IDENTIFICATION OF NOVEL INDUCERS FOR LL-37 EXPRESSION - ASSAY DEVELOPMENT AND MECHANISTIC STUDIES

Frank Nylén

Stockholm 2016
Identification of Novel Inducers for LL-37 Expression - Assay Development and Mechanistic Studies

THESIS FOR DOCTORAL DEGREE (Ph.D.)

By

Frank Nylén

Principal Supervisor:
Associate Professor Peter Bergman
Karolinska Institutet
Department of Laboratory medicine
Division of Clinical Microbiology

Opponent:
Professor Henk Haagsman
Utrecht University
Department of Infectious Diseases and Immunology
Division of Molecular Host Defense
Utrecht, The Netherlands

Co-supervisor(s):
Professor Birgitta Agerberth
Karolinska Institutet
Department of Laboratory medicine
Division of Clinical Microbiology

Examination Board:
Professor Petter Höglund
Karolinska Institutet
Department of Medicine
Center for Hematology and Regenerative Medicine

Professor Guðmundur Hrafn Guðmundsson
University of Iceland
Biomedical Center
Reykjavik, Iceland

Professor Jan Johansson
Karolinska Institutet
Department of Neurobiology, Care Sciences and Society
Division of Neurogeriatrics

Professor Ulrich Theopold
Stockholm University
Department of Molecular bioscience,
The Wenner-Gren Institute
Till mamma och pappa
ABSTRACT

The remaining reservoir of effective antibiotics is running out and more and more pathogens are gaining resistance, even to the last line of antibiotics. Therefore, novel treatment regimens against bacterial infections are urgently needed. Antimicrobial peptides (AMPs) are expressed at mucosal surfaces and protect the host against invasive microbes. Not only are they efficient microbial killers but they also orchestrate additional immune responses, including chemotaxis and wound healing. The focus in the current work has been on the human cathelicidin LL-37 (encoded by the CAMP gene), which can be considered as a marker for the AMP-system in general. Since several AMPs with different mechanism of action are released simultaneously, the risk of developing bacterial resistance is very low. Here we propose a concept where drug-like molecules that induce the expression of AMPs can be used to prevent or treat infections. We suggest that this novel approach designated ‘Host Directed Therapy’ can be used alone or together with traditional antibiotics for treatment of infections.

The overall aim with this thesis was to identify novel AMP inducing compounds and to dissect the regulatory pathways involved in their mechanism of action.

We set out to establish a cell-based reporter-gene-assay under the control of regulatory elements of the CAMP gene. Following a successful evaluation for high throughput screening purposes, six out of 1200 tested compounds were validated as novel inducers. One of these was the histone deacetylase (HDAC) inhibitor entinostat, which was found to be a potent inducer of LL-37 expression. Several derivatives of entinostat were synthesized and evaluated for LL-37 inducing activity. This strategy demonstrated which part of the molecule that was important for the induction.

Entinostat was shown to regulate LL-37 via activation of the transcription factors STAT3 and HIF-1α. Notably, HIF-1α was found to directly bind and initiate transcription of the CAMP gene. In contrast, no direct binding of STAT3 to the CAMP gene promoter could be demonstrated. However, STAT3 was clearly involved in the regulation of LL-37 expression, since macrophages from a patient carrying a mutated STAT3 gene were unable to upregulate LL-37 upon entinostat treatment.

Several additional regulatory pathways were discovered utilizing label free quantitative mass spectrometry, where protein levels were compared between unstimulated cells and cells treated with the AMP-inducing compounds lactose and phenylbutyrate. Next, the proteomic profiles were mapped to known pathways utilizing bioinformatics tools. The identified pathways were validated using the cell based reporter assay. The thyroid hormone receptor pathway, the eicosanoid signaling pathway and the steroid biosynthesis pathway could all be attributed to LL-37 expression.

Together these results indicate that AMPs are strictly regulated both in homeostasis and during infection. The novel AMP-inducing compounds identified here should be further evaluated as Host Directed Therapy against infections in cellular and animal models as well as in relevant disease cohorts.
LIST OF SCIENTIFIC PAPERS

The following articles form the basis of this thesis and are referred to with their corresponding Roman numeral.

   *These authors contributed equally to this study*


   *#These authors contributed equally to this study*

## CONTENTS

1 Background..........................................................................................................................1
  1.1 Introduction....................................................................................................................1
  1.2 The immune system........................................................................................................1
  1.3 Innate immunity.............................................................................................................1
    1.3.1 Innate sensing ..........................................................................................................2
    1.3.2 Innate effectors ......................................................................................................3
    1.3.3 Innate immune cells ...............................................................................................10
  1.4 Gene regulation.............................................................................................................12
    1.4.1 Regulation of antimicrobial peptides ....................................................................13
    1.4.2 Dysregulation of antimicrobial peptides in disease ...............................................17
  1.5 Transcription factors involved in the regulation of LL-37 ...........................................18
    1.5.1 Vitamin D Receptor ..............................................................................................18
    1.5.2 STAT3 ...................................................................................................................18
    1.5.3 HIF-1 ....................................................................................................................19
    1.5.4 PU.1 .....................................................................................................................19
    1.5.5 Thyroid Receptor .................................................................................................19
    1.5.6 Estrogen Receptor ...............................................................................................20
  1.6 HDACs and HDAC inhibitors ......................................................................................20
    1.6.1 Entinostat .............................................................................................................21
  1.7 Identification of novel drugs ..........................................................................................23
    1.7.1 Assay development ...............................................................................................23
    1.7.2 Chemical libraries .................................................................................................25
  2 Aims of this thesis ............................................................................................................27
  3 Methodology .....................................................................................................................29
    3.1 Molecular cloning .......................................................................................................29
    3.2 Chromatin Immuno Precipitation ..............................................................................32
    3.3 Proteomics coupled to pathway analyses ..................................................................33
  4 Results and Discussion ......................................................................................................35
    4.1 Paper I: Boosting innate immunity: Development and validation of a cell-based screening assay to identify LL-37 inducers .........................................................35
    4.2 Paper II: Highly potent inducers of endogenous antimicrobial peptides for host directed therapy of infections .............................................................................39
    4.3 Paper III: Entinostat up-regulates the CAMP gene encoding LL-37 via activation of STAT3 and HIF-1 transcription factors .........................................................42
    4.4 Paper IV: Label-free quantitative mass spectrometry reveals novel pathways involved in LL-37 expression .....................................................................................45
  5 Conclusions .......................................................................................................................47
  6 Future experiments or things I would have liked to do ....................................................48
  7 Acknowledgements ..........................................................................................................49
  8 References ........................................................................................................................52
## LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADME</td>
<td>Absorption, distribution, metabolism, and excretion</td>
</tr>
<tr>
<td>AMP</td>
<td>Antimicrobial peptide</td>
</tr>
<tr>
<td>APD</td>
<td>Aroylated phenylene diamine</td>
</tr>
<tr>
<td>ASA</td>
<td>Acetylsalicylic acid</td>
</tr>
<tr>
<td>Bcl2</td>
<td>B-cell lymphoma 2</td>
</tr>
<tr>
<td>bp</td>
<td>Base pairs</td>
</tr>
<tr>
<td>CAMP gene</td>
<td>Cathelicidin antimicrobial peptide gene</td>
</tr>
<tr>
<td>ChIP</td>
<td>Chromatin immuno precipitation</td>
</tr>
<tr>
<td>ChIP-seq</td>
<td>ChIP with subsequent deep sequencing</td>
</tr>
<tr>
<td>COX</td>
<td>Cyclooxygenase</td>
</tr>
<tr>
<td>DAMP</td>
<td>Damage associated molecular patterns</td>
</tr>
<tr>
<td>DMSO</td>
<td>Dimethyl sulphoxide</td>
</tr>
<tr>
<td>ER</td>
<td>Estrogen hormone receptor</td>
</tr>
<tr>
<td>Foxp3</td>
<td>Forkhead box P3</td>
</tr>
<tr>
<td>GR</td>
<td>Glucocorticoid receptor</td>
</tr>
<tr>
<td>HAT</td>
<td>Histone acetyl transferase</td>
</tr>
<tr>
<td>hBD</td>
<td>Human β defensin</td>
</tr>
<tr>
<td>hCAP-18</td>
<td>Human cationic antimicrobial protein 18 kDa</td>
</tr>
<tr>
<td>HD5-6</td>
<td>Human alpha defensin 5-6</td>
</tr>
<tr>
<td>HDAC</td>
<td>Histone deacetylase</td>
</tr>
<tr>
<td>HDACi</td>
<td>Histone deacetylase inhibitor</td>
</tr>
<tr>
<td>HIES</td>
<td>Hyper IgE syndrome</td>
</tr>
<tr>
<td>HIF-1</td>
<td>Hypoxia induced factor 1</td>
</tr>
<tr>
<td>HIV</td>
<td>Human immunodeficiency virus</td>
</tr>
<tr>
<td>HMGB1</td>
<td>High mobility group box 1</td>
</tr>
<tr>
<td>HNP1-4</td>
<td>Human neutrophil peptide-1-4, alpha defensin 1-4</td>
</tr>
<tr>
<td>HPD</td>
<td>HIF prolyl-hydroxylase</td>
</tr>
<tr>
<td>HTS</td>
<td>High throughput screening</td>
</tr>
<tr>
<td>IKB</td>
<td>Ingenuity knowledgebase</td>
</tr>
<tr>
<td>IL</td>
<td>Interleukin</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Definition</td>
</tr>
<tr>
<td>--------------</td>
<td>------------</td>
</tr>
<tr>
<td>IPA</td>
<td>Ingenuity pathway analyzer</td>
</tr>
<tr>
<td>JAK</td>
<td>Janus kinase</td>
</tr>
<tr>
<td>kb</td>
<td>kilo base pairs</td>
</tr>
<tr>
<td>KDAC</td>
<td>Lysine deacetylase</td>
</tr>
<tr>
<td>LCA</td>
<td>Lipocholic acid</td>
</tr>
<tr>
<td>LPS</td>
<td>Lipopolysaccharides</td>
</tr>
<tr>
<td>LTB4</td>
<td>Leukotriene B4</td>
</tr>
<tr>
<td>MAPK</td>
<td>Mitogen-activated protein kinases</td>
</tr>
<tr>
<td>MIC</td>
<td>Minimal inhibitory concentrations</td>
</tr>
<tr>
<td>MRSA</td>
<td>Methicillin-resistant <em>Staphylococcus aureus</em></td>
</tr>
<tr>
<td>MS</td>
<td>Mass spectrometry</td>
</tr>
<tr>
<td>Mtb</td>
<td>Mycobacterium tuberculosis</td>
</tr>
<tr>
<td>NET</td>
<td>Neutrophil extracellular trap</td>
</tr>
<tr>
<td>NF-κB</td>
<td>Nuclear factor κ light chain enhancer of activated B-cells</td>
</tr>
<tr>
<td>NK-cells</td>
<td>Natural Killer cells</td>
</tr>
<tr>
<td>NOD-like receptor</td>
<td>Nucleotide-binding oligomerization domain receptor</td>
</tr>
<tr>
<td>Nrf2</td>
<td>Nuclear related factor 2</td>
</tr>
<tr>
<td>P2X7</td>
<td>P2X purino-receptor 7</td>
</tr>
<tr>
<td>PAMP</td>
<td>Pathogen associated molecular patterns</td>
</tr>
<tr>
<td>PBA</td>
<td>Phenylbutyrate</td>
</tr>
<tr>
<td>pDC</td>
<td>Plasmacytoid dendritic cell</td>
</tr>
<tr>
<td>PGE2</td>
<td>Prostaglandin E2</td>
</tr>
<tr>
<td>PLS</td>
<td>Papillon-Lefèvre syndrome</td>
</tr>
<tr>
<td>PMN</td>
<td>Polymorphonuclear Granulocytes</td>
</tr>
<tr>
<td>PRR</td>
<td>Pattern recognition receptors</td>
</tr>
<tr>
<td>pVHL</td>
<td>von Hippel-Lindau tumor suppressor protein</td>
</tr>
<tr>
<td>qPCR</td>
<td>Quantitative PCR</td>
</tr>
<tr>
<td>Reg</td>
<td>Regenerating islet-derived protein</td>
</tr>
<tr>
<td>RNS</td>
<td>Reactive Nitrogen Species</td>
</tr>
<tr>
<td>ROS</td>
<td>Reactive Oxygen Species</td>
</tr>
<tr>
<td>RXR</td>
<td>Retinoid X receptor</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>shRNA</td>
<td>Short hairpin RNA</td>
</tr>
<tr>
<td>STAT</td>
<td>Signal transducer and activator of transcription</td>
</tr>
<tr>
<td>T3</td>
<td>Triiodothyronine</td>
</tr>
<tr>
<td>T4</td>
<td>Thyroxine</td>
</tr>
<tr>
<td>TLR</td>
<td>Toll like receptor</td>
</tr>
<tr>
<td>TR</td>
<td>Thyroid hormone receptor</td>
</tr>
<tr>
<td>VDR</td>
<td>Vitamin D receptor</td>
</tr>
</tbody>
</table>
1 BACKGROUND

1.1 INTRODUCTION

With the arrival of antibiotics during World War II, came a weapon against microbes and today, we are not used to people dying from ‘simple’ bacterial infections. However, the microbes are winning ground and more bacterial pathogens are gaining resistance even to our most advanced antibiotics. Therefore, to prevent the return of the pre-antibiotic era, it is of the utmost importance to develop novel antibiotics and treatment regimes. This thesis describes the field of ‘Host Directed Therapy’, where we aim to treat infections by aiding the immune system in order to eliminate pathogens. By treating the host with compounds that induce expression of a wide variety of antimicrobial peptides (AMPs), enhancing ROS and RNS production and by activating autophagy, the host can efficiently destroy invading pathogens, inhibit their growth and prevent attachment. Moreover, induction of multiple AMPs with different mechanisms of action are likely more effective and also less prone to cause the development of microbial resistance, which is a relevant point given the emerging spread of multidrug resistant bacteria (Marr et al., 2006; Singh et al., 2000).

1.2 THE IMMUNE SYSTEM

The human immune system is typically described as two separate parts, the innate immune system, which we share with all multicellular organisms and the adaptive immune system, which is exclusive to jawed vertebrates. The adaptive immune system developed from innate immunity and the two systems are intertwined with each other meaning that the separation between the two parts is not exact, and separately, neither would be enough to defend us nor function in an optimal manner. The immune system is a complex entity and has always been at the forefront of evolution to keep up with all microbes in the surrounding. The innate system is the first line of defense and stops most invaders at the barrier of our body. The microbes that breach the barrier will be recognized by receptors on the cells of the adaptive system and a clonal expansion will start. The adaptive immune system is slower in action than innate immunity but on the other hand is much more specific and efficient when it is fully matured. The expansion of T- and B- cells may take up to a week for full functionality and during that time the innate immune system needs to hold all invaders at bay.

