TARGETING THE FIRST BARRIER IN IMMUNE RESPONSE: DESIGN OF NOVEL TLR2 ANTAGONISTS AND BINDING MODE INVESTIGATION

Dissertation zur Erlangung des akademischen Grades
des Doktors der Naturwissenschaften (Dr. rer. nat.)

eingereicht im Fachbereich Biologie, Chemie, Pharmazie
der Freien Universität Berlin

vorgelegt von
Manuela S. Murgueitio
aus Gießen
2013

Die vorliegende Arbeit wurde von Januar 2011 bis August 2013 unter der Leitung von Prof. Dr. Gerhard Wolber am Institut für Pharmazie der Freien Universität Berlin angefertigt.

1. Gutachter: Prof. Dr. Gerhard Wolber
2. Gutachterin: PD Dr. Daniela Schuster

Disputation am: 28. Oktober 2013

Acknowledgements

Table of Contents

1 Introduction

1.1 Toll-like receptors

The Toll-gene was first discovered by Christiane Nüsslein-Volhard in *Drosophila melanogaster*. In the fruit-fly, the gene is responsible for embryonic dorsal-ventral polarity, i.e. embryos with a specific mutation in this gene show inversion or lateralization of the embryonic pattern [1, 2]. When Nüsslein-Volhard first discovered embryonic fly mutants with inversed dorsoventral-polarity, she exclaimed "Toll!" (German for great) and later named the mutated gene the Toll-gene [3]. Later on, the same gene was found to play a crucial role in the immune response of insects against fungi [4]. Nowadays, several members of this family have been identified in various organisms including vertebrates and form the family of Toll-like receptors (TLRs). In humans, ten isoforms of TLRs are known (TLR1 – TLR10), which play a crucial role in the activation of the innate immune response against invading microbial pathogens [5]. Recently, TLRs have been shown to act as key players in the development of autoimmune and chronic inflammatory diseases [6], as well as in pathogen sepsis [7] and cancer [8]. Because of this reason, the development of TLR2 antagonists has been suggested as a potential strategy for the treatment of these conditions [9]. In the following chapters an introduction into TLR biology (chapter 1.1.1), function and signaling (chapters 1.1.2 and 1.1.3) will be given, with focus on TLR2, the receptor studied in this work. Next, an overview on its role in the advent of several diseases will be outlined (chapter 1.1.4). Finally, TLR2 modulators with agonistic and antagonistic activity will be presented (chapter 1.1.5).

1.1.1 Toll-like receptors and innate immunity

The immune response of vertebrates can be divided into innate and acquired immunity [10]. Innate immunity represents the first line of host defense against invading pathogens. A rapid inflammatory response takes place and later leads to the activation of the more flexible but slower acquired immune response. Infectious nonself is distinguished from noninfectious self through a limited number of germline-encoded receptors, called pattern recognition receptors (PRRs) [11]. These are responsible for the recognition of so called pathogen associated molecular patterns (PAMPs) which are conserved molecular patterns present in microbes [10]. So far, four classes of PRRs have been described: Toll-like receptors (TLRs), nucleotide-

binding oligomerization domain leucine-rich repeat receptors (NLRs), C-type lectin receptors (CLRs) and retinoic acid-inducible gene I protein helicase receptors (RIG-I) [12]. The first family of PRRs to be discovered was the family of TLRs. They specifically recognize PAMPs like lipopeptides or dsRNA and subsequently trigger the initiation of a signaling cascade that leads to the activation of nuclear factor-κB (NF-κB) and the production of cytokines which lead to an inflammatory response [5, 10].

So far, ten and twelve functional TLRs have been characterized in humans and mice respectively. TLR1 – TLR9 are present in both species. While TLR10 is not functional in mice because of a retrovirus insertion, TLR11 – TLR13 have been lost from the human genome [10, 12]. TLRs can be subdivided into two main categories: extracellular TLRs (TLR1, TLR2 TLR4, TLR5, TLR6 and TLR10) that mainly recognize bacterial membrane components and intracellular TLRs responsible for the detection of viral particles (TLR3, TLR7, TLR8 and TLR9), which are expressed in intracellular compartments like endosomes [5].

