the cleavage point was residue 175 of HA2 (Compans, Klenk et al. 1970) (Brand and Skehel 1972) (Skehel and Waterfield 1975) (Dopheide and Ward 1981). The prepared fragment was antigenically (Brand and Skehel 1972) (Wrigley, Laver et al. 1977) and structurally (Flanagan and Skehel 1977) (Wiley, Skehel et al. 1977) intact and useful for crystallographic studies (Wiley and Skehel 1977) (Wilson, Skehel et al. 1981). In BHA, 328 residues of HA1 are disulphide-linked to 175 residues of HA2, and the disulphide bond is located between Cys14 of HA1 and Cys137 of HA2. The 135 Å long trimeric structure, revealed a globular region composed of HA1, placed on top of a ~76 Å long, triple-stranded, α-helical coiled coil, composed mainly of HA2. Residues 116-261 of each HA1 monomer are organized into 8-stranded antiparallel ß-sheet, known as "Swiss roll", which contains two antigenic sites in the form of loops (Figure 1.6). The

globular HA1 domains are attached to the stem of the molecule by two disulphide loops (residues 52-277 and 281-305). Each HA2 subunit contains two antiparallel helices: a short α-helix A (29 Å long) connected to a long α- helix C (76 Å long) via a loop B. The newly formed positively charged N-terminal fusion peptide (HA2 residues 1-23) is placed into a negatively charged cavity, which is considered a fusion-priming event (Skehel and Wiley 2000). At the viral membrane-proximal end, the molecule is arranged into a 5-stranded, antiparallel ß-sheet, composed of the N-terminus of HA1 and the C-terminus of HA2. The central super helix contains three HA2 helices twisted around each other by ~100°. The top of the trimeric structure is tightly packed and stabilized by nonpolar residues including Ile77, Leu80, Val84, Leu91, Leu98, Leu99 and Leu102. These residues form Van der Waals interactions around the 3-fold axis of the molecule. Polar and charged residues stabilize the viral membrane-proximal part of the coiled coil, and these are His106, Asp109, Asp112, Ser113, Asn116, Glu120 and Arg123. Intra-chain and inter-subunit salt bridges were shown to further stabilize this trimeric α-helical structure (Wilson, Skehel et al. 1981). The length of the observed molecule (~135 Å) and the buried position of the fusion peptide, closer to the viral envelope (~35 Å) than to the tip of the molecule (~100 Å) indicated, that in order to perform viral membrane fusion, the molecule must undergo a significant structural rearrangement. Thirteen years later, the crystal structure of a fragment of H3 HA2 ectodomain, isolated from the 1968 Hong Kong pandemic virus, was determined using X-ray crystallography (Bullough, Hughson et al. 1994), and its crystal structure is described in Chapter 1.6.3. A BHA trimer is shown in Figure 1.8.

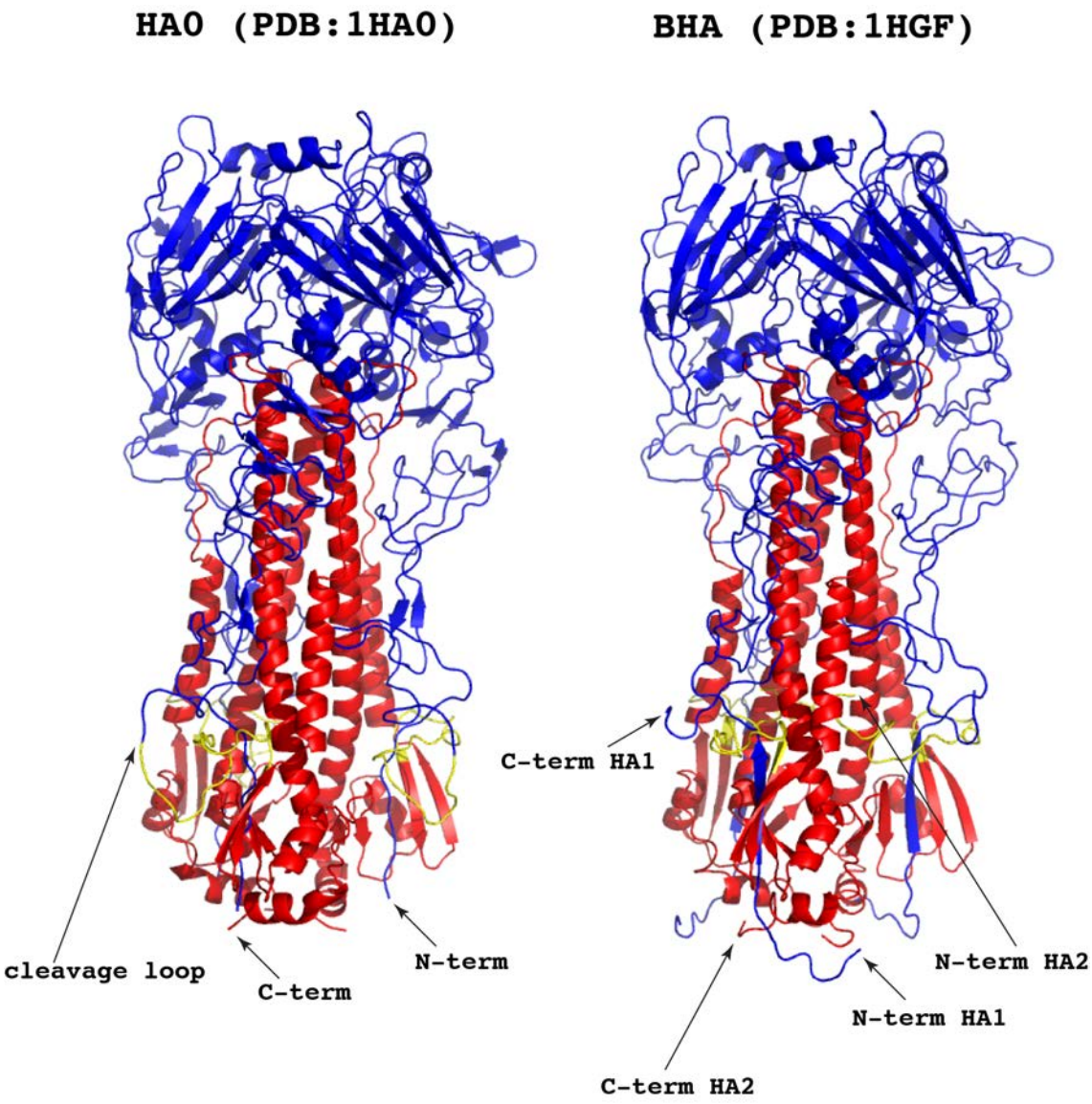


Figure 1.8 Crystal structures of the precursor HA0 and cleaved BHA. HA0 structure (PDB: 1HA0) (Chen, Lee et al. 1998) is shown on the left. Cleaved BHA structure (PDB: 1HGF) (Wilson, Skehel et al. 1981) is shown on the right. Upon cleavage by host cell proteases, HA is divided into receptor-binding (HA1) and fusion (HA2) subdomains, shown in blue and red, respectively. The first residue of the newly formed FP (yellow) becomes the new N-terminus of HA2.

1.6.3 Crystal structure of postfusion viral H3 TBHA2

A soluble fragment of H3 HA2 in a postfusion conformation was prepared by releasing HA from the viral envelope with bromelain (Brand and Skehel 1972) (Skehel and Waterfield 1975), low pH activation by incubation at pH 5.0, and subsequent digestions with trypsin and thermolysin to remove the hydrophobic subdomains of HA2. The structure of this proteolytic fragment of viral H3 HA2, referred to as H3 TBHA2, was determined using X-ray crystallography in 1994 (Bullough, Hughson et al. 1994). The previously observed irreversibility of the low pH-triggered structural rearrangement of HA, and its increased thermal stability suggested, that this refolded molecule is more thermodynamically stable than the molecule in its prefusion conformation (Ruigrok, Martin et al. 1986) (Chen, Wharton et al. 1995). As observed in the obtained crystal structure of H3 TBHA2, the enzymatically-prepared molecule contains the first 27 residues of HA1, and residues 40-162 of HA2 (Bullough, Hughson et al. 1994). Comparison of H3 TBHA2 to the structure of prefusion H3 BHA revealed, that the central α-helix extends at low pH. This elongation is achieved by a recruitment of residues 40-75 of H3 HA2 to the top of the long postfusion α-helix, and formation of an extended coiled coil is possible due to the presence of heptad repeats of hydrophobic residues in this region of the molecule (HA2 residues 38-125). Closer to the viral membrane, HA2 residues 106-112, which are helical in prefusion HA2, refold to form a reverse turn, which repositions the rest of the molecule by 180°. Residues 113-129 fold into a short α-helix E, which runs in an opposite direction to the central stem and forms a six-helix-bundle. The polypeptide chain is then arranged into a loop, a fragment of which forms a ß-hairpin (HA2 residues 131-140). Together with a disulphide-linked fragment of HA1, these residues are organized into an antiparallel ß-sheet. The ß-hairpin motif extends further to form another short helix G (residues 146-154), packed against the central stem, and the rest of the polypeptide chain exists as an extended ß-strand. Residues beyond HA2 residue 153 (one monomer) and residue 162 (two monomers) in viral H3 TBHA2 (Bullough, Hughson et al. 1994) seem to be disordered. The ~100 Å long postfusion core, comprises HA2 residues 40-105, which display the knobs-into-holes packing (Crick 1953). Compared to the crystal structure of H3 BHA, only 30 residues (HA2 residues 76-105), belonging to helix C, remain unchanged. This dramatic structural rearrangement as compared to BHA, serves to relocate the two hydrophobic subdomains of HA2 (the fusion peptide and the transmembrane anchor) towards the endosomal membrane (Bullough, Hughson et al. 1994). The two structures are compared in Figure 1.9.

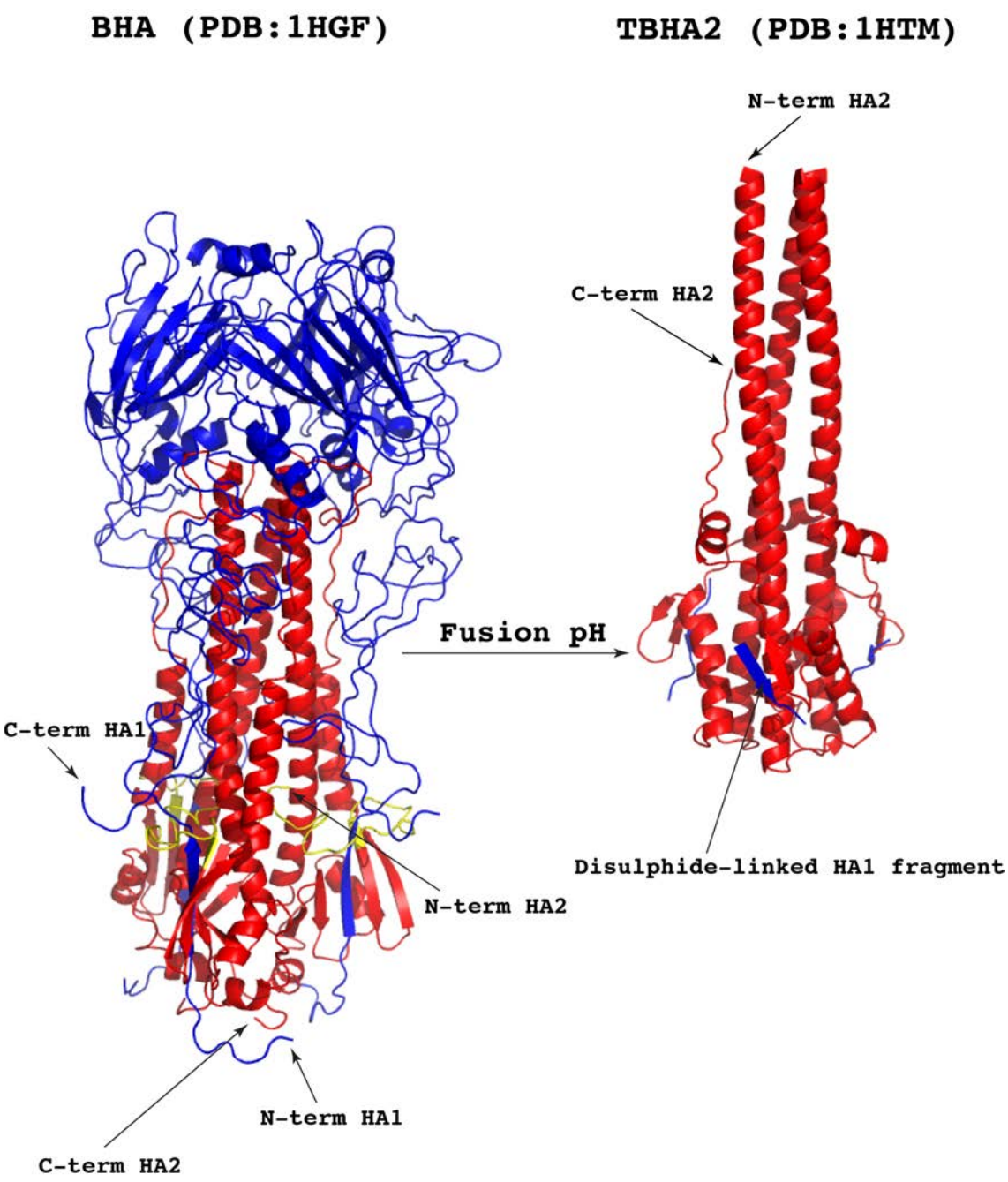


Figure 1.9 Crystal structures of prefusion and postfusion H3 HA. Trimeric cartoon representation of bromelain cleaved BHA is shown on the left (PDB: 1HGF) (Wilson, Skehel et al. 1981) and thermolysin solubilized, low pH-treated H3 TBHA2 is shown on the right (PDB: 1HTM) (Bullough, Hughson et al. 1994). HA1 domains are shown in blue, and HA2 domains are shown in red. The N- and C-termini are indicated.

1.6.4 Crystal structure of *E.coli*-expressed H3 HA2

Recombinantly-expressed fragments of H3 HA2 containing HA2 residues 38-175 (EBHA2) and HA2 residues 23-185 (EHA2), lacking the HA1 domains, the fusion peptide, and the transmembrane anchor, spontaneously folded into the same extended conformation as postfusion viral H3 TBHA2 (Bullough, Hughson et al. 1994) described in Section 1.6.3, when expressed in *E.coli* (Chen, Wharton et al. 1995). In order to prevent a potential aggregation of recombinantly-expressed molecules, HA2 Cys137, which forms a disulphide bond with Cys14 of HA1, was mutated to Serine. Cysteines 144 and 148, which form a disulphide bond within each HA2 monomer, were preserved. The expressed fragments were trimers of ~48 kDa, with high α-helical content and high melting temperatures, recorded previously for TBHA2 (Ruigrok, Martin et al. 1986). Their rod-like structure, ability to bind the low pH-specific antibodies, and the obtained proteolytic digestion products agreed with those previously observed for viral postfusion H3 TBHA2 (Ruigrok, Aitken et al. 1988) (Wharton, Calder et al. 1995).

The crystal structure of recombinantly-expressed H3 EHA2 (Chen, Skehel et al. 1999) suggests, that prefusion HA2 is metastable, and that low pH may be needed in order to induce the dissociation of HA1 subdomains *in vivo* (Chen, Wharton et al. 1995). Fragments of postfusion H3 HA2 (residues 1-185), containing the hydrophobic fusion peptide (HA2 residues 1-23), were previously expressed in *E.coli*. The expressed fragment was attached to a highly charged octapeptide FLAG (DYKDDDDK), which restored the solubility of the molecule (Chen, Skehel et al. 1998). This fusion peptide-containing fragment (known as F185) was shown to have a high α-helical content and high melting temperature, as previously observed for H3 EBHA2 and TBHA2 (Ruigrok, Martin et al. 1986). While removal of FLAG resulted in aggregation of the protein molecules into rosettes, removal of the hydrophobic fusion peptide (HA2 residues 1-23) using thermolysin, restored protein solubility (Chen, Skehel et al. 1998) (Skehel and Wiley 2000).

The crystal structure of H3 EHA2 revealed the previously unobserved N- and C-terminal segments of the postfusion molecule, and the N-terminal residues were organized into an N-cap domain (HA2 residues 34-37). These residues were shown to stabilize and terminate the α-helical coiled coil, by the formation of hydrogen bonds with three unpaired residues at the N-terminus of each α-helix (HA2 residues 38-40), and the last helical residue is Leu38 (Chen, Skehel et al. 1999). The capping residues (HA2 residues 34-37) interact with the extended C-termini of

the postfusion molecule (HA2 residues 174-178), stabilizing both membrane-interacting regions (Chen, Skehel et al. 1999). High conservation of Ala35, Ala36 and Asp37 suggests, that a similar N-cap exists in viral postfusion HA2 (Skehel and Wiley 2000). Given that beyond the observed N-cap the HA2 molecule is likely disordered, it is thought, that HA2 ectodomain and the fusion peptide may exist as two separately folded domains, linked by a flexible linker (HA2 residues 25-33) (Chen, Skehel et al. 1999). The stabilizing N-cap observed in the crystal structure of *E.coli*-expressed H3 HA2 (EHA2) (Chen, Skehel et al. 1999) is shown in Figure 1.10. A schematic representation of H3 HA, and fragments of viral and recombinantly-expressed H3 HA2 crystallized to date are shown in Figure 1.11.

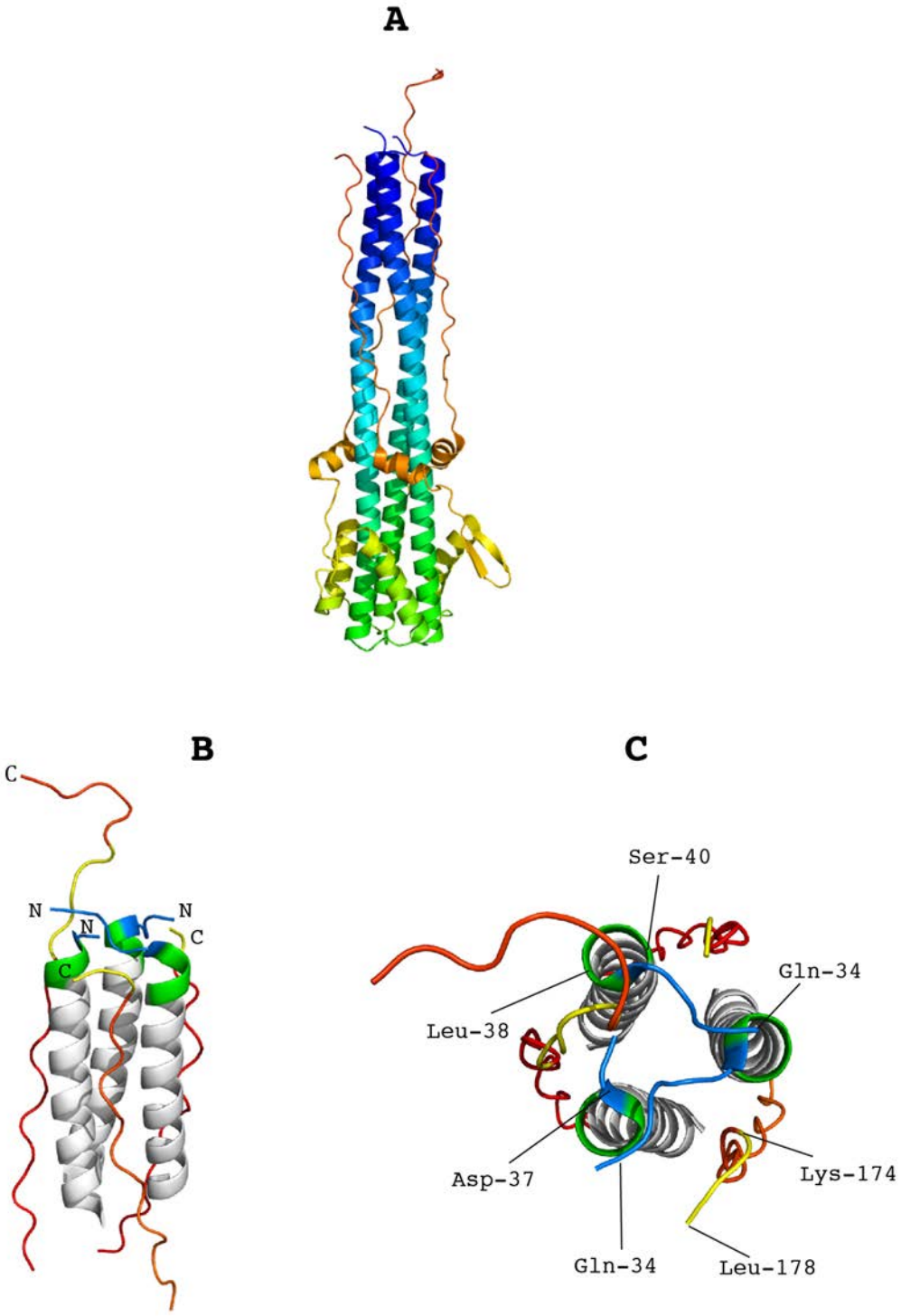


Figure 1.10 A trimer of H3 EHA2. Recombinantly-expressed trimer is shown in A. Viral membrane-distal end of the molecule viewed perpendicular to the molecular three-fold symmetry axis is shown in B. N-cap residues 34-37 are shown in blue, residues 38-40 in green, residues 174-178 in yellow. N-cap domain of EHA2 viewed down the molecular three-fold symmetry axis is shown in C. Colours as in B. (PDB: 1QU1) (Chen, Skehel et al. 1999).

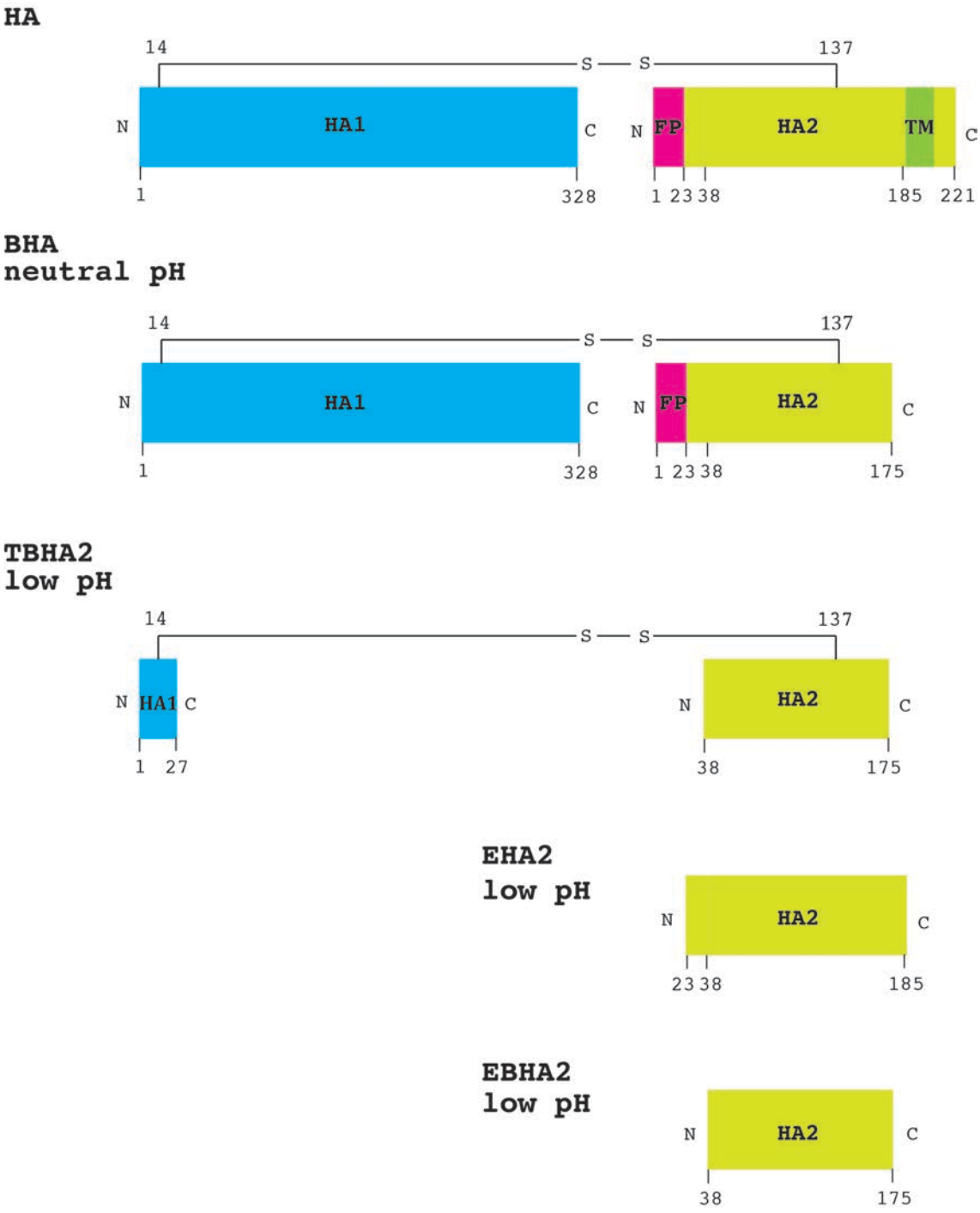


Figure 1.11 Schematic of H3 HA molecule and fragments crystallized to date. HA1 and HA2 domains are shown in blue and green, respectively. Top model: Cleaved HA, with the TM domain (HA2 residues 186-211) shown in dark green, the FP (HA2 residues 1-23) shown in magenta, and the disulphide bond linking HA1 residue 14 and HA2 residue 137 shown in black. Shown below are: BHA (Wilson, Skehel et al. 1981), TBHA2 (Bullough, Hughson et al. 1994), EHA2 (HA2 residues 23-185) (Chen, Skehel et al. 1999) and EBHA2 (HA2 residues 38-175) (Chen, Wharton et al. 1995). Figure adapted from (Chen, Skehel et al. 1999).

1.6.5 Crystal structure of *E.coli*-expressed Influenza B HA2

Influenza B viruses cause seasonal influenza outbreaks, and share low levels of sequence identity with influenza A (~39%). Influenza B viruses are divided into two co-circulating lineages: Yamagata and Victoria (Rota, Wallis et al. 1990). Recently, a crystal structure of *E.coli*-expressed, influenza B/Yamagata/73 (HA2 residues 31-181) in a postfusion conformation has been reported, revealing a hairpin-like structure as that of influenza A HA2 (Ni, Chen et al. 2014). As shown in Figure 1.12, the postfusion stem can be divided into three α-helical segments: helix A (residues 38-55), helix B (residues 56-75), and helix C (residues 76-105). Overall, the observed structure (Ni, Chen et al. 2014) is similar to that of postfusion influenza A H3 HA2 (Bullough, Hughson et al. 1994) (Chen, Skehel et al. 1999), and main structural differences can be seen in the regions of loops and in segment B (HA2 residues 56-75), belonging to the central, postfusion core of the extended molecule (Ni, Chen et al. 2014). The postfusion structure of the N-termini, containing residues of the N-cap domain (residues 34-37), and that of segment A of the long α-helix, were shown to be conserved between postfusion HA2 from influenza A and influenza B. The C-termini (HA2 residues 154-181) of the rearranged influenza B HA2 were shown to form contacts with the capping residues of the N-cap domain (Ni, Chen et al. 2014). Aligned monomers of *E.coli*-expressed influenza A H3 HA2 and influenza B HA2 in a postfusion conformation are shown in Figure 1.12.

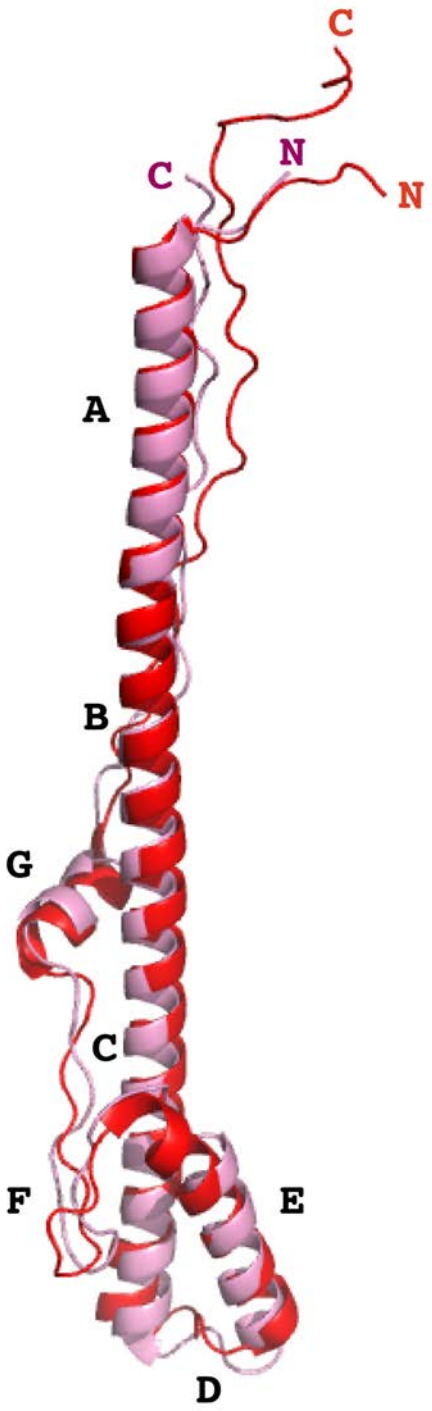


Figure 1.12 Structural alignment of X31 H3 EHA2 and influenza B HA2. A monomer of *E.coli*-expressed H3 EHA2 is shown in red (Chen, Skehel et al. 1999) (PDB: 1QU1) and monomer of influenza B HA2 in a postfusion conformation is shown in pink (Ni, Chen et al. 2014) (PDB: 4NKJ). The N- and C-terminal residues are indicated. Segments A-G undergo refolding at low pH, and the postfusion stem comprises helix A (residues 38-55), helix B (residues 56-75) and helix C (residues 76-105).

1.7 Group-specific structural features of HA

The 16 HA antigenic subtypes fall into two phylogenetic groups (Air 1981) (Russell, Gamblin et al. 2004), and sequence identity between them is 30-70% (Nobusawa, Aoyama et al. 1991) (Kawaoka, Yamnikova et al. 1990) (Rohm, Zhou et al. 1996). The primary sequence is a basis for classification of HA subtypes into group 1 (H1, H2, H5, H6, H8, H9, H11, H12, H13 and H16) and group 2 (H3, H4, H7, H10, H14 and H15) HA. Crystal structures of H1, H2, H3, H5, H7, H9, and H16 HA have to date been reported. Sequence alignment of representative subtypes from group 1 and group 2 HA in the region of HA2 (residues 1-221) is shown in Figure 1.13.

The two HA groups are further subdivided into four clades based on the phylogenetic characterization and sequence homology of HA gene. Group 1 HA is subdivided into two clades represented by H1 (H1, H2, H5, H6, H11, H13, H16) and H9 (H8, H9, H12) subtypes, respectively. Group 2 HA is subdivided into two clades represented by subtypes H3 (H3, H4, H14) and H7 (H7, H10, H15) (Nobusawa, Aoyama et al. 1991).

The characteristic structure and orientation of an interhelical loop B, present in each prefusion HA2 monomer, is one of the group-specific structural features of HA (Russell, Gamblin et al. 2004). Loop B (HA2 residues 56/59-74) connects the prefusion helices A (HA2 residues 38-55/58) and C (HA2 residues 75-105). Clades represented by H3 and H7 subtypes (group 2 HA) are characterised by a sharp nature of the turn at HA2 residue 75. The structure of this turn has been correlated with the presence of glycine at residue 75 of HA2 (the N-terminus of the conserved helix C), or with the presence of a carbohydrate side chain at residue 82 of HA2, in clades H3 and H7, respectively (group 2 HA) (Keil, Geyer et al. 1985). Clades H1 and H9 (group 1 HA) can be distinguished by a presence of tall turns at HA2 residue 75. Group-specific structures of prefusion loops B are shown in Figure 1.14. The characteristic nature of these fragments of prefusion HA2 (residues 56/59-75) is thought to affect the orientation of HA1 domains relative to the stem of the molecule (composed mainly of HA2), which is different in the two HA groups (Russell, Gamblin et al. 2004). Differences in rotation of the receptor binding domains about the three-fold symmetry axis of the molecule are the largest across groups, and difference between H1 and H7 is 31°. Within the group these differences are reduced to 6° (H3 to H7), and are the smallest within the subtype. For example the orientation of HA1 subdomains versus HA2 subdomains of H1 HA, isolated from three different virus strains: the 1918 pandemic virus (Reid,

Fanning et al. 1999), classical swine H1 A/swine/30 virus (Shope 1931), and from the prototype human H1 A/Puerto Rico/8/34 virus (Francis 1934), was previously shown to be very similar (rotation by 1°) (Russell, Gamblin et al. 2004). Rotation of HA1 domains relative to HA2 stem for representatives of the two HA groups is shown in Figure 1.15. Another group- and clade-specific structural feature of HA2 is the type of ionisable residues found in the cavity that accommodates the hydrophobic fusion peptide (at the N-terminus of HA2). Protonation of these residues after exposure of the virus to the low pH of the endosomes (pH~5.0), is thought to be a trigger for the conformational change of HA, required for haemagglutinin to perform viral membrane fusion by the formation of an extended intermediate (Chen, Skehel et al. 1999) (Ni, Chen et al. 2014). While some residues within this cavity (Asp109, Asp112 and Lys51) are conserved across all HA subtypes, other amino acids are group-specific. The conserved His17 of HA1, and His106 of HA2 are found in the primary sequences of group 2 HA (clades H3 and H7). Subtypes belonging to group 1 HA (clades H1 and H9) contain a conserved His111. Some amino acids belonging to this pocket are clade-specific, like HA2 Arg106 and Lys106, in clades H1 and H9, respectively (Ha, Stevens et al. 2002) (Russell, Gamblin et al. 2004).

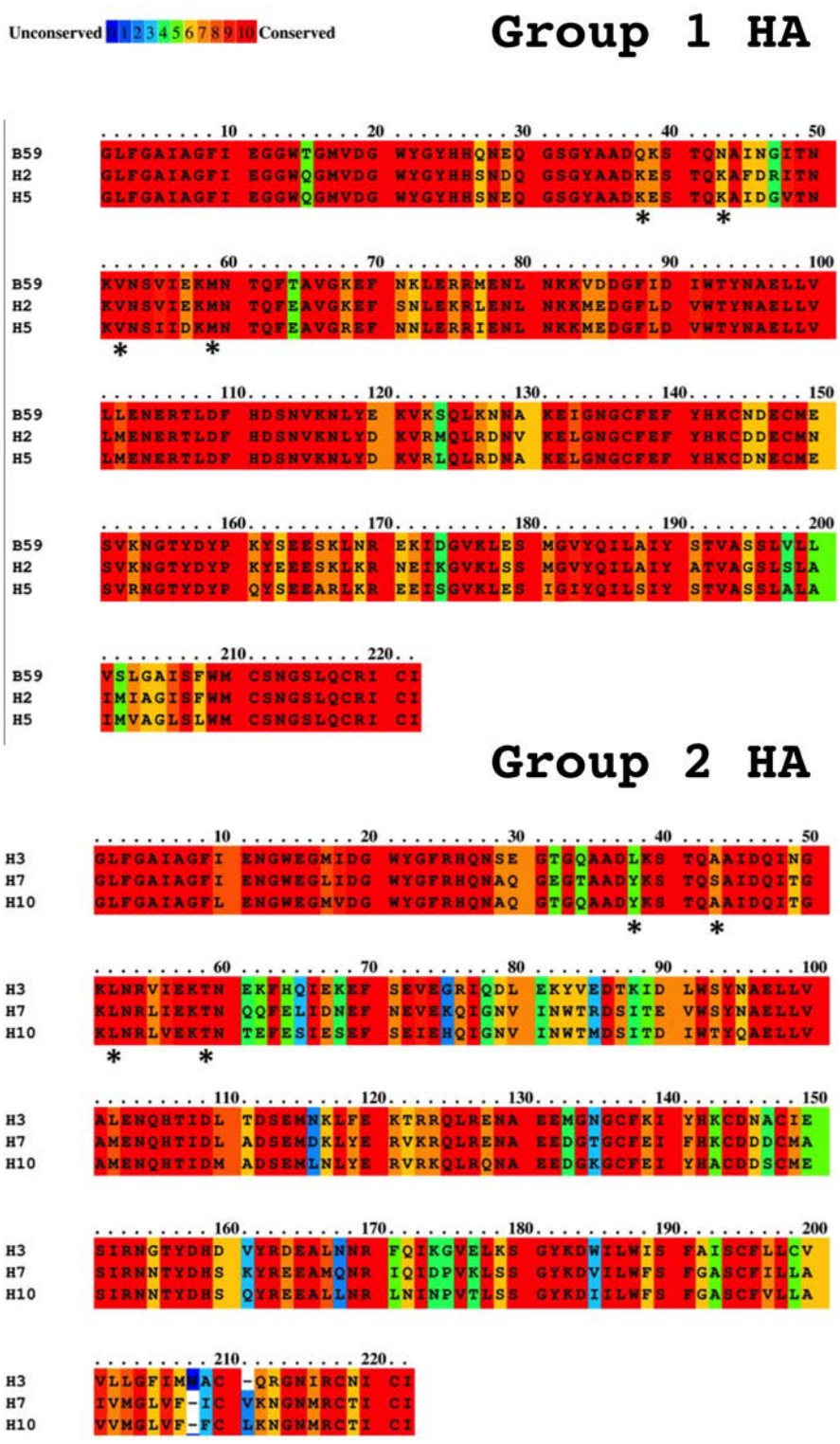


Figure 1.13 Sequence alignment of group 1 and group 2 HA in the region of HA2. Group 1 HA is represented by H1 (B59), H2 and H5. Group 2 HA is represented by H3, H7 and H10. Sequence alignment was carried out using PRALINE multiple sequence alignment (Simossis and Heringa 2003) (Simossis and Heringa 2005). Residues are coloured by conservation.

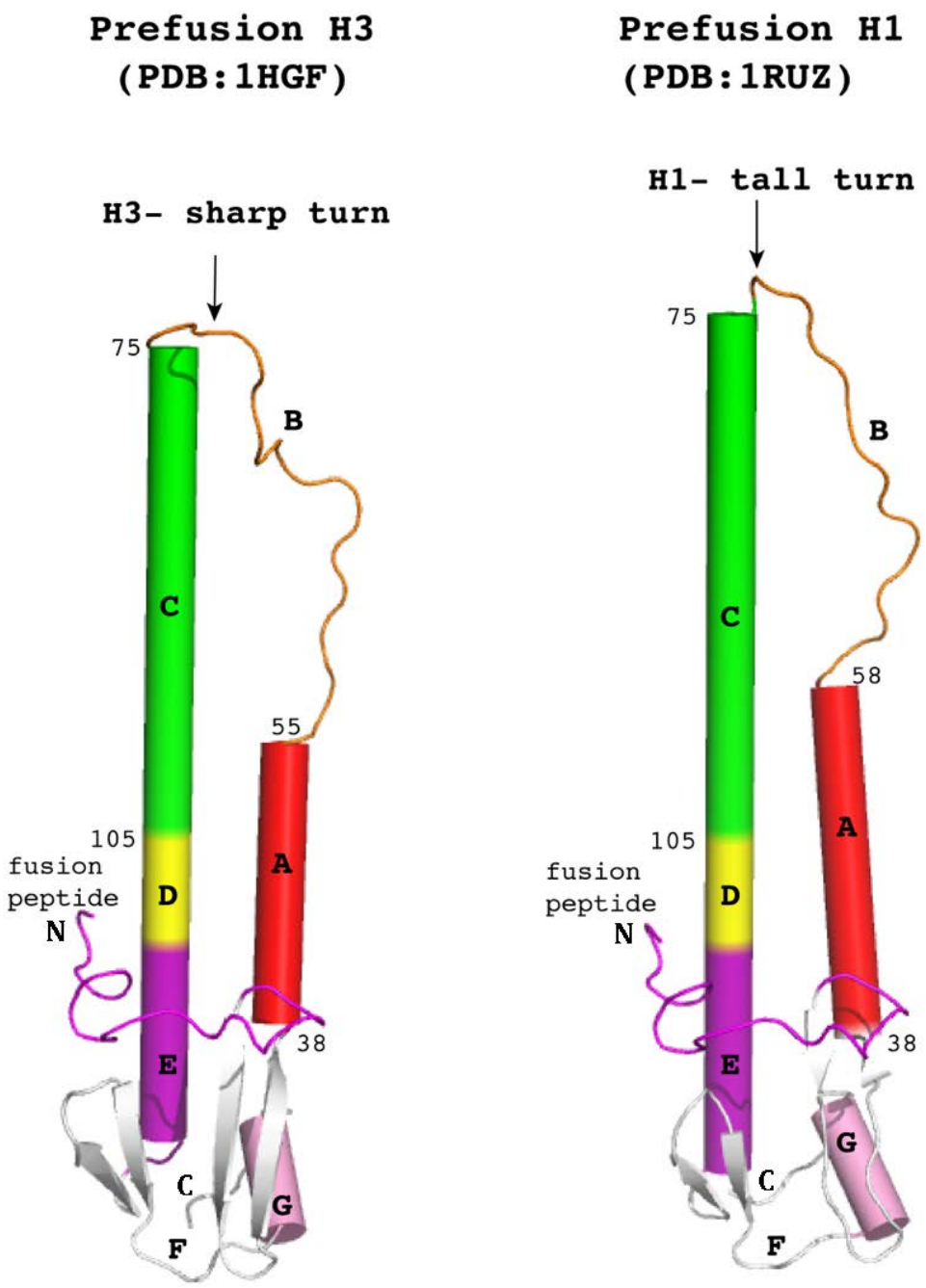


Figure 1.14 Group-specific structures of loop B. Monomeric cartoon representation of prefusion H3 HA2 (PDB: 1HGF) and H1 HA2 (PDB: 1RUZ). Helical segments A, C, D, E and G, which undergo refolding at low pH, are highlighted. Loop B, the fusion peptide, and other key residues are indicated.

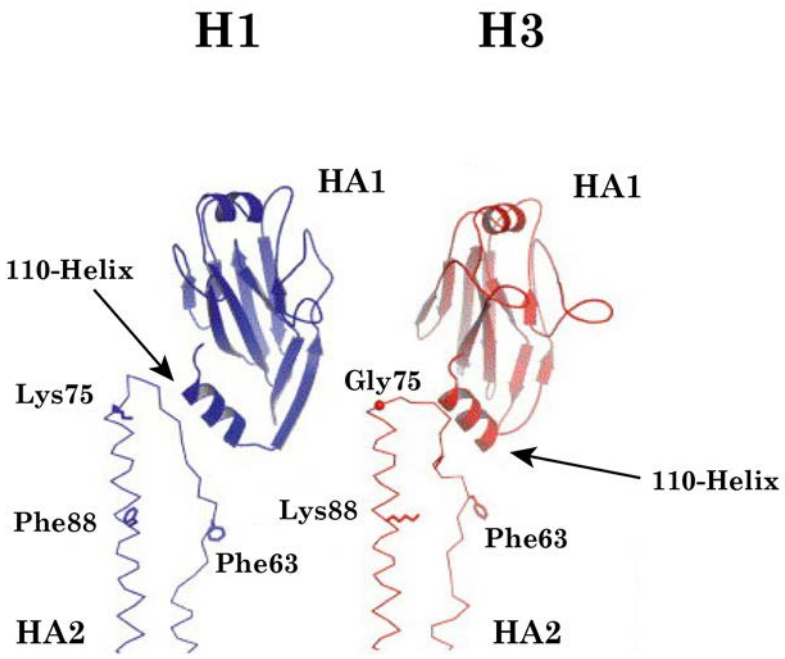


Figure 1.15 Rotation of HA1 domains relative to HA2 stem of HA. Molecules are shown as monomers. Group 1 HA is represented by subtype H1, and group 2 HA is represented by subtype H3. Figure from (Russell, Gamblin et al. 2004).

1.8 Anti-HA antibodies

1.8.1 Antibodies blocking receptor-binding

Majority of human anti-HA antibodies generated upon vaccination or infection, are directed against HA1 domains, which undergo a frequent antigenic change, and accumulated mutations resulting from RNA polymerase errors and selective immune pressure, lead to annual epidemics. As described in Section 1.4.3, five antigenic sites involved in antibody recognition have been characterized on the surface of globular head domains of H1 (Caton, Brownlee et al. 1982) and H3 HA (Wiley, Wilson et al. 1981). The serum antibody response is therefore subtype-specific, and cross-reactive antibodies are uncommon. The HA1-specific antibodies may block virus infectivity by interfering with the receptor-binding function of the molecule (Knossow and Skehel 2006). The antibody-binding epitopes located "below the tip" of HA were initially identified using electron microscopy (Wrigley, Brown et al. 1983). By binding to the viral membrane-distal end of HA, these antibodies block the attachment of the viral particle to the target cell (Outlaw and Dimmock 1991). The footprints of these antibodies range from 1200 A^2 -1500 A^2, and can therefore sterically block receptor binding, by covering the receptor-binding site, the surface area of which is ~800 A^2 (Skehel 2009). Structural studies on H3 HA in complex with Fab fragments from two neutralizing monoclonal antibodies, named HC19 and HC45, revealed HC19 directly blocking receptor-binding function of HA by interacting with the conserved residues of the receptor-binding site, and HC45 was found to neutralize infectivity indirectly, by binding 17 Å below the receptor-binding site (Bizebard, Gigant et al. 1995) (Fleury, Wharton et al. 1998) (Fleury, Barrere et al. 1999). Despite a much weaker affinity for HA, the HC19 antibody, which binds near the receptor-binding site, has been shown to neutralize infectivity more efficiently than HC45 (Fleury, Barrere et al. 1999). Human monoclonal antibody known as CH65 was also found to neutralize a broad spectrum of H1 viruses, by overlapping with the receptor-binding site (Whittle, Zhang et al. 2011). Cross-reactive antibodies directed at the conserved receptor-binding site and its surroundings have recently been identified, and these are C05 (neutralizes subtypes H1, H2 and H3) (Ekiert, Kashyap et al. 2012), and S139/1 (neutralizes various subtypes including H1 and H3) (Lee, Yoshida et al. 2012).

1.8.2 Antibodies interfering with the conformational change of HA

Antibodies, which do not block virus entry, can neutralize influenza viruses at a later stage of viral infection, for example by interfering with the low pH-triggered conformational change of HA, and thus viral membrane fusion (Corti, Voss et al. 2011). These antibodies block infection by binding to the conserved, viral membrane-proximal end of HA, near its fusion subdomain. It has been shown, that broadly-neutralizing antibodies are generally less effective than a strain-specific antibody *in vitro* (Haaheim and Schild 1980). These antibodies, however, can provide a significant protection *in vivo* (Mozdzanowska, Furchner et al. 1997) (Mozdzanowska, Feng et al. 2006). Such cross-reactive antibodies have been prepared, complexed with HA, and structures of the complexes were solved using X-ray crystallography. These monoclonal antibodies are known as CR6261 (Ekiert, Bhabha et al. 2009) and F10 (Sui, Hwang et al. 2009), and neutralize multiple influenza A subtypes of group 1 HA, by heavy chain-mediated binding to a conserved epitope composed of helix A of HA2 and few residues of HA1. Antibodies with broader specificity have also been identified, and these are known as FI6v3 (Corti, Voss et al. 2011) and CR9114 (Dreyfus, Laursen et al. 2012).

The FI6v3 antibody was isolated using a single plasma cell culture method, where supernatants were screened for the presence of antibodies able to neutralize both H5 and H7. The antibody was isolated from a donor, following a seasonal vaccination, and after infection with the 2009 pandemic SOIV (swine-origin influenza virus). The FI6v3 antibody recognizes an epitope similar to that recognized by CR9114, but binds at a different angle (90° rotation compared to CR9114) (Corti, Voss et al. 2011). The FI6v3 uses both heavy and light chain in its binding to HA, mainly by HCDR3 (heavy-chain variable complementarity determining region 3), and with LCDR1 (light-chain variable complementarity determining region 1), crosslinking two monomers of one haemagglutinin trimer by interacting with helix A of one HA monomer, and the fusion peptide of the neighbouring monomer (Corti, Voss et al. 2011). In the future, cross-reactive antibodies may be used for passive immunization as prophylaxis or therapy, and characterization of the conserved epitopes on HA, which may play a critical role in membrane fusion, can enable the development of broad influenza vaccines and antibody-based antiviral agents. The FI6 antibody bound to a trimer of prefusion H3 HA is shown in Figure 1.16.

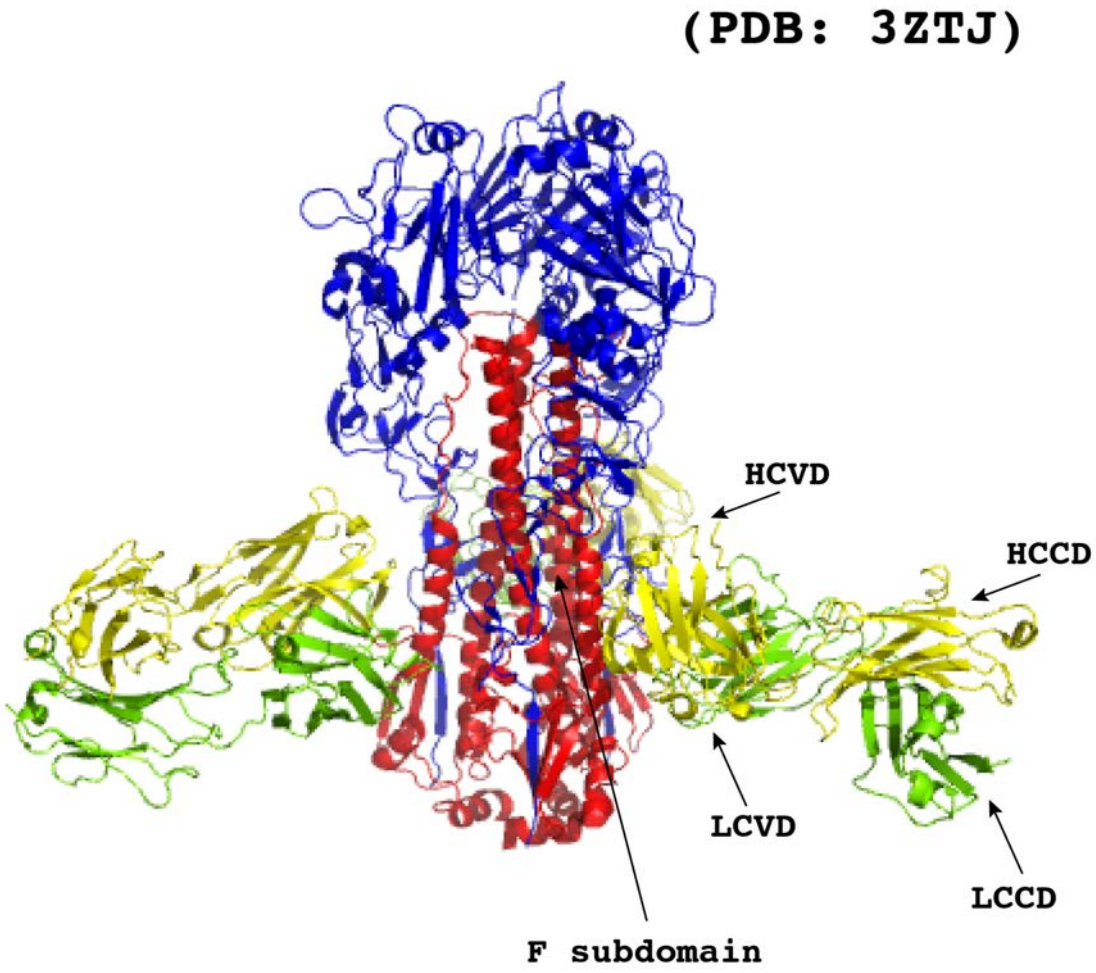


Figure 1.16 The FI6 antibody bound to a trimer of prefusion H3 HA. The antibody binds to the stem region using both light and heavy variable chains (LCVD and HCVD), crosslinking two HA monomers and interacting with helix A of one monomer and the fusion peptide of the neighbouring monomer (PDB: 3ZTJ) (Corti, Voss et al. 2011). Light and heavy constant domains (LCCD and HCCD) of FI6 are also shown. Antibody domains are shown in green and blue.

1.9 Current treatments

Infection can be prevented by acquisition of protective immunity through vaccination. Vaccine types currently available worldwide include: inactivated whole virion, split-products, purified surface antigens, and LAIV. The split-products vaccine is produced by disruption of the viral envelope with a detergent, followed by inactivation using formaldehyde. The surface purified surface antigen vaccine is composed of purified viral glycoproteins. Live attenuated influenza vaccine (LAIV) contains influenza A H1N1, influenza A H3N2 and two strains of influenza B virus (George, Farooq et al. 2010) (Kahn, Santibanez et al. 2015). The evolving nature of influenza virus, together with a reduction of vaccine-induced antibody response over time, result in a decline in vaccine-induced immunity, which last ~1 year. While immunity resulting from infection with the virus can last a lifetime (Yu, Tsibane et al. 2008), the seasonal inactivated vaccines need to be updated every year (Couch 2008), and their composition depends on the recommendations made by WHO Global Influenza Surveillance and Response System.

Antiviral drugs can reduce and shorten symptoms of viral infection when taken within two days of the onset of influenza symptoms (Montalto, Gum et al. 2000). Neuraminidase inhibitors, such as zanamivir and oseltamivir, block the receptor-destroying enzymatic activity of NA, by binding to the highly conserved enzyme active site, thus preventing the release of the newly assembled viruses from infected cells (Colman 1994). These competitive inhibitors of NA were developed using structure-based drug design, and closely resemble binding of the sialic acid. Oseltamivir (Lew, Chen et al. 2000) (Tamiflu) is taken orally, and zanamivir (Relenza) is administered by inhalation. Low frequency of viral resistance to NA inhibitors is associated with high conservation of NA active site, and resistance-causing mutations would most likely affect enzyme's activity (Ward, Small et al. 2005). Arbidol is a drug that inhibits viral membrane fusion by stabilizing the prefusion conformation of HA molecule, and thus preventing its low pH-induced structural rearrangement and blocking viral membrane fusion (Boriskin, Leneva et al. 2008) (Leneva, Russell et al. 2009).

1.10 Objectives for thesis

As described in Section 1.5, in order to assist viral membrane fusion, influenza virus HA must undergo a significant, low pH-induced structural rearrangement (Skehel, Bayley et al. 1982). The structural information concerning this conformational transition of influenza HA, has so far only been gained for subtype H3 of the 1968 pandemic virus (Bullough, Hughson et al. 1994) (Chen, Skehel et al. 1999), and this HA subtype belongs to the phylogenetic group 2 HA.

Group- and clade-specific structural differences between the prefusion HA molecules from the two groups, are the basis for classification of the emerging HA subtypes (Russell, Gamblin et al. 2004) and are summarized in Section 1.7.

Thus the primary objective for this thesis was to obtain a crystal structure of postfusion HA2 from group 1 HA. The obtained postfusion viral and recombinantly-expressed structures of HA2 representing the two phylogenetic groups will be compared to assess, whether group-specific characteristic features of prefusion HA cause structural differences in their postfusion form. Differences between the postfusion form of group 1 and group 2 HA molecules could be correlated to potential differences in function. Subtype H1 was chosen to represent group 1 HA, and crystallisation trials were performed using a low pH-activated viral, and recombinantly-expressed H1 HA2 derived from A/Brisbane/59/07 (B59) and A/PuertoRico/8/34 (PR8) H1N1 viruses.

In addition, the aim was to test the cross-reactive FI6 antibody (Corti, Voss et al. 2011) for its ability to disrupt the low pH-induced conformational change of HA, and for its effect on the infectivity of viruses containing HA from the two groups. Methods used for this part of the study included electron microscopy, circular dichroism, biolayer interferometry, limited proteolysis and *in vitro* MDCK1 cell infection.

2 Materials and Methods

2.1 Materials

2.1.1 Virus strains

1. A/Aichi/2/68 (H3N2) (X31)-recombinant influenza A virus (Kilbourne 1969). X31 is a reassortant that contains six internal genes of PR8, and the surface haemagglutinin and neuraminidase of the H3N2 virus
2. A/Brisbane/59/07 (H1N1) (B59)-influenza A vaccine strain
3. A/Puerto Rico/8/34 (H1N1) (PR8)-influenza A laboratory vaccine strain
4. A/Tokyo/3/67 (H2N2)- influenza A virus

Viruses were obtained from the WHO Influenza Centre (NIMR, London).

2.1.2 Antibodies

1. FI6–a cross-reactive anti-HA antibody obtained from Antonio Lanzavecchia (Institute for Research in Biomedicine, Bellinzona, Switzerland) (Corti, Voss et al. 2011)
2. Hc31-strain-specific anti-HA antibody obtained from Alan Douglas (NIMR, London) (Daniels, Douglas et al. 1983)

2.1.3 Reagents

Reagents used in this study were purchased from Sigma, Abgene, Fluka, GE Healthcare, Promega, Qiagen, Roche, Invitrogen, Hampton Research, Roche, Jena Bioscience and Molecular Dimensions. Specific chemicals and reagents are detailed in the relevant sections of this thesis.

2.1.4 Cell lines

MDCK1- cells from the kidney of a normal female adult Cocker Spaniel, which support growth of wide range of animal viruses. MDCK1 cells used in this study were obtained from Dr Nicole Friedrich (Marburg).

2.2 Protein expression

2.2.1 Viral H1 HA2 in a postfusion conformation

2.2.1.1 Virus infection and growth in embryonated eggs

The 9-10 day old embryonated chicken eggs were incubated at 37°C before infection with influenza viruses, diluted using 1X phosphate-buffered saline (PBS), containing 1% penicillin and 1% streptomycin, and 0.1 ml of a diluted inoculum was injected into the allantoic cavity using a 1 ml needle. The inoculation site was sealed with a melted wax/paraffin mixture, and the infected eggs were placed into an incubator, and left at 37 °C, for 48 h, at 60-65% humidity, in order to maintain moisture and prevent death of the embryos. Virus replicates in cells that make up the chorioallantoic membrane attached to the embryo, and which acts to remove soluble, insoluble and gaseous waste products of the developing embryo. New viral particles are released by budding into the allantoic fluid located in the allantoic cavity. Eggs were then placed into a cold room (4°C) for 12-24 h, according to The Code of Practice for the Humane Killing of Animals under Schedule 1 to the Animals (Scientific Procedures) Act 1986.

