Characterising Recent H1N1 Glycosylation and Its Implication for Influenza A Vaccine Effectiveness

Alexander Sivesgaard Bolotakis

Master's Thesis | Nanobiotechnology | 2016-2017

AALBORG UNIVERSITY
DENMARK

STATENS
SERUM
INSTITUT

AALBORG UNIVERSITY
Students' report

STATENS SERUM INSTITUT

Department of
Physics and Nanotechnology
Nanotechnology
Skjernvej 4A
9220 Aalborg E

Department of
Virological Surveillance
and Research
Artillerivej 5
2300 Copenhagen S

Title:

Characterising Recent H1N1 Glycosylation and Its Implication for Influenza A Vaccine Effectiveness

Theme:

Master Thesis

Project period:

P9-10, fall & autumn semesters, 2016-17

Author:

Alexander Sivesgaard Bolotakis

Supervisors:

Leonid Gurevich
Ramona Trebbien

Number of copies: 4
Number of pages: 100
Number of appendices: 2
Finished: 11.06.2017

Abstract

The flu season of 2015-2016 saw a decrease in the effectiveness of the influenza A vaccine, potentially caused by the introduction of new glycosylations in the hemagglutinin (HA) of H1N1.
23 seasonal influenza A virus (IAV) samples, spread across five seasons with root in the 2009 pandemic H1N1 outbreak (A(H1N1)pdm2009), are antigenically characterised by immunological assays and sequence analysis, to determine whether glycosylation is responsible for the decrease in vaccine effectiveness.
Similarly, characterisation with immunoblots and surface plasmon resonance (SPR) measurements is attempted to quantify antigenic differences.
Six out of eight of the surveyed IAV samples from the 2015-2016 season have new glycosylation sites at Ser179 of HA, caused by a substitution to Asn.
Investigating resistance towards antibodies raised to A(H1N1)pdm2009 by means of hemagglutiniation inhibition assays unveils a predominance for incremented fold-change factors in IAV samples harbouring the Ser179→Asn substitution, as observed in IAV samples A/Denmark/30/16 and A/Denmark/46/16.
Both samples are considered to be glycosylated, which makes it a probable explanation for the decrease in influenza A vaccine effectiveness.
Using SPR as a characterization technique of antigenic properties of IAVs requires refining. Studying analyte type and composition is the first step in implementing this method for influenza vaccine immunology.

AALBORG UNIVERSITY
Students' report

STATENS SERUM INSTITUT

Institut for
Fysik og Nanoteknologi
Nanoteknologi
Skjernvej 4A
9220 Aalborg Ø

Institut for
Virologisk Overvågning
og Forskning
Artillerivej 5
2300 København S

Titel:

Karakterisering af Nylig H1N1 Glykosylation og Dennes Implikation for Influenza A Vaccine Effektivitet

Projekt type:

Kandidatspeciale

Projekt periode:

P9-10, forår- og efterårs-semestrene, 2016-17

Forfatter:

Alexander Sivesgaard Bolotakis

Vejledere:

Ramona Trebbien
Leonid Gurevich

Antal kopier: 4
Antal sider: 100
Antal appendikser: 2
Afleveret: 11.06.2017

Synopsis

Influenza sæsonen 2015-2016 oplevede et fald i effektiviteten hos influenza A vaccinen, hvilket potentielt skyldes tilføjelsen af nye glykosylationer i hæmagglutinin (HA) hos H1N1.
23 influenza A virus (IAV) prøver, spredt over fem influenza sæsoner med rod i det pandemiske udbrud i 2009 (A(H1N1)pdm2009), bliver karakteriseret antigenisk vha. immunologiske eksperimenter og sekvensanalyser, for at bestemme hvorvidt nye glykosyleringer er ansvarlige for faldet i IAV vaccine effektiviteten.
Derudover bliver IAV prøverne karakteriseret med immunoblots og overfladeplasmonresonans målinger for at kvantificere deres antigene forskelle.
Seks ud af de otte undersøgte IAV prøver fra 2015-2016 sæsonen har tilføjet en ny glykosyleringsposition ved Ser179 hos HA, grundet et skift til Asn.
Undersøgelse af resistance mod antistoffer groet med udgangspunkt i A(H1N1)pdm2009 viste, at der var en tendens for højere fold-ændringsfaktorer i IAV prøver med Ser179→Asn substitutionen, hvilket er observeret i prøverne A/Denmark/30/16 og A/Denmark/46/16.
Begge prøver forventes glykosylerede, hvilket udgører en sandsynlig begrundelse for faldet i IAV vaccine effektiviteten.
Mere forståelse er nødvendigt, hvis overfladeplasmonresonans målinger skal benyttes til at analysere antigene egenskaber. Implementering af denne teknik inden for influenza vaccine immunologi kræver dybdegående analyse af analytens komposition og type.

Preface

This scientific report details the work of the author during a Master's thesis spanning two semesters, and is completed in collaboration with Statens Serum Institut (SSI) as a part of the Nanobiotechnology Master's Degree at Aalborg University (AAU). The supervisors assigned to the Master's thesis were Ramona Trebbien from SSI and Leonid Gurevich from AAU.

Numbers are assigned to each reference as listed in the bibliography, and appear in the report in square brackets, for instance: [5]. Should the reference be positioned after a full stop it implies that the reference is tied to all preceding text, until obstructed by another reference or paragraph. References found before a full stop or a comma signify adherence to the preceding sentence only. A reference mid-sentence serves as a placeholder for the full name of a paper. Figure captions without references are created by the author of the report.

The author wishes to thank Bente Andersen, Jesper Rønn and Dennis Jelsbak Schmidt for their help and guidance during experimental procedures. Furthermore, a special thanks goes to Anna Clarice Sotto Abacan for her help with corrections and spell checking.

Alexander Sivesgaard Bolotakis

Contents

1. Introduction

It is almost safe to assume that nearly everyone has become acquainted with the disease known as the flu, either directly or through others [1]. The disease is caused by the aptly named influenza virus, which has been around since B.C. according to historians [2]. It was, however, not until the beginning of the 20th century that influenza was first isolated from poultry, and then later from pigs, humans and waterfowl [3]. It has since been uncovered that the source of all influenza virus is waterfowl, although other interspecies transmissions are possible, for instance, between pigs and humans [3].

There are three recognised types of influenza viruses, namely influenza A virus (IAV), B and C. The latter two are only found in humans and pigs, while IAV is prevalent in all the previously mentioned species and others [2]. Of particular interest are the IAVs that circulate in humans on a seasonal basis, such as H1N1, commonly known as swine flu, and H3N2 [4]. H1N1 is the most well known of the two, seeing as it caused the devastating pandemic outbreak of 1918, commonly referred to as the Spanish flu [2], as well as the recent outbreak of 2009 [5]. Pandemic spread of the virus usually transpires due to large scale alterations in the whole genome, but the perpetuation of seasonal epidemics is handled by an inherent mutative system in the virus [6]. The unpredictablilty of influenza makes it difficult for the medicinal industry to adapt, and vaccines must continuously be developed every 1-3 years to contend with the elusiveness of the virus [7]. Coupling this with influenza-related hospitalizations and deaths during an epidemic, which can reach 300.000 and 40.000 during a single winter season, respectively [4], the financial burden can reach $10.4 billion in the United States alone [8]. At present, vaccination is highly recommended for high-risk demographics, defined by the World Health Organization (WHO) to encompass the elderly and people suffering from chronic and underlaying diseases, among others. Such groups are at a higher risk of developing severe and life threatening diseases from an influenza infection [9]. Although antiviral compounds that deal with immediate threats are available, such as Oseltamivir or Zanamivir, the ideal scenario would be to build up immunity against the virus through vaccination of the entire population, which would reduce the annual costs of dealing with this virus significantly [7].

Developing a universal vaccine is therefore of great interest, but the mutative fle-

xibility of IAV at antigenic sites makes it challenging [4]. Therefore, it is important to constantly monitor antigenic changes in the circulating virus population, in order to accommodate said changes in the next generation of vaccines. Thus, the populace is better prepared for future outbreaks. Antigenic changes are generally considered as amino acid substitutions at specified sites, but other features, such as glycosylation, also affect the antigenic properties of the virus [10].

The aim of this Master's thesis is to determine whether potentially newly acquired glycans on IAV are responsible for the decreased effectiveness of the currently employed vaccine, which is based on the 2009 pandemic virus (A(H1N1)pdm2009). This is accomplished by surveying IAV H1N1 samples, starting from the 2009 pandemic up until the influenza season of 2015-2016, to track alterations in the sequences and antigenic properties.

1.1 Influenza A virus

IAV is, along with influenza B and C viruses, a genus belonging to the *Orthomyxoviridae* family. They are all recognised as enveloped ribonucleic acid (RNA) viruses, with IAV being the most researched genus, since it regularly occurs as both pandemic and seasonal epidemics in humans, other mammals and birds [2]. The IAV virion is considered pleomorphic [3], meaning that it can have a variety of shapes, but is often occupying a filamentous shape upon isola-

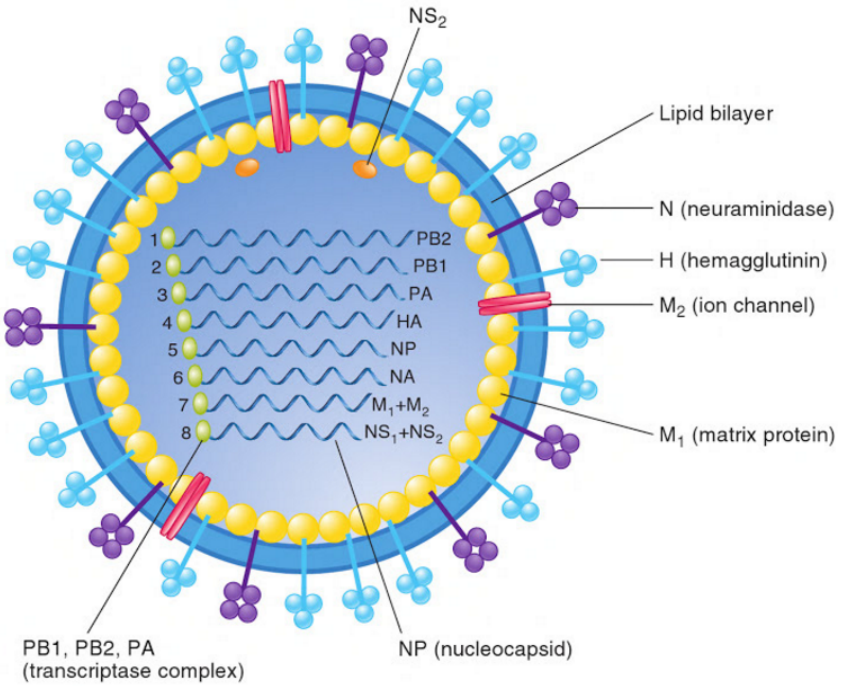

Figure 1.1: Schematic structure of the IAV virion. Eight segmented strands are depicted with a polymerase complex (green) at one end. The M1 proteins coat the inside of the lipid membrane which incorporates the surface proteins HA and NA. This image is part of a larger figure. [2]

tion [2]. However, after multiple passages through cell cultures they become generally spherical, with sizes ranging between 80-120 nm in diameter. The virion can be divided into three compartments, namely the core, which carries the genes, the viral matrix protein layer and finally the host-derived lipid membrane, called the envelope, all of which is illustrated in figure 1.1. [2]

1.1.1 Genome

The core holds eight segmented strands of negative-sense, single-stranded RNA ((-)ssRNA), each attached to several nucleoproteins (NPs) as well as a trimeric polymerase complex [2, 7]. The association of all five components is commonly referred to as a viral ribonucleoprotein (vRNP) complex [2, 7]. The remaining gene products are hemagglutinin (HA), neuraminidase (NA), matrix 1 (M1) and 2 (M2), non-structural 1 (NS1) and 2 (NS2), polymerase basic 1 (PB1) and 2 (PB2), and polymerase acidic (PA) [2, 3]. Additionally, there is another gene product, PB1-F2, encoded by an open reading frame within the PB1 reading frame [11]. This protein is involved in virus-induced cell death, targeted at the inner mitochondrial membrane [11]. Thus, the eight (-)ssRNA segments code for a total of 11 gene products, and are all listed in table 1.1 [2].

As evident from figure 1.1, M1 and M2 are both encoded segment 7. M1 is a matrix protein, and also the most abundant protein in the entire IAV virion. They are arranged in a shell, coating the inside of the lipid envelope, and together enclose all eight vRNP complexes as seen in figure 1.1. The M2 proteins assemble into a tetrameric membrane protein, which is used to form ion channels through the viral envelope. [2, 3]

Segment 8 encodes NS1 and NS2. NS2 is involved in the transfer of vRNPs out of the host cell nucleus, while NS1 is absent in the virion and can be considered non-essential for virulence. However, it is present in the host cell, where it combats host immune responses, such as interferons. [2]

The proteins involved in the vRNP, i.e. PB1, PB2, PA and NP, are encoded in segments 2, 1, 3 and 5, respectively [3, 7]. The vRNP complex plays a role in the transcription of viral RNA (vRNA) in the

Segment	Genes	Function	Length (nts)
1	PB2	Binding vRNP to host mRNA	2341
2	PB1 + PB1-F2	Catalyses transcription and cell death	2341
3	PA	Cleaves cap of host mRNA	2233
4	HA	Binding and fusion of virion and host membranes	2073
5	NP	Structural enhancements	1565
6	NA	Release of progeny virus	1413
7	M1 + M2	Forms the membrane and binds vRNA. Ion channel	1027
8	NS1 + NS2	vRNP transfer and resisting host immune responses	934

Table 1.1: The genes attributed to each segment, as well as their function and length in nts.

host cell nucleus [2]. NPs bind to the phosphate backbone of the RNA chain, approximately one unit per 24 nucleotides (nts), which exposes the RNA bases to other procedures [7].

Lastly, HA and NA, encoded in segments 4 and 6, respectively, are part of the surface glycoproteins integrated in the membrane of the virion. Three HA monomers are required to create a trimeric transmembrane protein. It is responsible for the joining of virion and host cell, as it binds to sialic acid receptors located on the surface of the host cell. Once anchored, it facilitates membrane fusion between virion and host cell by undergoing conformational changes. NA is a mushroom-shaped tetramer, vital to the process of releasing progeny virion from the host cell by separating sialic acid from the carbohydrates present on HA during budding. [2, 3]

There are approximately 4-5 times as many HAs embedded in the membrane compared to NAs [2, 12], which might make it an appropriate target for the formation of neutralizing antibodies (Abs). These two surface molecules are used to name and identify the various subtypes of influenza A, such as H1N1, H3N2 and H5N1, depending on which HAs and NAs are present on the surface. There are currently 18 variants of HA and 11 variants of NA according to the latest reports from Centers for Disease Control and Prevention [13]. The antigenic and structural attributes of HA can be used to divide the different variants into subtypes, from which two major groups emerge as seen in figure 1.2 [14].

Each group is divided into two clades, and the H1 clade of group 1 is further divided into two clades, H1a and H1b, giving a total of five clades across both groups. Each clade contains three HA subtypes, with the exception of H1a which contains four [14]. Figure 1.2 lacks two HA subtypes, namely H17 and H18, which were recently discovered in bats, however, both of these belong in the H1a clade [15, 16]. This clade contains not only the H1N1 strains responsible for the 1918 and 2009 pandemics, but also the H5 subtype found in the highly pathogenic H5N1 avian influenza strains [14], which have also been involved in zoonotic transmissions with severe disease and high mortality rate at approximately 50% in humans. The seasonal H1N1s are also con-

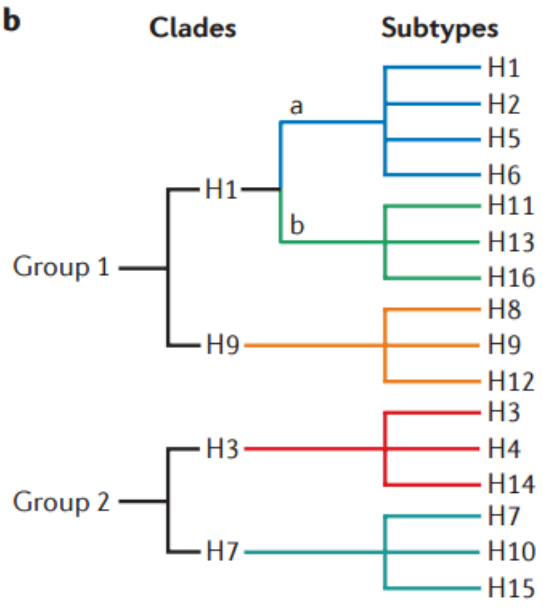

Figure 1.2: The distribution of HA subtypes. HAs discovered after august 2011 are not included. This image is part of a larger figure. [14]

tained in this group, but H3N2, which is another human seasonal flu, is part of group 2, and is the IAV responsible for substantial seasonal outbreaks in recent times. This can be attributed to the a higher rate of mutation in H3N2, which gives it a better opportunity to adapt to immune responses. [14]

1.1.2 Replication cycle

In order to replicate and survive, the IAV must infect target cells. Such is accomplished by binding to sialic acid compounds on the surface of target cells [2]. The HA found on IAVs infecting humans have a specificity towards sialic acid-containing sugars with the sialic acid bound to a galactose moiety with an α2-6 linkage [19]. As mentioned, HA is the glycoprotein responsible, not only for the initial binding of virus and cell, but also the fusion of their membranes. This occurs after the cell has taken in the virus by means of endocytosis, in which the virion is encapsulated in an endosome comprised of the target cell's lipid membrane, as illustrated in figure 1.4. [2]

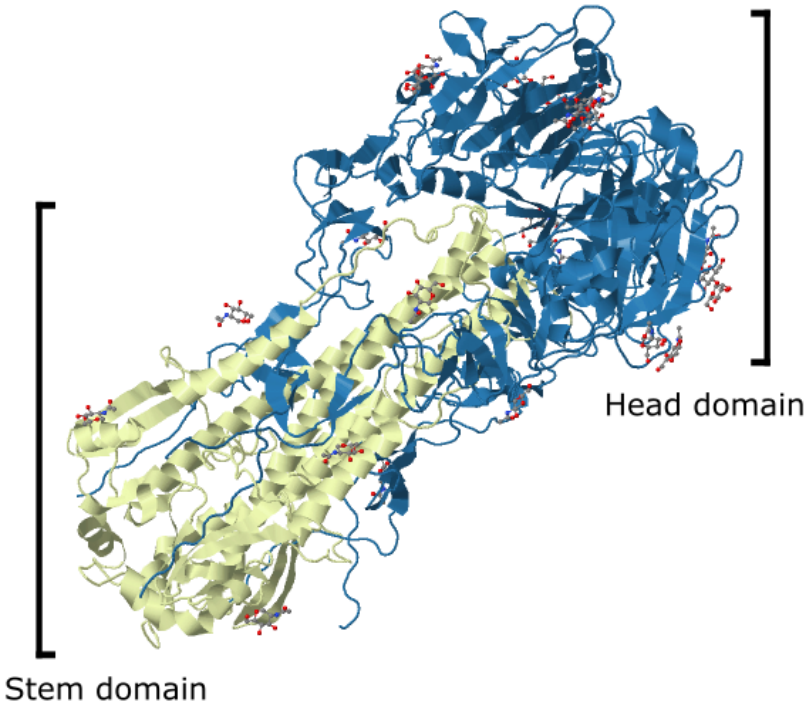

Figure 1.3: The HA trimer represented with ribbons and lines. The cream colored segment encompasses all three HA2 subunits, while the blue segment contains the HA1 subunits. The receptor-binding site (RBS) is found predominantly in the head domain, while the stem domain is responsible for membrane fusion. The vestigial esterase is overlapping both domains. Image is obtained from X-ray diffraction data of a crystalized H3N2 HA sample digested with Bromelain. The trace molecules are N-Acetyl-D-Glucosamine (GlcNAc) and β-D-Mannose. [17]

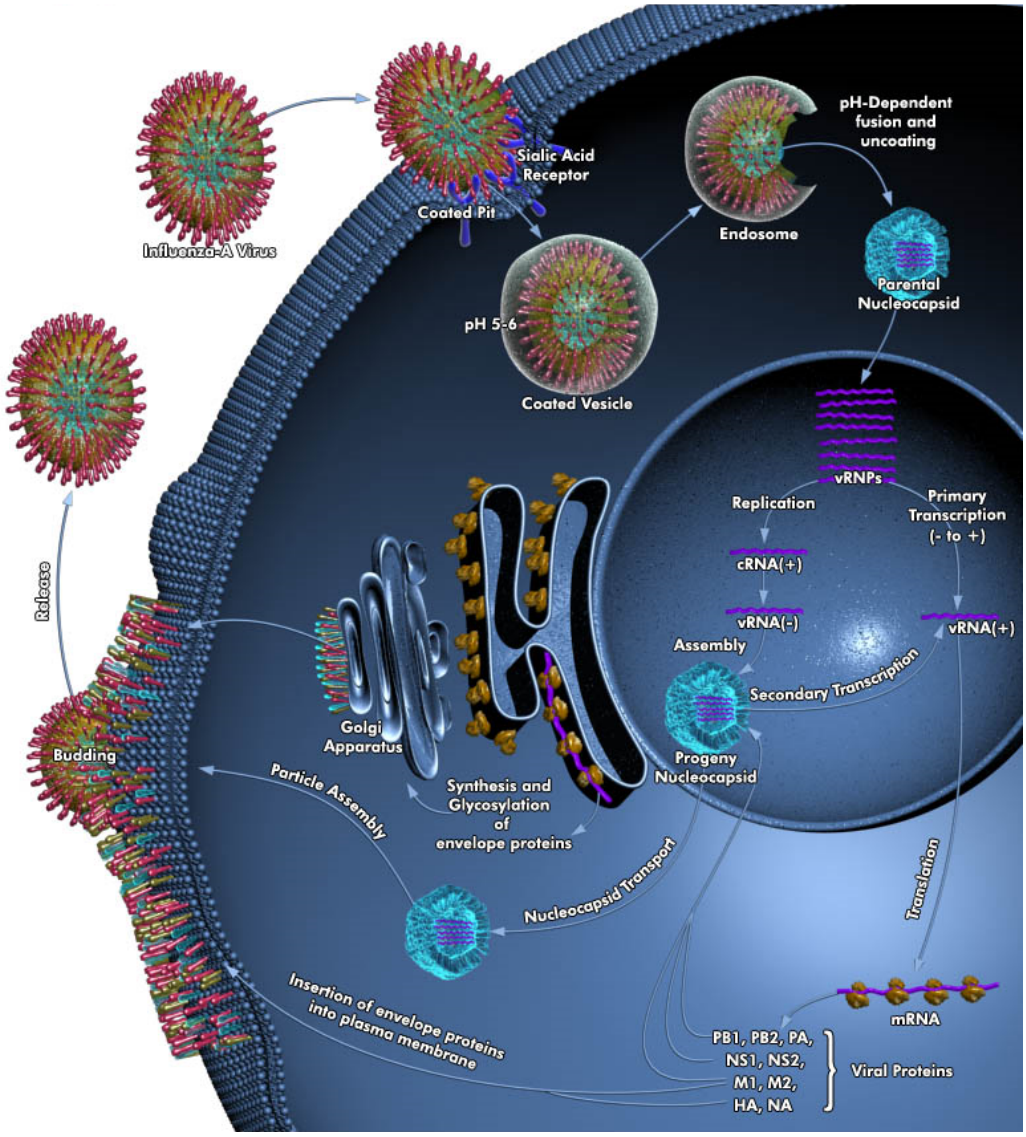

Figure 1.4: Replication cycle of IAV. It binds via a sialic acid receptor, HA, and enters the cell in an endosome. The difference in pH initiates fusion and uncoating of the virion, and the vRNPs are free to enter the host cell nucleus. In there, they undergo transcription or replication, depending on the purpose of the resultant RNA. Replicated vRNA becomes the genome of progeny virus, while transcribed vRNA is used to produce viral proteins. Most of these components are transported directly to the host cell membrane, awaiting particle assembly. However, some proteins, such as HA and NA, first pass through the secretory pathway, to obtain correct folding. As the only two glycoproteins associated with IAV, they also undergo glycosylation in the secretory pathway. These membrane proteins embed in the host membrane, where they become part of the progeny virus as they bud out and are released. This image is a modification of QIAGEN's original, copyrighted image by the author. The original image may be found at [18].

An overview of the HA trimer structure can be seen in figure 1.3. The initial precursor protein derived from the HA gene, the homo-trimer HA0, is not yet suitable for the virus, and therefore undergoes a number of posttranslational modifications, two of which are glycosylation and proteolytic cleavage. The latter results in segregation of HA0 into HA1 and HA2, illustrated in figure 1.3 with blue and cream colors, respectively. The protein consists of 549 amino acids (after cleavage of the signal peptide), where HA1, 327 residues, takes up a major part of the monomer, mainly by constituting the globular head domain which contains the receptor-binding site (RBS) for sialic acid at its tip [21]. The head domain is situated on top of the stem domain, and is predominantly made up of β-strands. [21] The stem domain, which is made up of HA2, 222 residues, and parts of HA1, consists mainly of α-helices and functions as an anchor which spans the lipid membrane, and is important in the membrane fusion process. HA1 and HA2 are held together by disulfide bridges. [3, 21] The stem domain constitutes the segment responsible for membrane fusion. The vestigial esterase overlaps both segments and functions as a flexible overlap between the two. [22]

Membrane fusion

Before mediating membrane fusion, HA0 must first be fully cleaved into HA1 and HA2, which can be done by cellular proteases [21]. Once cleaved, the difference in pH between the endosomal environment and the environment outside the cell, pH 5 and pH 7, respectively, is responsible for eliciting the fusion between their membranes as illustrated in figure 1.5. The pH change

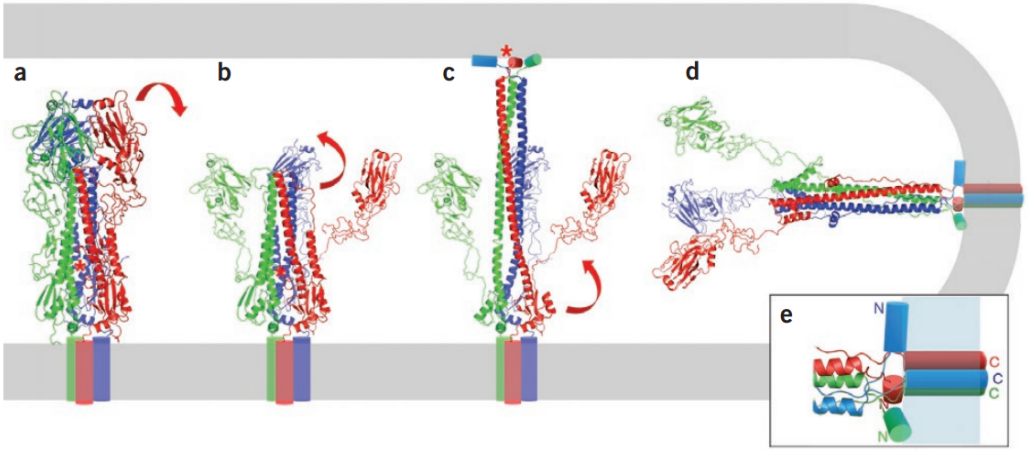

Figure 1.5: Conformational change of HA during membrane fusion. Each HA subunit is represented by a unique color. **a** - The conformation of HA before fusion is initiated. **b** - HA1 subunits dissociate due to changes in pH. **c** - The loop region between the two α-helices in the stem of a subunit becomes helical as well, which propagates the whole structure of HA into a triple α-helix, extending the fusion peptides towards the endosomal membrane as a result. **d** - The aforementioned loop region is reestablished, and the triple α-helix folds back onto itself, bringing the endosomal and the viral membrane together. **e** - The orientation of N- and C-termini after fusion. [20]

provokes HA1 of each subunit to open up, **a** to **b** in figure 1.5, thereby allowing for the extension of each HA2 subunit, **b** to **c** in figure 1.5. A loop region, which is located between two α-helices in the HA1 of a subunit, obtains a helical formation, which by extension forms a longer helix carrying a fusion peptide at the end. Together with the extended helices from the remaining subunits, it forms a triple α-helix, capped by the N-termini of the HA2 subunits, allowing the fusion peptides to reach the endosomal surface, as marked by the red asterisk in figure 1.5 **c**. Once both membranes are anchored, the partition of the stem region anchored in the viral membrane will attempt to fold back to obtain the unextended HA conformation, bringing it back into the proximity of the fusion peptides, **c** to **d** in figure 1.5. As evident from figure 1.5, it takes at least two such processes to occur in the the vicinity of one another, in order to form a pore and have the membranes fuse. [2, 21, 20, 23]

When the membranes are fused, the M2 transmembrane protein functions as an H^+ ion-channel, which allows for protons to trespass the viral envelope. The protons reduce the interaction between M1 and the vRNPs sufficiently to release the vRNPs, which can then be transported to the host cell nucleus, and begin one of two pathways tied to the viral replication cycle of influenza A as illustrated in figure 1.4. One pathway is regular replication, which produces the genes for progeny virus that therefore stays in the (-)ssRNA form. Mutations in the genome happen in this step, due to lack of proofreading in the RNA polymerase provided by the virus. The eventual progeny will thereby carry the mutations along when infecting other cells. The second pathway is transcription, in which the (-)ssRNA is converted to (+)ssRNA, a process which also occurs in the host nucleus. Generally, RNA viruses do not require transportation of the genome into the host cell nucleus, since positive sense, single stranded RNA ((+)ssRNA) can be translated directly into the respective proteins.