Inflammation does not only occur in response to pathogen invasion but also in response to sterile tissue injury. Considering this ambiguity, the so-called danger hypothesis was postulated by Matzinger in 1994 [13]. He stated that the immune system does not only distinguish between host and foreign molecules, but is responsible to combat danger regardless if it comes from exogenous or endogenous origin [13, 14]. Increasing evidence has shown that TLRs not only respond to PAMPs but also to endogenous molecules released after tissue damage or non-physiological cell death, so-called danger associated molecular patterns (DAMPs) [14-16].

TLR2 and its co-receptors TLR1 and TLR6 are part of the extracellular TLRs. They are mainly responsible for the recognition of bacterial cell wall components from gram-positive bacteria and fungi. Through heterodimerization with TLR1 or TLR6 different PAMPs such as lipopeptides [17], peptidoglycanes [18], lipoteichoic acid [18] or zymosan [19] are recognized by TLR2. In the case of lipopeptides, triacylated molecules are detected by the heterodimer formed between TLR2 and TLR1 [20] and diacylated lipopeptides through TLR2 in cooperation with TLR6 [21]. TLR10 is also believed to dimerize with TLR2 [22]. Several DAMPs like heat shock protein 70 (HSP70) [23] or hyaluronan fragments [24] are recognized by TLR2 as well.

1.1.2 Receptor structure and activation mechanism

The structural biology of TLRs has recently been reviewed by Botos *et al.* [25] and Manavalan *et al.* [26]. TLRs are membrane spanning glycoproteins with a conserved tertiary structure composed of a C-terminal intracellular and an N-terminal extracellular domain that are connected by a single transmembrane helix (Figure 1) [25, 27]. The ectodomain of the receptors is responsible for ligand binding and recognition, whereas the intracellular domain triggers signaling [10].
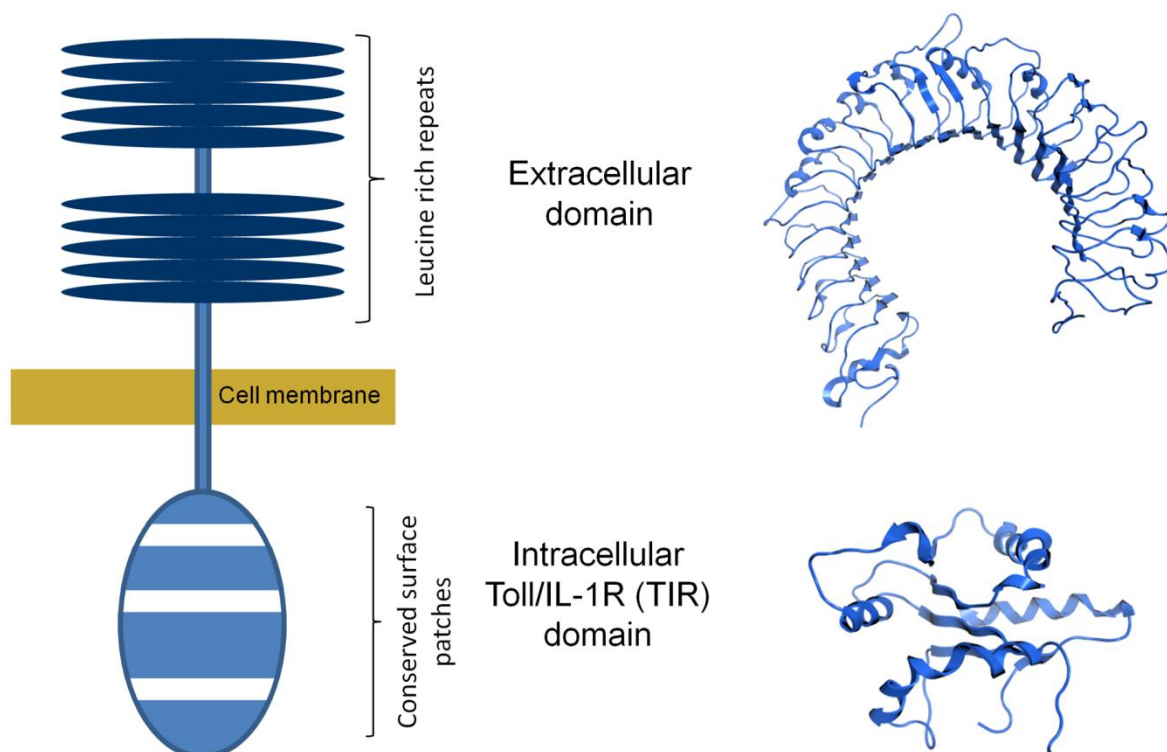
Figure 1: Overview on the tertiary structure of TLRs. Schematic representation of the receptor subunits and the cell membrane (left, [10]) and backbone representation of the crystal structure of the extracellular domain of TLR2 (PDB code: 2Z7X [20], top right) and its TIR domain (PDB code: 1FYW [28], bottom right).

