Biophysical Characterisation of the Involvement of Influenza A Virus Glycoproteins in Receptor Binding

Donald James Benton

July 2015

Division of Virology
MRC National Institute for Medical Research
The Ridgeway
Mill Hill
London
NW7 1AA

A thesis submitted to University College London for the degree of Doctor of Philosophy

I, Donald Benton confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Donald Benton

Date:

Acknowledgements

Firstly I would like to thank my supervisor Dr. John McCauley for taking me on as an inexperienced undergraduate, for helping me with his extensive knowledge of the influenza field and for allowing me to pursue a wide range of scientific interests while at the NIMR. I owe an enormous debt of gratitude to the Mill Hill institution that is the two Steves. Dr. Steve Martin, 'The Mill Hill Biophysics Oracle', for his immense knowledge and experience of all things related to biophysical techniques, the design of meaningful experiments and more importantly the interpretation of the results and also for putting up with me spending large amounts of time in his lab and teaching me so much about all things biophysical. Also Dr. Steve Wharton, otherwise known as 'The Whartonator' or 'The Wharton Movement', for passing on his wealth of experimental and theoretical knowledge of all things related to 'flu viruses and for teaching me the 'Wharton Methods' of virology, most of which include using a crusty piece of equipment from the 1980s or holding up a tube of a solution to the light and calibrating your eyes to determine the concentration of some substance or other. These techniques do, however, all work excellently. Without the two Steves I would have floundered a lot more experimentally and also my experimental interpretation would have been much diminished. Their combined >70 years working at NIMR and the experience accrued throughout this time and their willingness to help has been enormously important to me and many previous generations of PhD students.

I would like to thank the McCauley lab past and present: Dr. Haxia Xiao, Dr. Saira Hussain, Dr. Lauren Parker, Dr. Kerstin Beer, Dr. Steve Wharton and Michael Bennett for practical support, advice and helpful discussions. I would also like to

thank all of the lab members for making lab 274 such a fun place to work, with copious amounts of general hilarity.

I would like to thank Dr. Patrick Collins and Dr. Xiaoli (Alex) Xiong for a large amount of advice on protein expression and purification. I would also like to thank: Dr. Yi Pu Lin for providing me with plasmids and advice on reverse genetics; Dr. Rod Daniels for his excellent primer sequences and advice on viral RNA sequencing; the NIMR Large Scale Lab for help with protein expression; and all the members of the WHO Collaborating Centre for Research and Reference on Influenza at Mill Hill for the provision of viruses, reagents and advice.

I would like to thank the whole of the Division of Virology at NIMR for making it a fun and welcoming place to work, with neighbouring labs always willing to help with any sort of problem. I have had many a memorable evening on virology pub crawls, particularly on the three 'Twelve Pubs of Christmas', and would like to thank the pub crawl regulars: Steve Wharton, Patrick Collins, Nick Cattle, Darren Wight, Lauren Parker, Lesley Calder, Harriet Groom, Paula Ordóñez Suárez, Vicky Gregory, Marta Sanz-Ramos, Callum Donaldson and many others, too numerous to name.

Last but by no means least I would like to thank my family: my parents Mike and Mary for encouraging me into a career that involves never leaving university and to my girlfriend Georgie for her love and support throughout my PhD.

Abstract

Influenza A viruses (IAV) cause human infections in both seasonal and pandemic forms, which are responsible for significant levels of human infections. A major factor underlying IAV viral fitness and transmissibility, and therefore pandemic potential, is the interaction of the virus with the cell surface, which is mediated by the two surface glycoproteins haemagglutinin (HA) and neuraminidase (NA). These two glycoproteins have antagonistic activities with the HA responsible for receptor binding and the NA for receptor destroying activities. It has long been hypothesised that the balance between these antagonistic functions is important in altering virus fitness and transmissibility.

This thesis presents a new biophysical approach, based on biolayer interferometry, for determining the relative importance of the HA and NA in altering virus binding to a receptor coated surface. These experiments identified a number of fundamental characteristics of virus attachment and interaction with surfaces. This new technique has also been used to examine two other lineages of viruses that are at different stages of becoming transmissible in humans. H7N9 viruses first infected humans in 2013 and have caused a large number of zoonotic infections. The receptor binding characteristics of these H7N9 viruses are characterised with a particular focus on the NA, as N9 NAs have been previously shown to have receptor binding properties. In 2009 a new H1N1 pandemic emerged, which was formed from a reassortment of two swine viruses. These new viruses had acquired a number of substitutions in the HA, which could affect receptor binding, also the combination of HA and NA was also novel. A number of HA mutants and combinations of HA and NA were examined in terms of receptor binding in an attempt to understand what changes in swine viruses might make them transmissible in humans.

Table of Contents

List of Figures

List of Tables

Abbreviations

293T	Human embryonic kidney 293T cells
3SL-PAA	α2,3-sialyllactose sialoglycopolymer
3SLN	α2,3-sialyl-N-acetyllactosamine
3SLN-PAA	α2,3-sialyl-N-acetyllactosamine sialoglycopolymer
6SL-PAA	α2,6-sialyllactose sialoglycopolymer
6SLN	α2,6-sialyl-N-acetyllactosamine
6SLN-PAA	α2,6-sialyl-N-acetyllactosamine sialoglycopolymer
6SU-3SLN-PAA	α2,3-sialyl-N-acetyllactosamine-6' sulphated GlcNAc sialoglycopolymer
6SU-Sia Lex-PAA	Sialyl Lewis^{x}-6' sulphated GlcNAc sialoglycopolymer
Anhui13	A/Anhui/1/2013(H7N9)
AU	Arbitrary Units
BHA	Bromelain cleaved HA
BLI	Biolayer Interferometry
Brij-36T	Decaethylene glycol monododecyl ether
BSA	Bovine Serum Albumin
Cal4	A/California/4/2009(H1N1)
Cal7	A/California/7/2009(H1N1)
cDNA	Complementary DNA
CPNA	*Clostridium perfringens* neuraminidase
cRNA	Complementary RNA
csw	Classical swine
CV	Column Volume
dH_2O	Distilled water
DMEM	Dulbecco's Modified Eagle's Medium
DNA	Deoxyribonucleic Acid
dNTP	Deoxynucleoside Triphosphate
ECACC	European Collection of Cell Cultures
EDTA	Ethylenediaminetetraacetic acid
EHA2	*E. coli* Expressed HA2
ELISA	Enzyme-linked Immunosorbent Assay
f	Fractional saturation
FCS	Foetal Calf Serum

FPLC	Fast Protein Liquid Chromatography
GISAID	Global Initiative for Sharing All Influenza Data
GPI	Glycophosphatidylinositol
GlcNAc	N-Acetylglucosamine
HA	Haemagglutinin
HA0	Uncleaved HA precursor
HA1	N-terminal section of cleaved HA
HA2	C-terminal section of cleaved HA
Hb	Haemadsorption
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HPAI	Highly Pathogenic Avian Influenza
HRP	Horseradish Peroxidase
IAV	Influenza A Virus
Int	Fluorescence intensity (AU)
Inhib	NA inhibitors
Iowa06	A/Iowa/1/2006(H1N1)
IPTG	Isopropyl β-D-1-thiogalactopyranoside
k_{cat}	Enzyme turnover rate
K_d	Equilibrium dissociation constant
K_i	Equilibrium dissociation constant of enzyme inhibitor
K_m	Michaelis-Menten constant
k_{obs}	Observed rate constant
k_{off}	Dissociation rate constant
k_{on}	Association rate constant
LDH	Lactate Dehydrogenase
LED	Light Emitting Diode
LRT	Lower Respiratory Tract
LSTa	LS-Tetrasaccharide a
LSTc	LS-Tetrasaccharide c
M	Matrix
M1	Matrix protein 1
M2	Matrix protein 2
mAb	Monoclonal Antibody
mc	Multiplicity coefficient
MDBK	Madin-Derby Bovine Kidney

MDCK	Madin-Derby Canine Kidney
MDCK-SIAT	MDCK cell line stably transfected with plasmid expressing 2,6 sialyltransferase
MES	2-(N-morpholino)ethanesulfonic acid
MPK	Minipig Kidney
MRC NIMR	Medical Research Council National Institute for Medical Research (now known as the Francis Crick Institute, Mill Hill Laboratory)
mRNA	Messenger RNA
MST	Microscale Thermophoresis
MU	4-Methylumbelliferone
MUNANA	2'-(4-Methylumbelliferyl)-α-D-N-acetylneuraminic acid
n	Hill coefficient
NA	Neuraminidase
NADH	β-Nicotinamide adenine dinucleotide
NANA Aldolase	N-acetyl-neuramininic acid aldolase
NEP	Nuclear export protein
NIMR WHOCC	World Health Organization Collaborating Centre for Reference and Research on Influenza, Medical Research Council National Institute for Medical Research (now known as Francis Crick Institute Worldwide Influenza Centre)
NMR	Nuclear Magnetic Resonance
NP	Nucleoprotein
NS1	Non-Structural protein 1
NS2	Non-Structural protein 2
NTA	Nanoparticle Tracking Analysis
PA	Polymerase Acidic
PB1	Polymerase Basic 1
PB2	Polymerase Basic 2
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank
pdm09	2009 H1N1 pandemic
Pen/Strep	Penicillin and Streptomycin
PR8	A/Puerto Rico/8/34(H1N1)
RBS	Receptor Binding Site

RNA	Ribonucleic Acid
RNP	Ribonucleoprotein
RSL	Relative Sugar Loading
$RSL_{0.5}$	Relative Sugar Loading when fractional saturation = 0.5
RT-PCR	Reverse Transcription Polymerase Chain Reaction
SDS	Sodium Dodecylsulphate
SDS-PAGE	Sodium Dodecylsulphate Polyacrylamide Gel Electrophoresis
Sf9	Spodoptera frugiperda cell line
Sia Lex-PAA	Sialyl $Lewis^x$ sialoglycopolymer
SOC	Super Optimal broth with Catabolite repression
SPR	Surface Plasmon Resonance
ssRNA	Single Stranded RNA
TBHA2	BHA treated with low pH and subjected to trypsin and thermolysin digestion
TEV	Tobacco Etch Virus
TPCK	Tosyl Phenylalanyl Chloromethyl Ketone
TRBCs	Turkey Red Blood Cells
Tris	Tris(hydroxymethyl)aminomethane
TRS	Triple Reassortant Swine
URT	Upper Respiratory Tract
V_0	Initial enzymatic reaction rate
VASP	Vasodilator-Stimulation Phosphoprotein
v_i	Inhibited enzymatic reaction rate
V_{max}	Maximal reaction velocity
Vic361	A/Victoria/361/2012(H3N2)
vRNA	Viral RNA
WHO	World Health Organization
WSN	A/WSN/33(H1N1)
WT	Wild Type
X-31	High growth reassortant containing A/Aichi/2/68(H3N2) HA and NA with a PR8 background
X-31 HAM	X-31 horse adsorption mutant
X-31 NA Δ10	X-31 with residues 60 - 69 deleted from NA

1 Introduction

1.1 Influenza viruses

1.1.1 Classification

Influenza viruses are enveloped viruses with a segmented single-stranded negative-sense RNA genome. They are part of the family *Orthomyxoviridae*. There are 3 genera of Influenza viruses A, B and C, making up 3 of the 6 genera of *Orthomyxoviridae*. These are classified by antigenic differences between the M and NP proteins (King, 2011).

Influenza A, B and C viruses infect humans; however, only influenza A viruses (IAVs) have been responsible for causing pandemics. IAVs are capable of causing pandemics because of their wide host range, allowing between-host transmission events, which can allow antigenically novel viruses to infect an immunologically naïve population. Because of the pandemic potential of IAVs they are the focus of this thesis.

IAVs are classified into different subtypes by the antigenic properties of their surface glycoproteins haemagglutinin (HA) and neuraminidase (NA) (Gamblin & Skehel, 2010). There are currently 18 HA subtypes (H1 – H18) and 11 NA subtypes (N1 – N11) described. H1, H2 and H3 HA subtype-containing viruses have caused human transmissible infections and were responsible for causing all of the pandemics of the 20th and 21st century: 1918 H1N1, 1957 H2N2, 1968 H3N2 and 2009 H1N1 (Klenk, Garten & Matrosovich, 2011).

1.1.2 Virus reservoirs

The natural reservoir of most IAVs is held to be wild waterfowl, as all H1-H16 subtypes have been found in wild birds (Kawaoka *et al.*, 1988). Within these wild waterfowl the main hosts are from the orders *Anseriformes* (ducks, geese and swans) and *Charadriiformes* (gulls and shorebirds) (Olsen *et al.*, 2006; Vandegrift *et al.*, 2010; Yoon, Webby & Webster, 2014). This reservoir in wild birds has been found to infect either directly or indirectly, via an intermediate host, a wide range of different birds and mammals including humans, domestic poultry, pigs, horses, dogs, cats, seals and whales (Yoon, Webby & Webster, 2014).

The two most recently described HA subtypes, H17 and H18, and NA subtypes, N10 and N11, are only found in recently discovered influenza viruses found in bats. The H17N10 virus (Tong *et al.*, 2012) and H18N11 (Tong *et al.*, 2013) viruses were found by sequence analysis of bat samples, without isolation of the virus. Structural studies and binding analyses also showed unusual properties of the HA and NA of these viruses, with the HA showing no binding to sialic acid (Zhu *et al.*, 2013b; Sun *et al.*, 2013a) and the NA showing no sialidase activity (Zhu *et al.*, 2012b). After full synthesis of the genome the virus could not be generated by reverse genetics; however, when replacing the open reading frame of HA and NA, with those of the lab strain PR8, replicating virus could be generated (Zhou *et al.*, 2014).

1.1.3 The disease

Influenza is a highly transmissible respiratory pathogen, which in humans causes a significant disease burden. All influenza types A, B and C can infect humans although influenza C is held to cause a very low number of minor infections (Palese & Shaw, 2007). Seasonal human infections have been estimated to account for 3 – 5

million cases of influenza related illness annually, resulting in 250 000 – 500 000 deaths (World Health Organisation, 2014a).

A typical influenza infection, in humans, manifests itself as an acute respiratory infection, with symptoms such as high fever, cough, headache, weakness, myalgia and general malaise, with acute symptoms generally lasting in the order of 5 days (Taubenberger & Morens, 2008). In some severe cases the disease can develop into primary pneumonia directly from the viral infection or, in others, secondary bacterial pneumonia (Chertow & Memoli, 2013; Smith & McCullers, 2014). The transmission of IAVs between humans can occur via a number of different routes: aerosol, large droplets or via direct contact with substances carrying the virus such as secretions or bits of dust etc. (fomites), reviewed in (Tellier, 2006). Humans are thought to be contagious from around one day before showing symptoms up to 5 – 7 days after developing symptoms.

1.1.4 Virus structure

1.1.4.1 The virion

As described above, IAVs are enveloped viruses. The morphology of these viruses is variable, with some adopting spherical or elliptical morphology with particles 100 – 150 nm in diameter, whereas other viruses can form filamentous particles, stretching to over 1 μm in length (Calder *et al.*, 2010). Figure 1.1 shows a pictorial representation of an influenza virus particle. Figure 1.2 shows a cryo electron micrograph of a typical 'spherical' virus strain, X-31. Lab-adapted egg-propagated IAV strains are typically spherical, however clinical isolates and tissue culture-propagated viruses are more likely to be filamentous in nature and adaptation of

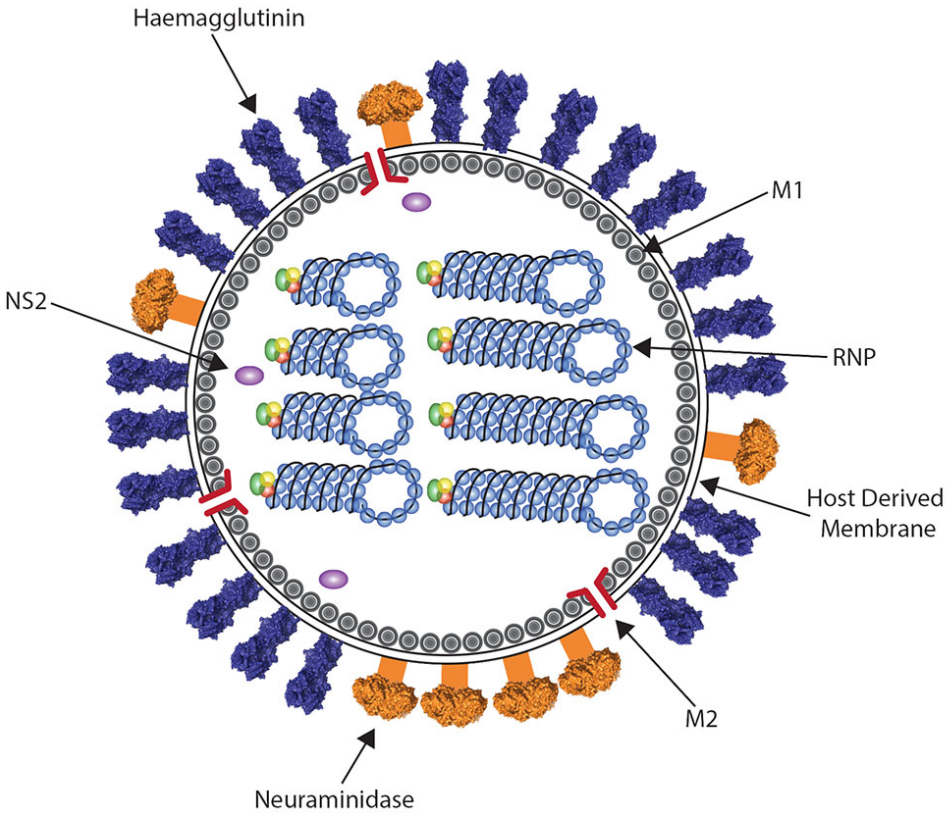


Figure 1.1: Schematic representation of influenza A virion. Virus consists of host-derived membrane studded with the surface glycoproteins haemagglutinin (HA) and neuraminidase (NA). Associated with the inside of the membrane is a layer of the matrix protein (M1). Within this membrane and layer of matrix protein are the RNPs, which consist of the single stranded negative sense RNA genome associated with NP and the viral polymerase complex. Components not drawn to scale.

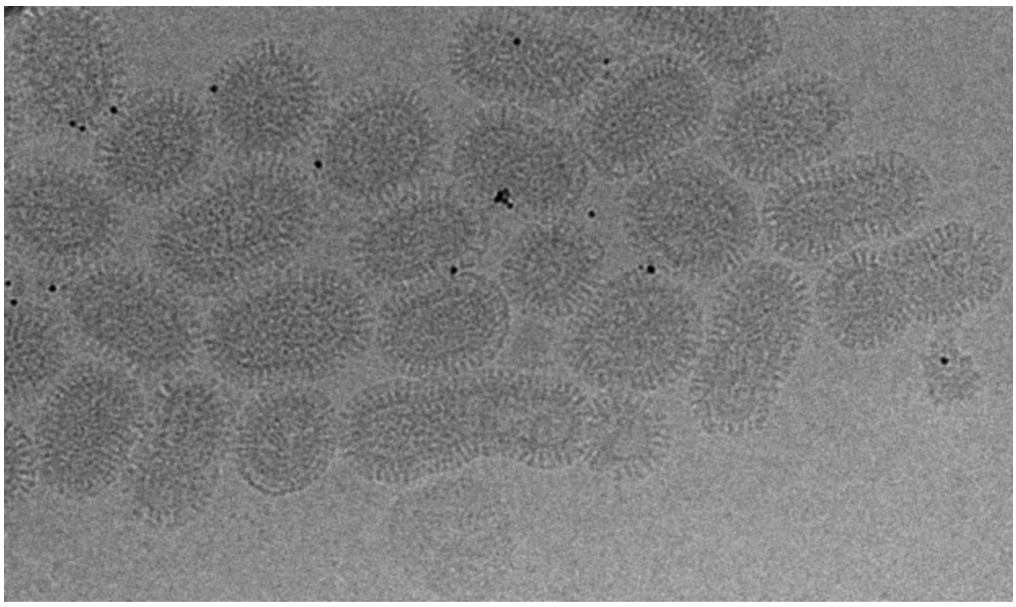


Figure 1.2: Cryo electron micrograph of X-31 virus. Image courtesy of Lesley Calder

clinical isolates causes a loss of filamentous morphology (Chu, Dawson & Elford, 1949; Kilbourne & Murphy, 1960). The role of this filamentous morphology in infection is unclear; however, the presence of this morphology in clinical isolates suggests that it could be relevant in terms of human infection. The filamentous morphology of IAVs has been found to be due to the M1 protein (Roberts, Lamb & Compans, 1998; Bourmakina & García-Sastre, 2003; Elleman & Barclay, 2004; Burleigh *et al.*, 2005).

The virus consists of a host-derived membrane, which is derived from the host plasma membrane upon virus budding, containing the surface glycoproteins HA and NA, with a dense layer of matrix (M1) protein lining the inside of the membrane. As well as the HA and NA there is also a small amount of the integral membrane protein M2. This protein is a low pH activated ion channel, which is responsible for allowing the transfer of protons into the virion to allow acidification of the virion and dissociation of the M1 protein, an important step in releasing the viral genome (Helenius, 1992). Inside the core of the virus are the ribonucleoproteins (RNPs). These RNPs consist of one of the eight segments of single stranded RNA (ssRNA) coated in nucleoprotein (NP) and with the associated heterotrimeric RNA dependant RNA polymerase, which consists of polymerase basic 2 (PB2), polymerase basic 1 (PB1) and polymerase acidic (PA) proteins.

The lipid bilayer is studded with the surface glycoproteins HA and NA. On a single spherical virion there are 300 – 400 HA trimers and 50 – 100 NA tetramers (Inglis *et al.*, 1976; Ruigrok & Andree, 1984; Harris *et al.*, 2006), although these numbers can vary dependent on the particular virus. NA has been shown to form patches of higher density at one end of the virus. This cluster of NA is thought to be at

the end of the virus that is proximal to the membrane at the end of virus budding and is consequently is at the opposite end to where the RNPs are present. This is particularly visible in filamentous viruses (Calder *et al.*, 2010).

The non-structural protein 2 (NS2), also known as nuclear export protein (NEP), is thought to be present at a low level in the interior of the virus (Richardson & Akkina, 1991). Non-structural protein 1 (NS1) is not thought to be present in the virion, rather it is only synthesised in the cell after infection and transcription of the viral genome. NS1 is known to have a wide range of functions including suppressing various antiviral responses, such as reducing the level of interferon released, reviewed in (Hale *et al.*, 2008).

1.1.4.2 Genomic structure

The IAV genome consists of 8 segments of negative sense RNA, ~14 kb in length. These 8 segments can produce 10 – 17 different proteins. The 10 proteins essential for viral replication are: PB2, PB1, PA, HA, NP, NA, M1, M2, NS1 and NS2/NEP (McCauley & Mahy, 1983). Two of these essential proteins are formed as gene splice products: M2 (Lamb, Lai & Choppin, 1981) and NS2 (Inglis, Gething & Brown, 1980). As well as these proteins there are a number of other proteins that can be produced, dependent on the particular strain of the virus: PB1-F2 (Chen *et al.*, 2001), PB1-N40 (Wise *et al.*, 2009), PA-X (Jagger *et al.*, 2012), M42 (Wise *et al.*, 2012), NS3 (Selman *et al.*, 2012), and PA-N155 and PA-N182 (Muramoto *et al.*, 2013).

1.2 Antigenic shift and drift

Influenza viruses have a vast potential for genetic variation. This occurs by two different methods of antigenic variation: antigenic shift and antigenic drift, reviewed in (Scholtissek, 1995; Taubenberger & Kash, 2010).

Antigenic shift occurs when a virus reassorts its segmented genome after coinfection of a cell with two different strains (intraspecies shift) or when the virus infects a different host (interspecies shift). The product can, in both cases, be a virus that is novel to the host immune system of a population of individuals and can therefore spread unimpeded.

Antigenic drift occurs due to point mutations in the viral genome, which result in an alteration in the antigenicity. Mutant viruses generated can be selected by immune pressure caused by antibody related immunity. The cause of these point mutations is thought to be the error prone method of viral genome replication. This is due to the low fidelity of the viral polymerase, which lacks proof-reading capabilities. These point mutations are then amplified by the large number of replicative cycles undergone by the virus.

1.3 Virus glycoproteins

Influenza A viruses are enveloped, with this membrane derived from the host cell upon virus budding. Within this membrane are the two viral encoded surface glycoproteins: haemagglutinin (HA) and neuraminidase (NA). The HA is responsible for binding the receptor, sialic (N-acetyl neuraminic) acid, and membrane fusion. The NA is a receptor destroying enzyme, with a number of functions in the viral life cycle, cleaving sialic acid, to which the virus may be bound, to allow efficient replication. Figure 1.3 shows phylogenetic trees for IAV HA subtypes H1 – H16 and IAV NA subtypes N1 – N9. Both are broadly classified into group 1 and group 2 based on phylogenic clustering.

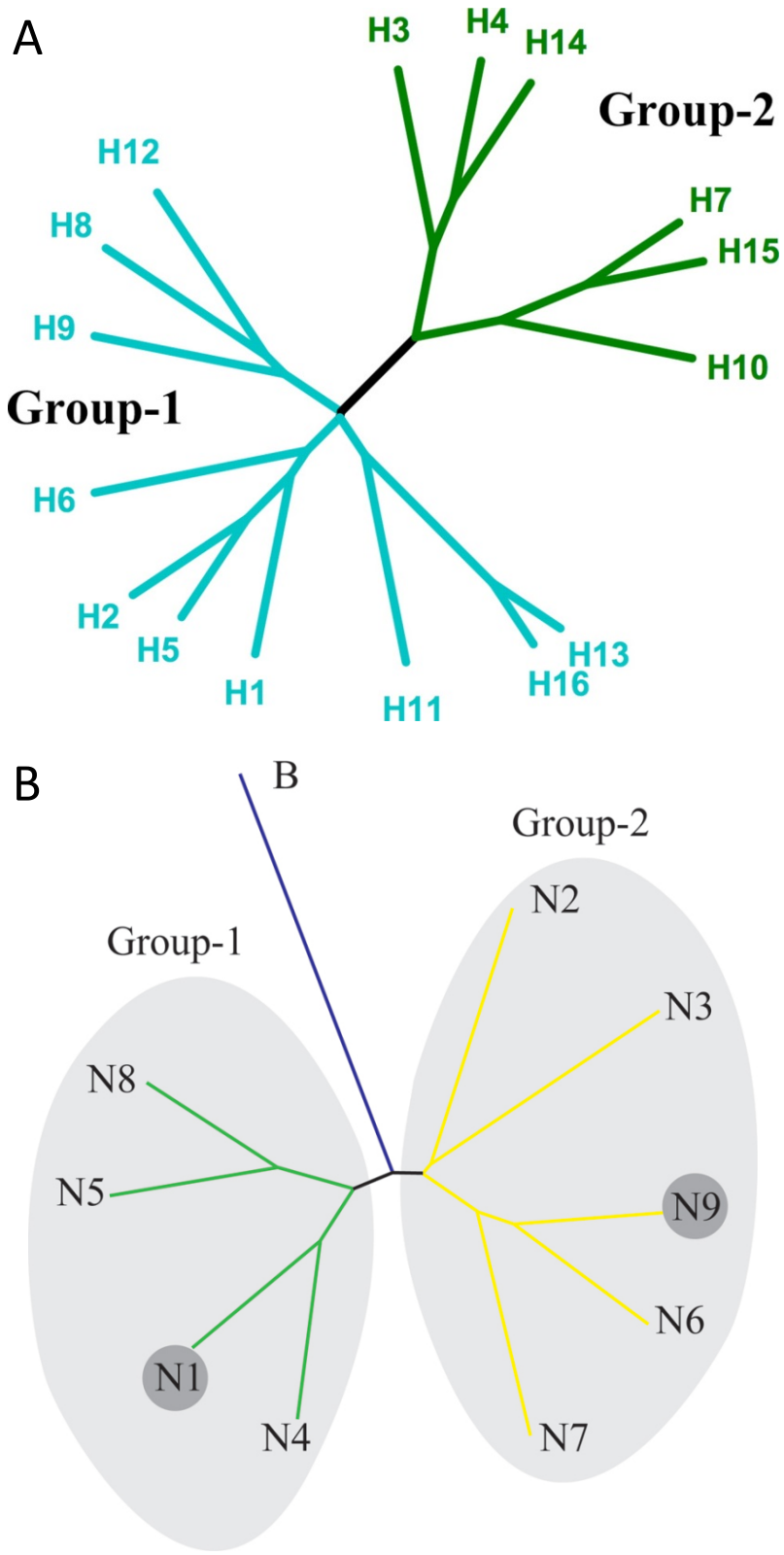


Figure 1.3: Phylogenetic trees of HA and NA subtypes. A) Phylogenetic tree of IAV HA subtypes H1 – H16, adapted from (Russell *et al.*, 2008). B) Phylogenetic tree of IAV NA subtypes N1 – N9 and Influenza B virus NA, adapted from (Russell *et al.*, 2006). Both HA and NA fall into two broad groups, Group 1 and Group 2, based on phylogenetic clustering.

1.3.1 Haemagglutintin

1.3.1.1 Structure of HA at neutral pH

The structure of IAV HA has been extensively studied by X-ray crystallography. For reviews on HA structure see (Wiley & Skehel, 1987; Skehel & Wiley, 2000). HA is a homotrimeric class I integral membrane glycoprotein. Each monomer consists of a protein of ~550 residues, ~75kDa monomeric mass, dependent on subtype. The C-terminal ~40 residues of each monomer of the trimer traverse the membrane, with a short cytoplasmic tail of ~15 residues.

The first structure of HA to be solved by X-ray crystallography was of a 1968 pandemic H3N2 HA in the high growth reassortant X-31, with the HA derived from the strain A/Aichi/2/68 (H3N2) (Wilson, Skehel & Wiley, 1981). This and many subsequent structures are of the soluble ectodomain of the HA. In this case the protein was generated by releasing the HA from purified virus by digestion with the protease bromelain, which cuts the protein at residue 503, producing a protein fragment known as BHA. The released ectodomain contains the majority of the protein, excluding the transmembrane domain and the cytoplasmic tail.

Figure 1.4 shows a structure obtained of this trimeric BHA ectodomain as well as an individual monomer. This structure revealed that the HA trimer is broadly made up of 2 different components: a fibrous stalk and a globular head region. The stalk of the HA is made up of a part of the N-terminal and C-terminal regions of HA1 and a coiled-coil, made up of HA2 subunits. The globular head region of the HA trimer is entirely made up of the HA1 subunit. This globular region of the HA1 consists of a large proportion of β-sheet, and a smaller proportion of α-helix.

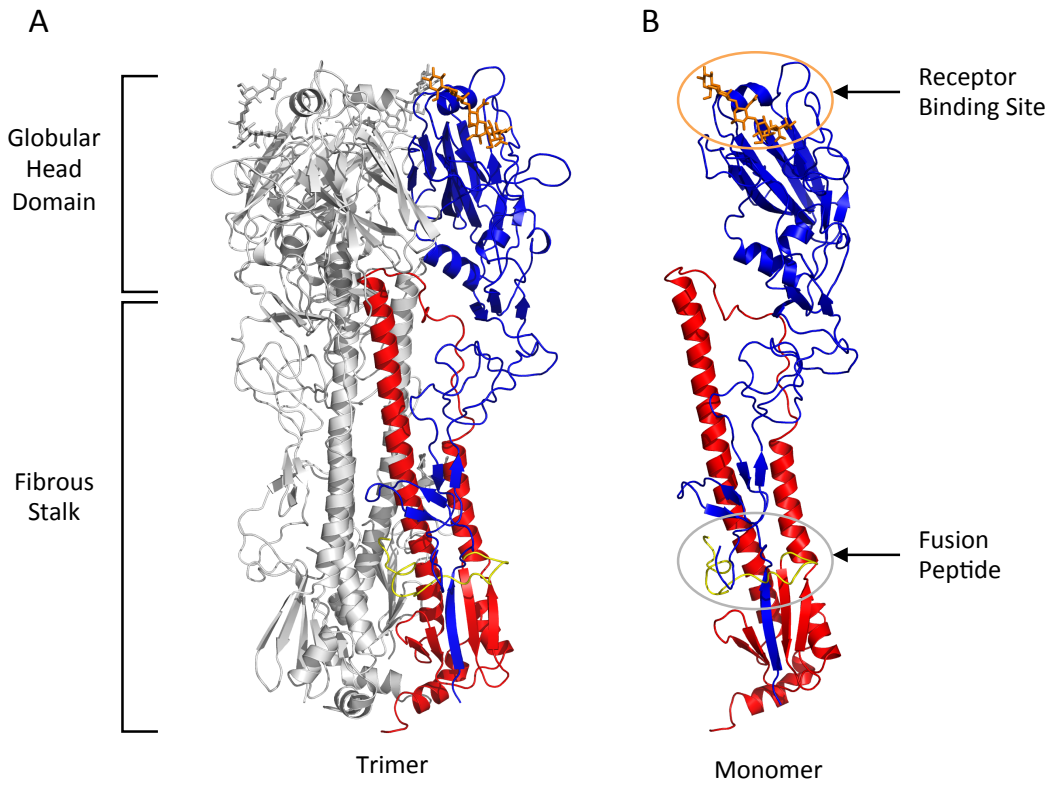


Figure 1.4: Structure of the HA at neutral pH. A) Trimeric HA with a single monomer coloured. B) Monomeric HA. HA1 shown in blue and HA2 shown in red. The fusion peptide is shown in yellow. Structure is of X-31 BHA with receptor analogue LSTa bound (coloured orange) (Eisen *et al.*, 1997) (Not in PDB). Figure made using PyMol (PyMol Molecular Graphics System, Schrödinger).

HA is synthesised as a single chain protein with an N-terminal signal peptide, which is co-translationally cleaved to form the nascent precursor protein HA0. Proteolytic cleavage, at position 329 (H3 numbering), separates each monomer into two disulphide-linked subunits HA1 and HA2. HA1 is the N-terminal domain, residues 1 – 328, and HA2 is the C-terminal domain, residues 330 – 550. The cleavage of HA0 into HA1 and HA2 is necessary for the HA to be able to undergo the conformational change needed to facilitate membrane fusion and therefore modulates virus infectivity (Klenk *et al.*, 1975; Lazarowitz & Choppin, 1975; Chen *et al.*, 1998).

The majority of HA0 molecules are cleaved at position 329 (H3 numbering) by host-derived extracellular serine proteases. Highly pathogenic avian influenza (HPAI) viruses have acquired a polybasic cleavage site. This reduces the protease specificity of the cleavage site, allowing cleavage by furin-like proteases in the golgi apparatus of the host cell (Stieneke-Gröber *et al.*, 1992). This cleavage of the HA allows these viruses to infect neighbouring cells without the need for control of extracellular cleavage of the HA. This allows these viruses to cause systemic infections, rather than localised to the site of HA cleavage. This acquired polybasic cleavage site is thought to have arisen in a number of different ways, including the slippage of the polymerase, which is the most common. However, recombination with other viral gene segments or recombination with ribosomal RNAs has also been seen, reviewed in (Böttcher-Friebertshäuser *et al.*, 2014).

As well as the structures of the ectodomain described above, where the virus attached HA is cleaved by proteases present in hens' eggs in which the virus had been propagated, structures have been solved of recombinant HA0, expressed using a vaccinia virus system, with the cleavage site removed by introduction of the

substitution R329Q (Chen *et al.*, 1998). This structure revealed that the cleavage site presents a projecting loop structure, enabling efficient cleavage. Comparison with the structure of the viral BHA showed a large movement of the cleaved ends of this loop. A comparison of uncleaved HA0 and cleaved HA1/HA2 shown in Figure 1.5.

At the distal tip of the HA there is a depression which, from the original structure obtained without bound receptor, was hypothesised to be the site of receptor binding (Fig. 1.4B) due to the conservation of a number of residues in different strains and also the structural similarity to another lectin, wheat germ agglutinin (Wilson, Skehel & Wiley, 1981). This conjecture was later shown to be correct when structures were solved with bound sialic acid receptor analogues (Weis *et al.*, 1988; Sauter *et al.*, 1992b; Watowich, Skehel & Wiley, 1994; Eisen *et al.*, 1997).

1.3.1.2 Receptor binding

HA is the protein primarily responsible for the virus binding to the receptor sialic acid. For reviews see (Skehel & Wiley, 2000; Xiong, McCauley & Steinhauer, 2014).

1.3.1.2.1 Structural basis of receptor binding

The receptor binding site (RBS) of HA is, as described above, formed of a shallow depression at the membrane distal end of the HA trimer, shown with a receptor bound in Figure 1.4. The RBS is made up of a number of structural regions of HA1: 1) the 130 loop; 2) the 150 loop; 3) the 190 helix; and 4) the 220 loop. There are many structures of HA with bound receptor of different HA subtypes and with different receptor analogues. For all of these structures the terminal sialic acid of the receptor is bound in roughly the same orientation, as shown in Figure 1.6A. The pyranose ring

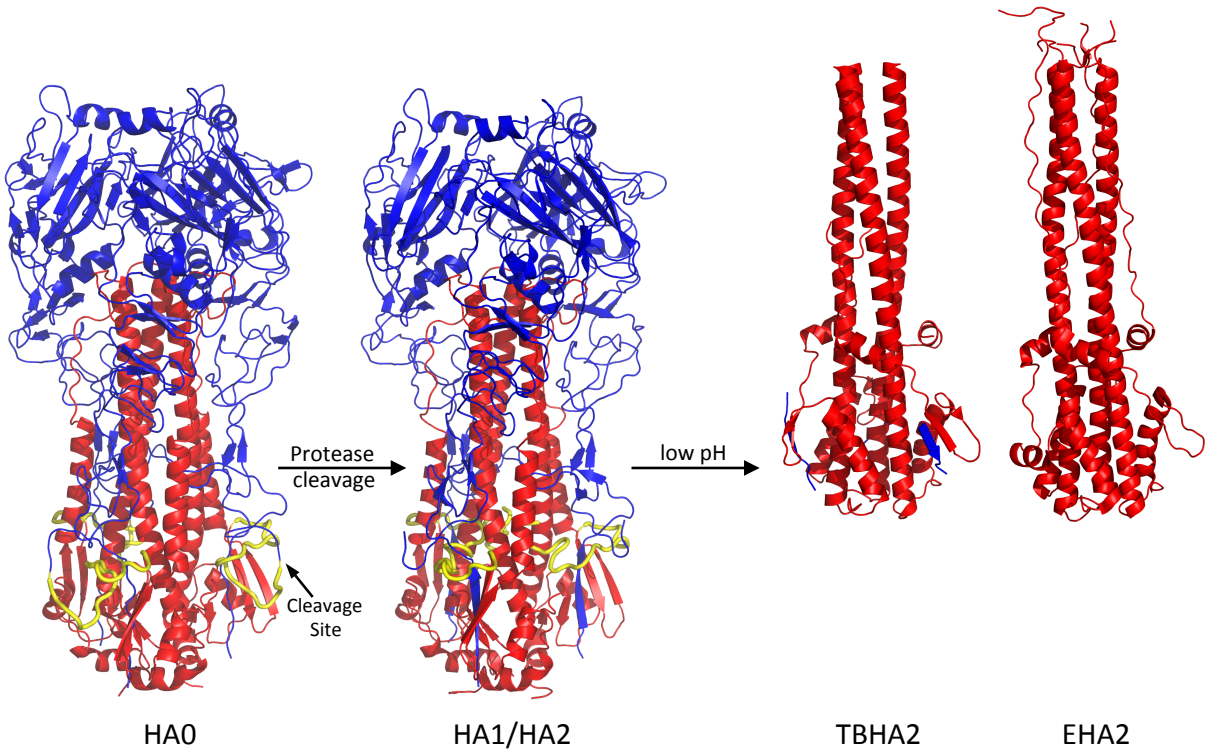


Figure 1.5: Comparison of HA0, neutral pH HA1/HA2, TBHA2 and EHA2 structures. HA1 is shown in blue and HA2 shown in red. The fusion peptide, when present in the structure, shown in yellow. Figure made using PyMol (PyMol Molecular Graphics System, Schrödinger) using structures of X-31 HA: HA0 (PDB: 1HA0) (Chen *et al.*, 1998); HA1/HA2 (PDB: 2YPG) (Eisen *et al.*, 1997); TBHA2 (PDB: 1HTM) (Bullough *et al.*, 1994a) and EHA2 (PDB: 1QU1) (Chen, Skehel & Wiley, 1999).

of the sialic acid lies flat against the RBS, with the glycerol side chain projecting towards the 190-helix, with the 8'-hydroxyl group forming a hydrogen bond with tyrosine 98 and the 9'-hydroxyl sometimes forming hydrogen bonds with glutamine 190, histidine 183 and serine 228 (depending on the virus strain). The acetamido nitrogen group forms a hydrogen bond with the carbonyl backbone residue of position 135 and the acetamido methyl group forms van der Waals contacts with the ring of tryptophan 153.

A secondary sialic acid binding site has also been identified near to the site described above, at the interface of two monomers (Sauter *et al.*, 1992a). By soaking HA crystals with different concentrations of receptor it was estimated that this secondary site has around a four-fold lower affinity for receptor. There is no evidence for this second site binding to the human-like receptor α2,6-linked sialic acid or for it having any influence on overall virus receptor binding.

1.3.1.2.2 Receptor specificity

As well as these conserved interactions of the sialic acid with the receptor binding site there are a number of other interactions that are more variable depending on the individual viral strain and the linkages of the terminal sugar residues. Sialic acid presents itself in three different linkages α2,3, α2,6 or α2,8. These linkages are formed when the 2' hydroxyl of sialic acid is linked to either the 3' or 6' carbon of the galactose or in the case of α2,8-linked sialic acid the 2' of one sialic acid is bound to the 8' carbon of another. 2,8-linked sialic acid can also form poly 2,8 sialic acid structures. The major linkages involved in IAV receptor binding are 2,6 and 2,3 linked, see Figure 1.7 for structures of two trisaccharide receptor analogues, one α2,6-linked (6SLN) and the other α2,3-linked (3SLN), used for experiments

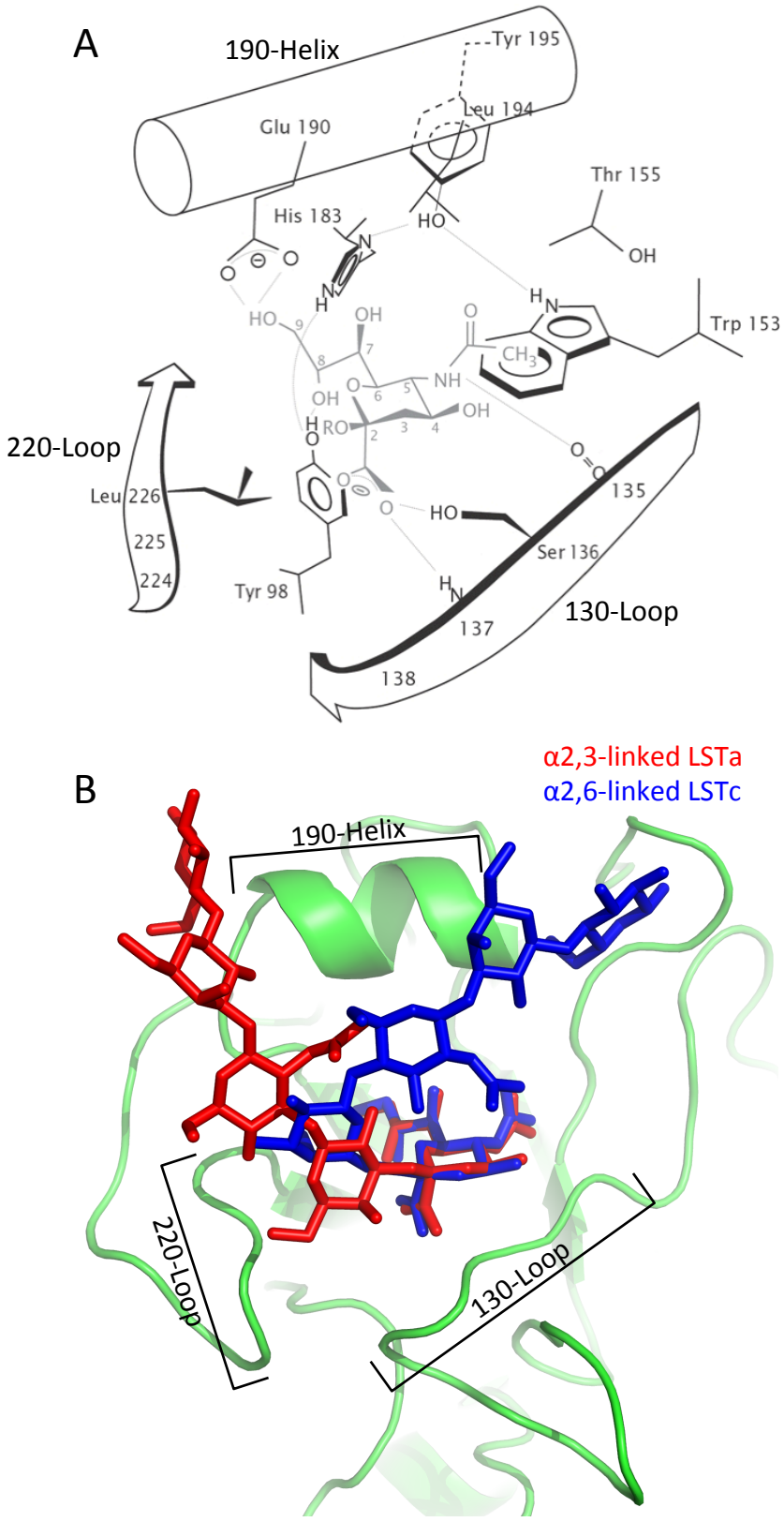
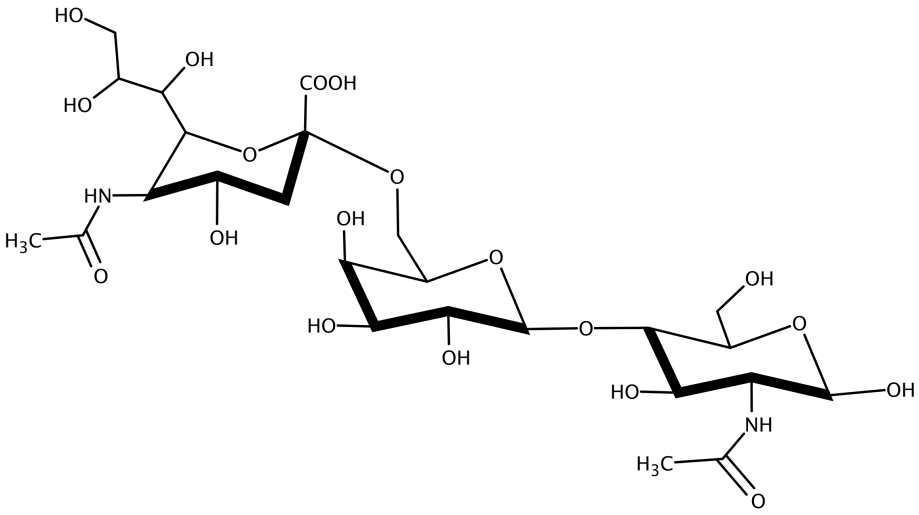
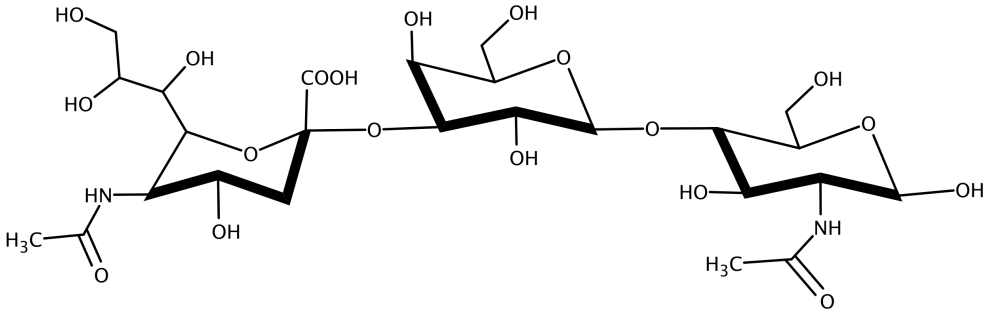


Figure 1.6: Structure of HA receptor binding site. A) Diagrammatic representation of HA residues involved in binding to sialic acid, adapted from (Skehel & Wiley, 2000). B) Structures of X-31 BHA bound to LSTa (Not in PDB) and LSTc (PDB: 2YPG) receptor analogues (Eisen *et al.*, 1997). Figure made using PyMol (PyMol Molecular Graphics System, Schrödinger).



α2,6-Sialyl-N-acetyllactosamine (6SLN)



α2,3-Sialyl-N-acetyllactosamine (3SLN)

Figure 1.7: Structure of receptor analogues used in these studies. MarvinSketch was used to draw structures (Marvin, ChemAxon, http://chemaxon.com).

presented in this thesis. These analogues consist of the terminal three sugars of the more complex glycan structures present on cell surface glycoproteins and glycolipids. There has been one report of IAVs binding to 2,8 linked sialic acid (Wu & Air, 2004), although this work was carried out using the mouse adapted neurotropic strain A/NWS/33, which has inherently unusual binding characteristics.

The ability of the virus to preferentially bind to α2,6- and α2,3-linked sialic acid is thought to be a major factor in determining host restriction of viruses. Rogers and Paulson were the first to observe that avian viruses have been found to typically bind to α2,3-linked sialic acid whereas human viruses typically bind to α2,6-linked sialic acid by selective resialylation of NA treated red blood cells (Rogers & Paulson, 1983). This receptor specificity of HA is thought to be a major determinant of the zoonotic potential of viruses from different animal reservoirs into humans and their subsequent transmissibility within the human population (Skehel & Wiley, 2000). In birds IAV infection is primarily an enteric infection, with transmission occurring typically via the faecal-oral route (Yoon, Webby & Webster, 2014). The sialic acid present in the alimentary canal of birds is primarily α2,3-linked (França, Stallknecht & Howerth, 2013) and consequently the receptor binding of avian infecting viruses is attuned to this sialic acid linkage. The primary site of infection in humans is in the upper respiratory tract (URT). The sialic acid present in the URT is primarily α2,6-linked (Couceiro, Paulson & Baum, 1993) and consequently viruses infecting humans must contain an HA that possess affinity for these receptors.

A number of viruses have been found to infect the lower respiratory tract (LRT) of humans. Present in the LRT is a higher density of α2,3-linked sialic acid receptors (Nicholls *et al.*, 2007). This higher density of α2,3-linked sugars allows

viruses to infect the LRT which would not typically infect the URT because of the incompatible receptor, such as avian H5N1 viruses (van Riel *et al.*, 2006).

It has also been shown that IAVs with different receptor preferences infect different cell types within the respiratory tract. Matrosovich *et al.* demonstrated that non-ciliated URT cells express higher levels of the human α2,6-linked sialic acid receptors and are more efficiently infected with viruses with an α2,6-linked sialic acid receptor preference. The inverse was also true with viruses with an avian receptor preference bound primarily to ciliated URT cells, which have higher levels of α2,3-linked sialic acid (Matrosovich *et al.*, 2004a). However, later work demonstrated that this partitioning of receptor specificity and infection tropism between different cell types is not as distinct as previously thought (Thompson *et al.*, 2006; Ibricevic *et al.*, 2006; Oshansky *et al.*, 2011).

The respiratory tract of pigs contains both α2,6- and α2,3-linked sialic acid. For this reason pigs have been postulated as a possible mixing vessel for viruses by allowing viruses with either avian or human specificities to reassort or to allow avian viruses to gain an HA gene with a preference for binding to the human α2,6-linked sialic acid receptor before entering the human population (Ito *et al.*, 1998).

Although the binding of the terminal sialic acid residue is highly conserved between different receptors and subtypes, there is more variation in the binding of distal sugar residues. Eisen *et al.* solved structures of pentasaccharides, LSTa (α2,3-linked sialic acid) and LSTc (α2,6-linked sialic acid) bound to X-31 BHA. These structures showed a folded conformation of the α2,6-linked LSTc and an extended conformation of the α2,3-linked LSTa. Figure 1.6B shows a structure of the HA receptor binding comparing the binding of LSTa and LSTc. These structures

demonstrated the importance of a number of residues in controlling the receptor specificity of the receptor-binding site. Residues 225, 226 and 227 are located at the bond between the sialic acid and the galactose of the receptor and are well characterised as being important in altering receptor specificity. The folded conformation of the α2,6 linked receptors also highlights the importance of interactions of the distal sugar residues with the 190-helix, labelled in Figure 1.6B.

It has also been hypothesised that the topology of glycans can play a role in host specificity of viruses, adding another level of complexity to the relative levels of α2,6 and α2,3 binding. Chandrasekaran *et al.* suggest that α2,3-linked receptors adopt a 'cone-like' topology and α2,6-linked receptors an 'umbrella-like' topology, suggesting an important role in binding of receptors further down the glycan chains (Chandrasekaran *et al.*, 2008).

The residue at position 226 has been identified as a major determinant of the receptor specificity of IAV HA. Evidence for the importance of this residue was first seen when Rogers *et al.* adapted the receptor binding preference of a 1968 H3 strain, X-31, to a more avian-like receptor binding preference by passaging the virus in the presence of horse serum (Rogers *et al.*, 1983), which is rich in α2,6-linked sialic acid (Pritchett & Paulson, 1989). This passage with excess α2,6-linked sialic acid present caused the substitution L226Q. This shifted the receptor binding specificity of the 'human virus' X-31 from a predominant preference for α2,6- to a preference for α2,3-linked sialic acid, shown by measuring the ability to agglutinate rederivatised red blood cells (Rogers *et al.*, 1983). Likewise, *in vitro* selection of an avian virus produced the reverse mutation Q226L, which made the avian virus bind predominantly to α2,6 rederivatised blood cells (Rogers *et al.*, 1985). This alteration,

Q226L, accompanied by G228S are thought to be important to the adaption of both the 1957 H2N2 and 1968 H3N2 pandemics to humans from wild birds, altering receptor binding characteristics (Rogers *et al.*, 1983; Naeve, Hinshaw & Webster, 1984; Bean *et al.*, 1992; Connor *et al.*, 1994; Matrosovich *et al.*, 2000). H1N1 viruses have quite different binding characteristics, with less reliance on the residues at position 226 and 228 with viruses with avian-like 226 and 228 residues not having such clear alterations in binding, with the important residues for altering receptor specificity appearing to be 225, 190 and 186 (Rogers & D'Souza, 1989; Matrosovich *et al.*, 1997), reviewed in (Xiong, McCauley & Steinhauer, 2014).

1.3.1.2.3 Receptor binding affinity

The receptor binding affinity of monomeric HA for monomeric sialic acid containing receptors has been found to be very weak, in the mM range, determined by NMR (Sauter *et al.*, 1989; 1992b) and more recently by microscale thermophoresis (MST) (Xiong *et al.*, 2013a). The strong binding of viruses to the cell surface to initiate infection must therefore be due to multivalency and cooperativity between HA binding sites on the same virus. The multivalency of IAV binding has been confirmed by experiments measuring unbinding events using optical tweezers and atomic force microscopy, which measured a number of sequential events of unbinding necessary to release a single virus bound to a cell (Sieben *et al.*, 2012). The necessary multivalent interaction of HA with receptors can also be seen when attempting to make sialic acid based inhibitors of virus binding, which need to be multivalent to have any significant affinity, reviewed in (Mammen, Choi & Whitesides, 1998; Matrosovich & Klenk, 2003).

The relationship between the monomeric binding of a single HA binding site to a monomeric receptor has been described by comparing monomeric protein-receptor affinities with data collected for viruses binding to receptor-coated surfaces. Xiong *et al.* related the dissociation constant of a monomeric interaction, $K_{d(receptor)}$, to that of a multivalent virus interaction, $K_{d(virus)}$, in terms of a multiplicity coefficient, mc ($K_{d(virus)} = K_{d(receptor)}^{mc}$). These studies identified that this multiplicity coefficient is in the range of 3 – 5.5 giving an effective lower limit on the number of monomeric interactions required for IAV interactions with a receptor-coated surface (Xiong *et al.*, 2013a).

1.3.1.3 Membrane fusion

HA is the virus surface protein responsible for causing membrane fusion. For reviews see (Skehel & Wiley, 2000; Cross *et al.*, 2009).

As previously described in section 1.3.1.1 HA is synthesised as a precursor, HA0, which must be cleaved by proteases into the two subunits HA1 and HA2, priming the membrane fusion potential of the HA. The cleavage of the HA into these two subunits generates the fusion peptide, a short stretch of 23 hydrophobic amino acids at the N terminus of HA2, which is then buried within the neutral pH conformation of HA.

Upon induction of fusion activity by low pH in the endosome the HA undergoes a large conformational change, which is thought to force this fusion peptide into the endosomal membrane, followed by another conformational change, which is thought to force the viral and endosomal membranes together, leading to membrane fusion.

There have been a large number of structural and biochemical studies that have shed light on the fusion process of HA. The structure of a protease treated

fragment of viral HA in a low pH treated conformation has been determined (Bullough *et al.*, 1994a). The BHA of X-31 was treated with low pH (5.0) and protease treated with trypsin and thermolysin, yielding a soluble protein fragment including residues 1 – 27 of HA1 and 38 – 175 of HA2 known as TBHA2 (Ruigrok *et al.*, 1988; Bullough *et al.*, 1994b). This low pH structure was found to have a very different conformation from that in the original neutral pH structure (Wilson, Skehel & Wiley, 1981). The HA2 undergoes a large structural rearrangement with the major feature being the formation of an extended helix at the N terminus of HA2 that moves the fusion peptide 100 Å away from its previously buried position near the viral membrane. As well as this movement of the fusion peptide there is rearrangement of the C-terminus of the protein, which appears to move the membrane bound C-terminus toward the fusion peptide, see Figure 1.5 for comparison of neutral and low pH structures.

A structure of HA2 has been obtained for protein expressed in *E. coli* (residues 38-175 HA2). This protein (EHA2) was found to adopt a similar conformation to that seen for the low pH treated TBHA2 (Chen *et al.*, 1995; Chen, Skehel & Wiley, 1999). The formation of this protein conformation in *E. coli* implies that this is the lowest energy conformation of HA2, suggesting that the neutral pH form of HA is held in an unstable form that, upon induction of fusion by HA0 cleavage and low pH, will refold to form the lowest energy conformation, similar to EHA2. For a comparison of structures of TBHA2 and EHA2 see Figure 1.5.

The precise mechanism by which the HA induces membrane fusion is unclear. The induction of low pH has been shown to cause insertion of residues 1 – 22 of HA2 into the membrane, indicating that the fusion peptide is the only part of the HA that

interacts with the membrane (Tsurudome *et al.*, 1992; Weber *et al.*, 1994; Durrer *et al.*, 1996). The interaction of the fusion peptide with the membrane and the associated conformational changes of the HA lead to fusion of the viral and endosomal membranes. The fusion of the two membranes is thought to happen via a hemifusion intermediate before the formation of a fusion pore owing to experiments carried out using GPI-anchor tethered BHA which allow fluorescent lipid dye to transfer from one membrane to another but does not allow the mixing of aqueous dyes separated by the two membranes (Kemble, Henis & White, 1993; Kemble, Danieli & White, 1994; Melikyan, White & Cohen, 1995; Floyd *et al.*, 2008).

The pH at which membrane fusion is initiated is variable between different HAs. The fusion pH of a typical human infecting HA is in the range 5.0 – 5.5 (Wharton, Skehel & Wiley, 1986; Galloway *et al.*, 2013). This pH of membrane fusion can, however, vary greatly and a large number of substitutions have been found to alter fusion pH, reviewed in (Russell, 2014). The pH of fusion is thought to be a determinant in the transmissibility of viruses from avian to human sources. Avian virus HAs typically have a higher pH of fusion, thus reducing their relative stability, reviewed in (Russell, 2014). Recent studies which involved the passage of H5N1 viruses in ferrets in order to increase transmissibility both introduced mutations which are thought to reduce the pH of fusion, thus stabilising the HA, suggesting that this increase in HA stability could be a necessary alteration for increased human transmissibility of some avian viruses (Imai *et al.*, 2012; Herfst *et al.*, 2012).

1.3.1.4 Glycosylation

IAV HA is a glycoprotein, which can have a variable number of N-linked glycosylations present. A number of these are conserved between subtypes and

considered vital for the structure of the HA, as they are involved in intra subunit interactions (Wilson, Skehel & Wiley, 1981). The number of theoretical glycosylation sites present, found by analyzing sequences for the Asn-Xaa-Ser/Thr motif (where Xaa is any amino acid but Pro), in HAs varies between 5 and 12 (Nobusawa *et al.*, 1991; Skehel & Wiley, 2000; Vigerust & Shepherd, 2007; Reading *et al.*, 2007). The presence of glycans near the RBS of HA have been found to reduce receptor binding of a number of different viruses (Aytay & Schulze, 1991; Ohuchi *et al.*, 1997; Gambaryan, Robertson & Matrosovich, 1999; Wagner *et al.*, 2000; Baigent & McCauley, 2001; Ohuchi *et al.*, 2002; Tsuchiya *et al.*, 2002; Abe *et al.*, 2004; Wang *et al.*, 2009). Acquisition of glycosylation is also a major mechanism of epitope shielding acquired to enhance immune evasion by IAVs (Blackburne, Hay & Goldstein, 2008), reviewed in (Reading *et al.*, 2007).

1.3.1.5 Antigenic variation

HA is the major surface protein component of the virus and consequently is the major antigenic target of the virus. This therefore selects for considerable variation within the HA, which is particularly focused around the globular head region of the protein. The major antigenic sites have been classified for H1 and H3 virus subtypes by epitope mapping (Wiley, Wilson & Skehel, 1981; Gerhard *et al.*, 1981; Caton *et al.*, 1982; Underwood, 1982). These epitopes are generally surface exposed flexible loops. Many of these loops are also part of the RBS and consequently antigenicity and receptor binding are inextricably linked.

1.3.2 Neuraminidase

NA is a sialidase and is responsible for cleaving terminal sialic acid residues from complex glycans to which the virus may be bound at different stages in the

infection cycle. For reviews see (Colman, 1994; Garman & Laver, 2005; Itzstein, 2007; Air, 2012).

1.3.2.1 Structure

NA is a homotetrameric class II integral membrane glycoprotein protein consisting of ~470 residues with a monomeric mass of ~50 kDa. When attached to the virus the protein is 'mushroom-shaped', with a globular head region sitting atop a narrow stalk (Harris *et al.*, 2006; Calder *et al.*, 2010).

The three dimensional structure of the head region of the IAV NA was solved by releasing the protein from purified virus using the enzyme mixture pronase, which releases a ~200 kDa tetramer from the virus, followed by crystallisation and X-ray diffraction (Varghese, Laver & Colman, 1983; Colman, Varghese & Laver, 1983). The head domain of the protein is made up of a structure largely consisting of β-sheet, consisting of six 4-stranded anti-parallel sheets. The sialidase site of NA is located centrally within the monomeric structure within a deep pocket, see Figure 1.8. The sialidase site has been defined owing to the highly conserved nature of the residues and their role in binding to sialic acid (Colman, Varghese & Laver, 1983; Varghese *et al.*, 1992). The residues which are involved in binding to sialic acid are shown in Figure 1.9A. The exact mechanism of enzyme catalysis is unknown. It has been suggested that the NA could cleave via a sialosyl cation transition-state complex (Chong *et al.*, 1992) but it has also been proposed that it could proceed via a covalently linked intermediate which attaches via Tyr 406. This has been proposed because a number of NA inhibitors that have been characterised that can inhibit by forming a covalent bond with the active site (Kim *et al.*, 2013; Vavricka *et al.*, 2013).

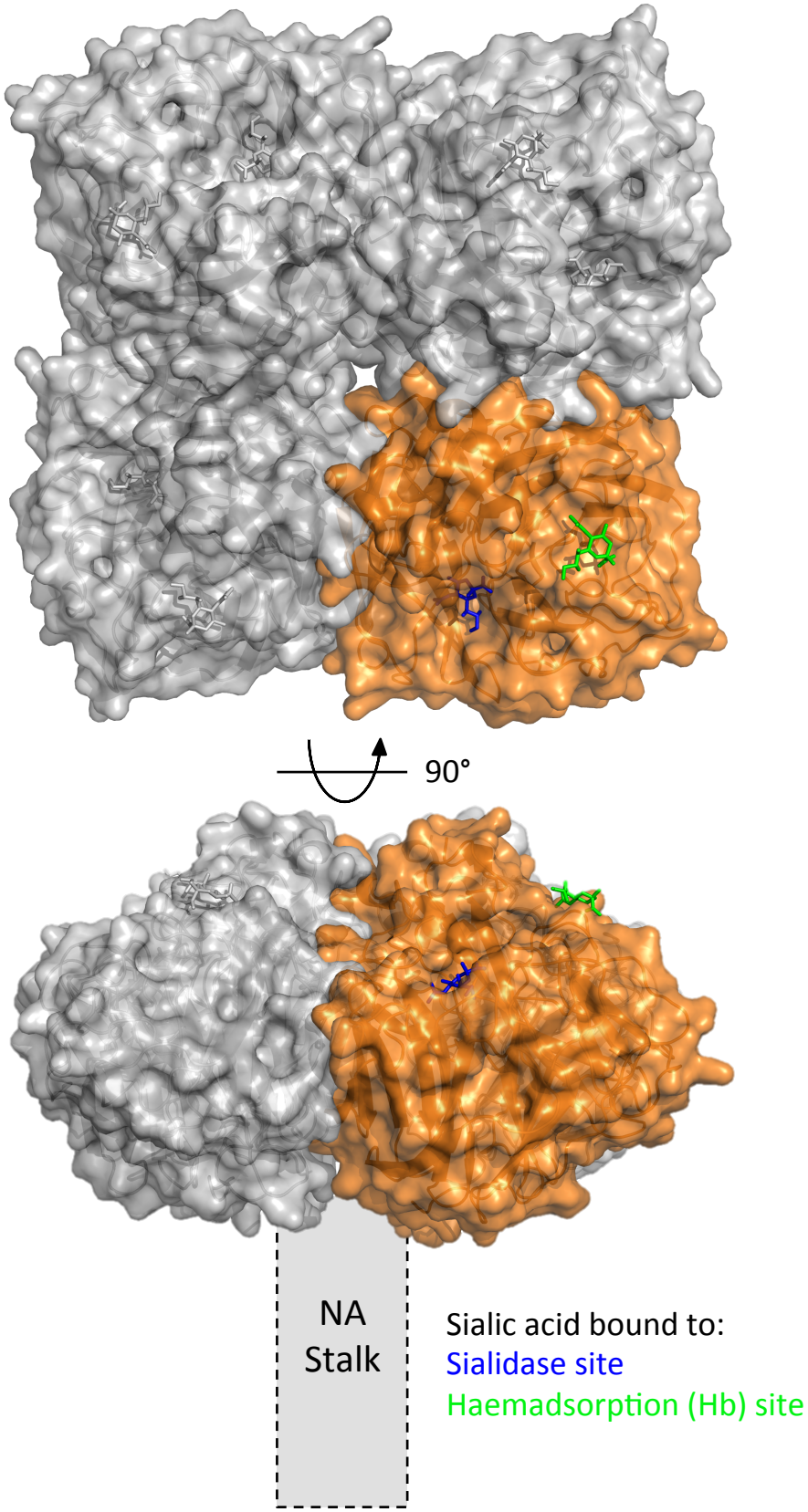


Figure 1.8: Structure of NA tetramer. N9 NA tetramer surface with one monomer coloured (A/tern/Australia/G70C/75) (PDB: 1MWE) (Varghese *et al.*, 1997). Sialic acid bound to sialidase site (blue) and haemadsorption (Hb) site (green) shown. Location of NA stalk indicated. Figure made using PyMol (PyMol Molecular Graphics System, Schrödinger).

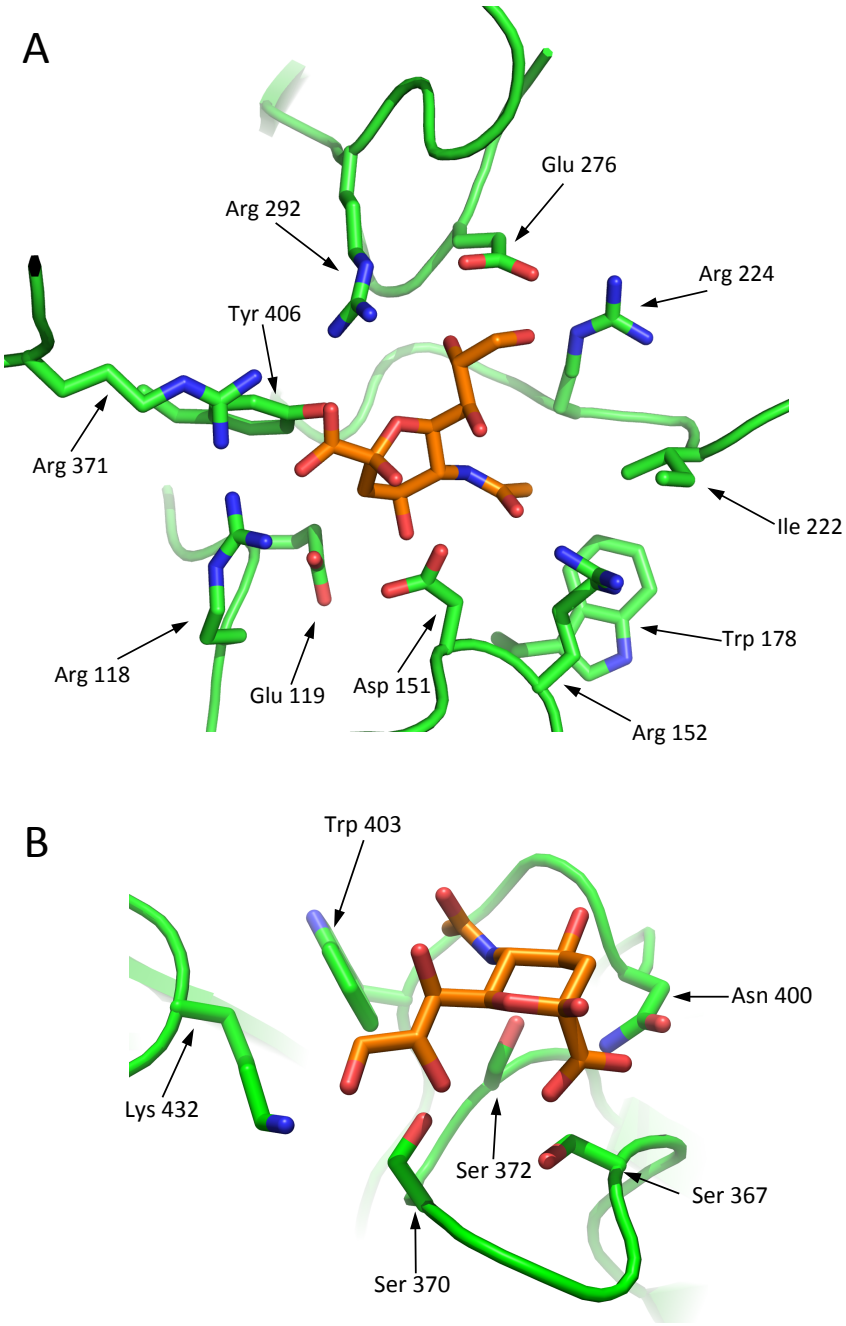


Figure 1.9: Structure of NA sialidase and haemadsorption sites. A) Sialidase site with conserved residues labelled. B) Haemadsorption (Hb) site with interacting residues labelled. Structures are of N9 NA (A/tern/Australia/G70C/75) (PDB: 1MWE) (Varghese *et al.*, 1997). Sialic acid shown in orange. Figure made using PyMol (PyMol Molecular Graphics System, Schrödinger).

A high affinity calcium binding site was identified in the first structure of NA (Varghese, Laver & Colman, 1983), as well as in a number of other subsequent structures. The exact role of this and other bound calcium ions is unknown; however, the presence of calcium ions has been found to be essential for full enzymatic activity (Carroll & Paulson, 1982; Chong, Pegg & Itzstein, 1991) and also for influenza B virus NA protein stability (Burmeister, Cusack & Ruigrok, 1994).

1.3.2.2 Sialidase activity

The sialidase activity of NA has been proposed to have a number of roles in the virus infection cycle. The NA is involved in releasing progeny virions from the cell surface after virus budding and also in preventing virus aggregation due to the binding of HA to sialic acid that is part of the glycosylation of HA and NA of other nearby virions. The role of NA in these activities can be seen when viruses are propagated with reduced NA activity, leading to aggregation of virus particles both to other viruses as well as to the cell surface. This has been seen with viruses either lacking NA activity (Palese *et al.*, 1974) or with their NA inhibited (Palese & Compans, 1976; Baigent, Bethell & McCauley, 1999). Another role of the sialidase activity is thought to be that of evading non-specific inhibition by sialylated substances such as mucins (Matrosovich *et al.*, 2004b; Cohen *et al.*, 2013; Yang *et al.*, 2014). NA has been found to have a low pH optimum of 5.5 – 6.5 and a temperature optimum of 37°C (Air, 2012).

A number of viruses have been found which appear to use the NA as a receptor binding protein via the sialidase site, rather than the haemadsorption site discussed below (Lin *et al.*, 2010; Zhu *et al.*, 2012a) (Stephen Wharton, unpublished data). This binding activity is mediated by a mutation around the sialidase site, D151G, which

lowers the catalytic turnover rate (k_{cat}) as well as reducing the K_m, for cleavage of the fluorogenic substrate 2'-(4-Methylumbelliferyl)-α-D-N-acetylneuraminic acid (MUNANA), indicating that this mutation makes the sialidase site adopt more of the characteristics of a receptor binding protein, increasing the affinity of the sialic acid interaction whilst simultaneously reducing the probability of substrate cleavage (Zhu *et al.*, 2012a) (Stephen Wharton, unpublished data).

NA has been found to cleave α2,3-linked sialic acid much more efficiently than α2,6-linked sialic acid. A number of studies have found that the ability of NAs to cleave α2,6-linked sialic acid derivatives increases as they adapt to the human host (Baum & Paulson, 1991; Kobasa *et al.*, 1999); however, 2,3-linked sialic acid remains the more efficiently cleaved substrate. This has also been measured more recently for wider ranges of sugars, with a similar pattern of an overall preference for α2,3-linked receptors (Li *et al.*, 2011b; Xu *et al.*, 2012b; Zhu *et al.*, 2012a).

1.3.2.3 Haemadsorption activity

Neuraminidases of the subtype N9 have been found to have a second sialic acid binding site. Laver *et al.* showed that full length detergent extracted N9 made into rosettes possessed haemagglutination activity and also that this haemagglutination activity was unaffected by inhibition of the sialidase site with a specific inhibitor, suggesting that this haemagglutination site was separate from the sialidase site (Laver *et al.*, 1984). NA from an H7N1 virus has also been found to have haemadsorbing activities (Hausmann *et al.*, 1995). This haemadsorption activity has been found in a wide range of other NAs, particularly those from avian viruses. Human virus NAs do not appear to possess these haemadsorbing properties (Kobasa *et al.*, 1997).