Part of the vRNP complex, namely PA, PB1 and PB2, mediates the conversion, by

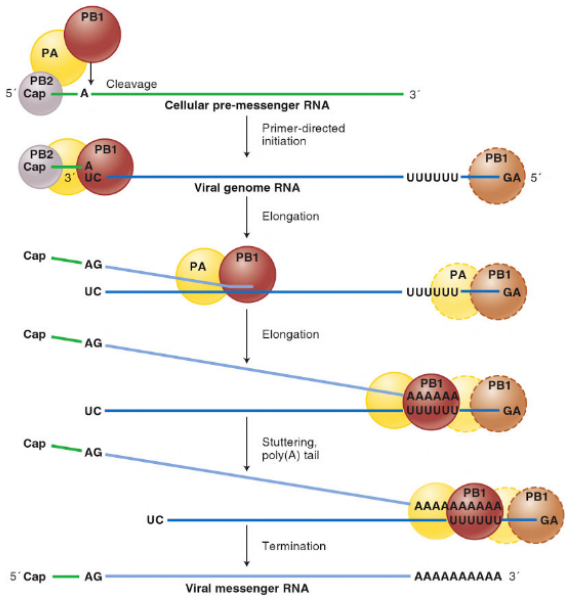

Figure 1.6: The cap-snatching mechanism. PB2 cleaves of a 10-13 nt fragment from the 5' end of cellular mRNA. The 3' end of the fragment attaches to the 3' end of the vRNA, and is utilized as a primer in order to elongate the product. After generating a poly-A-tail as encoded by the template, the transcription process is terminated. The resulting product is viral (+)ssRNA. [2]

the cap-snatching mechanism as seen in figure 1.6. [2]

PB2 in the trimeric polymerase complex steals a small segment of cellular messenger RNA (mRNA), about 10-13 nts, from the 5' end of the cellular mRNA, thereby creating the cap. The 3' end of the small segment attaches to the 3' end of the vRNA, shown with an adenine (A) binding to uracil (U), at which point it functions as a primer to produce the complementary strand pertaining to the vRNA. The transcription ensues and the product is elongated, until it reaches a gene sequence on the template with a large amount of repeated U, which produces a complementary poly-A-tail. The polymerase complex dissociates at that stage, resulting in (+)ssRNA ready for translation by host ribosomes. [2]

Replicated vRNA is not solely transported to the membrane for particle assembly. It also funnels back as secondary transcription, because high amounts of vRNA are required to ensure that there are adequate levels available for transcription into viral proteins. Once the ribosomes have translated the newly transcribed (+)vRNA, the proteins not requiring further modifications are transported to the membrane. However, membrane proteins, such as HA, NA and M2, must first pass through the secretory pathway in order to obtain their properly folded states. Additionally, the glycosylated surfaces of HA and NA are also acquired in the secretory pathway. Once glycosylated, cleaved and properly folded, these proteins are transported into the plasma membrane, where they are embedded.

Particle assembly is the final step happening inside the host cell as the progeny virus is generated. Antigenic shift can happen in this part of the process, if the host has been infected with two variants of influenza. Final release from the host occurs via the budding process. Once fully budded, NA facilitates release by detaching the sialic acid receptors from the sialic acids found on the surface of the host membrane. The progeny virus is now ready to continue infecting new host cells. [2] There are two main proposals as to what induces cell death during an IAV infection. One is simply that the production rate of regular proteins required by the host cell becomes too low, essentially robbing the cell of its functions, while the other is apoptosis or programmed cell death. [24, 25]

1.1.3 Immune responses

The defense system of an infected host consists of many levels, which are activated as a response to the adaptive evasiveness displayed by the pathogen [26]. In cases where the pathogen gets past the outer, preliminary defenses, such as our skin, it will initially illicit a response from the innate immune system [26]. Macrophages or natural killer cells could be part of such a response [26], but using IAV as an example, the reaction could be deployment of lectins, such as the mannose-binding lectin (MBL)

or surfactant protein D (SP-D). Both are examples of water soluble proteins, but membrane-bound lectins also exist. MBLs exploit the fact that HA and NA are both glycoproteins, rich with mannose containing N-linked glycans, which they demonstrate strong affinity towards [27, 28]. By binding to these specific glycans on HA MLBs might inhibit the infectivity of the virus, by blocking the binding to sialic acid residues or hindering membrane fusion. SP-Ds work similarly, but have a broader range of moieties they can bind to [27, 28]. Therefore, they are better at neutralizing highly glycosylated IAVs, which has been confirmed in several studies [29, 27, 30]. The degree of glycosylation is an important determinant regarding the sensitivity of MBLs and SP-Ds, meaning that IAV with HA containing few glycans will have a better change of escaping this type of immune response. [27, 28]

If the innate immune system is insufficient to deal with the pathogen the adaptive immune system is activated. This process can take several days, but the benefit is increased specificity in terms of recognition of the antigenic nature of the pathogen. This last level of the immune system is facilitated by T and B lymphocytes, the latter being responsible for the production of Abs. These Y-shaped immunoglobulins, as illustrated in figure 1.7, consist of two heavy and two light chains, coloured as blue and green, respectively. The two heavy chains are joined by two disulfide bridges, while the two light chains are attached to a heavy chain each with a single disulfide bridge. Abs have two distinct regions, namely the fragment, antigen-binding (Fab) regions, which is located at each of the two tips of the Y-shape, and the fragment, crystallisable (Fc) region which encompasses the bottom of the Y-shape. [26]

The Fab regions are mostly identical and are responsible for binding to the antigen, while the Fc region binds to receptors on cells that support the immune system. The Fab regions can also be referred to as paratopes, because they bind to epitopes which are specific for each antigen, meaning that the Fab regions must be equally specific.

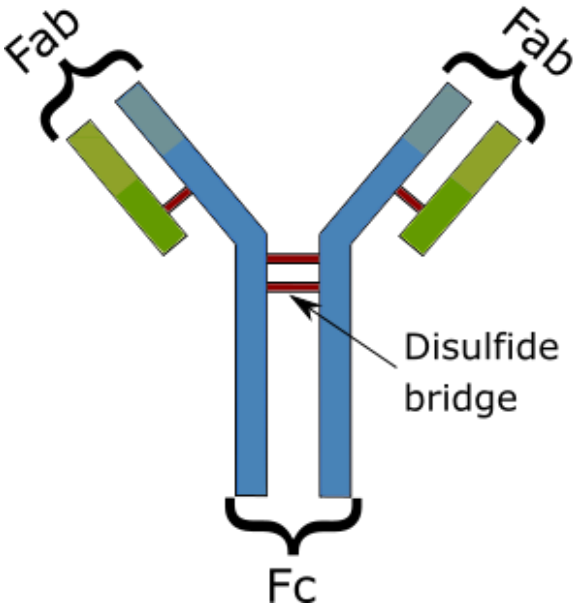

Figure 1.7: Schematic representation of an Ab. The heavy chains are coloured blue and the light chains green. The two heavy chains are linked by two disulfide bridges (red), and each light chain is linked to the nearest heavy chain with a single disulfide bridge each. The antigen-binding sites at the tips of the Y-shaped molecule are designated by a lighter color.

Abs can be poly- or monoclonal, depending on the ability to bind specifically to a single epitope (monoclonal), or if it consists of a mixture of Abs that each bind different but specific epitopes, effectively binding to multiple epitopes on the same antigen (polyclonal). The epitopes can be located at areas that do not inhibit the function of the antigen, but once the Abs are bound, the antigen is tagged for destruction by the immune system. Furthermore, epitopes can also be located at sites which are important for the function of the antigen, thereby inhibiting it directly. [26] With HA of A(H1N1)pdm2009 for instance, there are four such sites called antigenic sites, which are important for proper antigenic function [31]. These are named crossreactive *a* (Ca) and *b* (Cb), and strain specific *a* (Sa) and *b* (Sb) [21, 31]. However, Ca can be divided into Ca1 and Ca2 [21, 32]. Their position in the amino acid sequence can be read in table 1.2.

All five antigenic sites are located on the globular domain as illustrated in figures 1.8 and 1.9. These images are generated in YASARA [33], using 4GXX from the RCSB Protein Data Bank [34]. The HA1 is coloured white, and HA2 cyan. Sites Sa (magenta) and Sb (red) are located at the top of the globular domain, which means that they are a part of the RBS. The RBS (not depicted) is comprised of several, individual residues, e.g. Tyr91 and His180, as well as three distinct regions, namely the 130-loop, 190-helix and 220-loop [35]. Sites Ca1 (blue) and Ca2 (yellow) are located centrally, at the crossover between the RBS and the vestigial esterase, with the Cb site (green) just below, situated solely in the vestigial esterase. [21]

Vaccines

The specificity Abs have towards viral entities has made them the best option when it comes to preventing infections. This is evident from the reduction in mortality and growth in population the world has experienced since its introduction as disease prevention. The mechanisms of the adaptive immune system are exploited to provoke an immune response by supplying a non-infectious version of the antigen to the host, a procedure known as immunization or vaccination. This gives the host time to recognise the antigen, without being in danger of an infection, and is thus "taught" to combat subsequent infections. [2, 36]

There are a variety of approaches available when constructing a vaccine, such as inactivation, attenuation, cold-adaptivity, subunit or DNA vaccines. Inactivation has been utilized for the longest, and is still one of the primary methods of producing vaccines. The virus is first grown in embryonated eggs, then inactivated by detergents, also referred to as chemical inactivation. The vaccine should encompass virion parts from different strains, in order to be as effective as possible. [2, 36, 4]

The attenuated approach uses live virus, which has had parts or a part of its geno-

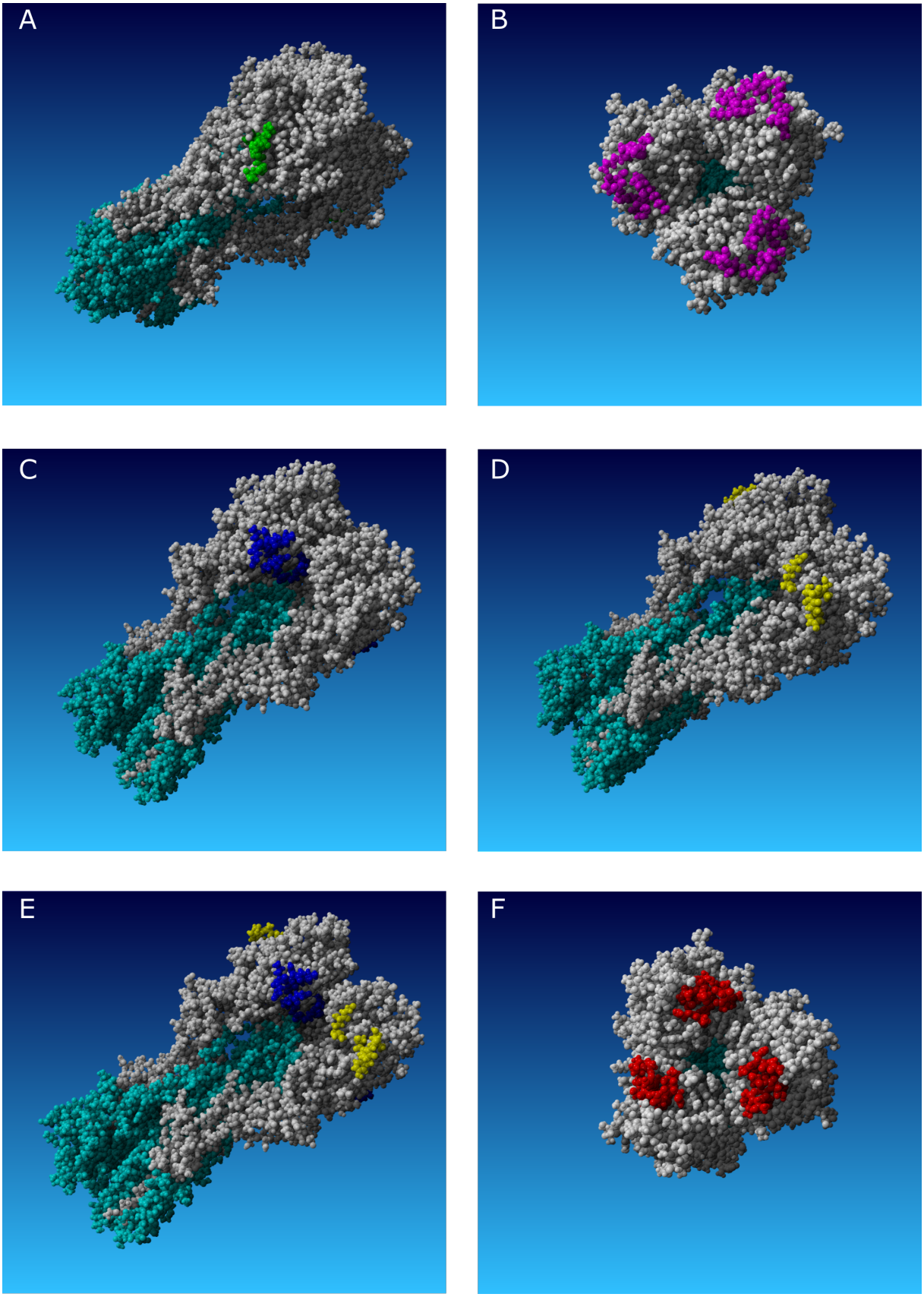

Figure 1.8: Placement of antigenic sites on the HA trimer. **A** - Cb site. **B** - Sa site. **C** - Ca1 site. **D** - Ca2 site. **E** - Ca1+Ca2 site. **F** - Sb site. The white part consists of all three HA1 subunits, while the cyan part is comprised of the HA2 subunits. Images are generated with YASARA using 4GXX from the RCSB Protein Data Bank [33, 34].

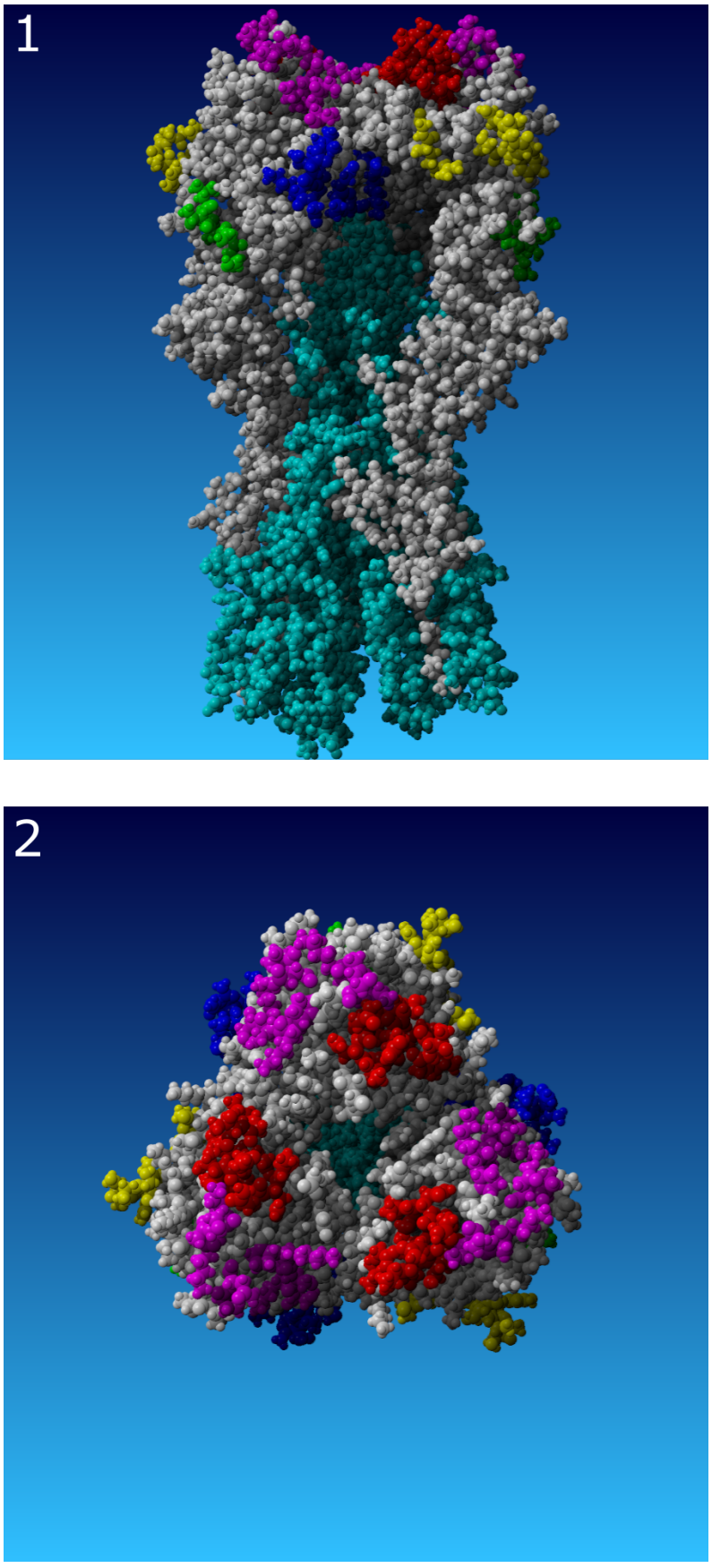

Figure 1.9: A collective overview of the antigenic sites on the HA trimer. **1** - Sideways view. **2** - Viewed from the top. **Green** - Cb site. **Magenta** - Sa site. **Blue** - Ca1 site. **Yellow** - Ca2 site. **Red** - Sb site. The white part consists of all three HA1 subunits, while the cyan part is comprised of the HA2 subunits. Images are generated with YASARA using 4GXX from the RCSB Protein Data Bank [33, 34].

me altered in order to attenuate its virulence. For instance, the cleavage site for HA1 and HA2 was mutated to achieve attenuated virulence, and the segment of the genome pertaining to these mutations were transferred to a another IAV strain called PR8, which is engineered to grow at a higher rate, thus producing more of the HA attenuated virus for the vaccine. [2, 36, 37]

Cold-adapted vaccines are made by optimizing the virus to grow at lower temperatures, for example 25 °C, which significantly weakens their virulence when introduced to the 37 °C environment of human and other mammalian hosts. The cells will develop Abs against the initial invaders, but since their growth is greatly attenuated, they cannot replicate at a rate sufficient for continuous infection. [4, 37] This approach has shown great protection rates in children [37].

Another way to make a vaccine is to produce antigenic subunits of the virus in question, separately. They are produced in other hosts, for instance by growing HA or NA in insect cells [37]. Using this method ensures that there is no risk of the virus self-assembling or becoming virulent again, since not all subunits are present in the vaccine. [38]

Finally, an approach still in the experimental phase is DNA vaccines [37]. DNA from a target virus is genetically engineered and injected into the host cells, for instance, via plasmids, in order to have the cells produce the antigen, and in turn the Abs, since it will be recognized by the host as a foreign entity [39]. This approach shows great prospects, seeing as it can be used for pathogens that are more dangerous, such as HIV, as it only requires the specific gene, and not the whole virus, to create the vaccine, essentially making it safer [4]. Additionally, DNA is considered easy to manufacture and store, which makes it useful under many circumstances [39].

Influenza vaccines Determining the composition of the seasonal vaccine for influenza requires extensive monitoring and reviewing of clinical and laboratory studies performed each year. WHO are responsible for organizing a meeting with their Collaboratory Centers twice a year, where they will conclude which influenza strains should be recommended for inclusion in the upcoming flu season. However, these meetings can only generate educated

Antigenic site	Position
Ca1	183-187, 220-222, 252-254
Ca2	154-159, 238-239
Cb	87-92
Sa	141-142, 170-174, 176-181
Sb	201-212

Table 1.2: The angtigenic sites of IAV H1N1 HA, and the positions of the amino acids they occupy in the primary structure.

forecasts for the impending flu season, seeing as the virus is highly changeable. Typically, the vaccine will contain components from 3-4 different influenza viruses that are estimated to be the ones most likely to spread in the coming season. This is based on data from the virus characterisation of the circulating influenza viruses and the vaccine effectiveness of the previous seasonal vaccine. The results of these meetings are only meant as guidelines, and each country has to choose whether or not to follow them. [40]

It is rarely possible to completely predict which strains will be circulating in the upcoming flu season. This presents a problem when the vaccine virus and the circulating virus are too dissimilar. A mismatch such as this can result in a low vaccine effectiveness. However, a good match will result in high vaccine effectiveness. [41] Mismatches can happen if the circulating strains do not correspond with the predicted strains for the vaccine, or due to antigenic changes that occur over the course of a season [42], which is a result of immune evasion mechanisms inherent in influenza viruses.

1.1.4 Immune evasion

Immune evasion can occur in multiple forms, but at the core it comes down to changes in the genome of the virus. IAV is a specialist in this area for two reasons, mainly because it can undergo antigenic drift and/or antigenic shift. The latter is the swapping of whole HA or NA genes between two influenza viruses, which therefore requires infection by two different strains of IAV in the replication cycle at the same time. This process, as exemplified in figure 1.11, where a gene segment of influenza X is swapped with the corresponding gene segment of influenza Y. Therefore, the potential onset of a pandemic spread with IAV is great, since it has a segmented genome and can infect a variety of species. For instance, the two distinct HA or NA genes can be swapped when the virions assemble, which produces three variants from the original two. This induction of genetic variation usually has a more severe effect, because a considerable amount of people will not have protection against the entirely new HA or NA, and pandemics can therefore arise. [2]

Antigenic drift

The other immune evasion mechanism, antigenic drift, is essentially point substitutions in the genome of IAV, which are caused by error-prone RNA polymerases that are part of the virus replication process. This is illustrated in figure 1.11, in which an influenza virion passes through the replication cycle in a host cell, to then be released with substitutions introduced into the genome, colored red. Compared to DNA polymerases, the RNA polymerase of IAV has neither proofreading nor editing, which can result in an accumulation of substitutions over time [2, 43, 44]. Thereby, each succesive replication of the

virus in the host creates a cluster of progeny with a multitude of genetic variations known as quasi-species [45]. These types of error-prone RNA polymerases have evolved to replicate near the error threshold, meaning that a small increase in the amount of introduced mutations would result in genetic instability of the virus [45].

Following the principle of "survival of the fittest", the species with the best phenotype in the cluster would be selected for further growth, but that is not the only determinant in the case of IAV [46]. Another principle, "survival of the flattest", is also involved in the selection process [46]. It is more advantageous for the virus to enhance the replication rate for the cluster of quasi-species, rather than select for the single, fittest genotype, because the variants in the immediate area, genotypically, can be insufficiently fit [45, 46]. By adhering to survival of the flattest, the virus has several groups of closely, genetically related variants that can be selected in case the dominant, i.e. fittest variant, is compromised by environmental pressure [45]. The principle is illustrated in figure 1.10 b.

As evident, the fitness maximum of the population in the high and narrow peak (red) is superior to the flat and broad peak (blue), but a high amount of mutational neighbours are situated far down on the narrow peak, meaning that they are not acceptably fit and could be non-infectious [46, 47]. The broad, flatter peak will include more progeny with adequate fitness levels, in turn resulting in enhanced survivability for the virus as a whole [46, 47]. A flatter peak might also allow the virus to adapt more readily to the host immune system, since it would have more options to choose from, potentially giving it a better chance of having a variant that can escape the Abs deployed by the immune system.

Antigenic drift can phenotypically manifest itself by directly altering the antigenic sites, for instance, by exchanging an amino acid, which could affect the binding affinity negatively for Abs aimed at these si-

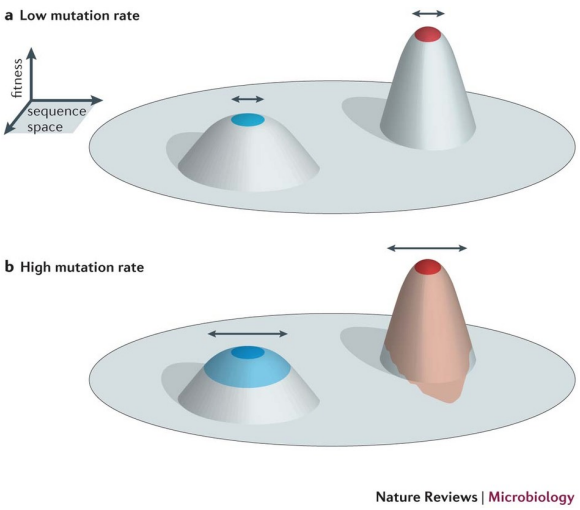

Figure 1.10: The fitness landscape has a 2-dimensional base called the sequence space where the different sequences are mapped. Each sequence is assigned a fitness level, represented by the vertical axis. **a** - Low mutation rates favour survival of the fittest (red dot), since the genetic variation in the population will be small. **b** - At high mutation rates, the genetic variety will be larger, which will then favour survival of the flattest since more variants maintain an adequate fitness (light blue dot) compared to the population concentrated around the fittest variant. As a result, the blue population will outcompete the red population when the mutation rate is high. [47]

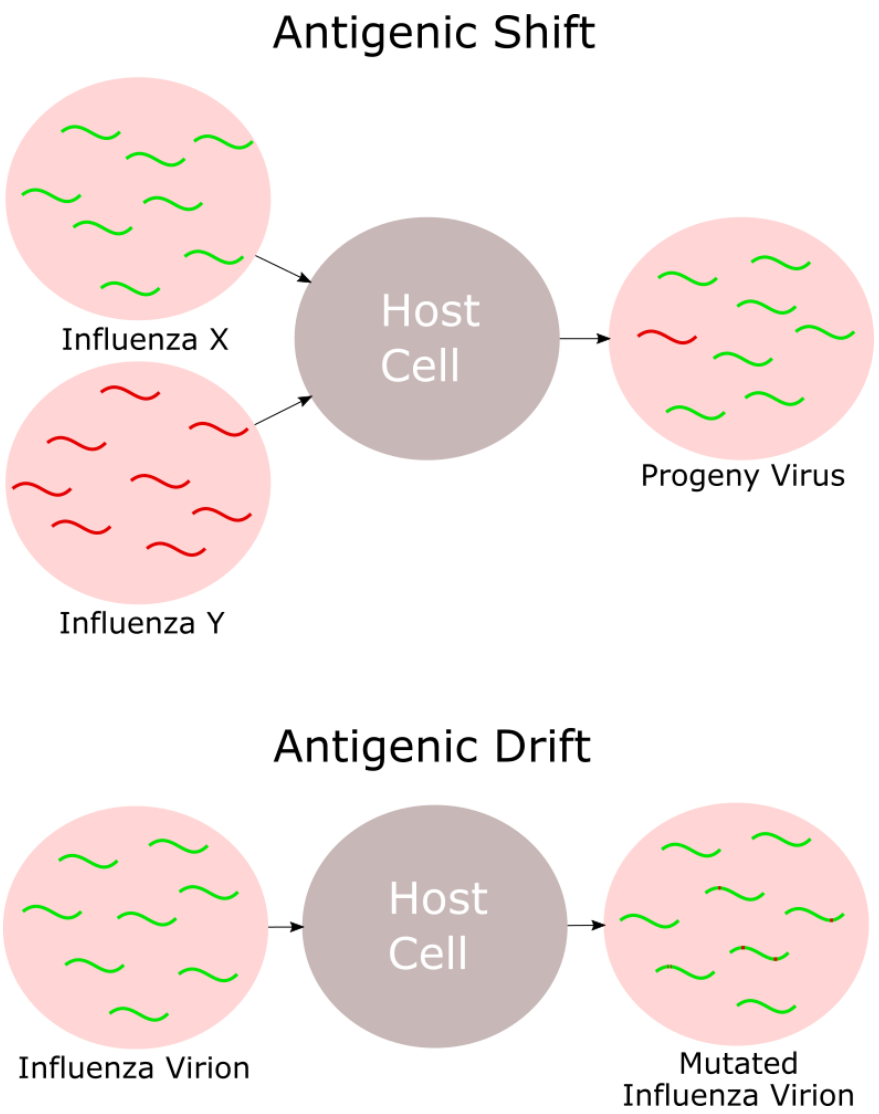

Figure 1.11: Schematic illustration of antigenic shift and antigenic drift. In the case of antigenic shift the host cell is infected with two different strains of influenza virus, here influenza X and Y. This might produce progeny virus with completely exchanged gene segments, exemplified here by influenza X switching out one of its gene segments, green, with a gene segment from influenza Y, red. Antigenic drift requires no simultaneous infection by two different strains, but is a result of error-prone RNA polymerases that introduce substitutions into the genome of the virus, depicted as red points.

tes [44, 48]. Especially substitutions in the epitopes near the RBS can have a great effect on the binding affinity of Abs [44, 48]. These sites have also been shown to have a higher rate of mutation [49]. However, an alternative mean of protection could be to introduce glycosylation sites in and around these epitopes and important functional sites.

Glycosylation

Glycosylation sites are found as specific sequons in an amino acid sequence, namely as Asp-X-Ser/Thr, where X can be any amino acid except for Pro [50]. The glycans attach to the amide of the Asp side-chain, hence they are termed *N*-linked glycosylations [51]. The structure of the glycans can take many forms, depending on what moie-

ties are involved and how they are branched [51]. A list of moeities and potential branchings can be seen in Appendix A.1. The individual moieties attach to each other via glycosidic bonds, while the attachment to the amide of the protein usually requires a specific moiety, e.g. GlcNAc in animals [51]. The inherent moieties and structures of the glycans depend on the position of the glycosylation site and how accessible it is for host derived saccharide-modifying enzymes [50, 51]. However, the amino acid sequence is not the only determinant in deciding which glycans attach, and glycosylation can also be impeded by steric hindrance from other glycans or certain amino acids occupying space in close proximity of the glycosylation site. The glycans found on each glycoprotein are heterogenous, but the most common ones are high-mannose or complex glycans. The former is defined as branched glycans terminating with the moiety mannose, while the latter has branches terminating in either *N*-acetylgalactosamine or galactose. [27]

Some of the glycans occurring on HA are vital during protein folding, and their absence can result in a non-functional protein [50, 52]. However, these conserved glycans are mostly positioned on the stem domain and do therefore not interfere with receptor binding [52], although a mutation in this area might affect membrane fusion. Glycosylation in the head domain is more variable [27], which is how it contributes to HA evolution [52]. There are consistently five recurring glycosylation sites located in the stem domain, per HA subunit, in nearly all HAs surveyed in the paper [52], while the head domain in general holds up to three glycosylation sites, although it is rare to have them all occupied simultaneously [52]. The conserved glycosylation sites in H1N1 can be found at residues 28, 40, 104, 304, 498 and 557, while the variables sites on the head domain are typically found at 142, 172, 177, 179, in close proximity of the Sa site [5, 53]. Three glycosylation sites on the head domain are mostly observed in recent strains of H1N1, i.e. last 2-3 years, and the early strains from the 1918 pandemic contained one glycosylation site at most [52], which is also the case for A(H1N1)pdm2009 [6]. This indicates that low levels of glycosylation on the head domain results in a strain with enhanced pathogenicity, yet the virus still evolves to obtain a higher amount of glycans, in order to evade immunity and thereby survive as a seasonal epidemic [5, 6].

