List of papers


III. Helin, A.S., Chapman, J.R., Tolf, C., Aarts, L., Bususu, I., Rosengren, K.I., Andersson, H.S., Waldenström, J. Relation between structure and function of three AvBD3b variants from mallard (*Anas platyrhynchos*). *Manuscript*


Lnu.se

Eco-immunological studies of innate immunity in Mallards (*Anas platyrhynchos*)
ECO-IMMUNOLOGICAL STUDIES OF
INNATE IMMUNITY IN MALLARDS
(*Anas platyrhynchos*)

Anu Helin

Linnaeus University Press
Eco-immunological studies of innate immunity in Mallards (Anas Platyrhynchos)
Doctoral dissertation, Department of Biology and Environmental Science, Linnaeus University, Kalmar, 2020

Published by: Linnaeus University Press, 351 95 Växjö
Printed by: Holmbergs, 2020
Abstract


This thesis comprises two sections, both of which explore eco-immunology of the innate immune system of mallards (Anas platyrhynchos). The innate immune system serves a pivotal role as the first line of defense against invading pathogens, and is comprised of physical and chemical barriers. Its main function is to inhibit and/or eliminate the pathogenic microorganisms while minimizing collateral damage to host cells.

The first section investigates the allelic variation and selective forces acting on five avian β-defensin (AvBD) genes. Showing that purifying selection is the predominant selective force, although one gene AvBD3b, appeared to be subject to balancing selection. Moreover, the solution structure of the AvBD3b peptide was solved in this work, revealing that it contains a typical β-defensin fold with three β-sheets. Linear and folded AvBD3b peptides were shown to exhibit similar antibacterial properties, indicating that the tertiary structure was not the primary determinant of antimicrobial activity. Moreover, testing the antimicrobial activity of synthetic AvBD peptides showed that they mostly had higher activity against Gram-negative than Gram-positive bacteria.

The second section investigates expression of two innate immune genes during avian influenza virus infection. Data quality in gene expression studies depends, in part, on the stability of the reference genes (RGs) used to normalize expression levels, meaning putative RGs must be validated prior to use. Eleven potential mallard RGs were tested and it was found that the stability varied across different tissue types, highlighting the importance of correct RG selection for the specific experimental conditions. Optimal RGs were then used in a gene expression study of retinoic acid inducible gene 1 (RIG-I) and myxovirus resistant gene (Mx) in mallards during a low pathogenic avian influenza (LPAI) infection. Upregulation of both genes was rapid and transient, returning back to basal levels two days post infection across most of the five tissue types analyzed.

This thesis provides new insights into the tertiary structure and antimicrobial activity of AvBDs, and how this relates to selective pressures exerted in natural populations. It also highlights the importance of RGs validation, and confirms that RIG-I and Mx are involved in the early stages of the mallard immune response to LPAI infection.

Keywords: Avian β-defensins (AvBDs), Avian gene expression, Host-pathogen interactions, Innate immunity, Mallard (Anas platyrhynchos), Myxovirus resistance gene (Mx), Retinoic acid inducible gene 1 (RIG-I)
Svensk sammanfattning

Den här avhandlingen består av två delar, vilka fokuserar på olika delar av det medfödda immunförsvaret hos gräsand (Anas platyrhynchos). Det medfödda immunförsvaret består av fysiska och kemiska barriärer vars uppgift är att hindra patogener att infektera kroppens celler, och att vid infektion eliminera dem.


Denna avhandling ger nya insikter om genetisk diversitet och evolution hos AvBD-gener i gräsand. Den tredimensionella strukturen av AvBD3b bidrar till ökad kunskap om AvBD-strukturer då endast några få defensinstrukturer från fåglar tidigare bestämts. Avhandlingen visar även på viken av att undersöka stabiliteten hos potentiella referensgener för att få pålitliga resultat vid expressionsstudier samt att generna RIG-I och Mx är involverade under den tidiga immunresponsen vid fågelinfluensavirusinfektion i gräsänder.

Nyckelord: Avian β-defensins (AvBD), Genuttryck, Gräsand (Anas platyrhynchos), Immunförsvar, Immungener, Myxovirus resistant gene (Mx), Patogen, Retinoic acid inducible gene 1 (RIG-I)
“Be like a duck. Remain calm on the surface and paddle like hell underneath.”

- Michael Caine
# Table of contents

List of papers............................................................................................................................... 3  
My contribution to the individual papers ...................................................................................... 4  
Abbreviations................................................................................................................................. 5  
Introduction..................................................................................................................................... 7  
Immune system............................................................................................................................... 8  
Physical defenses .......................................................................................................................... 8  
Innate immune system...................................................................................................................... 9  
  Pattern recognition receptors.......................................................................................................... 10  
  *Retinoic acid inducible gene 1* ..................................................................................................... 10  
  Signaling ........................................................................................................................................ 11  
  Effector mechanisms ..................................................................................................................... 11  
  Antimicrobial peptides .................................................................................................................. 12  
  *Myxovirus resistant gene* .............................................................................................................. 14  
Adaptive immune system.................................................................................................................. 15  
  B cells ............................................................................................................................................. 16  
  T cells ............................................................................................................................................. 16  
Key differences in immunity between mammals and birds ......................................................... 17  
Avian AMPs...................................................................................................................................... 19  
  Avian AMP families ....................................................................................................................... 19  
  Cathelicidins ................................................................................................................................. 19  
  LEAP-2 ......................................................................................................................................... 19  
  Avian β-defensins ......................................................................................................................... 20  
  Ovodefensins ................................................................................................................................. 21  
The mallard (*Anas platyrhynchos*) ............................................................................................... 21  
Diseases ........................................................................................................................................... 22  
  Avian influenza viruses ................................................................................................................ 22  
Aims .................................................................................................................................................. 25  
Brief methodological overview ....................................................................................................... 27  
Samples for allelic variation and selection study ......................................................................... 27  
Allelic variation and selection acting on AvBDs .............................................................................. 28  
Oxidative folding of AvBDs ............................................................................................................. 28  
Antimicrobial assays of AvBDs ........................................................................................................ 29  
Validation of reference genes for gene expression studies using qPCR ........................................ 30  
LPAI infection experiment in mallards ........................................................................................... 30  
Gene expression of *RIG-I* and *Mx*............................................................................................ 30
Results and discussion........................................................................................................33
Evolution and selection on waterfowl β-defensins.........................................................33
Antimicrobial activity of AvBDs.........................................................................................35
Gene expression and infection.........................................................................................38
Validation of reference genes.........................................................................................38
LPAI infection in mallards and gene expression..............................................................39
Conclusions.........................................................................................................................41
Acknowledgements........................................................................................................43
References.........................................................................................................................47
List of papers


III. Helin, A.S., Chapman, J.R., Tolf, C., Aarts, L., Bususu, I., Rosengren, K.J., Andersson, H.S., Waldenström, J. Relation between structure and function of three AvBD3b variants from mallard (*Anas platyrhynchos*). *Manuscript*


Paper I was reprinted with the permission from Oxford University press and copyright clearance center. Paper IV-VI are open access papers and were reprinted in accordance with the Creative Commons Attribution (CC BY) license.
My contribution to the individual papers

I. I did most of the laboratory work and participated in the analysis as well as in the writing process.

II. I designed the study with the help from some co-authors. I did all the laboratory work together with CT. I analyzed the data and wrote the first draft of the manuscript. All co-authors contribute in the writing process.

III. I designed the study with the help from some co-authors. I together with HSA and IB did the peptide folding where IB was the one leading the process at the beginning. I did all the microbiology laboratory work together with CT. I analyzed the data and wrote the first draft of the manuscript. Most co-authors contribute in the writing process.

IV. I did all the laboratory work, participated in the analysis and the writing process

V. I did all the laboratory work in Kalmar and most of the analysis. I together with JRC wrote the article.

VI. I did all the laboratory work in Kalmar and most of the analysis. I together with JRC wrote the article.
Abbreviations

aa: Amino acid
AIV: Avian influenza virus
AMP: Antimicrobial peptide
AvBD: Avian β-defensin
BCR: B cell receptor
CARD: Caspase recruitment domains
CFU: Colony forming unit
CI: Confidence interval
Ct: Cycle threshold value
Dpi: Day post infection
GOI: Gene of interest
GSH: Reduced glutathione
GSSG: Oxidized glutathione
HDP: Host defense peptide
HPAI: High pathogenic avian influenza
HPLC: High performance liquid chromatography
HSQC: Heteronuclear single quantum coherence spectroscopy
IC50: Half maximal inhibitory concentration
IFN: Interferon
ISG: Interferon stimulated gene
K: Kelvin
LA/LB: Lysogeny agar/broth
LC-MS: Liquid chromatography-mass spectrometry
LPAI: Low pathogenic avian influenza
MBC: Minimal bactericidal concentration
MHA/MHB: Müeller Hinton agar/broth
MHz: Mega hertz
MIC: Minimal inhibitory concentration
Mx: Myxovirus resistance gene
NaPB: Sodium phosphate buffer
NaPB+MHB: Sodium phosphate buffer containing 1% MHB
NMR: Nuclear magnetic resonance
NOE/NOESY: Nuclear Overhauser effect/NOE spectroscopy
nt: Nucleotide
PAMP: Pathogen-associated molecular patterns
PRR: Pattern recognition receptor
RG: Reference gene
RIG-I: Retinoic acid inducible gene I
RLR: RIG-I like receptor
RP-HPLC: Reverse phase- high performance liquid chromatography
Rpm: Revolutions per minute
TCR: T cell receptor
TOCSY: Total correlated spectroscopy
TSA/TSB: Tryptic soy agar/broth
UTR: Untranslated region
v/v: Volume per volume
w/v: Weight per volume
ZOI: Zone of inhibition
Introduction