The structure has been determined by X-ray crystallography of N9 NA with sialic acid bound in this secondary, haemadsorbtion (Hb), site (Varghese *et al.*, 1997). This structure revealed sialic acid bound both to the Hb site in a shallow depression on the top of the NA and the sialidase site, Figure 1.8. The sialic acid in the Hb site is bound in the chair conformation rather than the twisted boat of the sialidase site previously reported. The residues involved in binding the sialic acid at this Hb site are Ser367, Ser370, Ser372, Asn400 and Trp403 (N2 numbering), see Figure 1.9B. This original structure of sialic acid bound to the Hb site and subsequent structures of different subtypes have shown that the 2' hydroxyl of the bound sialic acid projects upwards, away from the structure, indicating that this Hb site is unlikely to have a sialic acid linkage specificity (Varghese *et al.*, 1997; Rudino *et al.*, 2006; Sun *et al.*, 2014). This is also backed up by the similar haemadsorption of erythrocytes from chicken and human sources (Kobasa *et al.*, 1997; Uhlendorff *et al.*, 2009), which are thought to contain different sialic acid linkages (Ito *et al.*, 1997).

Uhlendorff *et al.* introduced Hb properties into a 1957 N2 by site directed mutagenesis, introducing the mutation N367S. This Hb activity could then be removed by introducing the mutations S370L, N400S and W403R. This study also examined the influence of the presence of this Hb site on the activity of the NA. The Hb^+ virus allowed elution of the virus from erythrocytes bound at 4°C (to reduce sialidase activity), followed by elution at 37°C. This elution could not be achieved in the Hb^- virus, suggesting a role of the Hb site in the activity of the NA. Measurements of the Hb^+ and Hb^- NA cleaving small monomeric substrates showed no difference in activity between the two NAs; however, when measuring the cleavage of multivalent substrates there was enhanced cleavage when the Hb activity was present. This

suggested that the Hb site could be involved in tethering multivalent substrates to allow more efficient cleavage by the sialidase site (Uhlendorff *et al.*, 2009).

Lai *et al.* have measured by NMR the binding of sialic acid containing receptors to the Hb site of NA. They also suggest that the sialic acid can be competed out of the Hb site by adding the NA inhibitor oseltamivir carboxylate (Lai *et al.*, 2012).

1.3.2.4 Stalk length

The length of the stalk attaching the NA head region to the viral membrane has been found to vary. Deletions of stalk length have been found ranging from 1 – 36 amino acids in length. Changes in NA stalk length are found especially to occur as a deletion when viruses transmit from wild birds to domestic poultry (Li *et al.*, 2011a).

This deletion of NA stalk has been found to alter the propagation characteristics of the virus (Castrucci & Kawaoka, 1993; Luo, Chung & Palese, 1993; Baigent & McCauley, 2001; Zhou *et al.*, 2009; Sorrell *et al.*, 2010; Blumenkrantz *et al.*, 2013; Sun *et al.*, 2013b), reduce the ability to elute virus from erythrocytes (Els *et al.*, 1985; Castrucci & Kawaoka, 1993; Baigent, Bethell & McCauley, 1999; Matsuoka *et al.*, 2009; Zhou *et al.*, 2009; Blumenkrantz *et al.*, 2013; Sun *et al.*, 2013b), or alter virus transmissibility/pathogenicity (Matsuoka *et al.*, 2009; Zhou *et al.*, 2009; Sorrell *et al.*, 2010; Munier *et al.*, 2010; Blumenkrantz *et al.*, 2013). However, this stalk length reduction does not alter the NA cleavage characteristics for small monovalent substrates (Matsuoka *et al.*, 2009; Blumenkrantz *et al.*, 2013; Sun *et al.*, 2013b). The reduction in NA stalk length is likely to reduce the effective cleavage ability of the NA for natural substrates, as there will be reduced substrate accessibility; however, it is unclear as yet what the evolutionary pressure is for stalk truncation and its role in altering virus transmissibility. However, one can conjecture that the loss of NA stalk

length could be selected because of the nature and density of sialic acid receptors present at the site of infection, which could require a lower effective NA activity to allow the efficient binding of the HA for the initiation of infection.

1.4 Virus replication cycle

See Figure 1.10 for a schematic representation of the IAV replication cycle.

1.4.1 Attachment

Virus attachment is mediated by the HA and NA, as described in further detail in section 1.3, Figure 1.10 step 1.

1.4.2 Entry

The attachment of the virus to the cell membrane by HA leads to receptor-mediated endocytosis, Figure 1.10 step 2. This is thought to proceed either via a clathrin dependent or independent pathway because electron microscopy images of virus endocytosis show virions in both smooth and coated endosomes (Matlin *et al.*, 1981). The entry of filamentous viruses has been found to occur via macropinocytosis because of the increased size of these viruses (Rossman, Leser & Lamb, 2012).

After viral endocytosis the endosome matures and acidifies to lower the pH to a level that induces HA mediated membrane fusion (pH 5.0 – 5.5), as described in further detail in section 1.3.1.3. This membrane fusion forms a fusion pore that allows passage of the RNPs into the cytoplasm of the cell. Coupled with this membrane fusion is acidification of the interior of the virion by the ion channel M2, which allows the dissociation of the M1, allowing viral uncoating and also releasing the viral RNPs from interactions with the M1 (Pinto, Holsinger & Lamb, 1992; Bui, Whittaker & Helenius, 1996; Ivanovic *et al.*, 2012), Figure 1.10 step 3.

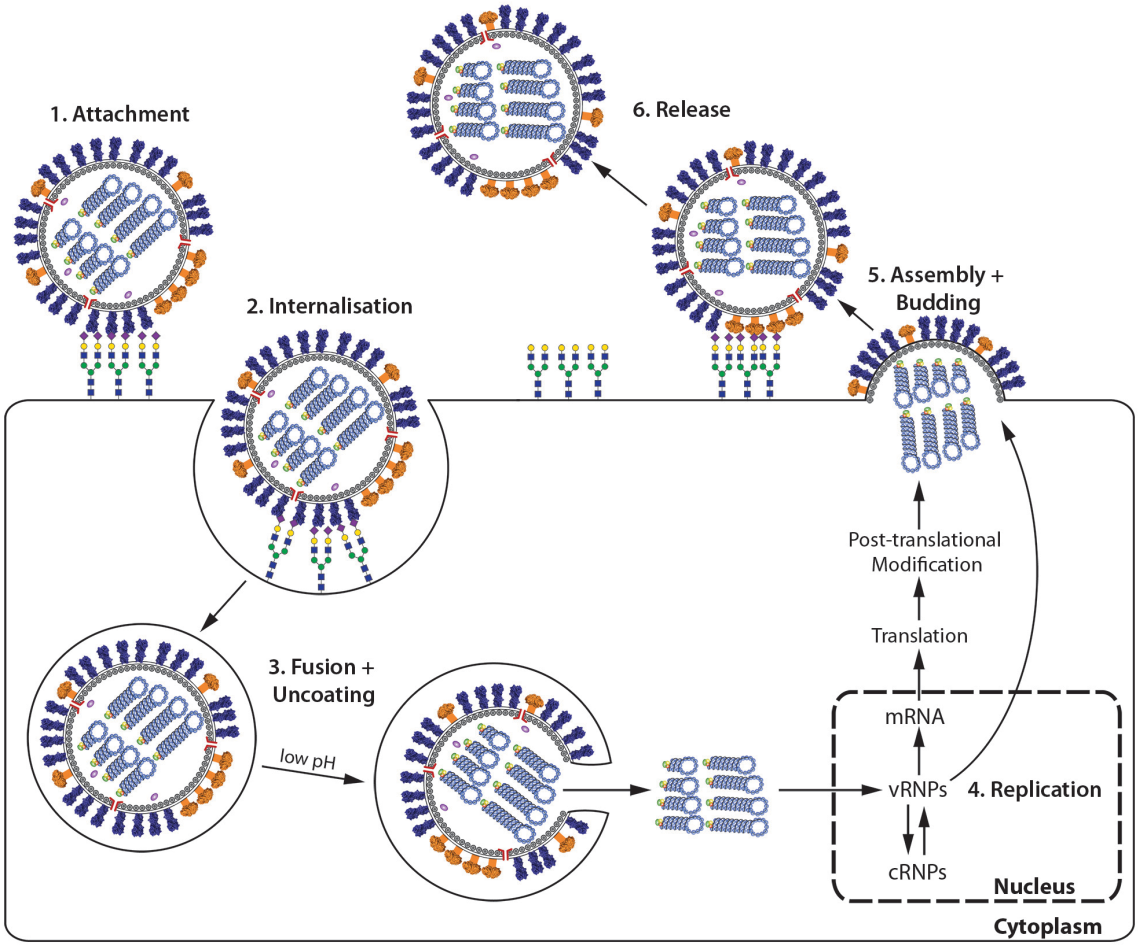


Figure 1.10: Schematic representation of IAV replication cycle. 1) Virus attaches to sialic acid containing cellular receptors before 2) internalisation by receptor mediated endocytosis. 3) Virus releases RNPs after HA mediated virus membrane fusion and M2 mediated virion acidification is induced by low pH in the late endosome. 4) RNPs enter the nucleus and undergo transcription and replication to generate more RNPs and to synthesise viral proteins. 5) RNPs and viral proteins assemble at the plasma membrane and bud to form new virus particles. 6) Virus release is facilitated by NA cleaving cell surface sialic acid receptors.

1.4.3 Replication

In order for replication of the viral RNPs to occur they must be trafficked into the nucleus. The movement of the RNPs from the cytoplasm to the periphery of the nucleus seems to simply rely on diffusion, as microtubules, intermediate filaments and actin filaments do not appear to be involved in the movement of RNPs (Babcock, Chen & Zhuang, 2004). Nuclear import of RNPs is aided by nuclear localisation signals present both in the NP and the polymerase subunits, which allows interaction with nuclear pore complexes. RNP entry into the nucleus is mediated by the classical importin pathway, reviewed in (Hutchinson & Fodor, 2013).

In order for the virus to replicate the RNPs must undergo replication, to generate more of the single stranded negative sense vRNA, and also transcription, to generate mRNAs to allow viral protein synthesis, reviewed in (Portela & Digard, 2002; Fodor, 2013), Figure 1.10 step 4.

vRNA transcription is carried out by a *cis* acting IAV polymerase attached to the RNP, which has both the 5' and 3' terminal sequences of the vRNA bound (Hagen *et al.*, 1994; Fodor, Pritlove & Brownlee, 1994). This binding is to RNA sequences which share a high level of homology in all IAV segments and between subtypes, a 13 base section at the 5' end of the vRNA and a 12 base section at the 3' end of the RNA (Skehel & Hay, 1978; Robertson, 1979; Desselberger *et al.*, 1980). The mRNAs generated by the polymerase complex require a 5' cap and polyadenylation of the transcript. Both of these are added to the mRNA by an unusual mechanism. The 5' 7-methylguanylate cap is taken from a cellular pre-mRNA by a process known as 'cap-snatching' (Plotch *et al.*, 1981). This process is thought to be aided by the association of the cellular RNA polymerase II, which is thought to aid the binding of a nascent

pre-mRNA by the PB2 subunit (Engelhardt, Smith & Fodor, 2005). The 5' cap is removed from the pre-mRNA by endonuclease activity of the PA subunit and the 3' end of this RNA primer is moved to the PB1 subunit in order to allow the initiation of transcription. Transcription proceeds until the polymerase meets a stretch of U residues, which acts as a polyadenylation signal (Robertson, Schubert & Lazzarini, 1981). This polyadenylation is carried out by the viral polymerase rather than a cellular poly(A) polymerase.

vRNA replication is carried out in two steps: first the negative sense vRNA is replicated into the positive sense cRNA, a full length copy of the vRNA (Hay, Skehel & McCauley, 1980), followed by the replication of cRNA back into vRNA. The replicated cRNA and vRNA are full length viral segments without a 5' cap or polyadenylation and therefore they require a separate initiation and termination mechanism, which is currently not well understood.

Initially the transcription activity of the polymerase predominates in order to produce large amounts of mRNA to generate the viral proteins required for efficient viral replication. Later on the vRNA replication predominates (Hay *et al.*, 1977; Smith & Hay, 1982), which could simply be due to the higher levels of NP present to prevent cleavage of the nascent vRNA by cellular endonucleases. Accumulation of the nuclear export protein (NEP), otherwise known as NS2, has been shown to be involved in the down regulation of transcription and an increase in replication (Robb, Smith & Vreede, 2009).

The newly replicated RNPs must be exported from the nucleus so that they can assemble new virus particles at the cell surface. This export mechanism is not well understood; however, it is believed to involve the binding of M1 to RNPs and also the

association of the NEP, which aids the interaction of RNPs with the nuclear pore complexes (O'Neill, Talon & Palese, 1998).

1.4.4 Assembly and budding

In order for virus assembly and budding to occur the various components of the virus must be made and trafficked to the cell surface. The vRNPs, as described in the previous section, are assembled and exported from the nucleus. The other viral components are translated and processed outside the nucleus. The HA and NA are synthesised in the endoplasmic reticulum. These proteins undergo N-linked core glycosylation within the ER. The HA and NA then undergo further glycan processing in the Golgi apparatus. The process of N-linked glycosylation is reviewed in (Weerapana & Imperiali, 2006).

Influenza viruses bud at the apical plasma membrane of cells (Rodriguez Boulan & Sabatini, 1978), Figure 1.10 step 5. The HA, NA and M2 are transported to the plasma membrane via the Golgi apparatus. Budding is thought to occur at lipid rafts, which are areas of the plasma membrane enriched in cholesterol and glycosphingolipids. It has been found that HA and NA embedded in the plasma membrane are located within these lipid rafts and the M2 ion channel is excluded (Rossman & Lamb, 2011). It is thought that the initiation of budding is caused by the recruitment of M1 by HA and/or NA to the plasma membrane. The M1 layer that forms under the plasma membrane forms an ordered helix (Calder *et al.*, 2010), which the newly replicated RNPs bind to via the NP (Bui, Whittaker & Helenius, 1996; Zhang & Lamb, 1996). The polymerisation of the M1 protein is thought to drive the formation of a new virus particle. After the virus particle has formed it needs to undergo membrane scission to form the virus particle. It is thought that M2 may play

a role in this scission (Rossman *et al.*, 2010a; 2010b). There is also thought to be a clustering of NA at the proximal end of the virus after membrane scission (Calder *et al.*, 2010). Whether this NA plays a role in allowing scission or whether it is there to enhance sialic acid cleavage in a localised area to facilitate virus release is unclear. It is thought that a filamentous morphology is generated either due to excessive polymerisation of M1 protein, generating extended structures or poor membrane scission, which terminates budding, reviewed in (Rossman & Lamb, 2011).

1.4.5 Release

Release of viruses from the cell surface is mediated by the action of the two surface glycoproteins HA and NA, particularly the activity of NA upon the cell surface, which removes terminal sialic acids from cell surface receptors, which could tether the virus to the cell surface, as described in further detail in section 1.3, Figure 1.10 step 6.

1.5 Current influenza infection prevention and therapy

The current approaches for IAV prevention and therapy can be grouped into two broad classes: vaccination and antivirals.

1.5.1 Vaccination

Currently there are two predominant vaccination strategies in use. These involve vaccination with influenza antigens derived from inactivated virus, typically administered as an intra-muscular injection, or vaccination with a live attenuated virus, typically administered via a nasal spray. The most common method for vaccine manufacture currently is the propagation of high growth reassortants of candidate virus strains in embryonated hens' eggs. As well as vaccines produced in eggs there is also a licensed vaccine, which is virus propagated in suspension MDCK cells; however,

the seed viruses used for manufacture are the same egg-propagated strains used for manufacture of the egg-propagated vaccine (Flucelvax, Novartis).

These vaccines currently contain two different IAV subtypes, a H1N1 2009 pandemic virus, a seasonal H3N2 virus and either one or two influenza B viruses, dependent on whether the vaccine is trivalent or quadrivalent. The composition of this vaccine is chosen to closely match the currently circulating viruses detected by the World Health Organization Global Influenza Surveillance and Response System. The vaccine recommendation is reviewed biannually, once a year for both northern and southern hemispheres, with the recommendations for the northern hemisphere in February and the southern hemisphere in September. The current recommendations for both the 2015 southern hemisphere season and the 2015 – 2016 northern hemisphere seasons trivalent vaccine are for an A/California/7/2009 (H1N1)pdm09-like virus, an A/Switzerland/9715293/2013 (H3N2)-like virus and a B/Phuket/3073/2013-like (Yamagata lineage) virus (and a B/Brisbane/60/2008-like [Victoria lineage] virus in addition for quadrivalent vaccines) (World Health Organisation, 2014b; 2015b). The primary purpose of these vaccines is inducing the production of antibodies, which provide a short-lived protection against infection. Consequently the vaccine has to be readministered annually to ensure adequate protection.

There are a number of alternative lesser-used vaccine strategies both in development and use including: the manufacture of antigen from a different source than propagated virus or vaccines designed to give a broader immunity to influenza infection, so termed 'universal vaccines', reviewed in (Wong & Webby, 2013; Krammer & Palese, 2015).

1.5.2 Antivirals

There are a number of IAV antivirals that are approved for use as well as a number in development. The approved antivirals fall broadly into four broad categories: M2 inhibitors, NA inhibitors, arbidol and polymerase inhibitors.

The M2 inhibitors belong to the class of molecules known as adamantanes. There are two adamantine derivatives licensed for anti-influenza use: amantadine and rimantidine. These inhibitors work by blocking the M2 ion channel, preventing acidification of the virion, thereby preventing the release of the RNPs into the cytoplasm. These antivirals have little therapeutic use for recent viruses as the majority of currently circulating viruses have developed resistance, reviewed in (Hayden, 2006). Adamantanes are not currently recommended for use by the WHO as an antiviral due to this high level of resistance (World Health Organisation, 2010).

There have been a number of NA inhibitors developed that have been approved for use either as a prophylactic treatment or to treat infection. These include the drugs oseltamivir (Tamiflu), zanamivir (Relenza) and the more recently approved peramivir (Rapivab) and laninamivir (Inavir). All of these NA inhibitors are designed to be transition state analogues of sialic acid containing glycans, which bind to the NA with low nM affinity, reviewed in (Itzstein, 2007). These inhibitors act as competitive inhibitors directly occupying the sialidase active site. This inhibition of sialidase activity causes viral aggregation by preventing the viruses from detaching from each other and the cell surface (Palese & Compans, 1976; Baigent, Bethell & McCauley, 1999).

Arbidol is an influenza therapeutic currently licensed only in Russia. Arbidol is thought to act by binding to the HA2, stabilising the HA, therefore preventing the

conformational change of HA associated with membrane fusion (Leneva *et al.*, 2009). The WHO currently does not recommend the use of arbidol, owing to a lack of sufficient clinical trial data (World Health Organisation, 2010).

There is a single RNA polymerase inhibitor which has recently been licensed for use in Japan, Favipiravir, reviewed in (Furuta *et al.*, 2013).

1.6 Balance between activities of haemagglutinin and neuraminidase

It has been hypothesized that the balance between the action of the two antagonistic glycoproteins, the receptor binding HA and the receptor destroying NA, can modulate virus infectivity and therefore transmissibility. A number of studies have shown the interrelated nature of the activities of the two proteins, reviewed in (Wagner *et al.*, 2000).

1.6.1 Passage of viruses with altered NA sialidase activity acquire compensatory mutations in the HA

There are a number of studies which have propagated IAVs in the presence of NA inhibitors, which led to escape mutations which alter the HA receptor binding (McKimm-Breschkin *et al.*, 1996; Blick *et al.*, 1998; Bantia *et al.*, 1998). These studies imply that reduction of NA activity can be compensated for by reducing HA binding, therefore facilitating virus release.

Rudneva *et al.* showed that when the NA from an H1N1 virus is reassorted with a number of different HA subtypes it can lead to a reduction in the propagation efficiency of virus in embryonated eggs due to the formation of aggregated virus (Rudneva *et al.*, 1993). Some of these low efficiency viruses were passaged to rescue higher propagation characteristics (Rudneva *et al.*, 1996). These viruses were found to have substitutions in their HA which reduced the overall receptor binding (Kaverin

et al., 1998), showing that the reduction of NA activity can be compensated for by a reduction of HA receptor binding, restoring some viral fitness.

Hughes *et al.* selected a virus that lacks NA activity by passaging the virus in the presence of anti-NA antibodies and supplementation with bacterial NA. This virus had gained a deletion in its open reading frame and also acquired a stop codon, which truncated the NA protein to little more than the cytoplasmic tail and the transmembrane domain. This virus was passaged in the presence of bacterial NA at reducing concentrations until a number of NA independent viruses were selected. These viruses had gained a number of HA substitutions which had greatly reduced the HA receptor binding (Hughes *et al.*, 2000).

1.6.2 Viruses reassorted with altered NA stalk require altered HA binding properties for efficient replication.

Castrucci and Kawaoka generated a number of viruses with both increased and decreased stalk lengths. There was a general pattern seen: the NA stalk deletion mutants had much less efficient elution of virus from erythrocytes and the stalk insertion mutants eluted faster. Stalk deletion viruses had impaired virus replication in eggs although there was no alteration of virus replication in tissue culture. A virus with the whole NA stalk removed, SD0 (24 amino acid deletion), caused less efficient infections when inoculated into mice (Castrucci & Kawaoka, 1993). Later work by Mitnaul *et al.* using the same stalk deletion virus, SD0, adapted the virus to propagation in eggs by serial passage. The resulting viruses had undergone a number of alterations. Some virus clones had acquired insertions in the stalk region of the NA, thus elongating the stalk to become more like the wild-type WSN virus. These viruses appeared to have obtained the insertion from the PB2, PB1 or NP. Some of the viruses

had also gained HA substitutions which reduced the virus avidity for binding to sialoglycopolymers (Mitnaul, Matrosovich & Castrucci, 2000).

Luo *et al.* also generated a panel of viruses with a number of different stalk deletions and insertions. Viruses with 18 and 28 residues deleted from the stalk of WSN virus had impaired virus replication in MDCK cell culture (Luo, Chung & Palese, 1993). This does not completely agree with experiments described above by Castrucci *et al.*, which showed little change in propagation characteristics of similar stalk deletion viruses in MDCK cells (Castrucci & Kawaoka, 1993). The reason for the difference in results could be due to slight differences in the stalk deletions made and also a difference in the MDCK line used, as receptor density could be an issue in carrying out these experiments, as Luo *et al.* do not see any differences in MDBK cells, which are likely to have different surface glycans than MDCK cells.

Wagner *et al.* generated a panel of reverse genetics viruses with different levels of glycosylation around the RBS on the HA as well as a more active N2 NA and a less active N1 NA. Viruses that contained the less active NA and had fewer glycosylations around the RBS saw a greater reduction in virus replication efficiency when compared to viruses with the more active N2. This is most likely due to the fact that the higher level of HA receptor binding caused by the loss of glycosylation is compensated for by the higher NA activity. Similar patterns were also seen of virus elution from erythrocytes and MDCK cells (Wagner *et al.*, 2000).

Baigent and McCauley also examined a panel of viruses with differences in HA glycosylation and NA activity, mediated by altering stalk length. It was found that in order for viruses lacking a glycosylation near the RBS to replicate efficiently there was a requirement for a long stalk NA for efficient replication in order to compensate

for the increased HA receptor binding. Viruses with inadequate NA activity, i.e. viruses lacking the glycosylation and with a short stalk NA, could have their replication enhanced by the addition of exogenous bacterial NA (Baigent & McCauley, 2001).

1.7 Non-specific inhibitors of influenza infection

There are a number of non-specific inhibitors of influenza replication. These typically act by binding to the HA, blocking receptor binding, reviewed in (Matrosovich & Klenk, 2003).

Non-specific inhibitors have been split into a number of different classes based on their particular properties (Hoyle, 1968). α (Francis) inhibitors are heat stable sialic acid containing mucoproteins, which are cleavable by NA. β (Chu) inhibitors are heat labile serum proteins, most likely to be antibodies or collagenous lectins (collectins). γ inhibitors are heat stable sialic acid containing proteins and are particularly prevalent in guinea pig and horse sera. They have been primarily linked to the protein α2-macroglobulin, which contains high levels of 4-O-acetyl-N-acetyl-neuraminic acid (Hanaoka *et al.*, 1989; Pritchett & Paulson, 1989; Ryan-Poirier & Kawaoka, 1991). This particular form of sialic acid can be bound by the HA, of some viruses, but not cleaved by the NA. C inhibitors are found in rat serum, however, they are not well characterised.

Mucus has long been known to inhibit influenza virus binding, shown by the inhibition of red blood cell binding (Fazekas de St Groth, 1952), and it is thought to be an important factor when IAVs infect the respiratory tract of humans. Mucus is secreted by the goblet cells in the respiratory epithelium, reviewed in (Thornton, Rousseau & McGuckin, 2008). Mucins are heavily O-glycosylated proteins and can

take a number of different forms including membrane bound, and secreted. They are able to form large multimeric structures, reviewed in (Perez-Vilar & Hill, 1999). The glycosylation of human mucins is thought to consist primarily of glycans terminated with α2,3-linked sialic acid (Lamblin *et al.*, 1984a; 1984b; Breg *et al.*, 1987; Royle *et al.*, 2008), however there have been reports of higher levels of α2,6-linked sialic acid present in porcine mucins, measured by binding of linkage specific lectins (Yang *et al.*, 2014). The specific linkages present in the human respiratory mucus remain unclear, as do their effective concentrations at the site of infection.

The role of NA in preventing non-specific inhibition of IAV infection has been seen by infecting human airway epithelial cell cultures, with active goblet cells that provide mucus conditions similar to that of the human airway. When the NA of viruses was inhibited there was much less efficient infection of these airway cultures, thought to be due to the virus particles being unable to remove bound mucus, therefore halting their progression towards the cell surface (Matrosovich *et al.*, 2004b). Cohen *et al.* also identified that IAVs bind to the layer of mucus above the airway epithelium of human trachea sections in a sialic acid dependent manner. Supplementing infected tissue culture samples with mucus also caused a reduction in virus replication in a mucus-dependent manner and the viral NA was capable of cleaving sialic acid from the mucus used (Cohen *et al.*, 2013). Biophysical studies that monitored the movement of fluorescently labelled viruses by microscopy found that swine viruses have a large reduction in their diffusion coefficient when porcine respiratory mucus was present, indicating a slower diffusion rate. The diffusion rate was further reduced upon the addition of NA inhibitors and increased when supplemented with exogenous NA (Yang *et al.*, 2014). Mucus inhibition of viral

infection has also been found to be lower when a virus possesses a longer stalk NA (Blumenkrantz *et al.*, 2013), again implying that increased effective NA activity enhances mucus cleavage and so the NA can play a role in the initiation of infection.

1.8 Human influenza A virus circulation history

It is believed that there have been a large number of influenza pandemics stretching back for many centuries (Taubenberger & Morens, 2009). There have been four well-documented pandemics of the past 100 years: 1918 H1N1, 1957 H2N2, 1968 H3N2 and 2009 H1N1.

1.8.1 Previous pandemics: 1918 - 2009

The 1918 H1N1 (Spanish 'flu) pandemic is thought to have been the most devastating of the past 150 years. The virus was thought to have killed ~50 million people worldwide (Taubenberger & Morens, 2006). The pandemic occurred in three separate waves with mortality rates estimated at ~3%, compared to the usual <0.1% for seasonal infections (Klenk, Garten & Matrosovich, 2011). The genetic origin of this pandemic is described in greater detail in the following section.

The 1957 H2N2 (Asian 'flu) pandemic emerged from a reassortment of the 1918 H1N1 virus lineage with an avian lineage, gaining the H2 HA, N2 NA and also PB1 from an avian source, with the remaining gene segments coming from the endemic H1N1 descendants of the 1918 pandemic (Scholtissek *et al.*, 1978; Kawaoka, Krauss & Webster, 1989). This pandemic is thought to have caused ~1 million deaths, making it less severe than the 1918 pandemic (Klenk, Garten & Matrosovich, 2011).

The 1968 H3N2 (Hong Kong 'flu) was caused by a reassortment of the endemic H2N2 virus with an avian H3 HA containing virus. The new H3N2 reassorted to gain a

new HA (Scholtissek *et al.*, 1978) and also a new PB1 segment (Kawaoka, Krauss & Webster, 1989). These viruses caused relatively mild infections compared to the two preceding pandemic strains (Klenk, Garten & Matrosovich, 2011). These H3N2 viruses continue to circulate to this day as seasonal strains, although they have undergone extensive antigenic drift in the intervening time.

The 2009 H1N1 pandemic, unlike the previous pandemics detailed above, was caused by a virus of swine rather than avian origin. It arose as a reassortment between two swine viruses (Garten *et al.*, 2009), as described in greater detail in the following section.

1.8.2 H1N1 circulation history

H1N1 viruses are known to have been circulating for much of the previous century in a number of different hosts including birds, pigs and humans. The avian reservoir of H1N1 is believed to have been constantly present for the previous century, leading to a number of new lineages by transferal into new hosts (Zimmer & Burke, 2009).

1.8.2.1 Human H1N1 1918 – 1957 and 1977 – 2009

H1N1 was responsible for causing the 1918 pandemic. This pandemic virus persisted as one of the major human seasonal viruses until the H2N2 pandemic appeared in 1957. These viruses are thought to be responsible for infecting pigs, leading to the 'classical swine' lineage of H1N1 viruses (Brown, 2000). The sequence of the 1918 virus HA has been determined from RNA extracted from both formalin fixed tissue (Taubenberger *et al.*, 1997) and from lung tissue preserved in permafrost (Reid *et al.*, 1999). This virus was reconstructed by gene synthesis and reverse genetics (Tumpey *et al.*, 2005). Analysis of these sequences suggest that the virus was

likely to have emerged from an avian source although it has been suggested that this may have not been a direct transfer, as fragments of RNA sequence extracted from preserved birds in 1917 are not very similar to the determined 1918 sequences (Fanning *et al.*, 2002). Computer modeling has also suggested that the avian virus that caused the 1918 pandemic moved into a mammalian host earlier than 1918, suggesting that it was not a direct avian-human zoonosis (Reis, Hay & Goldstein, 2009).

After the apparent disappearance of H1N1 viruses circulating in humans in 1957, upon the emergence of H2N2 viruses, there were a number of years where H1N1 viruses were not isolated from humans. H1N1 viruses reemerged in 1977. These H1N1 viruses co-circulated with H3N2 seasonal viruses until 2009 when the swine-origin pandemic H1N1 predominated. The reemergence of these H1N1 viruses is thought to be an accidental release of virus, as there is high sequence similarity between the newly emerged viruses and those isolated just prior to the 1957 pandemic (Scholtissek, Hoyningen & Rott, 1978; Nakajima, Desselberger & Palese, 1978; Raymond *et al.*, 1983). It has been suggested that the release of this virus could be a result of the improper inactivation of virus in a vaccine trial.

1.8.2.2 Classical and Eurasian swine viruses

There have been two predominant lineages of swine H1N1 viruses in the past century: the classical swine and Eurasian swine viruses.

Viruses from the classical swine lineage have a closely related origin to the 1918 pandemic H1N1 viruses and have been circulating in pigs in North America since that time (Brown, 2000). A classical swine virus was responsible for an outbreak in humans in 1976 at a US military base, Fort Dix (Goldfield *et al.*, 1977).

The virus was identified to be similar to the earliest known H1N1 virus isolates of the time, descended from the 1918 pandemic, consequently there was a mass vaccination campaign in the USA, mounted to attempt to prevent a further pandemic (Sencer & Millar, 2006).

Eurasian swine viruses are believed to have emerged directly from birds in around 1979 (Brown, 2000). These viruses retain a number of avian virus characteristics such as a general preference for α2,3-linked sialic acid, although this preference has reduced over time (Matrosovich *et al.*, 2000). These viruses circulate in pigs in Europe and Asia to the present day.

1.8.2.3 Triple-reassortant swine viruses

From 1998 onwards triple reassortant swine (TRS) viruses were found in pigs in the USA (Zhou *et al.*, 1999; Webby *et al.*, 2000). This virus was made up of a mixture of segments from avian, human and swine viruses. The PB2 and PA segments were from an avian H1N1 lineage, the PB1, HA and NA segments were from a human H3N2 lineage and the NP, M and NS segments were from the classical swine H1N1 lineage, see Figure 1.11. This TRS virus reassorted again with the H1N1 classical swine virus to form a number of different subtype viruses: H1N1, H1N2, H3N1 and H3N2. These viruses became enzootic in North American swine populations (Vincent *et al.*, 2008; Olsen, 2002; Garten *et al.*, 2009) and have been responsible for a number of zoonotic infections, reviewed in (Shinde *et al.*, 2009).

1.8.2.4 2009 pandemic H1N1

In 2009 a new human pandemic virus emerged, which was found to be a reassortant of a TRS virus of subtype H1N1 or H1N2 with an H1N1 swine virus from the Eurasian lineage. The NA and M genes came from the Eurasian swine virus with

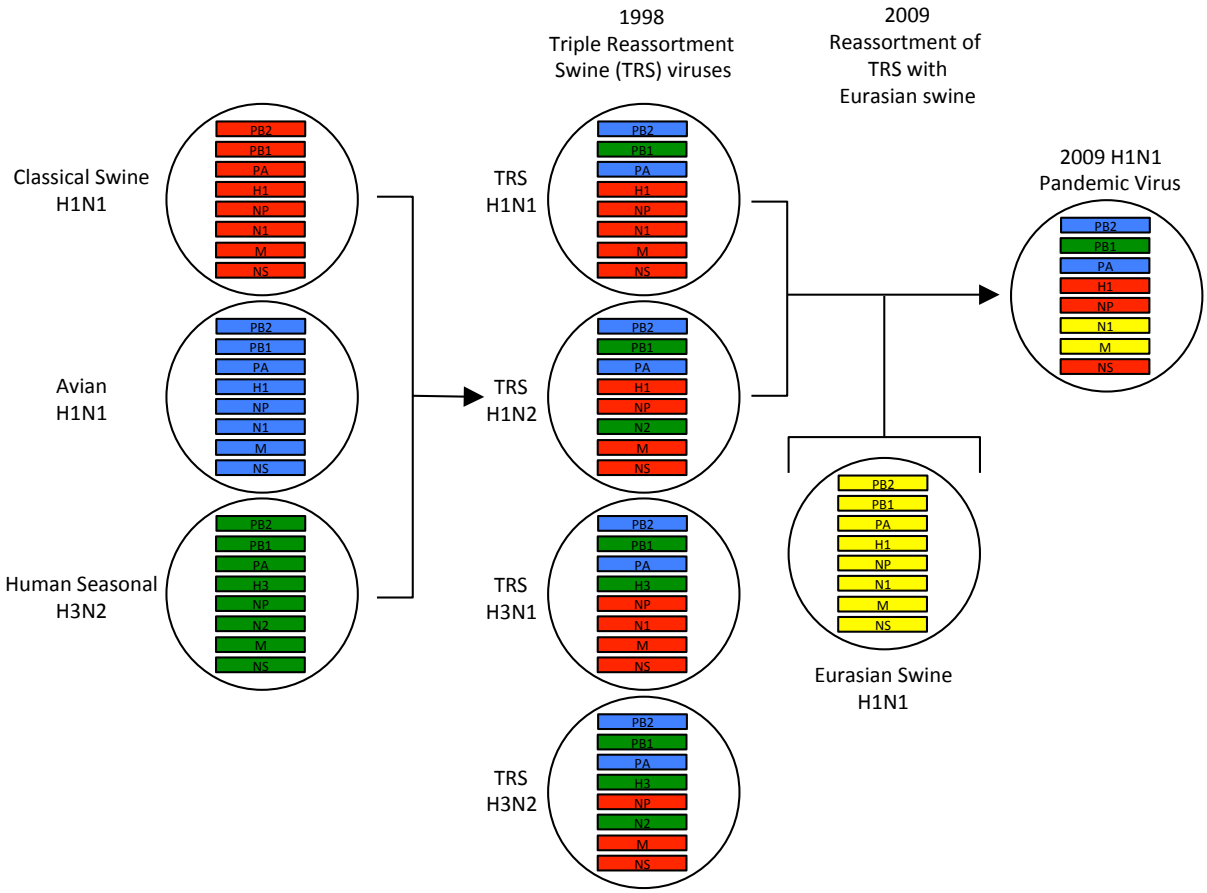


Figure 1.11: Reassortment events that led to the emergence of the 2009 H1N1 pandemic. In 1998 classical swine H1N1, avian H1N1 and human seasonal H3N2 viruses reassorted to produce H1N1, H1N2, H3N1 and H3N2 triple reassortant swine (TRS) viruses. In 2009 a reassortment of these TRS viruses with an Eurasian swine virus led to the 2009 H1N1 pandemic viruses.

the other six segments from the TRS virus (Garten *et al.*, 2009), see Figure 1.11. These 2009 pandemic viruses (pdm09) cause widespread infections, leading to the classification of the pdm09 viruses as a pandemic. The pdm09 viruses infected an estimated 60 million people between April 2009 and April 2010 in the USA alone, with an estimated 12500 deaths, which gives a death rate of ~0.02%, which is low in comparison with typical seasonal strains (Shrestha *et al.*, 2011).

There have been a number of studies that have examined the receptor binding of pdm09 and/or pre-pandemic TRS viruses. Both glycan microarray and other solid phase binding experiments have identified a predominance of affinity for α2,6-linked receptors in both the pdm09 and TRS H1 HAs, with little or no binding measurable for α2,3-linked receptors (Childs *et al.*, 2009; Maines *et al.*, 2009; Yang, Carney & Stevens, 2010; Chen *et al.*, 2011; Yen *et al.*, 2011; Bradley *et al.*, 2011; de Vries *et al.*, 2011; Xu *et al.*, 2010; 2012b). Glycan microarray experiments showed a similar binding profile for both TRS and pdm09 viruses (Childs *et al.*, 2009; Chen *et al.*, 2011; Yen *et al.*, 2011; Bradley *et al.*, 2011). de Vries *et al.* and Xu *et al.* found that the binding affinity of purified pdm09 HA was significantly weaker than that of purified TRS H1 HA (de Vries *et al.*, 2011; Xu *et al.*, 2012b). de Vries *et al.* attributed the difference seen in binding to the HA residues 200 and 227 (H3 Numbering) (de Vries *et al.*, 2011). Xu *et al.* attributed the differences they measured in HA binding to residues 134, 219 and 227 (H3 Numbering). This study also suggested that the emergence of the pandemic was due to the balance between the low binding pdm09 HA and the acquired low activity N1 NA from Eurasian swine (Xu *et al.*, 2012b). There is a difference in the receptor binding data between the pdm09 and TRS HA dependent on whether the experiments are done with purified proteins or virus. There are also a number of

other substitutions present in the H1 HA that could alter receptor binding but these were not investigated in the studies detailed above.

1.8.3 Major zoonotic viruses

As well as human seasonal and pandemic strains there have been a number of viruses that have transferred from animal reservoirs into the human population but which have not been capable of sustained human-to-human transmission. Two major recent zoonotic viruses are H5N1 and H7N9 viruses that have been found to infect humans through direct zoonoses from their reservoir host of wild and domesticated birds.

1.8.3.1 H5N1

High pathogenicity avian influenza (HPAI) H5N1 viruses were first found infecting humans in 2003 and are still causing a large number of infections to this day, with 842 confirmed cases worldwide and at least 447 mortalities, as of 23rd June 2015, giving a high case fatality ratio of ~55% (World Health Organisation, 2015a). These viruses remain a public health concern owing to their perceived pandemic potential, although there has been limited evidence of human-to-human transmission.

These viruses are classified as high pathogenicity as they cause mortality in poultry and also owing to the presence of a polybasic cleavage site in the HA, which allows cleavage of HA0 into HA1 and HA2 by a wider range of proteases, as described in section 1.3.1.1. H5N1 infection in humans is thought to be a primarily LRT infection, owing to the virus receptor binding being primarily to α2,3-linked sialic acid, as described in section 1.3.1.2.2. The location of these infections in the LRT is thought to be one of the reasons that these viruses have such a high mortality rate owing to the

higher incidence of pneumonia associated with H5N1 infection (Abdel-Ghafar *et al.*, 2008).

1.8.3.2 H7N9

H7N9 viruses were first found infecting humans in February 2013 in eastern China. These viruses were found to be a reassortment with the HA similar to that of a duck H7N3 virus, the NA was similar to that of an H7N9 virus found in wild birds and the remaining six internal genes from an H9N2 virus (Gao *et al.*, 2013). Until 23rd June 2015 there had been 672 confirmed cases with at least 271 deaths, giving a case fatality ratio of ~40%. There has been a wave of H7N9 every year since 2013, with the majority of infections taking place in the winter in both 2014 and 2015. These viruses are thought to have infected as direct zoonoses from avian populations. There is limited evidence of human-to-human transmission with reports of a small number of clusters of infected individuals (World Health Organisation, 2015c). These viruses also have poor aerosol transmission in ferrets (Zhu *et al.*, 2013a; Belser *et al.*, 2013; Richard *et al.*, 2013; Xu *et al.*, 2014a). These H7N9 viruses are now thought to be enzootic in the avian population of eastern China (Lam *et al.*, 2015) and are consequently a major risk to the human population.

These H7N9 viruses are quite different from the HPAI H5N1 viruses described above. They do not have the typical hallmarks of a high pathogenicity virus, lacking a polybasic cleavage site in the HA (Gao *et al.*, 2013) they also are of low pathogenicity in avian species. These viruses do, however, have a Q226L substitution present in the HA, which is associated with an increase in binding of the HA to the human α2,6-linked sialic acid (Rogers *et al.*, 1983), and also a five amino acid deletion in the NA stalk, which can alter the transmission characteristics, as described in section 1.3.2.4.

There have been a number of studies that have examined the structure and/or receptor binding properties of the these viruses, either examining virus binding (Xiong *et al.*, 2013b; Shi *et al.*, 2013; Belser *et al.*, 2013; Zhou *et al.*, 2013; Xu *et al.*, 2014b) or the binding of purified HA to immobilised receptors (Yang *et al.*, 2013; Shi *et al.*, 2013; Dortmans *et al.*, 2013; Ramos *et al.*, 2013). These studies have a broad agreement that the H7N9 viruses have a preference for binding to α2,3-linked sialic acid receptors, as would be expected for an avian virus; however, they also have increased affinity for α2,6-linked receptors in comparison with other avian viruses, although it is weaker than that of the α2,3-linked receptor binding.

The NA of the H7N9 virus has also been characterised, showing that the N9 NA has a poor catalytic efficiency at cleaving an artificial NA substrate (MUNANA) (Wu *et al.*, 2013). It has also been described that the N9 NA appears to have a low efficiency at cleaving α2,6-linked receptors (Xiong *et al.*, 2013b).

1.9 Summary

One of the major aims in IAV virology is attempting to anticipate which viruses will cause the next pandemic and attempt to either prevent the movement of this virus into the human population or prepare prophylactic or therapeutic approaches to control the virus infection. The defining factor in predicting the viral spread into and between humans is the transmissibility. This is a complex and multifactorial property that involves a large number of viral characteristics such as the fitness of the virus to replicate in human cells and the ability for efficient virus shedding and transmission to a new human host.

A major factor in determining viral fitness and therefore transmissibility is the initial stage of infection, receptor binding. This binding is controlled by both the HA

and NA surface glycoproteins which have antagonistic functions of receptor binding and receptor destroying. It has been hypothesised that for efficient virus infection and release from the cell there needs to be the correct balance between the functions of these two glycoproteins. Until now studies into the balance of these two proteins has involved altering the characteristics of one of the proteins and looking at escape mutations induced or looking at the effects that these changes have on virus characteristics. There have also been studies that have examined the two glycoproteins in isolation that then attempt to relate the antagonistic activities of the two proteins by correlating characteristics. There is consequently a requirement to understand the role of these two proteins in directly altering receptor binding in the context of a viral particle.

1.10 Objectives of thesis

This thesis aims to investigate the question of the balance between HA and NA by directly measuring the role of these two proteins in affecting virus binding to a receptor analogue coated surface by a biophysical technique. The methods developed to measure these characteristics are then used, in combination with a number of other biophysical techniques, to better understand a number of different combinations and properties of HA and NA, namely the properties of the H7N9 NA with respect to its role in receptor binding and also the novel combination of HA and NA seen upon the emergence of the 2009 H1N1 pandemic.

2 Materials and Methods

2.1 Materials

2.1.1 Viruses

Virus Name	Short Name	Origin	Description
X-31	-	NIMR WHOCC	A/Hong Kong/2/68(H3N2) high growth reassortant with PR8 background
X-31 HAM	HAM	NIMR WHOCC	X-31 L226Q Mutant
X-31 Δ10 NA	Δ10 NA	Reverse Genetics*	X-31 with residues 60-69 removed from NA stalk
A/Anhui/1/2013 (H7N9)	Anhui13	Reverse Genetics*	H7N9 6+2 reassortant with PR8
A/Anhui/1/2013 (H7N9) S367N NA	Anhui13 S367N	Reverse Genetics*	H7N9 6+2 reassortant with PR8 and S367N mutant NA
A/Anhui/1/2013 HA + PR8 NA Reassortant	H7N1	Reverse Genetics*	H7N1 7+1 reassortant with PR8
PR8 HA + A/Anhui/1/2013 NA Reassortant	H1N9 WT	Reverse Genetics*	H1N9 7+1 reassortant with PR8
PR8 HA + A/Anhui/1/2013 NA S367N Reassortant	H1N9 S367N	Reverse Genetics*	H1N9 7+1 reassortant with PR8 and S367N mutant NA
A/Victoria/361/2012 HA + A/Anhui/1/2013 NA Reassortant	H3N9 WT	Reverse Genetics*	H3N9 6+2 reassortant with PR8
A/Victoria/361/2012 HA + A/Anhui/1/2013 NA S367N Reassortant	H3N9 S367N	Reverse Genetics*	H3N9 6+2 reassortant with PR8 and S367N mutant NA
A/California/7/09 WT	Cal7 WT	Reverse Genetics*	Full 8 segment Cal7 (H1N1)
A/California/7/09 S132T	Cal7 S132T	Reverse Genetics*	Full 8 segment Cal7 with S132T mutant HA
A/California/7/09 K134R	Cal7 K134R	Reverse Genetics*	Full 8 segment Cal7 with K134R mutant HA
A/California/7/09 R149K	Cal7 R149K	Reverse Genetics*	Full 8 segment Cal7 with R149K mutant HA
A/California/7/09 P186S	Cal7 P186S	Reverse Genetics*	Full 8 segment Cal7 with P186S mutant HA
A/California/7/09 A189T	Cal7 A189T	Reverse Genetics*	Full 8 segment Cal7 with A189T mutant HA
A/California/7/09 K208R	Cal7 K208R	Reverse Genetics*	Full 8 segment Cal7 with K208R mutant HA
A/California/7/09 I219T	Cal7 I219T	Reverse Genetics*	Full 8 segment Cal7 with I219T mutant HA
A/California/7/09 I219A	Cal7 I219A	Reverse Genetics*	Full 8 segment Cal7 with I219A mutant HA
A/California/7/09 E227A	Cal7 E227A	Reverse Genetics*	Full 8 segment Cal7 with E227A mutant HA
A/California/7/09 I219A E227A	Cal7 I219A E227A	Reverse Genetics*	Full 8 segment Cal7 with I219A E227A mutant HA

A/California/7/09 I219T E227A	Cal7 I219T E227A	Reverse Genetics*	Full 8 segment Cal7 with I219T E227A mutant HA
A/Iowa/1/2006 WT	Iowa06 WT	Reverse Genetics*	Iowa06 HA and NA in a PR8 Background
A/Iowa/1/2006 T132S	Iowa06 T132S	Reverse Genetics*	Iowa06 T132S HA and NA in a PR8 Background
A/Iowa/1/2006 K134R	Iowa06 K134R	Reverse Genetics*	Iowa06 K134R HA and NA in a PR8 Background
A/Iowa/1/2006 R149K	Iowa06 R149K	Reverse Genetics*	Iowa06 R149K HA and NA in a PR8 Background
A/Iowa/1/2006 P186S	Iowa06 P186S	Reverse Genetics*	Iowa06 P186S HA and NA in a PR8 Background
A/Iowa/1/2006 T189A	Iowa06 T189A	Reverse Genetics*	Iowa06 T189A HA and NA in a PR8 Background
A/Iowa/1/2006 K208R	Iowa06 K208R	Reverse Genetics*	Iowa06 K208R HA and NA in a PR8 Background
A/Iowa/1/2006 A219I	Iowa06 A219I	Reverse Genetics*	Iowa06 A219I HA and NA in a PR8 Background
A/Iowa/1/2006 A219T	Iowa06 A219T	Reverse Genetics*	Iowa06 A219T HA and NA in a PR8 Background
A/Iowa/1/2006 A227E	Iowa06 A227E	Reverse Genetics*	Iowa06 A227E HA and NA in a PR8 Background
A/Iowa/1/2006 A219I A227E	Iowa06 A219I A227E	Reverse Genetics*	Iowa06 A219I A227E HA and NA in a PR8 Background
A/Iowa/1/2006 A219T A227E	Iowa06 A219T A227E	Reverse Genetics*	Iowa06 A219T A227E HA and NA in a PR8 Background
A/Iowa/1/2006 P186S T189A	Iowa06 P186S T189A	Reverse Genetics*	Iowa06 P186S T189A HA and NA in a PR8 Background
A/Iowa/1/2006 WT + Iowa Bkg	Iowa06 + Iowa bkg	Reverse Genetics*	Full 8 segment WT Iowa06
A/California/7/2009 HA A/Iowa/1/2006 NA + Cal7 Bkg	Cal7 HA Iowa06 NA + Cal bkg	Reverse Genetics*	WT Cal7 HA and Iowa06 NA in a Cal7 Background
A/California/7/2009 HA A/Iowa/1/2006 NA + Iowa06 Bkg	Cal7 HA Iowa06 NA + Iowa06 bkg	Reverse Genetics*	WT Cal7 HA and Iowa06 NA in an Iowa06 Background
A/WSN/33 + A/California/7/2009 NA	WSN + Cal7 NA	Reverse Genetics*	7 + 1 reassortant of WSN with Cal7 NA
A/WSN/33 + A/Iowa/1/2006 NA	WSN + Iowa06 NA	Reverse Genetics*	7 + 1 reassortant of WSN with Iowa06 NA

* Generated in these studies

2.1.2 Purified proteins

Protein Name	Production/Purification Method	Protein Source
X-31 NA	Bromelain cleaved from purified virus. Purified by ion exchange and gel filtration.	Patrick Collins
Cal7 HA	Insect cell expressed. Purified using cobalt column and gel filtration. Purification tag remains attached.	Generated during these studies
Iowa06 HA	Insect cell expressed. Purified using cobalt column and gel filtration. Purification tag remains attached.	Generated during these studies
Cal7 NA	Bromelain cleaved from purified virus. Purified by ion exchange and gel filtration.	Generated during these studies
Iowa06 NA	Detergent extracted from purified virus. Purified by ion exchange and gel filtration. Made into rosettes by removing detergent with hydrophobic beads.	Generated during these studies
Anhui13 NA WT	Insect cell expressed. Purified using cobalt column and gel filtration. Purification tag remains attached.	Generated during these studies
Anhui13 NA S367N	Insect cell expressed. Purified using cobalt column and gel filtration. Purification tag remains attached.	Generated during these studies

2.1.3 PCR primers

2.1.3.1 Primers used for reverse genetics cloning

Target	Primer Name	Sequence
Full Length HA + AarI restriction sites	AarI-HA-F1	TATTCACCTGCCTCAGGGAGCAAAAGCAGGGG
	AarI-NS-R890	ATATCACCTGCCTCGTATTAGTAGAAACAAGGGTGTTTT
Full length NA + AarI restriction sites	AarI-NA-F1	TATTCACCTGCCTCAGGGAGCAAAAGCAGGAGT
	AarI-NA-R1413	ATATCACCTGCCTCGTATTAGTAGAAACAAGGAGTTTTTT
Full length HA + BsmBI restriction sites	Bm-HA-F1	TATTCGTCTCAGGGAGCAAAAGCAGGGG
	Bm-NS-R890	ATATCGTCTCGTATTAGTAGAAACAAGGGTGTTTT
Cal7 HA splicing PCR	cswHAR873	CAGATCCAGCATTTCTTTCCATTGC
	cswHAF848	GCAATGGAAAGAAATGCTGGATCTG
	HAFUc	TATTCGTCTCAGGGAGCAAAAGCAGGGG
	HARUc	ATATCGTCTCGTATTAGTAGAAACAAGGGTGTTTT
Universal IAV primer	Uni12	AGCAAAAGCAGG

2.1.3.2 Primers for baculovirus subcloning

Target	Primer Name	Sequence
Anhui NA expression construct	N9_infusion_F	GGTCCGAAAACCATGGATGCTACTAGTAAATCAGTCACACCAAGGCTTCAATAA
	N9_infusion_R	GCTCGAATTCGGATCCTCATTAGAGGAAGTACTCGATCTTAGCGCCGTCA

2.1.3.3 Primers for viral RNA sequencing

2.1.3.3.1 H3N2 sequencing

X-31 sequencing

Target	Primer Name	Sequence
5' HA sequencing	HAFUc*	TATTCGTCTCAGGGAGCAAAAGCAGGGG
	H3A1R1*	GTCTATCATTCCCTCCCAACCATT
3' HA sequencing	H3A1F3*	TGCATCACTCCAAATGGAAGCATT
	HARUc*	ATATCGTCTCGTATTAGTAGAAACAAGGGTGTTTT
5' NA sequencing	NAFUc*	TATTGGTCTCAGGGAGCAAAAGCAGGAGT
	N2R3*	TTAGTATCAGCTCTTCCTGAAGCA
3' NA sequencing	N2F2*	CATGCGATCCTGGCAAGTGTTATC
	NARUc*	ATATGGTCTCGTATTAGTAGAAACAAGGAGTTTTTT

* Primers used in one-step RT-PCR

Vic361 HA sequencing

Target	Primer Name	Sequence
5' HA sequencing	HAFUc*	TATTCGTCTCAGGGAGCAAAAGCAGGGG
	H3A1R2	AATGCTTCCATTTGGAGTGATGCA
	H3A1R1*	GTCTATCATTCCCTCCCAACCATT
3' HA sequencing	H3A1F3*	TGCATCACTCCAAATGGAAGCATT
	HARUc*	ATATCGTCTCGTATTAGTAGAAACAAGGGTGTTTT

* Primers used in one-step RT-PCR

2.1.3.3.2 H7N9 sequencing

Target	Primer Name	Sequence
5' HA sequencing	H7HAF1*	AAAAGCAGGGGATACAAAATGAA
	H7HAR2	GGAGCTATGAARGCCCCATTGAA
	H7R3*	ATTGCCGATTGRGTGCTTTTGTA
3' HA sequencing	H7F3*	TCTGGAAGRATTGACTTTCAYTGG
	HARUc*	ATATCGTCTCGTATTAGTAGAAACAAGGGTGTTTT
5' NA sequencing	N9F1*	AAGCAGGGTCAAGATGAATCC
	N9R3*	GTGCTTATTGTYCTYCCTAGCCA
3' NA sequencing	N9F3*	CCAGGATGTCAATATGTATATC
	NARUc*	ATATGGTCTCGTATTAGTAGAAACAAGGAGTTTTTT

* Primers used in one-step RT-PCR

2.1.3.3.3 H1N1 sequencing

Cal7 sequencing

Target	Primer Name	Sequence
5' HA sequencing	cswHAF1*	TATTCGTCTCAGGGAGCAAAAGCAGGGG
	cswHAR873	CAGATCCAGCATTTCTTTCCATTGC
	cswHAR1264*	CCTACTGCTGTGAACTGTGTATTC
3' HA sequencing	cswHAF848*	GCAATGGAAAGAAATGCTGGATCTG
	HARUc*	ATATCGTCTCGTATTAGTAGAAACAAGGGTGTTTT
5' NA sequencing	NAFUc*	TATTGGTCTCAGGGAGCAAAAGCAGGAGT
	cswN1R1099*	CCTATCCAAACACCATTGCCGTAT
3' NA sequencing	cswN1F401*	GGAATGCAGAACCTTCTTCTTGAC
	NARUc*	ATATGGTCTCGTATTAGTAGAAACAAGGAGTTTTTT

* Primers used in one-step RT-PCR

Iowa06 sequencing

Target	Primer Name	Sequence
5' HA sequencing	IowaF1*	CCAAAATGAAGGCAATAATAGTAGTCT
	IowaHAR873	CAGATCCAGAGCCTCTTTTCATTGC
	IowaR1295*	AAACTCTCTATTCTTTTTTCCAATTG
3' HA sequencing	IowaHAF848*	GCAATGAAAAGAGGCTCTGGATCTG
	HARUc*	ATATCGTCTCGTATTAGTAGAAACAAGGGTGTTTT
5' NA sequencing	NAFUc*	TATTGGTCTCAGGGAGCAAAAGCAGGAGT
	IowaNAR953*	TCACCGAAAACCCCACTGCAT
3' NA sequencing	IowaNAF421*	GGAACCGTCAAGGACAGGAGT
	NARUc*	ATATGGTCTCGTATTAGTAGAAACAAGGAGTTTTTT

* Primers used in one-step RT-PCR

2.1.4 Plasmids

2.1.4.1 Background plasmids for reverse genetics

Plasmid Name	Source	Description
pHW2000	Laboratory of Robert Webster	Empty pHW2000 vector
pHW2000 PR8 PB2	Laboratory of Robert Webster	PR8 PB2 segment cloned into pHW2000 vector
pHW2000 PR8 PB1	Laboratory of Robert Webster	PR8 PB1 segment cloned into pHW2000 vector
pHW2000 PR8 PA	Laboratory of Robert Webster	PR8 PA segment cloned into pHW2000 vector
pHW2000 PR8 NP	Laboratory of Robert Webster	PR8 NP segment cloned into pHW2000 vector
pHW2000 PR8 M	Laboratory of Robert Webster	PR8 M segment cloned into pHW2000 vector
pHW2000 PR8 NS	Laboratory of Robert Webster	PR8 NS segment cloned into pHW2000 vector
pHW2000 WSN PB2	Laboratory of Robert Webster	WSN PB2 segment cloned into pHW2000 vector
pHW2000 WSN PB1	Laboratory of Robert Webster	WSN PB1 segment cloned into pHW2000 vector
pHW2000 WSN PA	Laboratory of Robert Webster	WSN PA segment cloned into pHW2000 vector
pHW2000 WSN NP	Laboratory of Robert Webster	WSN NP segment cloned into pHW2000 vector
pHW2000 WSN M	Laboratory of Robert Webster	WSN M segment cloned into pHW2000 vector
pHW2000 WSN NS	Laboratory of Robert Webster	WSN NS segment cloned into pHW2000 vector
pHW2000 Cal7 PB2	Previously cloned by Haxia Xiao	A/California/7/2009 PB2 segment cloned into pHW2000 vector
pHW2000 Cal7 PB1	Previously cloned by Haxia Xiao	A/California/7/2009 PB1 segment cloned into pHW2000 vector
pHW2000 Cal7 PA	Previously cloned by Haxia Xiao	A/California/7/2009 PA segment cloned into pHW2000 vector
pHW2000 Cal7 NP	Previously cloned by Haxia Xiao	A/California/7/2009 NP segment cloned into pHW2000 vector
pHW2000 Cal7 M	Previously cloned by Haxia Xiao	A/California/7/2009 M segment cloned into pHW2000 vector
pHW2000 Cal7 NS	Previously cloned by Haxia Xiao	A/California/7/2009 NS segment cloned into pHW2000 vector
pHW2000 Iowa06 PB2	Previously cloned by Haxia Xiao	A/Iowa/1/2006 PB2 segment cloned into pHW2000 vector
pHW2000 Iowa06 PB1	Previously cloned by Haxia Xiao	A/Iowa/1/2006 PB1 segment cloned into pHW2000 vector
pHW2000 Iowa06 PA	Previously cloned by Haxia Xiao	A/Iowa/1/2006 PA segment cloned into pHW2000 vector
pHW2000 Iowa06 NP	Previously cloned by Haxia Xiao	A/Iowa/1/2006 NP segment cloned into pHW2000 vector
pHW2000 Iowa06 M	Previously cloned by Haxia Xiao	A/Iowa/1/2006 M segment cloned into pHW2000 vector
pHW2000 Iowa06 NS	Previously cloned by Haxia Xiao	A/Iowa/1/2006 NS segment cloned into pHW2000 vector

2.1.4.2 HA and NA plasmids for reverse genetics

Plasmid Name	Source	Description
pHW2000 PR8 HA	Laboratory of Robert Webster	PR8 HA segment cloned into pHW2000 vector
pHW2000 PR8 NA	Laboratory of Robert Webster	PR8 NA segment cloned into pHW2000 vector
pHW2000 WSN HA	Laboratory of Robert Webster	WSN HA segment cloned into pHW2000 vector
pHW2000 X-31 HA	Cloned from viral cDNA*	X-31 HA segment cloned into pHW2000 vector
pHW2000 X-31 NA	Cloned from viral cDNA*	X-31 NA segment cloned into pHW2000 vector
pHW2000 X-31 NA Δ10	Mutagenesis of WT plasmid*	X-31 NA with residues 60 – 69 deleted cloned into pHW2000 vector
pHW2000 Anhui13 HA	Laboratory of Munir Iqbal	A/Anhui/1/2013 HA segment cloned into pHW2000 vector
pHW2000 Anhui13 NA	Laboratory of Munir Iqbal	A/Anhui/1/2013 NA segment cloned into pHW2000 vector
pHW2000 Anhui13 NA S367N	Mutagenesis of WT plasmid*	A/Anhui/1/2013 S367N NA segment cloned into pHW2000 vector
pHW2000 Vic361 HA	Previously cloned by Lauren Parker	A/Victoria/361/2012 HA segment cloned into pHW2000 vector
pHW2000 Cal7 HA	Cloned from viral cDNA*	A/California/7/2009 HA segment cloned into pHW2000 vector
pHW2000 Cal7 NA	Cloned from viral cDNA*	A/California/7/2009 NA segment cloned into pHW2000 vector
pHW2000 Cal7 HA S132T	Mutagenesis of WT plasmid*	A/California/7/2009 S132T HA segment cloned into pHW2000 vector
pHW2000 Cal7 HA K134R	Mutagenesis of WT plasmid*	A/California/7/2009 K134R HA segment cloned into pHW2000 vector
pHW2000 Cal7 HA K149R	Mutagenesis of WT plasmid*	A/California/7/2009 K149R HA segment cloned into pHW2000 vector
pHW2000 Cal7 HA S186P	Mutagenesis of WT plasmid*	A/California/7/2009 S186P HA segment cloned into pHW2000 vector
pHW2000 Cal7 HA A189T	Mutagenesis of WT plasmid*	A/California/7/2009 A189T HA segment cloned into pHW2000 vector
pHW2000 Cal7 HA R208K	Mutagenesis of WT plasmid*	A/California/7/2009 R208K HA segment cloned into pHW2000 vector
pHW2000 Cal7 HA I219T	Mutagenesis of WT plasmid*	A/California/7/2009 I219T HA segment cloned into pHW2000 vector
pHW2000 Cal7 HA I219A	Mutagenesis of WT plasmid*	A/California/7/2009 I219A HA segment cloned into pHW2000 vector
pHW2000 Cal7 HA E227A	Mutagenesis of WT plasmid*	A/California/7/2009 E227A HA segment cloned into pHW2000 vector
pHW2000 Cal7 HA I219A E227A	Mutagenesis of WT plasmid*	A/California/7/2009 I219A E227A HA segment cloned into pHW2000 vector
pHW2000 Cal7 HA I219T E227A	Mutagenesis of WT plasmid*	A/California/7/2009 I219T E227A HA segment cloned into pHW2000 vector
pHW2000 Iowa06 HA	Cloned from viral cDNA*	A/Iowa/1/2006 HA segment cloned into pHW2000 vector
pHW2000 Iowa06 NA	Cloned from viral cDNA*	A/Iowa/1/2006 NA segment cloned into pHW2000 vector
pHW2000 Iowa06 HA T132S	Mutagenesis of WT plasmid*	A/Iowa/1/2006 T132S HA segment cloned into pHW2000 vector
pHW2000 Iowa06 HA K134R	Mutagenesis of WT plasmid*	A/Iowa/1/2006 K134R HA segment cloned into pHW2000 vector

pHW2000 Iowa06 HA R149K	Mutagenesis of WT plasmid*	A/Iowa/1/2006 R149K HA segment cloned into pHW2000 vector
pHW2000 Iowa06 HA P186S	Mutagenesis of WT plasmid*	A/Iowa/1/2006 P186S HA segment cloned into pHW2000 vector
pHW2000 Iowa06 HA T189A	Mutagenesis of WT plasmid*	A/Iowa/1/2006 T189A HA segment cloned into pHW2000 vector
pHW2000 Iowa06 HA K208R	Mutagenesis of WT plasmid*	A/Iowa/1/2006 K208R HA segment cloned into pHW2000 vector
pHW2000 Iowa06 HA A219I	Mutagenesis of WT plasmid*	A/Iowa/1/2006 A219I HA segment cloned into pHW2000 vector
pHW2000 Iowa06 HA A219T	Mutagenesis of WT plasmid*	A/Iowa/1/2006 A219T HA segment cloned into pHW2000 vector
pHW2000 Iowa06 HA A227E	Mutagenesis of WT plasmid*	A/Iowa/1/2006 A227E HA segment cloned into pHW2000 vector
pHW2000 Iowa06 HA A219I A227E	Mutagenesis of WT plasmid*	A/Iowa/1/2006 A219I A227E HA segment cloned into pHW2000 vector
pHW2000 Iowa06 HA A219T A227E	Mutagenesis of WT plasmid*	A/Iowa/1/2006 A219T A227E HA segment cloned into pHW2000 vector
pHW2000 Iowa06 HA P186S T189A	Mutagenesis of WT plasmid*	A/Iowa/1/2006 P186S T189A HA segment cloned into pHW2000 vector

*Generated in these studies

2.1.4.3 Plasmids for insect cell protein expression

Plasmid Name	Source	Description	Expression System
pHAEM1	Sebastian Vachieri	Altered version of pAcGP67-A baculovirus expression plasmid, altered to include a trimerisation motif (foldon) a TEV protease cleavage site and a 6x His-tag	BaculoGold
pHAEM1 Cal7 HA	Generated in these studies	pHAEM1 with insect optimised gene encoding the ectodomain of Cal7 HA	BaculoGold
pHAEM1 Iowa06 HA	Generated in these studies	pHAEM1 with insect optimised gene encoding the ectodomain of Cal7 HA	BaculoGold
pAcGP67-A Anhui13 NA	Xiaoli Xiong	pAcGP67-A with expression construct for Anhui13 NA	BaculoGold
pFB-LIC-Bse	Xiaoli Xiong	Transfer vector compatible with the Bac-to-Bac system.	Bac to Bac
pFB-LIC-Bse Anhui13 NA	Generated in these Studies	pFB-LIC-Bse containing Anhui13 NA expression construct	Bac to Bac
pFB-LIC-Bse Anhui13 NA S367N	Generated in these Studies	pFB-LIC-Bse containing Anhui13 S367N NA expression construct	Bac to Bac

2.1.5 Small molecules

Molecule Name	Short Name	Source	Catalogue Number
α2,6-Sialyl-N-acetyllactosamine	6SLN	Dextra Laboratories Ltd., Reading, UK	SLN306
α2,3-Sialyl-N-acetyllactosamine	3SLN	Dextra Laboratories Ltd., Reading, UK	SLN302
α2,6-Sialyl-N-acetyllactosamine sialoglycopolymer (attached to polyacrylamide and biotin)	6SLN-PAA	Lectinity Holdings, Moscow, Russia	0997-BP
α2,3-Sialyl-N-acetyllactosamine sialoglycopolymer (attached to polyacrylamide and biotin)	3SLN-PAA	Lectinity Holdings, Moscow, Russia	0036-BP
α2,6-Sialyllactose sialoglycopolymer (attached to polyacrylamide and biotin)	6SL-PAA	Lectinity Holdings, Moscow, Russia	0063-BP
α2,3-Sialyllactose sialoglycopolymer (attached to polyacrylamide and biotin)	3SL-PAA	Lectinity Holdings, Moscow, Russia	0060-BP
α2,3-Sialyl-N-acetyllactosamine-6' sulphated GlcNAc sialoglycopolymer (attached to polyacrylamide and biotin)	6SU-3SLN-PAA	Lectinity Holdings, Moscow, Russia	0951-BP
Sialyl Lewis^x sialoglycopolymer (attached to polyacrylamide and biotin)	Sia Lex-PAA	Lectinity Holdings, Moscow, Russia	0062-BP
Sialyl Lewis^x-6' sulphated GlcNAc sialoglycopolymer (attached to polyacrylamide and biotin)	6SU-Sia Lex-PAA	Lectinity Holdings, Moscow, Russia	0020-BP
Oseltamivir carboxylate	-	Kind gift from Roche, Welwyn Garden City, UK	-
Zanamivir	-	Kind gift from GSK, Stevenage, UK	-

2.1.6 Bacterial media

L-broth:

1% (w/v) Bacto Tryptone
0.5% (w/v) Yeast Extract
1% (w/v) NaCl

L-agar:

L-broth + 1.5% (w/v) agar

SOC:

2% (w/v) Bacto Tryptone
0.5% (w/v) Yeast Extract
10 mM NaCl
2.5 mM KCl
10 mM $MgCl_2$
10 mM $MgSO_4$
20 mM Glucose

2.1.7 Buffers

PBS-A:

12 mM NaH_2PO_4/Na_2HPO_4 pH 7.4
172 mM NaCl
3.4 mM KCl

2.1.7.1 Electrophoresis buffers

50 x TAE:

2 M Tris-base pH 8.3 (pH adjusted using glacial acetic acid)
0.05 M EDTA

10 x SDS Running Buffer:

250 mM Tris-base
1.92 M Glycine
1% (w/v) SDS
Adjusted to pH 8.3 with HCl

5x SDS loading buffer:

0.2 M Tris-HCl (pH 7.4)
20% Glycerol
10% SDS
0.1% Bromophenol Blue
2 M β-Mercapto-ethanol (in reducing buffer only)

2.1.7.2 Buffer for making competent cells

TB Buffer:

10mM HEPES
15mM $CaCl_2$
250mM KCl
55mM $MnCl_2$
Adjusted to pH 6.7 with KOH

2.1.7.3 ELISA buffers

PBS-Tween:

PBS-A
0.05% (v/v) Tween-80

Block Buffer:

PBS-A
2.5% (w/v) BSA
20 mM EDTA
0.05% (v/v) Tween-20

Substrate Solution:

25 mM citrate buffer pH 4.5
0.1 mg/ml 3,3',5,5' – Tetramethylbenzidine dihydrochloride (TMB) (Sigma: T5525)
0.03% (v/v) H_2O_2

2.1.7.4 Protein purification buffers

His Buffer A:

25 mM NaH_2PO_4/Na_2HPO_4 pH 8.0
300 mM NaCl
10% (v/v) Glycerol
5 mM Imidazole
0.01% (w/v) NaN_3

His Buffer B:

25 mM NaH_2PO_4/Na_2HPO_4 pH 8.0
300 mM NaCl
10% (v/v) Glycerol
500 mM Imidazole
0.01% (w/v) NaN_3

Tris-Saline:

25mM Tris-HCl pH 8.0
150mM NaCl
0.01% (w/v) NaN_3

2.1.7.5 Virus binding assay buffers

HBS-EP:

10mM HEPES-NaOH pH 7.4
150mM NaCl
3mM EDTA
0.005% (v/v) Tween-20
0.01% (w/v) NaN_3

HBS-P + $CaCl_2$:

10mM HEPES-NaOH pH 7.4
150mM NaCl
0.005% (v/v) Tween-20
4mM $CaCl_2$
0.01% (w/v) NaN_3

2.1.7.6 Enzyme kinetics buffers

HBS-P + $CaCl_2$:

Same as above (section 2.1.7.5)

MES Buffer:

32.5mM MES-NaOH pH 6.5
4mM $CaCl_2$
0.01% (w/v) NaN_3

2.2 Methods

2.2.1 Cells

Madin-Derby Canine Kidney (MDCK) ECACC, MDCK Parent and Human Embryonic Kidney 293T (293T) cells were obtained from stocks held by the World Health Organisation Collaborating Centre for Reference and Research on Influenza, MRC National Institute for Medical Research (NIMR WHOCC), now known as the Francis Crick Institute Worldwide Influenza Centre. These cells were maintained in adherent cultures in Dulbecco's Modified Eagle's Medium (DMEM) (Sigma: D6429) supplemented with 10% (v/v) heat inactivated (1hr at 56°C) foetal calf serum (FCS) (Labtech: FCS-SA/500) and antibiotics (Pen/Strep [penicillin 100 U/ml, streptomycin 100 µg/ml]) (Sigma: P4333) at 37°C, 5% (v/v) CO_2.

Sf9 cells were obtained from stocks held by the Division of Molecular Structure, MRC NIMR. These cells were maintained in suspension culture in SF-900 II SFM (Life Technologies: 10902) at densities ranging from 5 x 10^5 – 1 x 10^7/ml. Cells were counted using Countess Automated Cell Counter (Life Technologies), according to the manufacturer's instructions. Cells were incubated at 28°C with shaking at 150 rpm. Protein expression was carried out by culturing the cells in Insect Xpress media (Lonza: 12-730Q)

2.2.2 Influenza virus propagation

2.2.2.1 Egg-propagation

Viruses were propagated in the allantoic cavity of 11-12 day old fertilised hens' eggs. Eggs were infected with 100 µl of virus diluted in PBS-A + Pen/Strep. Virus propagated from stocks of previously propagated virus was infected at a dilution

ranging from 10^{-3} – 10^{-6}-fold. Viruses propagated from a transfection supernatant were infected at dilutions ranging from 10^{0} – 10^{-4}-fold. Eggs were incubated at 34°C in a humidified incubator for a time ranging from 48 – 72 hrs. Eggs were chilled at 4°C overnight and allantoic fluid harvested.

Allantoic fluid used to generate virus stocks was subjected to a low speed centrifugation at 2500 x g for 15 mins to remove debris. Allantoic fluid was then filter sterilized through a 0.45 µm syringe filter and stored in 1 ml aliquots at -80°C.

2.2.2.2 Tissue culture-propagation

Confluent monolayers of MDCK cells were washed twice with PBS-A to remove residual serum. Virus diluted in serum-free medium (DMEM + Pen/Strep) was added to the monolayer in a small volume, typically 1 ml to every 25 cm^2 tissue culture surface area. Viruses were diluted in serum-free medium and inoculated at dilutions ranging from 10^{-3} – 10^{-6}-fold. Virus was allowed to attach for 1 hr in a 37°C, 5% CO_2 incubator. Inoculum was left on the cells and additional serum-free medium supplemented with 2.5 µg/ml TPCK Trypsin (Sigma: T1426) was added. Infected cells were incubated at 37°C, 5% CO_2 for 48 – 72 hrs.

Cell culture supernatant used for virus stocks was subjected to a low speed centrifugation 2500 x g for 15 mins and stored in 1 ml aliquots at -80°C.