Glycosylation has both a positive and a negative effect on pathogenicity [54]. As mentioned, Abs will attach to epitopes on HA, and adding more glycans can have the effect of masking said epitopes and blocking the Abs sterically [55], thereby neutralizing or attenuating the effect of the Abs [10, 56]. However, it might not mask the epitopes alone, but also the important functional parts, such as the RBS, which could result in lower binding affinity to sialic acid

receptors on the host cell [10, 54]. Additionally, the host has immune responses aimed at hyperglycosylated antigens, such as the previously mentioned lectins, which is why those are rarely found in propagated species [6]. Adding a high amount of glycans will therefore mainly have a negative effect, as both evasion of lectins and receptor-binding suffers as a result [10]. Getting the right amount and correct position of glycans is a balancing act in which factors such as pathogenicity, immune evasion, viral fitness and immune responses play vital roles [54, 55].

In essence, the main function of glycosylation on the head domain of HA is to avoid immune recognition by host-mediated Abs by changing antigenic properties, which is one way seasonal flu might overcome the immune system each year [53, 57]. This inevitable sacrifice of pathogenicity in favour of genetic diversity to increase survivability in immune population means that the currently employed vaccines require constant revision to effectively combat the immune adapted strains of IAV.

1.2 Theoretical methodology

There are several effective techniques when investigating IAVs, which can provide understanding of their adaptivity to the host immune system. The theoretical backgrounds of the methods used in this study are presented in this section.

1.2.1 Hemagglutination assay

A hemagglutination-assay (HA-assay) is commonly used to test a sample for the presence of IAV, and simultaneously provide a relative concentration in the form of HA-titers [58]. The method utilizes the fact that the HA surface molecules of IAV can bind sialic acid residues present on red blood cells (RBCs). This is usually performed in micro titer plate wells. Once bound, the RBCs agglutinate and form an opaque sheet, when suspended in rounded bottom wells, as illustrated in figure 1.12 **a** [58]. The appearance of agglutination depends on the shape of the microtiter plate wells and which animal species the RBCs originate from [59]. An insufficient amount of IAV will not be able to agglutinate the RBCs, and as a result they will drop to the bottom of the well, forming a dot- or halo-shape as seen in figure 1.12 **a**. The HA-titer is found by a dilution series, typically two-fold, of the virus sample in a microtiter plate, with addition of equivalent amounts of RBCs to each well. The more wells showcasing agglutination, the higher is the HA-titer of the IAV, which is exemplified in figure 1.12 **b**. The largest virus dilution exhibiting agglutination is defined as one HA-unit. [58]

Hemagglutination inhibition assay

The titers obtained from a HA-assay can be used to perform a hemagglutination inhibition-assay (HAI-assay), in which the

inhibition of the hemagglutinating activity of IAVs by HA-specific Abs can be measured. The samples are diluted according to their HA-titers to obtain four HA-units, which corresponds to the HA-titer two column left of the actual HA-titer. Taking row A in figure 1.12 as an example, 1:32 would be one HA-unit, and four HA-units would then be 1:8. These viral dilutions are then tested against a dilution series of an appropriate Ab/antiserum targeting the HA protein. Ab-inhibited IAV cannot agglutinate RBCs and as a consequence the RBCs aggregate at the bottom of the well. The HAI-assay result is reported as an HAI-titer and indicates the lowest antiserum concentration which inhibits agglutionation of RBCs. [60]

To investigate antigenic changes of circulating IAVs compared to e.g. a vaccine virus, Abs, raised in ferret against the vaccine virus, are used as a reference antiserum in the HAI-assay. To antigenically characterise the IAV samples, the HAI-titer of the vaccine virus is divided by the HAI-titers of the circulating IAVs, which gives the fold-change factor of the individual IAV sample. If the fold-change factor of the circulating IAVs has changed with 8 or more from the vaccine virus, the virus is considered antigenically different from the vaccine virus. [61]

1.2.2 Sequencing

The Sanger sequencing method [62, 63], or *dideoxyribonucleotide* (ddNTP) *chain termination sequencing*, is the current standard of DNA sequencing. There is a limit to the chain length of the DNA, approximately 800-1000 nts, so the gene of interest must be cut into fragments when exceeding said limit. The process takes place in a sequenator, where the DNA fragments are denatured, and appropriate primers subsequently attach at the 3' end of their respective fragments. A regular DNA synthesis is initiated incorporating ddNTPs conjugated with fluorescent labels. These molecules terminate the DNA chain growth, because the hydroxide at the 3' end is replaced by a proton. The result is a multitude of DNA chain fragments

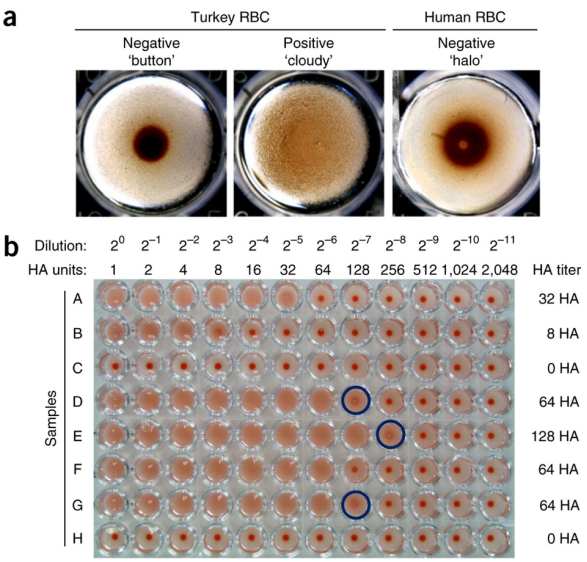

Figure 1.12: a - Three potential shapes the blood in a HA-assay might take. The dot and the opaque shapes are obtained with turkey RBCs, while the halo shape is formed with human RBCs. **b** - A full HA-titer plate. Virus samples are two-fold serial diluted left to right, and the different samples are labeled top to bottom. Rows A-B and D-G all have the ability to agglutinate blood, whereas C and H cannot. Partial agglutination is visible in the wells marked with blue rings. HA-titer for each sample row is displayed on the right. [58]

with varying lengths, which can then be separated by size. A capillary tube containing a gel is used for this purpose by letting the smallest fragments pass through at a faster rate as illustrated in figure 1.13. The gel is sensitive enough to separate fragments by a single nt, meaning that the DNA chain fragments consisting of a single ddNTP and primer will be the first to pass through the gel, then a ddNTP with a single nt and primer will be next, and so on.

A laser excites the fluorescent labels, which differ in color depending on the attached ddNTP, and the resulting fluorescence is registered by a detector. The product is a spectrogram which gives the sequence of the complementary strand to the template strand, and the fluorescence of each fragment pertaining to a specific size. [64]

1.2.3 Immunoblotting

Identification and analysis of a protein mixture can be done with western blotting, a standardized method in the field of molecular biology. The technique initially utilises gel electrophoresis, by applying an electric field across a gel loaded with protein samples. The negatively charged proteins will migrate towards the positive electrode, and the small pores, produced by the interaction of polyacrylamide with agarose, will promote faster movement for the smaller proteins. Thereby, the initial separation of the proteins is by size, with the smallest having travelled the farthest down the gel. The second separation process involves transferring the proteins to a membrane which is done by applying an electric field perpendicularly to the agarose gel, and have the proteins migrate to a membrane, for instance a nitrocellulose membrane, which is positioned between the gel and the positive electrode.

A series of Abs is then used to identify specific proteins. Each Ab incubation requires an initial blocking of the membrane with a blocking agent, for instance, nonfat dried milk or bovine serum albumin (BSA), to prevent unspecific binding on the membrane. Afterwards, the membrane is incubated with different Abs, which are specific to the target protein and one another. The last Ab must be labeled with a fluorescent marker, e.g. horseradish peroxidase (HRP). Adding hydrogen peroxide then produces a signal, which is only visible in places where the fluorescent marker has at-

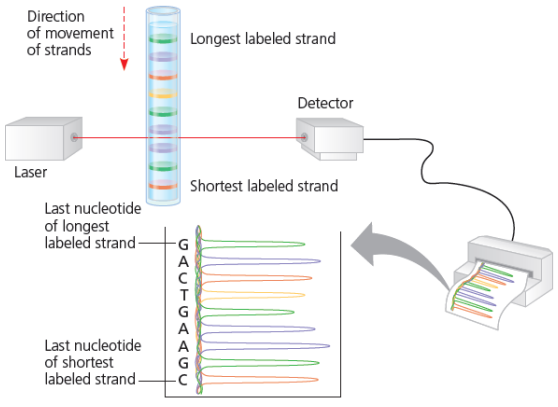

Figure 1.13: Sanger sequencing is accomplished by passing DNA chain fragments of increasing size through a gel-containing capillary tube. A laser illuminates the fluorescent label, and the signal is picked up by a detector. The sequence of the gene of interest and flourescence of each fragment pertaining to a specific size is presented with a spectrogram. [64]

tached, and thereby it becomes possible to identify and analyse the size of a specific protein in a complex mixture. [65]

1.2.4 Surface plasmon resonance

Surface plasmon resonance (SPR) is an optical detection method which can measure chemical interactions in real-time [66]. The four main components of an SPR set-up, as seen in figure 1.14, is a polarized light source, a prism, a detector and a sensorchip comprised of a dielectric layer and a thin metal film [67]. Electromagnetic waves pass through the prism and hit the metal film, typically gold, splitting up the light into reflected light and evanescent waves that propagates through the gold film [66, 67]. At the resonance angle the evanescent wave will transfer energy to the electrons at the interface between the gold film and the dielectric layer, essentially causing them to resonate at its own frequency [66, 67]. The resonating electrons are known as surface plasmons, which are highly sensitive to their surrounding environment [66]. In the case of biochemical SPR studies, an analyte, for instance an antigen, is introduced to an environment consisting of immobilized ligands, e.g. Abs [67]. The interaction between these two biomolecules will manifest itself as a change in the dielectric constant of the dielectric layer, due to increased mass, and hence the refractive index as well [66, 67]. Changes in the refractive index is thus registered by the detector as an offset of the resonance angle [66, 68]. Of interest to this thesis is a time-resolved measure of the kinetics of two biomolecules, which works by introducing free-flowing analyte molecules to immobilized ligands until they reach equilibrium, i.e. the same amount of analyte molecules attach and detach at any given point in time. Eventually, flow of analyte is terminated, and, depending on the inherent reactions, the molecules start detaching. [67] This approach produces a sensorgram which enables the measurement of affinity of an antigen-Ab interaction by determining the equilibrium constant. Information on binding kinetics is also attainable by establishing the association and dissociation constants. [68]

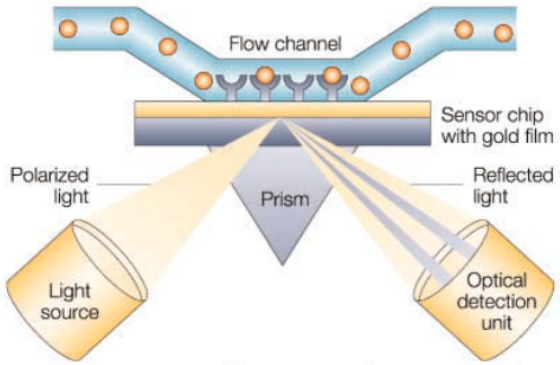

Figure 1.14: Schematic SPR set-up, where polarized light passes through a prism and hits the sensorship. Signals are recorded as angular shifts in the reflected light, which are dependent on chemical reactions occurring at the sensorchip surface, by the detector. This is a cropped and modified version of a larger image. [69]

2. Materials and Methods

Chemical	Description	Supplied by
β-mercaptoethanol	Lot no. BCBK8223V	Sigma-Aldrich
70% ethanol	Lot no. SE10049591	plum
Ethanolamine, Minimum 98%	Batch no. 045K0644	Sigma-Aldrich
2-Morpholinoethanesulfonic acid Monohydrate	Lot no. 1275145	Fluka
N-Hydroxysuccinimide	Lot no. BCBF6027V	Sigma-Aldrich
Glycine	Lot no. 0F008040	AppliChem
Sodium tetraborate decahydrate	Lot no. SZBG0500V	Sigma-Aldrich
Sodium dodecyl sulfate	Lot no. STBD6276V	Sigma-Aldrich
Disodium hydrogen phosphate dihydrate	Lot no. K33672780	Merck
Potassium chloride	Lot no. 1121871	Fluka
Potassium phosphate monobasic	Lot no. SLBR1363V	Sigma-Aldrich
Acetic acid $\geq$98%	Lot no. SZBF0220V	Sigma-Aldrich
N-(3-Dimethylaminopropyl)-N´-ethylcarboiimide hydrochloride	Lot no. BCBN0730V	Sigma-Aldrich
Tween 20	Lot no. S6299684 136	Merck
Sodium chloride	Lot no. SZBE0130V	Sigma-Aldrich
Methanol	Batch no. 8009104	Bie&Berntsen A-S
DAB tablets, 10 mg, pH 7.0	Lot no. 120222-3	Kem-En-Tec
Glycerol 85%	Lot no. Z277194 302	Merck
DL-Dithiothreitol	Lot no. SLBQ1154V	Sigma-Aldrich
Pyronin Y	Lot no. MKBW0735V	Sigma-Aldrich
Trizma® base	Lot no. SLBN7135V	Sigma-Aldrich
Trizma® hydrochloride	Lot no. SLBQ5915V	Sigma-Aldrich
Trypsin from bovine pancreas	Lot no. SLBF1700V	Sigma-Aldrich
Penicillin Streptomycin	Lot no. 1864846	gibco
IGEPAL CA-630	Prod. No. I8896	Sigma-Aldrich

Table 2.1: Raw chemicals used throughout the experiments of the thesis.

IAV isolate	Passages	Supplied by
Season 2009-2010		
A/Denmark/1302/09	6. MDCK 15.12.16	SSI
A/Denmark/1323/09	8. MDCK 15.12.16	SSI
A/Denmark/1325/09	4. MDCK 20.10.16	SSI
A/Denmark/1337/09	7. MDCK 24.10.16	SSI
A/Denmark/1338/09	6. MDCK 15.12.16	SSI
A/Denmark/1341/09	8. MDCK 24.10.16	SSI
Season 2010-2011		
A/Denmark/113/10	6. MDCK 15.12.16	SSI
A/Denmark/115/10	5. MDCK 04.11.16	SSI
A/Denmark/116/10	1. SIAT/5. MDCK 27.10.16	SSI
A/Denmark/120/10	7. MDCK 04.11.16	SSI
Season 2013-2014		
A/Denmark/106/13	3. MDCK 04.11.16	SSI
A/Denmark/05/14	5. MDCK 24.10.16	SSI
A/Denmark/11/14	6. MDCK 15.12.16.	SSI
A/Denmark/13/14	3. MDCK 24.10.16	SSI
A/Denmark/38/14	5. MDCK 04.11.16	SSI
Season 2014-2015		
A/Denmark/50/14	5. MDCK 15.12.16	SSI
A/Denmark/51/14	4. MDCK 22.11.16	SSI
A/Denmark/03/15	6. MDCK 15.12.16	SSI
Season 2015-2016		
A/Denmark/41/15	3. SIAT/2. MDCK 20.10.16	SSI
A/Denmark/41/15	4. SIAT 20.10.16	SSI
A/Denmark/43/15	5. MDCK 24.10.16	SSI
A/Denmark/48/15	4. SIAT 04.11.16	SSI
A/Denmark/61/15	4. SIAT 27.10.16	SSI
A/Denmark/20/16	4. MDCK 04.11.16	SSI
A/Denmark/27/16	3. MDCK 04.11.16	SSI
A/Denmark/30/16	1. SIAT/3. MDCK/1. SIAT 18.11.16	SSI
A/Denmark/41/16	3. SIAT 20.10.16	SSI
A/Denmark/46/16	3. SIAT 20.10.16	SSI
Reference virus		
A/California/07/09	2. SWL/4. MDCK	SSI

Table 2.2: A list of the IAV samples tested in this thesis. 2.SWL indicates use of secondary swine-like kidney cells.

Material	Description	Supplied by
SPR Sensorchip HC 200 M	Lot no. SC HC200M0217	XanTec bioanalytics
0.8% E-gel® w/ Ethidium Bromide	Lot no. Q10056	invitrogen
Nitrocellulose Membrane, 0.45 μm	Lot no. 1846345	invitrogen
Novex™ WedgeWell™ 10% Tris-Glycine gel	Lot no. 16071540	invitrogen

Table 2.3: Pre-made equipment and materials used in the experiments of this thesis.

Biological material	Description	Supplied by
MDCK cells	Cat. no. 85011435	ECACC
MDCK-SIAT1	Cat. no. 05071502	ECACC
Ferret antiserum A/California/07/09	RDE treated	SSI
Guinea pig blood		Dyrestalden SSI
2× Reaction mix (incl. dNTP)	Batch no. VOF-337	SSI
RNase free water	Batch no. VOF-286	SSI
DNase/RNase free water	Batch no. VOF-386	SSI
Big Dye Terminator v1.1 Cycle Sequencing RR-100	Lot no. 1408083	Applied Biosystems
Primers		
3.2 μM Forward primer A	Batch no. 06-04-2016 JRR	SSI
3.2 μM Reverse primer A	Batch no. 06-04-2016 JRR	SSI
3.2 μM Forward primer B	Batch no. 06-04-2016 JRR	SSI
3.2 μM Reverse primer B	Batch no. 06-04-2016 JRR	SSI
3.2 μM Forward primer C	Batch no. 06-04-2016 JRR	SSI
3.2 μM Reverse primer C	Batch no. 06-04-2016 JRR	SSI
Enzymes		
SuperScript III RT/Platinum Taq High Fidelity Enzyme Mix	Batch no. VOF-337	invitrogen
1 $\frac{\mu g}{mL}$ anti-ferret Igγ	Lot no. RI28521	Sigma-Aldrich
Precision Plus Protein™ All Blue Standard	Cat no. 161-0373	Bio-Rad
1 kb DNA ladder	Cat. no. N3232L	NEB
Anti-Rabbit Igγ(whole molecule)-Peroxidase antibody produced in goat	Batch no. 026M4782V	Sigma-Aldrich
Anti-Swine H1N1 hemagglutinin antibody produced in rabbit	Lot no. 523709031	Sigma-Aldrich

Table 2.4: Biological material including proteins, DNA and various mixes used in the experiments of this thesis. RDE indicates use of receptor destroying enzymes.

Component	Description	Supplied by
Dulbecco 1× PBS (pH 7.4)	Lot no. 1839839	gibco
5× Sequencing buffer	Lot no. 1407164	Applied Biosystems
PBS	Lot no. 6MB154	BioWhittaker
10 × TGS Buffer	Ref no. 161-0772	Bio-Rad
Minimum Essential Medium Eagle	Lot No. RNBF5679	Sigma-Aldrich
RNeasy Mini Kit		
RLT buffer	Lot no. 154032496	QIAGEN
RW1 buffer	Lot no. 154038591	QIAGEN
RPE buffer	Lot no. 154038250	QIAGEN
RNase free water	Lot no. 154037485	QIAGEN
illustra GFX PCR DNA and Gel Band Purification Kit		
Capture buffer type 3	Lot no. 9721121	GE Healthcare
Wash buffer type 1	Lot no. 9721121	GE Healthcare
Elution buffer type 6	Lot no. 9721121	GE Healthcare
MagNA Pure LC Total Nucleic Acid Isolation Kit		
Wash Buffer I	Ref no. 03038505001	Roche
Wash Buffer II	Ref no. 03038505001	Roche
Wash Buffer III	Ref no. 03038505001	Roche
Lysis/Binding Buffer	Ref no. 03038505001	Roche
Proteinase K	Ref no. 03038505001	Roche
Magnetic Glass Particles	Ref no. 03038505001	Roche
Elution Buffer	Ref no. 03038505001	Roche
BigDye XTerminator® Purification Kit		
SAM™ Solution	Lot no. 1609066	Applied Biosystems
Xterminator™ Solution Buffer™	Lot no. 1611076	Applied Biosystems
Pierce™ BCA Protein Assay Kit		
BCA Reagent A	Cat. No. 23225	ThermoFisher
BCA Reagent B	Cat. No. 23225	ThermoFisher
Albumin Standard Ampules, 2 mg/mL	Cat. No. 23225	ThermoFisher

Table 2.5: The kits and their components as used during this thesis.

A complete overview of the materials used over the course of this Master's thesis is featured in this chapter together with a concise methodology detailing the employed experimental procedures.

2.1 Materials

Growth medium

Growth medium for Madin-Darby canine kidney (MDCK) cells or MDCK cells transfected with cDNA of human 2,6-sialtransferase (SIAT) is prepared by adding 10 mL Penicillin Streptomycin, 10 mL of 200 mM L-Glutamin and 1 $\frac{\mu g}{mL}$ trypsin to 1 L of Minimum Essential Medium (MEM) Eagle.

RBC solution

A volume of 10 mL guinea pig blood is divided into three 50 mL Greiner tubes, and topped up with 1× Dulbecco's phosphate-buffered saline buffer (DPBS) to reach a total volume of 50 mL. The tubes are centrifuged at 2000 rounds per minute (rpm) for 10 min. The blood is then washed by discarding the supernatant, now containing the majority of serum, while avoiding remixing of the RBCs. The remaining RBCs are washed again by refilling each tube with 50 mL of 1× DPBS, after which they are centrifuged with equivalent settings. The supernatant is afterwards extracted with a Pasteur pipette to remove as much of it as possible. This procedure is done twice. The volume of the packed RBCs at the bottom of the tube is determined with a pipette and then diluted to a 10% RBC stock solution with 1× DPBS. 0.75 mL is extracted from the stock solution and mixed with 9.25 mL 1× DPBS to get a 0.75% RBC solution.

Lysis buffer C

The virus lysis buffer is a solution of 0.125 M NaCl, 20 mM Tris, pH 8.0 and 0.5% IGEPAL CA-630 dissolved in MilliQ-water.

2.1.1 Sequencing solutions

BigDye solution

Preparing a 1× BigDye solution for a single reaction requires 5 μL of DNase/RNase free water, 2 μL 5× Sequencing buffer and 1 μL Big Dye Terminator v1.1 Cycle Sequencing RR-100. The total volume of the solution can be scaled up depending on the amount of reactions required.

2.1.2 Western blot solutions

This section provides an overview of the preparation procedures of solutions and buffers required for western blotting.

Sample buffer

Preparing 40 mL sample buffer necessitates 4 mL of 0.5 M Tris buffer at pH 6.8, 4 mL glycerol, 0.8 g SDS, 0.6 g dithiothreitol (DTT) and 250 μL of 10% pyronin G. MilliQ-water is added last to obtain the final volume of 40 mL.

Western blot running buffer

A 1× western blot running (WBR) buffer is prepared by dissolving 100 mL of 10× Tris/glycine/SDS (TGS) buffer from Bio-Rad in 900 mL MilliQ-water.

Transfer buffer

Preparing 1 L of 1× transfer buffer solution requires 100 mL 10× TGS buffer, 286 mL 70% ethanol and 614 mL of MilliQ-water mixed together in a 1 L Blue-Cap bottle.

Tris buffer

Two Tris solution, an acidic and a basic, is required to prepare a 0.5 M Tris buffer at pH 7.5. The acidic 0.5 M Tris-hydrochloride is made my mixing 39.40 g of Tris-HCL with 500 mL MilliQ-water, while the basic 0.5 M Tris-base requires 15.14 g of Tris-base in 250 mL MilliQ-water. The 0.5 M Tris buffer is afterwards produced by adjusting the pH of the 0.5 M Tris-hydrochloride solution with the 0.5 M Tris-base until pH 7.5 is achieved.

Blocking buffer

The blocking buffer is prepared by mixing 100 mL of 0.5 M Tris buffer at pH 7.5 with 10 mL Tween 20 and 17.53 g NaCl. Lastly, MilliQ-water is added to obtain a final volume of 1000 mL blocking buffer. The buffer is stored at 5 °C.

Staining solution

Four DAB tablets are dissolved in 40 mL MilliQ-water in a 50 mL Greiner tube, which is covered in aluminium foil to reduce exposure of the solution to the light. 40 μL of 30% hydrogen peroxide is added shortly before use.

2.1.3 SPR solutions

Multiple solution are prepared to conduct the SPR experiments of this thesis.

SPR running Buffer

PBS from BioWhittaker is used as running buffer during these experiments.

Rinsing solutions

Rinsing Solution I has 25 mg of sodium dodecyl sulfate (SDS) dissolved in 5 mL MilliQ-water to obtain a 0.5% (w/w) concentration.

Rinsing Solution II has 37 mg of glycine dissolved in 10 mL MilliQ-water to obtain a 50 mM concentration. The pH is adjusted with sodium hydroxide to reach pH 9.5.

SPR elution buffer

A combined 1 M NaCl and 0.1 M sodium borate solution is prepared by mixing 3.81 g tetraborate decahydrate and 5.84 g NaCl in 100 mL MilliQ-water. The pH is adjusted with HCl to obtain pH 9.0. Precipitate is dissolved before each use by heating and stirring the solution.

Coupling buffer

A 5 mM coupling buffer contains 14.37 μL acetic acid mixed with MilliQ-water to obtain a final volume of 50 mL. The pH is adjusted with sodium hydroxide to pH 4.5. Afterwards, the solution is separated into 5 mL aliquots and stored at -20 °C.

Activation mix

A solution mix of 100 mM N-Hydroxysuccinimide (NHS) and 50 mM 2-Morpholinoethanesulfonic acid monohydrate (MES) is made by dissolving 173 mg NHS and 160 mg MES in 15 mL MilliQ-water. The pH is adjusted with sodium hydroxide to obtain pH 5.0. The solution is separated into aliquots of 1.5 mL and stored at -20 °C. 14 mg of N-(3-Dimethylaminopropyl)-N´-ethylcarboiimide HCl (EDC) is added to the unfrozen aliquot shortly before use, to make it a 100 mM NHS, 50 mM MES and 50 mM EDC activation mix.

Quenching buffer

Preparing a 1 M ethanolamine HCl solution with a total volume of 50 mL requires 3.07 mL ethanolamine and 46.93 mL MilliQ-water. HCl is used to lower the buffer pH to 8.5.

Regeneration solution

The glycine regeneration solution is prepared by dissolving 0.113 g of glycine in 15 mL of MilliQ-water to obtain a concentration of 100 mM. The solution is adjusted with HCl to get pH 2.0.

Capture solution

The solution of anti-ferret Abs used to capture the Abs present in the ferret antiserum raised against A/California/07/09 is made by mixing 15 μL of anti-ferret Igγ with 1485 μL coupling buffer.

2.2 Methods

2.2.1 Virus growth

IAV samples are grown in 48 well plates with confluent MDCK cells, i.e. approximately 50.000 adhering cells per well. The cell monolayer of each well is first washed with serum-free MEM Eagle. 100 μL growth medium, also serum-free, is added to each well afterwards. Cells are then inoculated with 100 μL of an IAV sample, resulting in a total of 200 μL in each well. Two wells containing no virus are used as controls. The finished plate is then incubated at 37 °C with 5% CO_2 for 1 hour. Afterwards, 800 μL of serum-free growth medium is added to each well. Infection of MDCK cells is then allowed to take place for 2-4 days, with regular monitoring of each culture. Once the MDCK cells are sufficiently lysed, the virus isolate is harvested and the product is ready for HA- and HAI-assays.

Virus lysate Virus isolates are inoculated as described in the previous section,

but infection is interrupted after 18-20 hours and the growth medium is removed from the wells (4 wells per virus isolate). 200 μL lysis buffer C per virus isolate is distributed among the four wells, which are then left to incubate for 5 min at ambient conditions. Once the adhering MDCK cells are free in solution, the contents of the wells are transferred to Eppendorf tubes and kept on ice for 5 min. The lysates are then centrifuged at 13000 rpm for 10 min at 4 °C. Lastly, the supernatant is transferred to another Eppendorf tube and stored at -80 °C.

2.2.2 HA/HAI-assay

The HA-assay is performed as a duplicate in a 96-well microtiter plate, exemplified in figure 2.1 **a**. 50 μL 1× PBS is added to each well, while column 1 has an additional 50 μL of the appropriate virus isolated added. A two-fold serial dilution of the virus isolate is carried out left to right, and the remainder is discarded. The dilution of each column can be read in figure 2.1 **a**. 50 μL of the 0.75% RBC solution is then added to each well, and the plate is lightly shaken to mix the contents in the individual wells. The assay is then incubated at room temperature for 1-2 hours.