Due to considerable homology of the intracellular domains, TLRs belong to a receptor superfamily that includes the interleukin-1-receptors (IL-1Rs), which is the reason why the highly conserved intracellular domain is known as the Toll/IL-1R (TIR) domain [10, 29]. It comprises a region of about 200 conserved amino acids, and is characterized by three conserved surface patches, which are crucial for signaling [30]. TIR domains are composed of

alternating β-strands and α-helices connected by loops, arranged as a five-stranded parallel β-sheet surrounded by five α-helices [28, 31, 32].

The ectodomains of TLRs are glycoproteins composed of 550-800 amino acid residues. They are characterized by 19 to 25 leucine-rich repeat (LRR) motifs. LRRs contain the leucine-rich sequence XLXXLXLXX and another conserved sequence of hydrophobic aminoacids XØXXØX$_4$FXXLX (X: any amino acid, Ø: hydrophobic amino acid) and consist of 24 to 29 residues [27]. In proteins, consecutive LRRs form a solenoid structure in which the conserved hydrophobic residues point to the inside and β-strands form a hydrogen-bonded parallel β-sheet. The closer packing of this structure causes the solenoid to bend towards the β-strands leading to the characteristic horseshoe like structure of TLRs [25, 26, 33].

In most TLRs ligand recognition occurs through binding sites located at the lateral surface of the ectodomain, which strongly differ in shape and ligand selectivity [20, 21, 34-36]. In the case of TLR2 and its co-receptors, TLR1 and TLR6, a big hydrophobic cavity is located at the border of the C-terminal and central subdomains [20, 21]. In chapter 4.1.1.1, a detailed description of these binding sites will be given.