2.2.3 Influenza virus purification

2.2.3.1 Egg-propagated virus purification

For purification, the allantoic fluid from >100 eggs was used. Virus was clarified by a low speed centrifugation 2500 x g for 15 mins. Virus was pelleted by centrifugation at 6400 x g 4°C overnight (Beckman Coulter F10 6x500 FiberLite

rotor). The virus pellet was resuspended, homogenised using a glass Dounce homogeniser and placed briefly (~2 mins) in a sonicating waterbath to ensure a homogeneous resuspension. This suspension was layered onto a 15 – 40% (w/v) continuous gradient of sucrose in PBS-A + 0.01% (w/v) NaN_3. Gradients were centrifuged at 110 000 x g for 45 mins at 4°C (Beckman Coulter SW32Ti rotor). The virus-containing band, identified by visual inspection, was removed from the gradient and diluted in PBS-A + 0.01% (w/v) NaN_3. Virus was pelleted at 165 000 x g for 60 mins at 4°C (Beckman Coulter SW32Ti rotor) and the pellet was resuspended in a small volume of PBS-A + 0.01% (w/v) NaN_3, typically 100 – 500 µl, and stored at 4°C.

2.2.3.2 Tissue culture-propagated virus purification

Cell supernatant from virus propagation was pooled and centrifuged twice at low speed (2500 x g) for 15 mins to remove debris. Virus was then pelleted either: at 6400 x g 4°C overnight (Beckman Coulter F10 6x500 FiberLite rotor) or 165 000 x g for 60 mins at 4°C (Beckman Coulter SW32ti rotor). Pellets were resuspended in a small volume of residual fluid, repelleted at 160 000 x g for 20 mins at 4°C (Beckman Coulter TLA-100.3 rotor), resuspended in a small volume of PBS + 0.01% (w/v) NaN_3, typically 100 – 500 µl, and stored at 4°C.

For virus binding assays tissue culture-propagated viruses were typically not purified further, as lack of purification was found not to affect virus binding to receptor coated surfaces and there was no measurable non-specific binding of cell debris. For biophysical binding analyses virus was grown in bulk tissue culture ~1 Litre in volume, which was typically the third passage after virus rescue.

2.2.4 Haemagglutination titre

2.2.4.1 Preparation of blood

Turkey red blood cells (TRBCs) were obtained stored in Alsever's solution (Public Health England Centre for Infectious Disease Surveillance and Control, Colindale, UK). Blood was pelleted by centrifugation at 850 x g for 15 mins and the supernatant was removed by aspiration. Blood was washed twice by resuspending the pellet in a large excess of PBS-A by gentle inversion of the tube and pelleting at 850 x g for 15 mins. The blood was stored as a pellet at 4°C.

2.2.4.2 Titration method

Haemagglutination titres were determined for all propagated viruses. 50 µl of allantoic fluid or tissue culture supernatant was diluted 2-fold across a v-bottom 96 well plate (Greiner Bio One: 651101) in 50 µl of PBS-A. 50 µl of 0.5% (v/v) TRBCs, diluted in PBS-A, was added to each well. The plate was incubated at room temperature for 30 mins. The haemagglutination titre was taken to be the reciprocal of the highest dilution which still showed agglutination. Lack of agglutination was determined to be a well where the pellet of blood formed a teardrop shape when the plate was tilted.

2.2.4.3 Haemagglutination assay elution

Elution of HA assays were measured by placing the HA assay microtitre plate, prepared as described in the previous section, covered with plastic film, in a 37°C incubator for ~18 hrs. The loss of agglutination was then measured by comparison with the haemagglutination assay prior to incubation.

2.2.5 Influenza virus reverse genetics

Reverse genetics was carried out using an 8 plasmid system described by Hoffmann *et al.* (Hoffmann *et al.*, 2000). In this method a cDNA of each of the 8 viral segments is cloned into the pHW2000 plasmid and all 8 plasmids are co-transfected into 293T cells to generate virus.

1 μg of each of the 8 plasmids, prepared in midiprep quantity using the QIAGEN Plasmid Plus Midiprep kit, used according to the manufacturer's instructions (Qiagen: 12945), were mixed together and diluted in 250 μl of Opti-MEM (Life Technologies: 31985070). This plasmid and Opti-MEM mixture was added to 10 μl of Lipofectamine 2000 (Life Technologies: 11668019), which had previously been diluted in 250 μl of Opti-MEM. This mixture was incubated at room temperature for 20 mins. 25 cm^2 flasks of 293T cells at 70 – 90% confluency were washed twice with Opti-MEM. The 500 μl mixture of plasmids and Lipofectamine were added to the flask of cells to a total volume of 3 ml with Opti-MEM, flasks were incubated at 37°C, 5% CO_2. After 5 – 6 hrs medium was changed for 3 ml of Opti-MEM + Pen/Strep. Flasks were further incubated for 24 hrs before adding TPCK trypsin (Sigma: T1426) to the cell supernatant at a concentration of 0.5 μg/ml. After further incubation for 18 – 24 hrs the cell supernatant was harvested and clarified by a low speed centrifugation (2500 x g for 10 mins). This cell supernatant was either directly inoculated into eggs or cell culture, or stored at -80°C.

2.2.6 Nucleic acid work

2.2.6.1 Agarose gel electrophoresis

DNA samples were diluted in 5 x Gelpilot DNA Loading dye (QIAGEN: 239901) and run on 1% agarose (Bioline: BIO-41027) gels were made and run using 1 x TAE buffer. A typical 10 x 10 cm gel was run for 20 - 40 mins at 120 V (12 V/cm).

2.2.6.2 DNA purification

DNA was purified from solution using the illustra GFX PCR DNA and gel band purification kit (GE Healthcare: 28-9034-71) as per the manufacturer's instructions. DNA was eluted from the column in 10 – 30 µl of dH_2O.

2.2.6.3 Gel purification of DNA samples

DNA was purified from an agarose gel by excising the appropriate band using a gel extraction tool (Starlab: N2000-0100) and purifying with the illustra GFX PCR DNA and gel band purification kit (GE Healthcare: 28-9034-71) as per the manufacturer's instructions. As above, DNA was eluted from the column in 10 – 30 µl of dH_2O.

2.2.6.4 Determination of DNA concentration

DNA was quantified using a Nanodrop 2000 spectrophotometer (Thermo Scientific) using the built in nucleic acid quantification software.

2.2.6.5 Oligonucleotide synthesis

Oligonucleotide PCR primers were synthesised by Eurofins Genomics (Ebersberg, Germany). They were synthesised and purified to HPSF (High purity, Salt-Free) standard and delivered in dH_2O at a concentration of 100 µM.

2.2.6.6 DNA sequencing

Sanger DNA sequencing was carried out by GATC Biotech (Konstanz, Germany). PCR products were sent at a concentration of 10 – 50 ng/µl and plasmids at 50 – 150 ng/µl. Primers were either synthesised and held by GATC biotech or were sent separately with DNA samples (synthesised by Eurofins Genomics).

2.2.6.7 Sequence analysis

DNA sequences were constructed from ABI trace files using the Staden package (Bonfield, Smith & Staden, 1995). Sequences of virus segments were checked by visual inspection for mixtures in sequence traces. Assembled sequences were aligned using BioEdit (Hall, 1999).

2.2.6.8 Virus library sequences

Library sequences of viral segments were obtained from the Global Initiative for Sharing All Influenza Data (GISAID) Epiflu database (http://gisaid.org).

2.2.6.9 Competent cells/transformation

Competent cells used for bacterial work were either commercial XL-10 Gold (Agilent: 200315) or cells made chemically competent with the following protocols:

A 3 ml overnight pre-culture of the desired strain was grown from a glycerol stock streaked on an L-agar plate. Cells for typical transformations were grown without antibiotics, cells generated for baculovirus generation [DH10EmBacY] were grown in the presence of 50 µg/ml kanamycin (Sigma: K0254); 30 µg/ml chloramphenicol (Sigma: C0378) and 10 µg/ml tetracycline (Sigma: T7660). The starter culture was inoculated into 500 ml L-broth (containing antibiotics if necessary). Cells were grown at 25°C until the cell density gave A_{600} of 0.4 – 0.6. The

culture was chilled on ice for ~15 mins and pelleted at 4000 x g for 15 mins. The pellet was resuspended in 100 ml of ice-cold TB buffer and incubated on ice for 10 mins. The suspension was pelleted at 4000 x g for 15 mins, resuspended in 18.4 ml ice-cold TB buffer and 1.4 ml DMSO and incubated on ice for 10 mins. The cells were flash frozen in an ethanol/dry ice bath and stored at -80°C in 250 µl aliquots.

Cells were transformed by the same method whether they were the commercial or in-house prepared cells. Competent cells were thawed on ice and mixed by gentle pipetting. 50 µl of cells were transferred to a prechilled 1.5 ml tube. DNA was added in a volume of up to 5 µl to the competent cells, mixed and incubated on ice for 15 mins. Cells were heat-shocked at 42°C for 1 min. 450 µl of SOC medium was added and tubes were incubated at 37°C with 250 rpm shaking for 1 hr.

2.2.6.10 RNA extraction

RNA was extracted from virus using the QIAamp Viral RNA Mini Kit (QIAGEN: 52906) according to the manufacturer's protocol. RNA was eluted from the column in 50 µl of dH_2O and stored at -20°C. RNA extraction was carried out either with 140 µl of allantoic fluid or cell culture supernatant or for a purified or concentrated virus the extraction was carried out on 1 µl of the virus suspension, diluted in 140 µl of PBS-A.

2.2.6.11 RT-PCR of viral RNA

Amplification of viral RNA segments for sequencing was carried out using a one-step RT-PCR protocol developed by the NIMR WHOCC. This protocol amplifies each viral segment in two separate overlapping parts, to allow segment specific amplification, as there is considerable sequence homology between the non-coding regions at the end of different segments (see section 2.1.3 for primer sequences).

50 µl PCR reactions contained the following components: 1.5 x Pfx buffer (Life Technologies), 1 mM $MgSO_4$, 0.5 mM of each of the dNTPs, 0.3 µM forward primer, 0.3 µM reverse primer, 200 U Superscript III RT (Life Technologies, 18080-085), 1.25 U Platinum Pfx polymerase (Life Technologies, 11708-039), 2 µl of extracted RNA.

RT-PCR cycling conditions were: 50°C: 30 mins, 94°C: 10 mins, 40x (94°C: 30 s, 55°C: 30 s, 68°C: 3 mins), 68°C: 10 mins, 4°C hold.

2.2.6.12 Cloning

2.2.6.12.1 Cloning for reverse genetics

Reverse genetics was carried out by cloning the HA and NA gene segments for the virus into the pHW2000 vector. Gene amplification was carried out with two-step RT-PCR, with the majority of genes amplified in a single reaction.

The cDNA synthesis step generated cDNAs for all of the viral segments using the Influenza A universal primer uni12 (see section 2.1.3.1 for sequence). cDNA synthesis was carried out using the Verso cDNA synthesis kit (Thermo: AB-1453). The reaction mixture was made as per the manufacturer's recommendations. The cycling conditions were: 42°C for 30 mins followed by 95°C for 5 mins.

Following cDNA synthesis PCR amplification was carried out using the following reaction mixture: 1.5 x Pfx buffer (Life Technologies), 1 mM $MgSO_4$, 0.5 mM of each of the dNTPs, 0.3 µM forward primer, 0.3 µM reverse primer, 1.25 U Platinum Pfx polymerase (Life Technologies, 11708-039), 2 µl of generated cDNA. PCR cycling conditions were: 94°C: 10 mins, 40x (94°C: 30 s, 55°C: 30 s, 68°C: 4 mins), 68°C: 10 mins, 4°C hold.

The majority of cloned HA and NA genes were amplified using a single PCR amplification, see section 2.1.3.1 for sequences of primers used. The primers amplify the full-length gene segment including non-coding regions whilst adding sites for the restriction enzymes BsmBI or AarI. Gene sequences were first analysed using RestrictionMapper (http://www.restrictionmapper.org/) to check for internal restriction sites and the primers used were chosen accordingly. HA amplification was typically inefficient owing to a similarity in primer sequence to the primers used for NS cloning and so two 50 µl reactions were carried out for HA amplification, which were then pooled at a later stage. A/California/7/2009 HA could not be amplified in a single PCR reaction and so a splicing PCR was carried out, where the gene was amplified in two separate overlapping sections and assembled. The HA was initially amplified using two pairs of primers: for the 5' section HAFUc and HAR873; and for the 3' section HAF848 and HARUc (see section 2.1.3.1 for primer sequences) using the same PCR conditions as listed above. The two PCR products were purified by agarose gel electrophoresis and a second PCR was set up with all of the components except the primers and using 50 ng of each product from the previous PCRs as templates. The PCR was carried out for 20 cycles as detailed above. The primers HAFUc and HARUc were then added to the mixture at the usual concentrations (0.3 µM of each) and the PCR was carried out for a further 20 cycles. The full length product was purified and used as template for a PCR using the typical HA cloning primers, as detailed above.

After PCR amplification the products were run on an agarose gel to confirm that the correct sized product was produced (~1800 bp for HA and ~1400 bp for NA). When the correct band was seen the remainder of the 50 µl reaction was gel purified.

The pHW2000 vector was cut using the enzyme BsmBI (NEB: R0580). Inserts were cut using either BsmBI or AarI (Thermo: ER1582) depending on the primers used for amplification. Up to 2 µg of DNA was cut in a 50 µl reaction according to the manufacturer's instructions. BsmBI reactions were incubated at 55°C for 1 hr; AarI reactions were incubated at 37°C for 1 hr. Digested pHW2000 was dephosphorylated using Antarctic Phosphatase (NEB: M0289) according to the manufacturer's instructions. The cut DNA was then purified, as described in section 2.2.6.2, and eluted in 20 µl of dH_2O.

Ligation was carried out using T4 DNA ligase (NEB: M0202). 20 µl Reactions contained: 1 µl of T4 DNA ligase (400 U), 2 µl of 10 x T4 DNA ligase buffer plus vector and insert at a molar ratio of 1:5 (vector:insert). The reaction was mixed and left for 30 mins at room temperature. A no insert control ligation was also carried out to assess the efficiency of ligation.

4 µl of ligation mixtures was transformed into competent bacterial cells and spread on L-agar plates containing 100 µg/ml ampicillin (Sigma: A9393). Colonies were picked from each ligation and cultured in 5 ml of L-broth containing 100 µg/ml ampicillin. The plasmid DNA was prepared using the QIAprep Spin Miniprep Kit (Qiagen: 27106). Prepared plasmids were run on an agarose gel alongside empty vector as well as a previously cloned HA or NA in pHW2000 and a Supercoiled DNA Ladder (NEB: N0472). This gel allowed the selection of clones that contained an insert of the correct size. 2 – 4 clones with the correct insert size were then sequenced.

2.2.6.12.2 Cloning for protein expression

Cloning of H1N1 HA

Genes encoding the ectodomain of the HA of Cal7 and Iowa06 were synthesised by GeneArt (Life Technologies). A construct was synthesised that encoded a protein that lacked the 17 residue signal peptide and was truncated to residue 502 (H1 Numbering) to remove the transmembrane domain. Two restriction sites were added to the ends of the construct, BamHI at the N terminal and NotI at the C terminal, to allow subcloning into the pHAEM1 vector, a modified version of the pAcGP67-A vector (BD Biosciences). This vector had been engineered to contain an N-terminal gp67 signal peptide to allow expression by secretion by Sf9 cells and a C-terminal protein purification tag containing a TEV protease cleavage site, T4 fibritin trimerisation motif (foldon) and a 6x His tag. The sequences of the H1 HA constructs are shown in the *Appendix*. These genes were codon optimised for insect cells using GeneArt software prior to synthesis.

The synthesised genes were delivered in a carrier vector, which was transformed and purified in miniprep quantities. 6 µg of this gene containing vector and the destination vector pHAEM1 was digested with BamHI-HF (NEB: R3136S) and also NotI-HF (NEB: R3189) in 100 µl reactions according to the manufacturer's recommendations. Digested pHAEM1 was dephosphorylated using Antarctic Phosphatase (NEB: M0289) according to the manufacturer's instructions. Both the desired insert and the pHAEM1 vector were purified by agarose gel electrophoresis. 100 ng of cut pHAEM1 was added to 75 ng of insert, (giving a molar ratio of ~1:5 vector:insert) in a 20 µl ligation mixture, including 2 µl 10x T4 DNA ligase buffer (NEB) and 1 µl T4 DNA Ligase (400 U)(NEB: M0202)**.** The reaction was carried out at room temperature for 30 mins. A no insert control ligation was also carried out to

assess the efficiency of ligation. This ligation was transformed into competent bacteria and plated on L-agar plates containing 100 μg/ml ampicillin. Colonies were picked, amplified and the plasmid DNA purified. These plasmids were tested to see that they contained the correct insert by agarose gel electrophoresis. Plasmids of the correct size were sequenced to verify the correct insertion site and sequence of the construct.

Subcloning H7N9 NA

The A/Anhui/1/2013 (Anhui13) NA ectodomain (residues 75 – 465) had been previously synthesised by Geneart (Life Technologies) and cloned into a modified version of the pHAEM1 vector, which included an N terminal 6 x His-tag, vasodilator-stimulator phosphoprotein (VASP) tetramerisation domain (Kühnel *et al.*, 2004; Xu *et al.*, 2008) and TEV cleavage site, for construct sequence see *Appendix*. The reagents used for the later stages of baculovirus expression using the BaculoGold system (BD Biosciences) had been discontinued; therefore the gene was subcloned into a vector, pFB-LIC-Bse, to use the Bac-to-Bac system (Life Technologies).

Anhui NA was subcloned using the In-Fusion HD Cloning Kit (Clontech) following the manufacturer's instructions. The full expression construct was amplified by PCR, from the previously cloned construct in the pAcGP67-A vector using oligonucleotides designed using the In-Fusion primer design tool (http://bioinfo.clontech.com/infusion/convertPcrPrimersInit.do) (for primer sequences see section 2.1.3.2). 4 μl of the transformation mixture was transformed into competent cells and plated on L-agar plates containing 100 μg/ml ampicillin. Colonies were picked and sequenced to check the insertion site and construct sequence.

2.2.6.13 Mutagenesis

2.2.6.13.1 Point mutations

Point mutations were introduced into plasmids using the Quikchange Lightning Site-Directed Mutagenesis Kit (Agilent: 210519) according to the manufacturer's instructions. Mutagenesis oligonucleotides were designed using the Quikchange primer design software (http://www.genomics.agilent.com/primerDesignProgram.jsp). All mutant plasmids were checked by sequencing to verify successful mutagenesis.

2.2.6.13.2 Large deletion mutagenesis

Large deletions were made in the X-31 NA to generate truncated NA stalks. These were generated in the reverse genetics plasmid using the Q5 Site-Directed Mutagenesis Kit (NEB: E0554S) according to the manufacturer's instructions. Mutagenesis oligonucleotides were designed using NEBaseChanger software (http://nebasechanger.neb.com/). All mutant plasmids were checked by sequencing to verify correct mutagenesis.

2.2.7 Insect cell protein expression

Insect cell protein expression was carried out using two different systems. Initially expression was carried out using the BaculoGold system (BD Biosciences); however, owing to the discontinuation of reagents for this system, later expression was done using a system based around the Bac-to-Bac system (Life Technologies).

2.2.7.1 BaculoGold baculovirus production

The desired protein expression constructs were cloned into the pHAEM1 vector (as described in section 2.2.6.12.2). The vector was amplified and purified

using the Plasmid plus Midi kit (Qiagen: 129450). Insect cells were transfected using Cellfectin II reagent (Life Technologies: 10362-100). Sf9 cells were plated at a density of 1 x 10^6 cells/well in 6-well plates. 5 µg of plasmid was mixed with 0.5 µg (5 µl) of BaculoGold Bright Baculovirus DNA (BD Biosciences: 552846) and 10 µl of Cellfectin II diluted in SF-900 II SFM (Life Technologies: 10902) and incubated at room temperature for 20 mins. DNA and Cellfectin mix was added to the cells and incubated at 28°C for 6 hrs. Medium was changed for 2 ml of fresh SF-900 II SFM and plates were incubated at 28°C for 72 – 96 hrs and baculovirus production was examined by looking for fluorescent cells using an inverted fluorescence microscope. Supernatant was harvested and clarified by low-speed centrifugation (2500 x g) and stored at 4°C in the dark.

2.2.7.2 Bac-to-Bac baculovirus production

Protein expression constructs were cloned into the pFB-LIC-Bse vector (as described in section 2.2.6.12.2). 0.5 – 1 µg of plasmid was transformed into competent cells made from DH10EmBacY glycerol stock (Geneva Biotech) (see section 2.2.6.9 for details). 100 µl of competent cells were incubated with DNA on ice for 30 mins, cells were heat shocked at 42°C for 1 min and 900 µl of SOC medium was added. Transformations were incubated at 37°C with shaking at 250 rpm for 4 – 6 hrs. Dilutions of transformations were plated out on L-agar plates containing: 50 µg/ml kanamycin (Sigma: K0254); 30 µg/ml chloramphenicol (Sigma: C0378); 10 µg/ml tetracycline (Sigma: T7660); 7 µg/ml gentamicin (Sigma: G1272); 100 µg/ml Bluogal (Life Technologies: 15519-028) and 40 µg/ml IPTG (Life Technologies: 15529-019). Plates were incubated at 37°C for 48 hrs. Bacterial colonies that had undergone recombination of the protein expression gene with baculovirus DNA gave a white

colony. White colonies were then picked and restreaked on a fresh plate to verify the white phenotype. White colonies were picked and grown in 10 ml L-broth containing 50 µg/ml kanamycin, 30µg/ml chloramphenicol, 10 µg/ml tetracycline and 7 µg/ml gentamicin and grown at 37°C with shaking at 200 rpm for ~18 hrs. The required baculovirus DNA (bacmid) is very large (>100 kb) and it was therefore prepared by isopropanol precipitation rather than a column spin kit. Cells were pelleted at 2500 x g for 10 mins. Pellets were resuspended in 300 µl of 15 mM Tris-HCl, pH 8.0, 10 mM EDTA, 100 µg/ml RNase A. Cells were lysed by adding 300 µl of 0.2 M NaOH + 1% (w/v) SDS. Lysis was incubated at room temperature for 5 mins. The mixture was neutralised with 300 µl of 3 M Potassium Acetate, pH 5.5. The sample was then incubated on ice for 10 mins and then centrifuged at 14 000 x g for 10 mins. The supernatant was added to 800 µl of isopropanol and incubated on ice for 10 mins. The sample was centrifuged at 14 000 x g for 15 mins, the supernatant was removed and the pellet was washed by adding 500 µl of 70% (v/v) ethanol followed by centrifugation at 14 000 x g for 5 mins. The washing step was repeated a further time. The supernatant was removed and the pellet was air dried for ~10 mins to remove residual ethanol. Finally the pellet was resuspended in 50 µl of TE buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA) and stored at 4°C.

Sf9 cells were transfected by plating 1 x 10^6 cells/well in a 6-well plate. 4 µl of Cellfectin II reagent (Life Technologies: 10362-100) was diluted in 100 µl of SF-900 II SFM (Life Technologies: 10902) and incubated at room temperature for 5 mins. 5 µg of the recombined baculovirus DNA (bacmid) was diluted in 100 µl of SF-900 II SFM and added to the Cellfectin mixture and incubated at room temperature for 20 mins. The DNA/Cellfectin mixture was added to the Sf9 cells and incubated at 28°C for 6 hrs.

Media was replaced with 2 ml of SF-900 II SFM and transfections were incubated at 28°C for 72 – 96 hrs. Baculovirus production was assayed by detecting fluorescent cells using an inverted fluorescence microscope. Supernatant was harvested and clarified by low-speed centrifugation and stored at 4°C in the dark.

2.2.7.3 Protein expression from generated baculovirus

Baculovirus propagation and protein expression was carried out by the same protocol whether initial virus was generated using the BaculoGold or Bac-to-Bac system. Passage 1 (P1) of baculoviruses was carried out by inoculating 50-500 µl of transfection supernatant into a 10 ml suspension culture of Sf9 cells at a density of 2 x 10^6 cells/ml. This culture was incubated at 28°C shaken at 150 rpm for 72 hrs. Passage 2 (P2) was initiated by inoculating a 20 ml suspension culture of Sf9 cells (2 x 10^6 cells/ml) with 50 – 500 µl of P1. The aim was for all cells to be fluorescent 72 hrs after infection. P2 virus was clarified by low speed centrifugation, filtered through a 0.2 µm syringe filter, supplemented with FCS to a final concentration of 2% (v/v) and stored at 4°C in the dark.

Large/Medium scale protein expression was carried out by infecting 2 – 3 litres of Sf9 cells grown in Insect Xpress media (Lonza: 12-730Q) at a density ranging from 3 x 10^6 – 4 x 10^6 cells/ml with sufficient of the P2 stock virus to allow total infection of the cells in the first virus replication cycle (all cells fluorescent by 48 hrs), typically 0.5 – 2 ml per 500 ml culture. Expression was carried out in 2 litre roller bottles, each containing 500 ml of cells, at 28°C, with shaking at 200 rpm for 72 hrs. The culture was harvested and cell debris was removed by centrifugation at 4000 x g for 20 mins.

2.2.8 Protein work

2.2.8.1 SDS-PAGE

SDS-PAGE was carried out using 4 – 20% Mini-PROTEAN TGX (Biorad: 456-1093) precast gels. Samples were diluted in 5 x SDS loading buffer, under either reducing or non-reducing conditions, and heated to 95°C for ~5 mins. Precision Plus All Blue Standards protein ladder (Biorad: 161-0373) was run as a molecular weight standard. Gels were run in 1 x SDS buffer at 150 V (~20 V/cm) for 50 mins. Gels were stained using InstantBlue protein stain (Expedion: ISB1L) according to the manufacturer's instructions.

2.2.8.2 Protein concentration determination

Purified protein concentration was determined using a measurement of A_{280}. Measurements were made using an ultra-micro volume quartz cuvette with a path length of 10 mm in a Jasco V-550 spectrophotometer. A baseline measurement was made with buffer alone and subtracted from the measured spectrum. Protein concentrations were determined using extinction coefficients calculated using the Expasy ProtParam software (http://web.expasy.org/protparam/).

2.2.8.3 Purification of N1 NA from purified virus

Cal7 NA purification

Proteins were digested from purified virus using bromelain (Sigma: B4882) at a 1 : 5 ratio (w/w) (assuming that 100 nM virus ≈ 10 mg/ml) in the presence of 50 mM β-mercaptoethanol for 1hr at 37°C. Virus cores were removed by centrifugation at 160 000 x g for 20 mins at 4°C (Beckman Coulter TLA-100.3 rotor). The buffer of the cleaved protein was exchanged into 25 mM Tris pH 8.0, 5 mM NaCl, 0.01% (w/v) NaN_3 using an Amicon Ultra – 15 ml 50 kDa cutoff filter (Millipore: UFC905024). The

protein was loaded onto an anion exchange 5 ml HiTrap Q Sepharose FF column (GE: 17-5156-01) using an ÄKTA FPLC system (GE), to remove contaminating HA. The flow through, which contained the NA was retained. Protein bound to the column was eluted with 25 mM Tris, pH 8.0, 1 M NaCl, 0.01% (w/v) NaN_3. The flow through was concentrated to a small volume using an Amicon Ultra – 15 ml 50 kDa cutoff filter to a volume of ~2 ml. A HiLoad Superdex 16/600 200 pg column was equilibrated in Tris-saline by washing through 150 ml at a flow rate of 1.5 ml/min using an ÄKTA FPLC system. The small volume of concentrated ion exchange flow through was loaded onto the gel filtration column using a 5 ml injection loop and the chromatography was carried out by passing buffer through the column at a flow rate of 1.5 ml/min for 1 column volume (CV) (120 ml) collecting 1ml fractions, while monitoring absorbance at 280 nm. Peak fractions were analysed by SDS-PAGE. Desired fractions were collected, pooled and concentrated using an Amicon Ultra – 15 ml 50 kDa cutoff filter. The protein was dispensed into aliquots, snap frozen in liquid N_2 and stored at -80°C.

Iowa06 NA purification

Proteins were detergent extracted from purified virus with 2% (v/v) Brij-36T (decaethylene glycol monododecyl ether)(Sigma: P9769) by incubating at room temperature for 30 mins, followed by centrifugation at 160 000 x g for 20 mins at 4°C (Beckman Coulter TLA-100.3 rotor) to remove the debris. The supernatant was concentrated using an Amicon Ultra – 15 ml 50 kDa cutoff filter and the buffer was exchanged for 25 mM Tris pH 8.0, 5 mM NaCl, 0.01% (w/v) NaN_3, 0.5% (v/v) Brij-36T. The protein was loaded onto a 5 ml HiTrap Q Sepharose FF column, previously equilibrated in buffer (GE: 17-5156-01) using an ÄKTA FPLC system, to remove contaminating HA. The flow through, which contained the NA, was concentrated to 1-

2 ml using an Amicon Ultra – 15 ml 50 kDa cutoff filter. The concentrated protein was loaded onto a HiLoad Superdex 16/600 200 pg column equilibrated in Tris-saline + 0.5% (v/v) Brij-36T. Peak fractions were analysed by SDS-PAGE and fractions containing NA were pooled and concentrated using an Amicon Ultra – 15 ml 50 kDa cutoff filter. NA rosettes were made by adding a ratio of 1:4 (v/v) of Biobeads SM-2 Adsorbent (Biorad: 152-3920) to the protein solution.

2.2.8.4 Purification of insect cell expressed proteins

Baculovirus expression supernatants were concentrated by the NIMR large-scale lab using a Quixstand hollow fiber filtration unit (GE) using a filter with a 10 kDa cutoff. This concentrated the expresed proteins into a volume of ~100 ml. The concentrated protein solution was supplemented with cOmplete protease inhibitor tablets, EDTA-free (Roche: 04693132001) and stored at 4°C. Concentrated buffer components were added to the solution to give final concentrations of: 25 mM phosphate pH 8.0, 300 mM NaCl, 5 mM Imidazole. Debris was removed by centrifugation at 44 000 x g for 30 mins (Beckman-Coulter JA-25.50 rotor). Four 5 ml HisTALON superflow cartridges (Clontech: 635682) were connected in series and washed in five column volumes (CV) of His Buffer B. Columns were then equlibrated with five CV of His Buffer A. Concentrated protein supernatant was loaded and recirculated through the columns for ~2 hrs using a peristaltic pump at a flow rate of ~5 ml/min. The column was attached to an ÄKTA FPLC system and the protein was eluted using the following protocol: the column was washed with 15 x CV of His buffer A; protein eluted with 10 x CV of a 0 – 100% gradient (100% His buffer A to 100% His buffer B, 5 – 500 mM imidazole) all at a flow rate of 5 ml/min, collecting 10 ml fractions. Small volumes of elution fractions were analysed by SDS-PAGE to verify

the presence of recombinant protein. Recombinant protein fractions were pooled and concentrated using an Amicon Ultra – 15 ml 50 kDa cutoff filter (Millipore: UFC905024) and buffer was exchanged using the same filter into Tris-saline by repeated washing with Tris-saline.

Proteins were further purified by size exclusion chromatography. The buffer exchange step after HisTALON purification was also used to concentrate the protein to a volume of 1 – 3 ml. A HiLoad Superdex 16/600 200 pg column was equilibrated in Tris-saline by washing through 150 ml at a flow rate of 1.5 ml/min using an ÄKTA FPLC system. The small volume of concentrated sample was loaded onto the column using a 5 ml injection loop and the gel filtration was carried out by washing the column with Tris-saline at a flow rate of 1.5 ml/min for 1 CV (120 ml), collecting 1.5 ml fractions. Peak fractions were analysed by SDS-PAGE and desired fractions were pooled and concentrated using an Amicon Ultra – 15 ml 50 kDa cutoff filter.

All purification of insect cell expressed proteins was carried out at room temperature.

2.2.8.5 Magnetic bead red blood cell pull down

100 µl of His-tag Isolation Dynabeads (Life Technologies: 10103D) were washed 3 x in 200 µl of HBS-P + $CaCl_2$ by pulling down the beads with a magnet, removing the supernatant and resuspending the beads in fresh buffer. The beads were resuspended in 100 µl of fresh buffer. 50 µl of 10 µM (monomeric concentration) purified NA was added, mixed and incubated at room temperature for 30 mins. The beads were washed 3 x with 200 µl of HBS-P + $CaCl_2$ and resuspended in 100 µl HBS-P + $CaCl_2$. 3 µl of coated beads were added to 100 µl of 0.5% (v/v) TRBCs suspended in HBS-P + $CaCl_2$ in a v-bottom 96 well plate (Greiner Bio One: 651101),

washed as described in section 2.2.4.1. Some experiments were supplemented with 100 μM oseltamivir carboxylate. The TRBC and bead mixture was incubated at room temperature for 15 mins, mixed thoroughly and the beads pulled down with a magnet. The supernatant was removed and transferred to another v-bottom 96 well plate and allowed to settle. Control experiments were carried out with no protein attached to the beads (a 'no protein' control), and a well where blood was allowed to settle in the absence of a magnetic bead pull down (a 'no bead' control). For quantitation of blood remaining in the supernatant the blood was lysed by the addition of SDS to a final concentration of 0.5% (w/v) (Sigma: L3771). Debris was removed by centrifugation at 3000 x g for 15 mins. Supernatant was removed and the absorbance of released haemoglobin measured at 540 nm using a Biotek synergy 2 plate reader. Absorbance values obtained were corrected for background absorbance of buffer and normalised to the no protein control values obtained. Data shown are a repeat of three independent experiments.

2.2.9 Influenza virus quantification

Virus was quantified using two different methods: SDS-PAGE or ELISA. SDS-PAGE quantification was used only to quantify egg-propagated gradient-purified viruses. ELISA quantification was used for both egg- and tissue culture-propagated viruses whether gradient-purified or not.

2.2.9.1 SDS-PAGE quantification

A small volume of egg propagated virus (typically 0.5 – 5μl) was analysed by SDS-PAGE under non-reducing conditions alongside a set of bovine serum albumin (BSA) standards made from a 2 mg/ml stock (Thermo: 23209) under reducing conditions. The standard amounts of protein used were typically 0.8, 1.6, 2.4, 3.2 and

4.0 µg of BSA. The gel was stained and scanned using an Epson Perfection V750 pro on transmission mode. Gel bands were quantified using Image J software (http://imagej.nih.gov/ij). The values for quantification of the BSA were used to plot a standard curve of band intensity against protein mass using Microsoft Excel. The virus NP band, which typically runs at 50 – 60 kDa was quantified and the amount of protein present was determined from the standard curve plotted for BSA. This amount of protein was then converted into a molar quantity, using a molecular weight of NP of ~55 kDa. This molar quantity was then converted to a molar concentration. The concentration of virus was determined using the assumption that a single virion contains 590 molecules of NP, based on a total genome size of 13588 nucleotides and 23 nucleotides bound per NP (Ruigrok, 1998).

2.2.9.2 Virus quantification by ELISA

Viruses were adsorbed to Immulon 2HB (Thermo: 10795026) plates in PBS-A overnight at 4°C. Each virus was diluted 1 in 100 in PBS-A and then diluted in a 2-fold dilution series across 24 wells. The plate was blocked with 200 µl of block buffer for 1 hr at room temperature. The virus membranes were permeabilised with 200 µl of 0.2% (v/v) Triton X-100 in PBS-A at room temperature for 30 mins. The plate was washed once with PBS-Tween and incubated with 100 µl of a 1 in 2000 dilution, in block buffer, of an anti-influenza A NP mAb (FR-51 Mouse mAb, Anti-influenza type A NP, USA Centers for Disease Control, Atlanta, GA, USA) and incubated at room temperature for 1 hr. The plate was then washed three times in PBS-Tween and incubated with a 1 in 1000 dilution of a HRP-conjugated goat anti-mouse IgG (Biorad: 172-1011) in block buffer at room temperature for 1 hr. The plate was washed three times in PBS-Tween and 100 µl of substrate solution was added and the plate

incubated at room temperature for 5 mins. 100 µl of 0.1 M H_2SO_4 was added to stop the reaction. The plate was read at 450 nm with a Biotek synergy 2 plate reader. The absorbance values for an unknown virus concentration were compared to a sample of a purified SDS-PAGE quantified control virus, X-31. A_{450} values for X-31 were plotted against Log_{10}[virus] using Graphpad Prism (GraphPad Software, San Diego, California, USA.). A curve was fitted using the sigmoidal dose-response (variable slope) function. Two absorbance values of each virus with unknown concentrations were chosen from the linear section of the sigmoidal values generated. These values were converted to a Log_{10}[virus] value by interpolation, using the X-31 standard curve. These values were converted back to a virus concentration, corrected for dilution and averaged to determine the virus concentration.

2.2.10 Biolayer interferometry influenza virus binding assays

Virus binding assays were carried out using an Octet RED system (Pall ForteBio Corp., Menlo Park, CA, USA). Binding was measured to receptor analogues, which consisted of trisaccharide sugar moieties attached to a biotinylated polyacrylamide backbone (Lectinity Holdings, Moscow, Russia). These receptor analogues consisted of 20% mol sugar and 5% mol biotin. A number of different receptor analogues were used, listed in section 2.1.5. Receptor analogue sialoglycopolymers were resuspended in HBS-EP at a concentration of 0.5 mg/ml and stored at -20°C. Working stocks were diluted to 10 µg/ml in HBS-EP and stored at 4°C.

2.2.10.1 Equilibrium binding assays

The purpose of the equilibrium binding assays is to generate data giving information on the relative avidity of the virus for particular receptor analogues. Virus equilibrium binding was carried out by measuring the binding of a constant

concentration of virus to a range of different immobilisation densities of receptor analogues loaded onto streptavidin-coated biosensors (Pall ForteBio Corp., Menlo Park, CA, USA). Receptor analogues were typically loaded to 8 biosensors simultaneously at concentrations starting at 0.5 μg/ml, serially diluted 1.75-fold across the 8 wells. The level of sugar loading was altered by changing the time of association to suit the avidity of the particular virus being studied (180 – 600 s). All experiments were carried out in HBS-EP. Virus was diluted to 100 pM in HBS-EP + 100 μM each of oseltamivir carboxylate (Roche, Welwyn Garden City, U.K.) and zanamivir (GSK, Stevenage, U.K.) to prevent the cleavage of receptor analogues by the viral NA. A typical experiment for measuring virus avidity was: 240 s baseline in HBS-EP, 180 – 600 s sugar loading, 240 s baseline in HBS-EP, 1800 s association in virus, 600 s dissociation in HBS-EP. Experiments were repeated at least twice for each virus-receptor combination.

Raw data of the experiments were extracted using the ForteBio Analysis package and the data were fitted using in-house software packages written by Dr. Stephen Martin (Francis Crick Institute, Mill Hill Laboratory). The amplitude of sugar loading was determined by measuring an average value of the last 30 s of the first baseline and subtracting it from an averaged number for the last 30 s of the second baseline. The association phase for virus binding was fitted to a double exponential function using non-linear least-squares methods to determine the offset, or extrapolated plateau value, and hence the amplitude for the curve.

The amplitude values were plotted as a function of sugar loading using Graphpad Prism (GraphPad Software, San Diego, California, USA.). A curve was fitted using the sigmoidal dose response (variable slope) function. Amplitude values were

converted to fractional saturation values by normalising the data, between the best fit values determined for Bottom and Top plateau, to values between 0 and 1. The normalised data were then fitted to a modified version of the Hill equation (Equation 2.1) using the custom non-linear least squares curve fitting programme in Graphpad Prism:

$$f = \frac{RSL^n}{RSL_{0.5}{}^n + RSL^n} \tag{2.1}$$

(Where f = fractional saturation, RSL = Relative sugar loading, $RSL_{0.5}$ = Relative sugar loading value where f = 0.5, n = a Hill coefficient).

These f values can be converted to apparent affinities for virus binding ($K_{d(Virus)}$) using Equation 2.2:

$$K_{d(Virus)} = \frac{[Virus](1-f)}{f} \tag{2.2}$$

Inspection of data obtained for more than ten different viruses showed that the ratio of the apparent affinities for any two receptor analogues (typically 3SLN-PAA and 6SLN-PAA) is given to a good approximation by Equation 2.3 (Stephen Martin, personal communication):

$$\frac{K_{d(Virus)}\,for\,6SLN}{K_{d(Virus)}\,for\,3SLN} = \left(\frac{RSL_{0.5}\,for\,6SLN}{RSL_{0.5}\,for\,3SLN}\right)^{4.5} \tag{2.3}$$

This equation with appropriate modification is also applicable in calculating the ratio of the apparent affinities for two different viruses binding to the same sugar.

2.2.10.2 Measurement of HA/NA balance

Experiments to measure HA/NA balance were carried out in HBS-P + $CaCl_2$. Experiments were carried out in the presence and absence of NA inhibitors (100μM both of oseltamivir carboxylate and zanamivir).

A typical experiment consisted of the following steps: 240 s baseline in buffer; 600 s loading of 0.5 μg/ml 6SLN-PAA/3SLN-PAA to give a saturated loading of ~0.6 nm; 240 s baseline in buffer; 3600 s virus association in the presence and absence of inhibitors at the desired virus concentration (ranging from 10 – 1000 pM); 600 s dissociation in buffer. Association measurements in the presence and absence of NA inhibitors for the same virus were made concurrently, with a no virus control to correct for drift in the signal throughout the measurement.

Data for the virus association step were extracted as a raw data file using the ForteBio data analysis package. Data were recorded with 5 data points per second (5 Hz) so consequently gives very large datasets. Datasets were reduced in size by taking every fifth value. Data were corrected for signal drift by subtracting the control dataset from all channels with virus binding data. These corrected data were then normalised to the extrapolated plateau point observed for 100 pM virus binding in the presence of NA inhibitors, as described in section 2.2.10.1. This normalization makes the response values for this curve range between values of 0 and 1, with the eventual plateau point reaching 1 (all data processing carried out using Microsoft Excel).

Normalised response data were plotted as a function of time using Graphpad Prism (GraphPad Software, San Diego, California, USA.). Curves were smoothed using

the built in function, carrying out moving average Savitsky-Golay smoothing, using 5 points on either side and smoothing with a 2nd order polynomial.

Measurements were repeated and were found to be highly consistent. All data presented were representative runs of experiments which had been repeated at least twice.

2.2.10.3 Measurement of sugar depletion

Sialic acid depletion of 6SLN-PAA and 3SLN-PAA was measured for 100 pM X-31. This was done by stopping the balance measurements in the absence of drugs at a number of different time points: 320 – 3600 s for 6SLN-PAA and 30 – 3600 s for 3SLN-PAA. At these time points the sensors were moved directly into virus of the same concentration with NA inhibitors present. The total response of virus binding was measured by determining: the initial binding response value before moving the sensor to the virus + inhibitors and the extrapolated plateau point for virus binding in the presence of inhibitors. These two response values were added together and used as the total response for virus binding. These response values were converted to fractional saturation values based on previous virus avidity measurements (see section 2.2.10.1). These fractional saturation values (f) were converted to values for relative sugar loading (RSL) using a rearranged version of Equation 2.1 (Equation 2.4):

$$RSL = \sqrt[n]{\frac{f.RSL_{0.5}{}^{n}}{1-f}} \qquad (2.4)$$

These determined RSL values were converted to a percentage of remaining sugar, assuming that for a saturated sensor RSL = 0.6 nm. Each data point determined is the mean of three independent repeats. Percentage sugar remaining was plotted as a function of time, showing the standard deviation from the mean of the repeats as error bars. A point was inserted at time = 0 as by definition 100% of sugar will remain. A curve was fitted using the two-phase decay function, using Graphpad Prism.

2.2.11 Enzyme kinetics

2.2.11.1 MUNANA kinetics

All MUNANA assays were carried out using a Jasco FP-6300 fluorimeter, fitted with an ETC-273T Peltier control unit, using an excitation wavelength of 365 nm and an emission wavelength of 450 nm (bandwidths 5 nm) at low sensitivity setting.

2.2.11.1.1 K_m determination

NA assays were carried out by fluorimetry for the small substrate 2'-(4-Methylumbelliferyl)-α-D-N-acetylneuraminic acid (MUNANA) (Sigma: M8639). MUNANA was dissolved in the analysis buffer at a concentration of 1 mM and stored in aliquots at -20°C. Experiments were carried out in disposable 2 ml cuvettes (Fisher: 10368834). 2 ml of buffer warmed to assay temperature was added to the cuvette with a dilution of protein or purified virus. Typical concentrations were 1 – 5 pM virus and 0.1 – 0.5 nM purified NA. Assays were carried out in a continuous fashion, with increasing amounts of MUNANA added to a single cuvette after measuring the initial rate of reaction for ~30 s at each stage. MUNANA was added at concentrations ranging from 5 – 200 μM. Data were corrected for dilution caused by the addition of further substrate. Data measured were in arbitrary units of dInt/min. These data were converted into μM/min using data relating fluorescence to cleaved

MUNANA concentration, determined by measuring the fluorescence of a known concentration of fully cleaved MUNANA. The difference in fluorescence measured was converted into a correction factor for each temperature and pH of buffer used (Table 2.1). These correction factors were used to convert values to µM/min values. These were then converted to an observed rate, when possible, by dividing by the monomeric concentration of NA present in the assay. The observed rate was plotted as a function of substrate concentration and fitted to the Michaelis-Menten equation (Equation 2.6) using Graphpad Prism to determine K_m and k_{cat} values.

Buffer	Temperature	Conversion factor (Int*/µM)
MES	25°C	2.49
	37°C	2.42
HBS + $CaCl_2$	25°C	11.35
	37°C	10.89

*Int = Fluorescence Intensity (Arbitrary Units)

Table 2.1: Conversion factors for fluorescence of MU released from MUNANA.

2.2.11.1.2 Determination of K_i values

K_i values for NA substrate inhibition were determined in a continuous assay. The NA used must have a known K_m value for MUNANA (determined as described in section 2.2.11.1.1). NA was used at a similar concentration to that used for K_m determination and MUNANA at a concentration of 50 µM. Initial reaction rates (v_0) were measured, followed by an inhibited rate (v_i) after the addition of inhibitors.

K_i values of inhibition by substrates 6SLN and 3SLN were determined using a 100 µl micro volume quartz fluorescence cuvette. v_0 was measured of a 100 µl before adding 1 – 10 µl of a 74.1 mM (50 mg/ml) stock of monomeric 6SLN or 3SLN (Dextra).

The inhibited rate was measured immediately, as the substrates are added at high concentration, consequently the binding of the substrate is rapid.

K_i was calculated by first correcting the v_i value for dilution of the enzyme, followed by a correction for MUNANA dilution by addition of the inhibitor using Equation 2.5:

$$v_0(corrected) = v_0 \frac{\left(\frac{[MU]_{Final}}{K_m + [MU]_{Final}}\right)}{\left(\frac{[MU]_{Initial}}{K_m + [MU]_{Initial}}\right)} \tag{2.5}$$

(Where [MU] = concentration of MUNANA)

The K_i value was determined using Equation 2.8, a rearrangement of Equations 2.6 and 2.7:

$$v_0 = \frac{V_{\max}[S]}{K_m + [S]} \tag{2.6}$$

$$v_i = \frac{V_{\max}[S]}{K_m\left(1 + \frac{[I]}{K_i}\right) + [S]} \tag{2.7}$$

$$K_i = \frac{K_m[I]}{(\frac{v_0}{v_i} - 1)\left(K_m + [S]\right)} \tag{2.8}$$

(Where [I] = concentration of inhibitor, [S] = Substrate (MUNANA) concentration)

All K_i values are a mean value of three independent measurements. K_i values for 6SLN and 3SLN are expressed as K_m values, as they are the same parameter for a competing substrate.

2.2.11.2 6SLN and 3SLN kinetics

Measurement of 6SLN and 3SLN cleavage was studied using a coupled enzyme system. Measurements were made using a Jasco V-550 spectrophotometer in an ultra-micro volume cuvette with a 10 mm path length. Measurements were made using the Buffer HBS-P + $CaCl_2$ supplemented with 1 mM $MgCl_2$ at 37°C. Each 100 µl reaction contained: 25 U β-galactosidase (Sigma: G5635); 0.25 U β-galactose dehydrogenase (Roche: 10662046001); 5 mM NAD^+ (Sigma: N7004). The substrates 6SLN and 3SLN (Dextra) were added at concentrations ranging from 50 – 7000 µM and purified NA at concentrations ranging from 0.5 – 25 nM. The 6SLN and 3SLN were found to have a small quantity (2 – 3%) of a contaminant that was found to generate a change in absorbance in the absence of NA. According to the manufacturer of the receptor analogues (Dextra Laboratories, Reading) this is likely to be N-acetyllactosamine. This contaminant was broken down by incubating each receptor at a concentration of 10 mg/ml in the presence of the reaction components listed above and incubated at 37°C for ~2 hrs. This was found to adequately remove this contaminating substance so that a change in absorbance was only seen upon the addition of NA. Absorbance at 340 nm was measured as a function of time and the maximal rate was measured. A_{340} rate measurements were converted into molar rates by using the extinction coefficient of NADH (ε = 6220 $M^{-1}cm^{-1}$). These molar rates were then converted to an observed rate by dividing values by the protein concentrations used in the assays. These observed rates were plotted as a function of

substrate concentration and fitted to the Michaelis-Menten equation to determine K_m and k_{cat} using the Graphpad Prism software.

2.2.11.3 Fetuin kinetics

Measurements of fetuin cleavage kinetics by NA were made an enzyme coupled system based on the measurement of released sialic acid by NA by using the two enzymes N-acetyl-neuramininic acid aldolase (NANA aldolase) and lactate dehydrogenase (LDH). Measurements were made in Jasco V-550 spectrophotometer using the buffer HBS-P + $CaCl_2$ in an ultra-microvolume 10 mm path cuvette at 37°C. 100 µl reactions contained 0.5 U NANA aldolase (Sigma: A6680), 5 U LDH from bovine heart (Sigma: L2625), 200 µM NADH (Sigma: N9410) and fetuin from foetal bovine serum (Sigma: F3004). The fetuin used was found to have a small quantity of contaminating free sialic acid and was pretreated under the assay conditions listed in the previous section at a stock fetuin concentration of 300 µM and incubated at 37°C for ~2 hrs. The readout of NA rate was the reduction in absorbance at 340 nm. The initial rate of reduction in absorbance determined using the supplied Jasco software. Rates of change in absorbance were changed to molar rates by using the extinction coefficient of NADH (ε = 6220 $M^{-1}cm^{-1}$). These molar rates were converted to observed rates by dividing values by the concentration of NA used in the assay. These observed rates were plotted as a function of fetuin concentration and fitted to the Michaelis-Menten equation to determine K_m and k_{cat} using Graphpad Prism software.

2.2.12 Low pH treatment of virus and trypsin digestion

Viruses were diluted to a concentration of 10 nM in PBS-A, based on NP ELISA quantitation, as described in section 2.2.9.2. The pH was lowered to 4.95 – 5.05 by the addition of 150 mM citrate buffer (citric acid/sodium citrate) pH 3.5 at ~5% of the

volume of virus. The pH was checked using a microscale pH electrode. Samples were incubated in a 37°C water bath for 15 mins. 5% sample volume of 1M Tris-HCl pH 8.0 was added to neutralise the pH of the sample which was then treated with TPCK trypsin (Sigma: T1426) at a 1:10 (w/w) ratio of trypsin:virus, assuming that a 10 nM virus suspension is ~1 mg/ml. Samples were incubated at room temperature for 1 hr and the trypsin was inhibited with an equal mass of soybean tryspin inhibitor (Sigma: T6522). Virus was pelleted 160 000 x g for 20 mins at 4°C (Beckman Coulter TLA-100 0.2 ml rotor). The supernatant was removed and the pellet was resuspended in the same volume of PBS-A + 0.01% (w/v) NaN_3.

2.2.13 Microscale thermophoresis (MST)

Microscale thermophoresis (MST) experiments were carried out using a Nanotemper Monolith NT.115 instrument. Protein labeling was carried out using the Monolith protein labeling kit RED-NHS (Nanotemper: MO-L001) according to manufacturer's instructions. Labeling was optimised to ensure a low level of labeling, ~1 dye molecule per protein multimer. The amine reactive dye was added at a concentration ~3-fold higher than the trimeric HA concentration or tetrameric NA concentration. The labeling efficiency was checked by comparing protein concentration, determined by A_{280}, as described in section 2.2.8.2, with the concentration of the dye by measuring A_{650}, as described in the manufacturer's instructions.

Experiments were carried out by diluting protein and adjusting the light-emitting diode (LED) power that gave a fluorescence of 500 – 2000 AU in the MST instrument. All experiments were carried out in Hydrophilic treated capillaries (Nanotemper). Experiments were carried out using a constant concentration of

fluorescently labelled protein in all capillaries and adding a varying concentration of either 6SLN or 3SLN (Dextra Laboratories Ltd., Reading, UK). Experiments measured thermophoresis of H1 HAs at 25°C using 30% excitation LED power and 50% infrared-laser power. Experiments consisted of: 5 s laser-off time, 30 s laser-on time, 5 s laser-off time. Measurements were made with at least 10 dilutions of the receptor analogues.

Raw thermophoresis fluorescence data were fitted and extracted using the provided Nanotemper software. Data were fitted to Equation 2.9 using the custom equation function in Graphpad Prism.

$$F = \frac{F_{\max}[L]}{[L] + K_d} + F_{\min} \tag{2.9}$$

(Where F is raw fluorescence value; [L] is concentration of ligand; and F_{max} and F_{min} are the maximum and minimum plateau values for the fitted curve)

Data were then converted to fractional saturation by subtracting F_{min} and dividing by F_{max}. Fractional saturation values (f) were plotted as a function of sugar concentration. Curves were fitted using Equation 2.10:

$$f = \frac{[L]}{[L] + K_d} \tag{2.10}$$

Previous studies (Xiong *et al.*, 2013a) using MST have shown that the major change in thermophoretic mobility occurs upon binding of the third sialic acid to an HA trimer. The experiment therefore primarily monitors the formation of this species and the observed binding constant, effectively the third stoichiometric dissociation

constant for a protein with three identical non-interacting sites, is equal to the intrinsic dissociation constant for a single site multiplied by three. Fitted K_d values were therefore corrected by dividing the value by three.

2.2.14 Nanoparticle tracking analysis (NTA)

Virus size was determined using nanoparticle tracking analysis (NTA) with a Nanosight LM-10 instrument (Malvern Instruments Ltd., Malvern, UK). Virus samples were introduced into the sample chamber at a dilution which allowed ~20 – 100 particles in a single field of view, typically a 1 in 5 000 – 50 000 dilution. Five individual 60 s measurements were made for each virus sample with the camera sensitivity set at 12. Particle sizes were fitted using the provided Nanosight NTA software using a detection cutoff of 5. Raw data for each of the 5 runs were combined for each virus, yielding 10 000 – 20 000 particle sizes for each sample. These data were converted to a frequency distribution with Graphpad Prism using the built in frequency distribution function with a bin size of 20 nm.

3 Development of Assays to Measure the Balance of Influenza Hemagglutinin and Neuraminidase Activities

3.1 Introduction

As described in the *Introduction*, section 1.6, the balance between the activities of IAV HA and NA is thought to be a major factor in determining virus fitness and transmissibility. The contributions of the activities of these two proteins, in altering virus receptor binding, has never been directly measured. In this chapter experiments are described which are devised to directly measure the effects of these two proteins on the interaction of viruses with receptor analogue coated surfaces.

The interaction of a virus with a receptor coated surface is dependent on a number of viral characteristics: 1) HA receptor binding affinity, K_d; 2) NA receptor destroying activity, K_m and k_{cat}; 3) distribution and relative proportions of HA and NA on the virus surface; 4) virus morphology, which can affect protein distributions and 5) NA stalk length, which can influence accessibility of the NA to the substrate.

In this chapter a new technique is described, which can make real-time measurements of virus binding both in the presence and absence of IAV NA inhibitors. This allows the quantitative analysis of the activity of the HA and NA in controlling interactions of the virus with surface immobilised receptor analogues. This chapter presents measurements performed to analyse the binding of the H3N2 high growth reassortant X-31 and related mutants in order to study the effects of making changes in the characteristics of the HA and NA.

3.2 Results

3.2.1 Choice of virus for binding studies

X-31 is a high growth reassortant of A/Aichi/2/68(H3N2) HA and NA combined with the remaining six gene segments from PR8 (A/Puerto Rico/8/34(H1N1)). This virus was chosen as a suitable candidate for establishing a new binding assay because it has been extensively characterised in many previous studies. X-31 HA was the first haemagglutinin to have its structure solved by X-ray crystallography (Wilson, Skehel & Wiley, 1981) and its binding characteristics have been extensively studied both in NMR experiments (Sauter *et al.*, 1989; 1992b) and more recently using microscale thermophoresis (MST) and biolayer interferometry (BLI) (Xiong *et al.*, 2013a). The virus also propagates to a high titre and is stable in purified form, as are the HA and NA glycoproteins purified from the virus. As well as extensive studies of wild-type X-31 there is also a well characterised receptor binding mutant of X-31, known as X-31 Horse Adsorption Mutant (X-31 HAM), which was generated by passaging X-31 in eggs in the presence of horse serum (Rogers *et al.*, 1983). X-31 HAM carries the receptor binding substitution L226Q and the HA has been shown to have significantly reduced affinity for α2,6-linked sialic acid receptors and a slightly increased affinity for α2,3-linked receptors (Rogers *et al.*, 1983; Sauter *et al.*, 1989; 1992b; Xiong *et al.*, 2013a). Shown in Figure 3.1 are BLI data for the binding of X-31 and X-31 HAM viruses to the receptor analogues 6SLN-PAA and 3SLN-PAA as a function of sugar loading. Dr. Stephen Martin has previously obtained these results. The parameters obtained from these data underpin a number of the calculations described in this chapter.

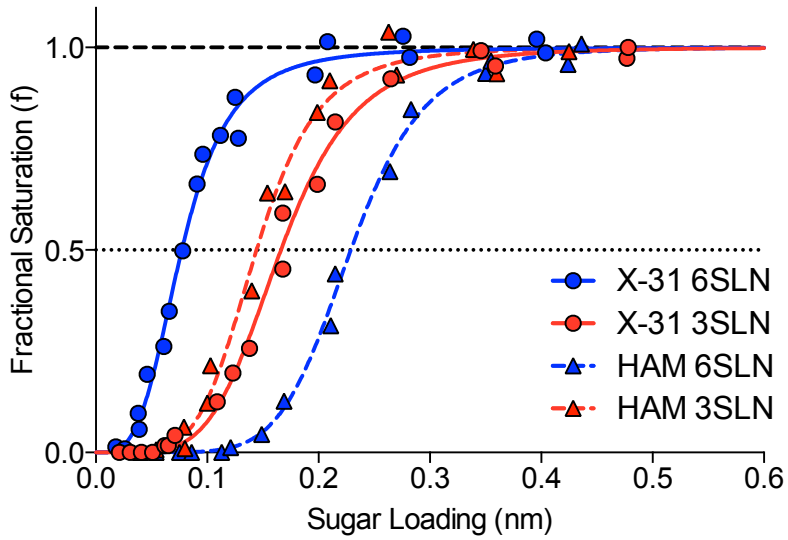


Figure 3.1: BLI data measuring HA mediated viral avidity of X-31 and X-31 HAM. Experiments measured fractional saturation by 100 pM X-31 and X-31 HAM in the presence of NA inhibitors as a function of sugar loading of 6SLN-PAA and 3SLN-PAA. X-31 HAM contains the substitution L226Q. Data obtained by Dr. Stephen Martin.

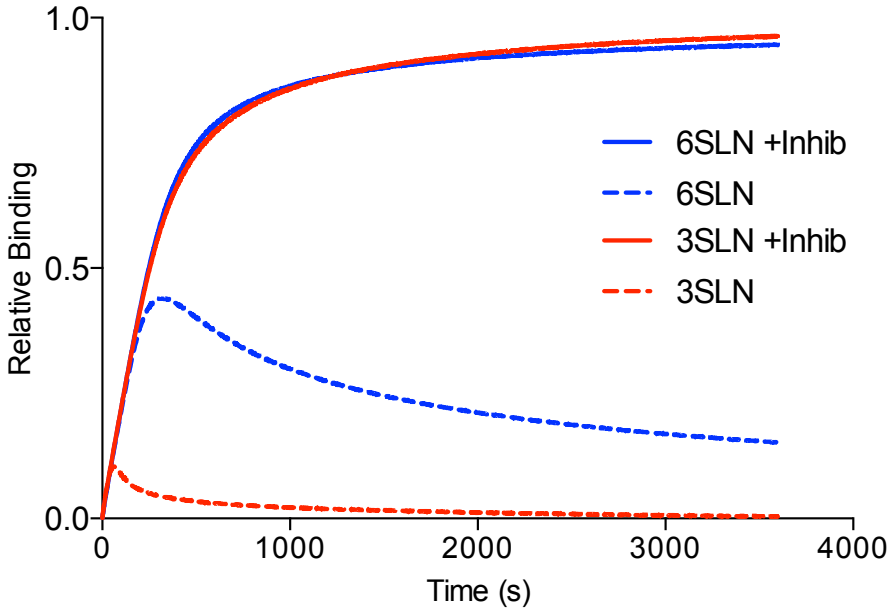


Figure 3.2: BLI data for 100 pM X-31 binding to sugar coated biosensors in the presence and absence of NA inhibitors. Biosensors were saturated with sialoglycopolymers bearing the receptor analogues 6SLN and 3SLN. Measurements are presented for virus binding in the presence (solid lines) and absence (dashed lines) of NA inhibitors (Inhib).

3.2.2 Development of binding assays to measure the balance of HA and NA activities

Experiments were carried out to measure the balance of HA and NA activities using biolayer interferometry (BLI). This approach has been used extensively to measure HA mediated virus binding in the presence of NA inhibitors (Lin *et al.*, 2012; Xiong *et al.*, 2013a; 2013b; 2013c; Crusat *et al.*, 2013; Vachieri *et al.*, 2014; Xiong *et al.*, 2014; Collins *et al.*, 2014). These experiments measured the binding of purified virus at a fixed concentration (typically 100 pM) to biosensors with different levels of immobilised sialoglycopolymers. The amount of sialoglycopolymer immobilised is characterised by the response observed during loading (known as relative sugar loading (RSL)). The total equilibrium response of virus binding to these different sugar loadings is measured, converted to a fractional saturation, and plotted as a function of the RSL. These measurements allow the quantification of virus binding affinity, as stronger binding viruses will reach saturation at lower relative sugar loadings.

The basis of the experiments developed to measure HA/NA balance is similar to the experiments described above. Two experiments are run concurrently: one with the NA inhibited with high concentrations of the two NA inhibitors, oseltamivir carboxylate and zanamivir, and the other uninhibited. The experiment with inhibited NA allows the measurement of the binding of the virus mediated by the HA only, whereas in measurements without inhibitors the combination of the effects of HA and NA activities on receptor binding can be seen.

As well as carrying out some experiments in the absence of NA inhibitors there were also a number of other changes in the experimental approach. Previous assays

were carried out using the buffer HBS-EP (10 mM HEPES-NaOH, pH 7.4, 150 mM NaCl, 3 mM EDTA and 0.005% Tween-20), which is unsuitable for assays where NA activity is important, as it contains EDTA and no Calcium, which can be important for NA activity (as described in *Introduction*, section 1.3.2.1). An altered buffer, HBS-P + $CaCl_2$ (10 mM HEPES-NaOH pH 7.4, 150 mM NaCl, 0.005% Tween-20, 4 mM $CaCl_2$) was used for all HA/NA balance measurements. Rather than carrying out experiments at different sugar loadings as with previous experiments, binding was measured to a biosensor saturated with sialoglycopolymers. These included sialoglycopolymers bearing the receptor analogues α2,6-Sialyl-N-acetyllactosamine (6SLN-PAA) or α2,3-Sialyl-N-acetyllactosamine (3SLN-PAA). The measurements are designed to look at the ability of NA to cleave sialic acid from the surface, and a fully saturated sensor allows for the highest virus binding capacity (and hence largest signal) before virus dissociates following cleavage of the sugars from the surface of the biosensor by the NA. The method for these measurements are described in greater detail in *Materials and Methods*, section 2.2.10.2.

Figure 3.2 shows a typical balance measurement for the binding of 100 pM X-31 to the α2,6 and α2,3 linked sialic acid receptor analogues 6SLN-PAA and 3SLN-PAA. The curves with inhibited NA show similar binding to both 6SLN-PAA and 3SLN-PAA, indicating a similar rate of binding to saturated sensors for both receptors. The binding curves for inhibited and uninhibited NA show a very similar initial binding phase. There is then a point at which the curves diverge, with a reduction in overall binding when the NA is active (without inhibitors present). The curves of virus binding in the absence of NA inhibitors then exhibit a turning point where the curve

flattens out and then starts to decrease slowly to very low levels of binding at later time points.

The overall shape of the curves for X-31 binding to 6SLN-PAA and 3SLN-PAA are similar; however, there is a substantial difference in the point at which the curves diverge. The 3SLN-PAA binding curves diverge at a much earlier time point (~30 s), whereas the 6SLN-PAA binding curves diverge at a later point (~320 s). The binding to 3SLN-PAA only reaches a total binding of ~10% of the total attainable virus binding capacity, whereas 6SLN-PAA reaches ~50% of total capacity. This difference in binding levels reflects the different properties of the HA and NA in binding to (HA) and cleaving (NA) the immobilised receptor. A striking feature of the 6SLN-PAA curve is that it does not reach a relative binding level of zero (no virus remaining bound) even at very long reaction times.

3.2.3 Effect of altering virus concentration on HA/NA balance measurements

The experiments to measure HA/NA balance described in the previous section (3.2.2) were all performed with the virus concentration (100 pM) used in the previously described avidity measurements. Experiments were therefore carried out to examine the effect that virus concentration has on the virus binding characteristics. Figure 3.3 shows X-31 binding profiles at the same concentration as before (100 pM) as well as 10-fold higher (1 nM) and 10-fold lower (10 pM) concentrations. The curves measured with NA inhibitors follow the expected pattern in terms of kinetics, with the rate of binding increasing with virus concentration. The level of saturated virus binding observed with 1 nM virus (normalised to that seen with 100 pM virus) is ~1.3, which is surprising, as both 100 pM and 1 nM are concentrations well in excess of the K_d of IAV binding to a saturated sugar coated surface, 10^{-12} – 10^{-13} M

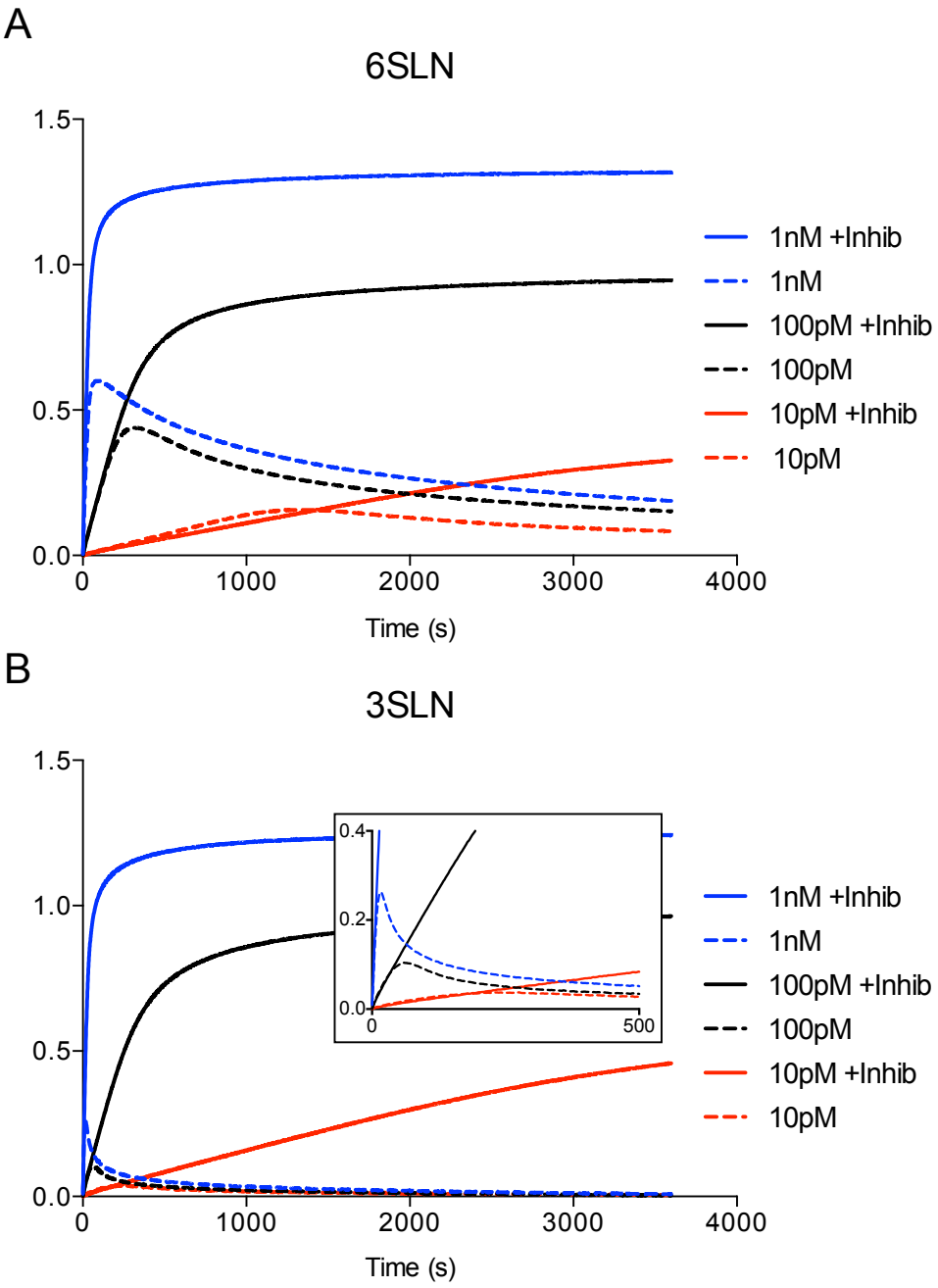


Figure 3.3: HA/NA balance curves to investigate effect of altering virus concentration. Curves show the binding of different concentrations of X-31 in the presence (solid lines) and absence (dashed lines) of NA inhibitors (Inhib). A) Virus binding to 6SLN-PAA and B) Virus binding to 3SLN-PAA with inlay showing an expanded section of initial binding.

(Stephen Martin, personal communication) and therefore they should give very similar saturation values. This increase in binding signal could be due to some form of virus aggregation on the sensor at high concentration, which would increase the signal in the way seen in these experiments.

The curve for the lowest virus concentration (10 pM) appears, at first sight, to saturate at a lower level (<1) but the reaction is very slow at this concentration and the binding is clearly not complete within the time scale over which measurements were made.

Without NA inhibitors present the curve for the highest concentration of virus (1 nM) shows faster initial binding with a higher peak of binding compared with the standard curve (100 pM). This binding then starts to reduce at an earlier time point and reduces to a level that is higher than that seen with 100 pM virus. The curve with the lowest virus concentration (10 pM) shows slower overall binding with a lower peak at a later time point and then lower amounts of virus remaining bound at long times. The data have the same pattern of differences for binding curves for both 6SLN-PAA and 3SLN-PAA; however, as seen in Figure 3.2, the loss of receptor binding in the absence of NA inhibitors happens much faster at all virus concentrations with the α2,3 linked receptor analogue. In the light of these results it was decided to carry out all further experiments using virus at 100 pM since it provides a good balance of adequate signal at early time points, without demanding the use of large amounts of virus.

3.2.4 Effect of assay temperature on measurements

As noted elsewhere, BLI experiments were typically carried out at 25°C. However, as the instrument has the facility for sample heating some experiments were carried out to examine the effect of performing measurements at more biologically relevant elevated temperatures. Experiments were not typically carried out at this temperature for practical reasons. Since, the BLI instrument holds the samples in an open 96-well plate significant evaporation from the wells occurs at higher temperatures. During long data accumulations this results in significant changes in sample concentrations.

Figure 3.4 shows a comparison of X-31 binding to 6SLN-PAA and 3SLN-PAA at 25˚C and 37˚C. Virus binding profiles in the presence of inhibitors are similar at the two temperatures, with slightly faster binding at the higher temperature (consistent with the normal effect of temperature on association rate constants). There does, however, appear to be significant loss of virus from the sensor at the later time points, particularly with 3SLN-PAA. This probably derives from two effects: the lower affinity of the virus and incomplete inhibition of the NA activity by the inhibitors at the elevated temperature.

In the absence of inhibitors the virus binding profiles are more different at the two temperatures; at 37^{o}C the maximal binding is reduced and the turning point of the curve occurs sooner. This reduction in virus binding at the higher temperature is due to two factors: a decrease in the affinity of the HA and an increase in the activity of the NA, resulting in more rapid depletion of the sugar from the sensor surface. The effect on the reduction of virus binding appears to be somewhat larger for 3SLN-PAA, with only very low levels of virus binding observed before sugar depletion.

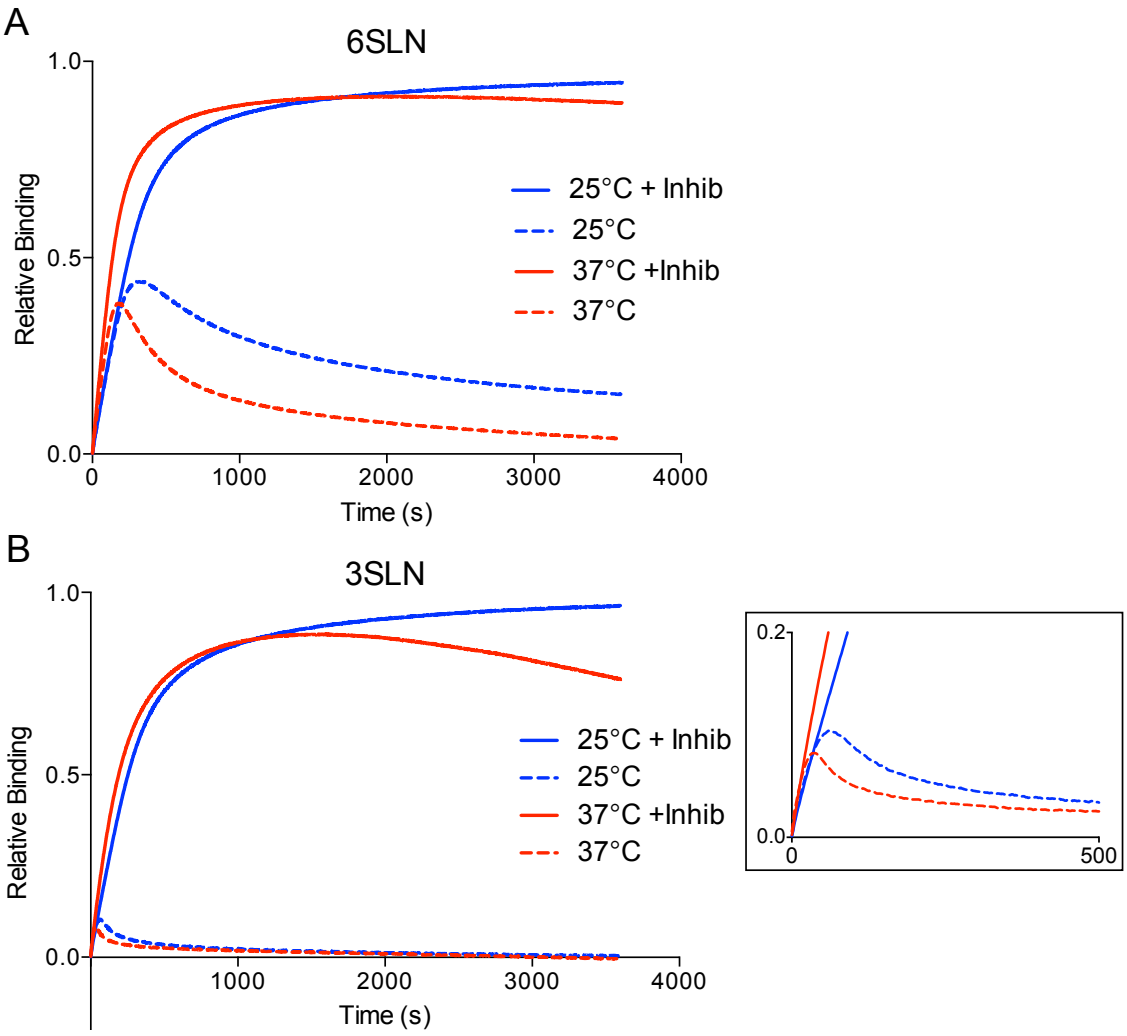


Figure 3.4: HA/NA balance measurements of X-31 binding at 25˚C and 37˚C. Measurements are made of 100pM X-31 binding to sensors saturated with the two receptor analogue sialoglycopolymers A) 6SLN-PAA and B) 3SLN-PAA in the presence and absence of NA inhibitors (Inhib).