The results of the HA-assay are evaluated in order to proceed with the HAI-assay, which is carried out as duplicate experiments with back titration for each virus isolate. The HA-titer corresponds to 1 HA-unit, and the virus isolate is diluted with 1× DPBS to 4 HA-units, reaching an approximate amount of 1.5 mL. A new microtiter plate is prepared with 25 μL 1× PBS in each well and 25 μL of 10% A/California/07/09 ferret antiserum in the first column, excluding the wells in the rows meant for back titration, i.e. rows C and F. These two rows will instead have 50 μL of the diluted virus isolate added to columns 1 and 2, and 50 μL of 1× DPBS added in columns 2-12. A two-fold serial

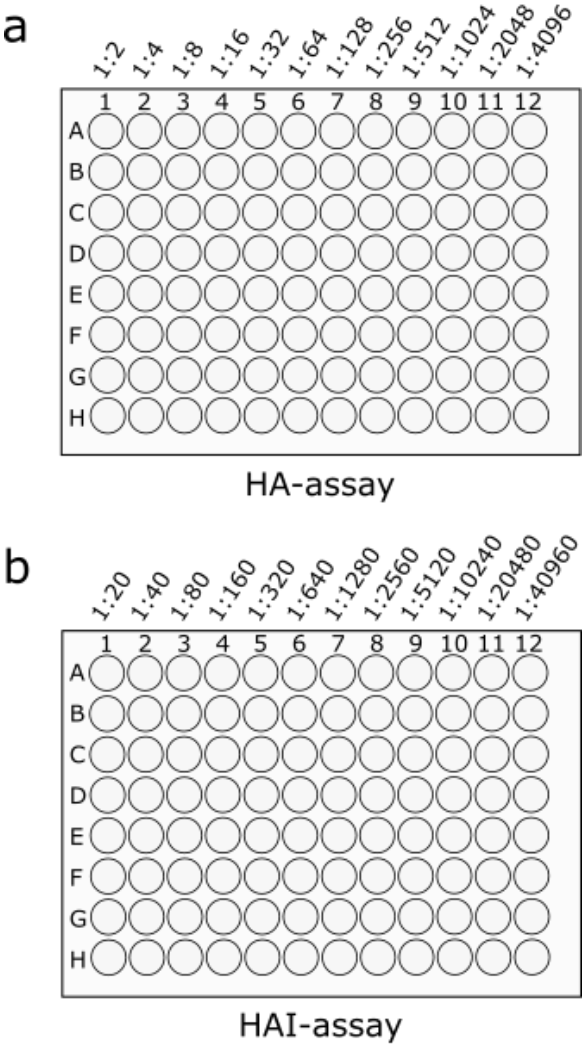

Figure 2.1: a. The dilution of each well in an HA-assay. **b.** The dilution of each well in an HAI-assay. The titers for both assays is recorded as the numbers after the colon.

dilution of rows C and F is initiated from column 2.

The ferret antiserum is also distributed as a two-fold serial dilution, starting with a 1:20 dilution in column 1, rows A, B, D and E, and the dilutions can be seen in figure 2.1 **b**. 25 μL of the diluted virus isolate added to each of these wells, and the plate is gently stirred. The Ab-antigen reaction is allowed to take place for 30 min at room temperature. Lastly, 50 μL of 0.75% RBC solution is added to each well and the plate is gently stirred again before incubating for 1-2 hours, after which the titers can be observed. Testing multiple IAV samples and the vaccine virus against the Abs raised to the vaccine virus gives an indication of antigenic changes the IAV samples have obtained compared to the vaccine virus.

2.2.3 Sequencing

Sequencing with Sanger's polymerase chain reaction (PCR) based method requires DNA. The starting point of these experiments is (-)ssRNA obtained from the IAV samples, extracted from the viral particles using the RNeasy Mini Kit, which is then reverse transcribed into cDNA.

RNA extraction

To extract the RNA, 400 μL of RLT buffer, a lysis buffer containing 10 $\frac{\mu L}{mL}$ of β-mercaptoethanol, is added to 200 μL of IAV isolate in a 1.5 mL Eppendorf tube and subsequently vortexed for 15 s. The solution is then centrifuged for 15 s at 8000× g and left to incubate at room temperature for 5 min. 600 μL of 70% ethanol is added to the 600 μL lysate, and carefully homogenised with a pipette. 600 μL of this solution is transferred to an RNeasy spin-column and centrifuged for 15 s at 8000× g. The flow-through is discarded, and the process is repeated with the remaining solution. Afterwards, 700 μL of RW1 buffer, a washing buffer, is added to the RNeasy spin-column, which is then centrifuged for 15 s at 8000× g. The flow-through is discarded. Then 500 μL of RPE buffer, another washing buffer, is added to the RNeasy spin-column, which is centrifuged at the identical settings, after which the flow-through is discarded. The same volume of RPE buffer is added again, and the RNeasy spin-column is centrifuged for 2 min at 8000×

RT-PCR		
Process	**Temperature**	**Time**
RT	55 °C	30 min
Denaturation	94 °C	2 min
Cycling 40×	94 °C	15 min
	60 °C	30 min
	68 °C	2 min
Extension	68 °C	5 min

Table 2.6: The PCR cycle programme used for RT-PCR.

g. This time, the collection tube is discarded along with the flow-through, and the RNeasy spin-column is moved to a new collection tube, after which is it centrifuged for an additional minute at 12.000× g. The RNeasy spin column is placed in a 1.5 mL Eppendorf tube and eluted by adding 60 μL of RNase free water to the RNeasy spin-column and centrifuging for 1 min at 8000× g. The final product is stored at -80 °C.

The MagNA Pure LC Total Nucleic Acid Isolation Kit is also used to extract RNA from the IAV samples. The procedure is handled by the MagNA Pure LC 1.0 instrument, which require a certain volume of the following reagents:

1. 17.6 mL Wash Buffer I
2. 8.20 mL Wash Buffer III
3. 5.80 mL Lysis/Binding Buffer
4. 2.60 mL Proteinase K
5. 3.40 mL Magnetic Glass Particles
6. 3.90 mL Elution Buffer
7. 8.70 mL Wash Buffer II

These reagents are poured into plastic trays and loaded into their appropriate positions in the MagNA Pure LC 1.0 instrument. The eluted volume is 100 μL.

The products from both RNA extraction methods are run with real-time PCR, to check for contamination and occurrence of other influenza strains, e.g. H3N2 or influenza B. The results are not included in this report, but CT-values of selected IAV samples pertaining to the SPR experiments are included in the result chapter.

Reverse transcriptase PCR

A reverse transcriptase (RT) PCR-mix, named H1 SuperScript III in the context of this report, is required to initiate RT-PCR. It is prepared for 25 μL or 50 μL reactions with the following content in each tube:

1. **(2×)** 12.5 μL 2× Reaction mix
2. **(2×)** 1.0 μL Forward primer A
3. **(2×)** 1.0 μL Reverse primer A
4. **(2×)** 0.8 μL SuperScript III RT-/Platinum Taq High Fidelity Enzyme Mix
5. **(2×)** 7.2 μL RNase free water

These 22.5 or 45 μL aliquots are stored at -80 °C.

Primer set-up of RT-PCR product A		
Well	**Sample**	**Primer**
1	1 μL RT-PCR product A	1 μL 3.2 μM Forward primer A
2	1 μL RT-PCR product A	1 μL 3.2 μM Reverse primer A
3	1 μL RT-PCR product A	1 μL 3.2 μM Forward primer B
4	1 μL RT-PCR product A	1 μL 3.2 μM Reverse primer B
5	1 μL RT-PCR product A	1 μL 3.2 μM Forward primer C
6	1 μL RT-PCR product A	1 μL 3.2 μM Reverse primer C

Table 2.7: The primer set-up for a single IAV sample. These 2 μL are mixed with the 8 μL of 1× BigDye solution.

2.5 or 5.0 μL of the appropriate RNA extract, obtained with the RNeasy Mini Kit or MagNa Pure LC, is added to 22.5 or 45 μL RT PCR-mix before beginning RT-PCR, giving a total volume of 25 or 50 μL for each reaction aliquot. Once dissolved, the RNA extract is spun down, and RT-PCR is initiated with either C1000™ or S1000™ Thermal Cycler from Bio-Rad, using the cycle programme outlined in table 2.6.

PCR product clean-up

Once the H1 gene has been PCR amplified it is run on a 0.8% agarose gel, containing ethidium bromide, for 30 min. The wells are loaded with 10 μL RT-PCR product and 10 μL RNase/DNase free water each. The DNA ladder is mixed with water in equal proportions an amount (1:1). The resulting DNA bands are imaged with UV light and the sizes of the RT-PCR products are determined.

The RT-PCR products are purified with the *illustra GFX PCR DNA and Gel Band Purification Kit*. First the entire volume of the remaining RT-PCR product is mixed with 500 μL of Capture buffer type 3 in a 1.5 mL Eppendorf tube and subsequently vortexed. It is centrifuged briefly to aggregate the solution at the bottom of the tube, after which is it transferred to a GFX MicroSpin column placed in a collection tube, and centrifuged at 16.000× g for 30 s to capture the RT-PCR product. Flow-through is discarded, 500 μL of Wash buffer type 1 is added to the column, and the reassembled column and collection tube is centrifuged at 16.000× g for 30 s. The collection tube and flow-through is discarded, and the GFX MicroSpin column is placed in a DNase-free 1.5 mL microcentrifuge tube. Elution is done by first adding 10-50 μL of Elution buffer type 6 centrally on the membrane of the column and incubating it for 1 min at room temperature. The amount of Elution buffer type 6 added to each sample is determined based on concentration estimates from the gel image, i.e. higher concentrations require more Elution buffer type 6. It is then centrifuged at 16.000× g for 1 min, and the flow-through containing the purified RT-PCR product is stored at -20 °C.

Sequencing PCR		
Process	**Temperature**	**Time**
Denaturation	94 °C	2 min
Cycling 25×	96 °C	10 s
	50 °C	5 s
	60 °C	4 min
Cooling	4 °C	15 min

Table 2.8: The PCR cycle programme used for sequencing PCR.

Sequencing PCR

The RT-PCR product is fragmented to prepare it for sequencing, by running an additional PCR with the product and six primers at different, predetermined positions. This procedure creates overlaps in the sequences, which provides a better base for the sequencing data. 8 μL of 1× BigDye solution is distributed into each well of a micro plate, together with 1 μL of appropriate RT-PCR product and 1 μL of an appropriate primer, resulting in 10 μL in each well. Each RT-PCR product therefore requires six wells to accommodate all the primers. An example of the primer set-up for a single sample is listed in table 2.7.

Once every reaction is prepared, the micro plate is briefly centrifuged to aggregate the solution at the bottom of the wells, after which the plate is run with the PCR cycle programme seen in table 2.8. The amplified fragments are stored at -20 °C.

Sequencing purification

The fragmented PCR products are additionally purified to optimise the reagents for sequencing. Each well has 45 μL of SAM solution and 10 μL of Xterminator™ solution buffer™ added, after which the wells are sealed and the plate is shaken on a micro plate shaker for 30 min at 2000 rpm. The plate is then centrifuged at 1000× g for 2 min, and the contents of the micro plate are thoroughly mixed with a pipette and transferred to another micro plate suitable for the sequenator. The wells are sealed and the micro plate is centrifuged at the same settings, after which it is loaded into the HITACHI 3500 Genetic Analyzer, and sequencing is initiated.

Sequence alignment

The total length of the H1 HA gene, 1777 base pairs (bps), is covered across six DNA fragments as seen in figure 2.2. These frag-

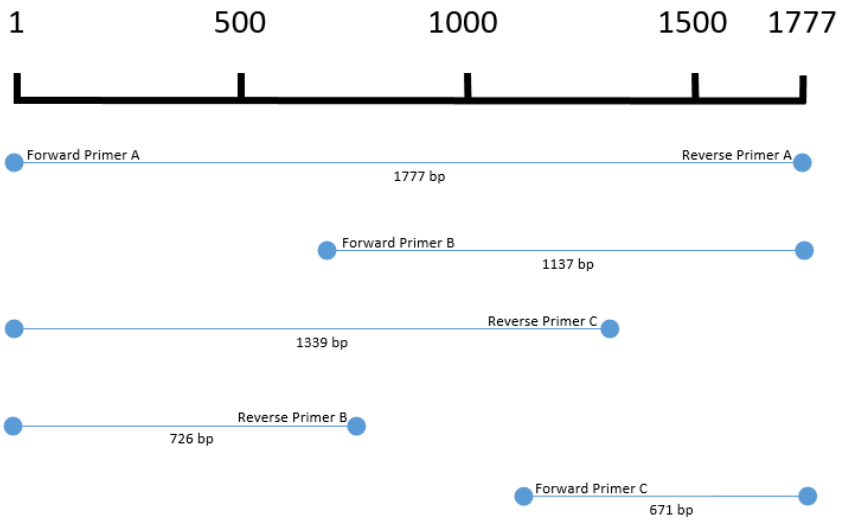

Figure 2.2: The five DNA fragments of the H1 gene as sequenced by the HITACHI 3500 Genetic Analyzer. The upper, 1777 bps fragment is sequenced twice from both ends. Image is courtesy of Jesper Rønn.

ments ensure at least triple coverage of the gene at any given point in the sequence. The upper, 1777 bps fragment is only depicted once, but is sequenced twice, once with Forward primer A and again with the Reverse primer A. However, the limitations of Sanger sequencing means that these two fragments will not be sequenced for the full 1777 bps, but between 800-1000 bps. [64]

The data is assembled and edited in BioNumerics, before being transferred to MEGA7, where the MUSCLE algorithm is used to align the cDNA sequences with the UPMGB clustering method and a maximum of 8 iterations [70]. The cDNA sequences are translated in MEGA7, and the antigenic sites of the protein sequences, as defined in table 1.2, are examined for substitutions. These are identified based on changed amino acid in the sequence when compared to the following source sequence: [71]. A total of 23 IAV H1 gene samples are sequenced, of which the full protein sequences can be found in the attached CD.

Phylogenic tree

A phylogenetic tree is generated from the fully sequenced IAV H1 cDNA sequences using the Tamura-Nei based Maximum Likelihood method in MEGA7 [72, 73]. Preliminary trees are acquired by applying algorithms BioNJ and Neighbor-Join to a matrix of pairwise distances, which were estimated with the Maximum Composite Likelihood approach. Included codon positions are 1st+2nd+3rd+Non-coding and the length of the branches indicate the amount of substitutions per nt. The percentage of trees with affiliated taxa that are clustered together is displayed at each branch. 1685 nt positions are included in the dataset, and gaps and missing data are omitted. A/California/07/09 is chosen as the root of the tree.

2.2.4 Glycosylation

Glycosylation analysis

The positions of glycosylation sites on the HA of the sequenced IAV samples are investigated by submitting the amino acid sequences to probability calculations with

NetNGlyc result	Parameters
Glycosylated sites	
+	Potential > 0.5
++	Potential > 0.5 AND Jury agreement (9/9) OR Potential > 0.75
+++	Potential > 0.75 AND Jury agreement
++++	Potential > 0.9 AND Jury agreement
Non-glycosylated sites	
−	Potential < 0.5
−−	Potential < 0.5 AND Jury agreement
− − −	Potential < 0.32 AND Jury agreement

Table 2.9: The NetNGlyc rankings and how they are determine based on the two parameters, namely potential and Jury agreement.

the NetNGlyc 1.0 Server [74]. This software can be used to predict N-linked glycosylation sites in human proteins by locating Asn-X-Ser/Thr sequons. The predicted glycosylation sites of each sequenced IAV sample are ranked on the scale seen in table 2.9, which is based on two parameters. First is the potential score of the glycosylation site, a value between 0.00 and 1.00, which is an average value based on calculations from nine neural networks. The second parameter is jury agreement, which shows how many of said networks agree with the potential score. [74]

Western blot

Preparing IAV lysates for SDS-PAGE requires 72 μL of IAV lysate sample and 18 μL sample buffer in a 1.5 mL Eppendorf tube, resulting in a total volume of 90 μL in each tube. The solutions are incubated at 100 °C for 3 min to denature the viral proteins. Afterwards, the solutions are centrifuged for 5 min at 12000 rpm to aggregate cell debris at the bottom of the tube.

The Novex™ WedgeWell™ 10% Tris-Glycine gel used for the SDS-PAGE is extracted from the bag and washed with deionised water. Comb and tape is removed from the gel, and the wells are carefully washed with 1× WBR buffer. The wells are then emptied and the gel is placed in the electrophoresis tank and mounted properly. The inner chamber of the electrophoresis tank is filled with 1× WBR buffer until the wells are covered, and then checked for leaks. Afterwards, the outer chamber is also filled with 1× WBR buffer. It is important to use clean equipment and gloves in all procedures to minimise contamination.

Control, an MDCK cell lysate, and sample solutions are loaded onto the Tris/Glycine gel, initially 60 μL per well, and a lane with 7.0 μL of Precision Plus Protein™ All Blue Standard. Electrophoresis is then initiated with a constant potential of 125 V, resulting in a current of 30 mA, and left to run for 7 min. The remaining 30 μL of each sample solution is then added to the appropriate wells and electrophoresis is resumed for a total of 90 min. Later, the gel is taken out of the electrophoresis tank and the surrounding cassette is carefully removed.

A sheet of nitrocellulose transfer membrane and two extra thick filter papers are soaked in a tray containing transfer buffer. The nitrocellulose membrane is aligned on top of a piece of filter paper, and subsequently squeezed with a cylindrical ob-

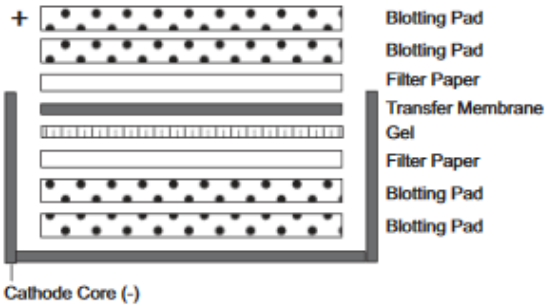

Figure 2.3: The setup for western blot experiments. The negatively charged proteins migrate from the gel towards the anode, at which point they are caught in the nitrocellulose membrane. The amount of blotting pads can vary. [75]

ject to ensure that no air bubbles are present between the sheets. The gel from the previous step is placed on top of the nitrocellulose membrane, and new air bubbles are squeezed out. Finally, the last piece of filter paper is placed on top of the gel, forming a sandwich. It is then rolled once again to remove air bubbles, and then interposed between an amount of blotting pads, soaked in transfer buffer, sufficient to ensure adequate contact between the two electrodes. The final membrane transfer set-up can be seen in figure 2.3. The blot is left to run for 2 hours at 125 V and 200 mA.

After blotting, the lane on the nitrocellulose membrane representing the protein ladder is cut out, while the rest of the membrane is left to incubate in a tray containing the blocking buffer for 30 min on an orbital shaker at room temperature.

The nitrocellulose membrane is then submerged in the primary Ab solution containing the Anti-Swine H1N1 hemagglutinin Ab diluted 1:600 in blocking buffer, and left to incubate for 2 hours at room temperature on an orbital shaker. It is afterwards washed with blocking buffer for 3×5 min on an orbital shaker.

The same procedure is used for the secondary Ab, HRP conjugated Anti-Rabbit Igγ(whole molecule)-Peroxidase Ab, diluted 1:5000. Afterwards, the blot is moved to a new tray containing 40 mL of a freshly made staining solution, which is allowed to stain the membrane for 5-30 min. The nitrocellulose membrane is then washed with MilliQ-water for 30 s, followed by a 10 min soaking in MilliQ-water on an orbital shaker. Finally the blot is left to dry between two new filter papers, where it can also be stored once dry. It is important to reduce the exposure to light of the nitrocellulose membrane after adding HRP conjugated Abs.

Protein concentration assay

The concentrations of the IAV lysates are assayed using the Pierce™ Bicinchoninic Acid Assay (BCA) Protein Assay Kit. First, the working reagent (WR) solution is prepared by mixing BCA Reagent A and BCA reagent B 50:1 with an adequate finale volume. The IAV lysates are distributed into a 96-well microtiter plate, each sample receiving two wells with 5 μL in each. Additionally, six standard (BSA) samples with concentrations ranging from 0-2000 $\frac{\mu g}{\mu L}$, are distributed similarly. 40 μL of WR solution is then added to each well, after which the plate is vortexed and then incubated at 37 °C for 30 min.

The absorbance of the content of each well is then measured at 562 nm, with the micro plate being vortexed briefly with every third measurement. Measurements are performed on the NanoDrop ND-1000 Spectrophotometer. Concentrations can then be estimated by first generating an absorbance standard curve from the measured BSA samples, and then com-

paring absorption values to obtain the concentration of the IAV lysates.

2.2.5 SPR

The sensorchip surface must undergo several binding event to facilitate Ab-antigen binding measurements. These binding events are illustrated in figure 2.4. **(1)** Initially, the sensorchip surface (green) is activated with NHS/EDC. **(2)** Afterwards, the the primary ligand (black) is injected onto the now activated surface (yellow). **(3)** The primary ligand then gets immobilised on the activated sensorchip surface. **(4)** The secondary ligand (blue) is injected onto the sensorchip surface. **(5)** The secondary ligand binds to the primary ligand by means of Ab-Ab interactions. **(6)** Lastly, the analyte is injected and binds to the secondary ligand. The binding between Ab-antigen is the interaction of interest.

Prior to SPR experiments, IAV samples are diluted in 1× DPBS buffer according to their CT-value. Afterwards, these solutions are inactivated by adding 10.5 μL of 7.5% formalin to 1989.5 μL of diluted IAV sample, resulting in a 0.04% concentration of formalin. These solutions are vortexed thoroughly, then incubated at 37 °C for 24 hours.

The SPR experiments are performed on a Reichert2SPR Dual Channel System SPR instrument. The flow cell is removed and the sapphire prism surface is rinsed with ethanol and subsequently swiped with a lens cleaning tissue until sufficiently clean. A droplet of immersion oil is distributed onto the sapphire prism surface and an

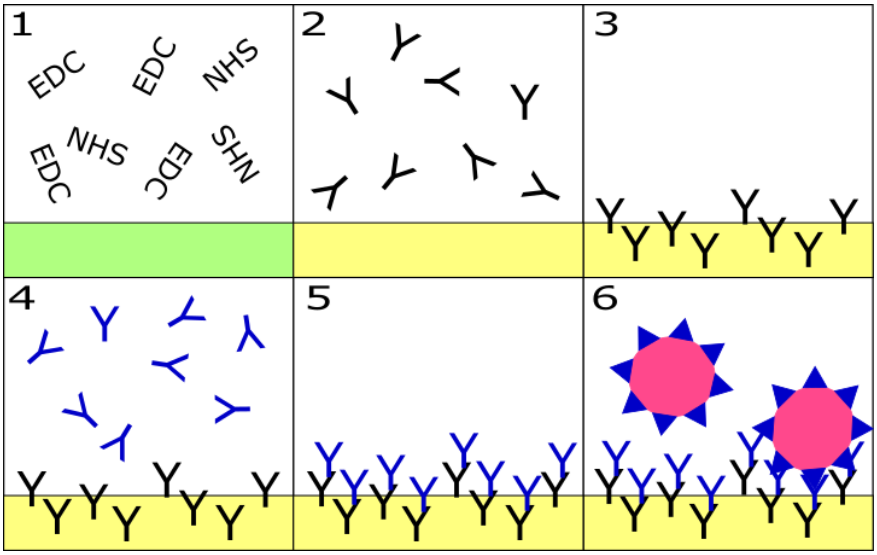

Figure 2.4: Schematic representation of the binding events occuring during SPR measurements. **(1)** The linear polycarboxylate hydrogel surface (green) is activated with NHS/EDC. **(2)** The capture solution containing the primary ligand (black) is injected. **(3)** The primary ligand is immobilised on the activated sensorchip surface (yellow). **(4)** The antiserum solution containing the secondary ligand (blue) is injected. **(5)** The secondary ligand binds to the primary ligand. **(6)** The analyte solution is is injected, and the viral particles bind to the secondary ligand.

SPR sensorchip HC 200 M is mounted in the slot. This sensorchip has a 200 nm thick linear polycarboxylate hydrogel coating with medium porosity. The dual channel systems allows for division of the sensorchip into a working partition, which is where primary experiments take place, and a reference partition which functions as a control. The flow cell is put back and the sensorchip is cleaned using the cleaning protocol seen in table 2.10, in order to obtain a stable baseline. The sensorchip is supplied with MilliQ-water during each "Wait" step, and a steady stream of 10 $\frac{\mu L}{min}$ MilliQ-water is applied in between protocol runs.

The cleaning protocol ensures that the majority of trace proteins in the flow system are removed, which is accomplished by the SDS in rinsing solution I. Rinsing solution II is used to clear out the SDS left behind from rinsing solution I, and finally the SPR elution buffer is used to condition the surface for subsequent steps.

Immobilisation of the primary ligand is managed with the Ab immobilisation protocol seen in table 2.11, and a sensorgram pertaining to the immobilisation of the primary ligand can be seen in figure 2.5, highlighting the most important steps. The sensorchip is made chemically active with a 250 μL injection of activation mix, which sees an increment in the signal. A subsequent 25 μL injection of coupling buffer is responsible for providing the best environment for the primary Abs to covalently bind to the activated sensorchip surface. The reference partition is then taken out of the flow loop, by diverting the valve and only supplying capture solution through

Cleaning protocol

Step	Description	Duration
1.	Refill pump at 25000 $\frac{\mu L}{min}$	N/A
2.	Infuse flow rate at 50 $\frac{\mu L}{min}$	N/A
3.	250 μL injection of rinsing solution I	5 min
4.	Wait	5 min
5.	250 μL injection of rinsing solution II	5 min
6.	250 μL injection of rinsing solution II	5 min
7.	Wait	5 min
8.	250 μL injection of SPR elution buffer	3 min
9.	Wait	5 min
10.	250 μL injection of SPR elution buffer	3 min

Table 2.10: An overview of the programme used for the cleaning protocol. The protocol has a total runtime of at least 36 min.

the left valve to the working partition, in order to avoid functionalisation of the reference partition. A total of 500 μL capture solution is supplied for 20 min in two separate instances, as evident from the significant increase of the signal. The reaction is allowed to take place for 20 min, after which some of the particles dissociate. The remaining NHS esters are then quenched with the two injections of 200 μL quenching buffer over 8 min, and a new baseline is acquired.

Before proceeding with the actual measurements, the filtered MilliQ-water supplied during "Wait" and between protocol runs is substituted with SPR running buffer. The measurement protocol is listed in table 2.12 and an example of a sensorgram pertaining to said protocol can be seen in figure 2.6. The now functionalised sensorchip is first injected with a 200 μL solution of ferret antiserum raised to A/California/07/09, diluted 1:10 in PBS buffer, for 5 min, which induces a steep increase in the signal. After dissociation and stabilisation, 200 μL of the antigen, i.e. inactivated IAV sample appropriately diluted in PBS buffer, is injected for 5 min and then dissociated again. The sensorchip is then injected with 200 μL of SPR running buffer for 5 min to make a blank measurement, before making a brief injection of 20

Ab immobilisation protocol

Step	Description	Duration
1.	Refill pump at 25000 $\frac{\mu L}{min}$	N/A
2.	Infuse flow rate at 25 $\frac{\mu L}{min}$	N/A
3.	250 μL injection of activation mix	10 min
4.	250 μL injection of coupling buffer	1 min
5.	Diverter valve (Left only: On)	N/A
6.	Refill pump at 25000 $\frac{\mu L}{min}$	N/A
7.	250 μL injection of capture solution	10 min
8.	250 μL injection of capture solution	10 min
9.	Wait	20 min
10.	Diverter valve (Left only: Off)	N/A
11.	Infuse flow rate at 50 $\frac{\mu L}{min}$	N/A
12.	250 μL injection of quenching solution	4 min
13.	250 μL injection of quenching solution	4 min
14.	Wait	10 min

Table 2.11: An overview of the programme used for the primary Ab immobilisation protocol. The protocol has a total runtime of 1 hour and 9 min.

Measurement protocol

Step	Description	Duration
1.	Refill pump at 25000 $\frac{\mu L}{min}$	N/A
2.	Infuse flow rate at 40 $\frac{\mu L}{min}$	N/A
3.	250 μL injection of ferret antiserum	5 min
4.	Dissociation	5 min
5.	Wait	2 min
6.	Refill Pump	N/A
7.	Wait	3 min
8.	250 μL injection of antigen	5 min
9.	Dissociation	5 min
10.	Wait	3 min
11.	250 μL injection of SPR running buffer	5 min
12.	Dissociation	5 min
13.	250 μL injection of regeneration solution	30 sec
14.	Dissociation	5 min

Table 2.12: An overview of the programme used for the measurement protocol. The protocol has a total runtime of at least 43 min and 30 s.

μL regeneration solution to clear the sensorchip surface of antigen and ferret antiserum, and prepare it for subsequent measurements.

The results of SPR are analysed with TraceDrawer.

