Avian Influenza Virus

Deciphering receptor interactions and their role in interspecies transmission

PER ERIKSSON
Influenza A virus (IAV) annually infects approximately 5–15% of the human population, causing ~500,000 deaths globally. Novel IAVs have emerged and spread pandemically in the human population, but have over time established endemic circulation with reduced pathogenicity causing seasonal influenza. The natural reservoir of IAVs is wild waterfowl. The past pandemics have been associated with host switch and have partly or entirely originated from birds, or adapted via passage through pigs (postulated IAV mixing vessel). Understanding IAV interspecies transmission mechanisms is essential for pandemic preparedness. Enzootic circulation of avian IAV (AIV) is concentrated to a few waterfowl species, while other bird species seldom are infected. A species barrier preventing IAV interspecies transmission has been suggested. To investigate IAV host range and mixing vessels, histochemistry studies were conducted with tissues from avian species, pigs, and humans. Virus adaptation to new hosts was studied by challenging tufted ducks and chickens with mallard-derived AIVs, together with AIV receptor tropism and glycoproteomic analysis of receptor distribution. Finally, receptor and tissue tropism in ducks was studied systematically for AIV (H1–16). More abundant AIV attachment to human than pig tissues was observed, questioning the pig mixing vessel theory. Attachment patterns of AIVs to bird tissues was generally broad with abundant attachment to trachea. However, among ducks, pronounced attachment was observed to colon of Anas spp., suggesting that intestinal infection might be restricted to Anas spp., whereas other species may be susceptible to respiratory infection. Tufted ducks and chickens could not be infected by intraesophageal inoculation further supporting this hypothesis. Glycan array analysis revealed 3'SLN, 3'STF, and their fucosylated and sulfated analogues as main AIV receptors. Moreover, AIV Neu5Acα2,6 recognition was widespread. Avian respiratory and intestinal tracts glycoproteomic analysis revealed that avian and mammalian receptor structures are much more similar than earlier thought. Furthermore, observed AIV subtype titer variation in challenged tufted ducks and chickens did not correlate with virus receptor tropism. In summary, this thesis suggests that IAV receptor recognition, in particular α2,3 vs. α2,6 sialylated receptor structures, is less important for the IAV interspecies barrier than previously thought.

Keywords: birds, glycobiology, glycovirology, host range, mixing vessel, virus attachment

Per Eriksson, Department of Medical Biochemistry and Microbiology, Box 582, Uppsala University, SE-75123 Uppsala, Sweden.

© Per Eriksson 2019

ISSN 1651-6206
ISBN 978-91-513-0797-8
urn:nbn:se:uu:diva-395407 (http://urn.kb.se/resolve?urn:nbn:se:uu:diva-395407)
To all the birds.

Greater flamingos (*Phoenicopterus roseus*), Parc Naturel Régional de Camargue, Provence-Alpes-Côte d’Azur, France, April 2018.
List of Publications

This thesis is based on the following papers, which are referred to in the text by their Roman numerals.


IV. Verhagen H. J, Eriksson P, Leijten L, Blixt O, Olsen B, Kuiken T et al. Host range of influenza A virus H1 to H16 in dabbling vs. diving ducks based on tissue and receptor binding studies. manuscript in preparation
Additional Publications

Publications not included in the thesis:


Contents

Introduction ................................................................. 13
Influenza A virus in humans ......................................... 13
  Pandemic influenza ...................................................... 13
  Seasonal influenza ...................................................... 14
The influenza virus ....................................................... 15
  Taxonomy and structure .............................................. 15
  Evolution ................................................................. 15
  Major surface proteins .............................................. 15
  The natural reservoir .................................................. 16
Interspecies transmission of influenza virus .................... 18
  Influenza virus in non-\textit{Anas}/non-\textit{Larus} or allies genera ........ 18
  Influenza in domestic poultry ...................................... 18
Zoonotic transmission of influenza A virus ...................... 20
  The mixing vessel ..................................................... 21
Glycobiology in the Perspective of Influenza A Virus ........... 23
  General introduction to glycans .................................... 23
  General introduction to sialic acid .............................. 25
  The influenza virus receptor ...................................... 26
  Distribution of influenza virus receptors ..................... 26
Wild birds in the perspective of IAV ............................... 28
  The vast diversity of wild birds ................................. 28
  Wild birds as vectors of IAV ...................................... 29
  Immune response to AIV infection in avian hosts .......... 29
In summary ............................................................... 31

Aims ........................................................................ 33
  General aim ............................................................... 33
  Specific aims ............................................................. 33
    Study I .................................................................. 33
    Study II ................................................................. 33
    Study III ............................................................... 34
    Study IV ............................................................... 34
Abbreviations

3’SLN  Neu5Acα2,3Galβ1,4GlcNAc (3’-sialyl N-acetyllactosamine)
6’SdiLN Neu5Acα2,6Galβ1,4GlcNAcβ1,3Galβ1,4GlcNAc (6’-sialyl N-acetyldilactosamine)
6’SLN Neu5Acα2,6Galβ1,4GlcNAc (N-acetyllactosamine)
aa  Amino acid
AIV  Avian influenza A virus
ANP32 Acidic leucine-rich nuclear phosphoprotein 32
Asn  Asparagine
CL  Cloacal
D  Aspartic acid
d.p.i.  Days post infection
E  Glutamic acid
EID₅₀  Egg infectious dose 50
ELISA  Enzyme-linked immunosorbent assay
ER  Endoplasmatic reticulum
FITC  Fluorescein isothiocyanate
Fuc  Fucose
Gal  Galactose
GalNAc  N-acetylgalactosamine
Glc  Glucose
GlcNAc  N-acetylglicosamine
GTP  Guanosine triphosphate
H  Histidine
HA  Hemagglutinin
HCD  Higher-energy collision dissociation
Hex  Hexose
HexNAc  N-acetyl hexosamine
HPAIV  Highly pathogenic AIV
IAV  Influenza A virus
INF  Interferon
ISG  Interferon-stimulated gene
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>IVPI</td>
<td>Intravenous pathogenicity index</td>
</tr>
<tr>
<td>K</td>
<td>Lysine</td>
</tr>
<tr>
<td>Kdn</td>
<td>2-keto-3-deoxynononic acid</td>
</tr>
<tr>
<td>L</td>
<td>Leucine</td>
</tr>
<tr>
<td>LacNAc</td>
<td>N-acetyl lactosamine</td>
</tr>
<tr>
<td>LC-MSn</td>
<td>Liquid chromatography mass spectrometry (MS part iterated n times)</td>
</tr>
<tr>
<td>LPAIV</td>
<td>Low pathogenic AIV</td>
</tr>
<tr>
<td>LRT</td>
<td>Lower respiratory tract</td>
</tr>
<tr>
<td>m/z</td>
<td>Mass to charge ratio</td>
</tr>
<tr>
<td>MAA-II</td>
<td><em>Maackia amurensis</em> lectin II</td>
</tr>
<tr>
<td>Man</td>
<td>Mannose</td>
</tr>
<tr>
<td>Man6P</td>
<td>Mannose-6-phosphate</td>
</tr>
<tr>
<td>MPR</td>
<td>Man6P receptor</td>
</tr>
<tr>
<td>MRCA</td>
<td>Most recent common ancestor</td>
</tr>
<tr>
<td>Mx1</td>
<td>Interferon-induced GTP-binding protein Mx1</td>
</tr>
<tr>
<td>MYA</td>
<td>Million years ago</td>
</tr>
<tr>
<td>N</td>
<td>Asparaginene</td>
</tr>
<tr>
<td>NA</td>
<td>Neuraminidase</td>
</tr>
<tr>
<td>Neu5Ac</td>
<td>N-acetyleneuraminic acid</td>
</tr>
<tr>
<td>Neu5Gc</td>
<td>N-glycolyneuraminic acid</td>
</tr>
<tr>
<td>NP</td>
<td>Nucleoprotein</td>
</tr>
<tr>
<td>NS1</td>
<td>Non-structural protein 1</td>
</tr>
<tr>
<td>OP</td>
<td>Oropharyngeal</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate buffer saline</td>
</tr>
<tr>
<td>PBS-T</td>
<td>PBS-Tween (0.5 % Tween)</td>
</tr>
<tr>
<td>PTM</td>
<td>Post translational modification</td>
</tr>
<tr>
<td>PVA</td>
<td>Pattern of virus attachment</td>
</tr>
<tr>
<td>R</td>
<td>Arginine</td>
</tr>
<tr>
<td>RBS</td>
<td>Receptor binding site</td>
</tr>
<tr>
<td>RdRP</td>
<td>RNA-dependent RNA polymerase</td>
</tr>
<tr>
<td>RIG-I</td>
<td>Retinoic acid-inducible gene I protein</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>RT-qPCR</td>
<td>Reverse transcriptase quantitative real-time polymerase chain reaction</td>
</tr>
<tr>
<td>S</td>
<td>Sialyl (indicating the presence of sialic acid)</td>
</tr>
<tr>
<td>S-Le^x</td>
<td>Neu5Acα2,3Galβ1,4[Fucα1,3]GlcNAc (sialyl Lewis x)</td>
</tr>
<tr>
<td>SA</td>
<td>Sialic acid</td>
</tr>
<tr>
<td>Ser</td>
<td>Serine</td>
</tr>
</tbody>
</table>
SNA  *Sambucus nigra* lectin

SNFG  Symbol nomenclature for graphical representations of glycans

SNP  Single nucleotide polymorphism

ssRNA  Single-stranded RNA

STF  Neu5Acα2,3Galβ1,3GalNAc (sialyl-Thomsen-Friedenreich antigen)

Su  Sulfate (SO$_4^{2-}$)

T  Threonine

TMA  Tissue micro array

Thr  Threonine

u  Unified atomic mass unit

URT  Upper respiratory tract

V  Valine

vPol  Virus polymerase

vRNP  Viral ribonucleoprotein
Introduction

Influenza A virus in humans

Influenza virus is known as the ethiological agent of influenza in humans and annually infect 5–15% of the human population, causing approximately 500,000 deaths globally [1]. The human disease influenza is caused by both influenza A and B viruses, where influenza A viruses (IAVs) are dominating the number of cases and evolve more rapidly than type B viruses [2]. Influenza A virus is circulating in the human population as an epidemic with human to human transmission. There is a continuous evolution of IAVs in the human population due to for example the development of immunity. Additionally, there is gene flow of IAV genetic material from birds and/or swine to humans, i.e. the origin of human adapted IAV is from the avian IAV (AIV) gene pool [3].

Pandemic influenza

Occasionally, novel IAVs enter the human population after reassortment in pigs between avian ancestry and mammalian adapted IAVs [4, 5]. If the human population is naïve to the new virus, it might spread rapidly within the human population stepping up from epidemic to pandemic scale. Additionally, these viruses tend to have increased virulence and pathogenicity, replicating in the lower respiratory tract (LRT), often followed by secondary bacterial pneumonia [6]. In the past century, there have been four pandemics: H1N1 in 1918/19, H2N2 in 1957, H3N2 in 1968, and H1N1 in 2009 [4, 5]. The 1918 H1N1 pandemic is known as the most devastating pandemic of all time causing approximately 50 million casualties [7]. However, over time these viruses adapt to their new host and decrease in virulence and pathogenicity and eventually either disappear (e.g. due to herd immunity or emergence of new competitive subtypes) or remain as endemic in the human population [1].
Seasonal influenza

As the virulence and pathogenicity of a pandemic IAV fades and host adaptation increases, the attenuated progeny will become part of the human adapted IAV gene pool and circulate as seasonal influenza [1]. Thus, the viruses that circulate as seasonal influenza in humans are the descendants of ancestor pandemic viruses, which in turn have their origin in the wild bird reservoir. Accordingly, common human IAV subtype combinations are H1N1 and H3N2, and historically H2N2. For seasonal influenza, there is a continuous battle between virus vs. host [1]. This is illustrated by the rapid evolution of H1N1 and H3N2 viruses in humans [8]. Yet, these two viruses have followed different trajectories, where the H1N1 virus has subdivided into separate lineages at multiple time points, whereas the H3N2 virus has mainly evolved within one lineage. Seasonal IAV generally infect the upper respiratory tract (URT) and the infection is commonly cleared within one week [6]. Accordingly, it is mainly the innate immune response that fights IAV infection, whereas developed adaptive immunity (e.g. antibodies) protects from reinfection with the same or antigenically similar viruses [1]. There is thus a continuous battle between the host innate immune system and virus anti-host cell anti-virus response proteins (including non-structural protein 1, NS1) that enable the virus to suppress the host immune response (e.g. suppress host anti-virus interferon response). Moreover, hemagglutinin (HA) is the most immunogenic protein on the IAV particle surface, thus there is high selective pressure on HA as IAV virus immunity develops. One approach for the virus to hide HA epitopes is by altering the glycosylation pattern of HA [9-11]. However, this is a delicate balance as both the most immunogenic part of HA and the receptor binding site (RBS) are located in the HA1 domain. Thus, it is critical for the virus to not impair the RBS functionality as it hides HA antibody epitopes.
The influenza virus

Taxonomy and structure

Influenza virus is an enveloped negative sense single-stranded RNA (ssRNA) virus within Orthomyxoviridae [12]. There are several types of influenza viruses [1]. The different types are denoted by capital letters A–D, where A is the most prevalent type in humans and birds (Figure 1). The typing of influenza A–D is based on the nucleoprotein. Influenza A virus is the topic of this thesis.