3.2.5 Investigation of complementing HA/NA balance measurements with soluble NA

The viral NA is tethered to the virion and therefore is limited in its ability to encounter substrate by the nature of the interaction of the virus with the sensor surface. Experiments were devised to remove this limitation by blocking the viral NA and complementing the solutions with soluble bacterial NA. Experiments consisted of measuring the binding of 100 pM X-31 in the presence of influenza NA specific inhibitors: oseltamivir carboxylate and zanamivir. The NA activity of the solution was then supplemented with different amounts of soluble bacterial NA from *Clostridium perfringens* (CPNA) added to the virus dilutions. This bacterial NA is not inhibited by inhibitors of the viral NA. The action of the CPNA is then that of random depletion of the sugar affected only by the ability of the soluble protein to meet the tethered substrate, with these interactions unaffected by other proteins and the complexities of multivalency.

Figure 3.5 shows typical experiments with this CPNA supplementation at different concentrations. It is apparent from these results that the curves are similar in shape to those seen with the uninhibited viral NA (see Fig. 3.2). Higher levels of CPNA cause lower overall virus binding, owing to faster depletion of the sugars on the surface.

The inclusion of soluble NA appears to make the curves reach lower overall binding levels more quickly. For example, the curve for binding to 3SLN-PAA with 33mU/ml CPNA (Fig. 3.5B, purple line) reaches a similar initial binding to that seen with X-31 binding to 3SLN-PAA in the absence of inhibitors, giving a similar height of turning point (Fig. 3.2). This implies that under both of these conditions a similar

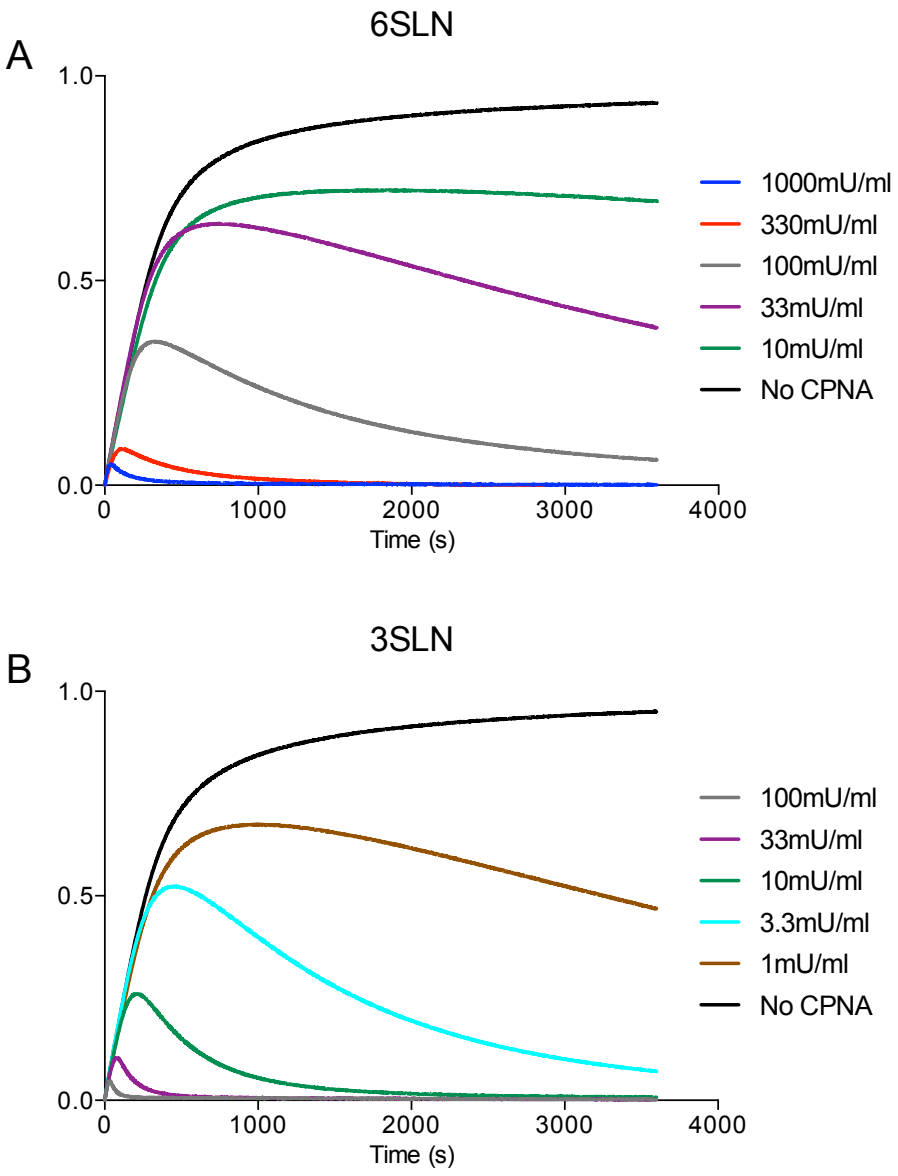


Figure 3.5: HA/NA balance curves investigating the effect of complementing X-31 receptor binding with bacterial NA. These BLI experiments were carried out with X-31 at 100 pM in the presence of NA inhibitors, with viral binding complemented with bacterial NA from *Clostridium perfringens* (CPNA) at different concentrations. Binding was measured to sialoglycopolymers bearing the two receptor analogues A) 6SLN-PAA and B) 3SLN-PAA.

amount of sugar is cleaved from the sensor surface in the initial part of the virus binding process by the viral NA or the CPNA. However, in the CPNA complemented experiment a relative binding value close to zero is reached much more quickly. This indicates that the cleavage at earlier time points by the virus is similar in nature to that of random depletion by a soluble NA and at later time points there are constraints upon the viral NA, reducing the effectiveness of cleavage compared to the soluble CPNA.

3.2.6 Determination of residence times of virus

As the virus cleaves sugar from the surface of the biosensor the number of interactions the viral HAs can make with the receptor surface will be reduced. Viruses will then be bound with lower affinity and this is most likely to be associated with an increased dissociation rate constant. When viruses are bound with very high affinity the dissociation rate constant will be very low and low dissociation rates are difficult to determine using BLI. However, dissociation rate constants (and hence average residence times) of the virus in the presence of NA inhibitors can be estimated for different sugar loadings if the virus affinity and association rate constant are known, or can be estimated. Assuming the following simple scheme for binding of a virus (V) to a surface (S), where k_{on} and k_{off} are the association and dissociation rate constants:

$$V + S \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} VS$$

An apparent dissociation constant (K_d) can be determined from the fractional saturation (f) values previously determined for X-31 and X-31 HAM binding to 6SLN-

PAA and 3SLN-PAA shown in Figure 3.1 using a rearrangement of Equation 3.1 to produce Equation 3.2:

$$f = \frac{[Virus]}{[Virus] + K_d} \tag{3.1}$$

$$K_d = \frac{[Virus](1 - f)}{f} \tag{3.2}$$

The association rate constant (k_{on}) was determined by carrying out experiments in which the binding of virus at different concentrations (125-1000 pM) was studied in the presence of NA inhibitors. In these experiments the sensors were all loaded with the same (saturated) amount of immobilised sugar. The binding curves for the different virus concentrations were fitted to a double exponential function and the fast observed rate component (k_{obs}) of the association phase, which accounted for the majority of the reaction amplitude, was plotted as a function of virus concentration, shown in Figure 3.6. This is a pseudo first-order reaction because the concentration of the virus is far greater than the concentration of immobilized sugar. A linear fit then gave an apparent association rate constant (k_{on}) of ~4 x 10^7 $M^{-1}s^{-1}$ for both 6SLN-PAA and 3SLN-PAA. Using this value and the K_d values, calculated using Equation 3.2, the dissociation rate constant, k_{off}, can be estimated for different sugar loadings ($k_{off} = k_{on}.K_d$). The apparent virus residence time is the reciprocal of the dissociation rate constant ($1/k_{off}$). These approximations are possible owing to the fact that the association rate constant for virus binding is effectively independent of sugar loading (Stephen Martin, personal communication).

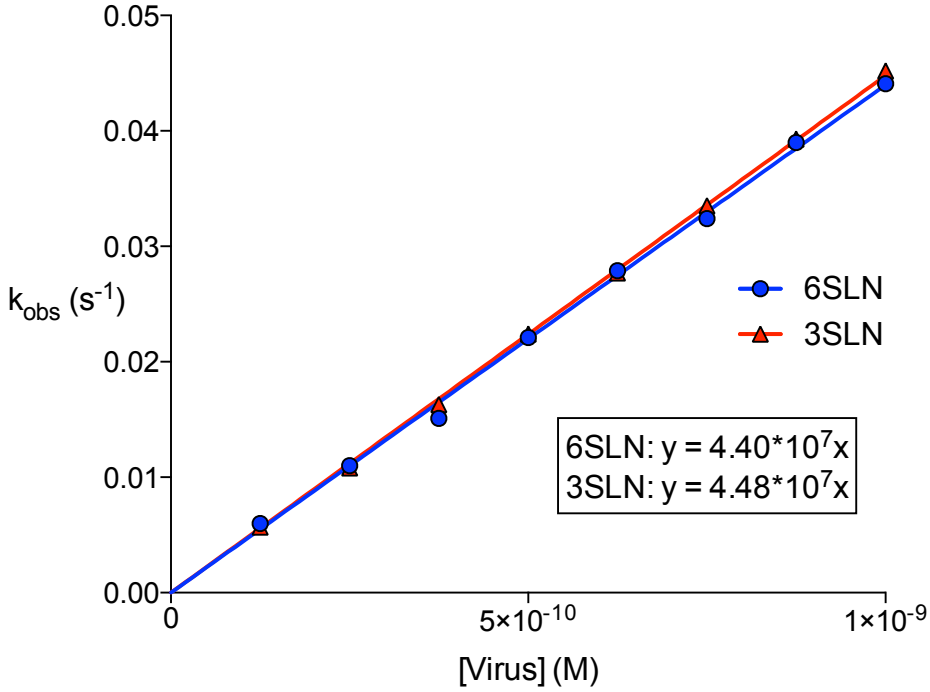


Figure 3.6: Determination of k_{on} for X-31 binding. Virus binding was measured to sensors saturated with sialoglycopolymers bearing 6SLN or 3SLN at different virus concentrations, in the presence of NA inhibitors. Binding curves were fitted to a double exponential function. The fast rate component (k_{obs}) was plotted as a function of virus concentration. Lines of best fit were plotted, forcing the line through the origin. Equations of the fitted lines are shown.

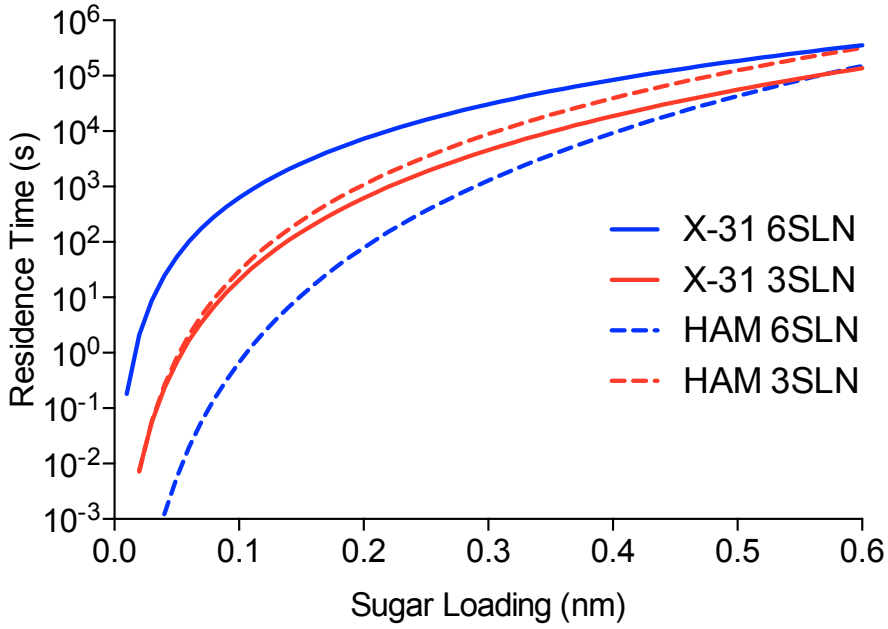


Figure 3.7: Estimated virus residence times as a function of sugar loading. k_{off} calculated for X-31 and X-31 HAM using the k_{on} value determined previously for X-31 and apparent K_d values determined at different sugar loadings ($k_{off} = k_{on}.K_d$). Residence time = $1/k_{off}$.

This k_{on} value of 4 x 10^7 $M^{-1}s^{-1}$ seems fast for a virus-surface interaction**.** The upper limit for a protein-protein interaction association rate constant, k_{on}, is ~10^9 $M^{-1}s^{-1}$, which is limited by the speed of diffusion. The determined k_{on} for virus binding is therefore very high for the interaction of a virus, which will have a much slower diffusion rate. However, k_{on} values also depend on what fraction of encounters are productive. For small molecules this fraction is small but in the case of influenza viruses it is likely to be close to 1. A study identified an association rate constant, k_{on}, value for viruses binding to cells of ~10^{12} $M^{-1}s^{-1}$ (Nunes-Correia, Ramalho-Santos & Nir, 1999). This value cannot be correct due to the constraint of the rate of diffusion. Two other studies by Meng *et al.* and Hidari *et al.* measured the k_{on} of a number of different influenza viruses binding to gangliosides or sialylated glycoproteins to give the more reasonable value of 1 x 10^6 $M^{-1}s^{-1}$ (Hidari *et al.*, 2007; Meng & Marriott, 2010).

Figure 3.7 shows the residence time for a bound virus plotted as a function of sugar loading over the range used in these experiments. It is apparent from these estimated values that the residence time of a virus at low sugar loadings is expected to be very short. The differences in HA receptor binding, for example the preference of X-31 for 6SLN-PAA over 3SLN-PAA is amplified at low sugar loadings and this is reflected in large differences in residence time. At saturated loading (RSL = 0.6 nm) the residence time of X-31 on 6SLN-PAA is 3.4 x 10^5 s, compared to 1.4 x 10^5 s on 3SLN-PAA, a difference of ~2.4-fold. At one fifth of the saturated loading value (RSL = 0.12) the residence times are reduced to 1.2 x 10^3 s for 6SLN-PAA and 51 s for 3SLN-PAA, a difference of ~23-fold. At very high sugar loadings the virus binds to both sugars with a similar (very low) K_d and the residence times are therefore similar. At

lower sugar loadings the K_d values for the two sugars increase and the difference between them becomes greater, accounting for the difference between residence times. The differences between receptor binding properties of X-31 and X-31 HAM (Fig. 3.2) can be seen in the calculated values for residence times. The slight increase in affinity of the virus for 3SLN-PAA causes an increase in residence time. The large reduction in affinity for 6SLN-PAA can be seen in a large reduction in virus residence time at all values, particularly at lower sugar loading values.

For both X-31 and X-31 HAM with both receptor types the slowest dissociation rate constant estimated for experiments performed at saturating sugar is probably of the order of 10^{-5} s^{-1}. This is similar to values determined for virus dissociation from a sugar-coated surface using surface plasmon resonance (SPR) (Stephen Martin, personal communication). These data were not replicated using BLI for practical reasons. Determining slow dissociation rate constants requires very long data collection times (>12 hrs) and too much sample evaporation occurs over such a long time, which will alter the virus concentration.

3.2.7 Kinetic characterisation of X-31 NA

It is clear that any explanation of the balance measurements described above must consider the interplay between the HA binding and NA sialidase activities. Although the binding properties of X-31 HA have been extensively characterized the kinetic properties of the NA from this virus are much less well understood. What is known is that the kinetic properties of the NA are highly dependent on pH and temperature. Experiments described in this section were designed to determine enzymatic properties for X-31 NA under different solution conditions. Purified X-31 NA was used in order to allow the measurement of an enzyme turnover rate (k_{cat}).

This protein was purified from bromelain digested purified egg-propagated virus and was kindly provided by Dr. Patrick Collins. Kinetic parameters were measured for the fluorogenic substrate 2'-(4-methylumbelliferyl)-α-D-N-acetylneuraminic acid (MUNANA) and also for the sugar substrates 6SLN and 3SLN in their monomeric form.

3.2.7.1 MUNANA kinetics

Kinetic parameters for MUNANA were measured using fluorimetry. This substrate consists of a sialic acid with a methylumbelliferyl (MU) group attached, which when cleaved by the NA releases the MU group, resulting in a large increase in fluorescence. Measurements were made using an unconventional continuous assay format. A very low concentration of NA was used (~0.2 nM monomeric concentration) and substrate was added to the reaction in a number of steps, with the rate of the reaction measured after each addition. This method of carrying out the assay works because the very low NA concentration ensures that the total depletion of the substrate over the whole time course is negligible and therefore one can assume that this does not affect the kinetic parameters determined. Measurements were made under two buffer conditions: MES buffer (32.5 mM MES-NaOH, pH 6.5, 4 mM $CaCl_2$), which is the buffer commonly used for NA assays; and HBS-P + $CaCl_2$ (HBS) (10 mM HEPES-NaOH pH 7.4, 150 mM NaCl, 0.005% Tween-20, 4 mM $CaCl_2$), which is the buffer used for the BLI balance measurements. Two different temperatures were also used for enzyme assays: 37˚C, the temperature commonly used for NA assays, as it gives maximal activity; and 25˚C, the temperature used for the majority of BLI experiments. Measurements were made and corrected for the effects of temperature and pH on fluorescence as described in *Materials and Methods*, section 2.2.11.1.1. Michaelis-Menten plots for X-31 NA MUNANA cleavage under the

different buffer and temperature conditions listed above are shown in Figure 3.8, and the fitted kinetic parameters are shown in Table 3.1, determined as described in *Materials and Methods*, section 2.2.11.1.1.

X-31 NA shows good activity both in HBS and MES buffers but surprisingly both the k_{cat} and the K_m are higher in the HBS buffer. The difference between these buffers most likely to affect MUNANA cleavage is the pH, with the MES buffer at pH 6.5 and the HBS-P + $CaCl_2$ at pH 7.4. NA is generally considered to be more active at lower pH and this is the reason that most NA assays are carried out below physiological pH. The HBS buffer also contains NaCl, at a concentration similar to that of an isotonic solution. In contrast, the MES buffer is hypotonic, containing no added NaCl and only small amounts of $CaCl_2$.

One possible explanation for the differences seen in the kinetic parameters could be instability of the NA in low salt buffers. Experiments were therefore carried out with an NA sample left diluted in the hypotonic MES buffer for extended periods(~1 hr) before use. No reduction in MUNANA cleavage was seen, indicating that X-31 NA truly does have a higher turnover rate and K_m at physiological pH in the presence of salt (data not shown). For both buffers the k_{cat} values at 25°C were ~2 – 3-fold lower than those determined at 37°C. This is a normal difference for this temperature range and is consistent with the fact IAV NA has a temperature optimum of 37°C. Likewise, for both buffers, the K_m values at 25°C were ~2-fold smaller than those determined at 37°C. This difference is typical for this temperature difference and indicates that the affinity of the NA for MUNANA increases as the temperature is decreased.

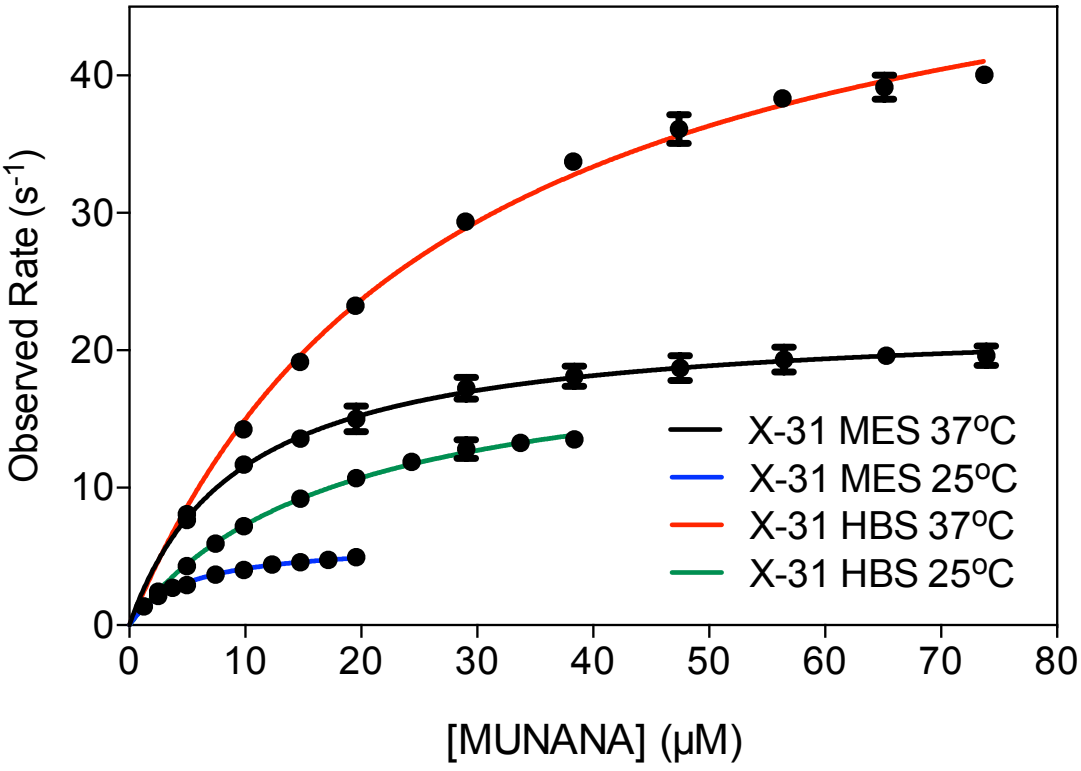


Figure 3.8: Michaelis-Menten plot for X-31 NA cleaving MUNANA. Carried out under different conditions of buffer (MES and HBS-P + $CaCl_2$) and temperature (25˚C and 37˚C). All measurements shown were made using the same NA concentration. Observed rate shown is determined by dividing the measured reaction rate by the concentration of protein used.

Substrate	Buffer/Temp	K_m ± S.E.[§] (μM)	k_{cat} ± S.E. (s^{-1})	k_{cat}/K_m ($\mu M^{-1}s^{-1}$)
MUNANA	MES[*] 37˚C	9.3 ± 0.6	22.4 ± 0.3	2.41
	MES 25˚C	4.8 ± 0.2	6.1 ± 0.1	1.27
	HBS[†] 37˚C	27.7 ± 1.1	56.6 ± 0.9	2.04
	HBS 25˚C	17.7 ± 0.8	20.3 ± 0.4	1.15
6SLN	HBS 37˚C	8070 ± 615	18.5 ± 0.9	0.0023
3SLN	HBS 37˚C	562.3 ± 20	97.5 ± 1.5	0.173

[*]MES = MES Buffer (32.5 mM MES-NaOH, pH 6.5, 4 mM $CaCl_2$)
[†]HBS = HBS-P + $CaCl_2$ Buffer (10 mM HEPES-NaOH, pH 7.4, 150 mM NaCl, 0.005% Tween-20)
[§]S.E. = Standard Error of the mean

Table 3.1: Steady state kinetic parameters determined for X-31 NA

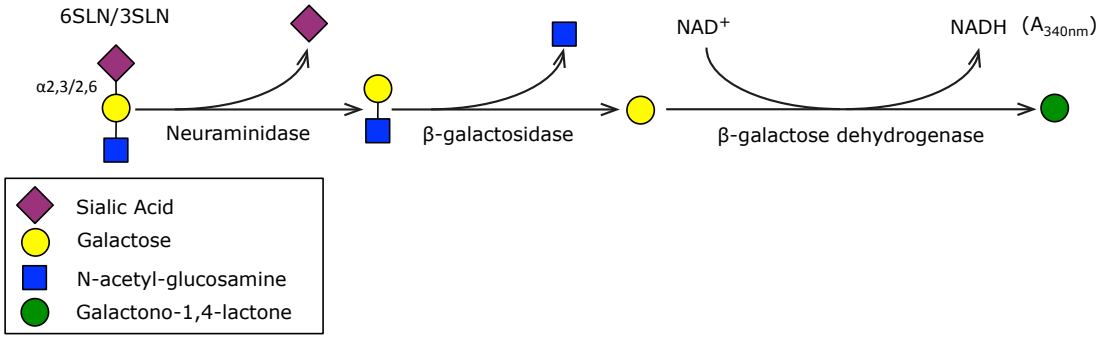


Figure 3.9: Reaction scheme of enzyme reporter system for measuring 6SLN and 3SLN kinetic parameters.

3.2.7.2 6SLN and 3SLN kinetics

In order to study the kinetics of X-31 NA cleavage of 6SLN and 3SLN an assay was developed based on an enzyme reporter system, which had been used in previous studies (Cabezas *et al.*, 1980). This system uses β-galactosidase and β-galactose dehydrogenase as reporter enzymes, with the measurement of activity based on the conversion of NAD^+ to NADH by β-galactose dehydrogenase, which can be monitored by measuring the increase in absorbance at 340 nm (reaction scheme shown in Figure 3.9). The assay was optimized to ensure that the reporter enzymes were not present at limiting concentrations, guaranteeing that the rate of reaction seen was that of the NA. This was verified by doubling the NA concentration and ensuring that it gave double the rate at both the highest and lowest concentrations of substrate used (data not shown). Measurements were all made using HBS-P + $CaCl_2$ supplemented with 1 mM $MgCl_2$ to allow for optimal activity of the β-galactosidase. All reactions were measured at 37°C, as this is the optimal temperature of enzyme activity for both the β-galactosidase and β-galactose dehydrogenase. Experiments were carried out as described in *Materials and Methods*, section 2.2.11.2. X-31 NA was used at a monomeric concentration of 8.3 nM for 6SLN kinetics and 0.83 nM for 3SLN kinetics.

Figure 3.10 shows Michaelis-Menten plots for X-31 NA cleaving 6SLN and 3SLN, and the enzymatic parameters obtained from the fits are shown in Table 3.1. It is clear from these results that X-31 NA is much more efficient at cleaving 3SLN than 6SLN, with a K_m ~15-fold lower and a k_{cat} ~5-fold higher (k_{cat}/K_m ~75-fold higher). Although it has been observed before that IAV NA has an overall preference for cleaving α2,3 linked sugars, as described in *Introduction*, section 1.3.2.2, it is not clear

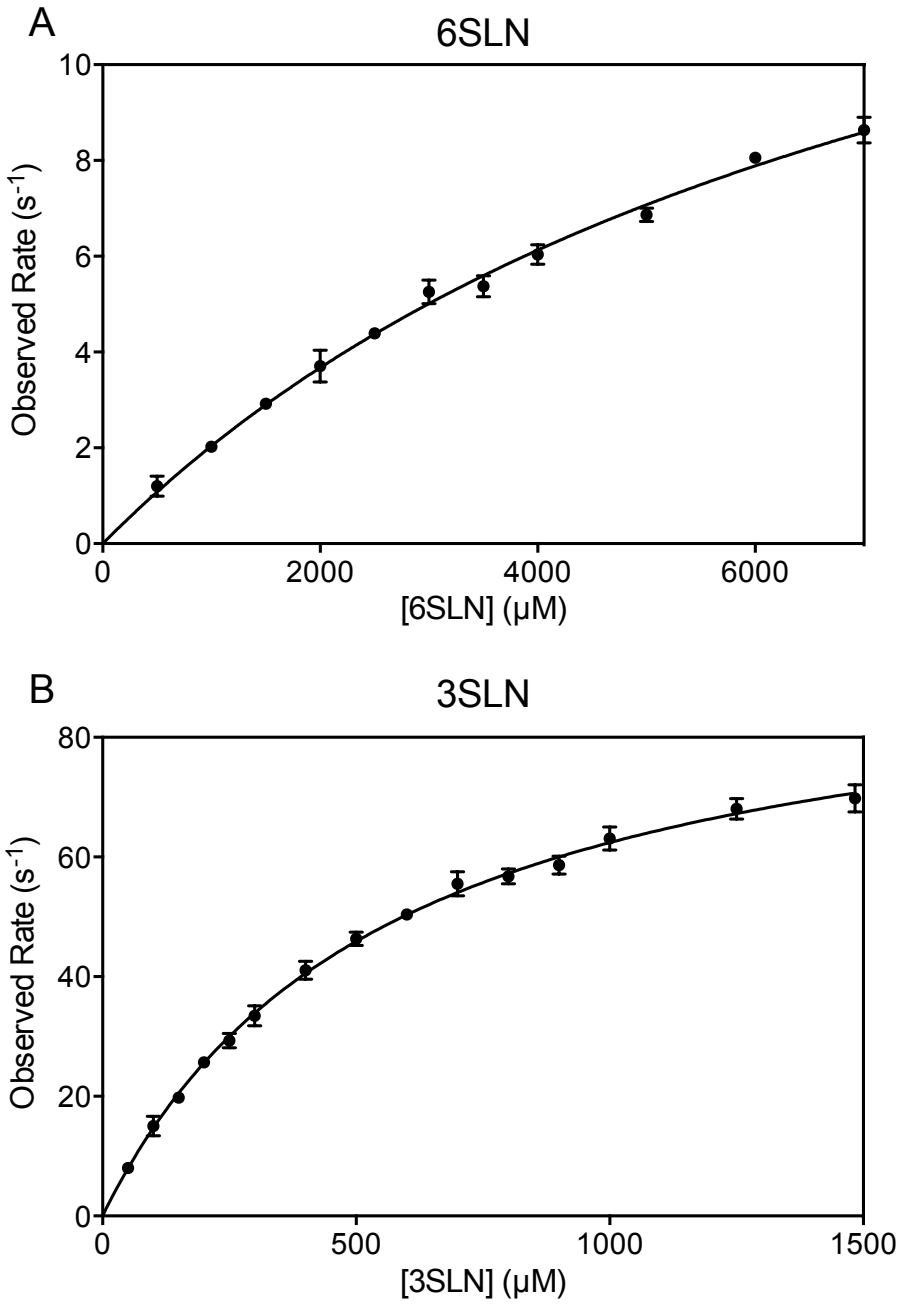


Figure 3.10: Michaelis-Menten plots of X-31 NA cleaving 6SLN and 3SLN. Measurements carried out in HBS-P + $CaCl_2$ + 1 mM $MgCl_2$ at 37˚C. Cleavage of A) 6SLN and B) 3SLN was monitored by the reporter enzymes: β-galactosidase and β-galactose dehydrogenase. Rate of reaction was measured as the rate of conversion of NAD^+ into NADH by the β-galactose dehydrogenase by monitoring the absorbance at 340 nm. Observed rate shown is determined by dividing the measured reaction rate by the concentration of protein used.

if this is a property of this particular enzyme or if it is due to the ketosidic bond being inherently more cleavable in the α2,3 linked form.

Comparing these results with the results obtained for MUNANA cleavage under the same buffer and temperature conditions (Table 3.1) shows that the K_m values for 6SLN and 3SLN are much higher than those for MUNANA. This suggests that MUNANA is bound more strongly than the natural sugars. This is likely to be due to the highly hydrophobic methylumbelliferyl group of MUNANA interacting in some way with the active site of NA (which contains a number of hydrophobic residues) and contributing to the free energy (and hence affinity) of the interaction. The turnover rate (k_{cat}) is, however, ~2-fold higher for 3SLN cleavage than for MUNANA cleavage, indicating that although the NA can bind MUNANA more efficiently it can be more efficient at cleaving the natural polysaccharide substrates. The k_{cat}/K_m values indicate that the NA of X-31 is much more efficient at cleaving MUNANA than 6SLN or 3SLN, with the value ~12-fold higher for MUNANA than the more efficiently cleaved of the natural substrates, 3SLN. These experiments highlight the requirement for NA kinetics to be measured for sugar substrates in order to fully interpret their influence on affecting the balance of HA and NA activities.

3.2.8 Characterisation of NA activity on sugar depletion

The differences between the curves presented above with inhibited and uninhibited NA must be due to the action of the NA on the surface immobilised sugars. Kinetic parameters determined for the cleavage of 6SLN and 3SLN presented in section 3.2.7.2 provide useful information on the relative abilities of the NA to cleave different substrates in solution. However, they do not give information on the effects of NA on a virus on these immobilised sugars, where the interaction of the NA with

the substrate will be influenced by the fact that HAs are bound to some of the sugars and by the fact that the NA may have restricted access to the surface attached substrate. Determining the cleavage of surface bound sugars is important in order to allow the understanding of NA cleavage in a more natural form of substrate presentation.

In order to determine the effects of NA on the surface, experiments were conducted to estimate the levels of sialic acid depletion from the sensor surface at different time points. These experiments consisted of starting a typical measurement of 100pM virus (X-31) binding in the absence of inhibitors and then stopping the measurement at specific times and transferring the sensors into wells containing the same concentration of virus in the presence of NA inhibitors. The remaining virus binding capacity of the sensor can then be estimated. 3SLN-PAA measurements were stopped at times ranging from 30 – 3600 s and 6SLN-PAA 320 – 3600 s. From these measurements the total response at long times is measured and converted to fractional saturation by comparison with previously measured binding curves with inhibited NA**.** These data were then converted to sugar loading values using the parameters obtained by fitting the data presented in Figure 3.2. These curves were fitted using a modified version of the Hill equation, Equation 3.3:

$$f = \frac{RSL^n}{RSL^n + RSL_{0.5}{}^n} \tag{3.3}$$

(Where: f = fractional saturation; RSL = Relative Sugar Loading; $RSL_{0.5}$ = RSL when f = 0.5; n = Hill coefficient)

Final fractional saturation values to determine sugar depletion at different time points were converted into sugar loadings using a rearrangement of Equation 3.3 (Equation 3.4):

$$RSL = \sqrt[n]{\frac{f.RSL_{0.5}{}^{n}}{1-f}} \tag{3.4}$$

These calculated sugar loading values were then converted to percentage of remaining sialic acid by assuming that at 100% saturation RSL = 0.6 nm. An example of the raw data for these measurements is shown in Figure 3.11, which shows measurements initiated in the absence of NA inhibitors before adding NA inhibitors at different set time points.

Figure 3.12 shows the values calculated for % remaining sugar as a function of time. The data fitted best to a double exponential function and that there are therefore two phases to the depletion. For both 6SLN-PAA and 3SLN-PAA there is an initial fast depletion of sugar followed by much slower depletion. Also shown in Figure 3.12 is the value of sugar loading, converted to a depletion value, at which the fractional saturation is 95% (f = 0.95), determined from data presented in Figure 3.1. This is the value at which HA mediated receptor binding is likely to be affected by sugar depletion. There is a broad agreement with the point that the depletion curves cross this line and the point that the X-31 binding curves with inhibited and uninhibited NA diverge. Shown in Figure 3.13 is a comparison of the HA/NA balance data shown in Figure 3.2 with the sugar depletion data from Figure 3.12.

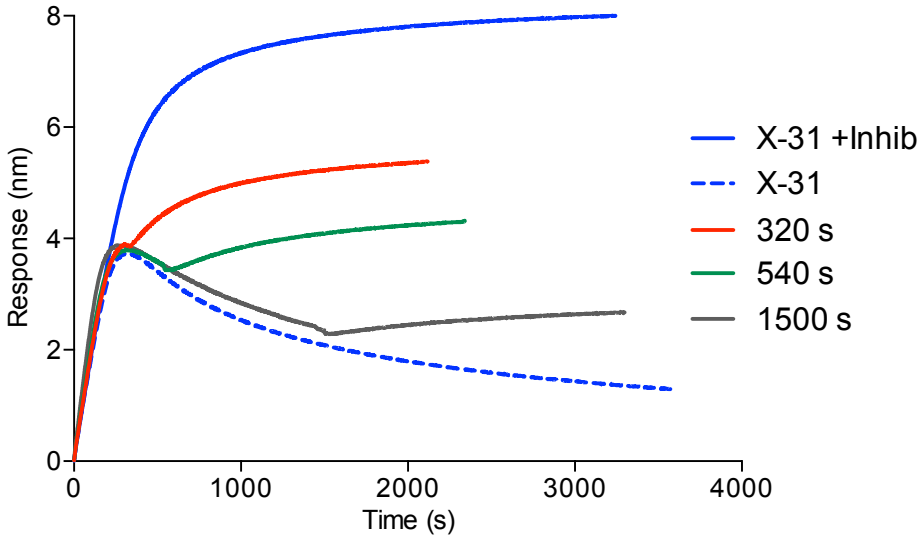


Figure 3.11: Raw data from X-31 sialic acid depletion experiments. X-31 binding measured in the absence of NA inhibitors was stopped at defined time points and the remaining receptor is saturated with virus in the presence of NA inhibitors. Data shown are for virus binding to sensors saturated with 6SLN-PAA with reactions stopped at a small number of illustrative time points.

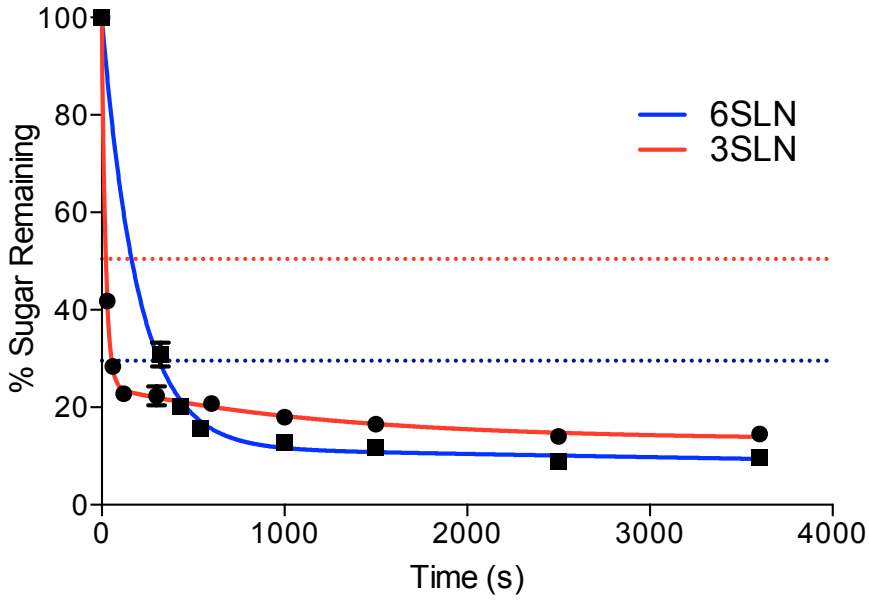


Figure 3.12: Measurement of the depletion of sialic acid from 6SLN-PAA and 3SLN-PAA by X-31. Data were collected to measure the total virus binding capacity of a sugar coated sensor at different time points during virus binding in the absence of NA inhibitors. Data shown are the mean of 3 independent measurements, with error bars showing standard deviation from the mean. Dashed lines indicate the sugar values at which the virus binds the different receptor at 95% binding capacity, therefore the level of depletion which will start to affect HA mediated receptor binding.

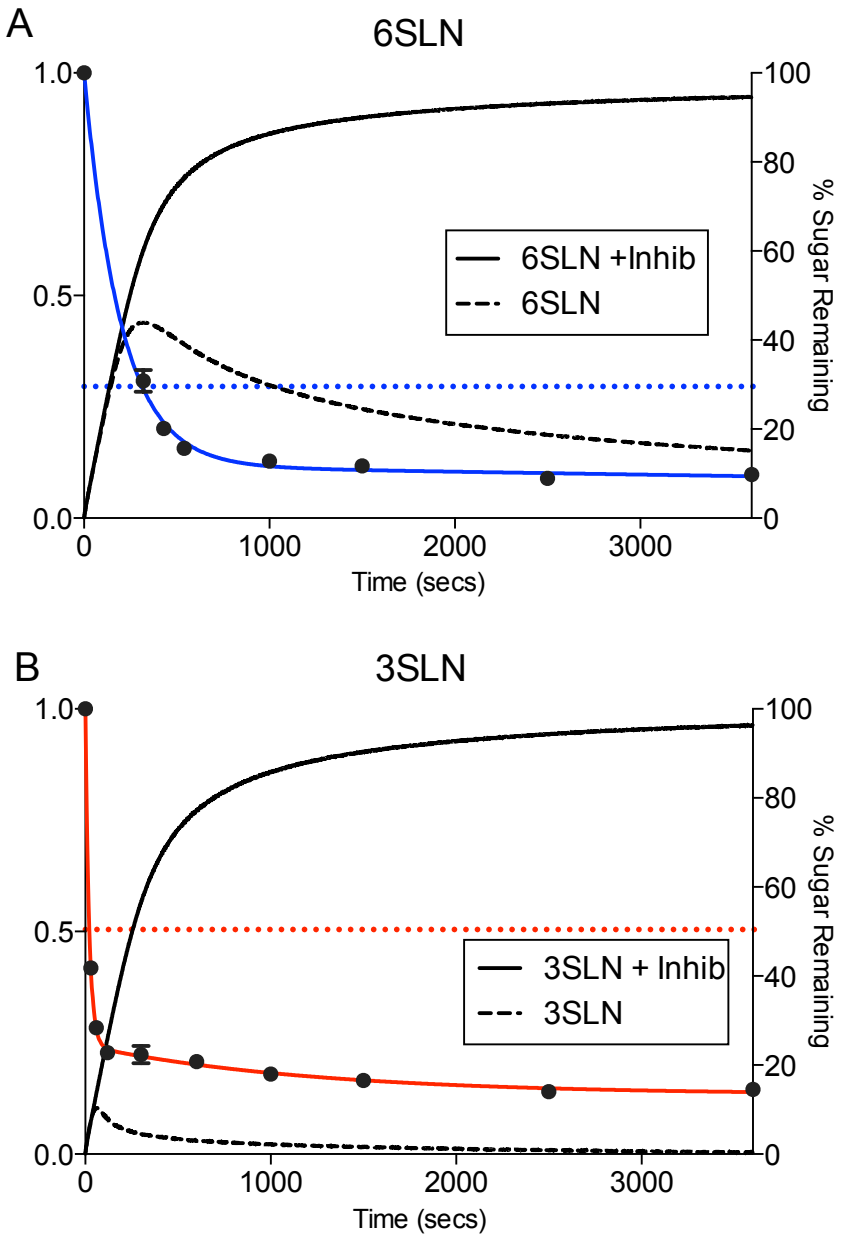


Figure 3.13: Comparison of X-31 sugar depletion data with HA/NA balance data. Both parameters measured for sialoglycopolymers bearing the two receptor analogues A) 6SLN and B) 3SLN.

From these data it is clear that the cleavage of 3SLN-PAA is much quicker than that seen for 6SLN-PAA, with >50% of sialic acid depleted by 30 s for 3SLN-PAA compared to 320 s for 6SLN-PAA. This is likely due to the much greater catalytic efficiency of X-31 NA in cleaving α2,3-linked sialic acid, as seen from the enzyme kinetic parameters shown in Table 3.1.

An interesting feature of the sugar depletion curves shown in Figure 3.12 is that the amount of 3SLN-PAA remaining at later time points is higher than that seen with 6SLN-PAA. This is presumably due to shorter-lived HA-mediated interactions of the virus with the 3SLN-PAA when the sugar loading has been reduced to low levels.

3.2.9 Effect of altering HA receptor binding affinity

All of the data presented above were obtained with the same virus, X-31. Experiments described in this section were carried out to examine the effect of altering HA receptor specificity. X-31 HAM is a mutant of X-31 containing the single substitution L226Q in the HA, selected by passaging the virus in the presence of horse serum (Rogers *et al.*, 1983). All other viral characteristics remain unchanged. This mutation causes major changes in the HA-mediated receptor binding characteristics of the virus; binding to 6SLN-PAA is significantly weakened and binding to 3SLN-PAA is somewhat strengthened, as seen in Figure 3.1 (Xiong *et al.*, 2013a). The affinities of monomeric sugars for isolated HA have been measured by NMR (Sauter *et al.*, 1989) and more recently by microscale thermophoresis (MST) (Xiong *et al.*, 2013a). The L226Q substitution decreased the affinity of the HA for monomeric 6SLN (K_d = 2.1 ± 0.3 mM (X-31) and K_d = 5.9 ± 0.7 mM (X-31 HAM)) whereas the substitution resulted in a small increase in affinity of HA for monomeric 3SLN (K_d = 3.2 ± 0.6 mM (X-31) and K_d = 2.9 ± 0.3 mM (X-31 HAM)) (Sauter *et al.*, 1989). These small differences in

the dissociation constants for monomeric interactions translate into the substantial affinity differences seen when studying the multivalent interaction between a virus and a sugar coated surface (Fig. 3.1).

Experiments to determine alterations of the HA and NA balance characteristics seen with a virus carrying this substitution were carried out. Figure 3.14 shows a comparison of the binding curves for the two viruses, X-31 and X-31 HAM. The largest difference can be seen in the experiments with 6SLN-PAA binding, where there is a large reduction in overall virus binding, with the inhibited and uninhibited curves diverging at much shorter times, for X-31 HAM. This must be associated with decreased affinity of the HA for 6SLN. This reduction in affinity results in the virus having shorter-lived interactions with the biosensor surface, as calculated in section 3.2.6. The change seen in experiments with 3SLN-PAA goes in the opposite direction, with X-31 HAM overall showing more virus binding than X-31. This is consistent with the observation that X-31 HAM has a slightly higher affinity for 3SLN than does X-31.

3.2.10 Effect of NA stalk length on HA/NA Balance

The NA of IAVs, upon adaptation of the virus to new hosts, can mutate to lose sections of the NA stalk, as described in *Introduction*, section 1.3.2.4. This reduction of NA stalk length will move the catalytic head region of the NA closer to the viral membrane. Such deletions would therefore be likely to restrict the ability of the NA to effectively encounter substrate.

In order to examine the effects of this NA stalk deletion on the HA/NA balance measurements presented in this chapter a number of stalk deletion mutant viruses were constructed. The HA and NA of X-31 were cloned into a reverse genetics vector,

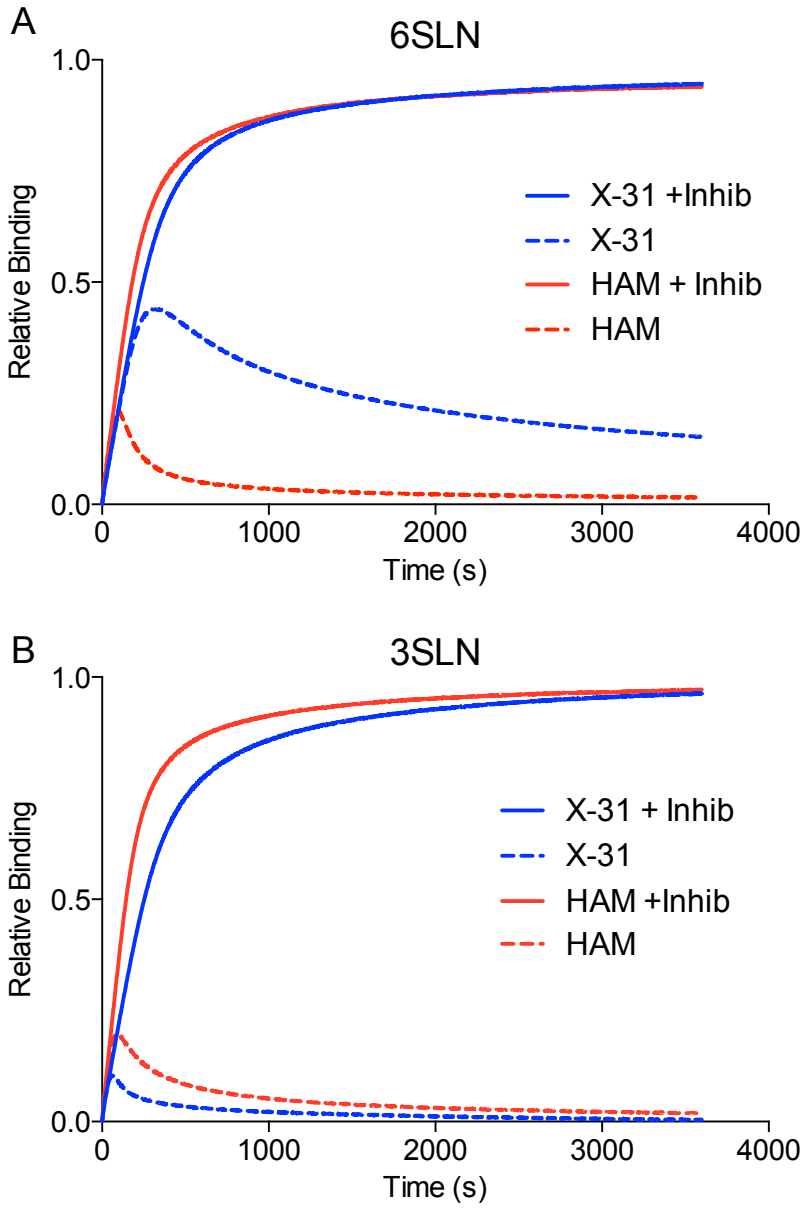


Figure 3.14: BLI HA/NA balance data for the binding of X-31 (100pM) and X-31 HAM. Measurements made in the presence and absence of inhibitors (inhib) to sensors saturated sialoglycopolymers bearing receptor analogues A) 6SLN and B) 3SLN.

as described in *Materials and Methods*, section 2.2.6.12.1. The NA was then subjected to deletion mutagenesis to make a number of different lengths of deletion, as described in *Materials and Methods*, section 2.2.6.13.2. Plasmids were constructed with deletions of: 5 (60 – 64), 10 (60 – 69), 15 (60 – 74), 20 (60 – 79), 21 (60 – 80) and 28 (54 – 81) residues. Numbers in brackets indicate amino acid numbers of the residues deleted. The deletions of 21 and 27 were based on previously reported stalk mutants generated in N2 NAs (Castrucci & Kawaoka, 1993; Sorrell *et al.*, 2010). Virus rescue was attempted for all of the stalk deletions listed above, as described in *Materials and Methods*, section 2.2.5. The transfections were inoculated directly into eggs. On passage the only virus that propagated efficiently in eggs was the virus with the 10 amino acid deletion, henceforth to be referred to as X-31 NA Δ10.

Figure 3.15 shows HA/NA balance measurements in which the binding of X-31 and X-31 NA Δ10 is compared. The binding of X-31 and X-31 NA Δ10 in the presence of NA inhibitors to 6SLN-PAA and 3SLN-PAA is very similar indicating no difference in the HA mediated binding characteristics. In the absence of NA inhibitors X-31 NA Δ10 shows more extensive virus binding, with the turning point at a later time point for both 6SLN-PAA and 3SLN-PAA. This shows that the NA with the stalk deletion is significantly less efficient at removing the surface immobilised sugars. The sialidase activities of X-31 and X-31 NA Δ10 cleaving small substrates were compared by measuring MUNANA cleavage activity of the two viruses at the same concentration. The cleavage rates were very similar (data not shown), indicating that there is no effect of the stalk deletion on the sialidase activity for small substrates. The differences seen in Figure 3.15 must therefore be due to a geometric difference in the positioning of the NA on the virion.

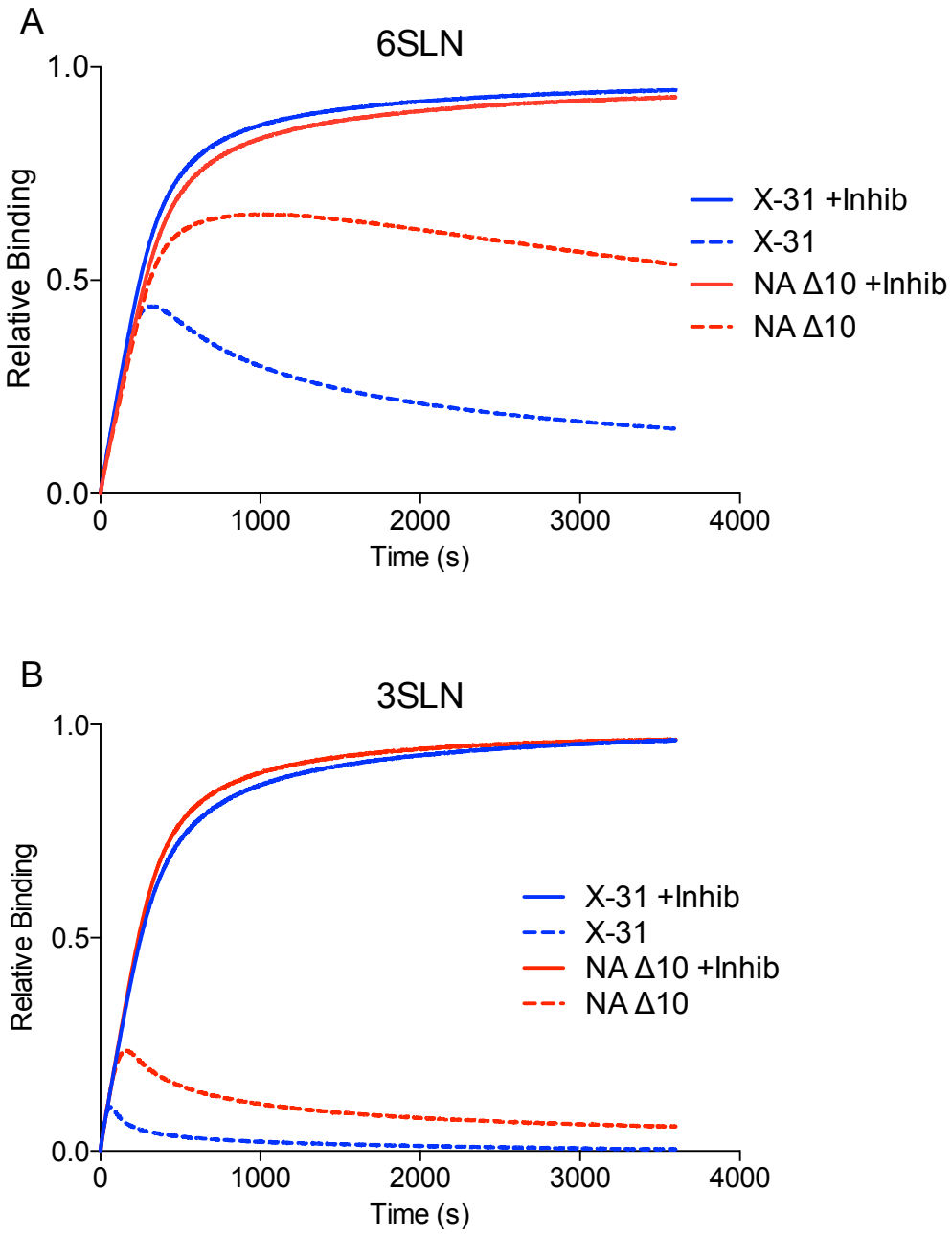


Figure 3.15: BLI HA/NA balance data for the binding of X-31 and X-31 NA Δ10 viruses. X-31 NA Δ10 virus has residues 60-69 removed from the NA stalk. Measurements made in the presence and absence of NA inhibitors (Inhib) for the binding to sialoglycopolymers bearing the receptor analogues A) 6SLN and B) 3SLN

3.3 Discussion

The experiments presented in this chapter are a new biophysical approach to answering questions regarding the balance between the activities of IAV HA and NA. This biophysical approach allows a better understanding of the processes involved in virus binding and release, which have previously been inferred from isolated kinetic and thermodynamic parameters for HA and NA, without consideration of the complexities associated with the multivalent nature of the interaction of a virus with a surface. It is clear from the data presented in this chapter that the binding and release of virus from surface immobilised sugars depends critically on the balance of activities: the ability of HA to bind receptors and the catalytic efficiency of the NA in depleting the sugars.

The interaction of HA with an isolated sialic acid is very weak, with K_ds in the mM range (Sauter *et al.*, 1989; 1992b; Xiong *et al.*, 2013a). Because these interactions are so weak the virus must form several HA-receptor interactions in a multivalent fashion in order to achieve stable attachment to a surface. If only a small number of interactions are formed the affinity is likely to be low and there is a high probability that the virus will dissociate before forming enough interactions to form a high affinity (and therefore long-lived) complex. In the case of a virus with inactive (inhibited) NA there are likely to be a large number of short-lived interactions with virus associating and rapidly dissociating. The $K_{d(receptor)}$ values determined for single receptor interactions with single sugars can be related to the multivalent apparent $K_{d(virus)}$ values. $K_{d(virus)} = K_{d(receptor)}^{mc}$, where mc = the multiplicity coefficient. The multiplicity coefficient was estimated to be in the range 3 – 5.5 (Xiong *et al.*, 2013a) for the range of sugar loadings covered in this study. For a receptor with $K_{d(Receptor)} = 3$

mM this would correspond to $K_{d(Virus)}$ values in the range 27 nM to 0.24 pM. It is important to note that that the multiplicity coefficient is not equal to the number of HA-receptor interactions that are formed. This would only be the case if each interaction contributed the same free energy, which is unlikely to be the case.

The combination of the results for X-31 binding in the presence and absence of inhibitors (Fig. 3.2), the results for the sugar depletion experiments (Fig. 3.12), and the values estimated for virus residence times (Fig. 3.7) allow a number of conclusions to be drawn concerning the interaction of the virus with the sensor surface in the absence of NA inhibitors. Initially the virus cleaves a very large amount of sugar from the surface surprisingly quickly, as shown in the sugar depletion experiments (Fig. 3.12). This rapid cleavage of sugars happens before a large amount of virus has bound to the sensor surface and has increased the signal measured. This is particularly clear when comparing the curves for X-31 binding to 3SLN-PAA in the presence and absence of inhibitors. The curves diverge after ~30 s giving a maximum response (fractional saturation) of ~0.1, indicating that only 10% of the surface is saturated before the NA has depleted so much of the sugar that the virus affinity has been dramatically reduced. This appears to be due to many short-lived interactions of the virus with the surface which, as reasoned above, are still long enough to allow time for the NA to cleave large amounts of sugar. This is particularly apparent for X-31 binding to 3SLN-PAA, likely due to the much higher efficiency of the sialidase at cleaving 2,3 linked sugars. At first sight the notion of transient virus interactions appears to be inconsistent with calculated residence times for a saturated sensor (RSL = 0.6 nm) of the order of ~10^5 s. These residence times will, however, only apply to viruses that have bound and have formed enough HA-receptor contacts to yield a

high affinity interaction. In practice this is likely to take some time as many such interactions must be formed and any particular low affinity HA-sugar contact is likely to be very short-lived.

These initial transient interactions could be considered to be similar to those at very low sugar loadings. For instance, if the interactions were similar to those at a low sugar loading of 0.02 nm the estimated residence times (Fig. 3.7) would be ~30 – 50 ms for 3SLN-PAA and ~2 – 5 s for 6SLN-PAA. What is important, however, is how much sugar the virus can cleave in these short-lived interactions. This will depend on a number of factors including: the positioning and amounts of NA on the virus, the position and distribution of substrate, and, what seems most important, the catalytic efficiency of the enzyme for cleaving substrate (k_{cat}/K_m). The much greater catalytic activity of X-31 NA for cleaving 3SLN, with K_m ~15-fold lower and k_{cat} ~5-fold higher than for 6SLN (Table 3.1), appears to be sufficient to give much more rapid sugar depletion in spite of the fact that the lower affinity of HA for 3SLN-PAA is associated with significantly shorter virus residence times.

After these initial fast cleavage events the rate of sugar depletion begins to decline very close to the time at which the curves show a turning point. This slow down in sugar depletion therefore begins at the point where the largest amount of virus is bound. This might be considered to be the best opportunity for the viral attached NA to cleave sugars; however, this does not appear to be the case. It appears that for the most efficient cleavage the virus needs to be undergoing the short-lived interactions present in the early parts of the binding assays, presumably as the viral attached NA is meeting more substrate. Virus particles will briefly interact with the sensor surface in a range of random orientations. The virus will then either dissociate

due to weak HA interactions or due to NA cleavage of the receptors. This fast turnover of virus appears to allow very efficient cleavage of the sugars.

The slow down in receptor cleavage could be due to a number of factors: 1) longer more established HA mediated virus interactions with the surface could be shielding the remaining sugars in the footprint of the virus; 2) long-lived virus interactions may have geometric constraints preventing the NA from meeting the substrate on the surface; or 3) shorter interaction times may prevent NA from interacting with the remaining surface sugars.

Some of the HA/NA balance curves recorded in the absence of NA inhibitors never reach a response value of zero. For example X-31 binding to 6SLN-PAA after 3600 s is still evident at a binding level (fractional saturation) of 0.2 or 20% of the saturation capacity when the NA is inhibited (Fig. 3.2). This is likely to be due to two possible scenarios: 1) a population of irreversibly attached viruses; or 2) viruses making interactions that are too short for the NA to efficiently meet any of the remaining (heavily depleted) substrate. The sialic acid depletion curves displayed in Figure 3.12 tend towards a level of remaining sugar of ~10% for both 6SLN-PAA and 3SLN-PAA. At this level of sugar loading (0.06 nm) the mean virus residence times would be ~100 s for 6SLN-PAA and ~1.7 s for 3SLN-PAA (see Fig. 3.7). These residence times do not appear to allow for the hypothesis of irreversibly attached viruses. This implies that the remaining binding of the virus must be due to short-lived interactions, which do not allow the NA to cleave significant amounts of the reaming sugar.

The results from the experiments with CPNA and inhibited viral NA indicate that the initial cleavage of sugars from the surface is similar whether the NA is

attached to the virus or not. The depletion of the sugars by the soluble CPNA, if measured, would be a simple single exponential decay, as the rate of sugar cleavage would reduce proportionally with the amount of sugar still remaining on the surface. The sugar depletion by the viral NA fits much better to a double exponential function (Fig. 3.12), although the number of data points is limited, suggesting two different phases of depletion: an initial fast depletion, followed by a much slower depletion. The effect of tethering the NA to the virus is to reduce its ability to cleave sugars when the amount of sugar on the sensor has been reduced to low levels.

When looking at the results comparing the binding of X-31 with X-31 HAM (Fig. 3.14) it is apparent that the reduction in overall receptor binding to 6SLN-PAA in the absence of NA inhibitors is directly attributable to the decrease in the affinity of this HA for 6SLN-PAA (Fig. 3.1). The turning points for X-31 HAM binding to both 6SLN-PAA and 3SLN-PAA are very similar (Fig. 3.14), indicating that the reduction in 6SLN receptor binding affinity balances out the reduced ability of the NA to cleave α2,6 linked sugars.

Comparing the stalk deletion mutant X-31 NA Δ10 with X-31 (Fig. 3.15) shows significantly increased overall binding to 6SLN-PAA and 3SLN-PAA, with the binding curves diverging at longer times. This is due to the reduced ability of the viral NA to interact with substrate, reducing the effective cleavage rate, despite the NA being equally active at cleaving small substrates. This is the first time that a stalk length reduction has been directly shown to reduce NA activity in a context similar to the interactions of IAVs with the cell surface.

The techniques developed in this chapter provide a good new tool for looking at HA/NA balance in a more quantitative manner, allowing better understanding of

this highly complex process involving many aspects of cooperativity and multivalency. This developed system comes with a number of caveats. The virus concentrations used (100 pM ≈ 10^{10} virus particles/ml) is high compared to virus concentrations present upon infection. The density of these surface immobilised sugars is thought to be similar to that of a cell surface (Xiong *et al.*, 2013a); however, these receptor analogues used are also short polysaccharides compared to complex glycosylation present on cell surface glycoproteins and glycolipids. The shortness of the glycan chains of these receptor analogues could alter the binding and cleavage characteristics measured in these experiments.

4 Biophysical Characterisation of the Involvement of Neuraminidase in H7N9 Receptor Binding

4.1 Introduction

H7N9 viruses were first isolated in February 2013 and reported by the Chinese Centre for Disease Control in March 2013. These viruses are thought to be a direct zoonosis from wild birds. They have caused extensive infections within mainland China and have been responsible for a large number of human infections, 672 laboratory-confirmed cases as of the 23rd June 2015. These viruses have been found do have a high case fatality ratio of ~40% but appear to be unable to readily transmit between humans (World Health Organisation, 2015c). Owing to the antigenically novel nature of these viruses and their ability to infect humans the cell interaction characteristics and consequently transmission properties of these viruses are of considerable importance.

Previous studies have shown that H7N9 viruses bind preferentially to the avian-like receptor, α2,3-linked sialic acid but also have strong binding to α2,6-linked sialic acid. This has been shown by BLI (Xiong *et al.*, 2013b; Yang *et al.*, 2013), surface plasmon resonance (SPR) (Shi *et al.*, 2013), glycan microarray (Belser *et al.*, 2013; Yang *et al.*, 2013; Dortmans *et al.*, 2013) and other solid phase binding assays (Zhou *et al.*, 2013; Ramos *et al.*, 2013; Shi *et al.*, 2013; Dortmans *et al.*, 2013; Xu *et al.*, 2014b). It is unclear, however, whether these binding characteristics are a reason that the human-human transmission of these viruses is limited.

N9 NAs have previously been shown to contain a secondary sialic acid binding site, known as the Haemadsorption (Hb) site, which is separate from the catalytic sialidase site (for more detail see *Introduction*, section 1.3.2.3). The NA of H7N9

isolates contains all the residues thought to be required to make this secondary binding site active, involving residues: 366 – 373, 399 – 404 and 430 – 433 (N2 Numbering) (Varghese *et al.*, 1997; Uhlendorff *et al.*, 2009).

This chapter examines the role of this secondary binding site in the receptor binding activity of the H7N9 virus. Receptor analogue binding assays were carried out with the WT virus and with mutant viruses in which the Hb site was knocked out in order to investigate the effects of the Hb site on solid phase receptor binding of the virus. Recombinant NAs in the WT and mutant forms were used for binding studies and to determine enzyme kinetic parameters.

4.2 Results

4.2.1 Construction of viruses used for studies

The viruses constructed for these studies were designed to assess the effect of altering the NA from the H7N9 virus in a number of different viral contexts. There were a number of safety considerations that had to be made when designing the viruses for these studies. WT H7N9 viruses are classed as a high-risk pathogen owing to their high case fatality ratio and also their perceived pandemic potential. Therefore, all viruses made in this study, excluding the 7 + 1 reassortants of PR8 with the NA from H7N9, were all generated and handled under BSL-4 conditions. Dr. Stephen Wharton was responsible for the rescue of the viruses by reverse genetics and their propagation and preparation for receptor binding studies under high containment. I was responsible for carrying out all of the molecular biology required to prepare the plasmids for transfection at lower containment levels, making the BLI measurements,

where the instrument was moved into the BSL-4 laboratory and the samples were analysed by controlling the instrument remotely.

H7N9 viruses were constructed using plasmids for the HA and NA of A/Anhui/1/2013 (H7N9) (Anhui13) cloned into the pHW2000 vector. Plasmids were a kind gift from the laboratory of Dr. Munir Iqbal (The Pirbright Institute). The H7N9 viruses were rescued with the HA and NA from Anhui13 and the remaining 6 gene segments from PR8. Viruses were rescued with both the WT and a point mutant NA, with the substitution S367N (N2 Numbering). This substitution had previously been found to abolish the haemadsorption activity of an NA with a binding Hb site (Uhlendorff *et al.*, 2009).

As well as the H7N9 viruses a pair of H3N9 viruses were constructed with the HA from an exclusively tissue culture-propagated virus, A/Victoria/361/2011 (H3N2), using a plasmid kindly donated by Dr. Lauren Parker. This is a seasonal H3N2 virus, which does not show any measurable binding to either 6SLN-PAA or 3SLN-PAA in BLI experiments (Lauren Parker, unpublished data). These H3N9 viruses were constructed with both the WT and S367N mutant NA in a PR8 background to determine if the presence of the Hb site in the NA was capable of rescuing receptor binding for viruses containing an HA with poor affinity for sialic acids. An H7N1 virus was also constructed with the HA from Anhui13 and the remaining 7 gene segments from PR8. This virus was constructed to examine the nature of H7 binding in a virus with an NA without an active Hb site. H1N9 viruses containing the NA from Anhui13 in the WT and S367N mutant forms and the remaining 7 gene segments from PR8 were constructed under BSL-2 conditions.

All viruses were rescued by transfecting 293T cells with the eight individual gene segment containing plasmids, as described in *Materials and Methods*, section 2.2.5. The H7N9 and H7N1 viruses were further propagated in parent MDCK cells whereas the H3N9 viruses were propagated in MDCK-SIAT cells, which are required for efficient propagation of viruses with the weakly binding H3 HAs (Lin *et al.*, 2012). These viruses were propagated in large-scale quantities (~750 ml) and the virus was concentrated by a long medium-speed centrifugation, as described in *Materials and Methods* section 2.2.3.2. H1N9 viruses were propagated in hens' eggs directly from the transfection supernatant. These viruses were concentrated by centrifugation of the allantoic fluid and purified through a sucrose gradient, as described in *Materials and Methods*, section 2.2.3.1.

4.2.2 N9 protein expression and purification

In order to allow detailed biochemical analyses of the N9 binding and enzymatic characteristics both the WT and S367N mutant Anhui13 NA were expressed in insect cells using a recombinant baculovirus system. The gene encoding the Anhui13 NA ectodomain (residues 75 – 465) had been previously synthesised and cloned into an insect expression vector for use with the BaculoGold system (BD Biosystems). The BaculoGold system had been discontinued, therefore the protein was expressed using the Bac-to-Bac system (Life Technologies). The expression construct was subcloned into a suitable vector, as described in *Materials and Methods*, section 2.2.6.12.2. The expression construct consisted of residues 75 – 465 with an N terminal tag consisting of a 6 x His-tag, a human vasodilator-stimulated phosphoprotein tetramerisation domain (Kühnel *et al.*, 2004; Xu *et al.*, 2008) and a TEV cleavage site under the control of the polyhedrin promoter with a gp67 secretion

signal peptide. See *Appendix* for construct sequence. The gene was prepared in a WT form and was also mutagenised to carry the S367N substitution, as described in *Materials and Methods*, section 2.2.6.13.1. The recombinant baculoviruses for WT and the S367N mutant were generated as described in *Materials and Methods*, section 2.2.7.2. The virus generated in this way was used to infect large volumes (~3 litres) of a suspension of Sf9 insect cells, as described in *Materials and Methods*, section 2.2.7.3.

The expressed protein was concentrated to a smaller volume using hollow fibre filtration and purified using a cobalt resin column, as described in *Materials and Methods*, section 2.2.8.4. Figure 5.1A shows SDS-PAGE analysis of fractions collected from the elution of the WT protein from the cobalt column. The initial four fractions that contained protein of the elution step (as numbered on Fig. 4.1A) were discarded because contaminants were visible on the gel. The remaining fractions were pooled (fractions 5 – 14 as shown in Fig. 4.1A), concentrated and buffer exchanged. The concentrated protein was then further purified by gel filtration to ensure that any protein used for later analyses was in the tetrameric form with the purification and tetramerisation tags still attached. This step ensured that any monomeric protein or aggregated protein was removed from the sample. The process is described in *Materials and Methods*, section 2.2.8.4. Figure 4.1B shows SDS-PAGE of fractions collected from the gel filtration and Figure 4.1C shows the elution profile from the column. The major peak of elution from the column was at 65 ml, which is the appropriate size for tetrameric NA, based on previous purifications carried out. The gel filtration profile indicated that the sample was of high purity, with only a very small amount of large aggregates, at lower retention volumes, and no visible peak corresponding to monomeric NA, at higher retention volumes. Fractions were pooled

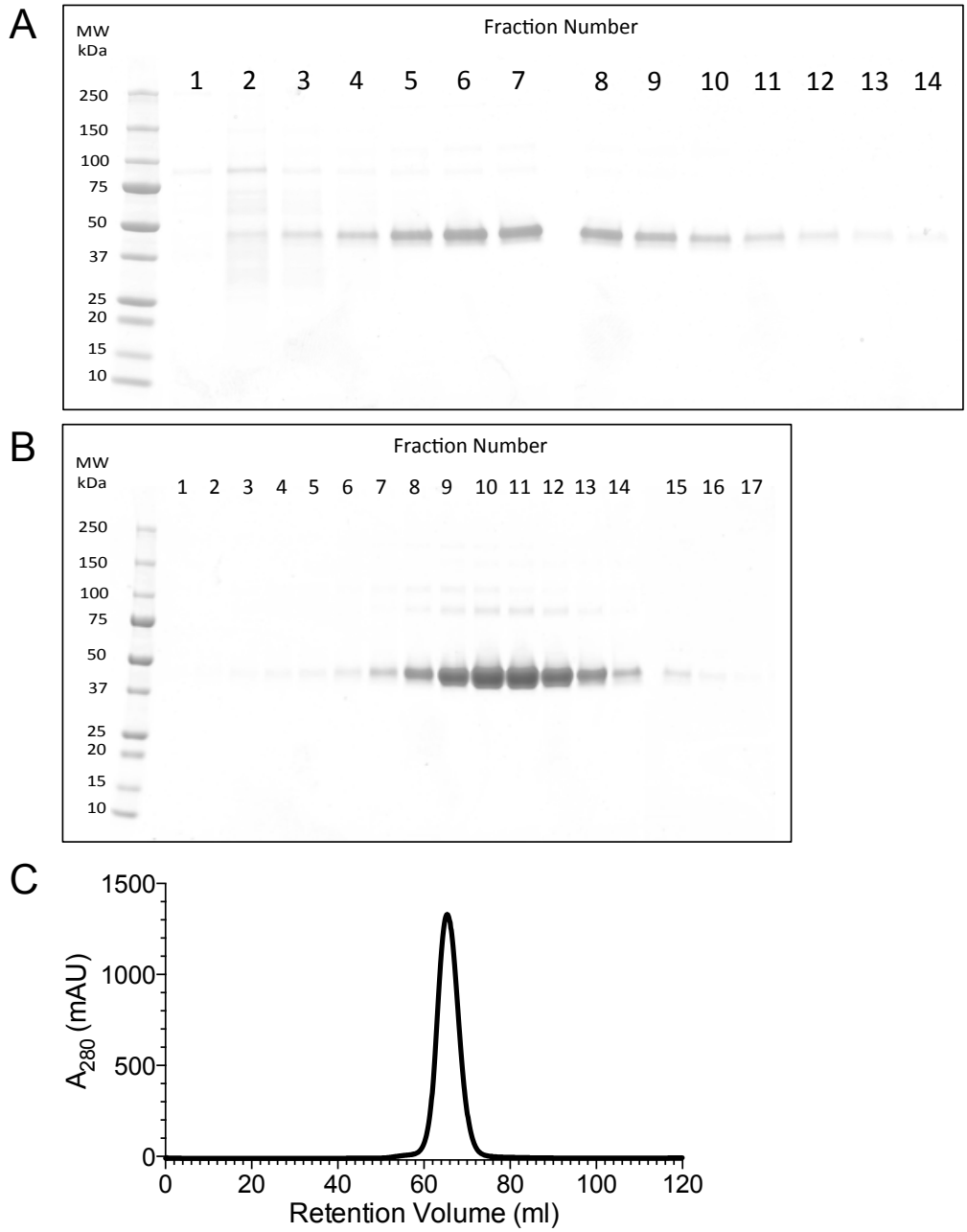


Figure 4.1: SDS-PAGE and gel filtration FPLC elution profile of N9 protein purification. A+B) SDS-PAGE analysis of fractions from FPLC protein purification of Anhui13 WT NA. Gels were run of 5 µl of elution fractions, run under non-reducing conditions. A) Protein containing elution fractions from cobalt column his-tag purification. B) Protein containing elution fractions from gel filtration purification. C) Elution profile of gel filtration purification of Anhui13 WT NA.

from the major tetrameric peak and concentrated to a smaller volume (fractions 7 – 15, as numbered in Figure 4.1B). SDS-PAGE and elution profiles shown in Figure 4.1 are of the WT protein; however, the elution profiles for the mutant S367N protein were very similar. All further experiments were carried out with the purification tag still attached to the N9 NAs in order to ensure the protein remains in a tetrameric form.