1.3 INNATE IMMUNITY

Innate immunity is the inherited defense against microbial threats towards the host and it is immediately activated when needed. It is evolutionary conserved and is the only immune-defense in plants, fungi and invertebrates. In vertebrates, including humans, innate immunity functions as the first line of defense but also activates and recruit cells of the adaptive immune system. All organisms need ways to discriminate between self and non-self, hence there is a need for sensing of foreign molecules, such as microbial patterns and the ability to distinguish them from self. The sensing is followed by a response with direct antimicrobial
action, tissue-repair or the recruitment of clotting cells. The innate system is therefore divided in a sensing part and an effector part.

1.3.1 Innate sensing

An array of recognition receptors is included in innate immunity with two main classes; the Toll-like receptors (TLRs), which are situated in the cell-membrane and in the membrane of the lysosomes and the nucleotide-binding oligomerization domain receptors (NOD-like receptors), which are located inside the cells. There are also additional innate receptors, such as the RAGE receptor and the RIG-1-like receptors (Kierdorf and Fritz, 2013; Loo and Gale, 2011; Xie et al., 2013). All receptors recognize certain structures or molecules, which signal infection or injury (Matzinger, 2002; Medzhitov, 2007, 2009). These receptors are jointly called pattern recognition receptors (PRRs) and they have specificity towards different molecular patterns. These patterns can be divided in pathogen associated molecular patterns (PAMPs) and damage associated molecular patterns (DAMPs) (Matzinger, 1994; Matzinger, 2002). PAMPs are generally unique and constitute vital parts of microbes, such as lipopolysaccharides (LPS), lipoteichoic acid, peptidoglycans, in addition to double stranded RNA, single stranded DNA and unmethylated CpGs originated from microbes (Mogensen, 2009).

DAMPs - also known as alarmins - are released from cells under stress and consist of several molecules or patterns. The classical DAMP is the protein High mobility group box 1 (HMGB1), which binds to the RAGE-receptor. Additional alarmins are self-DNA, heat-shock proteins, uric acid and S100 proteins (Foell et al., 2007; Lotze et al., 2007). In addition, AMPs can also function as alarmins. For example, LL-37 can bind to self DNA and form a complex, which is recognized by Toll like receptor 9 (TLR9) in plasmacytoid dendritic cells (pDCs), while the human neutrophil peptide-1 (HNP1) also known as α-defensin 1 regulates IL-1β post translationally via the P2X purino-receptor 7 (P2X7) (Chen et al., 2014; Lande et al., 2007).

In response to the different stimuli described above, pro-inflammatory pathways are triggered. The best described pathway involves activation of ‘nuclear factor κ light chain enhancer of activated B cells’ (NF-κB). In addition, there are several other important pathways in innate immunity, such as the mitogen-activated protein kinases (MAPK) P38 and ERK1/2 pathways. The signal cascade of the pro-inflammatory pathways leads to an increased expression and release of cytokines and chemokines but also triggers expression of additional defense molecules, such as AMPs and mucins (Enss et al., 2000; Perez-Vilar and Hill, 1999) (O'Neil et al., 1999). However, the specific response will be determined by two major factors; the ligand and in which cell-type the response occurs (Bianchi, 2007).
1.3.2 **Innate effectors**

1.3.2.1 **Reactive Oxygen Species (ROS) and Reactive Nitrogen Species (RNS)**

Both ROS and RNS are produced as a part of the innate defenses. The molecules exhibit a high reactivity by virtue of single electrons, which react with proteins, lipids and carbohydrates. NADPH oxidase is an enzyme, which is stored in the granules of neutrophils and can generate both ROS and RNS. In addition to direct microbial killing, both ROS and RNS are prominent signal molecules and activate several pathways of immunity, by activating redox sensitive proteins, such as tyrosine phosphatases which regulate the MAPK pathway (Neish, 2013). NF-κB is a redox sensitive transcription factor which is closely associated to many immune genes and functions (Gostner et al., 2013). The determination of Th1 or Th2 regulatory T cells is linked to the redox potential in the milieu around the naive T-cell, where an oxidative milieu promotes Th1 cells and versa reducing environment appears to promote Th2-cells (Gostner et al., 2013). Another major transcription factors regulated by ROS is the hypoxia induced factor 1 (HIF-1). To protect the host from damage, several antioxidant systems are in use, such as superoxide dismutases and the thioredoxins (Nathan and Cunningham-Bussel, 2013).

1.3.2.2 **Autophagy**

The mechanism of autophagy is how the cells dispose of old proteins and broken organelles in order to prevent them from intervening with the normal processes in the cell. Autophagy is a key component to keep the homeostasis of protein production and degradation. Interestingly, the process of autophagy recycles all material and it is therefore vital during starvation when less important proteins and organelles can be degraded and used for building parts and nutrients. However, it is also a very important pathway for clearance of intracellular pathogens, such as *Mycobacterium tuberculosis* (*Mtb*). In addition, autophagy have been demonstrated to modulate both MHC class I and class II antigen presentation (Crotzer and Blum, 2009; Glick et al., 2010).

1.3.2.3 **Antimicrobial peptides**

AMPs are often referred to as gene encoded antibiotics and are typically produced and released upon infection. They are cationic and amphipathic and as such have a preference for microbial cell-membranes and the main mode of action is disruption of cell-membranes (Zasloff, 2002). AMPs are ancient molecules in evolution and important effector molecules of the innate immune system. They are found throughout all branches of life, including archaea. The halocin S8 from *Haloarchaea* (Price and Shand, 2000) is an example of an AMP from a very primitive organism. Moreover, also bacteria express molecules with antimicrobial action against other bacterial species. These molecules are called bacteriocins and are important for the bacterial competition at mucosal surfaces (Dobson et al., 2012).

The first antimicrobial molecule found in humans was isolated by Alexander Fleming from nasal secretions in 1922 and this antimicrobial protein was named ‘lysozyme’ (Fleming,
1922). It is highly expressed in neutrophils (PMN’s) and also present in tears and breast milk. This finding set off an era of novel discoveries in the field. Between 1920 and 1950 several articles described the isolation of molecules with antimicrobial activity from many different organisms, Bombinin from the frog Bombina variegata, lactoferrin from cow milk and purothionins from wheat, (Fernandez de Caleya et al., 1972; Groves et al., 1965; Kiss and Michl, 1962). Technical limitations of the time hindered full characterization of the respective molecules but in 1980, the first AMPs were characterized and reported by the group of Hans G. Boman. These small peptides were isolated from the silk moth, Hyalophora cecropia and were subsequently named as ‘cecropins’ (Hultmark et al., 1980; Steiner et al., 1981). In the beginning of the eighties the defensins were isolated from rabbit macrophages by the group of Robert Lehrer (Selsted et al., 1983). At the end of the 80’s Michael Zasloff isolated and characterized the Magainins from the African clawed frog Xenopus laevis (Zasloff, 1987). These discoveries indicated that AMPs were widespread in nature. To date 2684 AMPs and proteins from different species and kingdoms have been described and in humans there are 112 reported antimicrobial peptides and proteins (Wang, 2014; Wang et al., 2016b).

AMPs share some common denominators across species. Typically, they are small (12–50 amino acids), have a positive net-charge with an amphipathic character and are usually defined by their secondary structure; α-helical, β-sheets or random coils, rich in arginine and prolines (Lai and Gallo, 2009).

1.3.2.3.1 Defensins

The defensins represent an ancient family of AMPs and in contrast to the cathelicidins they are more than 30 different defensin genes identified in the human genome. However, only ten of them have been characterized on the peptide level. Defensins have been found in animals, plants, fungi and in some species of myxobacteria (Lehrer and Ganz, 2002; Mygind et al., 2005; Vriens et al., 2014; Wong et al., 2007; Zhu, 2007, 2008). The defensins are synthesized as preproteins and are activated by cleavage into peptides with sizes between 18–45 amino acids. These peptides contain conserved cysteines, six in mammals and birds, eight in plants and vertebrates, forming disulphide bonds, which stabilize a β-sheet conformation (Carvalho Ade and Gomes, 2009; Ganz, 2003). The S-S bonds are however paired differently between the cysteine residues, giving rise to the division of the defensins into α, β and θ defensins (Ganz, 2003). The defensins, as a group, have been shown to have a broad and potent antimicrobial activity against both gram positive and gram negative bacteria, fungi and some enveloped viruses (Ganz, 2003). The concentration of enteric defensins can reach levels of above 10 mg/ml in the intestinal crypts and between 10 and 100 µg/ml in other epithelial cells, as determined by analysis of human skin and porcine tongue (Ganz, 2003). In a publication by Salzman et al. in 2003, defensins were for the first time shown to be important in vivo, where human α defensin-5 (HD5) enabled protection against enteric salmonellosis (Salzman et al., 2003). Like most AMPs, defensins act through the disruption of microbial cell membranes, but it has also been demonstrated that they interfere with both RNA and
DNA synthesis (Ganz, 2003). Interestingly, it has been reported that human β defensin-1 (hBD-1) increases its antimicrobial activity upon reduction of its S-S bridges, most likely via thioredoxin, which co-localizes with hBD-1 in epithelial cells (Schroeder et al., 2011). For HD5, on the other hand, reduction of its S-S bonds reduces its ability to kill microbes in vitro and instead induces the LPS neutralizing effect and turns the peptide into an effective zinc chelator (Wang et al., 2016a; Zhang et al., 2013). Defensins are more than just active antimicrobials; some also function as alarmins, regulators of cytokines and chemo-attractants of monocytes and T-cells (Chen et al., 2014; Chertov et al., 1996; Sakamoto et al., 2005; Semple and Dorin, 2012; Territo et al., 1989).

α-defensins

There are six α defensins expressed in humans. Four of them are known under the name of human neutrophil peptide (HNP1-4). The HNPs are highly concentrated in the primary granules of neutrophils and expressed in B-, T- and NK-cells as well as in monocytes (Agerberth et al., 2000; Ganz, 2003). The human defensins 5 and 6 (HD5 and HD6) are expressed by Paneth cells in the small intestine in a constitutive manner. They are stored in granules that upon stimulation are released into the intestinal crypts, where they can reach concentrations of more than 10 mg/ml (Bevins and Salzman, 2011; Ganz, 2003; Ouellette, 2011). HD5 and HD6 complement each other by virtue of their separate mechanisms in the innate immune system. HD5 has a direct antimicrobial action by lysing bacteria, while HD6, interestingly, can self-assemble or aggregate into what can be defined as “nanonets”, with the ability to ensnare bacteria (Bevins, 2013; Chu et al., 2012; Ouellette and Selsted, 2012).

HD5 is constitutively expressed in the female genital tract and induced in urethra by Neisseria gonorrhoeae infection. HD6 has also been implicated in the female genital tract upon the same infection (Klotman et al., 2008; Porter et al., 2005; Svinarich et al., 1997). Interestingly, HD5 and 6 as well as Neisseria gonorrhoeae infection have been suggested to facilitate HIV transmission (Klotman et al., 2008; Rapista et al., 2011).

β-defensins

The human β-defensins are mainly expressed in the skin and at mucosal barriers (Ganz, 2003; Lai and Gallo, 2009). Despite many genes encoding β-defensins in the human genome, only four peptide products have been characterized. hBD-1 is mainly constitutively expressed, while the other three β-defensins, hBD2-4, are regulated by PAMPs or DAMPs, e.g. by bacterial infection or injury (Nuding et al., 2013). However, we and others have demonstrated that also hBD-1 can be regulated, both by bacterial infection and by small molecular compounds (paper III) (Sorensen et al., 2005). In addition, β-defensins are expressed in the genital tract and interestingly hBD2-4 are down-regulated as a response to unprotected vaginal intercourse (Nakra et al., 2016).
θ-defensins

θ-defensins are interesting molecules that only exist in old world primates. θ-defensins are cyclic peptides and as such the only cyclic peptides found in mammals, where they seem to have evolved from α-defensins. The active θ-defensin is formed by ligation of two copies of the θ-defensin peptide; i.e. two copies of the precursor protein are needed for one θ-defensin. The precursor protein is interestingly closely related to the α-defensins (Tang et al., 1999).

Humans, chimpanzees and gorillas do not express θ-defensins, however the genes for θ-defensins can be found in the genomes of these species. In humans the mRNA is transcribed, but a mutation has introduced a premature stop codon and thus the active peptide cannot be translated. The θ-defensins have been studied for their potential as anti-viral compounds, especially in the treatment of HIV (Gallo et al., 2006; Munk et al., 2003).

1.3.2.3.2 Cathelicidins

Cathelicidins are generally considered mammalian AMPs and at least one cathelicidin gene have been identified within all mammals examined. However, they have also been found in snakes and catfish. They are synthesized as pre-pro-peptides and the signal sequence (pre) is directing the pre-pro-peptide towards the designated compartment, usually destined for exocytosis (Gudmundsson et al., 1996). The conserved cathelin domain (pro) has a high similarity to the cathepsin L inhibitor cathelin and the variable C-terminal domain is the part that confer the antimicrobial activity (Zanetti et al., 1995). This is conferred by a four exon gene which in human is entitled Cathelicidin AntiMicrobial Peptide or CAMP (Gudmundsson et al., 1996). It is located on the third chromosome and is only ~2 kilo base pairs (kb) long including introns (Figure 1). Interestingly, the variable active peptide is always located in the fourth exon, potentially explaining why the C-terminal and active peptides of the cathelicidins are not evolutionary conserved (Tomasinsig and Zanetti, 2005). Humans only carry one cathelicidin gene and this holds true also for mouse, rat and rabbit. The peptides in these species are all α-helical in their secondary structure, implying that these species can work as model systems for studies on the human cathelicidin. The α-helical secondary structure of the human C-terminal domain adopts this conformation upon contact with microbes, experimentally demonstrated using micelles (Li et al., 2006). Possibly avoiding host damage by undertaking a random coil structure until activity is needed. Furthermore it has been shown that the helical structure is also stabilized under mimicking of in vivo conditions (Johansson et al., 1998).

Cathelicidin is vital for resistance to several pathogens. For example, in mice it has been demonstrated that the mouse cathelicidin CRAMP is pivotal for resistance against C. albicans in the gut and against a sensitive strain of Group A Streptococcus in the skin (Fan et al., 2015; Nizet et al., 2001).

The human cathelicidin peptide was discovered in 1995 on cDNA level and the putative antimicrobial peptide was named FALL-39 (Agerberth et al., 1995). However, when the active peptide was isolated and characterized from granulocytes it was found to be two residues shorter. Thus, the peptide was named LL-37, after the two leucines in the N-terminal
end and from the length of 37 amino acid residues. The LL-37 peptide sequence is as follows. NH₂-LLGDFFRKSKKEKIGKEFKRIVQRIKDFLRNLVPRTES-COOH (Gudmundsson et al., 1996). The proform was designated as human cationic antimicrobial protein weighing 18 kDa or hCAP-18.