The influenza A virus particle itself is pleomorphic, with a diameter of approximately 80–100 nm, but IAV particles with diameters exceeding 1 μm have been reported [13]. The IAV has a segmented genome constituting eight segments that are individually transported into the host cell nucleus and can be alternatively spliced to twelve different proteins [1, 14]. Among these proteins the IAV encodes its own RNA-dependent RNA polymerase (RdRP) complex comprised by PB1, PB2, and PA, commonly referred to as just virus polymerase (vPol).

Evolution

Influenza A viruses evolve via two fundamentally different processes: antigenic drift and antigenic shift [1, 12]. Antigenic drift is caused by the virus’ own error prone vPol, which introduces approximately one point mutation per virus genome replication. Thereby, IAVs are continuously evolving, e.g. allowing for escape from the host immune system. Each IAV infection will generate a distribution of sequence variant progeny virions (quasispecies) [1, 15]. The theoretical sequence space is of high dimensionality (sometimes referred to as virus swarms), but in reality the available sequence space collapses down to a reduced dimensionality space, due to the fitness cost of introduced mutations. Selection then acts on subpopulations within the generated quasispecies. Antigenic shift is caused by the segmented nature of the virus genome [1, 12]. If IAVs X and Y co-infect the same host cell, genome segments and structural proteins might be shuffled upon virus assembly, which allows for the generation of virus Z. Antigenic shift allows for rapid changes in virus properties that might be beneficial in a new environment.

Major surface proteins

The IAV particle is enveloped within a lipid membrane originating from the previous host cell [12]. Protruding from the membrane are the two major surface virus glycoproteins HA and neuraminidase (NA). The stoichiometric ratio of HA vs. NA varies between virus particle size and geometry, but is
approximately 3–30, with higher ratio for larger IAV particles [13]. Hemagglutinin is the surface attachment protein that recognizes the host cell, whereas NA is the surface detachment protein that cleaves the virus–host cell bond. Hemagglutinin is a homotrimer, whereas NA is a homotetramer [12]. Additionally, HA is produced as a propeptide (HA0) that is cleaved by specific host proteases to HA1 and HA2. HA0 is both pH and structurally stable, not enabling virus–host cell membrane fusion. But as HA0 is cleaved, the propeptide is activated and becomes more flexible, allowing fusion between the virus and host cell membranes. Each monomer in the HA homotrimer consists of a globular head and a stalk. The globular head is constituted by HA1 and contains the receptor binding site (RBS), whereas the stalk comprises HA2 and part of HA1. The transmembrane domain of HA is within HA2, which plays a critical role during virus–host cell fusion. The adhesion of multiple HA1s to host cell receptors triggers the uptake of the virus via endocytosis [14]. As the pH decreases in late endosomes, the virus membrane fuses with the endosome membrane and uncoats (due to pH-induced conformational change of HA1 and proton influx through matrix protein 2) and there is a conformational change in HA2 that has a “spring loaded” mechanism that “pulls” out the viral ribonucleoproteins (vRNPs) into the host cell cytoplasm where they are further imported into the nucleus. Recent research has shown that virus uncoating takes place about 30 to 90 min post infection, but there is only a minority of the virus particles’ vRNPs that successfully reach the host cell nucleus [14]. Indeed, vRNP host cell nuclear import has been suggested as a barrier for IAV interspecies transmission as the vRNPs (and later produced NP and vPol) must be compatible with the host importin-α family members mediating the active import [16].

Mature HA1 and HA2 are linked by disulfide bridges. Moreover, the combination of antigenically distinct HA and NA of a particular IAV is used in the nomenclature of IAVs giving the subtype HxNy. There are currently 18 HA and 11 NA subtypes described [1, 12, 17, 18].

The natural reservoir
The natural reservoir of IAV is wild birds, especially the two aquatic orders Anseriformes and Charadriiformes [12, 19-21]. In these birds, avian influenza virus (AIV) causes asymptomatic infection with virus shedding for approximately 3–5 days, without any apparent loss in fitness for the host. The single most well described host species is mallard (Anas platyrhynchos) [22-24]. In the Northern Hemisphere, the prevalence of AIV in the wild bird population follows a cyclic pattern with peak prevalence during the fall, coinciding with fall migration and high numbers of first cycle AIV naïve birds. The exact time point of the peak AIV prevalence is fluctuating be-
tween years, but during the fall months, maximum AIV prevalence can be as high as 25% in mallards.

There are 16 HA and 9 NA subtypes described in wild birds [19]. However, all AIV subtypes are not equally common in the wild bird population; surveillance studies have shown fluctuating trends in subtype prevalence [22-24]. Subtypes H1–7 are generally much more common in wild waterfowl than subtypes H8–12 [25]. Generally, the most common HA subtypes in wild waterfowl include H3, H4, and H6 in Anseriformes and non-laridean Charadriiformes and H13 and H16 in Laridae, whereas the most common NA subtypes include N1, N2, N3, N4, N6, N7, N8, and N9 [22-24, 26]. Some of the most frequently detected HxNy combinations reported include: H1N1, H2N3, H3N8, H4N6, H6N2, and H7N7. There are differences between different HA subtypes in which NA subtype they tend to pair with [25]. This is mainly due to a phenomenon known as functional balance of HA vs. NA, as HA and NA are antagonistic where HA binds to the host cell receptor, and NA cleaves the HA–host cell receptor bond by cleaving off the SA from the penultimate sugar [13, 27]. This is an equilibrium between the \( K_{HA,ss} \) to \( K_{HA,dis} \) ratio vs. \( k_{NA,cat} \). To add further complexity to this issue, HA and NA interact not only with receptors on the host cell surface, but also with decoy receptors in the extracellular matrix and mucus coating the epithelia, which interfere with infection [28]. Thus, a too large disequilibrium in HA vs. NA activity would be devastating for the virus to fulfill the cycle of attachment-infection-replication-disperse.

**Tissue tropism in the natural reservoir**

Avian IAV is described to replicate in the epithelial cells lining the intestinal tract in ducks, and in mallards it has been shown to be accentuated towards the large intestine [12, 19, 29-31]. Progeny virus is shed in feces and the virus is transmitted via the fecal oral route, as shed virus can stay persistent in cold water for several months [32]. Additionally, virus histochemistry has shown that AIV is capable of attach to respiratory epithelium, especially in other avian genera than *Anas* and related ducks [33]. Indeed, field surveillance studies have shown that AIV can be isolated from both respiratory and cloacal samples [34, 35]. Occasionally, sampled wild birds have even been coinfected with different AIV subtypes isolated from the respiratory vs. cloacal samples from the same bird, suggesting intersubtype tissue tropism differences. Historically, the main focus of AIV surveillance has been directed towards mallards and other dabbling ducks (for instance mallards are ubiquitous and easy to catch, while other ducks tend to be more scattered and stay off-shore, thus being harder to catch), but more recently it has been suggested that the focus should be broadened to include a wider set of avian families to avoid biased sampling at certain avian stopover hotspots and “numerous and easy to catch” avian species [36].
Interspecies transmission of influenza virus

Influenza virus in non-\textit{Anas}/non-\textit{Larus} or allies genera

Surveillance studies have identified dabbling ducks (\textit{Anas} and allies) as the main AIV reservoir [22-24]. Additionally, Laridae is a well-known AIV host [19-21, 25, 37]. Especially, gulls are known to primarily harbor the two gull specific subtypes H13 and H16. In contrast to mallards, where peak AIV titers are found in the large intestine, studies using virus histochemistry have shown that in several other non-\textit{Anas} genera, AIV attachment is instead accentuated to trachea [33, 38]. Besides the well-characterized AIV reservoir host species, AIV occasionally transmit to other host species/genera/families/orders (Figure 1) [24, 39-41]. This can either be dead end spillover transmission, or the virus can be further transmitted. Moreover, the clinical outcome of AIV infection in non-AIV reservoir species may vary greatly [12, 42].

Influenza in domestic poultry

On the domestic side, poultry is highly susceptible to AIV infection and is an important reservoir (mainly chickens) of AIV in addition to the wild bird reservoir [43, 44]. Domestic fowl can be infected by AIV when in contact with infected wild birds. This is especially true for e.g. free-range poultry that might share feed/water with wild fowl, or transmission can happen at e.g. live bird markets. But, transmission can also occur if humans (or others) bring AIV contaminated material (e.g. avian feces) into a poultry stable. Transmission of AIV from wild birds to poultry can either cause limited outbreaks or establish long-term maintained poultry specific linages. The most common HA subtypes in poultry include H1, H2, H5, H6, H7, H8, H9, and H10, whereas the most common NA subtypes include N1, N2, N3, N4, N5, N6, N7, and N8 [26, 44]. The most common HxNy combinations include H1N5, H7N7, and H8N4. Additional subtype combinations of particular interest include H5N1, H5N5, H5N6, H5N8, H6N1, H6N2, H7N9, and H9N2 [45-49]. Avian influenza virus is a costly burden for the poultry industry as it causes loss in production, risk for zoonotic transmission, and very costly interventions [50].
Low pathogenic vs. highly pathogenic avian influenza virus

Domestic poultry is highly susceptible to AIV, but the clinical symptoms might vary greatly from reduced growth and/or egg productivity, to almost 100% mortality, the clinical outcome is intra/inter subtype lineage dependent [12, 42, 51]. Due to the high impact on the poultry industry, AIVs are sometimes classified based on their pathogenic properties in chickens (Gallus gallus). Most AIVs are low pathogenic AIVs (LPAIVs), but subtypes H5 and H7 may occasionally mutate to highly pathogenic AIVs (HPAIVs). Highly pathogenic AIV is clinically characterized by very high mortality, commonly assessed by the intravenous pathogenicity index (IVPI) where six-week old specific pathogen-free chickens are intravenously inoculated and monitored daily for ten days and clinical symptoms are scored: 0 if normal, 1 if sick, 2 if paralyzed, and 3 if dead. Commonly, HPAIV is defined as causing >75% mortality and having an IVPI >1.2 in chickens [42]. Genetically HPAIV is characterized by a polybasic amino acid (aa) insert at the cleavage site of the HA0 precursor that allows promiscuous host cell serine proteases to activate HA, resulting in that HPAIV HA can be activated in various tissues. The tissue tropism and pathobiology of LPAIV vs. HPAIV varies greatly in poultry [42, 51]. Low pathogenic AIV is restricted to the respiratory and digestive tract in poultry and might cause limited lesions. On the other hand, HPAIV might spread systemically and can cause necrosis and high mortality.
Zoonotic transmission of influenza A virus

Occasionally, AIV transmit from Aves to Mammalia (Figure 1) [3, 19, 52]. In a general sense, IAV is using a broad set of hosts. However, many maintained lineages of IAV are adapted to a particular host species/genus. Crossing the border to transmit to other genera/families/orders or even classes (e.g. Aves to Mammalia) can be challenging for the virus due to interhost differences in physiological and immunological properties [3, 53]. Critical bottlenecks pinpointed include host cell entry/exit, replication, and spread within tissues. Additionally, the new host species must be able to further transmit the virus and not just be a dead-end infection. Moreover, competent donors and recipients must come in contact to enable virus transmission. Host physiological and immunological barriers can be overcome by virus mutations and lack of close contact between donor and recipient hosts can be overcome by the virus changing from contact-borne transmission to airborne transmission. However, at the same time these might be costly mutations for the virus and it is critical that the required mutations do not impair the endogenous functionality of the virus.