2.2.1.2 Virus purification

Viruses were purified by the procedure of Skehel and Schild (Skehel and Schild 1971). To harvest the virus, the top of the egg was smashed with a knife handle to open the eggshell. The top of the eggshell covering the air sac was removed with tweezers. The air sac at the rounded end of an egg is responsible for respiration and pressure adjustments. The shell and chorioallantoic membranes were pierced with a pair of tweezers, and the allantoic fluid (~10 ml/egg) was aspirated using a vacuum pipette. The allantoic fluid was clarified by centrifugation at 3,000 rpm, for 15 min, at 4°C, in a Beckman swinging bucket centrifuge. The pellet composed of red blood cells was discarded, and the supernatant containing virus particles was centrifuged overnight at 6,000 rpm, 4°C, in a Beckman Aventi J25 centrifuge (rotor F10 or JLA 10500). The supernatant was decanted, and virus pellets were recovered and resuspended PBS buffer. Resuspended virus was homogenized using a glass homogenizer, and then sonicated for 1 min in a sonication bath. Virus (up to 8 ml/gradient) was layered on top of continuous 15-40% [v/v] sucrose density gradients, and centrifuged at 25,000 rpm, for 45 min, at 4°C, using a L90K or XL90 centrifuge with a SW32Ti rotor. In these gradients, virus particles travel through the gradient, until they reach a point, at which their density matches that of the surrounding sucrose. As influenza viruses have densities of ~1.1 g/cm^3 in an aqueous solution (Sharp 1945), the density of sucrose used for virus purification was 1.06-1.17 g/cm3, corresponding to 15-40% sucrose [v/v]. Virus-containing fractions were removed with a pipette, diluted with resuspension buffer containing 10 mM Tris pH 8.0 and 150 mM NaCl, to dilute the remaining sucrose, and centrifuged at 25,000 rpm, for 90 min, at 4°C, in a SW32Ti rotor. The supernatant was decanted, and virus pellets were diluted with a resuspension buffer (10 mM Tris pH 8.0, 150 mM NaCl).

2.2.1.3 Release of HA from purified virus with bromelain

Digestion of viral particles with bromelain releases a fragment of the HA ectodomain (Brand and Skehel 1972) (Waterfield, Espelie et al. 1979) by digesting once after Gly175 of HA2, and the released molecule is a soluble trimer (Dopheide and Ward 1981). After digestion with the enzyme, and centrifugation to remove virus cores, the soluble protein is present in the supernatant, and can be purified by sucrose density gradient centrifugation, or alternatively, by size-exclusion chromatography (Waterfield, Espelie et al. 1979). This enzymatically-released molecule is a transmembrane anchor-less fragment of prefusion HA, and therefore does not aggregate in solution (Waterfield, Espelie et al. 1979). This method has been used successfully on different virus strains, and susceptibility of HA to bromelain digestion was shown to be strain-dependent.

Experimental details:

Bromelain extract from Pineapple stem (Sigma) at 2% [w/w] in the presence of 10 mM ßME was added to virus suspensions, and incubated at 37°C, for 90 min, in a water bath. Proteolytic digestions were stopped by the addition of protease inhibitor cocktail tablets. Viruses were transferred to 3 ml polycarbonate centrifuge tubes, and centrifuged at 55,000 rpm, for 10 min, at 4°C, in a Beckman TL-100 ultracentrifuge with TLA 100.3 rotor. The supernatant containing BHA was purified by sucrose density gradient centrifugation followed by ion-exchange chromatography (2.3.4), or by ion exchange (2.3.4) and gel filtration (2.3.5).

In case of PR8 H1 BHA, the supernatant containing BHA was layered on top of a continuous 5-25% sucrose gradient [v/v], in 25 mM Tris pH 8.0, and centrifuged at 38,000 rpm, for 18 hours, at 10°C. Protein-containing fractions were identified using SDS-PAGE (2.3.3), and purified using ion-exchange chromatography (2.3.4). The amount of the bromelain-released protein was estimated using SDS-PAGE (2.3.3). As different virus strains vary in susceptibility of their HA to bromelain digestion (Waterfield, Espelie et al. 1979), some protein preparations required multiple bromelain digestions. Additional digestions were performed by incubating the resuspended pellets with bromelain as previously described, until enough BHA was released from virus particles and subsequent digestions were typically ~2 h, at 37°C.

2.2.1.4 Detergent extraction

An entire haemagglutinin molecule can be released from the viral particle by disruption of the lipid envelope with a detergent (Waterfield, Espelie et al. 1979). In this case, the released molecule contains the hydrophobic transmembrane (TM) anchor located at the C-terminus of HA2, and thus aggregates in solution. In order to obtain a soluble fragment of HA for structural studies, detergent extraction is usually followed by digestion of HA with trypsin, to remove the hydrophobic TM anchor (HA2 residues 185-211), and the resulting molecule is known as THA.

Experimental details:

Purified viruses were resuspended in 10 mM Tris pH 8.0, 150 mM NaCl, and 2% [v/v] Lauryldimethylamine-oxide (LDAO) or n-Octyl-ß-D-Glucoside (BOG). Viruses were incubated with a detergent at 4°C, for 30 min. Following incubation with a detergent, virus was centrifuged at 55,000 rpm, for 10 min, at 4°C, using a Beckman TL-100 ultracentrifuge with TLA 100.3 rotor, in order to remove virus cores. The HA-containing supernatant was collected and subjected to protein purification. Detergent-released HA was layered on top of a continuous 5-25% sucrose density gradient [v/v], containing 0.1% BOG [v/v], and centrifuged at 38,000 rpm, for 18 hours, at 10°C. Protein-containing fractions were identified by SDS-PAGE (2.3.3), pooled, and purified using ion-exchange chromatography (2.3.4) with a HiTrap Mono-Q HP column (GE Healthcare). Protein samples were concentrated using 50K MWCO concentrators (Vivaspin), and digested with trypsin, at a ratio 1:10 [w/w], trypsin: protein, for 1 hour, at RT, in order to remove the TM anchor at the C-terminus of H1 HA2. Digestions were stopped with an equal amount of soybean trypsin inhibitor (Sigma), and digestion products were separated on 5-25% sucrose gradients [v/v] in 10 mM Tris pH 8.0, at 38,000 rpm, for 18 hours, at 4°C. Protein samples were dialyzed into 25 mM Tris pH 8.0, 150 mM NaCl, concentrated using 6 ml 50K MWCO concentrators (Vivaspin), and subjected to a low pH treatment as described in Section 2.3.6.

Alternatively, following detergent solubilisation, the supernatant was diluted with 25 mM Tris pH 8.0, to contain 25mM NaCl, and the TM anchor removal was carried out directly by the addition of trypsin. Depending on virus strain and batch, removal of the TM anchor would last from 1.5 h to overnight. After the addition of the soybean trypsin inhibitor, the protein-containing solution was centrifuged at

25,000 rpm, for 15 min, at 4°C, in a BV52 rotor, and the collected supernatant was subjected to ion-exchange chromatography (2.3.4) and gel filtration (2.3.5).

2.2.2 Recombinant H1 HA2 in a low pH conformation

Recombinant baculovirus (*Autographa californica* multiple nuclear polyhedrosis virus) was used to express the ectodomain of H1 HA2 in cultured Sf9 (*Spodoptera frugiperda*) cells. The cDNA encoding HA2 residues 38-175 of A/Brisbane/59/07 H1N1 HA was subcloned into a modified pAcGP67A (BD Biosciences) baculovirus transfer vector (pHAEM) that carries a TEV protease cleavage site, and a 6x His tag (Lin, Xiong et al. 2012), and subcloning was performed by GeneArt. Cys137, which forms a disulphide bond with Cys14 of HA1 (H3 numbering) (Wilson, Skehel et al. 1981), was changed to Ser for expression (to prevent aggregation of HA2) (Chen, Wharton et al. 1995), and two cysteines (Cys144 and Cys148) that form disulphide bonds within HA2 were preserved. TEV protease cleavage sequence was included between the spacer and the N-terminus of HA2 (residue 38), to allow for a removal of the 6x His tag at the extreme N-terminus of the expressed polypeptide sequence, for use in affinity purification (Figure 2.1). Recombinant baculovirus was generated by cotransfection of BD BaculoGold™ Linearized Baculovirus DNA (BD Biosciences) and recombinant plasmid into Sf9 cells. Following virus amplification under ampicillin selection, large-scale expression was performed in 2.5 L of Sf9 insect cells suspension cultures, and the expressed protein is secreted into the cell culture medium. Protein expression was kindly carried out by Dr Phil Walker at NIMR. Cells were removed by centrifugation, 72 h post-infection, and the supernatant was concentrated using a large-scale concentration unit with a 10K MWCO membrane, and loaded onto a Talon cobalt column (2.3.8). Fractions containing recombinantly-expressed Bac B59 H1 HA2 were pooled, and protein buffer was exchanged, by concentrating into 20 mM Tris pH 8.0, 150 mM NaCl. Concentrated protein was digested with TEV protease [10:1 HA: TEV w/w, RT, overnight] to cleave the spacer and the 6x His tag. The His tag was removed by passage over the Ni-NTA resin (Qiagen), and Bac B59 H1 HA2 was further purified by size-exclusion chromatography (2.3.5) using the Superdex-200, 16/60-gel filtration column (GE), equilibrated with 20 mM Tris-HCl pH 8.0, 150 mM NaCl. Fractions containing Bac B59 H1 HA2 were collected, and protein buffer was exchanged to contain 10 mM Tris-HCl pH 8.0 and 50 mM NaCl. The expressed protein was concentrated to 6mg/ml (A_{280}) for crystallisation trials (2.3.9).

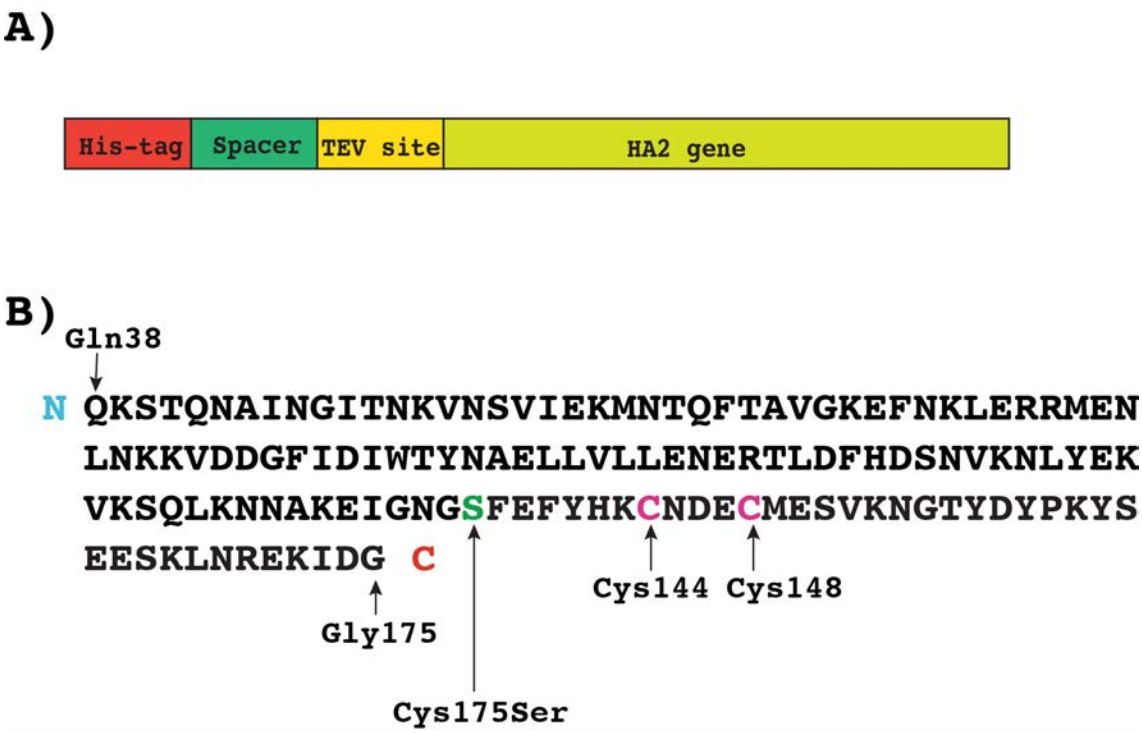


Figure 2.1 Bac B59 H1 HA2 expression construct. Construct containing the H1 HA2 gene is shown in A. The amino acid sequence of the encoded H1 HA2 fragment (HA2 residues 38-175) is shown in B. The N- and C-termini of the polypeptide fragment are indicated. Cysteines 144 and 148 that form a disulphide bond within each HA2 monomer are shown in yellow. Cysteine137, which forms a disulphide bond with HA1 and was changed to serine for expression and is shown in green.

2.3 Protein preparation

2.3.1 Quantification of viruses using SDS-PAGE

Quantification of X31 H3N2 and B59 H1N1 viruses was performed prior to biolayer interferometry experiments (2.4.6). Concentration of viruses was calculated from the estimated viral nucleoprotein (NP) contents, and standardized by comparison of the NP band intensity with the protein standard. The number of NP molecules/virion was calculated based on the influenza virion genome, consisting of 13,588 nucleotides, and binding of 23 nucleotides per one NP monomer (Ruigrok 1998). Bovine Serum Albumin (BSA) Standards (Sigma) were diluted to 200 µg/ml, and the reference samples were prepared by mixing 16, 12, 8 and 4 µl of BSA at 200 µg/ml with 10 µl of reducing 2X LDS, so the final amount of BSA was 3.2, 2.4, 1.6 and 0.5 µg, respectively. Virus stocks were diluted 2x with 20 mM Tris pH 8.0, 150 mM NaCl, and virus samples were prepared by mixing 2, 1 and 0.5 µl of viruses with 10 µl of non-reducing 2X LDS. All samples were boiled at 100˚C, spun in the centrifuge, and loaded onto Novex 4-12% Bis-Tris gels (Invitrogen). Gels were stained with InstantBlue Coomassie stain (Expedeon), scanned, and images were analysed using image-processing software (Image J). Following the addition of a correction factor, the intensity of viral NP bands was compared to that of BSA standards, and the approximate virus concentrations were calculated in nM. A typical quantification of viruses using SDS gel electrophoresis is shown in Figure 2.2.

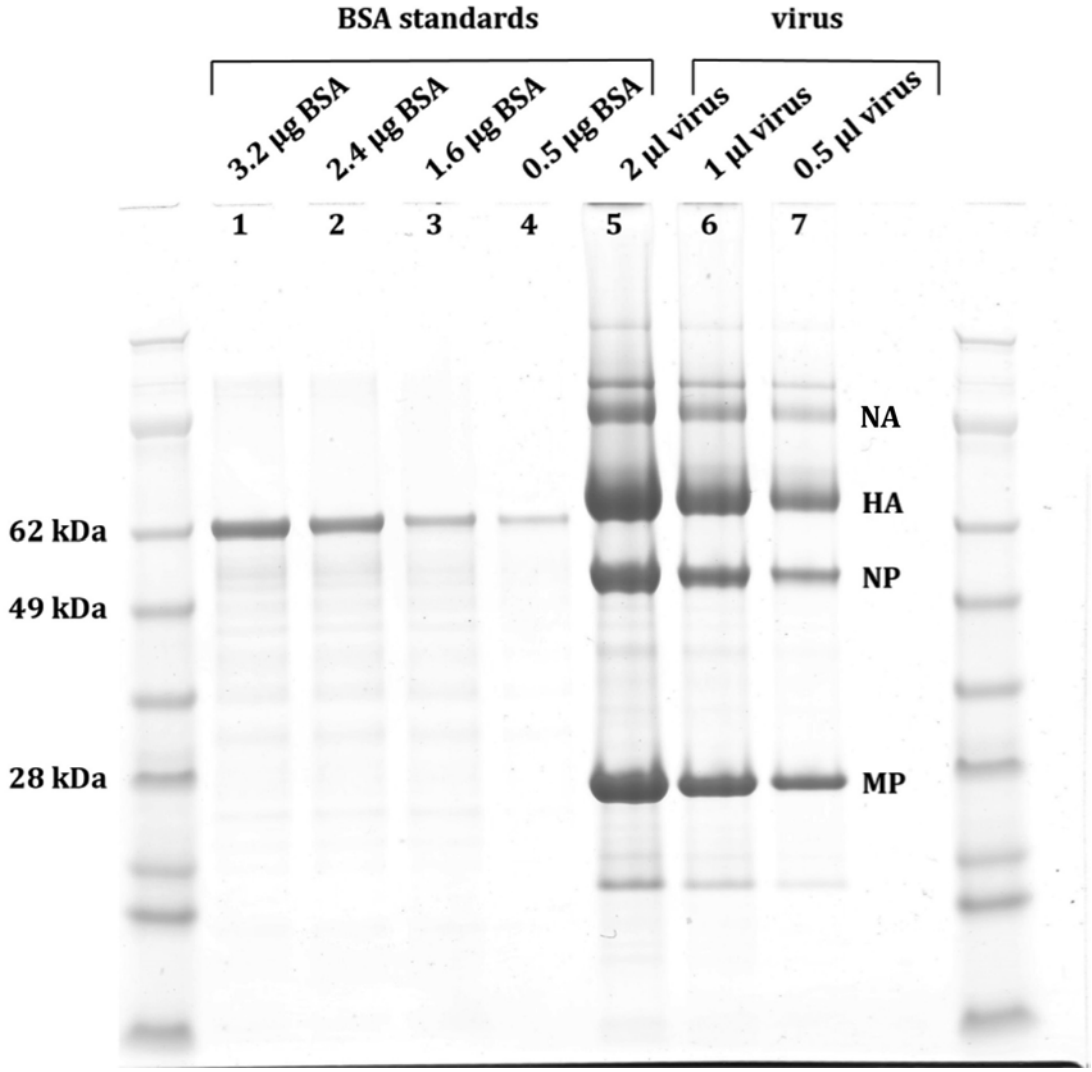


Figure 2.2 Quantification of viruses using SDS-PAGE. Lanes 1-4 contain BSA standards, and lanes 5-7 contain samples of B59 H1N1 virus. Virus concentration is determined by comparison of viral NP band intensity with the intensity of BSA standards. Details of SDS-PAGE are included in Section 2.3.3.

2.3.2 Determination of protein concentration

Protein concentration measurements were routinely carried out on the purified protein fractions, pooled, and concentrated using Vivaspin 20 50K, 30K and 10K MWCO concentrators (Sartorius) at various stages of protein preparation. Prior to measurements samples were usually centrifuged at 4,000 rpm, for 10 min, at 4°C. Protein concentration was determined using UV-Vis spectrophotometry, and the absorbance of protein samples was measured at a wavelength of 280 nm, using a NanoDrop Spectrometer (ND-1000) (Thermo Scientific).

2.3.3 SDS-PAGE

All Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis (SDS-PAGE) experiments were performed using Novex 4-12% Bis-Tris gels (Invitrogen), in 1X MES running buffer (Invitrogen), unless stated otherwise. Protein molecular weight was visualized using the SeeBlue® Plus2 Prestained Protein Standard (Invitrogen). Samples for non-reducing gels were prepared by mixing protein solutions (15 µl) with 5 µl of NuPAGE LDS sample buffer (Thermo Fisher Scientific). For reducing gels, 1 µl of 1M TCEP (tris (2-carboxyethyl) phosphine) at pH 8.0 was added to the samples. Samples were heated to 95°C, for 5 min, and loaded onto a gel. Gels were run in an XCell SureLock mini-cell (Invitrogen), at a constant 200 V, for 43 min, and stained with InstantBlue Coomassie stain (Expedeon). Throughout the thesis, gel electrophoresis images show SDS-PAGE under non-reducing (NR) and reducing (R) conditions.

2.3.4 Ion-exchange chromatography

Ion-exchange chromatography (IEX) was used in preparation of postfusion viral H1 HA2 following the release/extraction of H1 HA from viral envelope using the two methods (2.2.1.3 and 2.2.1.4). HA (pI 6.74) was selectively separated from NA (pI 6.09) and other contaminating viral proteins, using anion exchange chromatography. A HiTrap Mono-Q HP column (GE Healthcare) was first equilibrated with an equilibration buffer, containing 25 mM Tris pH 8.0 and 25 mM NaCl. Protein buffer was exchanged to 25 mM Tris pH 8.0, 100 mM NaCl, and protein sample was applied to the column using a P1 peristaltic pump (GE

Healthcare), at a flow rate of 2 ml/min. The column was attached to the AKTA purifier (GE Healthcare), and washed with 2 column volumes (CV) of binding buffer, to remove all non-bound proteins. HA was eluted by increasing the ionic strength of the elution buffer (25 mM Tris pH 8.0 and 1 M NaCl; 0-100% linear salt gradient), over a total volume of 250 ml. The 10 ml fractions were analysed using SDS-PAGE (Section 2.3.3), and HA-containing fractions were collected for further purification using size-exclusion chromatography (2.3.5).

2.3.5 Size-exclusion chromatography

Size-exclusion chromatography (SEC) was employed to separate contaminants and protein aggregates at various stages of protein preparation, as well as a final protein purification step prior to protein crystallisation. A Superdex S200 16/60 size exclusion column (GE Healthcare) was equilibrated with a gel filtration buffer, containing 25 mM Tris pH 8.0 and 150 mM NaCl, on the AKTA purifier (GE Healthcare), at a flow rate of 1.5 ml/min. Protein samples were concentrated to a volume of < 5 ml, injected onto the gel filtration column, and eluted at a flow rate of 1.5 ml/min, over a total volume of 120 ml (1.5 ml fractions). Fractions were analysed using SDS-PAGE (2.3.3), and the appropriate fractions were pooled for further analysis.

2.3.6 Low pH treatment

Background:

It is known, that the low pH treatment (~pH 5.0) results in an induction of structural refolding of influenza virus HA, which as part of this process, exposes its hydrophobic fusion peptide (FP) at the N-terminus of each HA2 monomer (residues 1-23) (Skehel, Bayley et al. 1982). Exposure of the FP subdomains of HA2 causes aggregation of multiple HA trimers (~8) into rosettes (Ruigrok, Aitken et al. 1988). Another feature of this low pH-induced refolding is the observed extension of the trimeric coiled coil (Bullough, Hughson et al. 1994) (Chen, Wharton et al. 1995) (Chen, Skehel et al. 1999), which exposes the protease cleavage sites normally buried at neutral pH, and the rearranged molecule becomes susceptible to proteolytic digestion (Skehel, Bayley et al. 1982). These two features of the observed structural reorganization of HA at low pH were exploited in preparation of postfusion viral H1 HA2, specifically in proteolytic removal of HA1 and the FP regions from the low pH-treated HA prior to crystallisation.

Experimental details:

The conformational change of viral, bromelain-released BHA (2.2.1.3) and detergent-extracted THA (2.2.1.4), was typically induced by incubation of BHA/THA samples at pH ~5.0. Incubation at low pH was preceded by purification of BHA/THA samples using ion-exchange chromatography (2.3.4) and gel filtration (Section 2.3.5). Purified BHA/THA samples were transferred to Sterilin™ 7 ml Polystyrene Bijou Containers (Thermo Scientific), and the pH of the samples was lowered to pH ~5.0 using 0.1 M citric acid, or alternatively, 0.1 M Sodium citrate pH 3.5. Following incubation of HA at pH ~5.0, the pH was readjusted to pH 7.8, using 1 M Tris pH 8.0. The progress of the expected conformational rearrangement of HA was monitored using limited proteolysis (2.3.7), and SDS-PAGE (2.3.3), utilizing the acquired susceptibility of the low pH-activated molecule to proteolytic cleavage, and samples containing the rearranged protein were pooled and subjected to further purification steps.

The procedure was also performed on purified viruses, and virus-antibody complexes prior to electron microscopy (2.4.5). The conformational change of virus-bound HA was induced by incubation of viruses and virus-antibody complexes at pH 4.8, and the pH was lowered using 0.1 M Sodium citrate pH 3.5, for 45 min, at RT. The low pH treatment was stopped by adjusting the pH of virus suspensions to pH 7.8, using 1 M Tris pH 8.0.

2.3.7 Limited proteolysis

Limited proteolysis was used at various stages of the presented research on the low pH-induced rearrangement of influenza virus HA:

- Bromelain from pineapple stem
    - to release HA from purified viruses (2.2.1.3)
- Trypsin
    - to remove the hydrophobic TM anchor at the C-terminus of HA2 (2.2.1.4)
    - to remove the HA1 domains from the refolded HA (3.4)
    - to remove the hydrophobic FP at the N-terminus of HA2 (3.5)
    - to assess the conformational state of Bac B59 H1 HA2 (4.2)
    - to assess the conformational state of virus-bound HA and to confirm the ability of the FI6 antibody to prevent the conformational change of virus-bound HA using electron microscopy (2.4.5)
- Thermolysin
    - to remove the hydrophobic FP at the N-terminus of HA2 (3.5)

Experimental details:

With the exception of thermolysin, all proteolytic digestions were preceded by adjusting the pH of protein solutions to pH 7.8, using 1 M Tris pH 8.0. Enzymes tested for the ability to remove the globular HA1 domains and the hydrophobic fusion peptide were:

- Trypsin, TPCK treated from Bovine Pancreas (Sigma). Trypsin was tested at concentrations of 0.1-10% [w/w], over 2.5 h, at RT. Samples were collected every 10 min, for the first 60 min, and every 30 min afterwards. Digestions were stopped by the addition of an equal amount of soybean trypsin inhibitor (Sigma).
- Thermolysin, from *Bacillus thermoproteolyticus* (Sigma). Thermolysin dissolved in PBS was tested at 2% [w/w], at 37°C, in the presence of 1 mM $CaCl_2$. The pH of the samples was adjusted to pH 5.0 using 0.1 M Sodium citrate, and the typical time course was 1.5-7 h. Reactions were stopped by the addition of 20 mM EDTA.
- α-Chymotrypsin, Elastase, Papain, Subtilisin and Endoproteinase Glu- C (Proti-Ace, Hampton Research). Used according to manufacturer's instructions.
- Proteinase K, Endoproteinase-Arg-C, Pepsin and Actinase E (Proti-Ace 2, Hampton Research). Used according to manufacturer's instructions.
- Endoproteinase Lys-C, sequencing grade (Promega). Endoproteinase Lys-C was used at 0.1-5% [w/w], for 30 min, at RT.
- Bromelain from pineapple stem (Sigma). Bromelain was used at 1% [w/w], for 30 min, at RT.

The Proti-Ace and Proti-Ace 2 proteases were solubilized in 10 mM Hepes pH 7.5, 500 mM NaCl, and tested at 0.1% [w/w], for 30 min, at RT. Papain, Elastase, Endoproteinase-Arg-C and Pepsin were additionally tested at 1% [w/w], for 1 hour, at 37°C. Reactions were stopped by the addition of 10 µl SDS-PAGE (10 µl).
Assessments of the conformational state of low pH-treated viral H1 HA2, and recombinantly-expressed Bac B59 H1 HA2, were carried out using 2% [w/w] trypsin (Sigma).

2.3.8 Immobilized metal affinity chromatography (IMAC)

IMAC is a group-specific affinity technique for separating proteins (Porath, Carlsson et al. 1975) based on a reversible interaction between certain amino acid side chains, here His, and immobilized divalent metal ions, in this case Co^{2+}. As Cobalt-based BD TALON™ IMAC Resin (Clontech) displays a reduced affinity for host proteins containing exposed His residues, it was used as an initial step in purification of recombinant polyhistidine-tagged Bac B59 H1 HA2.

Experimental details:

Baculovirus-infected cell extract was clarified by centrifugation at 4,000 rpm, for 30 min, at 4°C, in a Beckman swinging bucket centrifuge. The supernatant was concentrated to 100 ml using a 10K MWCO membrane. A solution containing 5 mM Imidazole pH 8.0, 10 mM Sodium phosphate pH 8.0, and 300 mM Sodium chloride was added to the cell extract. Following the addition of the complete, EDTA-free protease inhibitor tablets (Sigma), the concentrated supernatant was clarified by centrifugation at 25,000 rpm, for 30 min, using a BV52 rotor. Four 25 ml TALON Superflow Resins packed into HisTALON FPLC Cartridges (Clontech) were connected, and equilibrated with a binding buffer (5 mM Imidazole, 25 mM Sodium phosphate pH 8.0, 300 mM NaCl and 10% [v/v] glycerol). The soluble protein extract was loaded onto the Cobalt column using a peristaltic pump with a continuous flow, at a rate of 5 ml/min. The TALON column was attached to the AKTA purifier and the uncleaved protein was eluted in 10 ml fractions, using a gradient of an elution buffer, containing 500 mM Imidazole, 25 mM Sodium phosphate, 300 mM NaCl and 10% [v/v] glycerol (final pH adjusted to pH 8.0). Eluted protein fractions were analysed using SDS-PAGE (2.3.3), pooled, and concentrated, and the protein buffer was exchanged into 10 mM Tris pH 8.0, 150 mM NaCl, lowering the concentration of salt.

Polyhistidine-tag removal:

The His tag was removed using TEV protease at a ratio of 1:100 [w/w], at RT, O/N. The cleaved protein was recovered by passing the protein-containing solution through the Ni-NTA Agarose nickel-charged resin (Qiagen), pre-washed with PBS. While His tag residues bind to the nickel ions, cleaved HA passes through the affinity chromatography matrix in a gravity-flow column. His tag cleavage was confirmed using SDS-PAGE (2.3.3), and the cleaved protein was concentrated using a 10K MWCO concentrator (Sartorius) to < 5 ml, and further purified by gel filtration (2.3.5). The recombinantly-expressed Bac B59 H1 HA2 was present in 1.5 ml fractions D5-E8 eluted from the gel filtration column. The protein buffer was exchanged into 10 mM Tris pH 8.0 and 50 mM NaCl post-purification, and protein sample was concentrated to ~10 mg/ml (A_{280}) and subjected to crystallisation trials, as described in Section 2.3.9.

2.3.9 Protein crystallisation and crystal freezing

Protein crystallisation requires soluble, highly pure, homogenous, stable and concentrated (~10 mg/ml) protein samples, which in a typical crystallisation experiment are separated from a solution, in a process driven by a reduction in their solubility, driven by the addition of precipitants to the protein-containing solution. When a protein-precipitant drop is placed over a precipitant-containing well solution within a sealed compartment, concentration of the precipitant in the crystallisation drop will increase over time, in the process called vapour diffusion. This decrease in the availability of water in the equilibrating drop mimics an increase in the concentration of protein, and when its solubility limit is reached, and given correct conditions, protein molecules can separate from the supersaturated solution and form a crystal. Changes in protein solubility are temperature dependent, and can be described using a phase diagram (Figure 2.3). The solubility line defines the region below which crystallisation will not occur, and corresponds to a thermodynamic equilibrium between the concentrations of protein and the precipitating agent. Region above the solubility line is divided into three zones. Metastable zone (1) is a region, where crystals can grow however, a spontaneous nucleation does not occur, and induction of crystal growth may require seeding. Nucleation zone (2) is an area where protein molecules aggregate in a crystalline form. Precipitation zone (3) is a region where proteins separate from a solution, forming amorphous protein aggregates. Crystallisation techniques have been widely discussed in other sources (Bergfors 2009).

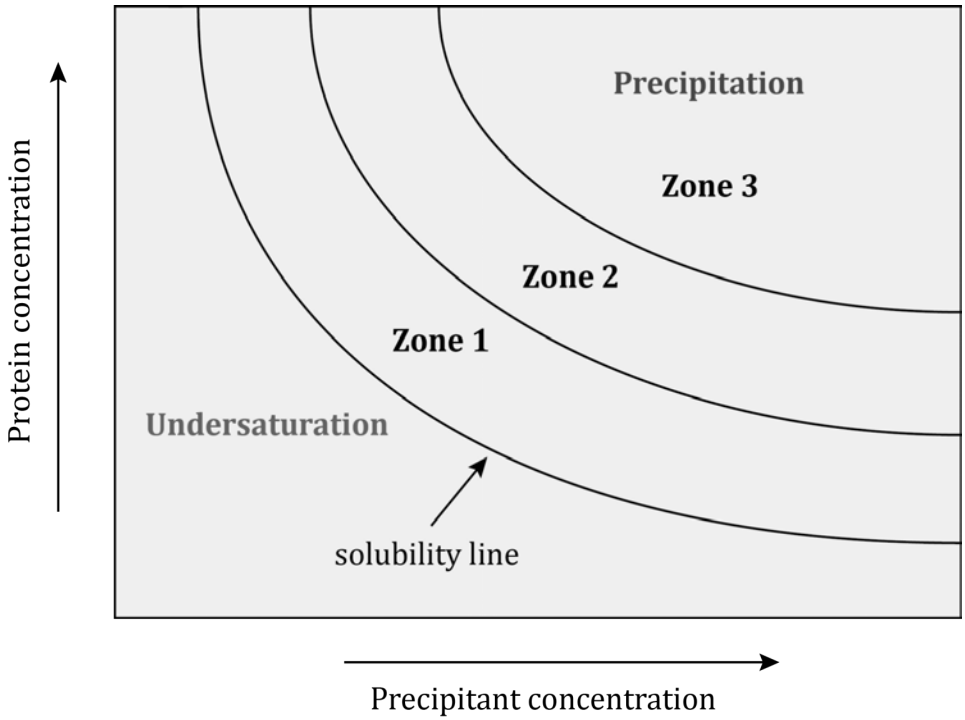


Figure 2.3 Protein solubility phase diagram for a given temperature. Zones 1-3 represent the metastable, nucleation and precipitation zones, respectively. Undersaturation and precipitation zones are indicated.

Experimental details:

Purified proteins were concentrated to ~5.5, 7.5 and ~10 mg/ml (A_{280}) using Vivaspin 6 ml and 500 µl concentrators (Sartorius), 10K MWCO, at 3,000 g, 4°C, in a benchtop centrifuge (Thermo Scientific). Homogeneity of protein samples was assessed using DLS (2.4.2) and SDS-PAGE (2.3.3). Protein crystallisation trials were set up at 18°C, using the commercially available crystallisation screens (Hampton Research, Qiagen, Molecular Dimensions and Jena Bioscience), as well as homemade screens designed by Dr Lesley F. Haire (NIMR). Proteins were crystallized using a sitting drop vapour diffusion technique. The volume of reservoir solution was typically 75 µl. Screen solutions were transferred from 96-well deep well blocks (Abgene) using the Liquidator 96 (Anachem), and dispensed into the MRC 2-well crystallisation plates (Swissci). Crystallisation droplets containing 100 nl of protein solution mixed with 100 nl of different reservoir solutions were set up using the Oryx liquid handling robot (Douglas Instruments). Crystallisation plates were sealed with Crystal Clear Sealing Tape (Hampton Research), and transferred to the Rock Imager 1000 automated imaging system (Formulatrix). The incubation temperature was set to 18°C, and high-resolution images were acquired according to the specified schedule, and over the course of a few weeks. Details of specific crystallisation trials are described in Sections 3.6 and 4.4. In order to improve the size and quality of the initially obtained crystals, the initial hits were optimised, and these trials included microseeding. Seed stocks were prepared from the initial crystalline material, by transferring the initially obtained crystals to 50 µl of the hit reservoir solution. The Seed Bead tube (Hampton Research) was vortexed for a minute, stopping to cool the seed stock on ice. Seeding was implemented robotically using the Oryx robot (Douglas Instruments), by simultaneously pipetting protein, reservoir solution and undiluted seed stocks, using a three-bore dispensing micro tip. Crystallisation drop contained 0.15 µl of protein, 0.10 µl of reservoir solution and 0.05 µl of undiluted seed stock. The reservoir solutions used in the microseeding experiments were closely related to the solutions from which the seeds were derived, typically varying the concentration of successful precipitant and in the pH of the accompanying buffer. The described procedure is based on the method of Allan D'Arcy (D'Arcy, Villard et al. 2007) (Luft and DeTitta 1999). In addition, seeding into unrelated crystallisation conditions was performed. This strategy, known as microseed matrix screening (MMS), can be used to obtain new crystal forms of different space groups, and better diffracting crystals (D'Arcy, Bergfors et al. 2014).

Crystals that grew to a suitable size of at least 50 µm in all dimensions, were selected for freezing. Crystals were cryoprotected using a two-step method, where the initial step was the addition of a cryoprotectant solution into the crystallisation drop, resulting in a cryoprotectant solution at half of its final concentration. Crystals were then transferred into the final, full-strength cryoprotectant solution. Crystals were harvested using an 18 mm Mounted CryoLoop™ - 20 micron (Hampton Research), plunged into liquid nitrogen, and stored under liquid nitrogen for data collection.

2.4 Analytical techniques

2.4.1 Haemagglutination assay

Haemagglutination assays were performed according to the standard protocol, using 0.5% suspensions of Turkey red blood cells (RBCs). Turkey red blood cells were obtained from Public Health England, under procedures regulated by the Animals (Scientific Procedures) Act 1986. In a V-bottom 96-well plate, 50 µl of 10^{-2}-10^{-5} allantoic fluid dilutions were added to 50 µl of PBS, and serially diluted 2-fold across the plate. Following the addition of 50 µl of 0.75% RBCs, the plate was incubated for 30 min, at RT, and HA titre for each virus dilution was calculated as the reciprocal of the last virus dilution where agglutination was observed.

2.4.2 Dynamic Light Scattering (DLS)

Polydispersity of protein samples was measured using DLS, which measures the size distribution profile of molecules in solution. This information was used to assess suitability of purified protein samples for the crystallisation trails (2.3.9). Prior to a DLS measurement, protein samples (~20 µl) at a concentration of ~0.5 mg/ml, were centrifuged at 10,000 g, at 20°C, for 10 min. Fluctuations in scattered light intensity due to the diffusing particles were measured using a Viscotek-802 instrument, with Omnisize 3.0 software (Malvern). The particle population was determined by calculating an average from 10 separate measurements, and samples containing a population of particles with a larger hydrodynamic radius, and indicating particle aggregation, were discarded. Samples, for which the

calculated molecular weight was within ~20% of the expected molecule size, and with polydispersity of less than 20%, were subjected to crystallisation trials.

As shown in Figure 2.4 and Table 2.1, the intensity distribution of B59 H1 TTHA2 sample shows a single peak with 100% contribution to the intensity-weighted scattering signal. The peak is slightly asymmetric, resulting in the position of the mean at 3.31nm Rh. The relative percentage standard deviation (% RSD) indicates the width of the peak relative to its height, at the half-height point. The estimated molecular weight of B59 H1 TTHA2 trimer (~63 kDa) correlates with the size of the molecule estimated using SDS-PAGE (2.3.3), and monodisperse samples were subjected to crystallisation trials (2.3.9).

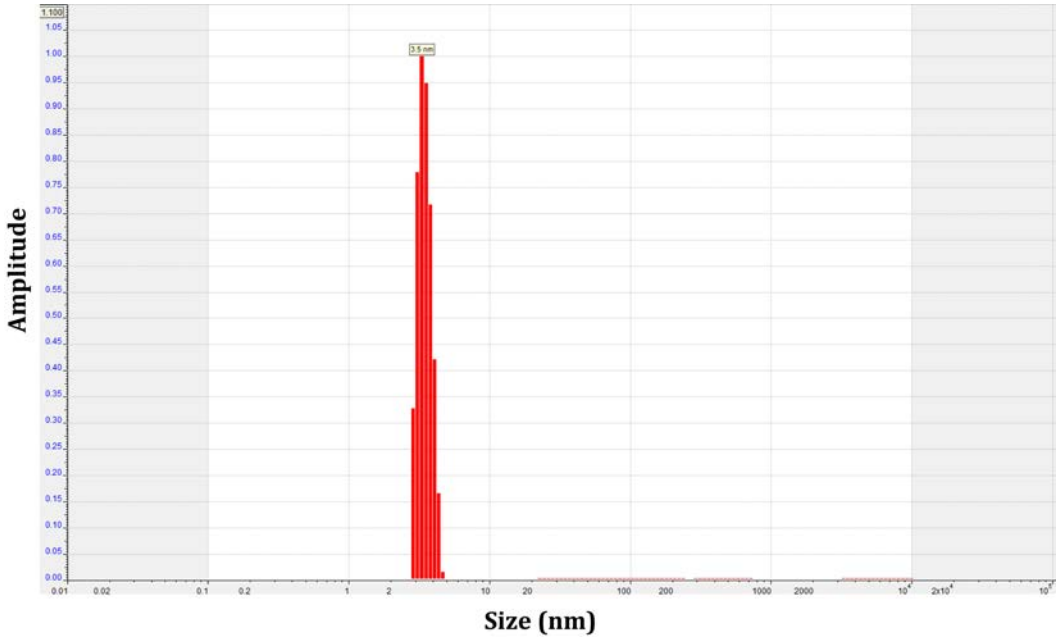


Figure 2.4 DLS profile for purified B59 H1 TTHA2. Mass distribution profile was derived from the raw intensities recorded during a DLS measurement using Viscotek-802.

Table 2.1 Summary of a DLS measurement using B59 H1 TTHA2.

Peak	% Area	Rh (nm)	Position	% RSD	MW (kDa)
1	100	3.49	3.31	10.9	63.41

2.4.3 Circular dichroism (CD) spectroscopy

CD measurements were used for estimation of the secondary structure content, and for assessment of thermal stability of Bac B59 H1 HA2, described in Chapter 4. Bac B59 H1 HA2 samples (2.2.2) were purified using IMAC (2.3.8) and diluted to ~0.15 mg/ml in two different dilution buffers. The two dilution buffers were:

- 10 mM phosphate pH 7.2, 150 mM NaCl
- 10 mM phosphate pH 5.2, 150 mM NaCl (titrated with citric acid)

CD measurements were performed using Jasco J-715 spectropolarimeter. The far-UV spectra were recorded at λ 195-260 nm, at 10°C and 25°C, using a 2 mm path length quartz cuvette, and the acquired spectra were used to estimate the secondary structure content of the expressed protein. Thermal denaturation measurements were performed at a fixed λ of 222 nm, over a 40-100°C range, at a rate of 2 °C/min. The final spectra were averages of 10 separate scans. The CD spectra were acquired and processed with the help of Dr Elizabeth Underwood (NIMR), and the results were analysed using software written by Dr Steve Martin. The recorded intensities (S) for far-UV data were expressed as:

$$\Delta\varepsilon_{MRW} = S \times MRW / 32980 \times C_{mg/ml} \times L$$

Where:

$\Delta\varepsilon_{MRW}$ *- mean residue CD extinction coefficient*

S - CD signal (millidegrees)

MRW - Mean Residue Weight (Da) $= \frac{molecular\ weight\ M_w}{number\ of\ peptide\ bonds\ N}$

$C_{mg/ml}$ *- concentration (mg/ml)*

L - path length (cm)

2.4.4 N-terminal sequencing

N-terminal sequencing was used to identify trypsin digestion products of PR8 H1 HA. Protein samples were digested with trypsin at 2% [w/w], as described in Section 2.3.7, and separated using SDS-PAGE (2.3.3). A ProBlott™ PVDF membrane (Applied Biosystems) was soaked in MeOH, and placed in a dish containing the electroblotting buffer (1X 100 mM CAPS, pH 11 in 10% v/v MeOH). The 4-12% Novex Bis-Tris gels (Invitrogen) were removed from the electrophoresis cell, and soaked in the electroblotting buffer for 5 min. The transblotting sandwich was assembled, and blotting of proteins onto the membrane was carried out using a BioRad Mini Trans-Blot Electrophoretic Transfer Cell, at a constant voltage of 50 V, RT, 45 min). The ProBlott™ was then removed from the transblotting sandwich, and rinsed with deionized water, prior to staining. The membrane was saturated with 100% MeOH for a few seconds and stained with 0.1% Coomassie® Blue R-250, in 40% MeOH/1% acetic acid [v/v]. The membrane was removed from the staining solution, destained using 50% [v/v] MeOH, and rinsed with deionized water. The membrane was dried between two sheets of filter paper and sent for N-terminal sequencing (Pepceuticals Ltd., Leicester, UK), and the obtained results are summarized in Section 3.3 and Figure 3.5.

2.4.5 Electron microscopy (EM)

Background:

The previously observed structural changes in molecular structure of HA at the pH of membrane fusion, include elongation and thinning of HA in its stem region, resulting from dissociation of HA1 domains and relocation of the amino-terminal fusion peptides of HA2 (residues 1-23) towards the target membrane (Ruigrok, Wrigley et al. 1986) (Ruigrok, Aitken et al. 1988). Exposure of these hydrophobic subdomains of HA2 causes aggregation of several HA trimers (~8) into rosettes (Skehel, Bayley et al. 1982) (Doms, Helenius et al. 1985). A radius of ~15 nm (Ruigrok, Aitken et al. 1988) and a terminal knob (Ruigrok, Martin et al. 1986) (Ruigrok, Wrigley et al. 1986) have been previously observed using EM. Removal of HA1 domains and the hydrophobic fusion peptides (HA2 residues 1-23) by limited proteolysis have also been previously observed using EM, and removal of these regions with trypsin or thermolysin resolubilizes the ectodomain of viral HA, and is

manifested by dissociation of rosettes (Daniels 1983). Dissociation usually starts after 10 min of incubation with thermolysin (2% w/w), and proteolytic digestion was previously shown to be complete after 40 min (Ruigrok, Aitken et al. 1988). In this thesis, EM was one of the techniques used to examine the effect of the FI6 antibody (Corti, Voss et al. 2011) on the low pH-induced conformational change of both virus-bound, and isolated viral HA of H1 and H3 subtypes. The obtained results are discussed in Sections 6.1.2 and 6.2, respectively.

Experimental details:

Virus-antibody complexes were prepared by mixing 50 µl of purified X31 (H3N2) and B59 (H1N1) viruses at ~10 mg/ml with 150 µl of the FI6 antibody at 9.8 mg/ml, O/N, at 4°C. A buffer containing 150 µl of 20 mM Tris pH 8.0, 150 mM NaCl was added to the samples, which were then pelleted at 55,000 rpm, for 10 min, at RT, to remove the excess antibody. Pellets were gently resuspended to the final volume of 350 µl, using 20 mM Tris pH 8.0, 150 mM NaCl. The pH of viral suspensions was lowered to pH 4.8, using 0.1 M Sodium citrate pH 3.5, for 5 min, at 37°C. The pH was then readjusted to pH 7.8, using 1 M Tris pH 8.0. The virus-antibody complexes were digested with 2% trypsin [w/w], for 45 min, at RT, and reactions were stopped by the addition of the same amount of the soybean trypsin inhibitor. Conformation of virus-bound HA was assessed using both SDS-PAGE (2.3.3) and EM. Samples were taken for analysis at each step of virus-antibody complex preparation, and control virus samples (without antibody) were included.

The effect of the FI6 antibody binding on the structural reorganization of HA at low pH by EM was also studied using detergent-extracted molecules at 0.5 mg/ml (A_{280}). Various FI6 concentrations were tested, and the optimum concentration determined by observation was 0.16 mg/ml.

All samples for EM were prepared on copper discs (grids) cast with a fine mesh. A thin layer of carbon was deposited onto mica sheets by evaporating carbon graphite, using the EMITECH K950X High Vacuum Carbon Evaporator with a small graphite rod. Current applied across the graphite rod caused it to incandesce, and carbon was evaporated off the rod and deposited on the surface. Protein samples were pipetted between the mica and carbon film, and placed (carbon film on top) into a float, containing 1% neutral SST (Sodium silicotungstate) negative stain. Sticky copper grids prepared by soaking in chloroform with sellotape were placed on top of the carbon film, and grids were blotted dry using filter paper, placed into vacuum sample holder and viewed using a JEOL 1200 EX TEM. Digital images of protein samples and protein-antibody complexes were recorded using

TECNA1, with 150000x magnification. EM experiments were carried out with the help of Dr Lesley Calder, and the obtained results are included in Section 6.

2.4.6 Biolayer interferometry

Biolayer interferometry measures the interference from the reflection of white light, as a function of the optical layer thickness at the biosensor surface. A shift in the interference pattern is caused by binding of molecules to the tip of the biosensor, resulting in a change in the wavelength (λ) of light. The technique was used to measure binding of H1N1 and H3N2 viruses to human and avian receptor analogs in the presence of the FI6 (Corti, Voss et al. 2011) and Hc31 (Daniels, Douglas et al. 1983) antibodies. In order to account for the host-determined preference for sialic acid receptors, sugars with both linkages to galactose (α-2,3- or α-2,6-) were used, and measurements were recorded using the Octet RED biolayer interferometer (Pall ForteBio Corp.). The technique was used to quantize binding of the two influenza viruses to the immobilized sialic acid receptor analogs in real time, and the obtained results are shown in Section 6.3.

Experimental details:

Biotinylated α-2,3- (α-2, 3-SLN) and α-2,6- (α-2, 6-SLN)- linked sialyl lactosamine sugars linked to a polyacrylamide backbone and composed of 30 kDa polymers, containing 20% sugar and 5% biotin [w/w], were obtained from Lectinity Holding, Inc. (Moscow, Russia). The polymers were immobilized on streptavidin biosensors (ForteBio), at a concentration of ~0.5 µg/ml. Four sensors were used to test the effect of each antibody (FI6 and Hc31) on binding of viruses to the sialic acid receptor analogs. The streptavidin sensors (ForteBio) were rinsed with water, and sugars at concentrations of 0.01-0.5 µg/µl were applied onto the sensors for 5 min. The sensors were washed with a buffer containing 10 mM Hepes pH 7.4, 150 mM NaCl, 3 mM EDTA and 0.005% [v/v] Tween 20. In order to prevent viral NA from cleaving the sensor-bound receptor analogs, oseltamivir carboxylate at 10 µM (Roche, Welwyn Garden City, UK), and zanamivir at 10 µM (GSK, Stevenage, UK) were added to the buffer. The initial concentrations of viruses were 62 nM for X31 (H3N2), and 34 nM for B59 (H1N1), as determined using virus quantification based on SDS-PAGE (2.3.1). Viruses were diluted to 0.1-1 nM using the same buffer, and after an initial measurement, the desired concentration of viruses was determined to be ~100 pM. Binding of the two viruses was measured in the

presence of the two antibodies at different concentrations. Stock concentrations of the FI6 and Hc31 antibodies were 8.5 and 14.6 mg/ml, respectively. The FI6 and Hc31 antibody solutions (2-8 µl) were added to the wells containing a buffer, α-2,3- and α-2,6- sugar analogs and virus samples. Contents of the wells were mixed, and measurements were repeated. Binding of viruses to the immobilized sialic acid analogs was measured in real time, over 30-50 min association step, at 25˚C. The relative amounts of viruses bound to the sensors in the presence of varying concentrations of the two antibodies, were calculated from the amplitude of the response at the end of the association step, and normalized by dividing by the maximum response (usually 5-6 nm). The normalized response was plotted as a function of antibody concentration. Biolayer interferometry experiments were performed with Dr Steve Martin, and the obtained results are summarized in Section 6.3.

2.4.7 MDCK1 cell infection

2.4.7.1 Estimation of the amount of infectious virus in the prepared samples using the plaque assay

The ability of the cross-reactive F16 antibody (Corti, Voss et al. 2011) to neutralize the infectivity of the X31 (H3N2) and B59 (H1N1) viruses *in vitro* was tested using the MDCK1 cell infection assay. A parallel cell infection assay using the H3 specific Hc31 antibody (Daniels, Douglas et al. 1983) was carried out as a control. Cells were infected with viruses in the presence and absence of the two antibodies, and viral replication was assessed by quantification of viral NP production. The X31 and B59 viruses from frozen stocks were diluted in a ten-fold dilution series, from 10^{-1} to 10^{-5} in PBS buffer, containing 100 U/ml Penicillin-Streptomycin (Sigma), and viruses were propagated as described in Section 2.2.1.1. Allantoic fluid was harvested, and centrifuged at 3,000 rpm, for 30 min, at 4°C. The supernatant was collected, and HA titre was determined using the HA assay (2.4.1). The allantoic fluid samples (0.5 ml) with best HA titres were stored at -80°C to be used in future experiments. Confluent monolayers of MDCK1 cells (70-80%), grown in 6-well or 12-well plates were washed with a pre-warmed PBS to remove the foetal calf serum (FCS). Cells were then washed again three times with 200 µl of virus growth medium (VGM), and 50 µl of VGM was dispensed into the plate. Cells were infected with 50 µl of allantoic fluid (10^{-1}-10^{-8}) dilutions, for 3 hours, at RT. Plates were placed briefly onto a platform shaker (30-50 rpm), every 10 min, to avoid drying.

Compositions of media used in the plaque assay are shown in Table 2.2. Following removal of the inoculum, cells were overlaid with an avicel overlay (200 µl), containing 1.25 µg/ ml trypsin (Sigma) (2.5 ml/well for a 6-well plate, and 1 ml/ well for a 12-well plate), and incubated at 37°C, 5% CO2, for up to 22 hours. Avicel (FMC Biopolymer) provides a semi-solid environment, and prevents the spread of the virus to the surrounding cells. The overlay was prepared by mixing the plaque medium and avicel 1:1 [v/v], with the addition of 1.25 µg/ml [w/v] trypsin, TPCK treated (Sigma). While avicel overlay enables quantification of individual plaques, addition of trypsin is essential for cleavage of HA into HA1 and HA2, and therefore infectivity of influenza virus. The overlay was removed at various time points, and cells fixed using 4.1% [v/v] formaldehyde (Fisher) in PBS, for 30 min, at 4°C. Initially, the number of plaques at different time points was assessed by staining with toluidine blue (Sigma), which was added to 4.1% [v/v] formaldehyde in PBS, at 0.2% [v/v]. Plaques were stained overnight at RT. Stain-containing solution was then removed, and plaques counted using a light microscope.

Having determined the appropriate concentration of X31 and B59 viruses to be used in the cell infection assay (based on the ability to resolve individual plaques), the experiments were repeated using determined time points. After removal of the avicel overlay, cells were fixed, washed three times with PBS, and incubated at 4°C, O/N. In order to detect viral NP, cells were permeabilized with 0.2% Triton-100 [v/v] in PBS (100 µl/well, 30 min, RT, shaking), and washed with PBS, containing 0.1% Tween 20 [v/v]. Plaques were then visualized by immunostaining for influenza NP (Table 2.3). Primary and secondary antibodies used for immunostaining were prepared in ELISA buffer, containing 10% horse serum and 0.1% Tween 80 [v/v]. Plaques were visualized using the HRP on tetra methyl benzidine (TMB) based substrate (TrueBlue™), counted (Sullivan, Kloess et al. 2012), and their size determined to 0.1 mm by Dr Yan Gu, using a house-assembled scanner, with IX70 Olympus microscope and Surveyor software, by counting the number of pixels in positively stained areas of 20 single NP positive cells, and determination of a mean pixel number.

Table 2.2 Composition of media and reagents used in the plaque assay.

Plaque medium (500 ml)	100 ml 10X MEM (Sigma) (2X final)
	5 ml 200mM Glutamax (Sigma) 5 ml Penicillin-Streptomycin (Sigma)- (10000 U/ml penicillin, 10 mg streptomycin/ml) 25 ml 7.5% [v/v] Na2CO3 10 ml 7% [v/v] Bovine serum albumin (BSA) 355 ml dH2O
Avicel	2.4% [w/v] avicel (FMC Biopolymer) in dH2O
MDCK cell growth medium	Minimum essential medium (MEM) (Sigma)
	10% [v/v] Foetal Calf Serum (Perbio) 100 U/ml Penicillin/Streptomycin (Sigma) 100 U/ml Glutamax (Sigma)
Virus growth medium (VGM)	Minimum essential medium (MEM) (Sigma) 0.14% [w/v] Bovine serum albumin (BSA) 100 U/ml Penicillin/Streptomycin (Sigma) 100 U/ml Glutamax (Sigma)

Table 2.3 Immunostaining for influenza NP.

Immunostaining for influenza NP	
1- addition of primary antibody	mouse monoclonal influenza A NP-specific (AbCam) 1:4000 [v/v],1 h, RT, 50 µl/well
2- cell washing	3X with PBS, 0.1% [v/v] Tween 20
3- addition of secondary antibody	goat anti-mouse IGG, HRP conjugate (Biorad) 1:1000 [v/v], 1 h, RT, 50 µl/well
4- cell washing	3X with PBS, 0.1% [v/v] Tween 20
5- substrate addition	TrueBlue™ Peroxidase Substrate (KPL) 1:1000 [v/v]in 30% Hydrogen peroxide, 10 min, RT
6-substrate removal and plate drying	2X 200 µl H20/well

MOI calculation:

The appropriate multiplicity of infection (MOI) to be used in influenza virus cell infection assays was previously determined to 0.1 PFU/cell by Dr Yan Gu, and after determination of the infectious virus using the plaque assay, the number of PFU/ml of virus inoculum was calculated. A well of a 96-well plate containing a confluent monolayer of MDCK cells was previously determined to have ~5.6 x 10^4 cells by Dr Saira Hussain (NIMR). The concentration of viruses at 10^{-6} dilution, required for preparation of 50 µl of inoculum/well was calculated, and inocula were prepared by diluting virus stocks with serum-free media. MOI calculations were performed according to the example:

Determination of the infectious virus titre of X31 by plaque assay:

Average number of plaques from replicate wells at 10^{-6} dilution (PFU): 17
200 µl of inoculum used to infect a 12-well plate contains 17x 10^6 plaques
To calculate PFU/ml: 17 x 10^6 x 5= 8.5 x 10^7 PFU/ml
Infect cells at MOI= 0.1 PFU/cell
Confluent monolayer of cells has ~5.6 x 10^4 cells/ well
Therefore, for 0.1 PFU/ml, 5.6 x 10^3 PFU/ well
For a 96-well plate use 50µl inoculum: 5.6 x 10^3 PFU/ 50 µl
As the infectious X31 stock has a titre of 8.5 x 10^7 PFU/ml, therefore 50 µl inoculum has 8.5x 10^7/ 20= 0.425 x 10^7 PFU/ 50 µl.
For MOI of 0.1 PFU/ cell, the X31 virus needs to be diluted 0.425 x 10^7/ 5.6 x 10^3= 1:760.
Virus was diluted 1:10 (100 µl virus: 900 µl serum-free medium) initially, and then 1:76.

2.4.7.2 Infection of MDCK1 cells in the presence of FI6 and Hc31 antibodies

The MDCK1 cells were seeded into 96-well plates containing the cell growth medium. The X31 (H3N2) and B59 (H1N1) viruses were pre-incubated with serial dilutions of the FI6 and the Hc31 antibodies, for 1 h, at RT. The initial concentration of the FI6 antibody was 8.5 mg/ml, and the antibody was diluted 1:2 with virus growth medium to the final concentration of 4.25 mg/ml. The initial concentration of the Hc31 antibody was 14.6 mg/ml, and the antibody was diluted 1:100 to the final concentration of 0.146 mg/ml. Virus growth medium (100 µl) was dispensed into the deep well blocks. Final antibody stocks (50 µl) were added to the first well, and serially diluted across each plate. The X31 and B59 viruses were added to the wells (150 µl), for 1 hour, at RT. Viruses were pre-incubated with the two antibodies and then used for infection of MDCK1 cells. Cells were infected with 100 µl of viruses pre-incubated with the two antibodies, in a series of two-fold dilutions across the plate. Viral control was prepared by mixing 150 µl of virus growth medium with 150 µl of viruses. Plates were incubated at 37°C and 5% CO_2, fixed at different time points (3, 4 and 6 h), and immunostained for influenza NP (Table 2.3). Cell infection assays were performed with Dr Yipu Lin. Inhibition of viral infection following the addition of the two antibodies was compared using the average number of NP positive cells at the corresponding antibody dilution. Cell counting analysis was performed by Dr Yan Gu, and the obtained results are shown in Section 6.4.