**Figure 1.** Schematic representation of the *CAMP* gene and its translational product. Transcription factors activate RNA polymerase II and mediate transcription. The *CAMP* gene is transcribed from four exons, where Exon 1-3 encode the cathelin domain including a sequence that signals for exocytosis. Exon four encodes the peptide with the antimicrobial activity, LL-37, shown by dual colors, representing the amphipathicity of the peptide. The signal sequence is removed after direction for exocytosis. The remaining pro-peptide is designated hCAP-18 and the active peptide is cleaved off by protease 3 in neutrophils and adopts an α-helical conformation.

LL-37 has been shown to be potent against many different microbial species (Turner et al., 1998), and is considered especially potent against gram negative bacteria because of the high affinity to LPS. LL-37 actually binds LPS with a 2:1 stoichiometry that is cooperative to its nature (Turner et al., 1998). Minimal inhibitory concentrations (MIC) values of LL-37 varies with bacteria and method used, it is however not unusual with sub-µM concentrations (Larrick et al., 1995). Furthermore, the activity based on the ability to achieve α-helical structure of LL-37 is enhanced in the presence of anions and by increasing the pH value (Johansson et al., 1998). It has been demonstrated that apart from the microbial killing, LL-37 is responsible for additional effects, such as angiogenesis, apoptosis and as already mentioned, endotoxin neutralizing (Table 1).
### Table 1. Additional functions of LL-37 apart from the antimicrobial effects.

<table>
<thead>
<tr>
<th>Function</th>
<th>Mechanism</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chemotaxis</td>
<td>By activation of formyl peptide receptor-like 1 (FPRL1) and CXCR2 (IL-8 receptor)</td>
<td>(De et al., 2000; Zhang et al., 2009)</td>
</tr>
<tr>
<td>Apoptosis</td>
<td>Both inducing apoptosis and increase life span of cells. Action suggested via FPRL1, P2X7, P53, Bax ,Bak and Bcl-2</td>
<td>(Barlow et al., 2006; Lau et al., 2006; Nagaoka et al., 2006; Ren et al., 2012)</td>
</tr>
<tr>
<td>Angiogenic</td>
<td>Via FPRL1</td>
<td>(Koczulla et al., 2003)</td>
</tr>
<tr>
<td>Neutralization of endotoxin</td>
<td>Direct LPS binding and competitive binding to P2X7</td>
<td>(Hu et al., 2014; Turner et al., 1998)</td>
</tr>
<tr>
<td>Wound healing</td>
<td>LL-37 induce keratinocyte migration via EGFR activation</td>
<td>(Heilborn et al., 2003; Shaykhiev et al., 2005; Tokumaru et al., 2005)</td>
</tr>
<tr>
<td>Histamine release</td>
<td>Via G protein-phospholipase C-dependent manner</td>
<td>(Niyonsaba et al., 2001)</td>
</tr>
<tr>
<td>Leukotriene B4 release</td>
<td>Via activation of FPR2/ALX</td>
<td>(Wan et al., 2011; Wan et al., 2007; Wan et al., 2014a)</td>
</tr>
<tr>
<td>Increase in ROS production and release of α defensins</td>
<td>By enhanced production of IL-8</td>
<td>(Zheng et al., 2007)</td>
</tr>
<tr>
<td>Modulation of dendritic cell function</td>
<td>Via regulation of TLR ligands</td>
<td>(Kandler et al., 2006)</td>
</tr>
<tr>
<td>Macrophage differentiation</td>
<td></td>
<td>(van der Does et al., 2010)</td>
</tr>
<tr>
<td>Induction of autophagy</td>
<td></td>
<td>(Rekha et al., 2015)</td>
</tr>
<tr>
<td>Increase phagocytosis</td>
<td>Via activation of integrin Mac-1</td>
<td>(Wan et al., 2014b; Zhang et al., 2016)</td>
</tr>
</tbody>
</table>

1.3.2.3.3 **Antimicrobial proteins**

The S100 calcium proteins are a large group of structurally related proteins of which several have demonstrated to exhibit antimicrobial activity. Calprotectin consists as a dimer of
S100A8 and A9 and is highly abundant in the cytoplasm of neutrophils (Edgeworth et al., 1991). The mechanism of action is attributed to its chelating properties, where it deprives the microbes of essential minerals (Clohessy and Golden, 1995; Damo et al., 2013; Sohnle et al., 2000). Another S100 protein with antimicrobial activity is psoriasin, which was first isolated from psoriatic lesions, where it is drastically induced. Interestingly, psoriatic lesions are seldom infected despite the prolonged exposure to possible invasive pathogens, likely due to chelating of divalent zinc ions (Glaser et al., 2005). Psoriasin is also expressed during differentiation of mammary epithelial cells, and it is associated with poor prognosis in estrogen negative breast cancer tumors (Emberley et al., 2003; Vegfors et al., 2012).

Lactotransferrin is another chelating protein that was first identified to act as a transporter for iron in milk and blood. However, it is now known to be active against viruses, bacteria and fungi, and the main mechanism is sequestering of irons from the milieu and hence to starve the microbes (Bullen et al., 1972; Jenssen and Hancock, 2009).

Contrary to AMPs, many antimicrobial proteins are negatively charged, hence the chelating properties, which seem to be a general route of antimicrobial activity. In fact all metal sequestering proteins can be considered to be antimicrobial by nature, disregarding their actual function.

Regenerating islet-derived proteins (Regs) from the superfamily of C-type lectins have been detected in many different cell-types and are often associated with differentiation and proliferation (Parikh et al., 2012). Generally, theRegs are thought of as intestinal proteins and have been demonstrated to keep the small intestinal epithelial cells sterile (Vaishnava et al., 2011). In mice it has been demonstrated that signal transducer and activator of transcription 3 (STAT3) is a key regulator of Reg3γ in both intestinal and lung epithelia, and that Reg3γ is vital in MRSA pneumonia clearance (Choi et al., 2013; Lee et al., 2012; Murano et al., 2014). Also in humans it has been demonstrated that several Reg genes are upregulated upon STAT3 activation (Murano et al., 2014). The killing mechanism of Reg3γ action have not yet been determined (Gallo and Hooper, 2012). Reg3α is bactericidal for Gram-positive but not Gram-negative bacteria, which is explained by the fact that LPS blocks the ability of Reg3α to form pores in bacterial membranes (Mukherjee et al., 2014). Reg4 is less studied but is strongly associated with colorectal cancer (Kawasaki et al., 2015).

Histones are proteins that are known to aid in the packaging of DNA in the cell nucleus. In addition, they also have antibacterial properties, where peptides derived from the larger histone proteins have been demonstrated to inhibit bacteria \textit{in vitro}. They are also a major part of the neutrophil extracellular traps (NETs) and as such also inhibit infection (Kawasaki and Iwamuro, 2008).

Ribonucleases are responsible for degradation of RNA in the cell and six out of eight have been reported to be involved in host defenses, RNases 2, 3, 5, 7, and 8 (Wang, 2014).
1.3.2.4 Linking innate and adaptive immunity

In addition to their capacity to directly kill microbes, the innate effectors (AMPs, ROS and RNS) can all be described as linkers to the adaptive immune system, which via the activation of cytokines, chemokines and specific AMPs continue to strengthen defenses by recruiting cells of both the innate and adaptive immune system.

1.3.3 Innate immune cells

1.3.3.1 Epithelial cells

The first innate defenses that any microbe will face are the physical barriers of the host, i.e. the skin, the eyes, the mouth, lungs and the lining of the gut. A part of this defense is epithelial cells, which are able to mount an effective defense against any microbial attack. Epithelial cells are covered by a layer of mucins, building the mucous membrane, as in the gut lining (Perez-Vilar and Hill, 1999). The mucin layer harbors several effector molecules, such as AMPs that are able to eliminate potential pathogens. The mucin layer consists of the highly glycosylated mucin proteins, which forms a virtually impassible gel-like structure. The mucus layer is replenished by special epithelial cells designated goblet cells. The result of this replenishment from below is a near sterile environment closest to the epithelial cells. The fact is that very few of the microbes we encounter every day will ever be in proximity to the epithelial cells. Even in the small intestine, where the mucus layer is thin, it has been demonstrated that bacteria is kept at distance from the epithelial cells by secretion of the antimicrobial protein Reg3γ (Vaishnava et al., 2011). If the epithelial layer or the mucosa layer is damaged, a pathogen may reach a site, where it can infect or invade. Epithelial cells are decorated with PRRs, which recognize PAMPs or DAMPs and expression and release of cytokines and chemokines is initiated. This leads to recruitment and activation of other immune cells, such as antigen presenting cells. Interestingly, TLR4 in intestinal epithelial cells are located on the basal side of the cells to avoid unnecessary activation by LPS (Hornef et al., 2003). Most of our barriers are covered by commensal bacteria specific to the place of colonization, helping the host by competing with pathogens for both nutrients and habitation. Moreover, it has been demonstrated that many of the commensal bacteria are inducing the expression of AMPs to which the commensal bacteria are resistant to and in that way keep pathogenic competitors at bay (Bevins, 2005; Kamada et al., 2013). The number of species in the natural flora of the gut increases with the distance from the stomach and in the colon the number of commensal bacteria reaches 10 to the power of 13 (Sartor, 2008). In contrast, the human skin barrier is more arid of commensals. Instead of the mucus layer seen in eyes, intestine and lung, the skin has the apoptotic layer of keratinocytes, which stops invading pathogens with its hard exterior. In skin, the AMPs are mainly produced when the barrier is broken, damaged or infected. For example, cathelicidin and hBD1 is constitutively expressed at low levels but may be induced together with hBD2-4 upon infection or damage (Kenshi and Richard, 2008).
1.3.3.2 Polymorphonuclear Granulocytes (PMNs)

PMNs makes up 50-70% of all leukocytes in humans and are named for their small vesicles or granules, which can be detected upon Hematoxylin staining of blood smear (Borregaard et al., 1987). Of the granulocytes around 99% are neutrophils and 1% consists of eosinophils, basophils and mast cells. The neutrophils are the first immune cells recruited to a site of infection or injury, where they immediately destroy microbes by phagocytosis or by releasing their granular content. The granules contain high concentrations of AMPs ($\alpha$-defensins and hCAP-18) and additional proteins with antimicrobial activity, such as bactericidal/permeability-increasing protein and Lysozyme. The AMP concentration in the phagocytic vacuoles reaches concentration above 10 mg/ml (Ganz, 2003). The granules also contain proteases such as protease 3, which has been shown to cleave the human cathelicidin pro-form hCAP-18 into the cathelin domain and the active peptide LL-37 (Sorensen et al., 2001). The neutrophils, besides their granules, carry disassembled NADPH oxidase 2 separated in the cytosol and in the membrane which upon stimuli readily can be assembled and generate an oxidative burst as a general defense. Neutrophils can undergo what is called NETosis, where they in a suicidal fashion spew out NETs, a sticky entity made up of chromatin and decorated with AMPs, trapping microbes like flies in a spider web (Brinkmann and Zychlinsky, 2007; Sorensen et al., 2001).

1.3.3.3 Macrophages

Monocytes are circulating in the blood stream and lymphatic system, and depending on stimuli they differentiate into either macrophages or myeloid dendritic cells.

Macrophages are found in many tissues as specialized cells with phagocytic properties. They have different names depending on the tissue; e.g. Kupffer cells in the liver, Langerhans cells in the skin and alveolar macrophages in the lungs. The macrophages are professional phagocytes of bacteria and present antigens from engulfed and degraded bacteria. Thus, macrophages can sense danger signals and convey immunological signals as well as engulf and kill bacteria. The circulating macrophages traverse the body to sites of infection upon stress signals. There are also stationary macrophages present in the tissues, which are often the first cells that a pathogen encounters after invasion. Upon recognition they start to express and release cytokines, leading to recruitment of PMNs and cells from the adaptive immune system (Medzhitov, 2007). Macrophages can be divided in two subtypes, the M1 and M2 types, where the M1 is more pro-inflammatory and M2 ameliorate inflammation and drive tissue repair (tolerogenic phenotype) (Mills, 2012). Macrophages can also be used as intracellular reservoirs for certain bacteria, such as Mtb, which after phagocytosis is able to block the intracellular killing machinery. Notably, it has been shown by our group that Mtb downregulates the expression of the human cathelicidin LL-37 in macrophages (Rekha et al., 2015).
1.3.3.4 Th17 cells

Th17 cells are a subset of T-helper cells with the ability to produce IL-17, hence the name of the cell. In addition to IL-17, Th17 cells also produce IL-22 and together these cytokines orchestrate the mucosal immune-response to many pathogens. Th17 cells have been implicated in both induced clearance of pathogens and autoimmune diseases. Interestingly, various microbial species seem to induce the differentiation of naive T-cells into the Th17 cell type where, *Propionibacterium acnes* and *Klebsiella pneumoniae* in addition to the fungi *Candida albicans* are three examples (Korn et al., 2009). In psoriatic plaques it has been demonstrated that Th17 cells are increased and considered to correlate with the severity of the disease (Pène et al., 2008). For Th17 differentiation to take place, STAT3 is central. Interestingly, patients with Hyper IgE syndrome (HIES) have a loss of function mutation in the gene encoding STAT3 and consequently lack IL-17 producing T helper cells (Bevins, 2005). Since Th17-cytokines have been shown to induce AMP-expression in epithelial cells, a lack of these cytokines may explain why infections with specific bacteria, including *Staphylococcus aureus* are so frequent in HIES patients (Milner et al., 2008).

1.3.3.5 Natural Killer cells, Dendritic cells, Eosinophils, Basophils and Mast cells

NK-cells are responsible for elimination of autologous (endogenous) cells that have been infected with a virus or turned into dangerous tumor cells (Vivier et al., 2008). Dendritic cells are antigen presenting cells that constantly patrol the tissues and sample antigens, which are presented. DCs express most PRR’s and are excellent antigen presenting cells, which rapidly may reach a lymph node, where they present the antigens to B- and T-cells (Medzhitov, 2007). Eosinophils, basophils and mast cells are part of PMNs but carry other defensive molecules in their granules. They are all closely associated to inflammation and allergy and they are regulators of both innate and adaptive immunity. They are also important in the defense against parasites such as helminths and ticks. Interestingly, eosinophils as well as mast cells are able to produce NETs (Rosenberg et al., 2013; Stone et al., 2010; Voehringer, 2013; Wernersson and Pejler, 2014).

1.4 GENE REGULATION

In each cell, DNA is stored that have the blueprint to make up an entirely new, albeit “cloned” individual. In humans this blueprint contains some 20 000 genes and not all of them can be expressed at the same time. The active genes in a cell are quite different if you compare for example a macrophage to an adipocyte. Subsequently the question arrives, what determines, which genes should be transcribed in time and in space?

DNA is packed in the nucleus wrapped around histones. This histone-DNA complex is designated nucleosomes, which can be further condensed until it reaches its most compact form, which can be seen during mitosis, the metaphase chromosome. During this phase no transcription can take place.
The active genes are sitting in the relaxed chromatin structure designated as euchromatin, while the more rigid structure with inactive genes is named heterochromatin or the 30 nm fiber.