TLRs are believed to exist as pre-assembled dimers in ligand unbound state. After ligand binding, the extracellular domains are thought to undergo conformational changes that bring their C-termini and thus the TIR domains closer to each other, creating an interaction platform for the recruitment of adaptor proteins [29]. The existence of a pre-shaped dimer and its rearrangement after ligand binding was recently confirmed for TLR8 by crystallization of the unbound homodimer and the dimer bound to small molecule antagonists (Figure 2) by Tanji *et al.* [36]. In the unliganded form of TLR8, the C-terminal regions are separated by ~53 Å making the association of the cytoplasmic domains impossible (PDB code: 3W3G). In contrast, in the agonist bound form the dimers are located in a rearranged manner causing the C-termini to be in closer proximity to each other (~30 Å), which enables the TIR-domains to interact with each other (PDB codes: 3W3J, 3W3K, 3W3N, 3W3L) [36]. Neither the unliganded form of TLR2/1 nor the one of TLR2/6 have been crystallized so far. Anyway it is likely that the rearrangement and activation of the receptors functions in a similar manner in which a pre-formed dimer exists and undergoes structural change after ligand binding (Figure 3).
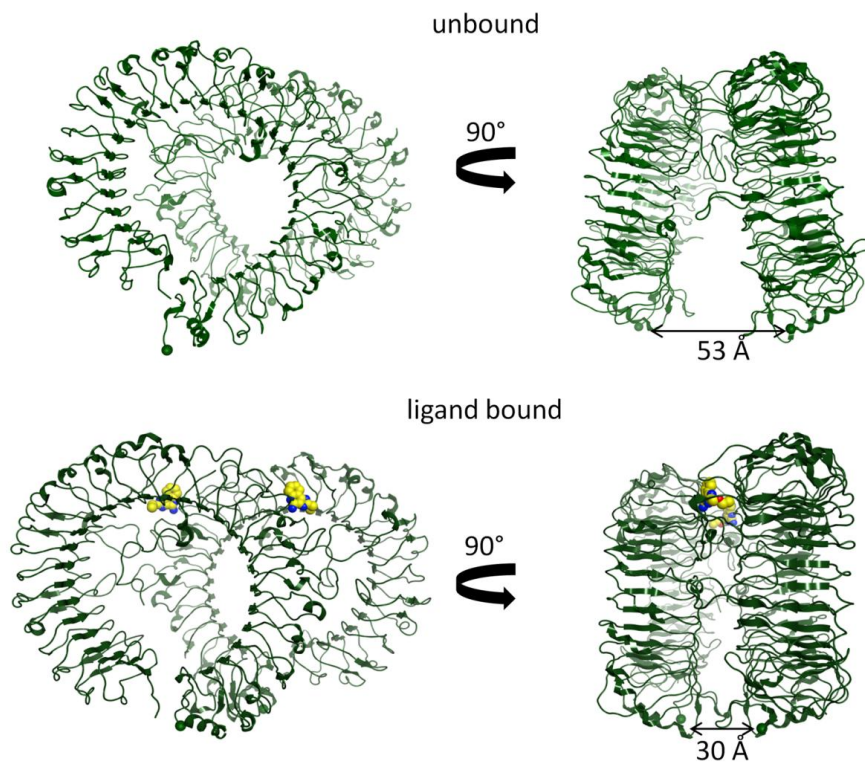
Figure 2: Crystal structure of the TLR8 homodimer in ligand unbound and bound state. On the top the crystal structure of the ligand unbound TLR8 homodimer is shown (PDB code 3W3G). On the bottom the bound form is depicted with the ligand in yellow (PDB code 3W3J). Structural rearrangement causing the C-terminal domains to come closer to each other (from 53 Å to 30 Å) is clearly visible [36].
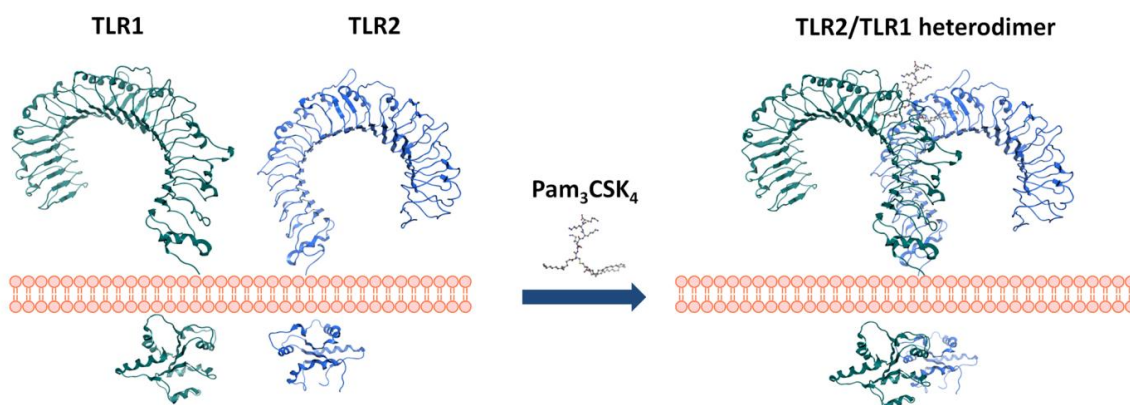
Figure 3: Schematic depiction of ligand induced dimerization of TLR2 with TLR1. The shown extracellular domains belong to theTLR2 dimer with TLR1 and Pam$_3$CSK$_4$ a triacylated lipopeptide (PDB-code: 2Z7X [20]). The crystal structure of the TLR10 TIR-domains (PDB-code: 2J67 [37]) were used to depict the intracellular domains.