3 Crystal structure of viral H1 TBHA2

3.1 Virus growth and purification

Details of virus infection, growth in embryonated eggs, purification and isolation of viral H1 HA from purified viruses are included in Section 2.2.1. Typically, the yield of virus from 2000 eggs was 80-100 mg.

3.2 Isolation of H1 HA

***Bromelain release*:**

The TM anchor-less H1 BHA was typically released from the H1N1 viruses using 2% [w/w] bromelain. Release of H1 BHA required several bromelain digestion steps, and majority of BHA was released only after a second bromelain digestion (2.2.1.3). The released soluble fragment was purified using sucrose density gradient centrifugation. The supernatant containing H1 BHA was layered on top of a 5-25% [v/v] sucrose density gradient, and centrifuged at 38,000 rpm, for 18 h, at 10°C, in a Beckman XL-90 ultracentrifuge with a SW41 rotor. A tube connected to a peristaltic pump was gently inserted into the gradient, and 1 ml fractions were collected starting from the bottom of each sucrose density gradient. Purification of H1 BHA by sucrose density gradient centrifugation is shown in Figure 3.1. The protein-containing fractions were pooled and subjected to ion exchange chromatography (2.3.4) in order to remove bromelain. Protein eluted at 150-200 mM NaCl, and the supernatant containing H1 BHA was further purified by ion-exchange chromatography (2.3.4) followed by gel filtration (2.3.5). The yield of H1 BHA obtained using the enzymatic method was typically ~15 mg from ~100 mg of virus.

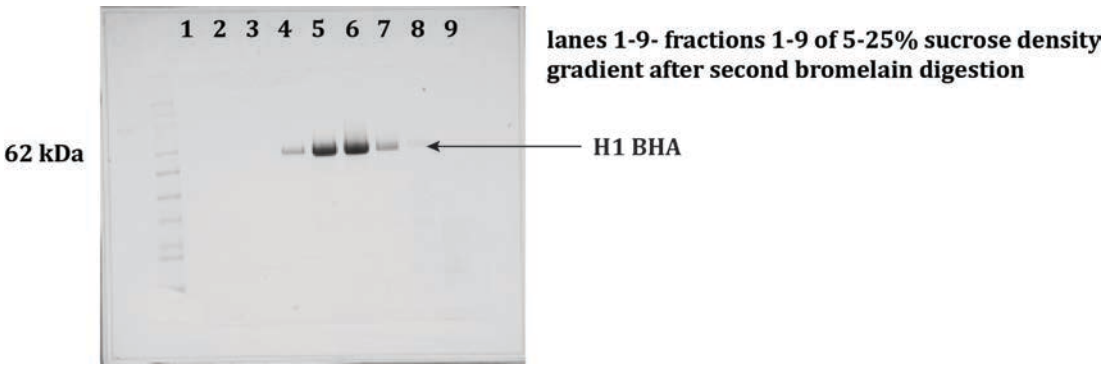


Figure 3.1 Purification of H1 BHA by sucrose density gradient centrifugation. Shown are 1ml fractions 1-9 of 5-25% [v/v] sucrose density gradient (10ml) analysed by non-reducing SDS-PAGE. Details of SDS-PAGE are included in Section 2.3.3. Soluble H1 BHA was purified from fractions 4-7.

***Detergent extraction*:**

Detergent extraction of viral H1 HA with LDAO or BOG (Section 2.2.1.4) was typically followed by purification of the released protein using detergent-containing sucrose density gradients. Detergent-released H1 HA was purified from 1 ml fractions (4-7) of 5-25% [v/v] sucrose density gradients at 38,000 rpm, 18 h, and 10°C.

Transmembrane anchor removal

Prior to the removal of the hydrophobic TM anchor with trypsin, the HA-containing fractions (4-7, Figure 3.2, A) were pooled, and H1 HA were further purified by ion-exchange chromatography (2.3.4). The TM insert was successfully removed with trypsin (1:10 w/w, trypsin: A) after 1 h, at RT, and digestions were stopped with an equal weight of soybean trypsin inhibitor (Sigma). Removal of the hydrophobic TM anchor is very difficult to observe using SDS-PAGE under non-reducing conditions, due to the small size of this subdomain of HA2. The progress of TM anchor removal can only be observed under reducing conditions. The observed decrease in the size of the band corresponding to HA2 correlates with the removal of the ~ 3 kDa TM anchor (Figure 3.2, B, lane 4), and the TM-less HA molecules (H1 THA) were further purified by either sucrose density gradient centrifugation (5-25% sucrose gradient without detergent, at 38,000 rpm, 18 h and 10°C), or ion-exchange chromatography (2.3.4) and gel filtration (2.3.5). In case of sucrose density gradient centrifugation, the final preparation step involved removal of sucrose from the H1 THA sample by centrifugation and buffer exchange. Alternatively, the TM anchor was removed by adding trypsin to the diluted supernatant, directly after

solubilisation of the virus with a detergent. H1 THA was then subjected to ion-exchange chromatography (2.3.4) and gel filtration (2.3.5). As this method proved to be very effective and less time consuming, it was used routinely in H1 THA preparations. Figure 3.2, C, shows a H1 THA band of ~63 kDa. Experimental details related to detergent extraction can be found in Section 2.2.1.4. For protein released from the virus using detergent extraction, the yield of the protein was higher, typically ~40 mg. The two methods for isolation of viral H1 BHA/THA were used with both, PR8 and B59 H1N1 viruses, and in combination with various purification techniques used in this study.

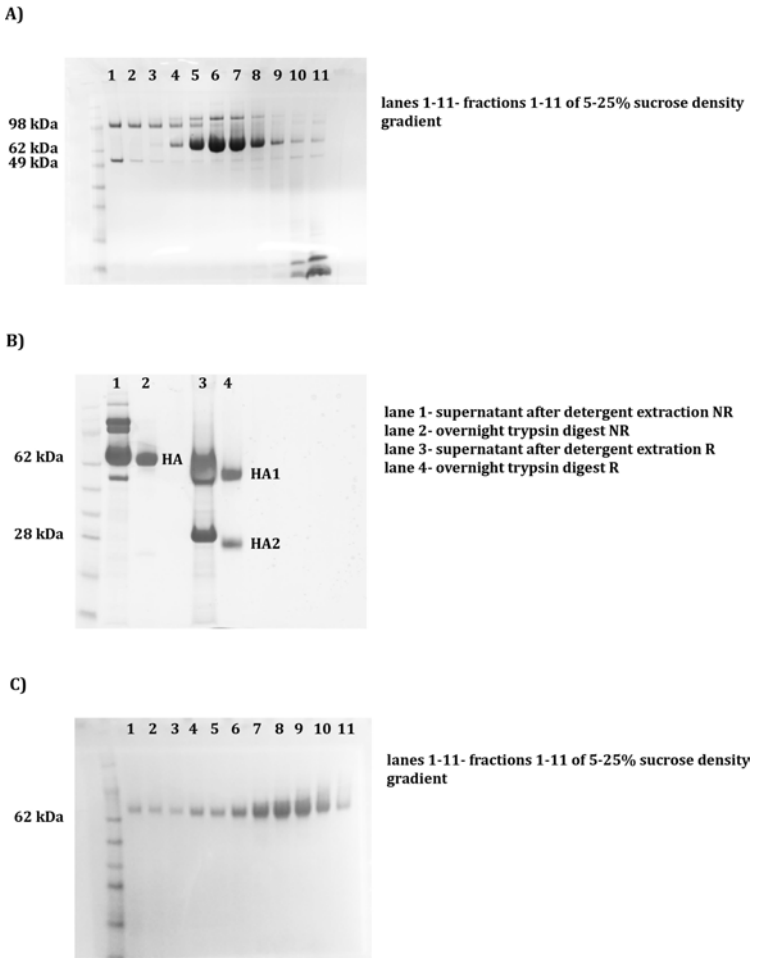


Figure 3.2 Release of viral H1 HA with BOG. The release of B59 H1 HA with 2% BOG is shown in A. B59 H1 HA is present in fractions 5-8 (1 ml) of a 5-25% sucrose density gradient. The TM anchor removal is shown in B. C shows purification of B59 H1 THA (fractions 6-10 of a 5-25% v/v sucrose density gradient). Details of non-reducing (NR) and reducing (R) SDS-PAGE are included in Section 2.3.3.

3.3 Induction of H1 BHA conformational change by incubation at low pH

Low pH treatment:

Conformational change of H1 HA was induced by incubation at pH 5.0 (Section 2.3.6), and the release of the hydrophobic fusion peptides at the N-termini of HA2 (residues 1-23) is known to cause aggregation. The aggregated molecules are normally present in the bottom fractions of a 5-25% [v/v] sucrose density gradient, with a 1ml sucrose cushion at 60% [w/v] (Bullough, Hughson et al. 1994). This FP-mediated aggregation of the TM anchor-less molecule was not observed for PR8 H1 BHA. Only a fraction of PR8 H1 BHA molecules aggregated in the bottom fractions of a sucrose density gradient. However, more protein was present in middle fractions, where soluble BHA would normally be present (Figure 3.3). This suggested that either not all of the PR8 H1 BHA molecules have undergone the low pH-induced conformational change, or that the protein was contaminated with proteases that cleaved the N-terminal fusion peptide of HA2 (residues 1-23) after exposure of the molecule to low pH (~5.0). In order to compare the unusual solubility of the low pH-treated PR8 H1 BHA to that of other HA subtypes, the conformational change of HA of subtypes H3 (X31 strain) and H2 (Tokyo strain) was induced in a similar manner, and the low pH-activated proteins were separated on a 5-25% sucrose density gradients, with a 1ml 60% [w/v] sucrose cushion. Figure 3.3 shows that in contrast to PR8 H1 HA, the low pH-treated X31 H3 BHA (A) and Tokyo H2 BHA (B) molecules aggregate at the bottom of a continuous 5-25% sucrose density gradient as previously expected.

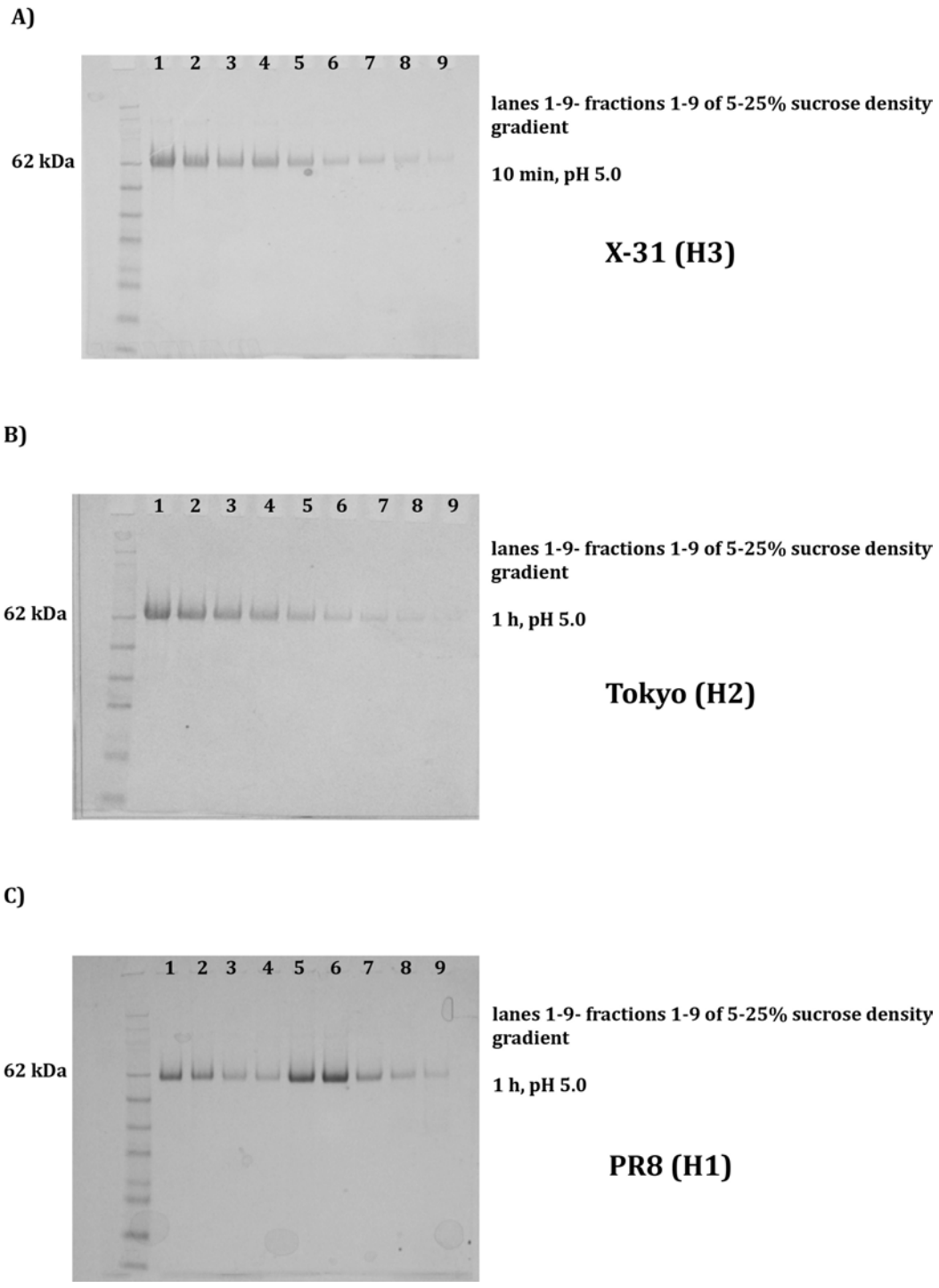


Figure 3.3 Induction of the conformational change of influenza virus HA. Figures A-C show 1 ml fractions of a 5-25% v/v sucrose density gradient (with a 60% w/v sucrose cushion), containing low pH-treated X31 H3 BHA (A), Tokyo H2 BHA (B) and PR8 H1 BHA (C), analysed by SDS-PAGE. The H3 and H2 rosettes aggregate in the bottom fraction of a sucrose density gradient (A and B). C shows the presence of both, aggregated (fractions 1 and 2) and soluble (fractions 5 and 6) PR8 H1 BHA, after low pH activation. Details of SDS-PAGE are included in Section 2.3.3.

Determination of a conformational state of low pH-treated H1 BHA by limited proteolysis:

In order to assess the conformational state of bromelain-released, low pH-treated PR8 H1 BHA (Figure 3.3, C) the protein was tested for its susceptibility to proteolytic cleavage. Trypsin digestions were carried out both before and after the low pH treatment (pH ~5.0). Limited proteolysis was carried out using 10:1 BHA: trypsin [w/w], for 20 min, at RT. Results confirmed, that prior to the low pH treatment, H1 BHA was not susceptible to trypsin digestion, and that protein became susceptible to trypsin only after incubation of the sample at pH ~5.0. Trypsin digestion products of low pH-treated H1 HA are present in fractions 11-15 (0.5 ml) of a continuous 5-25% [v/v] sucrose density gradient (10 ml) (Figure 3.4, B). This observation was consistent with the results previously reported for H3 BHA (Ruigrok, Martin et al. 1986), and suggested, that the obtained protein was in fact a fusion peptide-less PR8 H1 TBHA2.

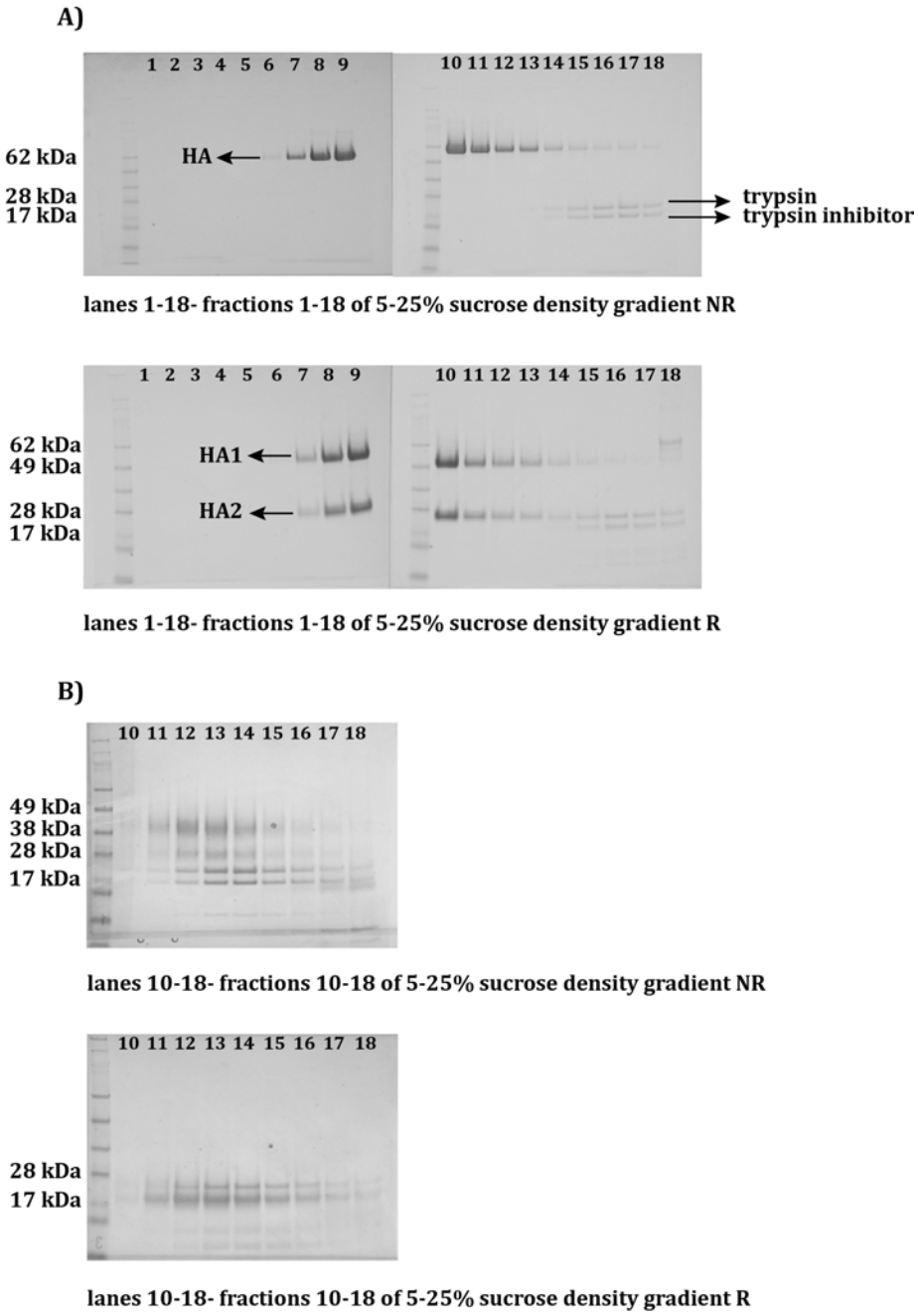
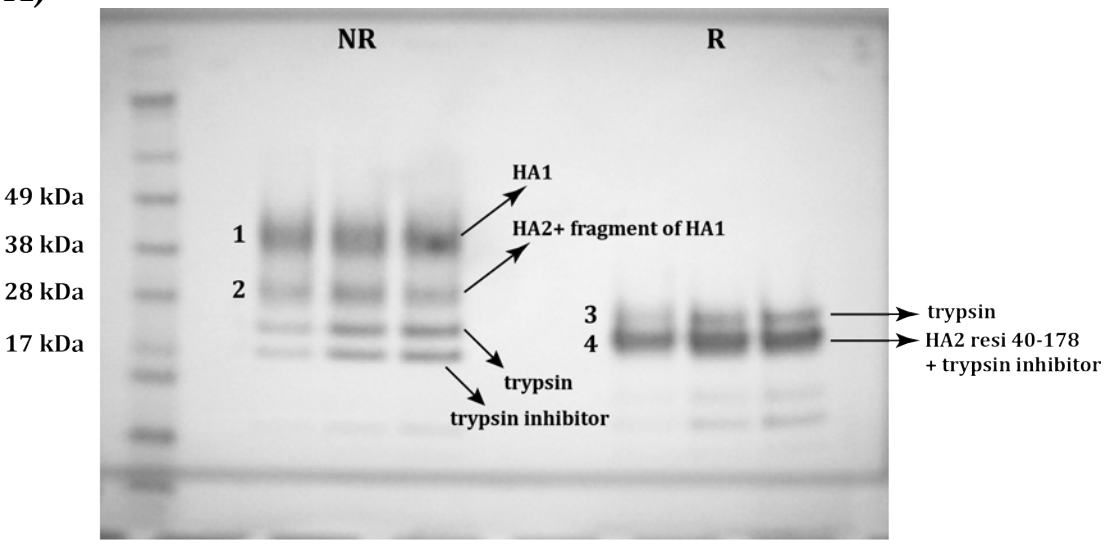


Figure 3.4 Assessment of a conformational state of low pH-treated H1 BHA by limited proteolysis. Trypsin digestion of neutral pH H1 BHA is shown in A. Trypsin digestion of low pH- treated H1 BHA shown in B. Fractions 1-18 are 0.5 ml fractions of a 10 ml 5-25% [v/v] sucrose density gradient, with fraction 1 being the bottom fraction. Details of non-reducing (NR) and reducing (R) SDS-PAGE are included in Section 2.3.3.

N-terminal sequencing:

In order to identify the trypsin digestion products of low pH-treated (pH ~5.0), viral PR8 H1 BHA, samples were transferred onto a ProBlott™ membrane and analysed by N-terminal sequencing according to a protocol described in Section 2.4.4. The N-terminal sequencing results confirmed that the obtained product was in fact a fusion peptide-less PR8 H1 TBHA2. Bands 1 and 2 gave the same sequences and showed the presence of both HA1 and HA2 disulphide-linked fragments (Figure 3.5). Band 1 is likely to be a fusion peptide-less HA2 (residue 40 onwards) with a larger fragment of HA1. Under reducing conditions (Figure 3.5, band 4) the protein band contains only the HA2 fragment, which corresponds to a removal of disulphide-linked HA1 domains (Figure 3.5). The N-terminal sequencing results presented in Figure 3.5, show, that the low pH-activated H1 BHA isolated from PR8 H1N1 virus is very susceptible to proteolytic cleavage, and digestion with trypsin results in a removal of both, the hydrophobic fusion peptide (at the extreme N-terminus of HA2) and most parts of HA1 domains, at the same time. The obtained fragment was named PR8 H1 TBHA2, and following a final purification by sucrose density gradient centrifugation and gel filtration (2.3.5), the purified protein was subjected to crystallisation trials (Sections 2.3.9 and 3.6).

A)



B)

MKANLLVLLC ALAAADADTI CIGYHANNST DTVDTVLEKN VTVTHSVNLL 050
EDSHNGKLCR LKGIAPLQLG KCNIAGWLLG NPECDPLLPV RSWSYIVETP 100
NSENGICYPG DFIDYEELRE QLSSVSSFER FEIFPKESSW PNHNTNGVTA 150
ACSHEGKSSF YRNLLWLTEK EGSYPNLKNS YVNKKGKEVL VLWGIHHPSN 200
SKEQQNLYQN ENAYVSVVTS NYNRRFTPEI AERPKVRDQA GRMNYYWTLL 250
KPGDTIIFEA NGNLIAPMYA FALSRGFGSG IITSNASMHE CNTKCQTPLG 300
AINSSLPYQN IHPVTIGECP KYVRSAKLRM VTGLRNIPSI QSRGLFGAIA 350
GFIEGGWTGM IDGWYGYHHQ NEQGSGYAAD QKSTQNAING ITNKVNTVIE 400
KMNIQFTAVG KEFNKLEKRM ENLNKKVDDG FLDIWTYNAE LLVLLENERT 450
LDFHDSNVKN LYEKVKSQLK NNAKEIGNGC FEFYHKCDNE CMESVRNGTY 500
DYPKYSEESK LNREKVDGVK LESMGIYQIL AIYSTVASSL VLLVSLGAIS 550
FWMCSNGSLQ CRICI 565

C)

Band 1	**DTICIGYHANN**	**STQNAINGITN**
Band 2	**DTICIGYHANN**	**STQNAINGITN**
Band 3	**IVGGYTXAA- TRYPSIN**	
Band 4	**DFVLDNEGN-TRYPSIN INHIBITOR**	**STQNAINGITN**

Figure 3.5 Identification of trypsin digestion products of low pH-treated H1 BHA by N-terminal sequencing. Digestion of low pH-treated H1 BHA with trypsin is shown in A. Details of SDS-PAGE are included in Section 2.3.3. B shows the primary sequence of PR8 H1 HA: the signal peptide (green), HA1 (blue), the N-terminal HA2 fusion peptide (yellow) the rest of HA2 (red). Sequences corresponding to HA2 identified by N-terminal sequencing are underlined. Composition of bands 1-4 (A) identified by N-terminal sequencing is shown in C.

3.4 Removal of HA1 domains by limited proteolysis

Following the low pH treatment (pH ~5.0), the dissociated HA1 domains of B59 H1 BHA and B59 H1 THA were removed using trypsin at a range of 0.1–5% [w/w] (protein: trypsin), for 1 h, at RT. Digestions were stopped using an equal amount of soybean trypsin inhibitor (Sigma). Removal of HA1 domains was successful using 5% [w/w] trypsin. The fusion peptide-aggregated protein (Bullough, Hughson et al. 1994) required for further protein preparation steps was purified on a continuous 5-25% [v/v] sucrose density gradient, with a 60% [w/v] sucrose cushion (1ml) at 35,000 rpm, 16 h, 20°C, or by ion-exchange chromatography (2.3.4) and subsequent gel filtration (2.3.5). Figure 3.6 shows B59 H1 THA before activation by low pH (A), and B59 H1 THA2 aggregates (rosettes) following removal of HA1 domains with 5% [w/w] trypsin (B and C).

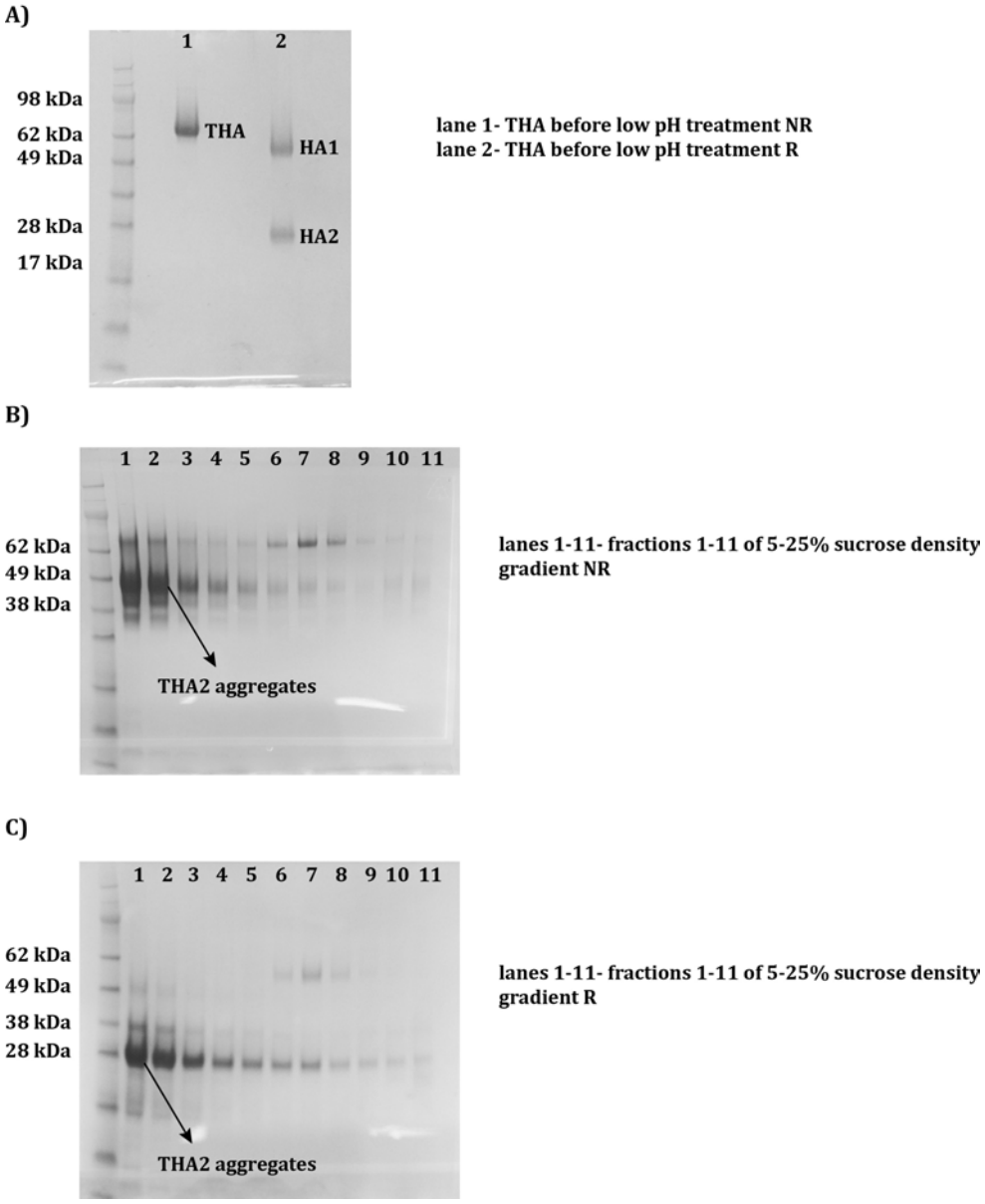
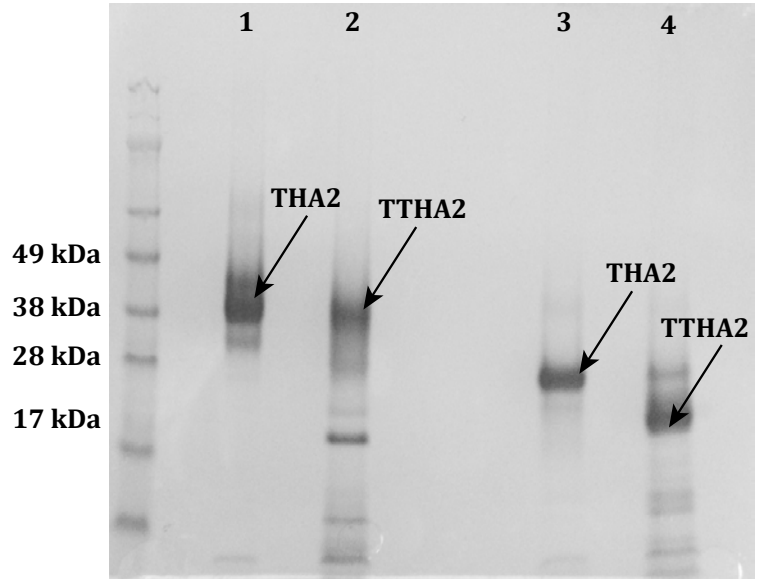


Figure 3.6 Purification of H1 THA2 aggregates. Detergent-released B59 H1 THA before the low pH treatment is shown in A. B and C show the removal of HA1 domains from the low pH-activated H1 THA using trypsin. Protein aggregates are present in the bottom fractions of 5-25% sucrose density gradient with a 60% sucrose cushion. Details of SDS-PAGE are included in Section 2.3.3.

3.5 Removal of the fusion peptide by limited proteolysis

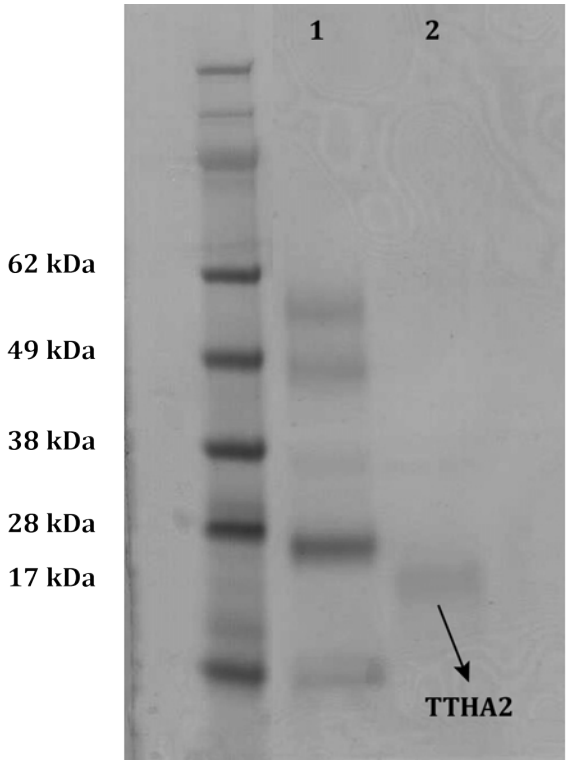
High susceptibility of PR8 H1 HA to proteolytic cleavage resulted in a removal of both, HA1 domains, and the hydrophobic fusion peptide at the N-terminus of HA2 in a single digestion step. At the N-terminus of HA2 residues 1-39, which include the fusion peptide, were removed, as determined by N-terminal sequencing (Section 3.3 and Figure 3.5). The soluble H1 TBHA2 isolated from this virus strain using the enzymatic method was subjected directly to crystallisation trials (Sections 2.3.9 and 3.6).

In case of B59 H1 HA, digestion using 0.1-1% [w/w] trypsin, for 30 min, at RT, did not remove the fusion peptide. Removal of the fusion peptide using 2-10% [w/w] trypsin, for 10-20 min, at 37°C, resulted in an incomplete digestion, with approximately half of the protein still aggregating in the bottom fractions of a 5-25% [v/v] sucrose density gradient, following centrifugation at 38,000 rpm, for 18 h, at 4°C. Digestion with 10% trypsin [w/w], for 1 h, at 37°C resulted in a removal of the hydrophobic fusion peptides. The truncated protein can be observed using SDS-PAGE (2.3.3) under reducing conditions, as presented in Figure 3.7, lane 4. Solubility of the obtained B59 H1 TTHA2 was confirmed using sucrose density gradient centrifugation, and the size of the observed band (~17 kDa) (Figure 3.7, lane 4) is the same, as the size of a fusion peptide-less product of proteolytic digestion of PR8 H1 TBHA2, the sequence of which was confirmed to start with STQNAINGITN (HA2 residue 40 onwards) by N- terminal sequencing (Section 3.3, Figure 3.5). Because the resulting band was the same size, regardless of the concentration of trypsin used to remove the fusion peptide at 37°C, all future preparations of postfusion H1 HA2 were carried out using 10% [w/w] trypsin, at 37°C, which removed both HA1 domains and the fusion peptide at the same time, and resulted in a consistent fragment of postfusion viral H1 HA2. Both B59 H1 TBHA2 and B59 H1 TTHA2 were the same size as that of PR8 H1 TBHA2, as determined by SDS-PAGE (2.3.3). Thermolysin removed the fusion peptide at a concentration of 2% [w/w], at pH 5.0, in the presence of 1mM $CaCl_2$. The pH was adjusted to 5.0 using 0.1 M Sodium citrate pH 3.5. Digestion was set up at 37°C, over a 1.5-7 h time course. Thermolysin removed the fusion peptide only after a 7 h digestion at 37°C, and the resulting TTHA2 band can be observed by SDS-PAGE under reducing conditions. The size of the obtained TTHA2 band is similar to that obtained using trypsin (~17 kDa) and is shown in Figure 3.8, lane 2.



lanes 1 & 2- non-reducing SDS-PAGE
lanes 3 & 4- reducing SDS-PAGE

Figure 3.7 Removal of the fusion peptide from H1 rosettes with trypsin. Fusion peptide-less B59 H1 TTHA2 is shown in lanes 2 and 4, under non-reducing and reducing conditions. Details of SDS-PAGE are included in Section 2.3.3.



lane 1- THA2 R
lane 2- TTHA2 R

Figure 3.8 Removal of the fusion peptide from H1 rosettes with thermolysin. Fusion peptide-less B59 H1 TTHA2 is shown in lane 2. Details of SDS-PAGE are included in Section 2.3.3.

Testing of other proteases for the ability to remove the fusion peptide:

In an attempt to obtain a longer fusion peptide-less fragment of B59 H1 TTHA2, proteolytic screening was carried out as described in Section 2.3.7. Limited proteolysis of aggregated H1 THA2 was performed versus a number of proteases. The proteases were selected based on their ability to cleave hydrophobic amino acids, and included: α-Chymotrypsin, Papain, Elastase, Subtilisin, Endoproteinase Glu-C, Proteinase K, Clostripain (Endoproteinase-Arg-C), Pepsin and Actinase. Proteases (Proti-Ace Kit and Proti-Ace 2 Kit, Hampton Research) and used according to manufacturer's instructions at 0.01 mg/ml. The analysis of protein digests using SDS-PAGE (2.3.3) presented in Figure 3.9 shows, that while α-Chymotrypsin, Subtilisin and Proteinase K completely digested B59 H1 THA2, enzymes such as Endoproteinase Glu-C and Actinase partially degraded the protein at 0.01 mg/ml. Digestions with Papain, Elastase, Clostripain and Pepsin (Figure 3.10, lanes 3, 4, 8 and 9) were repeated using a higher concentration of proteases (0.1 mg/ml), over a period of 60 min, at 37°C. Figure 3.10 shows, that none of the enzymes removed the fusion peptide at 0.1 mg/ml. Endoproteinase Lys-C was tested for its ability to remove the fusion peptide at 0.1-5% [w/w], for 30 min, at RT. Reactions resulted in a complete digestion of H1 THA2.

Summary of protein digestion experiments:

Viral H1 HA2 in a postfusion conformation has been challenging to prepare proteolytically. Only trypsin and thermolysin were successful in removing the hydrophobic fusion peptide at the N-terminus of HA2 (HA2 residues 1-23). Thermolysin removed the fusion peptide at 2% [w/w], and digestion was complete after 7 h, at 37°C. Removal of the fusion peptide and HA1 domains with trypsin required at least 5% [w/w] of the enzyme, and 1 h, at 37°C. Both activities resulted in a similar fragment of ~17 kDa. This proteolytically prepared fragment of influenza virus HA (H1 subtype) in a postfusion conformation was subjected to crystallisation trials as described in Section 3.6.

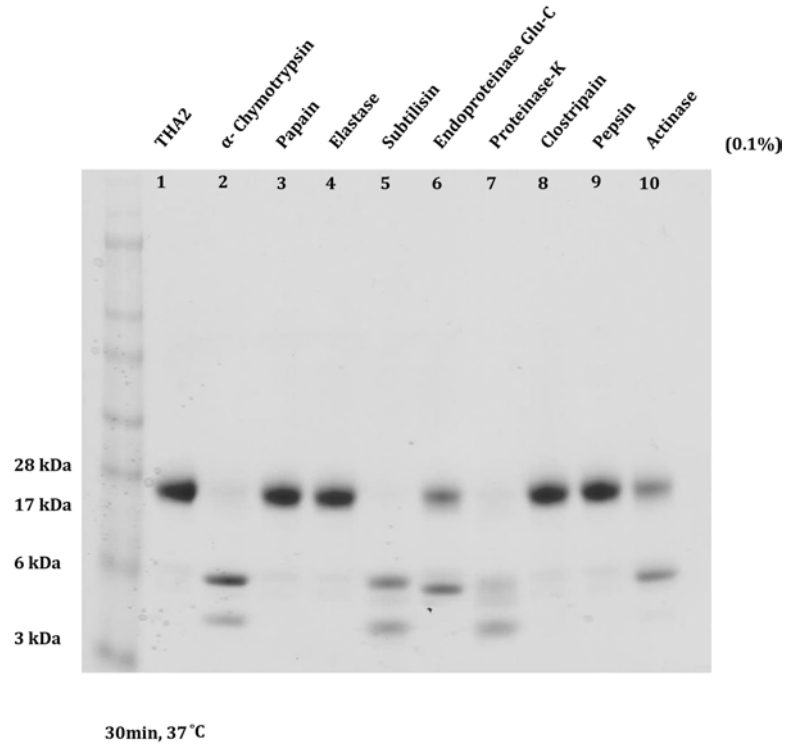


Figure 3.9 Removal of the hydrophobic fusion peptide using Proti-Ace proteases. Proteases (Hampton Research) were used at 0.1% [w/w], for 30 min, at 37°C . Completely degraded B59 H1 THA2 is shown in lanes 2, 5 and 7. Partially degraded protein is shown in lanes 6 and 10. No degradation is observed in lanes 1, 3, 4, 8 and 9. Reducing SDS-PAGE (2.3.3).

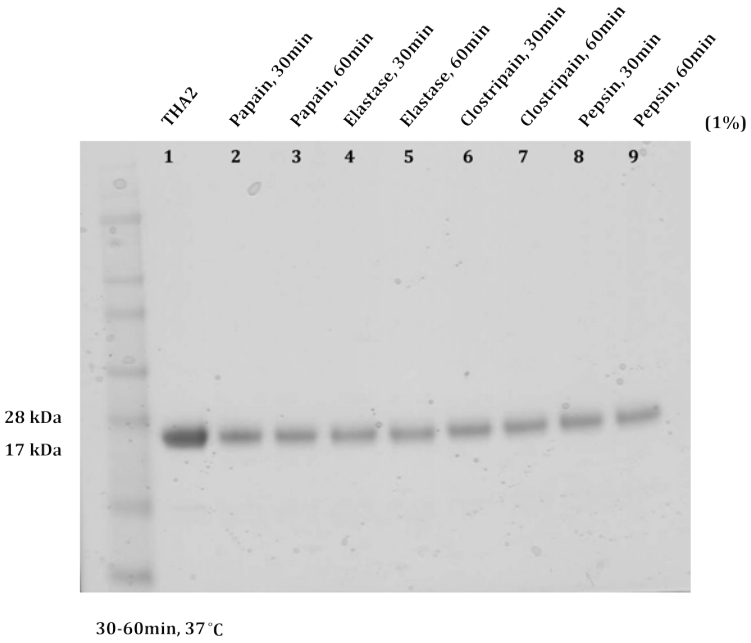


Figure 3.10 Removal of the hydrophobic fusion peptide using other enzymes. Papain, Elastase, Clostripain and Pepsin at 1% [w/w] (Hampton Research) were used for 30-60 min, at 37°C. Reducing SDS-PAGE (2.3.3).

3.6 Crystallisation

Crystals of PR8 H1 TBHA2:

Suitability of the prepared H1 TBHA2 and TTHA2 samples for crystallisation trials was assessed using DLS as described in Section 2.4.2. The PR8 H1 TBHA2 sample was concentrated to ~5 mg/ml (A_{280}), using a Vivaspin 10K MWCO concentrator (Sartorius) at 3,000 g, 4°C, in a benchtop centrifuge (Thermo Scientific). Crystallisation experiments were set up using a vapour diffusion technique as described in Section 2.3.9. Protein crystallized from various PEGs at pH 6.0-7.5, with seeding. Microseed stock was prepared from an initial crystalline material grown from 16% PEG 3350 and 0.1M MES pH 6.5. Conditions successful in producing crystals of PR8 H1 TBHA2 contained 0.1M buffers, such as: Bis-tris pH 5.5, MES pH 6.0-6.5, Hepes pH 7.5, Pipes pH 6.8-7.0, Sodium citrate pH 6.5, Sodium malonate pH 6.0, and 8% tacsimate pH 6.0. Initial hits were obtained from: 12-25% PEG 3350 (pH 6.0-7.5), 20% PEG 4000 (pH 6.5), 16% PEG 8000 (pH 5.5), 20-30% PEG 600 (pH 6.0-7.0), 25% PEG 550MME (pH 6.5). Although crystals were of suitable size and good morphology (Figure 3.11), they diffracted weakly, with best crystal diffracting to about ~5 Å. In order to shrink the crystal lattice and improve crystal packing, dehydration experiments were performed (by serial transfer into drops with an increasing concentration of precipitant), prior to crystal freezing. Despite extensive optimisation trials and testing of various cryoprotectants (Ethylene glycol, Glycerol and PEG 400), no improvement in diffraction resolution was observed. Details related to protein crystallisation including preparation of the seed solution and crystal freezing, are included in Section 2.3.9.

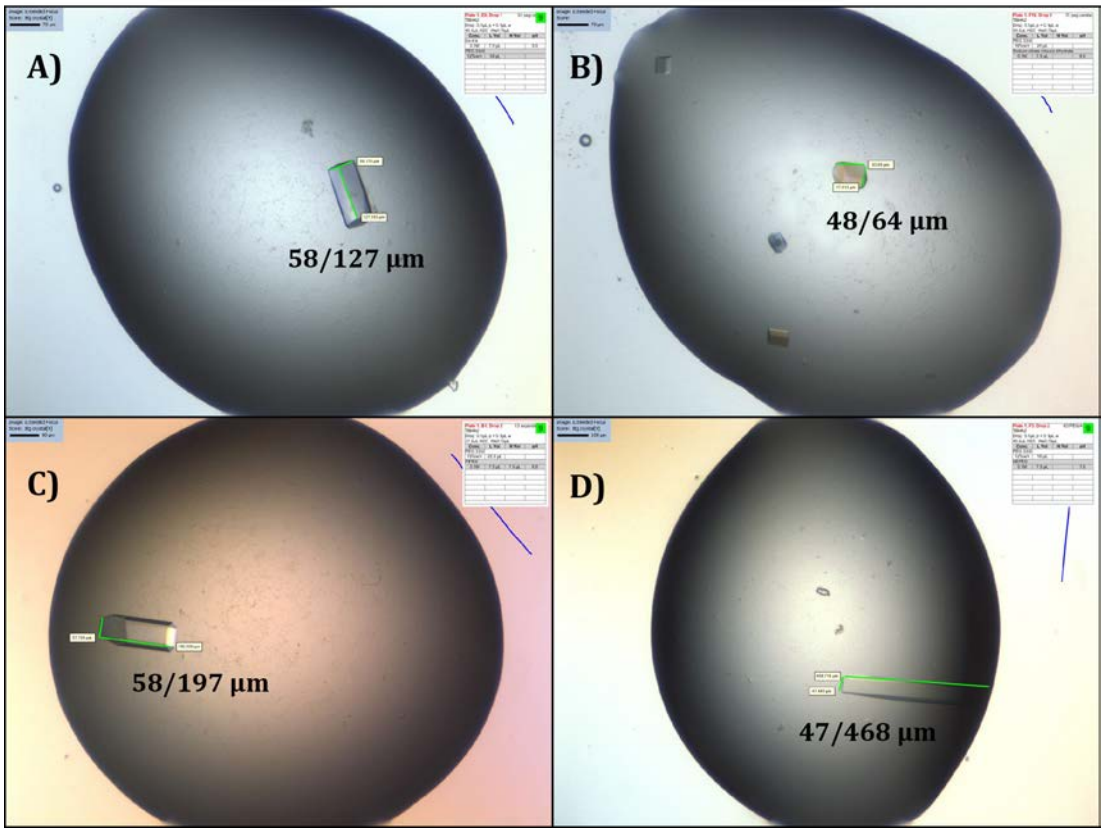


Figure 3.11 Crystals of PR8 H1 TBHA2. Crystals were grown from 12 % PEG 3350, 0.1M Bis-tris pH 5.5 (A), 16% PEG 3350, 0.1M citrate pH 6.5 (B), 15% PEG 3350, 0.1M Pipes pH 6.8 (C) and 12% PEG 3350, 0.1M Hepes pH 7.5 (D). Crystal dimensions are shown in µm. Images were obtained using Rock Imager 1000 with Rock Maker software (Formulatrix).

Crystals of B59 H1 TBHA2:

B59 H1 TBHA2 and B59 H1 TTHA2 were concentrated to ~7.5 mg/ml (A_{280}), and ~5.5 mg/ml (A_{280}), respectively, using a Vivaspin 10K MWCO concentrator (Sartorius) at 3,000 g, 4°C, in a benchtop centrifuge (Thermo Scientific). The concentrated protein was subjected to crystallisation trials (2.3.9). Protein crystallized from conditions containing various PEGs. Buffers successful in producing crystals of B59 H1 TBHA2 and B59 H1 TTHA2 included 0.1M solutions of Pipes pH 6.8, Hepes pH 7.5, MES pH 6.0, Potassium phosphate monobasic and Bicine pH 9.0. The initial crystals obtained from 25% PEG 3350 and 0.1M Hepes pH 7.5 were used to prepare a microseed solution. Although microseeding resulted in crystals with improved size and morphology, crystals of B59 H1 TBHA2 and TTHA2 were generally smaller (Figure 3.12), and took longer to grow than those of PR8 H1 TBHA2 (Figure 3.11). Initial hits were obtained from: 15-25% PEG 3350 (pH 5.5-7.5), 10% PEG 6000 (pH 9.0), 20% PEG 600 (pH 6.0). Diffraction quality crystals of B59 H1 TBHA2 were obtained from 16% PEG 3350, 0.1M MES pH 6.0 with seeding (Figure 3.12, C), and from 15% PEG 3350, 0.1M MES pH 6.0+1.5% [v/v] MPD, 0.01M Hepes pH 7.5 and 0.02M Sodium citrate (Figure 3.12, D).

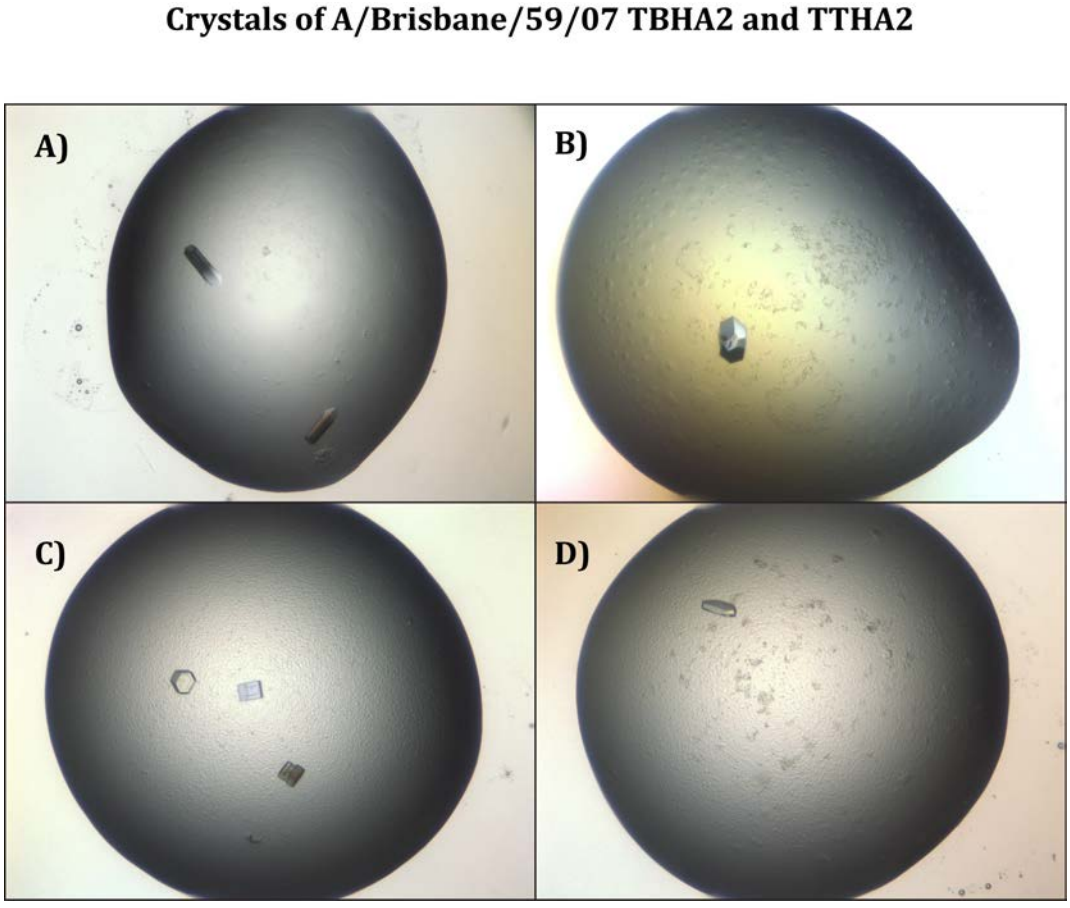


Figure 3.12 Crystals of B59 H1 TTHA2 and B59 H1 TBHA2. Crystals of TTHA2 were grown from 20% PEG 3350, 0.1M Pipes pH 6.8 (A), and from 20% PEG 3350, 0.1M MES pH 6.0 (B). Crystals of TBHA2 were obtained from 16% PEG 3350, 0.1M MES pH 6.0 (C), and from 15% PEG 3350, 0.1M MES pH 6.0+1.5% v/v MPD, 0.01M Hepes pH 7.5, 0.02M Sodium citrate (D). Images obtained using Rock Imager 1000 with Rock Maker Software (Formulatrix).

3.7 Data collection and processing

Thirteen crystals of B59 H1 TBHA2 in a postfusion conformation were tested. Two crystals gave good diffraction and datasets were collected. Dataset A was from a crystal grown from 15% PEG 3350 and 0.1M MES pH 6.0, with the addition of 1.5% MPD, 0.01M Hepes pH 7.5, 0.02M Sodium citrate. The cryoprotectant solution was identical to the crystallisation condition with a slightly higher concentration of PEG (~2% v/v), and 20% [v/v] glycerol. Dataset B was from a crystal grown from a condition containing 16% PEG 3350, 0.1M MES pH 6.0, with seeding. In this case the crystal was cryoprotected with a solution containing 20% PEG 3350, 0.1M MES pH 6.0, and 20% [v/v] ethylene glycol as a cryoprotectant. Details related to protein crystallisation and crystal freezing are included in Sections 2.3.9 and 3.6.

Diffraction data were collected using an oscillation method (angle increment 0.5°/frame, exposure period 0.5 sec/frame, λ=0.97 Å), and recorded using a Pilatus detector by Dr Phil Walker (NIMR), at the IO4 beam line at Diamond Light Source Synchrotron (Harwell, UK). The two datasets were indexed and integrated using Xia2 system (Kabsch 2010) (Winter, Lobley et al. 2013). The indicated space group for both data sets was P 3 2 1 (150). Dataset A diffracted to 3.4 Å, and dataset B to 3.3 Å. However, analysis of dataset B indicated an issue with anisotropy. The Aimless (Evans 2006) (Evans and Murshudov 2013) log file output prepared using Loggraph (CCP4) (Figure 3.13) shows the anisotropy of dataset B. Analysis of signal/noise (I/σ) against resolution for dataset B is shown in Figure 3.14. In contrast, the equivalent analysis of dataset A indicated that there was no issue with anisotropy. Anisotropy analysis of dataset A is shown in Figure 3.15. Analysis of signal/noise (I/σ) against resolution is shown in Figure 3.16. Therefore, dataset A was used to solve the structure of B59 H1 TBHA2.

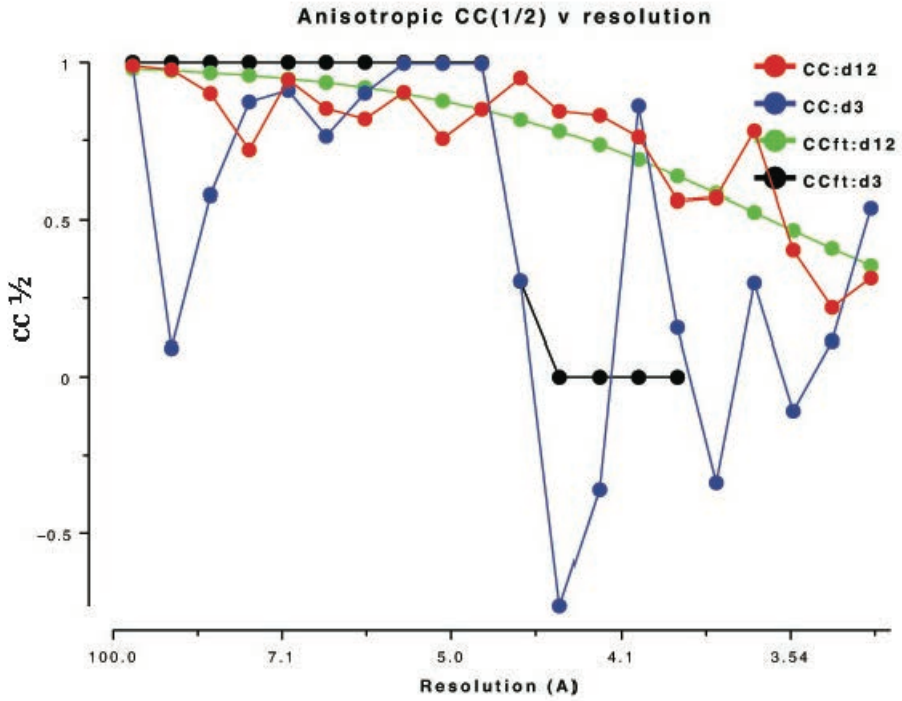


Figure 3.13 Anisotropy analysis of dataset B obtained from a crystal grown from 20% PEG 3350 and 0.1M MES pH 6.0. Graph prepared from Aimless log file (Evans and Murshudov 2013) using Loggraph (CCP4) (Winn 2003). Shown is the half-dataset correlation coefficient ($CC_{1/2}$) for the directions d12 and d3 versus resolution.