How tightly wrapped the DNA is around the histones depends on epigenetic factors, usually acetyl- and methyl-groups on the different histones. For DNA to be transcribed it needs to be accessible to transcription factors (TFs), which are proteins that bind DNA and initiate or repress transcription of genes. Therefore, almost all active genes in a cell are located in stretches of euchromatin and the promoters of the genes are located between two nucleosomes (Consortium, 2012; Thurman et al., 2012).

DNA is transcribed into RNA by RNA polymerases I-III. RNA polymerase I and III produce the different RNA subunits of the ribosomes, while RNA polymerase II is the one responsible for transcription of all genes destined for translation.

RNA polymerase II is not able to initiate transcription by itself instead it needs additional proteins to coordinate and activate transcription, including TFs. TFs contain a DNA binding domain, recognizing specific nucleotide-sequences on the DNA string, so called TF binding motifs. In the simplest scenario, a TF binds the designated site and recruits RNA polymerase II and additional co-activators, which will initiate transcription. Less simple scenarios include TFs functioning as repressors or recruitment of large complexes that can both unwind and condense the chromatin.

Most TFs are activated through posttranslational modifications or simply by their induced expression. Some TFs, however, need an endogenous or exogenous ligand to induce binding/transcription. These are called nuclear receptors and several of them seem to be essential for CAMP gene expression (paper IV) (Gombart et al., 2005; Kulkarni et al., 2016; Park et al., 2014).

What ultimately decides which genes should be transcribed to mRNA are a complex system, consisting of chromatin condensation, transcription factors and various stimuli.

1.4.1 Regulation of antimicrobial peptides

The identified regulatory pathways of AMP expression is depending on whether the peptide is needed constantly or only during challenge. Thus, some AMPs are constitutively expressed, some are inducible and some are both. Several mechanisms that govern AMP-expression will be discussed below.

The regulatory pathways of AMPs depend on several factors, for instance, which PRR is activated, which microbe is infecting and the responsive tissue. For example, hBD1 is kept at a constitutive expression by HIF-1 in the intestine, where oxygen pressure often reaches 25 % of normoxia (Kelly et al., 2013). Another tissue specific pathway is the activation of the taste receptor for bitterness, T2R38, which results in induced expression of hBD2 in the upper airway epithelium and in the oral cavity (Gil et al., 2015).
Several post-translational modifications are responsible for AMP regulation. Processing is for instance a vital way of regulating AMPs. LL-37 needs to be cleaved from the pro-peptide hCAP-18 to reach its active state. Recently Koro et al. demonstrated that LL-37 can be carbamylated at 7 different residues and that the carbamylation has profound effects on the peptide and its signaling pathway (Koro et al., 2016).

Citrullination is another post translational modification that can alter the effects of LL-37. It has been demonstrated to reduce the binding capacity to LPS, hence reduce the beneficial effects of LL-37 during sepsis (Koziel et al., 2014). Further, the antimicrobial effect against Staphylococcus aureus, Streptococcus pneumoniae, and Haemophilus influenzae is reduced by citrullination. Citrullinated LL-37 has been found to be more efficient in recruiting monocytes and is also more likely to be degraded by proteases (Kilsgard et al., 2012).

There is a difference in storage and expression of AMPs between neutrophils and epithelial cells. In neutrophils there are minor transcription and translation going on after maturation, instead all active polypeptides are stored in granules and are released upon specific stimuli. In epithelial cells there are no granules that can keep the peptides from harming the cell, so instead they are expressed and continuously released. AMP activation usually occurs upon infection, inflammation and wounding (Hancock and Diamond, 2000; Sorensen et al., 2003). This is facilitated by recognition of PAMP and/or DAMP molecules (Matzinger, 2002). This recognition will trigger a pro-inflammatory response, recruiting immune cells and initiate the expression of several immune genes. It has also been demonstrated that AMPs can be induced by wound-healing factors, such as insulin-like growth factor I and transforming growth factor α, probably in an effort to keep the damaged barrier sterile until the repair is finalized (Sorensen et al., 2003).

LL-37 is also regulated by parathyroid hormone (Muehleisen et al., 2012) and by several steroid hormones, such as vitamin D (Wang et al., 2004b), thyroid hormone (paper IV), estrogen (Park et al., 2014) and glucocorticoids (paper I) (Jensen et al., 2011; Kulkarni et al., 2016). The short chain fatty acid butyrate, a bacterial fermentation product, is readily produced in the gut and able to induce cathelicidin expression both in vitro and in vivo in a rabbit model (Raqib et al., 2006; Schauber et al., 2003).

1.4.1.1 Host Directed Therapy

A highly effective immune-system needs to be tightly regulated, in order to prevent attack on “self” or unnecessary wasting of energy after the intruder is eliminated. Therefore we have several regulatory pathways that keep the AMPs at bay or release them upon invading pathogens. Usually this regulation is exactly what the body requires, however, the peptides can be dysregulated in various diseases. We work on a concept that we designate “host directed therapy” based on the application of small molecular compounds that can induce AMPs in various tissues.

Conceptually it has been shown that supplementation of vitamin D to primary immune deficient patients significantly reduced the use of antibiotics in the treated group (Bergman et
Furthermore, adjunctive therapy with phenylbutyrate (PBA) and vitamin D to tuberculosis patients significantly reduced the time to sputum conversion and the clinical score (Mily et al., 2015). Recently, it was shown that the levels of AMPs was increased in saliva and plasma after 14 weeks of vitamin D supplementation in athletes (He et al., 2016).

Host directed therapy is meant to be an alternative or complement to classical antibiotic treatment and hence the question arise, what happens if the pathogens gain resistance to the AMPs? This is a valid question which should be contemplated. The fact is that many pathogens today are resistant to specific AMPs, at least during in vitro conditions, when a single peptide is tested alone (Cole and Nizet, 2016). However, the system is evolutionary ancient and multi-resistance towards the combined action of many AMPs has not been found. Our concept of inducing multiple AMPs with different mechanisms of action, rather than administering a single peptide would be more efficient from a resistance point of view. This is supported in several reports showing that the peptides are co-regulated (paper III) (Aberg et al., 2008; Hau et al., 2013; Ong et al., 2002; Sarker et al., 2011).

Small molecular compounds that are known to regulate AMP expression are not only vitamin D and PBA but also, for example, mimosine extracted from the Leucaena leucocephala seed, which stabilizes HIF-1α and thereby increase LL-37 expression (paper III) (Fan et al., 2015). Endoplasmic reticulum stress is also a potent inducer of LL-37 expression and molecules like tunicamycin induce endoplasmic reticulum stress and LL-37 (Park et al., 2011). Small molecule AMP regulators are listed in Table 2.
Table 2. List of small molecular compounds regulating AMPs.

<table>
<thead>
<tr>
<th>Compound</th>
<th>↑AMP</th>
<th>Mechanism</th>
<th>Reference</th>
<th>Type of compound</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mimosine</td>
<td>↑LL-37</td>
<td>Stabilizes HIF-1α</td>
<td>(Fan et al., 2015)</td>
<td>Alkaloid</td>
</tr>
<tr>
<td>Vitamin D</td>
<td>↑LL-37, ↑hBD1</td>
<td>Activates VDR</td>
<td>(Wang et al., 2004b)</td>
<td>Steroid hormone</td>
</tr>
<tr>
<td>Butyrate, PBA and trichostatin</td>
<td>↑LL-37</td>
<td>HDAC inhibition, Activates PU.1 and VDR. Discussed in paper IV</td>
<td>(Kulkarni et al., 2015; Schauben et al., 2003)</td>
<td>HDAC inhibitors, butyrate is a fermentation product of gut microbiota</td>
</tr>
<tr>
<td>Tunicamycin and thapsigargin</td>
<td>↑LL-37</td>
<td>Activation of endoplasmic reticulum stress response</td>
<td>(Park et al., 2011)</td>
<td></td>
</tr>
<tr>
<td>T3 and T4</td>
<td>↑LL-37</td>
<td>Activation of the thyroid hormone pathway.</td>
<td>Paper IV</td>
<td>Active form of thyroid hormone and its precursor form, respectively</td>
</tr>
<tr>
<td>Tiratricol, clobetasol, auranofin, isovaleric- and isobutyric-acid</td>
<td>↑LL-37</td>
<td></td>
<td>Paper I</td>
<td>T3 analogue, corticosteroid, antirheumatic agent, PBA derivatives</td>
</tr>
<tr>
<td>Litochoic acid</td>
<td>↑LL-37</td>
<td>VDR and FXR activation</td>
<td>(Termen et al., 2008)</td>
<td>Secondary bile acid</td>
</tr>
<tr>
<td>Lactose</td>
<td>↑LL-37</td>
<td>Mechanism discussed in paper IV</td>
<td>(Cederlund et al., 2013)</td>
<td>Sugar</td>
</tr>
<tr>
<td>Entinostat</td>
<td>↑LL-37 and ↑hBD1</td>
<td>Activation of STAT3 and HIF-1α</td>
<td>Paper II and III</td>
<td>Aroylated phenylene diamine, HDAC inhibitor</td>
</tr>
<tr>
<td>Betamethasone valerate, triamcinolone acetonide</td>
<td>↑hBD2-3, ↑psoriasin, ↑RNase 7 and ↑LL-37</td>
<td></td>
<td>(Jensen et al., 2011)</td>
<td>Corticosteroid</td>
</tr>
<tr>
<td>Pimecrolimus</td>
<td>↑hBD3 and ↑LL-37</td>
<td></td>
<td>(Jensen et al., 2011)</td>
<td>Immunomodulatory</td>
</tr>
<tr>
<td>Dexamethasone</td>
<td>↓LL-37</td>
<td></td>
<td>(Kulkarni et al., 2016)</td>
<td>Corticosteroid</td>
</tr>
</tbody>
</table>
1.4.2 Dysregulation of antimicrobial peptides in disease

Antimicrobial peptides have been reported to be dysregulated in diseases ranging from autoimmune diseases to bacterial infections as well as in various cancers. During shigellosis the Shigella bacteria downregulate the expression of LL-37 and hBD1 and 3 and thereby evade killing and growth inhibition, facilitating further invasion (Islam et al., 2001; Sperandio et al., 2008). This downregulation can, however, be overcome by the use of an inducer, such as PBA (Sarker et al., 2011). *Mtb* is an intracellular pathogen able to inhibit autophagy, a mechanism important for elimination of intracellular bacteria. Rekha et al. reported that the active form of vitamin D and PBA re-activate autophagy in an LL-37 dependent fashion (Rekha et al., 2015). *Neisseria gonorrhoeae* is also able to downregulate the expression of LL-37 and the effect seems to correlate with the virulence of the strain (Bergman et al., 2005). Patients with Papillon-Lefèvre syndrome (PLS) syndrome have a mutation in the cathepsin C gene. Cathepsin C is activating neutrophil elastase, cathepsin G, neutrophil serine protease 4, and protease 3. Protease 3 is the main processing enzyme of LL-37 in neutrophils (Sorensen et al., 2001). These patients are characterized by palmar hyperkeratosis and by severe periodontitis. Interestingly, PLS patients have the same number of neutrophils recruited to infections as healthy controls. However, there is a complete lack of LL-37 in the neutrophils of PLS patients but the inactive precursor protein, hCAP-18 is abundant (Eick et al., 2014). Morbus Kostmann disease is also characterized with severe periodontitis and it has been demonstrated that there is reduced levels of both LL-37 and HNP1-3 in the neutrophils of these patients (Putsep et al., 2002).

Patients with autoimmune diseases often have increased expression of AMPs, which is considered to be a consequence of the increase of DAMPs from the injured tissue. For example, in psoriatic lesions the cathelicidin LL-37 is upregulated, possibly as a response to tissue stress (Frohm et al., 1997). As discussed earlier, AMPs are also acting as DAMPs or alarmins and LL-37 can bind to self DNA in pDCs and as a complex act as an activator of TLR9, inducing the expression and release of type 1 interferon and hence drive the inflammation (Lande et al., 2007). In rosacea and systemic lupus, LL-37 is also upregulated and is proposed to drive the inflammation by recruitment of mast cells (Muto et al., 2014; Yamasaki et al., 2007). Another disorder with dysregulation of AMPs is Crohn's disease. This is especially associated with the expression of enteric α-defensins in the gut mucosa (Wehkamp et al., 2004) and to an altered microbiota in addition to a genetic disorder of the nucleotide-binding oligomerization domain-containing protein 2 (NOD2) receptor (Hugot et al., 2001; Kobayashi et al., 2005). The sensing mechanism exerted by NOD2 is malfunctioning and the normal antibacterial response is downregulated. The barrier function is disturbed and activation of compensatory pro-inflammatory pathways occur, however leading to an overcompensation, with a chronic inflammation in the gut as a result (Philpott et al., 2014).
1.5 TRANSCRIPTION FACTORS INVOLVED IN THE REGULATION OF LL-37

1.5.1 Vitamin D Receptor

The vitamin D receptor (VDR) is a nuclear receptor which dimerizes with retinoid X receptor (RXR) to initiate transcription. The main activating molecule is 1,25-dihydroxyvitamin D₃ or calcitriol, a metabolite of the pro-hormone, vitamin D (Miller, 2016). It has also been demonstrated that VDR can use the secondary bile acid, lithocholic acid (LCA) as well as PBA as a ligand (Adachi et al., 2005; Kulkarni et al., 2015). VDR is the most studied TF when it comes to regulation of the antimicrobial peptide LL-37 and binds to the promoter of the CAMP gene that has been shown in several chromatin immune precipitation (ChIP) experiments (Gombart et al., 2005; Seuter et al., 2013; Termen et al., 2008; Wang et al., 2004b) and CAMP expression is induced in vitro upon 1,25-dihydroxyvitamin D₃ treatment (Gombart et al., 2005; Wang et al., 2004b). Vitamin D has also been demonstrated to induce the gene encoding human β-defensin 2 (hBD2) (Wang et al., 2004b).

Interestingly, only the human cathelicidin gene CAMP is equipped with a VDR binding site, a late evolutionary trait that possibly could reflect the connection between immunity, ambient UV-exposure and the subsequent increase in vitamin D levels that the loss of fur facilitated.

1.5.2 STAT3

Signal transducer and activator of transcription (STAT) 3 has been associated with various cancers, including solid tumors and blood-cancer (lymphoma, leukemia). Since STAT3 is activated in a number of tumor types, it has been associated with cancer in general. A list of STAT3 associated cancer types can be found in Buettner et.al. (Buettner et al., 2002).