1.1.3 TLR2 signaling

As mentioned before, ligand recognition and binding induces the dimerization of TLRs, which triggers an inflammatory signaling cascade. Different TLRs activate specific biological responses. Cell surface TLRs like TLR2/1, TLR2/6 or TLR5 mainly induce the production of cytokines through the activation of NF-κB, whereas TLR4 or TLR3 initiate the production of type I interferon in addition to the cytokine response [5, 38]. Evidence has been found for cell type specific signaling of the receptors [39]. In this manner, TLR2 has been shown to activate type I interferon production in response to the infection with vaccinia virus in inflammatory monocytes, but not in dendritic cells or macrophages

The differences in signaling response induced by TLRs can be explained by the existence of the TIR-domain containing adaptor molecules. These are recruited by the activated TLRs and initiate distinct signaling cascades leading to a specific response. All TLRs except TLR3 signal through the MyD88-dependent signaling pathway. In this pathway, MyD88 (myeloid differentiation primary response protein) is the first molecule recruited to the TLR TIR domain that leads to the activation of NF-κB and the mitogen activated protein kinases (MAPKs) to trigger the release of inflammatory cytokines. In contrast, TLR3, but also TLR4, first recruit TRIF (toll-like receptor adapter molecule 1), which is the first step of the TRIF-dependent pathway leading to the activation of type I interferon and inflammatory cytokines [5, 10, 38].

The response of TLR2 and its co-receptors (TLR1 and TLR6) to bacterial PAMPs mainly occurs through the MyD88-dependent pathway, of which a brief overview will be given here. More detailed information on the signaling cascade or on the alternative TRIF-dependent pathway is described in two extensive reviews of TLR signaling authored by Akira and Takeda 2004 [10] or Kawai and Akira 2010 [5]. After interaction with the receptor, MyD88 recruits the IL-1 receptor-associated kinases IRAK4, IRAK1, IRAK2 and IRAK-M. The sequential activation of the kinases starting with IRAK4 leads to the interaction with TRAF6 (tumor-necrosis-factor-receptor-associated factor 6). TRAF6 and IRAK1 then dissociate from the receptor and form a signaling complex with TAK1 (transforming-growth-factor-β-activated kinase), TAB1 (TAK1-binding protein 1) and TAB2. In the next step, IRAK1 is degraded and the complex translocated to the cytosol where it binds to the ubiquitin ligases UBC13 (ubiquitin-conjugating enzyme 13) and UEV1A (ubiquitin-conjugating enzyme E2