Multicellular organisms possess an innate immune system that helps to defend against microorganisms. In particular, it targets pathogenic microbes, being those that are harmful to the organism (host) and therefore need to be rapidly controlled before causing too much harm to host tissues, cells and metabolic processes. Infectious diseases occur when pathogens evade the host’s initial immune defenses and begin replicating and causing damage. Higher organisms have a second, more advanced system called the adaptive immune system. Although often treated separately in the literature, there is in fact a constant interplay between cells and proteins from both branches of the immune system. As anyone who has opened an immunology textbook can attest to, the immune system is very diverse and complicated to understand. It is simultaneously essential for survival and a powerful tool that can cause self-harm if not controlled; as such it is highly regulated in an intricate web of genes, molecules, peptides and proteins and can therefore be deployed at a level appropriate to the circumstance facing the organism, such as the virulence of an infection.

Much of our current understanding of the immune system stems from studies in humans and model system animals such as mice, rats, chickens, pigs and Drosophila (Chou et al., 2013; Jennings, 2011; Kaiser, 2010; Meurens et al., 2012). In contrast, much less is known about wild animal species. In this thesis, I have chosen to study one wild bird species – the mallard (Anas platyrhynchos) – and different components of its innate immune system. This duck species has a wide Holarctic distribution and is often common in areas occupied by humans. It is also a reservoir species for influenza A viruses, from which viruses may be seeded into domestic poultry and cause outbreaks of highly pathogenic avian influenza. Some of the other pathogens carried by mallards also have zoonotic potential, therefore understanding mallard immune processes can be helpful in understanding epidemiology of diseases in other species including humans.
Immune system molecules can signal when invading microorganisms have entered the host and help to initiate an immune response by recruiting specialized cells and proteins that eliminate the invader. Other immune molecules are involved in the direct inhibition of pathogens, or in other immune processes. In this thesis, I have investigated several innate immune genes involved in immune defense against different pathogens. To this end, I have used a variety of chemical and microbial assays and genetic tools to address my research questions.

**Immune system**

As this thesis involves immune processes, a recapitulation of the general outline of the immune system may be a good place to start. The main function of the immune system is to provide a protective barrier against harmful agents such as pathogens and toxins, while minimizing collateral damage in the host. The immune system is highly coordinated via different molecules used for signaling between cells and recruiting cells to infected areas (Fig. 1) (Male et al., 2012).

![Fig. 1. Cell types involved in innate and adaptive immunity. Reprinted from Yamauchi and Moroishi (2019) under CC BY 4.0 license.](image)

**Physical defenses**

The very first line of protection against pathogens are the physical barriers of the host which prevent microbes from entering the body, and/or adhering to tissues. Such barriers include the skin, mucus membranes and other mechanical defenses such as eye lashes and eyelids which, together with blinking and tear
production can remove unwanted microbes from the eye. These barriers provide not only a physical obstacle, but can also produce different chemical substances that directly kill the pathogens (Abbas et al., 2016). Additionally, the collection of microorganisms that inhabits a healthy animal – the microbiome – can help keep harmful microorganisms at bay, for example via competitive interactions such as adherence to intestinal cells (Abbas et al., 2016; Liévin-Le Moal & Servin, 2006). Many microorganisms are stopped by these barriers, but if a microbe does manage to evade this first line of defense, then additional parts of the immune system are activated to try and eliminate the threat.

**Innate immune system**

A vital component of host defense is the innate immune system. It is comprised of a range of different immune cells and chemical defenses. It is a nonspecific defense that acts within minutes or hours of infection and is present in all living organisms. The most abundant innate immune cells are the neutrophils, which are produced in the bone marrow and which circulate in the blood. Neutrophils are phagocytic cells that engulf pathogens and which can easily leave the blood and enter infected tissues, though they are short-lived once inside tissues. Furthermore, neutrophils can release different substances from granular storage upon infection (Abbas et al., 2016; Gasteiger et al., 2017). Additional innate immune cell types include 1) monocytes - similar to neutrophils, though when they migrate into tissues they can develop into macrophages which are longer-lived and engulf pathogens; 2) dendritic cells - starts the inflammatory process and stimulates the adaptive immune system; 3) mast cells - stimulates inflammation and can release different chemical substances such as cytokines (Abbas et al., 2016); and 4) natural killer (NK) cells - cytotoxic lymphocytes that can directly lyse infected cells and contain granules with proteins that are used to kill intracellular pathogens (Abbas et al., 2016; Abel et al., 2018).

Another important component of the innate immune system is the complement system. Complement is comprised of a variety of proteins that can act as proteolytic enzymes and regulatory control proteins. Additionally, complement proteins can mark pathogens for destruction by phagocytes (Abbas et al., 2016; Sarma & Ward, 2011). These proteins are mainly produced by hepatocytes in the liver, though a variety of innate immune cells can produce a subset of them (Lubbers et al., 2017).

Furthermore, there are different types of plasma proteins and cytokines. A rapid immune response relies on pattern recognition receptors (PRRs) in and on the different immune cells to recognize the structures of the invading pathogens.
Pattern recognition receptors

Pattern recognition receptors (PRRs) recognize pathogen-associated molecular patterns (PAMPs), and damage-associated molecular patterns (DAMPs) (Abbas et al., 2016; Newton & Dixit, 2012; Takeuchi & Akira, 2010). PAMPs are evolutionarily conserved features of microbes that are crucial for their survival and are involved in infection and colonization of the host (Abbas et al., 2016; Akira et al., 2006). Commonly, part of the PRR structure possesses PAMP affinity, whereas another is involved in a downstream signaling pathway that stimulates the production of different molecules needed to fight off the invading pathogen. PRRs are abundant and diverse, with several broad families including Toll-like receptors (TLRs), NOD-like receptors (NDRs), RIG-I like receptors (RLRs) and C-type lectin receptors (CLRs). PRRs are expressed in or on a variety of immune cells and have evolved to recognize different PAMPs (Akira et al., 2006; Ishii et al., 2008; Takeuchi & Akira, 2010). Some PRRs, such as TLR2, can recognize PAMPs from a variety of pathogens such as Gram-positive bacteria, viruses and fungi (Coates et al., 2018), whereas others are more specific, such as RLRs, which only recognize viral RNA (Loo & Gale, 2011).

Retinoic acid inducible gene 1

The RLR family of PRRs comprises three different genes: melanoma differentiation associated factor 5 (MDA5), laboratory of genetics and physiology 2 (LGP2) and retinoic acid inducible gene 1 (RIG-I). In vertebrates, expression of these genes is upregulated during RNA virus infections, though RIG-I is apparently missing in some fish and bird species, including chicken (Gallus gallus) (Barber et al., 2010; Zou et al., 2009). In this thesis RIG-I was studied. It is an RNA sensor which recognizes RNA molecules from viruses such as orthomyxoviruses, including influenza virus. RIG-I is located in the cytoplasm of the cell and is comprised of two caspase activation and recruitment domains (CARDs) at the N-terminus, a DEAD box helicase/ATPase domain in the middle that is involved in binding RNA and a repressor domain at the C-terminal domain that is involved in autoregulation (Fig. 2, left panel) (Loo & Gale, 2011; Takeuchi & Akira, 2010). RIG-I preferentially targets short RNA molecules with uncapped 5’ triphosphate (5’ppp), compared to MDA5 which preferentially binds longer fragment (Fig. 2, right panel) (Loo & Gale, 2011). Upon binding of viral RNA, RIG-I is activated and the closed structure “opens”, and interacts with other molecules involved in the downstream signaling cascade that eventually leads to increased expression of type 1 interferons (IFNs, namely INFα/β, see Fig. 2, right panel) and cytokines (Loo & Gale, 2011; Takeuchi & Akira, 2010). RIG-I is an IFN inducible protein, thereby stimulating the production of RIG-I via a positive feedback loop (Akira et al., 2006).
Signaling

Interactions between PRRs and PAMPs or DAMPs leads to a conformational change in the receptor, followed by oligomerization, and assembly of several different sub-units to form a large signaling complex. This initiates the release of different factors that promote recruitment of leucocytes to the inflamed area, release of cytokines and production of antimicrobial peptides (AMPs) and interferon stimulated genes (ISGs) (Takeuchi & Akira, 2010). For example, TLRs are important transmembrane PRRs containing leucine-rich-repeat (LRR) motifs in the plasma membrane which bind the PAMP, and a Toll/Interleukin 1 receptor (TIR) domain in the cytosol, involved in signal transduction to an adaptor protein such as myeloid differentiation primary response 88 (MyD88). MyD88 then initiates a signaling cascade that culminates in the activation of NF-κB, an important transcription factor that regulates inflammatory responses (Newton & Dixit, 2012).