Studies of the 2014–2015 outbreak and epizootic of avian-origin H10N7 IAV in seals in northwest Europe found that the virus’ genetic diversity was the highest just after the host jump [54, 55]. As beneficial mutations are favored and fixed, the genetic diversity decreases over time and stabilizes, i.e. reduced genetic plasticity. The main sequence variation was observed in the HA gene, indicating the importance of host cell receptor recognition/entry. The importance of the divergence in host cell receptor expression across species as a barrier for interspecies transmission is well documented [56-58]. However, recently the importance of the virus encoded RdRP has been highlighted in the context of zoonotic transmission of IAVs [16]. Particularly, for successful replication the PB2 unit of the RdRP complex is dependent on interaction with the host cell protein ANP32 family, a group of host chromatin regulators [59, 60]. The ANP32 protein family is involved in many events in the host cell, including transcriptional regulation, protein phosphorylation, intracellular transport, and cell-death pathways [61]. In avian cells, AIVs’ PB2 interact with avian host cell ANP32A, whereas human adapted IAVs interact with ANP32A and ANP32B in human cells [62-64]. In human cells, AIV type PB2 has a malfunctional interaction with ANP32 proteins. However, a mutation in the AIV type PB2 giving an aa substitution E627K has been shown to mediate functional interaction with human host cell ANP32A and ANP32B. The biochemical explanation as to why both the A and B ANP32 analogues are used as co-factors in human cells, but only ANP32A in avian cells may be explained by that both avian ANP32A and human ANP32A and ANP32B have asparagine and aspartate at residues 129 and 130, whereas avian ANP32B has isoleucine and asparagine at the corre-
sponding sites. Moreover, avian type ANP32A has 33 additional aas (after aa 175 and onwards) neighboring the C-terminus, in contrast to the shorter human type ANP32A. Thus, the selection for the PB2 E627K mutation of avian vPol in human cells, as an adaptation to interact with the shorter human ANP32 co-factors. However, sequence analysis of avian ANP32A has revealed that there exist at least three different splice variants of avian ANP32A [60]. Among several investigated species, chickens were reported to express three different splice variants; a full length ANP32A_X1 (33 aa insert), a somewhat truncated isoform ANP32A_X2 (29 aa insert), and a mammalian-like isoform ANP32A_X3 (lacking any aa insert), in the stoichiometric ratio 66:25:9 %. Notably, quail (phylogenetically close to chicken) was reported to express the three splice variants in an approximate stoichiometric ratio of 60:40:0 %. Moreover, gulls were reported to solely express the full length ANP32A_X1 isoform, whereas mallards were reported to express the three isoforms in an approximate stoichiometric ratio of 85:10:5 %. Geese were reported to have a shifted splice variant ratio of approximately 70:5:25 %. On the other hand, among the passerines investigated, swallows were found to have a dominance of the human ANP32A isoform, by the approximate stoichiometric ratio 15:25:60 %. Additionally, ratites are reported to only express mammalian-like ANP32A_X3. Furthermore, ANP32A expression levels were reported to be uniform across tissues and not affected by AIV infection using chicken cell lines. The complexity of ANP32A splice variants distribution across avian species and the interaction with PB2 add further complexity to the mechanisms underlying interspecies transmission and the ecology of AIV. Clearly, presence of human-like ANP32A_X3 in birds allows for existence (or in some cases even selection) for mammalian adapted PB2. Thus, making the step smaller from a (ANP32A_X3 dominant) avian to mammalian cell.

The mixing vessel

It was early acknowledged that despite the broad IAV host acceptance, lineages of IAVs are mainly specific for certain groups of animals (including birds vs. domestic mammals vs. humans etc.) and seldom jump between these groups [56, 57, 65]. Thus, it was postulated that there exists a species barrier that IAV must overcome to be able to transmit across species or higher taxonomic levels. This barrier might be of varying magnitude between different species. Early IAV research on the IAV species barrier focused on NP, and Scholtissek et al. identified that the nucleoprotein (NP) of avian vs. human adapted IAVs are distinct, but IAV isolated from pigs have a NP of either avian or human type and pigs were postulated as a mixing vessel for IAVs [65, 66]. However, during the past 20 years the focus of the species barrier and the molecular explanation to pigs being a mixing vessel for IAV has been attributed to the receptor tropism of avian vs. human IAVs.
and the dual display of avian and human type receptors in the pig respiratory tract (see *The influenza virus receptor* paragraph) [58]. Due to the nature of IAVs having a segmented genome, co-infection of pigs with IAVs of e.g. human and avian origin, allows for *reassortment* (antigenic shift) yielding reassortant progeny virions with new properties [3]. Influenza A viruses commonly do not jump between different classes (e.g. bird to human host) in one step. Rather, this is a stepwise process involving multiple hosts. The nature of IAV having an error prone RdRP yields a distribution of quasi-species progeny virions with various features [13]. If an IAV with certain beneficial properties end up in a susceptible, competent, and naïve new host species, the virus can spread rapidly.
Glycobiology in the Perspective of Influenza A Virus

General introduction to glycans

Glycan is a general term for any molecule (irrespective of size) that either is solely a carbohydrate, or comprises a carbohydrate subdomain [67, 68]. Glycans are both excreted and found intracellularly and can be both soluble and membrane bound (Figure 2). Glycans are essential in e.g. cell-cell signaling and cell recognition. For instance, glycosylation (the addition of a glycan) is the most common type of post-translational modification (PTM) and is often critical for correct protein folding and/or functionality. In contrast to protein synthesis which is deterministic in the sense of one codon giving one aa, glycan synthesis is much more complex because the theoretical number of possible combinations are almost endless. Moreover, synthesis of glycan structures is an *ad hoc* process in that added glycans are often a function of what building blocks are available at the moment and the synthesis is made by complexes of glycosyltransferases that depend on each other. Protein glycosylation commonly occurs via N- (Asn–X–Ser/Thr sequon, where X is any aa except proline) or O-linked (Ser/Thr sequon) glycosylation. In eukaryotes, N-linked glycosylation takes place in the ER lumen where a ready-made leister-shaped glycan bundle is added to the nitrogen atom of the Asn side chain. In humans, the glycan bundle is referred to as a high mannose structure, due to the high content of Man (two basal GlcNAcs and then 1+2 “arms” with 9 Mans). As part of the downstream PTM, the added glycan can be further modified in the Golgi apparatus, where it e.g. can be trimmed down and/or additional other monosaccharides can be added, including additional branching. If only one “leister arm” is modified, the glycan is known as hybrid type, whereas if both “arms” are modified, the glycan is known as complex type. O-linked glycosylation is less explored than N-linked glycosylation but takes place in the Golgi apparatus. The glycosylation pattern is unique between individuals, but general features are shared within one species. This is a very important tool for the body to distinguish between self vs. non-self. There are many pathogens that utilize glycans, both to localize target host cells and to mimic host cells to escape clearance by the immune system; among these pathogens are IAVs [69].
Figure 2. Diversity of vertebrate glycosylation adopted with permission from Moremen, Tiemeyer & Nairn 2012.

Figure 3. N-acetylneuraminic acid (sialic acid).
General introduction to sialic acid

A hallmark of human glycosylation is the addition of Neu5Ac (sialic acid) at the terminal position of complex glycans. Sialic acid is a family of glycans based on the nine carbon Kdn structure [67]. Common types of SA are Neu5Ac and Neu5Gc that are mainly located at the termini of complex glycans in animals. Neu5Gc is mammalian specific, but a few species lack Neu5Gc including humans, ferrets, dogs, and seals [70]. In contrast to other primates, humans have an inactivation in the gene encoding the hydroxylase that converts Neu5Ac to Neu5Gc [71]. Thus, a hallmark of humans is the display of only Neu5Ac at the termini of complex glycans and inability to produce Neu5Gc. Accordingly, in human glycobiology the abbreviations SA and Neu5Ac are often used interchangeably and SA refers to Neu5Ac if not else explicitly stated (Figure 3) [72]. In the molecular world of biochemistry, stereochemistry is essential [68]. Two monosaccharides can be linked via either an α or a β glycosidic bond (ether bridge). Somewhat simplified, if the two monosaccharides that are to be linked are visualized as two planes (Haworth projection), the α glycosidic bond will point below the plane, whereas the β glycosidic bond will point above the plane. However, the formal definition is based on the Rectus/Sinister (R/S) conformation at the stereocentra. An α glycosidic bond is formed when the two stereocentra have the same conformation and a β glycosidic bond is formed when the two stereocentra have different conformations. Sialic acid is commonly bond by α glycosidic bonds in positions 2,3/6 when bond to other monosaccharides and 2,8(/9) when bond to another SA molecule (numbering starts at the anomeric carbon) [67]. When linked to another monosaccharide, SA is commonly linked to Gal that in turn can be part of a longer glycan chain. Based on the linkage conformations, these chains will not only be different in chemical composition, but will be different in space. Additionally, any substitutions (e.g. addition of a sulfur group or fucose) or branching will change the chemical environment. These glycans can in turn be added to a plethora of different macromolecules including proteins and lipids at the cell membrane. Moreover, the glycans at the cell membrane are not homogenously, but heterogeneously distributed over the cell membrane forming glycan patches. Together, the glycans at the surface of the cell membrane form the glycocalyx. The glycocalyx is essential for e.g. cell recognition and determining self vs. non-self. However, many pathogens including IAV have evolved to make use of the glycocalyx for host cell recognition [69].
The influenza virus receptor

The major surface protein of the IAV particle is HA, which is the surface attachment protein that binds to SA on the host cell surface [12]. Influenza HA preferentially bind Neu5Ac, whereas binding to Neu5Gc is very rare [70]. The affinity of association between a single HA–Neu5Ac binding is weak ($K_{\text{diss}} > 0.1$ mM), but as the virus particle surface is covered with HAs and Neu5Ac is ubiquitous on the host cell surface, there is a strong avidity effect [28]. A historical landmark of IAV research was the identification of AIV binding to α2,3-linked Neu5Ac, whereas mammalian adapted IAVs bind α2,6-linked Neu5Ac [56-58]. Since then, acquisition of the ability to bind α2,6-linked Neu5Ac has been regarded as a zoonotic prerequisite for AIV to be able to cross the species barrier to mammals. Additional to the linkage conformation of the terminal Neu5Ac to the penultimate Gal bond, the downstream glycans and their substituents have been shown to play an important role in the HA–Neu5Ac interaction process [73, 74]. Particularly, 3’STF has been postulated as key receptor for AIVs of duck origin, whereas isolates from gulls prefer S-Le$^x$ [73]. Further investigation of gull-specific H13 and H16 isolates has shown that the preferred receptor structure can vary within a single subtype [74]. For example, H13 isolates had maximum binding to either 3’S LN or Su-S-Le$^x$ and H16 isolates had maximum binding to either S-Le$^x$ or Su-3’S LN. On the other hand, human isolates prefer longer sialylated structures including 6’SdiLN and Neu5Acα2,6Galβ1,4GlcNAcβ1,3Galβ1,4(Fucα1,3)GlcNAcβ1, 3Galβ1,4(Fucα1,3)GlcNAc [75, 76]. Moreover, it has been suggested that preferred receptor structure may vary based on host phylogeny and virus subtype. For instance HPAIVs H5 and H7 isolated from poultry have been shown to have tropism for sulfated 3’S LN and S-Le$^x$ respectively [73]. Additionally, further investigation of H7 AIVs has reported shared tropism for sulfated 3’S LN and/or sulfated S-Le$^x$ despite host species of virus isolation [77]. Furthermore, it has been shown that various LPAIVs have an intrinsic property to bind α2,6-linked Neu5Ac, especially H9N2 viruses isolated from members of Galliformes [74, 78, 79]. In a recent study, it was found that IAVs (of avian, human, and swine origin) could bind mannos-6-phosphate (Man6P) containing non-sialylated glycan structures, derived from human lung, in a SA-independent manner [80].