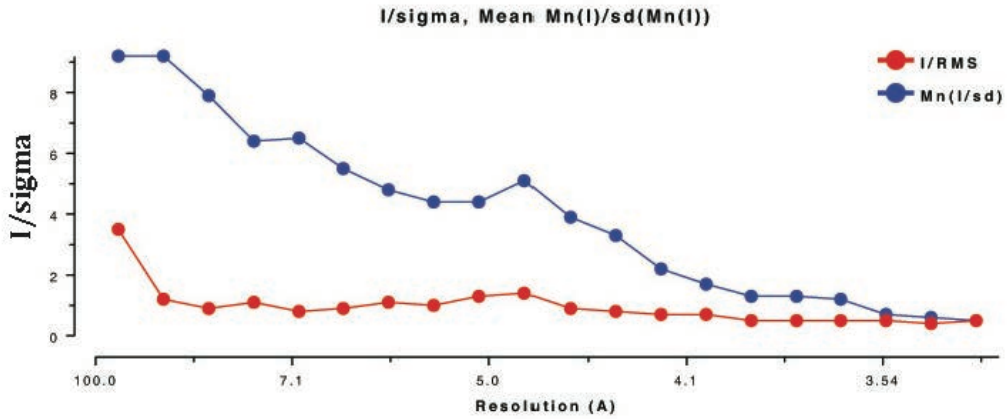


Figure 3.14 Analysis of I/sigma and Mean (I/sd) versus resolution for dataset B obtained from a crystal grown from 20% PEG 3350 and 0.1M MES pH 6.0. Graph prepared from Aimless log file (Evans and Murshudov 2013) using Loggraph (CCP4) (Winn 2003).

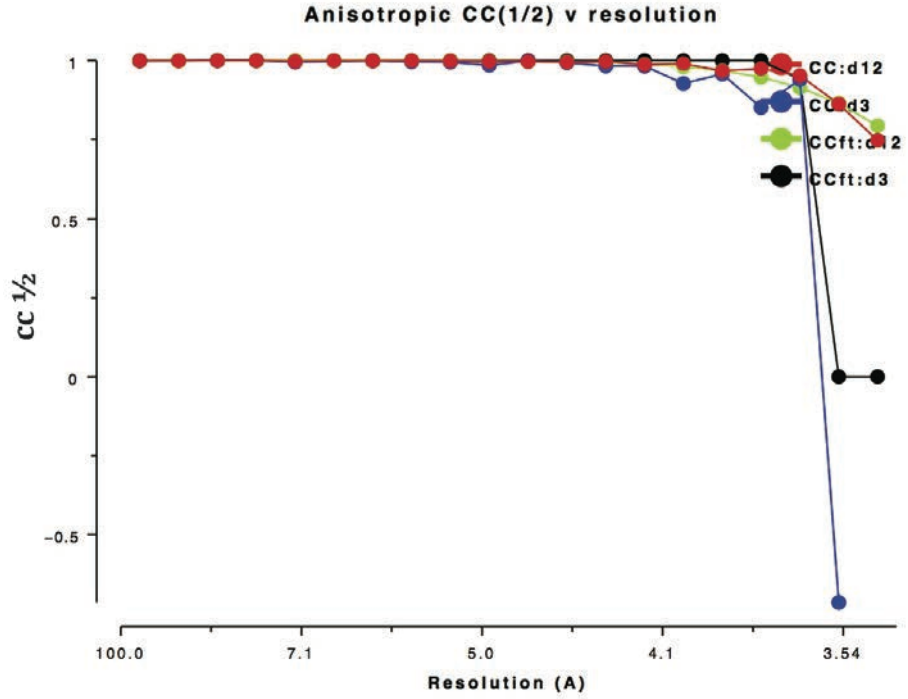


Figure 3.15 Anisotropy analysis of dataset A obtained from a crystal grown from 15% PEG 3350, 0.1M MES pH 6.0, 1.5% MPD, 0.01M Hepes pH 7.5 and 0.02M Sodium citrate. Graph prepared from Aimless log file (Evans and Murshudov 2013) using Loggraph (CCP4) (Winn 2003). Shown is the half-dataset correlation coefficient ($CC_{1/2}$) for the directions d12 and d3 versus resolution.

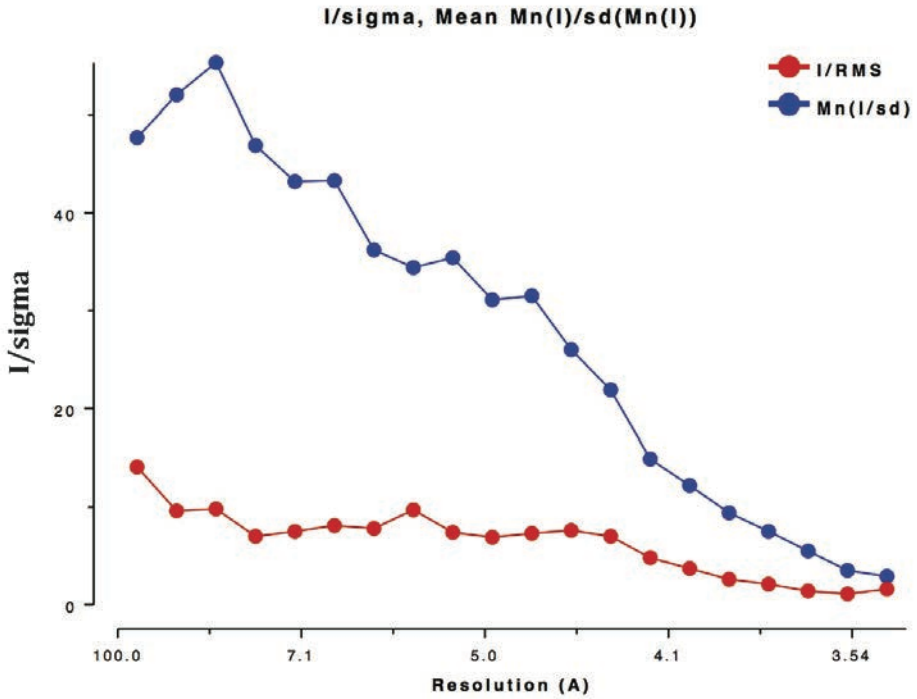


Figure 3.16 Analysis of I/sigma and Mean (I/sd) versus resolution for dataset A obtained from a crystal grown from 15% PEG 3350, 0.1M MES pH 6.0, 1.5% MPD, 0.01M Hepes pH 7.5 and 0.02M Sodium citrate. Graph prepared from Aimless log file (Evans and Murshudov 2013) using Loggraph (CCP4) (Winn 2003).

Molecular replacement and model refinement:

The unmerged XDS ASCII reflection file was used for space group assignment using Pointless (CCP4) (Evans 2006) (Evans 2011) (Winn, Ballard et al. 2011). Pointless confirmed that the most likely space group was P 3 2 1 (150). Space group P 3 2 1 was assigned with a total space group probability estimate of 0.999, and systematic absences-based estimate of 1.00. Scores for the possible Laue groups are shown in Table 3.1, which includes scores for a likelihood measure (Lklhd), a "significance" score (Z-score), a correlation coefficient (CC) and the multiplicity-weighted R factor (Rmeas).

Table 3.1 Scores for possible Laue groups for the B59 H1 TBHA2 crystal. Scores obtained using Pointless (CCP4).

	Laue Group	Lklhd	NetZc	CC	Rmeas
1	P -3 m 1	0.999	8.43	0.97	0.06
2	C 1 2/m 1	0.00	5.40	0.98	0.04
3	C 1 2/m 1	0.00	5.19	0.97	0.05
4	C 1 2/m 1	0.00	5.20	0.97	0.05
5	P -3	0.00	5.24	0.97	0.06
6	P -1	0.00	4.78	0.99	0.03

The molecular weight of each B59 H1 TBHA2 monomer was estimated to ~10,000 Da by N-terminal sequencing (Section 3.3). The number of molecules in the asymmetric unit of the P 3 2 1 unit cell was calculated using Matthews-Coeff (CCP4) (Kantardjieff and Rupp 2003). Based on the calculated % solvent content, there is most likely one monomer of postfusion B59 H1 TBHA2 in the asymmetric unit of the P 3 2 1 crystal (Table 3.2). Data collection and refinement statistics are shown in Table 3.3.

Table 3.2 Number of B59 H1 TBHA2 monomers in the asymmetric unit of the P321 (150) unit cell and the corresponding % solvent content. The Matthews coefficient (V_m) is calculated as $\frac{volume\ of\ unit\ cell}{M_w \times Z \times X}$, where Z is the number of asymmetric units in the unit cell, and X is the number of molecules in the asymmetric unit.

Cell volume: 196972.97 ~10kDa/monomer P321 (150)			
Nmol/asym	Matthews Coeff	% solvent	P (tot)
1	3.28	62.56	0.98
2	1.64	25.11	0.02

Molecular replacement using a monomer of recombinantly-expressed H3 HA2 (residues 33-173) (PDB: 1QU1) (Chen, Skehel et al. 1999) or viral postfusion H3 HA2 (residues 40-162) (PDB: 1HTM) (Bullough, Hughson et al. 1994) as a search model using MOLREP (Vagin and Teplyakov 2010) and PHASER (McCoy, Grosse-Kunstleve et al. 2007), did not find a solution initially. Given the length of the shorter molecular replacement model (PDB: 1HTM), which is ~98 Å long (Ser40-Gln105), and the dimensions of the unit cell: 59.60, 59.60, 64.03 (a, b, c), it was concluded, that even the shorter molecule (1HTM) would not fit into the P 3 2 1-unit cell. For this reason, the original model (1HTM) was truncated, according to trypsin cleavage sites predicted by PeptideCutter (ExPASy). The subsequent deletions from the N-terminus of H3 HA2 (1HTM) model are shown in Figure 3.17. A corresponding number of residues running antiparallel to the main α-helix were removed from the C-terminal region of the search model. The final trimmed model of H3 TBHA2 contained one H3 HA2 monomer (residues 69-156). The trimmed postfusion helix of H3 TBHA2 was pruned to carbon beta using CHAINSAW (CCP4 suite) (Winn, Ballard et al. 2011). In this procedure, a PDB file was mutated according to an input sequence alignment between target and model sequences, by pruning non-conserved residues to C_β atoms. Pruned H3 TBHA2 monomer (PDB: 1HTM) was used as a new search model for molecular replacement using MOLREP (Lebedev, Vagin et al. 2008) (Vagin and Teplyakov 2010) and PHASER (McCoy, Grosse-Kunstleve et al. 2007) (CCP4 suite).

Molecular replacement was successful using a model of viral H3 TBHA2 (PDB: 1HTM), trimmed to contain HA2 residues 69-159 (Figure 3.17) and pruned to beta atom using CHAINSAW (Winn, Ballard et al. 2011). Both MR programs predicted the same packing of molecules in the crystal lattice. Molecular replacement was also carried out using a monomer of postfusion B59 H1 TBHA2, solved previously in a lower symmetry space group. This strategy resulted in lower R factors than those obtained using H3 TBHA2 monomer (1HTM) (Bullough, Hughson et al. 1994), thus a monomer of H1 TBHA2 was used as a search model for molecular replacement.

The model of postfusion B59 H1 TBHA2 was built into the σ- weighted $2F_O$-F_C electron density map using Coot (Emsley and Cowtan 2004), and refined using Buster (Bricogne G. 2011). Figures for structural snapshots were generated using PyMol (Molecular Graphics System, Version 1.2r3pre, Schrödinger, LLC.). Data collection and refinement statistics are presented in Table 3.3. Ramachandran

analysis of the final refined model is shown in Figure 3.18. Only two residues, Arg 106 and Asp 146, lie outside the preferred regions.

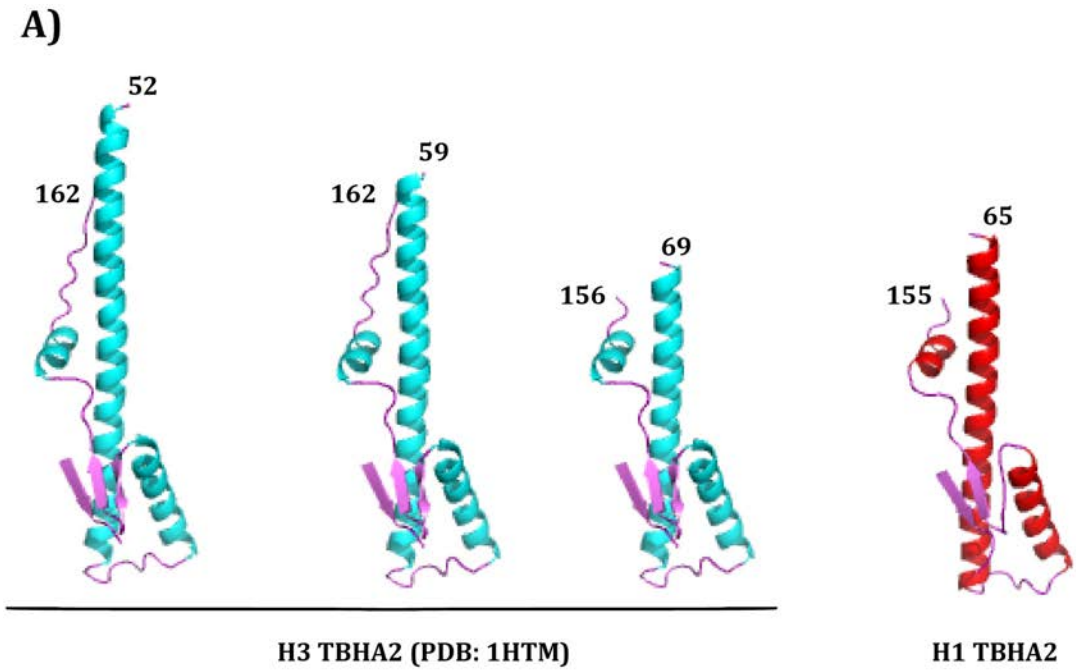
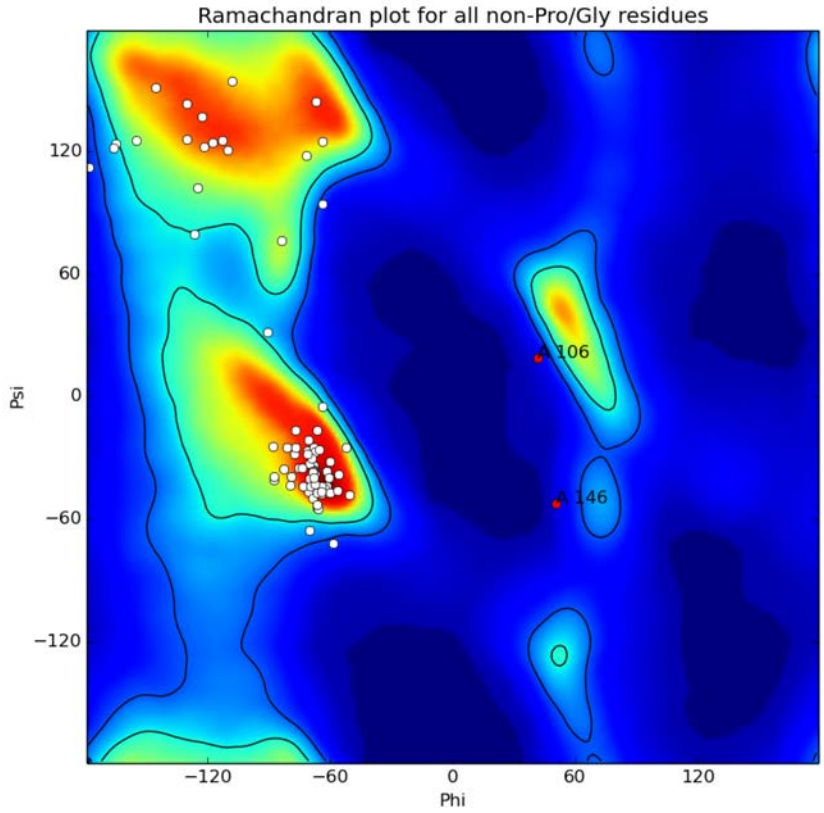


Figure 3.17 Truncations of a search model for molecular replacement. Trimmed H3 TBHA2 monomers (PDB: 1HTM) (Bullough, Hughson et al. 1994) are shown in blue, and a model of B59 H1 TBHA2 solved in P 3 2 1 space group is shown in red (A). The primary sequence of H3 TBHA2 (residues 40-162) aligned to the corresponding fragment of H1 HA2 is shown in B. Molecular replacement model truncations are indicated. The non-conserved residues were additionally pruned to carbon beta using Chainsaw (CCP4) (Winn, Ballard et al. 2011). Sequences were aligned using PRALINE (IBIVU Server) (Simossis and Heringa 2003).

Table 3.3 Data collection and refinement statistics for B59 H1 TBHA2 (molecular replacement).

B59 H1 TBHA2	
Data collection	
Space group	P 3 2 1
Cell dimensions	
a, b, c (Å)	59.60, 59.60, 64.03
α, β, γ (◦)	90.00, 90.00, 120.00
Resolution (Å)	40.18-3.41 (3.50-3.41)*
R_{merge}	0.069 (0.465)*
Mean I/σ(I)	23.3 (2.9)*
Completeness (%)	97.7 (84.1)*
Redundancy	13.2 (6.1)*
Refinement	
Resolution (Å)	32.01-3.41
No.reflections	1912
R_{work}/R_{free}	0.31/0.33
No. protein atoms	613
Protein B-factors	120.5
Rms deviations	
Bond lengths (Å)	0.009
Bond angles (Å)	1.09

*Values in parentheses are for the highest-resolution shell.



Chain	Residue	Score	Phi	Psi
A	ARG 106	0.01	41.7	19.2
A	ASP 146	0.01	50.5	-52.3

Figure 3.18 Ramachandran plot for B59 H1 TBHA2 model. The plot was prepared using Phenix GUI, Comprehensive validation (Afonine, Grosse-Kunstleve et al. 2012).

3.8 Crystal structure of B59 H1 TBHA2

The refolded B59 H1 TBHA2 was prepared by releasing HA from viral particles using bromelain (2.2.1.3) and incubation at pH ~5.0 to induce the conformational change of HA (3.3), after which the HA1 domains and the fusion peptide were removed by digestion with trypsin. Crystals grew from 15% PEG 3350, 0.1M MES pH 6.0, with the addition of 1.5% [v/v] MPD, 0.01M Hepes pH 7.5 and 0.01M Sodium citrate, and the structure was solved by molecular replacement using a monomer of H1 TBHA2 solved previously in a lower symmetry space group, as a search model. There is one monomer of postfusion B59 H1 TBHA2 in the asymmetric unit of the P 3 2 1-unit cell. The postfusion structure of B59 H1 TBHA2 is 62.29 Å long, and the central stem comprises HA2 residues Ala65-Glu105. The HA2 chain was labelled A and is a continuous polypeptide chain containing HA2 residues 65-155. The connecting loop (HA2 residues 106-112) turns the long central helix (Ala65-Glu105) into an antiparallel shorter α-helix (HA2 residues 113-130), which then extends to form a loop. Within this loop residues 131-140 of HA2 form a ß-hairpin, and together with a short strand of HA1 containing Asp11-Tyr17 (H3 numbering), this region of the molecule folds into an antiparallel ß-sheet. The HA1 chain was named D. Most parts of the extended C-terminal fragments of the postfusion molecule are unobserved in the crystal structure of B59 H1 TBHA2. Figure 3.19 shows a trimeric (A) and monomeric (B) cartoon representation of the obtained structure. Loop containing the ß-hairpin (Lys131- Phe140) with a disulphide-linked fragment of HA1 (Asp11-Tyr17) is magnified in Figure 3.19 C. The obtained structure of postfusion viral B59 H1 TBHA2 is compared to other known structures of postfusion HA2 in Chapter 5.

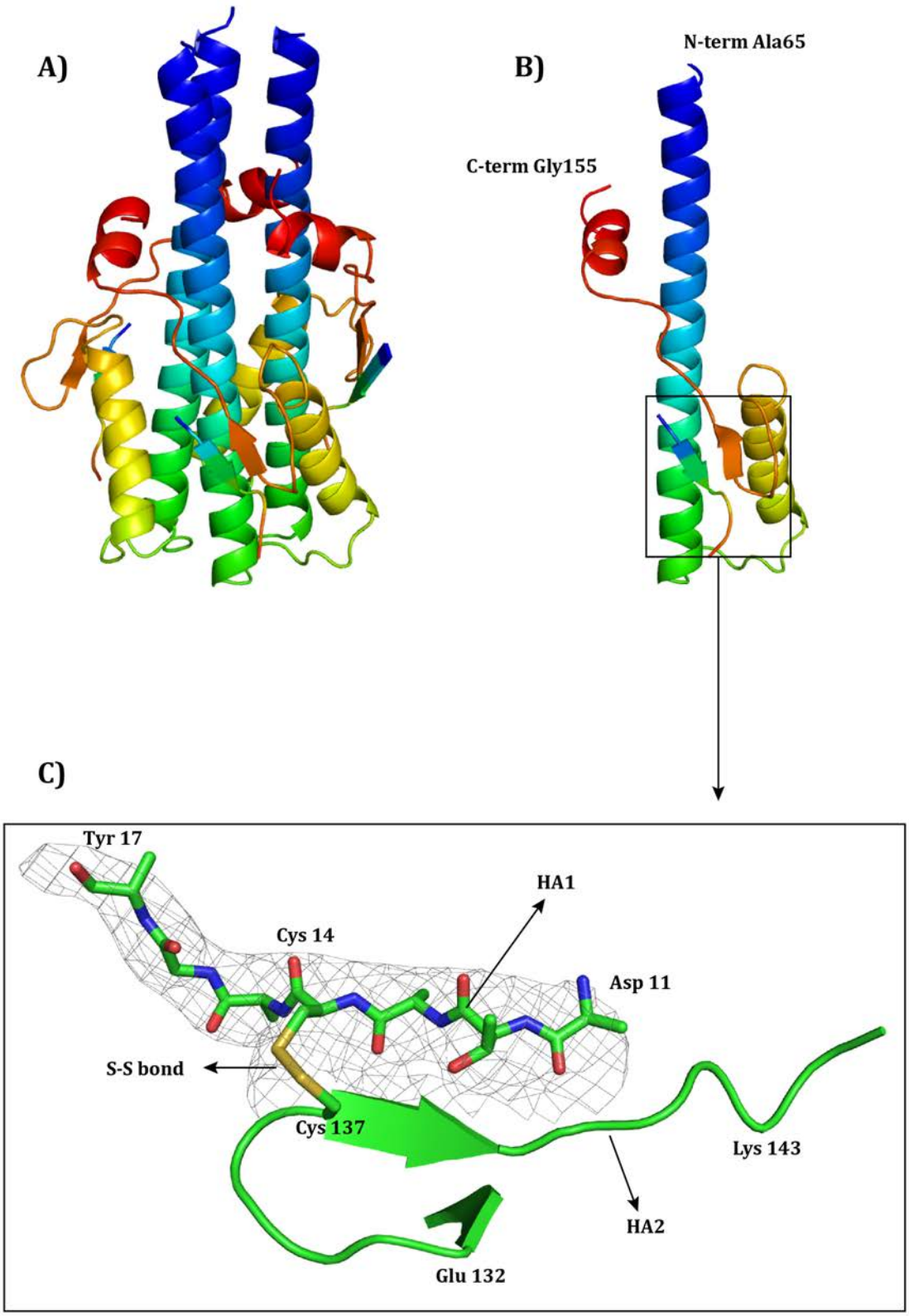


Figure 3.19 Crystal structure of B59 H1 TBHA2. A trimer of postfusion B59 H1 TBHA2 is shown in A. A B59 H1 TBHA2 monomer is shown in B. Loops at HA2 residues 132-143 are shown in C. HA1 chains are shown as sticks and S-S bond is indicated (H3 numbering).

4 Crystal structure of recombinant H1 HA2

Summary:

Two fragments of postfusion H3 HA2 (group 2 HA), containing HA2 residues 38-175 and 23-185, referred to as EBHA2 and EHA2, respectively, and a fragment of postfusion influenza B HA2, have been previously expressed in *E.coli* (Chen, Wharton et al. 1995) (Ni, Chen et al. 2014).

In an attempt to prepare a longer than enzymatically prepared (Chapter 3) fragment of postfusion H1 HA2, HA2 residues 38-175 of A/Brisbane/59/07 H1 HA were expressed using Baculovirus/insect cell system (Section 2.2.2). Conformational assessment of the recombinantly-expressed protein was carried out using limited proteolysis (2.3.7). The secondary structure content and thermal stability of the expressed protein were estimated using circular dichroism (2.4.3). The recombinantly-expressed postfusion H1 HA2 (residues 38-175) spontaneously folded into the postfusion conformation, and after a final purification by gel filtration (2.3.5), the protein was subjected to crystallisation trials (2.3.9). Diffraction quality crystals were obtained from 18% PEG 3350, 0.1M Sodium citrate pH 5.6 and 2% tacsimate. Crystals were cryoprotected with the crystallisation solution containing 25% [v/v] Ethylene glycol.

4.1 Protein expression and purification

Recombinant H1 HA2 (residues 38-175) was expressed using Baculovirus/insect cell system as described in Section 2.2.2. The expressed protein was purified using IMAC and size-exclusion chromatography, as described in Sections 2.3.8 and 2.3.5, respectively. Protein containing fractions A12-B5 (10 ml) eluted from a cobalt column are shown in Figure 4.1, A. Cleavage and removal of the polyhistidine tag using TEV protease and Ni beads are shown B and C, respectively. The expressed protein folded into a soluble trimer of ~48 kDa (calculated M_w) as determined by a final size-exclusion chromatography (2.3.5) and SDS-PAGE (2.3.3) (Figure 4.2).

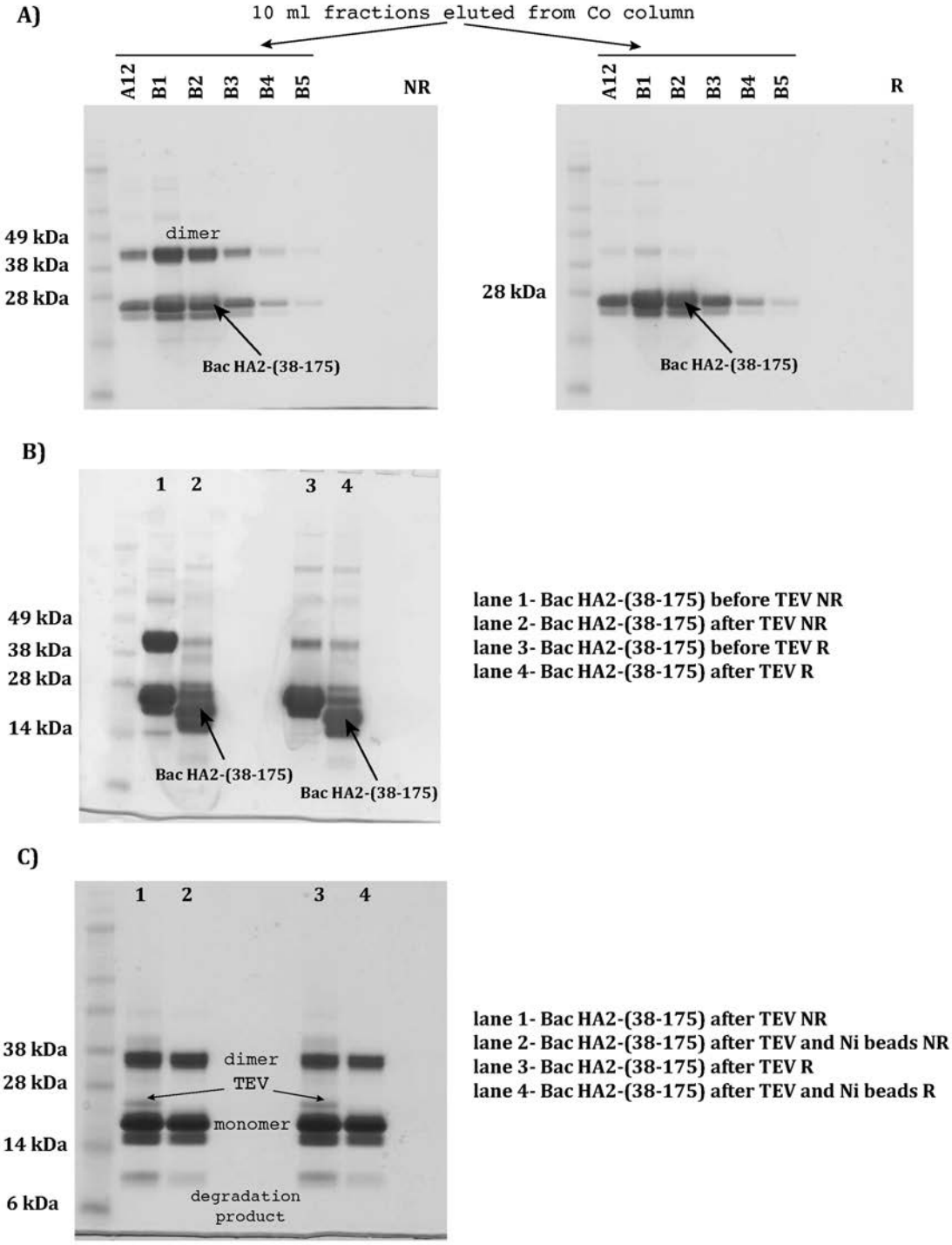


Figure 4.1 Purification of Bac 59 H1 HA2. Fractions A12-B5 (10 ml) eluted from a cobalt column and analysed using SDS-PAGE, under non-reducing (NR) and reducing (R) conditions, are shown in A. Removal of the polyhistidine tag using TEV protease is shown in B, under non-reducing (lanes 1 and 2), and reducing (lanes 3 and 4) conditions. Removal of TEV using nickel beads (Qiagen) is shown in C. Details of SDS-PAGE are included in Section 2.3.3.

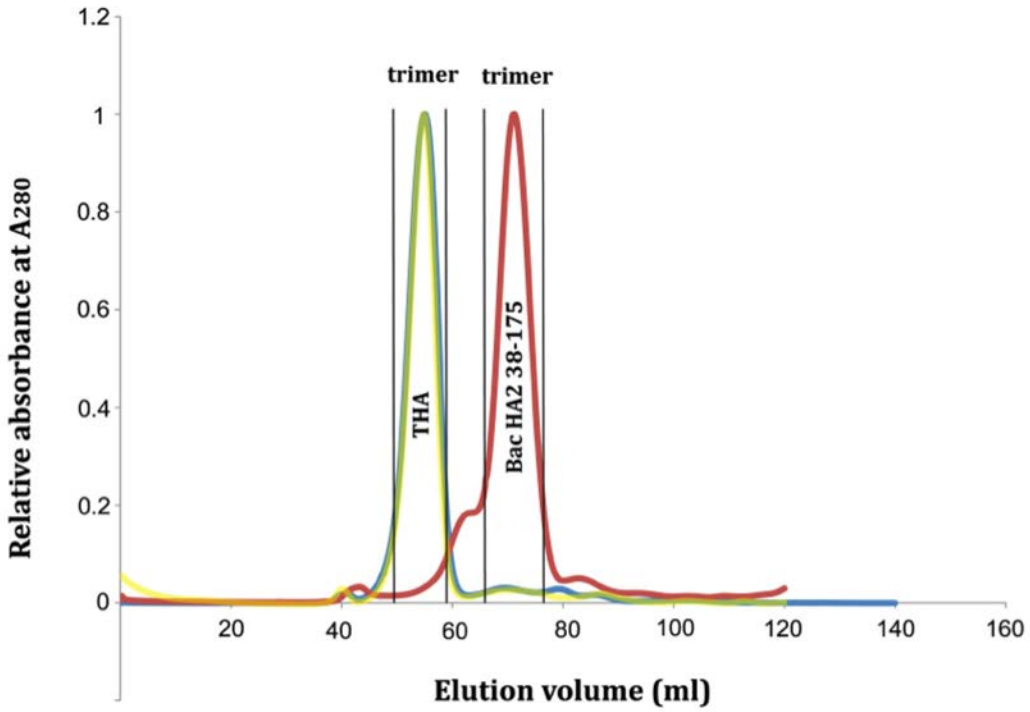


Figure 4.2 Chromatogram from final size-exclusion chromatography of Bac B59 H1 HA2. Recombinantly-expressed H1 HA2 is shown in red, and H1 THA is shown in yellow. The y-axis plots the relative absorbance at 280 nm. The x-axis plots the retention times in ml. Vertical lines indicate fractions pooled for crystallisation trials and electron microscopy. The reference curve for a THA trimer is shown in blue (H5 subtype).

4.2 Conformational assessment of recombinantly-expressed H1 HA2

Limited proteolysis:

Conformation of the recombinantly-expressed Bac B59 H1 HA2 (residues 38-175) was assessed using limited proteolysis, as described in Section 2.3.7. The recombinantly-expressed protein at 0.15 mg/ml, was digested with trypsin at 2% [w/w] and 5% [w/w], for 30min, at RT. Digestions were stopped using an equal amount of soybean trypsin inhibitor (Sigma), and tryptic digestion products were analysed using SDS-PAGE, as described in Section 2.3.3. As presented in Figure 4.3, digestion of the expressed Bac 59 H1 HA2 with trypsin, resulted in a ~6 kDa fragment as observed using SDS-PAGE (2.3.3). The observed trypsin digestion product is smaller (~6 kDa) than those previously obtained by proteolytic digestion of H3 TBHA2 and H3 EBHA2, and the size of the previously obtained fragments was ~10 kDa (Wharton, Calder et al. 1995) (Chen, Wharton et al. 1995). Fragments of similar size (~6 kDa) were obtained by proteolytic digestion of viral B59 H1 HA2 using α-Chymotrypsin, Subtilisin, Endoproteinase Glu-C and Proteinase K at 0.1% [w/w], as observed using SDS-PAGE (2.3.3), and the obtained fragments are shown in Figure 3.9.

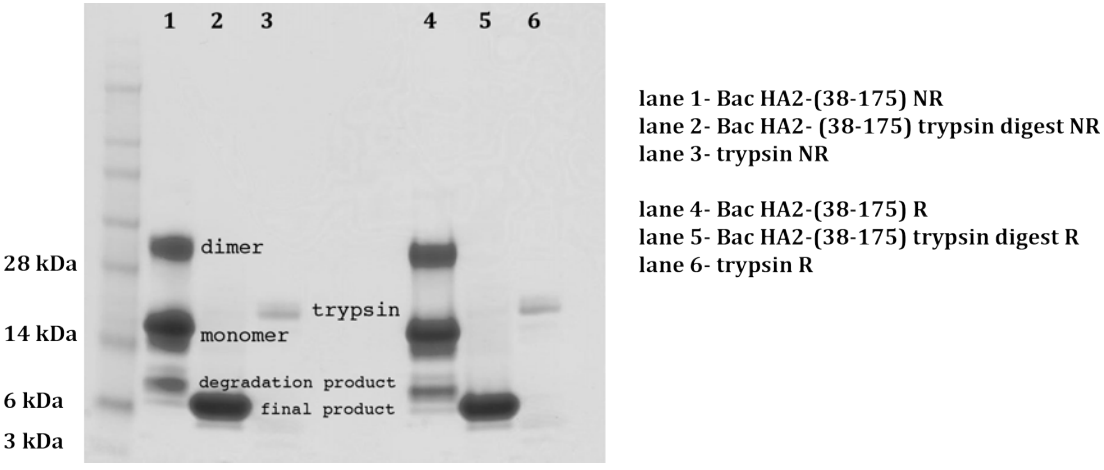


Figure 4.3 Assessment of the conformational state of Bac B59 H1 HA2 by limited proteolysis. Details of SDS-PAGE are included in Section 2.3.3.

Circular dichroism and thermal stability measurements:

Conformation of the recombinantly-expressed Bac B59 H1 HA2 (residues 38-175) was also assessed using circular dichroism spectroscopy (CD). Experimental details relating to CD are included in Section 2.4.3. Far UV spectra recorded at λ 195-260 nm allowed estimation of the secondary structure content. No significant differences in secondary structure at pH 5.2 and pH 7.2 were detected (Figure 4.4), suggesting a stable secondary structure in different buffer conditions. The observed high α-helical content (Tables 4.1 and 4.2) is characteristic of the extended α-helical coiled coil of postfusion H3 HA2 (Ruigrok, Aitken et al. 1988) (Chen, Wharton et al. 1995). The secondary content predictions shown in Tables 4.1 and 4.2 were calculated using three different algorithms: SELCON3 (Sreerama, Venyaminov et al. 1999), CONTINLL (van Stokkum, Spoelder et al. 1990), and CDSSTR (Sreerama and Woody 2000), and then averaged. The averaged content of secondary structure elements, turns and unstructured regions estimated for the recombinantly-expressed H1 HA2 (residues 38-175) is shown in % (Tables 4.1 and 4.2).

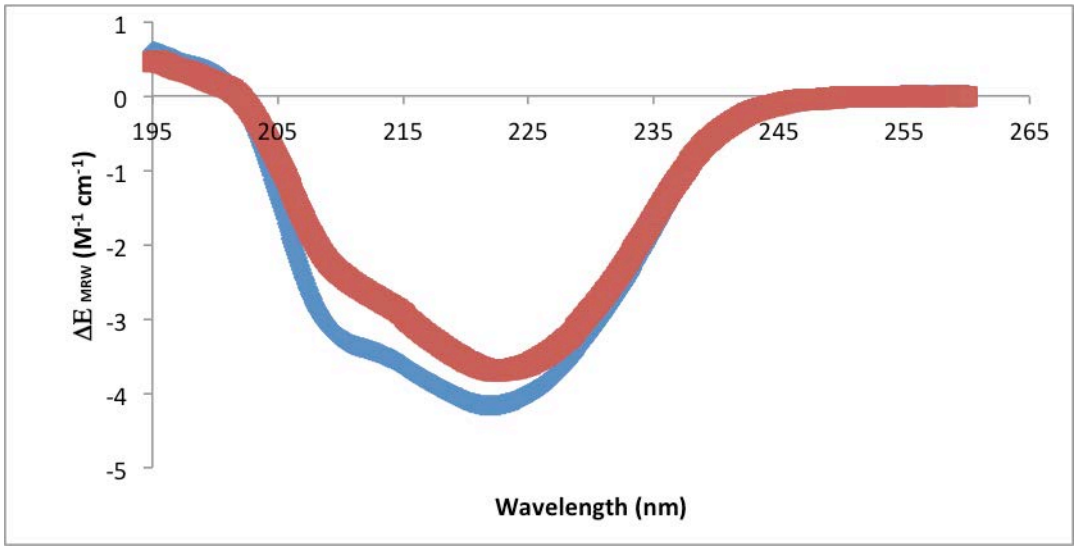


Figure 4.4 Far UV CD spectrum of Bac B59 H1 HA2. CD spectra obtained at pH 5.2 (red) and pH 7.2 (blue). Minima at 208 and 222 nm are characteristics of α-helices.

Table 4.1 Secondary structure content predictions for Bac B59 H1 HA2 at pH 5.2 by CD spectroscopy at λ=195-260 nm.

	a-helix	b-sheet	turn	unstructured	algorithm
pH 5.2	26.9	18.6	21.3	32.3	selcon 3
	25.8	12.4	18.6	43.2	continll
	34.5	17.3	17.5	30.2	cdsstr
	29.07	**16.10**	**19.13**	**35.23**	average

Table 4.2 Secondary structure content predictions for Bac B59 H1 HA2 at pH 7.2 by CD spectroscopy at λ=195-260 nm.

	a-helix	b-sheet	turn	unstructured	algorithm
pH 7.2	30.9	18.6	21	30.08	selcon 3
	33.3	7.3	20.5	38.8	continll
	40.9	17.1	15.6	26.5	cdsstr
	35.03	**14.33**	**19.03**	**32.03**	**average**

Thermal denaturation experiments performed at a fixed λ=222 nm, as described in Section 2.4.3, allowed assessing protein stability at pH 5.2 and pH 7.2, by observing protein unfolding transitions in relation to thermal denaturation. The melting temperature (T_m) was determined as a midpoint between folded and unfolded conformation. As presented in Figure 4.5, the expressed protein is more stable with respect to thermal denaturation at pH 5.2 (protein unfolds at T_m=72 °C) than pH 7.2 (T_m=67 °C). The temperature ranged from 40 to 100 °C. High thermal stability and pH dependence of unfolding transitions are also characteristics of H3 HA2 in a low pH conformation (Ruigrok, Aitken et al. 1988) (Chen, Wharton et al. 1995).

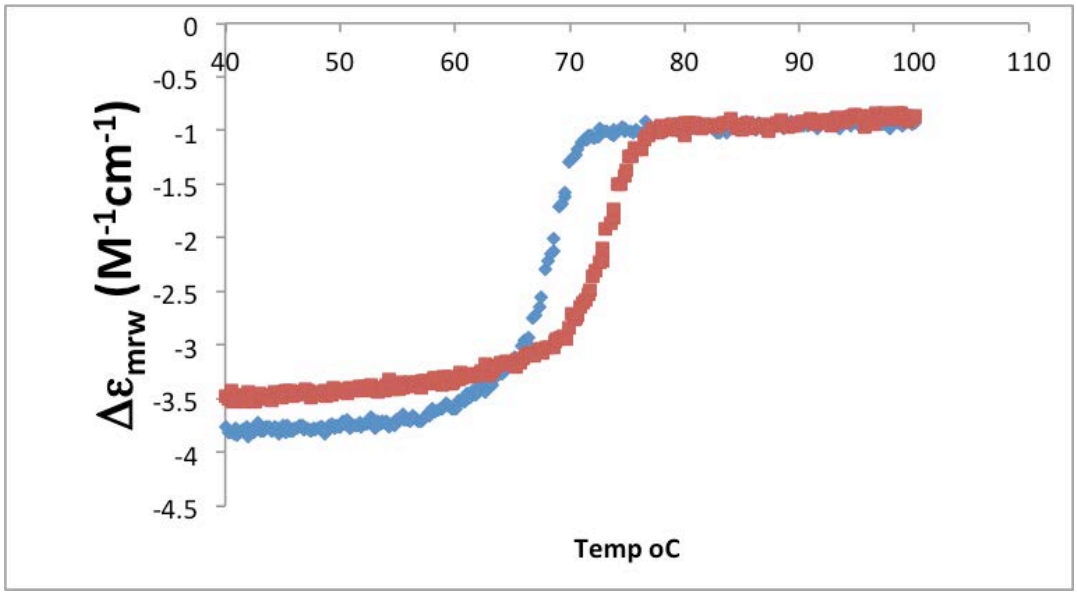


Figure 4.5 Thermal denaturation of Bac B59 H1 HA2. Denaturation curves at pH 5.2 (red) and pH 7.2 (blue) were obtained by CD spectroscopy at a fixed λ=222 nm over a temperature range 40-100 °C.

4.3 Comparison of recombinantly-expressed postfusion H1 HA2 and H3 HA2 using CD

4.3.1 Secondary structure

The far UV (λ 195-260 nm) spectra of recombinantly-expressed Bac B59 H1 HA2 (residues 38-175) (Section 4.2) predicted a high α-helical content associated with the extension of the trimeric stem of HA at low pH, and corresponding with the characteristic CD profile obtained for an equivalent fragment of *E.coli*-expressed H3 HA2 (residues 38-175) referred to as EBHA2 (Chen, Wharton et al. 1995), and for enzymatically prepared H3 TBHA2 (HA2 residues 40-162) (Ruigrok, Aitken et al. 1988).

4.3.2 Thermal stability and pH dependence of unfolding transitions

Thermal stability measurements described in Section 4.2 and shown in Figure 4.5 indicate, that Bac B59 H1 HA2 has high thermal stability, characteristic of EBHA2 (Chen, Wharton et al. 1995) and TBHA2 (Ruigrok, Aitken et al. 1988). Melting temperatures of unfolding transitions for the three constructs are compiled in Table 4.3.

Table 4.3 Melting temperatures for X31 H3 EBHA2, X31 H3 TBHA2 and Bac B59 H1 HA2 by CD spectroscopy at λ= 222 nm. Melting temperatures (T_m) for the three constructs, obtained at low (5.0-5.2) and high (7.0-7.2) pH are shown.

Construct	pH 5.0/5.2	pH 7.0/7.2
EBHA2	Tm=84°C	Tm=70°C
TBHA2	Tm=95°C	Tm=76°C
H1 HA2(38-175)	Tm=72°C	Tm=67°C

The X31 H3 EBHA2 is more stable to thermal denaturation at pH 5.2 than pH 7.2 by 14°C. Viral X31 H3 TBHA2 denatures at a temperature 19°C higher at pH 5.0 than at pH 7.0. In contrast, Bac B59 H1 HA2 unfolds at a temperature only 5°C higher at pH 5.2 than at pH 7.2. Although the high thermal stability of Bac B59 H1 HA2, and the observed pH dependence of unfolding transitions (Figure 4.5) are similar to those of X31 H3 EBHA2 and X31 H3 TBHA2, the recombinantly-

expressed postfusion H1 HA2 was shown to be less stable than H3 HA2 with respect to thermal denaturation.

4.4 Crystallisation

The expressed Bac B59 H1 HA2 was concentrated to ~10 mg/ml (A_{280}), using a Vivaspin 10K MWCO concentrator (Sartorius) at 3,000 g, 4°C, in a benchtop centrifuge (Thermo Scientific), and crystallized using a sitting drop vapour diffusion technique, as described in section 2.3.9. Crystals grew from conditions containing 12-20% PEG 3350, 0.1M Sodium citrate pH 5.6 and 2-7% [v/v] tacsimate. Optimisation trials involving varying the concentration of protein and precipitant resulted in diffraction quality crystals. Diffracting crystals were obtained from 18% PEG 3350, 0.1M Sodium citrate pH 5.6 and 2% [v/v] tacsimate pH 5.0 (Hampton Research). Concentration of protein varied between 5.6-8 mg/ml, and drops were seeded with a microseed solution, obtained from 15% PEG 3350 and 0.1M MES pH 6.0. Crystals are shown in Figure 4.6. Crystals were cryoprotected with a crystallisation solution containing 25% [v/v] ethylene glycol, and frozen by direct immersion in liquid nitrogen.

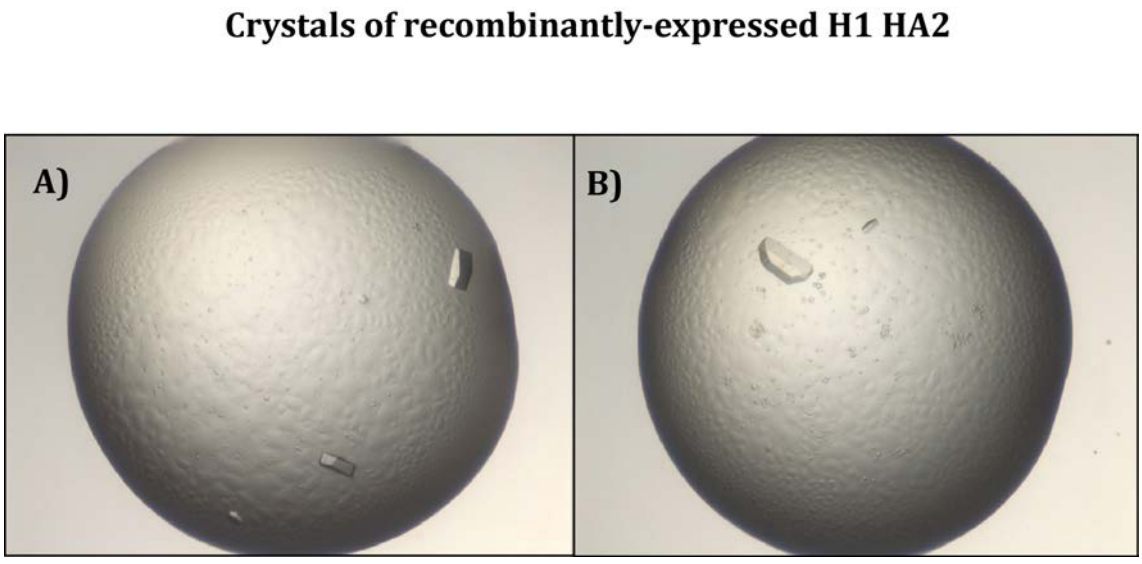


Figure 4.6 Crystals of Bac B59 H1 HA2. Crystals grown from 18% PEG 3350, 0.1M Sodium citrate pH 5.6, 2% [v/v] tacsimate are shown in A, and crystals grown from 14% PEG 3350, 0.1M Sodium citrate pH 5.6 and 7% tacsimate are shown in B.

4.5 Data collection and processing

Diffraction data were collected using an oscillation method (angle increment 0.15° and 0.5°/frame, exposure period 0.1 and 0.5 sec/frame, wavelength (λ) 0.9200, 0.96861 and 0.97625 Å), and recorded using a Pilatus detector by Dr Phil Walker (NIMR), at the IO4 beam line at Diamond Light Source Synchrotron (Harwell, UK).

Five complete datasets were collected, four of these were indexed and integrated using Xia2 system (Kabsch 2010) (Winter, Lobley et al. 2013). One dataset could not be indexed due to the poor quality of data and low resolution (~5.6 Å). The four other crystals gave an indicated space group of R 3 2 (155). The dimensions of the R 3 2-unit cell were: 55.97, 55.97, 431.16 (a, b, c) and 90.00, 90.00, 120.00 (α, ß, γ). The best dataset was chosen based on crystallographic statistics and data analysis using Aimless and Loggraph (CCP4) (Winn 2003). 1834 frames were collected (angle increment 0.15°/frame, exposure 0.1s/frame, λ=0.97625 Å), and the crystal diffracted to 3.4 Å. Resolution estimate and the anisotropic half-dataset correlation coefficient ($CC_{1/2}$) calculated by Aimless (CCP4), and shown graphically using Loggraph (CCP4) (Winn, Ballard et al. 2011), are shown in Figures 4.7 and 4.8, respectively. Data collection statistics are shown in Table 4.6.

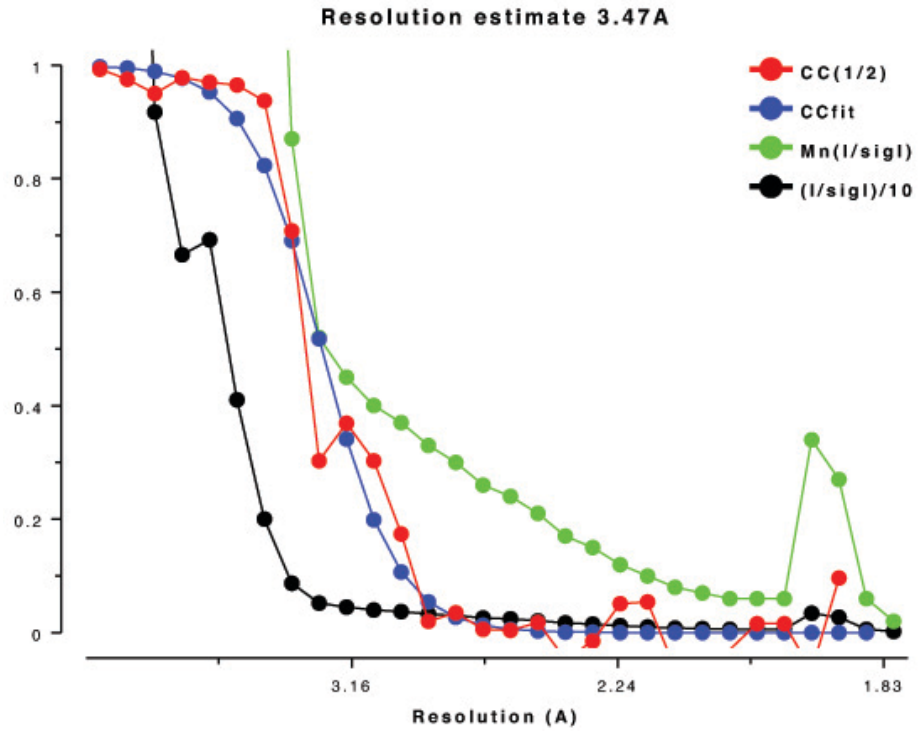


Figure 4.7 Resolution estimate for data collected from a crystal of Bac B59 H1 HA2 grown from 18% [v/v] PEG 3350, 0.1M Sodium citrate pH 5.6 and 2% [v/v] tacsimate. Data collected by oscillation method (0.15°/frame, exposure 0.1s/frame, λ=0.97625). Shown is the graphical output from Pointless log file obtained using Loggraph (CCP4) (Winn, Ballard et al. 2011).

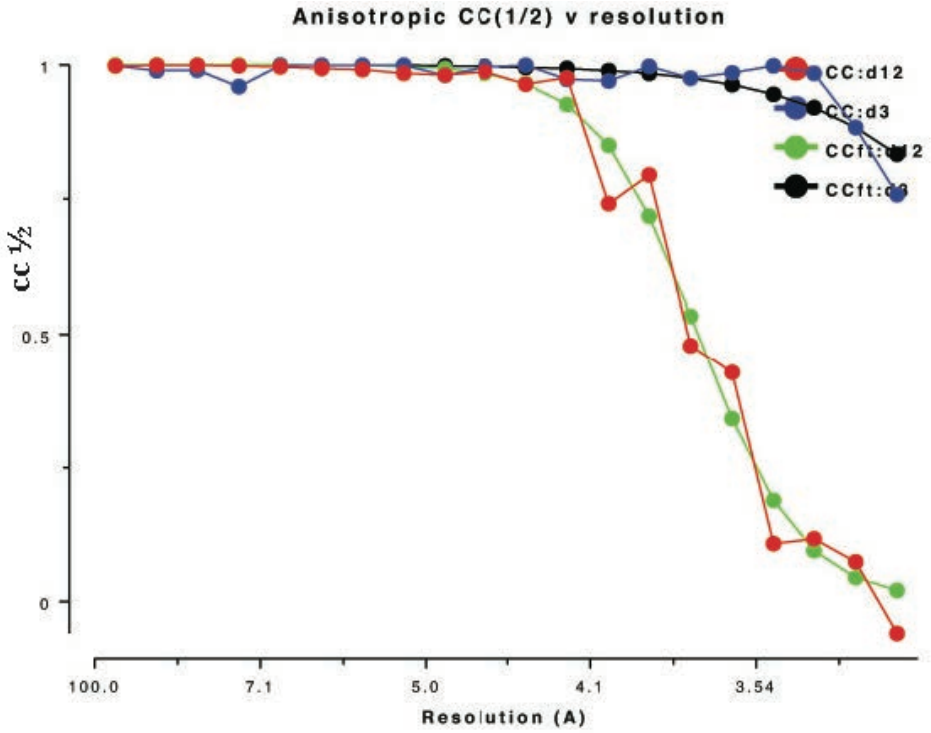


Figure 4.8 Anisotropy analysis of the 3.4 Å dataset obtained from a crystal of Bac B59 H1 HA2 grown from 18% PEG 3350, 0.1M Sodium citrate pH 5.6 and 2% [v/v] tacsimate. Graph prepared from Aimless log file (Evans and Murshudov 2013) using Loggraph (CCP4) (Winn 2003). Shown is the half-dataset correlation coefficient (CC1/2) for the directions d12 and d3 versus resolution.

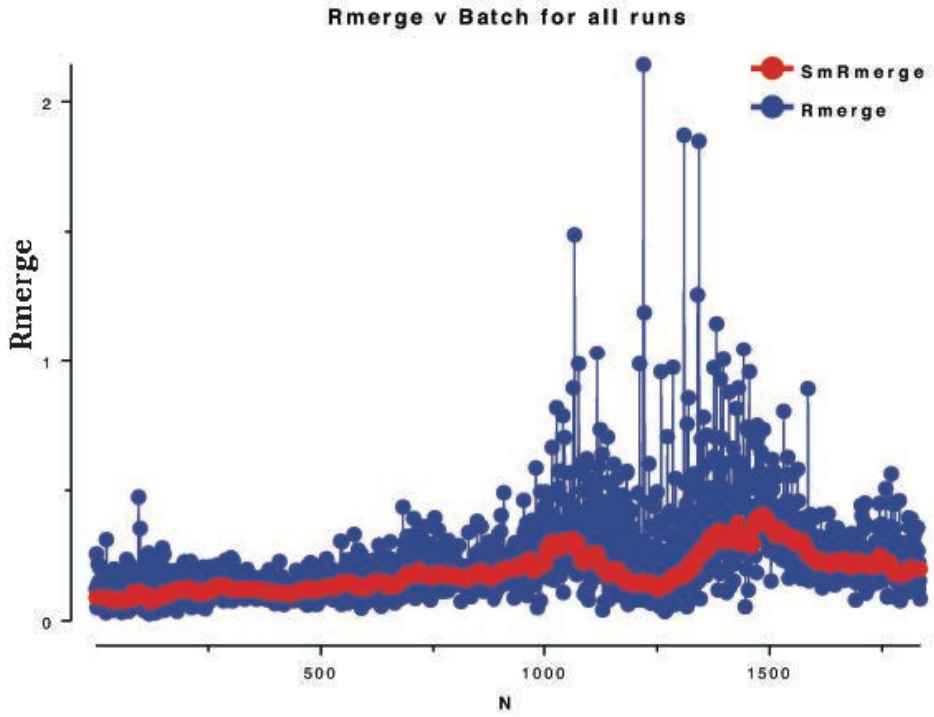


Figure 4.9 Distribution of Rmerge over 1834 frames collected at 0.15°/frame, 0.1s/frame exposure and λ=0.97625 from a crystal of Bac B59 H1 HA2, grown from 18% PEG 3350, 0.1M Sodium citrate pH 5.6 and 2% [v/v] tacsimate. Shown is the Rmerge versus batch calculated using Aimless and displayed using Loggraph (CCP4) (Winn, Ballard et al. 2011).

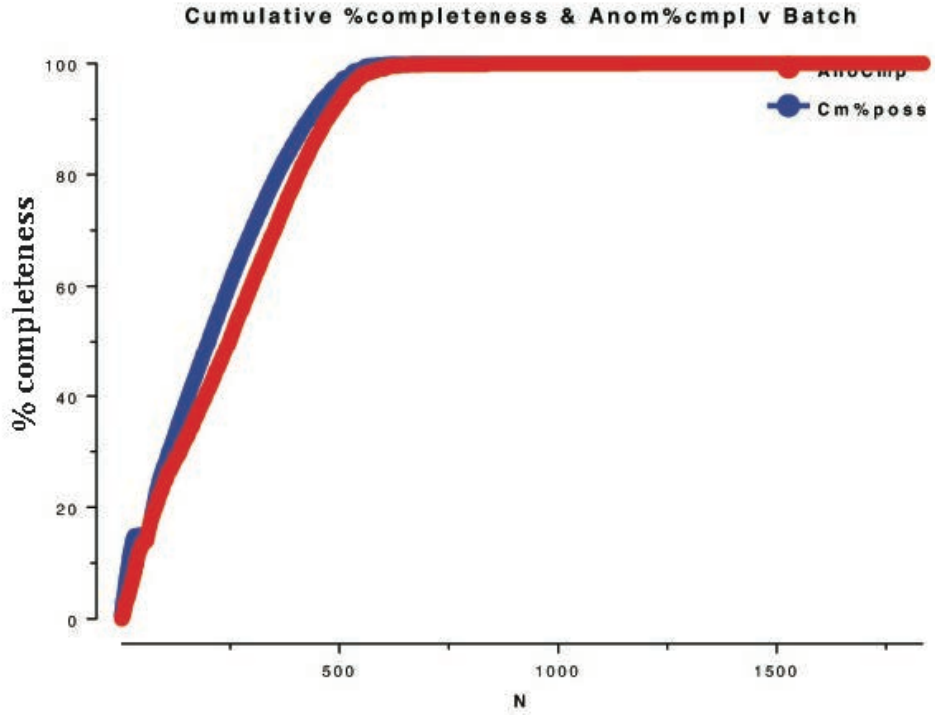


Figure 4.10 Data completeness versus the 1834 frames collected at 0.15°/frame, 0.1s/frame exposure and λ=0.97625 from a crystal of Bac B59 H1 HA2 grown from 18% PEG 3350, 0.1M Sodium citrate pH 5.6 and 2% [v/v] tacsimate. Shown is the % completeness versus batch N calculated using Aimless and displayed using Loggraph (CCP4) (Winn, Ballard et al. 2011).