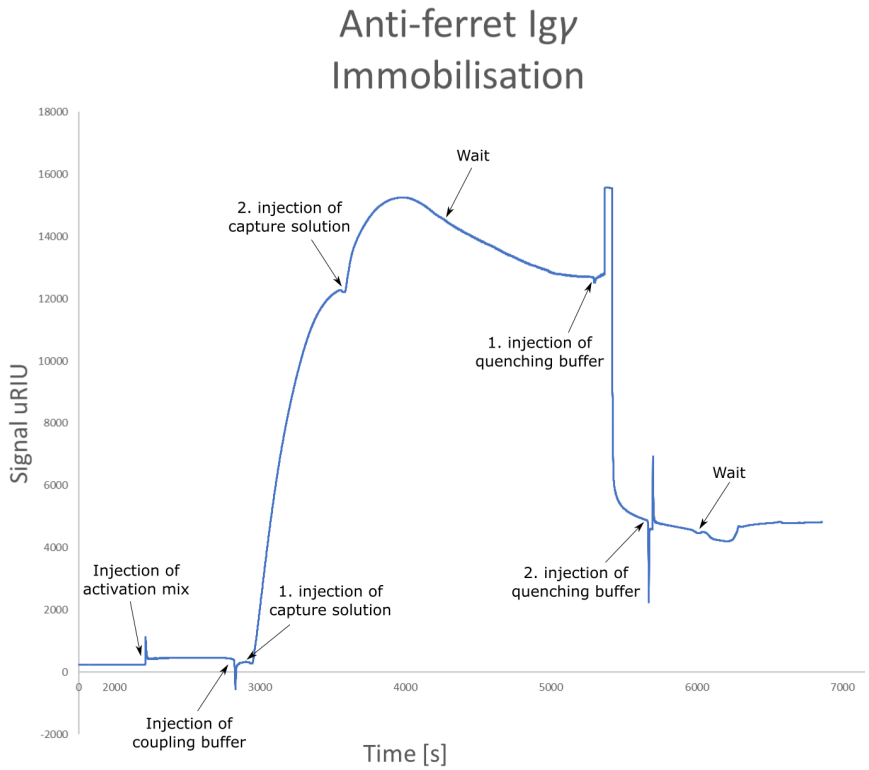

Figure 2.5: Immobilisation of anti-ferret Igγ. Injection points are marked with arrows, and correlate with the steps listed in 2.11. This is the signal with the reference partition data subtracted.

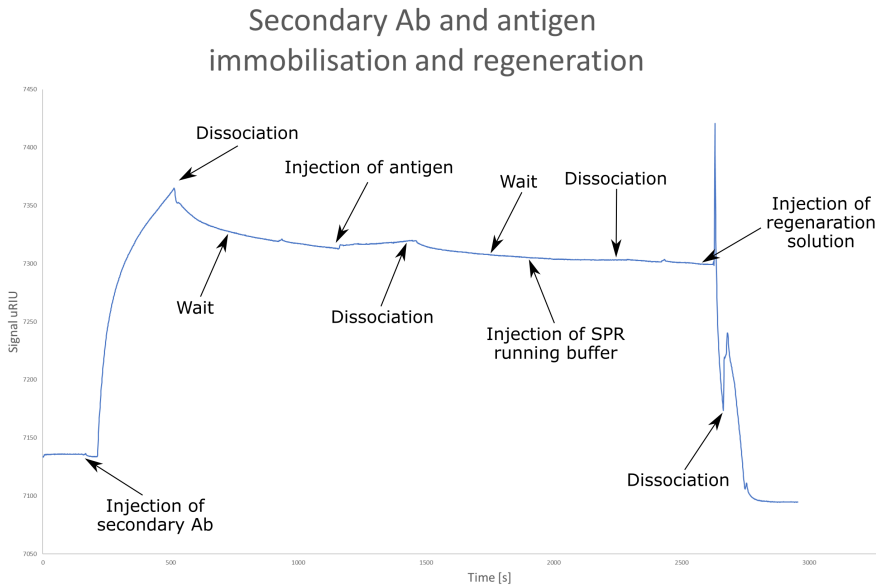

Figure 2.6: Immobilisation of ferret antiserum and antigen. Injection points and dissociations are marked with arrows, and correlate with the steps listed in 2.12. The reference partition data is subtracted in this example.

3. Results

3.1 IAV sample sequencing

All 24 complete IAV H1 sequences, aligned and translated in MEGA7, were manually investigated at the residues specified in table 1.2. This analysis is summarised in table 3.1, where no substitutions appear as N/M. Amino acids abbreviations can be found in Appendix B.1. Figure 3.2 is included for reference, to better illustrate which sites, e.g. receptor-binding and cleavage site, are in sequential proximity of the antigenic sites. The sequences in this figure start at a different residue, meaning that there will be some differences in the numbering.

It is evident from table 3.1 that substitutions were recurring at certain sites, such as Ser220→Thr in the Ca1 site, Lys180→Gln in the Sa site from season 2013-2014 onwards, and Ser202→Thr in the Sb site after the 2009-2010 season. Additionally, Ser179→Asn was introduced along with Lys180→Gln in the 2015-2016 season. The remaining antigenic sites had few, sporadic substitutions, and there were two unknown amino acids at positions 172 and 239 for A/Denmark/120/10 and A/California/07/09, respectively. Investigating the DNA sequence of both IAV H1 samples revealed that A/California/07/09 had a purine (R) in position 716 of the H1 gene, which according to the IUPAC nt codes, listed in Appendix B.2, translates to either A or guanine (G). This means that the unknown amino acid was either Gly or Asp. The unknown nt of A/Denmark/120/10 was also an R, and the corresponding amino acid was thus either Gly or Asp.

No substitutions occurred in the CB site. It is placed solely on the vestigial esterase, as evident in figure 3.2, and is the furthest from the tip of the head domain out of all the antigenic sites. However, it is situated 16 amino acids away from a single residue belonging to the RBS.

The phylogenetic tree generated from the fully sequenced IAV H1 samples can be seen in figure 3.1. The taxa pertaining to the same season are coloured alike, e.g. green for the 2009-2010 season. The root is coloured red. The figure displays the evolution of the sequenced IAVs, which showcases increased distance in the sequences as the virus is allowed to persist in the population.

The phylogenetic tree depicts an expected coherence within IAV H1 samples of the

IAV samples		Sequence analysis							
	Antigenic site	Cb site	Sa site		Ca1 site		Ca2 site		Sb site
	Amino acid No.	87-92	141-142	170-174	176-181	183-187	220-222	252-254	154-159	238-239	201-212
	Amino acid Sequence	L-S-T-A-S-S	P-N	K-K-G-N	P-K-L-S-K-S	I-N-D-K-G	S-S-R	E-P-G	P-H-A-G-A-K	R-D	T-S-A-D-Q-Q-S-L-Y-Q-N-A
Season 2009-2010	A/California/07/09	N/M	N/M	N/M	N/M	N/M	N/M	N/M	N/M	Asp239→?	N/M
	A/Denmark/1323/09	N/M	N/M	N/M	N/M	N/M	Ser220→Thr	Glu252→Lys	N/M	N/M	Thr201→Asn
	A/Denmark/1325/09	N/M	N/M	Lys171→Gln	N/M	N/M	Ser220→Thr	N/M	N/M	N/M	N/M
	A/Denmark/1337/09	N/M	N/M	N/M	N/M	N/M	Ser220→Thr	N/M	N/M	N/M	N/M
	A/Denmark/1338/09	N/M	N/M	N/M	N/M	N/M	Ser220→Thr	N/M	N/M	N/M	N/M
Season 2010-2011	A/Denmark/113/10	N/M	N/M	N/M	N/M	N/M	Ser220→Thr	N/M	N/M	N/M	Ser202→Thr
	A/Denmark/115/10	N/M	N/M	N/M	N/M	N/M	Ser220→Thr	N/M	N/M	N/M	Ser202→Thr
	A/Denmark/116/10	N/M	Asn141→Asp	N/M	N/M	N/M	Ser220→Thr	N/M	N/M	N/M	N/M
	A/Denmark/120/10	N/M	N/M	Gly172→?	N/M	N/M	Ser220→Thr	N/M	N/M	Asp239→Gly	Ser202→Thr
Season 2013-2014	A/Denmark/106/13	N/M	N/M	N/M	Lys180→Gln	N/M	Ser220→Thr	N/M	N/M	N/M	Ser202→Thr
	A/Denmark/05/14	N/M	N/M	N/M	Lys180→Gln	N/M	Ser220→Thr	N/M	Ala156→Asp	N/M	Ser202→Thr
	A/Denmark/11/14	N/M	N/M	N/M	Lys180→Gln	N/M	Ser220→Thr	N/M	N/M	N/M	Ser202→Thr
	A/Denmark/38/14	N/M	N/M	N/M	Lys180→Gln	N/M	Ser220→Thr	N/M	N/M	N/M	Ser202→Thr
Season 2014-2015	A/Denmark/51/14	N/M	N/M	N/M	Lys180→Gln	N/M	Ser220→Thr	Glu252→Asp	N/M	N/M	Ser202→Thr
	A/Denmark/03/15	N/M	N/M	N/M	Lys180→Gln	N/M	Ser220→Thr	N/M	N/M	N/M	Ser202→Thr
Season 2015-2016	A/Denmark/41/15	N/M	N/M	N/M	Lys180→Gln+ Ser179→Asn	N/M	Ser220→Thr	N/M	N/M	N/M	Ser202→Thr
	A/Denmark/43/15	N/M	N/M	N/M	Lys180→Gln+ Ser179→Asn	N/M	Ser220→Thr	N/M	N/M	N/M	Ser202→Thr
	A/Denmark/48/15	N/M	N/M	N/M	Lys180→Gln	N/M	Ser220→Thr	N/M	N/M	N/M	Ser202→Thr
	A/Denmark/20/16	N/M	N/M	N/M	Lys180→Gln	N/M	Ser220→Thr	N/M	N/M	N/M	Ser202→Thr
	A/Denmark/27/16	N/M	N/M	N/M	Lys180→Gln+ Ser179→Asn	N/M	Ser220→Thr	N/M	N/M	N/M	Ser202→Thr
	A/Denmark/30/16	N/M	N/M	N/M	Lys180→Gln+ Ser179→Asn	N/M	Ser220→Thr	N/M	N/M	N/M	Ser202→Thr
	A/Denmark/41/16	N/M	N/M	N/M	Lys180→Gln+ Ser179→Asn	N/M	Ser220→Thr	N/M	N/M	N/M	Ser202→Thr
	A/Denmark/46/16	N/M	N/M	N/M	Lys180→Gln+ Ser179→Asn	Gly187→Glu	Ser220→Thr	N/M	N/M	N/M	Ser202→Thr

Table 3.1: Substitutions for all fully sequenced IAV H1 samples and the corresponding antigenic site listed according to season. Substitutions are indicated with the original amino acid and position pointing to the newly introduced amino acid, e.g. Gly187→Glu. No substitutions are indicated with N/M.

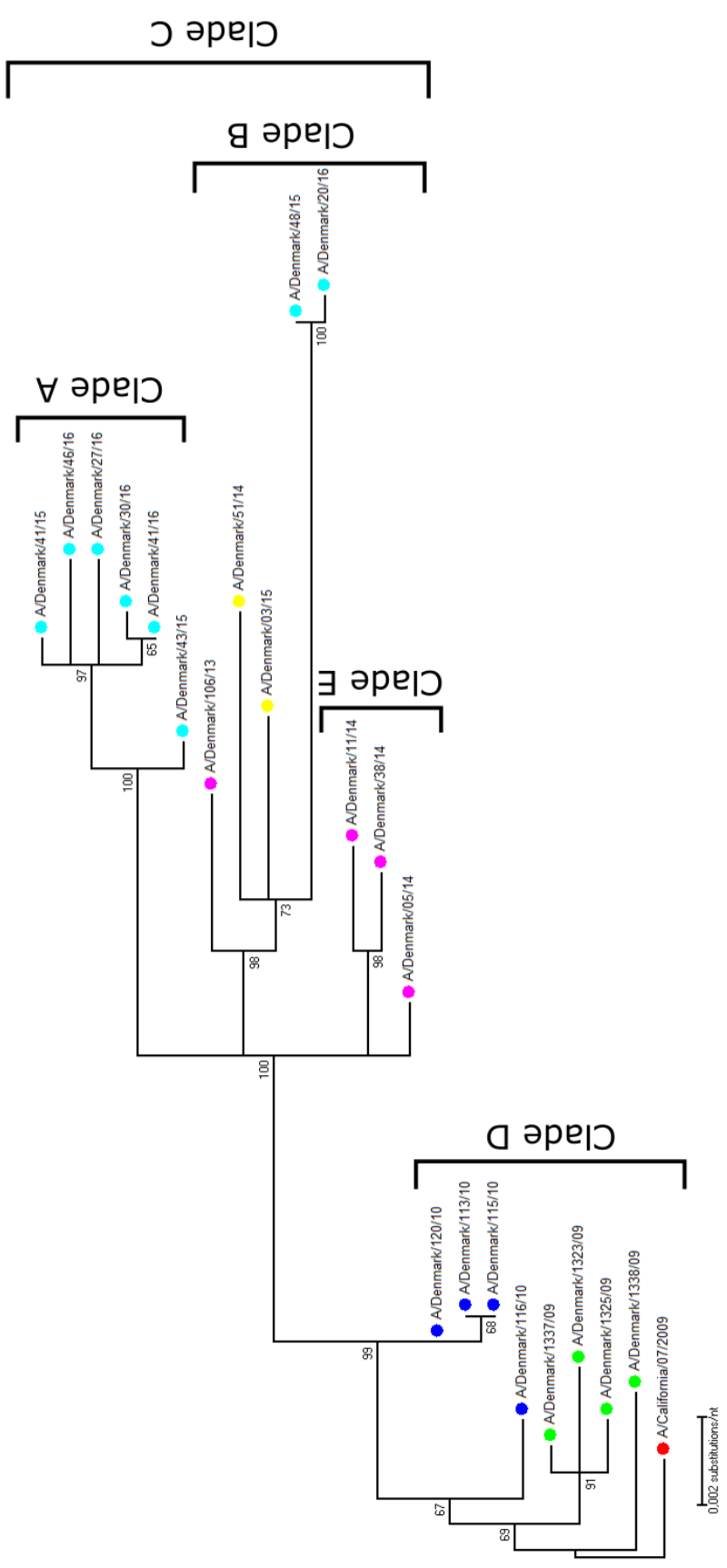

Figure 3.1: Phylogenetic tree of sequenced IAV samples using the Maximum Likelihood method based on the model of Tamura-Nei [72]. Algorithms BioNJ and Neighbor-Join were applied to a matrix of paired sequence distances, which were estimated utilizing the Maximum Composite Likelihood approach, to obtain the preliminary phylogenetic trees. From these, the superior tree, with a log likelihood of -3243.6986, was selected. The analysis of the 23 IAV H1 sequences was generated in MEGA7 [73], with included codon positions being 1st+2nd+3rd+Non-coding. The branch lengths indicate number of substitutions per nt. Each branch displays the percentage of trees with affiliated taxa that have clustered together. The final dataset includes 1685 positions with gaps or missing data being omitted. The designated root of the tree is A/California/07/09. The taxa are coloured according to season, i.e. season 2009-2010 (green), 2010-2011 (blue), 2013-2014 (magenta), 2014-2015 (yellow) and 2015-2016 (cyan). The root is coloured red.

same season, and adequate separation between the seasons, the exceptions being A/Denmark/106/13, A/Denmark/48/15 and A/Denmark/20/16. Looking at the clade C, which contains seasons 2013-2014, 2014-2015 and 2015-2016, it was evident that all included taxa are distanced relatively far, evolutionary, from the root when compared to taxa from the two previous seasons. It is, however, worth noting that there was also a significant time-wise separation (two seasons) of clades C and D, which will inevitably have an effect on the evolutionary distance.

The IAV samples from season 2009-2010 were all in close evolutionary proximity of the root sequence. The 2010-2011 season (blue) samples were also clustered together, although A/Denmark/116/10 was a closer relative to the IAV H1 samples from the previous season. The 2013-2014 season (magenta) and subsequent seasons, i.e. 2014-2015 (yellow) and 2015-2016 (cyan), were all distanced relatively far from the root. Despite the general clustering for each season, A/Denmark/48/15 and A/Denmark/20/16 were more closely related to the 2014-2015 season. Additionally, A/Denmark/106/13 was more related to the 2014-2015 season compared to the

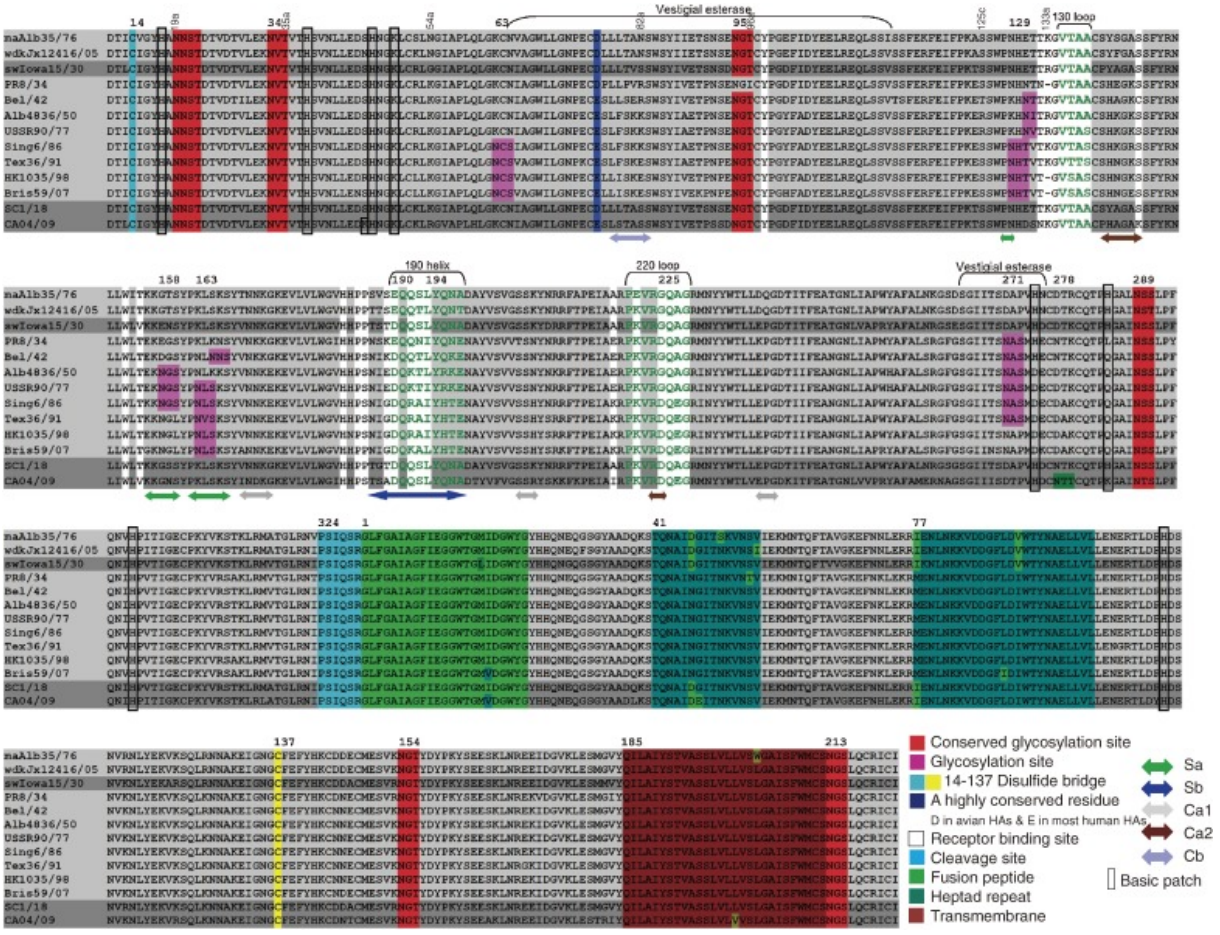

Figure 3.2: Alignment of H1N1 IAV sequences and positions of various sites important to HA function. Of specific interest are the antigenic sites, illustrated with coloured, double-headed arrows, the RBS, coloured white, and the position of the vestigial esterase. [21]

2013-2014 season. Of special interest is the clade A, consisting solely of IAV samples 2015-2016 which were isolated from the other taxa. The support of the nodes, i.e. the percentile number located at each node, was satisfactory, although at some nodes it dropped down between 65-70%. This indicates that the tree representing these IAV H1 samples could be considered satisfactory.

Comparison of the phylogenetic tree to the sequence analysis of antigenic sites in table 3.1 underlined the consistency between substitution and clade separation. For instance, the introduction of Lys180→Gln from the season 2013-2014 and onwards fitted decently with the separation between clades C and D.

3.2 Antigenic characterization

HA- and HAI-assays were used to characterise changes in antigenic nature of the included IAV samples by comparing them to the immunogenic performance of the reference virus, A/California/07/09. Both assays used guinea pig blood harvested in-house, and all HAI-assays used the same ferret antiserum raised to A/California/07/09. All IAV samples were tested across five days, meaning that five different controls were measured. The results of both assays can be seen in table 3.2.

The individual IAV samples and controls are colour-coded according the date of the experiment took place, due to the significant variation of control titers between experiments. Colours and dates are affiliated as: Green-yellow - 08/11/2016, yellow - 14/11/2016, green - 15/11/2016, blue - 10/03/2017 and orange - 14/03/2017. Each IAV sample had both titers measured in duplicates, and back titration was also performed for each sample in the HAI-assays. Dilutions were determined based on the HA-titer, to get the required 4 HA-units. Back titration, a measure of correct dilution, is given as + or (+), i.e. a + for each fully agglutinated well and a (+) in cases with low levels of agglutination. Only measurements with HAI-titers showcasing a back titration agglutination level of 3-4 wells were accepted as credible, with (+) counting as either 0 or 1 well. Fold-change factors were calculated by dividing the HAI-titer of the appropriate control with the HAI-titer of an IAV sample. An IAV sample with a fold-change factor of 8 or more can be considered antigenically different, and three such samples are found in the table, specifically A/Denmark/05/14, A/Denmark/30/16 and A/Denmark/46/16. All three had relatively low HAI-titers between 160-320, while their respective reference viruses had HAI-titers of 1280 and 2560. Additionally, three samples have fold-change factors of 4, namely A/Denmark/1323/09, A/Denmark/113/10 and A/Denmark/51/14. The remainder all had a fold-change fa-

HA- and HAI-titers

IAV sample	HA-titer	HAI-assay dilution IAV isolate/PBS	HAI-titer	Back titration	Fold change
A/Denmark/1323/09	16	1/1.5	640	+++	4
A/Denmark/1325/09	32	1/4	640	+++(+)	2
A/Denmark/1337/09	16	1/1.5	320	+++	2
A/Denmark/1338/09	32	1/4	1280	+++	2
A/Denmark/1341/09	32	1/4	320	+++	2
A/Denmark/113/10	8	1/1	640	+++	4
A/Denmark/115/10	16-32	1/1	640	+++(+)	1
A/Denmark/116/10	8-16	1/1	160	+++	2
A/Denmark/120/10	32	1/4	640	+++(+)	1
A/Denmark/106/13	8-16	1/1	320	+++(+)	2
A/Denmark/05/14	32	1/4	160	+++(+)	8
A/Denmark/11/14	4	2/1	1280	+++	0.5
A/Denmark/13/14	16-32	1/3	160	++(+)	2
A/Denmark/38/14	16	1/1.5	640	++++	1
A/Denmark/50/14	32	1/4	1280	+++	2
A/Denmark/51/14	32	1/4	640	+++	4
A/Denmark/03/15	16	1/1.5	1280	+++	2
A/Denmark/41/15	64	1/8	640	++++	1
A/Denmark/43/15	8-16	1/1	160	++(+)	2
A/Denmark/48/15	64-128	1/8	320	+++(+)	2
A/Denmark/20/16	64-128	1/8	640	++(+)	1
A/Denmark/27/16	16-32	1/3	320	+++(+)	2
A/Denmark/30/16	32	1/4	320	+++	8
A/Denmark/41/16	32-64	1/6	2560	+++(+)	0.5
A/Denmark/46/16	64	1/8	160	+++(+)	8
IAV controls					
A/California/07/09	128-256	1/20	1280	+++(+)	N/A
A/California/07/09	128-256	1/16	320	+++(+)	N/A
A/California/07/09	512-1024	1/16	640	++++(+)	N/A
A/California/07/09	512	1/60	2560	+++(+)	N/A
A/California/07/09	2560	1/30	640	+++	N/A

Table 3.2: HA- and HAI-assay results of IAV samples tested against reference virus A/California/07/09. Results are colour-coded for titer comparison between individual IAV samples and the reference virus data pertaining to the specific date where those results were acquired: **Green-yellow** - 08/11/2016. **Yellow** - 14/11/2016. **Green** - 15/11/2016. **Blue** - 10/03/2017. **Orange** - 14/03/2017.

ctors between 0.5-2 and can be considered antigenically equivalent to the reference virus.

The results of the HAI-assays were varying with no obvious trends. There was a generally lower HAI-titer in the 2013-2014 and 2015-2016 seasons, although some of the highest HAI-titers were found there as well. There was no clear decrease in HAI-titers as the IAV samples distance themselves, time-wise, from the reference virus, but there were some IAV isolates with HAI-titers of 160 scattered across the seasons. It is worth noting that the back titration of A/California/07/09 from 15/11/2016 was on the threshold of acceptable, which indicated that results from that date might require further testing. Additionally, a considerable amount of IAV samples were assayed more than once, in cases where back titration was outside the

acceptable area.

Comparing the fold-change factors to the phylogenetic tree showed a few interesting relationships. For instance, A/Denmark/113/10 and A/Denmark/115/10 had taken the same evolutionary path and were equally distanced from the root virus, but their fold-change factors of 4 and 1, respectively, were contrasting. There were no apparent differences in their antigenic sites, so the antigenic diversity must be attributed to unknown factors. It could potentially be caused by the questionable dilution of the 15/11/2016 control (as evident from the back titration), but it is doubtful whether that would elevate the fold-change factor of A/Denmark/115/10 to match the fold-change factor of A/Denmark/113/10.

The two IAV samples from the 2015-2016 season with fold-change factors of 8 were both situated in clade A, but showed no other apparent relation. Another IAV sample with a fold-change factor of 8 is A/Denmark/05/14, which was on a branch of its own. It was closer to the root than other samples from the same season, but it carried a unique substitution at Ala→Asp in the Ca2 site. This substitution could be decisive for the antigenic properties of this strain, seeing as it was antigenically different from the reference virus.

The 2009-2010 season showcased consistent fold-change factors of 2, with the exception of A/Denmark/1323/09, which was 4. This corresponded to some extent with the fact that this IAV sample was the furthest removed from the root, of the samples from this particular season, although not by a great deal. It did however carry a substitution at Glu252→Lys, which was specific for this sample, and is therefore a possible explanation for the higher fold-change factor.

It is interesting to note how the last three season all diverged into three separate clades from a single common ancestor of the 2010-2011 season. Clade E did not seem to have become the basis for further evolution, which is what happened in the clade B. Clade A, evolved from the same ancestor, was clearly separated from the other clades and forms a good potential base for further evolution due to the increased probability of finding IAV samples with incremented fold-change factors in this clade.

Lastly, two IAV samples, A/Denmark/48/15 and A/Denmark/20/16, had substitution Val190→Ile in the RBS. A hydrophobic side chain was exchanged with another, and the difference in size between Val and Ile is not significant either. The substitution did not seem to have any impact on the antigenicity, as the fold-change factors of both did not exceed 2. However, considering the phylogenetic tree, it was clear that these two stood out, as they were both removed from other IAV samples of their respective season. Additionally, both taxa were the ones furthest from the root.

3.3 Glycosylation

NetNGlyc software was used to locate and evaluate glycosylation sites in the surveyed IAV H1 protein sequences. The potential score of each identified glycosylation site is displayed for each sequenced IAV H1 sample in figures 3.4 and 3.5. In the graphs, the potential score is on the Y-axis, while the X-axis is the position in the protein sequence. The red horizontal line represents the 0.5 threshold.

There were at least eight glycosylation sequons for each IAV sample, and at least six of these, at positions 28, 40, 104, 304, 498 and 557, obtained a potential score of 0.5 or more. Only in the case of A/Denmark/05/14 did position 293 get a potential above 0.5, whereas it was below 0.5 for the remainder of the IAV sequences. It is worth remarking that the potential scores at position 498 only just pass the threshold, but does so in every surveyed IAV sample. Reviewing the jury agreement for this particular sequon revealed that only 4/9 networks agreed across all surveyed IAV samples.

A clear coherence was observable between the NetNGlyc-located glycosylation sequons and the expected glycosylation sites of A(H1N1)pdm2009-derived IAVs discussed in section 1.1.4. The 104 sequon was the only conserved glycosylation site residing on the head domain of the HA trimer.

Of special interest is the appearance of a new glycosylation sequon at position 179, which occured in all sequenced IAV samples from the 2015-2016 influenza season, save A/Denmark/48/15 and A/Denmark/20/16. The potential scores for this glycosylation sequon were higher than 0.5 in all sequenced IAV samples harbouring said sequon.

By considering both the potential scores and the jury agreements (not shown), a complete overview of the NetNGlyc results, evaluated according to table 2.9, for all fully sequenced IAV H1 samples can be seen in table 3.3. As evident, none of the special cases mentioned above obtained a NetNGlyc result higher than + due to insufficient jury agreement. There were at least four glycosylation sequons at positions 28, 40, 304 and 557 with NetNGlyc results of ++ or above, while positions 104

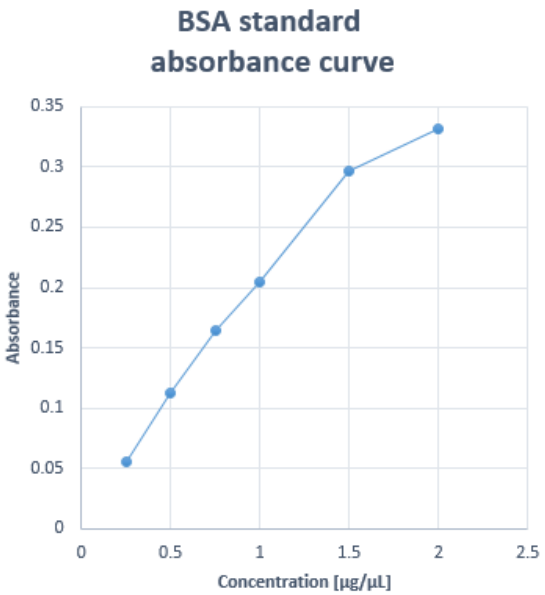

Figure 3.3: BSA standard absorbance curve generated from six BSA samples. The concentration is given in $\frac{\mu g}{\mu L}$. Absorbance was measured with NanoDrop ND-1000 Spectrophotometer.