Distribution of influenza virus receptors

It was originally believed that α2,3-linked Neu5Ac was exclusive to birds and α2,6-linked Neu5Ac characteristic to humans, whereas pigs displayed mixed α2,3,6-linked Neu5Ac in their respiratory epithelium [56-58]. However, recent glycomic characterizations of the pig and human respiratory tracts have revealed that humans also display α2,3-linked Neu5Ac and in-
deed, the glycomic repertoire of the human vs. pig respiratory tracts are very similar [78, 79, 81]. In the human respiratory epithelium, there is a dominance of α2,6-linked Neu5Ac in the URT, and a gradual transition to dominance of α2,3-linked Neu5Ac in the LRT. Additionally, pigs display Neu5Gc in their respiratory epithelium and the role of Neu5Gc as potential IAV receptor in pigs has been questioned. Moreover, several studies of tissues from birds have reported the expression of α2,6-linked Neu5Ac in avian URT epithelia in various avian species [38, 82-84]. In conclusion, α2,3 and α2,6-linked Neu5Ac are expressed both in humans and birds.
Wild birds in the perspective of IAV

The vast diversity of wild birds

Birds belong to the class Aves and are the most diverse group of extant tetrapods comprising over 10,000 species [85, 86]. Feathers are a mutual characteristic for all birds that otherwise display great variation in morphology, ecology, and behavior. Despite the partially shared ecology, Anseriformes and Charadriiformes are two distinctive orders with a most recent common ancestor (MRCA) occurring at least 72 million years ago (MYA), and frequently share aquatic habitats, in contrast to many other avian orders that are mainly terrestrial [85, 87]. Mallards are dabbling ducks (*Anas* genus and allies) feeding in shallow surface waters and the single most well described AIV host [12, 22-24, 86]. The closest relatives (MRCA 5 MYA) to dabbling ducks are the diving ducks that instead feed in somewhat deeper waters by diving [85-87]. One of the most common diving ducks widely spread across Eurasia is tufted duck (*Aythya fuligula*) [86, 88]. Despite mallards and tufted ducks partly occupying similar habitats, AIV is reported much more frequently in mallards than tufted ducks [22-24].

Birds are commonly divided as nonpasserines and passerines [89-94]. Passeriformes is the largest avian order comprising more than half of all extant bird species. Surveillance studies have occasionally reported LPAIVs in passerines [39-41], but dabbling ducks and shorebirds are still the major reservoir of AIV [22, 24]. However, other avian taxa (including members of Passeriformes) can potentially act as bridge species/vectors e.g. transmitting AIV to domestic poultry as free-range poultry and passeriform bird species often co-feed [36].
Wild birds as vectors of IAV

Low pathogenic AIV is commonly asymptomatically shed by members of Anseriformes and Charadriiformes [19]. Several of these species undergo medium to long distance migration from tropic/subtropic to boreal regions and back annually, and some species even annually migrate between the Southern and Northern Hemispheres. Thus, there is an annual flow of AIV along avian migratory flyways (commonly originating from the northern breeding grounds and spreading south) [95]. Viruses can be exchanged at stopover sites among migratory birds where the avian migratory flyways intersect, as well as potential virus transmission from migratory birds to the local avifauna. Infected wild birds might then transmit the virus to domestic poultry that enables massive propagation and mutations that eventually allow virus transmission to mammals, including humans [3, 46, 47, 95-97]. In addition to transmission to mammals, domestic poultry mutant viruses with increased virulence and pathogenicity may also transmit back to the wild bird population that can spread the viruses on a large geographical scale and/or causing high morbidity and mortality in wild birds. Thus, it is critical to prevent AIVs to enter large poultry holdings as this in the worst case may generate novel IAVs with pandemic potential.

Immune response to AIV infection in avian hosts

In natural AIV hosts (e.g. dabbling ducks), AIV infection is over within commonly one week, with peak shed virus titers at about 2 d.p.i. [29, 30]. Thus, acquired adaptive immunity will protect from reinfection with the same or antigenically related AIVs [98]. However, the adaptive immune response is too slow to combat AIV infection in an immunologically naïve bird. Rather, it is the innate immune response that fights infection of novel AIVs [16]. Especially, retinoic acid-inducible gene I protein (RIG-I, also known as DDX58) has been shown as a major host cell response to AIV infection [99]. RIG-I is a cytoplasmic pattern recognition receptor sensing virus replication RNA products [16]. RNA binding by RIG-I induces a conformational change that exposes the caspase recruitment domain (CARD) allowing interaction with mitochondrial antiviral signaling protein (MAVS), leading to the expression of type I (e.g. INF-α and INF-β) and III (e.g. INF-λ) interferons. Produced interferons are secreted and bound by the infected and neighboring cells. This induces a signaling cascade that triggers the expression of a plethora of interferon-stimulated genes (ISGs) that provide antiviral effects. Additionally, interferon pathway sensors (including RIG-I and tripartite motif-containing protein 25 [TRIM25]) can directly bind AIV vRNPs, thereby blocking their nuclear entry or RNA synthesis. RIG-I has been shown to be active in several avian families including ducks, geese, and pigeons. However, chickens do not have a functional RIG-I homologue and
vRNPs are thus not targeted by RIG-I in chickens. Experimental studies of mallards have shown that RIG-I expression levels are elevated within the first 24 h post infection, but return to baseline after 48 h [99]. Moreover, another host interferon-induced factor targeting AIV vRNPs is interferon-regulated resistance GTP-binding protein MxA (also known as Mx1) that can inhibit vRNP nuclear import [16]. However, the Mx proteins in chickens are reported to not restrict AIV replication. Furthermore, there are many additional host cell factors interfering with the AIV vRNPs (including interferon-induced transmembrane protein 3 [IFITM3], which prevents the fusion of the virus envelope with the endosomal membrane, and thus inhibiting vRNP release) limiting vRNP nuclear import and restricting early stages of AIV infection. To combat the extensive antiviral (mainly interferon-induced) host cell response to AIV infection, AIVs have evolved interferon antagonists within the AIV genome, particularly NS1. NS1 can act through several mechanisms, including (1) binding to RIG-I and hinder RIG-I activation, and (2) binding to cleavage and polyadenylation specific factor 30 kDa subunit (CPSF30 [also known as CPSF4]) preventing the processing of newly synthesized mRNAs, including interferon-β mRNAs. In conclusion, there is a continuous arms race between host innate antivirus response factors (including RIG-I) and interferon-induced antivirus responses vs. the virus antagonistic anti-host cell response factors (including NS1).
In summary

Influenza virus continues to be a pathogen of global concern annually affecting humans, livestock and wildlife [19, 50, 100]. It was early recognized that the natural reservoir of the AIVs is wild waterfowl and the pathogenesis and virus ecology is well described in mallards (the most common AIV host species) [29, 30]. Later it was discovered that gulls harbor two specific AIV subtypes not reported from ducks and it has recently been proposed to widen the perspective of AIV research to better understand the ecology of AIVs [19, 36]. Despite the well-characterized pathogenesis in mallards, a lot remains unknown about AIV susceptibility and properties in other avian species, including closely related duck species. Pilot studies have indicated that AIVs might have different properties in non-mallard hosts in contrast to mallards [33, 101]. Similarly, IAVs where initially described as either α2,3-linked SA (avian type) or α2,6-linked SA (human type) binders and the pig was suggested as a mixing vessel for IAV due to its dual display of both α2,3 and α2,6-linked SA in its epithelium [56-58]. However, more recent studies have revealed that the receptor tropism is much more complex than initially thought and recent glycomic analysis of the pig and human respiratory tracts has found limited differences in the receptor repertoire of the human vs. pig respiratory tracts [78, 79]. Moreover, studies of the available receptors in birds using lectins have revealed shared structures between the human and avian respiratory tract, adding further complexity and question to the postulate of the linkage conformation of the terminal SA to penultimate Gal as the major determinant of the postulated species barrier [38, 82-84]. Yet, a comprehensive characterization of the available receptor structures in birds is still missing, which will be essential to fully understand what receptor structures are of biological relevance. Finally, a hallmark of IAV is the ability to jump between different host species, which ultimately may cause pandemics [3, 4]. Essential in this process is the transmission of AIVs to domestic poultry. Within this thesis, interspecies transmission of AIV to both ducks and chickens was studied to investigate potential barriers (both virus and host related) for interspecies transmission, a critical step in IAV ecology.
Aims

General aim
The general aim of this thesis was to investigate bases and barriers for inter-species transmission of IAVs between different bird species and mammals with focus on the virus–host cell attachment process.

Specific aims

Study I
Receptor incompatibility has been suggested as a species barrier that usually prevents IAV from transmitting between distantly related species. Pigs have been suggested as a mixing vessel facilitating the spread of IAVs from birds to humans. Yet, occasionally AIVs transmit directly from avian to human hosts. The aim of this study was to explore any potential differences in the pattern of virus attachment (PVA) to human and pig tissues between IAV isolates of avian and human origin and to study how such patterns relate to the receptor tropism of the different viruses.

Study II
Long-term surveillance studies have identified Anseriformes and Charadriiformes as the main reservoirs of AIVs. Yet, AIVs have been isolated from other avian orders and it is of interest to widen the focus of AIV host species research to better understand its zoonotic ecology. The aim of this study was to identify avian species potentially susceptible to AIV by the degree of virus attachment in target organs. In addition, we wanted to reveal potential differences in PVA to respiratory vs. intestinal tracts across investigated species and their geographical origin.
Study III

Avian influenza virus is commonly found in mallards, but seldom reported from tufted ducks, despite these two avian species partly occupying the same habitat. Additionally, AIV is a huge burden for the poultry industry and poultry can serve as an intermediate host for spread of AIV from wild birds to humans. The aim of this study was to investigate potential host and/or virus related barriers for interspecies transmission of mallard origin LPAIVs to tufted ducks and chickens.

Study IV

There have been many attempts to characterize AIV tissue attachment in wild birds and AIV receptor tropism, but these studies commonly include just a few virus isolates/subtypes. The aim of this study was to investigate receptor tropism and tissue attachment pattern of representatives of all HA subtypes (H1–16) known to circulate in wild birds, and how the sixteen different HA relate to each other in terms of their receptor tropism.
Methods

Ethics statement

Study I
Handling of human tissues was approved by the Regional board of the ethical committee at Uppsala University (reference #2002-577). Human tissue donors had all given informed consent for non-commercial research use of tissue biopsies.

Study II
Wild bird tissue sampling procedures were approved by the Swedish Environmental Protection Agency (permit numbers 412-6267-08NV/412-5977-08NV), the Swedish Board of Agriculture (permit numbers 74-08/43-09), the Chilean Agriculture Ministry (permit number 1-25-2008), and the Ethics Committee of the Veterinary University of Concepción (permit number CE1-2006).

Study III
All animal experiments were approved by the Uppsala regional board of the animal ethics committee, Sweden (permission number 5.8.18-07998/2017). All animal experiments were conducted in BSL2 animal facilities at the Swedish National Veterinary Institute.

Study IV
Paraffin embedded tissue blocks were kindly donated by the Erasmus Medical Center, Rotterdam, Netherlands. The handling of avian tissues was approved by the Dutch Animal Ethical Committee.