Effector mechanisms

The interaction between PRRs and PAMPs and the release of proinflammatory cytokines results in activation of cellular and molecular effector mechanisms to neutralize invading pathogens. Effector cells include different leucocytes such
as phagocytes, natural killer cells (NK cells) and neutrophils, whereas the molecular responses include, for example, antimicrobial peptides (AMPs) and complement activation. Leucocytes are important components of the innate immune defense in birds and mammals, with granulocytes containing granules with different chemical substances that can either have an antimicrobial activity or function to recruit other immune cells to the infected area (Male et al., 2012). NK cells are cytotoxic cells that can directly lyse infected cells. Additionally, NK cells release inflammatory cytokines such as interleukins (ILs) and tumor necrosis factors (TNFs) upon stimulation of their receptors (Abel et al., 2018; Male et al., 2012). The complement system is comprised of proteins that bind pathogens and cells not protected by specific surface molecules, and mark them for degradation. Thereafter, phagocytic cells engulf the pathogen and destroy it (Merle et al., 2015; Sarma & Ward, 2011). Arguably, the most important effector molecules during certain types of microbial infection, particularly bacterial and fungal infections, are the antimicrobial peptides.

**Antimicrobial peptides**

AMPs, also known as host defense peptides (HDPs), are short (less than 50 aa) cationic peptides. They are amphipathic, meaning that they contain both hydrophobic and hydrophilic regions. AMPs have broad-spectrum antimicrobial activity, and can either directly neutralize invading microorganisms or inhibit their protein and/or DNA synthesis (Coates et al., 2018; De Smet & Contreras, 2005). They are evolutionarily ancient peptides, found in the vast majority of living organisms, from unicellular to vertebrates. In unicellular organisms, AMPs play a role in nutrient acquisition, whereas in multicellular organisms they comprise an important component of the innate immune system (Ageitos et al., 2017). AMPs may also have regulatory and chemotactic effects on various immune cells (Lai & Gallo, 2009). These peptides can either be constitutively expressed, induced upon infection, or stored in secretory granules in granulocytes (Zhang & Gallo, 2016). In birds and mammals, the two main families of AMPs are cathelicidins and defensins. Cathelicidins are less diverse than defensins, and most organisms have fewer cathelicidin genes than defensin genes (Cheng et al., 2015; Zhang & Gallo, 2016). Both cathelicidins and defensins are produced as inactive three-part precursors by keratinocytes, epithelial cells and circulating phagocytic cells. This precursor contains a signal peptide at the N-terminus which tags the peptide for secretion, a propiece (which is sometimes absent) followed by the amino acid sequence of the active mature peptide (Fig. 3). The propiece can act as a chaperone that enables correct peptide folding, as well as inhibiting the activity of the mature peptide before it is cleaved off (Ganz, 2003; Lai & Gallo, 2009; Tennessen, 2005).
Cathelicidins are characterized by the cathelin-like domain (named after cathelin, a protein that can inhibit cathepsin L). The cathelin-like domain, as well as the signal peptide, are highly conserved within and between species, whereas the mature peptide is highly variable, even within the same species (Fig. 3A). Consequently, there is considerable variation in size and structure of mature cathelicidin peptides (Zanetti, 2005).

![Gene organization of (A) cathelicidin and (B) human β-defensin 2 (hBD2).](image)

Defensins are mainly characterized by their mature peptide, which includes six conserved cysteine (C) residues that form three intramolecular disulfide bonds (Ganz, 2003). Defensins are further divided into three subfamilies based on their shape and disulfide bond connectivity, α-defensins connecting C1-C6, C2-C4, C3-C5, β-defensins connecting C1-C5, C2-C4, C3-C6 and γ-defensins being circular peptides formed from two truncated α-defensins (Ganz, 2003; Selsted & Ouellette, 2005). β-defensins are the largest family of defensins and are found in all vertebrates, whereas α-defensins are restricted to mammals. γ-defensins have only been found in old world monkeys (i.e. baboons and macaques), although truncated γ-defensin genes have been found in human, gorilla and chimpanzee genomes (Selsted & Ouellette, 2005; Zhu & Gao, 2013).
ancestral proto-defensin that was present before the evolutionary split of birds and mammals is hypothesized to have been β-defensin-like (Cuperus et al., 2013; van Dijk et al., 2008; Xiao et al., 2004). Different models of how β-defensins interact with and inhibit pathogens have been proposed. Perhaps the most widely accepted is the ‘carpet model’ whereby cationic regions of the peptide which initially interacts with the negatively charged microbial membrane via electrostatic interactions, and where hydrophobic peptide regions then enable incorporation of peptides into the hydrophobic interior of the microbial membrane. The defensins aggregate in the membrane, causing tension and leakage; this inhibits and ultimately kills the microbe (Fig. 4) (Ganz, 2003).

![Proposed modes of action of antimicrobial peptides. From Wikipedia Commons user Ymahn under a CC BY-SA 3.0](https://commons.wikimedia.org/w/index.php?curid=15850987)

**Myxovirus resistant gene**

When IFNα/β production is activated via a PAMP or a DAMP binding to a PRR, a signaling cascade is initiated that leads to the activation of ISGs. There are several important ISGs, one of which is the *Myxovirus resistant* (*Mx*) gene. The Mx protein is a large GTPase which is involved in the inhibition of RNA viruses, such as influenza virus, by interfering with assembly of the viral ribonuclease protein assembly (Verhelst et al., 2013). *Mx* is present in most vertebrate species. Depending on species, Mx proteins are located either in the nucleus and/or in the cytoplasm where they inhibit virus replication (Lee & Vidal, 2002; Verhelst et al., 2013). Nuclear Mx inhibits viruses belonging to the family *Orthomyxoviridae* such as influenza virus, whereas cytosolic Mx inhibits
viruses from the family *Rhabdoviridae* such as vesicular stomatitis virus (Fig. 5) (Verhelst et al., 2013).

**Adaptive immune system**

The second major branch of the immune system is the adaptive immune system. It is activated via signals from the innate immune system and by binding of antigens to receptors on antigen presenting cells (APCs) which are lymphocytes in the adaptive immune system. The adaptive immune system takes longer to activate upon encounter with a ‘new’ pathogen. However, once activated the adaptive immune response is more specific than the innate immune response. Additionally, because the adaptive immune system has memory (see below), at the next encounter with the same pathogen, the immune response is much more rapid and robust. Broadly speaking, the adaptive immune system can be divided into two parts, the humoral (B cells and antibodies) and the cell-mediated (T cells and macrophages) pathways. Both B and T cells are produced in the bone marrow and go through a maturation process involving proliferation, expression of receptor genes and selection of the receptors (Abbas et al., 2016; Alberts, 2002).
**B cells**

B cells are lymphocytes (a type of white blood cell) that produce antibodies, also known as immunoglobulins or Igs. Antibodies are large Y-shaped proteins that neutralize pathogens via recognition of antigens. Antibodies can be presented as receptors on the B cell surface, called B cell receptors (BCRs), or can be released into circulation. Each B cell produces only one type of antibody with affinity for a specific antigen (Alberts, 2002). B cell development in the bone marrow includes cell proliferation and BCR expression. BCRs contain a heavy and a light chain, and expression of the receptor starts with DNA recombination of the heavy chain and then of the light chain. The receptors are highly variable, due to different combinations of three gene segments being generated during the recombination process, and via junctional diversity when the different segments are joined together. During this process, nucleotides can be removed or random nucleotides can be added, leading to variation that is not coded in the genome (Market & Papavasiliou, 2003). This process means that collectively, BCRs can recognize a wide variety of antigens. B cell clones carrying BCRs with an affinity for the hosts own antigens are eliminated in a negative selection process to avoid autoimmune responses, while remaining immature B cells migrate to the spleen, where the final step of the maturation occurs (Abbas et al., 2016). Here, B cells encounter their first foreign antigen, presented by a dendritic cell, and differentiate into plasma cells or enter the germinal center reaction, which is where memory cells are developed (LeBien & Tedder, 2008). Different antibodies can have affinity for many different antigens, such as proteins, carbohydrates, nucleic acids or lipids. Antibodies in the blood and the mucus layer of epithelial cells can neutralize and eliminate pathogens before they have the opportunity to colonize host cells (Abbas et al., 2016).

**T cells**

T cells are lymphocytes that are produced in the bone marrow and migrate to the thymus to develop into mature T cells. The T cell receptor (TCR) on T cells is comprised of two chains, called the α and β chains, which are produced in a process involving DNA recombination, similar to that described above for BCRs. In the thymus, immature T cells undergo negative and positive selection involving recognition of the major histocompatibility complex (MHC) molecule. T cells that successfully interact with the MHC survive while remaining cells die via apoptosis. T cells with an MHC affinity that is too high will also die. Additionally, TCRs have co-receptors called cluster of differentiation 4 (CD4) or CD8 and these determine which MHC molecules are recognized. CD4 recognizes class II MHC and CD8 recognizes class I MHC molecules. CD4 positive T cells are also known as T helper cells (Th cells) and recognize both extra- and intra-cellular pathogens (Fig. 6) (Alberts, 2002). They
help to defend against extracellular pathogens via a B cell mediated humoral response, Th cells are involved in the activation of antibody production by pathogen specific B cells (Alberts, 2002). During defense against intracellular pathogens, Th cells are involved in the activation of pathogen-engulfing phagocytes, which present pathogen protein fragments on a class II MHC molecules on the cell surface (Abbas et al., 2016; Alberts, 2002). Th cells also stimulate CD8 positive T cells, also known as cytotoxic T cells, that recognize cells infected with intracellular microbes by presentation of pathogen associated antigens on cell surface class I MHC molecules. Once activated, a cytotoxic T cell releases the content from its secretory vesicles, and this induces the infected cell to undergo apoptosis (Abbas et al., 2016; Alberts, 2002).