Molecular replacement and model refinement:

The merged XDS ASCII reflection file was used for space group assignment using Pointless (CCP4) (Evans 2006) (Evans 2011) (Winn, Ballard et al. 2011). Pointless confirmed that the most likely space group was R 3 2 (155). Space group R 3 2 was assigned with a total space group probability estimate of 1.00, and systematic absences-based estimate of 1.00. Scores for the H -3 m Laue group are shown in Table 4.4, which includes scores for a likelihood measure (Lklhd), a "significance" score (Z-score), a correlation coefficient (CC), and the multiplicity-weighted R factor (Rmeas).

Table 4.4 Scores for H-3 m Laue group for a crystal of Bac B59 H1 HA2 calculated using Pointless (CCP4) (Winn, Ballard et al. 2011).

	Laue Group	Lklhd	NetZc	CC	Rmeas
1	H -3 m	1.00	3.49	1.00	0.00

The number of molecules in the asymmetric unit of the R 3 2 unit cell was calculated using Matthews-Coeff program (CCP4) (Kantardjieff and Rupp 2003). Based on the molecular weight, estimated from the number of residues in the PDB file (137 residues/monomer), and the calculated % solvent content there are most likely one or two monomers of postfusion H1 HA2 in the asymmetric unit of the R 3 2 crystal (Table 4.5).

Table 4.5 Number of Bac B59 H1 HA2 monomers in the asymmetric unit of the R 3 2 (155) unit cell and the corresponding % solvent content.

Cell volume: 1166135.5 13388 Da/monomer H32 (155)			
Nmol/asym	Matthews Coeff	% solvent	P (tot)
1	4.84	74.60	0.01
2	2.42	49.2	0.99
3	1.61	23.79	0

Molecular Replacement was initially attempted using a monomer of the *E.coli*-expressed H3 EHA2 (HA2 residues 33-173) (PDB: 1QU1) (Chen, Skehel et al. 1999) as the search model using MOLREP (Vagin and Teplyakov 2010) and PHASER (McCoy, Grosse-Kunstleve et al. 2007) however, no unique solution was found. A partial solution was found in a lower symmetry space group (P1), using a search model of H1 HA2 (residues 42-166), which gave interpretable density. This model was edited using the electron density as a guide to prepare a new search model for molecular replacement with the data processed in R 3 2. The new search model contained one long (HA2 residues 45-164), and one short (HA2 residues 75-128) monomer of H1 HA2, Figure 4.11. Molecular replacement with this search model was successful using both, PHASER and MOLREP (CCP4) (Winn, Ballard et al. 2011). The model of postfusion Bac B59 H1 HA2 was completed by building into the σ-weighted $2F_O$-F_C electron density map using Coot (Emsley and Cowtan 2004), and refined using Buster (Bricogne G. 2011). Figures for structural snapshots were generated using PyMol (Molecular Graphics System, Version 1.2r3pre, Schrödinger, LLC.). Data collection and refinement statistics are presented in Table 4.6. Ramachandran analysis of the final refined model is shown in Figure 4.12. Only three residues, Asn 154, Thr 156 and Thr 107 are outside the preferred regions.

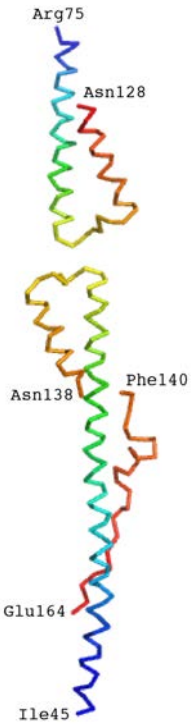
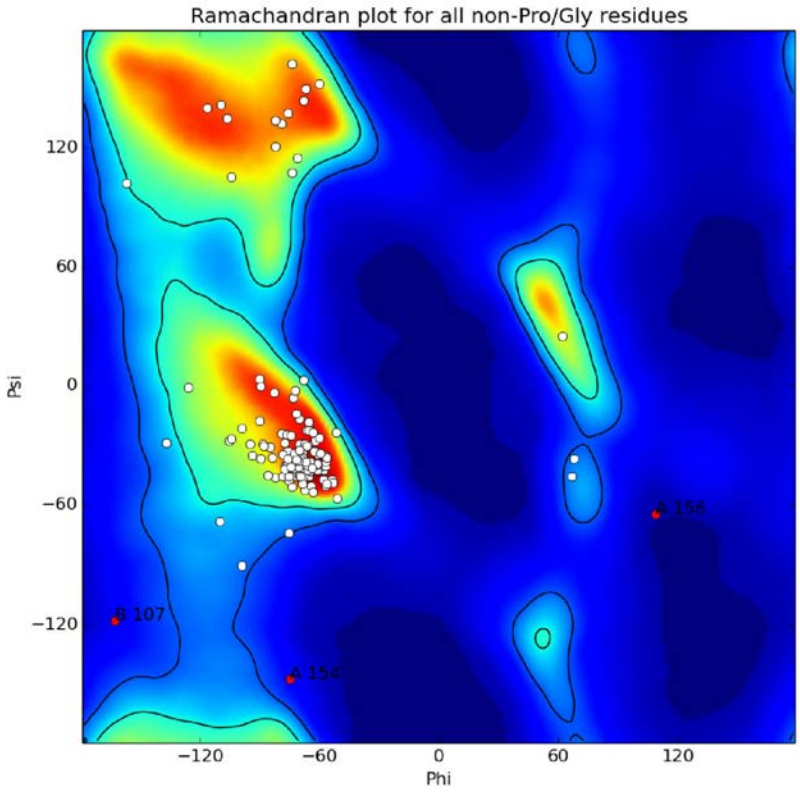


Figure 4.11 Model of postfusion H1 HA2 used as a search model for molecular replacement. The model was prepared after investigation of the electron density map obtained using PHASER and MOLREP (CCP4).

Table 4.6 Data collection and refinement statistics for Bac B59 H1 HA2 (molecular replacement).

Bac B59 H1 HA2	
Data collection	
Space group	R 3 2
Cell dimensions	
a, b, c (Å)	59.97, 59.97, 431.19
α, β, γ (◦)	90.00, 90.00, 120.00
Resolution (Å)	71.86-3.17 (3.25-3.41)*
R_{merge}	0.168 (3.36)*
Mean I/σ(I)	8.6 (0.7)*
Completeness (%)	100.0 (100.0)*
Redundancy	14.3 (14.4)*
Refinement	
Resolution (Å)	15.59-3.40
No.reflections	4809
R_{work}/R_{free}	0.29/0.30
No. protein atoms	1213
Protein B-factors	147.40
Rms deviations	
Bond lengths (Å)	0.009
Bond angles (Å)	1.18

*Values in parentheses are for highest-resolution shell.



Ramachandran plot outliers:

Chain	Residue	Score	Phi	Psi
A	ASN 154	0.02	-74.90	-146.70
A	THR 156	0.00	108.90	-64.10
B	THR 107	0.01	-163.00	-117.50

Figure 4.12 Ramachandran plot for the model of Bac B59 H1 HA2. Model obtained using Phenix GUI, Comprehensive validation (Afonine, Grosse-Kunstleve et al. 2012).

4.6 Crystal structure of Bac B59 H1 HA2

The sequence identity between influenza H3 and H1 HA2 is 52.5 % as calculated using ClustalW2 (Larkin, Blackshields et al. 2007). The aligned sequences of HA2 domains from the two HA subtypes are shown in Figure 4.13 A.

There are two monomers in the asymmetric unit of the R 3 2-unit cell, labelled A and B. Chain A is ordered from Ile45 (N-terminus of HA2) onwards. Chain B is shorter and starts at Arg75. The postfusion H1 HA2 adopts the hairpin-like conformation observed previously for H3 HA2 (Bullough, Hughson et al. 1994) (Chen, Wharton et al. 1995) (Chen, Skehel et al. 1999). The N-terminal region (residues 31-37), which forms an N-cap domain in the crystal structure of *E.coli*-expressed H3 HA2 (residues 31-37) (Chen, Skehel et al. 1999), and part of helix A (residues 38-41), are unstructured in the obtained crystal structure. Three segments form the postfusion stem of H1 HA2. These are: prefusion helix A (HA2 residues 42-58), prefusion loop B (HA2 residues 59-74), and prefusion helix C (HA2 residues 75-105). As for the postfusion H3 HA2 structures (Bullough, Hughson et al. 1994) (Chen, Skehel et al. 1999), region constant between the pre- and postfusion structures of H1 HA2 is helix C (HA2 residues 75-105). The long central stem of the refolded molecule is followed by a loop (D) (HA2 residues 106-112). This fragment is helical in the prefusion H1 HA2, and connects the central stem of the molecule to helix E (HA2 residues 113-127). A fragment of a loop (F) (HA2 residues 130-145) is unstructured in the crystal structure of Bac B59 H1 HA2, and weaker electron density for this fragment (HA2 residues 129-139) indicates greater structural flexibility. These fragments are present in the proteolytically prepared B59 H1 TBHA2 described in Chapter 3, and are stabilized by disulphide-linked fragments of HA1 domains (HA1 residues 11-17) (Figure 3.19). Loop F of the postfusion H1 HA2 extends to form another short helix G (HA2 residues 146-155), which runs in an antiparallel orientation to the central stem of the molecule. The C-terminal residues observed in the postfusion monomers of H1 HA2 are Glu164 in chain A, and Asn128 in chain B.

The core of Bac B59 H1 HA2 trimer is formed by a number of hydrophobic residues, which are either identical, or conserved in hydrophobicity between the two HA subtypes. One exception is Phe88, which is buried in postfusion H1 HA2. The corresponding residue in H3 HA2 is Lysine, the side chain of which appears to be exposed to solvent (PDB: 1QU1). The side chain of Phe88 in postfusion H1 HA2 points towards the 3-fold axis of H1 HA2 trimer, and is surrounded by two large

hydrophobic residues, Ile91 and Trp92. Trp92 appears to make a hydrophobic interaction with Val122 of the antiparallel helix E (HA2 residues 113-127), possibly stabilizing the six-helix-bundle. Hydrophobic interactions in the core of postfusion H1 HA2, and in the region of the six-helix-bundle are shown in Figure 4.13.

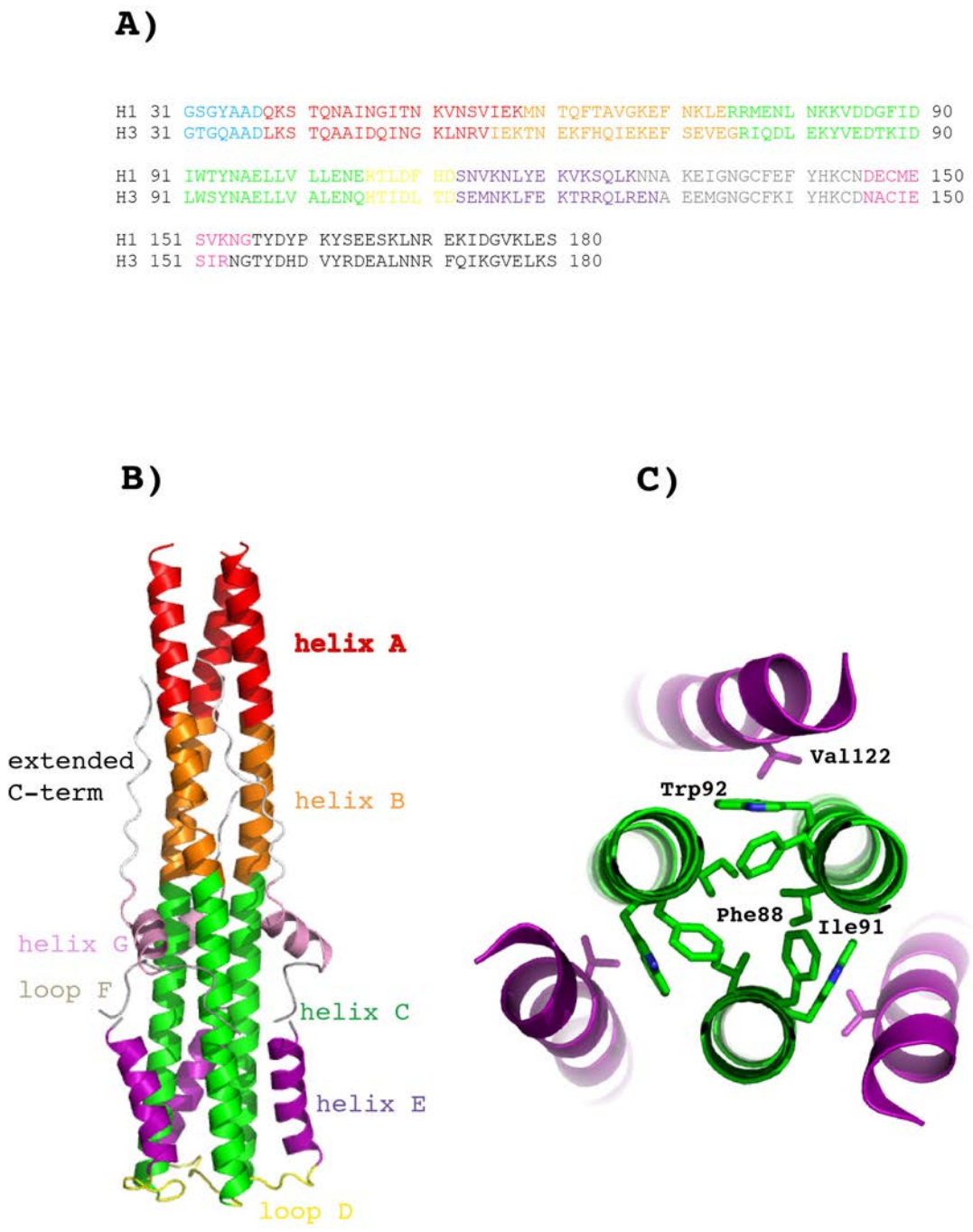


Figure 4.13 Crystal structure of Bac B59 H1 HA2. (A) Sequence alignment of H3 and H1 HA2 fragments containing residues 31-180. Residues are coloured by segments, which undergo refolding at low pH. (B) A trimer of H1 HA2 generated by symmetry operations using PyMol. (C) Burial of Phe88 in the protein core of refolded Bac B59 H1 HA2 (residues 38-175).

5 Postfusion H1 HA2 and its comparison to other HA2 structures

5.1 Viral and recombinantly-expressed H1 HA2

The crystal structure of postfusion viral B59 H1 TBHA2 is described in Chapter 3. One monomer (chain A) in the asymmetric unit (space group P 3 2 1) contains a fragment of a postfusion molecule (HA2 residues 65-155). Loop F (HA2 residues 128-145) is ordered, and connected to a short fragment of HA1 domain. The two domains are linked via a disulphide bond, formed between Cys137 of HA2 and Cys14 of HA1.

The crystal structure of postfusion, recombinantly-expressed Bac B59 H1 HA2 is described in Chapter 4. There are two monomers in the asymmetric unit (space group R 3 2). Electron density is discontinuous for both longer (chain A, residues 45-164), and shorter (chain B, residues 75-128) monomer, and the unobserved fragment comprises residues 129-139.

Differences in lengths of the observed fragments in the two crystal structures of postfusion H1 HA2, result most likely from proteolytic cleavage of the prepared protein post purification. Enzymatically prepared B59 H1 TBHA2 was most likely cleaved at Lys58. Although missing in the obtained structure of B59 H1 TBHA2, the crystal structure of Bac B59 H1 HA2 suggests, that the side chain of HA2 Lys58 points out from the trimer surface, and does not seem to be protected by the extended C-terminal fragment of the refolded molecule (Figure 5.1, A). Superposition of the longer monomer (chain A) of the Bac B59 H1 HA2 and a monomer of B59 H1 TBHA2 gives an overall rmsd of 0.6 Å (Figure 5.1, C).

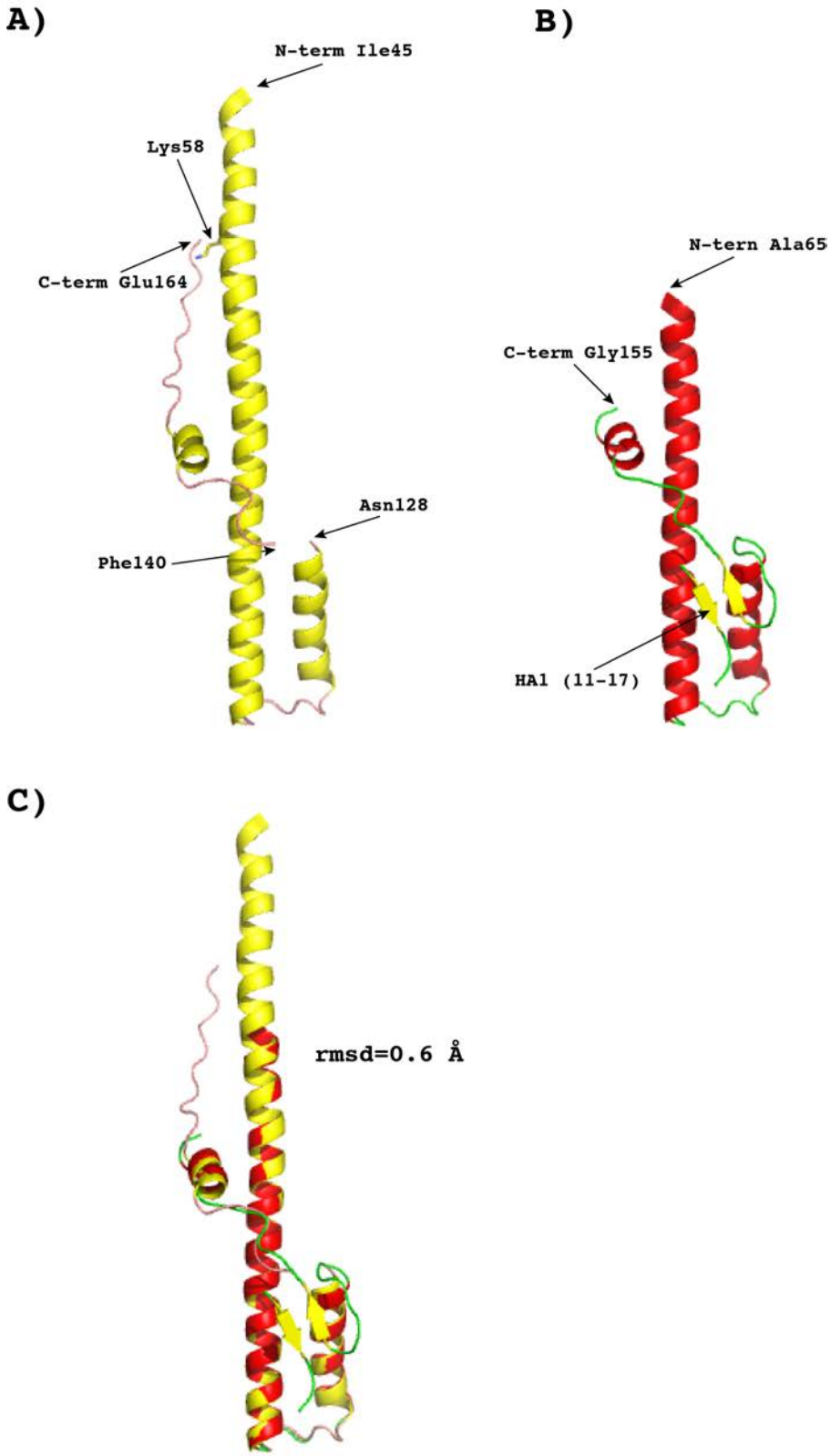


Figure 5.1 A model of postfusion influenza H1 HA2. A monomer of Bac B59 H1 HA2 is shown in A. A monomer of B59 H1 TBHA2 is shown in B. Superposition of the two monomers gives and overall rmsd of 0.6 Å and is shown in C. Key residues are indicated.

Segments for structural comparison of viral and recombinantly-expressed H1 HA2 in a postfusion conformation are shown in Table 5.1. Conformational changes of segments A to G, and the extended C-terminal fragment of H1 HA2 are summarized in Table 5.1. The appropriate residue range is shown. Segment A (residues 45-58) is helical in both pre- and postfusion H1 HA2, and is recruited to the central coiled coil following the structural refolding at low pH. Segment B (residues 59-74) is a random coil in prefusion H1 HA2 and adopts a helical structure at low pH. The constant region C (HA2 residues 75-105) is shown in green, and fragment of H1 HA2 that undergoes a 180° helix-to-turn transition (HA2 residues 106-112) is shown in yellow.

The corresponding regions of postfusion H1 HA2, which undergo the low pH-induced structural refolding, were aligned, and rmsd between the two structures was calculated using PyMol (Molecular Graphics System, Version 1.2r3pre, Schrödinger, LLC.). The calculated rmsd values are shown in Table 5.2.

Table 5.1 Refolding of H1 HA2 in response to low pH. Both pre- and postfusion conformations of segments A-G, and the extended C-terminal fragment of H1 HA2 are shown.

H1 HA2 segment (residue range)	H1 prefusion	H1 postfusion
A (38-58)	α-helix	α-helix
B (59-74)	random coil	α-helix
C (75-105)	α-helix	α-helix
D (106-112)	α-helix	turn
E (113-127)	α-helix	α-helix
F (128-145)	loop	loop
G (146-153)	α-helix	α-helix
extended C-term (154-167)	random coil	random coil

Table 5.2 Structural comparison of B59 H1 TBHA2 and Bac B59 H1 HA2. Rmsd between the corresponding fragments A-G, and the extended C-terminal fragment was calculated using PyMol.

Viral to recombinantly expressed H1 HA2 segments (residue range)	Rmsd (Å)
A (45-58)	missing in TBHA2
B (65-74)	0.536
C (75-105)	0.388
D (106-112)	0.357
E (113-127)	0.258
F (130-145)	missing in recombinant
G (146-153)	0.887
extended C-term (154-164)	missing in TBHA2
central helices (65-105)	0.437
Overall rmsd (monomers)	0.6

The overall reorganisation of H1 HA2 after its exposure to low pH is similar to that observed previously for H3 HA2 (Bullough, Hughson et al. 1994) (Chen, Skehel et al. 1999). Figure 5.2 shows a monomer of prefusion H1 HA2 (PDB: 1RUZ) and the two obtained structures of postfusion H1 HA2. Regions that undergo transition at the pH of membrane fusion are coloured according to colour scheme used in Tables 5.1 and 5.2. A short fragment of HA1 remains disulphide-linked to HA2, and the disulphide bond is between Cys14 of HA1 and Cys137 of HA2. Relocation of helix A to the top of the central coiled coil extends the structure towards the target membrane. Loop B characterized by a tall turn at residue 75 in the prefusion structure of group 1 HA, is also recruited to the central coiled coil. Short helical fragment D (HA2 residues 106-112), which makes contacts with the fusion peptide in the prefusion conformation, refolds to form a loop, which reverses direction of the HA2 polypeptide chain by 180°, and results in positioning of the C-terminal fragment of postfusion H1 HA2 near the N-terminus of the postfusion molecule (Figures 5.1 and 5.2).

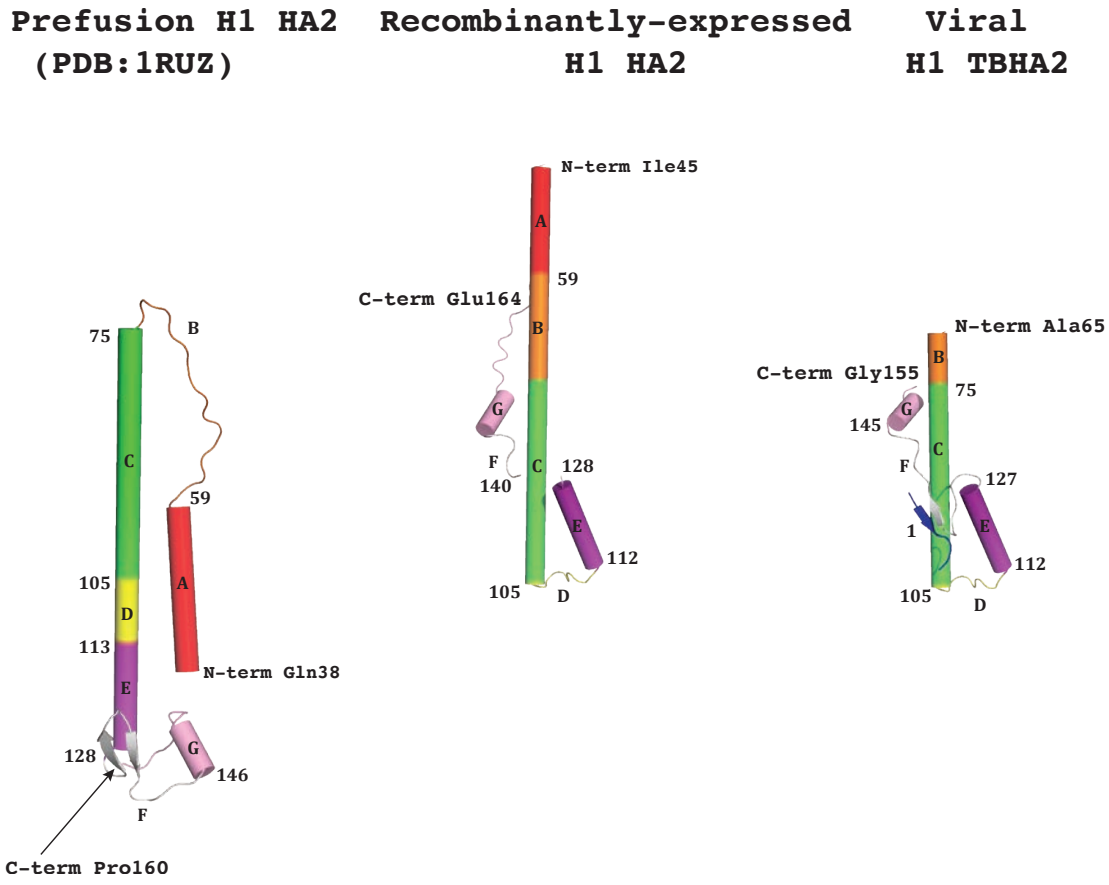


Figure 5.2 The low pH-induced refolding of H1 HA2. A prefusion monomer of H1 HA2 (PDB: 1RUZ) is shown on the left, a monomer of recombinantly-expressed H1 HA2 is shown in the middle, and a monomer of viral postfusion H1 HA2 is shown is on the left. Segments A-G, which undergo refolding at low pH are shown as cylinders.

5.2 Postfusion viral H1 and H3 TBHA2

A monomer of postfusion viral H1 HA2 (residues 65-155) present in the crystal structure of B59 H1 TBHA2 and described in Chapter 3, is 25 amino acid residues shorter at its N-terminus than the previously solved structure of X31 H3 TBHA2, which starts at residue 40 of HA2, and contains residues 1-27 of HA1 (Bullough, Hughson et al. 1994). Helix A, together with first ~5 residues of postfusion helix B are unobserved in the crystal structure of low pH-activated B59 H1 TBHA2. The unstable nature of the N-terminal fragment of the refolded B59 H1 TBHA2 may be related to the exposure of additional, previously hidden protease cleavage sites after exposure of the molecule to low pH (~5.0). The appropriate cleavage sites at the N- and C-termini of H1 HA2 are shown in Appendix 2. The pre- and postfusion conformations of segments A-G and the C-terminal fragment of H3 HA2 (Bullough, Hughson et al. 1994) (Chen, Skehel et al. 1999) (Gamblin, Haire et al. 2004) are summarized in Table 5.3. Four helices and three ß-strands are conserved between the pre- and postfusion H1 and H3 TBHA2 (Bullough, Hughson et al. 1994). Helix A of prefusion H3 HA2 (Table 5.3) is three amino acids shorter than that of H1 HA2 (Table 5.1). In prefusion H1 HA2, Ile56, Glu57 and Lys58 belong to helix A, and are then recruited to the top of the postfusion molecule. Structural alignments were carried out using PyMol (Molecular Graphics System, Version 1.2r3pre, Schrödinger, LLC.), and alignments were preceded by a removal of the atomic coordinates corresponding to fragments of HA1 from the two PDB files. The root-mean-square deviation (rmsd) between the equivalent atoms was calculated according to the formula:

$$RMSD = \sqrt{\frac{SUM\ (dii)^2}{N}}$$

Where:

RMSD- root-mean-square deviation of the Ca atomic coordinates after rigid body superposition

dii- distance between equivalent atoms

N- number of atoms matched in each structure

Alignment of prefusion H3 and H1 monomers gives an rmsd of ~1 Å and is shown in Figure 5.3. Structural alignment of prefusion H3 and H1 HA2 monomers and the corresponding rmsd values are presented in Table 5.4. Rmsd for the given regions was calculated by superposing the specified residues of prefusion H3 HA2 (PDB: 1HGF) (chain B) and prefusion H1 HA2 (PDB: 1RUZ) (chain I).

Due to differences in lengths of helices A between the prefusion monomers of H3 (PDB:1HGF) and H1 HA2 (PDB:1RUZ) shown in Figure 5.3, and the lengths of enzymatically prepared X31 H3 TBHA2 (PDB: 1HTM) (Bullough, Hughson et al. 1994) and B59 H1 TBHA2 presented in this thesis (Chapter 3), the postfusion structures of viral HA2 from the two groups HA were compared, by calculating rmsd for the corresponding regions present in both crystal structures (HA2 residues 65-155). Structural alignment of postfusion H3 (PDB: 1HTM) (chain F) and B59 H1 TBHA2 (chain A) monomers, and the corresponding rmsd values, are presented in Table 5.5. With rmsd of ~4 Å, loops B that connect helices A to central helices C in prefusion HA2, are structurally different between the two HA subtypes (Table 5.4), and adopt a very similar conformation (rmsd=0.46 Å) after refolding of the molecules at low pH (Table 5.5 and Figure 5.4). Helices C (HA2 residues 76-105) remain similar after the low pH-induced conformational change, with a slight increase in the rmsd value from ~0.4 Å to ~0.6 Å. Segments D (HA2 residues 106-112) are helical in the prefusion structures, and can be aligned very well between prefusion H3 and H1 HA2 (rmsd=~0.3 Å). Turns formed by a helix-to-turn transition of these residues are different in the two postfusion viral HA2 structures, which is reflected by an increase in rmsd to ~0.8 Å (Tables 5.4 and 5.5). Threonine at position 107 of H1 HA2 packs towards the 3-fold axis of the HA2 trimer, and may form weak hydrogen bonds with carbonyl oxygen atom of Thr107 belonging the neighbouring HA2 subunit (Figure 5.5). Leucine at position 108 of H1 TBHA2 is located in the middle of two adjacent helices C and E of each monomer. Although the exact orientation of all amino acid side chains in this region of the refolded B59 H1 TBHA2 cannot be determined precisely due to low resolution, it appears, that this part of the rearranged molecule is held mainly by hydrophobic interactions between methyl groups of Threonine residues of the neighbouring monomers and weak hydrogen bonds. Interhelical loops D (HA2 residues 106-112) change the direction of the polypeptide chain in crystal structures of both H3 (Bullough, Hughson et al. 1994) and H1 TBHA2. The observed postfusion structures of these regions of viral, enzymatically prepared, and low pH-activated H3 and H1 TBHA2, are compared in Figure 5.5.

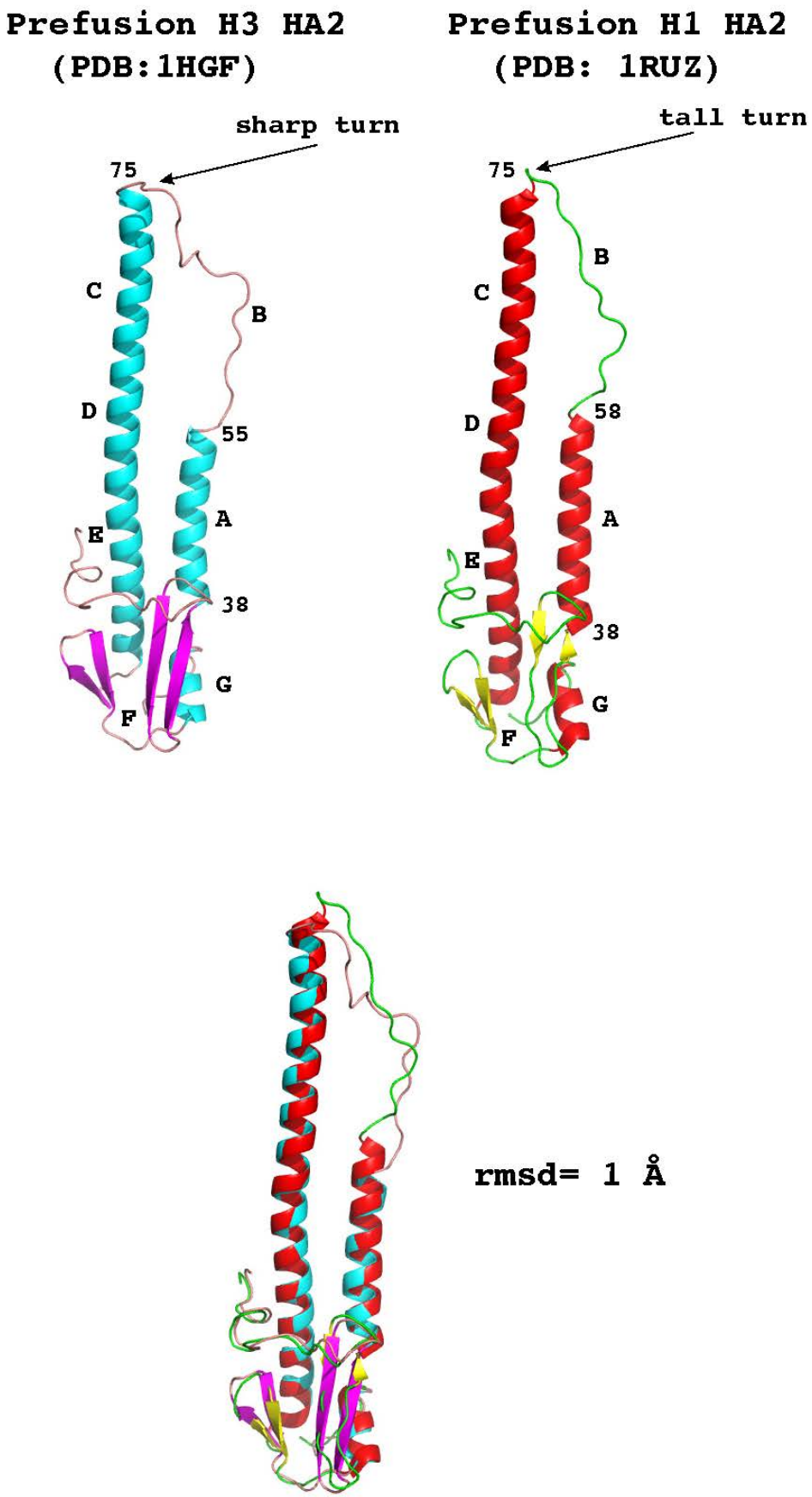


Figure 5.3 Structural alignment of prefusion H3 and H1 HA2. Monomers of H3 HA2 (PDB: 1HGF) and H1 HA2 (PDB: 1RUZ) are shown in blue and red, respectively. Segments A-G, key residues and group-specific differences in turns are indicated. Alignment of the corresponding fragments gives an overall rmsd of 1 Å.

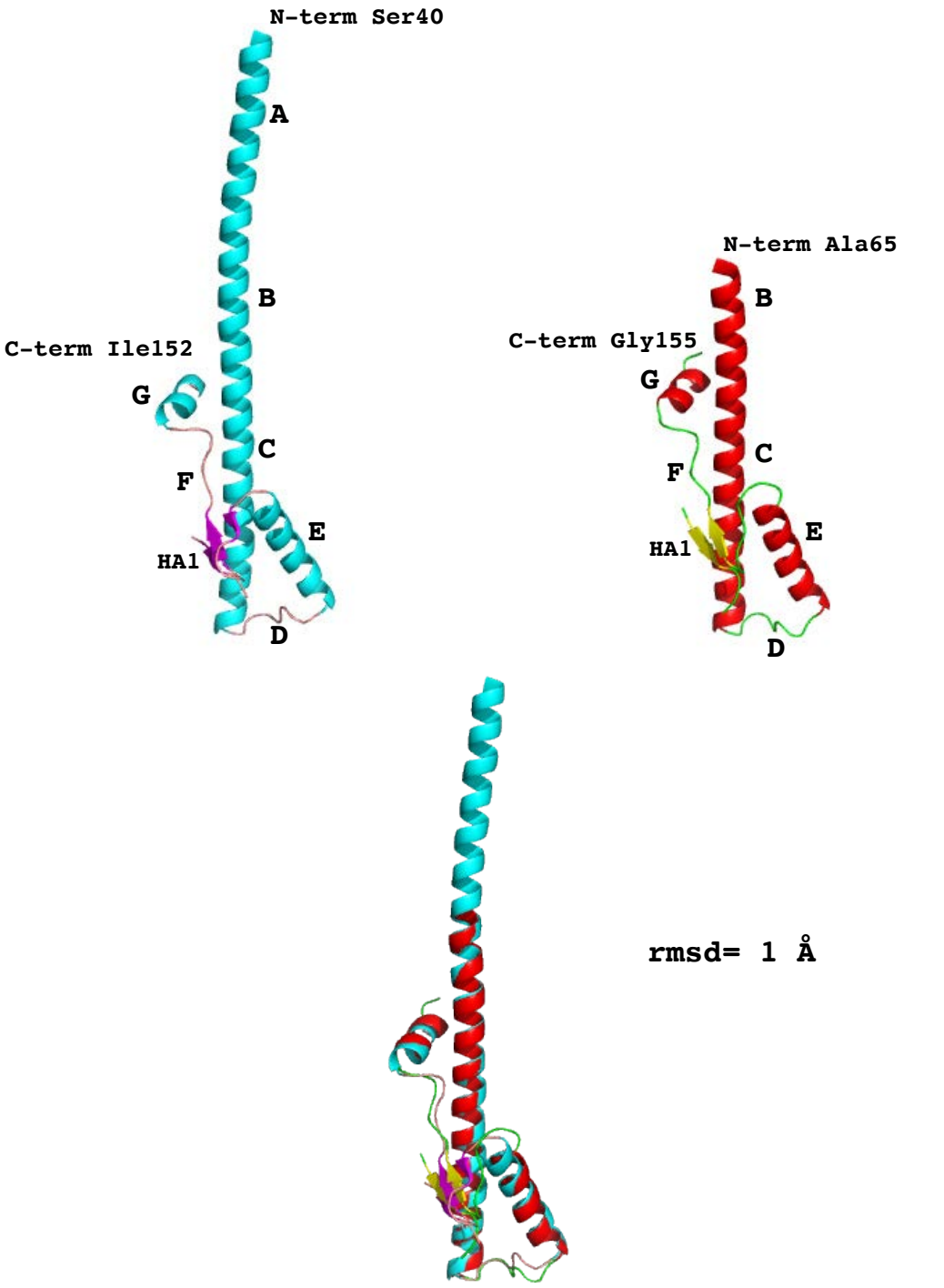


Figure 5.4 Structural alignment of postfusion X31 H3 TBHA2 and B59 H1 TBHA2. Alignment of X31 H3 TBHA2 (PDB: 1HTM) and B59 H1 TBHA2 monomers gives an overall rmsd of 1 Å. The X31 H3 TBHA2 monomer was trimmed at its C-terminus.

Table 5.3 Refolding of H3 HA2 in response to low pH. Both pre- and postfusion conformation of segments A-G and the extended C-terminal fragment of H3 HA2 are shown.

H3 HA2 segment (residue range)	H3 prefusion	H3 postfusion
A (38-55)	α-helix	α-helix
B (56-75)	random coil	α-helix
C (76-105)	α-helix	α-helix
D (106-112)	α-helix	turn
E (113-129)	α-helix	α-helix
F (130-145)	loop	loop
G (146-153)	α-helix	α-helix
extended C-term (154-185)	random coil	random coil

Table 5.4 Structural comparison of prefusion H3 HA2 (PDB: 1HGF) and H1 HA2 (PDB: 1RUZ). Rmsd between the corresponding fragments A-G was calculated using PyMol.

H3 to H1 prefusion HA2 segments (residue range)	Rmsd (Å)
A (38-55)	0.508
B (59-74)	**4.138**
C (76-105)	0.423
D (106-112)	0.299
E (113-127)	0.314
F (130-145)	0.880
G (146-153)	0.398
extended C-term (154-160)	0.993
Overall rmsd (monomers)	1.005

Table 5.5 Structural comparison of X31 H3 TBHA2 (PDB: 1HTM) and B59 H1 TBHA2. Rmsd between the corresponding fragments A-G was calculated using PyMol.

H3 to H1 postfusion TBHA2 segments (residue range)	Rmsd (Å)
B (65-74)	0.458
C (76-105)	0.650
D (106-112)	0.812
E (113-127)	0.408
F (130-145)	1.541
G (146-155)	0.647
central helices (65-105)	0.492
Overall rmsd (HA2 monomers+ HA1 chains)	1.056

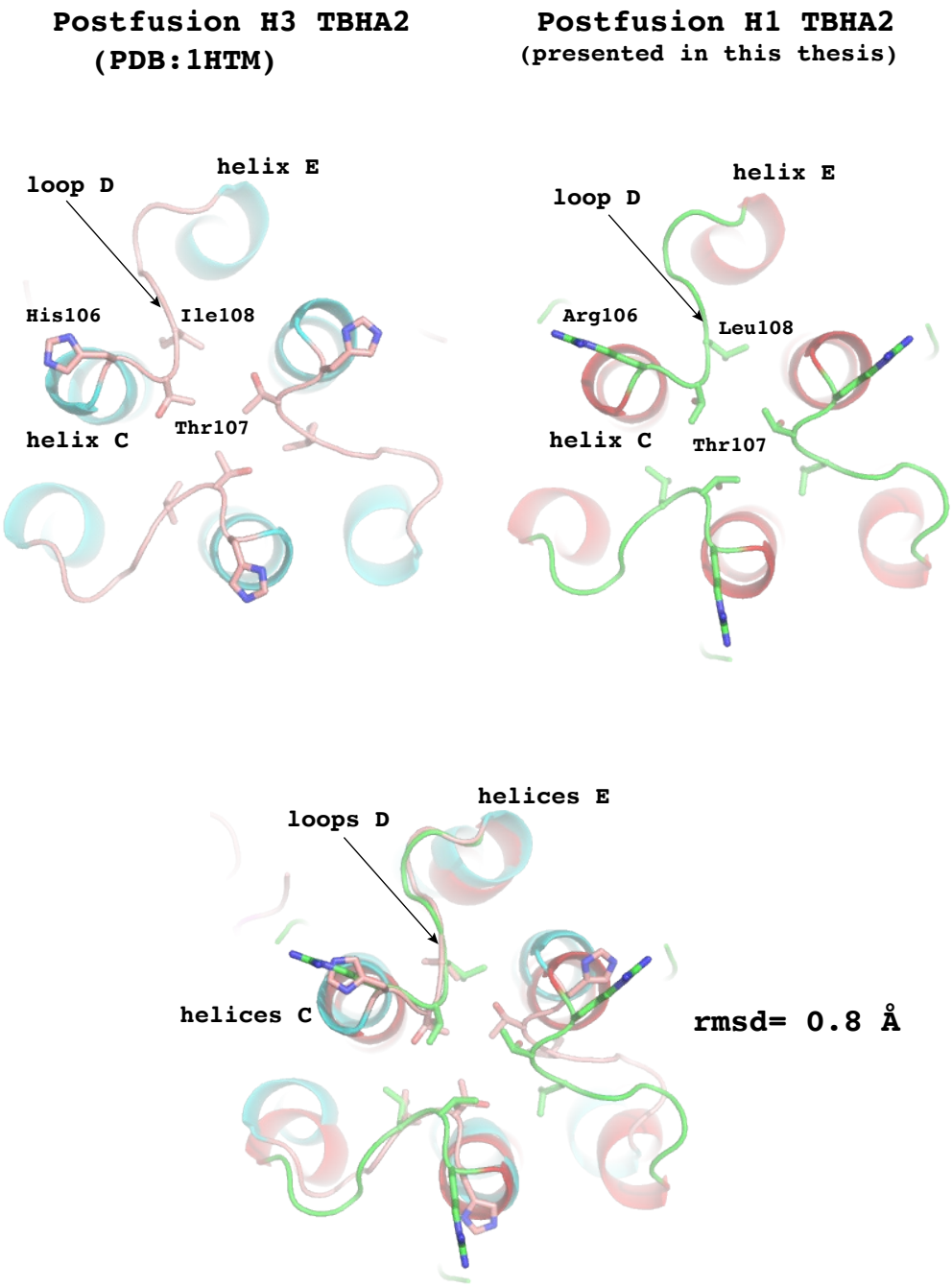


Figure 5.5 Comparison of turns in X31 H3 TBHA2 and B59 H1 TBHA2. View down the 3-fold symmetry axis of H3 (PDB: 1HTM) and H1 TBHA2 trimers generated by symmetry operations in PyMol. Key residues are indicated. Alignment of equivalent regions comprising the C-terminal part of helix C, loop D (HA2 residues 106-112) and the N-terminal end of helix E is shown. Threonine 107 packs closely towards the 3-fold axis of both refolded molecules indicating, that this region is held mainly by hydrophobic interactions. Superposition was carried out by an alignment of HA2 residues 106-112 of the two monomers.

The antiparallel helices E comprise residues 113-129 in H3 TBHA2, and residues 113-127 in B59 H1 TBHA2. These helices pack into the groove of two neighbouring helices C in both structures (Figure 5.5). The structural deviation between helices E of viral, pre- and postfusion H3 and H1 HA2 is low (~0.3 and ~0.4 Å, respectively), when calculated by superposing the entire HA2 monomers. Alignment of central helices (HA2 residues 65-105), revealed, that angles between helices E and helices C in the postfusion structures of H3 and H1 TBHA2 are different, and that helices E of postfusion B59 H1 TBHA2 sit above the equivalent helices in postfusion viral X31 H3 TBHA2. The observed relative position of helices E in the two structures of postfusion viral TBHA2 is shown in Figure 5.6.

Residues 130-145 form a loop (F) in postfusion HA2 of both subtypes, and these loops are disulphide linked to short fragments of HA1. The connecting loops F are different in both pre- (rmsd ~0.9 Å), and postfusion (rmsd ~1.5 Å) viral H3 and H1 HA2 structures, indicating a flexible nature of these fragments. The disulphide bond is formed between Cys137 of HA2 and Cys14 of HA1 in both subtypes (H3 numbering). Loops at residues 130-145 of HA2, and the attached fragments of HA1, are shown in Figure 5.6.

The shortest helices G (HA2 residues 146-153) observed in the two postfusion TBHA2 (H3 and H1) monomers can be aligned with an rmsd of ~0.65 Å. Residues beyond residue 155 are disordered in the B59 H1 TBHA2 structure, and were excluded from the structural comparison. While the rmsd value obtained by superposing the entire monomers of postfusion H1 and H3 TBHA2 (Bullough, Hughson et al. 1994) is ~1.06 Å, alignment of central helices (HA2 residues 65-105) reduces this value to ~0.5 Å.

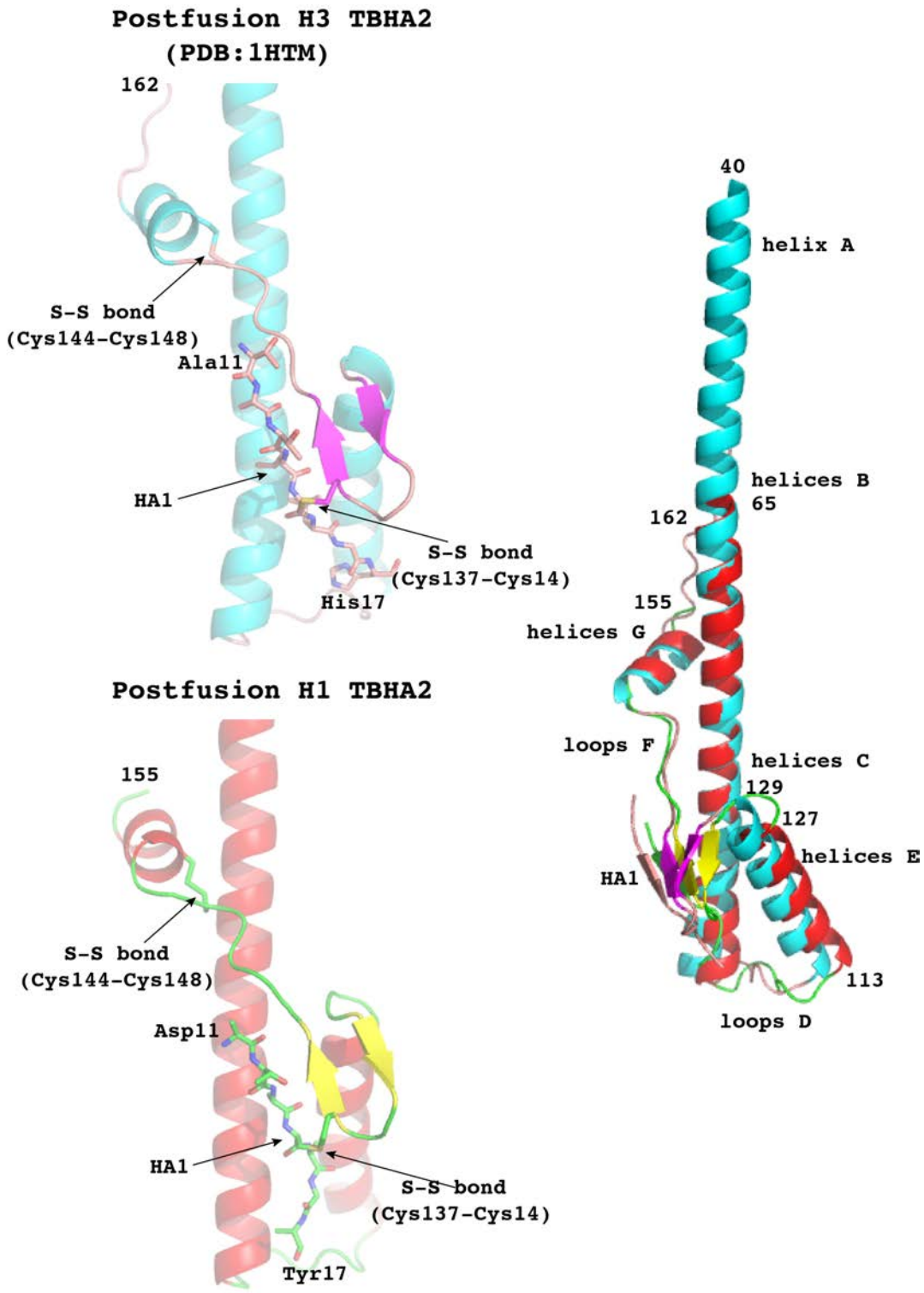


Figure 5.6 Loops at residues 130-145 of X31 H3 TBHA2 and B59 H1 TBHA2. Monomers of H3 TBHA2 (PDB: 1HTM) and B59 H1 TBHA2 are shown in blue and red, respectively. Loops at HA2 residues 130-145 are shown as cartoon (purple and yellow), HA1 chains are shown as sticks, and S-S bonds are indicated (H3 numbering). The two monomers superposed by aligning residues HA2 residues 65-105 (main helices) of the two HA subtypes are shown on the right.

A monomer of enzymatically prepared B59 H1 TBHA2 (Chapter 3) contains HA2 residues 65-155. The polypeptide chain starts at Ala65 (N-terminus), indicating cleavage by trypsin at Lys58 or cleavage by contaminating proteases near Thr64. In prefusion H1 HA2, Lys58 belongs to helix A, which is three amino acids longer than helix A of prefusion H3 HA2 (Figure 5.3). The proteolytically-prepared postfusion B59 H1 TBHA2 may have been cleaved after induction of the conformational change of H1 HA by incubation at the pH of fusion, and upon exposure of the previously buried protease cleavage sites, or possibly post-purification. In comparison, a monomer of postfusion X31 H3 TBHA2 (PDB: 1HTM) (Bullough, Hughson et al. 1994) prepared using a similar method is 25 amino acids longer than that of B59 H1 TBHA2, and contains HA2 residues 40-162.

The observed increased susceptibility of B59 H1 TBHA2 to proteolytic digestion in the region of HA2 residues 40-65 was analysed in relation to the probability of coiled coil formation. Primary sequences of H1 and H3 HA2 were compared to a database of known coiled coils, and the probability of their formation was calculated using COILS/PCOILS (Bioinformatics Toolkit; Max-Planck Institute for Developmental Biology). The program uses a matrix specific for left-handed coiled coils, and produces a set of probabilities of coiled coil forming potential for a given sequence. As shown in Figure 5.7, coiled coil prediction profiles for both H3 and H1 HA2, correlate directly with the lengths of the monomers observed in crystal structures of X31 H3 TBHA2 (PDB: 1HTM) (Bullough, Hughson et al. 1994) and B59 H1 TBHA2 (Chapter 3). The predicted propensity of HA2 residues 65-90 to form a stable coiled coil correlates with the length of the molecules observed in the crystal structures of both enzymatically-prepared TBHA2. The observed length of B59 H1 TBHA2 (Ala65-), is consistent with the low coiled coil propensity predicted based on the primary sequence of this fragment of H1 HA2 (residues 1-64). Surprisingly, a relatively low coiled coil forming potential calculated for the equivalent fragment of H3 HA2, is sufficient for H3 HA2 residues 1-64 to remain folded as a stable super helix, as observed in the crystal structure of X31 H3 TBHA2 (PDB: 1HTM) (Bullough, Hughson et al. 1994).

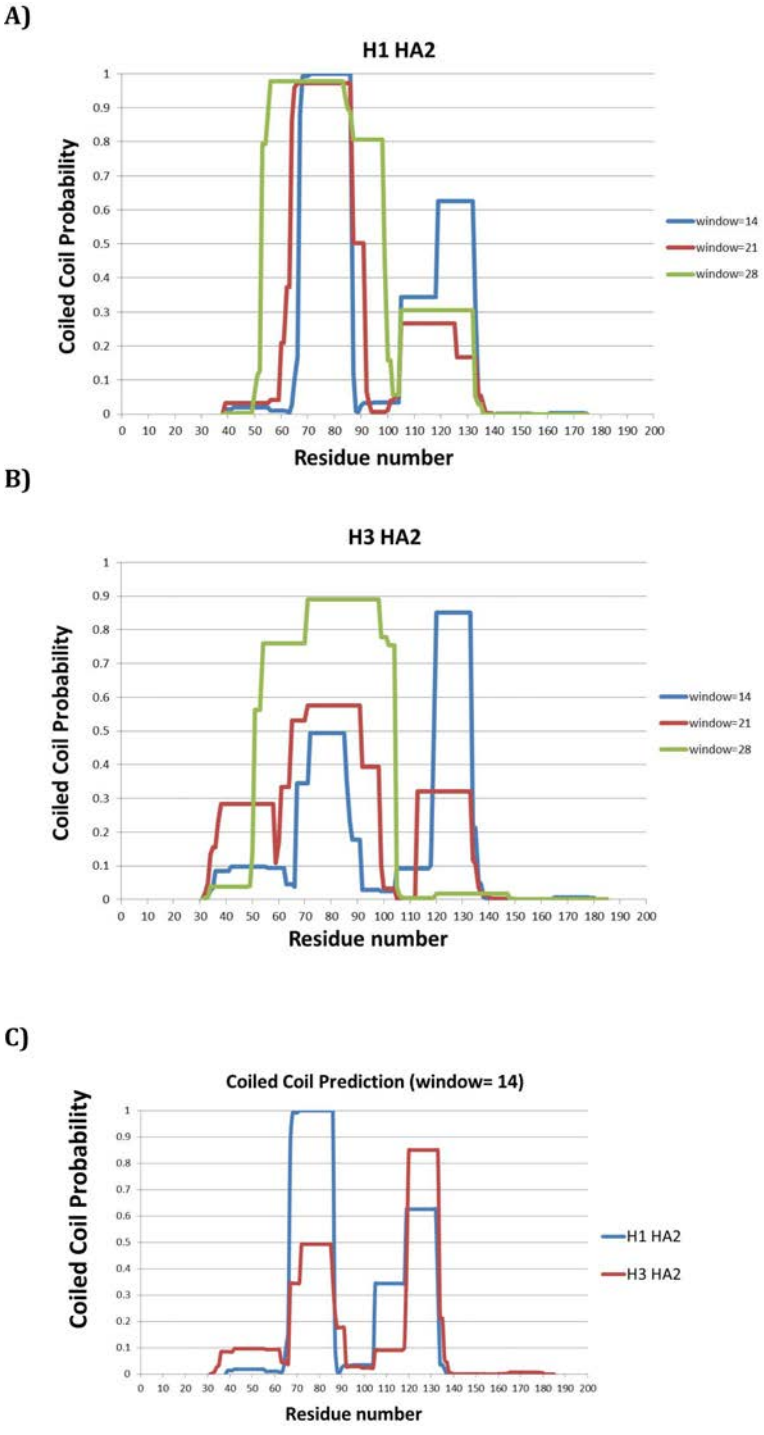


Figure 5.7 Coiled coil prediction profiles for H1 and H3 HA2. The helical propensity of H1 HA2 is shown in A, and coiled coil prediction for H3 HA2 is shown in B. An overlap of coiled probability for H1 and H3 HA2 (window=14) is shown in C. Figure prepared using COILS/PCOILS (Bioinformatics Toolkit; Max-Planck Institute for Developmental Biology).