4.2.3 N9 red blood cell pull down

The NA from H7N9 viruses contains the motifs in the NA that suggest that the Hb site of this NA should bind to sialic acids. Assays were carried out to verify the presence of haemagglutination properties in the recombinant N9 and the loss of these properties upon the mutation S367N. These assays involve testing the ability of the recombinant protein attached to magnetic beads to bind to and pull down turkey red blood cells (TRBCs). This assay was developed to study the NA of some H3N2 viruses, which can use the sialidase site of the NA for receptor binding (Stephen Wharton, unpublished data).

Recombinant N9 protein was attached to His-tag binding magnetic beads. Unbound protein was washed away and beads were mixed with TRBCs and incubated to allow time for binding. The beads were washed and then mixed with TRBCs and the magnetic beads immediately pulled down with a magnet. The supernatant was removed and any remaining blood allowed to settle in a 96-well microplate, as in a standard haemagglutination assay. If the protein shows haemagglutination ability the blood will be attached to the magnetic beads and therefore there will be no blood remaining in the supernatant to settle out and *vice versa*. The procedure is described in more detail in *Materials and Methods*, section 2.2.8.5.

Assays were carried out with the WT and mutant S367N N9 proteins. Figure 4.2A shows an image of a representative experiment, showing the supernatant of the pulldown after settling in a microplate. A control experiment was also carried out using beads without protein attached, as a test for non-specific binding. This 'no protein control' could then be compared to the 'no bead control' with blood at the same concentration as that used in the pull down.

It is clear from these experiments that the WT N9 has the ability to pull down a large quantity of TRBCs compared to the no-protein control, effectively removing any visible settled blood. The mutant S367N, however, appears to not pull down any blood, leaving an amount of blood in the supernatant similar to that seen with the no protein control. This indicates that the haemagglutination activity present in the WT protein is mediated via the Hb site and that the S367N substitution efficiently removes this activity. There is no clear visual difference in the settled bloods for experiments performed in the presence and absence of the NA inhibitor oseltamivir carboxylate, which will block the sialidase site, therefore suggesting that there is no contribution from the sialidase site to blood pull down.

In order to quantitatively measure the amount of blood present in the pull down supernatant blood was lysed in 0.5% SDS, debris removed by centrifugation and the amount of haemoglobin assayed by measuring the absorbance at 540 nm. For experimental detail see *Materials and Methods*, section 2.2.8.5. Figure 4.2B shows results for the percentage of blood remaining in the pull down supernatant from three replicate experiments. The results have been normalised to the no protein control. These results give a similar pattern to that seen with the blood settled in the wells, showing a reduction of ~80% of blood present with the WT protein with

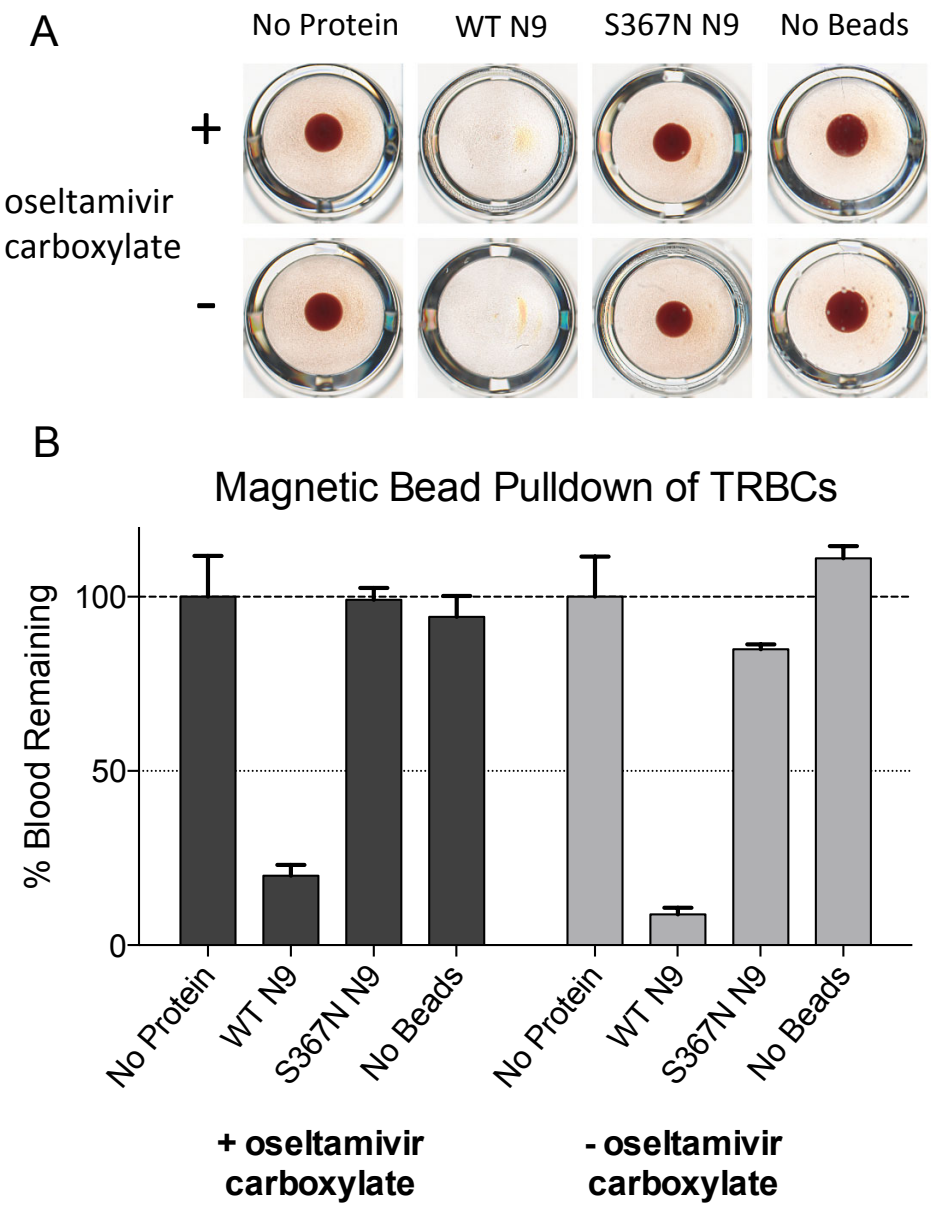


Figure 4.2: Turkey red blood cell pull down by N9 proteins. His-tagged WT and S367N N9 proteins were attached to magnetic beads. Beads were used to pull down turkey red blood cells (TRBCs) both in the presence and absence of 100 μM of the NA inhibitor oseltamivir carboxylate. No protein control consists of a pull down with beads with no protein attached and no bead control is a sample of blood of the same concentration used in the pulldown. Both controls allow a measurement of non-specific binding of TRBCs to the beads rather than the attached proteins. A) Supernatant left to settle in microtitre plate to quantify remaining TRBCs in supernatant by visual inspection. Results shown are from a single representative experiment. B) TRBCs in supernatant quantified by lysis in 0.5% SDS and measurement of A_{540nm}. Converted to % remaining blood by normalising to no protein control. Values shown are for a mean of 3 independent measurements with error bars indicating standard deviation from the mean.

oseltamivir carboxylate present, with the S367N showing no significant reduction in blood pull down compared to both the no protein and no bead controls. There are differences between the experiments carried out in the presence and absence of oseltamivir. For both the WT and S367N proteins there is more pull down of the TRBCs when oseltamivir carboxylate is absent, giving a reduction of remaining blood cells by a further 10 – 15% compared to when oseltamivir carboxylate is present. These observations are counterintuitive, as these experiments are carried out at room temperature where the sialidase activity would be thought to reduce the overall binding to TRBCs. This suggests that the sialidase site could be involved in enhancing the binding of the NA to the TRBCs under the conditions of these experiments, which could not be seen with the settled blood (Fig. 4.2A).

4.2.4 Removal of HA1 from H1N9 viruses

H1N9 viruses were constructed with the HA of PR8 and the NA of Anhui13, as described in section 4.2.1. These viruses were constructed in order to attempt to remove the HA1 by proteolysis whilst leaving the NA intact and unaffected. Since this virus would then only have NA remaining, with an inactive remnant of HA, which is incapable of receptor binding, it could be used to see if the NA was capable of acting as a receptor binding protein in a multivalent situation. In principle, such an interaction might be measurable by BLI and this would allow a measure of the relative contribution of the NA to the overall receptor binding affinity.

HA was cleaved from the surface by first exposing the virus to low pH (4.95 – 5.05) in order to induce the conformational change in the HA which makes the HA1 susceptible to protease cleavage. The pH of the solution was then neutralised and the virus treated with trypsin. Protease digestion was inhibited and the remaining virus

pelleted out of solution, the supernatant was discarded and the pellet resuspended, using a technique similar to that used in previous studies (Skehel *et al.*, 1982; Ruigrok *et al.*, 1986), see *Materials and Methods*, section 2.2.12, for experimental details. Figure 4.3 shows SDS-PAGE of viruses before and after the low pH treatment. This gel shows that under non-reducing conditions the HA band is removed by the trypsin digestion. There are also bands at lower molecular weights which are likely to be parts of the HA which are still attached to the virus, as seen in (Skehel *et al.*, 1982). The gel run under reducing conditions shows a loss of the HA1 band upon protease digestion, while the HA2 band remains unchanged by the treatment. The density of the NA band also appears unchanged by the protease treatment.

The NA activity of the virus was checked by measuring the cleavage of the soluble NA substrate MUNANA. The cleavage was found to be unaffected by the low pH treatment of the virus. Figure 4.4 shows the rate of MUNANA cleavage measured for a small quantity of the native virus, virus after low pH treatment and digestion and also the resuspended pellet and the supernatant from the subsequent ultracentrifugation. For both the WT and S367N H1N9 viruses there is no reduction in MUNANA cleavage after the low pH treatment and digestion, indicating that the NA is unaffected by the procedure. The supernatant from ultracentrifugation has very low MUNANA cleavage activity, indicating that the NA is still attached to the virus, in the pellet. There is loss of NA activity of ~20% when looking at the resuspended pellet; however, this is likely to be due to incomplete pellet resuspension or virus aggregation, rather than a real loss of NA activity.

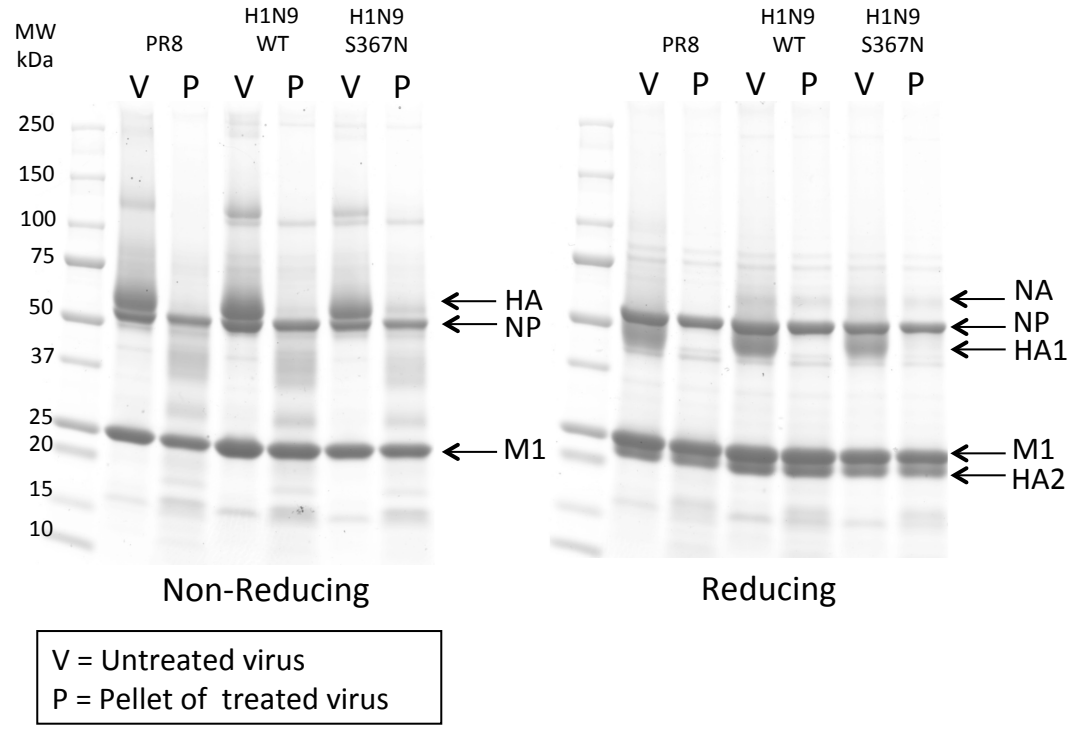


Figure 4.3: SDS-PAGE of low pH and trypsin treated H1N9 virus to remove HA1. Viruses were treated with low pH (4.95 – 5.05) to induce the HA conformational change before pH neutralisation and trypsin digestion to remove HA1; virus was then pelleted by ultracentrifugation. A small quantity of original untreated virus (V) and treated virus pellet (P) were analysed by SDS-PAGE. PR8 and H1N9 viruses with WT and S367N mutant NA are shown under both reducing and non-reducing conditions.

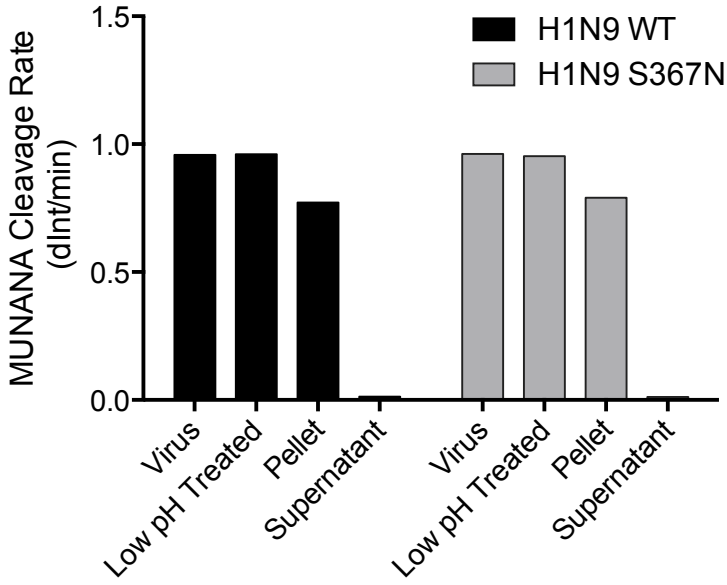


Figure 4.4: Residual NA activity after low pH treatment. Measurements made of untreated virus (virus), virus after low pH treatment and trypsin digestion (low pH treated) and virus pellet and supernatant after ultracentrifugation of the low pH treated virus. Rate of MUNANA cleavage measured for a small volume of each sample with a set concentration of 50 μM of MUNANA. The value shown is rate of change of fluorescence intensity (AU).

	HA Titre (log_2(Titre))		**Elution** (log_2(Titre Difference))	
Virus	**Native Virus**	**Low pH Virus***	**Native Virus**	**Low pH Virus**
PR8	17	6	- §	-
H1N9 WT	17	6	↑† 7	↓‡ 1
H1N9 S367N	17	6	↑ 5	↓ 1

* Low pH treated and trypsin digested virus prior to ultracentrifugation
§ - denotes no visible elution of virus from TRBCs.
† ↑ denotes elution from the highest virus concentration
‡ ↓ denotes elution from the lowest virus concentration

Table 4.1: HA titres and TRBC elution properties of H1N9 viruses before and after low pH treatment.

The HA titre of the viruses post treatment was measured. Table 4.1 shows the HA titres of viruses binding to TRBCs. The binding of the native virus is high (HA titre of 2^{17}) owing to the high concentration of virus used in the low pH treatment (~10 nM). Upon the low pH treatment and trypsin digestion the HA titre drops by 11 wells, a reduction of 2048-fold. The residual binding detected (HA titre = 2^6) is present for the PR8 control as well as the H1N9 viruses, showing that the residual binding is the same for both the binding (N9 WT) and non-binding (PR8 and N9 S367N) NAs, indicating that it is most likely from residual HA or non-specific binding rather than NA binding. There could be a difference in HA titre with the WT and S367N NA; however, this may not be seen, as the HA assay is relatively insensitive, with haemagglutination titres being separated by 2-fold dilutions of virus.

Measurement of the binding of the virus with HA1 removed was attempted using BLI. There was very little binding signal to sensors saturated with either 6SLN-PAA or 3SLN-PAA. Binding studies were performed with virus at 1 x (100 pM), 10 x (1 nM) and 50 x (5 nM) virus concentration. The signal increased at higher concentrations; however, the binding measured to the sialoglycopolymers was not much larger than that of non-specific binding to an unloaded sensor. The HA1 removal procedure was repeated and the same results were obtained (data not shown).

This lack of binding signal above background non-specific binding indicates that if the NA does act in receptor binding it is not of adequate binding strength to allow significant virus binding in the absence of HA binding. This could be due to there being only low levels of NA present on the viral surface, as NA is the minor glycoprotein constituent of the viral membrane. The Hb site of the NA could also have

a very low affinity for receptor analogues. It was noted that the low pH treatment increased the non-specific binding of the virus to unloaded sensors. The low pH and trypsin treated HA would likely be in a conformation similar to a structure obtained of the post-fusion form of HA, EHA2 (see *Introduction*, Figure 1.5), which could have altered binding properties to the biosensor surface. There was also a similar level of binding measured with full 8-segment PR8, included as a negative control, indicating that the binding seen was either from residual uncleaved HA1 or non-specific interactions with the sensor. Attempts were made to try and reduce the non-specific binding by adding BSA and increasing the amount of Tween-20 present in the buffer; however, neither of these approaches improved the data (not shown). The H1N9 viruses were also rescued with a WSN background, which has previously been found to increase the amount of NA present in a virus, in order to increase the NA binding signal above that of non-specific binding. These viruses also did not yield significantly improved data (not shown).

4.2.5 Elution of H1N9 viruses from TRBCs

The ability of the H1N9 viruses to cleave sialic acid in both native and low pH/protease digested forms was assessed by measuring the ability of the viruses to elute from TRBCs. Elution was carried out by covering the microtitre plate and placing it in a 37°C incubator for ~18 hrs. The difference in HA titre from the values measured previously was noted. The results for these assays are shown in Table 4.1. The PR8 control virus did not show any elution from TRBCs either in the native or low pH treated forms, which is surprising, as PR8 NA would be thought to be highly active due to its efficient propagation in eggs. The low pH and trypsin digested H1N9 viruses both showed a 1 well reduction in HA titre from the highest dilution that

previously showed haemagglutination activity. The native viruses, however, eluted from the lowest virus dilutions (highest virus concentration). The WT NA virus eluted 7 wells from the lowest dilution to the 2^7 virus dilution whereas the S367N virus eluted 5 wells, a 4-fold difference in the concentration of virus required to elute from the TRBCs. This suggests that the loss of the Hb site reduces the ability of the NA to cleave the sialic acid from red blood cells.

4.2.6 Biolayer interferometry receptor binding studies

A number of experiments were carried out to quantitatively assess the impact of the presence of sialic acid binding capabilities via the Hb site. These involve measuring equilibrium virus binding as a function of sugar loading in the presence of NA inhibitors as well as measurements to determine the balance between the activities of HA and NA in altering virus receptor binding.

4.2.6.1 H7N9 and H1N9 BLI equilibrium measurements

The equilibrium receptor binding of the various constructed viruses was measured using BLI in the presence of NA inhibitors, as described in *Materials and Methods*, section 2.2.10.1. These measurements were made on a pair of viruses with Anhui13 HA and NA (H7N9) and also viruses with PR8 HA and Anhui13 NA (H1N9) as well as full 8 segment PR8 as a control. Both sets of viruses were made with the NA both in WT and S367N mutant forms in order to determine if there was any contribution of the secondary binding site to equilibrium virus binding in the presence of NA inhibitors used to block the sialidase site.

Figure 4.5 shows the results for the binding assays of the two H7N9 viruses to the receptor analogues 6SLN-PAA and 3SLN-PAA. The binding of the WT virus is

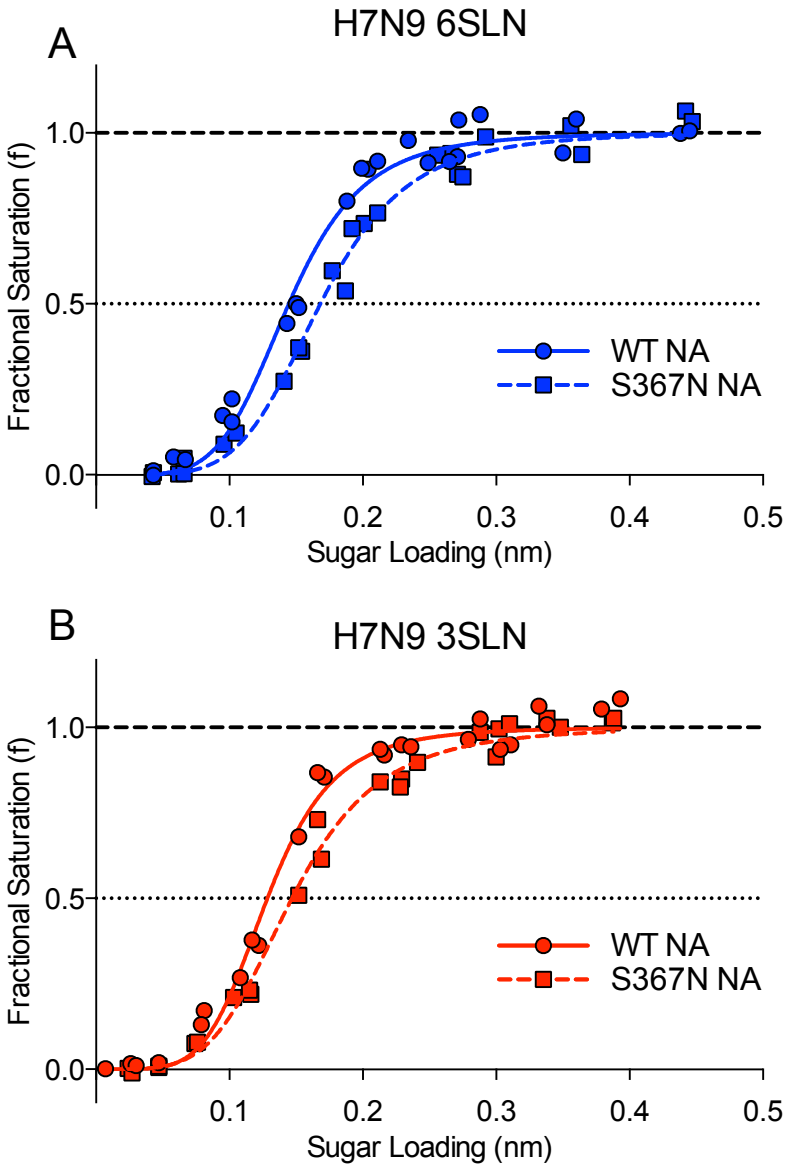


Figure 4.5: H7N9 virus binding to receptor analogues. Measurements made in the presence of NA inhibitors. 100 pM virus binding to A) 6SLN-PAA and B) 3SLN-PAA. Binding shown of virus with WT NA and Hb site mutant S367N.

stronger for the avian-like receptor analogue 3SLN-PAA with slightly weaker binding to the human-like receptor analogue 6SLN-PAA. These results are similar to those seen when the binding of the original egg-propagated clinical isolate virus was measured by BLI (Xiong *et al.*, 2013b). It is clear from these experiments that there is a slight decrease in receptor binding to both 6SLN-PAA and 3SLN-PAA when the NA is mutated (S367N). This drop is an increase in relative K_d of ~2-fold.

Figure 4.6 shows the same measurements for H1N9 binding. This PR8 HA used has a receptor binding preference for the human-like receptor analogue 6SLN-PAA with weaker binding to the avian-like receptor analogue 3SLN-PAA. The mutation of the N9 from WT to S367N gives a similar reduction in receptor binding for both 6SLN-PAA and 3SLN-PAA as seen for the H7N9 viruses (Fig. 4.5). It has a similar effect on the relative K_d increasing it ~2-fold. The loss of the Hb activity makes the virus binding, to both 6SLN-PAA and 3SLN-PAA, indistinguishable from that of the full 8 segment PR8, which does not have Hb properties (Fig. 4.6).

The results for the H7N9 and H1N9 viruses showed that there is a consistent loss of virus receptor binding when any binding through the Hb site is eliminated by mutagenesis. This indicates that the NA of H7N9 viruses can have an effect on altering equilibrium receptor binding of virus to receptor analogues, indicating a role of the secondary binding site of NA in strengthening the receptor binding affinity of the virus as a whole in the absence of sialidase activity. This increase in binding does not appear to have an α2,6-/α2,3-linked sialic acid preference with the alterations in binding strength similar for binding of the H7N9 and H1N9 viruses to 6SLN-PAA and 3SLN-PAA when the S367N substitution was introduced.

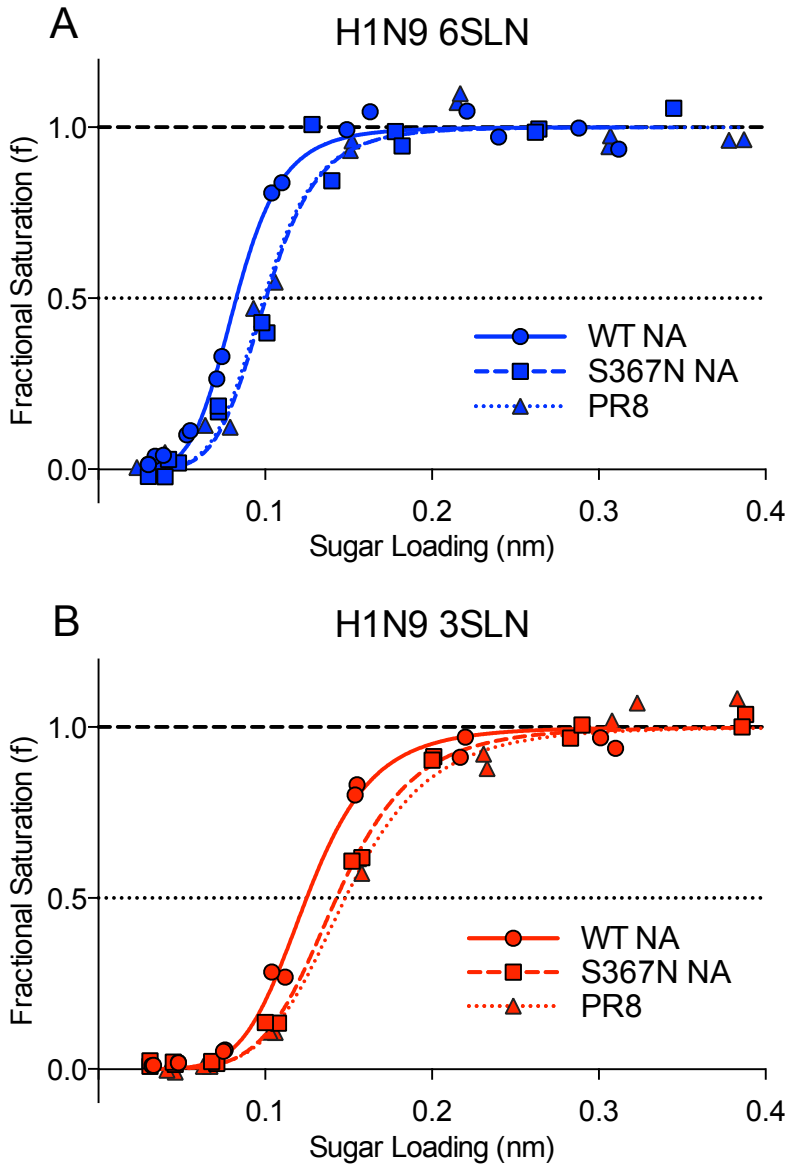


Figure 4.6: H1N9 virus binding to receptor analogues. Measurements made in the presence of NA inhibitors. 100 pM virus binding to A) 6SLN-PAA and B) 3SLN-PAA. Binding shown of virus with WT NA and Hb site mutant S367N. Binding of full 8-segment PR8 shown for comparison.

4.2.6.2 H3N9 Receptor Binding

Viruses were generated with the HA of tissue culture-propagated A/Victoria/361/2012 (H3N2) with both the WT and S367N N9 NA. Viruses containing this HA have no detectable binding to receptor analogues in BLI experiments (Lauren Parker, unpublished data). These viruses were constructed to determine if the Hb site of the N9 would be capable of restoring binding capacity to a previously unmeasurable binding interaction. Binding to sensors saturated with a range of different sialoglycopolymers was measured, in the presence of NA inhibitors, at the usual virus concentration of 100 pM and at 10 x concentration (1nM) and no binding was measurable for either WT or S367N. Virus concentration was increased to 50 x concentration (5 nM) and a small amount of binding was measurable but only to 6SLN-PAA, shown in Figure 4.7. The binding of the virus with WT NA to 6SLN-PAA is very weak with a saturation value of binding of ~0.8 nm. This is very weak compared to the binding measured of H7N9 viruses, which saturate the sensor at ~3.5 nm at these saturated sugar loading levels. This binding is reduced when the S367N mutation is introduced into the NA to a saturation level of ~0.25 nm. No significant binding was visible to any of the other receptor analogues tested (Fig. 4.7).

These data for H3N9 binding again highlight the fact that the presence of the Hb site in the N9 NA can enhance virus receptor binding. The presence of the Hb site in the WT N9 appears to allow binding of this virus which otherwise would be much weaker. The virus has a receptor preference for 6SLN-PAA; however, this is not necessarily due to a receptor specificity of the Hb site, rather that the H3 has a preference for binding to 6SLN-PAA over the other analogues, which is below the threshold of measurement, and the inclusion of the N9 with the Hb site increases

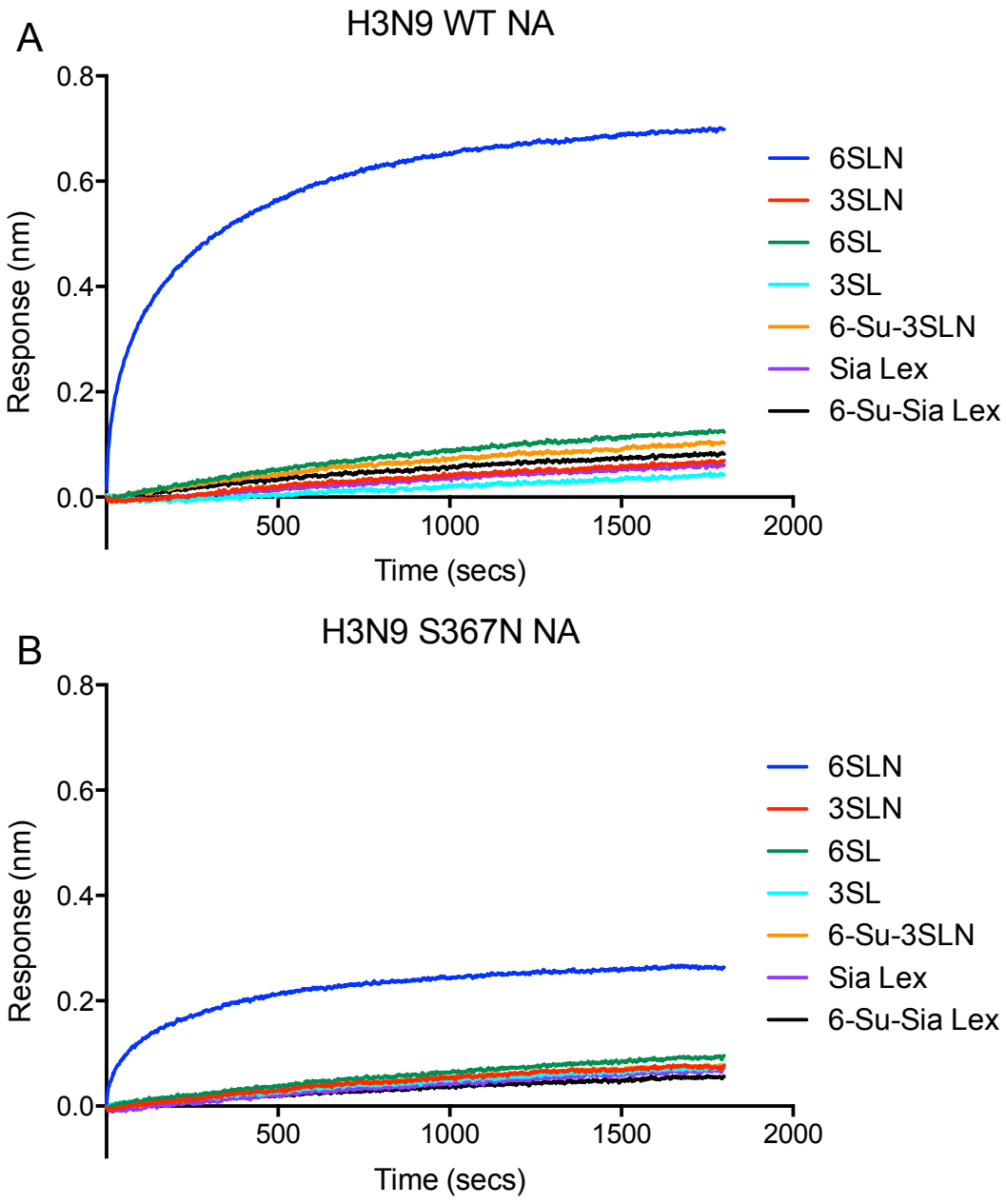


Figure 4.7: BLI data of H3N9 virus binding to sialoglycopolymers. Viruses consist of the low binding H3N2 HA Vic361 coupled with the N9 NA in WT and S367N forms in a PR8 background. Binding measured in the presence of NA inhibitors at 50 x typical virus concentration (5 nM). Binding measured to sensors saturated with sialoglycopolymers bearing the receptor analogues: 6SLN, α2,6-Sialyl-N-acetyllactosamine; 3SLN, α2,3-Sialyl-N-acetyllactosamine; 6SL, α2,6-Sialyllactose; 3SL, α2,3-Sialyllactose; 6'-Su-3SLN, 3SLN 6' sulphated on GlcNAc; SiaLe^x, Sialyl-Lewis X; 6-Su-SiaLe^x, SiaLe^x 6' sulphated on GlcNAc. Response is not normalised as proportion of saturation value.

overall sialic acid binding and allows the virus to pass over this threshold of binding, with the receptor specificity still determined by the HA. This binding was still very weak with a virus concentration of 5 nM necessary in order to gain measurable receptor binding and the overall response of binding very low. These results correspond well with the lack of measurable binding of low pH treated H1N9 in BLI experiments, see section 4.2.4.

4.2.6.3 Measurement of H7N9 and H1N9 HA/NA balance

The previous BLI experiments (sections 4.2.6.1 and 4.2.6.2) explored the binding of virus in the presence of NA inhibitors. The role of the secondary binding site is thought to be more complex than simply enhancing virus receptor binding, possibly also being responsible for an increase in NA activity due to tethering of multivalent substrates, allowing more efficient cleavage by the sialidase site (Uhlendorff *et al.*, 2009) (see *Introduction*, section 1.3.2.3**,** for further details). In order to understand the role of the NA in controlling receptor binding in a more dynamic system, experiments were carried out to measure the role of HA and NA in receptor binding, using assays previously developed (described in Chapter 3) where BLI experiments are carried out to measure the binding of virus to sensors saturated with sialoglycopolymer receptor analogues both in the presence and absence of NA inhibitors. Figures 4.8 and 4.9 show HA/NA balance measurements for the binding of 100 pM H7N9 viruses to 6SLN-PAA and 3SLN-PAA coated surfaces. Measurements were made for viruses with the WT NA and with the S367N haemadsorption mutation. The binding curve for 3SLN-PAA is similar for both the WT and S367N NA with very low overall levels of virus binding with the NA uninhibited, before the

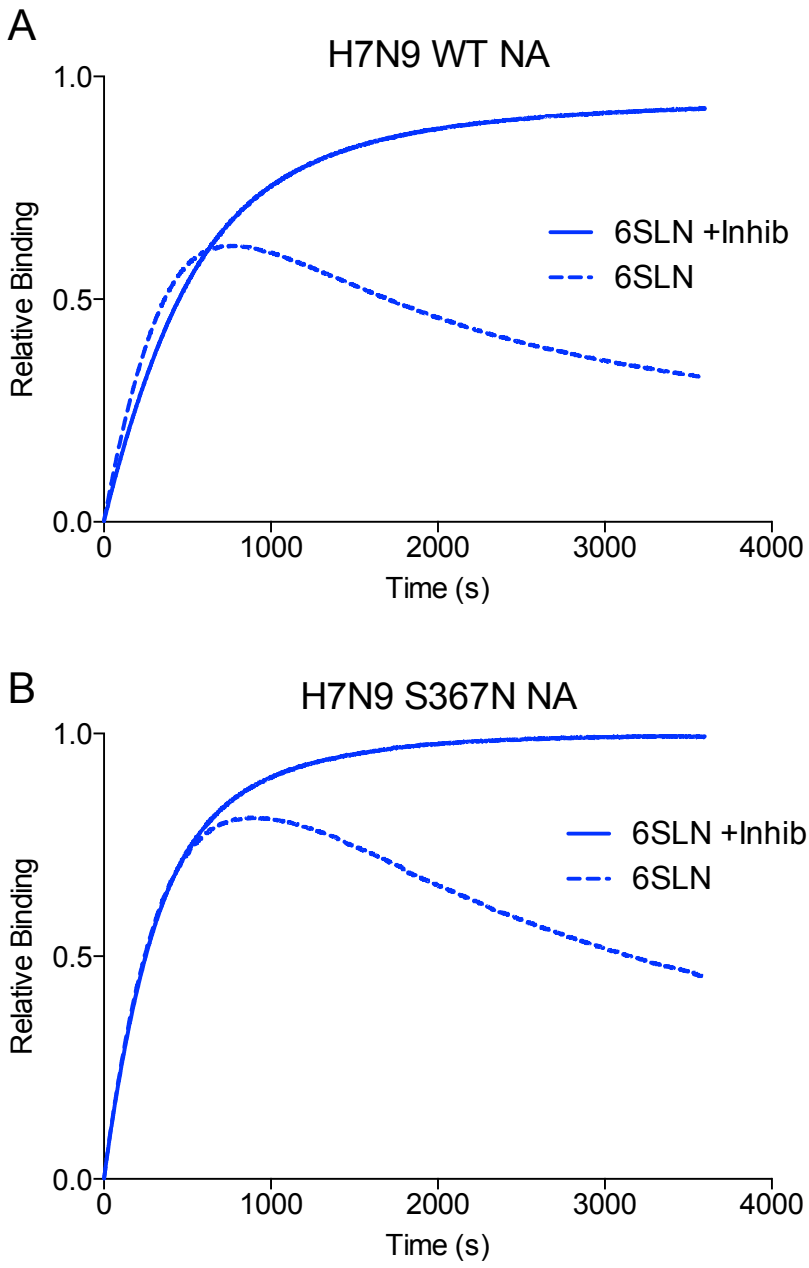


Figure 4.8: HA/NA balance measurements of H7N9 virus binding to 6SLN-PAA. 100 pM virus binding measured to sialoglycopolymers bearing receptor analogue 6SLN in the presence and absence of NA inhibitors (Inhib). Binding measured of virus with A) WT NA and B) S367N NA.

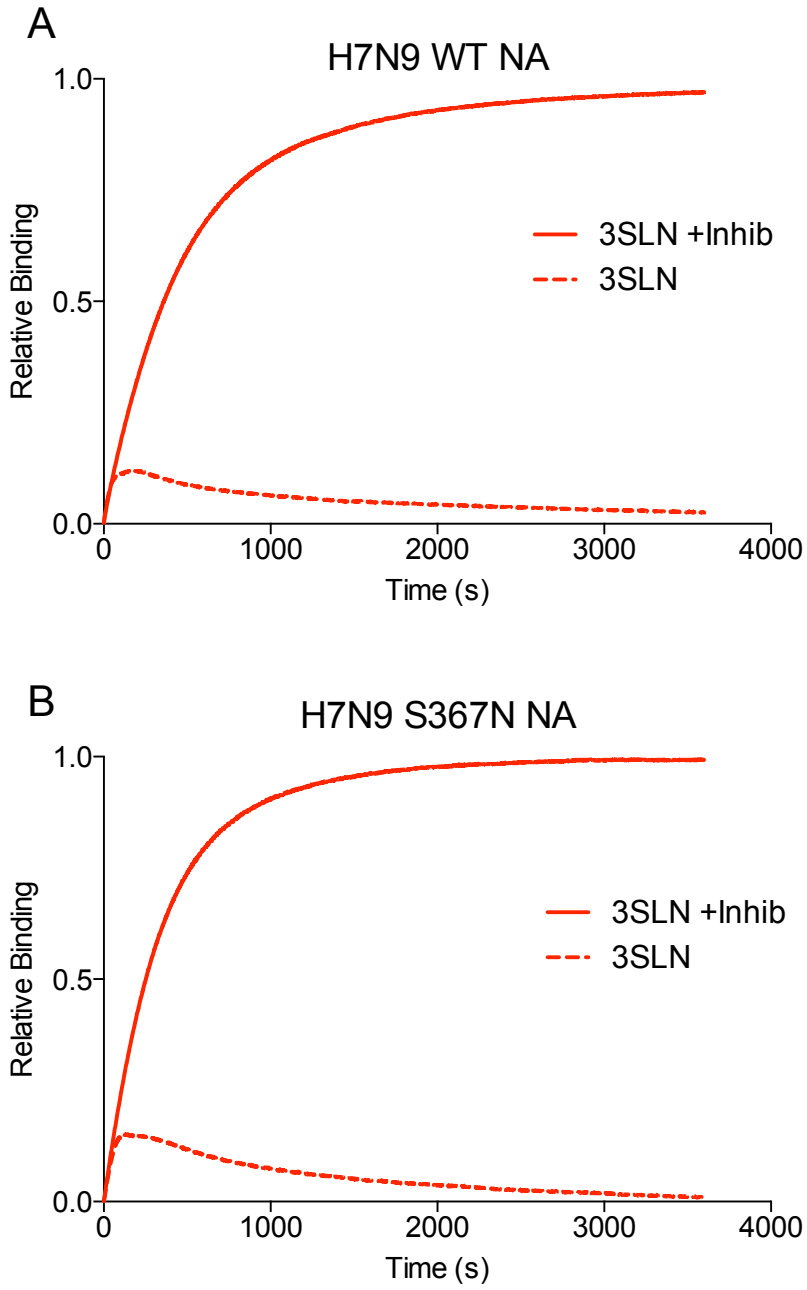


Figure 4.9: HA/NA balance measurements of H7N9 virus binding to 3SLN-PAA. 100 pM virus binding measured to sialoglycopolymers bearing receptor analogue 3SLN in the presence and absence of NA inhibitors (Inhib). Binding measured of virus with A) WT NA and B) S367N NA.

curves diverge at an early time point and then decrease to very low binding levels at later time points. These results are similar to those seen of X-31 binding to 3SLN-PAA (Fig. 3.2). Binding of the H7N9 viruses to 6SLN-PAA coated surfaces show an interesting effect with the WT NA, where there is an initial enhancement of receptor binding in the absence of NA inhibitors. This enhancement of receptor binding is lost when the S367N mutation is introduced. This increase of binding with the WT NA implies that there is cooperativity between the sialidase site, which is blocked by the NA inhibitors, and the active haemadsorption site, as the increase in binding is only seen when the sialidase site is uninhibited and the Hb site present. This enhancement is not seen when the viruses are binding to 3SLN-PAA (Fig. 4.9). The overall balance between HA and NA activities of these viruses must, therefore, be considered in greater detail in order to understand these effects.

The NA of the H7N9 appears to be efficient at cleaving the α2,3 linked 3SLN-PAA, as the virus does not show a large binding signal in the absence of NA inhibitors (Fig. 4.9), as seen for the same experiment with X-31 (Fig. 3.2). The activity against the α2,6 linked sialic acid seems low, as there is extensive virus binding in the absence of NA inhibitors (Fig. 4.8), with a similar level of virus binding to both 6SLN-PAA and 3SLN-PAA when measured in equilibrium experiments with NA inhibitors present (Fig. 4.5). This poor enzymatic cleavage of 6SLN by the N9 NA could be responsible for the enhanced receptor binding seen with the H7N9 WT NA binding to 6SLN-PAA (Fig. 4.8), as an enzyme with an enhanced substrate binding capacity (low K_m) and/or low catalytic rate (k_{cat}) can have a higher proportion of stronger binding events which do not result in substrate cleavage, turning the NA into more of a substrate binding protein. This substrate binding by the sialidase site could then act

cooperatively with the haemadsorption site to enhance receptor binding to give the results as seen in Figure 4.8.

As well as the phenomenon of enhanced initial receptor binding to 6SLN-PAA there also appears to some increase in the ability of the NA to cleave 6SLN when the haemadsorption site is present, as the peak of maximal virus binding in the absence of NA inhibitors is lower with the WT NA than with the virus carrying the S367N substitution. This could be due to the ability of the secondary binding site to enhance NA activity against large substrates.

Figures 4.10 and 4.11 show the same HA/NA balance experiments carried out with the H1N9 viruses, which consist of a 7 + 1 reassortment with PR8. The curves show similar patterns to those of the H7N9 viruses. The H1N9 viruses with the two different NAs bind similarly to 3SLN-PAA (Fig. 4.11), with low overall virus binding decaying to very low levels at later time points. The binding to 6SLN-PAA (Fig. 4.10) shows the similar initial enhancement of binding when the Hb site is present and NA inhibitors are absent. This enhancement of binding is then lost with the virus carrying the S367N substitution in the NA. When comparing the enhancement of binding to 6SLN-PAA of H7N9 (Fig. 4.8) and H1N9 (Fig. 4.10) it is apparent that there is a smaller enhancement of binding for the H1N9. This may be due to the fact that the HA of PR8 has a higher affinity for 6SLN-PAA (Fig 4.6) than the HA of H7 (Fig 4.5), with a K_d for virus binding ~12-fold lower. This would mean that NA binding via the Hb and sialidase sites is making a smaller proportional contribution to the overall binding strength, making the enhancement appear smaller. It is not apparent from these results for the H1N9 viruses that the N9 has an enhanced cleavage rate when the Hb site is present as can be seen by the greater rate of virus dissociation at later time

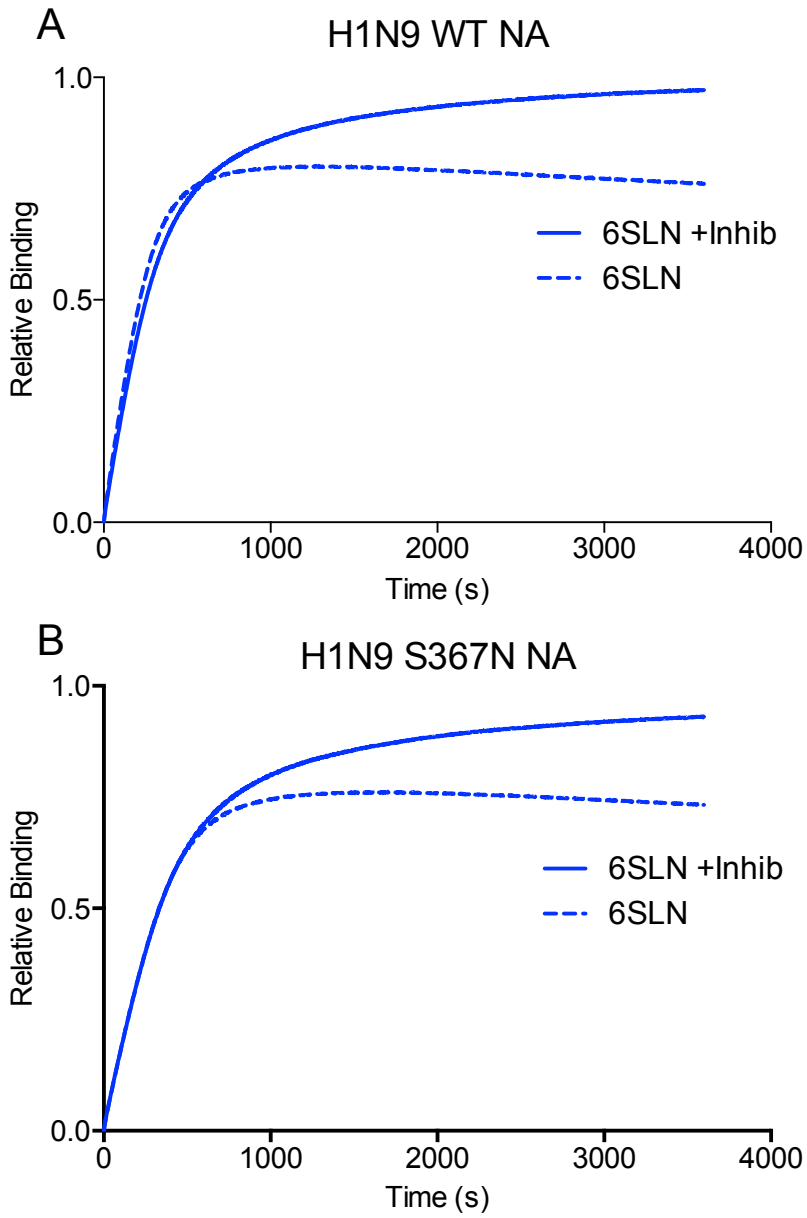


Figure 4.10: HA/NA balance measurements of H1N9 virus binding to 6SLN-PAA. 100 pM virus binding measured to sialoglycopolymers bearing receptor analogue 6SLN in the presence and absence of NA inhibitors (Inhib). Binding measured of virus with A) WT NA and B) S367N NA.

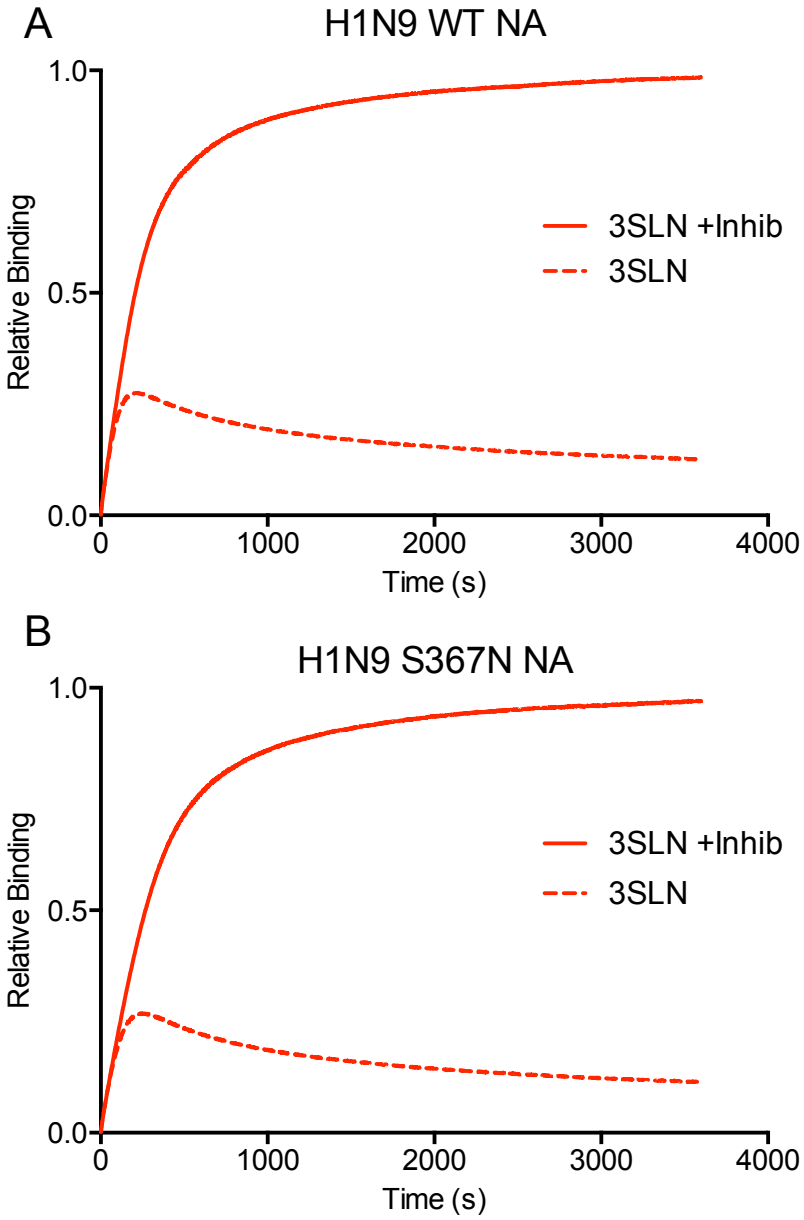


Figure 4.11: HA/NA balance measurements of H1N9 virus binding to 3SLN-PAA. 100 pM virus binding measured to sialoglycopolymers bearing receptor analogue 3SLN in the presence and absence of NA inhibitors (Inhib). Binding measured of virus with A) WT NA and B) S367N NA.

points seen with the H7N9 viruses (Fig. 4.8). This could be due to the greater affinity of the virus for 6SLN-PAA (Fig. 4.6), making the enhanced NA activity less noticeable.

4.2.7 Measurement of N9 receptor analogue affinity by microscale thermophoresis

The Hb site of the NA has been shown in BLI experiments to enhance virus receptor binding (Figs. 4.5 and 4.6). Experiments were carried out to attempt to measure a binding constant of monomeric receptor for this second binding site. The affinity of HA for monomeric receptor analogues has recently been determined using a new technology, microscale thermophoresis (MST) (Xiong *et al.*, 2013a). This technique works by measuring the diffusion of a fluorescently labelled analyte at a constant concentration in a temperature gradient in the presence of different concentrations of a binding partner. This technique has been used for HA binding, as the sample volume required is very low, consequently the weak binding of HA to receptor analogues, K_d in mM range, can be measured without the need to use excessive quantities of the monomeric receptor analogues. A similar protocol was carried out to attempt to measure the binding of N9 to receptor analogues via the Hb site. Baculovirus expressed recombinant WT and S367N N9, expressed as described in section 4.2.2, was fluorescently labelled with a reactive amine linked dye, as described in *Materials and Methods*, section 2.2.13. The labeling procedure was optimised to allow a low level of loading, ~ 1 fluorophore per NA tetramer.

These labelled proteins were analysed for binding to both 6SLN and 3SLN using 2-fold serial dilutions of receptor analogues from a starting concentration of 150 mM in the presence of 100 μM oseltamivir carboxylate. The assay buffer used was altered from HBS-P + $CaCl_2$ to include 50 mM HEPES-NaOH pH 7.4. The increased

concentration of HEPES was included in order to counteract the pH perturbation caused by the high concentration of sialic acid present. Thermophoresis was measured at a number of different strengths of temperature gradients by altering the instrument laser power. Unfortunately, there was no measurable difference in thermophoresis at the different sugar concentrations used (data not shown).

4.2.8 N9 enzyme kinetics

HA/NA balance data presented above (section 4.2.6.3) suggest that the binding of H7N9 and H1N9 viruses bearing the H7N9 NA have enhanced initial binding to α2,6-linked sialic acid when the sialidase site of the NA is uninhibited. This suggests a possible role for the sialidase site in receptor binding. Characterising the kinetics of substrate cleavage of the N9 is consequently of interest. Experiments are described below which characterise the cleavage of the soluble monomeric substrates MUNANA, 6SLN and 3SLN. There are also suggestions in HA/NA balance experiments above (Fig. 4.8), TRBC elution results (Table 4.1) and also in previous reports (Uhlendorff *et al.*, 2009) that having the Hb site present enhances the ability of the NA to cleave large sialic acid containing substrates such as the heavily glycosylated protein fetuin or sugar coated surfaces as in BLI experiments. The kinetics of fetuin cleavage by the N9 proteins have therefore also been characterised. Recombinant N9s were used for these kinetic measurements. The expression and purification of these proteins has been described in section 4.2.2.

4.2.8.1 MUNANA kinetics

The kinetics of cleavage of MUNANA by WT and S367N N9 NA was characterised by fluorimetry, correcting fluorescence data for the different buffers as

previously described, in section 3.2.7.1 and *Materials and Methods*, section 2.2.11.1.1. Measurements were made with both WT and S367N NA (at a monomeric concentration of 0.25 nM) using the MUNANA substrate at concentrations in the range 10 – 175 μM in both MES and HBS-P + $CaCl_2$ buffers. Michaelis-Menten plots for the WT and S367N proteins are shown in Figure 4.12 and the fitted parameters are shown in Table 4.2.

The WT and S367N N9 proteins have very similar K_m and k_{cat} values for measurements made with MUNANA in both buffers. This indicates that the S367N mutation does not affect the sialidase activity of the NA for small monomeric substrates. The WT NA has a K_m value of 38.5 ± 2.1 μM in MES buffer, with a K_m ~ 4-fold higher in the HBS-P buffer (137.9 ± 12.8 μM). This indicates that the NA binds to substrate better at a lower pH, although this is not the only alteration in the two buffers, as there is also a difference in ionic strength. These K_m values are similar to those previously reported for similar expressed N9 protein (Wu *et al.*, 2013). Under both buffer conditions the enzyme has a similar turnover rate with a k_{cat} value of ~90 s^{-1}. Comparing these kinetic parameters with those obtained for X-31 NA indicate that this N9 NA is less efficient, having a higher K_m and a lower k_{cat}.

4.2.8.2 6SLN and 3SLN kinetics

The kinetic parameters for MUNANA cleavage determined above are useful in confirming that the WT and S367N proteins have identical sialidase activities for cleaving small substrate. These values, however, do not allow understanding of the role of linkage specificity for the NA. Experiments were carried out based on the system previously described in Chapter 3 (section 3.2.7.2) using the coupled enzymes

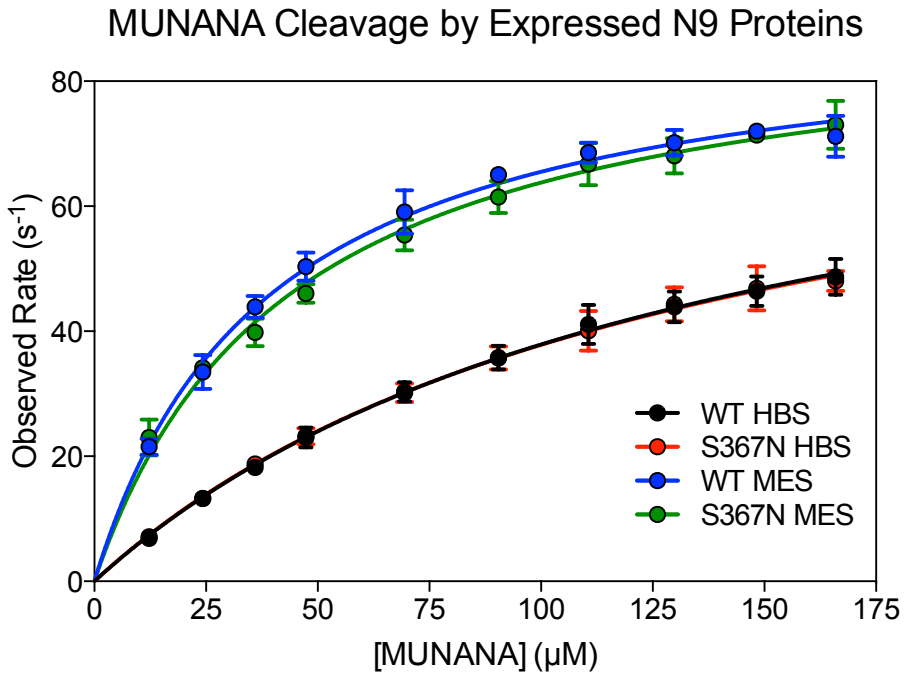


Figure 4.12: Michaelis-Menten plots of MUNANA cleavage by WT and S367N N9 NA. Measurements made at 37°C in either HBS (HBS-P + $CaCl_2$: 10 mM HEPES pH 7.4, 150 mM NaCl, 0.005% Tween-20, 4 mM $CaCl_2$) or MES (32.5mM MES pH 6.5, 4 mM $CaCl_2$) buffers. Each point is the mean of three replicates with error bars showing standard deviation from the mean. Least squares fit of Michaelis-Menten equation is shown. Raw reaction rate data are converted to an observed rate by dividing rates by NA concentration used in assay.

Substrate	Protein	Buffer/Temp	K_m ± S.E.[§] (µM)	k_{cat} ± S.E. (s^{-1})	k_{cat}/K_m ($\mu M^{-1}s^{-1}$)
MUNANA	WT N9	MES* 37˚C	38.5 ± 2.1	90.7 ± 1.6	2.36
		HBS† 37˚C	137.9 ± 12.8	90.1 ± 4.7	0.654
	S367N N9	MES 37˚C	43.4 ± 2.9	91.5 ± 2.1	2.11
		HBS 37˚C	134.2 ± 12.9	88.6 ± 4.7	0.660
6SLN	WT N9	HBS 37˚C	4179 ± 615	2.2 ± 0.2	0.000537
3SLN		HBS 37˚C	810.0 ± 100	159.5 ± 8.5	0.197
Fetuin	WT N9	HBS 37˚C	197.5 ± 37.3	71.6 ± 8.3	0.363
	S367N N9	HBS 37˚C	346.4 ± 95.6	73.0 ± 14.5	0.211

*MES MES Buffer (32.5 mM MES-NaOH, pH 6.5, 4 mM $CaCl_2$)
†HBS HBS-P + $CaCl_2$ Buffer (10 mM HEPES-NaOH, pH 7.4, 150 mM NaCl, 0.005% Tween-20, 4 mM $CaCl_2$)
§S.E. Standard Error of the mean

Table 4.2: Steady state kinetic parameters determined for N9 WT and S367N NA.

β-galactosidase and β-galactose dehydrogenase, using the production of NADH as a reporter for the NA cleavage rate of the monomeric sugars 6SLN and 3SLN (reaction scheme shown in Figure 3.9), experimental details described in *Materials and Methods*, section 2.2.11.2. It was decided that the MUNANA cleavage kinetic parameters determined were adequate evidence for the fact that the activity of the NA was unaltered by the S367N substitution and that this was also likely to be the case for the monomeric sugars (6SLN and 3SLN). The experiments detailed below were therefore only performed with the WT protein in order to limit the amount of sugar consumed in the experiments.

Figure 4.13A shows the Michaelis-Menten plot for the cleavage of the 3SLN substrate by Anhui13 NA, with the fitted kinetic parameters shown in Table 4.2. These parameters for 3SLN were measured with NA at a monomeric concentration of 0.5 nM. The cleavage of 3SLN by the WT NA appears to be efficient with a K_m of 810.0 ± 100 μM and a k_{cat} of 159.5 ± 8.5 s^{-1}. These values are similar to those determined for X-31 (K_m = 562.3 ± 20 μM and k_{cat} = 97.5 ± 1.5 s^{-1}). The N9 has a slightly higher K_m, indicating that it is less efficient at binding to the substrate, but also has a higher k_{cat}, indicating a higher rate of substrate turnover. Despite these differences the overall catalytic efficiency(k_{cat}/K_m) is similar (0.197 $\mu M^{-1}s^{-1}$ for the N9 and 0.173 $\mu M^{-1}s^{-1}$ for X-31). When comparing 3SLN cleavage with that of MUNANA it is apparent that there is a similar pattern to that seen with X-31 NA with a higher K_m for 3SLN, indicating weaker substrate binding, but a higher k_{cat} for substrate cleavage. This indicates, as discussed previously in section 3.2.7.2, that the sialidase site of the NA can bind better to the artificial hydrophobic MUNANA but has a lower cleavage rate than the natural substrate 3SLN. These 3SLN cleavage data agree well with the HA/NA balance BLI

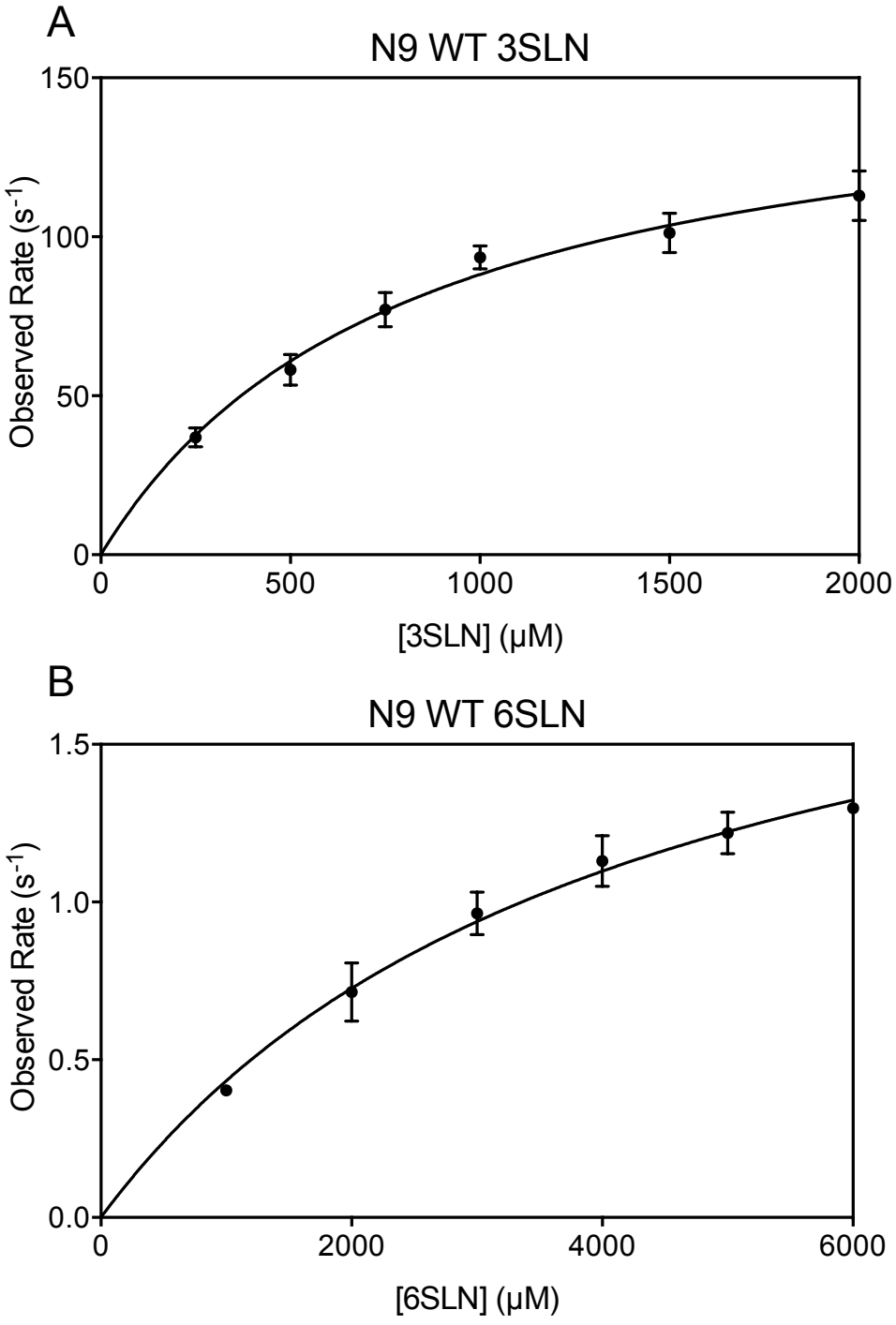


Figure 4.13: Michaelis-Menten plots of the kinetics of 6SLN and 3SLN cleavage by WT N9 NA. Plots are shown for A) 3SLN and B) 6SLN. Cleavage rate measured by measuring rate of N-acetyllactosamine release using coupled enzyme system of β-galactosidase and β-galactose dehydrogenase using the absorbance of generated NADH as the reporter of rate. Each point is mean of three repeats with error bars showing standard deviation from the mean. Least squares fit of Michaelis-Menten equation is shown. Raw reaction rate data are converted to an observed rate by dividing rates by NA concentration used in assay.

experiments for H7N9 and H1N9 viruses with WT NA binding to 3SLN-PAA (Figs. 4.9 and 4.11), where the viruses do not bind extensively to the sensor in the absence of NA inhibitors despite strong binding to 3SLN-PAA by both the H1 and H7 HAs (Figs. 4.5 and 4.6), in a fashion similar to that seen with X-31 binding to 3SLN-PAA (Fig. 3.2).

Figure 4.13B shows a Michaelis-Menten curve determined for the WT N9 cleaving 6SLN and Table 4.2 shows the fitted parameters. Cleavage of 6SLN was measured with NA at a 50-fold higher monomeric concentration (25 nM). The K_m value determined was 4179 ± 615 μM, which is ~5-fold higher than that determined for 3SLN (810 ± 100 μM). The turnover rate (k_{cat}) for 6SLN is very low (2.2 ± 0.2 s^{-1}) compared with that measured for 3SLN (159.5 ± 8.5 s^{-1}). When compared to the results obtained for X-31 NA the N9 has ~2-fold lower K_m and ~10-fold lower k_{cat} for 6SLN cleavage, indicating that the sialidase site has greater affinity for binding to 6SLN but has a much lower likelihood of this interaction resulting in substrate cleavage, making it potentially more effective as a receptor binding site.

If one assumes the simplest possible enzymatic reaction and takes a reasonable value for the association constant (k_1) of sialic acid for the NA active site of 5 x 10^5 M^{-1} s^{-1} (Sung, Wynsberghe & Amaro, 2010) one can calculate the dissociation constant (k_{-1}) using the K_m and k_{cat} values determined for the N9 using a simple rearrangement of Equation 4.1 to produce Equation 4.2:

$$K_m = \frac{k_{-1} + k_{cat}}{k_1} \tag{4.1}$$

$$k_{-1} = K_m.k_1 - k_{cat} \tag{4.2}$$

From the calculated k_{-1} and the experimentally determined k_{cat} one can determine the probability of the formation of an enzyme substrate complex leading to a cleavage event or to the substrate dissociating by determining the ratio between k_{cat} and k_{-1}. The N9 is ~1000-fold more likely to dissociate, rather than cleave 6SLN compared to a ~1.5-fold likelihood for 3SLN. The comparable probabilities for X-31 are a ~200-fold higher probability of 6SLN dissociation rather than cleavage and a ~1.8-fold likelihood for 3SLN. These kinetic parameters therefore show that the N9 has a ~5-fold increased likelihood of 6SLN dissociation from the N9 NA when compared to X-31 NA. This coupled with the observation of a lower K_m, compared to X-31 NA, reinforces the hypothesis that substrate binding via the sialidase site is responsible for the enhanced initial binding of H7N9 and H1N9 viruses with the WT NA to 6SLN-PAA (Figs 4.8 and 4.10).

4.2.8.3 Multivalent substrate kinetics

It has been previously shown that the presence of the Hb site in an NA enhances the cleavage of large multivalent substrates such as heavily glycosylated glycoproteins (Uhlendorff *et al.*, 2009). The elution of H1N9 viruses from TRBCs also suggests an enhanced NA activity when the Hb site is absent (Table 4.1). Experiments were therefore carried out to investigate whether this N9 NA also shows the same enhancement of cleavage. An enzyme assay was devised to monitor the rate of cleavage of the glycoprotein fetuin. The previously designed system used to quantify 6SLN and 3SLN cleavage using β-galactosidase and β-galactose dehydrogenase did not work in preliminary experiments carried out using fetuin. This is most probably due to problems with the β-galactosidase cleavage step. In the previous system the β-galactosidase was used to cleave the disaccharide N-acetyllactosamine, which is

released from both 6SLN and 3SLN after NA cleavage. This released sugar consists of a galactose connected by a β1,4 linkage to an N-acetylglucosamine moiety. There are two potential problems with using fetuin as the NA substrate which could affect the efficiency of the β-galactosidase used in the assay. The linkages present on the fetuin (β1,4 or β1,3) could be different and the substrate will also be presented in forms longer than a trisaccharide, either of which scenarios could alter β-galactosidase activity.

A new system was devised to measure the released sialic acid, rather than the remaining glycans of the fetuin. This system was based on the two coupling enzymes N-acetyl neuraminic acid aldolase (NANA aldolase) and lactate dehydrogenase (LDH). The NANA aldolase catalyses the release of pyruvate from sialic acid and the LDH converts the released pyruvate into lactate, while converting NADH into NAD^+, the conversion of which can be monitored by measuring the reduction of absorbance at 340 nm, see reaction scheme in Figure 4.14. The quantities of enzyme were optimised to ensure that enough of both coupling enzymes were present to correctly report NA rate, as was done for the development of the previous system (section 3.2.7.2), by ensuring that at the highest and lowest concentrations of substrate the initial rate doubled when the amount of NA was doubled (data not shown). The reaction components in a 100 µl reaction were: 0.5 U NANA aldolase, 5 U LDH and 200 µM NADH. Fetuin concentrations were calculated assuming a molecular mass of 48.4 kDa. Experiments were done in the BLI assay buffer HBS-P + $CaCl_2$. See *Materials and Methods*, section 2.2.11.3 for full experimental details.