**Fig. 6.** CD4 positive T cell-mediated activation of B cells and CD8 positive T cells as well as phagocytic macrophages.

### Key differences in immunity between mammals and birds

Overall, the avian immune system shares many features with the mammalian immune system, both in terms of innate and adaptive immunity. However, as the avian and mammalian lineages split more than 200 million years ago, the avian immune system has features not seen in mammals (Fig. 7) (Kaspers & Schat, 2013). Firstly, avian species lack neutrophils, the most abundant leucocyte in the human immune system. Instead, avian species have heterophils (granulocytes) that are thought to be functional equivalents to neutrophils. Neutrophils and heterophils differ in that heterophils have a comparably weak oxidative burst release of reactive oxygen compounds, and depend instead on
antimicrobial peptides released from granules (Harmon, 1998). Secondly, birds have a unique lymphoid organ that is lacking in mammals, namely the bursa of Fabricius, which is a sack-like organ near the cloaca that regresses during sexual maturation (Kaspers & Schat, 2013). Third, the B cell maturation occurs in the bone marrow in mammals, and in the bursa of Fabricius in birds. This maturation occurs during a brief period, from the late embryonic stage to shortly after hatching. In mammals, diversity in the immunoglobulin (Ig) repertoire is due to rearrangement of the different segments building up the Ig light- and heavy-chains, as described earlier. In birds, Ig variation is mainly generated by somatic gene conversion, where nucleotide segments are replaced between the functional gene and pseudogenes; this is possible since they have several pseudogenes upstream of the functional gene (Arakawa & Buerstedde, 2004; Kaspers & Schat, 2013). In birds, the antibody repertoire only expands until the chick is about 5-7 weeks of age. Once expansion halts and the antibody repertoire has matured the bursa starts to regress. Older birds rely on B cells derived from extra-bursal precursors that are most likely produced by post-bursal stem cells in the bone marrow (Kaspers & Schat, 2013). Because antibody expansion only occurs early in chick development, if any pathogens infect the bursa during this time, of the chick, these birds will be limited in their antibody repertoire and thereby limited in their defense against the pathogens (Kaspers & Schat, 2013). Fourth, the location of the thymus, the site of T cells maturation, differs between mammals and birds. The mammalian thymus is located near the heart whereas the avian thymus is located in the neck (Fig. 7), present as 7-8 lobes on each side of the third cervical vertebra (Kaspers & Schat, 2013).

Fig. 7. Schematic representation of anatomical location of thymus (yellow) and bursa of Fabricius (red) in birds.
**Avian AMPs**

Avian species have a similar AMP repertoire as other vertebrates, with a few exceptions. Ovodefensins are unique to birds and reptiles. Whereas, α- and β-defensin genes found in mammals and primates are not present in birds. The first avian AMPs were isolated from avian heterophils in 1994 by Evans et al. (1994), they isolated five defensins from chicken and turkey leucocytes. Since then, the entire defensin and cathelicidin clusters of several bird species, have been characterized (Cheng et al., 2015).

**Avian AMP families**

The two major families of AMPs in birds are avian β-defensins (AvBDs) and cathelicidins (CATH). Additionally, there are two avian AMP families: ovodefensins found in the oviduct and egg white, and liver-expressed antimicrobial peptide-2 (LEAP-2).

**Cathelicidins**

Four cathelicidin genes have been identified in the chicken genome: CATH1-3 and CATH-B1 (Cheng et al., 2015). Cathelicidins have also been identified in other avian species, including turkey (Meleagris gallopavo), crested ibis (Nipponia nippon) and mallards (Anas platyrhynchos), with copy number variation (Cheng et al., 2015; Gao et al., 2015). Genes encoding chicken cathelicidins are located at the end of chromosome 2 (Zhang & Sunkara, 2014). These genes are comprised of four exons, with the mature peptide coded on the last exon together with a minor part of the highly conserved cathelin-like domain and the 3’UTR. The first three exons code for the 5’UTR, the highly conserved signal peptide, and the majority of the cathelin-like domain (Zhang & Sunkara, 2014). The mature peptides of cathelicidins vary considerably in sequence composition, peptide length and tertiary structure, so cathelicidins are mainly characterized by their propiece, a large cathelin-like domain (Zanetti, 2005).

**LEAP-2**

LEAP-2 is an antimicrobial peptide containing four conserved cysteine residues forming two intramolecular disulfide bonds. It was discovered in the chicken genome by Lynn et al. (2003), and is located on chromosome 13 (Townes et al., 2004). A study showed that LEAP-2 expression is induced in the liver and small intestine by Salmonella infection and that LEAP-2 can inhibit the growth of Salmonella spp. in vitro (Townes et al., 2004). LEAP-2 has also been identified in mallards and shown to have >85% sequence identity with other avian LEAP-
2 peptides (Hong et al., 2019). Disulfide bonds appear to be important for activity, and mallard LEAP-2 seems to be more active against Gram positive (G+) than Gram negative (G-) bacteria (Hong et al., 2019).

**Avian β-defensins**

The first section of this thesis specifically focuses on avian β-defensin genes (AvBDs), which are found in a gene cluster on chromosome 3 (Cuperus et al., 2013; Huang et al., 2013). AvBD genes have been described in all bird species for which genomes have been analyzed (e.g. Hellgren & Ekblom, 2010; Huang et al., 2013; Xiao et al., 2004). The copy number of AvBD genes in different species is somewhat variable, ranging from 13 in the Japanese quail (*Coturnix japononica*) (Morris et al., 2019) to 22 in the zebra finch genome (*Taeniopygia guttata*) (Cheng et al., 2015; Hellgren & Ekblom, 2010). AvBDs are generally comprised of 3-4 exons, where the first exon contains the 5’ UTR, the second exon contains the hydrophobic signal peptide (rich in leucine) and parts of the propiece, and the third exon contains the rest of the propiece and a large section of the mature peptide and the fourth exon contains the remainder of the mature peptide as well as the 3’ UTR (Fig. 8) (van Dijk et al., 2008).

![AvBD Gene Structure](image)

**Fig. 8.** The AvBD gene structure including four exons (E1-E4) that encode the signal peptide (blue), the propiece (red) and the mature peptide (black).

The mature AvBD contains a conserved six cysteine (C) residue motif, CX4-7 CX3-6 CX7-10 CX5-6 CC (where X denotes any other amino acid residue), forming three intramolecular disulfide bonds (Fig. 9) (Cheng et al., 2015). AvBDs are either constitutively expressed or induced upon stimulation (van Dijk et al., 2008; Zhao et al., 2001). Both modes of expression have been shown by Ebers et al. (2009) in primary chicken oviduct epithelial cells. This study showed for example that infection with *Salmonella* rapidly shifted AvBD expression from low basal expression levels to strong upregulation during infection (Ebers et al., 2009). Indeed, low basal expression but rapid infection-induced upregulation seems to be the general pattern for many AMPs across taxa (Chong et al., 2008;
Lemaitre et al., 1997). This pattern of AMP induction only when required for an immune response is likely due to the energetic costs of production, as well as avoiding unwanted effects on commensal organisms in the microbiome (Hanson et al., 2019).

**Fig. 9.** AvBD amino acid sequence of the mature AvBD3b:1 including conserved cysteine residues (boldface) and characteristic β-defensin disulfide bond connectivity (C1-C5, C2-C4 and C3-C6) depicted with the black lines.

**Ovodefensins**

Ovodefensins are cysteine rich peptides highly expressed in the bird oviduct and also in the egg white (albumin). They are present in avian and reptilian species (Hervé et al., 2014; Whenham et al., 2015; Zhang et al., 2019). Ovodefensins have, like β-defensins, a conserved six-cysteine motif, though the number of intervening amino acids differs. Furthermore, like AvBDs, ovodefensins are located on chromosome 3 in the chicken genome (Zhang & Sunkara, 2014). One chicken egg ovodefensin, named gallin, has been synthesized and the three-dimensional structure solved, revealing high structural similarity to avian β-defensins (Hervé et al., 2014). Based on genetic and structural similarities it has been speculated that ovodefensins arose from a β-defensin gene duplication event (Zhang et al., 2019). Additionally, the gallin peptide has been shown to have antimicrobial activity against Gram-negative but not against Gram-positive bacteria (Hervé et al., 2014).