Activation of STAT3 usually occurs through activation of Janus Kinases (JAKs) by cytokines, and JAK in turn activates STAT3 through phosphorylation. Phosphorylated STAT3 will dimerize and translocate to the nucleus, where it regulates genes important for proliferation, apoptosis and differentiation (Hirano et al., 2000; Li et al., 2010; Sherry et al., 2009). STAT3 is also regulated via many different cytokines and is thereby involved in the regulation and control of both adaptive and innate immunity. The general consensus is that STAT3 exhibits anti-inflammatory properties and thereby promotes tumor evasion by directing the anti-tumor immune response towards tolerance. Inhibition of STAT3 has been proposed as a novel anti-cancer-treatment by activating the immune-response against the tumor (Wang et al., 2004a). More recent data have indicated that gain of function mutations in STAT3 leads to an early onset of autoimmune diseases and that loss of function results in immunodeficiency (Casanova et al., 2012; Forbes et al., 2016; Milner et al., 2015). One example of loss of function mutations in the STAT3 gene is the causative mutation of HIES (Holland et al., 2007). This disease is characterized with skeletal malformations, recurrent staphylococcal skin abscesses, eczema and pulmonary infections as well as intestinal problems (Freeman and Holland, 2008, 2010). In a report from our group, Cederlund et.al found that HIES patient failed to produce and release AMP during colonization and bacterial infection in the nasopharynx (Cederlund et al., 2011).
STAT3 is known to regulate several AMPs, including the Regs from the superfamily of C-type lectins, as well as the S100A7 protein (Choi et al., 2013; Hulse et al.; Lee et al., 2012; Murano et al., 2014). We report the novel finding that entinostat-mediated LL-37 expression is regulated via STAT3 (paper III).

1.5.3 HIF-1

Hypoxia inducible factor 1 is a transcription factor, which is activated upon oxygen deprivation and is responsible for the cells emergency response. It has also been linked to immunity, wound repair and regeneration (Andrikopoulou et al., 2011; Cramer et al., 2003; Peyssonnaux et al., 2005; Zhang et al., 2015). HIF-1 functions as a heterodimer with α and β subunits. The β -subunit is constitutively expressed and not activated upon oxygen deprivation. The α-subunit is on the other hand inducible on transcriptional level and also stabilized during low oxygen levels. The α-subunit has a proline residue that is hydroxylated by HIF prolyl-hydroxylases (HPDs). The hydroxylated proline residue will readily be ubiquitinated and, hence, HIF-1α will be degraded. HPD uses oxygen as a cofactor and, therefore, during hypoxia, HIF-1α will be stabilized, e.g. not degraded, and hypoxia induced transcription will be initiated. HIF-1 can however be activated also when the oxygen level is sufficient. Usually this occurs through inhibition of the HPDs, which is the case for the HPD-inhibitor mimosine (Warnecke et al., 2003). Binding of STAT3 to HIF-1α will competitively stop ubiquitination by the von Hippel-Lindau tumor suppressor protein (pVHL) and hence stabilizes HIF-1α, as discussed in paper III and in Jung et al. (Jung et al., 2008). We also reveal that STAT3 activation increases the transcription of HIF-1α. Notably HIF-1α is further stabilized by acetylation of the lysine residue, K709 (Geng et al., 2012), a posttranslational modification that possibly may facilitate through the use of histone deacetylase (HDAC) inhibitors (HDACi). The constitutive expression of intestinal hBD1 has been demonstrated to be entirely dependent on HIF-1(Kelly et al., 2013).

1.5.4 PU.1

PU.1 is a transcription factor encoded by the SPI1 gene and is tissue specific to myeloid cells and generally only active during immune cell development, such as during macrophage differentiation from human peripheral blood monocytes (Celada et al., 1996). This differentiation usually takes place after PRR signaling, when there is a need for macrophages with elevated AMP expression (Mosser and Edwards, 2008).

PU.1 binds to an element designated as a PU-box, which can be found in the promoter of the CAMP gene and ChIP analysis revealed binding of PU.1 to this site (Termen et al., 2008).

1.5.5 Thyroid Receptor

Thyroid hormones and their receptors have been described as a part in the regulation of immunity (De Vito et al., 2011), including regulation of the CAMP gene (paper IV). It also regulates chemotaxis, phagocytosis, reactive oxygen species (ROS) production, and cytokine synthesis and release (De Vito et al., 2011). The thyroid hormone receptor (TR) is most
active, when it dimerizes with RXR (Zhang and Lazar, 2000). Interestingly, the TRα subunit can be acetylated at three lysine residues in the DNA binding domain, which increases binding affinity. The acetylation occurs upon treatment with the ligand T3 (Sanchez-Pacheco et al., 2009) a posttranslational modifications that possibly can be facilitated through the use of HDAC inhibitors.

1.5.6 Estrogen Receptor

There are two genes encoding the estrogen receptor (ER) subunits, ESR1 and ESR2 encoding ERα and ERβ. These subunits are tissue specifically expressed and the functioning receptor can be both a homo- and hetero- dimer. The active form of estrogen and the ligand of ER is 17β-estradiol.

Estrogen hormone and ERβ have been demonstrated to be involved in the regulation of the cathelicidin LL-37 and also regulate immune cells (Kovats, 2015; Park et al., 2014). In paper IV, the ERα pathway was identified, suggesting an involvement in the regulation of the CAMP gene.

1.6 HDACs AND HDAC INHIBITORS

Post translational modifications of proteins are another way of regulating protein turnover and function. The most common modifications are phosphorylation, acetylation, N-linked glycosylation, amidation, hydroxylation and methylation (Khoury et al., 2011). However, there are many other post translational modifications, just to mention a few, SUMOylation, ubiquitination and adenylation (Falkenberg and Johnstone, 2014). Acetylation either occurs at the N-terminal end of a protein or at a lysine residue. Removing the lysine acetylation is carried out by a class of enzymes designated as HDACs or KDACs (histone deacetylases or lysine deacetylases) (Falkenberg and Johnstone, 2014). HDACs have the main function of removing an acetyl group on the lysine residue of proteins as opposed to histone acetyl transferases (HATs). A majority of the known acetylated lysine containing proteins is de facto histones and the deacetylation of these proteins condenses the chromatin structure, resulting in decreased transcription. HATs work on the same proteins but in the opposite way, leading to equilibrium of acetylated and non-acetylated histones (Falkenberg and Johnstone, 2014). Typically, inhibition of the HDACs is thought to increase gene expression by altering the equilibrium towards a more open chromatin structure. This is however not the case since almost the same amount of genes are both up- and down-regulated upon treatment with an HDAC inhibitor (Falkenberg and Johnstone, 2014). The interplay of histone modifications is intricate and HDACs would probably be viewed in the context as a part of chromatin remodeling multiprotein complexes, where the action is dependent on the complex rather than on the specific HDAC. These remodeling complexes are often recruited by TFs and hence initiate histone remodeling (Lai and Wade, 2011; Muchardt and Yaniv, 1999) (Figure 2). There are 18 mammalian HDACs, which are subdivided into class I, IIA and b, III and IV depending on sequence homology to the original yeast enzymes. All but the class III enzymes are Zinc dependent for their activity, instead class III is dependent on NAD⁺. The protein
names are HDAC followed by a number except for the class III again, which are designated as sirtuins. Interestingly, the gene regulation is specific for the different HDACs, for example, HDACs of class I are generally associated with regulation of innate immune genes, while class IIa regulates adaptive immune genes. The enzymes HDAC1-3 of class I are inhibiting pro-inflammatory cytokine production and TLR signaling. HDAC3 for instance inhibits NF-κB signaling through deacetylation of the p65 NF-κB subunit (Chen et al., 2001). In contrast class I HDACs are associated with interferon signaling. Interestingly, this is tightly interwoven with the STAT enzymes (see paper III) (Falkenberg and Johnstone, 2014). A specific HDAC6 inhibitor of class IIb induces ROS production and increase bacterial clearance in macrophages (Ariffin et al., 2015). The inhibition of HDACs can hence have opposite effects with both pro- and anti-inflammatory activities. Moreover, HDAC inhibitors do not only act on histones of the chromatin but also modifies other proteins. Transcription factors are often acetylated, and one example is STAT3, which is discussed in paper III and in (Yuan et al., 2005). Acetylated Lys685 in STAT3 will aid in dimerization and thereby increase transcription of STAT3 regulated genes. ERα as well as HIF-1α also contain lysine residues, which have been shown to increase transcription if acetylated (Wang et al., 2001).

HDAC inhibitors are being evaluated as adjunctive cancer treatment by virtue of their potentiation of cytostatic drugs. In addition, they are relevant for other indications, such as neurodegeneration and inflammation (Falkenberg and Johnstone, 2014). HDAC inhibitors are also implicated in attempts to cure HIV, due to their potential to activate viral replication in cellular reservoirs (Shan et al., 2014).

1.6.1 Entinostat

Entinostat is an HDAC inhibitor of the benzamide-type. However, chemically it can also be defined as an aroylated phenylene diamine (APD), which is further discussed in paper II. Entinostat is selective against class I HDACs, i.e. HDAC1-3, with an IC₅₀ of sub µM concentration (Lauffer et al., 2013). Entinostat is currently in clinical trials as adjunctive treatment in a variety of cancer indications, such as blood cancers and solid tumors.
Figure 2. Gene regulation and transcription factors regulating the CAMP gene. Heterochromatin inhibits transcription by hiding the genes in tightly packed nucleosomes (a). A transcription factor can bind the DNA and facilitates de-condensation by recruiting special chromatin remodeling complexes. Inhibition of HDACs can also facilitate the less dense heterochromatin (b). Transcription requires recruitment of several factors to form an initiation complex that directs RNA polymerase II (c). Transcription factors known to regulate CAMP gene expression are STAT3, thyroid receptor (TR), estrogen receptor (ER), vitamin D receptor (VDR), retinoid X receptor (RXR), hypoxia inducible factor 1 (HIF-1), PU.1, C/EBP and glucocorticoid receptor (GR) (d).
1.7 IDENTIFICATION OF NOVEL DRUGS

Novel small molecular compounds to be utilized as drugs are difficult to find. The number of molecules that can be synthesized is vast and the numbers of targets of these compounds are large. Therefore, combining the large amount of compounds with the large number of targets in a rapid, simple and efficient way is desirable. Drugs are being developed in an ‘intelligent’ way designed to fit the active site in a given enzyme. This is, however, a difficult task, where a crystal structure is needed and thus does not represent the dynamic structure of a protein. With computer-aided drug design the limitation of the crystal structure can to some extent be circumvented by modeling the structure of a specific enzyme in a computer program and even apply dynamic properties to the structure (Sliwoski et al., 2014; Van Drie, 2007).

Although these computer programs are getting better, protein modeling takes huge processing power and the result is difficult to validate. However, when the structure is available, it is possible with additional programs to analyze a massive number of compounds for binding to the enzyme of interest. This is designated as virtual high throughput screening (HTS) (Sliwoski et al., 2014; Van Drie, 2007). For information about methods of computer-aided drug design and virtual HTS and the drugs that have been discovered by these methods, please see the excellent overviews by Van Drie and Sliwoski et al. (Sliwoski et al., 2014; Van Drie, 2007). The use of various bioinformatic tools is a necessity in research today, but nothing beats ‘reality’ and the majority of drugs discovered in the last thirty or forty years and also today is through the use of ‘reality’ HTS or by serendipity. However, a good and reliable assay is necessary in order to search for novel compounds with biological activity.

1.7.1 Assay development

Assays come in many shapes, but there are some general requirements outlined below. Overall, the main aim of the selected assay is to reach a result in an easy, fast and cheap manner.

The choice of assay also depends on the amount of information available about the final target. For example, in the current work we aimed for induction of a specific peptide without knowing which intracellular signaling pathways that are involved. We therefore designed an assay as close to in vivo conditions as possible; in this case a so called ‘cell-based assay’.

Alternatively, if we had been searching for inhibitors of a specific enzyme, it would have been better to first produce the enzyme in large amounts. Next, the enzyme would have been mixed with compounds from a chemical library to search for inhibition of the chemical reaction, as measured by a chromogenic substrate changing the color or absorbance, i.e. a ‘biochemical assay’ and in this case a ‘colorimetric assay’.

All techniques can be streamlined and the method chosen must be weighed against a number of considerations. What is the biology of the target? Is it an enzyme, a receptor or a gene? What equipment is available in the laboratory? What is the background and experience of the scientist who is developing the assay? Should the target be inhibited or induced? How many compounds shall be screened? The following considerations in assay development are
adopted from Huges et al. (Hughes et al., 2011) and they were used when the Luciferase construct was designed in paper I:

1. Consider whether the assay is relevant from a pharmacological point of view. Are there known compounds with activity on the target, that can be used for evaluation? If so, use these compounds to evaluate whether the assay is responding as in previous methods and can be entrusted to identify novel compounds within the desired potency range.

2. Consider that the assay must be stable enough to work within the program of a HTS environment. HTS is often done over the time of several days, and a drug screening program often lasts for years. It is therefore vital that the assay is reproducible enough to give the same response to the controls across both time and plates.

3. Consider the cost per compound, reagents are expensive and so are plates. The assay should be able to be run in microtiter plates. Plates with 96-wells up to 1536-wells are available and with more wells less reagents per compound are needed. However, with the higher number of wells per plate comes a higher cost in form of automation and plate readers.

4. Consider to evaluate the quality of the assay, which is usually done, by assessing the Z’-factor (Zhang et al., 1999). The Z’-factor method measures the span, where a hit compound can be detected and contemplates the variance in both positive and negative control signals or simply put, it gives a statistical value on the chance of detecting a hit. The Z’ factor is a number between zero and one, where 0.5-1 denotes an excellent assay. Before the screening endeavor the Z’-factor would be determined but also during the run of a HTS the quality is assessed using the Z’-factor by adding controls to each plate. This adds the opportunity to follow the quality of the assay in time and between plates. The use of the Z’-factor to determine assay quality is also discussed in paper I and in the methodology section.

5. Consider the concentration of solvents and of the library compounds, and how these factors influence the assay. Compounds of the library are usually stored and shipped in a solvent. This means that in the assay there will be two external parameters affecting the system. Therefore the solvent needs to be tested for effect prior to the HTS. In the case of dimethyl sulphoxide (DMSO) potential effects are usually not observed up to 1 % in cell based assays, while up to 10 % can be used in biochemical assays. To test compounds in a library at different concentrations are inefficient and the concentration used must be carefully considered. Typically, 1 or 10 µM is used as the final concentration in the assay, however, both higher and lower concentrations may be used depending on the library, the target and the expected outcome. Lower concentrations often give fewer hits, albeit stronger activity upon the target (more potent) and reciprocally, higher concentrations give more hits with less activity (less potent).
1.7.2 Chemical libraries

The number of possible conformations of molecules almost reaches infinity and hence there are very large compound libraries available within the pharma-industry but also commercial libraries are available. What library to use is a matter of where the screening takes place, the cost per screened compound and whether the intellectual property is well covered. The use of a ‘proper’ HTS library challenges the target with compounds far from each other in chemical space and often includes a huge number of molecules. More often libraries of focused screens are used. A ‘Lipinski’s rule of five’ library is an example of a focused library and is often utilized to remove the cost of testing molecules that most likely never will be used as drugs (Lipinski et al., 1997). The Lipinski’s rule of five hints whether a molecule is drug-like or falls outside the boundaries set for oral drugs regarding absorption, distribution, metabolism, and excretion (ADME). Other focused libraries can be made up of molecules that are known to act on targets of similar background or derivatives of a compound known to act on the target of interest. The latter case is often the step after the HTS, the hit-to-lead phase, where a hit compound is slightly altered with combinatorial chemistry to find an even better hit or a lead compound (Hughes et al., 2011). In paper I we used the Prestwick chemical library of which all compounds are known to activate various targets. All compounds have been through regulatory controls and the safety profile is known for each compound. The compounds have also been selected to be of high diversity in molecular space. This combination will give the library a high probability for hits but also for hits that can be further developed into drugs. Importantly, the Prestwick compounds can also be directly turned into repurposed drugs. Repurposing of a drug is efficient from a medical point of view, since all the safety studies already have been performed. The perspective of patenting the re-purposed drug is however more difficult.
2 AIMS OF THIS THESIS

The main aim of this thesis has been to identify novel inducers and to elucidate the mechanism of action of AMP expression in general and more specifically the expression of the human cathelicidin LL-37.