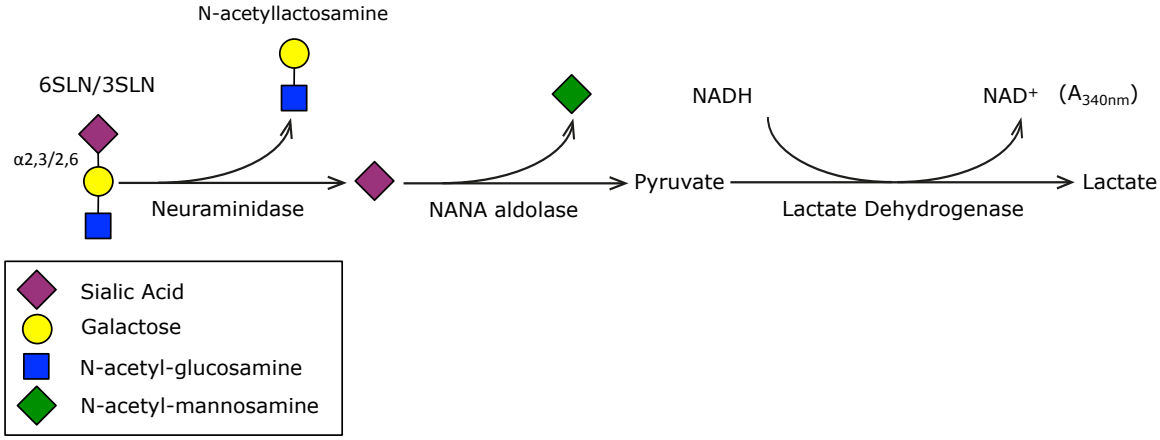


Figure 4.14: Reaction scheme of coupled enzyme assay used to monitor cleavage of sialic acid from fetuin.

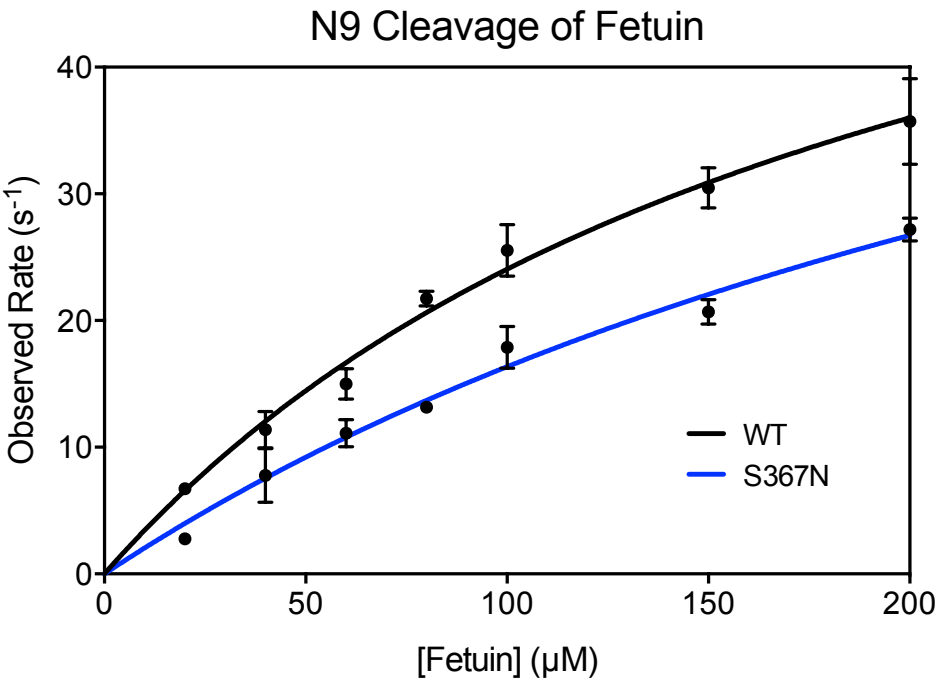


Figure 4.15: Michaelis-Menten plot of fetuin cleavage by WT and S367N N9 NAs. Rate measured by assaying released sialic acid using NANA aldolase and lactate dehydrogenase as coupled reporter system, measuring rate of conversion of NADH to NAD^+ by absorbance as reporter of NA rate. Each point shown is mean of three repeats with error bars indicating standard deviation from the mean. Least squares fit of Michaelis-Menten equation is shown. Raw reaction rate data are converted to an observed rate by dividing rates by NA concentration used in assay.

Figure 4.15 shows the Michaelis-Menten plot for fetuin cleavage by the expressed WT and S367N NA N9 proteins at a monomeric NA concentration of 5 nM. Table 4.2 shows the fitted parameters for these plots. Both plots give sufficient data to allow determination of K_m and k_{cat} values. Results at higher fetuin concentrations would have been desirable but the maximum concentration used (200 μM) is ~ 10 mg/ml. It was considered likely that there could be problems with fetuin solubility at higher concentrations and also that the pH could be perturbed by the large amounts of sialic acid present. The pH of the 200 μM fetuin was checked and this protein concentration did not affect the pH of the buffered solutions used.

The WT N9 protein has a low K_m (197.5 ± 37.3 μM) compared to the comparable monomeric sugar substrates 6SLN and 3SLN (K_m = 4173 ± 615 μM [6SLN] and 810.0 ± 100 μM [3SLN]). This does not necessarily indicate a stronger binding affinity for the sialic acid on the fetuin but rather that the fetuin is likely to contain multiple sialic acid moieties on each protein, thus lowering the K_m. The k_{cat} is, however, a direct measure of the amount of released sialic acid and is therefore directly comparable to the values for 6SLN and 3SLN. The k_{cat} value for the WT protein is 71.6 ± 8.3 s^{-1}, which ~2-fold lower than the value for 3SLN (159.5 ± 8.5 μM) and ~30-fold higher than the value for 6SLN (2.2 ± 0.2 s^{-1}). This difference in k_{cat} values could be due to there being a mixture of the more efficiently cleaved α2,3-linked sialic acid and the less well cleaved α2,6-linked sialic acid, thus making the rate somewhere in the middle of the two. It could also be due to the altered substrate characteristics when it is attached to a larger protein rather than a freely diffusing trisaccharide moiety. The sialic acid will have a slower diffusion rate being attached to a protein, as well as the other possible confounding factors such as the way the

substrate is presented on the protein surface. There are also the complications of multivalency. The expressed protein used is in a tetrameric form and consequently will have active sites that could be spaced in a way that allows sialic acid residues from the same fetuin molecule to bind multiple sialidase sites simultaneously.

The data obtained for the cleavage of fetuin by the S367N mutant N9 show a higher K_m value of 346.4 ± 95.6 μM (WT K_m = 197.5 ± 37.3) an increase of ~1.75-fold; however the k_{cat} is very similar (k_{cat} = 71.6 ± 8.3 s^{-1} [WT] and 73.0 ± 14.5 s^{-1} [S367N]). This indicates that the presence of the Hb site enhances binding of fetuin to the N9 but does not alter the overall cleavage rate when substrate concentration is not limiting. The Hb site is therefore likely to be involved in tethering the substrate in order that the sialidase site can bind and cleave the substrate more efficiently.

4.3 Discussion

The experiments using the TRBC pull down assays presented in this chapter confirm that the N9 of H7N9 viruses has sialic acid binding properties, and that this binding can be eliminated by inserting the substitution S367N into the N9 NA (Fig 4.2). The removal of the Hb binding site by mutagenesis was found to reduce overall virus binding in the presence of NA inhibitors. This was seen when measuring overall virus binding as a function of sugar loading by BLI for H7N9 (Fig 4.5) and for H1N9 (Fig 4.6), and also in the reduction in binding amplitude for 6SLN-PAA binding for the very weak binding H3N9 viruses constructed in this work (Fig 4.7). This loss of binding is solely due to this S367N substitution eliminating binding at the second binding site.

HA/NA balance measurements made of H7N9 and H1N9 viruses binding to 6SLN-PAA indicate that the sialidase site has a role in binding to α2,6 linked sialic acid, seen by the presence of an initial enhancement of binding to 6SLN-PAA when the Hb site is present and the sialidase site is uninhibited (Figs. 4.8 and 4.10), indicating that there is a cooperativity in receptor binding via the two binding sites. These observations of the involvement of the sialidase site in α2,6 binding was reinforced by NA kinetic measurements made for the cleavage of 6SLN, indicating that the WT N9 NA has ~2-fold lower K_m and ~10-fold lower k_{cat} when compared to the previously measured parameters for X-31 NA (Table 3.1), indicating that the NA has stronger interactions with the receptor but a much lower probability of this binding resulting in a cleavage event.

Understanding the potential impact of these changes on overall virus receptor binding is more difficult. The inclusion of the secondary binding site clearly causes an increase in the apparent K_d for virus binding (~2-fold) but estimating the energetic contribution from the secondary site is difficult without more information. The binding constant for the second binding site is not known and attempts to measure this using the MST experiments detailed in section 4.2.7 were unsuccessful. The binding is most likely to be very weak, with a K_d most likely in the mM range, as previously determined crystal structures of a similar N9 with receptor bound in the secondary site required 20 mM sialic acid in order to see electron density (Varghese *et al.*, 1997) and crystal structures of the Anhui13 NA studied here required 50 mM of receptor in order to see good electron density (Xiaoli Xiong, unpublished data). It is worth noting that if the K_d for the interaction is indeed in the low mM range then it is

similar to values determined for the interaction of HA with receptors measured using NMR and MST (Sauter *et al.*, 1989; 1992b; Xiong *et al.*, 2013a).

The presence of this site in the NA rather than the HA adds a number of geometrical complications to receptor binding. NA is present on viruses in clusters, which are thought to be at the area of the virus closest to the host cell membrane upon budding and membrane scission (Calder *et al.*, 2010), see *Introduction*, section 1.4.4, for more details. The effect of this clustering of NA upon binding to a surface is unclear. When experiments are carried out with the sialidase site inhibited the clustered NA could increase overall binding if the affinity for sialic acid is higher than the NA or the binding sites could be more accessible to the surface immobilised receptor either due to the wider spacing of binding sites, compared to HA, or due to the binding mode of sialic acid by the Hb site, where the sialic acid binds to allow the attached sugars to attach perpendicular to the surface of the protein, rather than the more integrated binding of HA (Varghese *et al.*, 1997) (Fig. 1.9B).

As well as this complication of NA clustering the height of the NA 'head region' compared to the height of the HA can be mediated by altering the NA stalk length. A reduction in stalk height leads to a lower activity of NA at cleaving surface immobilised sialic acid, as shown in section 3.2.10. The stalk of the N9 reduced by five residues when compared to other similar N9 proteins from wild birds, deleting residues 68 – 72 (Chen *et al.*, 2015). This reduction in stalk length would likely reduce the effective NA sialidase activity, similar to the 10 residue deletion measured of X-31 NA Δ10 (Fig. 3.15). A 5 residue deletion from the stalk would likely bring the globular head of the NA to a similar height as the HA, as a typical stalk length NA protrudes above the level of HA (Harris *et al.*, 2006). The reduction in stalk length

would alter the geometry of the NA binding properties measured in this chapter. It could, logically, either strengthen or weaken the overall virus binding via HA and NA. The reduction of stalk length could prevent the interaction of NA which is present in a neighbouring area to the virus binding 'footprint' which could be enhancing the overall virus binding. The reduction of stalk length could also increase virus binding by making the surface of the virus flatter, without the NA protruding above the HA, therefore allowing better binding of all proteins.

A reduction in NA stalk is typically considered a replicative adaptation of viruses to poultry (see *Introduction*, section 1.3.2.4). The reduction of stalk length in these H7N9 viruses appears to be an adaptive change in the NA which is required for human infection, as there have been very few H7N9 human isolates containing a full length stalk (2 of 174 full length NA sequences, A/Guandong/3/2013 and A/Shanghai/5/2013, GISAID (http://gisaid.org) accessed 22.4.15). It is also worth noting that there have been no avian H7N9 viruses isolated since 2011 with a full length NA stalk including viruses isolated during the period of much enhanced surveillance of H7N9 viruses in avian species since the outbreak began in 2013. It appears counterintuitive that a reduction in stalk length would be a selective advantage in terms of reducing sialidase activity for a virus with an NA so poor at cleaving the human α2,6-linked receptors. It therefore appears that it could be a selective advantage for the receptor binding properties of the NA.

Overall virus binding affinity has been previously related to the affinity of HA for monomeric sugars using a multiplicity coefficient (Xiong *et al.*, 2013a). Trying to include energetic contributions from a binding site on an additional protein (the NA) would clearly be very difficult. On the one hand the presence of additional binding

partners would obviously be expected to increase the overall affinity for the virus and the fact that the NA is generally taller than the HA could effectively increase the surface of the virus able to make contacts with the surface. On the other hand any binding by the NA might actually reduce the number of HA molecules in the virus 'footprint' that can interact with the surface sugars thus weakening the interaction.

There have been suggestions in previous reports that oseltamivir carboxylate can bind to the Hb site of NA (Lai *et al.*, 2012). There is no apparent evidence of this binding in the experiments carried out in this chapter at the concentrations of the drugs used. The TRBC pull down assay showed strong binding in the presence of the high concentration of 100 μM of oseltamivir carboxylate, indicating no inhibition of the Hb site, (Fig. 4.2). These experiments are consistent with previous work that concluded that there was no effect of NA inhibitors on the haemadsorption activity (Laver *et al.*, 1984).

The enhanced binding of the N9 to 6SLN through the sialidase site is associated with a low K_m and a very low k_{cat} and is similar to variants of the NA from H3N2 viruses, which gain substitutions upon adaptation to propagation in tissue culture. These H3N2 viruses have substitutions in the NA at positions 151 which typically occur when viruses are propagated in MDCK cells rather than MDCK-SIAT cells which have elevated levels of α2,6 sugars (Lin *et al.*, 2010). Expressed N2 proteins with these mutations have MUNANA cleavage kinetics with low K_m and low k_{cat} and reduced overall receptor cleavage (Zhu *et al.*, 2012a) (Stephen Wharton, unpublished data). The sialidase site in the N9 does not have the same properties as that of the binding N2 with a 151 substitution, as the N9 retains efficient MUNANA and 3SLN cleavage but only has poor 6SLN cleavage. It is unclear whether this 6SLN

binding activity is an evolved trait or requirement for human infection or whether it is simply a property that is inherent to the avian NA.

What are the implications of the binding properties of these H7N9 viruses for human infection? The H7N9 virus is from an avian source and the binding properties are similar to those of other avian viruses. It has an overall preference for the surface immobilised avian-like receptor analogue 3SLN-PAA (Fig 4.5) and the NA is highly active against the monomeric sugar substrate 3SLN (Table 4.2). There is also, however, strong binding for the human-like receptor analogue 6SLN-PAA, which is not a common feature in a number of other avian H7 and H5 viruses. The presence of binding to 6SLN-PAA is likely to be a result of the Q226L mutation present in the H7 HA (Xiong *et al.*, 2013b; Dortmans *et al.*, 2013).

When looking at the binding properties to the receptors more pertinent to the majority of human infections, α2,6 linked sialic acid, the viruses show interesting properties, with the overall virus binding in the presence of NA inhibitors being strong for 6SLN-PAA, although not as strong as for 3SLN-PAA (Fig 4.5). The NA is also ineffective at cleaving 6SLN, as seen when measuring the enzyme kinetics (Table 4.2) and the virus release is comparatively slow in the HA/NA balance experiments (Fig. 4.8), when compared to X-31 (Fig 3.2). The kinetic parameters for the NA for 6SLN cleavage show a very low overall catalytic rate, k_{cat}, and an enhanced binding, low K_m, when compared to the parameters determined for X-31. This suggests that the N9 NA sialidase site has enhanced binding properties, making it contribute to receptor binding to α2,6-linked sialic acids.

Consideration of all the data presented here suggests that the NA of H7N9 viruses enhances overall virus binding to α2,6-linked sialic acids, which are the

receptors present in the primary site of human infection, the upper respiratory tract epithelial cells. This enhancement of binding from the dual effects of the Hb site and the sialidase site could aid initial virus infection but may act as a restrictive factor when the virus is released from the cell to initiate infection elsewhere.

This combination of binding factors could explain why H7N9 viruses appear to efficiently infect humans, whilst having limited human-human transmissibility. The viruses manage to infect humans aided by their enhanced binding to α2,6-linked sialic acid mediated by the sialidase and Hb sites and these viruses cause typical infections, which are unlikely to be adversely affected by the poor release of viruses from α2,6-linked sugars owing to the large number of viruses produced by each infected cell and the close proximity of other cells to infect. This lack of release from α2,6 linked sugars, however, could be the reason that there is limited human-human transmissibility of these viruses, as the requirements for virus release to enable aerosol transmission are likely to be much more stringent than those for cell-to-cell transmission within the same host, thus requiring efficient NA to release large quantities of virus from aggregates, to promote detachment from cells and to evade non-specific inhibitors such as mucins.

5 Biophysical Receptor Binding Studies of H1N1 2009 Pandemic Emergence

5.1 Introduction

A new lineage of H1N1 IAVs emerged in 2009, causing a pandemic, which is thought to have infected an estimated 60 million people between April 2009 and April 2010 in the USA alone (Shrestha *et al.*, 2011), as well as similarly widespread infections worldwide. These 2009 pandemic (pdm09) viruses were the result of a reassortment of viruses from four different lineages, as described in *Introduction,* section 1.8.2.4 (Fig. 1.11). This novel lineage of reassortant viruses is thought to have infected humans as a zoonosis from swine. Preceding the emergence of the 2009 pandemic virus there had been viruses circulating in swine that contained viral segments from three of the sources in the pdm09 viruses, known as triple-reassortant swine (TRS). These TRS viruses differ from the pdm09 in the source of the M and NA segments as well as a number of amino acid substitutions in the HA. TRS viruses were responsible for a number of sporadic human zoonoses (Shinde *et al.*, 2009); however, there was no evidence for direct human-to-human transmission of these viruses.

As described previously a major factor in determining the transmissibility of IAVs is the interactions of the surface glycoproteins HA and NA with the cell surface. Experiments presented in this chapter were carried out with both TRS and pdm09 viruses in an attempt to increase our understanding of which factors increased the transmissibility of the virus and led to the pandemic. Experiments were carried out to investigate the HA mediated receptor binding of pre-pandemic TRS viruses and pdm09 to receptor analogues, which vary at a number of positions owing to a number of drift mutations of the csw H1 HA. Investigations were also carried out into the HA

and NA balance characteristics of these two lineages of viruses, examining both the HA, from classical swine (csw) sources, and the NA, which is from Eurasian swine sources.

5.2 Results

5.2.1 Choice of viruses for studies

Prototypic viruses were chosen to represent the pre-pandemic TRS viruses and the pdm09 virus. A/Iowa/1/2006 (Iowa06) was chosen as a representative TRS virus. This virus was isolated from a human zoonotic infection of a swine H1N1 (Shinde *et al.*, 2009). A/California/7/2009 (Cal7) was chosen as a representative pdm09 virus. It was one of the first isolates obtained upon the emergence of pdm09 viruses. Cal7 has been the H1N1 vaccine virus since the first inclusion of a pdm09 virus in the trivalent/quadravalent inactivated vaccine. Both Iowa06 and Cal7 also have full reverse genetics background systems (6 segments PB2, PB1, PA, NP, M and NS) previously cloned in the McCauley laboratory by Dr. Haixia Xiao.

5.2.2 Analysis of HA sequences pertaining to pdm09 emergence

Sequence alignments of pdm09 and pre-pandemic TRS virus HA sequences were carried out. A selection of TRS viruses containing an H1 glycoprotein from the classical swine lineage and also a selection of the first tissue culture-propagated pdm09 isolates were gathered from the GISAID database (http://gisaid.org) and aligned. These alignments were carried out to identify conserved substitutions in the HA, which occurred between TRS viruses, carrying a classical swine (csw) lineage H1 HA, and the emergence of the first strains of pdm09 viruses. The conserved substitutions found during these alignments were mapped onto a crystal structure of

a pdm09 HA (Xu *et al.*, 2012a). The common substitutions occurring around the receptor binding site of the HA were selected for further investigation. The sites of interest, as well as other residues known to play a major role in determining receptor specificity are shown in Table 5.1. The location of these residues with respect to the receptor binding site is shown in Figure 5.1. An alignment of the sequences of Cal7, Iowa06 and a number of strains used in previous studies is also included in Figure 5.2. The conserved differences (Table 5.1) thought to have the potential to alter receptor binding of viruses are generally in regions of the HA which are known to alter sialic acid receptor binding, such as the 130-loop, the 190-helix and the 220-loop. Conserved amino acid substitutions, between TRS and pdm09 H1 HAs, were found at the following positions: 132, 149, and 227 (H3 numbering). Residues 134, 186, 219 and 225 show a conserved residue in early pdm09 viruses but show a variation in csw HA sequences. Residue 134 is typically a Lys in pdm09 and an Arg in csw viruses, however the Iowa06 virus used here, and a number of other csw H1 containing isolates, contain the pdm09-like Lys. Residue 186 and 189 are conserved as a Ser 186 and Ala 189 in pdm09 viruses; however, there appears to be polymorphism at these positions in csw HA sequences, with viruses either carrying the Ser 186 and Ala 189 as with pdm09 viruses or Pro 186 and Thr 189. Residue 219 is a conserved Ile in pdm09 viruses but shows polymorphism between Thr and Ala in csw viruses.

5.2.3 Construction of reverse genetics viruses

The HA and NA of Cal7 and Iowa06 were cloned into the reverse genetics plasmid pHW2000, as described in *Materials and Methods*, section 2.2.6.12.1. The Cal7 virus that the HA was cloned from was egg-adapted and carried the adaptive substitution Q226R, which is not present in the sequences of tissue culture-

	HA Residue													
	132	**134**	**149**	**186**	**189**	**190**	**200**	**208**	**219**	**225**	**226**	**227**	**228**	**H3 Numbering**
Virus	128	130	146	183	186	187	197	205	216	222	223	224	225	H1 Numbering
pdm09														
A/California/7/09	S	K	K	S	A	**D**	A	**R**	I	D	**Q**	E	**G**	
A/California/4/09	S	K	K	S	A	**D**	T	**R**	I	D	**Q**	E	**G**	
A/Texas/05/09	S	K	K	S	A	**D**	A	**R**	I	D	**Q**	E	**G**	
A/England/195/09	S	K	K	S	A	**D**	A	**R**	I	D	**Q**	E	**G**	
Classical Swine														
A/Ohio/1/07	T	R	R	S	A	**D**	A	**R**	T	D	**Q**	A	**G**	
A/Iowa/1/06	T	K	R	P	T	**D**	A	**K**	A	N	**Q**	A	**G**	
A/swine/Iowa/00239/04	T	R	R	P	T	**D**	A	**K**	A	D	**Q**	A	**G**	
A/swine/Illinois/100084/01	T	R	R	S	A	**D**	A	**R**	T	D	**Q**	A	**G**	
A/swine/Ohio/891/01	T	R	R	S	A	**D**	A	**R**	T	D	**Q**	A	**G**	
A/swine/Indiana/P12439/00	T	R	R	S	A	**D**	A	**R**	A	D	**Q**	A	**G**	
A/New Jersey/8/76	T	X	R	P	T	**D**	A	**K**	A	G	**Q**	A	**G**	

Table 5.1: Table of the sequences of HA residues under investigation for a number of representative csw and pdm09 HAs.

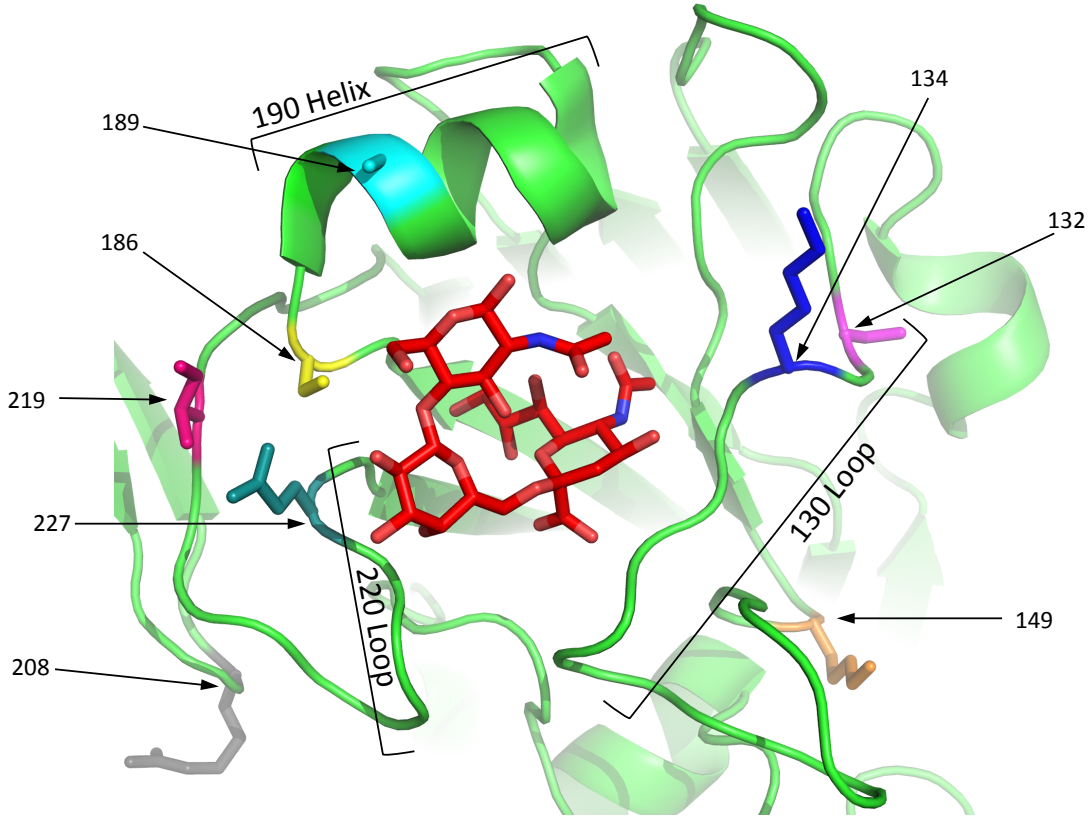


Figure 5.1: Structural location of H1 HA residues studied. Plotted on to structure of A/California/4/2009 HA complexed with 6SLN (PDB: 3UBN) (Xu *et al.*, 2012a). Shown in red is the receptor analogue 6SLN. Diagram made using PyMol (PyMol Molecular Graphics System, Schrödinger).

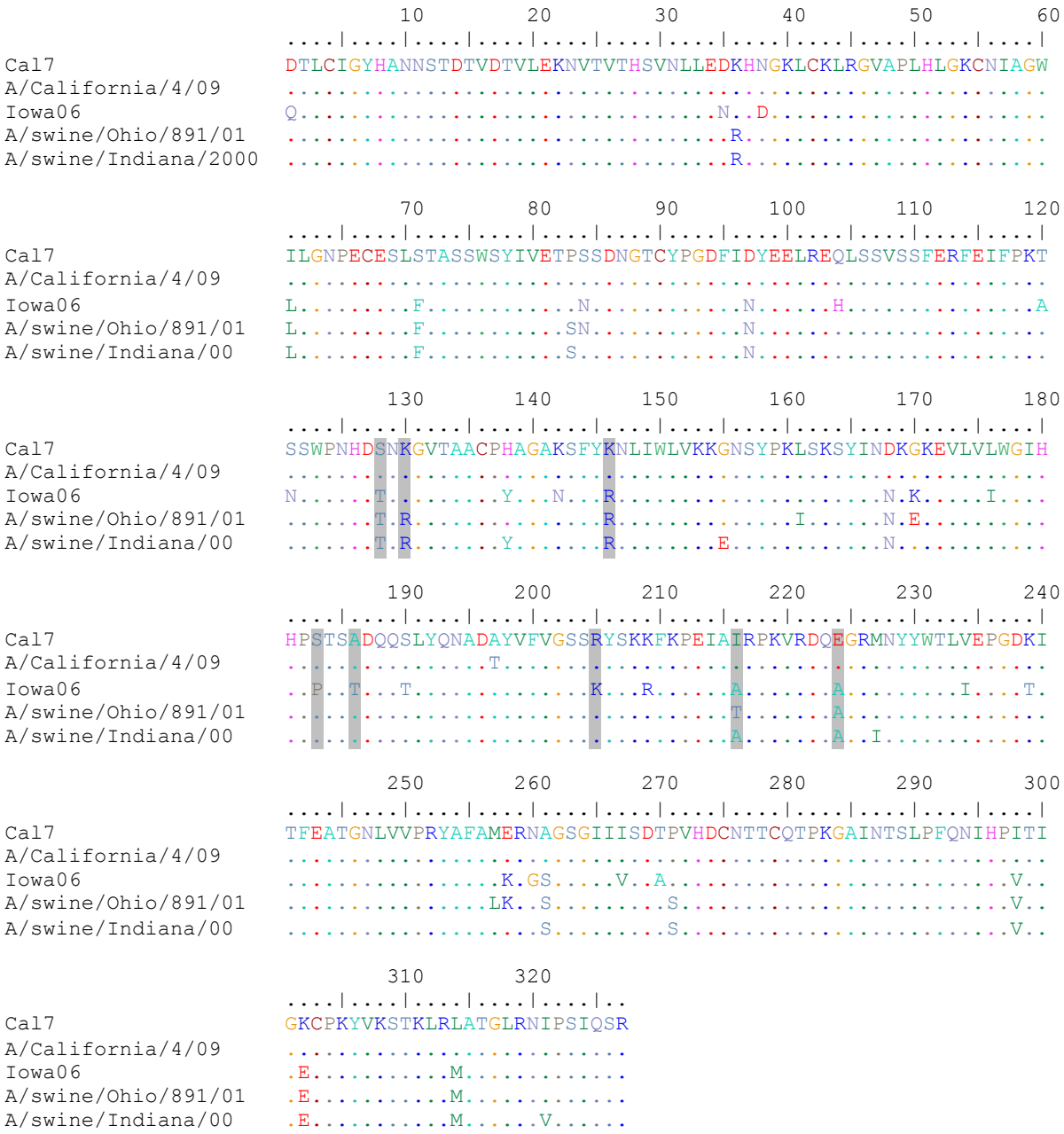


Figure 5.2: Alignment of H1 HA1 sequences. HA1 sequences are aligned for the two viruses under study (Cal7 and Iowa06) and viruses used for other receptor binding studies of pdm09 emergence. A/California/4/09 and A/swine/Ohio/891/01 used in (de Vries *et al.*, 2011). A/California/4/09 and A/swine/Indiana/P12439/2000 (A/swine/Indiana/00) used in (Xu *et al.*, 2012b). Residues are shaded which are studied in this chapter. Alignment made is in H1 numbering. 132 (H3 numbering) is 128 (H1 numbering), 134 (H3) is 130 (H1), 149 (H3) is 146 (H1), 186 (H3) is 183 (H1), 189 (H3) is 186 (H1), 208 (H3) is 205 (H1), 219 (H3) is 216 (H1), 227 (H3) is 224 (H1).

propagated isolates of Cal7 deposited in GISAID (http://gisaid.org). This mutation was reversed by site directed mutagenesis, as described in *Materials and Methods*, section 2.2.6.13.1. Iowa06 HA was cloned from an egg-propagated stock of the virus. Upon sequencing the cloned HA gene of Iowa06 all clones analysed contained the substitutions D131E and G157E (H3 numbering), when compared to the published sequence deposited in GISAID (http://gisaid.org). These changes were converted to the wild type sequence by site directed mutagenesis. The HA of Iowa06 also contains an Asn at position 225, which is different from other classical swine HA isolates. It was decided to mutate the HA to a more classical swine-like sequence by mutagenising the plasmid to contain the substitution N225D. This mutagenised form of Iowa06 will be referred to as wild type in the studies presented here.

Recombinant viruses were generated by transfecting the cloned HA and NA plasmids, with the remaining six gene segments from PR8, into 293T cells, as described in *Materials and Methods*, section 2.2.5. Iowa06, rescued with a PR8 background and subsequently propagated well in tissue culture. Cal7 did not propagate when the transfections were carried out using a PR8 background; however, the virus rescued and propagated successfully when using a full 6 segment Cal7 background. Viruses used for HA avidity receptor binding studies were: Iowa06 based viruses in a PR8 background and Cal7 based viruses with a full 6 segment Cal7 background.

Viruses were initially propagated in minipig kidney (MPK) cells, as they were found to give the best titres for WT and mutant versions of both Iowa06 and Cal7 viruses. Iowa06 viruses propagated well and had the correct sequence after passage; however, although the Cal7 viruses propagated well they were found upon

sequencing to have gained polymorphisms at positions 157 – 159 with the substitutions: K157E, G158E or N159E. These mutations have been found previously upon pdm09 virus passage (Klimov *et al.*, 2012), particularly in MDCK-SIAT cells, which are MDCK cells stably transfected with a plasmid expressing additional α2,6 sialyltransferases (Matrosovich *et al.*, 2003). A range of MDCK cell lines were tested for optimal virus titre, without altering the HA sequence. The cells that gave the best viral yield were ECACC MDCK cells. These were used for further propagation of these viruses.

Biophysical analyses required large amounts of concentrated virus. Viruses were passaged three times, from the transfection supernatant, with the final passage in a large volume (~1 Litre), as described in *Materials and Methods*, section 2.2.3.2. Viruses were concentrated by centrifugation and quantified as described in *Materials and Methods*, section 2.2.9.2. Virus HA and NA genes in these preparations were sequenced to verify the absence of undesired mutations and sequence polymorphisms, as described in *Materials and Methods*, section 2.2.6.11.

5.2.4 Biolayer interferometry measurements of virus affinity

Receptor binding assays were carried out based on a previously established assay, where virus binding is measured by BLI to sensors loaded with different amounts of the receptor analogues 6SLN-PAA and 3SLN-PAA. As described previously data for virus fractional saturation were then plotted as a function of sugar loading (*Materials and Methods*, section 2.2.10.1).

Binding for all H1N1 viruses studied was only detectable to the human-like receptor analogue 6SLN-PAA. Although measurements of binding to 3SLN-PAA were

carried out there was no measurable signal for any of the viruses constructed at saturated sugar loading (~0.6 nm) both at 1 x (100 pM) and 10 x (1 nM) virus concentrations.

Results are presented (Figures 5.3, 5.4, 5.6 and 5.8) for the various mutant viruses constructed. The difference in relative K_d of virus binding to 6SLN-PAA has also been determined, as described in *Materials and Methods*, section 2.2.10.1. These are presented in Table 5.2 and are expressed relative to the relevant wild type virus, which is given a normalised K_d of 1. Values with an increase or decrease greater than 2-fold are considered significant.

5.2.4.1 Wild-type pdm09 Cal7 and TRS Iowa06 viruses

Figure 5.3 shows the BLI data for the WT viruses (at 100 pM) binding to both receptor analogues 6SLN-PAA and 3SLN-PAA. Both Iowa06 and Cal7 show very strong binding to 6SLN-PAA, similar to that seen with X-31 (Fig. 3.1), with no measurable binding to 3SLN-PAA. The 6SLN-PAA binding is very similar for both Iowa06 and Cal7.

5.2.4.2 Mutant viruses: HA residues 132, 134 and 149

Figure 5.4 shows the binding of Iowa06 and Cal7 viruses with amino acid substitutions at positions 132, 134 and 149. The substitution S132T in Cal7 HA reduces overall binding of the virus to 6SLN-PAA, with an increase in relative K_d of ~4-fold, whereas the equivalent mutation in the Iowa06 HA has no significant effect on receptor binding. Position 134 displays polymorphism between Lys and Arg in classical swine HA sequences but is always Lys in pdm09 viruses. Both Cal7 and Iowa06 WT viruses have a Lys at this position. Cal7 K134R and Iowa06 K134R show

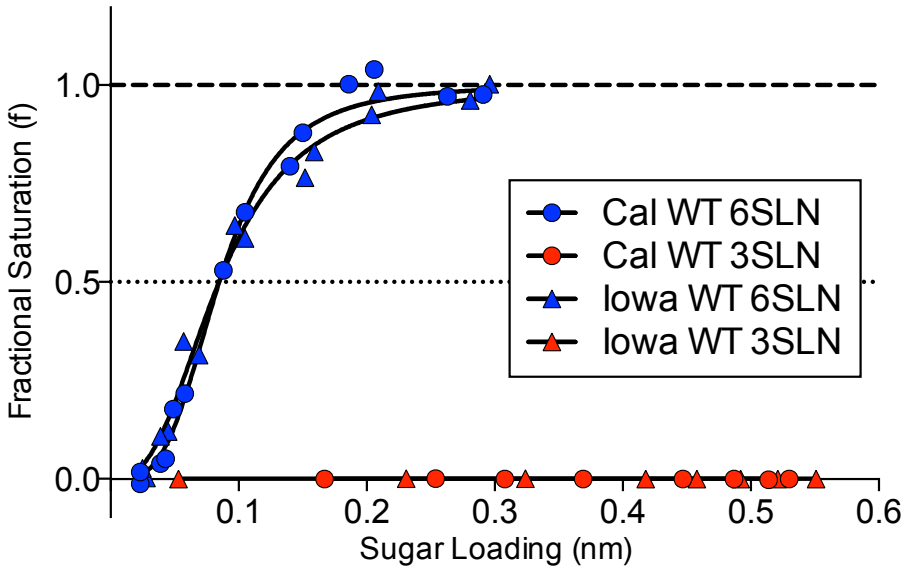


Figure 5.3: Equilibrium binding of WT Cal7 and Iowa06 to 6SLN-PAA and 3SLN-PAA. Fractional saturation of virus binding measured as a function of sugar loading in the presence of NA inhibitors.

Compared to Cal7 WT		Compared to Iowa06 WT	
Virus	**Relative K_d**	**Virus**	**Relative K_d**
Cal7 WT	1	Iowa06 WT	1
Cal7 S132T	4.37	Iowa06 T132S	**0.61**
Cal7 K134R	**1.25**	Iowa06 K134R	**0.63**
Cal7 K149R	7.03	Iowa06 R149K	8.79
Cal7 S186P	0.10	Iowa06 P186S	4.16
Cal7 A189T	2.98	Iowa06 T189A	0.13
Cal7 R208K	**0.94**	Iowa06 K208R	15.66
Cal7 I219T	0.15	Iowa06 A219I	0.14
Cal7 I219A	0.33	Iowa06 A219T	**1.33**
Cal7 E227A	0.07	Iowa06 A227E	**1.19**
Cal7 I219A E227A	0.04	Iowa06 A219I A227E	3.79
Cal7 I219T E227A	0.05	Iowa06 A219T A227E	3.91
Iowa06 WT	**0.99**	Cal7 WT	**1.01**

Table 5.2: Relative K_d values of mutant Cal7 and Iowa06 viruses binding to 6SLN-PAA. Values are calculated relative to the relevant wild type virus, as described in *Materials and Methods*, section 2.2.10.1. Values shown in bold are differences <2-fold and therefore considered not significantly different from the wild-type.

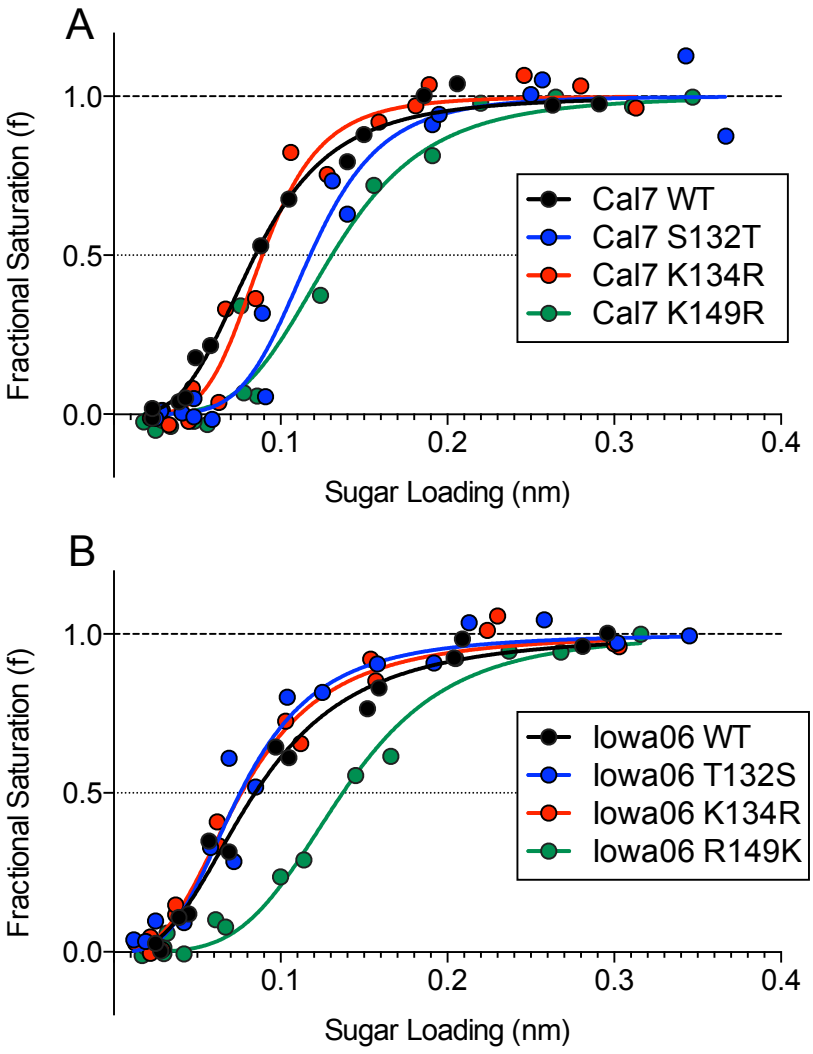


Figure 5.4: Equilibrium virus binding to 6SLN-PAA of Cal7 and Iowa06 viruses with HA residue substitutions at positions 132, 134 and 149. Fractional saturation of virus binding measured as a function of sugar loading in the presence of NA inhibitors.

no significant difference in the binding to 6SLN-PAA compared to the WT viruses. The substitution K149R in Cal7 HA causes a decrease in 6SLN-PAA receptor binding affinity, with an increase in relative K_d of ~7-fold. The opposite substitution R149K in Iowa06 HA also causes a reduction in receptor affinity of a similar magnitude, with a reduction in relative K_d of ~9-fold.

Differences in binding affinity are seen when substitutions are introduced at positions 132 and 149. Neither of these residues are directly involved in interacting with the receptor, see Figure 5.5A. The role of these residues in altering receptor binding is likely to be due to alterations in the coordination of the 130-loop. This could interfere with a number of interactions such as the backbone carboxyl of Val 135 with the nitrogen of the acetamido group of sialic acid and the interactions of the hydroxyl of Thr 136 and the backbone amine of Ala 137 with the carboxyl group of the sialic acid (Fig. 5.5A). The alterations at position 149 could also be responsible for interfering with the interaction of Lys 145 with the 4' hydroxyl of the sialic acid (Fig. 5.5A).

5.2.4.3 Mutant viruses: HA residues 186, 189 and 208

Figure 5.6 shows the BLI receptor binding data for Iowa06 and Cal7 viruses binding to 6SLN-PAA with substitutions introduced at positions 186, 189 and 208. The substitution S186P when introduced into the Cal7 HA causes an increase in overall binding to 6SLN-PAA (decrease in relative K_d of ~10-fold). The opposite substitution introduced into Iowa06 (P186S) leads to a reduction in 6SLN-PAA receptor binding (increase in relative K_d of ~4-fold). The substitution A189T in Cal7 HA leads to a reduction in receptor binding affinity for 6SLN-PAA (increase in

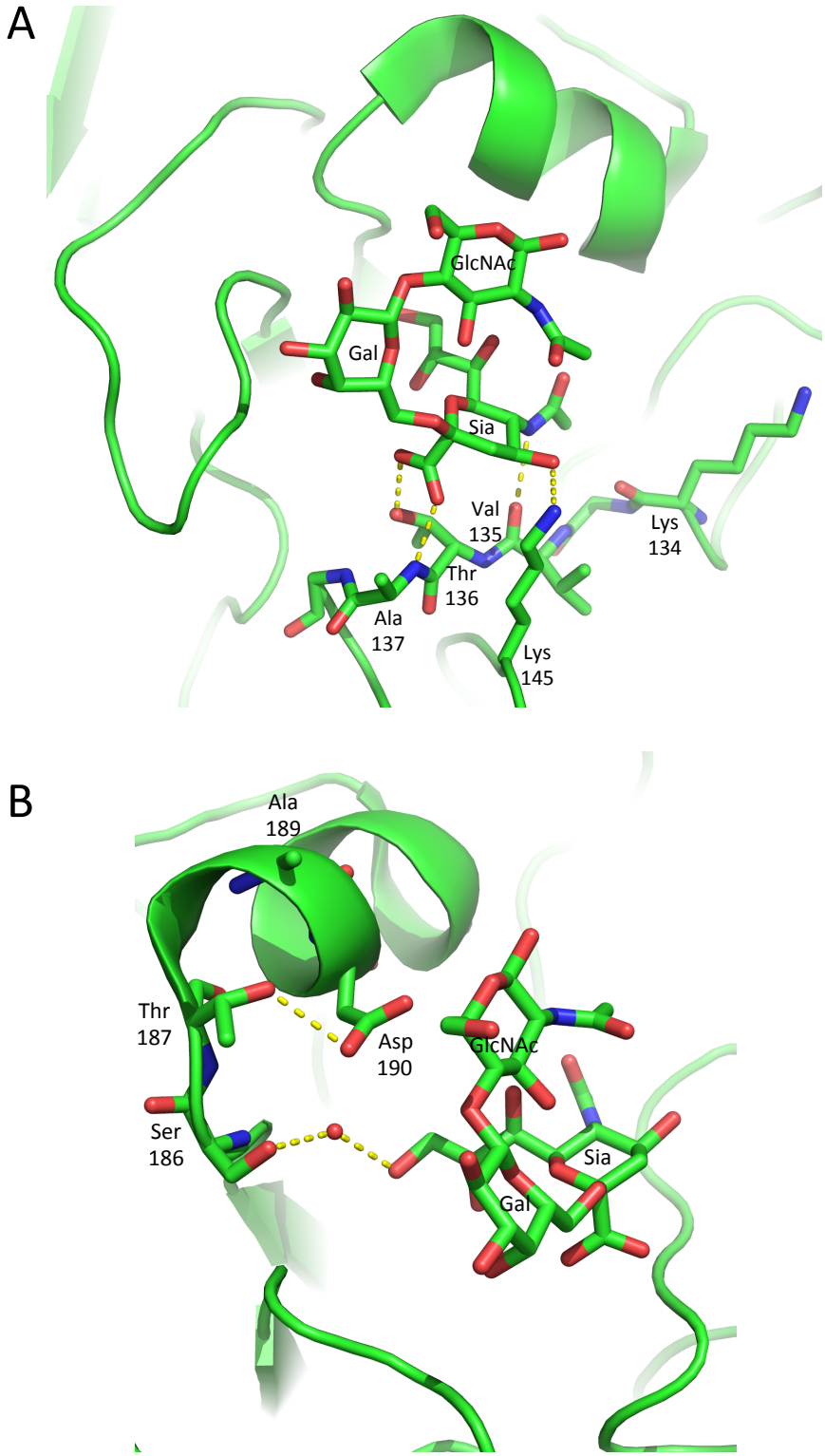


Figure 5.5: Structure of interactions of the 130-loop and 190-helix of A/California/4/2009 HA with 6SLN. HA structure of A/California/4/2009 HA bound to receptor 6SLN (PDB: 3UBN) (Xu *et al.*, 2012a) with interactions of A) 130-loop residues and B) 190-helix residues with the receptor. Figures made using PyMol (PyMol Molecular Graphics System, Schrödinger).

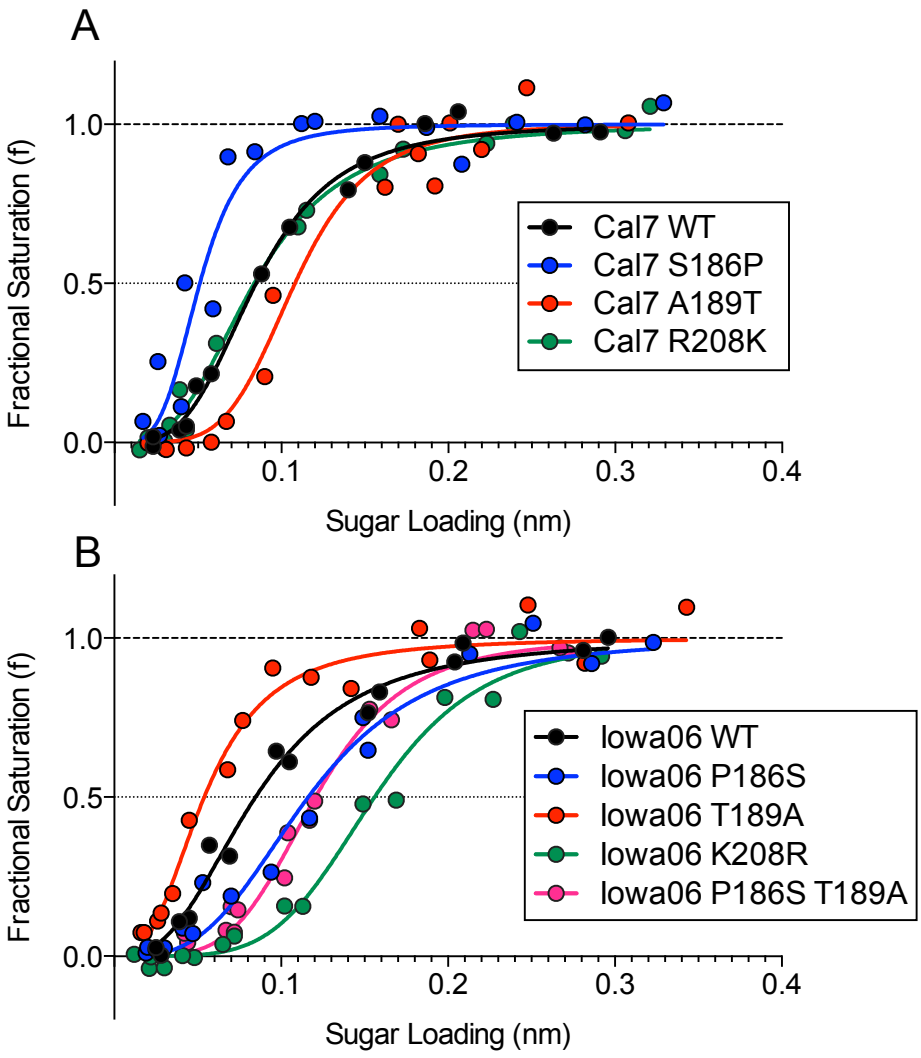


Figure 5.6: Equilibrium virus binding to 6SLN-PAA of Cal7 and Iowa06 viruses with HA residue substitutions at positions 186, 189 and 208. Fractional saturation of virus binding measured as a function of sugar loading in the presence of NA inhibitors.

relative K_d of ~3-fold). The substitution T189A introduced into Iowa06 HA causes an increase in 6SLN-PAA receptor binding (reduction in relative K_d of ~7.5-fold). The substitution R208K does not cause a significant alteration in binding characteristics to 6SLN-PAA when introduced into Cal7 HA; however, the opposite substitution (K208R) introduced into Iowa06 HA causes a large decrease in receptor binding to 6SLN-PAA, equivalent to an increase in the relative K_d of ~15-fold.

Residue 186, when a Ser, is involved in forming a hydrogen bond with the 9' hydroxyl of sialic acid through a water molecule, as can be seen in Figure 5.5B. The change of this Ser to Pro in Cal7 causes an increase in receptor binding affinity whereas the mutation of Pro to Ser in Iowa06 causes a decrease in receptor binding affinity, implying that for optimal receptor binding a Pro needs to be present at this site. It seems counterintuitive that the loss of a hydrogen bond should be accompanied by an increase in receptor binding affinity; however, the presence of a Pro will likely perturb the backbone arrangement due to the increased rigidity, owing to restricted movement around the amide bond. This could then interfere with the hydrogen bond between Thr187 and Asp190 and allow the Asp190 to bind to the acetamido nitrogen of the GlcNAc, forming a new hydrogen bond that would be stronger than the lost bond from Ser186, as it would be a direct hydrogen bond rather than via bound water, see Figure 5.5B.

Residue 189 needs to be an Ala for optimal 6SLN-PAA receptor binding characteristics. The A189T substitution reduces 6SLN-PAA binding in Cal7 and the T189A substitution increases receptor binding in Iowa06. The presence of a Thr189 could also be responsible for perturbing a hydrogen bond formed between Asp190 and the sialic acid. The linkage between positions 186 and 189 can be seen with the

Iowa06 double mutant P186S T189A that has a very similar binding profile to P186S alone (Fig. 5.6), indicating that the enhanced binding due to the loss of Thr189 is lost when the Ser186 is present.

The difference seen in 6SLN-PAA binding when the substitution K208R is introduced in Iowa06 HA is difficult to explain, as it is a considerable distance from the receptor binding site, see Figure 5.7. The location of the residue is at the interface between HA monomers. This residue could therefore be involved in altering interactions between subunits and possibly changing the positioning of the 220-loop, thus altering receptor binding properties. The reason that the opposite mutation (R208K) makes no difference to Cal7 receptor binding could be due to the nature of the other residues in the 220-loop, such as the alterations in 219 and 227 described below. Amino acid substitutions at this interface between monomers, residue 218, has been previously found to cause alterations in receptor binding in H3 HA (Daniels *et al.*, 1987).

5.2.4.4 Mutant viruses: HA residues 219 and 227

Residue 219 is polymorphic in classical swine H1 viruses. Early pdm09 viruses always have an Ile at that position and classical swine viruses being polymorphic between Ala and Thr. Substitutions I219T and I219A when introduced into Cal7 cause an increase in receptor binding to 6SLN-PAA (reductions in relative K_d of ~6.5-fold and ~3-fold respectively), see Figure 5.8A. The substitution A219I introduced into Iowa06 causes an increase in binding affinity to 6SLN-PAA (reduction in relative K_d of ~7-fold) whereas substitution A219T does not significantly alter the binding characteristics, see Figure 5.7C. The substitution E227A introduced into Cal7 HA causes an increase in binding affinity (reduction in relative K_d of ~14-fold); however,

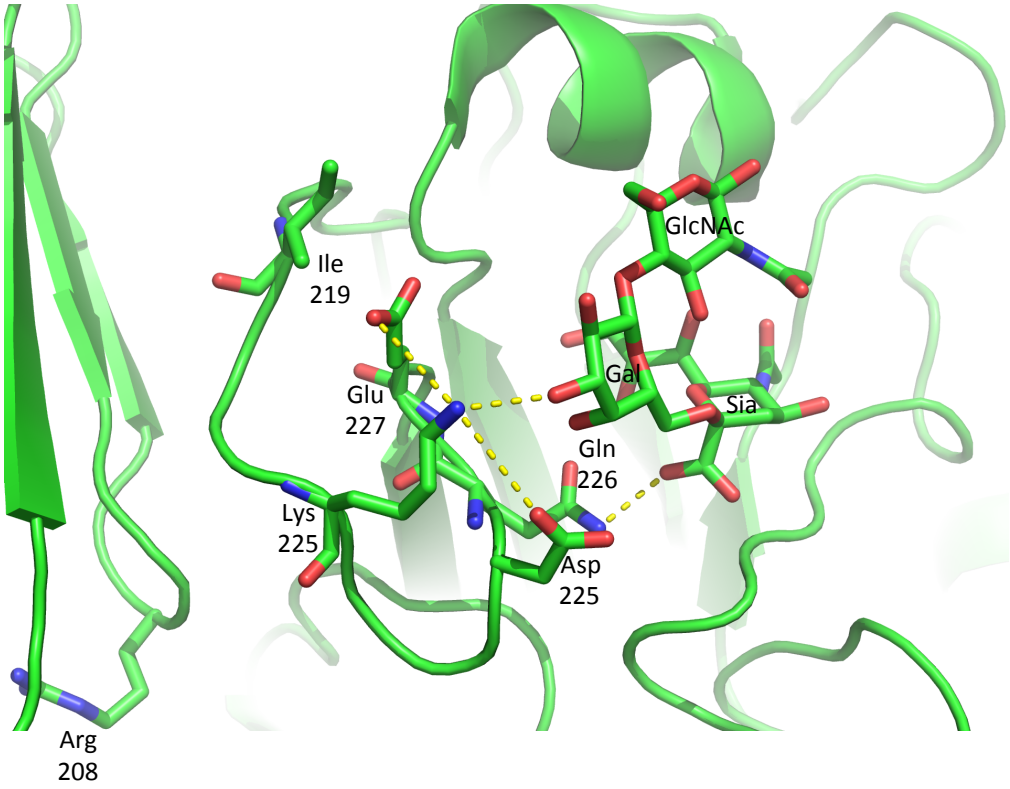


Figure 5.7: Structure of interactions of the 220-loop and 190-helix of A/California/4/2009 HA with 6SLN. HA structure of A/California/4/2009 HA bound to receptor 6SLN (PDB: 3UBN) (Xu *et al.*, 2012a) with interactions of the 220-loop with the receptor. Figures made using PyMol (PyMol Molecular Graphics System, Schrödinger).

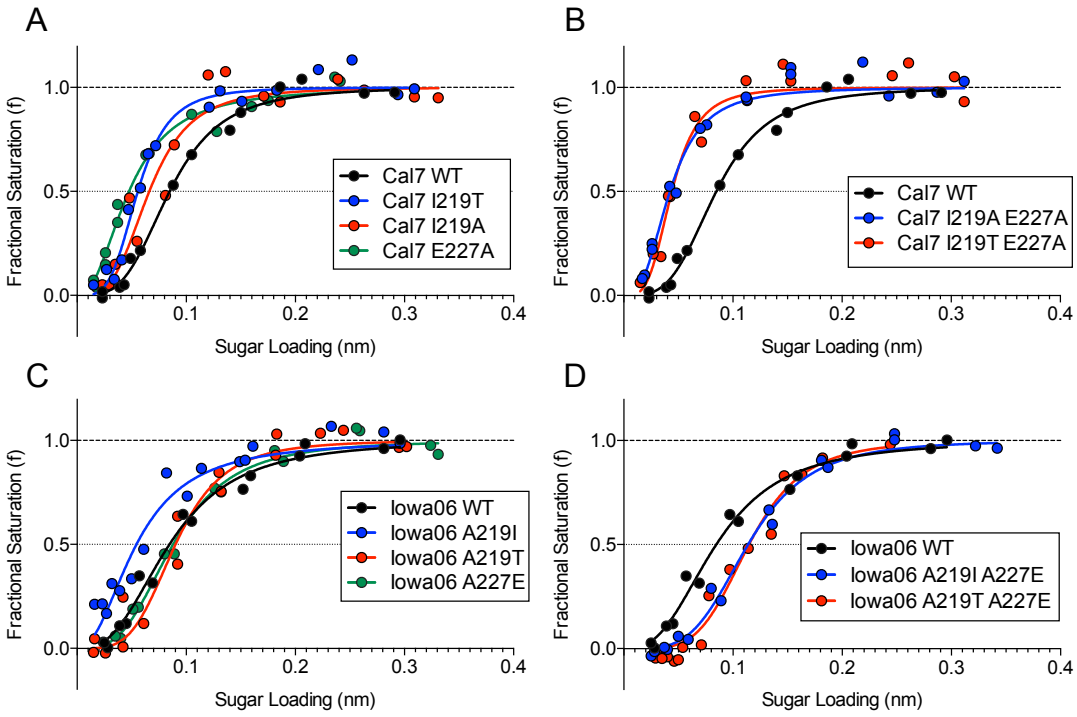


Figure 5.8: Equilibrium virus binding to 6SLN-PAA of Cal7 and Iowa06 viruses with HA residue substitutions at positions 219 and 227. Fractional saturation of virus binding measured as a function of sugar loading in the presence of NA inhibitors.

the opposite mutation (A227E) does not have a significant effect on the binding of Iowa06 (Figs. 5.8A and C).

Owing to the proximity of the 219 and 227 residues combination mutants were constructed with substitutions at both positions. The substitutions I219T E227A and I219A E227A increase binding to 6SLN-PAA when introduced into Cal7 HA (Fig. 5.8B). When either 219 substitution is combined with E227A, in Cal7 HA, there is an additive effect of enhancement of receptor binding to 6SLN-PAA (a decrease in relative K_d of ~25-fold for I219A E227A and ~20-fold for I219T E227A) (Fig. 5.8B). The substitution A219I in Iowa06 HA causes an increase in receptor binding to 6SLN-PAA; however, when combined with A227E there is a decrease in 6SLN-PAA receptor binding (increase in relative K_d of ~4-fold) (Figs. 5.8C and D). The substitution A219T in Iowa06 does not have a significant effect on 6SLN-PAA receptor binding when compared to wild type; however, when combined with A227E there is a reduction in receptor binding to 6SLN-PAA (increase in relative K_d of ~4-fold) (Fig. 5.8D).

Neither residue (219 or 227) is directly involved in interacting with the receptor. Residue 219 is situated behind the 220 loop (Fig. 5.7). The role of 219 may be that of forming a hydrophobic wall to ensure that the 227 residue is correctly located in order that it can coordinate Lys222 to ensure that it interacts correctly with the 3' hydroxyl of the galactose. No difference is seen in Iowa06 receptor binding with the alteration of Iowa06 HA from A227E unless 219 is an Ile or Thr upon which there is a reduction in binding to 6SLN-PAA compared to the WT. Cal7 behaves slightly differently with the E227A mutation causing an increase in receptor binding to a similar extent whether the 219 is an Ala, Ile or Thr. This difference could be due

to altered 220-loop coordination due to differences at 208, which altered Iowa06 6SLN-PAA binding.

5.2.5 H1N1 HA expression and purification

In order to study the binding of individual trimers Cal7 and Iowa06 HAs were expressed in insect cells using a recombinant baculovirus system. Expression constructs of Cal7 and Iowa06 HA ectodomain were synthesised and cloned into the HA expression plasmid pHAEM1, as described in *Materials and Methods*, section 2.2.6.12.2. The HA gene was synthesised for residues 1 – 502 (H1 numbering), without the 17 residue N-terminal signal peptide and the cytoplasmic tail. The pHAEM1 plasmid encodes a gp67 secretion signal peptide at the N-terminus, which is removed post translationally, and a C-terminal purification tag consisting of a TEV protease cleavage site, a foldon (T4 Fibritin) trimerisation motif and a 6 x His tag. The sequences of expression constructs are shown in the *Appendix*. Recombinant baculoviruses were generated containing these cloned genes using the Baculogold system, as described in *Materials and Methods*, section 2.2.7.1. These baculoviruses were used to infect bulk cultures of Sf9 cells for protein expression, as described in *Materials and Methods*, section 2.2.7.3. The expressed protein was concentrated by hollow fibre filtration and purified using a cobalt column as described in *Materials and Methods*, section 2.2.8.4. Figure 5.9A shows SDS-PAGE analysis of elution fractions from this his-tag purification. There was significant contamination in earlier elution fractions, visible by SDS-PAGE. Fractions 7 – 17, as numbered in Figure 5.9A, were pooled and concentrated. The typical purification procedure for recombinant HA then includes a digestion with TPCK-trypsin, followed by gel filtration. This

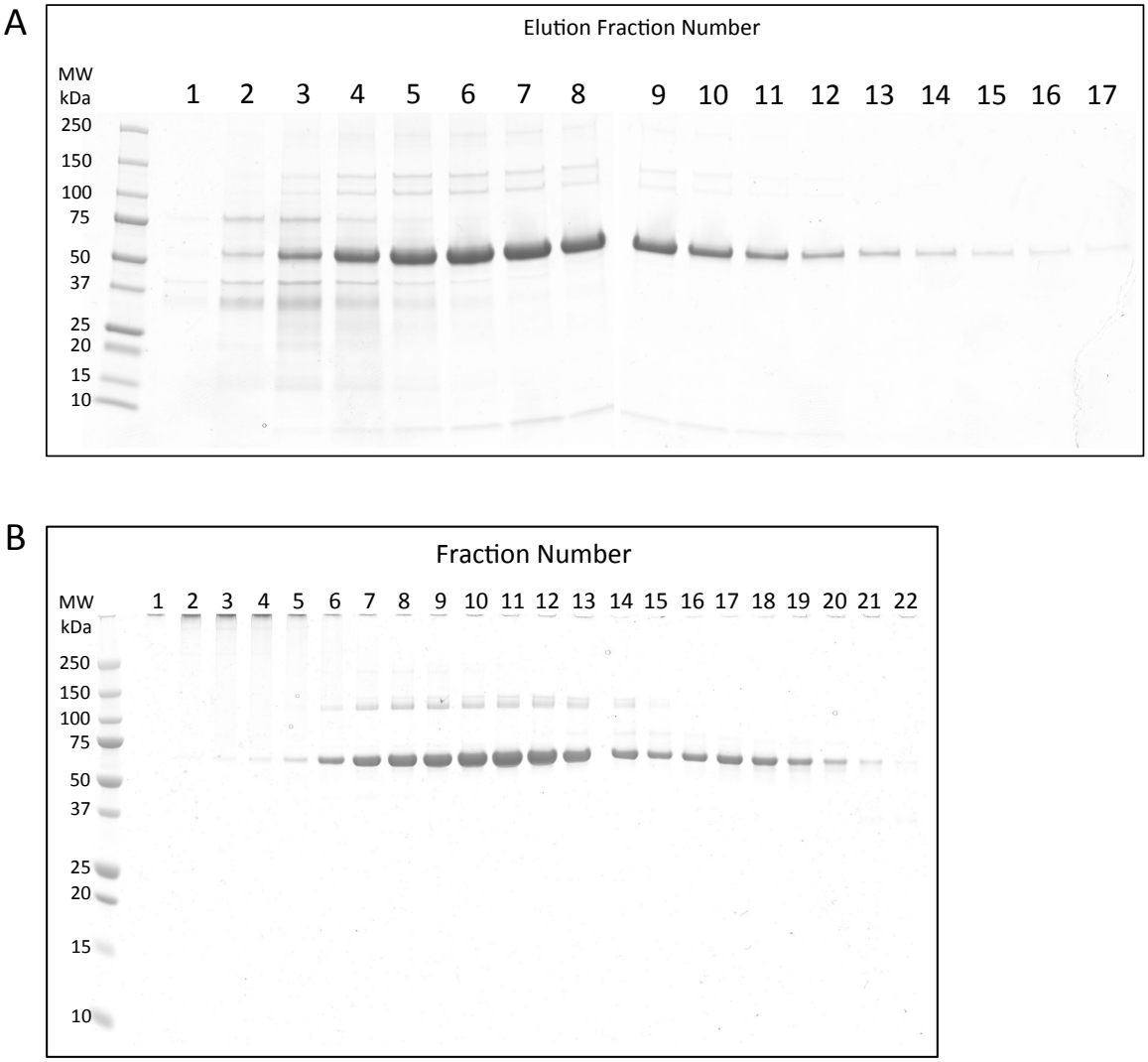


Figure 5.9: SDS-PAGE analysis of Cal7 HA purification. A) Elution fractions from cobalt column purification. B) Peak fractions of gel filtration purification. All gels were run under non-reducing conditions.

digestion cleaves the purified HA0, into HA1 and HA2, and simultaneously removes the purification-tag by cleaving at the TEV protease recognition sequence. Figure 5.10 shows SDS-PAGE analysis of trypsin and TEV protease digested Cal7 and Iowa06 HA as well as a positive control of an insect expressed H5 HA, kindly provided by Dr. Xiaoli Xiong. There is a small amount of cleaved HA, giving HA1 (~50 kDa) and HA2 (~25 kDa) bands in the undigested samples for the H1 and H5 proteins, most likely due to cleavage of the HA during expression by insect cell proteases. It is clear from this gel that both H1 proteins are digested into incorrectly sized products by trypsin, compared to the H5 control, suggesting protein instability. When the proteins are digested with TEV protease there is a slight reduction in the size of the HA0 due to the removal of the purification tag.

Owing to the instability of the H1 proteins to trypsin digestion protein purification was carried out by removing the purification tag with TEV protease followed by gel filtration. When this was carried out the protein eluted from the gel filtration column at a retention volume consistent with the protein being monomeric (~70ml) (data not shown). These results suggested that the instability of the protein was caused by removal of the trimerisation domain engineered into the purification tag. It was therefore decided to proceed with gel filtration purification of the uncleaved protein, without removal of the purification tag. Figure 5.9B shows SDS-PAGE analysis of elution fractions from this purification. There was a significant amount of large molecular mass impurities present in the earlier fractions (45 - 50 ml retention volume), visible in fractions 1 - 5 in Figure 5.9B. The major peak pertains to the trimeric form of the protein with a peak at ~57ml retention volume, fractions 6 – 15 as numbered in Figure 5.9B, which is the retention volume seen for trimeric HA

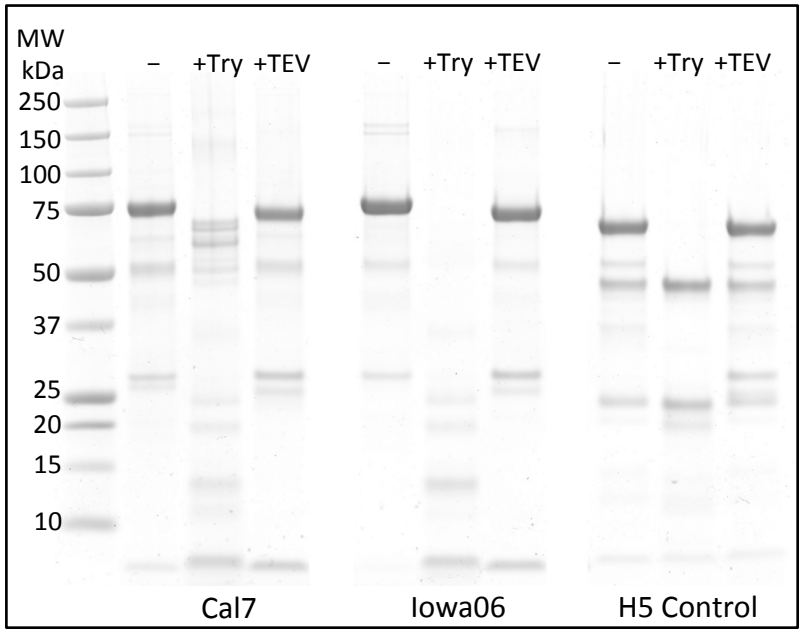


Figure 5.10: SDS-PAGE analysis of protease susceptibility of Cal7 and Iowa06 expressed HA. Proteins digested with both trypsin (Try) and TEV proteases. All samples run under reducing conditions.

based on previous experiments. As additional proof of the oligomeric nature of these fractions there are also multimeric HA bands present at higher molecular weights on the non reducing gel shown, at 100 – 150 kDa. There is also a minor peak at higher retention volumes (~70ml), which is a small amount of monomer present in the sample, fractions 16 – 22 Fig 5.9B. Fractions 8 – 14 were pooled and concentrated. Purification gels shown are for the purification of Cal7 HA; however, Iowa06 HA purified in a very similar manner.

5.2.6 Measurement of H1 binding by microscale thermophoresis

The monomeric binding of Cal7 and Iowa06 was measured by microscale thermophoresis (MST). This technique has been previously used to measure the affinity of purified HA for monomeric receptor analogues (Xiong *et al.*, 2013a). This method measures the movement of a constant concentration of a fluorescently labelled analyte in a temperature gradient in the presence of a range of concentrations of a binding partner.

The purified HA, generated as described in section 5.2.5, was labelled with an amine-reactive fluorescent dye, as described in *Materials and Methods*, section 2.2.13. This labeling was optimised in order to attach 1 – 2 dye molecules per HA trimer. MST was performed with the labelled protein at a constant concentration and using a range of different concentrations of 6SLN (0.05 – 30 mM), and the K_d values calculated as described in *Materials and Methods*, section 2.2.13. MST measurements were only made of binding of the HA to 6SLN. Experiments using 3SLN were omitted due to the lack of measurable binding of the viruses containing these HAs in BLI experiments.

Figure 5.11 shows the plotted data and Table 5.3 the plotted K_d values for the MST of Cal7 and Iowa06 HA binding to 6SLN. It is apparent from these MST data that the monomeric binding of both Cal7 and Iowa06 to 6SLN are very similar, as with the BLI data for receptor binding (Figure 5.3), giving a K_d value of ~3.5 mM. The data obtained are not of high quality, with significant variability, as the protein appeared to denature when greater temperature gradients were applied to the samples. The labeling of the reactive amines of the protein also appears to be problematic, as there is a Lys residue present in both the Cal7 and Iowa06 receptor binding sites at position 134. This Lys is also likely to be highly accessible to solvent and will consequently be one of the more easily labelled free lysine residues, possibly interfering with the measurements made. It is notable that the presence of a lysine in the vicinity of the receptor binding site (Lys 137) also caused similar issues for measuring MST binding parameters for an H10 HA (Vachieri *et al.*, 2014). These MST results reinforce the similarity of binding of Cal7 and Iowa06 NA to 6SLN similar to the results obtained for BLI (Fig. 5.3).

5.2.7 H1N1 NA purification

Purification of soluble NA from viruses containing the NA of Cal7 or Iowa06 was attempted in order to allow detailed kinetic analyses of these proteins. Viruses were constructed containing the NA of Iowa06 or Cal7 in a 7+1 reassortment with the lab-adapted WSN strain. Reassortment with this virus, rather than PR8, has been found to give higher yields of NA when extracted from purified virus (Patrick Collins, personal communication). These recombinant viruses were propagated in bulk using hens' eggs (~1500 eggs). Both viruses were then purified from the allantoic fluid, as described in *Materials and Methods*, section 2.2.3.1. Figure 5.12 shows SDS-PAGE

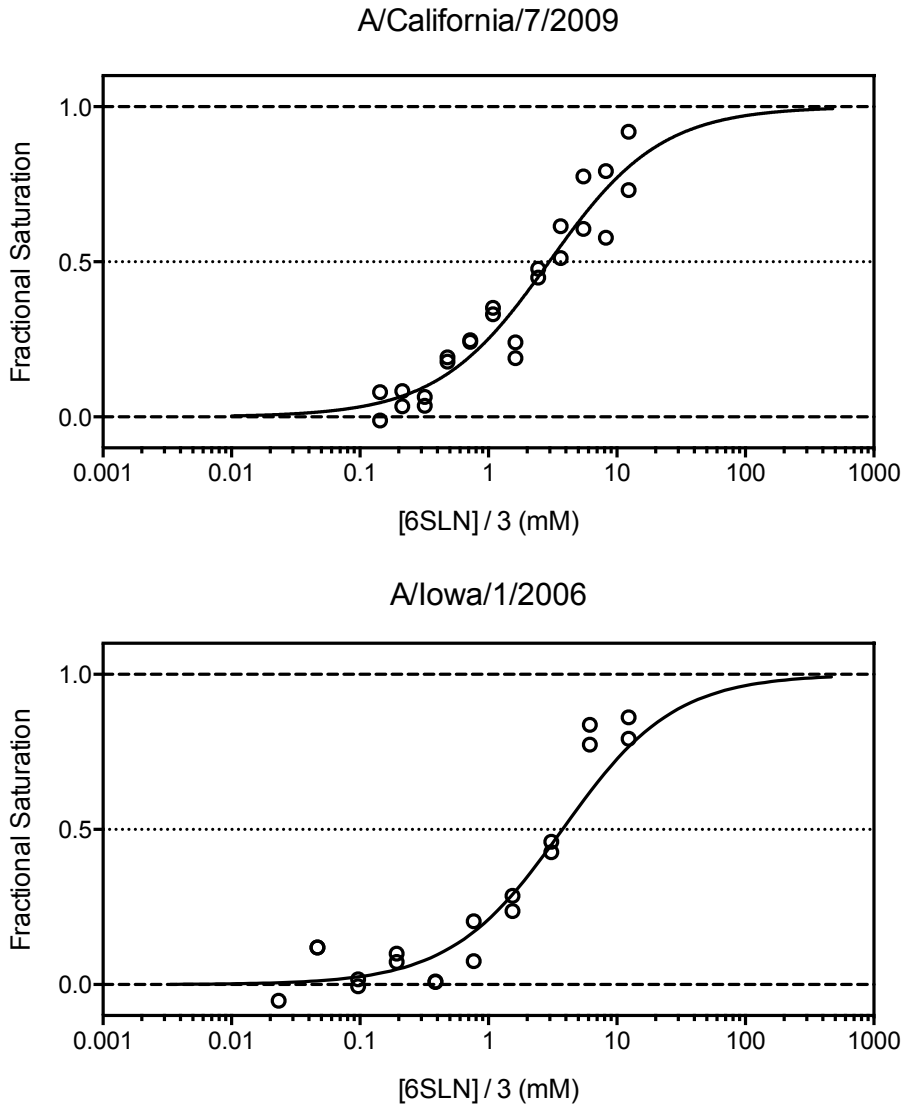


Figure 5.11: Microscale thermophoresis data for Cal7 and Iowa06 HA binding to 6SLN. Data obtained for monomeric 6SLN binding to insect cell expressed protein. Data shown are for duplicate experiments. Fit to determine K_d shown. 6SLN concentrations are presented as [6SLN]/3 to correct K_d. See *Materials and Methods*, section 2.2.13 for more details and (Xiong *et al.*, 2013a).