variant 1). Through ubiquitinilation of TRAF6 the activation of TAK1 is induced. This kinase then phosphorylates the MAP kinases and the IKK complex (inhibitor of nuclear factor-κB (IκB)-kinase complex). The IKK complex (composed of IKK-α, IKK-β and IKK-γ) then phosphorylates IκBα what leads to its degradation and allows NF-κB to translocate to the cell nucleus and start the expression of its target genes and the production of inflammatory cytokines [5, 10, 38]. An overview on the signaling pathway is given in Figure 4.
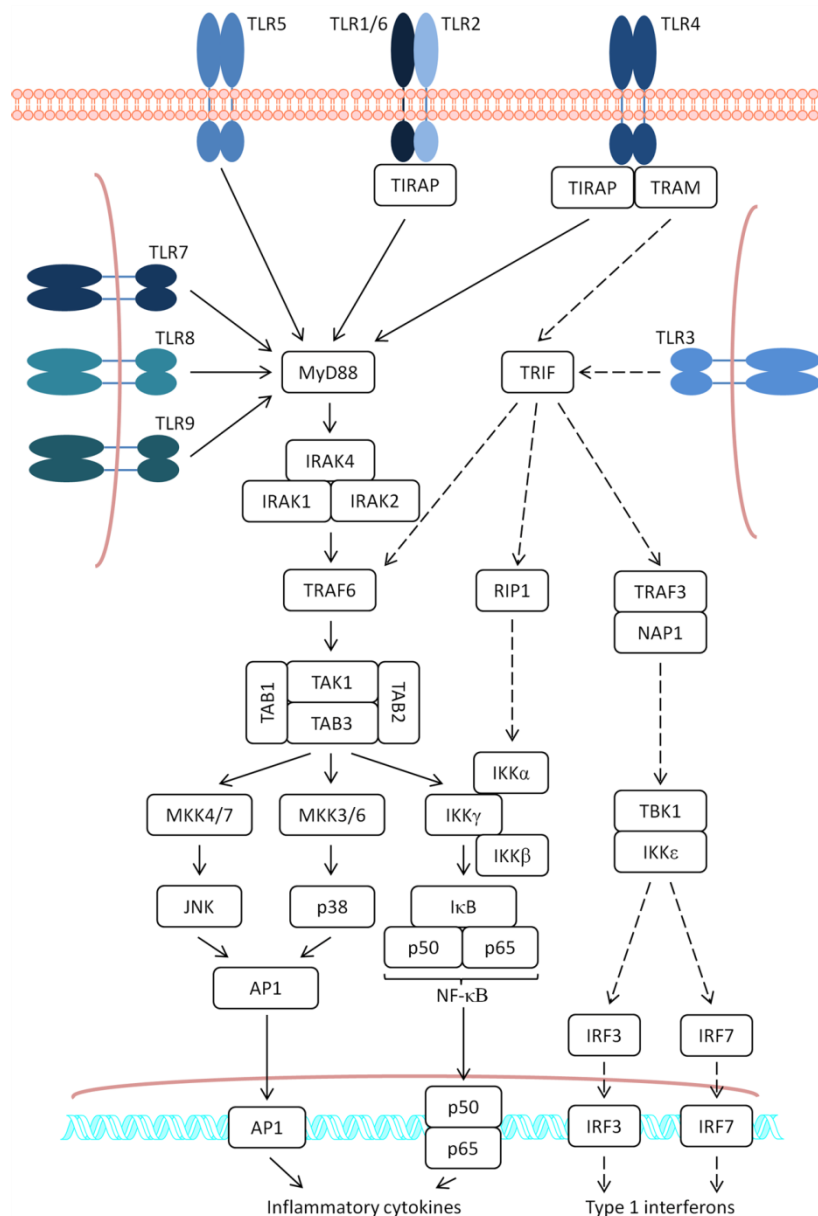
Figure 4: Overview on TLR signaling pathways. TLRs are schematically drawn in blue. The MyD88-dependent pathway is shown through continuous lines, the TRIF-dependent pathway through dashed lines [5, 10].

1.1.4 Diseases related to TLR2

TLR2 has recently been reported to be a key player in the onset and development of several disorders of the immune system [6] and other severe diseases like cancer [40]. Here a brief overview on the evidence for the involvement of TLR2 in the progression of these conditions and the potential therapeutic application of TLR2 modulators will be given.

The immune response to bacteria triggered by TLRs is crucial for the protection of the host organism. Nevertheless, sometimes this response to an external attack takes place in an uncontrolled way, leading to excessive production of cytokines and nitric oxide called "cytokine storm". This can cause severe tissue damage, organ failure and even systemic collapse [41]. TLR2/1 and TLR2/6 signaling have been reported to directly mediate the onset of Group B streptococcus neonatal sepsis through the recognition of bacterial lipoproteins [42]. Due to this fact, TLR antagonists have been suggested and investigated as suitable therapeutic tools against sepsis. Unfortunately, so far, test results for TLR4 antagonists on patients with severe sepsis have not shown positive outcomes [43-45]. Still, results for TLR2 antagonists could be different and the application of TLR2 modulators for other inflammatory disorders remains promising.

Targeting TLR2 for therapeutic purposes has also been suggested for several autoimmune and inflammatory diseases [46, 47]. It has been clearly proven that TLR2 plays a crucial role in the progression of rheumatoid arthritis by the activation of synovial fibroblasts, a crucial step in the disease process [48]. Targeting TLR2 has shown promising results in an *ex vivo* model of this condition and has accordingly been proposed for its treatment [49]. Furthermore, TLR2 has been linked to autoimmune diabetes, as the disease occurs to a much lower extent in TLR2 deficient mice compared to the wild type [50]. In addition, the modulation of TLR2 has been suggested as a therapeutic strategy against this disease after encouraging outcomes in a mouse model [51]. Moreover, the inhibition of TLR2 with the anti-TLR2 antibodies has successfully been shown to reduce infarct size and enhance cardiac function in murine and porcine models of ischemia/reperfusion injury [52, 53]. One of the tested antibodies (OPN-305) has been granted orphan status for the prevention of this clinical condition [47]. A connection between different TLR2 polymorphisms and the susceptibility to asthma and allergies, as well as to the appearance of atherosclerosis have been observed, so the modulation of TLR2 could prove to be useful for the treatment of such diseases [9].