**The mallard (Anas platyrhynchos)**

The main study species in this thesis is the mallard (Fig. 10), a dabbling duck with broad Holarctic distribution (Reeber, 2015). In the northern hemisphere this species is partially migratory, meaning that its propensity to migrate varies between populations and individuals. Northern breeding populations are almost exclusively migratory, while southern populations, to a larger extent are resident year-round (Reeber, 2015). Migration of mallards is usually performed as short flights (1-7 hours), interspersed with long stopovers (days to several weeks) in suitable areas with wetlands (Kleyheeg et al., 2019; van Toor et al., 2018). During stopovers and wintering, aggregation of mallards and other waterfowl is common, providing opportunities for microbial transmission. Mallard feeding
behavior (i.e. drinking, feeding and defecating in the same environment), makes it a good potential host for different microorganisms that can be transmitted in water.

![Male mallard. Picture taken by Michelle Wille, used with permission.](image)

**Diseases**

Microorganisms cause pathogenic infections in mallards just as they do in other animals, including bacterial, viral and fungal diseases. Due to its importance as a game species, and as the origin of domestic ducks, mallard diseases have received more research interest than many other wild bird species. Of particular concern is avian influenza, caused by influenza A virus, which is associated with large disease outbreaks in poultry (Chatziprodromidou et al., 2018; Khokhar et al., 2015). Mallards can also harbor important enterobacterial pathogens such as Escherichia coli (*E. coli*), *Salmonella* spp. and *Campylobacter* spp., as well as other important poultry pathogens such as Newcastle disease virus (Alexander, 2000; Benskin et al., 2009; Fallacara et al., 2001). However, in this thesis one focus is on avian influenza viruses (AIVs).

**Avian influenza viruses**

AIVs are single stranded, negative-sense RNA viruses with segmented genomes. There is a large variety of AIV subtypes, defined by the two surface proteins hemagglutinin (HA, 16 types in birds) and neuraminidases (NA, 9 types in birds) (Kosik & Yewdell, 2019). The combination of HA and NA makes up the subtype, for example H1N5 has type 1 HA and type 5 NA.
Additionally, AIVs are classified as highly pathogenic or low pathogenic, based on their pathogenicity in chicken (Swayne & Suarez, 2000). Mallard and other waterfowl are commonly infected with low pathogenic avian influenza (LPAI) viruses, where they cause no or only mild clinical symptoms in the host (Jourdain et al., 2010) and is manifested primarily as an infection of the lower gastrointestinal tract. This contrasts markedly to highly pathogenic avian influenza (HPAI) viruses, which can cause systemic infections in the respiratory tract often associated with high mortality in domestic poultry, as well as in many wild bird species, and occasionally humans. This shift in pathogenicity follows changes in the HA cleavage site from monobasic in LPAI to polybasic in HPAI (Swayne & Suarez, 2000).
**Aims**

The overall aim of my thesis was to acquire a deeper understanding of the innate immune response in mallards against bacterial and viral pathogens, with an emphasis on β-defensins and the viral response genes RIG-I and Mx.

- **Paper I**, characterizes five AvBD genes in mallards and other waterfowl and investigates diversity and selection acting on these genes.

- **Paper II**, examines the antimicrobial activity of three AvBDs against Gram-positive and Gram-negative bacteria and compares antimicrobial activity between three naturally occurring variants of the same gene.

- **Paper III**, compares the structure and function of three naturally occurring AvBD3b variants.

- **Paper IV**, highlights the importance of validation of reference genes in gene expression studies.

- **Papers V and VI**, investigates gene expression of RIG-I and Mx during LPAI infection in mallards.
Brief methodological overview

Samples for allelic variation and selection study

As part of my studies, I investigated genetic variation of AvBD genes from waterfowl (see Paper I). In order to do so, 274 blood samples were collected from a population of migratory mallards trapped (Fig. 11) for ringing purposes at Ottenby Bird Observatory on the southern tip of Öland (56°12’N, 16°24’E), Sweden from 2012 - 2014. Additionally, 190 samples from 16 other mallard populations across the global range were collected and two samples from each of 43 different species from the Anatidae family (ducks, geese, swans) were collected. For more details see Paper I.

Fig. 11. Duck trap at Ottenby Bird Observatory, Öland, Sweden. Picture taken by Michelle Wille, used with permission.
Allelic variation and selection acting on AvBDs

To determine the allelic variation and to infer mode of selection acting on five AvBD genes (AvBD3b, AvBD4, AvBD5, AvBD10 and AvBD13) in mallards and other waterfowl species (Paper I) DNA was extracted using a standard ammonium acetate precipitation protocol. DNA was then amplified using polymerase chain reaction (PCR) with gene-specific primers and Sanger sequenced. Sequences were analyzed using the Geneious software (Biomatters, Auckland, New Zealand). Amplicons from individuals that were heterozygous at two or more positions were cloned using pGEM®-T Easy Vector System (Promega) and re-sequenced. Sequences were aligned and trimmed to only contain the nucleotide sequence encoding the mature peptide, which was used for the evolutionary analyses. DNA sequences were translated to amino acid sequences and analyzed using analytical tools available on the DataMonkey server (Pond & Frost, 2005), DnaSP (Librado & Rozas, 2009) and PAML (Xu & Yang, 2013). These analyses were performed at three increasingly broad levels: Swedish (Ottenby) mallards; global mallards; all waterfowl.

Oxidative folding of AvBDs

Using mallard AvBD amino acid sequences elucidated in Paper I, we then obtained synthetic AvBDs and proceeded to fold them into their native conformation and test their antimicrobial activity (Papers II and III). Synthetic linear forms of AvBD3b (three variants), AvBD4, AvBD10 and AvBD13 (three variants) were purchased, and exposed to a series of different oxidative folding experiments. The final protocol for folding of peptides was conducted using reduced and oxidized glutathione in a Tris-HCl buffer overnight with slow agitation. Peptides were then purified using high performance liquid chromatography (HPLC). Peptide purity and mass were determined via liquid chromatography–mass spectrometry (LC-MS), thereafter the peptides were lyophilized and frozen for later use. The three-dimensional structure was assessed using nuclear magnetic resonance (NMR) and 2D NMR. Because of technical challenges, only AvBD3b variants were successfully folded and subjected to NMR, allowing to solve the structure for one peptide encoded by an allele that is common among mallards globally, while the structures of the two other AvBD3b variants were inferred with reference to the previously solved structure, using I-TASSER protocols (Yang & Zhang, 2015). Several attempts were made to optimize the oxidative folding of remaining peptides using different concentrations and ratios of reduced and oxidized glutathione as well as supplementing the reaction buffer with substances such as dimethylsulfoxide (DMSO). However, despite these efforts, remaining peptides
could not be folded within the time frame of this project, why these were evaluated in their linear form.

Antimicrobial assays of AvBDs

Antimicrobial assays are experimental setups employed to evaluate biological activity of chemical substances. Two different assays were used to evaluate the antimicrobial effect of AvBDs from mallard. The first assay was a modified version of the radial diffusion assay (also known as the zone of inhibition assay, ZOI) (Lehrer et al., 1991). Radial diffusion assay is an agar-based assay where the bacteria grows within a bottom layer of nutrient-poor agar and AvBDs, at different concentrations, are added to small holes punched through the agar (Fig. 12A). The plates are incubated for three hours to allow bactericidal activity, after which a nutrient-rich top layer of agar is then added and the plates are incubated overnight. If the peptide is able to inhibit bacteria from growing on the plate, inhibition zones will be observed, and the diameter of zones provides a measure of antibacterial activity. As such, inhibition zones were measured in triplicate using calipers and plates were photographed after ~12-18 hours of incubation. For a more detailed description, see Paper III. The second method was a microdilution assay based on Varkey and Nagaraj (2005). This assay allows determination of the minimal inhibitory concentration (MIC) of the AvBDs. It is a buffer-based assay where the peptide and bacteria are incubated in a volume of buffer for 2 hours with slow agitation in a 96-well low attachment microtiter plate (Fig. 12B). Thereafter the content of the wells is diluted (such that a countable number of colonies will grow) and spread on agar plates. Plates are incubated overnight and colonies counted. For a more detailed description, see Paper II.

Fig. 12. Overview of the experimental setup including sample distribution and dilution in (A) radial diffusion assay and (B) micro dilution assay, respectively. (A) Colored dots represent different peptide concentrations. (B) Blue dots show the evaporation barrier. Plates were divided in two to enable simultaneous analysis of different AvBD peptides represented by the line in both A and B.
Validation of reference genes for gene expression studies using qPCR

When quantifying gene expression with real-time quantitative PCR (hereafter qPCR), it is important to use stable reference genes (RGs) to normalize the expression level of the tested gene. Therefore, potential RGs should be validated prior to use in experiments. A good RG should be stably expressed in the tissue independent of treatment (e.g. infection status). Furthermore, more than one RG is required for optimal experimental results, and the number of RGs needed to obtain a reliable gene expression base line should be evaluated experimentally (Bustin et al., 2009). Validation of RGs is an often overlooked step in many gene expression studies (Chapman & Waldenström, 2015). Different algorithms can be used to analyze gene expression data, such as GeNorm (Vandesompele et al., 2002) and NormFinder (Andersen et al., 2004), which calculate stability of the RGs; GeNorm additionally assesses the minimal number of RGs required to provide a reliable reference. More detailed descriptions regarding why and how to validate RGs, along with RG stability results for LPAI infected duck tissues, are provided in Paper IV.