Protein	Corrected K_d ± S.E.* (mM)
Cal7 HA	3.3 ± 0.8
Iowa06 HA	3.8 ± 0.9

* S.E.: Standard Error of the mean

Table 5.3: Fitted kinetic parameters of MST experiments measuring the binding of Cal7 and Iowa06 HA to 6SLN.

analysis of these purified viruses. Both Cal7 and Iowa06 NA containing viruses have a band corresponding to NA, running at ~70 kDa.

Purification of NA from the virus was attempted using a number of different methods. The first method was digestion with the protease bromelain, followed by removal of the virus core by ultracentrifugation, as described in *Materials and Methods*, section 2.2.8.3. The Cal7 NA was successfully cleaved from the virus using bromelain, as had been previously done with pdm09 NA (van der Vries *et al.*, 2012), whereas the Iowa06 virus showed no significant cleavage of NA from the virus even at high bromelain concentrations. Figure 5.13 shows SDS-PAGE of the bromelain digestion of the viruses containing Cal7 or Iowa06 NA. In Figure 5.13A the cleavage of NA from the Cal7 virus leads to two bands in the supernatant, which are both NA. This was shown after gel filtration experiments described below where the two products were partially separated by gel filtration (Fig. 5.14). Fractions containing a majority of the larger or smaller product had a similar NA activity proportional to the amount of protein present in the fraction (data not shown). These two bands seen by SDS-PAGE are likely to be different bromelain cleavage products. One is at ~70 kDa, similar in size to that seen in the undigested virus NA, and the other is ~40 kDa. They are likely to have different amounts of the NA stalk removed, with the 70 kDa band corresponding to soluble protein with the majority of the stalk still attached and the smaller 40 kDa band with more of the stalk removed. The two bands are of a similar intensity. Also present in the supernatant is some HA2 (~25 kDa) but no visible HA1, which should run at ~50 kDa.

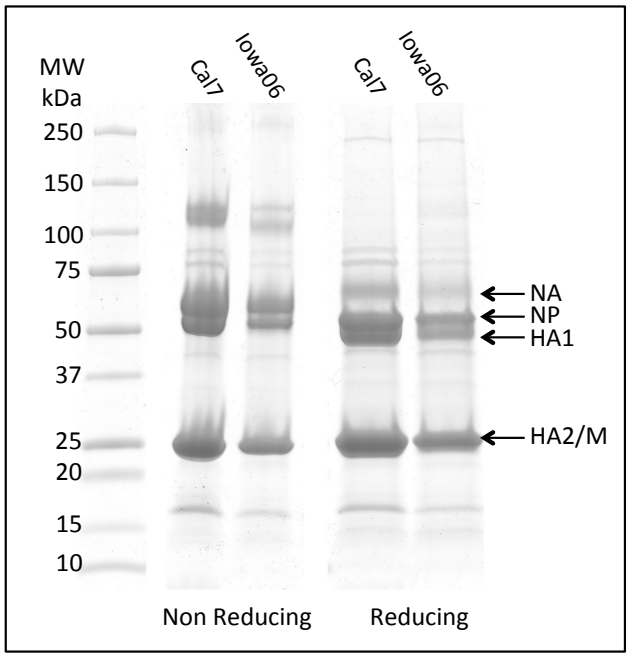


Figure 5.12: SDS-PAGE of 7+1 reassortant viruses used for Cal7 and Iowa06 NA purification. Viruses contain Cal7 or Iowa06 NA with the remaining 7 segments from WSN. Viruses were propagated in hens' eggs and purified through sucrose gradients. Samples were analysed under both reducing and non-reducing conditions.

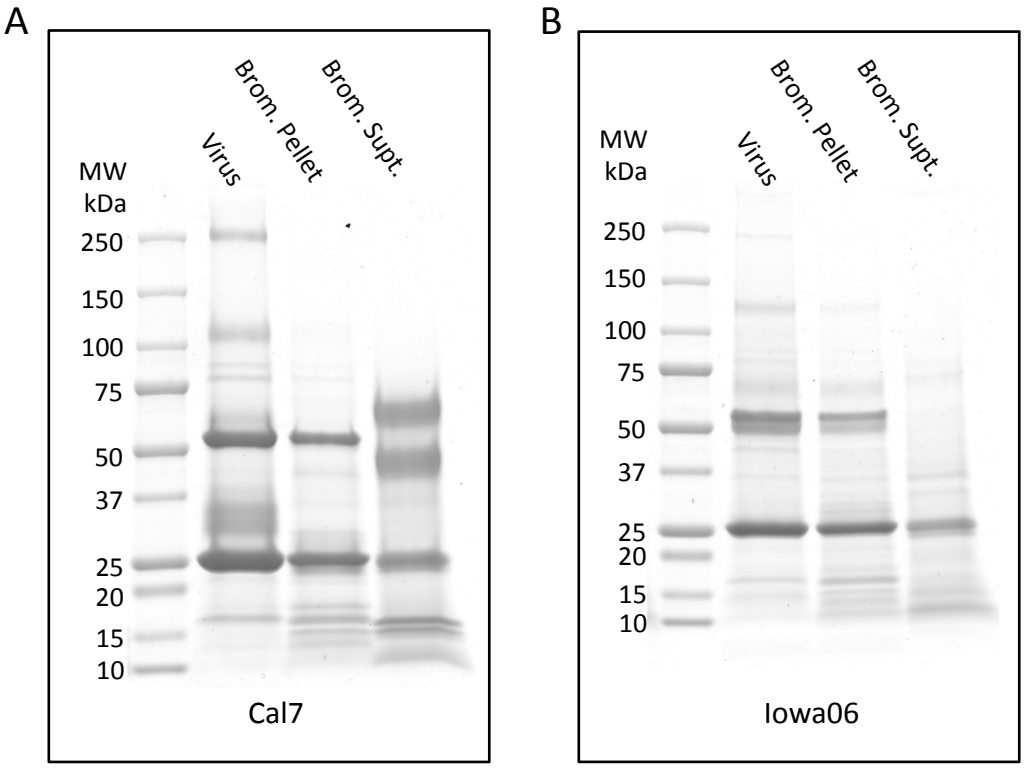


Figure 5.13: SDS-PAGE analysis of bromelain digestion of Cal7 and Iowa06 containing purified viruses. Bromelain cleavage of A) WSN + Cal7 NA and B) WSN + Iowa06 NA. Lanes analysing the bromelain cleavage supernatant (Brom. Supt) contain 5 x the volume of sample when compared to pellet (Brom. Pellet) and virus lanes. All samples run under reducing conditions.

The Cal7 NA extracted by bromelain digestion was further purified by anion exchange chromatography to remove contaminating HA, as described in *Materials and Methods*, section 2.2.8.3. The column flow through was retained, which contained the NA. This flow through was concentrated and purified further by gel filtration. Figure 5.14 shows SDS-PAGE analysis of the peak fractions. The protein eluted in a broad single peak ranging from ~55 – 70 ml. The gel filtration partially separated the two bromelain digestion products of NA seen previously (Fig. 5.13A). The larger ~70 kDa form of the NA formed what appears to be a dimer, which runs at ~120 kDa when the SDS-PAGE is performed under non-reducing conditions. It was concluded that the cleavage of the NA into these different products is unlikely to affect kinetic measurements. Fractions 2 – 18 were pooled and concentrated. Aliquots of the concentrated NA were snap frozen in liquid nitrogen and stored at -80°C.

The Iowa06 NA could not be removed from the virus by bromelain digestion (Fig. 5.13B). It was therefore decided to extract the virus glycoproteins using detergent and then protease digest the NA ectodomain from the cytoplasmic tail and stalk. Detergent extraction of the viral glycoproteins was attempted in a number of different detergents and the most efficient detergent extraction was found by extracting the proteins in 2% Brij-36T. This condition did not extract any M1 or NP into the ultracentrifugation supernatant, monitored by SDS-PAGE (data not shown). The HA was removed from the detergent extracted protein by anion exchange chromatography in the presence of detergent. The flowthrough fraction was applied to a gel filtration column to remove contaminating lipid and also to buffer exchange to a lower concentration of detergent (0.5% Brij-36T). The fractions containing NA were pooled. Trypsin and bromelain cleavage of the detergent solubilised NA was

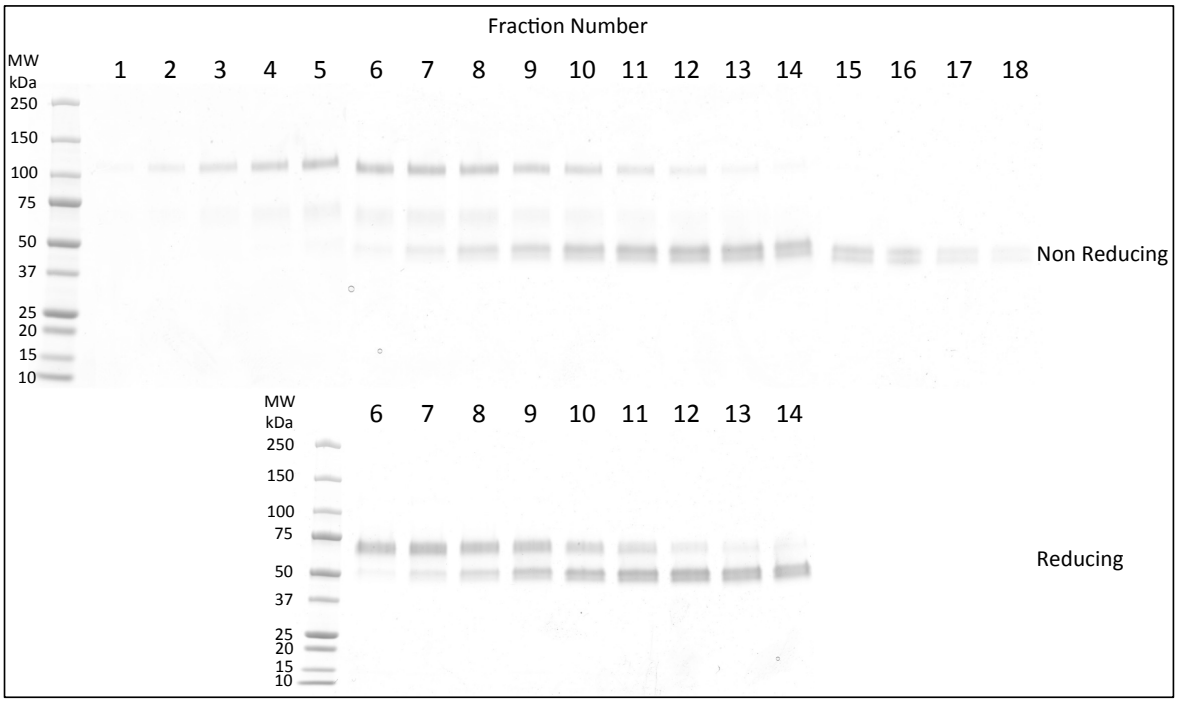


Figure 5.14: SDS-PAGE analysis of Cal7 NA gel filtration purification. Flow through fraction of ion exchange of bromelain cleaved WSN + Cal7 NA purified by gel filtration. Samples of fractions run under both non-reducing and reducing conditions.

attempted under a number of different conditions, varying ratios of proteases, different salt concentrations and concentrations of β-mercaptoethanol (for bromelain digestions); however, cleavage was unsuccessful. The detergent extracted protein was converted into rosettes by removal of the detergent by hydrophobic adsorbent beads, as described in *Materials and Methods*, section 2.2.8.3. Protease cleavage of the NA rosettes was also attempted under a range of different conditions and protein cleavage could not be achieved. It was decided to use the protein rosettes directly for measuring NA kinetics; however, when carrying out preliminary experiments the NA rosettes had largely lost NA activity and the activity degraded further upon brief incubation at 37°C. It was decided to measure the NA kinetics for Iowa06 only using purified virus.

5.2.8 H1N1 NA kinetic characterisation

NA kinetic characterisation was carried out both Cal7 NA and Iowa06 NA. The Cal7 NA was purified from virus, as described above in section 5.2.7. This purified protein allows the determination of k_{cat}, as the protein concentration can be determined. The Iowa06 NA could not be purified from the virus whilst still retaining activity comparable to when it was attached to the virus. Purified virus was used for kinetic characterisation. The purified virus used was a 7 + 1 reassortant of WSN with Iowa06 NA. The concentration of the NA present in the purified virus cannot be easily measured and consequently the k_{cat} value cannot be determined; however. the V_{max} value can be determined, which allows the comparison of measurements made with the same virus concentration under different reaction conditions.

5.2.8.1 MUNANA kinetics

Kinetic measurements were made for Cal7 and Iowa06 NA cleaving the fluorogenic substrate MUNANA, as described in *Materials and Methods*, section 2.2.11.1, and section 3.2.7.1. Measurements were made both in MES buffer (32.5 mM MES-NaOH, pH 6.5, 4 mM $CaCl_2$) and HBS buffer (10 mM HEPES-NaOH pH 7.4, 150 mM NaCl, 0.005% Tween-20, 4 mM $CaCl_2$) and at two different temperatures, 25°C and 37°C. Figure 5.15 shows Michaelis-Menten plots of MUNANA cleavage determined for Cal7 and Iowa06 NA and shown in Table 5.4 are the fitted kinetic parameters.

Cal7 NA has low K_m and high k_{cat} values and consequently high k_{cat}/K_m values for MUNANA cleavage under all conditions tested, indicating high enzymatic efficiency. The K_m and k_{cat} values determined in MES buffer are similar to those measured for X-31 (Table 3.1), indicating a highly active NA at cleaving MUNANA. There is, however, a difference seen when comparing the two buffer conditions between Cal7 and X-31 NA. Both Cal7 and X-31 NA show a 2 – 3 fold increase in K_m when measurements are made in the HBS buffer, however X-31 has an increased k_{cat} in HBS, indicating an increased turnover rate, whereas Cal7 NA has similar k_{cat} values under both buffer conditions. This indicates a possible difference in pH dependency of NA activity, although the difference between the two buffers is not simply pH, as there are also differences in the concentration of NaCl.

Iowa06 NA also has low K_m values, indicating high affinity binding of MUNANA by this NA. The K_m values measured in MES buffer are ~1.5-fold higher, under both temperature conditions, when compared to those measured for Cal7, indicating slightly weaker binding of MUNANA; however, when measured in HBS the Iowa06 K_m

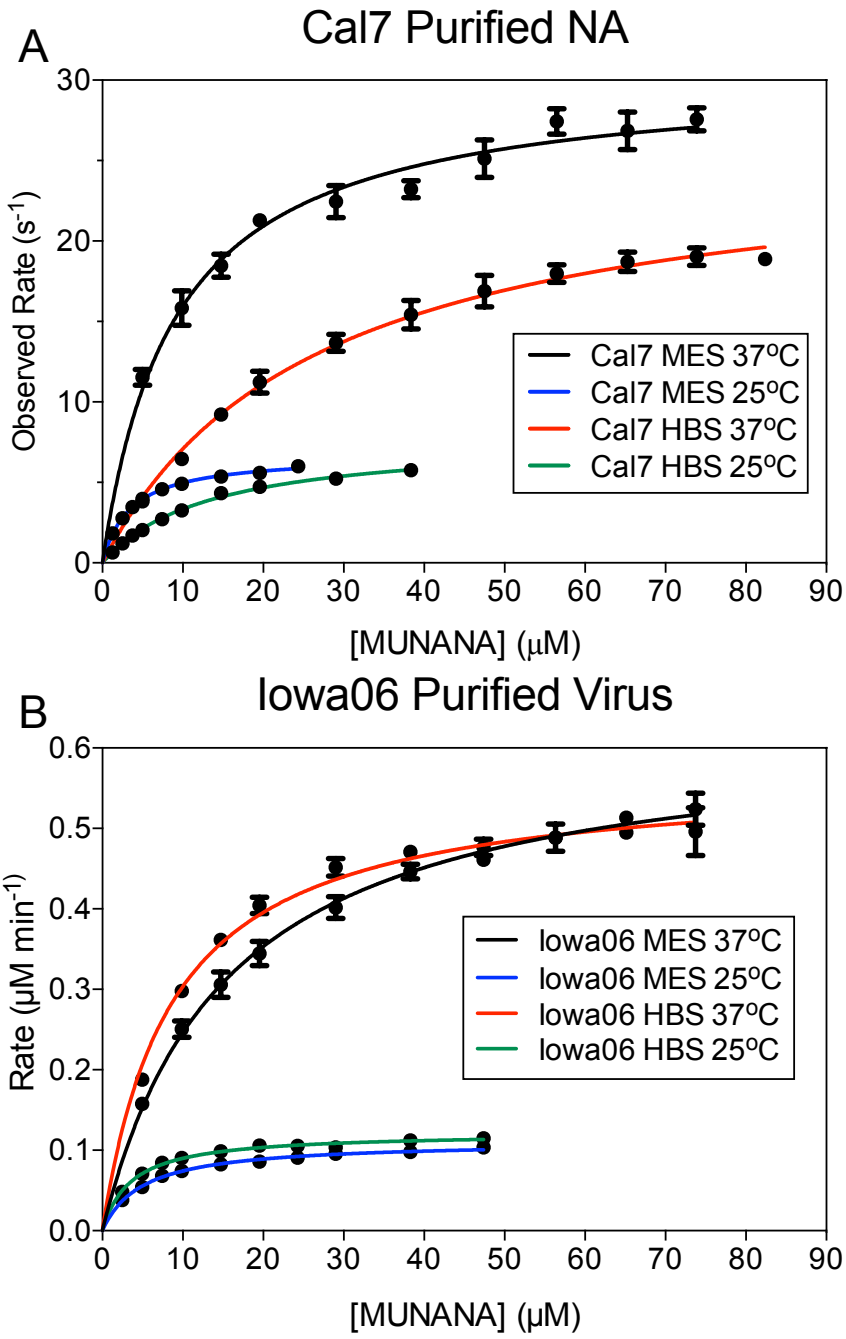


Figure 5.15: Michaelis-Menten plots of Cal7 and Iowa06 NA cleavage of MUNANA. Measurements made using: A) Cal7 purified NA, converting obtained rate values to observed rate by dividing by the rate by the concentration of NA used; B) 7+1 reassortant of WSN + Iowa06 NA purified virus, data could not be converted to observed rate, as the concentration of NA present is not known. All graphs show least squares fit of Michaelis-Menten equation.

Protein	Substrate	Buffer/ Temp	K_m ± S.E. [§] (μM)	K_{cat} ± S.E. (s^{-1})	V_{max} ± S.E. (μM min^{-1})	k_{cat}/K_m ($\mu M^{-1}s^{-1}$)
Cal7 Purified NA	MUNANA	MES* 37°C	9.0 ± 0.6	30.4 ± 0.5	-	3.38
		MES 25°C	3.5 ± 0.1	6.7 ± 0.1	-	1.91
		HBS† 37°C	26.6 ± 1.6	26.0 ± 0.6	-	0.97
		HBS 25°C	13.3 ± 0.7	7.8 ± 0.2	-	0.59
	6SLN	HBS 37°C	600 ± 45	-	-	-
	3SLN	HBS 37°C	42.8 ± 0.6	-	-	-
Iowa Purified Virus	MUNANA	MES 37°C	15.4 ± 0.2	-	0.625 ± 0.007	-
		MES 25°C	5.0 ± 0.3	-	0.111 ± 0.001	-
		HBS 37°C	8.7 ± 0.5	-	0.568 ± 0.007	-
		HBS 25°C	3.6 ± 0.2	-	0.122 ± 0.002	-
	6SLN	HBS 37°C	463 ± 14	-	-	-
	3SLN	HBS 37°C	105 ± 5	-	-	-

*MES MES Buffer (32.5 mM MES-NaOH, pH 6.5, 4 mM $CaCl_2$)
†HBS HBS-P + $CaCl_2$ Buffer (10 mM HEPES-NaOH, pH 7.4, 150 mM NaCl, 0.005% Tween-20 + 4 mM $CaCl_2$)
§S.E. Standard Error of the mean
- Value not determined

Table 5.4: Steady state NA kinetic Parameters for Cal7 and Iowa06 NA.

values are ~3-fold lower, indicating a more efficient binding of MUNANA. There are no k_{cat} data for Iowa 06 NA. These cannot be determined because the measurements were made using purified virus and the NA concentration is not known. Consequently, no comparison of overall catalytic efficiency can be made for the Cal7 and Iowa06 NAs. The V_{max} values determined for Iowa 06 NA show a similar pattern to those seen with the k_{cat} values determined for Cal7 with similar values when the experiments were carried out at 37°C when compared to 25°C, similar to the k_{cat} values determined for X-31 and Cal7.

5.2.8.2 6SLN and 3SLN kinetics

Owing to the lack of available purified Iowa06 NA it was decided not to carry out the detailed measurements of NA cleavage of 6SLN and 3SLN, as described in section 3.2.7.2, as the k_{cat} for Iowa06 NA could not be determined. The K_m of 6SLN and 3SLN was determined by measuring the K_i of 6SLN or 3SLN for inhibition of MUNANA cleavage. By measuring the uninhibited and inhibited rate, at a set concentration of 6/3SLN, the K_i can be determined, as described in *Materials and Methods*, section 2.2.11.1.2. The K_i of MUNANA inhibition by 6SLN or 3SLN is directly equivalent to the K_m, as it is the concentration required to half saturate the sialidase site. These K_i values for MUNANA inhibition were found to be similar to measurements of K_m made using the enzyme-coupled assay, previously described in section 3.2.7.2.

The K_m values determined for 6SLN and 3SLN are shown in Table 5.4. Both Cal7 and Iowa06 NA exhibit a relatively low K_m for 6SLN (Cal7 = 600 ± 45 μM, Iowa06 = 463 ± 14 μM) when compared to X-31 (8070 ± 615 μM), with both N1 proteins exhibiting a K_m >10-fold lower, indicating a much higher affinity binding of 6SLN. Iowa06 NA has a K_m for 6SLN which is ~1.3-fold lower than Cal7 indicating a slightly

higher affinity. The K_m values for 3SLN are also lower (Cal7 = 42.8 ± 0.6 μM, Iowa06 = 105 ± 5 μM) when compared to X-31 (562.3 ± 20 μM), with both N1 proteins giving K_m values >5-fold lower. Cal7 NA has a K_m value which is ~2-fold lower than that measured for Iowa06 indicating enhanced binding to 3SLN.

The K_m data show that the N1 NAs studied have high binding affinities for 6SLN and 3SLN compared to X-31; however, this information is not enough to fully understand the NA characteristics, as there is no information on the relative rates of enzymatic turnover.

5.2.9 Measurement of HA and NA balance

The emergence of the pdm09 viruses was accompanied with a novel combination of HA and NA, consequently there is likely to be an impact on the balance between the activities of these two glycoproteins. Assays were carried out to measure the activities of the different proteins in altering the interaction of virus with a receptor-coated surface, using the system previously developed in section 3.2.2.

Virus interaction studies carried out in the presence of NA inhibitors, described in section 5.2.4, were performed using viruses with the full 8-segment Cal7 virus; however, Iowa06 viruses were a 6 + 2 reassortant using the HA and NA from Iowa06 and the background from PR8. It is known that the background of a virus can alter the morphology and also the proportions of HA and NA present on the virus surface and consequently the virus used for HA/NA balance studies was the full 8-segment Iowa06 virus. Viruses based on the Iowa06 background were not used previously for studies of viruses with mutant HAs, as they have much more inefficient propagation characteristics. Reassortant viruses were constructed by generating

viruses with the HA and NA of Cal7 and Iowa06 in both the Cal7 and Iowa06 backgrounds by reverse genetics. A virus with Cal7 HA and Iowa06 NA rescued and propagated well with both Cal7 and Iowa06 genetic backgrounds and both have very similar equilibrium binding characteristics, when compared to WT Cal7 and Iowa06 viruses, measured in the presence of NA inhibitors, shown in Figure 5.16. Reassortant viruses with Iowa06 HA and Cal7 NA rescued and propagated well; however, when the binding of these viruses was measured they had unusual kinetics of binding, which were very slow compared to the WT and the other reassortant viruses (data not shown). These viruses were not used further because of these unusual characteristics.

Figure 5.17 shows the HA and NA balance measurements for viruses binding to biosensor surface saturated with 6SLN in the presence and absence of NA inhibitors. The measurements made of the wild type viruses made in the absence of NA inhibitors show a number of differences. Iowa06 WT has more extensive virus binding, with a peak virus binding of ~0.4 whereas Cal7 has a lower peak of virus binding of ~0.25. This indicates that the Cal7 virus has either more powerful NA activity for cleaving 6SLN or that the particular HA of the virus causes shorter residence times. The reduction in binding seen after the peak virus binding shows a similar shape for both Cal7 and Iowa06 viruses; however, the Iowa06 virus appears to have a higher rate of decline in binding after the turning point. The level of virus remaining on the sensor at the end of the measurement is higher for Iowa06 despite this enhanced rate of virus dissociation after the peak of virus binding.

When examining the measurements made for the reassortant virus with Cal7 HA and Iowa06 NA with a Cal7 background the initial peak of binding is similar to

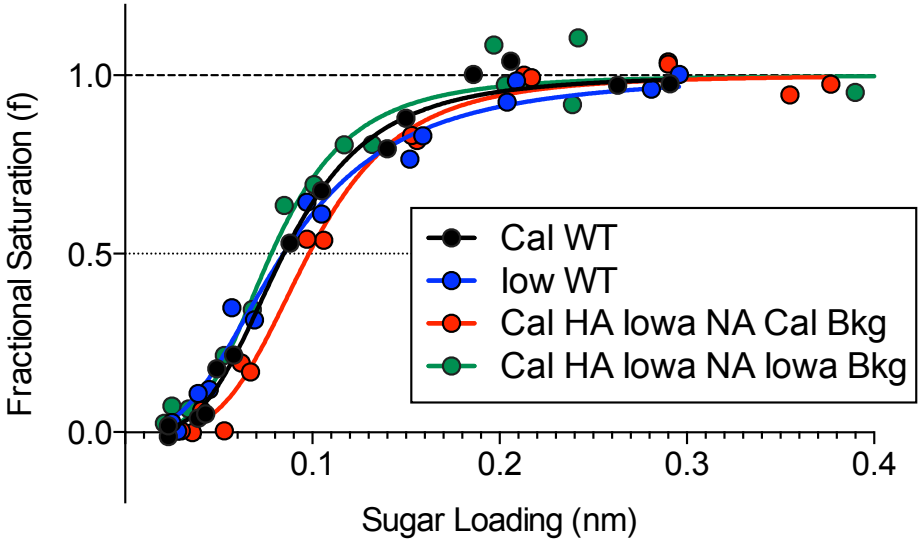


Figure 5.16: Comparison of equilibrium virus binding to 6SLN-PAA of Cal7, Iowa06 and reassortant viruses. Fractional saturation of 100 pM virus binding measured for a constant concentration of virus as a function of sugar loading. All measurements made in the presence of NA inhibitors.

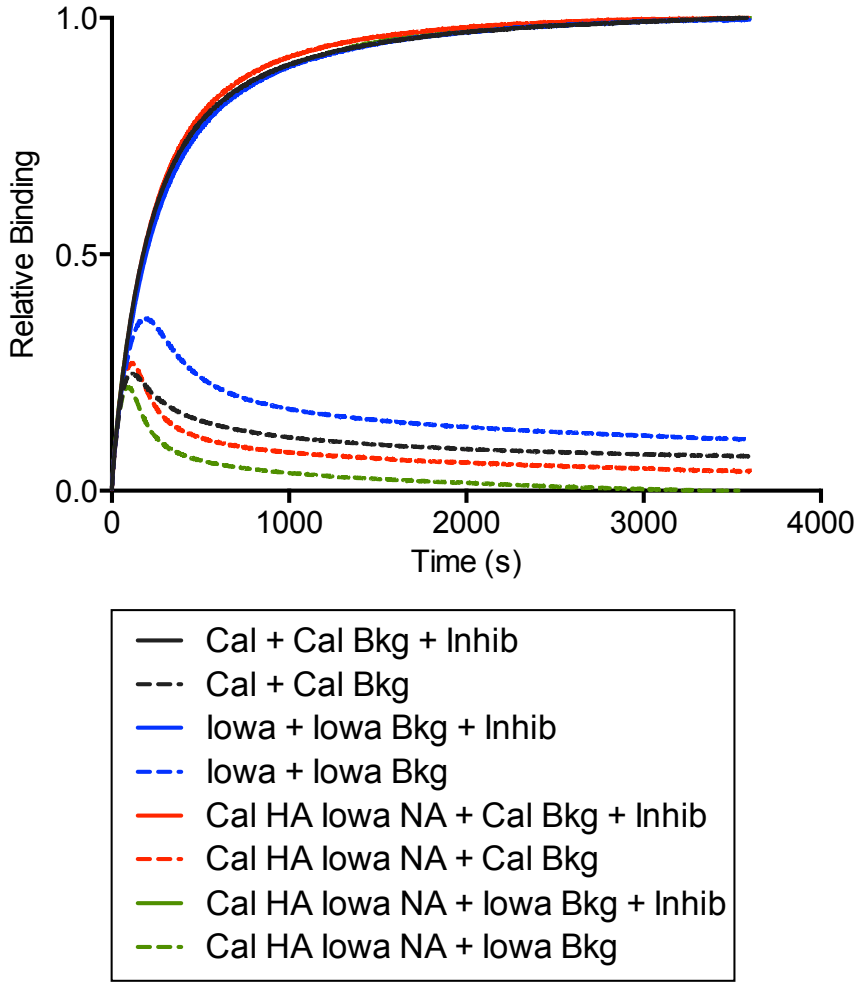


Figure 5.17: HA/NA balance binding data for Cal7, Iowa06 and reassortant viruses binding to 6SLN-PAA. Measurements made of 100 pM virus in the presence and absence of NA inhibitors (Inhib).

that seen with the Cal7 WT virus indicating similar levels of initial sugar depletion and virus residence times; however, the binding decays to a lower level than that seen with the full Cal7 virus suggesting that the reduction in binding at later time points is due to the substitution of the NA from Cal7 for that of Iowa06. This virus also appears to have the same curve shape of that seen with the full Iowa06 WT virus with a higher rate of virus binding decay, suggesting that this higher rate of virus dissociation seen is due to the presence of the Iowa06 NA within the virus. The same reassortant virus constructed with the Iowa06 genetic background has different binding characteristics from those seen with the Cal7 background. The peak of virus binding is lower and the end point of virus binding is much lower in the presence of the Iowa06 background. This implies that the background can alter the virus binding characteristics. This is likely to be due either to a change in virus morphology or the amount or proportions of HA and NA present on the virus particle.

5.2.10 Measurement of virus size by nanoparticle tracking analysis

The previous data presented regarding HA and NA balance of the studied H1N1 viruses suggests that the background of the reassortant HA and NA viruses were responsible for a difference in virus binding characteristics. One factor that could be responsible for these differences could be virus morphology, which is mainly determined by the M1 protein, as described in *Introduction***,** section 1.1.4.1.

Nanoparticle tracking analysis (NTA) was carried out on the virus used for the receptor binding assays. This technique records a video of the light scattered by particles in a suspension illuminated by a laser. Software tracks the movement of particles over time and based on the speed of movement the diffusion coefficient and

therefore the particle size can be determined. The method used is described in *Materials and Methods*, section 2.2.14.

The particle size profile for the viruses used in HA/NA balance experiments is very similar, with a mode peak particle size of ~160 nm and a mean particle size of ~180 nm, Figure 5.18. The distribution is not Gaussian with a skew towards higher particle sizes. The similarity of these data suggest that virus morphology is not responsible for the differences seen in the balance measurements for viruses with different backgrounds presented in the previous section. These virus-sizing measurements do have a number of caveats due to the analysis methods and the virus preparation techniques used. The virus used for these measurements and previous HA/NA balance measurements are simply pelleted from the tissue culture supernatant and used without further purification. This form of NTA relies only on visualising particles based on their ability to scatter light without any way of knowing if the particles measured are virus or some other particle of the same size, as any other particles will also be biological material and will therefore scatter light in a similar manner to virus. A portion of the particles measured could be exosomes, reviewed in (Théry, Zitvogel & Amigorena, 2002).

5.3 Discussion

Receptor binding information for WT Cal7 and Iowa06 viruses described in section 5.2.4 shows that both viruses possess high affinity for 6SLN-PAA but only very low affinity for 3SLN-PAA. This is in broad agreement with a number of studies carried out using microarrays or other solid-phase binding assays, which examine

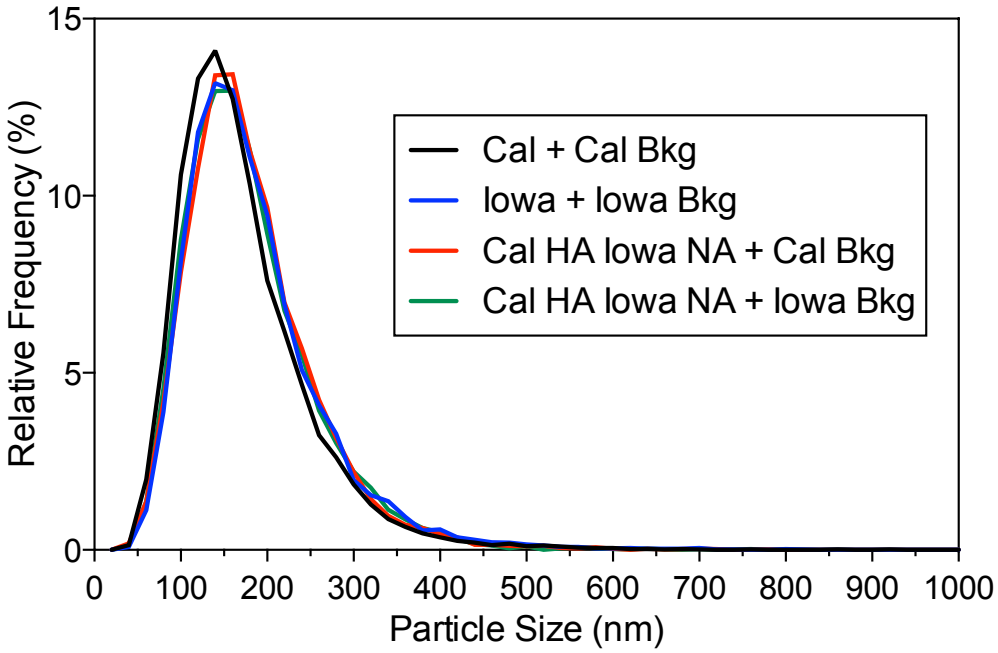


Figure 5.18: Nanoparticle tracking analysis of Cal7, Iowa06 and reassortant viruses. Measurements made of 10 000 – 20 000 independent particle tracks. Particle size made into a frequency distribution using a bin size of 20 nm.

either pdm09 or TRS H1 binding (Childs *et al.*, 2009; Maines *et al.*, 2009; Yang, Carney & Stevens, 2010; Chen *et al.*, 2011; Yen *et al.*, 2011; Bradley *et al.*, 2011; de Vries *et al.*, 2011; Xu *et al.*, 2012a; 2012b).

Somewhat different results were obtained when binding of pdm09 and TRS H1 containing viruses was measured using a sialylated glycan microarray. Childs *et al.* used very similar viruses to those used in this chapter, using two pdm09 viruses: Cal4 (A/California/4/2009), and A/Hamburg/5/2009, which are both very similar to Cal7 and also the same TRS Iowa06 virus, as well as the historical csw A/New Jersey/8/76. When measuring virus binding to sialylated glycan microarrays the results show a similar binding pattern for both the pdm09 and TRS/csw viruses; however, virus binding was detectable for a number of the α2,3 linked receptors in the array, which is not seen in the experiments presented in this chapter (Childs *et al.*, 2009). Later studies by Chen *et al.*, Yen *et al.* and Bradley *et al.* examined the binding of a range of other pdm09 H1N1 viruses and TRS H1 containing viruses by glycan microarray. All of these papers show a similar pattern to that seen in the earlier work by Childs *et al.*, where the early pdm09 isolates show a major preference for α2,6-linked sialic acid. There is, however little or no evidence for binding to α2,3-linked sugars. The contemporary TRS H1 containing viruses appear to bind to a wider range of α2,6-linked receptors when compared to the pdm09 viruses, but there is still little measurable binding to α2,3-linked receptors (Chen *et al.*, 2011; Yen *et al.*, 2011; Bradley *et al.*, 2011).

There have been two studies carried out measuring the binding of antibody complexed expressed H1 HA from TRS and pdm09 sources. de Vries *et al.* show a much-reduced binding of pdm09 HA compared to csw HA to immobilised fetuin. This

is also combined with glycan microarray data, which show binding only to α2,6-linked receptors with the same pattern of reduced binding of pdm09 HA when compared to the csw HA. There was also no measurable binding to any α2,3-linked sialic acid receptors in the microarrays for either pdm09 or csw HAs (de Vries *et al.*, 2011). Xu *et al.* also show a reduced binding of pdm09 virus HA when compared to that of csw HA measured both by solid phase binding assay to the receptor analogue 6SLNLN and also glycan microarrays (Xu *et al.*, 2012b).

The results presented in this chapter for the binding of Cal7 and Iowa06 WT viruses (Fig. 5.3) agree more with those measured for virus binding, (Childs *et al.*, 2009; Chen *et al.*, 2011; Yen *et al.*, 2011; Bradley *et al.*, 2011) rather than purified protein binding (de Vries *et al.*, 2011; Xu *et al.*, 2012a; 2012b). These previous virus binding experiments give results that imply similar binding characteristics for both csw H1 containing and early pdm09 viruses, rather than the large difference in receptor binding characteristics seen when measured with purified protein (de Vries *et al.*, 2011; Xu *et al.*, 2012a; 2012b).

The difference seen when measuring purified HA and virus binding could be explained by a number of different factors. The purified HAs are expressed as a soluble ectodomain, similar to those expressed for this chapter. These soluble ectodomains are then complexed with two different antibodies giving a maximum of 4 HA trimers present in each complex, which as well as being a different presentation to that of a virus, with different flexibility between proteins, could also provide insufficient cooperativity, thus lowering the sensitivity of the assays. The soluble H1 HAs under study in this chapter also appear to be highly unstable, as can be seen in the trypsin susceptibility assays (Fig. 5.10). This instability of the proteins could be

responsible for altering the receptor binding characteristics of the proteins seen when the HA is not stabilised by being attached to a virus. There could be a difference in stability between the TRS and pdm09 HAs, as there are a significant number of differences in the protein sequence in areas which could be responsible for altering subunit interactions such as those at the interface of the subunits both in HA1 and HA2 subunits. The presentation of the HA on the viruses is clearly the most biologically relevant form and consequently the results obtained for whole virus binding are likely to be more representative of virus-cell interactions.

de Vries *et al.* use unaltered HA proteins for their binding studies but they do not provide any details about the relative stability of the different proteins used (de Vries *et al.*, 2011). Xu *et al.* adopt a different approach where they engineer a disulphide bond into the HA1 between residues 205 and 220 in order to stabilise the HAs under study (Xu *et al.*, 2012b). Although this could partially alleviate the problems of protein stability it could also alter the receptor binding characteristics of the proteins studied by altering the flexibility of the 220 loop. This disulphide is also very close to residue 208, which was found to alter receptor binding characteristics of Iowa06 in this chapter (Fig. 5.5B).

The receptor binding studies of mutant Cal7 and Iowa06 viruses presented in this chapter, section 5.2.4, demonstrate that substituting a number of residues can have an effect on the 6SLN-PAA receptor binding properties of the H1N1 viruses under study. Differences in 6SLN-PAA binding were found when altering residues 132, 134, 149, 186, 189, 208, 219 and 227. The combination of the differences between Cal7 and Iowa06 imply that the various residues that are changed, which individually do alter receptor binding, when combined result in viruses with similar receptor

binding properties, suggesting a compensatory network of mutations present, which leaves overall virus binding characteristics unchanged.

The two studies detailed above which analyse the binding of different expressed H1 proteins with respect to pdm09 emergence both suggest a number of mutations responsible for the alterations in receptor binding that they measure for wild-type proteins. de Vries *et al.* suggest that the double substitution T200A and A227E is responsible for the difference in receptor binding that they measure. The reported T200A substitution is a peculiarity of the particular pdm09 strain used for those studies, Cal4, and is not present in Cal7 or the vast majority of other early pdm09 strains. This T200A substitution is the only difference between Cal7 and Cal4 and consequently the panel of mutants made by de Vries *et al.* bearing this T200A substitution in combination with other mutations is directly comparable to the results obtained for Cal7 mutant viruses constructed in this chapter. The only mutation present in combination with the T200A, reported by de Vries et al., that gives a difference in binding assays to fetuin is A227E, which increases the receptor binding of the pdm09 HA, which is similar to the results for Cal7 binding to 6SLN-PAA presented in this chapter (Fig. 5.7) (de Vries *et al.*, 2011). There are two other mutations which are made in the pdm09 virus HA by de Vries *et al.* which are in common with the studies presented above: K149R and I219T. Neither of these mutations in combination with T200A show a difference in pdm09 HA receptor binding (de Vries *et al.*, 2011). This does not agree with results presented in section 5.2.4 which show that these mutations introduced into Cal7 HA alter receptor binding. The pre-pandemic swine HA used in de Vries *et al.*, A/swine/Ohio/891/01 (H1N2) does not contain a number of the substitutions seen in Iowa06 when compared to

pdm09 viruses, including substitutions at 134, 186, 189 208 and 219. The differences in the swine HA used when compared to Iowa06 (R134K, P186S, T189A, K208R and T219A, see Figure 5.2) one could consider likely to decrease the overall virus binding, as the 186 and 189 mutations in combination and the 208 mutation alone would cause a large decrease in 6SLN-PAA binding, decreasing the overall binding of the 'wild-type' TRS HA (Fig. 5.6) rather than the increase in binding reported by de Vries *et al.* (de Vries *et al.*, 2011).

Xu *et al.* examine the binding of Cal4 HA and mutants thereof as well as the pre-pandemic swine HA from A/swine/Indiana/P12439/2000 (H1N2) and associated mutations. The Cal4 differs from Cal7 at position 200 as detailed above. The swine virus HA used in this previous study varies from the Iowa06 HA at a number of positions, carrying the following substitutions which are located in close proximity to the receptor binding site: K134R, G158E, P186S, T189A and K208R (H3 Numbering) (Fig. 5.2). This large number of mutations makes it difficult to compare the swine virus used with Iowa06. The majority of the substitutions present in the swine virus used by Xu *et al.* are typical of other isolated strains. The G158E substitution is, however, atypical in csw H1 proteins. This substitution has been selected in experiments carried out when viruses were passaged in an inappropriate cell line, see section 5.2.3. This mutation also appears in isolated pdm09 viruses, particularly when they are passaged in MDCK-SIAT cells as well as the substitutions K157E and N159E. The selection of this mutant in tissue culture implies that it is a receptor binding alteration and consequently likely to perturb the HA binding results of this previous study. Xu *et al.* attribute the reduction of binding that they see in pdm09 HA to the following mutations: K134R, I219A and E227A, which increase

binding to 6SLNLN and make the binding pattern more like that of a TRS H1. All of the mutations listed above were reported to give an increase of binding when introduced alone and in combination gave the greatest increase in receptor binding. The results from Xu *et al.* regarding I219A and E227A introduced into a pdm09 HA both alone and in combination agree with the results of this chapter, with both mutations increasing the binding of Cal7 to 6SLN-PAA (Fig 5.8); however, the substitution K134R did not cause a difference in binding of Cal7 (Fig. 5.4). Xu *et al.* also report that the increased binding that they measure for the TRS H1 can be reduced to pdm09 levels by introducing the inverse mutations (R134K, A219I and A227E). The results presented in this chapter show, as with Cal7, substitution at position 134 (K134R) does not affect 6SLN-PAA binding of Iowa06. The A227E substitution does not affect 6SLN-PAA binding of Iowa06 when introduced alone. The A219I mutation increases Iowa06 binding when introduced alone, however this increase in binding is reversed to a reduction of binding when combined with A227E, agreeing in part with the results of Xu *et al.* (Xu *et al.*, 2012b).

The previous studies discussed above (de Vries *et al.*, 2011; Xu *et al.*, 2012b) attempt to explain the receptor binding of pandemic emergence by a small number of differences in receptor binding residues and both present a combination of a small number of substitutions which give receptor binding profiles similar to the pdm09 or TRS H1 that they have mutated their proteins towards; however, the mutagenesis studies presented in this chapter demonstrate that it is unlikely to be so simple as two or three substitutions, rather that there is an incomplete picture of other receptor binding changes present with a lack of understanding of the interrelated nature of the residues of the receptor binding site.

The characterisation of the NA from pdm09 and/or TRS has been carried out in a number of different studies (Yen *et al.*, 2011; van der Vries *et al.*, 2012; Xu *et al.*, 2012b; Gerlach *et al.*, 2012; Blumenkrantz *et al.*, 2013; Abed *et al.*, 2014; Campbell *et al.*, 2014). The results described in this chapter of the K_m for MUNANA cleavage by Cal7 NA are similar to a number of other studies of pdm09 NAs, which all gave a low K_m (Yen *et al.*, 2011; van der Vries *et al.*, 2012; Blumenkrantz *et al.*, 2013; Abed *et al.*, 2014; Campbell *et al.*, 2014). The measured low K_m value for Iowa06 NA cleavage of MUNANA is also similar to a reported value for a TRS N1 NA (Yen *et al.*, 2011). Xu *et al.*, however, obtained kinetic values for purified pdm09 NA which give both a high K_m and low k_{cat} (Xu *et al.*, 2012b) when compared to the results presented in this chapter for Cal7 NA, see Table 5.4, suggesting a partially inactive protein.

The K_m values for 6SLN and 3SLN of Cal7 and Iowa06 presented in this chapter, Table 5.4, indicate that both NAs bind the 3SLN more efficiently than the 6SLN, likely indicating a greater efficiency of cleavage of 3SLN. The K_m values obtained are similar to values determined for both pdm09 NA and TRS NA cleaving 6SL and 3SL (Yen *et al.*, 2011). Yen *et al.* also measures a lower relative V_{max} for TRS N1 NA when compared to pdm09 NA. There are two other studies that compare the overall efficiency of cleavage by measuring sialic acid removal from a microarray (Xu *et al.*, 2012b) or an end point measure of released sialic acid from a single fixed concentration of substrate (Gerlach *et al.*, 2012). Both of these studies demonstrate an overall preference for α2,3-linked sialic acid and a lower overall activity of TRS NA, as is found in experiments presented in section 5.2.8.

Measurements made for HA/NA balance presented in this chapter, section 5.2.9, demonstrate that there is a difference in overall virus binding to 6SLN-PAA

between Cal7 and Iowa06 in the absence of NA inhibitors, indicating an overall lower binding capacity of pdm09 viruses. The binding capacity appears to be reduced further when the NA from Iowa06 is introduced (in the same Cal7 background). This appears counterintuitive, as the WT Iowa06 virus has a higher binding level than Cal7, with a very similar binding of the HA in the presence of NA inhibitors, indicating lower overall NA activity. This difference in binding of the reassortant virus could be due to two different factors: 1) different characteristics of k_{on} and k_{off} between Cal7 and Iowa06 HAs, which could lead to similar equilibrium binding (K_d) characteristics but alter the residence times of the virus; and 2) the amounts and proportions of proteins on the virus surface could be altered due to the different virus background, possibly caused by the different M protein.

The involvement of the M protein in the binding characteristics of the virus appears to be likely owing to the different binding characteristics seen for the reassortant (Cal7 HA Iowa NA) virus when rescued with an Iowa06 background. The inclusion of the Iowa background leads to a further reduction in virus binding indicating either greater levels of NA activity or a reduction in HA mediated residence time; however, without further reassortment studies and biochemical data on the levels of protein present on the different reassortant viruses it is not possible to draw firm conclusions as to the relative protein contents and the importance of the gene segments in altering the HA/NA balance data obtained. It does, however, not appear to be a matter of altered virus morphology, which is very similar for all viruses (Fig. 5.18).

The involvement of the M protein in the pdm09 virus emergence has been studied previously. Campbell *et al.* showed that introducing a pdm09 M protein into

PR8 increases the amount of NA present and also leads to more viruses with an extended filamentous morphology (Campbell *et al.*, 2014). This study, however, does not include any information on the activity of a TRS M protein in altering virus characteristics. A number of studies have found that the presence of both the M and NA from pdm09 viruses enhances virus transmission in guinea pigs (Chou *et al.*, 2011), ferrets (Lakdawala *et al.*, 2011) and pigs (Ma *et al.*, 2012). These differences found in virus transmissibility, in a range of hosts upon altering NA and M segments, are not possible to relate to the biophysical data obtained in this chapter.

Xu *et al.* described the HA and NA balance of TRS and pdm09 viruses by plotting relative 6SLNLN affinity of the HA against the catalytic efficiency of the NA (k_{cat}/K_m) at cleaving MUNANA and drawing the conclusion that there is a required correlation of relative HA and NA characteristics in order for efficient human infection and transmissibility that the TRS viruses appear not to have (Xu *et al.*, 2012b). This technique of relating HA and NA activity appears to be over simplistic, as results obtained in this chapter demonstrate that: 1) the HA binding is very similar between pdm09 and TRS viruses, Fig. 5.3; 2) the NA characteristics at cleaving MUNANA do not appear to be significantly different for pdm09 and TRS NAs, Table 5.4; 3) MUNANA cleavage by NA is likely not to be a good model for predicting the NA activity on cellular glycans, as the MUNANA cleavage kinetic parameters vary from those for natural substrates, 6SLN and 3SLN, in an unpredictable fashion, see Tables 3.1, 4.2 and 5.4; 4) HA/NA balance measurements of the WT Cal7 and Iowa06 viruses are not very different from each other, Fig. 5.17; and 5) the genetic background of the virus could be important in altering the HA/NA balance characteristics, as seen when looking at the reassortant (Cal HA Iowa NA) viruses with either Cal7 or Iowa06

backgrounds, Figure 5.17**,** which is non seen experimentally when simply looking at purified proteins.

The emergence of the 2009 H1N1 pandemic appears to be a multifactorial process of gene drift and shift, with the HA undergoing a large number of substitutions, which do not dramatically alter the equilibrium binding characteristics of the virus. The new NA obtained by reassortment appears to have slightly increased catalytic efficiency. The genetic background, likely to be due to substitution of M segments by reassortment, also affects the relative HA and NA balance. The importance of the HA mutations and NA and M alterations in changing the biochemical properties of the virus and therefore transmissibility remains unclear.

6 Final Discussion

One of the major aims of influenza virus research is to be able to predict the next pandemic strain. The major factor underlying pandemic emergence, apart from immune novelty, is the need for the virus to have the high fitness required for efficient transmission. The process of virus transmission is complex and involves a number of different factors. Currently the methods for understanding IAV transmissibility rely on a number of different *in vivo* and *in vitro* methods, which all have their own advantages and disadvantages and when used in combination provide the most complete picture of transmissibility. A major factor underlying the viral fitness and therefore transmissibility is the initial stage of IAV attachment to the cell surface, which is controlled by the two surface glycoproteins HA and NA.

This thesis presents a new *in vitro* biophysical method for investigating the relationship between the antagonistic activities of the two IAV surface glycoproteins. This technique has been implemented to examine a number of different viruses both in order to understand fundamental characteristics of virus attachment and release from a receptor coated surface and to be able to investigate a number of viruses, with a view to relating binding characteristics to virus fitness and transmissibility.

The binding of H7N9 viruses was examined, particularly with a view to examining the role of the NA in receptor binding. These experiments demonstrate that the N9 NA present in these viruses enhances receptor binding to both α2,6- and α2,3-linked sialic acid receptors via the secondary Hb site. An initial enhancement of virus receptor binding to α2,6-linked receptors was also seen when the Hb site was present and the NA uninhibited, which could be explained by the kinetic characteristics of the sialidase site that favour α2,6-linked sialic acid binding at a

higher affinity accompanied with a lower cleavage probability upon substrate binding, turning the sialidase site into more of a receptor binding site. The binding characteristics of a number of H1N1 viruses were also examined in order to attempt to understand the differences in virus binding which could be responsible for the emergence of the 2009 H1N1 pandemic. The effect of a number of conserved amino acid substitutions between the pre-pandemic TRS H1 HAs and the pdm09 H1 HAs on receptor binding was examined as well as the novel combination of HA and NA seen upon the emergence of this novel virus lineage.

IAV receptor binding is a complex multivalent interaction that is often viewed as a rather static process of virus attaching irreversibly to a cell surface, with the interaction resulting in virus internalisation and initiation of infection. The reality is much more complex. The interaction of a virus with a surface will require a number of interactions of an HA subunit with single sialic acids to allow a residence time long enough to allow virus internalisation. The role of the NA, if any, in altering initial virus attachment is unclear. Although there is extensive evidence that adequate NA activity is required for virus release and disaggregation, there is little evidence that too much viral NA activity at the stage of virus attachment is detrimental to infectivity. This is most likely a matter of virus concentration. Virus infection is typically initiated by a small number of particles, which can attach to a large area of cell surface, with the viruses at ~100 nm in size and the cell ~10 – 50 μm in diameter. The probability that a small number of virus particles could deplete enough sugar from such a large cell surface and thereby have a major effect on HA binding is almost certainly low. A more active NA may reduce the residence time of individual viruses, but it appears unlikely that this would truly impinge on the internalisation of the virus, other than by slightly

extending the attachment stage. It appears that the role of the NA would be most important for virus release where virus will be at a much higher concentration. It would not be unreasonable for localised virus concentration upon release to reach well into nM levels (1 nM ≈ 0.6 virus/μm^3). There will also be a quantity of unincorporated NA present in the apical cell membrane at the site of budding which will also help virus release. The lack of importance of too much NA activity compared to the HA binding characteristics can be seen when looking at recent H3N2 seasonal viruses which have very low binding to α2,6-linked receptors and maintain high NA sialidase activity and yet they remain epidemic strains, indicating sufficient virus fitness and transmissibility.

The only example of reduced NA activity being a selective advantage is the NA stalk deletion that occurs upon the adaptation of viruses from wild birds to domestic poultry. Experiments carried out in this thesis have shown that a stalk deletion reduces effective NA activity (section 3.2.10). This reduction in NA activity must be necessary for enhancing virus attachment, as it does not appear that reduced NA activity would ever be an advantage for virus release. There appear to be only two plausible explanations for this evolutionary advantage of reduced NA activity: 1) the density of the receptors on the cell surface in domestic poultry may be low and consequently the increased NA activity dramatically impinges on the virus attaching for enough time to initiate infection; or 2) the number of virus particles present at the site of infection in domestic poultry could be higher than that for wild bird and mammalian infection, therefore depleting the sugar at a greater rate and reducing overall receptor binding.

The understanding of HA and NA balance in controlling cell infection becomes far more complicated when the NA also possesses receptor binding characteristics, as seen for some H3N2 viruses and also as shown for the N9 from H7N9 viruses in this thesis (Chapter 4). The N2 binding via the sialidase site is somewhat simpler than for the N9 owing to the absence of the complication of the binding contribution from the Hb site. For the N9 with the sialidase site acting as a receptor binding protein for α2,6-linked receptors and with the Hb site also enhancing receptor binding and substrate cleavage simultaneously it makes it difficult to dissect the relative contributions of these various roles of the NA and their energetic contributions to the binding characteristics of the virus. In some respects, however, it is irrelevant how the different proteins contribute to the binding and/or release characteristics of the virus, as it is only the combination of these activities that will alter the infectious characteristics of the virus.

The big question that arises from these experiments measuring HA and NA balance is: what is the combination of HA and NA activities required for optimal transmissibility between humans and therefore the emergence of a human pandemic virus? One way to examine this question would be to use the techniques described in this thesis to examine the HA and NA balance characteristics of previous viruses that have made a successful jump from animal to human reservoirs and which therefore formed pandemic strains. Of the four pandemics, for which virus samples are available, only the 1918 H1N1 and 2009 H1N1 pandemic viruses could be relied upon to be representative of the circulating pandemic strains of the time. The 1918 H1N1 virus was reconstructed from gene sequences obtained from infected individuals and should therefore be the correct sequence and the 2009 H1N1 virus has a large

number of samples isolated by tissue culture, which does not have such a selective pressure for altering HA receptor binding properties as egg-propagation. There are also a large number of sequences of pdm09 viruses obtained directly from clinical isolates, without virus propagation. The 1957 H2N2 and 1968 H3N2 virus isolates available are all egg-adapted, as this was the customary method of virus propagation at the time, and consequently it is quite possible that the viruses have acquired a number of egg-adaptive substitutions in the HA. These substitutions will be acquired to allow efficient propagation of the virus by enhancing binding to the spectrum of receptors present in the allantoic membrane of hens' eggs, which is thought to have a high density of α2,3-linked receptors. Apart from the egg propagated nature of the isolates the earliest RNA sequencing methods required large quantities of viral RNA. The viruses therefore had to be propagated on a large scale and clinical isolates could not be analysed. This lack of samples to analyse for pandemic emergence characteristics makes it difficult, therefore, to draw robust conclusions about the required HA and NA characteristics for optimal fitness and transmissibility in the human context.

There are a number of implications for virus binding when one considers the range of virus morphologies that can occur. IAV morphology is variable in a spectrum of sizes from the 'spherical' viruses of ~100 nm in diameter to filamentous viruses >1 μm in length. It has been previously suggested that the filamentous viruses may be important in the infection cycle within humans. The receptor binding characteristics of these elongated viruses has not been individually characterised in comparison to the smaller spherical particles. In terms of the simple HA mediated receptor binding characteristics of these filaments, one would expect that the viruses would bind with

greater affinity with the filaments lying flat to the cell surface, as this would allow a larger number of HA interactions with the cell surface. The way that these viruses would interact with the BLI sensors used for the studies shown in this thesis should be similar to that of a cell membrane; however, it is unclear how the binding of filamentous virus would affect the amplitude of the measurement made. If the filaments attached flat against the surface of the sensor it could lower the overall density of virus binding compared to the measurements made for spherical virus, as the overall density of bound virus could be reduced. If a population of the viruses attached perpendicular to the sensor surface with only the tip of the filament bound it could lead to a much-increased signal, as the effective thickness of the surface of the BLI sensor would increase greatly. The role of the NA in the filaments is harder to predict. It is not thought that the proportions of HA and NA are different for filamentous viruses, when compared to 'spherical' viruses, and consequently the same surface area of virus membrane would likely deplete the sugar from the sensor surface to the same extent as for a spherical virus. The difference, however, is in the interaction kinetics of the virus. The filamentous virus particles are much greater in size and will consequently have a slower diffusion rate compared to spherical viruses. If there are transient interactions as shown in this thesis they are likely to be longer in duration owing to this slower diffusion, therefore giving the filamentous particles a greater probability of forming enough interactions with the surface, causing longer virus residence times, which would be sufficient to initiate virus infection. One does, however, need to consider the different modes and therefore possible different speeds of virus entry: spherical particles enter via endocytosis whereas filamentous virus enters by macropinocytosis. The longer interactions of filamentous viruses could also increase the depletion of sugar from the surface or, as appears more likely

from the experiments shown in this thesis, the longer interactions would reduce the overall NA cleavage, as the most efficient sialic acid depletion is seen in the early stages of virus binding when the interactions are short-lived.

An interesting feature of the HA/NA balance measurements made for both X-31, in Chapter 1, and H7N9 viruses, in Chapter 2, is the large difference in virus binding to the α2,3- and α2,6-linked receptors when the NA is not inhibited. There is a general pattern of lower overall virus binding to the receptor analogue 3SLN-PAA, when the NA is not inhibited. This appears to be mainly due to the preference of the NA for cleaving the α2,3-linked receptors. The reason for this overall preference of NA for α2,3-linked sugars is unclear, whether it is an evolved characteristic or an inherent property of the naturally higher cleavability of α2,3-linked sugars. The implications for this linkage preference are interesting when comparing human and avian virus characteristics. Human viruses typically infect URT cells which mainly express α2,6-linked receptors (as described in *Introduction*, section 1.3.1.2.2). Also present in the URT is mucus, which is thought to be rich in α2,3-linked receptors (as described in *Introduction*, section 1.7). In order to escape inhibition by this mucus human viruses must have low avidity for α2,3-linked receptors and/or an NA that cleaves α2,3-linked receptors efficiently. Whereas for efficient infection of URT cells the virus must have a strong avidity for α2,6-linked receptors and also an NA that is efficient at cleaving α2,6-linked receptors. The adaptation of IAV HA to α2,6-linked receptor preference has been extensively characterised and linked to a number of individual substitutions in the receptor binding site; however, the adaption of the receptor specificity of the NA is less well understood. There has not been any in-depth kinetic analysis of other avian virus NAs which would allow comparison with the N9

data presented in this thesis making it unclear whether NA from wild birds typically possesses activity for cleaving α2,6-linked receptors. The N9 characterised in this thesis (Chapter 4), which is similar to NAs found in wild birds, has very poor enzymatic turnover rates and strong substrate binding characteristics for α2,6-linked receptors, which make it act as a receptor binding protein. Whether this is the case for a number of other avian NAs is unclear. It would, therefore, be of interest to characterise the cleavage specificity and kinetic parameters of a number of NAs from different hosts and at different stages of adaptation to the human host to fully understand the consequences of this linkage specificity.

Another virus characteristic that has been poorly characterised and could be of great importance in affecting the HA and NA balance characteristics of viruses is the relative amounts, proportions and distribution of the HA and NA on the virus surface. The proportions of these proteins on different viruses have been reported anecdotally to vary significantly for viruses of different subtypes and with different genetic backgrounds. It would be of interest to carry out a systematic analysis of the quantities of the two glycoproteins on the virus surface and also to attempt to understand which viral component is responsible for altering these characteristics by measuring the molar quantities of the two proteins and relating these back to the number of viral particles present.

The technique described in this thesis for measuring HA and NA balance is useful for determining virus binding characteristics in more detail and in more mechanistic detail than seen before. However, there are a number of caveats with this particular technique that need to be considered. The virus concentration used in these experiments is high (100 pM ≈ 10^{10} particles/ml) and the sugars used are an

artificial system of the terminal trisaccharide moiety from much more complex long chain glycans, as described in Chapter 3. It would therefore be interesting to carry out these balance experiments using possibly more relevant receptors. There are a large number of alternative receptors available, such as those used on glycan microarrays; however, they are not available in the multimeric polyacrylamide attached variety required for the BLI experiments described in this thesis.

The sialylated glycans present on the cell surface present a wide range of different structures and linkages which have been characterized to an extent by mass spectrometry (Bateman *et al.*, 2010; Walther *et al.*, 2013; Bern *et al.*, 2013; Chan *et al.*, 2013; Jia *et al.*, 2014). Glycan microarrays have also been constructed from glycans directly isolated from tissues relevant to IAV infection (Song *et al.*, 2011; Byrd-Leotis *et al.*, 2014), which as well as allowing the characterisation of different glycans present in these tissues allow a semi-quantitative measurement of the affinity of each of these species in the binding of viruses and can therefore be used to define putative receptors. Both the mass spectroscopy and glycan microarray methods detailed above do, however, have a number of problems in giving a full picture of the importance of different glycans. Conventional mass spectrometry studies do not give reliable quantitative measures of the relative amounts of each glycan species and the differentiation of different linkages (α2,6- and α2,3-linked sialic acid) is difficult by conventional mass spectrometry techniques, as glycan species with these two different linkages have identical masses. Studies are required which quantify the glycans of the cell surface of defined cell types from specific tissues, with a method that can differentiate different sialic acid linkages, coupled with microarrays to define which are most relevant to the receptor binding of different viruses.

The technique for measuring the balance of HA and NA activities, presented in this thesis, combined with the NA kinetic characterisation methods, provide a new way to understand the role of these two proteins in receptor binding in an unprecedented quantitative manner. These techniques can then be used to infer the biochemical characteristics of viruses that are responsible for mediating transmission and therefore pandemic potential of a virus. It is clear that there is a requirement for both the relevant biochemical and biophysical studies to be combined with *in vivo* experiments of virus transmission and adaption in order to get the fullest picture of virus transmission characteristics.

Bibliography

Abdel-Ghafar, A.-N., Chotpitayasunondh, T., Gao, Z., Hayden, F.G., Nguyen, D.H., de Jong, M.D., Naghdaliyev, A., Peiris, J.S.M., Shindo, N., Soeroso, S. & Uyeki, T.M. (2008) Update on avian influenza A (H5N1) virus infection in humans. *The New England journal of medicine*. 358 (3), pp. 261–273.

Abe, Y., Takashita, E., Sugawara, K., Matsuzaki, Y., Muraki, Y. & Hongo, S. (2004) Effect of the addition of oligosaccharides on the biological activities and antigenicity of influenza A/H3N2 virus hemagglutinin. *The Journal of Virology*. 78 (18), pp. 9605–9611.

Abed, Y., Pizzorno, A., Bouhy, X., Rhéaume, C. & Boivin, G. (2014) Impact of potential permissive neuraminidase mutations on viral fitness of the H275Y oseltamivir-resistant influenza A(H1N1)pdm09 virus in vitro, in mice and in ferrets. *The Journal of Virology*. 88 (3), pp. 1652–1658.

Air, G.M. (2012) Influenza neuraminidase. *Influenza and other respiratory viruses*. 6 (4), pp. 245–256.

Aytay, S. & Schulze, I.T. (1991) Single amino acid substitutions in the hemagglutinin can alter the host range and receptor binding properties of H1 strains of influenza A virus. *The Journal of Virology*. 65 (6), pp. 3022–3028.

Babcock, H.P., Chen, C. & Zhuang, X. (2004) Using single-particle tracking to study nuclear trafficking of viral genes. *Biophysical journal*. 87 (4), pp. 2749–2758.

Baigent, S.J. & McCauley, J.W. (2001) Glycosylation of haemagglutinin and stalk-length of neuraminidase combine to regulate the growth of avian influenza viruses in tissue culture. *Virus research*. 79 (1-2), pp. 177–185.

Baigent, S.J., Bethell, R.C. & McCauley, J.W. (1999) Genetic analysis reveals that both haemagglutinin and neuraminidase determine the sensitivity of naturally occurring avian influenza viruses to zanamivir in vitro. *Virology*. 263 (2), pp. 323–338.

Bantia, S., Ghate, A.A., Ananth, S.L., Babu, Y.S., Air, G.M. & Walsh, G.M. (1998) Generation and characterization of a mutant of influenza A virus selected with the neuraminidase inhibitor BCX-140. *Antimicrobial agents and chemotherapy*. 42 (4), pp. 801–807.

Bateman, A.C., Karamanska, R., Busch, M.G., Dell, A., Olsen, C.W. & Haslam, S.M. (2010) Glycan analysis and influenza A virus infection of primary swine respiratory epithelial cells: the importance of NeuAc{alpha}2-6 glycans. *The Journal of Biological Chemistry*. 285 (44), pp. 34016–34026.

Baum, L.G. & Paulson, J.C. (1991) The N2 neuraminidase of human influenza virus has acquired a substrate specificity complementary to the hemagglutinin receptor specificity. *Virology*. 180 (1), pp. 10–15.

Bean, W.J., Schell, M., Katz, J., Kawaoka, Y., Naeve, C., Gorman, O. & Webster, R.G. (1992) Evolution of the H3 influenza virus hemagglutinin from human and nonhuman hosts. *The Journal of Virology*. 66 (2), pp. 1129–1138.

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

Bern, M., Brito, A.E., Pang, P.-C., Rekhi, A., Dell, A. & Haslam, S.M. (2013) Polylactosaminoglycan glycomics: enhancing the detection of high-molecular-weight N-glycans in matrix-assisted laser desorption ionization time-of-flight profiles by matched filtering. *Molecular & Cellular Proteomics*. 12 (4), pp. 996–1004.

Blackburne, B.P., Hay, A.J. & Goldstein, R.A. (2008) Changing selective pressure during antigenic changes in human influenza H3. Bruce Levin (ed.). *PLoS pathogens*. 4 (5), pp. e1000058.

Blick, T.J., Sahasrabudhe, A., McDonald, M., Owens, I.J., Morley, P.J., Fenton, R.J. & McKimm-Breschkin, J.L. (1998) The interaction of neuraminidase and hemagglutinin mutations in influenza virus in resistance to 4-guanidino-Neu5Ac2en. *Virology*. 246 (1), pp. 95–103.

Blumenkrantz, D., Roberts, K.L., Shelton, H., Lycett, S. & Barclay, W.S. (2013) The short stalk length of highly pathogenic avian influenza H5N1 virus neuraminidase limits transmission of pandemic H1N1 virus in ferrets. *The Journal of Virology*. 87 (19), pp. 10539–10551.

Bonfield, J.K., Smith, K.F. & Staden, R. (1995) A new DNA sequence assembly program. *Nucleic acids research*. 23 (24), pp. 4992–4999.

Bourmakina, S.V. & García-Sastre, A. (2003) Reverse genetics studies on the filamentous morphology of influenza A virus. *The Journal of General Virology*. 84 (Pt 3), pp. 517–527.

Böttcher-Friebertshäuser, E., Garten, W., Matrosovich, M. & Klenk, H.-D. (2014) The hemagglutinin: a determinant of pathogenicity. *Current topics in microbiology and immunology*. 385 (Chapter 384), pp. 3–34.

Bradley, K.C., Jones, C.A., Tompkins, S.M., Tripp, R.A., Russell, R.J., Gramer, M.R., Heimburg-Molinaro, J., Smith, D.F., Cummings, R.D. & Steinhauer, D.A. (2011) Comparison of the receptor binding properties of contemporary swine isolates and early human pandemic H1N1 isolates (Novel 2009 H1N1). *Virology*. 413 (2), pp. 169–182.

Breg, J., van Halbeek, H., Vliegenthart, J.F., Lamblin, G., Houvenaghel, M.C. & Roussel, P. (1987) Structure of sialyl-oligosaccharides isolated from bronchial mucus glycoproteins of patients (blood group O) suffering from cystic fibrosis. *European journal of biochemistry / FEBS*. 168 (1), pp. 57–68.

Brown, I.H. (2000) The epidemiology and evolution of influenza viruses in pigs. *Veterinary microbiology*. 74 (1-2), pp. 29–46.

Bui, M., Whittaker, G. & Helenius, A. (1996) Effect of M1 protein and low pH on nuclear transport of influenza virus ribonucleoproteins. *The Journal of Virology*. 70 (12), pp. 8391–8401.

Bullough, P.A., Hughson, F.M., Skehel, J.J. & Wiley, D.C. (1994a) Structure of influenza haemagglutinin at the pH of membrane fusion. *Nature*. 371 (6492), pp. 37–43.

Bullough, P.A., Hughson, F.M., Treharne, A.C., Ruigrok, R.W., Skehel, J.J. & Wiley, D.C. (1994b) *Crystals of a fragment of influenza haemagglutinin in the low pH induced conformation.* 236 (4), pp. 1262–1265.

Burleigh, L.M., Calder, L.J., Skehel, J.J. & Steinhauer, D.A. (2005) Influenza a viruses with mutations in the m1 helix six domain display a wide variety of morphological phenotypes. *The Journal of Virology*. 79 (2), pp. 1262–1270.

Burmeister, W.P., Cusack, S. & Ruigrok, R.W. (1994) Calcium is needed for the thermostability of influenza B virus neuraminidase. *The Journal of General Virology*. 75 (Pt 2)pp. 381–388.

Byrd-Leotis, L., Liu, R., Bradley, K.C., Lasanajak, Y., Cummings, S.F., Song, X., Heimburg-Molinaro, J., Galloway, S.E., Culhane, M.R., Smith, D.F., Steinhauer, D.A. & Cummings, R.D. (2014) Shotgun glycomics of pig lung identifies natural endogenous receptors for influenza viruses. *Proceedings of the National Academy of Sciences of the United States of America*. 111 (22), pp. E2241–E2250.

Cabezas, J.A., Calvo, P., Eid, P., Martin, J. & Perez, N. (1980) Neuraminidase from influenza virus A (H3N2). Specificity towards several substrates and procedure of activity determination. *Biochimica et biophysica acta*. 616pp. 228–238.

Calder, L.J., Wasilewski, S., Berriman, J.A. & Rosenthal, P.B. (2010) Structural organization of a filamentous influenza A virus. *Proceedings of the National Academy of Sciences of the United States of America*. 107 (23), pp. 10685–10690.

Campbell, P.J., Danzy, S., Kyriakis, C.S., Deymier, M.J., Lowen, A.C. & Steel, J. (2014) The M Segment of the 2009 Pandemic Influenza Virus Confers Increased Neuraminidase Activity, Filamentous Morphology, and Efficient Contact Transmissibility to A/Puerto Rico/8/1934-Based Reassortant Viruses. *The Journal of Virology*. 88 (7), pp. 3802–3814.

Carroll, S.M. & Paulson, J.C. (1982) Complete metal ion requirement of influenza virus N1 neuraminidases. *Archives of Virology*. 71 (3), pp. 273–277.

Castrucci, M.R. & Kawaoka, Y. (1993) Biologic importance of neuraminidase stalk length in influenza A virus. *The Journal of Virology*. 67 (2), pp. 759–764.

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

Chan, R.W.Y., Karamanska, R., Van Poucke, S., Van Reeth, K., Chan, I.W.W., Chan, M.C.W., Dell, A., Peiris, J.S.M., Haslam, S.M., Guan, Y. & Nicholls, J.M. (2013) Infection of swine ex vivo tissues with avian viruses including H7N9 and correlation with glycomic analysis. - PubMed - NCBI. *Influenza and other respiratory viruses*. 7 (6), pp. 1269–1282.

Chandrasekaran, A., Srinivasan, A., Raman, R., Viswanathan, K., Raguram, S., Tumpey, T.M., Sasisekharan, V. & Sasisekharan, R. (2008) Glycan topology determines human adaptation of avian H5N1 virus hemagglutinin. *Nature Biotechnology*. 26 (1), pp. 107–113.

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

Chen, J., Skehel, J.J. & Wiley, D.C. (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. *Proceedings of the National Academy of Sciences of the United States of America*. 96 (16), pp. 8967–8972.