All experimental methods were performed in accordance with relevant guidelines and regulations.
In vitro Virus histochemistry

Virus histochemistry is a convenient technique to assess virus attachment to tissues. Tissue specimens of interest were formalin fixed and paraffin embedded [102, 103]. Embedded tissue paraffin blocks were then cut (commonly 4–5 μm thick) and mounted on glass microscopy slides. Investigated viruses were FITC-labeled to allow for detection with the same primary antibody across virus subtypes [102]. In study I and II two AIVs isolates from mallards (Anseriformes; Anas platyrhynchos) 1. A/Mallard/Sweden/68619/2007 (H3N2) and 2. A/Mallard/Sweden/81/2002 (H6N2), one isolate from ruddy turnstone (Charadriiformes; Arenaria interpres) A/Turnstone/Delaware/15/2007 (H12N5) and one isolate from black-headed gull (Charadriiformes; Chroicocephalus ridibundus) A/Black-headed gull/Sweden/2/1999 (H16N3) were investigated. Additionally, a human seasonal IAV of the same subtype combination as one of the mallard-derived viruses were studied A/Netherlands/213/2003 (H3N2). In study I, human tissue specimens (examined by a pathologist as healthy non-malignant tissue) were obtained from the Department of Pathology at the Uppsala University Hospital, Sweden, directed by the Uppsala Biobank and pig tissues were kindly donated by a small abattoir in central Sweden and formalin fixed at slaughter. In study II, to include tissue material from bird species from both the New and Old Worlds, trachea and colon (to represent respiratory and intestinal tracts) samples were collected in Chile and Sweden. All sampled individuals were adult birds. The birds were not tested for AIV prior to sample collection. In brief, all formalin fixed tissues were mounted and baked for 45 minutes at 60 °C prior to virus histochemistry. Tissue slides were deparaffinized with xylene, hydrated in graded alcohols to distilled water, and blocked for endogenous peroxidases in 0.3 % hydrogen peroxide. Each tissue section was incubated over night at 4 °C with 50 hemagglutinin units of purified formalin fixed FITC-labeled IAV or PBS (Medicago AB, Uppsala, Sweden) as negative control. FITC-labeled viruses were detected by a peroxidase labeled α-FITC rabbit polyclonal antibody (#ab19492, Abcam). The signal was amplified by a tyramide signal amplification kit (PerkinElmer AB, Upplands-Väsby, Sweden). Peroxidase signal was revealed with 3-amino-9-ethyl-carbazole (Sigma-Aldrich AB, Stockholm, Sweden). Tissues were counterstained with hematoxylin (Sigma-Aldrich), mounted with vision mount (Thermo Scientific) and scanned using Aperio Image Scope (Aperio Technologies, CA., U.S.A. using a 20x (human and pig slides) or 40x (bird slides) objective and scored by two independent observers. To visualize the linkage conformation of SA in the investigated tissues, each tissue type was stained with MAA-II binding to α2,3-linked SA and SNA binding to α2,6-linked SA. Tissue slides were prepared similarly as
described for the virus histochemistry. Tissue slides were stained with either 4 μg/mL MAA-II (BioNordika AB, Stockholm, Sweden) or 2 μg/mL SNA (BioNordika) from Vector Laboratories. Lectin stained tissues were counterstained, mounted, scanned, and scored following the same procedure as described for the virus histochemistry.

**Tissue micro array**

Histochemistry stainings are commonly done using full sections on microscopy glass slides. However, this is a tedious process with low throughput and difficult to scale up. Tissue micro array (TMA) is a miniaturized version of traditional full section histochemistry that allows for high throughput [103]. In TMA technology tissue samples are collected, embedded, and reviewed according to the same procedure as for traditional histology. Embedded blocks are cut in full section format and reviewed to localize the cell type(s) of interest and the areas of interest of the blocks are marked. The subarea of interest of each block is then punched with a puncher collecting a 1 mm diameter core from the block. This core is then transferred to a new empty block were collected cores are aligned into a micro array format according to a defined template, now comprising a new block that instead of one tissue specimen constitutes an array of aligned cores. This block is then cut, mounted, and stained following the same procedure as for full section histology staining. This allows for mounting of n rows × m columns (12×10 cores used in this thesis) in one block. Thus, the need of individual microscopy slides is greatly reduced. Multiple cores originating from the same tissue block or related samples can be incorporated in the same TMA to allow for replicate samples.

**Glycan array**

It was early recognized that IAV uses SA as receptor for host cell attachment and that different kinds of SAs were utilized depending on the type of virus [56-58, 104, 105]. Especially, it has been shown that the glycan structure downstream of the Neu5Ac–Gal in the vicinity of the terminus creates an environmental effect affecting the HA–Neu5Ac interaction [73-75]. Synthetic glycan structures can be printed in a defined pattern on glass microscopy slides [106]. A glycan array including different α2,3 and α2,6-linked sialylated glycans, as well as non-sialylated glycans was used to investigate virus glycan attachment. In brief, glass slides with immobilized glycans (each structure spotted in triplicate) were blocked with 50 mM sodiumtetraborate and 84 mM HCl (Sigma-Aldrich A/S, Brøndby, Denmark) followed by incubation on shake in a humidity chamber with 20-500 hemagglutinin units IAVs together with 10 μM Oseltamivir carboxylate (Sigma-Aldrich) per well at room temperature for one hour, reagents were diluted in 1×PBS-T (Sigma-Aldrich). After washing with 1×PBS three times, bound viruses were
detected by α-FITC rabbit polyclonal antibody (Abcam) (1 μg/mL) in combination with Alexa555-labeled α-rabbit polyclonal antibody (1 μg/mL) (#150078, Abcam). Negative controls without any virus were included for determination of any unspecific signal. The fluorescence signal was measured by a Scan Array G, Microarray Scanner (PerkinElmer A/S, Skovlunde, Denmark) and analyzed by ProScanArray Express Version 4.0 (PerkinElmer).

Epithelial cell isolation and sialome characterization

**Determination of Neu5Ac–Gal linkage stereochemistry**

Essential for the virus replication process is the initial host cell recognition and attachment [12]. However, available receptors vary between different tissues and hosts [3, 58]. Lectin studies are limited to the linkage conformation of Neu5Ac bond to the penultimate Gal, but complete glycan structures (Neu5Ac–Gal and downstream) can be characterized by reversed-phase liquid chromatography tandem mass spectrometry (LC-MSⁿ) analysis [78, 79]. Liquid chromatography-MSⁿ is a powerful and versatile technique where molecules are ionized and separated by their mass to charge ratios, and their structure can be deduced from their (iterated) fragmentation patterns [107-110]. Thereby, it is possible to determine the glycan structure, the site-specific attachment site of the glycopeptide and identity of the protein, all extracted from the same molecule, and include the possibility to perform relative quantification of co-occurring glycoforms. In general, such glycoproteomic strategies are based on 1) protease digestion of protein samples (generally using trypsin, which cleaves peptides on the carboxyl side of arginine or lysine), 2) the enrichment of glycopeptides from the vast amount of non-glycosylated peptides and 3) reversed-phase LC-MSⁿ. In LC-MSⁿ, eluting precursor ions are subjected to fragmentation, often with the use of higher-energy collision dissociation (HCD) to provide a stepwise decomposition of both the peptide backbone and the monosaccharide residues of the glycan for sequence determinations of both in protein/glycan database searches. One drawback is that the determination of saccharide identities and glycosidic linkage positions is not normally possible in LC-MSⁿ, for instance a mass sequence of 291 u + 162 u + 203 u + 1 u (m/z 657) may be composed of Neu5Acα2,3Galβ1-3GalNAc+H⁺ or Neu5Acα2,6Galβ1,4GlcNAc+H⁺ or both. However, with the use of HCD, glycan derived saccharide oxonium ions are very prominent in the MS/MS (MS²) spectra, and the relative abundances of critical oxonium ions can be used to distinguish GalNAc from GlcNAc [109]. Lately it is now possible to distinguish Neu5Acα2,3 vs. Neu5Acα2,6 using Equation 1, which is based on the LacNAc-derived ions over Neu5Ac-derived ions times a stoichiometric compensatory factor based
on the number of Neu5Ac vs. HexNAc residues [111]. The ratio is 0.4–0.6 for Neu5Aca2,3 and 0.8–1.5 for Neu5Aca2–6 terminated glycopeptides.

\[
\frac{\ln n}{Nn} = \frac{I_{204} + I_{366}}{I_{274} + I_{292}} \times \frac{n(\text{Neu5Ac})_{\text{GalGlcNAc}} + 0.5n(\text{Neu5Ac})_{\text{GalGalNAc}}}{n(\text{GlcNAc}) + 0.5n(\text{GalNAc})}
\]

*Equation 1. The formula for Neu5Ac stereochemistry determination. I = intensity and n = number of recorded ions. m/z ratios 204 – GlcNAc; 274 – Neu5Ac–H2O; 292 – Neu5Ac; 366 – GlcNAc-Gal.*

**Isolation of epithelial cells**

Biologically available receptor structures were determined by post mortem sampling of trachea, lung, ileum, and colon from healthy uninfected chickens (*Gallus gallus domesticus*), mallards (*Anas platyrhynchos*), and tufted ducks (*Aythya fuligula*) (five individuals/species). The necropsies were rinsed in cold PBS upon collection to remove any debris. Adipose tissue, blood vessels, etc. were removed and the intestinal specimens were cut open and intestinal content and mucus was carefully removed. Cleaned samples were incubated at 37 °C for 1 h in 10 mL DPBS with 1 mM DTT and 3 mM EDTA. After incubation, the liquid was carefully aspirated and replaced with PBS. The samples were vortexed for 5×20 s and the tissues were then removed. The remaining cell slurries were spun at 1000 g for 5 min at 4 °C and the supernatant was discarded. The pellets were snap-frozen in liquid nitrogen and stored at −80 °C until LC-MS analysis.

**Protein digestion, enrichment and fractionation**

The samples were homogenized using the lysis matrix D on FastPrep®-24 instrument (MP Biomedicals, OH, U.S.A.) in lysis buffer (50 mM triethylammonium bicarbonate (TEAB), 2% sodium dodecyl sulfate (SDS)) and 5 cycles 40 s each. The samples were centrifuged at maximum speed for 15 min, and supernatants harvested. The protein concentration was determined using Pierce™ BCA Protein Assay (Thermo Scientific, Waltham, MA, U.S.A) and the Benchmark Plus microplate reader (Bio-Rad, Hercules, CA, U.S.A.) with BSA solutions as standards.

Sample aliquots (500–1000 µg) were trypsin digested using the filter-aided sample preparation (FASP) according to Wisniewski et al. [112] with small modifications. Briefly, samples were reduced with 100 mM dithiothreitol at 60 °C for 30 min, spin-filtered (30 kDa MWCO Pall Nanosep centrifugation filters, Sigma-Aldrich), washed repeatedly with 8 M urea and followed by digestion buffer (1% sodium deoxycholate (SDC) in 50 mM TEAB) prior to alkylation with 10 mM methyl methanethiosulfonate in digestion buffer for 20 min. Protein digestion was performed in digestion buffer by addition of 5 µg Pierce MS grade Trypsin (Thermo Fisher Scientific) at 37 °C overnight,
followed by an additional 2 hour incubation with a new trypsin addition the consecutive day. Peptides were isolated by centrifugation and SDC was removed by acidification with 10 % trifluoroacetic acid. Glycopeptides were enriched with hydrophilic interaction liquid chromatography (HILIC) according to Parker et al. [113], with slight modifications. In short, peptides were loaded onto an in-house zwitterionic Zic-HILIC SPE cartridge containing 20 mg of Zic-HILIC particles (10 μm, 200 Å; Sequant/Merck, Burlington, MA, U.S.A.). The flow-through was collected and recirculated through the column an additional three times. The column was washed with totally 1.2 mL of 80 % (v/v) acetonitrile and 1 % (v/v) trifluoroacetic acid. Enriched glycopeptides were eluted with 4 times 50 μL 0.1 % (v/v) trifluoroacetic acid followed by 50 μL of 25 mM NH₄HCO₃ and finally 50 μL of 50 % (v/v) acetonitrile and dried by vacuum centrifugation. The samples were fractionated into six fractions (5-17.5 % acetonitrile in 0.1% trimethylamine), using Pierce High pH Reversed-Phase Peptide Fractionation Kit (Thermo Scientific) according to the manufacturer’s protocol and pooled into 3 samples. Pooled samples were dried in a vacuum centrifuge and reconstituted in 15 μL of 3 % acetonitrile, 0.1 % formic acid for LC-MS analysis.