LPAI infection experiment in mallards

Thirty-three naïve juvenile male mallards were used in this experiment, which was conducted at the National Veterinary Institute (SVA) in Uppsala. Three mallards were artificially inoculated with an H1N1 LPAI strain, and three days later placed in a room containing 25 mallards to allow semi-natural infection to occur via food and water shared between birds (LPAI is transmitted via the fecal-oral route). An additional five mallards were housed separately and kept pathogen-free to serve as controls, more details of the experimental set-up are provided in Paper V. In order to investigate expression changes of two innate immune genes during a LPAI infection in different tissues, the mallards were euthanized over the time-course of infection, different tissues were collected and snap frozen in liquid nitrogen to preserve the mRNA. Five mallards were euthanized at 0.5, 1, 2, 4, and 7 days post infection (dpi, where \( t = 0 \) was the point at which the three inoculated mallards were introduced to the room housing the naïve mallards).

Gene expression of \textit{RIG-I} and \textit{Mx}

Gene expression of \textit{RIG-I} and \textit{Mx} in different mallard tissues during a LPAI infection was performed by first extracting mRNA using the RiboPure Kit
(Ambion) from tissues and blood thawed in RNAlater-ICE (Ambion) to keep the RNA from degrading. To remove contaminating genomic DNA (gDNA), extracted mRNA was treated with DNase and DNA degradation verified on an agarose gel containing 1% bleach (Aranda et al., 2012). cDNA was synthesized from the mRNA via reverse transcription using Superscript III (Invitrogen) and random hexamers (Invitrogen). Gene expression levels were measured using qPCR and amplification values normalized against two to three previously established RGs (in Paper IV). For more details of methods and results see Papers V and VI.
Results and discussion

One way of understanding the role of an innate immune gene is to investigate the genetic variation (allele frequencies) in a population and thereafter testing the antimicrobial activity of detected variants. To this end, I quantified genetic variation in five AvBD genes from wild mallards (and other Anatidae) (Paper I) and tested the antimicrobial activity of peptides encoded by these naturally occurring gene variants (Papers II and III). In addition, the importance of AvBD peptide structure conformation for antimicrobial activity was investigated (Paper III).

Evolution and selection on waterfowl β-defensins

In natural mallard populations, genetic variation is expected to be reasonably high due to large population sizes and migratory connectivity (Čížková et al., 2012; Kraus et al., 2011). Diploid organisms have two copies of the same gene, on autosomal chromosomes, with one located on each chromosome. If the nucleotide sequence in the two gene copies is identical, the individual is homozygous (i.e. it possesses a single allele of that gene), but if there is sequence variation between the gene copies due to mutations (i.e. it possesses two alleles of the gene), the individual is heterozygous. Because AvBDs are relatively short, the main types of mutations that accumulate are single nucleotide polymorphisms (SNPs, one to several per gene), although longer insertions and deletions (indels) can occur. When SNPs occur in the first or second position of a codon, substitutions are often non-synonymous, in other words a different amino acid is encoded at that position. Alternatively, SNPs can be synonymous, especially those that occur at the third position of a codon. Such substitutions are also sometimes known as silent mutations because they do not change the amino acid sequence, although this is increasingly recognized as a misnomer (Parmley & Hurst, 2007). For the remainder of this thesis unless
otherwise specified, ‘allele’ and ‘variant’ will be used to denote alternate forms of the same gene which differ at the amino acid level, with allele being used for genes and variant being used for the synthetic protein (peptide) products of genes.

In Paper I, the aim was to characterize naturally occurring genetic diversity in five different AvBD genes, \( AvBD3b, AvBD4, AvBD5, AvBD10 \) and \( AvBD13 \), in mallards sampled across 16 populations in Europe, Asia and North America, as well as from 43 other species of Anatidae (ducks, geese and swans). Moreover, this diversity was investigated in the “local” mallard population (sampled at Ottenby) alone and compared to the combined diversity in all mallards to identify population genetic structure in AvBD genes (Paper I). First, sequence analyses of AvBD genes in the local mallard population, revealed a substantial nucleotide (nt) allelic variation, ranging from seven to as many as 44 nt alleles per gene (Fig. 13). However, the majority of observed genetic variation consisted of synonymous mutations, resulting in lower sequence diversity in AvBD amino acid (aa) alleles. For example, analyses revealed 44 different \( AvBD10 \) nt alleles among sampled mallards, but only seven of these mutations were nonsynonymous, thus \( AvBD10 \) aa allelic diversity was quite low (7 alleles), and on par with \( AvBD3b \) (6 alleles), \( AvBD4 \) (4 alleles), \( AvBD5 \) (2 alleles) and \( AvBD13 \) (5 alleles).

Next, observed AvBD allele frequencies were used to infer the mode of selection acting on these five AvBD genes. Four of the genes (\( AvBD4, AvBD5, AvBD10 \) and \( AvBD13 \)), had a strong signal of purifying selection. Indeed, for each of these four genes, the vast majority of mallard individuals, both locally and globally, was found to be homozygous for a single aa allele at each locus. In contrast, evidence was found for balancing selection acting on \( AvBD3b \) such that two aa alleles were maintained at similar frequencies in local and global mallard populations. In Swedish mallards, the nt allele coding for the most common aa allele variant, was present in 69% of individuals whereas the second common allele was present in 42%. Only 46% and 25% of individuals were

![Fig. 13. Allelic variation in the Swedish mallards including most frequently detected alleles of each AvBD gene indicated in red and blue, respectively. The number on the right-hand side of each pie chart represent the total number of alleles detected of that particular gene. Figure modified from Paper I.](image-url)
homozygous for the first and second allele variant, respectively. Balancing selection acting on AMPs has been shown previously (e.g. Chapman et al., 2019; Hellgren & Sheldon, 2011; Hollox & Armour, 2008). AvBD3b belongs to the AvBD3 gene cluster which has been suggested to have undergone recent lineage specific duplication events, and contains six different genes in mallards (Cheng et al., 2015). For all five mallard AvBD genes investigated, hydrophobicity, based on amino acid residue composition, was broadly similar between alleles of a particular gene, indicating that there might be functional constrains actively maintaining this feature.

Furthermore, PCR primers designed for five AvBD genes in mallards were used to amplify corresponding genes in 43 waterfowl species. For this, two individuals per species were used (see Paper I). Amongst sampled species, a minimum of two out of five AvBD genes were amplified. For all five genes identical aa alleles were found across species. Furthermore, the most common mallard aa allele was also the most widely shared between species of waterfowl for AvBD4, AvBD5 and AvBD10. The most common mallard AvBD13 aa allele was only present in one closely related species (Anas crecca) and the most widely shared AvBD13 aa allele across waterfowl was absent from mallards. AvBD3b had the lowest level of allele sharing across species, with the most common mallard allele only being found in two other species, both of which were closely related dabbling ducks. The other common (balanced) allele was not found in any other species.

The fact that specific AvBD alleles were found in multiple waterfowl species, and in globally dispersed mallard populations, suggests that these alleles are strongly selected, perhaps due to them having optimal immune activity. The conservation of certain alleles across discrete mallard populations, and even across waterfowl species, could be due to similarities in selective pressures, such as an overlapping pathogen fauna. To determine whether these differences in selective pressures acting on AMPs could be attributed to the bactericidal efficacy of specific alleles, antimicrobial properties were investigated in vitro (see Papers II and III below).

**Antimicrobial activity of AvBDs**

Findings in Paper I led to the investigation of whether different naturally occurring aa alleles of AvBDs differ in their antimicrobial properties, and whether there is a relationship between peptide structure and function. In vitro assays were used to measure antimicrobial activity of synthetic AvBDs to Gram-negative (G-) and Gram-positive (G+) bacteria. The first of these studies (Paper II), relied on reduced, linear forms of AvBD4, AvBD10 and three
variants of AvBD13 (the most common peptide variant and two rare variants). Peptides were tested against two G- bacteria (*Escherichia coli* and *Salmonella enterica* serovar Typhimurium) and two G+ bacteria (*Staphylococcus aureus* and *Micrococcus luteus*), using a microdilution assay to determine the concentration range where peptides are able to inhibit bacterial growth. Assay results revealed distinct differences in antibacterial activity for AvBDs, where AvBD4 inhibited growth of both G- and G+ bacteria at concentrations below 3 µM (Fig. 14), whereas AvBD10 did not have any effect on these bacterial species, even at concentrations as high as 128 µM.

![Peptide conc. (µM)](image)

*Fig. 14. Antibacterial activity of the AvBD4 against two G- (dark blue) *E. coli* and (light blue) *S. Typhimurium*, and two G+ (dark green) *S. aureus* and (light green) *M. luteus*. Dashed line indicates IC50 point. Picture from Paper II.*

Interestingly, theoretical models for the mode of action of β-defensins, whereby positively charged residues are attracted to negatively charged bacterial membranes, appear to be consistent with results in Paper II, such that the peptide with the highest net positive charge (AvBD4, +7.8) was the most potent and the peptide with the lowest net charge (AvBD10, +2.7) showed no activity. Of the AvBD13 variants, the one with the highest net positive charge (AvBD13:2, +5) was also the one that had the strongest activity against bacteria as compared to the other two variants (AvBD13:1 and AvBD13:3, both +4). It is somewhat surprising that the peptide variant most commonly found in mallards (AvBD13:1) did not display the strongest bactericidal effect on bacteria. However, it is possible that the selective pressure maintaining this allele at high frequencies is exerted by a pathogen that was not tested here. Lastly, the results showed that AvBD4 and AvBD13 were more effective against G- than G+ bacteria. It has previously been shown that AvBDs are active against both G- and G+ bacteria, but also, in agreement with results presented in Papers I and II, that different AvBDs have distinct activity profiles against different bacterial species (Hellgren et al., 2010; Lee et al., 2016; Yacoub et al.,
In a previous study of chicken AvBDs, antimicrobial activity testing included the same *E. coli* strain as that used in Paper II, and reported results similar to those for AvBD4 and AvBD10 peptides from mallards, with AvBD4 being more active than AvBD10 (Lee et al., 2016).