As shown in Figure 5.8, the N-terminal region of H3 HA2 (residues 38-86), comprising helix A, loop B and part of helix C, contains more hydrophobic (green), more charged (grey), and less polar (red) amino acids, than the equivalent fragment of H1 HA2. The overall distribution of amino acids in this fragment of H3 HA2 a higher coiled coil forming potential than that of the equivalent fragment of H1 HA2, which may be related to the observed differences in lengths of the two molecules when prepared proteolytically. Amino acid substitutions that may be responsible for differences in strengths of the hydrophobic component of H3 and H1 HA2 in the region of residues 38-64 are marked (*). Residues 52 and 59 of HA2 are group-specific, and these are Valine and Methionine in H1, H2 and H5 (group 1 HA), and Leucine and Threonine in H3, H7 and H10 (group 2 HA). Sequence alignments of representative subtypes from two groups HA are included in Figure 1.13. Differences in amphipathicity of the corresponding regions of H1 and H3 HA2 (residues 38-64) were also visualized as helical wheels. As presented in Figure 5.9, the apolar stripe defined by hydrophobic residues in positions *a* and *d* of H3 HA2 (A), contains a very strong cluster of hydrophobic residues. The opposite face of H3 HA2 helix contains a strong cluster of hydrophilic residues (Q42, N49, N53 and N60), and in total there are four heptad repeats in this fragment of H3 HA2. The corresponding residues of H1 HA2 appear to form three heptads, the frame of which is shifted by 3 amino acids (B).

	H1 (PDB: 1RUZ)	coiled coil probability	H3 HA2 (PDB: 1HGF)	coiled coil probability	
	38 Q *	0.001	38 L *	0.085	
Lys39 ←	**39 K**	**0.014**	**39 K**	**0.085**	→ Lys39
	40 S	0.014	40 S	0.085	→ Ser40
	41 T	0.014	41 T	0.085	
	42 Q	0.019	42 Q	0.097	
	43 N *	0.019	43 A *	0.097	
	44 A	0.019	44 A	0.097	
	45 I	0.019	45 I	0.097	
	46 N	0.019	46 D	0.097	
	47 G	0.019	47 Q	0.097	
	48 I	0.019	48 I	0.097	
	49 T	0.019	49 N	0.097	
	50 N	0.019	50 G	0.097	
	51 K	0.019	51 K	0.097	
	52 V *	0.019	52 L *	0.097	
	53 N	0.019	53 N	0.097	
	54 S	0.019	54 R	0.097	
	55 V	0.019	55 V	0.097	
	56 I	0.011	56 I	0.094	
	57 E	0.011	57 E	0.094	
Lys58 ←	**58 K**	**0.011**	**58 K**	**0.094**	
	59 M *	0.011	59 T *	0.094	
	60 N	0.011	60 N	0.094	
	61 T	0.011	61 E	0.094	
	62 Q	0.009	62 K	0.094	
	63 F	0.005	63 F	0.046	
Thr64 ←	**64 T**	**0.027**	**64 H**	**0.046**	
Ala65 ←	65 A	0.101	65 Q	0.046	
	66 V	0.167	66 I	0.038	
	67 G	0.88	67 E	0.345	
	68 K	0.992	68 K	0.345	
	69 E	0.992	69 E	0.345	
	70 F	0.992	70 F	0.345	
	71 N	1	71 S	0.345	
	72 K	1	72 E	0.493	
	73 L	1	73 V	0.493	
	74 E	1	74 E	0.493	
	75 R	1	75 G	0.493	
	76 R	1	76 R	0.493	
	77 M	1	77 I	0.493	
	78 E	1	78 Q	0.493	
	79 N	1	79 D	0.493	
	80 L	1	80 L	0.493	
	81 N	1	81 E	0.493	
	82 K	1	82 K	0.493	
	83 K	1	83 Y	0.493	
	84 V	1	84 V	0.493	
	85 D	1	85 E	0.493	
	86 D	1	86 D	0.355	

Figure 5.8 Comparison of H1 and H3 sequences in the region of HA2 residues 38-86 in terms of their coiled coil forming potential. Sequences with coiled coil forming potential are shown in salmon, and sequences with no coiled coil forming potential are shown in blue. Polar, hydrophobic and charged residues are shown in red, green and grey, respectively. Key residues are indicated. Amino acid substitutions, which may account for differences in coiled coil propensity of these regions of H1 and H3 HA2, are marked (*). The N-termini of enzymatically prepared H3 and H1 TBHA2 (Ser40 and Ala65) are marked. Potential cleavage sites are indicated (Lys39, Lys58, Thr64).

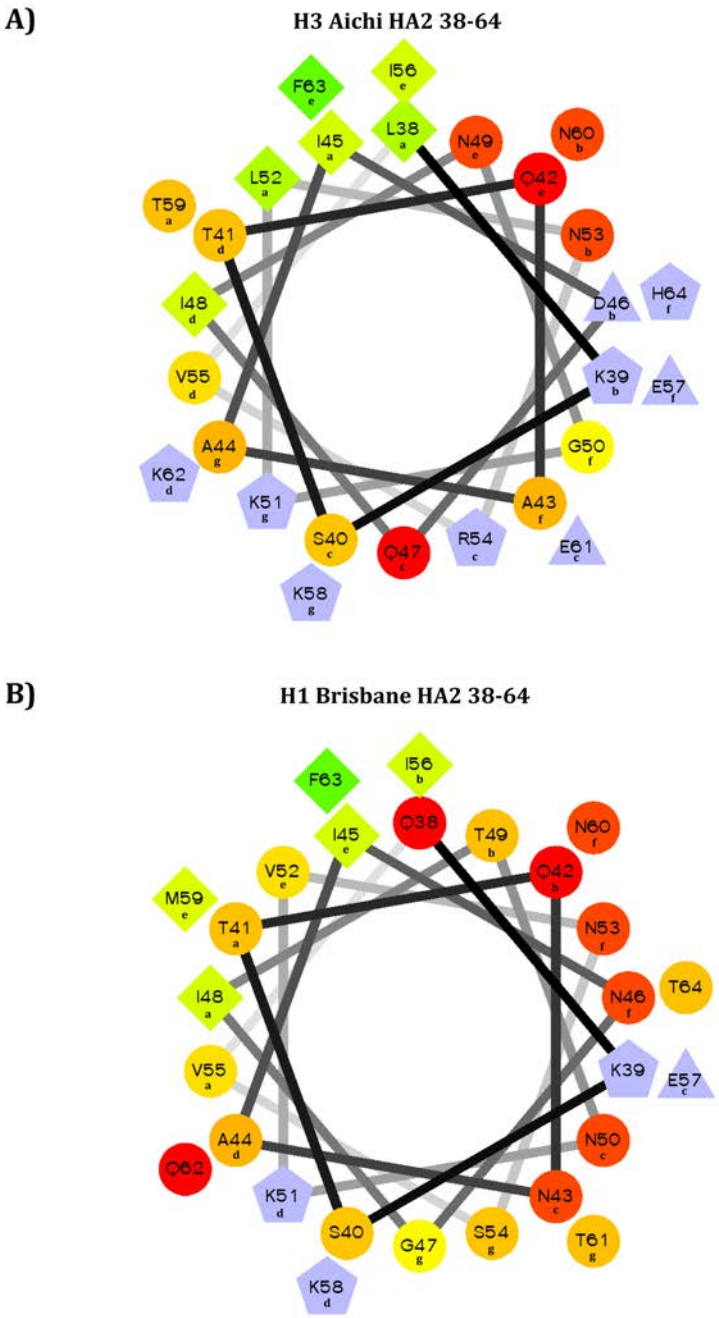


Figure 5.9 Helical wheel representation of H3 and H1 sequences in the region of HA2 residues 38-64. H3 HA2 is shown in A, and H1 HA2 is shown in B. The most hydrophobic residues are shown in green as diamonds, zero hydrophobicity in yellow. The most hydrophilic residues are shown in red. Potentially charged residues are shown in blue, with potentially negatively charged residues as triangles, and potentially positively charged as pentagons (rzlab.ucr.edu/scripts/wheel).

Comparison of ~10 kDa H3 and H1 HA2 proteolytic digestion products:

The structural features of B59 H1 TBHA2 appear to be different to those of X31 H3 TBHA2 when tested proteolytically. Incubation of X31 H3 TBHA2 and X31 H3 EBHA2 with trypsin and thermolysin yields different ~10 kDa fragments (Ruigrok, Aitken et al. 1988) (Chen, Wharton et al. 1995) shown in Figure 5.10. While incubation of X31 H3 TBHA2 with thermolysin, yields a fragment of HA2 containing residues 38-125, digestion of the same fragment with trypsin results in two fragments, one containing HA2 residues 40-124, and another containing HA2 residues 40-127. Incubation of X31 H3 EBHA2 with trypsin results in two fragments, identical to those obtained by digestion of X31 H3 TBHA2, and an additional intermediate containing HA2 residues 40-139, and as previously determined using N-terminal sequencing and mass spectrometry. As the resulting fragments are similar after digestion with different enzymes it was proposed, that cleavage at these positions is due to the exposure of these parts on the surface of H3 HA2. In all cases, the final structure contains the helical hairpin composed of a long α-helix (residues 40-105), followed by a turn at residues 106-112 and a short helix E (residues 113-128) (Chen, Wharton et al. 1995), and the resulting products are all shorter at their C-termini than a monomer of enzymatically-prepared B59 H1 TBHA2 described in this thesis. As described in Chapter 3, B59 H1 TBHA2 is a trimer of postfusion H1 HA2, containing short fragments of disulphide-linked HA1 domains (residues 11-17). The calculated molecular weight of HA1-less monomer is 10677.03 Da. Despite being exposed on the molecule, loops at HA2 residues 130-145 are unaffected by limited proteolysis, and shortening of the postfusion molecule results from a removal of amino acid residues from its N- and C-termini. Fragments resulting from proteolytic digestion of X31 H3 TBHA2, X31 H3 EBHA2 and B59 H1 TBHA2 are shown in Figure 5.10.

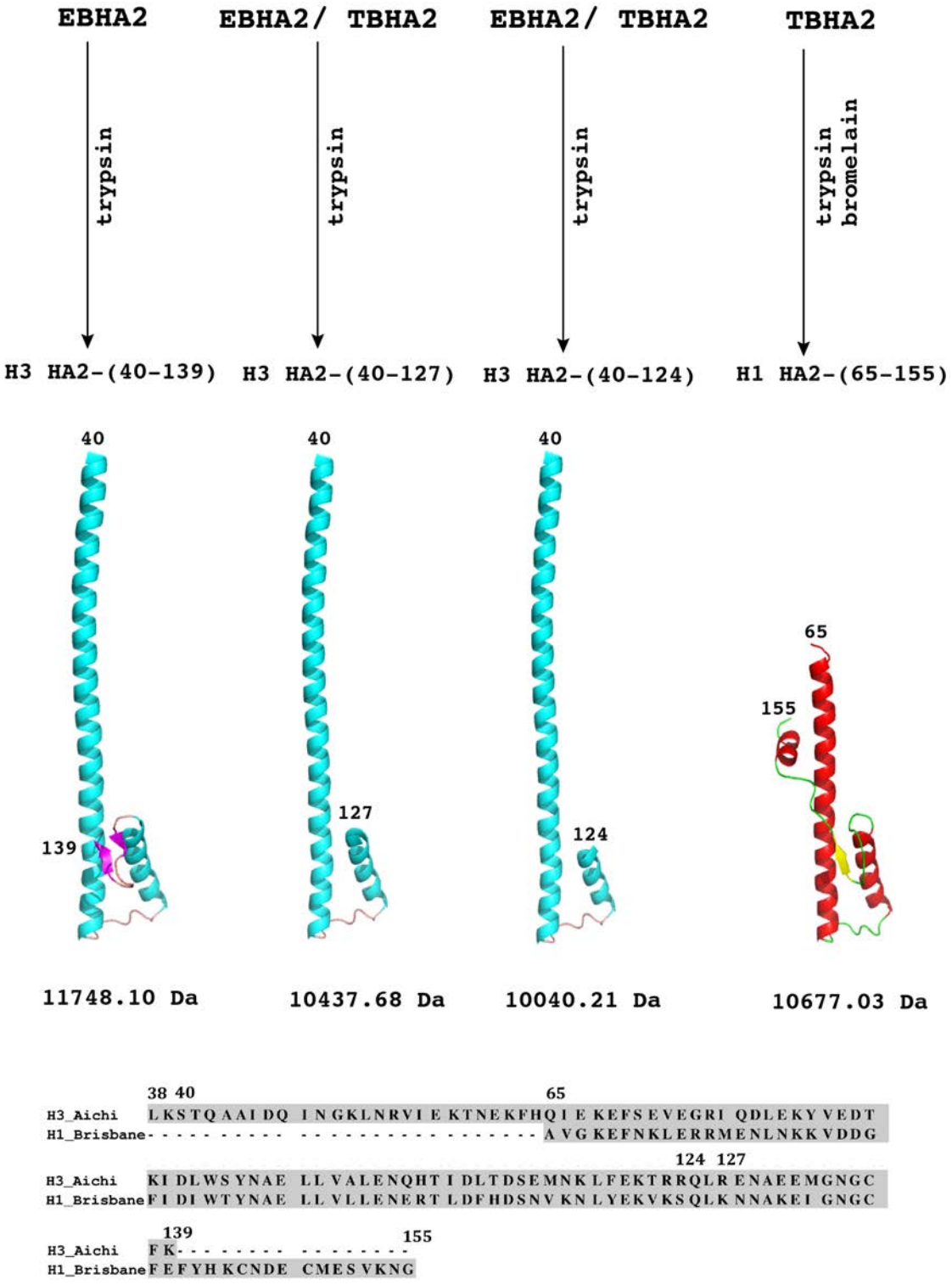


Figure 5.10 X31 H3 TBHA2, X31 H3 EBHA2 and B59 H1 TBHA2 proteolytic digestion products. Three postfusion H3 monomers are shown in blue, and a monomer of B59 H1 TBHA2 is shown in red. Molecular weight of each monomer was calculated from the primary sequence using Compute pI/Mw tool (ExPASy). The sequence alignment of H3 and H1 HA2 in the region of interest is shown (PRALINE; IBIVU server), and positions of key residues are indicated.

5.3 Recombinantly-expressed H1 and H3 HA2

The sequence identity between HA2 domains (residues 1-222) of B59 (UniProt: D5F1Q8) and X31 (UniProt: P03437) HA is 55% as calculated using PRALINE multiple sequence alignment (IBIVU Server) (Simossis and Heringa 2003) (Simossis and Heringa 2005). The crystal structures of Bac B59 H1 HA2 (Chapter 4) and X31 H3 EHA2 & EBHA2 (Chen, Skehel et al. 1999) were compared, by calculating rmsd for equivalent segments (A-G), and according to residue ranges given in Table 5.3. Structural alignment of the equivalent regions that undergo the low pH-induced refolding was carried out using chain B of X31 H3 EHA2 (PDB: 1QU1) (Chen, Skehel et al. 1999), and chain A of Bac B59 H1 HA2 (residues 45-164), present in the asymmetric unit of the R 3 2 unit cell . While the alignment of the entire monomers gives an overall rmsd of ~1.1 Å, superposition of the central helices (HA2 residues 42-105) reduces the rmsd value to ~0.37 Å. Distribution of structural deviation across the HA molecule was calculated as described in Chapter 5.2, and the obtained rmsd values are given in Table 5.6. Residues unobserved in the obtained structure of Bac B59 H1 HA2 (chains A and B) could not be compared to the corresponding fragment of X31 H3 EHA2, and residues excluded from structural comparison were HA2 residues 129-139, and the extended C-terminal fragment beyond Glu164 of the longer HA2 monomer (chain A). Distribution of structural deviation across the prefusion H3 HA2 (PDB: 1HGF) (Sauter, Hanson et al. 1992) and H1 HA2 (PDB: 1RUZ) (Sauter, Hanson et al. 1992, Gamblin, Haire et al. 2004), and across postfusion X31 H3 EHA2 (Chen, Skehel et al. 1999) and Bac B59 H1 HA2 structures was calculated, and the obtained rmsd values are shown in Table 5.6. With an rmsd of ~4 Å, interhelical loops B connecting prefusion helices A to central helices C, are completely different in the prefusion structures of H3 and H1 HA2, but adopt an almost identical conformation in the refolded molecules, and the structural deviation between them decreases to ~0.3 Å. While structural alignment of recombinantly-expressed H1 and H3 HA2 monomers gives an rmsd of ~0.8 Å, superposition of main helices (HA2 residues 45-105) reduces the rmsd value to 0.6 Å. The superposed monomers of X31 H3 EHA2 and Bac B59 H1 HA2 are shown in Figure 5.11. Trimers were generated by symmetry operations are these are shown in Figure 5.12. The two molecules are coloured by segments that undergo refolding at low pH.

Table 5.6 Structural comparison of prefusion H3 and H1 HA2 (PDB: 1HGF and 1RUZ), and postfusion X31 H3 EHA2 (PDB: 1QU1) and Bac B59 H1 HA2 fragments. Rmsd between the corresponding fragments A-G was calculated using PyMol.

	Prefusion	Postfusion
H3 to H1 HA2 segments (residue range)	**Rmsd (Å)**	**Rmsd (Å)**
A (45-55)	0.508	0.484
B (59-74)	**4.138**	0.291
C (76-105)	0.423	0.331
D (106-112)	0.299	0.704
E (113-127)	0.314	0.397
F (130-145)	0.88	missing in H1
G (146-153)	0.398	0.848
extended C-term (154-164)	0.993	1.542
central helices (45-105)		0.578
Overall rmsd (monomers)	1.005	0.796

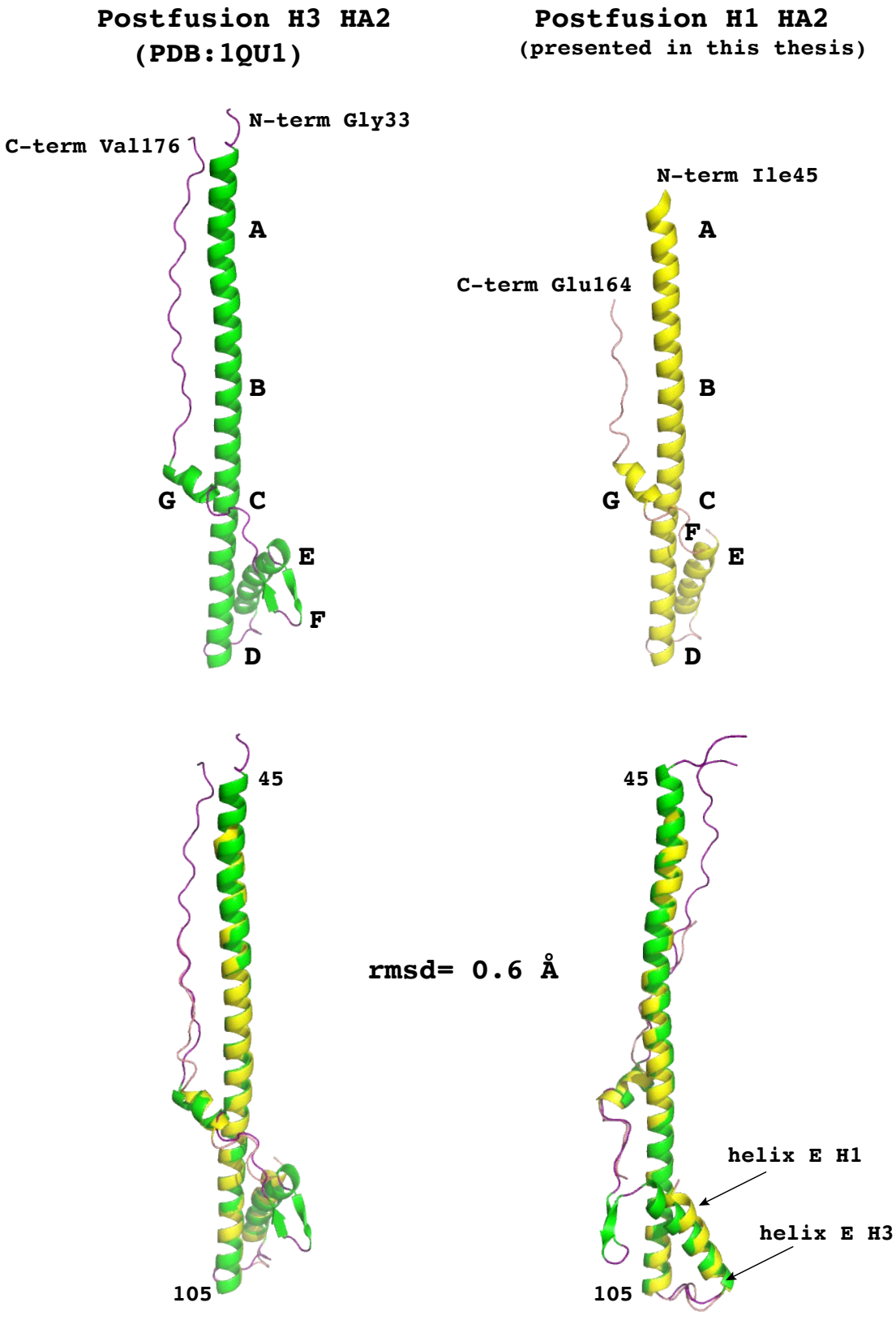


Figure 5.11 Recombinantly-expressed X31 H3 EHA2 and Bac B59 H1 HA2. The *E.coli*-expressed H3 EHA2 (PDB: 1QU1) (Chen, Skehel et al. 1999), is shown in green and purple. The Baculovirus/insect cell expressed B59 H1 HA2 is shown in yellow and salmon. Alignment of central helices (HA2 residues 45-105) gives an overall rmsd of ~0.6 Å. Position of helices E in both refolded HA2 is shown.

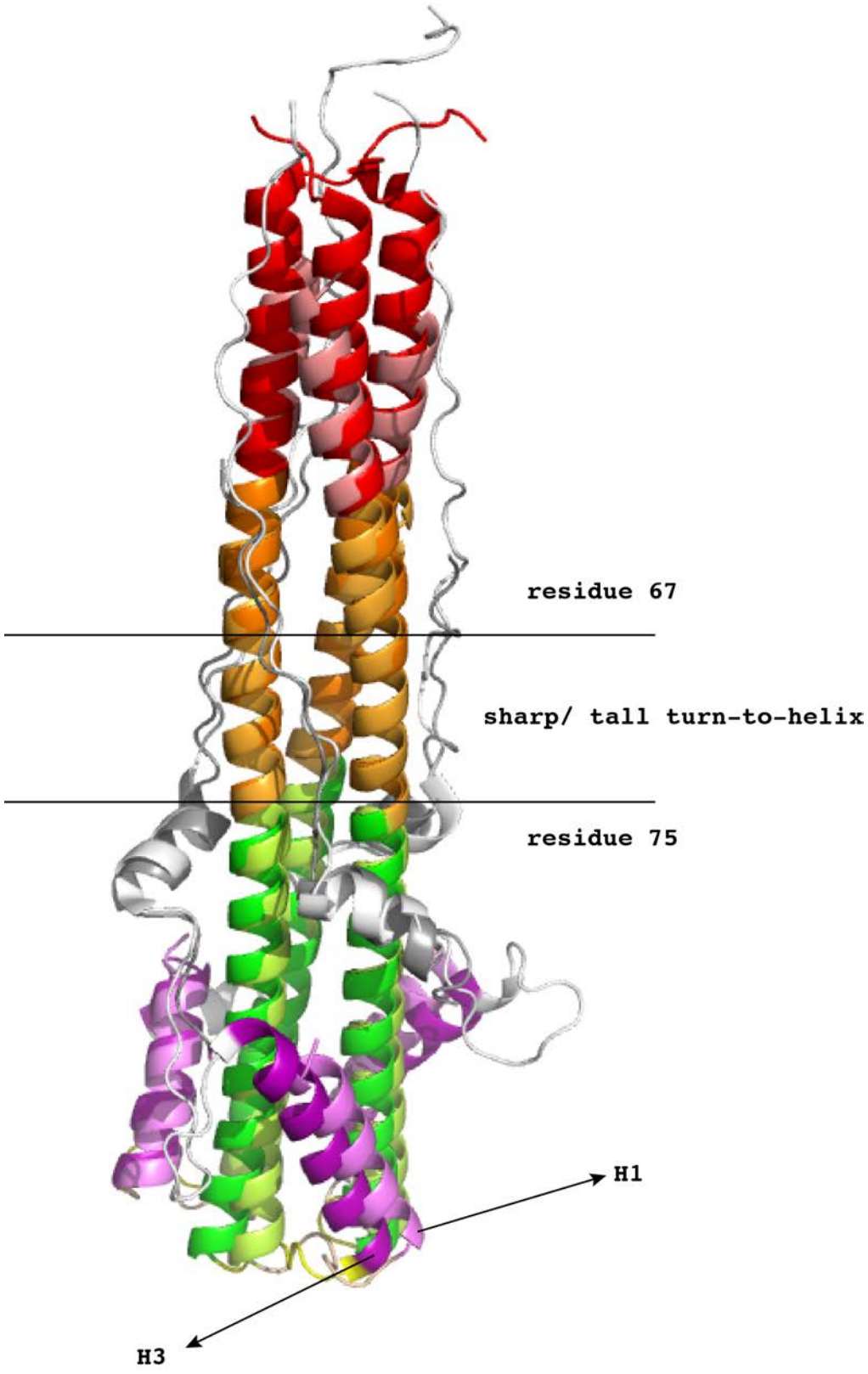


Figure 5.12 A symmetry generated trimer of Bac B59 H1 HA2 aligned to a trimer of X31 H3 EHA2. Structures are coloured by segments, which undergo refolding at low pH. Bac B59 H1 HA2 is shown in salmon, bright orange, lime, wheat and violet. X31 H3 EHA2 (PDB: 1QU1) is shown in red, orange, green, yellow and purple.

5.4 Recombinantly-expressed H1 HA2 and Influenza B HA2

The structure of *E.coli*-expressed influenza B HA2 (PDB: 4NKJ) (Ni, Chen et al. 2014) has been recently obtained using X-ray crystallography. The structure was solved in space group R 3 2. One monomer in the asymmetric unit contains HA2 residues 34-175. The sequence identity between HA2 fragments (residues 1-222) of B59 (UniProt: D5F1Q8) and Influenza B virus (B/Texas/3394/2013) (UniProt: S5DRP1) HA is 39%, as calculated using PRALINE multiple sequence alignment (IBIVU Server) (Simossis and Heringa 2003) (Simossis and Heringa 2005). The rmsd calculated by aligning the two monomers in PyMol is ~1.5 Å. Superposition was carried out using chain A of postfusion influenza B HA2 (PDB: 4NKJ) and chain A of Bac B59 H1 HA2 (Chapter 4). Alignment of central helices (residues 45-105) reduces the rmsd value to ~1.1 Å. As shown in Table 5.7, the central postfusion helix of influenza B HA2 (up to HA2 residue 105) is almost identical to that of Bac B59 H1 HA2. Fragments most similar between the two postfusion structures are helices B (HA2 residues 59-74) and helices E (HA2 residues 113-127), and the respective rmsd values are 0.27 Å and 0.29 Å. The superposed monomers labelled by segments (A-G) that undergo transition at low pH are shown in Figure 5.13. Differences in structures of turns D at HA2 residues 106-112 (rmsd ~2 Å) arise most likely from differences in their primary sequence. Structures of turns formed by a low pH-induced helix-to-turn transition are compared in Figure 5.14.

Table 5.7 Structural comparison of recombinantly-expressed influenza B (PDB: 4NKJ) and Bac B59 H1 HA2. Rmsd between the corresponding fragments A-G was calculated using PyMol.

H1 to Flu B postfusion HA2 segments (residue range)	Rmsd (Å)
A (45-55)	0.409
B (59-74)	0.270
C (76-105)	0.468
D (106-112)	2.049
E (113-127)	0.293
F (130-145)	missing in H1
G (146-153)	0.852
extended C-term (154-164)	2.075
central helices (45-105)	1.098
Overall rmsd (monomers)	1.475

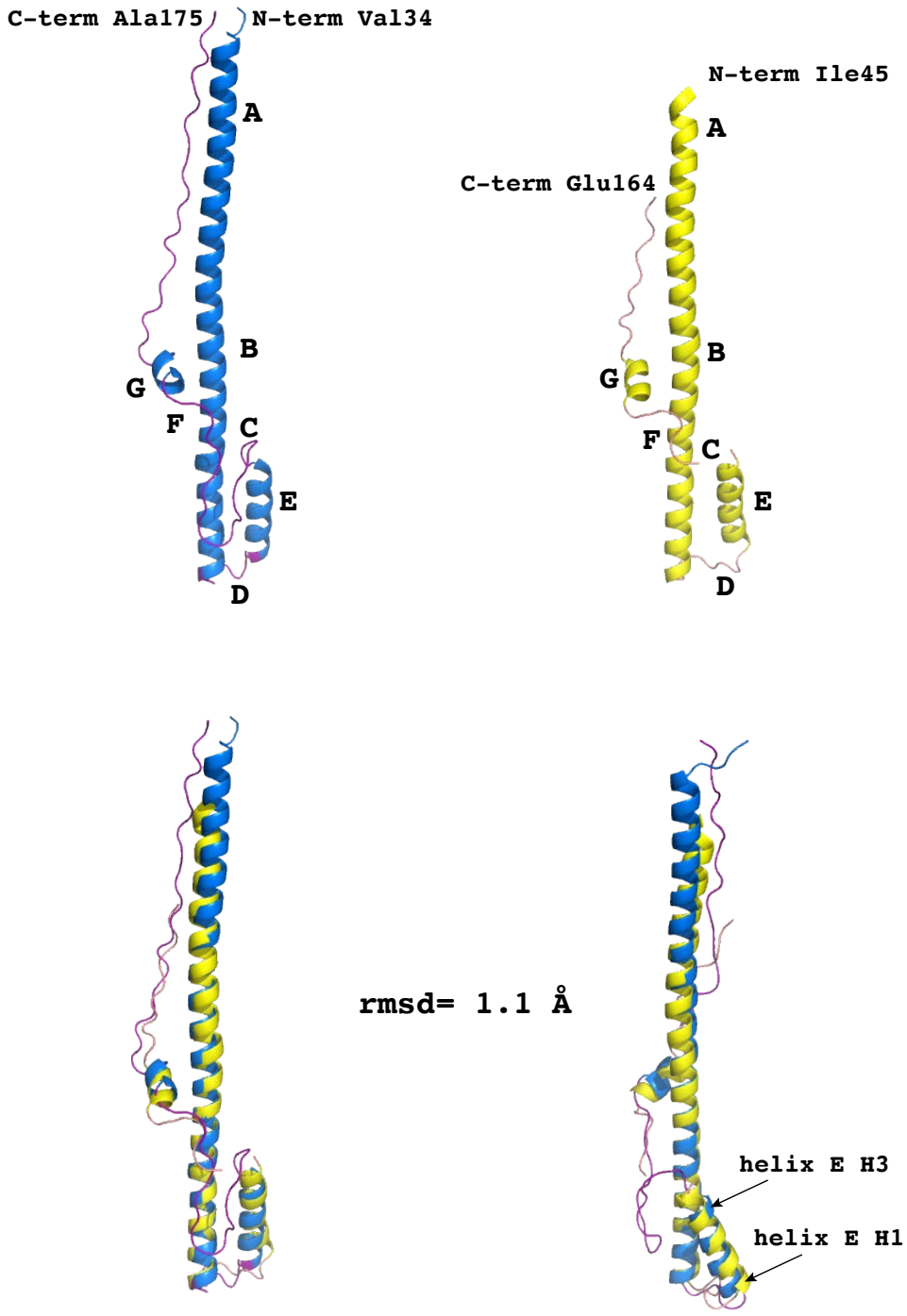


Figure 5.13 Comparison of recombinantly-expressed influenza B HA2 and Bac B59 H1 HA2. Recombinantly-expressed influenza B HA2 (PDB: 4NKJ) is shown in blue, and Bac B59 H1 HA2 is shown in yellow. Alignment of central helices (HA2 residues 42-105) gives an overall rmsd of 1.1 Å.

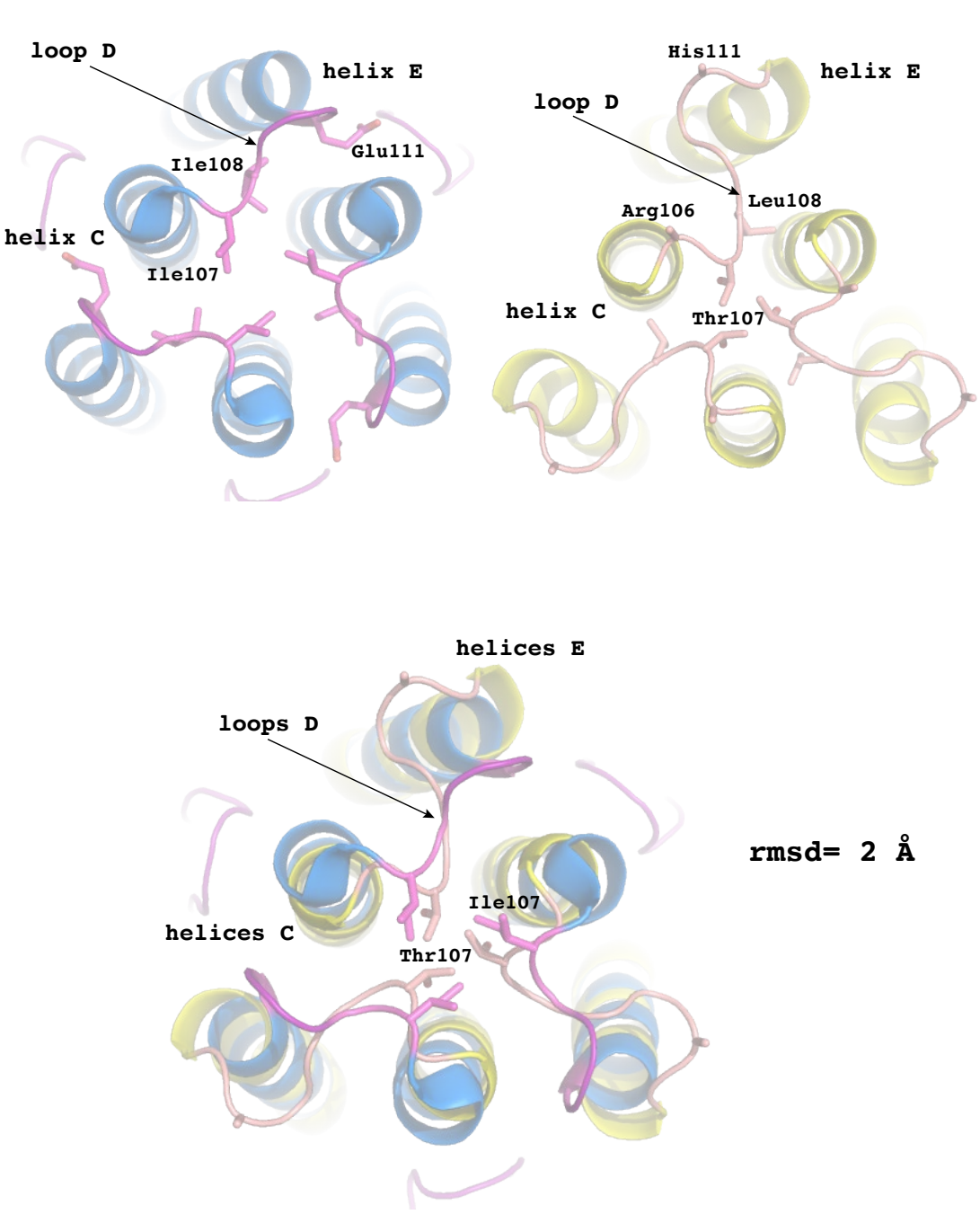


Figure 5.14 Comparison of turns at residues 106-112 of recombinantly-expressed influenza B HA2 and Bac B59 H1 HA2. View down the 3-fold symmetry axis of postfusion influenza B HA2 trimer (PDB: 4NKJ) and a symmetry generated Bac B59 H1 HA2 trimer. Key residues are indicated. Alignment of regions gives an rmsd of ~2 Å. Threonine 107 packs closely towards the 3-fold axis of the postfusion H1 HA2 structure. Residues 107 and 108 of influenza B HA2 are Isoleucine.

6 Characterisation of FI6 antibody binding to the two groups HA

The cross-reactive FI6 antibody has been previously shown to bind to a conserved region in the fusion (F) subdomain of HA2 present in both HA groups (Corti, Voss et al. 2011). In this thesis, the antibody was tested for its ability to prevent the low pH-induced conformational change of HA, and therefore viral membrane fusion. The specifications of the FI6 with relation to HA binding, and group-specific structural features of influenza virus HA, which may play a role in the efficiency of the FI6-mediated neutralization, are discussed in Chapters 1.8.2 and 1.7.

6.1 Binding of the FI6 antibody to virus-bound HA

To investigate the ability of the cross-reactive FI6 antibody (Corti, Voss et al. 2011) to prevent the conformational change of virus-bound HA, two viruses with HA belonging to two distinct HA groups were chosen. Strain B59 was chosen to represent group 1 HA, and strain X31 was chosen to represent group 2 HA. In order to assess the ability of FI6 to prevent the conformational change of the virus-bound molecule, virus-antibody complexes were prepared, treated with low pH to induce the conformational change of virus-bound HA (Section 2.3.6) and digested with trypsin, to confirm the conformation of the low pH-treated molecule. The obtained samples were analysed by SDS-PAGE (Section 2.3.3) and electron microscopy (Section 2.4.5).

Preparation of H1N1- and H3N2-FI6 antibody complexes:

The X31 and B59 viruses were grown in embryonated chicken eggs and purified as described in 2.2.1.1 and 2.2.1.2. The purified viruses were incubated with an excess of the FI6 antibody O/N. Viruses were pelleted to remove the antibody that has not bound, and resuspended in 25 mM Tris pH 8.0, 150 mM NaCl.

6.1.1 Assessment of the ability of the FI6 antibody to prevent the conformational change of virus-bound HA by limited proteolysis

Virus-antibody complexes were pre-incubated at 37°C, for 2 min. The pH of virus suspensions was lowered to pH 4.8, using 0.1 M citric acid (37°C, 5 min) in order to induce the conformational change of virus-bound HA (2.3.6). The pH was then readjusted to pH 7.8, using 1 M Tris pH 8.0, and virus-antibody complexes were digested with trypsin (50:1 w/w, complex: trypsin), for 45 min, at RT. Digestions were stopped by the addition of an equal amount of soybean trypsin inhibitor (Sigma), and samples were analysed by SDS-PAGE (Section 2.3.3). The SDS-PAGE analysis of tryptic digestion products is shown in Figure 6.1. The analysed samples included virus-antibody complexes, complexes treated with low pH, and their trypsin digests. The antibody-less virus samples were used as a control. As shown in Figure 6.1, A, lane 3, almost all of the antibody-less H3 HA (X31) was digested with trypsin after the low pH treatment. In case of H3 with the FI6 antibody bound, a proportion of HA stays intact following the low pH treatment and trypsin digestion (Figure 6.1, A, lane 6). As shown in Figure 6.1, B, while all of antibody-less H1 HA was digested with trypsin after low pH treatment (lane 3), H1 HA complexed with the FI6 antibody stays intact after exposure to trypsin. The obtained results confirm that binding of the FI6 to the fusion subdomain of H3 HA2 is not as tight as its binding to the corresponding region of H1 HA2.

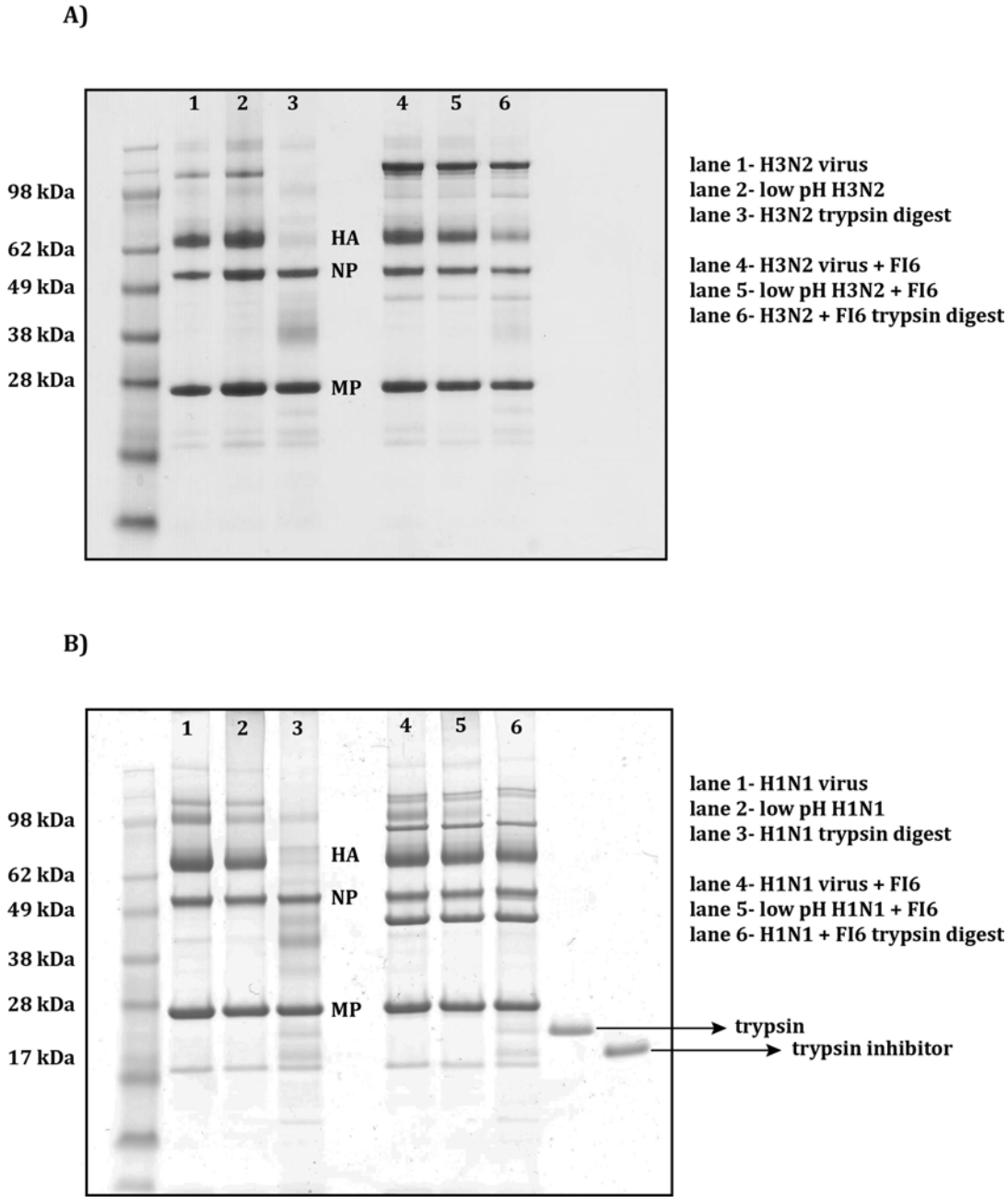


Figure 6.1 Assessment of the ability of the FI6 antibody to prevent a conformational change of virus-bound HA by limited proteolysis. Binding of the FI6 antibody (Corti, Voss et al. 2011) to X31 and B59 viruses is shown in A and B, respectively. Bands corresponding to influenza nucleoprotein (NP) and influenza matrix protein (MP) are labelled. Details of SDS-PAGE are included in Section 2.3.3.

6.1.2 Binding of the FI6 antibody to virus-bound HA by electron microscopy

The X31 and B59 viruses were complexed with the FI6 antibody, treated with low pH and trypsin as described in Section 6.1.1, and analysed using electron microscopy. The FI6 antibody-less viruses were imaged at each step as a control. Experimental details related to electron microscopy are included in Section 2.4.5. Figure 6.2 shows electron micrographs of the FI6-bound H3 HA on the surface of X31 virus. Images of B59 virus complexed with FI6 could not be obtained at satisfactory resolution, due to the observed instability of virus-bound H1 HA. As observed, virus particles were separate at pH 7.0, with straight and even HA spikes. Following the addition of the FI6 antibody, viral particles appeared more crowded, and observation of individual HA spikes became difficult. The excess of used FI6 antibody is visible in the background. As opposed to antibodies that bind to the globular HA1 domains, causing cross-linking between virus particles, binding of the FI6 antibody to the stem region of HA, does not cause aggregation of virus particles. Following incubation at pH 5.0, the antibody-less X31 particles aggregated, which is consistent with HA on the viral particles undergoing the low pH-induced conformational change. Particle aggregation was not observed for X31-FI6 complexes at pH 5.0, and in this case HA spikes appeared to stay intact, suggesting that the conformational change of X31 HA was prevented by the FI6 antibody binding. Subsequent digestion with trypsin aimed to remove the dissociated HA1 domains from the low pH-treated sample, and these were successfully removed from the antibody-less X31 HA, which reduced particle aggregation. Trypsin digestion of the X31-FI6 complexes did not remove the HA1 domains, and HA spikes stayed intact. Some short, thin HA spikes were observed, which indicated, that a fraction of X31 HA on the virus surface had undergone a conformational change in the presence of FI6. The obtained results suggested that sufficient amount of the FI6 antibody binds to H3 HA on the surface of the X31 virus to prevent a proportion of HA from undergoing the low pH-induced conformational rearrangement.

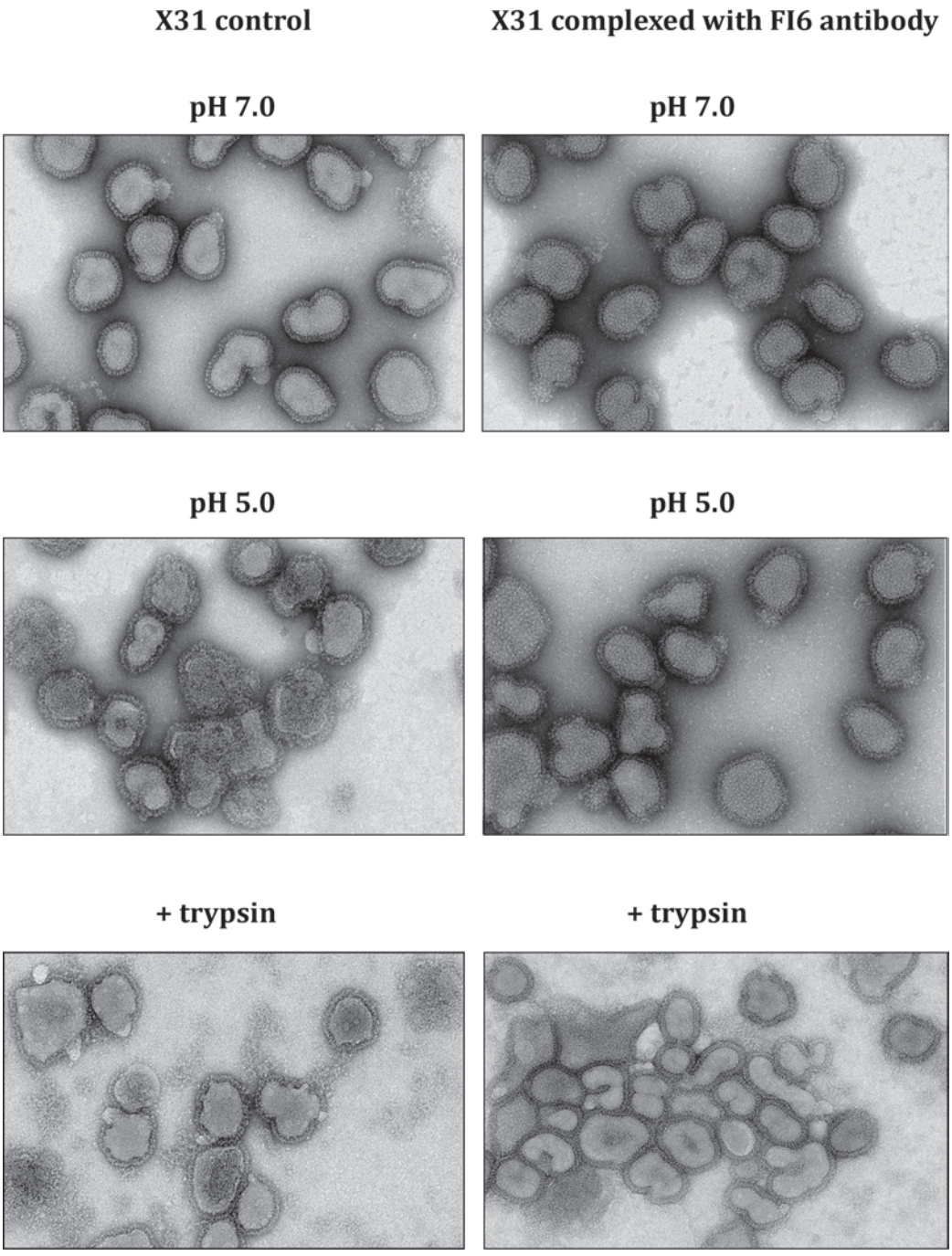


Figure 6.2 Electron microscopy of the FI6 antibody binding to the X31 virus. Images acquired by Dr Lesley Calder.

6.2 Binding of the FI6 antibody to THA by electron microscopy

The number of HA spikes on each viral particle is ~400, and each HA trimer has three FI6 binding sites (Corti, Voss et al. 2011). In order to determine, how many FI6 molecules need to bind to an individual HA trimer in order to block the conformational change of HA, binding of FI6 was studied using a detergent-extracted THA, released from X31 and B59 viruses using βOG or LDAO. The TM anchor-less THA was prepared as described in 2.2.1.4, and purified by gel filtration (2.3.5). A chromatogram from a final THA purification is shown in Figure 4.2 (Chapter 4). The THA-FI6 antibody complexes were prepared as described in 6.1.1. Following incubation at pH 5.0, THA molecules would normally aggregate into rosettes of 6-8 molecules, due to exposure of the hydrophobic fusion peptides at the pH of membrane fusion (Skehel, Bayley et al. 1982) (Doms, Helenius et al. 1985) (Ruigrok, Aitken et al. 1988). This property of THA was exploited to examine the ability of FI6 to prevent the conformational change of the detergent-released X31 H3 THA and B59 H1 THA. As expected, THA rosettes were unobserved at pH 7.0 using electron microscopy, and majority of the observed molecules were THA-FI6 complexes. The antibody appeared to bind to two binding sites on a THA trimer. Some THA trimers had only one binding site occupied by FI6, and very few THA molecules had all three sites bound to the antibody. Some single THA molecules were also visible. After incubation of THA-FI6 complexes at pH 5.0, mainly individual complexes were observed (Figure 6.3, B). As molecules lacking the FI6 antibody formed rosettes at pH 5.0, it can be concluded, that the FI6 antibody prevents the formation of rosettes at the pH of membrane fusion, by blocking the conformational change of HA. Individual THA-FI6 complexes of different morphology, resulting from the flexibility of variable arms of the FI6 are magnified, and presented schematically in Figure 6.3, and the observed complexes adopt similar structures at pH 7.0 and pH 5.0.

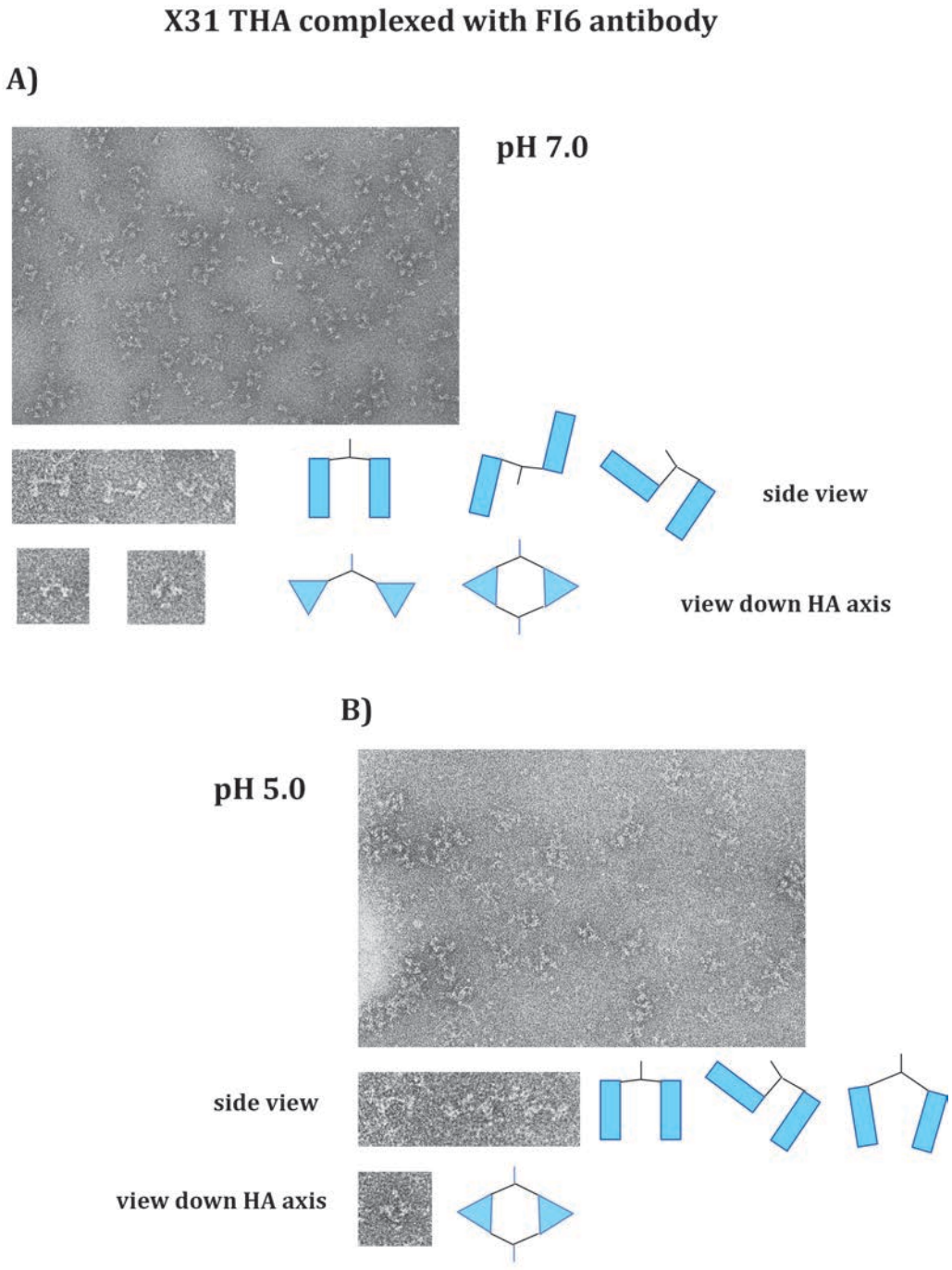


Figure 6.3 Electron microscopy of X31 H3 THA-FI6 antibody complexes. Images at neutral pH (A) and at the pH of membrane fusion (B) are shown. Images were acquired by Dr Lesley Calder.

6.3 Binding of viruses to human and avian receptor analogs in the presence of FI6 by biolayer interferometry

Binding of X31 and B59 viruses complexed with the FI6 antibody (Corti, Voss et al. 2011) to human and avian receptor analogs was tested using biolayer interferometry. The principles of biolayer interferometry and experimental details are described in 2.4.6. The technique was used to estimate the concentration range of the FI6 antibody to be used in the MDCK1 cell infection assay described in (6.4) and also to verify, whether high concentrations of FI6 interfere with binding of viral particles to the sialic acid receptors on the surface of infected cells. The results indicate, that binding of FI6 antibody to X31 virus does not stop binding of X31 to sialic acid receptor analogs (Figure 6.4, B, yellow). The Hc31 strain-specific antibody is expected to completely block binding of X31 virus to sialic acid, and the observed weak binding (Figure 6.4, B, blue), at low concentration of Hc31, may result from overestimation of the initial antibody concentration. In case of B59 virus (group 1 HA) neither of the antibodies prevents the virus from binding to sialic acid receptor analogs, as shown in Figure 6.4, A, and the results confirm that even at very high concentrations, the FI6 antibody does not sterically block binding of viruses to sialic acid-containing receptors.

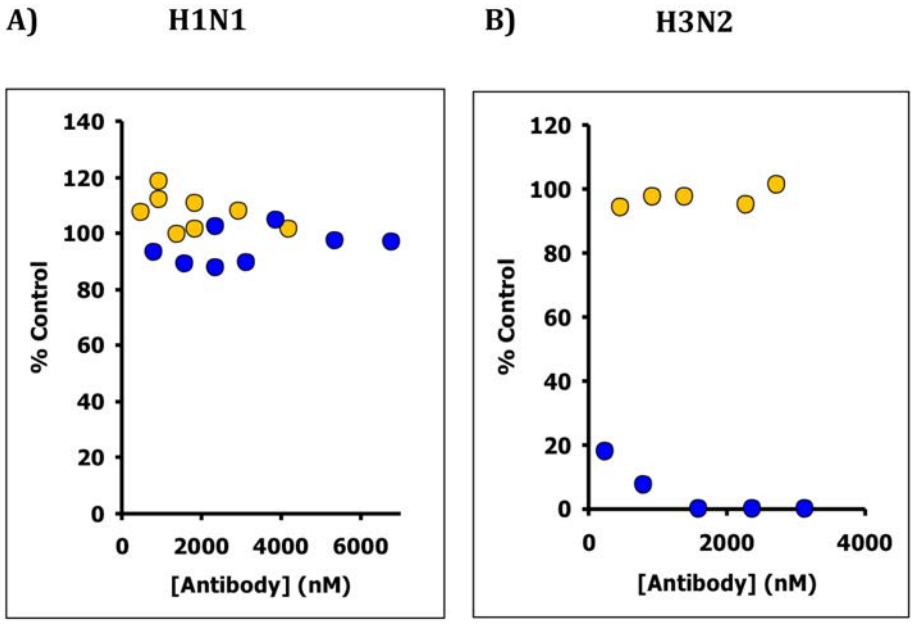


Figure 6.4 Binding of B59 H1N1 and X31 H3N2 virus-antibody complexes to sialic acid receptor analogs by biolayer interferometry. Shown is a response on the sensor (% of the control). Virus-FI6 complexes are shown in yellow, and virus-Hc31 complexes are shown in blue.