The specific aims are:

- To establish a methodology for rapid identification of AMP inducers and to identify novel inducers and validate them in vitro. (Paper I)

- To evaluate a novel group of AMP inducing compounds and to elucidate how their effectiveness are related to the chemical structures. (Paper II)

- To study the LL-37 inducing mechanism of entinostat, a compound identified in paper II. (Paper III)

- To investigate the mechanism and signaling pathways involved in PBA and lactose mediated induction of the CAMP gene through the use of label free mass spectrometry coupled to proteomic analysis. (Paper IV)
3 METHODOLOGY

General aspects: Most methods utilized in this thesis are well established and described elsewhere. Western blot and quantitative real-time PCR, immunohistochemistry and cell culture work is described in (Kai-Larsen et al., 2007; Sarker et al., 2011) as well as in papers I, II, III and IV. Isolation and differentiation of mononuclear cells to macrophages is described in paper III as well as how cell lysate fractionation was performed.

3.1 MOLECULAR CLONING

Molecular cloning is a collective terminology of methods used in the manipulation of genetic material, often combined from different species and by utilizing enzymes with the ability to specifically alter DNA. Replication of the genetic material is crucial and almost always performed in a host organism, such as Escherichia coli. In this thesis molecular cloning has been used to produce the MN8CampLuc cell line and is described in detail in paper I.

Before early 1970’s, molecular biology and genetics were laden with the fact that genes and other genetic information could not be isolated. This notion changed dramatically with the finding of bacterial restriction endonucleases and with the identification of the DNA ligases. These findings enabled the first joining of recombinant DNA in 1972 and 1973 (Cohen, 2013; Cohen et al., 1973; Jackson et al., 1972). The amount of discoveries and advances that this methodology has brought to science in general cannot be overestimated.

The starting point of any molecular cloning experiment is the vector or the plasmid, a standard DNA backbone with all essential elements needed for replication in bacteria. To this backbone other traits are inserted, such as a selection cassette, i.e. an antibiotic resistance gene, and a ‘multiple cloning’ site for easy insertion of novel DNA. The variants of the backbone are almost endless; there are special backbones for experiments on microRNA and special backbones for overexpression experiments and others to be used for bacterial protein expression. We utilized a plasmid, pGL-4.26, with a luciferase cassette used for reporter gene assays. The plasmid includes an ampicillin\textsuperscript{R} cassette for selection of clones in the bacterial host and a hygromycin\textsuperscript{R} cassette for selection in mammalian cells.

Normally, for the construction of a reporter gene assay, only a small part of the gene promoter is inserted in front of the luciferase reporter gene. We, however, had clear evidence that part of the regulation of the CAMP gene was located to intron 1. There was also minor information on the regulatory elements upstream of the gene and to cover as much of the regulatory elements as possible it was decided to keep the reporter gene under as much endogenous CAMP gene control as possible. Molecular cloning, especially the steps regulating transformation (bacterial uptake) and transfection (mammalian cell uptake) is dependent on the size of the plasmid. Transfection efficiency declines exponentially with plasmid size above 10 kb. The backbone of the pGL-4.26 plasmid is ~ 6 kb, which left us with 4 kb to work with. Fortunately, the CAMP-gene is a rather small gene of only 2 kb, including introns. Therefore we decided to incorporate the entire gene into the plasmid and
we included a 3 kb stretch of promoter DNA upstream of the gene (Figure 3) and although this reduced the transfection efficiency (final size ~10.7 kb), we did not want to miss potentially important regulatory elements. For the cloning we utilized a seamless cloning kit allowing for insertion of fragments into a plasmid without the use of restriction enzymes. Instead we introduced long overhangs by PCR on both the insert and the plasmid (A1-A4 in Figure 3). These fragments were mixed with an enzyme that facilitated recombination, i.e. seamless cloning. This technique allowed for insertion of the CAMP gene fragment in frame with the luciferase. It also permitted us to easily remove the stop codon of the CAMP gene and the start codon from the luciferase gene. The final product would produce a fusion enzyme consisting of proLL37-Luciferase under control of all regulatory elements of the CAMP gene (Figure 3). By removing the start codon of the luciferase gene we asserted a low background luminescence signal, since all other possible start codons were located in the CAMP gene. Before the genomic integration procedure we transiently transfected the, human colonic epithelial, cell-line HT-29 with the construct and verified that the construct indeed was able to produce functioning luciferase. Stable integration demands a linear DNA fragment and a circular plasmid will randomly be linearized prior to integration. By actively linearizing the plasmid prior to transfection the chances of a successful experiment were highly enhanced. The vector utilized could be linearized within a 3.1 kb stretch of DNA. If the linearization would have been allowed to occur randomly, only 29% of all integration events would have been functioning and 68% of all hygromycin resistant clones would be identified as false positives. We used the restriction enzyme SpeI for linearization (Figure 3).
Figure 3. Schematic representation of the vector used in paper I, including genomic integration, transcriptional and translational end product. The pGL-CampLuc vector was produced by fusion of the CAMP gene into a pGL-4.26 vector (a). After stable transfection and selection using hygromycin, the fusion gene produces fusion CampLuc mRNA under endogenous-like CAMP gene control (b). The ensuing fusion protein was designated proLL37-Luciferase and expression can be quantified utilizing Luciferin and a luminometer (b).

At the end of our experiment we obtained a high success rate and could choose between several clones with different phenotypes. The MN8CampLuc cell line was propagated from a clone of moderate background and with a potent response to known LL-37 inducers.

The MN8CampLuc cell line has been invaluable to all of the following work in this thesis. However, any reporter cell line has certain drawbacks, which needs to be taken into consideration. The stable integration is not under endogenous control and the place of genomic insertion is random. Thus, stable integration can alter the phenotype of the cells and affect specific responses. It has therefore been important to always verify our results in the
parental cell line HT-29. At the time of the construction, a novel technique of zinc finger nucleases, which allowed for exact insertion into the genome, arrived on the market. This method enabled the possibility to insert luciferase in frame with the endogenous \textit{CAMP} gene and actually have full endogenous control over the reporter gene insertion site. The technique was however very expensive and time consuming. Therefore, we did not continue with site specific integration of the luciferase-reporter construct. Today, on the other hand, with the rapid development of the CRISPR/Cas9 technology, site specific integration would be possible in a fast and easy way (Cong et al., 2013; Jinek et al., 2012; Jinek et al., 2013; Mali et al., 2013).

3.2 CHROMATIN IMMUNO PRECIPITATION

Chromatin Immuno Precipitation (ChIP) coupled to qPCR is a technique developed to analyze the recruitment of proteins to DNA or to additional proteins bound to DNA. In short, the method allows for discriminating between increased or decreased binding of a protein to a general specified part of DNA. The prerequisites demand that a protein of interest and a stretch of DNA is selected for analysis of differential binding.

In this work we used ChIP in paper III for determining whether STAT3 and HIF-1α were recruited to the \textit{CAMP} gene promoter upon entinostat stimulation. Cells were treated with entinostat or the vehicle control. In this case cells were fixed using formaldehyde, which caused covalent binding of DNA and proteins in proximity to each other. Next, cells were lysed and the lysates, which contained a chromatin mixture of DNA and bound proteins, was sonicated. The sonication broke DNA into random fragments of approximately 400-500 bp. The size of the fragments can be regulated by altering sonication cycles, pulse lengths and sound intensity. These 400-500 bp DNA fragments carried bound proteins of various sorts and by utilizing a specific antibody, in this case against HIF-1α and STAT3, only fragments that carried these proteins were fished out. This procedure resulted in an enrichment of fragments with bound HIF-1α or STAT3. Next, the fixation was reversed and the DNA fragments were purified and separated from the proteins. By utilizing primers close to the putative binding site of the particular protein, we determined if there was an increase in binding in treated cells compared to unstimulated cells.

The technique is invaluable when it comes to determine if a recruited protein occurs at the promoter of your gene of interest. However, there are some important limitations to this technique. First, the protein of interest must be known and there has to be specific antibodies against it. The general binding site would also have been identified. There is the possibility to construct several primer pairs that cover a large part of the DNA if the binding site is unknown. The method does require some material though and therefore the number of primer pairs cannot be too large. Another drawback with ChIP-analysis is that it cannot distinguish between stationary and recruited transcription factors. For example, the VDR is considered to be stationary on the \textit{CAMP} gene, i.e. it acts as an inhibitor of transcription until it is activated by the ligand calcitriol (Seuter et al., 2013). This means that although the gene is regulated by VDR, the recruitment seen as the difference between stimulated and unstimulated cells will
not be visible, since there is no increase of VDR binding. Finally, ChIP-analysis is highly sensitive to methodological errors, due to many washing steps, small volumes and the risk of losing material during the many experimental parts. Thus, there is a need to be careful and to include proper controls for all steps in order to obtain reliable and reproducible results.

Utilizing ChIP with subsequent deep sequencing (ChIP-seq) will increase costs, but the problem with stationary TFs will be circumvented, since the difference is compared to both the average signal and to the unstimulated cells. Interestingly, read densities, utilizing ChIP-seq is indicative of the protein/DNA binding affinity. This allows for comparison of preferential protein binding to different promoters and could possibly be used to compare the binding of, for example, VDR to different genes encoding AMPs. The binding sites that are needed to be predicted before ChIP-qPCR can instead be observed in the results from ChIP-seq, and the exact binding site can be determined within 10 bp (Jothi et al., 2008). ChIP-seq is limited to sequencing read lengths and to repetitive regions in the genome and of course to the cost. The limitations are likely to be circumvented with the new generation of sequencing techniques in addition to more powerful computers and software.

Finally, most datasets from deep sequencing projects are uploaded to public repositories and are free of charge for other scientists to use. Much information can be gained from these repositories and should be utilized.

3.3 PROTEOMICS COUPLED TO PATHWAY ANALYSES

In paper IV we utilized label-free quantitative mass spectrometry (MS) to illuminate the proteome of cells and how it differs with treatment. The benefit of analyzing proteins and protein levels are to circumvent all method errors that follow transcription dependent methods. With quantitative MS there is no translational regulation or transcript stability to account for and moreover, proteins are the functional end product. Utilizing quantitative MS over gel based methods, such as 2D-DIGE is rapidly winning ground. The hands on time and the time to obtain results are significantly shorter and 2D-DIGE is unable to map hydrophobic, high molecular weight proteins with extreme isoelectric points. However, the ability of quantitative MS to identify full length proteins and especially to separate between post translational modifications is still difficult (Arentz et al., 2015). Another drawback is the inability to identify ‘absence’ in quantitative MS; i.e. it is very difficult to prove the presence of ‘on-off’ proteins. Comparing label free MS method to labeling MS methods, such as stable isotope labeling by amino acids in cell culture (SILAC) or isotope tags for relative and absolute quantification (iTRAQ) mostly favors label free MS. As with any labeling chemistry there is always the preference for certain reactions, hence an introduced error in labeled entities and in results. The number of experimental steps is also much lower in the label free setting, which reduces costs. The normality in growing cells with ‘unnatural’ isotopes can also be discussed. The only drawback of label free quantitative MS is the need of an instrument with high enough resolution (Bantscheff et al., 2007; Distler et al., 2014).
Finally, the results we obtained were analyzed by the use of QIAGEN’s Ingenuity® Pathway Analysis. This way of allowing a computer program to map and analyze the vast data set streamlined the whole process and rapidly allowed for identification of potential regulatory pathways. We utilized the MN8CampLuc cell line to rapidly sort the pathways into LL-37 unrelated but affected pathways and of signaling pathways regulating the expression of the CAMP gene.
4 RESULTS AND DISCUSSION

Antimicrobial peptides were early appreciated for their involvement against infection and inflammation. Therefore, much effort has been focused on possible therapeutic applications using the peptides as treatment per se. However, this approach has so far not been very successful. There have been several attempts to administer the peptides topically in various clinical settings, however, despite several attempts none has reached the market so far (Fox, 2013; Moore, 2003).

If the body instead could be ‘forced’ to produce more of the peptides at the site of infection many drawbacks of exogenous administration could be avoided. Since many AMPs are regulated together (Sarker et al., 2014) it is tempting to search for inducers as a novel approach against infections. It is known that PBA can induce LL-37 and was initially developed to treat urea cycle disorders and thus has been evaluated for safety in humans. This means that PBA can be ‘repurposed’ on a new indication without the need of safety evaluation and in that manner be introduced on the market.

The exact mechanisms behind the regulation of AMPs are still largely unknown. With a better understanding of the mechanisms we can find more effective inducers and choose the inducer that activates a preferred pathway and hence optimize effects and reduce side effects.

4.1 PAPER I: BOOSTING INNATE IMMUNITY: DEVELOPMENT AND VALIDATION OF A CELL-BASED SCREENING ASSAY TO IDENTIFY LL-37 INDUCERS

Rationale: To search for novel inducers of LL-37 expression in an efficient manner, therefore it was important to develop a simple, fast and cheap screening-system.

Aim: To develop a cell-based luciferase reporter-system, which could be used for HTS of chemical libraries.

Results: The colonic epithelial cell line HT-29 was stably transfected with a Luciferase construct, where the CAMP gene was seamlessly integrated with the luciferase gene; i.e. luciferase will only be transcribed together with the CAMP gene. The regulation will resemble the endogenous regulatory machinery, since a large (3 kb) part of the promotor as well as all introns were preserved. The construct was verified to be fully integrated and with only one copy into the genome. The expected protein product was a fusion protein of proLL-37 (hCAP-18) and Luciferase. This new cell line was designated MN8CampLuc and was tested for responsiveness to known inducers of the CAMP gene. The cell line responded in a similar manner as the parental cell line HT-29 to stimulation by vitamin D, LCA, lactose, butyrate and PBA. During the work, a new paper by Park et.al showed that endoplasmic reticulum stress induced LL-37 (Park et al., 2011). Thus, addition of tunicamycin, an endoplasmic reticulum stress inducing compound, to the MN8CampLuc cell line resulted in induced expression of LL-37, which was in line with the reported result.
After having confirmed that the cell line indeed was responding as expected to known inducers, a small panel of HDAC inhibitors and derivatives thereof was screened. All tested compounds were able to induce the expression of proLL37-Luciferase to the same extent or more prominent than the positive control PBA. To further confirm the validity of these novel inducers, the parental cell-line HT-29 was screened with these compounds and similar results were obtained.