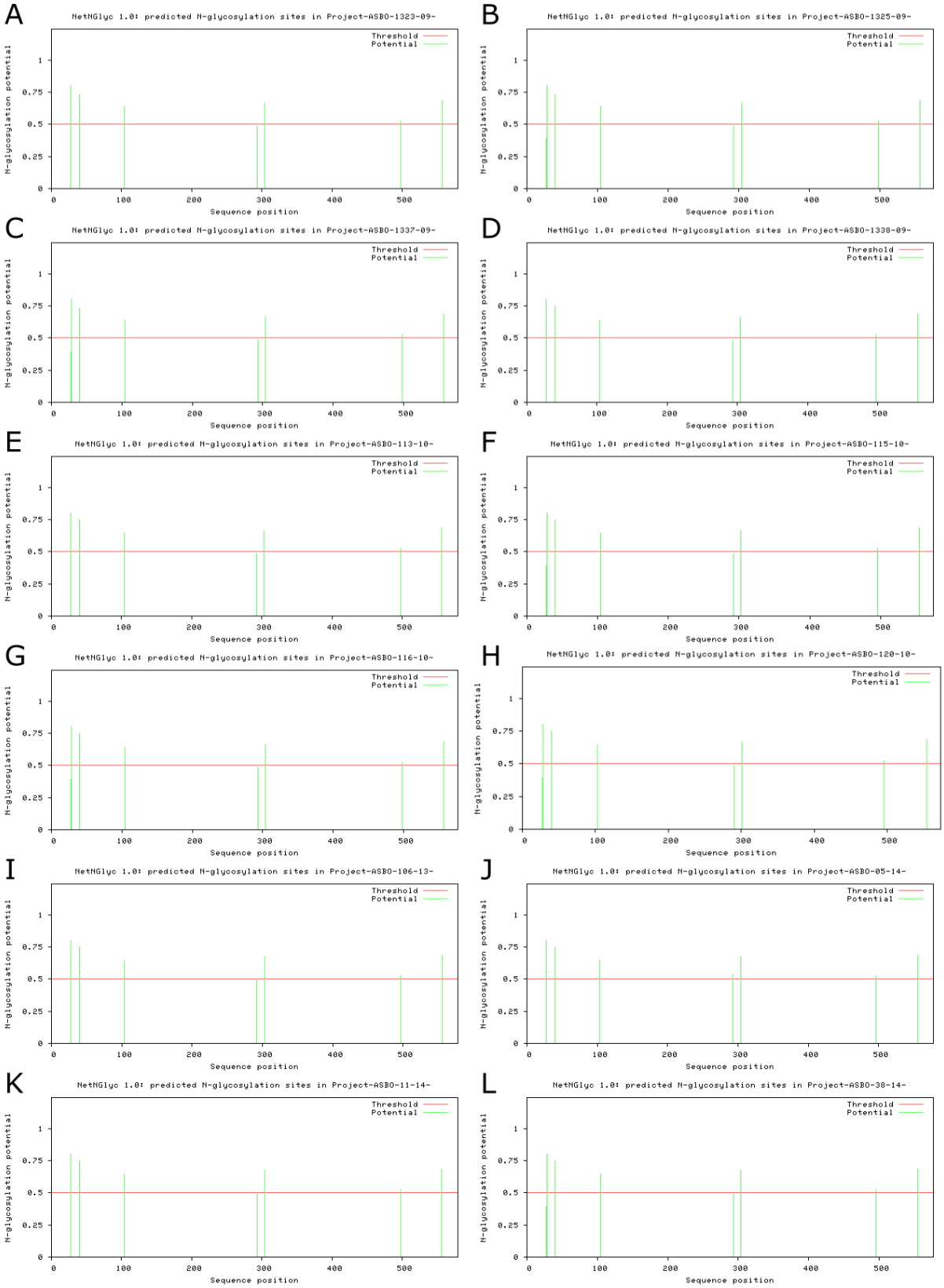

Figure 3.4: Average potential score of glycosylation sequons for the following 12 IAV samples: **A** - A/Denmark/1323/09. **B** - A/Denmark/1325/09. **C** - A/Denmark/1337/09. **D** - A/Denmark/1338/09. **E** - A/Denmark/113/10. **F** - A/Denmark/115/10. **G** - A/Denmark/116/10. **H** - A/Denmark/120/10. **I** - A/Denmark/106/13. **J** - A/Denmark/05/14. **K** - A/Denmark/11/14. **L** - A/Denmark/38/14. The Y-axis shows the potential score, while sequence position is on the X-axis. The 0.5 threshold is shown with a red horizontal line.

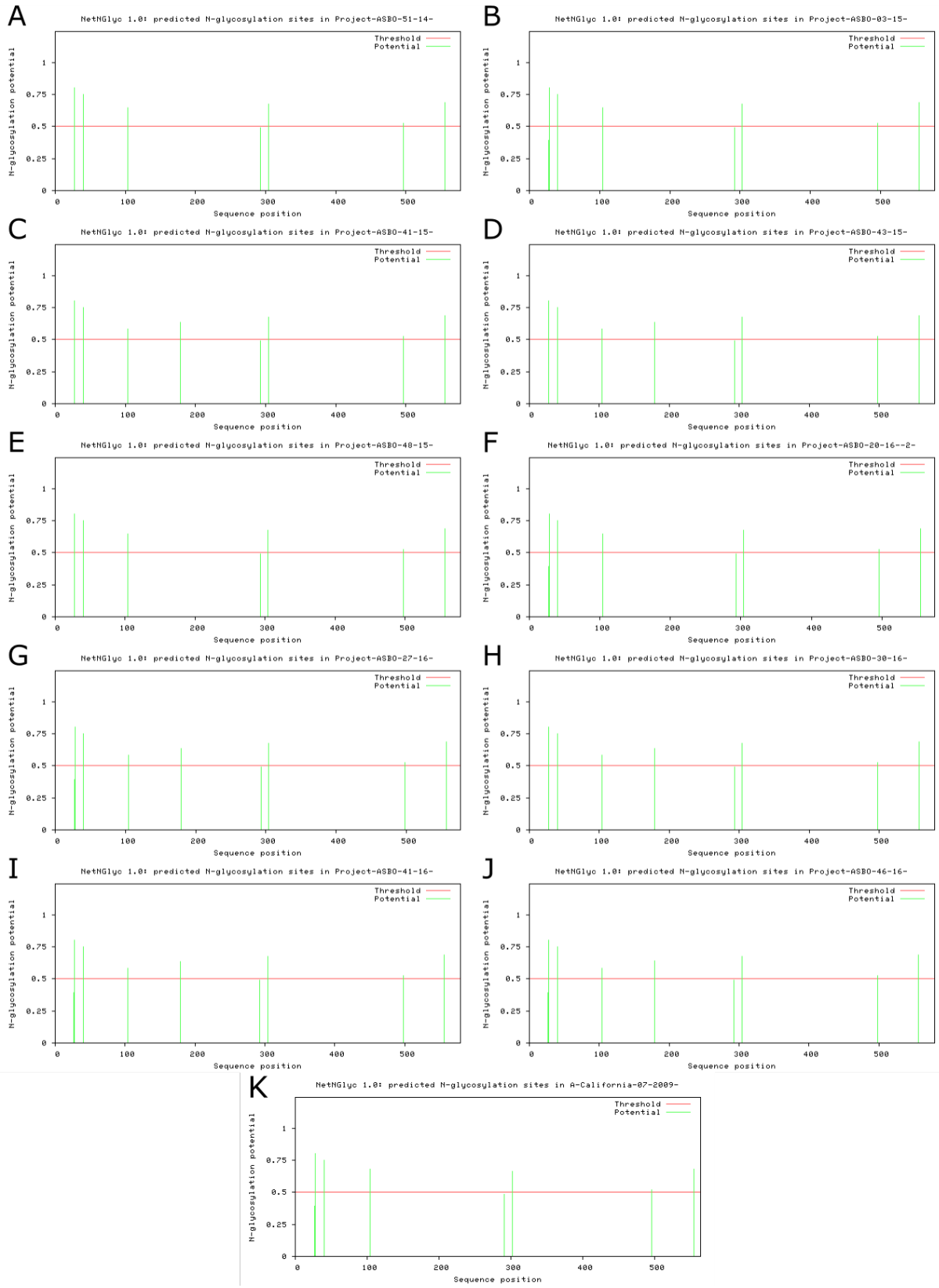

Figure 3.5: Average potential score of glycosylation sequons for the following 11 IAV samples: **A** - A/Denmark/51/14. **B** - A/Denmark/03/15. **C** - A/Denmark/41/15. **D** - A/Denmark/43/15. **E** - A/Denmark/48/15. **F** - A/Denmark/20/16. **G** - A/Denmark/27/16. **H** - A/Denmark/30/16. **I** - A/Denmark/41/16. **J** - A/Denmark/46/16. **K** - A/California/07/09. The Y-axis shows the potential score, while sequence position is on the X-axis. The 0.5 potential score threshold is shown with a red horizontal line.

NetNGlyc 1.0 results									
Glycosylation position	27	28	40	104	179	293	304	498	557
IAV sample									
A/Denmark/1323/09	-	+++	++	+	N/G	-	++	+	++
A/Denmark/1325/09	-	+++	++	+	N/G	-	++	+	++
A/Denmark/1337/09	-	+++	++	+	N/G	-	++	+	++
A/Denmark/1338/09	-	+++	++	+	N/G	-	++	+	++
A/Denmark/113/10	-	+++	++	+	N/G	-	++	+	++
A/Denmark/115/10	-	+++	++	+	N/G	-	++	+	++
A/Denmark/116/10	-	+++	++	+	N/G	-	++	+	++
A/Denmark/120/10	-	+++	++	+	N/G	-	++	+	++
A/Denmark/106/13	-	+++	++	+	N/G	-	++	+	++
A/Denmark/05/14	-	+++	++	+	N/G	+	++	+	++
A/Denmark/11/14	-	+++	++	+	N/G	-	++	+	++
A/Denmark/38/14	-	+++	++	+	N/G	-	++	+	++
A/Denmark/51/14	-	+++	++	+	N/G	-	++	+	++
A/Denmark/03/15	-	+++	++	+	N/G	-	++	+	++
A/Denmark/41/15	-	+++	++	+	+	-	++	+	++
A/Denmark/43/15	-	+++	++	+	+	-	++	+	++
A/Denmark/48/15	-	+++	++	+	N/G	-	++	+	++
A/Denmark/20/16	-	+++	++	+	N/G	-	++	+	++
A/Denmark/27/16	-	+++	++	+	+	-	++	+	++
A/Denmark/30/16	-	+++	++	+	+	-	++	+	++
A/Denmark/41/16	-	+++	++	+	+	-	++	+	++
A/Denmark/46/16	-	+++	++	+	+	-	++	+	++
A/California/07/09	-	+++	++	+	N/G	-	++	+	++

Table 3.3: The NetNGlyc 1.0 evaluation results. Determining parameters are potential score and jury agreement. No glycosylation sequon appears as N/G.

and 498 obtained +. The glycosylation sequon at position 179 for IAV samples from the 2015-2016 season all had NetNGlyc results of + with agreement between 8/9 networks.

Comparing NetNGlyc results to the fold-change factors imparted two points of significance. Although all sequences harbouring the 179 glycosylation sequon showcased a favourable potential score, the corresponding fold-change factors only correlated in 2 out of 6 cases, namely for A/Denmark/30/16 and A/Denmark/46/16. Furthermore, A/Denmark/05/14, with a fold-change factor of 8, was the only IAV H1 sample with a potential score above 0.5 at the 293 glycosylation sequon.

Glycosylation was further assayed by performing a western blot on a select group of IAV sample lysates listed in the table 3.4. Regular IAV isolates were used initial, but the lack of visible bands prompted the switch to IAV lysates in order to increase

IAV lysate	Absorbance (562 nm)	Concentration $[\frac{\mu g}{\mu L}]$
A/Denmark/1323/09	0.072	0.322
A/Denmark/1337/09	0.049	0.214
A/Denmark/113/10	0.077	0.343
A/Denmark/116/10	0.072	0.318
A/Denmark/120/10	0.064	0.288
A/Denmark/05/14	0.066	0.293
A/Denmark/38/14	0.051	0.222
A/Denmark/51/14	0.052	0.227
A/Denmark/48/15	0.094	0.441
A/Denmark/27/16	0.062	0.277
A/Denmark/30/16	0.075	0.331
A/Denmark/46/16	0.081	0.357
A/California/07/09	0.078	0.347
MDCK cell lysate	0.079	0.348

Table 3.4: The absorbance at 562 nm measured on the NanoDrop ND-1000 for the selected IAV lysates, and the concentration given in $\frac{\mu g}{\mu L}$ calculated by comparing individual absorbance values to the BSA standard absorbance curve seen in figure 3.3.

the concentration of unbound HA. However, visual protein bands were still absent and as a result, the concentrations of a select number of IAV lysates were determined by measuring absorbance at 562 nm and comparing them to the BSA standard curve in figure 3.3. The resulting concentrations can be seen in table 3.4.

The results of the final western blot can be seen in figure 3.6, displaying a blotted nitrocellulose membrane. As evident there were several bands, the most notable ones at around 70 kDa in lanes 2-10. According to another study the expected size of WT H1 is 80 kDa [54], while the addition of a single glycan should see a slight, albeit noticeable increment in size. Furthermore, the similarly sized band visible in the lane representing the MDCK cell lysate control indicated that the 70 kDa bands were derived from the MDCK cell lysate. The unedited image of the blot only displayed the 70 kDa bands, as well as two thicker bands specific to A/California/07/09 and A/Denmark/1337/09 located at 50 kDa. However, image editing revealed further bands at increased molecular weights.

IAV isolate	CT-value	Dilution
A/Denmark/1337/09	16.15	1:2
A/Denmark/116/10	16.18	1:2
A/Denmark/05/14	16.93	1:0
A/Denmark/38/14	17.30	1:0
A/Denmark/27/16	17.94	1:0
A/Denmark/46/16	15.23	1:4
A/California/07/09	14.69	1:9

Table 3.5: Dilution of IAV isolates prior to addition of formalin, based on CT-values.

Most are unspecific, such as the bands at 150 kDa in lanes 4-9, but a few IAV samples, such as A/Denmark/30/16 and A/Denmark/30/16, were represented by bands around 85-90 kDa. Regrettably there were no similar visible bands representing the other IAV samples, and comparison was therefore not possible.

3.4 SPR

Investigation of the relative binding affinity between HA in the IAV samples and the Abs in the ferret antiserum was done with SPR measurements. To ensure similar concentrations of virus in each analyte solution, the isolates were diluted according to their CT-values, which can be seen in table 3.5.

Initially, measurements of the control analyte, A/California/07/09, were made at varying dilutions to validate the consistency and versatility of the method with regard to this specific type of analyte. The results can be seen in figure 3.7, which showcases the binding of A/California/07/09 diluted in PBS 1:0

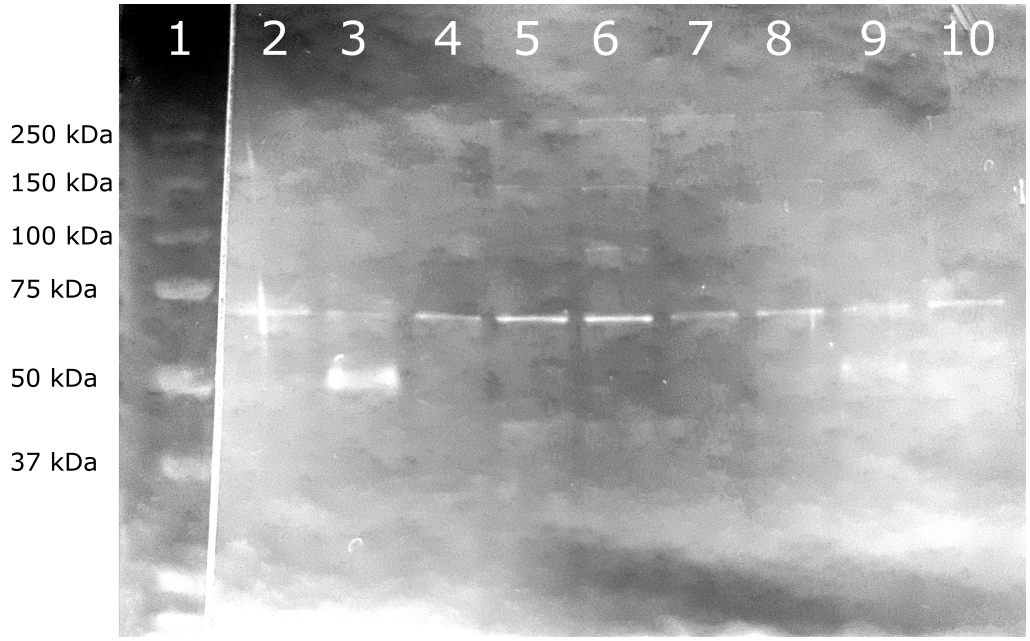

Figure 3.6: Western blot of selected IAV lysates. Each lane represents the following contents: **1)** 7 μL Precision Plus Protein™ All Blue Standard. **2)** 72 μL MDCK cell lysate + 18 μL Sample buffer. **3)** 72 μL A/California/07/09 cell lysate + 18 μL Sample buffer. **4)** 72 μL A/Denmark/27/16 cell lysate + 18 μL Sample buffer. **5)** 72 μL A/Denmark/46/16 cell lysate + 18 μL Sample buffer. **6)** 72 μL A/Denmark/30/16 cell lysate + 18 μL Sample buffer. **7)** 72 μL A/Denmark/51/14 cell lysate + 18 μL Sample buffer. **8)** 72 μL A/Denmark/05/14 cell lysate + 18 μL Sample buffer. **9)** 72 μL A/Denmark/1337/09 cell lysate + 18 μL Sample buffer. **10)** 72 μL A/Denmark/1323/09 cell lysate + 18 μL Sample buffer. The blot was run for 2 hours at 125 V and 200 mA. Afterwards, it was incubated with anti-swine H1N1 hemagglutinin Ab diluted 1:600 in block buffer for 2 hours, followed by incubation with anti-rabbit Igγ(whole molecule)-peroxidase Ab diluted 1:5000 in blocking buffer. Lastly, the blot was exposed to the staining solution for 10 min and subsequently washed with MilliQ-water.

(black), 1:1 (red), 1:4 (blue) and 1:9 (green). These binding curves represent the measured signal of the working partition with the signal of the reference partition subtracted. A measurement of a 1:49 dilution of A/California/07/09 was omitted and instead subtracted as a blank measurement. The best fit, outlined at every binding curve, was obtained using the OneToTwo interaction model with a local fit, because it had a higher level of flexibility in terms of binding types. Antigen-Ab binding interactions with different binding constants will happen due to the use of polyclonal Abs, which is likely what can be observed in the sensorgram. Initially there was a steep increase in the signal, which was followed by a significantly halted increase, indicating that two or more different types of binding occurred. However, using the local fit meant that maximum signal values (B_{max}) must be chosen locally for each binding event, which with the OneToTwo models gives two B_{max} values. The fraction of these two values should be close to equal for each measured concentration, which was not the case.

The lack of knowledge regarding the concentrations of Ab and antigen made it impractical to calculate association and dissociation constants, but still allowed for a relative comparison between the concentration. As expected, there was a noticeable decrease in the signal as the concentration of analyte was decreased. Additionally, the signal did not return to the baseline after dissociation, which was evidence of specific binding. Based on these results it should be possible to do the same experiments on the remaining IAV isolates, but as evident from figure 3.8, this was not

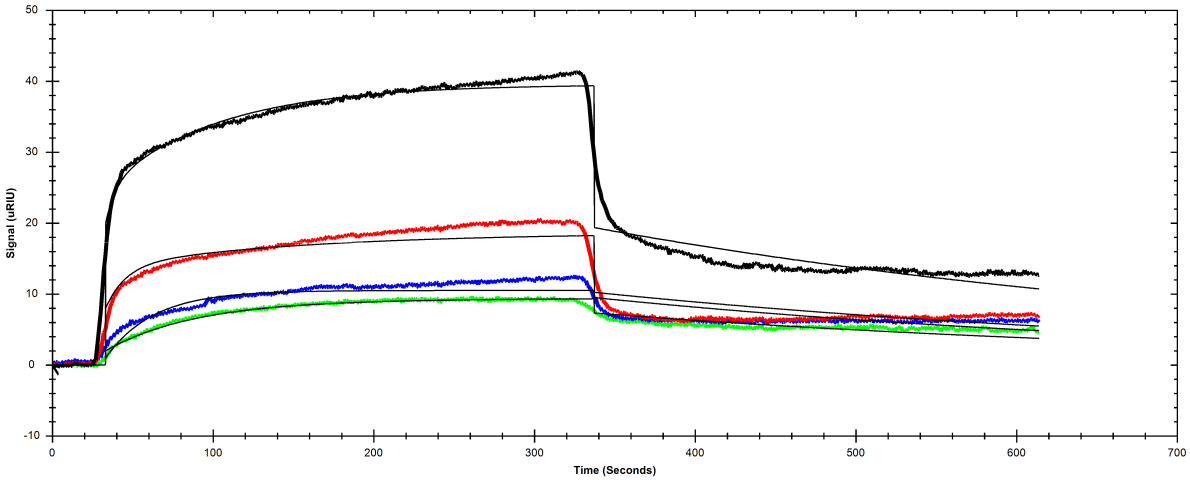

Figure 3.7: Sensorgram containing the binding curves of A/California/07/09 diluted in PBS. **Black** - 1:0. **Red** - 1:1. **Blue** - 1:4. **Green** - 1:9. Data from the reference partition is subtracted, and a measurement of a 1:49 dilution of A/California/07/09 is also subtracted as a blank measurement. The best fit, outlined at every binding curve, was obtained using the OneToTwo interaction model with a local fit. The Y-axis shows the signal in μRIU, while the X-axis represents time in s.

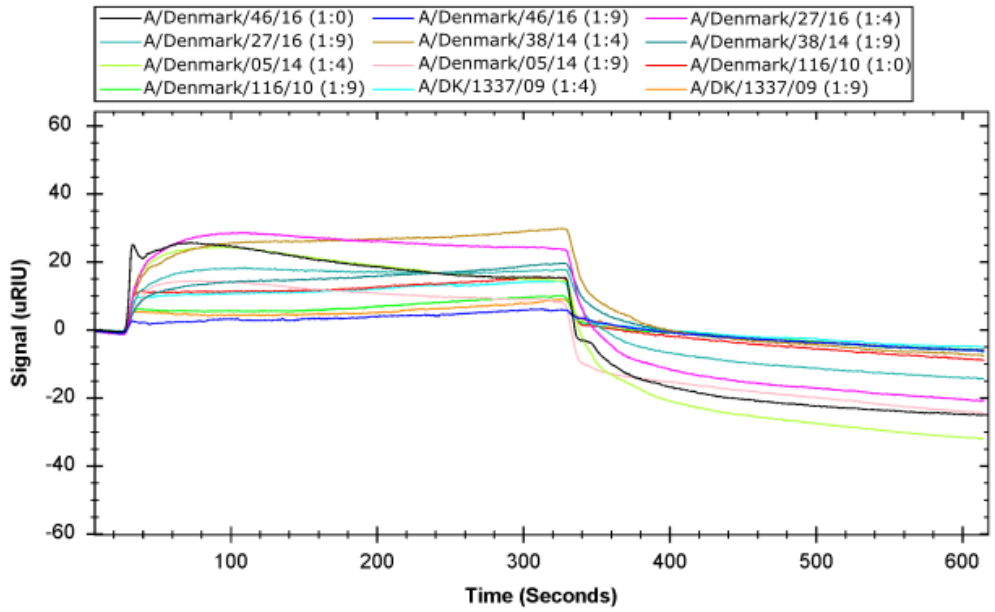

Figure 3.8: Sensorgram of binding curves for the IAV isolates diluted in PBS as listed in the legend. Data from the reference partition is subtracted. The Y-axis shows the signal in μRIU, while the X-axis represents time in s.

the case. Here, all the selected IAV isolates were tested with SPR at two different concentrations, all of which is listed in the figure legend. Although the signals reached expected values, around 20 refractive index units (μRIU), there were no signs of specific binding in any of the measurements. All signals returned to the baseline, sometimes

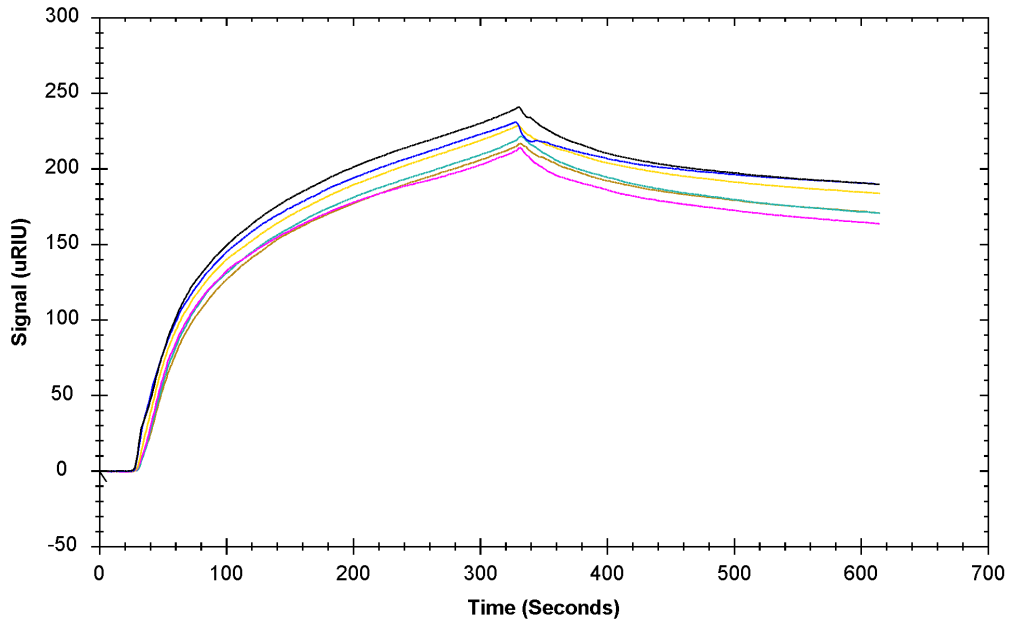

Figure 3.9: Arbitrarily selected binding curves showing association and dissociation of the secondary ligand, i.e. ferret antiserum raised to A/California/07/09. Data from the reference partition is subtracted. The Y-axis shows the signal in μRIU, while the X-axis represents time in s.

going even lower, but this was likely caused by the drift experienced in the SPR running buffer during this measurement run. It was not possible to correct the effects of this drift in TraceDrawer, but it is possible that doing so would give a clearer idea of specific binding in these IAV samples.

Lastly, figure 3.9, a small collection of arbitrarily selected secondary ligand binding curves, is included to underline that immobilisation of the secondary ligand was successful, since relatively high signal values, between 200-250 μRIU, were recorded with low amount of dissociation, indicating ample levels of specific binding. Furthermore, the immobilisation was reproducible, seeing as nearly equal signal values were attainable with each measurement run.

4. Discussion

23 IAV samples were surveyed for added glycosylation outside the conserved sequons native to A(H1N1)pdm2009 at positions 28, 40, 104, 304, 498 and 557. According to NetNGlyc results, the six IAV H1 samples carrying the Ser179→Asn substitution, all of which are from the 2015-2016 season, acquire a new glycosylation site at position 179. Residue 179 is located inside the Sa site on top of the HA trimer, where antigenic masking by means of glycosylation is shown to have a decreasing effect on neutralization by polyclonal Abs [8]. Glycosylation here may also mask the Ca2 and Sb site of the neighbouring subunit [76]. However, this is not the case for all glycosylations located in the Sa site, since glycans emerging as a result of substitutions at residues 177 and 179 proved to have little effect on antigenicity [8, 76]. Particularly, 179 might block the RBS and have negative implications on the binding between the IAV and the host cell receptors [76].

The antigenic analysis using HAI-tests highlights two IAV samples of season 2015-2016, namely A/Denmark/30/16 and A/Denmark/46/16 with substitutions in the Sa site. Both carry two substitutions in succession, namely Ser179→Asn and Lys180→Gln, but are not the only IAV samples doing so, seeing as A/Denmark/41/15, A/Denmark/43/15, A/Denmark/27/16 and A/Denmark/41/16 all have the same substitutions. However, A/Denmark/46/16 has an additional, unique substitution at Gly187→Glu, which sets it apart from other IAVs of the 2015-2016 season. The IAV H1 samples carrying the Ser179→Asn and Lys180→Gln substitutions are exclusively found in clade A of the phylogenetic tree. The two samples from that season *not* carrying this substitution, i.e. A/Denmark/48/15 and A/Denmark/20/16, are found in clade B together with IAV H1 samples from the previous season, albeit with greater evolutionary distance from the root than their counterparts. Both A/Denmark/46/16 and A/Denmark/30/16 demonstrate a fold-change factor of 8, which qualifies them as antigenically different, but do not seem to be distinctly related from the other IAV H1 samples of their respective season. Both have relatively low HAI-titers between 160-320, while their respective reference viruses had HAI-titers of 1280 and 2560. The fold-change factors of the remaining IAV H1 sample from that season are between 0.5-2, thereby disqualifying them as being antigenically different.

Factoring in the NetNGlyc potential scores of these 8 IAV H1 samples, which are close the the 0.5 threshold, it is possible that the existence of this glycosylation site does not guarantee glycosylation. Inconsistent glycosylation might be the reason why not all of them demonstrate antigenically altering fold-change factors. Even though added glycosylation was not determined categorically by immunoblotting, it is still probable that A/Denmark/46/16 and A/Denmark/30/16 are both glycosylated at position 179 based on coherency between sequence analysis, HAI-titers and fold-change factors.