6.4 Determination of infectivity of H1N1 and H3N2 viruses neutralized by the FI6 antibody

Infectivity of H3N2 and H1N1 viruses neutralized by the broadly neutralizing FI6 antibody was measured using the MDCK1 cell infection assay, described in 2.4.7. Experiments were performed using X31 and B59 viruses. The ability of the antibody-bound viruses to enter the infected cells was tested using biolayer interferometry (6.3) prior to cell infection. The HA titre in the collected allantoic fluid was determined using HA assay (2.4.1). The HA titre in the allantoic fluid containing B59 virus was estimated to ~512 HAU for 10^{-2} – 10^{-4} inoculum dilutions, and 1024 HAU for 10^{-5} dilution. The HA titre in the allantoic fluid containing X31 virus was lower (256 HAU for 10^{-2}- 10^{-3}, and 512 HAU for 10^{-4} – 10^{-5} initial allantoic fluid dilutions). Samples with the highest HA titre were used for cell infection experiments. Preparation of allantoic fluid and the FI6 antibody dilutions, as well as preparation of the Hc31 antibody dilutions (used as a control), are described in 2.4.7. Cell infection assay (2.4.7) was carried out over 6 hours to ensure only one cycle of viral infection. Results indicate that the Hc31 antibody blocks infection of MDCK1 cells with the X31 virus, and as expected, infectivity of B59 virus remained unaffected (Figure 6.5). This result is consistent with the specificity of Hc31, which specifically binds to the globular head domains of H3 HA, blocking receptor-binding (Daniels, Douglas et al. 1983). The FI6 antibody was found to reduce viral infection with both viruses. At high concentrations, the FI6 appeared to completely block infection of MDCK1 cells with the B59 virus, but was found to be less effective in neutralisation of the infectivity of the X31 virus (Figure 6.6).

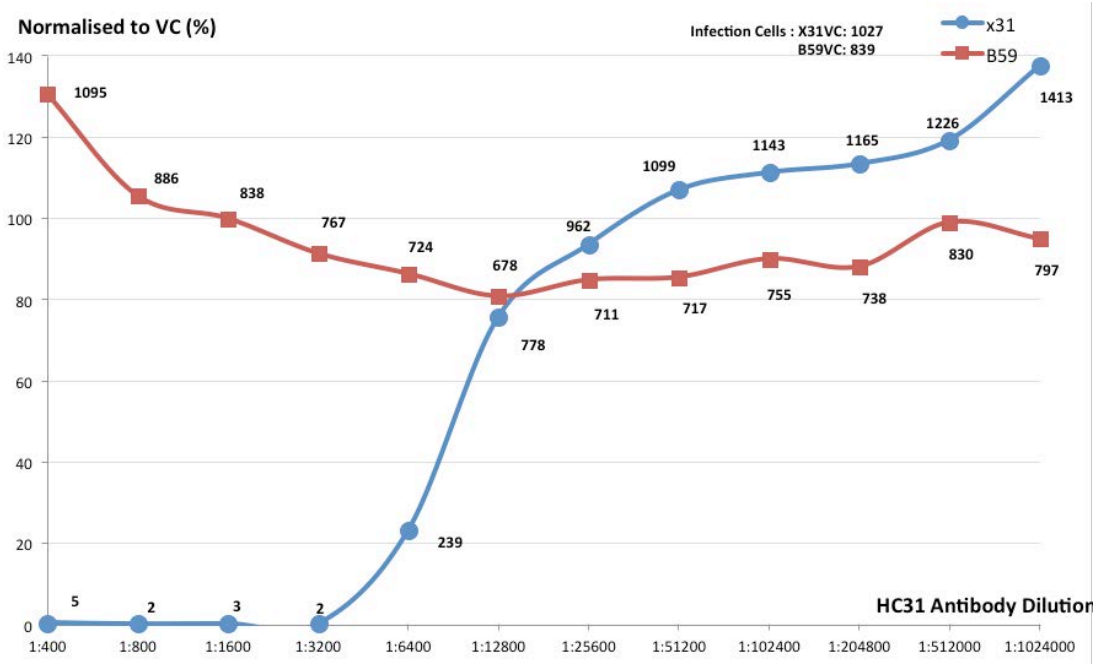


Figure 6.5 Quantification of viral infection in the presence of the Hc31 antibody. Virus strains include B59 (H1N1) shown in red, and X31 (H3N2) shown in blue. Labels on the curve indicate the average number of positive cells at corresponding Hc31 dilution.

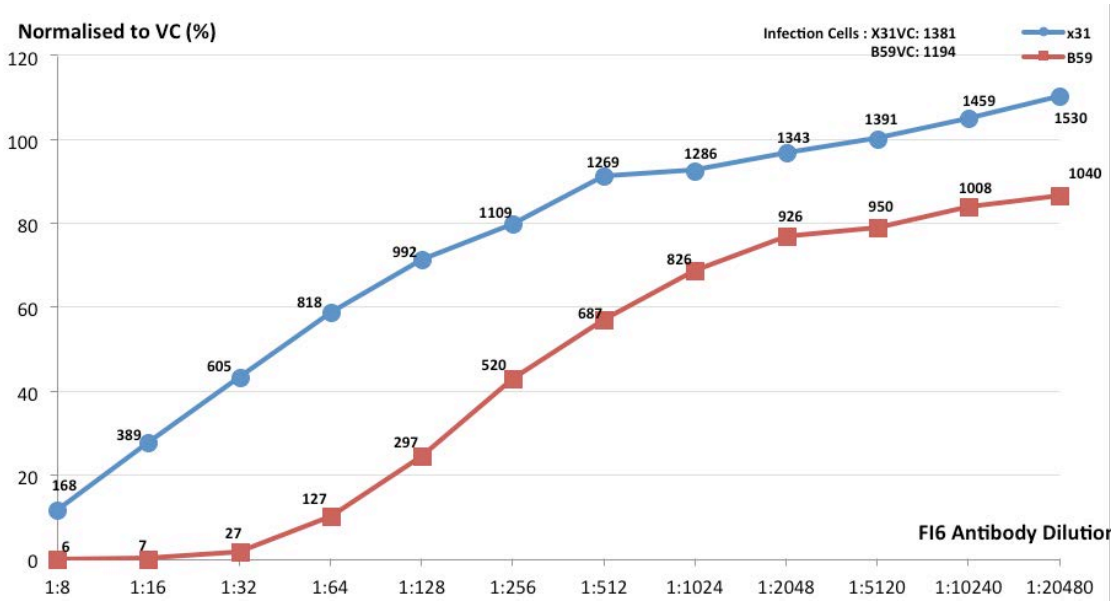


Figure 6.6 Quantification of viral infection in the presence of the FI6 antibody. Virus strains include B59 (H1N1) shown in red, and X31 (H3N2) shown in blue. Labels on the curve indicate the average number of positive cells at corresponding FI6 dilution.

7 Discussion

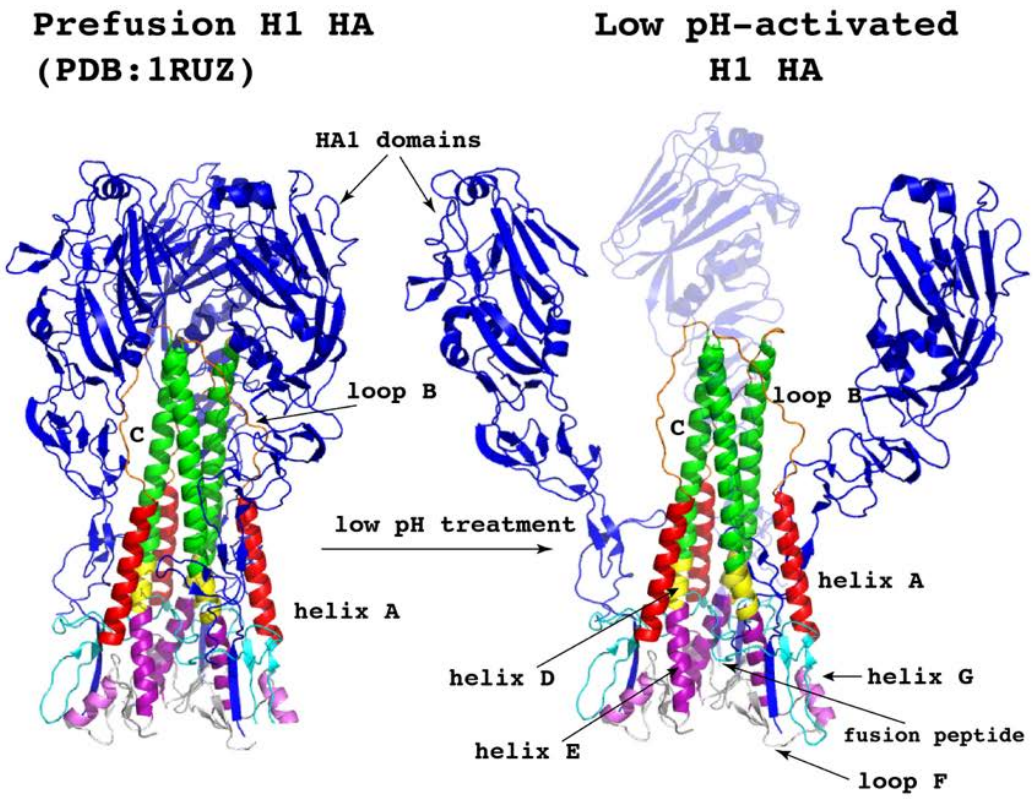
7.1 Purpose of study

Influenza HA is a molecule with three major functions. These are receptor binding, viral membrane fusion and antigenic variation. In this thesis, the molecule was studied with respect to its role in viral membrane fusion, which leads to the release of viral genome into the host cytoplasm, allowing for its replication. To perform this critical step of virus cycle, the molecule must undergo a significant rearrangement, which eventually brings the viral and endosomal membranes into proximity. In contrast to the abundance of prefusion influenza A HA structures obtained to date using X-ray crystallography, the number of available postfusion structures is limited. Two crystal structures of postfusion H3 HA2 (group 2 HA) were previously obtained, and these are enzymatically-prepared, viral H3 TBHA2 (Bullough, Hughson et al. 1994) and *E.coli*-expressed H3 EHA2 & EBHA2 (Chen, Skehel et al. 1999) fragments of the ectodomain.

The purpose of this study was to obtain a structure of postfusion HA2 from group 1 HA using X-ray crystallography, and its comparison to the previously reported structures of postfusion H3 HA2 (Bullough, Hughson et al. 1994) (Chen, Skehel et al. 1999), and to the recently reported structure of influenza B HA2 in a postfusion conformation (Ni, Chen et al. 2014). The aim of this thesis was to use structural comparison of these molecules, in order to investigate, whether group-specific structural features of the prefusion molecules (discussed in 1.7), result in differences in their postfusion structures. Subtype H1 was chosen to represent group 1 HA, and HA of PR8 and B59 H1N1 viruses was used. Subtype H1 is responsible for the 1918 pandemic (Gamblin, Haire et al. 2004), and given its place in the same HA group as the highly pathogenic H5 subtype, structural information related to the low pH-induced structural rearrangement of H1 HA2, could be related to other HA subtypes from group 1 HA.

7.2 Crystal structure of B59 H1 TBHA2

Postfusion H1 TBHA2 from the PR8 and B59 H1N1 viruses were prepared proteolytically, and subjected to crystallisation trials as described in Chapter 3. Crystals of best morphology and size were obtained using refolded HA2 released from PR8 H1N1 virus with bromelain. Protein crystallized from solutions containing various PEGS as precipitants. The pH of buffers ranged from 5.5-7.5. Crystals diffracted weakly, despite extensive optimisation trials, testing of different cryoprotectants and dehydration experiments. Diffraction quality crystals of B59 H1 TBHA2 were obtained from 15% PEG 3350, 0.1M MES pH 6.0, with the addition of 1.5% MPD, 0.01M Hepes pH 7.5 and 0.02M Sodium citrate. Crystals diffracted to ~3.4 Å, and the structure was solved by molecular replacement in space group P 3 2 1. There was one monomer of B59 H1 TBHA2 in the asymmetric unit. The triple-stranded, α-helical coiled coil comprises HA2 residues 65-105. This trimeric structure is followed by a connecting loop (D), formed by a helix-to-turn transition of HA2 residues 106-112, upon exposure of the molecule to acidic pH. (Figure 4.3, B). Loop D reverses residues beyond Asp112 by 180°, and results in an antiparallel packing of a short helix E (Ser113-Lys127). Residues 130-145 of B59 H1 TBHA2 are arranged into a loop (F), and majority of residues within this loop (Lys131-Phe140) form a ß-hairpin structure. Loop F is connected to a short fragment of HA1 (Asp11-Tyr17) via a disulphide bond between Cys137 of HA2 and Cys14 of HA1. Loop F is followed by a short helix G comprising HA2 residues 146-155, and the C-terminal residue is Gly155. Preparation of B59 H1 TBHA2 for crystallographic studies and the obtained crystal structure are shown in Figure 7.1. Regions of the molecule that undergo the low pH-induced structural refolding are labelled.



Postfusion H1 TBHA2

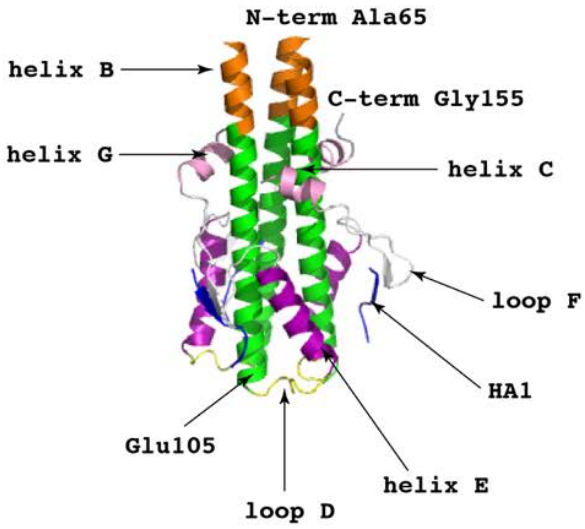


Figure 7.1 Preparation of B59 H1 TBHA2 for crystallographic studies. Shown are: a prefusion trimer of H1 HA (PDB: 1RUZ) (Gamblin, Haire et al. 2004), dissociation of HA1 domains in response to low pH, and the obtained structure of B59 H1 TBHA2. Structures are coloured by segments, which undergo refolding at low pH.

7.3 Analysis of H1 TBHA2 fragment

The N-terminal sequencing of enzymatically-prepared fusion peptide-less PR8 H1 TBHA2 confirmed its N-terminus to be Ser40 of HA2. The purified protein runs as a ~17 kDa band under reducing conditions (Figures 3.7 and 3.8), corresponding to the size of HA1-less HA2 monomer. Resuspended crystals of PR8 H1 TBHA2 also ran as a ~17 kDa band under reducing conditions, indicating presence of the prepared fragment. The B59 H1 TBHA2 prepared using the same method was not sequenced, but also ran as a ~17 kDa band, indicating the N-terminus of the purified protein to be Ser40, as in case of PR8 H1 TBHA2. However, only residues 65-155 of HA2, disulphide-linked to a short fragment of HA1 (residues 11-17), can be seen in the crystal structure of B59 H1 TBHA2. The molecular weight of B59 H1 TBHA2 monomer calculated from the built structure is ~10677 Da per monomer. Missing electron density for ~25 residues at the N- and C-termini of the prepared B59 H1 HA2 would have accounted for the rest of the initial molecular weight of the prepared protein (~17 kDa/monomer) (Figures 3.7 and 3.8) indicating, that the prepared protein was cleaved post-purification.

Overall, the HA of B59 H1N1 virus was found to be very unstable and highly susceptible to proteolytic cleavage. Conservation of proteolytic cleavage sites between HA2 segments of B59 and PR8 influenza virus HA suggests some additional instability in B59 HA2 in the region comprising residues 40-64. Cleavage of this region and the ~25 residues at the C-terminus, would result in a fragment of ~10 kDa, which correlates with the molecular weight of a degradation product observed by SDS-PAGE under reducing conditions (Figure 3.7, lanes 2 and 4), and with the molecular weight of a monomer calculated from the obtained structure. Despite cleavage of the postfusion molecule at both N- and C-termini, the six-helical bundle and the connecting loops F characteristic of the postfusion molecule are present in the obtained crystal structure of B59 H1 TBHA2, and this part of the refolded viral molecule can be considered very stable.

7.4 Crystal structure of Bac B59 H1 HA2

Due to the difficulties encountered in preparation of postfusion viral H1 TBHA2 described in Chapter 3, residues 38-175 of B59 H1 HA2 were expressed using Baculovirus/insect cell system (2.2.2). Protein was purified using IMAC (2.3.8) and gel filtration (2.3.5), and its conformation was assessed in two ways. Firstly, using limited proteolysis (2.3.7 and 4.2), and digestion of the purified protein with trypsin resulted in a ~6 kDa fragment (Figure 4.3). The susceptibility of the expressed fragment to trypsin digestion indicated exposure of previously buried trypsin cleavage sites, associated with the low pH-induced refolding of the HA2 molecule. Secondly, the secondary structure content was determined using CD spectroscopy (2.4.3 and 4.2), and far UV spectra recorded at λ=195-260 nm indicated a high α-helical content of Bac B59 H1 HA2. The predicted α-helical content of the expressed protein containing HA2 residues 38-175 was ~ 30%, and is characteristic of the postfusion H3 HA2 (Ruigrok, Aitken et al. 1988) (Chen, Wharton et al. 1995). Estimation of the secondary structure content of Bac B59 H1 HA2 at neutral and acidic pH using CD is shown in Tables 4.1 and 4.2. Another feature of postfusion HA2 is its high stability with respect to thermal denaturation (Ruigrok, Aitken et al. 1988) (Chen, Wharton et al. 1995). CD was therefore used to monitor the thermal stability of the recombinant protein. The melting temperature (T_m) of recombinantly-expressed Bac B59 H1 HA2 was shown to be pH-dependent, and the T_m was 72°C for pH 5.2, and 67°C for pH 7.2 (Table 4.3). Use of the Baculovirus/insect cell expression system resulted in high levels of recombinant protein expression, and spontaneous folding of recombinant H1 HA2 into a low pH-induced conformation makes it a preferred eukaryotic system for the expression of postfusion HA2 for structural analysis.

Diffraction quality crystals of Bac B59 H1 HA2 were grown from 18% PEG 3350, 0.1 M Sodium citrate pH 5.6 and 2% tacsimate [v/v]. Crystals diffracted to ~3.4 Å and the structure was solved by molecular replacement (4.5) using a specially prepared model in space group R 3 2. The obtained crystal structure of Bac B59 H1 HA2 is described in Chapter 4. There are two monomers of postfusion H1 HA2 in the asymmetric unit of the R 3 2 crystal. The two monomers vary in length. Chain designated A, contains HA2 residues 45-164, and chain B contains HA2 residues 75-128. Chain B is even shorter than that observed in the crystal of the enzymatically prepared B59 H1 TBHA2 described in Chapter 3. Molecular weight calculated from the primary sequence according to the built model is 6511.37 Da,

and corresponds to the size of a degradation product (~6 kDa) observed using SDS-PAGE, and shown in Figure 4.3. Interestingly, ~6 kDa fragments were also obtained by treating viral H1 HA2 with various proteases (Figure 3.9). The observed six-helix-bundle appears to be stabilized by a hydrophobic interaction between Val122 belonging to helix E and Trp92 belonging to the central stem, and is shown in Figure 4.13.

7.5 Viral and recombinantly-expressed H1 HA2

The two monomers of postfusion H1 HA2 are compared in Section 5.1. Structural comparison between the two molecules was carried out using a monomer of B59 H1 TBHA2 (chain A; HA2 residues 65-155) present in the asymmetric unit of the P 3 2 1-unit cell, and the longer monomer of recombinantly-expressed Bac B59 H1 HA2 (chain A; HA2 residues 45-164) present in the asymmetric unit of the R 3 2-unit cell. The overall rmsd of 0.6 Å obtained by aligning the two monomers indicates a high overall level of structural similarity. Alignment of just the central helices (HA2 residues 65-105) reduces this value to ~0.44 Å. The stems of the two molecules, including the six-helical-bundle regions, are almost identical, and this part of the postfusion molecule is composed of segments C (HA2 residues 75-105), loop D (HA2 residues 106-112) and helix E (HA2 residues 113-127). The most variable region between the two structures was helix G (HA2 residues 146-153), and differences may be a result of crystal packing. Stability of these fragments may have been affected by the observed increased flexibility at the N- and C-termini of the postfusion H1 HA2, and the observed progressive shortening of the refolded molecule (Chapters 3 and 4). Structural deviations between different regions of the postfusion molecules are shown in Table 5.2. Given that viral B59 H1 TBHA2 contains a disulphide-linked fragment of HA1, and that conformational change of HA was induced by its incubation at the pH of fusion (pH ~5.0), this molecule most likely represents the conformation of the rearranged H1 HA2 *in vivo*. As the top part (helix A and part of helix B) and the extended C-terminus of the postfusion molecule are unobserved in the crystal structure of B59 H1 TBHA2, a model of postfusion H1 HA2 was generated, by combining the two postfusion H1 HA2 structures obtained using X-ray crystallography (Chapters 3 and 4). The composite model was generated by superposing central helices (HA2 residues 65-105) of the two monomers in Coot (Emsley and Cowtan 2004) and symmetry operations is shown in Figure 7.2. The middle and viral membrane-proximal part of the

molecule, including the six-helical-bundle and short fragments of HA1 domains, corresponds to the structure of B59 H1 TBHA2 obtained at a resolution of 3.4 Å. The top part of the postfusion model (HA2 residues 45-64) and regions C-terminal to the antiparallel helix G (HA2 residues 146-153) were reconstructed using the equivalent fragments of H1 HA2 present in the obtained crystal structure of recombinantly-expressed Bac B59 H1 HA2.

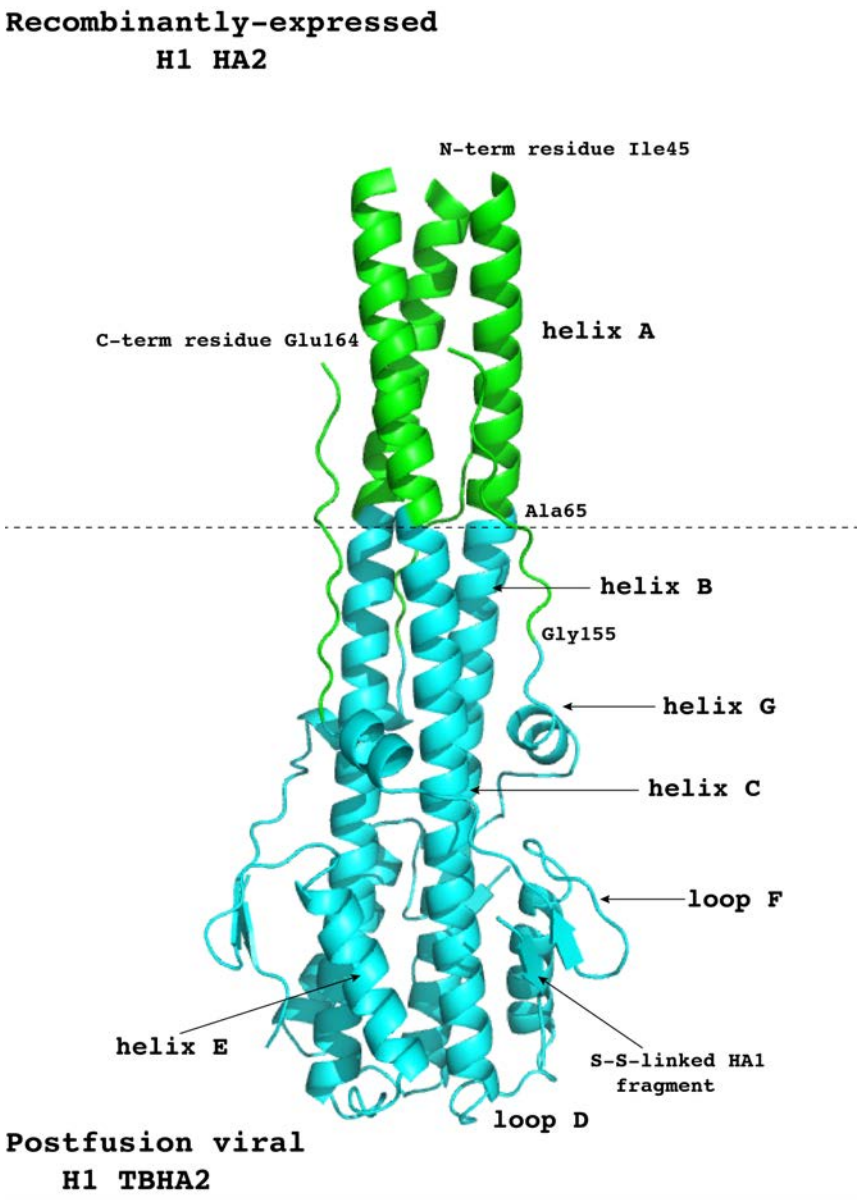


Figure 7.2 Composite model of H1 HA2 in a postfusion conformation. A trimer of B59 H1 TBHA2 is shown in cyan. The N- and C-terminal residues (Ala65 and Gly155) are indicated. The top part of the molecule was modelled using HA2 residues 45-64 present in the obtained crystal structure of Bac B59 H1 HA2 (green). Fragments of loops unobserved in B59 H1 TBHA2 were reconstructed using HA2 residues 156-164 of Bac B59 H1 HA2. Model prepared using LSQ Superpose (Coot) (Emsley and Cowtan 2004).

7.6 Comparison of postfusion H3 and H1 HA2

It has been previously suggested, that the acidic pH of endosomes is only required for dissociation of HA1 domains, and that subsequent conformational changes leading to membrane fusion follow spontaneously (Carr and Kim 1993). It has been previously shown, that structural refolding of HA2 can also occur at neutral pH (Chen, Wharton et al. 1995), and *E.coli*-expressed H3 EBHA2 spontaneously folded into a low pH-induced conformation (Chen, Skehel et al. 1999). The recombinantly-expressed fragment of Bac B59 H1 HA2 and described in Chapter 4, spontaneously folded into a low pH-induced conformation in the absence of HA1 domains. Bac B59 H1 HA2 was susceptible to digestion with trypsin, which indicated presence of the extended structure in the expressed molecule (4.2 and Figure 4.3). Thermal stability of Bac B59 H1 HA2 was evaluated using CD spectroscopy (4.3.2), and the recorded elevated melting temperatures confirmed the presence of a postfusion super helix (Table 4.3). Although the registered melting temperatures were lower than those previously recorded for H3 TBHA2 (Ruigrok, Aitken et al. 1988) and H3 EBHA2 (Chen, Wharton et al. 1995), this feature of Bac B59 H1 HA2 agrees with the idea that a prefusion HA2 is metastable.

The postfusion structures of viral B59 H1 TBHA2 and X31 H3 TBHA2 are compared in 5.2 (Figures 5.3, 5.4, 5.5 and 5.6), and the structural deviations calculated for different regions of the two postfusion molecules are shown in Table 5.5. Although the structure of B59 H1 TBHA2 presented in Chapter 3 is 25 amino acids shorter than that of X31 H3 TBHA2 prepared using a similar method (Bullough, Hughson et al. 1994), reorganization by a formation of an antiparallel six-helix-bundle, characteristic of postfusion X31 H3 TBHA2, is observed in crystal structure of viral postfusion B59 H1 TBHA2. Region comprising residues 1-64 of H1 HA2 is unstable and prone to proteolytic digestion. This observation correlates with a weak coiled coil forming potential predicted for H1 HA2 (residues 1-64) calculated based on the primary sequence (Figures 5.7 and 5.8). Despite a very low probability of coiled coil formation predicted for an equivalent region of H3 HA2, a strong cluster of hydrophobic residues in the region of Leu38-Phe63 of H3 HA2 is most likely responsible for the ability of postfusion X31 H3 TBHA2 to retain the trimeric conformation of this segment. A shorter length of the enzymatically-prepared H1 TBHA2 (both B59 & PR8), when compared to X31 H3 TBHA2, most likely results from differences in the ability of postfusion H1 HA2 to form a stable

super helix, and may indicate an increased flexibility of the N- and C-terminal regions of postfusion H1 HA2.

Interestingly, the equivalent sequences of other subtypes belonging to group 1 HA (H2 and H5) show a very similar coiled coil probability profile, when calculated using COILS/PCOILS (Bioinformatics Toolkit; Max-Planck Institute for Developmental Biology) suggesting, that other enzymatically-prepared fragments of postfusion HA2 from group 1 of HA, are also likely to start in the region of HA2 residue 65. The helical propensity of H1, H2 and H5 HA2 (group 1 HA) is shown in Figure 7.3. Coiled coil probability profiles calculated for the same regions of H7 and H10 (group 2 HA) are also very similar, and the expected helical propensity of the equivalent N-terminal fragments of H7 and H10 HA2 (residues 40-65) is even higher, than that of H3 HA2. The sequence-based coiled coil prediction for H7 and H10 HA2 (group 2 HA) is shown in Figure 7.4.

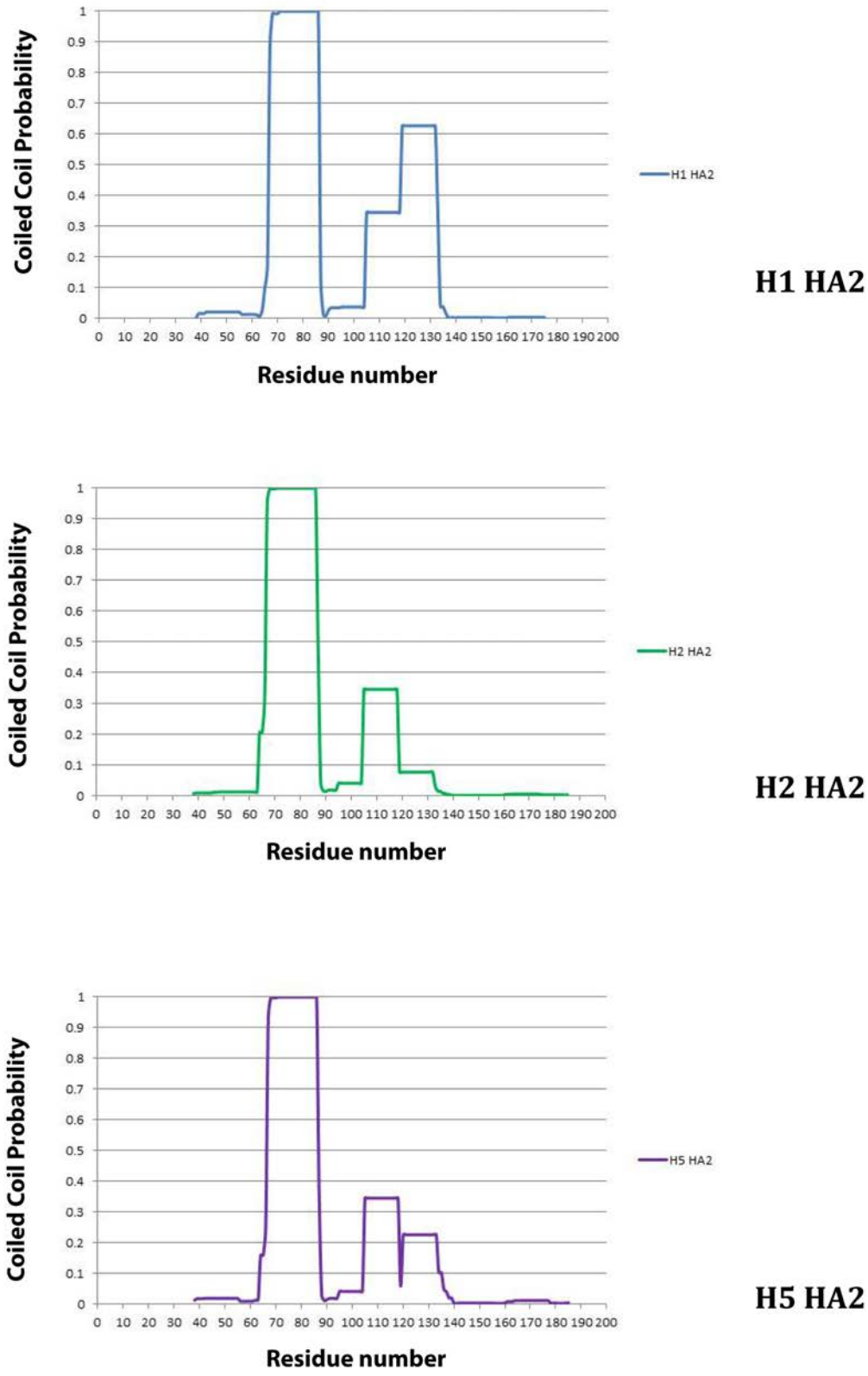


Figure 7.3 Coiled coil prediction profiles for group 1 HA. Profiles for H1, H2 and H5 HA2 are shown. Figure prepared using COILS/PCOILS (Bioinformatics Toolkit; Max-Planck Institute for Developmental Biology).

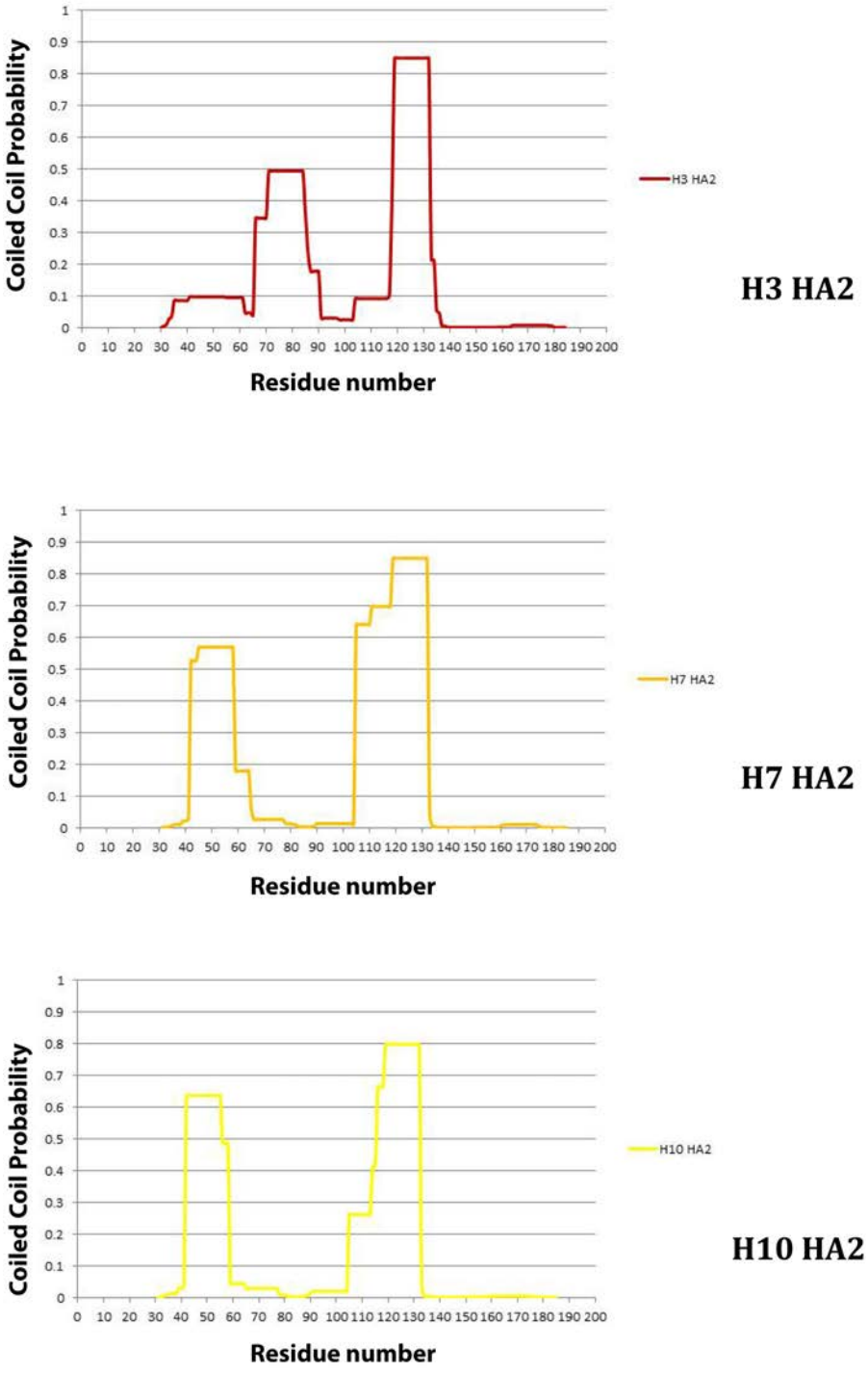


Figure 7.4 Coiled coil prediction profiles for group 2 HA. Profiles for H3, H7 and H10 HA2 are shown. Figure prepared using COILS/PCOILS (Bioinformatics Toolkit; Max-Planck Institute for Developmental Biology).

Differences in type and distribution of amino acid residues in group 1 and group 2 HA within the corresponding HA2 fragments, comprising prefusion helices A, loops B and parts of helices C, are shown in Figures 5.8 and 5.9. Distribution of amino acids in these fragments of prefusion HA2 appears to affect the size and shape of internal cavities, located between loops B and helices C of prefusion H3 and H1 HA2 (Figures 7.5 and 7.6). The molecular surface areas of H3 and H1 HA are 165649.563 $Å^2$ and 164987.031 $Å^2$, respectively, as calculated using PyMol (Molecular Graphics System, Version 1.2r3pre, Schrödinger, LLC.). Although the overall solvent accessible area is slightly greater for H3 HA (97536.930 $Å^2$) when compared to H1 HA (91539.172 $Å^2$), this cavity, together with a high number of polar residues in this region of H1 HA2, seem to directly affect the length of proteolytically prepared of H1 TBHA2 presented in this thesis. The internal cavity at the interface of loop B (HA2 residues 59-74) and helix C (HA2 residues 75-105) of H1 HA2 is shown in Figure 7.6, and its presence may also be related to the position of HA1 domains with respect to HA2 stem in prefusion H1 HA. Figure 7.5 shows the conformations of key amino acid residues within helices A and loops B in the two prefusion molecules (H3 and H1 HA2). It has been previously shown that proteolytically prepared H3 TBHA2 is cleaved at Lys39 (shown in blue), and residues at the interface of loop B and helix C (Lys58 and Lys62) are protected from proteolytic cleavage, and present in the crystal structure of X31 H3 TBHA2 (Bullough, Hughson et al. 1994). In case of enzymatically-prepared B59 H1 TBHA2 described in Chapter 3, these fragments are unobserved, and missing fragments are helix A, and fragment of loop B, up to Thr64 (shown in red). Cleavage of these residues at the interface of loop B (HA2 residues 59-74) and helix C (HA2 residues 75-105) resulting in a short length of the obtained B59 H1 TBHA2 fragment, may be related to the presence of an internal cavity at the interface of these parts of prefusion H1 HA2. Potential proteolytic cleavage sites are shown.

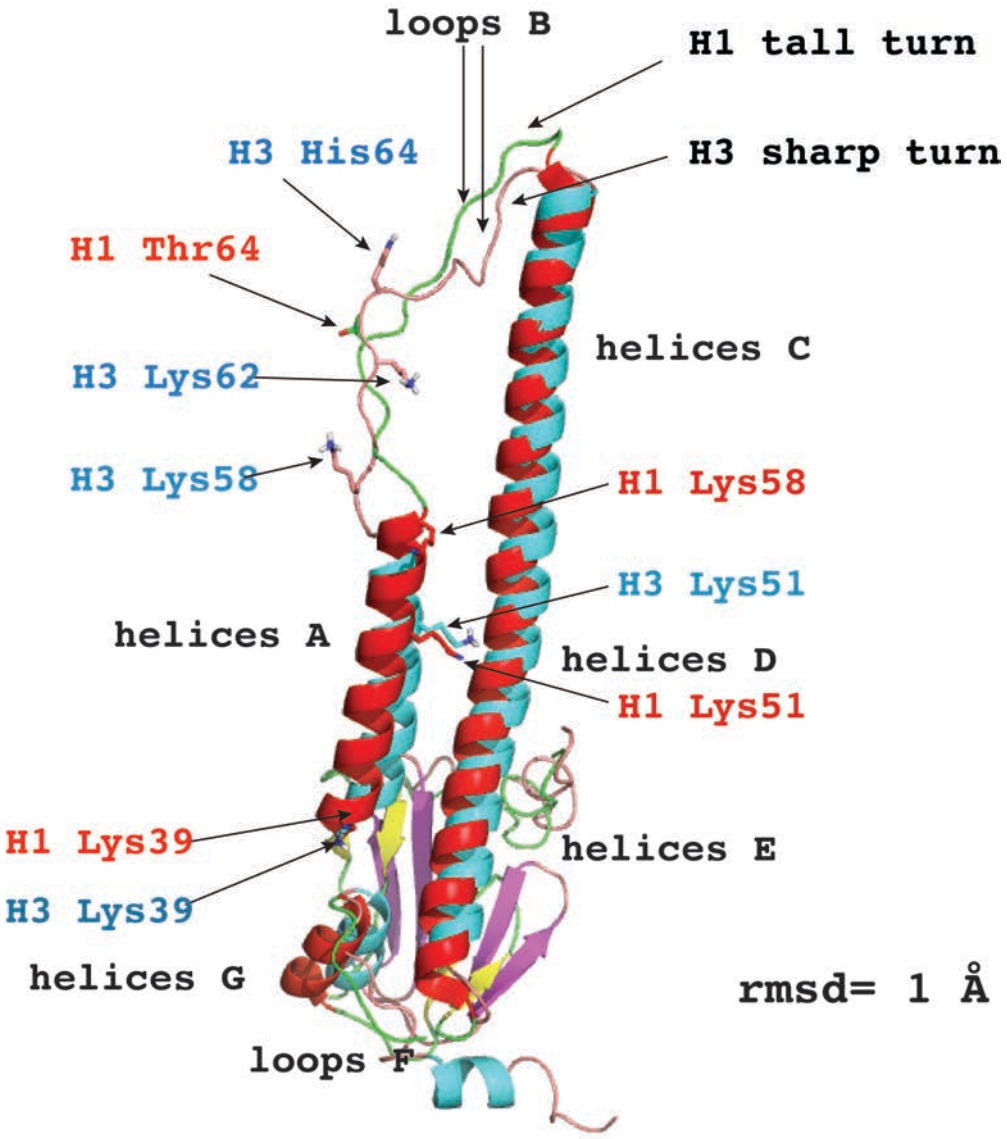


Figure 7.5 Interhelical loops in prefusion H3 and H1 HA2 structures. Monomer of prefusion H3 HA2 (PDB: 1HGF) is shown in blue, purple and salmon. Monomer of H1 HA2 (PDB: 1RUZ) is shown in red, yellow and green. Conformation of key residues at the interface of loops B and helices C in the two subtypes is shown.

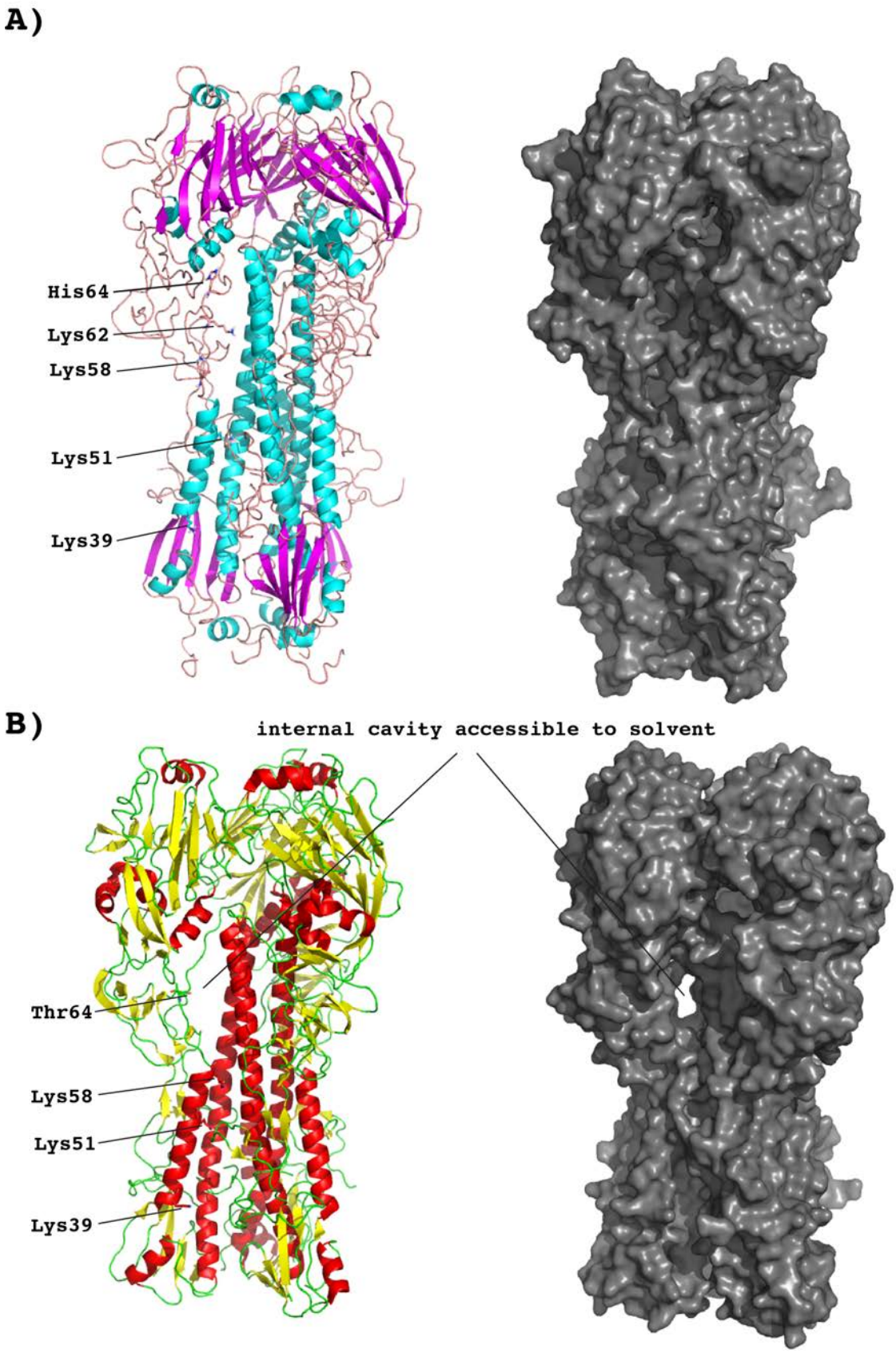


Figure 7.6 Solvent accessible areas in prefusion H3 and H1 HA. Molecules are coloured by secondary structure elements (helix-sheet-loop). H3 HA (PDB: 1HGF) is shown in blue, purple and salmon (A), and H1 HA (PDB: 1RUZ) is shown in red, yellow and green (B). Internal cavity at the interface of HA1 and HA2, and loop B and helix C, present in prefusion H1 HA is shown.

7.7 Final conclusions in relation to the low pH-induced refolding of H3 and H1 HA2

The sequence identity between residues 1-222 of B59 H1 HA2 (UniProt: D5F1Q8) and X31 H3 HA2 (UniProt: P03437) is 55% as calculated using PRALINE multiple sequence alignment (Simossis and Heringa 2003) (Simossis and Heringa 2005). The two crystal structures of postfusion H1 HA2 (group 1 HA), described in Chapters 3 and 4, were compared to the previously solved structures of postfusion H3 HA2 (group 2 HA) (Bullough, Hughson et al. 1994) (Chen, Skehel et al. 1999). In addition, the obtained postfusion structures of postfusion H1 HA2 and already known structures of postfusion H3 HA2 were analysed in relation to their prefusion form (Gamblin, Haire et al. 2004) (Sauter, Hanson et al. 1992), and detailed structural analysis is presented in Sections 5.1, 5.2 and 5.3. The data show the structures of postfusion HA2 to be conserved between the two HA groups with minor differences. Prefusion helices A and interhelical loops B, are recruited to the extended trimeric coiled coil, adopting an almost identical conformation in postfusion HA2 of the two subtypes. In both cases, the postfusion HA is a trimer of molecules arranged into a hairpin-like conformation. Despite variations in lengths of the obtained molecules, the central trimeric coiled coil, comprising fragments of helices A (HA2 residues 45-55), helices B (HA2 residues 59-74) and helices C (HA2 residues 76-105) is almost the same in the postfusion HA2 structures of H1 and H3 HA subtypes. The structural deviation between turns D (HA2 residues 106-112) in the two molecules changes from ~0.3 Å in their prefusion form, to ~0.8 Å postfusion. These turns change the direction of the polypeptide chain by 180°, and the six-helix-bundle is formed by a sideways positioning of helices E (HA2 residues 113-127). Although when aligned directly these helices are very similar (rmsd ~0.4 Å) in postfusion H1 and H3 HA2, superposition of main helices revealed that helices E of the H1 subtype sit on top of the equivalent helices E in postfusion H3 HA2. Side chain packing at the interface of helices E (HA2 residues 113-127) and helices C (HA2 residues 76-105) in postfusion structures of H1 and H3 HA2 is shown in Figure 7.7. Residue 122 of helix E is a polar Threonine in H3 HA2 (Figure 7.7, A), and a hydrophobic Valine in H1 HA2. A second substitution is at residue 115, which is Methionine in H3 HA2 and Valine in H1 HA2 (Figure 7.7, A). Approximate angles between helices E (HA2 residues 113-127) and helices C (HA2 residues 76-105) are slightly different in the postfusion structures of H1 and H3 HA2, and these are shown in Figure 7.7, B.

The stabilizing N-cap that holds the C-terminus in the proximity to the N-terminus observed previously in the structure of *E.coli*-expressed H3 EHA2 (Chen, Skehel et al. 1999) is absent in both structures of postfusion H1 HA2, described in Chapters 3 and 4 of this thesis. Some residues implicated in the formation of the helix-terminating N-cap, are conserved in the primary sequences of H1 and H3 HA2, and these are the N-terminal Ala35 and Ser40, and the C-terminal Ile173 and Val176. One exception is HA2 residue 174, which in H1 HA2 is Aspartic acid. Nonpolar Leu38 at the N-terminus of H3 HA2 is substituted with a polar Glutamine in H1 HA2. Even though the structural information relating to these fragments of refolded H1 HA2 is unavailable, the above substitutions suggest weaker interactions between the capping residues and the extended C-terminal fragments of H1 HA2.

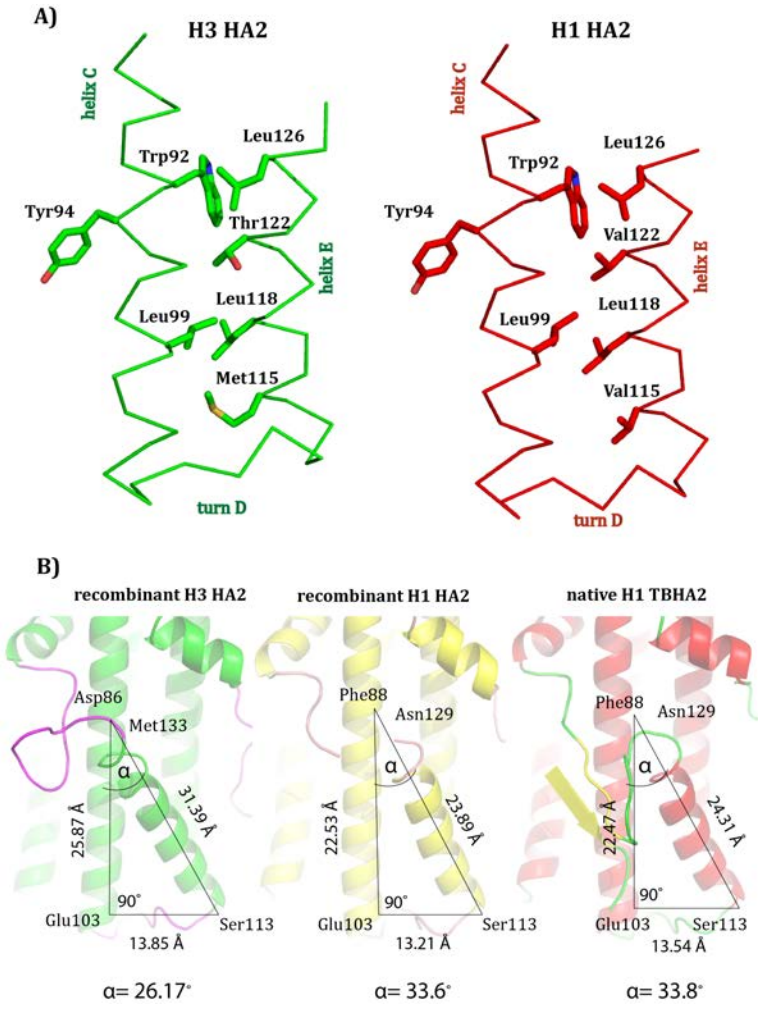


Figure 7.7 Differences in packing of helices E against central helices C in postfusion H3 and H1 HA2. Side chain packing at the interface of the two helices in H3 EHA2 (PDB: 1QU1) and B59 H1 TBHA2 is shown in A. Approximate angles between helices C and E of X31 H3 EHA2 (PDB: 1QU1), Bac B59 H1 HA2 and B59 H1 TBHA2 are shown in B.

7.8 Comparison of postfusion H1 HA2 and Influenza B HA2

The sequence identity between the corresponding fragments (residues 1-222) of B59 HA2 (UniProt: D5F1Q8) and Influenza B virus (B/Texas/3394/2013) HA2 (UniProt: S5DRP1) is 39% as calculated using PRALINE multiple sequence alignment (IBIVU Server) (Simossis and Heringa 2003) (Simossis and Heringa 2005). The postfusion structure of recombinantly-expressed Bac B59 H1 HA2 was compared to the recently published postfusion structure of influenza B HA2 (Ni, Chen et al. 2014) (Section 5.4). Structural deviations between the corresponding segments A-G that undergo refolding at low pH are summarized in Table 5.7, and the two structures are compared in Figures 5.13 and 5.14. In summary, despite low levels of sequence identity, the stems of the two molecules, comprising HA2 residues 75-105 and spanning helices A, B and C, are almost identical (Table 5.7). Superposition of turns D at residues 106-112 of HA2 gives an rmsd of ~2 Å, and the observed differences are most likely related to differences in the primary sequence. With rmsd of ~0.3 Å, helices E (HA2 residues 113-127) are almost the same in the two postfusion structures.

7.9 Implications for viral membrane fusion

Analysis and interpretation of the structural data on postfusion HA2 from group 1 and group 2 HA in relation to viral membrane fusion is problematic, due to the lack of information regarding position of the refolded molecule with respect to the two interacting membranes during viral membrane fusion. The two structures of postfusion H1 HA2 discussed in this thesis, contain features conserved in the postfusion structures of HA2 from the two HA groups, and in the postfusion structure of influenza B HA2, which shares a very low level of sequence identity with influenza A (~39%). The data show that viral membrane fusion requires formation of a trimeric hairpin-like structure, as previously found for the rearranged H3 HA2 (Bullough, Hughson et al. 1994) (Chen, Skehel et al. 1999). The shorter lengths of the obtained fragments of postfusion H1 HA2 may indicate an increased flexibility of the N- and C-termini of postfusion H1 HA2 (Figures 5.10 and 5.11). As these regions would terminate with the hydrophobic fusion peptide

and the transmembrane anchor *in vivo*, this structural flexibility of the N- and C-termini of postfusion H1 HA2 can be considered as beneficial in terms of the ability of H1 HA2 to perform viral membrane fusion suggesting, that the observed flexibility of these regions of postfusion H1 HA2 may allow for a close apposition of the two interacting membranes. The shortest fragment of postfusion B59 H1 TBHA2 shown in Figure 5.4 may indicate the core of the molecule, which must remain trimeric in order to facilitate membrane fusion. The ~6 kDa degradation product of postfusion H1 HA2, comprising the six-helical-bundle (HA2 residues 75-128) was obtained by proteolytic digestion of Bac B59 H1 HA2 (Chapter 4), which has not been exposed to proteases (other than TEV), and the obtained fragment was observed using SDS-PAGE (Figure 4.3).

Overall, the viral membrane-distal part of postfusion H1 HA2 of B59 H1N1 virus appears less likely to form a tight hydrophobic core characteristic of postfusion H3 HA2 (Bullough, Hughson et al. 1994) (Chen, Skehel et al. 1999). Although formation of the extended super helix in order to deliver the extruded fusion peptide to the target membrane is possible, as shown by the extension of the triple-stranded, α-helical coiled coil in the crystal structure of Bac B59 H1 HA2 (Chapter 4), this state appears to be unstable. The observed progressive shortening of postfusion H1 HA2 may indicate, that this segment extends initially to deliver the fusion peptide to the target membrane, but remains flexible in order to facilitate membrane fusion. In this sense, the elongated molecule observed in crystal structure of Bac B59 H1 HA2 might represent a short-lived intermediate in the process of viral membrane fusion. This finding seems interesting, given that coiled coil prediction profiles for H2 and highly pathogenic H5 are identical to that of H1 HA2, and may indicate a correlation in the way that these subtypes perform viral membrane fusion. In a hypothetical scenario that HA2 remains positioned vertically in relation to the two interacting membranes during viral membrane fusion, the stability features of the refolded HA2 of two HA groups may affect the distance of apposition between viral and endosomal membranes (H3 TBHA2 ~98 Å long/ H1 TBHA2 ~60 Å long). In case of a 'fold-back' hypothesis of membrane fusion (Skehel and Wiley 2000), the observed flexibility of postfusion H1 HA2 near its N- and C-termini may play a role in the overall efficiency of viral membrane fusion. In this model, positioning of the HA2 trimer would bring the interacting membranes into a proximity of ~4 nm in the region of six-helix-bundle, and less than 1 nm in the region of fusion peptide and transmembrane anchor. The shortest postfusion molecule present in the crystal of recombinantly-expressed H1 HA2 (Chapter 4) containing HA2 residues 75-128 measures ~45 Å (between the N-terminal Arg75

and Glu105), and vertical positioning of the molecule during viral membrane fusion, would bring the two membranes within a distance of ~4 nm.