**NanoLC MS analysis**

Peptide samples were analyzed on an Orbitrap Fusion Tribrid mass spectrometer interfaced with Easy-nLC1200 liquid chromatography system (Thermo Fisher Scientific). Peptides were trapped on an Acclaim Pepmap 100 C18 trap column (100 μm x 2 cm, particle size 5 μm, Thermo Fisher Scientific) and separated on an in-house packed analytical column (75 μm x 30 cm, particle size 3 μm, Reprosil-Pur C18, Dr. Maisch) using a linear gradient (solvent A; 0.2 % formic acid in water and solvent B; 80 % acetonitrile, 0.2 % formic acid in water) from 7 % to 35 % B over 45 min followed by an increase to 100% B in 5 min, and finally 100% B for 10 min at a flow of 300 nL/min. MS¹ scans were performed at 120 000 resolution, m/z range 600-2000, the most abundant double or multiply charged precursors from the MS¹ scans were selected with a duty cycle of 3 s, isolated with a 3 Da window, fragmented via higher-energy collision induced dissociation (HCD) at the energy setting of 30 with a maximum injection time of 118 ms, and the MS² spectra were detected in the Orbitrap at 30 000 resolution. Dynamic exclusion was enabled with 10 ppm tolerance and 10 s duration. For some selected samples representing glycan structures with sialic acid, MS³ was performed on the m/z 657.23 ion after HCD 20 using a second energy of HCD 20 and detection in the ion trap.

**Glycoproteomic analysis**

The LC-MS/MS raw files were analyzed with the Byonic software (Protein metrics, Cupertino, CA, U.S.A.) using a modified list of the glycan modifications “182 human N-glycans” and “6 most common O-glycans”. For in-
stance, glycoforms with Hex$_4$HexNAc$_5$ core structure, differing from the normal Hex$_3$HexNAc$_4$ complex biantennary core, was added. A Byonic score cut off >300 was used for glycopeptide hits and Neu5Ac containing hits were manually verified with the following inclusion criteria 1) presence of the correct peptide+HexNAc ion with respect to the precursor mass and identified glycan mass; 2) presence of the Neu5Ac oxonium ions m/z 274 and m/z 292; and 3) for hits including fucose, presence of peptide+HexNAc+dHex ion for the identification of a core fucose, and/or presence of m/z 512 ion (HexHexNAcdHex) for identification of Fuc on the antennae. Further Byonic search criteria included: the FASTA database was *Anas platyrhynchos* (Organism ID 8840; 27089 sequences) and *Gallus gallus* (Organism ID 9031); C-terminal cleavage after lysine and arginine; accuracy for MS$^1$ was 10 ppm and for MS$^2$ it was 20 ppm; static modification was methylthio on cysteine (+45.9877 u) and variable modification, apart from glycans, was methionine oxidation (+15.9949 u). Extracted ion chromatograms were traced at diagnostic MS$^2$ ions including chosen peptide+HexNAc ions to identify additional glycoforms sharing the same peptide and saccharide oxonium ions for instance m/z 274.09 for tracing Neu5Ac; m/z 290.09 for tracing Neu5Gc and m/z 495.18 for tracing Neu5AcHexNAc. Tufted duck glycopeptides were identified using the mallard genome as template, due to non-available tufted duck genome.

**Virus quantification by RT-qPCR**

Viral RNA was extracted from the collected swab samples from chickens and tufted ducks using the Maxwell 16 Viral Total Nucleic Acid Purification Kit together with a Maxwell 16 extraction robot (Promega, Nacka, Sweden). Quantification of virus titer was made relative to the mean of standard series dilutions of the eight LPAIV isolates used to prepare the bird inocula and expressed as egg infectious dose 50 (EID$_{50}$) equivalents. Analysis of swab samples was done using a reverse transcriptase real-time polymerase chain reaction (RT-qPCR) targeting the IAV matrix gene and using the AgPath-ID one-step RT-PCR reagent kit (Thermo Fisher Scientific) using a Bio-Rad CFX96 machine (Bio-Rad) [114, 115].

**Whole genome sequence and variant(s) detection**

Total viral RNA was extracted from the virus inocula, as well as all organs with a Cq value < 30 and positive swabs at 3 d.p.i. with a Cq value < 30 using QIAamp viral RNA Mini kit (Qiagen, Hilden, Germany) according to the kit protocol. Virus whole-genome amplification was made using an RT-PCR protocol and two primer sets, as described previously by Pohlmann *et al.* 2018 [116]. Sample quality control was made using Qubit 4 Fluorometer (using Qubit™ dsDNA BR assay Kit) and Agilent 2200 TapeStation (using
Agilent D1000 ScreenTape assay kit). Library construction and sequencing were made using an Illumina platform (PE150) at Novogene (Beijing, China).

For variant(s) detection, generated reads were analyzed using Geneious Prime work package (Biomatters, Auckland, New Zealand) with the following pipeline. Initially, the primer sequences were trimmed off from the raw reads using the “Trim Ends” Geneious Prime plugin. Further, the trimmed reads were mapped against the whole genome sequence obtained from the used homologous inoculum, and a consensus sequence was generated for each sample. Finally, SNP (single nucleotide polymorphism) analysis was conducted using “variation/SNPs” plugin implemented in Geneious Prime work package on the assembled contigs with a maximum variant P-value of $10^{-6}$.

**In vivo**

The animal model system

**Birds**

To investigate the infectivity of mallard origin LPAIVs in tufted ducks and white leghorn chickens (representing domestic poultry), captive tufted ducks obtained from a commercial breeder (Snavelhof, Veeningen, the Netherlands) and white leghorn chickens were inoculated with a panel comprising eight different mallard origin LPAIV isolates in a controlled BSL2 vivarium. Prior to experimental start, all birds were serologically tested to ensure that they were IAV naïve by an α-nucleoprotein antibody ELISA (Idexx, Hoofddorp, Netherlands) according to the manufacturer’s instructions.

**Viruses**

All viruses used for the inoculation of experimental birds were isolates from the Ottenby Bird Observatory IAV repository at the Linnaeus University. Isolates were selected on HA phylogeny to include representatives from the five main avian HA clades [98]. Selected HA subtypes were: H3, H4, H6, H8, H9, H10, H11, and H15. Isolates were grown in specific pathogen free embryonated chicken eggs according to standard procedures.

**Pathogenicity and transmissibility in chickens and tufted ducks**

Four birds were inoculated oculonasally and four birds were inoculated intraesophageally (same experimental setup for both chickens and tufted ducks) with one of the eight studied LPAIV isolates (inoculum $10^6$ EID$_{50}$/) or PBS as negative control. The birds were monitored daily from 0–3 days post infection (d.p.i.) for clinical signs and mortality. All birds were sampled
daily by oropharyngeal (OP) and cloacal (CL) swabs to study any virus shedding via the respiratory or digestive tract. At 3 d.p.i. the birds were euthanized and necropsies were collected from lung, spleen and colon to study virus tissue distribution.

**In silico**

Data science toolbox

A number of open source tools were used for data analysis all written in R language implemented by RStudio [117, 118]. Especially, data handling was made by magrittr and the tidyverse [119, 120]. Data visualization was made by ggplot2, pheatmap, heatmaply, and plotly [121-129].
Results and Discussion

Study I

A species barrier has been postulated to prevent interspecies transmission of IAV between distant taxa [58, 65]. Yet, occasionally AIVs transmit directly from avian to human hosts [130-134]. It was thus of interest to further investigate if and how AIVs attach to human vs. pig tissues. In study I, an AIV panel constituting four isolates (mallard H3N2, mallard H6N1, ruddy turnstone H12N5, and black-headed gull H16N3) were characterized in their PVA to a panel of human and pig tissues with a seasonal human H3N2 isolate as reference (Figure 4 and 5). The panel included human eye (conjunctiva and cornea), nasopharynx, bronchus, lung, and colon and pig trachea, bronchiole, and lung tissues. Additionally, the SA distribution in these tissues was investigated by the use of the lectins MAA-II and SNA. Moreover, the receptor tropism of the viruses was characterized using a glycan array.

Figure 4. Example of histochemistry stainings of human tissues. Each column represents staining with one staining agent, red color indicates positive staining. The tissues were counterstained using hematoxylin (blue).
The most important observation was that AIV attachment was more widespread in human than in pig tissues (Figure 4 and 5). Interestingly, presence of Neu5Acα2,3-Gal could not be detected in the pig tissues using MAA-II. The studied tissue and virus panel was a small sample set and potential heterogeneity in glycosylation pattern between different breeds of domestic pigs might add further complexity to the issue. Still, the reported results provide an interesting indication that AIV might not require passage through pigs to acquire adaptations to human host cell receptors. Indeed, glycomic analyses using LC-MS^n have reported that in terms of glycans expressed in the respiratory tract, human and pigs are highly similar [78, 79]. Yet, based on epidemiological data AIV transmission from bird to human is still rare [130-134]. Histochemistry only provides information about host cell attachment, which is an essential step in the infection process, but successful infection, replication, and transmission is dependent on several downstream factors apart from attachment. Most likely, as suggested by Scholtissek et al. already in 1985, host cell specificity is governed by more than a single factor [65].

In terms of glycan receptor tropism, the human seasonal H3N2 isolate had a “classical human” receptor tropism only binding to α2,6-linked Neu5Ac and the AIVs had a “typical avian” pattern mainly binding to α2,3-linked SA. However, several of the investigated AIVs also bond α2,6-linked Neu5Ac, which could theoretically explain the extensive attachment observed to the investigated human tissues. These viruses showed medium to very strong binding to particular α2,6-linked Neu5Ac structures. Clearly, classifying AIVs as limited to only binding α2,3-linked Neu5Ac and stating α2,6-linked Neu5Ac as a hallmark of humans is an oversimplification, indicating the complexity of the species barrier.
Study II

Most surveillance studies and animal model systems of AIV have been based on mallards [12, 19, 22-24, 29, 30]. However, it is well known that AIV infect several distinct taxa of wild birds. Given the vast diversity of wild birds, it would be unexpected if AIV behaves identical in all wild bird host species. Indeed, pilot studies have reported varying PVA in non-mallard species [33, 82, 101]. The aim of this study was to expand the histochemistry analysis from just a few bird species to a larger panel comprising both “classical” and “non-classical” AIV hosts. In this study, four AIVs (mallard H3N2, mallard H6N1, ruddy turnstone H12N5, and black-headed gull H16N3) and one human seasonal H3N2 isolate were characterized in their PVA to trachea and colon from 26 different bird species. Additionally, the SA distribution in the tissues was assessed using the lectins MAA-II and SNA. In general, there was more virus attachment to trachea than colon of the birds, suggesting the importance of trachea and URT as target organs for AIV replication in non-mallard species (Figure 6). This finding corroborates findings of earlier smaller studies reporting increased attachment to trachea in non-mallard species [33, 101]. Additionally, extensive attachment to tissues from chicken and mallard corroborate the importance of these two species in the ecology of AIV. Moreover, extensive virus attachment was observed to several “non-classical” AIV host species, including terrestrial members of Passeriformes, indicating that these bird species have the potential to act as bridge vectors between wild aquatic birds and domestic poultry. Especially, maximum staining with all staining reagents (including human H3N2 virus) was observed to the trachea of European robin (Erithacus rubecula). European robin is a common ground dwelling bird in rural, peri-urban, and urban environments. This suggests that European robin (or related species in other parts of the world that occupies a similar niche) could be an important bridge species, as this species commonly come in contact with both humans, as well as other wild bird species, and potentially free-range poultry as well. Once again highlighting the importance of IAV sampling schemes beyond the traditionally surveyed aquatic species.
Figure 6. Examples of histochemistry stainings of avian epithelia. Each column represents staining with one staining agent, red color indicates positive staining. The tissues were counterstained using hematoxylin (blue). A. trachea; B. colon.
Study III

The natural reservoir of AIV is wild birds, or more precisely wild aquatic birds [12, 19]. Yet, there is a large difference in AIV prevalence between different aquatic avian taxa [22-24]. On the other hand, transmission of AIV to poultry has a massive negative effect on poultry industry [50]. Moreover, transmission of AIV from wild birds to poultry may ultimately lead to the generation of novel IAV human pandemics [3, 4, 46]. Therefore, to study mechanisms of AIV interspecies transmission a virus panel comprising eight phylogenetically distinct mallard-derived LPAIVs (H3N8, H4N6, H6N2, H8N4, H9N2, H10N1, H11N9, and H15N5) were inoculated in chickens (representing domestic poultry) and tufted ducks (representing diving ducks) to study any establishment of infection (Figure 7). Birds were either inoculated oculonasally or intraesophageally. In wild aquatic birds, AIV is considered to transmit via the fecal–oral route, which has been clearly demonstrated for mallards [30]. However, the results of this study clearly demonstrated that for chickens and tufted ducks, the oculonasal route of infection is required for successful infection as this route yielded infection by several of the studied LPAIV isolates in both chickens and tufted ducks. In contrast, all intraesophageally inoculated chickens were negative and only two intraesophageally inoculated tufted duck individuals were positive in a single CL sample (H8N4 1 d.p.i. and H11N9 2 d.p.i.).