It is possible that immunoblotting was unsuccessful due to low concentrations of HA in the IAV lysates. From the measurements of concentration, it is clear that the concentrations were low, since they barely reach the concentration of the least concentrated BSA standard sample ($0.25\ \frac{\mu g}{\mu L}$). On top of that, the IAV lysates contain all viral proteins, as well as some MDCK related proteins, which means that HA only make up a small fraction of the measured proteins. It therefore might not be possible to obtain visible bands representing HA from any of the lysates.

Although the introduction of Asn179 is novel to the 2015-2016 season of the surveyed IAV H1 samples, it is not the case for H1N1 in general, since a few papers already investigated this specific substitution and the effects it has on antigenic properties [6, 54, 77, 76]. When including literature pertaining to the same substitution in HA of H3N2 with Ser165→Asn, it becomes clear that this is a recurrent substitution. For instance, Asn179 was present in H1N1 from 1933-1949, but was lacking in seasonal strains from 1977 until the pandemic outbreak in 2009 [54]. After its reintroduction in 2009, the substitution of Ser179→Asn has, based on a dataset of >5000 A(H1N1)pdm2009 strains, seen a decline in appearance, but the sequence analysis of this thesis indicate it could be facing a resurgence. However, the amount of surveyed IAV samples is too small to make any conclusive statements. [54]

From surveying IAV H1 samples post the 1918 H1N1 pandemic, it becomes evident that the virus acquired new glycosylation sites, in some cases as many as four sites on the head domain alone. However, glycosylation sites at random positions were never sustained, and it never became hyperglycosylated, indicating that there is a delicate balance and purpose to the glycosylations. The recurrence of an IAV with low levels of glycosylation on the head domain, i.e. A(H1N1)pdm2009, has reset the glycosylation counter in seasonal IAVs, and it is likely that following influenza seasons will undergo a similar evolutionary pattern as was observed after the 1918 H1N1 pandemic. [6, 8] The occurrence of glycosylation sites at Asn179 in six out of eight IAV H1 samples from the 2015-2016 season might be the first indicator that this is the case. It seems paradoxical, conside-

ring that the successful pandemic strains carry few glycosylation sites, that it is beneficial for the virus to develop additional glycosylation sites. However, because glycosylation can help the virus avoid immune recognition, it is allowed to sustain its status as a seasonal virus, which is preferable when considering survival [57]. It corresponds well with the idea of survival of the flattest, which predicts that an evolutionary diverse progeny virus pool with reduced pathogenicity is favourable to an evolutionary similar, yet highly pathogenic progeny virus pool.

Due to the low level of glycosylation on the head domain of A(H1N1)pdm2009 (Asn104), its transition to a seasonal influenza virus has made it resistant to collectins such as SP-D and MBL. Moreover, the presence of a second glycan at position 179 did not increase sensitivity towards either collectin during an investigation of two IAV samples carrying said glycan [54]. The same group investigated the sensitivity of a Ser179→Asn substituted H1N1 IAV toward Ab-mediated neutralization by testing it against polyclonal Abs raised to an A(H1N1)pdm2009 equivalent (Auck/09), using A/California/07/09 as a reference virus. It was found that the Ser179→Asn substituted H1N1 IAV and the reference virus expressed similar sensitivity towards the Abs, further indicating that this glycosylation site is not associated with resistance to A(H1N1)pdm2009 raised Abs. Additionally, they tried to introduce Asn179 into their Auck/09 H1N1, but immunoblots confirmed that despite the introduction of the amino acid, the resulting HA molecules were not glycosy-

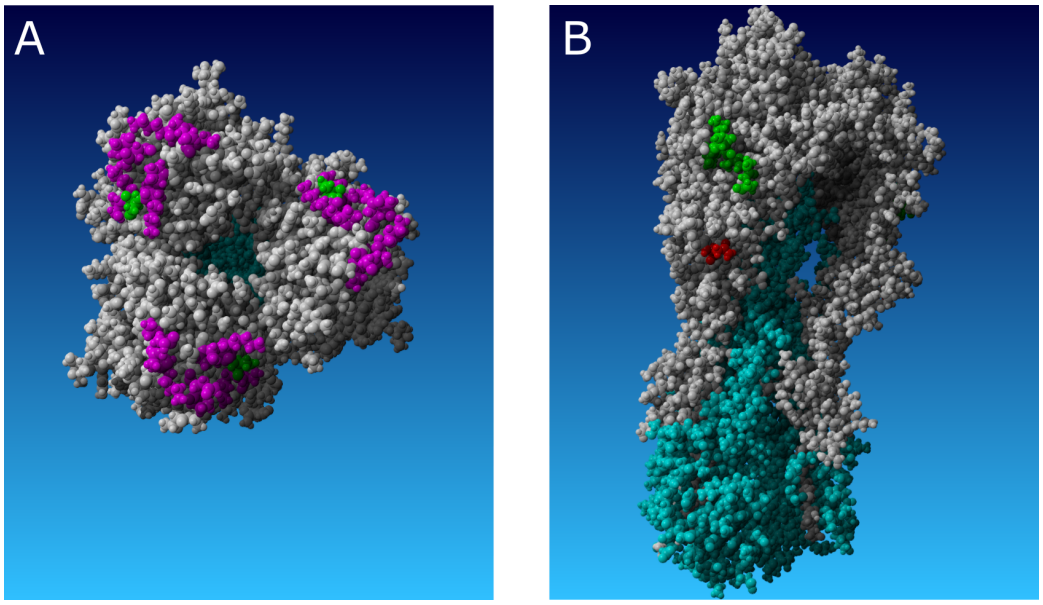

Figure 4.1: Placement of residues 179 (A) and 293 (B) with respect to their nearest antigenic sites on the HA trimer. **A** - Position of residue 179 (green) within the Sa site (magenta). **B** - Position of residue 293 (red) in proximity of the Cb site (green). White part consists of all three HA1 subunits, while the cyan part is comprised of the HA2 subunit. Images are generated with YASARA using 4GXX from the RCSB Protein Data Bank [33, 34].

lated. The lack of glycosylation at that site is attributed to structural differences between pandemic Auck/09 and the seasonal Asn179 glycosylated strains they acquired for the study, which suggests that either the structure of A(H1N1)pdm2009 is not optimal for successful glycosylation of Asn179, or that it does not alter the antigenic properties of virus significantly. [54] This inconsistent glycosylation at position 179 correlates with the observed fold-change factors of season 2015-2016, which indicate that only two out of six IAV samples carrying this substitution are in fact glycosylated. However, confirmation of glycosylation is still required to fully conclude whether the glycosylation is inconsistent at position 179.

Another group found that, through homology modelling predictions, Asn179 is not sterically hindered and that glycosylation here is likely to shield not only the Sa site, but also the Ca2 and Sb site of the neighbouring HA subunit [57]. Later studies revealed that glycosylation at Asn179 might disrupt the interaction of sialic residues on the surface of the host cell and the RBS of neighbouring HA subunit [76], which might explain why IAVs containing a glycan at this position are not showing larger increase in fold-change factors, since the positive effect of antigenic shielding might be negated by reduced binding to the host cell receptors.

Looking at figure 4.1 **A**, which is an image of 4GXX generated in YASARA, it is evident that residue 179 is within the confines of the Sa site. It is placed at the edge of the site towards the neighbouring HA subunit, which gives further credibility to the argument that a glycan here masks the Ca2 and Sb sites as well. However, the inherent location and the fact that the glycan leans towards those sites might mean that the Sa site is not benefitting substantially from this glycosylation. This could be an issue, since polyclonal Abs raised against A(H1N1)pdm2009 are primarily aimed at the Sa site [8]. This might be the reason why this glycosylation sequon was shifted to position 177 in the seasonal IAVs following the 1918 pandemic [54].

Verifying the presence of additional glycosylation is therefore necessary in order to explicitly conclude whether it is the emergence of further glycosylation that affect the antigenic properties of A/Denmark/46/16 and A/Denmark/30/16. The immunoblotting of this thesis did not show any meaningful results, largely due to low concentrations of HA in the IAV lysates, although the inexperience of the investigators concerning this specific method also played its part. However, the marked difference in fold-change factors within the 2015-2016 season serves as a strong indicator that glycans are present in these two IAV samples.

Another potential glycosylation was observable in IAV sample A/Denmark/05/14, at position 293. As evident in figure 4.1 **B**, this residue (red) is situated at the bor-

der between the head and stem domain, with the Cb site being the closest antigenic site. The placement of this glycosylation site seems ineffective, as it is not in or close to any antigenic sites, and therefore should not play a pivotal role in Ab recognition. Additionally, the five residues surrounding residue 293 on either side are conserved, which makes it difficult to conclude why this particular IAV sample has a potential score of >0.5. Immunoblotting is therefore needed to determine whether the site is glycosylated. This glycosylation site has existed since the introduction of A(H1N1)pdm2009 [57], but it does no appear to be glycosylated.

Other antigenic factors must therefore play a role, seeing as IAV sample A/Denmark/05/14 has a fold-change factor of 8, the same as A/Denmark/30/16 and A/Denmark/46/16. It might be explained by the unique substitution Ala156→Asp in the Ca2 site, which, when compared to the fold-change factor, might induce a change in antigenic properties. Comparing the substitution Ala156→Asp of A/Denmark/05/14 to other sequences from the same season, which are all identical to A/Denmark/05/14 in other antigenic sites, one finds that there is a stark contrast between the fold-change factors within this season. This might suggest that the said substitution has a profound effect on the antigenic properties of the virus. The substitution does, however, not reappear in subsequent or previous seasons, and it might therefore be detriment to the fitness and/or stability of the virus itself. From figure 4.2 it is clear that the residue (blue) is found jutting out on the surface of the protein, which is uncharacteristic for hydrophobic residues such as Ala. Changing into Asp could make the surface of the Ca2 site more hydrophilic given that this amino acid is charged. However, Asp is also slightly bulkier than Ala, which might affect the binding of Abs targeting the Ca2 site. Figuring out the overall effect of this substitution would require additional investigation into the alterations it causes in the protein structure.

Two residues important to the RBS, i.e. Val190 and Lys225, were also investigated

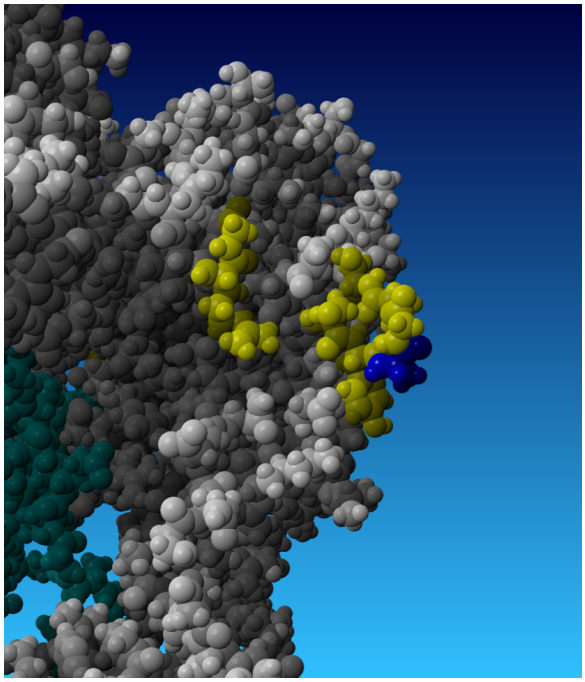

Figure 4.2: Locations of residue 156 (blue) in the Ca2 site (yellow). The white part consists of all three HA1 subunits, while the cyan part is comprised of the HA2 subunit. Images are generated with YASARA using 4GXX from the RCSB Protein Data Bank [33, 34].

in regard to substitutions. A single IAV sample, A/Denmark/51/14, has a substitution of interest, namely Lys225→Arg. This substitution does not promote a great difference in terms of side chain structure and properties, as both are positively charged side chains of similar size. However, this substitution is accompanied by an increase in the fold-change factor, although it is unclear whether it is caused by a substitution in the RBS or a substitution in the Ca1 site at Glu252→Asp. Both factors are unique to this IAV sample, but the latter seems more likely, seeing as A/Denmark/1323/09 also has a similar substitution with Glu252→Lys and a fold-change factor matching that of A/Denmark/51/14.

Lastly, measuring binding affinity between ferret antiserum raised to A/California/07/09 and a select number of IAV isolates with SPR was largely unsuccessful. Only the control showed signs of specific binding, while the signal of the remaining IAV samples all returned or surpassed the baseline. The majority of the measured analytes failed to saturate the secondary ligand on the surface, seeing as equilibrium was not reached during the injection of analyte. Secondly, the binding curves of the non-control samples could not be fitted to any predetermined models, due to the drift in the supplied running buffer.

There are several possible explanations as to why this occurred. The analyte could have a consequential effect, which could be attributed to one of several reasons. In order to perform measurements on Class II pathogens in a Class I laboratory, it was necessary to inactivate the IAVs, which was done using formalin. The active component of formalin, formaldehyde, inactivates viruses by cross-linking primary amino groups with neighbouring nitrogen atoms present in the protein. Although the HA molecules should maintain their overall structure, and thus epitopes, it is possible that slight alterations in their topology could affect binding between ligand and analyte. The sensorchip surface and the ligand molecules might also be affected upon exposure to the formalin in the analyte solution. Cross-linking within an Ab could be more damaging to the measurements, as they might lose function completely. The concentration and activity of formalin at this stage is arguably low, but it cannot be considered negligible when considering the sensitivity of the SPR instrument.

The other components present in the analyte solutions also visibly affected the results, which was evident from a brief, yet steep, increment in each signal as the analyte solution was injected onto the sensorchip surface. This problem was solved by subtracting the data from the reference partition. However, the PBS used to dilute the IAV isolates was the same one used to form the baseline of the SPR measurements, which means that this increment must be attributed to a difference

in solvents, likely caused by the unknown components in the IAV isolates. To counteract this effect it is necessary to either purify the analyte or incorporate the virus growth medium into the SPR running buffer. However, the latter solution is difficult to implement, as the IAV samples are initially diluted to obtain equivalent CT-values, meaning that the amount of MEM Eagle will vary from sample to sample. An attainable approach would be to increase the concentration of viral particles in order to allow for higher dilution during SPR measurements. This could be done by using smaller volumes of growth medium during virus growth.

Primarily, it is worth considering that in using intact IAVs as the analyte for SPR measurements, the amount of available HA molecules is limited substantially upon initial binding to the ligand. An intact IAV virion will only be able to interact with a few of the immobilized ferret Abs at a time, meaning that an excessive amount of HA molecules are not bound. Furthermore, the sheer size of the analyte, compared to the ligand, could have certain implications. For instance, bound virions will sterically block the approach of other virions to free ligand molecules, thereby decreasing the overall signal. This might be the reason why saturation was rarely achieved, since not all ligand molecules could be occupied. Morever, the individual viral particles may not necessarily weigh the same, and the effect each binding will have on the dielectric constant would vary as a result. Possible solutions would be to either immobilize the IAVs instead of the ferret Abs, or to use pure H1N1 HAs as the analyte. The former would increase the amount of reactive sites significantly, i.e. places where the analyte can bind, but at the same time it would require a lot of sensorchips, since they would need replacement more frequently. Additionally, the signal of analyte-ligand binding might be larger than the binding signal between primary and secondary ligands, as a single, immobilized virion will be able to bind multiple HA molecules. Isolating and purifying the analyte seems to be the best option, as it would increase the amount of reactions happening, since more HA molecules could bind to the immobilized Abs.

It is also possible that using polyclonal Abs gave an inconsistent signal, because the binding affinities will differ for the different Abs within the polyclonal group. Although the bound masses should be similar, the rate of binding can vary, which is what can lead to an inconsistent signal. Using the OneToTwo fitting model with local fit resulted in a decent fit, but it might be better to use monoclonal Abs to get a more stable signal that is easier to fit.

The amount of papers concerning IAVs and SPR is scarce, which means that this is still a largely uncharted area of research. Most papers in this area of research concern themselves with measuring binding affinity between HA and a variety

of glycan-based ligands, with the majority using certain forms of purified HA molecules [78, 79, 80, 81]. It is possible that SPR can be useful when measuring the interactions of IAV antigens and their corresponding Abs, but great efforts are required to find the correct approach which will give worthwhile data while maintaining the integrity and purpose of the experiment. Although the fitting of the control samples seemed adequate, difficulties were still encountered regarding local fitting parameters. It is possible that binding events change with concentration, but knowing the concentration of analyte and ligand beforehand would make it easier to control these parameters, which in turn might fix the disproportionate fractions of the B_{max} values. However, the binding of secondary analyte from the ferret antiserum showed promising results, and is a good basis for further investigation.

5. Conclusion

Six out of 23 sequenced IAV H1 samples were projected to be glycosylated at residue 179 based on NetNGlyc results. All six lie within the 2015-2016 influenza season, and acquire a new glycosylation site due to the substitution Ser179→Asn. Two of these six IAV samples, A/Denmark/30/16 and A/Denmark/46/16, displayed fold-change factors of 8, which is sufficient to declare them antigenically different from the reference virus, A/California/07/09. When correlating those two results it seems probable that those two IAV samples are glycosylated, and that their newly acquired glycosylations might have decreased the effectiveness of the 2015-2016 IAV vaccine, based on their performance in the immunological assay. However, it is not possible to make any explicit conclusions in this regard, seeing as the presence of new glycans could not be categorically determined through immunoblotting.

Other studies have shown that glycosylation at residue 179 is not of significance to antigenicity, which is coherent with the fact that both IAV samples only *just* qualify as antigenically different. Similarly, not all sequenced IAV samples harbouring the Ser179→Asn substitution were glycosylated, based on fold-change factors, which also correlates with findings in other studies regarding the suitability of residue 179 as a glycosylation site.

A/Denmark/05/14 is the only IAV sample outside the 2015-2016 season with a fold-factor of 8. There are two potential explanations to this, the most likely being the introduction of Ala156→Asp in the Ca2 site seeing as it is unique to this IAV sample. By altering the structural properties of the Ca2 site, this substitution might be responsible for the increase in fold-change factor. The second explanation is the NetNGlyc determined glycosylation of a dormant glycosylation site at residue 293. However, residue 293 is not in close proximity of any antigenic sites, which makes it unlikely to alter any antigenic properties. Making any explicit conclusions about the presence of this glycan requires credible immunoblotting.

Lastly, using SPR techniques to assay antigenetic alterations in IAVs is a promising approach, but requires substantial investigation into the correct composition and combination of analyte and ligand in order to establish it as a viable method. Signals were detectable, but due to mismatches in size and/or type of analyte/ligand, and lacking information regarding concentrations of analyte/ligand, the results became impractical to evaluate.

6. Limitations and outlook

Experimental limitations

A few aspects in each experimental procedure were limited by varying factors, most of which are presented here.

A feature common for all experiments was the lack of statistically relevant results, in part due to the amount of available samples and the allotted time. With a larger investigative time frame it would be optimal to obtain a minimum of $N = 3$ repetitions for every condition/parameter influencing the results.

Immunological assays As evident from the results, both HA- and HAI-titers of the reference virus were varying by a relatively large margin between experimental dates, even though each sample originated from the same stock solution. It is easiest to observe in control sample from 15/11/2016 (green), which had a high HA-titer, but was diluted according to previous experiments (1/16). From the corresponding back titration it was evident that the concentration of the reference virus was too high in the HAI-assay, and that this could affect the HAI-titer by lowering it from 1280 to 640. This would impact fold-change factors of IAV samples tested on that date, elevating it to 4 in some instances. Regardless, none of the IAV samples would qualify as antigenically different if this was the case.

Varying RBC concentrations could play a role in the titer disparity, and also incubation time is an influential variable. Potentially, it could be down to inexperience of the investigator, but ultimately it would be ideal for the control titers to be more similar, which would give a better basis for comparison of antigenic properties. Performing the experiment on all IAV samples in a single day would remove doubts regarding RBC concentrations, but incubation times would still vary, as overseeing this many IAV samples would take considerable amounts of time.

Inspecting the results of the HA-assays also uncovered certain trends. The HA-titers appear to increase with each passing season, although outliers are present. Particularly, the first four seasons did not display HA-titers above 32, while the 2015-2016 season had at least four IAV sample with a HA-titer of 64 or above. The reason for increased HA-titers in IAV samples of the most recent season could be caused by a decline in the hemagglutinating potency of the IAVs as a result of longer, dormant periods for the older samples compared to samples from the most recent season. The effect of time on HA-titers is insufficiently investigated, but a study on

avian IAVs of the H4N2 subtype demonstrated retained HA-titer upto 42 months for samples stored at -20 °C [82]. However, these IAV samples were inactivated with formalin and concentrated in polyethylene glycol [82], which is vastly different from plain dilution in PBS as used in this thesis. Regardless, the virion is not required to be active in order to agglutinate RBCs, and one would therefore expect the agglutinating potency of the active and inactive IAV to be similar. Having an overview of the HA-titers initially measured for the surveyed IAV samples of this thesis would show whether the HA-titers have changed over time.

Not all of the preselected IAV samples could obtain an acceptable HA-titer, meaning that certain seasons had a small sample size, and comparison between seasons suffered as a result. For instance, the 2014-2015 season had three assayed samples, whereas the 2015-2016 season had eight. This could be attributed to the concentration of the IAV in the individual isolate or a loss of hemagglutinating potency in older samples.

Lastly, all IAV samples were assayed in doublets, but ideally it should be triplets, since variation in a single row will render the sample assay inconclusive. However, this would increase both time and expenses, and the amounts of IAV sample and antiserum are also limiting factors.

Sequencing The only limitations encountered during sequencing was evidence of contamination, which rendered a number of the sequences incomplete. The incomplete sequences were therefore omitted from the final sequence analysis of antigenic sites. Repeating the experiments would solve this problem, but time was a limiting factor in this scenario.

Western blot Using western blot to study glycosylation only provides confirmation on the relative presence of the glycan, but no information regarding the position or structure/type. Experiments using IAV isolates were performed multiple times with no visible protein bands, and several techniques and approaches were employed as a result. Regrowing some of the isolates as lysates and increasing the amount of viral sample added to each well of the Tris-Glycine gel (30μL to 90 μL) made it clear that the viral concentrations of the IAV isolates were too low to perform western blots using HRP as a colorimetric marker. There were visible bands at these concentrations, but not any specific to HA. Lower amounts of lysis buffer C during lysis did not alleviate the problem, but it is possible that using a different visualization technique, e.g. chemiluminescence, would solve this problem, but this requires an instrument able to detect these signals. Although the NetNGlyc results gave an adequate indication of potential glycosylations, it would be interesting to examine all IAV samples by western blot

to underline the coherence between NetNGlyc results and the actual nature of the protein.

SPR measurements Getting the optimal results from SPR measurements generally requires well-defined ligands and analytes, in the sense that the concentration and content of the these are known. Additionally, it is preferable to use purified analytes and ligands to reduce unspecific and incomplete binding. Although the primary ligand (anti-ferret Igγ) was sufficiently purified, the same is not the case for the used ferret antiserum (secondary ligand) nor the IAV isolates (analytes). Although the ferret antiserum showed ample specific binding, it still contains a high amount of albumin, which will bind to the sensorchip surface, and thus the primary ligand, in high amounts, thereby possibly reducing the amount of bound secondary ligands. This will diminish the signal of the binding between secondary ligand and analyte, which is not desirable considering that analyte concentrations are already relatively low. Furthermore, the concentration of Abs in the secondary ligand was also unknown, which makes it difficult to obtain any constants related to binding events. Additionally, the amount of available antiserum solution was a limiting factor in terms of how many measurements could be done. Therefore, knowing the concentration might have made further dilutions possible without significant loss of specific binding, in turn allowing for more primary measurements.

Improving experimental procedures

There are ample opportunities for improving the experimental procedures given enough time and resources. Examining the HA- and HAI-assays to begin with, it would be ideal to repeat HAI-assay of the IAV samples not demonstrating back titrations of $+++$, to have a better basis for comparison between their HAI-titers and fold-change factors. Repeats should also be done in order to lessen the degree of variety in the titers of the reference virus controls. The initial set of selected IAV samples for surveying was 35, but only 25 of those could be grown to express sufficiently high HA-titers, and from those only 23 could be sequenced and had acceptable back titration. With more time it might be possible to obtain results for the remaining IAV samples, and also repeat the HAI-assays to acquire acceptable back titration values.

Increasing the IAV sample size, particularly for the earlier seasons, could provide a better basis for comparison between the results across all experimental methods. It would have a great impact on the sequence analysis, as it would give a better idea of how the virus has evolved in the antigenic sites since the 2009 pandemic. This way, it would be possible to assemble a broader comparison of the evolutionary pathway of A(H1N1)pdm2009 to that of the 1918

pandemic. Having more time to redo the sequencing of those IAV samples that were not fully sequences would also improve the situation, albeit only slightly. In spite of all the advantages of manual RNA extraction it was still a time consuming procedure, and several repeats were required due to continuous contamination, which is why further sequencing was ceased, and another isolation method was used for the remaining samples. With enough time, it would be possible to obtain IAV H1 sample sequences with similar concentration and purity through the RNeasy Mini Kit, since the MagNA Pure LC isolation kit provided IAV H1 samples of lower concentration and purity.

SPR measurements were only performed on a select group of IAV samples, and doing a full assay of all samples would be beneficial concerning comparison between the results. The experiments performed on the selected IAV samples should be repeated for two reasons, but mainly to eliminate the buffer drift which impeded the proper evaluation of the results. Secondly, it would be interesting to assess the increased fold-change factors of A/Denmark/05/14, A/Denmark/30/16 and A/Denmark/46/16 from the perspective of binding affinity, since it was difficult to conclude whether additional glycosylation is the determinant of antigenic differences in those three samples. This might reveal information regarding the origin of the antigenic change.

Purifying the IAV samples for the SPR measurements would be ideal, because the inessential contents of the IAV isolates, i.e. growth medium and MDCK cell debris, could affect the highly sensitive components of the SPR instrument negatively, thereby impacting the results. Passing the IAV isolate through a lectin column, thereby retaining only the glycosylated components of the isolate, might serve as an intermediate solution. This could also be useful for the western blot, as it is would decrease the likelihood of unspecific bands on the final blot.

Lastly, performing the SPR measurements in a Class II laboratory would allow the use of active IAVs, which would be more fitting when constructing an analysis aimed at an active version of the reference virus. It would lessen the potential damage formalin can have on the antigens and the ligands as well. Inactivation imparts an uncertainty upon the structure of the antigen, which requires extensive research to fully understand.

Future outlook

Expanding the investigative scope of the thesis would be interesting, and there are several approached which could shed some light on the lesser known aspects of the results.

SPR could be utilized to explore the kinetics of binding between virus and antiserum, but perhaps also with sialic acid receptors

or RBCs. This would allow for quantification of the antigen-Ab binding beyond that which is attainable by immunological assays, thereby giving a more comprehensive insight into their interactions. There are still obstacles to overcome in terms of the quality of the used analytes and ligand, i.e. their condition, concentration and purity, but other studies prove that it is possible to perform SPR measurements on IAV HA given the right circumstances [78, 83]. For instance, it was possible to monitor steady-state equilibrium binding betwen mouse-derived Abs and recombinant trimerized HA molecules [83]. This indicates that it would be ideal to first isolate the HA molecule from the virus and then measure the binding kinetics. The fastest method would be to break up the virion and pass the viral proteins through a lectin column containing lectins specific to HA. However, this would require extensive knowledge of the glycans on HA and NA, since it would be necessary to find a lectin that is exclusive towards the glycans of HA. A safer, albeit slower method would to transfer the genetic material of HA from the IAV samples onto plasmids and grow them separately in other organisms. However, this experiment almost warrants a study of its own, especially considering the amount of recombinant HA samples that must be grown and isolated.

Having purified HA molecules could open up possibilities in other areas, such as using mass spectrometry (MS) to locate glycosylated motifs. Glycoamidases such as Peptide-N-Glycosidase F can be used to enzymatically cleave the glycans at the bond between GlcNAc and the Asn residue, thereby transforming Asn to Asp. MS makes it possible to discern differences in molecular weight of the fragments obtained by digesting the HA protein with trypsin, since there is a difference of 1 Da in molecular weight between the two amino acids, i.e. 132 Da for Asn and 133 Da for Asp. Heavier fragments, when compared to the theoretical weight, will then be an indicator of glycosylated Asn residues. [84] This method of mapping glycosylation would provide more information and certainty of glycosylation than western blotting can provide. The position and amount of glycans on the HA samples would be known, which would give more qualitative results.

To obtain better understanding of the antigenic properties of the virus it would be interesting to simulate the HA protein in certain situations. For instance, it would be possible to introduce the substitutions of interest, such as Ser179→Asn and Ala156→Asp, to observe how the antigenic sites are affected structurally. It might also be interesting to look into the stability of A/Denmark/05/14 to see whether it was affected by Ala156→Asp substitution. Although this variant had larger fold-change factor than other IAV sample prior to the 2015-2016 season it never reappears, which might mean that there are other fitness factors, e.g. stability or receptor

binding, that have a negative effect. Simulating the HA protein with an additional glycan at position 179 would be of great significance, since it might be possible to conclude where and how it could potentially block neutralization by Abs. Docking simulations with the Abs as the ligand would further this research as well. This idea could be coupled with MS studies into the structure of the glycan in order to present a more complete picture of their interactions.

The IAV samples of this thesis are tested exclusively against the reference virus, but experimenting with different Abs could potentially uncover whether an updated vaccine is required. For instance, an antiserum could be raised against the IAV samples harbouring the Ser179→Asn substitution to see how it matches with IAV strains *not* harbouring said substitution. Furthermore, doing enzyme-linked immunosorbent assays with monoclonal Abs aimed at the Sa site could potentially reveal whether the substitution at residue 179 has a significant effect on the antigenicity of the IAV samples. This could also be done with SPR, which again would be able to quantify any change in the binding affinity Abs have towards the Sa site.