Prior to HTS, the assay needed to be quality assessed, utilizing the industrial standard Z’-factor (described in 1.7.1 Assay development). The MN8CampLuc cells scored 0.7 as Z’-factor, which indicated that the assay was well adjusted for HTS purposes. We decided to use the assay in 96 well plate format in order to save time, reagents and hands-on-time.

The Prestwick library consists of 1200 compounds and was delivered in 15 plates with column 1 and 12 empty for controls. Negative and positive controls were added to the empty columns. Cells in 15 different 96-well plates were treated with the compounds in a final concentration of 10 µM and incubated for 24 hours. The next day each well in the 15 plates were inspected under the microscope and marked in case for signs of cytotoxicity. The cells were then washed, lysed and measured for luciferase activity. Fold induction for each compound was calculated for each plate by dividing the signal of the compound with the average of the four vehicle controls present on each plate. Hits were determined to be outside of three standard deviations of the average fold induction for all compounds. The temperature in the plate reader slowly increased with time, which correlated to an increase in luciferase signal intensity. To circumvent this, the calculation of the hit threshold was performed on the fold inductions, as described in 1.7.1 Assay development, rather than on the raw signal. We got 18 positive hits which is a 1.5 % hit rate, a typical number from this type of focused libraries.

The continued quality assessment of the assay during the screening-procedure gave 15 Z’-factor values ranging between 0.7 and 0.91 which was more than adequate (Figure 4).
HTS is often associated with some false positive signals and therefore it is prudent to re-run the hits. A dose-response experiment would allow both optimal concentration and false positive determinations. Out of the 18 hits, 7 could not be reproduced. At first this was considered to be false positives, however, the second experiment was run with 1% DMSO in order to have the high concentrations of the compounds in the dose-response experiment. In the HTS, on the other hand, the DMSO concentration was kept at 0.1%, therefore a re-run with the “false positives” at 10 µM with 0.1 % DMSO showed increased proLL37-Luciferase expression. Notably, cathelicidin is induced by ER-stress as shown by Park et.al (Park et al., 2011), and it has also been demonstrated that DMSO can act as an endoplasmic reticulum stress reliever (Song et al., 2012). In account of these experiments the notion that DMSO can attenuate endoplasmic reticulum stress induced cathelicidin expression is a likely possibility for the obtained results.
Since the assay is a cell based system with a reporter construct there is a chance that the responses are due to a new phenotype introduced by the transfection. Therefore it is important that the hits can be verified in the parental cell line. The validation was performed on seven of the 11 confirmed hits. Four hits, 5-Azacytidine, Azathioprine, Thioguanosine and Mercaptopurine, were rejected for further analysis since their safety profile would make them unsuitable to treat infections. Azathioprine and Mercaptopurine are used to induce immune suppression and the other two for treatment of late stage cancers, in which the safety profile is not that restricted. The seven remaining compounds were used to stimulate the parental HT-29 cell line in a dose-range around the optimal concentrations obtained in the reporter cell line dose response experiment. Three out of these seven compounds were confirmed as inducers of the \textit{CAMP} gene also in the parental cells.

In total, the two small screening-experiments utilizing the Prestwick library and selected HDAC-inhibitors produced five novel compounds with potent \textit{CAMP}-inducing capacity: tiratricol, clobetasol and auranofin from the Prestwick library along with isovaleric- and isobutyric-acid from the screening of HDAC inhibitors.

Tiratricol is an analog of T3, the active form of thyroid hormone, and the thyroid hormone pathway was found to be upregulated as described in paper IV, where we also could show that T3 in fact induced \textit{CAMP} gene expression. We know that many immune genes are regulated by thyroid hormones, however despite several attempts we have not been able to reproduce the T3 results on additional cells responding to T3, such as primary PBMC’s and macrophages.

Clobetasol is a corticosteroid also known to activate the glucocorticoid receptor (GR), which has been shown to downregulate AMP expression (Kulkarni et al., 2016). Notably, out of 17 compounds in the Prestwick library with glucocorticoid activity, only clobetasol induced LL-37. Thus, the induction exerted by clobetasol is likely very specific. The mechanism is still unknown and whether the action is mediated through the GR is not determined at this point and will be the subject of follow-up experiments in our laboratory.

Lastly, we also identified auranofin, an organo-gold compound used to relieve symptoms in rheumatoid arthritis. Auranofin is known to act on both nuclear related factor 2 (Nrf2) and NF-κB (Kim et al., 2010; Nakaya et al., 2011), however, neither of these pathways are likely to induce LL-37 to any excess. NF-κB has been demonstrated to regulate cathelicidin (Li et al., 2009; Park et al., 2011), however, in paper III, the pTRAF vector was used to identify entinostat as an inducer of HIF-1α. The same experimental setup also enabled determination of NF-κB and Nrf2-activation. Utilizing additional compounds that activated these pathways more than auranofin, we could not observe any induction of \textit{CAMP} gene expression (data not shown), hence auranofin most likely increases LL-37 expression via a pathway separate from NF-κB and Nrf2. Interestingly, auranofin has been found to be effective against both methicillin-resistant \textit{Staphylococcus aureus} (MRSA) and \textit{Mtb}. A likely mode of action of auranofin against MRSA is through the activation of the redox pathway (Harbut et al., 2015),
but increased AMP-expression may be a possible explanation (Dean et al., 2011; Pence et al., 2015; Sass et al., 2008).

4.2 **PAPER II: HIGHLY POTENT INDUCERS OF ENDOGENOUS ANTIMICROBIAL PEPTIDES FOR HOST DIRECTED THERAPY OF INFECTIONS**

**Rationale:** Most compounds initially tested in the MN8CampLuc cell line showed moderate inducing ability, and somewhat higher than for PBA. However, one compound, Entinostat, exhibited potent AMP-inducing properties. This prompted further investigation if compounds with similar structural features, (aroylated phenylene diamines, APDs) also were potent inducers. Therefore it was of importance to investigate structural requirements and to produce inducers with alternative structures in order to be able to select candidates with a suitable activity vs toxicity profile as well as to obtain proof of concept in an animal infection model.

**Aims:** 1) To develop a number of different APDs and related compounds. 2) To obtain information on how structural features affect the ability to induce LL-37 by utilizing the MN8CampLuc cell line as the first assay. 3) To obtain information if the activity is reproducible in the parental cell line HT-29 and to investigate if addition of vitamin D affects the induction. 4) To obtain in vivo data in a rabbit model of Shigellosis, showing the potential efficacy of entinostat.

*(Numbers refer to compounds depicted in paper II and in Figure 5.)*

During the evaluation process of the MN8CampLuc cells in paper I, several experiments with different possible inducers were performed. Butyrate, PBA and the two newly identified HDAC inhibitors from paper I proved that HDAC-inhibition can be associated to LL-37 induction. Therefore, the structurally unrelated second generation HDAC inhibitors, vorinostat and entinostat (5), were tested in the MN8CampLuc cell line. Despite being an HDACi that is several orders of magnitude more potent than PBA, vorinostat (not depicted) induced proLL37-Luciferase expression to approximately the same level as PBA. Entinostat (5), however, outperformed all previously identified inducers of LL-37 expression. Furthermore, the effect occurred at a concentration about 1000 times lower than any other inducer hitherto tested.

A number of different compounds with similar structures were synthesized with variation of substituents and other structural features. Thereafter these were analyzed for proLL37-Lucifase induction and the consensus was that the general structure needed for induction was the APD moiety (boxed molecule in Figure 5). Thus, a structurally new class of inducers, APDs, was discovered. It was found that many of the novel APD-compounds were 15-20 times more potent than PBA (1) to induce proLL37-Lucifase and at 100-1000 times lower concentration.

From the structure-activity studies it became clear that the substituents of the aroyl moiety (2-4, 6-9) influenced the inducing ability. Also the diamine moiety was crucial for induction
since neither nitrogen could be exchanged for oxygen without complete loss of activity (10, 11, 13). Neither could the ring structure be exchanged for an aliphatic diamine (14, 15).

Substitutions in the aroyl moiety were more forgiving with respect to loss of activity. On the other hand, removal of the carbonyl group resulted in a pronounced reduction in activity (16).

One of the derivatives was able to induce expression of proLL37-Luciferase to somewhat higher levels than entinostat, albeit at a higher concentration. This could be due to a less toxic substituent for this compound, allowing the dose response to rise above that of entinostat.

Compound 8 lacked the aroylated substituent and had reduced potency, and was less toxic than 5. However, potency is often rejected for less toxic compounds and both compound 8 and 9 were considered to be possible leads for future drug development.

To obtain *in vivo* data, entinostat was used to treat rabbits in a model of shigellosis. The results clearly demonstrated that entinostat improved clinical symptoms in the rabbit, reduced the infectious burden and induced cathelicidin-expression in the gut epithelial layer, as shown by immunohistochemistry. It should be mentioned that entinostat also is known to activate other parts of the immune system i.e. to reduce the levels of forkhead box P3 (Foxp3) in regulatory T-cells and hence activate a more pro-inflammatory and possibly a more antibacterial response (Shen et al., 2012). Entinostat has also been demonstrated to regulate immunity by inducing the unfolded protein response pathway, also called endoreticulum stress (Gameiro et al., 2016). Whether the mechanism behind the beneficial effects is due to induction of AMPs or involves a combination of additional immune functions needs to be explored. However, our comparison with even more potent HDAC inhibitors suggests that HDAC inhibition is not the major mechanism behind the induced expression of LL-37.
Figure 5. Chemical structures of the APD compounds tested for induction of LL-37. Number 1 is PBA and number 5 is entinostat. The boxed structure in the left top corner indicates the general structure of an APD.
4.3 PAPER III: ENTINOSTAT UP-REGULATES THE CAMP GENE ENCODING LL-37 VIA ACTIVATION OF STAT3 AND HIF-1 TRANSCRIPTION FACTORS

Rationale: Entinostat was found to be our most potent inducer of LL-37 expression. However, the mechanism behind this induction has not been investigated. If the involved transcription-factors could be identified, we could learn more details of LL-37 induction but also obtain insight how additional AMPs are regulated by the same pathway.

Aim: To identify transcription factors involved in entinostat-mediated induction of LL-37 and potential co-regulation with other AMPs.

Results: The entinostat inducing effect was discovered during the evaluation of the MN8CampLuc cells in paper I. These results were confirmed and further evaluated for dose-response and kinetics in the parental cell line HT-29. The optimal time for induction was found to be 24 h, while the optimal dose was defined as 2.5 µM, based on the balance between optimal effect and minimal toxicity. We also hypothesized that entinostat would upregulate several AMPs and therefore we investigated expression-levels for human β defensins. A combined action of several AMPs with different mechanism of action, leading to a multiple attack to the pathogen would minimize the development of bacterial resistance. Notably, we found that entinostat induced hBD1, an important peptide in epithelial defense (Prado-Montes de Oca, 2010), whereas HBD2 was not affected.

Entinostat has been shown to have potent immunomodulatory effects. Shen et al. reported that entinostat inhibits the effect of T regulatory cells by acetylation and activation of STAT3, which in turn de-activated Foxp3, the master regulator of T regulatory cell development (Hori et al., 2003; Shen et al., 2012). To elucidate whether STAT3 could be involved in AMP-expression, the CAMP gene was scanned for STAT3 binding elements. Interestingly, several STAT3 binding sites were identified in the promoter of the CAMP gene.

Further, the induction of CAMP gene expression could be repressed using both short hairpin RNA (shRNA) and a STAT3-inhibitor. Additional evidence for STAT3-activation by entinostat was obtained by the observation of increased expression of ‘B-cell lymphoma 2’ (Bcl2), a STAT3 downstream target gene (Carpenter and Lo, 2014) and by Western blot analysis on nuclear extracts, where STAT3 appeared to be accumulated in the nucleus after entinostat-treatment. However, no binding of STAT3 to the CAMP gene could be detected, utilizing the ChIP technique. Hence, it was concluded that STAT3 was a part of the regulatory pathway of LL-37-expression, albeit it was acting as a second mediator rather than binding directly to the CAMP gene promoter.

Given that HIF-1α, a transcription factor involved in cellular response to hypoxia, has been shown to regulate LL-37 expression, we hypothesized that HIF-1α might be activated by entinostat treatment and thereby increase expression of LL-37. Indeed, we found binding elements for HIF-1α in the CAMP gene promoter and stimulation by entinostat caused induced expression of HIF-1α mRNA in HT-29 cells. To test whether this increase in mRNA was associated with an increase in active HIF-1, the vector pTRAF was utilized (Johansson et
This vector contains the gene for a fluorescent protein under the control of tandem repeats of HIF-1 transcription factor binding sites. Therefore, entinostat was used to stimulate HEK293 cells transfected with the pTRAF vector, which revealed that HIF-1 activation was indeed increased upon entinostat treatment. In addition, utilizing ChIP, HIF-1α was shown to be enriched on the CAMP gene upon stimulation with entinostat.

Furthermore, shRNA experiments indicated that knockdown of HIF-1α completely abrogated entinostat elicited induction of the CAMP gene. Interestingly, shRNA directed towards STAT3 inhibited the entinostat induced HIF1A gene expression, suggesting an important cross-talk between STAT3 and HIF-1α.

The following mechanism of LL-37 induction elicited by entinostat is suggested. Entinostat inhibits deacetylation of STAT3 and hence increase the amount of active STAT3, which translocates to the nucleus where it acts as a dual activator by binding and activating transcription of the HIF1A gene but also by stabilization of the HIF-1α subunit, which then in turn will bind to the promoter of the CAMP gene and initiate transcription (see Figure 7).

Finally, we had access to whole-blood derived macrophages from a HIES-patient, with a causative STAT3 mutation. Interestingly, entinostat did not induce CAMP gene-expression in these cells as compared to two healthy control individuals, which lend further support to the observation that STAT3 is important for LL-37 expression (Figure 6).

Figure 6. Entinostat elicited induction of CAMP gene expression is reduced in macrophages derived from a HIES patient compared to healthy individuals.
Figure 7. Mechanism of entinostat induced LL-37 expression. Entinostat inhibit HDAC activity, which shifts STAT3 towards a more active state (a). Acetylated STAT3 is translocated to the nucleus and activates transcription of the HIF1A gene (b). Increased transcription of HIF1A will lead to more HIF-1α translated protein (c). Acetylated STAT3 inhibits ubiquitination and hence degradation of HIF-1α (d). Stabilized HIF-1α then binds to the CAMP gene and initiates transcription (e), leading to downstream translation of the pro-peptide of LL-37, hCAP-18 (f).
4.4 PAPER IV: LABEL-FREE QUANTITATIVE MASS SPECTROMETRY REVEALS NOVEL PATHWAYS INVOLVED IN LL-37 EXPRESSION

Rationale: It was found that co-stimulation with lactose and PBA, exhibited a synergistic effect on AMP-inducing properties. However, there was no information on the mechanism governing this potent induction. Identifying the mechanism could potentially lead to novel inducing compounds.