In oculonasally-inoculated chickens, AIV was detected in OP swabs of all subtypes, but only H3N8, H4N6, H6N2, H10N1, and H11N9 viruses could be detected from CL swabs. Virus could be detected by RT-qPCR in tissue samples from H6N2, H8N4, H11N9, and H15N5-inoculated chickens. The highest organ titer was measured in colon from H6N2-inoculated chickens (7.27 EID₅₀ equivalents), but positive samples were also obtained from lung and spleen. This indicates not only local replication at the site of inoculation, but systemic spread. Indeed, the H6N2 virus was also the virus that established the most prominent infection in chickens as measured by the OP swab samples.

In tufted ducks, all oculonasally-inoculated viruses were detected in OP swab samples. However, the H6N2 virus was the only virus that was detect-
Inoculated H6N2, H8N4, H9N2, H10N1, and H11N9 viruses were all detected by RT-qPCR in tufted duck spleen samples indicating systemic spread. The H6N2 virus was the only virus positive in all three organs (lung, spleen, colon), with the maximum titer in spleen (4.76 EID₅₀ equivalents). Hence, of the studied eight different mallard-derived LPAIVs H6N2 was the most potent virus in both chickens and tufted ducks in terms of establishing infection and yielding high virus titers.

H9N2 AIV is endemic in chickens in several parts of the world. It was thus unexpected that hardly any virus was detected from the H9N2 inoculated chickens. To address this, a phylogenetic analysis of the HA gene from the investigated mallard-derived H9N2 LPAIV and 681 H9N2 isolates from domestic fowl and wild birds was conducted. The analysis revealed that wild type H9 isolated from wild waterfowl (including the studied mallard-derived H9N2 LPAIV) were phylogenetically distinct from chicken endemic H9N2 isolates. The SNPs recorded were spread across the complete HA gene, suggesting functional and immunological differences between these wild type wild waterfowl H9 vs. chicken endemic H9 isolates.

Putative viral genetic alterations during virus infection of chickens and tufted ducks were investigated in swab and tissue samples obtained from the H3N8, H6N1, H8N4, H9N2, and H10N1 groups. Whole genome sequences from the positive samples were compared to sequences from the corresponding virus inocula. A number of synonymous and non-synonymous SNPs were observed in viruses recovered from different groups compared to the homologous virus inocula, but no deletions/insertions were recorded. Viruses obtained from the OP swabs from the H3N8 chicken inoculated group revealed three SNPs with one non-synonymous mutation at the PB2 gene segment (aa substitution mutation T16N. In the H6N2 inoculated chicken group, a common aa variation (V63L) in the PA was observed in viruses obtained from colon, spleen, and lung. Additionally, the same aa substitution was recorded in H6N2 viruses recovered from CL swabs at 3 d.p.i. from tufted ducks. The HA gene segment obtained from the virus recovered from the OP swab/3d.p.i./H6 showed a non-synonymous mutation at position 1300 (aa: D434N). Moreover, many additional SNPs were seen in the M gene segment in viruses recovered from Ck/H6/spleen, TD/H6/colon, and TD/H6/ spleen. Previously described amino acid variations, associated with host specificity, were detected in the NP protein (N319K) of TD/H6/spleen and the HA protein (N183H) of the TD/H6/ CL3d.p.i. Further, in the H8N4 tufted ducks inoculated group, the virus recovered from TD/H8/3OP3d.p.i. revealed several SNPs in its polymerase gene segments (PB2, PB1, and PA) and the HA. Finally, the viruses recovered from H9N2 and H10N1 inoculated groups showed no SNPs among their gene segments. In summary, rather few non-synonymous SNPs were detected and it is interesting that for some
of the sampled viruses (including H9N2 and H10N1) no SNPs were detected. The phenotypic effect of the non-synonymous SNPs detected requires further functional tests.

It has earlier been postulated that [6Su]-3'SLN is a receptor for chicken adapted H5N1 viruses [73]. Indeed, several of the studied mallard-derived LPAIVs that established infection in chickens did bind [6Su]-3’SLN, including the studied H6N2 virus. However, binding to [6Su]-3’SLN was not predictive for successful establishment of infection in chickens, as for example the H9N2 virus also bound strongly to this receptor structure, but hardly established any infection in the inoculated chickens. On the other hand, the H10N1 and H15N5 viruses had much more restricted glycan binding profiles, in contrast to the other studied viruses. Yet, these viruses could be detected in OP swab samples from both chickens and tufted ducks for all three experimental days and one chicken CL swab sample was positive for H10N1 virus at 2 d.p.i. Additionally, the H10N1 virus was detected in tufted duck spleen samples and the H15N5 virus was detected in chicken lung samples, indicating that the virus had spread from the site of inoculation to other organs.

None of the studied LPAIVs had tropism for Neu5Gc in the glycan array, which is in accordance with reports of Neu5Gc tropism being extremely rare in wild type AIVs and the lack of Neu5Gc in birds [70]. Tissue samples were collected from trachea, lung, ileum, and colon from chickens, mallards, and tufted ducks. Mallards were characterized as well, as mallards are the most well described host of LPAIVs [19]. Indeed, Neu5Gc was not detected in the glycoproteomic analysis of any of the avian tissue samples. It was earlier thought that avian glycosylation produced shorter glycans, whereas human glycosylation produce longer glycans [135]. However, glycoproteomic analysis of avian tissues clearly demonstrated that longer sialylated glycan structures were present in all sampled avian tissues as well. Additionally, sialylated Lewis structures were present in all sampled avian tissues as well. Additionally, sialylated Lewis structures were present in all sampled avian tissues (Figure 8) [136]. Concordantly, maximum binding of the studied H3N8, H4N6, H6N2, H8N4, and H9N2 viruses was observed to sialyl Lewis structures in the glycan array, highlighting the importance of these structures as receptors for AIVs in mallards. Neu5Acα2,3 and fucosylated sialylated glycan structures were found approximately equally distributed across the investigated species and tissues, except in mallard ileum that had an elevated portion of sialylated Lewis structures. Conversely, the least sialylated fucosylated glycan structures were detected in tufted duck ileum. Given the widespread tropism for S-Leα in the glycan array, it can be speculated that these observations are part of the explanation of AIV widespread tropism for mallard intestine, but seldom isolates from tufted ducks.
N-glycopeptides terminating with α2,3-linked Neu5Ac were found in membrane proteins from all sampled avian tissues. Additionally N-glycopeptides terminating with α2,6-linked Neu5Ac were found in secreted serum proteins from all sampled avian tissues. However, a few N-glycopeptides terminating with α2,6-linked Neu5Ac were detected in predicted membrane proteins. Identified O-glycopeptides had core 1 structure (Galβ1,3GalNAcα1-O-Ser/Thr) with Neu5Acα2,6 on the GalNAc for serum glycoproteins and Neu5Acα2,3 on the Gal, or both on the Gal and the GalNAc, for membrane glycoproteins.

All studied mallard-derived LPAIVs did bind to α2,6-linked Neu5Ac in the glycan array as well. Originally, α2,6-linked Neu5Ac was regarded as a mammalian IAV receptor [12, 58], but over time this concept has been challenged [38, 82-84]. Indeed, medium to strong binding was observed by most of the studied viruses to Neu5Acα2,3/6-decorated core 1 O-glycans in the glycan array. O-glycans are commonly found on secreted mucins comprising a protective extracellular matrix on the apical side of epithelia [137]. The role of sialylated O-glycans as receptors (i.e. decoy vs. functional), as well as the role of α2,6-linked Neu5Ac as receptor for AIV in wild birds, and the interaction of AIVs with sialylated secreted glycoproteins vs. membrane associated sialylated glycoproteins in a biological setting needs further investigation. The available receptor repertoire of fowl, swine, and humans were demonstrated to be much more similar than previously known and binding of Neu5Acα2,6 seem to be an intrinsic property among many wild type AIVs. This has implications for the earlier postulated receptor barrier

---

**Figure 8.** Examples of fucosylated (red triangles) structures terminating with Neu5Acα2,3 (purple rhombi) from tracheae of mallards (A), chickens (B), and tufted ducks (C). Glycan symbols are according to the SNFG format [136].
for interspecies transmission of IAVs [58]. Without doubt, recognition of Neu5Acα2,6 is a hallmark of human adapted IAVs. However, it seems as rather than gain of Neu5Acα2,6 tropism, loss of Neu5Acα2,3 tropism is a necessary step for AIVs to become human adapted [80]. It can be speculated that one of the reasons favoring the loss of Neu5Acα2,3 tropism in humans is the high content of Neu5Acα2,3 in secreted mucins in the human respiratory tract, acting as decoy receptors protecting the underlying epithelium.

Sulfated 3’SLN and S-Le^x have been suggested as main receptors for H7 AIVs, HPAIV H5 isolated from chickens and Laridae-specific H13 and H16 LPAIVs [73, 74, 77]. However, no sulfated glycopeptides were identified in the glycoproteomic analysis of investigated avian tissues, suggesting that sulfated sialylated glycans (including Su-3’SLN and Su-S-Le^x) are not present in avian epithelia and thus do not act as receptors for IAVs in birds. Phosphorylated high-Man N-glycopeptides were identified in the investigated mallard and chicken tissues (not tested for tufted duck). High-Man structures have historically not been considered in the context of IAV-host cell receptor interactions, due to their non-sialylated nature. But, Byrd-Leotis et al. recently reported that IAV (including AIVs) have an intrinsic property to bind Man6P containing non-sialylated glycan structures in a SA-independent manner [80]. Indeed, it’s intriguing that such structures were identified in the investigated mallard and chicken epithelia. Byrd-Leotis et al. showed that the binding of Man6P containing glycans was not mediated by the canonical HA RBS, but through unidentified phosphorylated glycan binding sites of the virus particle [80]. Indeed, they identified host cell-derived Man6P receptors (MPRs) on the surface of the virus particle. However, binding of Man6P was shown to not be mediated by host cell-derived MPRs. Thus, the mechanism of IAV binding to Man6P needs further investigation, but the finding of phosphorylated high-Man structures in target cells in mallards (an AIV natural reservoir host species) and chickens brings biological meaning to the principle proven by Byrd-Leotis et al. Moreover, in the investigated mallard and chicken epithelia phosphorylated high-Man structures were mainly identified on N-glycopeptides derived from lysosomal enzymes including carboxypeptidase and aminopeptidase (phosphorylation of such enzymes directs them to the lysosome [138]). A possible explanation to this phenomenon is if IAV binding to Man6P takes place during late phase cellular entry, it could facilitate the virus uncoating process.