Lastly, it seems prudent to resume this survey of IAV samples after the next seasonal outbreak of H1N1, mainly to see whether the Ser179→Asn substitution has been maintain and if the glycosylation profile has changed. As mentioned, the substitution at residue 179 was also introduced in the wake of the 1918 pandemic, but was later exchanged for a substitution at residue 177 in 1951 [76]. It is probable that A(H1N1)pdm2009 will take a similar path in terms of glycosylation and that a 177 substitution will emerge in the coming seasons, but it might take several years before it is observed again. However, the heightened immune pressure derived from constantly updated vaccines might accelerate this process. It did take H1N1 more than a decade to develop the 179 substitution after the 1918 pandemic, and the early appearance in A(H1N1)pdm2009 might be an indicator of the heightened immune pressure. [54]

Bibliography

[1] Natalie J. McDonald, Catherine B. Smith, and Nancy J. Cox. Antigenic drift in the evolution of H1N1 influenza A viruses resulting from deletion of a single amino acid in the haemagglutinin gene. *Journal of General Virology*, **88**(1):3209–3213, 2007.

[2] Teri Shors. *Understanding Viruses*. Number ISBN: 978-0-7637-2932-5 in 1st ed. Jones and Bartlett Publishers, 2008.

[3] Robert G. Webster, William J. Bean, Owen T. Gorman, Thomas M. Chambers, and Yoshihiro Kawaoka. Evolution and Ecology of Influenza A Viruses. *Microbiological Reviews*, **56**(1):152–179, 1992.

[4] Peter Palese and Adolfo García-Sastre. Influenza vaccines: present and future. *The Journal of Clinical Investigation*, **110**(1):9–13, 2002.

[5] Chih-Jen Wei, Jeffrey C. Boyington, Kaifan Dai, Katherine V. Houser, Melissa B. Pearce, Wing-Pui Kong, Zhi yong Yang, Terrence M. Tumpey, and Gary J. Nabel. Cross-Neutralization of 1918 and 2009 Influenza Viruses: Role of Glycans in Viral Evolution and Vaccine Design. *Science Translational Medicine*, **2**(24):1–8, 2010.

[6] Jin Il Kim, Ilseob Lee, Sehee Park, Min-Woong Hwang, Joon-Yong Bae, Sangmoo Lee, Jun Heo, Mee Sook Park, Adolfo García-Sastre, and Man-Seong Park. Genetic Requirement for Hemagglutinin Glycosylation and Its Implications for Influenza A H1N1 Virus Evolution. *Journal of Virology*, **87**(13):7539–7549, 2013.

[7] Amie J. Eisfeld, Gabriele Neumann, and Yoshihiro Kawaoka. At the centre: influenza A virus ribonucleoproteins. *Nature Reviews Microbiology*, **13**(1):28–41, 2014.

[8] Rafael A. Medina, Silke Stertz, Balaji Manicassamy, Petra Zimmerman, Xiangjie Sun, Randy A. Albrecht, Hanni Uusi-Kerttula, Osvaldo Zagordi, Robert B. Belshe, Sharon E. Frey, Terrence M. Tumpey, and Adolfo García-Sastre. Glycosylation in the Globular Head of the Hemagglutinin Protein Modulate the Virulence and Antigenic Properties of the H1N1 Influenza Virus. *Science Translational Medicine*, **5**(187):1–12, 2013.

[9] *Influenza (Seasonal) Fact sheet.* `http://www.who.int/mediacentre/factsheets/fs211/en/`. Downloaded: 03-03-2017.

[10] Suman R. Das, Scott E. Hensley, Alexandre David, Loren Schmidt, James S. Gibbs, Pere Puigbò, William L. Ince, Jack R. Bennink, and Jonathan W. Yewdell. Fitness costs limit influenza A virus hemagglutinin glycosylation as an immune evasion strategy. *Proceedings of the National Academy of Sciences*, **108**(51):1417–1422, 2011.

[11] James S. Gibbs, Daniela Malide, Felicita Hornung, Jack R. Bennink, and Jonathan W. Yewdell. The Influenza A PB1-F2 Protein Targets the Inner Mitochondrial Membrane via a Predicted Basic Amphipathic Helix That Disrupts Mitochondrial Function. *Journal of Virology*, **77**(13):7214–7224, 2003.

[12] Wentian Chen, Yaogang Zhong, Yannan Qin, Shinseng Sun, and Zheng Li. The Evolutionary Pattern of Glycosylation Sites in Influenza Virus (H5N1) Hemagglutinin and Neuraminidase. *PLoS - One*, **7**(11):1–11, 2012.

[13] *Types of Influenza Viruses.* `https://www.cdc.gov/flu/about/viruses/types.htm`. Downloaded: 07-02-2017.

[14] Rafael A. Medina and Adolfo Garcia-Sastre. Influenza A viruses: new research developments. *Nature Reviews Microbiology*, **9**(1):590–603, 2011.

[15] Suxiang Tong, Yan Li, Pierre Rivailler, Christina Conrardy, Danilo A. Alvarez Castillo, Li-Mei Chen, Sergio Recuenco, James A. Ellison, Charles T. Davis, Ian A. York, Amy S. Turmelle, David Moran, Shannon Rogers, Mang Shi, Ying Tao, Michael R. Weil, Kevin Tang, Lori A. Rowe, Scott Sammons, Xiyan Xu, Michael Frace, Kim A. Lindblade, Nancy J. Cox, Larry J. Anderson, Charles E. Rupprecht, and Ruben O. Donis. A distinct lineage of influenza A virus from bats. *Proceedings of the National Academy of Sciences of the United States of America*, **109**(11):4269–4274, 2012.

[16] Suxiang Tong, Xueyong Zhu, Yan Li, Mang Shi, Jing Zhang, Melissa Bourgeois, Hua Yang, Xianfeng Chen, Sergio Recuenco, Jorge Gomez, Li-Mei Chen, Adam Johnson, Ying Tao, Cyrille Dreygus, Wenli Yu, Ryan McBride, Paul J. Carney, Amy T. Gilbert, Jessie Chang, Zhu Guo, Charles T. Davis, James C. Paulson, James Stevens, Charles E. Rupprecht, Edward C. Holmes, Ian A. Wilson, and Ruben O. Donis. New World Bats Habour Diverse Influenza A Viruses. *PLoS - Pathogens*, **9**(10):1–12, 2013.

[17] D. Fleury, S. A. Wharton, J. J. Skehel, M. Knossow, and T. Bizebard. Antigen distortion allows influenza virus to escape neutralization. *Nature structural biology*, **5**(2):119–123, 1998.

[18] *Influenza-A Virus Replication Cycle.* `www.QIAGEN.com/fi/shop/genes-and-pathways/pathway-details/?pwid=247`. Downloaded: 07-02-2017.

[19] Rui Xu, Ryan McBride, Corwin M. Nycholat, James C. Paulson, and Ian A. Wilson. Structural Characterization of the Hemagglutinin Receptor Specificity from the 2009 H1N1 Influenza Pandemic. *Journal of Virology*, **86**(2):982–990, 2012.

[20] Stephen C. Harrison. Viral membrane fusion. *Nature Structural & Molecular Biology*, **15**(7):690–698, 2008.

[21] Nongluk Sriwilaijaroen and Yasuo Suzuki. Molecular basis of the structure and function of H1 hemagglutinin in influenza virus. *Proceedings of the Japan Academy, Series B*, **88**(1):226–249, 2012.

[22] Rui Xu and Ian A. Wilson. Structural Characterization of an Early Fusion Intermediate of Influenza Virus Hemagglutinin. *Journal of Virology*, **85**(10):5172–5182, 2011.

[23] Guangxiang Luo, Al Torri, William E. Harte, Stephanie Danetz, Christopher Cianci, Laurence Tiley, Susan Day, David Mullaney, Kuo-Long Yu, Carl Ouellet, Pierre Dextraze, Nicholas Meanwell, Richard Colonno, and Mark Krystal. Molecular Mechanism Underlying the Action of a Novel Fusion Inhibitor of Influenza A Virus. *Journal of Virology*, **71**(5):4062–4070, 1997.

[24] Juan J. Sanz-Ezquerro, Thomas Zürcher amd Susana de la Luna, Juan Ortín, and Amelia Nieto. The Amino-Terminal One-Third of the Influenza Virus PA Protein Is Responsible for the Induction of Proteolysis. *Journal of Virology*, **70**(3):1905–1911, 1996.

[25] S. Tripathi, J. Batra, W. Cao, K Sharma, J. R. Patel, P. Ranjan, A. Kumar, J. M. Katz, N. J. Cox, R. B. Lal, S. Sambhara, and S. K. Lal. Influenza A virus nucleoprotein induces apoptosis in human airway epithelial cells: implications of a novel interaction between nucleoprotein and host protein Clusterin. *Cell death and disease*, **4**(562):1–12, 2013.

[26] Harvey Lodish, Arnold Berk, Chris A. Kaiser, Monty Krieger, Anthony Bretscher, Hidde Ploegh, Angelika Amon, and Matthew P. Scott. *Molecular Cell Biology*. Number ISBN: 978-1-4292-3413-9 in 7th ed. W. H. Freeman and Company, 2012.

[27] Michelle D. Tate, Emma R. Job, Yi-Mo Deng, Vithiagaran Gunalan, Sebastian Maurer-Stroh, and Patrick C. Reading. Playing Hide and Seek: How Glycosylation of the Influenza Virus Hemagglutinin Can Modulate the Immune Response to Infection. *Viruses*, **6**(1):1294–1316, 2014.

[28] Marine L. B. Hillaire, Martin van Eijk, Nella J. Nieuwkoop, Stella E. Vogelzang van Trierum, Ron A. M. Fouchier, Albert D. M. E. Osterhaus, Henk P. Haagsman, and Guus F. Rimmelzwaan. The number and position of N-linked glycosylation sites in the hemagglutinin determine differential recognition of seasonal and 2009 pandemic H1N1 influenza virus by porcine surfactant protein D. *Virus Research*, **169**(1):301–305, 2012.

[29] Kevan L. Hartshorn, Richard Webby, Mitchell R. White, Tesfaldet Tecle, Clark Pan, Susan Boucher, Rodney J. Moreland, Erika C. Crouch, and Ronald K. Scheule. Role of viral hemagglutinin glycosylation in anti-influenza activities of recombinant surfactant protein D. *Respiratory Research*, **9**(65):1–12, 2008.

[30] Michelle D. Tate, Andrew G. Brooks, and Patrick C. Reading. Specific Sites of *N*-Linked Glycosylation on the Hemagglutinin of H1N1 Subtype Influenza A Virus Determine Sensitivity to Inhibitors of the Innate Immune System and Virulence in Mice. *The Journal of Immunology*, **187**(1):1884–1894, 2011.

[31] Jhang-Wei Huang, Wei-Fan Lin, and Jinn-Moon Yan. Antigenic sites of H1N1 influenza virus hemagglutinin revealed by natural isolates and inhibition assays. *Vaccine*, **30**(1):6327–6337, 2012.

[32] Yoko Matsuzaki, Kanetsu Sugawara, Mina Nakauchi, Yoshimasa Takahashi, Taishi Onodera, Yasuko Tsunetsugu-Yokota, Takayuki Matsumura, Manabu Ato, Kazuo Kobayashi, Yoshitaka Shimotai, Katsumi Mizuta, Seiji Hongo, Masato Tashiro, and Eri Nobusawa. Epitope Mapping of the Hemagglutinin Molecule of A/(H1N1)pdm09 Influenza Virus by Using Monoclonal Antibody Escape Mutans. *Journal of Virology*, **88**(21):12364–12373, 2014.

[33] Elmar Krieger and Gert Vriend. YASARA View - molecular graphics for all devides - from smartphones to workstations. *Bioinformatics*, **30**(20):2981–2982, 2014.

[34] T. Tsibane, D. C. Ekiert, J. C. Krause, O. Martinez, J. E. Crowe, I. A. Wilson, and C. F. Basler. Influenza Human Monoclonal Antibody 1F1 Interacts with Three

Major Antigenic Sites and Residues Mediating Human Receptor Specificity in H1N1 Viruses. *PloS Pathogens*, **8**(12):1–9, 2012.

[35] Hua Yang, Paul Carney, and James Stevens. Structure and Receptor binding properties of a pandemic H1N1 virus hemagglutinin. *PLoS Currents*, **1**(1), 2010.

[36] David O. White and Frank J. Fenner. *Medical Virology*. Number ISBN: 0-12-746642-8 in 4th ed. Academic Press, 1994.

[37] World Health Organization. *State of the art vaccine research and development*, 2006.

[38] W. E. P. Beyer, A. M. Palache, and A. D. M. E. Osterhaus. Comparison of Serology and Reactogenicity between Influenza Subunit Vaccines and Whole Virus or Split Vaccines. *Clinical Drug Investigation*, **15**(1):1–12, 1998.

[39] Bernadette Ferraro, Matthew P. Morrow, Natalie A. Hutnick, Thomas H. Shin, Colleen E. Lucke, and David B. Weiner. Clinical Applications of DNA Vaccines: Current Progress. *Vaccines*, **53**(1):296–302, 2011.

[40] *Selecting Viruses for the Seasonal Influenza Vaccine*. `https://www.cdc.gov/flu/about/season/vaccine-selection.htm`. Downloaded: 06-03-2017.

[41] *Vaccine Effectiveness - How Well Does the Flu Vaccine Work?* `https://www.cdc.gov/flu/about/qa/vaccineeffect.htm`. Downloaded: 06-03-2017.

[42] *WHO Questions and Answers - Vaccine effectiveness estimates for seasonal influenza vaccines*. `http://www.who.int/influenza/vaccines/virus/recommendations/201502_qanda_vaccineeffectiveness.pdf`. Downloaded: 06-03-2017.

[43] Scott E. Hensley, Suman R. Das, Adam L. Bailey, Loren M. Schmidt, Heather D. Hickman, Akila Jayaraman, Karthik Viswanathan, Rahul Raman, Ram Sasisekharan, Jack R. Binnink, and Jonathan W. Yewdell. Hemagglutinin Receptor Binding Avidity Drives Influenza Virus Antigenic Drift. *Science*, **326**(5953):734–736, 2009.

[44] Stephen J. Stray and Lindsey B. Pittman. Subtype- and antigenic site-specific differences in biophysical influences on evolution of influenza virus hemagglutinin. *Virology Journal*, **9**(91):1–16, 2012.

[45] Esteban Domingo and Celia Perales. Viral Quasispecies and Lethal Mutagenesis. *European Review*, **24**(1):39–48, 2016.

[46] Claus O. Wilke, Jia Lan Wang, Charles Ofria, Richard E. Lenski, and Christoph Adami. Evolution of digital organisms at high mutation rates leads to survival of the flattest. *Nature*, **412**(19):331–333, 2001.

[47] Adam S. Lauring, Judith Frydman, and Raul Andino. The role of mutational robustness in RNA virus evolution. *Nature reviews*, **11**(1):327–336, 2013.

[48] Scott E. Hensley, Suman R. Das, Adam L. Bailey, Loren M. Schmidt, Heather D. Hickman, Akila Jayaraman, Karthik Viswanathan, Rahul Raman, Ram Sasisekharan, Jack R. Bennink, and Jonathan W. Yewdell. Hemagglutinin Receptor Binding Avidity Drives Influenza A Virus Antigenic Drift. *Science*, **326**(5953):734–736, 2009.

[49] Jihui Ping, Chengjun Li, Guohua Deng, Yongping Jiang, Guobin Tian, Shuxia Zhang, Zhigao Bu, and Hualan Chen. Single amino-acid mutation in the HA alters the recognition of H9N2 influenza virus by a monoclonal antibody. *Biochemical and Biophysical Research Communications*, **371**(1):168–171, 2008.

[50] Xiangjie Sun, Akila Jayaraman, Pavithra Maniprasad, Rahul Raman Kathrine V. Houser, Claudia Pappas, Hui Zeng, Ram Sasisekharan, Jacqueline M. Katz, and Terrence M. Tumpey. N-linked Glycosylation of the Hemagglutinin Protein Influences Virulence and Antigenicity of the 1918 Pandemic and seasonal H1N1 Influenza A Virus. *Journal of Virology*, **87**(15):8756–8766, 2013.

[51] Maureen E. Taylor and Kurt Drickamer. *Introduction to Glycobiology*. Number ISBN: 978-0-19-956911-3 in 3rd ed. Oxford University Press, 2011.

[52] Suman R. Das, Pere Puigbò, Scott E. Hensley, Darrel E. Hurt, Jack R. Bennink, and Jonathan W. Yewdell. Glycosylation Focuses Sequence Variation in the Influenza A H1 Hemagglutinin Globular Domain. *PLoS Pathogens*, **6**(11):1–13, 2010.

[53] Yan Zhang, Jiping Zhu, Yongtao Li, Konrad C. Bradley, Jiyue Cao, Huanchun Chen, Meilin Jin, and Hongbo Zhou. Glycosylation on Hemagglutinin Affects the Virulence and Pathogenicity of Pandemic H1N1/2009 Influenza A Virus in Mice. *PloS ONE*, **8**(4):1–9, 2013.

[54] Emma R. Job, Yi-Mo Deng, Kenneth K. Barfod, Michelle D. Tate, Natalie Caldwell, Scott Reddiex, Sebastian Maurer-Stroh, Andrew G. Brooks, and

Patrick C. Reading. Addition of Glycosylation to Influenza A Virus Hemagglutinin Modulates Antibody-Mediated Recognition of H1N1 2009 Pandemic Viruses. *The Journal of Immunology*, **190**(1):2169–2177, 2013.

[55] David J. Vigerust and Virginia L. Shepherd. Virus glycosylation: role in virulence and immune interactions. *TRENDS in Microbiology*, **15**(5):211–218, 2007.

[56] Irene T. Schulze. Effects of Glycosylation on the Properties and Function of Influenza Virus Hemagglutinin. *The Journal of Infectious Diseases*, **176**(1):24–28, 1997.

[57] Shisheng Sun, Qinzhe Wang, Fei Zhao, Wentain Chen, and Zheng Li. Glycosylation Site Alteration in the Evolution of Influenza A (H1N1) Viruses. *PloS ONE*, **6**(7):1–9, 2011.

[58] Amie J. Eisfeld, Gabriele Neumann, and Yoshihiro Kawaoka. Influenza A virus isolation, culture and identification. *Nature Protocols*, **9**(11):2663–2681, 2014.

[59] Inna G. Ovsyannikova, Sarah J. White, Randy A. Albrecht, Adolfo García-Sastre, and Gregory A. Poland. Turkey Versus Guinea Pig Blood Cells: Hemagglutination Differences Alter Hemagglutinin Inhibition Responses Against Influenza A/H1N1. *Viral Immunology*, **27**(4):174–178, 2014.

[60] World Health Organization Collaborating Centres. *Manual for the laboratory diagnosis and virological surveillance of influenza*. Number ISBN: 978-92-4-154809-0 in 1st ed. WHO Press, 2011.

[61] European Centre for Disease Prevention and Control. *Influenza virus characterisation guidelines for the northern hemisphere influenza season 2015-2016*, 2015.

[62] F. Sanger and A. R. Coulson. A Rapid Method for Determining Sequences in DNA by Primed Synthesis with DNA Polymerase. *Journal of Molecular Biology*, **94**(1):441–448, 1975.

[63] F. Sanger, S. Nicklen, and A. R. Coulson. DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences*, **74**(12):5463–5467, 1977.

[64] Jane B. Reece, Lisa A. Urry, Michael L. Cain, Steven A. Wasserman, Peter V. Minorsky, and Robert B. Jackson. *Campbell Biology*. Number ISBN: 978-0-32-177565-8 in 10th ed. Pearson, 2014.

[65] Tahrin Mahrmood and Ping-Chang Yang. Western Blot: Technique, Theory, and Trouble Shooting. *North American Journal of Medical Sciences*, **4**(9):429–434, 2012.

[66] Yan Mao, Yu Bao, Wei Wang, Zhenggang Li, and Li Niu. Development and Application of Time-Resolved Surface Plasmon Resonance Spectrometer. *American Journal of Analytical Chemistry*, **2**(1):589–604, 2011.

[67] Anna J. Tudos and Richard B. M. Schasfoort. *Handbook of Surface Plasmon Resonance*, pages 1–14. Number ISBN: 978-0-85404-267-8 in Volume 1. RSC Publishing, 2008.

[68] Anupam Singhal, Charles A. Haynes, and Carl L. Hansen. Microfluidic Measurement of Antibody-Antigen Binding Kinetics from Low-Abundance Samples and Single Cells. *Analytical Chemistry*, **82**(20):8671–8679, 2010.

[69] Matthew A. Cooper. Optical Biosensors in Drug Discovery. *Nature reviews*, **1**(7):515–528, 2002.

[70] Robert C. Edgar. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, **32**(5):1792–1797, 2004.

[71] D. Jernigan, D. Wentworth, J. Katz, J. Barnes, R. Garten, and X. Xu. Influenza A virus (A/California/07/2009(H1N1)) segment 4 hemagglutinin (HA) gene, complete cds. *WHO Collaborating Center for Surveillance, Epidemiology and Control of Influenza, Influenza Division, Centers for Disease Control and Prevention*, 2016.

[72] Tamura K. and Nei M. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular biology and Evolution*, **10**(1):512–526, 1993.

[73] Kumar S., Stecher G., and Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*, **33**(1):1870–1874, 2016.

[74] R. Gupta, E. Jung, and S. Brunak. *Prediction of N-glycosylation sites in human proteins*, 2004.

[75] invitrogen. *XCell II™ Blot Module*, 2009.

[76] Shisheng Sun, Qinzhe Wang, Fei Zhao, Wentian Chen, and Zheng Li. Prediction of Biological Functions on Glycosylation Site Migration in Human Influenza H1N1 Viruses. *PloS ONE*, **7**(2):1–8, 2012.

[77] Emma R. Job, Yi-Mo Deng, Michelle D. Tate, Barbara Bottazzi, Erika C. Crouch, Melinda M. Dean, Alberto Mantovani, Andrew G. Brooks, and Patrick C. Reading. Pandemic H1N1 Influenza A Viruses Are Resistant to the Antiviral Activities of Innate Immune Proteins of the Collectin and Pentraxin Superfamilies. *The Journal of Immunology*, **185**(1):4284–4291, 2010.

[78] Emi Suenaga, Hiroshi Mizuno, and Kumar K. R. Penmetcha. Monitoring influenza hemagglutinin and the glycan interactions using surface plasmon resonance. *Biosensors and Bioelectronics*, **32**(1):195–201, 2012.

[79] Carl-Fredrik Mandenius, Ronghui Wang, Anna Aldén, Gunnar Bergström, Sabine Thébault, Charles Lutsch, and Sten Ohlson. Monitoring of influenza virus hemagglutinin in the process samples using weak affinity ligands and surface plasmon resonance. *Analytica Chimica Acta*, **623**(1):66–75, 2008.

[80] Darin K. Takemoto, John J. Skehel, and Don C. Wiley. A Surface Plasmon Resonance Assay for the Binding of Influenza Virus Hemagglutinin to Its Sialic Acid Receptor. *Virology*, **271**(1):452–458, 1996.

[81] Emi Suenaga, Hiroshi Mizuno, and Kumar K. R. Penmetcha. Influenza virus surveillance using surface plasmon resonance. *Virulence*, **3**(5):464–470, 2012.

[82] Olufemi O. Fatunmbi, John A. Newman, David A. Halvorson, and V. Sivanandan. Effect of Temperature on the Stability of Avian Influenza Virus Antigens under Different Storage Conditions. *American Association of Avian Pathologists*, **37**(3):639–646, 1993.

[83] Javier G. Magadán, Surender Khurana, Suman R. Das, Gregory M. Frank, James Stevens, Hana Golding, Jack R. Bennink, and Jonathan W. Yewdell. Influenza A Hemagglutinin Trimerization Completes Monomer Folding and Antigenicity. *Journal of Virology*, **87**(17):9742–9753, 2013.

[84] Ziv Roth, Galit Yehezkel, and Isam Khalaila. Identification and Quantification of Protein Glycosylation. *Internation Journal of Carbohydrate Chemistry*, **2012**(1):1–10, 2012.

[85] Dilusha S. Dalpathado and Heather Desaire. Glycopeptide analysis by mass spectrometry. *Analyst*, **133**(6):731–738, 2008.

[86] Dan Cojocari. *Department of Medical Biophysics,* ***University of Toronto***, 2010.

[87] Andrew D. Johnson. An extended IUPAC nomenclature code for polymorphic nucleic acids. *Bioinformatics*, **26**(10):1386–1389, 2010.

Appendix A

1. Glycan moeties and branching

An overview of the glycan structures mentioned in this thesis can be seen in figure A.1.

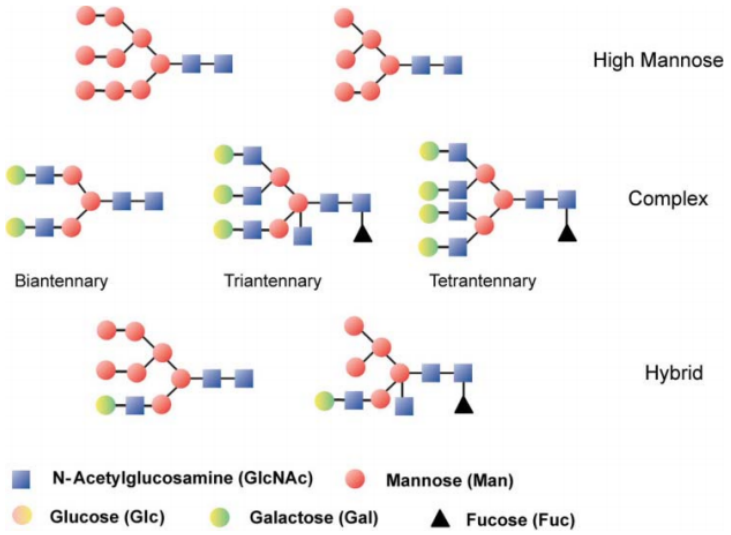

Figure A.1: Common types of glycans and some of their structures. [85]

Appendix B

1. Amino acids

The structure and abbreviations of amino acids can be seen in figure B.1.

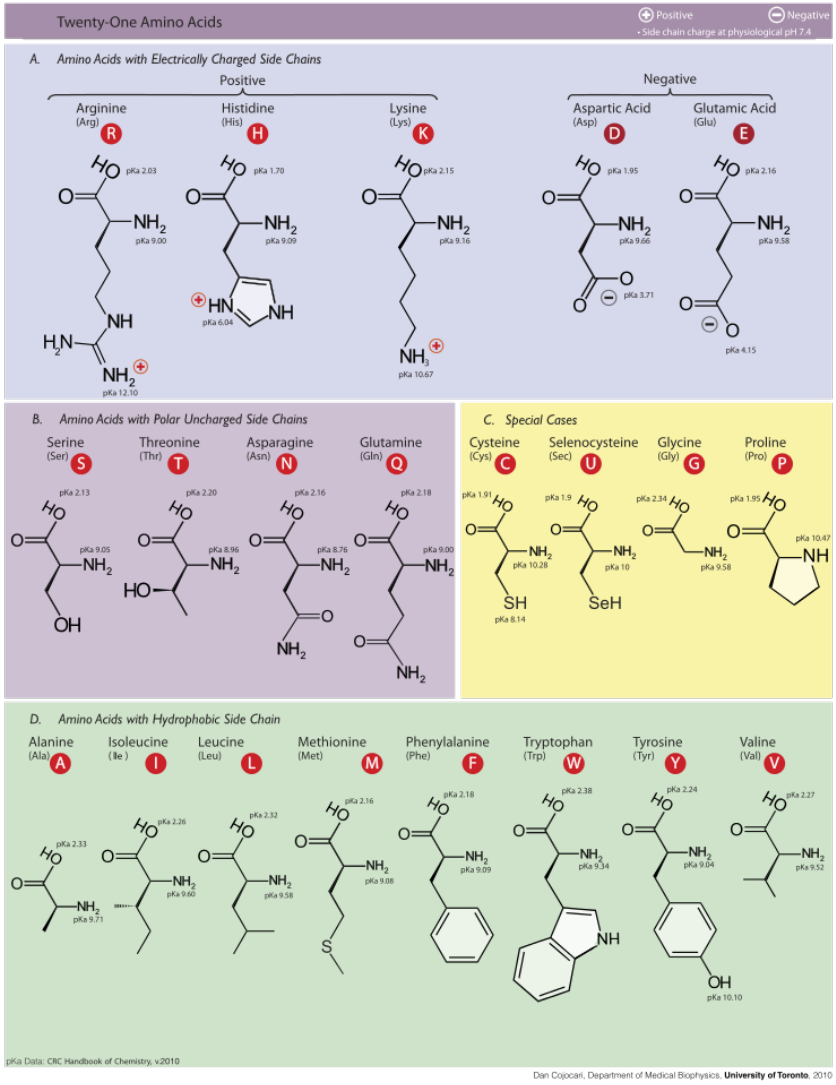

Figure B.1: Table displaying the structure and charge of 21 amino acids. Their abbreviations are also available. [86]

2. Nucleotides

Abbreviations ambiguity codes for nucleotides are seen in figure B.2.

Symbol	Mnemonic	Translation
A		A (adenine)
C		C (cytosine)
G		G (guanine)
T		T (thymine)
U		U (uracil)
R	puRine	A or G (purines)
Y	pYrimidine	C or T/U (pyrimidines)
M	aMino group	A or C
K	Keto group	G or T/U
S	Strong interaction	C or G
W	Weak interaction	A or T/U
H	not G	A, C or T/U
B	not A	C, G or T/U
V	not T/U	A, C or G
D	not C	A, G or T/U
N	aNy	A, C, G or T/U

Figure B.2: A table of the symbols used to describe nucleotides and ambiguous nucleotides in regards to sequencing results. [87]