Chen, J., Wharton, S.A., Weissenhorn, W., Calder, L.J., Hughson, F.M., Skehel, J.J. & Wiley, D.C. (1995) A soluble domain of the membrane-anchoring chain of influenza virus hemagglutinin (HA2) folds in Escherichia coli into the low-pH-induced conformation. *Proceedings of the National Academy of Sciences of the United States of America*. 92 (26), pp. 12205–12209.

Chen, L., Zhu, F., Xiong, C., Zhang, Z., Jiang, L., Chen, Y., Zhao, G. & Jiang, Q. (2015) Could a deletion in neuraminidase stalk strengthen human tropism of the novel avian influenza virus H7N9 in China, 2013? *International journal of environmental research and public health*. 12 (1), pp. 1020–1028.

Chen, L.-M., Rivailler, P., Hossain, J., Carney, P., Balish, A., Perry, I., Davis, C.T., Garten, R., Shu, B., Xu, X., Klimov, A., Paulson, J.C., Cox, N.J., Swenson, S., et al. (2011) Receptor specificity of subtype H1 influenza A viruses isolated from swine and humans in the United States. *Virology*. 412 (2), pp. 401–410.

Chen, W., Calvo, P.A., Malide, D., Gibbs, J., Schubert, U., Bacik, I., Basta, S., O'Neill, R., Schickli, J., Palese, P., Henklein, P., Bennink, J.R. & Yewdell, J.W. (2001) A novel influenza A virus mitochondrial protein that induces cell death. *Nature medicine*. 7 (12), pp. 1306–1312.

Chertow, D.S. & Memoli, M.J. (2013) Bacterial coinfection in influenza: a grand rounds review. *The Journal of the American Medical Association*. 309 (3), pp. 275–282.

Childs, R.A., Palma, A.S., Wharton, S., Matrosovich, T., Liu, Y., Chai, W., Campanero-Rhodes, M.A., Zhang, Y., Eickmann, M., Kiso, M., Hay, A., Matrosovich, M. & Feizi, T. (2009) Receptor-binding specificity of pandemic influenza A (H1N1) 2009 virus determined by carbohydrate microarray. *Nature Biotechnology*. 27 (9), pp. 797–799.

Chong, A.K., Pegg, M.S. & Itzstein, von, M. (1991) Influenza virus sialidase: effect of calcium on steady-state kinetic parameters. *Biochimica et biophysica acta*. 1077 (1), pp. 65–71.

Chong, A.K., Pegg, M.S., Taylor, N.R. & Itzstein, von, M. (1992) Evidence for a sialosyl cation transition-state complex in the reaction of sialidase from influenza virus. *European journal of biochemistry / FEBS*. 207 (1), pp. 335–343.

Chou, Y.-Y., Albrecht, R.A., Pica, N., Lowen, A.C., Richt, J.A., García-Sastre, A., Palese, P. & Hai, R. (2011) The M segment of the 2009 new pandemic H1N1 influenza virus is critical for its high transmission efficiency in the guinea pig model. *The Journal of Virology*. 85 (21), pp. 11235–11241.

Chu, C.M., Dawson, I.M. & Elford, W.J. (1949) Filamentous forms associated with newly isolated influenza virus. *The Lancet*. 1 (6554), pp. 602–605.

Cohen, M., Zhang, X.-Q., Senaati, H.P., Chen, H.-W., Varki, N.M., Schooley, R.T. & Gagneux, P. (2013) Influenza A penetrates host mucus by cleaving sialic acids with neuraminidase. *Virology journal*. 10 (1), pp. 321.

Collins, P.J., Vachieri, S.G., Haire, L.F., Ogrodowicz, R.W., Martin, S.R., Walker, P.A., Xiong, X., Gamblin, S.J. & Skehel, J.J. (2014) Recent evolution of equine influenza and the origin of canine influenza. *Proceedings of the National Academy of Sciences of the United States of America*. 111 (30), pp. 11175–11180.

Colman, P.M. (1994) Influenza virus neuraminidase: structure, antibodies, and inhibitors. *Protein science : a publication of the Protein Society*. 3 (10), pp. 1687–1696.

Colman, P.M., Varghese, J.N. & Laver, W.G. (1983) Structure of the catalytic and antigenic sites in influenza virus neuraminidase. *Nature*. 303 (5912), pp. 41–44.

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

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

Cross, K., Langley, W., Russell, R., Skehel, J. & Steinhauer, D. (2009) Composition and Functions of the Influenza Fusion Peptide. *Protein & Peptide Letters*. 16 (7), pp. 766–778.

Crusat, M., Liu, J., Palma, A.S., Childs, R.A., Liu, Y., Wharton, S.A., Lin, Y.P., Coombs, P.J., Martin, S.R., Matrosovich, M., Chen, Z., Stevens, D.J., Hien, V.M., Thanh, T.T., et al. (2013) Changes in the hemagglutinin of H5N1 viruses during human infection--influence on receptor binding. *Virology*. 447 (1-2), pp. 326–337.

Daniels, R.S., Jeffries, S., Yates, P., Schild, G.C., Rogers, G.N., Paulson, J.C., Wharton, S.A., Douglas, A.R., Skehel, J.J. & Wiley, D.C. (1987) The receptor-binding and membrane-fusion properties of influenza virus variants selected using anti-haemagglutinin monoclonal antibodies. *The EMBO journal*. 6 (5), pp. 1459–1465.

de Vries, R.P., de Vries, E., Moore, K.S., Rigter, A., Rottier, P.J.M. & de Haan, C.A.M. (2011) Only two residues are responsible for the dramatic difference in receptor binding between swine and new pandemic H1 hemagglutinin. *The Journal of Biological Chemistry*. 286 (7), pp. 5868–5875.

Desselberger, U., Racaniello, V.R., Zazra, J.J. & Palese, P. (1980) The 3‘and 5’-terminal sequences of influenza A, B and C virus RNA segments are highly conserved and show partial inverted complementarity. *Gene*. 8, pp. 315–328.

Dortmans, J.C.F.M., Dekkers, J., Wickramasinghe, I.N.A., Verheije, M.H., Rottier, P.J.M., van Kuppeveld, F.J.M., de Vries, E. & de Haan, C.A.M. (2013) Adaptation of novel H7N9 influenza A virus to human receptors. *Scientific reports*. 3, pp. 3058.

Durrer, P., Galli, C., Hoenke, S., Corti, C., Glück, R., Vorherr, T. & Brunner, J. (1996) H+-induced membrane insertion of influenza virus hemagglutinin involves the HA2 amino-terminal fusion peptide but not the coiled coil region. *The Journal of Biological Chemistry*. 271 (23), pp. 13417–13421.

Eisen, M.B., Sabesan, S., Skehel, J.J. & Wiley, D.C. (1997) Binding of the influenza A virus to cell-surface receptors: structures of five hemagglutinin-sialyloligosaccharide complexes determined by X-ray crystallography. *Virology*. 232 (1), pp. 19–31.

Elleman, C.J. & Barclay, W.S. (2004) The M1 matrix protein controls the filamentous phenotype of influenza A virus. *Virology*. 321 (1), pp. 144–153.

Els, M.C., Air, G.M., Murti, K.G., Webster, R.G. & Laver, W.G. (1985) An 18-amino acid deletion in an influenza neuraminidase. *Virology*. 142 (2), pp. 241–247.

Engelhardt, O.G., Smith, M. & Fodor, E. (2005) Association of the influenza A virus RNA-dependent RNA polymerase with cellular RNA polymerase II. *The Journal of Virology*. 79 (9), pp. 5812–5818.

Fanning, T.G., Slemons, R.D., Reid, A.H., Janczewski, T.A., Dean, J. & Taubenberger, J.K. (2002) 1917 avian influenza virus sequences suggest that the 1918 pandemic virus did not acquire its hemagglutinin directly from birds. *The Journal of Virology*. 76 (15), pp. 7860–7862.

Fazekas de St Groth, S. (1952) Nasal mucus and influenza viruses. I. The haemagglutinin inhibitor in nasal secretions. *The Journal of hygiene*. 50 (4), pp. 471–490.

Floyd, D.L., Ragains, J.R., Skehel, J.J., Harrison, S.C. & van Oijen, A.M. (2008) Single-particle kinetics of influenza virus membrane fusion. *Proceedings of the National Academy of Sciences of the United States of America*. 105 (40), pp. 15382–15387.

Fodor, E. (2013) The RNA polymerase of influenza a virus: mechanisms of viral transcription and replication. *Acta virologica*. 57 (2), pp. 113–122.

Fodor, E., Pritlove, D.C. & Brownlee, G.G. (1994) The influenza virus panhandle is involved in the initiation of transcription. *The Journal of Virology*. 68 (6), pp. 4092–4096.

França, M., Stallknecht, D.E. & Howerth, E.W. (2013) Expression and distribution of sialic acid influenza virus receptors in wild birds. *Avian pathology : journal of the W.V.P.A*. 42 (1), pp. 60–71.

Furuta, Y., Gowen, B.B., Takahashi, K., Shiraki, K. & Smee, D.F. (2013) Favipiravir (T-705), a novel viral RNA polymerase inhibitor. *Antiviral research*. 100 (2), pp. 446–454.

Galloway, S.E., Reed, M.L., Russell, C.J. & Steinhauer, D.A. (2013) Influenza HA subtypes demonstrate divergent phenotypes for cleavage activation and pH of fusion: implications for host range and adaptation. *PLoS pathogens*. 9 (2), pp. e1003151.

Gambaryan, A.S., Robertson, J.S. & Matrosovich, M.N. (1999) Effects of egg-adaptation on the receptor-binding properties of human influenza A and B viruses. *Virology*. 258 (2), pp. 232–239.

Gamblin, S.J. & Skehel, J.J. (2010) Influenza hemagglutinin and neuraminidase membrane glycoproteins. *The Journal of Biological Chemistry*. 285 (37), pp. 28403–28409.

Gao, R., Cao, B., Hu, Y., Feng, Z., Wang, D., Hu, W., Chen, J., Jie, Z., Qiu, H., Xu, K., Xu, X., Lu, H., Zhu, W., Gao, Z., et al. (2013) Human infection with a novel avian-origin influenza A (H7N9) virus. *The New England journal of medicine*. 368 (20), pp. 1888–1897.

Garman, E. & Laver, G. (2005) The structure, function, and inhibition of influenza virus neuraminidase. In: Wolfgang Fischer (ed.). *Viral Membrane Proteins Structure, Function, and Drug Design*. New York: Viral Membrane Proteins: Structure. doi:10.1007/0-387-28146-0_17.pdf.

Garten, R.J., Davis, C.T., Russell, C.A., Shu, B., Lindstrom, S., Balish, A., Sessions, W.M., Xu, X., Skepner, E., Deyde, V., Okomo-Adhiambo, M., Gubareva, L., Barnes, J., Smith, C.B., et al. (2009) Antigenic and genetic characteristics of swine-origin 2009 A(H1N1) influenza viruses circulating in humans. *Science*. 325 (5937), pp. 197–201.

Gerhard, W., Yewdell, J., Frankel, M.E. & Webster, R. (1981) Antigenic structure of influenza virus haemagglutinin defined by hybridoma antibodies. *Nature*. 290 (5808), pp. 713–717.

Gerlach, T., Kühling, L., Uhlendorff, J., Laukemper, V., Matrosovich, T., Czudai-Matwich, V., Schwalm, F., Klenk, H.D. & Matrosovich, M. (2012) Characterization of the neuraminidase of the H1N1/09 pandemic influenza virus. *Vaccine*. 30 (51), pp. 7348–7352.

Goldfield, M., Bartley, J.D., Pizzuti, W., Black, H.C., Altman, R. & Halperin, W.E. (1977) Influenza in New Jersey in 1976: isolations of influenza A/New Jersey/76 virus at Fort Dix. *The Journal of infectious diseases*. 136 Suppl pp. S347–S355.

Hagen, M., Chung, T.D., Butcher, J.A. & Krystal, M. (1994) Recombinant influenza virus polymerase: requirement of both 5‘ and 3’ viral ends for endonuclease activity. *The Journal of Virology*. 68 (3), pp. 1509–1515.

Hale, B.G., Randall, R.E., Ortín, J. & Jackson, D. (2008) The multifunctional NS1 protein of influenza A viruses. *The Journal of General Virology*. 89 (10), pp. 2359–2376.

Hall, T.A. (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic acids symposium series*. 41.

Hanaoka, K., Pritchett, T.J., Takasaki, S., Kochibe, N., Sabesan, S., Paulson, J.C. & Kobata, A. (1989) 4-O-acetyl-N-acetylneuraminic acid in the N-linked carbohydrate structures of equine and guinea pig alpha 2-macroglobulins, potent inhibitors of influenza virus infection. *The Journal of Biological Chemistry*. 264 (17), pp. 9842–9849.

Harris, A., Cardone, G., Winkler, D.C., Heymann, J.B., Brecher, M., White, J.M. & Steven, A.C. (2006) Influenza virus pleiomorphy characterized by cryoelectron tomography. *Proceedings of the National Academy of Sciences of the United States of America*. 103 (50), pp. 19123–19127.

Hausmann, J., Kretzschmar, E., Garten, W. & Klenk, H.D. (1995) N1 neuraminidase of influenza virus A/FPV/Rostock/34 has haemadsorbing activity. *The Journal of General Virology*. 76 (7), pp. 1719–1728.

Hay, A.J., Lomniczi, B., Bellamy, A.R. & Skehel, J.J. (1977) Transcription of the influenza virus genome. *Virology*. 83 (2), pp. 337–355.

Hay, A.J., Skehel, J.J. & McCauley, J. (1980) Structure and synthesis of influenza virus complementary RNAs. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*. 288, (1029), pp. 341–348.

Hayden, F.G. (2006) Antivirals for influenza: historical perspectives and lessons learned. *Antiviral research*. 71, pp. 372–378.

Helenius, A. (1992) Unpacking the incoming influenza virus. *Cell*. 69 (4), pp. 577–578.

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

Hidari, K.I.P.J., Shimada, S., Suzuki, Y. & Suzuki, T. (2007) Binding kinetics of influenza viruses to sialic acid-containing carbohydrates. *Glycoconjugate journal*. 24 (9), pp. 583–590.

Hoffmann, E., Neumann, G., Kawaoka, Y., Hobom, G. & Webster, R.G. (2000) A DNA transfection system for generation of influenza A virus from eight plasmids. *Proceedings of the National Academy of Sciences of the United States of America*. 97 (11), pp. 6108–6113.

Hoyle, L. (1968) *The Influenza Viruses*. New York: Springer-Verlag.

Hughes, M.T., Matrosovich, M., Rodgers, M.E., McGregor, M. & Kawaoka, Y. (2000) Influenza A viruses lacking sialidase activity can undergo multiple cycles of replication in cell culture, eggs, or mice. *The Journal of Virology*. 74 (11), pp. 5206–5212.

Hutchinson, E.C. & Fodor, E. (2013) Transport of the influenza virus genome from nucleus to nucleus. *Viruses*. 5 (10), pp. 2424–2446.

Ibricevic, A., Walter, M.J., Newby, C., Battaile, J.T., Brown, E.G., Holtzman, M.J. & Brody, S.L. (2006) Influenza virus receptor specificity and cell tropism in mouse and human airway epithelial cells. *The Journal of Virology*. 80 (15), pp. 7469–7480.

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

Inglis, S.C., Carroll, A.R., Lamb, R.A. & Mahy, B.W. (1976) Polypeptides specified by the influenza virus genome I. Evidence for eight distinct gene products specified by fowl plague virus. *Virology*. 74 (2), pp. 489–503.

Inglis, S.C., Gething, M.J. & Brown, C.M. (1980) Relationship between the messenger RNAs transcribed from two overlapping genes of influenza virus. *Nucleic acids research*. 8 (16), pp. 3575–3589.

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

Ito, T., Suzuki, Y., Mitnaul, L., Vines, A., Kida, H. & Kawaoka, Y. (1997) Receptor specificity of influenza A viruses correlates with the agglutination of erythrocytes from different animal species. *Virology*. 227 (2), pp. 493–499.

Itzstein, von, M. (2007) The war against influenza: discovery and development of sialidase inhibitors. *Nature reviews. Drug discovery*. 6 (12), pp. 967–974.

Ivanovic, T., Rozendaal, R., Floyd, D.L., Popovic, M., van Oijen, A.M. & Harrison, S.C. (2012) Kinetics of proton transport into influenza virions by the viral M2 channel. Paul Digard (ed.). *PloS one*. 7 (3), pp. e31566.

Jagger, B.W., Wise, H.M., Kash, J.C., Walters, K.A., Wills, N.M., Xiao, Y.L., Dunfee, R.L., Schwartzman, L.M., Ozinsky, A., Bell, G.L., Dalton, R.M., Lo, A., Efstathiou, S., Atkins, J.F., et al. (2012) An overlapping protein-coding region in influenza A virus segment 3 modulates the host response. *Science*. 337 (6091), pp. 199–204.

Jia, N., Barclay, W.S., Roberts, K., Yen, H.-L., Chan, R.W.Y., Lam, A.K.Y., Air, G., Peiris, J.S.M., Dell, A., Nicholls, J.M. & Haslam, S.M. (2014) Glycomic characterization of respiratory tract tissues of ferrets: implications for its use in influenza virus infection studies. *The Journal of Biological Chemistry*. 289 (41), pp. 28489–28504.

Kaverin, N.V., Gambaryan, A.S., Bovin, N.V., Rudneva, I.A., Shilov, A.A., Khodova, O.M., Varich, N.L., Sinitsin, B.V., Makarova, N.V. & Kropotkina, E.A. (1998) Postreassortment changes in influenza A virus hemagglutinin restoring HA-NA functional match. *Virology*. 244 (2), pp. 315–321.

Kawaoka, Y., Chambers, T.M., Sladen, W.L. & Webster, R.G. (1988) Is the gene pool of influenza viruses in shorebirds and gulls different from that in wild ducks? *Virology*. 163 (1), pp. 247–250.

Kawaoka, Y., Krauss, S. & Webster, R.G. (1989) Avian-to-human transmission of the PB1 gene of influenza A viruses in the 1957 and 1968 pandemics. *The Journal of Virology*. 63 (11), pp. 4603–4608.

Kemble, G.W., Danieli, T. & White, J.M. (1994) Lipid-anchored influenza hemagglutinin promotes hemifusion, not complete fusion. *Cell*. 76 (2), pp. 383–391.

Kemble, G.W., Henis, Y.I. & White, J.M. (1993) GPI- and transmembrane-anchored influenza hemagglutinin differ in structure and receptor binding activity. *The Journal of cell biology*. 122 (6), pp. 1253–1265.

Kilbourne, E.D. & Murphy, J.S. (1960) Genetic Studies of Influenza Viruses I. Viral Morphology and Growth Capacity as Exchangeable Genetic Traits. Rapid *in Ovo* Adaptation of Early Passage Asian Strain Isolates by Combination with PR8. *The Journal of Experimental Medicine*. 111 (3), pp. 387–406.

Kim, J.-H., Resende, R., Wennekes, T., Chen, H.-M., Bance, N., Buchini, S., Watts, A.G., Pilling, P., Streltsov, V.A., Petric, M., Liggins, R., Barrett, S., McKimm-Breschkin, J.L., Niikura, M., et al. (2013) Mechanism-based covalent neuraminidase inhibitors with broad-spectrum influenza antiviral activity. *Science*. 340 (6128), pp. 71–75.

King, A. (2011) *Virus Taxonomy: Ninth report of the International Committee on Taxonomy of Viruses*. Edited by: A King, A King, E Lefkowitz, M Adams, & E Carstens (eds.). Elsevier.

Klenk, H.D., Garten, W. & Matrosovich, M. (2011) Molecular mechanisms of interspecies transmission and pathogenicity of influenza viruses: Lessons from the 2009 pandemic. *BioEssays : news and reviews in molecular, cellular and developmental biology*. 33 (3), pp. 180–188.

Klenk, H.D., Rott, R., Orlich, M. & Blödorn, J. (1975) Activation of influenza A viruses by trypsin treatment. *Virology*. 68pp. 426–439.

Klimov, A.I., Garten, R., Russell, C., Barr, I.G., Besselaar, T.G., Daniels, R., Engelhardt, O.G., Grohmann, G., Itamura, S., Kelso, A., McCauley, J., Odagiri, T., Smith, D., Tashiro, M., et al. (2012) WHO recommendations for the viruses to be used in the 2012 Southern Hemisphere Influenza Vaccine: epidemiology, antigenic and genetic characteristics of influenza A(H1N1)pdm09, A(H3N2) and B influenza viruses collected from February to September 2011. *Vaccine*. 30 (45), pp. 6461–6471.

Kobasa, D., Kodihalli, S., Luo, M., Castrucci, M.R., Donatelli, I., Suzuki, Y., Suzuki, T. & Kawaoka, Y. (1999) Amino acid residues contributing to the substrate specificity of the influenza A virus neuraminidase. *The Journal of Virology*. 73 (8), pp. 6743–6751.

Kobasa, D., Rodgers, M.E., Wells, K. & Kawaoka, Y. (1997) Neuraminidase hemadsorption activity, conserved in avian influenza A viruses, does not influence viral replication in ducks. *The Journal of Virology*. 71 (9), pp. 6706–6713.

Krammer, F. & Palese, P. (2015) Advances in the development of influenza virus vaccines. *Nature reviews. Drug discovery*. 14 (3), pp. 167–182.

Kühnel, K., Jarchau, T., Wolf, E., Schlichting, I., Walter, U., Wittinghofer, A. & Strelkov, S.V. (2004) The VASP tetramerization domain is a right-handed coiled coil based on a 15-residue repeat. *Proceedings of the National Academy of Sciences of the United States of America*. 101 (49), pp. 17027–17032.

Lai, J.C.C., Garcia, J.M., Dyason, J.C., Böhm, R., Madge, P.D., Rose, F.J., Nicholls, J.M., Peiris, J.S.M., Haselhorst, T. & Itzstein, von, M. (2012) A secondary sialic acid binding site on influenza virus neuraminidase: fact or fiction? *Angewandte Chemie International Edition*. 51 (9), pp. 2221–2224.

Lakdawala, S.S., Lamirande, E.W., Suguitan, A.L., Wang, W., Santos, C.P., Vogel, L., Matsuoka, Y., Lindsley, W.G., Jin, H. & Subbarao, K. (2011) Eurasian-origin gene segments contribute to the transmissibility, aerosol release, and morphology of the 2009 pandemic H1N1 influenza virus. Ron A M Fouchier (ed.). *PLoS pathogens*. 7 (12), pp. e1002443.

Lam, T.T.-Y., Zhou, B., Wang, J., Chai, Y., Shen, Y., Chen, X., Ma, C., Hong, W., Chen, Y., Zhang, Y., Duan, L., Chen, P., Jiang, J., Zhang, Y., et al. (2015) Dissemination, divergence and establishment of H7N9 influenza viruses in China. *Nature*. 522 (7554), pp. 102–105.

Lamb, R.A., Lai, C.J. & Choppin, P.W. (1981) Sequences of mRNAs derived from genome RNA segment 7 of influenza virus: colinear and interrupted mRNAs code for overlapping proteins. *Proceedings of the National Academy of Sciences of the United States of America*. 78 (7), pp. 4170–4174.

Lamblin, G., Boersma, A., Klein, A., Roussel, P., van Halbeek, H. & Vliegenthart, J.F. (1984a) Primary structure determination of five sialylated oligosaccharides derived from bronchial mucus glycoproteins of patients suffering from cystic fibrosis. The occurrence of the NeuAc alpha(2----3)Gal beta(1----4)[Fuc alpha(1----3)] GlcNAc beta(1----.) structural element revealed by 500-MHz 1H NMR spectroscopy. [online]. *The Journal of Biological Chemistry*. 259 (14), pp. 9051–9058.

Lamblin, G., Lhermitte, M., Klein, A., Roussel, P., Van Halbeek, H. & Vliegenthart, J.F. (1984b) Carbohydrate chains from human bronchial mucus glycoproteins: a wide spectrum of oligosaccharide structures. *Biochemical Society transactions*. 12 (4), pp. 599–600.

Laver, W.G., Colman, P.M., Webster, R.G., Hinshaw, V.S. & Air, G.M. (1984) Influenza virus neuraminidase with hemagglutinin activity. *Virology*. 137 (2), pp. 314–323.

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

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

Li, J., Zu Dohna, H., Cardona, C.J., Miller, J. & Carpenter, T.E. (2011a) Emergence and genetic variation of neuraminidase stalk deletions in avian influenza viruses. Maciej F Boni (ed.). *PloS one*. 6 (2), pp. e14722.

Li, Y., Cao, H., Dao, N., Luo, Z., Yu, H., Chen, Y., Xing, Z., Baumgarth, N., Cardona, C. & Chen, X. (2011b) High-throughput neuraminidase substrate specificity study of human and avian influenza A viruses. *Virology*. 415 (1), pp. 12–19.

Lin, Y.P., Gregory, V., Collins, P., Kloess, J., Wharton, S., Cattle, N., Lackenby, A., Daniels, R. & Hay, A. (2010) Neuraminidase receptor binding variants of human influenza A(H3N2) viruses resulting from substitution of aspartic acid 151 in the catalytic site: a role in virus attachment? *The Journal of Virology*. 84 (13), pp. 6769–6781.

Lin, Y.P., Xiong, X., Wharton, S.A., Martin, S.R., Coombs, P.J., Vachieri, S.G., Christodoulou, E., Walker, P.A., Liu, J., Skehel, J.J., Gamblin, S.J., Hay, A.J., Daniels, R.S. & McCauley, J.W. (2012) Evolution of the receptor binding properties of the influenza A(H3N2) hemagglutinin. *Proceedings of the National Academy of Sciences of the United States of America*. 109 (52), pp. 21474–21479.

Luo, G., Chung, J. & Palese, P. (1993) Alterations of the stalk of the influenza virus neuraminidase: deletions and insertions. *Virus research*. 29pp. 141–153.

Ma, W., Liu, Q., Bawa, B., Qiao, C. & Qi, W. (2012) The NA and M genes of the 2009 pandemic influenza H1N1 virus functionally cooperate to facilitate efficient replication and transmissibility in pigs. *Journal of general virology*. 93, pp. 1261–1268.

Maines, T.R., Jayaraman, A., Belser, J.A., Wadford, D.A., Pappas, C., Zeng, H., Gustin, K.M., Pearce, M.B., Viswanathan, K., Shriver, Z.H., Raman, R., Cox, N.J., Sasisekharan, R., Katz, J.M., et al. (2009) Transmission and pathogenesis of swine-origin 2009 A(H1N1) influenza viruses in ferrets and mice. *Science*. 325 (5939), pp. 484–487.

Mammen, M., Choi, S.K. & Whitesides, G.M. (1998) Polyvalent interactions in biological systems: implications for design and use of multivalent ligands and inhibitors. *Angewandte Chemie …*.

Matlin, K.S., Reggio, H., Helenius, A. & Simons, K. (1981) Infectious entry pathway of influenza virus in a canine kidney cell line. *The Journal of cell biology*. 91 (3), pp. 601–613.

Matrosovich, M. & Klenk, H.-D. (2003) Natural and synthetic sialic acid-containing inhibitors of influenza virus receptor binding. *Reviews in medical virology*. 13 (2), pp. 85–97.

Matrosovich, M., Matrosovich, T., Carr, J., Roberts, N.A. & Klenk, H.-D. (2003) Overexpression of the alpha-2,6-sialyltransferase in MDCK cells increases influenza virus sensitivity to neuraminidase inhibitors. *The Journal of Virology*. 77 (15), pp. 8418–8425.

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

Matrosovich, M.N., Gambaryan, A.S., Teneberg, S., Piskarev, V.E., Yamnikova, S.S., Lvov, D.K., Robertson, J.S. & Karlsson, K.A. (1997) Avian influenza A viruses differ from human viruses by recognition of sialyloligosaccharides and gangliosides and by a higher conservation of the HA receptor-binding site. *Virology*. 233 (1), pp. 224–234.

Matrosovich, M.N., Matrosovich, T.Y., Gray, T., Roberts, N.A. & Klenk, H.-D. (2004a) Human and avian influenza viruses target different cell types in cultures of human airway epithelium. *Proceedings of the National Academy of Sciences of the United States of America*. 101 (13), pp. 4620–4624.

Matrosovich, M.N., Matrosovich, T.Y., Gray, T., Roberts, N.A. & Klenk, H.-D. (2004b) Neuraminidase is important for the initiation of influenza virus infection in human airway epithelium. *The Journal of Virology*. 78 (22), pp. 12665–12667.

Matsuoka, Y., Swayne, D.E., Thomas, C., Rameix-Welti, M.-A., Naffakh, N., Warnes, C., Altholtz, M., Donis, R. & Subbarao, K. (2009) Neuraminidase stalk length and additional glycosylation of the hemagglutinin influence the virulence of influenza H5N1 viruses for mice. *The Journal of Virology*. 83 (9), pp. 4704–4708.

McCauley, J.W. & Mahy, B.W. (1983) Structure and function of the influenza virus genome. *The Biochemical journal*. 211 (2), pp. 281–294.

McKimm-Breschkin, J.L., Blick, T.J., Sahasrabudhe, A., Tiong, T., Marshall, D., Hart, G.J., Bethell, R.C. & Penn, C.R. (1996) Generation and characterization of variants of NWS/G70C influenza virus after in vitro passage in 4-amino-Neu5Ac2en and 4-guanidino-Neu5Ac2en. *Antimicrobial agents and chemotherapy*. 40 (1), pp. 40–46.

Melikyan, G.B., White, J.M. & Cohen, F.S. (1995) GPI-anchored influenza hemagglutinin induces hemifusion to both red blood cell and planar bilayer membranes. *The Journal of cell biology*. 131 (3), pp. 679–691.

Meng, B. & Marriott, A.C. (2010) The receptor preference of influenza viruses. *Influenza and other respiratory viruses*.

Mitnaul, L.J., Matrosovich, M.N. & Castrucci, M.R. (2000) Balanced hemagglutinin and neuraminidase activities are critical for efficient replication of influenza A virus. *The Journal of Virology*. 74 (13), pp. 6015–6020.

Munier, S., Larcher, T., Cormier-Aline, F., Soubieux, D., Su, B., Guigand, L., Labrosse, B., Cherel, Y., Quéré, P., Marc, D. & Naffakh, N. (2010) A genetically engineered waterfowl influenza virus with a deletion in the stalk of the neuraminidase has increased virulence for chickens. *The Journal of Virology*. 84 (2), pp. 940–952.

Muramoto, Y., Noda, T., Kawakami, E., Akkina, R. & Kawaoka, Y. (2013) Identification of novel influenza A virus proteins translated from PA mRNA. *The Journal of Virology*. 87 (5), pp. 2455–2462.

Naeve, C.W., Hinshaw, V.S. & Webster, R.G. (1984) Mutations in the hemagglutinin receptor-binding site can change the biological properties of an influenza virus. *The Journal of Virology*. 51 (2), pp. 567–569.

Nakajima, K., Desselberger, U. & Palese, P. (1978) Recent human influenza A (H1N1) viruses are closely related genetically to strains isolated in 1950. *Nature*. 274, pp. 334–339.

Nicholls, J.M., Bourne, A.J., Chen, H., Guan, Y. & Peiris, J.S.M. (2007) Sialic acid receptor detection in the human respiratory tract: evidence for widespread distribution of potential binding sites for human and avian influenza viruses. *Respiratory research*. 8 (1), pp. 73.

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

Nunes-Correia, I., Ramalho-Santos, J. & Nir, S. (1999) Interactions of influenza virus with cultured cells: detailed kinetic modeling of binding and endocytosis. *Biochemistry*. 38 (3), pp. 1095–1101.

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

Ohuchi, M., Ohuchi, R., Feldmann, A. & Klenk, H.D. (1997) Regulation of receptor binding affinity of influenza virus hemagglutinin by its carbohydrate moiety. *The Journal of Virology*. 71 (11), pp. 8377–8384.

Ohuchi, M., Ohuchi, R., Sakai, T. & Matsumoto, A. (2002) Tight binding of influenza virus hemagglutinin to its receptor interferes with fusion pore dilation. *The Journal of Virology*. 76 (24), pp. 12405–12413.

Olsen, B., Munster, V.J., Wallensten, A., Waldenström, J., Osterhaus, A.D.M.E. & Fouchier, R.A.M. (2006) Global patterns of influenza a virus in wild birds. *Science*. 312 (5772), pp. 384–388.

Olsen, C.W. (2002) The emergence of novel swine influenza viruses in North America. *Virus research*. 85 (2), pp. 199–210.

Oshansky, C.M., Pickens, J.A., Bradley, K.C., Jones, L.P., Saavedra-Ebner, G.M., Barber, J.P., Crabtree, J.M., Steinhauer, D.A., Tompkins, S.M. & Tripp, R.A. (2011) Avian influenza viruses infect primary human bronchial epithelial cells unconstrained by sialic acid α2,3 residues. *PloS one*. 6 (6), pp. e21183.

Palese, P. & Compans, R.W. (1976) Inhibition of Influenza Virus Replication in Tissue Culture by 2-deoxy-2,3-dehydro-N-trifluoroacetylneuraminic acid (FANA): Mechanism of Action. *The Journal of General Virology*. 33 (1), pp. 159–163.

Palese, P. & Shaw, M. (2007) Orthomyxoviridae: The Viruses and Their Replication. In: D M Knipe, Edited by: D M Knipe, & P M Howley (eds.). *Fields Virology*. Fields Virology. 5 edition. Philadelphia: Lippincott, Williams & Wilkins. pp. pp. 1647–1689.

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

Perez-Vilar, J. & Hill, R.L. (1999) The structure and assembly of secreted mucins. *The Journal of Biological Chemistry*. 274 (45), pp. 31751–31754.

Pinto, L.H., Holsinger, L.J. & Lamb, R.A. (1992) Influenza virus M2 protein has ion channel activity. *Cell*. 69 (3), pp. 517–528.

Plotch, S.J., Bouloy, M., Ulmanen, I. & Krug, R.M. (1981) A unique cap (m 7 GpppXm)-dependent influenza virion endonuclease cleaves capped RNAs to generate the primers that initiate viral RNA transcription. *Cell*. 23, pp. 847–858.

Portela, A. & Digard, P. (2002) The influenza virus nucleoprotein: a multifunctional RNA-binding protein pivotal to virus replication. *The Journal of General Virology*. 83 (Pt 4), pp. 723–734.

Pritchett, T.J. & Paulson, J.C. (1989) Basis for the potent inhibition of influenza virus infection by equine and guinea pig alpha 2-macroglobulin. *The Journal of Biological Chemistry*. 264 (17), pp. 9850–9858.

Ramos, I., Krammer, F., Hai, R., Aguilera, D., Bernal-Rubio, D., Steel, J., García-Sastre, A. & Fernandez-Sesma, A. (2013) H7N9 influenza viruses interact preferentially with α2,3-linked sialic acids and bind weakly to α2,6-linked sialic acids. *The Journal of General Virology*. 94 (Pt 11), pp. 2417–2423.

Raymond, F.L., Caton, A.J., Cox, N.J., Kendal, A.P. & Brownlee, G.G. (1983) Antigenicity and evolution amongst recent influenza viruses of H1N1 subtype. *Nucleic acids research*. 11 (20), pp. 7191–7203.

Reading, P.C., Tate, M.D., Pickett, D.L. & Brooks, A.G. (2007) Glycosylation as a target for recognition of influenza viruses by the innate immune system. In: *Advances in Experimental Medicine and Biology*. New York, NY: Springer New York. pp. pp. 279–292. doi:10.1007/978-0-387-71767-8_20.

Reid, A.H., Fanning, T.G., Hultin, J.V. & Taubenberger, J.K. (1999) Origin and evolution of the 1918 'Spanish' influenza virus hemagglutinin gene. *Proceedings of the National Academy of Sciences of the United States of America*. 96 (4), pp. 1651–1656.

Reis, dos, M., Hay, A.J. & Goldstein, R.A. (2009) Using non-homogeneous models of nucleotide substitution to identify host shift events: application to the origin of the 1918 'Spanish' influenza pandemic virus. *Journal of molecular evolution*. 69 (4), pp. 333–345.

Richard, M., Schrauwen, E.J.A., de Graaf, M., Bestebroer, T.M., Spronken, M.I.J., van Boheemen, S., de Meulder, D., Lexmond, P., Linster, M., Herfst, S., Smith, D.J., van den Brand, J.M., Burke, D.F., Kuiken, T., et al. (2013) Limited airborne transmission of H7N9 influenza A virus between ferrets. *Nature*. 501 (7468), pp. 560–563.

Richardson, J.C. & Akkina, R.K. (1991) NS2 protein of influenza virus is found in purified virus and phosphorylated in infected cells. *Archives of Virology*. 116 (1-4), pp. 69–80.

Robb, N.C., Smith, M. & Vreede, F.T. (2009) NS2/NEP protein regulates transcription and replication of the influenza virus RNA genome. *The Journal of General Virology*. 90 (6), pp. 1398–1407.

Roberts, P.C., Lamb, R.A. & Compans, R.W. (1998) The M1 and M2 Proteins of Influenza A Virus Are Important Determinants in Filamentous Particle Formation. *Virology*. 240 (1), pp. 127–137.

Robertson, J.S. (1979) 5‘ and 3’ terminal nucleotide sequences of the RNA genome segments of influenza virus. *Nucleic acids research*. 6 (12), pp. 3745–3757.

Robertson, J.S., Schubert, M. & Lazzarini, R.A. (1981) Polyadenylation sites for influenza virus mRNA. *The Journal of Virology*. 38 (1), pp. 157–163.

Rodriguez Boulan, E. & Sabatini, D.D. (1978) Asymmetric budding of viruses in epithelial monlayers: a model system for study of epithelial polarity. *Proceedings of the National Academy of Sciences of the United States of America*. 75 (10), pp. 5071–5075.

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

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

Rogers, G.N., Daniels, R.S., Skehel, J.J., Wiley, D.C., Wang, X.F., Higa, H.H. & Paulson, J.C. (1985) Host-mediated selection of influenza virus receptor variants. Sialic acid-alpha 2,6Gal-specific clones of A/duck/Ukraine/1/63 revert to sialic acid-alpha 2,3Gal-specific wild type in ovo. *The Journal of Biological Chemistry*. 260 (12), pp. 7362–7367.

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

Rossman, J.S. & Lamb, R.A. (2011) Influenza virus assembly and budding. *Virology*. 411 (2), pp. 229–236.

Rossman, J.S., Jing, X., Leser, G.P. & Lamb, R.A. (2010a) Influenza virus M2 protein mediates ESCRT-independent membrane scission. *Cell*. 142 (6), pp. 902–913.

Rossman, J.S., Jing, X., Leser, G.P., Balannik, V., Pinto, L.H. & Lamb, R.A. (2010b) Influenza virus m2 ion channel protein is necessary for filamentous virion formation. *The Journal of Virology*. 84 (10), pp. 5078–5088.

Rossman, J.S., Leser, G.P. & Lamb, R.A. (2012) Filamentous influenza virus enters cells via macropinocytosis. *The Journal of Virology*. 86 (20), pp. 10950–10960.

Royle, L., Matthews, E., Corfield, A., Berry, M. & Rudd, P.M. (2008) Glycan structures of ocular surface mucins in man, rabbit and dog display species differences. *Glycoconjugate journal*. 25pp. 763–773.

Rudino, P.E., Tunnah, P., Crennell, S.J., Webster, R.G., Laver, W.G. & Garman, E.F. (2006) *The crystal structure of type A influenza virus neuraminidase subtype N6 reveals the existence of two separate Neu5Ac binding sites.* [Accessed 13 July 2015].

Rudneva, I.A., Kovaleva, V.P., Varich, N.L. & Farashyan, V.R. (1993) Influenza A virus reassortants with surface glycoprotein genes of the avian parent viruses: effects of HA and NA gene combinations on virus aggregation. *Archives of Virology*. 133 (3-4), pp. 437–450.

Rudneva, I.A., Sklyanskaya, E.I., Barulina, O.S., Yamnikova, S.S., Kovaleva, V.P., Tsvetkova, I.V. & Kaverin, N.V. (1996) Phenotypic expression of HA-NA combinations in human-avian influenza A virus reassortants. *Archives of Virology*. 141 (6), pp. 1091–1099.

Ruigrok, R. & Andree, P.J. (1984) Characterization of three highly purified influenza virus strains by electron microscopy. *The Journal of General Virology*.

Ruigrok, R.W. (1998) Structure of influenza A, B and C viruses. In: Edited by: K G Nicholson, K G Nicholson, R G Webster, & A J Hay (eds.). *Textbook of Influenza*. Blackwell Science, Oxford. pp. pp. 29–42.

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

Ruigrok, R.W.H., Aitken, A., Calder, L.J., Martin, S.R., Skehel, J.J., Wharton, S.A., Weis, W. & Wiley, D.C. (1988) Studies on the Structure of the Influenza Virus Haemagglutinin at the pH of Membrane Fusion. *The Journal of General Virology*. 69 (11), pp. 2785–2795.

Russell, C.J. (2014) Acid-induced membrane fusion by the hemagglutinin protein and its role in influenza virus biology. In: *Influenza Pathogenesis and Control - Volume I*. Springer International Publishing. pp. pp. 93–116.

Russell, R.J., Haire, L.F., Stevens, D.J., Collins, P.J., Lin, Y.P., Blackburn, G.M., Hay, A.J., Gamblin, S.J. & Skehel, J.J. (2006) The structure of H5N1 avian influenza neuraminidase suggests new opportunities for drug design. *Nature*. 443 (7107), pp. 45–49.

Russell, R.J., Kerry, P.S., Stevens, D.J., Steinhauer, D.A., Martin, S.R., Gamblin, S.J. & Skehel, J.J. (2008) Structure of influenza hemagglutinin in complex with an inhibitor of membrane fusion. *Proceedings of the National Academy of Sciences of the United States of America*. 105 (46), pp. 17736–17741.

Ryan-Poirier, K.A. & Kawaoka, Y. (1991) Distinct glycoprotein inhibitors of influenza A virus in different animal sera. *The Journal of Virology*. 65 (1), pp. 389–395.

Sauter, N.K., Bednarski, M.D., Wurzburg, B.A., Hanson, J.E., Whitesides, G.M., Skehel, J.J. & Wiley, D.C. (1989) Hemagglutinins from two influenza virus variants bind to sialic acid derivatives with millimolar dissociation constants: a 500-MHz proton nuclear magnetic resonance study. *Biochemistry*. 28 (21), pp. 8388–8396.

Sauter, N.K., Glick, G.D., Crowther, R.L., Park, S.J., Eisen, M.B., Skehel, J.J., Knowles, J.R. & Wiley, D.C. (1992a) Crystallographic detection of a second ligand binding site in influenza virus hemagglutinin. *Proceedings of the National Academy of Sciences of the United States of America*. 89 (1), pp. 324–328.

Sauter, N.K., Hanson, J.E., Glick, G.D., Brown, J.H., Crowther, R.L., Park, S.J., Skehel, J.J. & Wiley, D.C. (1992b) 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), pp. 9609–9621.

Scholtissek, C. (1995) Molecular evolution of influenza viruses. *Virus Genes*. 11 (2), pp. 209–215.

Scholtissek, C., Hoyningen, Von, V. & Rott, R. (1978) Genetic relatedness between the new 1977 epidemic strains (H1N1) of influenza and human influenza strains isolated between 1947 and 1957 (H1N1). *Virology*. 89 (2), pp. 613–617.

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

Selman, M., Dankar, S.K., Forbes, N.E., Jia, J.-J. & Brown, E.G. (2012) Adaptive mutation in influenza A virus non-structural gene is linked to host switching and induces a novel protein by alternative splicing. *Emerging Microbes & Infections*. 1 (11), pp. e42.

Sencer, D.J. & Millar, J.D. (2006) Reflections on the 1976 swine flu vaccination program. *Emerging Infectious Diseases*. 12 (1), pp. 29–33.

Shi, Y., Zhang, W., Wang, F., Qi, J., Wu, Y., Song, H., Gao, F., Bi, Y., Zhang, Y., Fan, Z., Qin, C., Sun, H., Liu, J., Haywood, J., et al. (2013) Structures and receptor binding of hemagglutinins from human-infecting H7N9 influenza viruses. *Science*. 342 (6155), pp. 243–247.

Shinde, V., Bridges, C.B., Uyeki, T.M., Shu, B., Balish, A., Xu, X., Lindstrom, S., Gubareva, L.V., Deyde, V., Garten, R.J., Harris, M., Gerber, S., Vagasky, S., Smith, F., et al. (2009) Triple-reassortant swine influenza A (H1) in humans in the United States, 2005-2009. *The New England journal of medicine*. 360 (25), pp. 2616–2625.

Shrestha, S.S., Swerdlow, D.L., Borse, R.H., Prabhu, V.S., Finelli, L., Atkins, C.Y., Owusu-Edusei, K., Bell, B., Mead, P.S., Biggerstaff, M., Brammer, L., Davidson, H., Jernigan, D., Jhung, M.A., et al. (2011) Estimating the burden of 2009 pandemic influenza A (H1N1) in the United States (April 2009-April 2010). *Clinical infectious diseases*. 52 (Suppl 1), pp. S75–S82.

Sieben, C., Kappel, C., Zhu, R., Wozniak, A., Rankl, C., Hinterdorfer, P., Grubmüller, H. & Herrmann, A. (2012) Influenza virus binds its host cell using multiple dynamic interactions. *Proceedings of the National Academy of Sciences of the United States of America*. 109 (34), pp. 13626–13631.

Skehel, J.J. & Hay, A.J. (1978) Nucleotide sequences at the 5′ termini of influenza virus RNAs and their transcripts. *Nucleic acids research*. 5 (4), pp. 1207–1219.

Skehel, J.J. & Wiley, D.C. (2000) Receptor binding and membrane fusion in virus entry: the influenza hemagglutinin. *Annual review of biochemistry*. 69 (1), pp. 531–569.

Skehel, J.J., Bayley, P.M., Brown, E.B., Martin, S.R., Waterfield, M.D., White, J.M., Wilson, I.A. & Wiley, D.C. (1982) Changes in the conformation of influenza virus hemagglutinin at the pH optimum of virus-mediated membrane fusion. *Proceedings of the National Academy of Sciences of the United States of America*. 79 (4), pp. 968–972.

Smith, A.M. & McCullers, J.A. (2014) Secondary bacterial infections in influenza virus infection pathogenesis. In: *Influenza Pathogenesis and Control - Volume I*. Springer International Publishing. pp. pp. 327–356.

Smith, G.L. & Hay, A.J. (1982) Replication of the influenza virus genome. *Virology*. 118, pp. 96–108.

Song, X., Lasanajak, Y., Xia, B., Heimburg-Molinaro, J., Rhea, J.M., Ju, H., Zhao, C., Molinaro, R.J., Cummings, R.D. & Smith, D.F. (2011) Shotgun glycomics: a microarray strategy for functional glycomics. *Nature methods*. 8 (1), pp. 85–90.

Sorrell, E.M., Song, H., Pena, L. & Perez, D.R. (2010) A 27-amino-acid deletion in the neuraminidase stalk supports replication of an avian H2N2 influenza A virus in the respiratory tract of chickens. *The Journal of Virology*. 84 (22), pp. 11831–11840.

Stieneke-Gröber, A., Vey, M., Angliker, H., Shaw, E., Thomas, G., Roberts, C., Klenk, H.D. & Garten, W. (1992) Influenza virus hemagglutinin with multibasic cleavage site is activated by furin, a subtilisin-like endoprotease. *The EMBO journal*. 11 (7), pp. 2407–2414.

Sun, X., Li, Q., Wu, Y., Wang, M., Liu, Y., Qi, J., Vavricka, C.J. & Gao, G.F. (2014) Structure of influenza virus N7: the last piece of the neuraminidase 'jigsaw' puzzle. *The Journal of Virology*. 88 (16), pp. 9197–9207.

Sun, X., Shi, Y., Lu, X., He, J., Gao, F., Yan, J., Qi, J. & Gao, G.F. (2013a) Bat-derived influenza hemagglutinin H17 does not bind canonical avian or human receptors and most likely uses a unique entry mechanism. *Cell reports*. 3 (3), pp. 769–778.

Sun, Y., Tan, Y., Wei, K., Sun, H., Shi, Y., Pu, J., Yang, H., Gao, G.F., Yin, Y., Feng, W., Perez, D.R. & Liu, J. (2013b) Amino acid 316 of hemagglutinin and the neuraminidase stalk length influence virulence of H9N2 influenza virus in chickens and mice. *The Journal of Virology*. 87 (5), pp. 2963–2968.

Sung, J.C., Wynsberghe, A. & Amaro, R.E. (2010) Role of secondary sialic acid binding sites in influenza N1 neuraminidase. *Journal of the American Chemical Society*.

Taubenberger, J.K. & Kash, J.C. (2010) Influenza virus evolution, host adaptation, and pandemic formation. *Cell host & microbe*. 7 (6), pp. 440–451.

Taubenberger, J.K. & Morens, D.M. (2006) 1918 Influenza: the mother of all pandemics. *Rev Biomed*.

Taubenberger, J.K. & Morens, D.M. (2009) Pandemic influenza--including a risk assessment of H5N1. *Revue scientifique et technique*. 28 (1), pp. 187–202.

Taubenberger, J.K. & Morens, D.M. (2008) The pathology of influenza virus infections. *Annual review of pathology*. 3 pp. 499–522.

Taubenberger, J.K., Reid, A.H., Krafft, A.E., Bijwaard, K.E. & Fanning, T.G. (1997) Initial genetic characterization of the 1918 'Spanish' influenza virus. *Science*. 275 (5307), pp. 1793–1796.

Tellier, R. (2006) Review of aerosol transmission of influenza A virus. *Emerging Infectious Diseases*. 12 (11), pp. 1657–1662.

Théry, C., Zitvogel, L. & Amigorena, S. (2002) Exosomes: composition, biogenesis and function. *Nature reviews. Immunology*. 2 (8), pp. 569–579.

Thompson, C.I., Barclay, W.S., Zambon, M.C. & Pickles, R.J. (2006) Infection of human airway epithelium by human and avian strains of influenza a virus. *The Journal of Virology*. 80 (16), pp. 8060–8068.

Thornton, D.J., Rousseau, K. & McGuckin, M.A. (2008) Structure and function of the polymeric mucins in airways mucus. *Annu Rev Physiol*. 70 (1), pp. 459–486.

Tong, S., Li, Y., Rivailler, P., Conrardy, C., Castillo, D.A.A., Chen, L.-M., Recuenco, S., Ellison, J.A., Davis, C.T., York, I.A., Turmelle, A.S., Moran, D., Rogers, S., Shi, M., et al. (2012) A distinct lineage of influenza A virus from bats. *Proceedings of the National Academy of Sciences of the United States of America*. 109 (11), pp. 4269–4274.

Tong, S., Zhu, X., Li, Y., Shi, M., Zhang, J., Bourgeois, M., Yang, H., Chen, X., Recuenco, S., Gomez, J., Chen, L.-M., Johnson, A., Tao, Y., Dreyfus, C., et al. (2013) New world bats harbor diverse influenza A viruses. *PLoS pathogens*. 9 (10), pp. e1003657.

Tsuchiya, E., Sugawara, K., Hongo, S., Matsuzaki, Y., Muraki, Y., Li, Z.-N. & Nakamura, K. (2002) Effect of addition of new oligosaccharide chains to the globular head of influenza A/H2N2 virus haemagglutinin on the intracellular transport and biological activities of the molecule. *The Journal of General Virology*. 83 (Pt 5), pp. 1137–1146.

Tsurudome, M., Glück, R., Graf, R., Falchetto, R., Schaller, U. & Brunner, J. (1992) Lipid interactions of the hemagglutinin HA2 NH2-terminal segment during influenza virus-induced membrane fusion. *The Journal of Biological Chemistry*. 267 (28), pp. 20225–20232.

Tumpey, T.M., Basler, C.F., Aguilar, P.V., Zeng, H., Solórzano, A., Swayne, D.E., Cox, N.J., Katz, J.M., Taubenberger, J.K., Palese, P. & García-Sastre, A. (2005) Characterization of the reconstructed 1918 Spanish influenza pandemic virus. *Science*. 310 (5745), pp. 77–80.

Uhlendorff, J., Matrosovich, T., Klenk, H.-D. & Matrosovich, M. (2009) Functional significance of the hemadsorption activity of influenza virus neuraminidase and its alteration in pandemic viruses. *Archives of Virology*. 154 (6), pp. 945–957.

Underwood, P.A. (1982) Mapping of antigenic changes in the haemagglutinin of Hong Kong influenza (H3N2) strains using a large panel of monoclonal antibodies. *The Journal of General Virology*. 62 (1), pp. 153–169.

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

van der Vries, E., Collins, P.J., Vachieri, S.G., Xiong, X., Liu, J., Walker, P.A., Haire, L.F., Hay, A.J., Schutten, M., Osterhaus, A.D.M.E., Martin, S.R., Boucher, C.A.B., Skehel, J.J. & Gamblin, S.J. (2012) H1N1 2009 pandemic influenza virus: resistance of the I223R neuraminidase mutant explained by kinetic and structural analysis. *PLoS pathogens*. 8 (9), pp. e1002914.

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

Vandegrift, K.J., Sokolow, S.H., Daszak, P. & Kilpatrick, A.M. (2010) Ecology of avian influenza viruses in a changing world. *Annals of the New York Academy of Sciences*. 1195pp. 113–128.

Varghese, J.N., Colman, P.M., van Donkelaar, A., Blick, T.J., Sahasrabudhe, A. & McKimm-Breschkin, J.L. (1997) Structural evidence for a second sialic acid binding site in avian influenza virus neuraminidases. *Proceedings of the National Academy of Sciences of the United States of America*. 94 (22), pp. 11808–11812.

Varghese, J.N., Laver, W.G. & Colman, P.M. (1983) Structure of the influenza virus glycoprotein antigen neuraminidase at 2.9 Å resolution. *Nature*. 303 (5912), pp. 35–40.

Varghese, J.N., McKimm-Breschkin, J.L., Caldwell, J.B., Kortt, A.A. & Colman, P.M. (1992) The structure of the complex between influenza virus neuraminidase and sialic acid, the viral receptor. *Proteins*. 14 (3), pp. 327–332.

Vavricka, C.J., Liu, Y., Kiyota, H., Sriwilaijaroen, N., Qi, J., Tanaka, K., Wu, Y., Li, Q., Li, Y., Yan, J., Suzuki, Y. & Gao, G.F. (2013) Influenza neuraminidase operates via a nucleophilic mechanism and can be targeted by covalent inhibitors. *Nature communications*. 4, pp. 1491.

Vigerust, D.J. & Shepherd, V.L. (2007) Virus glycosylation: role in virulence and immune interactions. *Trends in microbiology*. 15 (5), pp. 211–218.

Vincent, A.L., Ma, W., Lager, K.M., Janke, B.H. & Richt, J.A. (2008) Swine influenza viruses: a North American perspective. *Advances in virus research*. 72, pp. 127–154.

Wagner, R., Wolff, T., Herwig, A., Pleschka, S. & Klenk, H.D. (2000) Interdependence of hemagglutinin glycosylation and neuraminidase as regulators of influenza virus growth: a study by reverse genetics. *The Journal of Virology*. 74 (14), pp. 6316–6323.

Walther, T., Karamanska, R., Chan, R.W.Y., Chan, M.C.W., Jia, N., Air, G., Hopton, C., Wong, M.P., Dell, A., Malik Peiris, J.S., Haslam, S.M. & Nicholls, J.M. (2013) Glycomic Analysis of Human Respiratory Tract Tissues and Correlation with Influenza Virus Infection. *PLoS pathogens*. 9 (3), pp. e1003223.

Wang, C.-C., Chen, J.-R., Tseng, Y.-C., Hsu, C.-H., Hung, Y.-F., Chen, S.-W., Chen, C.-M., Khoo, K.-H., Cheng, T.-J., Cheng, Y.-S.E., Jan, J.-T., Wu, C.-Y., Ma, C. & Wong, C.-H. (2009) Glycans on influenza hemagglutinin affect receptor binding and immune response. *Proceedings of the National Academy of Sciences of the United States of America*. 106 (43), pp. 18137–18142.

Watowich, S.J., Skehel, J.J. & Wiley, D.C. (1994) Crystal structures of influenza virus hemagglutinin in complex with high-affinity receptor analogs. *Structure*. 2 (8), pp. 719–731.

Webby, R.J., Swenson, S.L., Krauss, S.L., Gerrish, P.J., Goyal, S.M. & Webster, R.G. (2000) Evolution of swine H3N2 influenza viruses in the United States. *The Journal of Virology*. 74 (18), pp. 8243–8251.

Weber, T., Paesold, G., Galli, C., Mischler, R., Semenza, G. & Brunner, J. (1994) Evidence for H(+)-induced insertion of influenza hemagglutinin HA2 N-terminal segment into viral membrane. *The Journal of Biological Chemistry*. 269 (28), pp. 18353–18358.

Weerapana, E. & Imperiali, B. (2006) Asparagine-linked protein glycosylation: from eukaryotic to prokaryotic systems. *Glycobiology*. 16 (6), pp. 91R–101R.

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

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

Wiley, D.C. & Skehel, J.J. (1987) The structure and function of the hemagglutinin membrane glycoprotein of influenza virus. *Annual review of biochemistry*. 56 (1), pp. 365–394.

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

Wilson, I., Skehel, J. & Wiley, D.C. (1981) Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3Å resolution. *Nature*. 289 pp. 366–373.

Wise, H.M., Foeglein, A., Sun, J., Dalton, R.M., Patel, S., Howard, W., Anderson, E.C., Barclay, W.S. & Digard, P. (2009) A complicated message: Identification of a novel PB1-related protein translated from influenza A virus segment 2 mRNA. *The Journal of Virology*. 83 (16), pp. 8021–8031.

Wise, H.M., Hutchinson, E.C., Jagger, B.W., Stuart, A.D., Kang, Z.H., Robb, N., Schwartzman, L.M., Kash, J.C., Fodor, E., Firth, A.E., Gog, J.R., Taubenberger, J.K. & Digard, P. (2012) Identification of a novel splice variant form of the influenza A virus M2 ion channel with an antigenically distinct ectodomain. *PLoS pathogens*. 8 (11), pp. e1002998.

Wong, S.-S. & Webby, R.J. (2013) Traditional and new influenza vaccines. *Clinical microbiology reviews*. 26 (3), pp. 476–492.

World Health Organisation (2014a) *Influenza (Seasonal)*. Available from: http://www.who.int/mediacentre/factsheets/fs211/en/ [Accessed 19 April 2015a].

World Health Organisation (2015a) *Influenza at the human-animal interface*. Available from: http://www.who.int/influenza/human_animal_interface/Influenza_Summary_IRA_HA_interface_23_June_2015.pdf?ua=1 [Accessed 2 July 2015a].

World Health Organisation (2014b) *Recommended composition of influenza virus vaccines for use in the 2015 southern hemisphere influenza season*. Available from: http://www.who.int/influenza/vaccines/virus/recommendations/201409_recommendation.pdf?ua=1 [Accessed 17 April 2015b].

World Health Organisation (2015b) *Recommended composition of influenza virus vaccines for use in the 2015–2016 northern hemisphere influenza season.* Available from: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=25771542&retmode=ref&cmd=prlinks [Accessed 17 April 2015b].

World Health Organisation (2010) *WHO Guidelines for Pharmacological Management of Pandemic Influenza A(H1N1) 2009 and other Influenza Viruses*. Available from: http://www.who.int/csr/resources/publications/swineflu/h1n1_guidelines_pharmaceutical_mngt.pdf?ua=1 [Accessed 17 April 2015].

World Health Organisation (2015c) *WHO risk assessment of Human infections with avian influenza A(H7N9) virus*. Available from: http://www.who.int/influenza/human_animal_interface/influenza_h7n9/RiskAssessment_H7N9_23Feb20115.pdf?ua=1 [Accessed 15 April 2015c].

Wu, W. & Air, G.M. (2004) Binding of influenza viruses to sialic acids: reassortant viruses with A/NWS/33 hemagglutinin bind to alpha2,8-linked sialic acid. *Virology*. 325 (2), pp. 340–350.

Wu, Y., Bi, Y., Vavricka, C.J., Sun, X., Zhang, Y., Gao, F., Zhao, M., Xiao, H., Qin, C., He, J., Liu, W., Yan, J., Qi, J. & Gao, G.F. (2013) Characterization of two distinct neuraminidases from avian-origin human-infecting H7N9 influenza viruses. *Cell research*. 23 (12), pp. 1347–1355.

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

Xiong, X., Martin, S.R., Haire, L.F., Wharton, S.A., Daniels, R.S., Bennett, M.S., McCauley, J.W., Collins, P.J., Walker, P.A., Skehel, J.J. & Gamblin, S.J. (2013b) Receptor binding by an H7N9 influenza virus from humans. *Nature*. 499 (7459), pp. 496–499.

Xiong, X., McCauley, J.W. & Steinhauer, D.A. (2014) Receptor binding properties of the influenza virus hemagglutinin as a determinant of host range. In: *Influenza Pathogenesis and Control - Volume I*. pp. pp. 63–91.

Xiong, X., Tuzikov, A., Coombs, P.J., Martin, S.R., Walker, P.A., Gamblin, S.J., Bovin, N. & Skehel, J.J. (2013c) Recognition of sulphated and fucosylated receptor sialosides by A/Vietnam/1194/2004 (H5N1) influenza virus. *Virus research*. 178 (1), pp. 12–14.

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

Xu, L., Bao, L., Deng, W., Dong, L., Zhu, H., Chen, T., Lv, Q., Li, F., Yuan, J., Xiang, Z., Gao, K., Xu, Y., Huang, L., Li, Y., et al. (2014a) Novel avian-origin human influenza A(H7N9) can be transmitted between ferrets via respiratory droplets. *The Journal of infectious diseases*. 209 (4), pp. 551–556.

Xu, L., Bao, L., Deng, W., Zhu, H., Li, F., Chen, T., Lv, Q., Yuan, J., Xu, Y., Li, Y., Yao, Y., Gu, S., Yu, P., Chen, H., et al. (2014b) Rapid adaptation of avian H7N9 virus in pigs. *Virology*. 452-453pp. 231–236.

Xu, R., Ekiert, D.C., Krause, J.C., Hai, R., Crowe, J.E. & Wilson, I.A. (2010) Structural basis of preexisting immunity to the 2009 H1N1 pandemic influenza virus. *Science*. 328 (5976), pp. 357–360.

Xu, R., McBride, R., Nycholat, C.M., Paulson, J.C. & Wilson, I.A. (2012a) Structural characterization of the hemagglutinin receptor specificity from the 2009 H1N1 influenza pandemic. *The Journal of Virology*. 86 (2), pp. 982–990.

Xu, R., Zhu, X., McBride, R., Nycholat, C.M., Yu, W., Paulson, J.C. & Wilson, I.A. (2012b) Functional balance of the hemagglutinin and neuraminidase activities accompanies the emergence of the 2009 H1N1 influenza pandemic. *The Journal of Virology*. 86 (17), pp. 9221–9232.

Xu, X., Zhu, X., Dwek, R.A., Stevens, J. & Wilson, I.A. (2008) Structural characterization of the 1918 influenza virus H1N1 neuraminidase. *The Journal of Virology*. 82 (21), pp. 10493–10501.

Yang, H., Carney, P. & Stevens, J. (2010) Structure and Receptor binding properties of a pandemic H1N1 virus hemagglutinin. *PloS Currents*. 2pp. RRN1152.

Yang, H., Carney, P.J., Chang, J.C., Villanueva, J.M. & Stevens, J. (2013) Structural analysis of the hemagglutinin from the recent 2013 H7N9 influenza virus. *The Journal of Virology*. 87 (22), pp. 12433–12446.

Yang, X., Steukers, L., Forier, K., Xiong, R., Braeckmans, K., Van Reeth, K. & Nauwynck, H. (2014) A beneficiary role for neuraminidase in influenza virus penetration through the respiratory mucus. *PloS one*. 9 (10), pp. e110026.

Yen, H.-L., Liang, C.-H., Wu, C.-Y., Forrest, H.L., Ferguson, A., Choy, K.-T., Jones, J., Wong, D.D.-Y., Cheung, P.P.-H., Hsu, C.-H., Li, O.T., Yuen, K.M., Chan, R.W.Y., Poon, L.L.M., et al. (2011) Hemagglutinin-neuraminidase balance confers respiratory-droplet transmissibility of the pandemic H1N1 influenza virus in ferrets. *Proceedings of the National Academy of Sciences of the United States of America*. 108 (34), pp. 14264–14269.

Yoon, S.-W., Webby, R.J. & Webster, R.G. (2014) Evolution and ecology of influenza A viruses. In: *Influenza Pathogenesis and Control - Volume I*. Current Topics in Microbiology and Immunology. Springer International Publishing. pp. pp. 359–375. doi:10.1007/82_2014_396.

Zhang, J. & Lamb, R.A. (1996) Characterization of the membrane association of the influenza virus matrix protein in living cells. *Virology*. 225 (2), pp. 255–266.

Zhou, B., Ma, J., Liu, Q., Bawa, B., Wang, W., Shabman, R.S., Duff, M., Lee, J., Lang, Y., Cao, N., Nagy, A., Lin, X., Stockwell, T.B., Richt, J.A., et al. (2014) Characterization of uncultivable bat influenza virus using a replicative synthetic virus. Volker Thiel (ed.). *PLoS pathogens*. 10 (10), pp. e1004420.

Zhou, H., Yu, Z., Hu, Y., Tu, J., Zou, W., Peng, Y., Zhu, J., Li, Y., Zhang, A., Yu, Z., Ye, Z., Chen, H. & Jin, M. (2009) The Special Neuraminidase Stalk-Motif Responsible for Increased Virulence and Pathogenesis of H5N1 Influenza A Virus. *PloS one*. 4 (7), pp. e6277.

Zhou, J., Wang, D., Gao, R., Zhao, B., Song, J., Qi, X., Zhang, Y., Shi, Y., Yang, L., Zhu, W., Bai, T., Qin, K., Lan, Y., Zou, S., et al. (2013) Biological features of novel avian influenza A (H7N9) virus. *Nature*. 499 (7459), pp. 500–503.

Zhou, N.N., Senne, D.A., Landgraf, J.S., Swenson, S.L., Erickson, G., Rossow, K., Liu, L., Yoon, K.J., Krauss, S. & Webster, R.G. (1999) Genetic reassortment of avian, swine, and human influenza A viruses in American pigs. *The Journal of Virology*. 73 (10), pp. 8851–8856.

Zhu, H., Wang, D., Kelvin, D.J., Li, L., Zheng, Z., Yoon, S.-W., Wong, S.-S., Farooqui, A., Wang, J., Banner, D., Chen, R., Zheng, R., Zhou, J., Zhang, Y., et al. (2013a) Infectivity, transmission, and pathology of human-isolated H7N9 influenza virus in ferrets and pigs. *Science*. 341 (6142), pp. 183–186.

Zhu, X., McBride, R., Nycholat, C.M., Yu, W., Paulson, J.C. & Wilson, I.A. (2012a) Influenza virus neuraminidases with reduced enzymatic activity that avidly bind sialic Acid receptors. *The Journal of Virology*. 86 (24), pp. 13371–13383.

Zhu, X., Yang, H., Guo, Z., Yu, W., Carney, P.J., Li, Y., Chen, L.-M., Paulson, J.C., Donis, R.O., Tong, S., Stevens, J. & Wilson, I.A. (2012b) Crystal structures of two subtype N10 neuraminidase-like proteins from bat influenza A viruses reveal a diverged putative active site. *Proceedings of the National Academy of Sciences of the United States of America*. 109 (46), pp. 18903–18908.

Zhu, X., Yu, W., McBride, R., Li, Y., Chen, L.-M., Donis, R.O., Tong, S., Paulson, J.C. & Wilson, I.A. (2013b) Hemagglutinin homologue from H17N10 bat influenza virus exhibits divergent receptor-binding and pH-dependent fusion activities. *Proceedings of the National Academy of Sciences of the United States of America*. 110 (4), pp. 1458–1463.

Zimmer, S.M. & Burke, D.S. (2009) Historical perspective--Emergence of influenza A (H1N1) viruses. *The New England journal of medicine*. 361 (3), pp. 279–285.

Appendix

N9 NA expression construct sequence

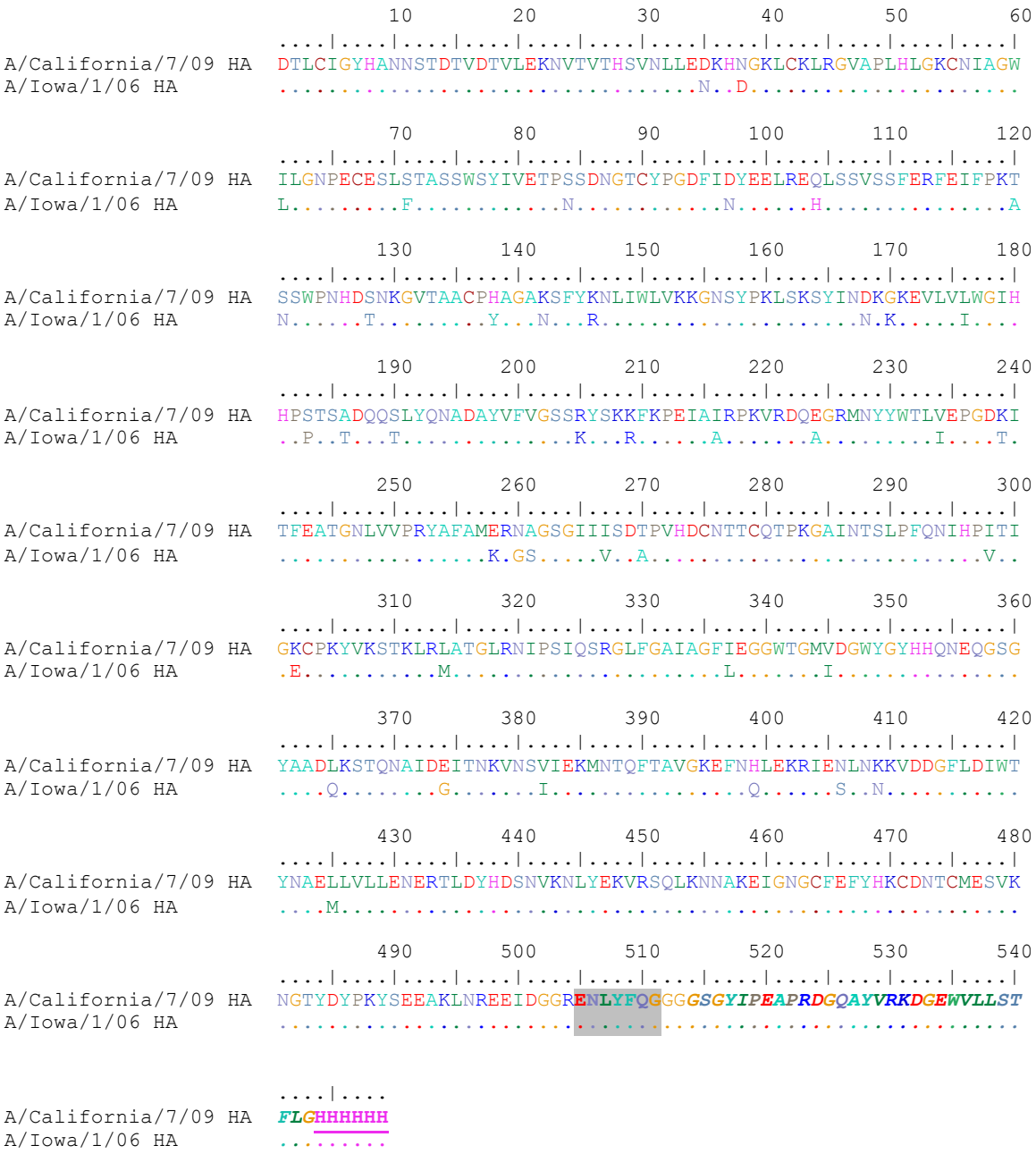
```
                      10        20        30        40        50        60
                ....|....|....|....|....|....|....|....|....|....|....|....|
A/Anhui/1/2013  HHHHHHSSSDYSDLQRVKQELLEEVKKELQKVKEEIIEAFVQELRKRGENLYFQGPSRRT

                      70        80        90       100       110       120
                ....|....|....|....|....|....|....|....|....|....|....|....|
A/Anhui/1/2013  SRNFNNLTKGLCTINSWHIYGKDNAVRIGESSDVLVTREPYVSCDPDECRFYALSQGTTI

                     130       140       150       160       170       180
                ....|....|....|....|....|....|....|....|....|....|....|....|
A/Anhui/1/2013  RGKHSNGTIHDRSQYRALISWPLSSPPTVYNSRVECIGWSSTSCHDGKSRMSICISGPNN

                     190       200       210       220       230       240
                ....|....|....|....|....|....|....|....|....|....|....|....|
A/Anhui/1/2013  NASAVVWYNRRPVAEINTWARNILRTQESECVCHNGVCPVVFTDGSATGPADTRIYYFKE

                     250       260       270       280       290       300
                ....|....|....|....|....|....|....|....|....|....|....|....|
A/Anhui/1/2013  GKILKWESLTGTAKHIEECSCYGERTGITCTCRDNWQGSNRPVIQIDPVAMTHTSQYICS

                     310       320       330       340       350       360
                ....|....|....|....|....|....|....|....|....|....|....|....|
A/Anhui/1/2013  PVLTDNPRPNDPNIGKCNDPYPGNNNNGVKGFSYLDGANTWLGRTISTASRSGYEMLKVP

                     370       380       390       400       410       420
                ....|....|....|....|....|....|....|....|....|....|....|....|
A/Anhui/1/2013  NALTDDRSKPIQGQTIVLNADWSGYSGSFMDYWAEGDCYRACFYVELIRGRPKEDKVWWT

                     430       440
                ....|....|....|....|....|....
A/Anhui/1/2013  SNSIVSMCSSTEFLGQWNWPDGAKIEYFL
```

Sequence of expression construct for A/Anhui/1/2013 NA ectodomain. Residues 1 – 6 are the hexa-his tag (underlined), residues 7 – 47 are the vasodilator-stimulated phosphoprotein (VASP) tetramerisation domain (italics) and residues 49 – 55 are a TEV protease recognition sequence (shaded). Residues 59 – 449 correspond to residues 75 – 465 of the viral NA open reading frame.

Construct was under the control of the polyhedrin promoter and also contained an N-terminal gp67 secretion signal peptide* not shown in construct sequences as it is post-translationally cleaved.

* gp67 signal peptide: MLLVNQSHQGFNKEHTSKMVSAIVLYVLLAAAAHSAFA

H1 HA expression construct sequences



Sequence expression constructs for Cal7 and Iowa06 HA ectodomain expression. Residues 1 – 502 equate to residues 1 – 502 (H1 numbering) of the HA after the removal of the signal peptide (first 17 residues of the H1 HA open reading frame). Residues 505 – 511 consist of a TEV protease cleavage site (shaded); residues 514 – 543 consist of a foldon trimerisation motif (italics) and residues 544 – 549 consist of a hexa-his-tag (underlined).

Construct was under the control of the polyhedrin promoter and also contained an N-terminal gp67 secretion signal peptide* not shown in construct sequences as it is post-translationally cleaved.

* gp67 signal peptide: MLLVNQSHQGFNKEHTSKMVSAIVLYVLLAAAAHSAFA