In the second study on antimicrobial activity (Paper III), the relationship between structure and activity of three AvBD3b variants was investigated. Results presented in Paper I revealed that the *AvBD3b* gene is subject to balancing selection in mallards. Therefore, two peptides encoded by the most frequently detected alleles in mallard (AvBD3b:1 and AvBD3b:2) and a third peptide encoded by a rare allele (AvBD3b:3) were included. Peptides were synthesized commercially and oxidized to induce folding. Tertiary structure and correct cysteine connectivity of AvBD3b:1 was confirmed using nuclear magnetic resonance (NMR) (Fig. 15). The tertiary structure of the other two variants were inferred using the structure of AvBD3b:1 as reference. To date, three-dimensional structures of only seven different AvBDs, from two bird species, have been determined (www.rcsb.org/pdb). Therefore, the structure of mallard AvBD3b described in Paper III adds to the overall understanding of AvBD structure. Antibacterial activity of folded and linear variants of AvBD3b:1, AvBD3b:2 and AvBD3b:3 peptides were analyzed by radial diffusion and microdilution assays. The radial diffusion assay showed that the zones of inhibition (ZOIs) on plates with *E. coli* (G-) were consistently larger for assays including folded peptides than in assays with linear peptides. This difference in activity between folded and linear peptides was not observed in assays with *S. aureus* (G+). In general, both linear and folded AvBD3b:1, AvBD3b:2 and AvBD3b:3 were more active against *E. coli* than *S. aureus*. Somewhat similar results were observed in the microdilution assay, such that the peptides were more active against *E. coli*. However, in this assay, differences between linear and folded peptides were not detected. A higher net positive charge appeared to result in higher antibacterial activity, with AvBD3b:3 (+6.7) generally being more active than AvBD3b:1 and AvBD3b:2 (+5.6), in agreement with observations made in Paper II.
The results from Papers II and III shed some new light to our understanding of the structure-activity relationships of AvBDs, and correspond to findings in previous work, which have showed that the activity of β-defensins against bacteria seems to depend less on tertiary structure including disulfide bond configuration than on other factors such as charge (Yang et al., 2016). Disulfide bonds/overall fold may, however, be important for other immunological functions, such as AvBD chemotactic activity (Wu et al., 2003; Yang et al., 2016), or simply to stabilize the peptide once it is secreted to the extracellular matrix (Martin et al., 1998), and to protect against enzymatic degradation.

Gene expression and infection

Another way to understand the role of specific genes in innate immune responses is to follow how their expression changes during the course of an infection. There are several ways one can approach this, such as microarray, qPCR or RNA-seq. In the second section of this thesis, qPCR was used to follow the expression of two innate immune genes in mallards after infection with AIV. Paper IV is a methodological paper describing the selection of appropriate reference genes, while studies of expression of RIG-I and Mx following LPAI H1N1 infection are described in Papers V and VI.

Validation of reference genes

When performing gene expression studies using quantitative real-time PCR (qPCR), a crucial and often overlooked first step is to validate the reference genes (RGs) used for normalization of expression levels. qPCR-based gene expression studies in which a single, unvalidated RG is used, across tissue types and experimental conditions, is commonly seen in the literature (Chapman &
Waldenström, 2015). According to the MIQE guidelines (Bustin et al., 2009), it is not recommended to normalize gene expression signals against only one RG, and selection of RGs should be based on experimental evidence of stable expression (measured as Ct value) of RGs in different sample types of interest (e.g. uninfected and infected individuals, sexes, tissue types).

In Paper IV, the stability of 11 different candidate RGs was assessed by qPCR using mRNA extracted from six different tissue types in mallards during the time-course of LPAI infection. Gene expression based on mRNA amplification of RGs was evaluated using the GeNorm and NormFinder algorithms, enabling quantification of RG stability, as well as the optimal number of RGs, per tissue. Interestingly, the two RGs most widely used (and usually unvalidated) in qPCR studies, ActB and GAPDH (Chapman & Waldenström, 2015), was not among the genes found most suitable in the different tissue types in Paper IV.

LPAI infection in mallards and gene expression

As described in the introduction, mallards are reservoir hosts for a variety of pathogens, one which is avian influenza viruses (AIVs). Mallards harbor many different AIV subtypes, with the majority being low pathogenic avian influenza (LPAI) viruses; these circulate in wild waterfowl all year round without causing any major overt clinical symptoms. LPAI is mainly a gastrointestinal viral infection in mallard where it replicates in the cells lining the intestinal tract and is excreted with the feces. In order for the mallard host to mount an immune response to AIV, the pathogen must first be detected upon host cellular entry. An important component in this process is the PRR gene RIG-I. It is an RNA sensor found in the cytoplasm that recognizes viral RNA and initiates a signaling cascade that ultimately leads to the activation of an important ISG. One of these genes, Mx, encodes a protein that is important during AIV infection by inhibiting the transcription of virus particles in the nucleus.

In Papers V and VI, the expressional changes of RIG-I and Mx in mallards infected with LPAI over a seven-day period was investigated. Two to three of the RGs described in Paper IV were used, per tissue type, to normalize mRNA expression levels of RIG-I and Mx as measured by qPCR. Gene expression in blood, spleen, upper and lower gastrointestinal tract, and colon were measured. RIG-I showed similar results in all tested tissues with moderate (9-21 fold) upregulation 1 day post infection (dpi), thereafter returning to basal expression levels by 2-4 dpi. Similar results were found for expression of Mx, with strong upregulation detected 1 dpi followed by a return to basal expression levels by 2-4 dpi. Interestingly, upregulation of Mx expression was much stronger in spleen (107-fold, Fig. 16) compared to other tissues (<16-fold), while mRNA
from *Mx* was not detected at all in blood samples from mallard (neither in infected nor uninfected individuals). In birds, the spleen is the site of lymphocyte production and storage (Smith & Hunt, 2004). Lymphocytes are responsible for both humoral and cellular responses. As such, strong upregulation of *RIG-I* in spleen may be tied to a rapid increase in lymphocyte production after LPAI infection. The fact that the timing of upregulation of *RIG-I* and *Mx* was tightly synchronized was not unexpected, given that *RIG-I* expression triggers *Mx* expression. However, it was somewhat surprising that upregulation of these two genes was so transient, given that mallards continued to shed viral particles throughout the course of the experiment, suggesting a possible need for a sustained immune response. It may be that after the earliest stage of infection, other genes related to innate immunity are activated.

![Graph](image)

**Fig. 16.** *Mx* gene expression in spleen of mallards infected with LPAI. Expression was measured as fold change (y axis) during the course of infection (x axis). Error bars show the upper 95% confidence interval of the mean expression value and the asterisks show significant difference as compared to time point 0 (control).
Conclusions

The findings presented in the first section of this thesis adds to our knowledge regarding genetic diversity in five AvBD genes in mallards and other waterfowl, including allele frequencies and modes of selection that shaped these population frequencies. Identification of AvBD alleles that are shared in Anatidae species which split from each other as much as 49 million years ago, is particularly interesting, and implies that these specific alleles have a considerable functional importance. Moreover, results presented in this thesis show that peptides derived from AvBD alleles found in wild waterfowl inhibit growth of G- and G+ bacteria in a dose dependent manner. Interestingly, peptide encoded by the most frequently detected AvBD13 allele in mallards showed weaker antibacterial activity compared to peptides derived from rare alleles. This thesis also demonstrates that peptides with higher net positive charge are more active against G- and G+ bacteria, and also that AvBD peptides were, in general, more potent against G- bacteria.

This thesis also provides the first three-dimensional structure determined for mallard AvBD. AvBD3b:1 linear peptide was subjected to oxidative folding and the structure was determined by NMR. The structure includes characteristic β-defensin features including antiparallel β-sheets and disulfide bond configuration. However, structure-functional comparisons of AvBD3bs showed that folding, including disulfide bond formation, is not crucial for the antibacterial activity of these peptides.

The second section of this thesis provides new insights concerning expression of immune-related RIG-I and Mx genes in different tissue types in mallard during LPAI infection. These genes are part of the innate immune system, and their expression from these was upregulated in mallards within 24 hours of infection. However, this upregulation was highly transient, indicating that RIG-I and Mx proteins are only required during the initial stages of an immune
response against LPAI. In addition, expression levels of these immune-related genes were normalized to a number of reference genes, and a methodology for finding and evaluating these genes was established in this thesis.