Concluding the study, there was correlation between some of the main receptors bound in the glycan array and sialylated glycans identified in the avian tissue samples. However, there was no clear correlation between the glycan binding profile and obtained virus swab titers from the inoculated birds or establishment of infection, indicating that there are many obstacles to overcome for a virus when introduced into a new host species, additional
to host cell receptor recognition. Still, several of the viruses in the studied virus panel were detected in both tufted ducks and chickens, including tissue specimens. Among these, the mallard-derived LPAIV H6N2 virus had the highest intrinsic property of interspecies transmission.
Study IV

Historically, field surveillance data has been heavily biased towards especially mallards [22-24]. However, it has recently been suggested that the focus of AIV research should be broadened to include more species to increase the understanding of AIV ecology [36]. In this study, one isolate from each of the H1-16 LPAIVs isolated from wild birds was characterized using glycan array analysis together with virus histochemistry using avian tissues. Four dabbling duck species; gadwall (Mareca strepera), Eurasian wigeon (M. penelope), mallard (Anas platyrhynchos), and Eurasian teal (A. crecca), and two diving duck species common pochard (Aythya ferina), and tufted duck (A. fuligula) were investigated together with White Leghorn chickens (Gallus gallus domesticus). Subtypes H3, H4, and H6 are dominating in mallards, whereas the other HA subtypes are less common or even just infrequently reported [25]. It was thus hypothesized that the subtype variability reported from field surveillance studies would be mirrored in the HA subtype attachment pattern observed to duck colon. This turned out not to be the case, as the attachment pattern of H3, H4, and H6 was not distinguishable from the other studied subtypes (Figure 9). Instead, the studied H5 virus (an intermediately reported subtype in wild ducks) dominated the attachment to studied cola, followed by the studied H7 and H14 viruses [25]. However, in studied cola mallard and Eurasian teal highly dominated with medium to intense attachment of all studied viruses except H13, H15, and H16. Mallard and Eurasian teal are both members of the Anas genus [86]. Gadwall and Eurasian wigeon are former members of Anas, but have been moved to the new genus Mareca. The prominent difference in observed virus attachment to Anas ducks vs. Mareca and Aythya ducks clearly corroborates the importance of Anas ducks as hosts of AIVs.

Figure 9. Histochemistry scores of trachea (left) and colon (right); 1. no stained cells; 2. <10 % stained cells; 3. 10–50 % stained cells; 4. >50 % stained cells. The row annotation illustrates investigated avian species’ taxonomy. Column annotations illustrate avian species of virus isolation, AIV HA group, and clade.
The observed attachment pattern to trachea was clearly different from the attachment pattern to colon. Most studied viruses attached poorly to all studied bird species, except the studied H5, H7, and H14 viruses that instead attached extensively to most of the studied bird species. This is interesting from several aspects as first, these viruses are all moderately to infrequently reported in wild birds and H14 AIV is not reported from Scandinavia [25]. Secondly, colon was early described as the site of AIV infection in ducks and has since been regarded as the major organ of interest in AIV research in wild ducks [30]. However, this study suggests that trachea should not be neglected as an organ of interest in AIV research in wild birds, especially for the detection of the so-called rare subtypes.

It was early recognized that the linkage conformation of the bond between Neu5Ac and the penultimate Gal is essential for the HA–Neu5Ac interaction [58]. Additionally, later studies demonstrated the importance of the glycan moieties downstream of the Neu5Ac penultimate Gal for the HA–Neu5Ac interaction process [73-75]. Moreover, it has been shown that the tissue tropism between different LPAIVs is highly variable in different host species [33, 101]. Common for all viruses investigated in this study were that they only bond Neu5Ac and no attachment was observed to non-sialylated structures or Neu5Ac structure analogues such as Neu5Gc or Kdn, except for the studied H7N7 virus that bound Neu5Gcα2,6Galβ1,4GlcNAc (Figure 10). The H7N7 virus distinguished as having the most restricted glycan binding profile in the virus panel and medium intensity binding to Neu5Gcα2,6Galβ1,4GlcNAc. Binding to Neu5Gc is extremely uncommon among wild type non-H7 IAVs and has only been reported as weak binding for H7 AIVs and strong for equine H7N7 virus [70, 77]. The equine isolate caused a highly pathogenic outbreak in horses, but was not maintained in horses. The equine H7N7 virus was reported to bind specifically to 3’Neu5Gc–LN, but unfortunately no 3’Neu5Gc–LN was incorporated in the current glycan array analysis [70]. Taken together, these findings suggests that tropism of Neu5Gc might be an intrinsic property unique to H7 viruses. Neu5Gc is commonly expressed in many mammals including horses and pigs, but not in humans, ferrets, seals, and dogs nor birds [70]. It has thus been suggested that tropism for Neu5Gc is selected against. Still, viruses with Neu5Gc tropism have zoonotic potential for other mammals including horses and pigs.
For the non H7-viruses, main binding was observed to α2,3-linked Neu5Ac structures, especially sulfated 3’SLN and sialyl-Lewis structures. Interestingly, binding to α2,6-linked Neu5Ac showed to be much more widespread than earlier thought. Particular α2,6-linked structures bound included 6’SdiLN, [6Su]Galβ1,3[Neu5Aca2,6]GalNAc, and Neu5Aca2,6GalNAc. Widespread binding was also observed to the chimeric Neu5Aca2,3Galβ1,3[Neu5Aca2,6]GalNAc structure. It has been postulated that different HA subtypes prefer different receptor structures [73, 74]. However, among the 16 studied LPAIV isolates from subtypes H1–16 the observed glycan binding profiles were rather similar across subtypes. Moreover, HA phylogeny or host species of virus isolation did not correlate with the observed glycan binding profile. Instead, the investigated viruses could be grouped into viruses with a restricted vs. promiscuous glycan binding profile. Thus, reported results did not support the “one HA subtype, one
receptor structure” hypothesis, but illustrate a much more complex relation between HA subtype and the observed glycan binding pattern.

The tropism for α2,3 vs. α2,6-linked Neu5Ac has for long been described as a species barrier between AIVs and human adapted IAVs [58]. However, all the studied AIVs did to some extent bind α2,6-linked Neu5Ac. Yet, direct transmission of AIV to humans from an avian host is rare, suggesting that there are other barriers to overcome for successful zoonotic transmission of AIVs apart from host cell receptor recognition.
Conclusion

Study I
This study corroborates earlier reports suggesting that AIV attachment is more extensive in human tissues in contrast to pig tissues. The results suggest that not all AIVs would require passage through pigs to adapt to mammalian receptor structures but have intrinsic tropism to such structures. Given the rare incidence of direct AIV transmission between birds and humans, the reported results suggest that the observed species barrier (preventing direct AIV transmission between birds and humans) is comprised of multiple factors additional to the virus receptor tropism, such as e.g. variance in host cell endogenous biophysiological environment and response to virus infection.

Study II
The observed AIV attachment to tissues from various avian species, including both classical and non-classical AIV hosts, together with more extensive virus attachment to trachea than colon in non-chicken/non-mallard species, suggest the importance of the respiratory tract in AIV replication. Moreover, any differences attributed to the geographical origin of virus or tissue donor species could not be observed.

Study III
Successful infection in tufted ducks and chickens via the oculonasal and not the intraesophageal route, suggests that the tropism of AIV to the gastrointestinal tract, as earlier observed in mallards, does not apply to all duck species. Additionally, this study provides the first glycoproteomic analysis of avian respiratory and intestinal tissues, it was clearly demonstrated that birds do produce longer sialylated glycans, including sialylated Lewis structures. Moreover, it was clearly demonstrated that birds produce both Neu5Aca2,6-decorated N and O-linked glycans further questioning the rationale of α2,6-linked Neu5Ac as a mammalian-specific IAV receptor.
Study IV

The study corroborates the importance of Anas ducks as hosts of AIV together with trachea as a tissue of interest for the detection of rare AIV subtypes. Moreover, rather than “one HA subtype, one receptor”, AIV receptor tropism is much more complex with shared receptors across HA subtypes. Main binding to earlier described variants of $\alpha_{2,3}$-linked Neu5Ac together with widespread binding to particular forms of $\alpha_{2,6}$-linked Neu5Ac suggest that Neu5Aca2,6 tropism might be an intrinsic property of many wild type AIVs.

Final conclusion

The natural reservoir of IAV and virus replication in the mallard intestine was described in the 1970’s and the pig was described as a mixing vessel for IAVs during the 1980/90’s. Since then, field surveillance has mainly targeted dabbling ducks, but later it has been shown that gulls carry the gull-specific H13 and H16 AIV subtypes. More recently it has been suggested that the focus of AIV field surveillance should be broadened to cover more avian taxa to improve the understanding of AIV ecology and transmission chains.

Study I demonstrated that wild type AIVs attach poorly to pig tissues and in fact attach more to human tissues. Moreover, study I, III, and IV demonstrated that tropism for Neu5Aca2,6 seem to be a widespread intrinsic property among many wild type AIVs. Taken together these two findings challenge the dogma as pigs as a “mixing vessel” for IAVs. Indeed, from a receptor point of view pigs appear to not be pivotal for IAV interspecies transmission. However, rather than structural reasons (e.g. receptor display) pigs might still be a “mixing vessel” for IAVs due to ecological reasons (e.g. kept in high numbers and densities and commonly in contact with other livestock facilitating virus exchange) and host cell environmental factors.

Study II and IV demonstrated that AIV attachment in many bird species is shifted towards trachea and not colon, except in Anas ducks (including mallards). Indeed, in the infection experiment in study III, tufted ducks and chickens were more susceptible to virus inoculation in the respiratory tract than esophagus. Additionally, detected virus titers were consistently higher in oropharyngeal than cloacal swabs. Together these data suggest that in most bird species AIV target organs are in the respiratory tract, whereas in Anas ducks AIV targets the large intestine.
Study III provided the first glycoproteomic characterization of the avian respiratory and intestinal tracts. The reported results partly corroborates some of the suggested main AIV receptors in wild birds (including 3’STF, 3’SLN and S-Le structures), but showed that other suggested main AIV receptors (e.g. sulfated sialylated glycans) are not present in mallards, tufted ducks or chickens. Moreover, the presence of recently identified non-sialylated phosphorylated IAV receptors were confirmed in mallards and chickens. Concluding the avian sialome and glycome in a larger context, mammalian (e.g. human) and avian sialylated glycan structures seem to be much more similar than earlier thought, implying that available receptor structures alone are not a major barrier for IAV interspecies transmission. Yet, based on epidemiological data direct transmission of AIV from birds to humans is very rare. Rather than the plethora of displayed receptor structures (here suggested to be similar across birds and mammals) as a major barrier per se, barriers may exist in the role of decoy vs. functional receptors across different species. Furthermore, additional suggested barriers for interspecies transmission are host cell biophysiological properties (including core temperature, pH, and available transcription factors such as ANP32) and host cell innate immune response to infection.
Acknowledgment

Patrik Ellström, my main supervisor, thanks for understanding the importance of balance between work and spending time in the outdoors.

Björn Olsen many thanks to you for employing this “random” newly graduated engineer calling you saying “I know birds and I would like to work for you. Please give me a PhD student position!”.

Åke Lundkvist, thanks for believing in the potential of students.

Jonas Waldenström, I highly appreciate your feedback on manuscripts improving my written texts.

Josef Järhult, thanks for solving all administrative work with authorities and SVA regarding the animal studies.

Cecilia Lindskog Bergström thanks for hosting me and allowing me to do wet lab work in your lab.

Ola Blixt thanks for hosting me and allowing me to do wet lab work in your lab.

Mahmoud Naguib, thanks for your excellent supervision in handling the birds in the vivarium.

Michelle Wille, thanks for being an influenza virus peer.

Marcin Kierczak, Thomas Källman, Roy Francis, Sebastian DiLorenzo, Bengt Sennblad, and Nikolay Oskolkov, thanks for organizing the most useful courses within my PhD training with excellent pedagogics and professional execution.

Erik Enbody and Grace Sprehn for thoughtful comments on wording of this thesis and sharing bird love and talking about running over lunch.

My parents, for emphasizing the importance of proper education and your continuous support.
References


[64] Baker SF, Mehle A. ANP32B, or not to be, that is the question for influenza virus. Elife 2019;8.


[67] Cohen M, Varki A. The sialome--far more than the sum of its parts. OMICS 2010;14(4):455-64.


Shirihai LSH. Handbook of Western Palearctic Birds. Helm; 2018.


[131] Fouchier RA, Schneeberger PM, Rozendaal FW, Broekman JM, Kemink SA, Munster V, et al. Avian influenza A virus (H7N7) associated with human conjunctivitis and a fatal case of acute respiratory distress syndrome. Proc Natl Acad Sci U S A 2004;101(5):1356-61.


Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine 1609

Editor: The Dean of the Faculty of Medicine

A doctoral dissertation from the Faculty of Medicine, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine. (Prior to January, 2005, the series was published under the title “Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine”.)

Distribution: publications.uu.se
urn:nbn:se:uu:diva-395407