7.10 Functional studies using the FI6 antibody

The structural data on postfusion H1 HA2 presented in this thesis are supplemented by functional studies involving the cross-reactive FI6 antibody (Corti, Voss et al. 2011) that binds near the conserved fusion subdomain of HA2 (Corti, Voss et al. 2011), and its binding interferes with the low pH-induced conformational change of HA after the uptake of the virus into the endosomes. Binding of FI6 to purified virus and to a detergent-released HA of the two subtypes was studied using the acquired susceptibility of postfusion HA to proteolytic cleavage. In this thesis, the FI6 antibody was shown to block the conformational change of H1 and H3 influenza A HA using limited digestion of a refolded, virus-bound HA with trypsin (Section 6.1.1) and electron microscopy (Section 6.1.2). The ability of the FI6 antibody to prevent the low pH-induced structural rearrangement of a detergent-released molecule (THA) was shown by a decrease in rosette formation using electron microscopy (Section 6.2). The antibody does not appear to block binding of H1N1 and H3N2 viruses to human and avian receptor analogs even at very high concentrations, as shown using surface biolayer interferometry (Section 6.3). Cell infection experiments described in Section 6.4 were performed to assess the ability of FI6 to neutralize infection of MDCK1 cells with H1N1 and H3N2 viruses. Results suggest that FI6 reduces infection with the X31 virus (H3N2), and can completely block infection with B59 (H1N1) virus at high concentrations. This observation agrees with the fact, that cross-reactive antibodies are less efficient than the strain-specific antibodies, and it is possible, that FI6 would completely block viral infection with the X31 virus at concentrations higher than used in this thesis. The FI6 antibody was shown to be more effective in blocking the conformational change of H1 HA using all of the above methods. The observed reduced sensitivity of X31 virus to the FI6-mediated neutralization, by blocking the conformational change of endosome-internalized HA, may be related to the discussed group-specific features within the antibody-binding regions of prefusion HA. These include a sugar present at HA2 Asn38 in four out of five group 2 HA subtypes, and a different conformation of Trp21, which could interfere with the accessibility of the antigenic site on helix A of H3 HA. These features are described in Section 1.8.2 on antibodies interfering with the

conformational change of HA. The overall infectivity of the X31 virus in the presence of FI6 was found to be higher than that of the B59 virus using the MDCK1 cell infection assay (6.4)

.

7.11 Future work

The understanding of HA-mediated viral membrane fusion will be facilitated when structures of HA, including the fusion peptide and the transmembrane anchor, become available. Stabilization of these hydrophobic domains of HA2 for crystallographic studies can be achieved using detergents or lipids, and crystallisation trials can be carried out using a lipidic cubic phase technique. However, the expected flexibility of those regions of the refolded HA2 molecule suggests, that the potential crystal structures of full length HA2 might only represent an intermediate in the process of membrane fusion.

Possible unfolding of the postfusion H1 HA2 super helix should be tested at various pH values and incubation times, employing the acquired susceptibility of the refolded molecule to protease digestion, and digestion products analysed by N-terminal sequencing. Additionally, the postfusion viral HA2 from other group 1 HA subtypes can be prepared using the enzymatic method and subjected to crystallisation trials, and their crystal structures can be compared to the available structures of postfusion HA2 from the two groups HA.

Interpretation of the available structural data will also be facilitated once the orientation of HA relative to the interacting membranes, and at various stages of viral membrane fusion is determined, and this may possible using a functional membrane fusion assay combined with high-resolution electron microscopy. Fusogenic properties of various constructs can be measured as a decrease in FRET between two liposome membrane-bound fluorescent probes upon fusion (Stegmann, Hoekstra et al. 1985) (Wharton, Skehel et al. 1986), and this method could be used to test the ability of the FI6-bound HA to perform viral membrane fusion.

Bibliography

(WHO 1980). A revision of the system of nomenclature for influenza viruses: a WHO memorandum. Bull World Health Organ. **58:** 585-591.

Afonine, P. V., R. W. Grosse-Kunstleve, N. Echols, J. J. Headd, N. W. Moriarty, M. Mustyakimov, T. C. Terwilliger, A. Urzhumtsev, P. H. Zwart and P. D. Adams (2012). "Towards automated crystallographic structure refinement with phenix.refine." Acta Crystallogr D Biol Crystallogr **68**(Pt 4): 352-367.

Air, G. M. (1981). "Sequence relationships among the hemagglutinin genes of 12 subtypes of influenza A virus." Proc Natl Acad Sci U S A **78**(12): 7639-7643.

Area, E., J. Martin-Benito, P. Gastaminza, E. Torreira, J. M. Valpuesta, J. L. Carrascosa and J. Ortin (2004). "3D structure of the influenza virus polymerase complex: localization of subunit domains." Proc Natl Acad Sci U S A **101**(1): 308-313.

Aytay, S. and I. T. Schulze (1991). "Single amino acid substitutions in the hemagglutinin can alter the host range and receptor binding properties of H1 strains of influenza A virus." J Virol **65**(6): 3022-3028.

Barbey-Martin, C., B. Gigant, T. Bizebard, L. J. Calder, S. A. Wharton, J. J. Skehel and M. Knossow (2002). "An antibody that prevents the hemagglutinin low pH fusogenic transition." Virology **294**(1): 70-74.

Baum, L. G. and J. C. Paulson (1990). "Sialyloligosaccharides of the respiratory epithelium in the selection of human influenza virus receptor specificity." Acta Histochem Suppl **40**: 35-38.

Belser, J. A., K. M. Gustin, M. B. Pearce, T. R. Maines, H. Zeng, C. Pappas, X. Sun, P. J. Carney, J. M. Villanueva, J. Stevens, J. M. Katz and T. M. Tumpey (2013). "Pathogenesis and transmission of avian influenza A (H7N9) virus in ferrets and mice." Nature **501**(7468): 556-559.

Bergfors, T. M. (2009). Protein Crystallization. La Jolla, California, International University Line.

Bizebard, T., B. Gigant, P. Rigolet, B. Rasmussen, O. Diat, P. Bosecke, S. A. Wharton, J. J. Skehel and M. Knossow (1995). "Structure of influenza virus haemagglutinin complexed with a neutralizing antibody." Nature **376**(6535): 92-94.

Boriskin, Y. S., I. A. Leneva, E. I. Pecheur and S. J. Polyak (2008). "Arbidol: a broad-spectrum antiviral compound that blocks viral fusion." Curr Med Chem **15**(10): 997-1005.

Bos, T. J., A. R. Davis and D. P. Nayak (1984). "NH2-terminal hydrophobic region of influenza virus neuraminidase provides the signal function in translocation." Proc Natl Acad Sci U S A **81**(8): 2327-2331.

Brand, C. M. and J. J. Skehel (1972). "Crystalline antigen from the influenza virus envelope." Nat New Biol **238**(83): 145-147.

Breschkin, A. M., J. Ahern and D. O. White (1981). "Antigenic determinants of influenza virus hemagglutinin. VIII. Topography of the antigenic regions of influenza virus hemagglutinin determined by competitive radioimmunoassay with monoclonal antibodies." Virology **113**(1): 130-140.

Bricogne G., B. E., Brandl M., Flensburg C., Keller P., Paciorek W., Roversi P., Sharff A., Smart O. S., Vonrhein C., Womack T. O. (2011). BUSTER version X. Y. Z. Cambridge, United Kingdom, Global Phasing Ltd.

Bullough, P. A., F. M. Hughson, J. J. Skehel and D. C. Wiley (1994). "Structure of influenza haemagglutinin at the pH of membrane fusion." Nature **371**(6492): 37-43.

Bullough, P. A., F. M. Hughson, A. C. Treharne, R. W. Ruigrok, J. J. Skehel and D. C. Wiley (1994). "Crystals of a fragment of influenza haemagglutinin in the low pH induced conformation." J Mol Biol **236**(4): 1262-1265.

Burnet, F. M. and J. D. Stone (1947). "The receptor-destroying enzyme of V. cholerae." Aust J Exp Biol Med Sci **25**(Pt 3): 227-233.

Calder, L. J., S. Wasilewski, J. A. Berriman and P. B. Rosenthal (2010). "Structural organization of a filamentous influenza A virus." Proc Natl Acad Sci U S A **107**(23): 10685-10690.

Carr, C. M. and P. S. Kim (1993). "A spring-loaded mechanism for the conformational change of influenza hemagglutinin." Cell **73**(4): 823-832.

Castells, M. T., J. Ballesta, L. M. Pastor, J. F. Madrid and J. A. Marin (1990). "Histochemical characterization of glycoconjugates in the epithelium of the extrapulmonary airways of several vertebrates." Histochem J **22**(1): 24-35.

Caton, A. J., G. G. Brownlee, J. W. Yewdell and W. Gerhard (1982). "The antigenic structure of the influenza virus A/PR/8/34 hemagglutinin (H1 subtype)." Cell **31**(2 Pt 1): 417-427.

Chen, J., K. H. Lee, D. A. Steinhauer, D. J. Stevens, J. J. Skehel and D. C. Wiley (1998). "Structure of the hemagglutinin precursor cleavage site, a determinant of influenza pathogenicity and the origin of the labile conformation." Cell **95**(3): 409-417.

Chen, J., J. J. Skehel and D. C. Wiley (1998). "A polar octapeptide fused to the N-terminal fusion peptide solubilizes the influenza virus HA2 subunit ectodomain." Biochemistry **37**(39): 13643-13649.

Chen, J., J. J. Skehel and D. C. Wiley (1999). "N- and C-terminal residues combine in the fusion-pH influenza hemagglutinin HA(2) subunit to form an N cap that terminates the triple-stranded coiled coil." Proc Natl Acad Sci U S A **96**(16): 8967-8972.

Chen, J., S. A. Wharton, W. Weissenhorn, L. J. Calder, F. M. Hughson, J. J. Skehel and D. C. Wiley (1995). "A soluble domain of the membrane-anchoring chain of influenza virus hemagglutinin (HA2) folds in Escherichia coli into the low-pH-induced conformation." Proc Natl Acad Sci U S A **92**(26): 12205-12209.

Chen, L. M., O. Blixt, J. Stevens, A. S. Lipatov, C. T. Davis, B. E. Collins, N. J. Cox, J. C. Paulson and R. O. Donis (2012). "In vitro evolution of H5N1 avian influenza virus toward human-type receptor specificity." Virology **422**(1): 105-113.

Chen, W., J. Helenius, I. Braakman and A. Helenius (1995). "Cotranslational folding and calnexin binding during glycoprotein synthesis." Proc Natl Acad Sci U S A **92**(14): 6229-6233.

Colman, P. M. (1994). "Influenza virus neuraminidase: structure, antibodies, and inhibitors." Protein Sci **3**(10): 1687-1696.

Compans, R. W. (1973). "Influenza virus proteins. II. Association with components of the cytoplasm." Virology **51**(1): 56-70.

Compans, R. W., J. Content and P. H. Duesberg (1972). "Structure of the ribonucleoprotein of influenza virus." J Virol **10**(4): 795-800.

Compans, R. W., H. D. Klenk, L. A. Caliguiri and P. W. Choppin (1970). "Influenza virus proteins. I. Analysis of polypeptides of the virion and identification of spike glycoproteins." Virology **42**(4): 880-889.

Connor, R. J., Y. Kawaoka, R. G. Webster and J. C. Paulson (1994). "Receptor specificity in human, avian, and equine H2 and H3 influenza virus isolates." Virology **205**(1): 17-23.

Copeland, C. S., K. P. Zimmer, K. R. Wagner, G. A. Healey, I. Mellman and A. Helenius (1988). "Folding, trimerization, and transport are sequential events in the biogenesis of influenza virus hemagglutinin." Cell **53**(2): 197-209.

Corti, D., J. Voss, S. J. Gamblin, G. Codoni, A. Macagno, D. Jarrossay, S. G. Vachieri, D. Pinna, A. Minola, F. Vanzetta, C. Silacci, B. M. Fernandez-Rodriguez, G. Agatic, S. Bianchi, I. Giacchetto-Sasselli, L. Calder, F. Sallusto, P. Collins, L. F. Haire, N. Temperton, J. P. Langedijk, J. J. Skehel and A. Lanzavecchia (2011). "A neutralizing antibody selected from plasma cells that binds to group 1 and group 2 influenza A hemagglutinins." Science **333**(6044): 850-856.

Couceiro, J. N., J. C. Paulson and L. G. Baum (1993). "Influenza virus strains selectively recognize sialyloligosaccharides on human respiratory epithelium; the role of the host cell in selection of hemagglutinin receptor specificity." Virus Res **29**(2): 155-165.

Couch, R. B. (2008). "Seasonal inactivated influenza virus vaccines." Vaccine **26 Suppl 4**: D5-9.

Crick, F. H. C. (1953). "The packing of alpha-helices: simple coiled-coils." Acta. Crystallogr **6**: 689-697.

Cross, K. J., W. A. Langley, R. J. Russell, J. J. Skehel and D. A. Steinhauer (2009). "Composition and functions of the influenza fusion peptide." Protein Pept Lett **16**(7): 766-778.

Cross, K. J., S. A. Wharton, J. J. Skehel, D. C. Wiley and D. A. Steinhauer (2001). "Studies on influenza haemagglutinin fusion peptide mutants generated by reverse genetics." EMBO J **20**(16): 4432-4442.

D'Arcy, A., T. Bergfors, S. W. Cowan-Jacob and M. Marsh (2014). "Microseed matrix screening for optimization in protein crystallization: what have we learned?" Acta Crystallogr F Struct Biol Commun **70**(Pt 9): 1117-1126.

D'Arcy, A., F. Villard and M. Marsh (2007). "An automated microseed matrix-screening method for protein crystallization." Acta Crystallogr D Biol Crystallogr **63**(Pt 4): 550-554.

Daniels, R. S., A. R. Douglas, J. J. Skehel and D. C. Wiley (1983). "Analyses of the antigenicity of influenza haemagglutinin at the pH optimum for virus-mediated membrane fusion." J Gen Virol **64 (Pt 8)**: 1657-1662.

Daniels, R. S., Douglas, A. R., Skehel, J. J., Waterfield, M. D., Wilson, L. A., Wiley, D. C. (1983). Studies of the influenza virus hemagglutinin in the pH5 conformation. The Origin of Pandemic Influenza Viruses. W. G. Laver. Amsterdam, Elsevier**:** 1-7.

Domingo, E., D. Sabo, T. Taniguchi and C. Weissmann (1978). "Nucleotide sequence heterogeneity of an RNA phage population." Cell **13**(4): 735-744.

Doms, R. W., A. Helenius and J. White (1985). "Membrane fusion activity of the influenza virus hemagglutinin. The low pH-induced conformational change." J Biol Chem **260**(5): 2973-2981.

Dopheide, T. A. and C. W. Ward (1981). "The location of the bromelain cleavage site in a Hong Kong influenza virus Haemagglutinin." J Gen Virol **52**(Pt 2): 367-370.

Dowdle, W. R., M. T. Coleman, E. C. Hall and V. Knez (1969). "Properties of the Hong Kong influenza virus. Antigenic relationship of the Hong Kong virus haemagglutinin to that of other human influenza A viruses." Bull World Health Organ **41**(3): 419-424.

Dreyfus, C., N. S. Laursen, T. Kwaks, D. Zuijdgeest, R. Khayat, D. C. Ekiert, J. H. Lee, Z. Metlagel, M. V. Bujny, M. Jongeneelen, R. van der Vlugt, M. Lamrani, H. J. Korse, E. Geelen, O. Sahin, M. Sieuwerts, J. P. Brakenhoff, R. Vogels, O. T. Li, L. L. Poon, M. Peiris, W. Koudstaal, A. B. Ward, I. A. Wilson, J. Goudsmit and R. H. Friesen (2012). "Highly conserved protective epitopes on influenza B viruses." Science **337**(6100): 1343-1348.

Dunham, E. J., V. G. Dugan, E. K. Kaser, S. E. Perkins, I. H. Brown, E. C. Holmes and J. K. Taubenberger (2009). "Different evolutionary trajectories of European avian-like and classical swine H1N1 influenza A viruses." J Virol **83**(11): 5485-5494.

Ekiert, D. C., G. Bhabha, M. A. Elsliger, R. H. Friesen, M. Jongeneelen, M. Throsby, J. Goudsmit and I. A. Wilson (2009). "Antibody recognition of a highly conserved influenza virus epitope." Science **324**(5924): 246-251.

Ekiert, D. C., A. K. Kashyap, J. Steel, A. Rubrum, G. Bhabha, R. Khayat, J. H. Lee, M. A. Dillon, R. E. O'Neil, A. M. Faynboym, M. Horowitz, L. Horowitz, A. B. Ward, P. Palese, R. Webby, R. A. Lerner, R. R. Bhatt and I. A. Wilson (2012). "Cross-neutralization of influenza A viruses mediated by a single antibody loop." Nature **489**(7417): 526-532.

Emsley, P. and K. Cowtan (2004). "Coot: model-building tools for molecular graphics." Acta Crystallogr D Biol Crystallogr **60**(Pt 12 Pt 1): 2126-2132.

Evans, P. (2006). "Scaling and assessment of data quality." Acta Crystallogr D Biol Crystallogr **62**(Pt 1): 72-82.

Evans, P. R. (2011). "An introduction to data reduction: space-group determination, scaling and intensity statistics." Acta Crystallogr D Biol Crystallogr **67**(Pt 4): 282-292.

Evans, P. R. and G. N. Murshudov (2013). "How good are my data and what is the resolution?" Acta Crystallogr D Biol Crystallogr **69**(Pt 7): 1204-1214.

Flanagan, M. T. and J. J. Skehel (1977). "The conformation of influenza virus haemagglutinin." FEBS Lett **80**(1): 57-60.

Fleury, D., B. Barrere, T. Bizebard, R. S. Daniels, J. J. Skehel and M. Knossow (1999). "A complex of influenza hemagglutinin with a neutralizing antibody that binds outside the virus receptor binding site." Nat Struct Biol **6**(6): 530-534.

Fleury, D., S. A. Wharton, J. J. Skehel, M. Knossow and T. Bizebard (1998). "Antigen distortion allows influenza virus to escape neutralization." Nat Struct Biol **5**(2): 119-123.

Fouchier, R. A., V. Munster, A. Wallensten, T. M. Bestebroer, S. Herfst, D. Smith, G. F. Rimmelzwaan, B. Olsen and A. D. Osterhaus (2005). "Characterization of a novel influenza A virus hemagglutinin subtype (H16) obtained from black-headed gulls." J Virol **79**(5): 2814-2822.

Francis, T., Jr. (1934). "Transmission of Influenza by a Filterable Virus." Science **80**(2081): 457-459.

Francis, T., Jr. (1940). "A New Type of Virus from Epidemic Influenza." Science **92**(2392): 405-408.

Fraser, C., C. A. Donnelly, S. Cauchemez, W. P. Hanage, M. D. Van Kerkhove, T. D. Hollingsworth, J. Griffin, R. F. Baggaley, H. E. Jenkins, E. J. Lyons, T. Jombart, W. R. Hinsley, N. C. Grassly, F. Balloux, A. C. Ghani, N. M. Ferguson, A. Rambaut, O. G. Pybus, H. Lopez-Gatell, C. M. Alpuche-Aranda, I. B. Chapela, E. P. Zavala, D. M. Guevara, F. Checchi, E. Garcia, S. Hugonnet, C. Roth and W. H. O. R. P. A. Collaboration (2009). "Pandemic potential of a strain of influenza A (H1N1): early findings." Science **324**(5934): 1557-1561.

Gallagher, P. J., J. M. Henneberry, J. F. Sambrook and M. J. Gething (1992). "Glycosylation requirements for intracellular transport and function of the hemagglutinin of influenza virus." J Virol **66**(12): 7136-7145.

Gambaryan, A. S., V. P. Marinina, A. B. Tuzikov, N. V. Bovin, I. A. Rudneva, B. V. Sinitsyn, A. A. Shilov and M. N. Matrosovich (1998). "Effects of host-dependent glycosylation of hemagglutinin on receptor-binding properties on H1N1 human influenza A virus grown in MDCK cells and in embryonated eggs." Virology **247**(2): 170-177.

Gambaryan, A. S., A. B. Tuzikov, V. E. Piskarev, S. S. Yamnikova, D. K. Lvov, J. S. Robertson, N. V. Bovin and M. N. Matrosovich (1997). "Specification of receptor-binding phenotypes of influenza virus isolates from different hosts using synthetic sialylglycopolymers: non-egg-adapted human H1 and H3 influenza A and influenza B viruses share a common high binding affinity for 6'-sialyl(N-acetyllactosamine)." Virology **232**(2): 345-350.

Gamblin, S. J., L. F. Haire, R. J. Russell, D. J. Stevens, B. Xiao, Y. Ha, N. Vasisht, D. A. Steinhauer, R. S. Daniels, A. Elliot, D. C. Wiley and J. J. Skehel (2004). "The structure and receptor binding properties of the 1918 influenza hemagglutinin." Science **303**(5665): 1838-1842.

Gamblin, S. J. and J. J. Skehel (2010). "Influenza hemagglutinin and neuraminidase membrane glycoproteins." J Biol Chem **285**(37): 28403-28409.

Garten, W., F. X. Bosch, D. Linder, R. Rott and H. D. Klenk (1981). "Proteolytic activation of the influenza virus hemagglutinin: The structure of the cleavage site and the enzymes involved in cleavage." Virology **115**(2): 361-374.

Garten, W., S. Hallenberger, D. Ortmann, W. Schafer, M. Vey, H. Angliker, E. Shaw and H. D. Klenk (1994). "Processing of viral glycoproteins by the subtilisin-like endoprotease furin and its inhibition by specific peptidylchloroalkylketones." Biochimie **76**(3-4): 217-225.

George, M., M. Farooq, T. Dang, B. Cortes, J. Liu and L. Maranga (2010). "Production of cell culture (MDCK) derived live attenuated influenza vaccine (LAIV) in a fully disposable platform process." Biotechnol Bioeng **106**(6): 906-917.

Gething, M. J., K. McCammon and J. Sambrook (1986). "Expression of wild-type and mutant forms of influenza hemagglutinin: the role of folding in intracellular transport." Cell **46**(6): 939-950.

Gething, M. J., J. M. White and M. D. Waterfield (1978). "Purification of the fusion protein of Sendai virus: analysis of the NH2-terminal sequence generated during precursor activation." Proc Natl Acad Sci U S A **75**(6): 2737-2740.

Gibbs, J. S., D. Malide, F. Hornung, J. R. Bennink and J. W. Yewdell (2003). "The influenza A virus PB1-F2 protein targets the inner mitochondrial membrane via a predicted basic amphipathic helix that disrupts mitochondrial function." J Virol **77**(13): 7214-7224.

Godley, L., J. Pfeifer, D. Steinhauer, B. Ely, G. Shaw, R. Kaufmann, E. Suchanek, C. Pabo, J. J. Skehel, D. C. Wiley and et al. (1992). "Introduction of intersubunit disulfide bonds in the membrane-distal region of the influenza hemagglutinin abolishes membrane fusion activity." Cell **68**(4): 635-645.

Gottschalk, A. (1959). Chemistry of virus receptors. The viruses: biochemical, biological, and biophysical properties. S. W. M. Burnet F.M. New York, Academic Press, Inc.

Gottschalk, A. (1959). "On the mechanism underlying initiation of influenza virus infection." Ergeb Mikrobiol Immunitatsforsch Exp Ther **32**: 1-22.

Graves, P. N., J. L. Schulman, J. F. Young and P. Palese (1983). "Preparation of influenza virus subviral particles lacking the HA1 subunit of hemagglutinin: unmasking of cross-reactive HA2 determinants." Virology **126**(1): 106-116.

Guo, X. L., L. Li, D. Q. Wei, Y. S. Zhu and K. C. Chou (2008). "Cleavage mechanism of the H5N1 hemagglutinin by trypsin and furin." Amino Acids **35**(2): 375-382.

Ha, Y., D. J. Stevens, J. J. Skehel and D. C. Wiley (2002). "H5 avian and H9 swine influenza virus haemagglutinin structures: possible origin of influenza subtypes." EMBO J **21**(5): 865-875.

Ha, Y., D. J. Stevens, J. J. Skehel and D. C. Wiley (2003). "X-ray structure of the hemagglutinin of a potential H3 avian progenitor of the 1968 Hong Kong pandemic influenza virus." Virology **309**(2): 209-218.

Haaheim, L. R. and G. C. Schild (1980). "Antibodies to the strain-specific and cross-reactive determinants of the haemagglutinin of influenza H3N2 viruses. Antiviral activities of the antibodies in biological systems." Acta Pathol Microbiol Scand B **88**(6): 335-340.

Harrison, S. C. (2008). "Viral membrane fusion." Nat Struct Mol Biol **15**(7): 690-698.

Hay, A. J. (1974). "Studies on the formation of the influenza virus envelope." Virology **60**(2): 398-418.

Hay, A. J., V. Gregory, A. R. Douglas and Y. P. Lin (2001). "The evolution of human influenza viruses." Philos Trans R Soc Lond B Biol Sci **356**(1416): 1861-1870.

Hay, A. J., B. Lomniczi, A. R. Bellamy and J. J. Skehel (1977). "Transcription of the influenza virus genome." Virology **83**(2): 337-355.

Herfst, S., E. J. Schrauwen, M. Linster, S. Chutinimitkul, E. de Wit, V. J. Munster, E. M. Sorrell, T. M. Bestebroer, D. F. Burke, D. J. Smith, G. F. Rimmelzwaan, A. D. Osterhaus and R. A. Fouchier (2012). "Airborne transmission of influenza A/H5N1 virus between ferrets." Science **336**(6088): 1534-1541.

Hinshaw, V. S., G. M. Air, A. J. Gibbs, L. Graves, B. Prescott and D. Karunakaran (1982). "Antigenic and genetic characterization of a novel hemagglutinin subtype of influenza A viruses from gulls." J Virol **42**(3): 865-872.

Hirst, G. K. (1941). "The Agglutination of Red Cells by Allantoic Fluid of Chick Embryos Infected with Influenza Virus." Science **94**(2427): 22-23.

Hirst, G. K. (1942). "The Quantitative Determination of Influenza Virus and Antibodies by Means of Red Cell Agglutination." J Exp Med **75**(1): 49-64.

Horimoto, T. and Y. Kawaoka (2005). "Influenza: lessons from past pandemics, warnings from current incidents." Nat Rev Microbiol **3**(8): 591-600.

Huang, Q., R. Opitz, E. W. Knapp and A. Herrmann (2002). "Protonation and stability of the globular domain of influenza virus hemagglutinin." Biophys J **82**(2): 1050-1058.

Huang, R. T., R. Rott and H. D. Klenk (1981). "Influenza viruses cause hemolysis and fusion of cells." Virology **110**(1): 243-247.

Imai, M., T. Mizuno and K. Kawasaki (2006). "Membrane fusion by single influenza hemagglutinin trimers. Kinetic evidence from image analysis of hemagglutinin-reconstituted vesicles." J Biol Chem **281**(18): 12729-12735.

Imai, M., T. Watanabe, M. Hatta, S. C. Das, M. Ozawa, K. Shinya, G. Zhong, A. Hanson, H. Katsura, S. Watanabe, C. Li, E. Kawakami, S. Yamada, M. Kiso, Y. Suzuki, E. A. Maher, G. Neumann and Y. Kawaoka (2012). "Experimental adaptation of an influenza H5 HA confers respiratory droplet transmission to a reassortant H5 HA/H1N1 virus in ferrets." Nature **486**(7403): 420-428.

Ito, T. (2000). "Interspecies transmission and receptor recognition of influenza A viruses." Microbiol Immunol **44**(6): 423-430.

Ito, T., J. N. Couceiro, S. Kelm, L. G. Baum, S. Krauss, M. R. Castrucci, I. Donatelli, H. Kida, J. C. Paulson, R. G. Webster and Y. Kawaoka (1998). "Molecular basis for the generation in pigs of influenza A viruses with pandemic potential." J Virol **72**(9): 7367-7373.

Ito, T., Y. Suzuki, T. Suzuki, A. Takada, T. Horimoto, K. Wells, H. Kida, K. Otsuki, M. Kiso, H. Ishida and Y. Kawaoka (2000). "Recognition of N-glycolylneuraminic acid linked to galactose by the alpha2,3 linkage is associated with intestinal replication of influenza A virus in ducks." J Virol **74**(19): 9300-9305.

Itoh, Y., K. Shinya, M. Kiso, T. Watanabe, Y. Sakoda, M. Hatta, Y. Muramoto, D. Tamura, Y. Sakai-Tagawa, T. Noda, S. Sakabe, M. Imai, Y. Hatta, S. Watanabe, C. Li, S. Yamada, K. Fujii, S. Murakami, H. Imai, S. Kakugawa, M. Ito, R. Takano, K. Iwatsuki-Horimoto, M. Shimojima, T. Horimoto, H. Goto, K. Takahashi, A. Makino, H. Ishigaki, M. Nakayama, M. Okamatsu, K. Takahashi, D. Warshauer, P. A. Shult, R. Saito, H. Suzuki, Y. Furuta, M. Yamashita, K. Mitamura, K. Nakano, M. Nakamura, R. Brockman-Schneider, H. Mitamura, M. Yamazaki, N. Sugaya, M. Suresh, M. Ozawa, G. Neumann, J. Gern, H. Kida, K. Ogasawara and Y. Kawaoka (2009). "In vitro and in vivo characterization of new swine-origin H1N1 influenza viruses." Nature **460**(7258): 1021-1025.

Kabsch, W. (2010). "Xds." Acta Crystallogr D Biol Crystallogr **66**(Pt 2): 125-132.

Kahn, K. E., T. A. Santibanez, Y. Zhai and J. A. Singleton (2015). "Influenza vaccination type, live, attenuated influenza vaccine (LAIV) versus inactivated influenza vaccine (IIV), received by children, United States, 2011-12 through 2013-14 influenza seasons." Vaccine **33**(39): 5196-5203.

Kantardjieff, K. A. and B. Rupp (2003). "Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals." Protein Sci **12**(9): 1865-1871.

Kates, M., A. C. Allison, D. A. Tyrell and A. T. James (1962). "Origin of lipids in influenza virus." Cold Spring Harb Symp Quant Biol **27**: 293-301.

Kawaoka, Y., S. Krauss and R. G. Webster (1989). "Avian-to-human transmission of the PB1 gene of influenza A viruses in the 1957 and 1968 pandemics." J Virol **63**(11): 4603-4608.

Kawaoka, Y., S. Yamnikova, T. M. Chambers, D. K. Lvov and R. G. Webster (1990). "Molecular characterization of a new hemagglutinin, subtype H14, of influenza A virus." Virology **179**(2): 759-767.

Keil, W., R. Geyer, J. Dabrowski, U. Dabrowski, H. Niemann, S. Stirm and H. D. Klenk (1985). "Carbohydrates of influenza virus. Structural elucidation of the individual glycans of the FPV hemagglutinin by two-dimensional 1H n.m.r. and methylation analysis." EMBO J **4**(10): 2711-2720.

Kilbourne, E. D. (1969). "Future influenza vaccines and the use of genetic recombinants." Bull World Health Organ **41**(3): 643-645.

Klenk, H. D., W. Garten and M. Matrosovich (2011). "Molecular mechanisms of interspecies transmission and pathogenicity of influenza viruses: Lessons from the 2009 pandemic." Bioessays **33**(3): 180-188.

Klenk, H. D., R. Rott, M. Orlich and J. Blodorn (1975). "Activation of influenza A viruses by trypsin treatment." Virology **68**(2): 426-439.

Klenk, H. D., C. Scholtissek and R. Rott (1972). "Inhibition of glycoprotein biosynthesis of influenza virus by D-glucosamine and 2-deoxy-D-glucose." Virology **49**(3): 723-734.

Knossow, M. and J. J. Skehel (2006). "Variation and infectivity neutralization in influenza." Immunology **119**(1): 1-7.

Kochs, G., A. Garcia-Sastre and L. Martinez-Sobrido (2007). "Multiple anti-interferon actions of the influenza A virus NS1 protein." J Virol **81**(13): 7011-7021.

Lamb, R. A. and P. W. Choppin (1983). "The gene structure and replication of influenza virus." Annu Rev Biochem **52**: 467-506.

Lamb RA, K. R. (2001). Orthomyxoviridae: The viruses and their Replication. Philadelphia, Lippincott.

Larkin, M. A., G. Blackshields, N. P. Brown, R. Chenna, P. A. McGettigan, H. McWilliam, F. Valentin, I. M. Wallace, A. Wilm, R. Lopez, J. D. Thompson, T. J. Gibson and D. G. Higgins (2007). "Clustal W and Clustal X version 2.0." Bioinformatics **23**(21): 2947-2948.

Laver, W. G. (1971). "Separation of two polypeptide chains from the hemagglutinin subunit of influenza virus." Virology **45**(1): 275-288.

Laver, W. G. and E. D. Kilbourne (1966). "Identification in a recombinant influenza virus of structural proteins derived from both parents." Virology **30**(3): 493-501.

Laver, W. G. and R. C. Valentine (1969). "Morphology of the isolated hemagglutinin and neuraminidase subunits of influenza virus." Virology **38**(1): 105-119.

Lazarowitz, S. G. and P. W. Choppin (1975). "Enhancement of the infectivity of influenza A and B viruses by proteolytic cleavage of the hemagglutinin polypeptide." Virology **68**(2): 440-454.

Lazarowitz, S. G., R. W. Compans and P. W. Choppin (1971). "Influenza virus structural and nonstructural proteins in infected cells and their plasma membranes." Virology **46**(3): 830-843.

Lebedev, A. A., A. A. Vagin and G. N. Murshudov (2008). "Model preparation in MOLREP and examples of model improvement using X-ray data." Acta Crystallogr D Biol Crystallogr **64**(Pt 1): 33-39.

Lee, P. S., R. Yoshida, D. C. Ekiert, N. Sakai, Y. Suzuki, A. Takada and I. A. Wilson (2012). "Heterosubtypic antibody recognition of the influenza virus hemagglutinin receptor binding site enhanced by avidity." Proc Natl Acad Sci U S A **109**(42): 17040-17045.

Leneva, I. A., R. J. Russell, Y. S. Boriskin and A. J. Hay (2009). "Characteristics of arbidol-resistant mutants of influenza virus: implications for the mechanism of anti-influenza action of arbidol." Antiviral Res **81**(2): 132-140.

Lew, W., X. Chen and C. U. Kim (2000). "Discovery and development of GS 4104 (oseltamivir): an orally active influenza neuraminidase inhibitor." Curr Med Chem **7**(6): 663-672.

Lin, Y. P., X. Xiong, S. A. Wharton, S. R. Martin, P. J. Coombs, S. G. Vachieri, E. Christodoulou, P. A. Walker, J. Liu, J. J. Skehel, S. J. Gamblin, A. J. Hay, R. S. Daniels and J. W. McCauley (2012). "Evolution of the receptor binding properties of the influenza A(H3N2) hemagglutinin." Proc Natl Acad Sci U S A **109**(52): 21474-21479.

Luft, J. R. and G. T. DeTitta (1999). "A method to produce microseed stock for use in the crystallization of biological macromolecules." Acta Crystallogr D Biol Crystallogr **55**(Pt 5): 988-993.

Maeda, T. and S. Ohnishi (1980). "Activation of influenza virus by acidic media causes hemolysis and fusion of erythrocytes." FEBS Lett **122**(2): 283-287.

Matrosovich, M., A. Tuzikov, N. Bovin, A. Gambaryan, A. Klimov, M. R. Castrucci, I. Donatelli and Y. Kawaoka (2000). "Early alterations of the receptor-binding properties of H1, H2, and H3 avian influenza virus hemagglutinins after their introduction into mammals." J Virol **74**(18): 8502-8512.

Matrosovich, M., N. Zhou, Y. Kawaoka and R. Webster (1999). "The surface glycoproteins of H5 influenza viruses isolated from humans, chickens, and wild aquatic birds have distinguishable properties." J Virol **73**(2): 1146-1155.

Matrosovich, M. N., T. Y. Matrosovich, T. Gray, N. A. Roberts and H. D. Klenk (2004). "Human and avian influenza viruses target different cell types in cultures of human airway epithelium." Proc Natl Acad Sci U S A **101**(13): 4620-4624.

McCauley, J., J. Bye, K. Elder, M. J. Gething, J. J. Skehel, A. Smith and M. D. Waterfield (1979). "Influenza virus haemagglutinin signal sequence." FEBS Lett **108**(2): 422-426.

McCoy, A. J., R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni and R. J. Read (2007). "Phaser crystallographic software." J Appl Crystallogr **40**(Pt 4): 658-674.

McGeoch, D., P. Fellner and C. Newton (1976). "Influenza virus genome consists of eight distinct RNA species." Proc Natl Acad Sci U S A **73**(9): 3045-3049.

Mir-Shekari, S. Y., D. A. Ashford, D. J. Harvey, R. A. Dwek and I. T. Schulze (1997). "The glycosylation of the influenza A virus hemagglutinin by mammalian cells. A site-specific study." J Biol Chem **272**(7): 4027-4036.

Montalto, N. J., K. D. Gum and J. V. Ashley (2000). "Updated treatment for influenza A and B." Am Fam Physician **62**(11): 2467-2476.

Mozdzanowska, K., J. Feng, M. Eid, D. Zharikova and W. Gerhard (2006). "Enhancement of neutralizing activity of influenza virus-specific antibodies by serum components." Virology **352**(2): 418-426.

Mozdzanowska, K., M. Furchner, G. Washko, J. Mozdzanowski and W. Gerhard (1997). "A pulmonary influenza virus infection in SCID mice can be cured by treatment with hemagglutinin-specific antibodies that display very low virus-neutralizing activity in vitro." J Virol **71**(6): 4347-4355.

Murti, K. G. and R. G. Webster (1986). "Distribution of hemagglutinin and neuraminidase on influenza virions as revealed by immunoelectron microscopy." Virology **149**(1): 36-43.

Ni, F., X. Chen, J. Shen and Q. Wang (2014). "Structural insights into the membrane fusion mechanism mediated by influenza virus hemagglutinin." Biochemistry **53**(5): 846-854.

Nobusawa, E., T. Aoyama, H. Kato, Y. Suzuki, Y. Tateno and K. Nakajima (1991). "Comparison of complete amino acid sequences and receptor-binding properties among 13 serotypes of hemagglutinins of influenza A viruses." Virology **182**(2): 475-485.

Nobusawa, E. and K. Sato (2006). "Comparison of the mutation rates of human influenza A and B viruses." J Virol **80**(7): 3675-3678.

O'Neill, R. E., J. Talon and P. Palese (1998). "The influenza virus NEP (NS2 protein) mediates the nuclear export of viral ribonucleoproteins." EMBO J **17**(1): 288-296.

Osterhaus, A. D., G. F. Rimmelzwaan, B. E. Martina, T. M. Bestebroer and R. A. Fouchier (2000). "Influenza B virus in seals." Science **288**(5468): 1051-1053.

Outlaw, M. C. and N. J. Dimmock (1991). "Insights into neutralization of animal viruses gained from study of influenza virus." Epidemiol Infect **106**(2): 205-220.

Palese, P., K. Tobita, M. Ueda and R. W. Compans (1974). "Characterization of temperature sensitive influenza virus mutants defective in neuraminidase." Virology **61**(2): 397-410.

Perdue, M. L., M. Garcia, D. Senne and M. Fraire (1997). "Virulence-associated sequence duplication at the hemagglutinin cleavage site of avian influenza viruses." Virus Res **49**(2): 173-186.

Pinto, L. H., L. J. Holsinger and R. A. Lamb (1992). "Influenza virus M2 protein has ion channel activity." Cell **69**(3): 517-528.

Popova, L., K. Smith, A. H. West, P. C. Wilson, J. A. James, L. F. Thompson and G. M. Air (2012). "Immunodominance of antigenic site B over site A of hemagglutinin of recent H3N2 influenza viruses." PLoS One **7**(7): e41895.

Porath, J., J. Carlsson, I. Olsson and G. Belfrage (1975). "Metal chelate affinity chromatography, a new approach to protein fractionation." Nature **258**(5536): 598-599.

Qiao, H., R. T. Armstrong, G. B. Melikyan, F. S. Cohen and J. M. White (1999). "A specific point mutant at position 1 of the influenza hemagglutinin fusion peptide displays a hemifusion phenotype." Mol Biol Cell **10**(8): 2759-2769.

Reid, A. H., T. G. Fanning, J. V. Hultin and J. K. Taubenberger (1999). "Origin and evolution of the 1918 "Spanish" influenza virus hemagglutinin gene." Proc Natl Acad Sci U S A **96**(4): 1651-1656.

Rogers, G. N. and B. L. D'Souza (1989). "Receptor binding properties of human and animal H1 influenza virus isolates." Virology **173**(1): 317-322.

Rogers, G. N. and J. C. Paulson (1983). "Receptor determinants of human and animal influenza virus isolates: differences in receptor specificity of the H3 hemagglutinin based on species of origin." Virology **127**(2): 361-373.

Rogers, G. N., J. C. Paulson, R. S. Daniels, J. J. Skehel, I. A. Wilson and D. C. Wiley (1983). "Single amino acid substitutions in influenza haemagglutinin change receptor binding specificity." Nature **304**(5921): 76-78.

Rohm, C., N. Zhou, J. Suss, J. Mackenzie and R. G. Webster (1996). "Characterization of a novel influenza hemagglutinin, H15: criteria for determination of influenza A subtypes." Virology **217**(2): 508-516.

Rosenthal, P. B., X. Zhang, F. Formanowski, W. Fitz, C. H. Wong, H. Meier-Ewert, J. J. Skehel and D. C. Wiley (1998). "Structure of the haemagglutinin-esterase-fusion glycoprotein of influenza C virus." Nature **396**(6706): 92-96.

Rossman, J. S., G. P. Leser and R. A. Lamb (2012). "Filamentous influenza virus enters cells via macropinocytosis." J Virol **86**(20): 10950-10960.

Rota, P. A., T. R. Wallis, M. W. Harmon, J. S. Rota, A. P. Kendal and K. Nerome (1990). "Cocirculation of two distinct evolutionary lineages of influenza type B virus since 1983." Virology **175**(1): 59-68.

Ruigrok, R. W. (1998). Structure of influenza A, B and C viruses. Textbook of Influenza. K. G. Nicholson, R. G. Webster and A. J. Hay, Blackwell Science, Oxford**:** 29-42.

Ruigrok, R. W., A. Aitken, L. J. Calder, S. R. Martin, J. J. Skehel, S. A. Wharton, W. Weis and D. C. Wiley (1988). "Studies on the structure of the influenza virus haemagglutinin at the pH of membrane fusion." J Gen Virol **69 (Pt 11)**: 2785-2795.

Ruigrok, R. W., A. Barge, P. Durrer, J. Brunner, K. Ma and G. R. Whittaker (2000). "Membrane interaction of influenza virus M1 protein." Virology **267**(2): 289-298.

Ruigrok, R. W., S. R. Martin, S. A. Wharton, J. J. Skehel, P. M. Bayley and D. C. Wiley (1986). "Conformational changes in the hemagglutinin of influenza virus which accompany heat-induced fusion of virus with liposomes." Virology **155**(2): 484-497.

Ruigrok, R. W., N. G. Wrigley, L. J. Calder, S. Cusack, S. A. Wharton, E. B. Brown and J. J. Skehel (1986). "Electron microscopy of the low pH structure of influenza virus haemagglutinin." EMBO J **5**(1): 41-49.

Russell, R. J., S. J. Gamblin, L. F. Haire, D. J. Stevens, B. Xiao, Y. Ha and J. J. Skehel (2004). "H1 and H7 influenza haemagglutinin structures extend a structural classification of haemagglutinin subtypes." Virology **325**(2): 287-296.

Sauter, N. K., M. D. Bednarski, B. A. Wurzburg, J. E. Hanson, G. M. Whitesides, J. J. Skehel and D. C. Wiley (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**(21): 8388-8396.

Sauter, N. K., J. E. Hanson, G. D. Glick, J. H. Brown, R. L. Crowther, S. J. Park, J. J. Skehel and D. C. Wiley (1992). "Binding of influenza virus hemagglutinin to analogs of its cell-surface receptor, sialic acid: analysis by proton nuclear magnetic resonance spectroscopy and X-ray crystallography." Biochemistry **31**(40): 9609-9621.

Schild, G. C., R. W. Newman, R. G. Webster, D. Major and V. S. Hinshaw (1980). "Antigenic analysis of influenza A virus surface antigens: considerations for the nomenclature of influenza virus." Comp Immunol Microbiol Infect Dis **3**(1-2): 5-18.

Scholtissek, C., W. Rohde, V. Von Hoyningen and R. Rott (1978). "On the origin of the human influenza virus subtypes H2N2 and H3N2." Virology **87**(1): 13-20.

Schulze, I. T. (1970). "The structure of influenza virus. I. The polypeptides of the virion." Virology **42**(4): 890-904.

Schulze, I. T. (1997). "Effects of glycosylation on the properties and functions of influenza virus hemagglutinin." J Infect Dis **176 Suppl 1**: S24-28.

Sharp, D. G., Taylor, A. R., McLean, I. W., Beard D., Beard, W. (1945). "Densities and sizes of the influenza viruses A (PR8 strain) and B (Lee strain) and the swine influenza virus." J. Biol. Chem. **159**: 29-44.

Shope, R. E. (1931). "Swine Influenza : Iii. Filtration Experiments and Etiology." J Exp Med **54**(3): 373-385.

Shortridge, K. F., P. Gao, Y. Guan, T. Ito, Y. Kawaoka, D. Markwell, A. Takada and R. G. Webster (2000). "Interspecies transmission of influenza viruses: H5N1 virus and a Hong Kong SAR perspective." Vet Microbiol **74**(1-2): 141-147.

Simonsen, L., M. J. Clarke, L. B. Schonberger, N. H. Arden, N. J. Cox and K. Fukuda (1998). "Pandemic versus epidemic influenza mortality: a pattern of changing age distribution." J Infect Dis **178**(1): 53-60.

Simossis, V. A. and J. Heringa (2003). "The PRALINE online server: optimising progressive multiple alignment on the web." Comput Biol Chem **27**(4-5): 511-519.

Simossis, V. A. and J. Heringa (2005). "PRALINE: a multiple sequence alignment toolbox that integrates homology-extended and secondary structure information." Nucleic Acids Res **33**(Web Server issue): W289-294.

Skehel, J. (2009). "An overview of influenza haemagglutinin and neuraminidase." Biologicals **37**(3): 177-178.

Skehel, J. J., P. M. Bayley, E. B. Brown, S. R. Martin, M. D. Waterfield, J. M. White, I. A. Wilson and D. C. Wiley (1982). "Changes in the conformation of influenza virus hemagglutinin at the pH optimum of virus-mediated membrane fusion." Proc Natl Acad Sci U S A **79**(4): 968-972.

Skehel, J. J., R. S. Daniels, A. R. Douglas and D. C. Wiley (1983). "Antigenic and amino acid sequence variations in the haemagglutinins of type A influenza viruses recently isolated from human subjects." Bull World Health Organ **61**(4): 671-676.

Skehel, J. J., R. S. Daniels, A. J. Hay, R. Ruigrok, S. A. Wharton, N. G. Wrigley, W. Weiss and D. C. Willey (1986). "Structural changes in influenza virus haemagglutinin at the pH of membrane fusion." Biochem Soc Trans **14**(2): 252-253.

Skehel, J. J. and G. C. Schild (1971). "The polypeptide composition of influenza A viruses." Virology **44**(2): 396-408.

Skehel, J. J., D. J. Stevens, R. S. Daniels, A. R. Douglas, M. Knossow, I. A. Wilson and D. C. Wiley (1984). "A carbohydrate side chain on hemagglutinins of Hong Kong influenza

viruses inhibits recognition by a monoclonal antibody." Proc Natl Acad Sci U S A **81**(6): 1779-1783.

Skehel, J. J. and M. D. Waterfield (1975). "Studies on the primary structure of the influenza virus hemagglutinin." Proc Natl Acad Sci U S A **72**(1): 93-97.

Skehel, J. J., M. D. Waterfield, J. W. McCauley, K. Elder and D. C. Wiley (1980). "Studies on the structure of the haemagglutinin." Philos Trans R Soc Lond B Biol Sci **288**(1029): 335-339.

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

Smith, G. J., X. H. Fan, J. Wang, K. S. Li, K. Qin, J. X. Zhang, D. Vijaykrishna, C. L. Cheung, K. Huang, J. M. Rayner, J. S. Peiris, H. Chen, R. G. Webster and Y. Guan (2006). "Emergence and predominance of an H5N1 influenza variant in China." Proc Natl Acad Sci U S A **103**(45): 16936-16941.

Smith, G. J., D. Vijaykrishna, J. Bahl, S. J. Lycett, M. Worobey, O. G. Pybus, S. K. Ma, C. L. Cheung, J. Raghwani, S. Bhatt, J. S. Peiris, Y. Guan and A. Rambaut (2009). "Origins and evolutionary genomics of the 2009 swine-origin H1N1 influenza A epidemic." Nature **459**(7250): 1122-1125.

Smith, W., Andrewes, C.H. & Laidlaw, P.P. (1933). "A virus obtained from influenza patients." Lancet **ii**(66).

Sreerama, N., S. Y. Venyaminov and R. W. Woody (1999). "Estimation of the number of alpha-helical and beta-strand segments in proteins using circular dichroism spectroscopy." Protein Sci **8**(2): 370-380.

Sreerama, N. and R. W. Woody (2000). "Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set." Anal Biochem **287**(2): 252-260.

Stegmann, T., D. Hoekstra, G. Scherphof and J. Wilschut (1985). "Kinetics of pH-dependent fusion between influenza virus and liposomes." Biochemistry **24**(13): 3107-3113.

Steinhauer, D. A. (1999). "Role of hemagglutinin cleavage for the pathogenicity of influenza virus." Virology **258**(1): 1-20.

Stevens, J., O. Blixt, T. M. Tumpey, J. K. Taubenberger, J. C. Paulson and I. A. Wilson (2006). "Structure and receptor specificity of the hemagglutinin from an H5N1 influenza virus." Science **312**(5772): 404-410.

Stieneke-Grober, A., M. Vey, H. Angliker, E. Shaw, G. Thomas, C. Roberts, H. D. Klenk and W. Garten (1992). "Influenza virus hemagglutinin with multibasic cleavage site is activated by furin, a subtilisin-like endoprotease." EMBO J **11**(7): 2407-2414.

Sugrue, R. J. and A. J. Hay (1991). "Structural characteristics of the M2 protein of influenza A viruses: evidence that it forms a tetrameric channel." Virology **180**(2): 617-624.

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

R. O. Donis, R. C. Liddington and W. A. Marasco (2009). "Structural and functional bases for broad-spectrum neutralization of avian and human influenza A viruses." Nat Struct Mol Biol **16**(3): 265-273.

Sullivan, K., J. Kloess, C. Qian, D. Bell, A. Hay, Y. P. Lin and Y. Gu (2012). "High throughput virus plaque quantitation using a flatbed scanner." J Virol Methods **179**(1): 81-89.

Taylor, R. M. (1949). "Studies on survival of influenza virus between epidemics and antigenic variants of the virus." Am J Public Health Nations Health **39**(2): 171-178.

To, K. K., A. K. Tsang, J. F. Chan, V. C. Cheng, H. Chen and K. Y. Yuen (2014). "Emergence in China of human disease due to avian influenza A(H10N8)--cause for concern?" J Infect **68**(3): 205-215.

Tong, S., Y. Li, P. Rivailler, C. Conrardy, D. A. Castillo, L. M. Chen, S. Recuenco, J. A. Ellison, C. T. Davis, I. A. York, A. S. Turmelle, D. Moran, S. Rogers, M. Shi, Y. Tao, M. R. Weil, K. Tang, L. A. Rowe, S. Sammons, X. Xu, M. Frace, K. A. Lindblade, N. J. Cox, L. J. Anderson, C. E. Rupprecht and R. O. Donis (2012). "A distinct lineage of influenza A virus from bats." Proc Natl Acad Sci U S A **109**(11): 4269-4274.

Tong, S., X. Zhu, Y. Li, M. Shi, J. Zhang, M. Bourgeois, H. Yang, X. Chen, S. Recuenco, J. Gomez, L. M. Chen, A. Johnson, Y. Tao, C. Dreyfus, W. Yu, R. McBride, P. J. Carney, A. T. Gilbert, J. Chang, Z. Guo, C. T. Davis, J. C. Paulson, J. Stevens, C. E. Rupprecht, E. C. Holmes, I. A. Wilson and R. O. Donis (2013). "New world bats harbor diverse influenza A viruses." PLoS Pathog **9**(10): e1003657.

Vachieri, S. G., X. Xiong, P. J. Collins, P. A. Walker, S. R. Martin, L. F. Haire, Y. Zhang, J. W. McCauley, S. J. Gamblin and J. J. Skehel (2014). "Receptor binding by H10 influenza viruses." Nature **511**(7510): 475-477.

Vagin, A. and A. Teplyakov (2010). "Molecular replacement with MOLREP." Acta Crystallogr D Biol Crystallogr **66**(Pt 1): 22-25.

van Riel, D., V. J. Munster, E. de Wit, G. F. Rimmelzwaan, R. A. Fouchier, A. D. Osterhaus and T. Kuiken (2006). "H5N1 Virus Attachment to Lower Respiratory Tract." Science **312**(5772): 399.

van Stokkum, I. H., H. J. Spoelder, M. Bloemendal, R. van Grondelle and F. C. Groen (1990). "Estimation of protein secondary structure and error analysis from circular dichroism spectra." Anal Biochem **191**(1): 110-118.

Ward, P., I. Small, J. Smith, P. Suter and R. Dutkowski (2005). "Oseltamivir (Tamiflu) and its potential for use in the event of an influenza pandemic." J Antimicrob Chemother **55 Suppl 1**: i5-i21.

Waterfield, M. D., K. Espelie, K. Elder and J. J. Skehel (1979). "Structure of the haemagglutinin of influenza virus." Br Med Bull **35**(1): 57-63.

Webster, R. G., W. J. Bean, O. T. Gorman, T. M. Chambers and Y. Kawaoka (1992). "Evolution and ecology of influenza A viruses." Microbiol Rev **56**(1): 152-179.

Webster, R. G. and R. Rott (1987). "Influenza virus A pathogenicity: the pivotal role of hemagglutinin." Cell **50**(5): 665-666.

Weis, W., J. H. Brown, S. Cusack, J. C. Paulson, J. J. Skehel and D. C. Wiley (1988). "Structure of the influenza virus haemagglutinin complexed with its receptor, sialic acid." Nature **333**(6172): 426-431.

Wharton, S. A. (1987). "The role of influenza virus haemagglutinin in membrane fusion." Microbiol Sci **4**(4): 119-124.

Wharton, S. A., L. J. Calder, R. W. Ruigrok, J. J. Skehel, D. A. Steinhauer and D. C. Wiley (1995). "Electron microscopy of antibody complexes of influenza virus haemagglutinin in the fusion pH conformation." EMBO J **14**(2): 240-246.

Wharton, S. A., J. J. Skehel and D. C. Wiley (1986). "Studies of influenza haemagglutinin-mediated membrane fusion." Virology **149**(1): 27-35.

Whittle, J. R., R. Zhang, S. Khurana, L. R. King, J. Manischewitz, H. Golding, P. R. Dormitzer, B. F. Haynes, E. B. Walter, M. A. Moody, T. B. Kepler, H. X. Liao and S. C. Harrison (2011). "Broadly neutralizing human antibody that recognizes the receptor-binding pocket of influenza virus hemagglutinin." Proc Natl Acad Sci U S A **108**(34): 14216-14221.

Wiley, D. C. and J. J. Skehel (1977). "Crystallization and x-ray diffraction studies on the haemagglutinin glycoprotein from the membrane of influenza virus." J Mol Biol **112**(2): 343-347.

Wiley, D. C. and J. J. Skehel (1987). "The structure and function of the hemagglutinin membrane glycoprotein of influenza virus." Annu Rev Biochem **56**: 365-394.

Wiley, D. C., J. J. Skehel and M. Waterfield (1977). "Evidence from studies with a cross-linking reagent that the haemagglutinin of influenza virus is a trimer." Virology **79**(2): 446-448.

Wiley, D. C., I. A. Wilson and J. J. Skehel (1981). "Structural identification of the antibody-binding sites of Hong Kong influenza haemagglutinin and their involvement in antigenic variation." Nature **289**(5796): 373-378.

Wilson, I. A. and N. J. Cox (1990). "Structural basis of immune recognition of influenza virus hemagglutinin." Annu Rev Immunol **8**: 737-771.

Wilson, I. A., J. J. Skehel and D. C. Wiley (1981). "Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 A resolution." Nature **289**(5796): 366-373.

Winn, M. D. (2003). "An overview of the CCP4 project in protein crystallography: an example of a collaborative project." J Synchrotron Radiat **10**(Pt 1): 23-25.

Winn, M. D., C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A. Vagin and K. S. Wilson (2011). "Overview of the CCP4 suite and current developments." Acta Crystallogr D Biol Crystallogr **67**(Pt 4): 235-242.

Winter, G., C. M. Lobley and S. M. Prince (2013). "Decision making in xia2." Acta Crystallogr D Biol Crystallogr **69**(Pt 7): 1260-1273.

Wrigley, N. G., E. B. Brown, R. S. Daniels, A. R. Douglas, J. J. Skehel and D. C. Wiley (1983). "Electron microscopy of influenza haemagglutinin-monoclonal antibody complexes." Virology **131**(2): 308-314.

Wrigley, N. G., W. G. Laver and J. C. Downie (1977). "Binding of antibodies to isolated haemagglutinin and neuraminidase molecules of influenza virus observed in the electron microscope." J Mol Biol **109**(3): 405-421.

Xiong, X., P. J. Coombs, S. R. Martin, J. Liu, H. Xiao, J. W. McCauley, K. Locher, P. A. Walker, P. J. Collins, Y. Kawaoka, J. J. Skehel and S. J. Gamblin (2013). "Receptor binding by a ferret-transmissible H5 avian influenza virus." Nature **497**(7449): 392-396.

Xiong, X., H. Xiao, S. R. Martin, P. J. Coombs, J. Liu, P. J. Collins, S. G. Vachieri, P. A. Walker, Y. P. Lin, J. W. McCauley, S. J. Gamblin and J. J. Skehel (2014). "Enhanced human receptor binding by H5 haemagglutinins." Virology **456-457**: 179-187.

Yewdell, J. W., R. G. Webster and W. U. Gerhard (1979). "Antigenic variation in three distinct determinants of an influenza type A haemagglutinin molecule." Nature **279**(5710): 246-248.

Yu, X., T. Tsibane, P. A. McGraw, F. S. House, C. J. Keefer, M. D. Hicar, T. M. Tumpey, C. Pappas, L. A. Perrone, O. Martinez, J. Stevens, I. A. Wilson, P. V. Aguilar, E. L. Altschuler, C. F. Basler and J. E. Crowe, Jr. (2008). "Neutralizing antibodies derived from the B cells of 1918 influenza pandemic survivors." Nature **455**(7212): 532-536.

Zamarin, D., M. B. Ortigoza and P. Palese (2006). "Influenza A virus PB1-F2 protein contributes to viral pathogenesis in mice." J Virol **80**(16): 7976-7983.

Zambon, M. C. (1999). "Epidemiology and pathogenesis of influenza." J Antimicrob Chemother **44 Suppl B**: 3-9.

Zebedee, S. L. and R. A. Lamb (1988). "Influenza A virus M2 protein: monoclonal antibody restriction of virus growth and detection of M2 in virions." J Virol **62**(8): 2762-2772.

Zhang, W., Y. Shi, X. Lu, Y. Shu, J. Qi and G. F. Gao (2013). "An airborne transmissible avian influenza H5 hemagglutinin seen at the atomic level." Science **340**(6139): 1463-1467.