Taken together, this thesis provides several new insights concerning the innate immune system of mallards, and have identified a number of interesting avenues for future investigation. It has also helped to demonstrate that it is possible to move beyond model species when studying innate immune genes. Indeed, increasing our understanding of immunity in wild species, especially those that act as vectors for pathogenic zoonotic diseases, is an important step towards controlling future outbreaks of disease.
Acknowledgements

“The main thing in life is to know your own mind”
- Snufkin

This thesis would not have been possible without the help and support from a selection of people, within and outside of this academic bubble.

I want to start by thanking my supervisor Jonas for giving me the opportunity to take on this adventure called a PhD. This journey has been somewhat of a roller-coaster, I think you can agree! It really has had its ups and downs all from nice publications and conferences to peptides and bacteria not behaving in the lab as wished, and not to forget a little hurdle at a critical point for me, the move, but I finally made it! Jonas you have given me the opportunity to grow academically and to learn to work independently.

Jo my co-supervisor, what should I say? If it hadn’t been for you I probably would not have continued with a MSc and then a PhD. It was during my last year as a bachelor when I decided that I wanted to continue studying and do a PhD, thanks to all the cool lab work I was doing during my exam work. Even though some people advised me to not take on a PhD before really thinking it over, which I did, and still decided to go on and do a MSc and then apply for a PhD position. In the beginning you were there to support me, in person, but then it was time for you to move on and you moved to the US. This was a bit of an eye wakening for me! Suddenly I was standing on my own without you as a safety net. We continued to skype at (more or less) regular basis and I had to grow and be even more independent, thank you for believing in me!

Then we have Håkan my second co-supervisor, you have helped me with the peptide chemistry parts, and believe me, I really do appreciate it! It has not been easy, for any of us (I guess) with all the setbacks. This peptide folding was supposed to be “easy” and straight forward but guess what? It wasn’t! Many long hours figuring out why things don’t work or what to change to get them to
work. This took so much longer time then what was expected and made me so furious so many times, but in the end, we managed to fold and solve a peptide structure! So, it wasn’t for nothing, though it has been a long uphill run.

I really want to show my appreciation for the rest of the zoonotic ecology and epidemiology group (or whatever the name of our research group is nowadays), previous and present people, for listening to my issues and trying to cheer me up when things felt tuff and for eating all the baked goods a dragged with me to work so I didn’t need to eat them on my own. Previous group members; Abtesaim, Alexis, Daniel, Michelle, Jo, Sara, Jacintha and Sjoerd, Martin, Richard, Josanne and Neus (even though your kind of are still in our group) and then we have the present lab group Conny, Håkan, Marielle and Jenny, thank you all for nice conversations, fikas and everything what you have done for me. A special THANKS to Conny for all your support through my PhD and most of all during the last year or two, I would have never made it without you! Jenny “I declare!” thank you for all your support in the lab and outside. I know I haven’t always been the happiest and easiest person, but you two (Conny and Jenny) have always kept me going and dragging me to fika and lunches so thank you!

Then we have the morning coffee superheroes, Carina, Christofer, Elias, Emil, Fredrik, Håkan and Mariëlle, what a way to start the morning! This crew is fluctuating, with people coming and going. This awesome crew has contributed to a very important start of the day for me during my years as a PhD student containing very interesting discussion about all kinds of things ;) The core crew, that always was there, Emil and Fredrik (and Carina on WhatsApp) you have made the tuff times the last year or two at least a bit easier! We have had great discussion about aquariums, Star Wars, horses and all kind of stuff that I never heard of and I’ve learned stuff I didn’t know I even needed ;) Christofer and Emil, you guys have really been there for me the past year in both happy and annoying discussions about finishing a PhD and all what that includes, it feels good to have had you with me through this roller-coaster of emotions and energy.

The 9 o’clock fika group with mainly people from our lab group and then of course Camilla and Sabina that always have a smile to share, thank you for always being there with a smile and something nice to talk about.

Carina you deserve an additional thanks for everything you have done for me through the years. You know me very well and have supported me in so many ways and I really wish you had been here the past year, but WhatsApp is better than nothing! ;)

I declare!
All the cool people that worked here when I started that moved away one after the other, started by Michelle leaving and then Anna S, Jo, Mireia, Carina and even Stephan left me! You all made my start here a great place and allowed me to join for Friday beers so thank you all!

I would also like to give a thanks to all the members in EEMiS and the PhD council group, these are awesome ways to interact with other PhD student at the University. There have been many people making a mark, in one way or another, during my years as a PhD and making it an overall great experience. Thank you: Andreas, Ben, Camilla, Clara PM, Clara VC, Laura S, Stephanie, Sabina, Hanna B, Hanna Bench, Hanna F, Elin, Emelie, Johanna, Henrik, Gaofeng, Fredrik, Emil, Eva, Mariëlle, Håkan, Oscar, Richard, Maurice, Sina, Laura F, Diego, Yeserin, Daniela, George, Sören, Furrong, Devi, Stina, Xiafen, Stephan, Anders M, Henrik H, Leon, Stefan, Nina, Therese, Fatima and everyone that I might have forgot to write.

A special thanks to the mentor program, MP7, it was NOT always easy but it definitely has helped me to talk about all kinds of things and it’s good to talk to people that are in the same situation. We all share similar issues and difficulties and to listen to others and discuss things makes it easier to cope with them by yourself. Thanks to the group from here that was sharing this with me; Ben, Christofer, Eva, Gaofeng and Emelie from Kalmar and Oscar, Shakil and Yihan from Lund. I also would like to show my appreciation to my mentor, Päivi, who I had very nice meetings with discussing all kinds of issues that I had and didn’t always feel I could talk to anyone else about.

Then we have the people in Lund during GENECO stuff and dissertations, thank you for always making me feel welcome, and what a good way to interact with people that are facing the same problems as you. Anna Drews we have had a lot of fun! I’m so glad to have learned to know you, thank you for being there for me and letting me ventilate when I felt like exploding!

This has been a long journey and there are few people that have been through the whole journey with me, more or less, since the bachelor times to now! I really would like to thank you Christofer (mentioned for the fourth time but you have been very helpful and cheerful during these years) and Natalie that you both have been around and supported me through it all. Not to forget Fredrik you have been there too through it all and I’m really grateful for that.

I want to show my appreciation to all the Auntnado girls that I got to know during my short period in Uppsala. You girls have been there for me through all that’s happen the past years and supported and encouraged me to finish the PhD.
So, thank you: Denise, Isa, Lena, Livia, Maike, Melissa and Kjersti for all the fun we have had and all the support you have given me!

My warmest thanks to Emma and Daniel for taking me away from Kalmar and doing a lot of fun things, even when I didn’t think I had time or energy, very much appreciated! Thanks also to the whole music quiz team “Lindas med Z” what a way to spend many Thursday evenings and to forget about the life of a PhD for a while.

Ultimately, I would like to thank my family without them I would have never made it so far. They have supported me in everything I’ve ever done!
Tärkeimmät jätetään viimeiseksi! Nyt on aika kiittää rakasta perhettäni! Kiitos että olette aina ollut tukena ja turvana, vaikka ette aina ole ihan ymmärtäneet mitä minä oikein hommaan mutta kuitenkin antaneet minun tehdä mitä minä teen. Rakas äiti ja isi, mitä minä olisin tehnyt ilman teidän tukea, olette aina olleet tukena ja tosi positiivisina minun valintaa jatkaa ”opiskelua” näin pitkään. Olette saaneet kuunnella minun motkotusta, kun kaikki työhommatt on mennyt päin honka. Suuri kiitos että olette jaksaneet olla minun tukena!

Rakas siskoni Tiina, mitäs sitä kirjottaisi? Kiitos että olet jaksanut kuunnella monta pitkää tuntia minun valitustani kaikesta ja ei aina vaan työhommista. Olen tosi kiitollinen, että olet uskonut minun ja siihen että minä pystyn suorittaa väitöskirjani.

Jani, rakas veljeni sinä et aina ole ymmärtänyt miten minä voi tehdä niin paljon tötä niin vähällä palkalla, vaikka olen yrittänyt selostaa, että ei tätä hommaa tehdä sen takia että tulisi rikkaaksi. Kuitenkin olet kuunnellut ja antanut tukea ja yrittänyt rohkeutta, että kyllä se siitä vielä vaikka rankalta tuntuu ;) Suuri kiitos myös sinulle Pauliina! Me olemme tehneet paljon juttuja yhdessä pienenä pitäen ja sinä olet myös ollut mukana tässä reissussa ;) vaikka vähän etäisyydellä. Olen kiitollinen, että olet jaksanut kuunnella minun ongelmia, vaikka omiakin on välillä ollut.

Haluan myös kiittää muuta perhettä ja pyydän anteeksi siitä, että suomenkielenkirjoitus ei enää suju ihan niin hyvin kuin ennen.
References


Hellgren, O., Sheldon, B. C., & Buckling, A. (2010). *In vitro* tests of natural allelic variation of innate immune genes (avian β-defensins) reveal functional differences in


Huang, Y., Li, Y., Burt, D. W., Chen, H., Zhang, Y., Qian, W., . . . Li, N. (2013). The duck genome and transcriptome provide insight into an avian influenza virus reservoir species. *Nat Genet*, 45(7), 776-783. doi:10.1038/ng.2657