Aim: To perform a proteomic screen on cell lysates from HT-29 cells stimulated with vehicle, lactose, PBA or a combination of PBA-lactose and to use bioinformatic methods to map the proteomic data on regulatory pathways.

In previous work from our group it was demonstrated that a synergistic induction of CAMP gene expression occurred upon combined stimulation with PBA and lactose (Cederlund et al., 2013).

To elucidate the mechanism behind this induction, cells were treated with a combination of PBA and lactose or with each compound, separately. Next, the cell lysates were prepared and analyzed, utilizing Label-Free Quantitative MS to determine the complete proteome of these cells. More than 1300 proteins were identified, quantified and protein quantities were normalized to the protein content of untreated cells. Proteins recognized to be up- or downregulated were mapped to the ingenuity knowledgebase (IKB) and pathways were predicted, utilizing the ingenuity pathway analyzer (IPA). IPA also assigned significance values to the suggested pathways and ranked them accordingly. Utilizing the MN8CampLuc cell line developed in paper I 11 pathways were suggested as activated by IPA and were evaluated by treating the cells with either agonists or antagonists, to crucial components of the identified pathways. Three pathways were induced or blocked with specific agonists and antagonists, utilizing MN8CampLuc cells: i.e. the thyroid hormone receptor/retinoid X receptor (TR/RXR) pathway, the eicosanoid signaling pathway and the steroid biosynthesis pathway. Stimulation of MN8CampLuc cells with the active thyroid hormone (T3) and the proform (T4) resulted in increased luminescent signals. Moreover, the combination of lactose with T3 and T4 indicated an additive effect, whereas PBA and T3 exhibited a more synergistic effect. Notably, the parental cell line HT-29, did not respond to T4-stimulation, while T3 induced CAMP gene expression within the short span of 6 h of stimulation. Thus, these results indicated an important difference between the reporter cell-line and the parental cells. mRNA degradation is a regulatory pathway often applied to immune genes and it is possible that the CAMP gene is regulated this way and that the extended mRNA of the CampLuc fusion gene is more stable and hence prolong the optimal time of induction (Schott and Stoecklin, 2010). Also consider the time lag between transcription and translation as well as the stability of mRNA versus protein. At this point, we cannot explain why T4 failed to produce a response in the parental cells, but we plan to elucidate the details in future projects.

T3 is a ligand of the thyroid hormone receptor (TR), and it is known that HDAC inhibitors can acetylate various transcription factors. Sanchez-Pacheco et al. reported that the TR can be
activated via acetylation, which could explain the synergistic effect of T3 combined with the HDAC inhibitor PBA (Sanchez-Pacheco et al., 2009).

The Eicosanoid signaling pathway was examined using several antagonists of eicosanoid synthesis. MN8CampLuc cells were treated with PBA and lactose separately or in combination and mixed with the antagonists. We found that inhibition of cyclooxygenase-2 (COX-2) was enough to significantly reduce PBA/lactose induction of proLL37-Luciferase expression and this could be validated in the parental cell line HT-29, utilizing the COX-1/2 antagonist acetylsalicylic acid (ASA). Given that COX-2 inhibition would deplete the levels of PGE2, addition of downstream prostaglandin E2 (PGE2) would rescue CAMP gene expression. However, we could not prove this hypothesis in our experiments. One explanation for this unexpected result is that the eicosanoid signaling pathway is a downstream target with an AMP inducing feed-back mechanism since cathelicidin has been demonstrated to induce the production of both COX-2 and PGE2 (Chamorro et al., 2009). Another explanation is that there are other COX-2 generated lipid products that are responsible for the regulation. Bernard and Gallo have demonstrated that hBD2 and hBD3 also are COX-2 regulated and that the lipid products PGD2 and 15-deoxy-Δ12,14-PGJ2 is inducing these defensins (Bernard and Gallo, 2010). Interestingly, Park et al. introduced sphingosine-1-phosphate as a key inducer of human cathelicidin and there are indications that sphingosine-1-phosphate increases expression of COX-2 (Hsu et al., 2015; Park et al., 2016; Park et al., 2014; Rumzhum et al., 2016; Volzke et al., 2014). There are additional links between lipid mediators and LL-37, including the interconnected circuits of activation and production of LL-37 and leukotriene B4 (Wan et al., 2011; Wan et al., 2007; Wan et al., 2014a).

The steroid biosynthesis pathway was also probed by inhibition studies on the MN8CampLuc cells. PBA and lactose, separately and in combination, were co-incubated with Simvastatin, an inhibitor of HMG-CoA reductase, which would block the production of cholesterol and isoprenoids. Simvastatin inhibited CAMP gene induction by PBA/lactose and the effect of PBA per se, but not the effect of lactose. The results were validated in the parental cell line HT-29, with the difference that the effect of lactose mediated induction also could be inhibited by simvastatin. The pathway was independent of the downstream products mevalonate and cholesterol since addition of these compounds did not rescue induction of the CAMP gene. These experiments suggest a HMG-CoA dependent but cholesterol-independent regulation of LL-37 expression. More experiments must be performed to prove this fact but there are indications that statins may exhibit anti-inflammatory effects. Statins have been shown to affect several signaling pathways, including the NF-κB, AP-1, and HIF-1 pathways, which all are important for the regulation of different AMPs. Interestingly, all these pathways are down-regulated by statins, which is somewhat in line with our results (Ganz, 2003; Kida et al., 2006; Peyssonnaux et al., 2005; Rius et al., 2008; Tousoulis et al., 2014).
5 CONCLUSIONS

In conclusion, our results indicate that LL-37 is part of a complex regulatory network and is not merely a response to infection and wounding. It seems that LL-37 and other AMPs are carefully regulated also during non-stress situations in order to reach homeostasis between the host, the commensal and the pathogenic microbiota.

- Paper I: A well functioning assay was produced and 5 novel inducers of LL-37 expression were reported.

- Paper II: Entinostat was found to be a potent inducer of LL-37 expression, however, the inducing mechanism was only in part due to HDAC inhibition. Altering the chemical structure of entinostat displayed which parts of the molecule that was important for achieving induction.

- Paper III: By combining several classical molecular techniques we could conclude that LL-37 was regulated by entinostat through the transcription factors HIF-1 and STAT3. We also showed that entinostat induced LL-37 expression was impaired in a HIES patient with a known STAT3-mutation.

- Paper IV: Utilizing label free mass spectrometry we have gained full insight into how different stimuli alter the proteome of the cell populations. Strategic use of proteomic pathway analysis tools and quick examination of pathways utilizing the MN8CampLuc cells showed that this methodology is a resonable alternative to large HTS. We could demonstrate that three of several proposed pathways indeed regulated LL-37 expression.
6 FUTURE EXPERIMENTS OR THINGS I WOULD HAVE LIKED TO DO

There are several experiments that could be performed to gain more insight into the regulating pathways of AMPs in general and for LL-37 expression in particular.

1. Since HDAC inhibitors are potent inducers of LL-37 but the detailed mechanism of action remains elusive, I suggest that a screen of a larger HDAC inhibitor library is investigated. Thereby, a better and more specific understanding of the role of HDAC-inhibition in the regulation of AMP-expression can be gained.

2. Another way to further dissect the mechanism could be to alter the chemical structure of entinostat to obtain a compound with retained LL-37 inducing activity but with no HDAC-inhibitory activity.

3. The development of CRISPR/CAS9 is a true revolution in molecular biology and enables rapid gene-editing. This technique could be used to both knockdown and knock in genes important for AMP expression, which will greatly advance the knowledge of AMP regulation.

4. To obtain additional information on transcription-factor of relevance for AMP-expression, there are several new and powerful techniques available. For example, The use of enChIP coupled to MS, a method for capturing specific regions of DNA, utilizing CRISPR/CAS9 and ChIP, followed by analysis of the protein composition in this region could also be of great value to deciphering which factors drive or inhibit AMP transcription (Fujita and Fujii, 2013). In addition, ChIP-seq is another powerful method to find transcription-factors. Many experiments using this method have been performed where different stimuli and different antibodies have been used, all relevant for AMP expression. Notably, these data are often made available and can be collected from various web-based depositories. These data-sets can be re-evaluated with regard to AMP-expression. Most likely, there is a wealth of relevant information available in silico.
7 ACKNOWLEDGEMENTS

The work in this thesis has been carried out at the Karolinska Institutet, Department of Medical Biochemistry and Biophysics, Chemistry I and II and at the Department of Laboratory Medicine, Clinical Microbiology. I would like to thank the institute and both departments for providing a stimulating environment for research and aiding me in my work and thesis writing.

I would like to thank all co-authors, donors and patients, their contributions have been invaluable to my work.

I was told, I should start with my acknowledgements, “you don’t want to forget anyone because you are feeling stressed in the end”. Well, sorry if I have left anyone out!

I’ve been fortunate to have two main supervisors during my PhD, and even though Birgitta Agerberth isn’t my main supervisor on paper any more, she has not lost interest in me. She took me in and has smoothly guided me through these last 5? years and sculpted me in her own image. Now everything is ‘novel’ to me and ‘furthermore’, I have done more ChIPs than I want to remember, ‘however’ with a great mind and always being friendly, she is the image of the scientist I want to be! I will always be grateful to you for all your help, guidance and support, many thanks!

My main super-dupervisor Peter Bergman’s picture is in Wikipedia filed under ‘Effective’. Apparently, he published a paper on amyloids and wrote another about fruit flies while I was writing chapter 1.6.1. No need for Google when he’s around, he knows everything. Additionally, he’s professional, friendly and importantly, always takes to new ideas with vigor! Your help and guidance has been invaluable. Thank you!

I would also like to thank my Icelandic Viking-like co-supervisor Guðmundur Hrafn Guðmundsson. A man teeming with ideas and like all respectable scientists, he knows his way around molecular cloning. My only regret is that I never got to work in his lab in Iceland, apparently they don't use incubators there; they just have flasks hanging at predetermined lengths over volcanos.

Erica Miraglia, I could only have dreamt of a better co-worker. I wouldn’t have wanted to do this without you and you made coming to work so much more fun (you other guys were also fun)! We were a great team, miss you! Oh, I once made her blow into a USB connector on the back of her computer to get her internet going…

Rekah Sultana Rokeya or Rokeya Rekah Sultana or Rokeya Sultana Rekah, we just call her Rekha. Or Rokeya. We don’t really know. She’s the best of us. Runs ELISAs because they’re fun!

From Torpet, Monica Lindh, the one to lean on and ask how to calculate molarity. Now she mostly grows stuff or when I think about, she mostly throws stuff. Matte had to help her for
an entire day. Well she too needs some help because she helps me the rest of the time. Thank you Monica!

We got a new roommate and group member, Sim Salah ‘Zangenah’ Bim var han där. A man with many interests, a hairy chest and he only listens to KENT. Thanks for all the lunches and interesting discussions about science, but mostly about society. And flowers. And hedges. And razorblades. And excavators, cameras, cars, trailers, and people.

Andreas, haxball and wordfeud master. Most of all, thank you for everything. Kobi lomba.

Terese, Terese, Terese.

All the other people at the AMP lab. Ylva Kajjan Larsson, whisperer of not so secret stuff. Protim, teacher of bad words and he only eats yellow snow food. Cina, Linda, Mislav, Savitha, Christophf, Sadik, and a few more I’m sure I’ve missed.

From Preusseen, Michael LandrEh, will be remembered for the thickest, thesis on family guy science, ever. But he thought It was going to say something different. Its nice that you are coming home. I’ve missed you.

Limpan, computer genius who has probably been infected with a computer virus at some point.

Åke and Essam, best friends and forever my friends.

Åke, Susi, Susanne and fellow teachers at MBB.

Elias, Marcus and Katarina from MBB

Min, Jia, Dolores and the rest who were present at Chemistry 2 when I was there.

Some administrative people who are special to me are Svante, Chad, Torbjörn och Johanna så klart. Tack för festerna och all dans.

Thanks Ella, Carina, Marie and Gunvor, and to the rest of you at former Chemistry 1 and MBB, thanks.

Our friends from the physiology department, Lennart, Anne and Ellinor.

Roger and Håkan from BioNut.

Everyone at the new place, LabMed, and thanks to the welcoming committee of Las Frelin, Matti Sällberg, Anders Sönnerborg, Andrej Weintraub, Anna and Annica Karlsson.

I also want to mention some people I’ve talked to a little extra, Shuba, Marjan, Xioshan, Kerstin, Sępöjde, Mohammed, Markus, Sophie, Marita, and many more I realize now. Thanks to you all.

Jovanaka, for helping out in the very last minute.
My comrade Babbi and friend BomBomLas from Whatsapp. Thanks for all the selfies and laughs and for being my friends, I hope to see you both IRL someday soon.

A big hug and a big thank you, to my mentor and friend Ylva Terelius. She is the reason why I got in to science in the first place. I was hoping that maybe you wanted to continue mentoring me?

To all my friends from and around Fyfa, Anna P, or R, let us call her Kråkan or Kajjan. She is dear to me and therefore has many names. Magda with the greatest laugh in the world! Gukken, världens bästa Gukken! Gukkens friend Larsa, fun at parties. Björn, who left for Göteborg. Torun, also left for Göteborg. Peder Peppar! Calle, Maria, Aki, hon i flickan, Annica, Sara, Klas, Marcus, Lullan, Pixi, Etienne, Isa, and many more. I forgot Alvin.

Thanks!

Tack till alla mina andra kompisar, och mina kära kusiner. Ni är bäst!

Alla andra i hela släkten också!

Christine och Per som försett mig med en sån fin och väluppförd kvinna.

Min tuffa syster Katja som fortfarande är min stora idol och mina fans Ragnar och Birger. Vi ska nog se till att det blir vetenskapsmän av er också.

Mina älskade föräldrar, Lisbeth och Jarl eller mamma och pappa som jag kallar dom, är mina största beundrare och utan er så hade det så klart inte blivit nånting. Jag älskar er!


Jag avslutar min lista med att stolt nämna, Sara, min vackra kvinna! Utan dig vill jag inte vara. Tack för all din uppmuntran och hjälp under den här tiden, Jag älskar dig och ser fram emot vår bästa sommar någonsin. Puss!
REFERENCES


Hirano, T., Ishihara, K., and Hibi, M. (2000). Roles of STAT3 in mediating the cell growth, differentiation and survival signals relayed through the IL-6 family of cytokine receptors. Oncogene 19, 2548-2556.


epidermal antimicrobial protein expression in atopic dermatitis and in EFAD mice by pimecrolimus compared to corticosteroids. Experimental dermatology 20, 783-788.


Johansson, K., Cebula, M., Rengby, O., Dreij, K., Carlstrom, K.E., Sigmundsson, K., Piehl, F., and Arner, E.S. (2015). Cross-talk in HEK293 cells between Nrf2, HIF and NF-kappaB activities upon challenges with redox therapeutics characterized with single-cell resolution. Antioxid Redox Signal.