A further area, in which increasing evidence for the important role of TLRs for disease progression has been gathered, is carcinogenesis [8]. Already at the end of the 18th century a positive correlation between bacterial infection and the remission of malignant disease (cancer) was observed [54]. Later on, the activation of the innate immunity through the TLR4 agonist lipopolysaccharides (LPS) from the outer cell membrane of bacteria was found to be the cause for tumor regression [8, 54]. This knowledge was used to establish the treatment of bladder tumors with Bacillus Calmette Guérin in 1976 [55]. Currently, SPM-105, a preparation of cell wall skeleton components of mycobacterium bovis BCG Tokyo with agonistic activity on TLR2, is being studied in preclinical studies for the treatment of cancer [47]. The preparation had previously been shown to reduce tumor growth in mice in a TLR2 dependent manner [56].

1.1.5   TLR2 modulators

As described in the previous chapter, TLR2 has vastly been proposed as a promising drug target for the treatment of various severe diseases. On the one side, TLR2 antagonists promise to be useful tools for the treatment of inflammatory or autoimmune conditions, whereas agonists could be employed in cancer therapy. To date, only very few drug-like TLR2 modulators are known. Most known agonists or antagonists are lipids that mimic endogenous ligands. However, these are less suitable for therapeutic use due to their high molecular weight, low solubility and ligand efficiency. Due to this fact, the interest in the discovery of novel small molecule TLR2 modulators is rising [57]. In the following chapters an overview on currently known synthetic TLR2 agonists and antagonists will be given.

1.1.5.1   TLR2 agonists

Most synthetic TLR2 agonists developed so far have been derived from endogenous ligands of the receptor. Synthetic tri- (Pam$_3$CSK$_4$) and diacylated lipopeptides (Pam$_2$CSK$_4$ or FSL-1) are routinely used to activate TLR2/1 or TLR2/6 signaling respectively in experimental setups [58-60]. Extensive work has been done for the understanding of necessary chemical features of synthetic lipopeptides and to optimize them for their potential therapeutic use, as well as to find alternative chemical classes of lipopeptide mimics. Buwitt-Beckmann and coworkers performed extensive work on the structure-activity relationship (SAR) of synthetic lipopeptides [61, 62]. Two ester bound lipid chains with a minimum of eight carbon atoms

were found to be necessary for a stable TLR2 response [61]. In a further study, the authors stress the importance of the peptide sequence as it heavily influences the potency of the activation, especially the exchange of the cysteine leads to loss of activity [63]. Furthermore, the authors state that TLR6 independent signaling with diacylated lipopeptides is possible depending on the sequence of the peptide part of the agonist [62]. In an independent study Okusawa *et al*. analyze the SAR of lipopeptides derived from the *Mycoplasma salivarium* lipoproteins and come to the conclusion that small changes in the peptide sequence have strong influence on the activity of the agonists [64]. Recently, several groups have made efforts to synthesize simpler and more water-soluble lipopeptides. This led to series of mono-acylated compounds that specifically activate TLR2 signaling in human but not in mice cells (**a44** and **s16** Figure 5) [65, 66]. An exemplary overview on TLR2 agonists derived from the lipopeptide structure is shown in Figure 5.

So far, only one study describing the discovery of small molecule TLR2 agonists structurally unrelated to lipopeptides has been published by Guan *et al*. [58]. Five small molecules (Figure 6) with agonistic activity on TLR2 were discovered by high-throughput screening of a compound collection comprising 24,000 compounds, among which 9,000 pertained to a private collection of the Department of Chemistry at the University of Illinois and the remaining were from the Chembridge Screening Library [67]. Compounds A, B, and C show specificity for the activation of TLR2/1 signaling, compound E activated TLR2/1 and TLR2/6 signaling, whereas compound F only stimulates TLR2/6 response in an TLR2 dependent IL-8-driven luciferase reporter assay.
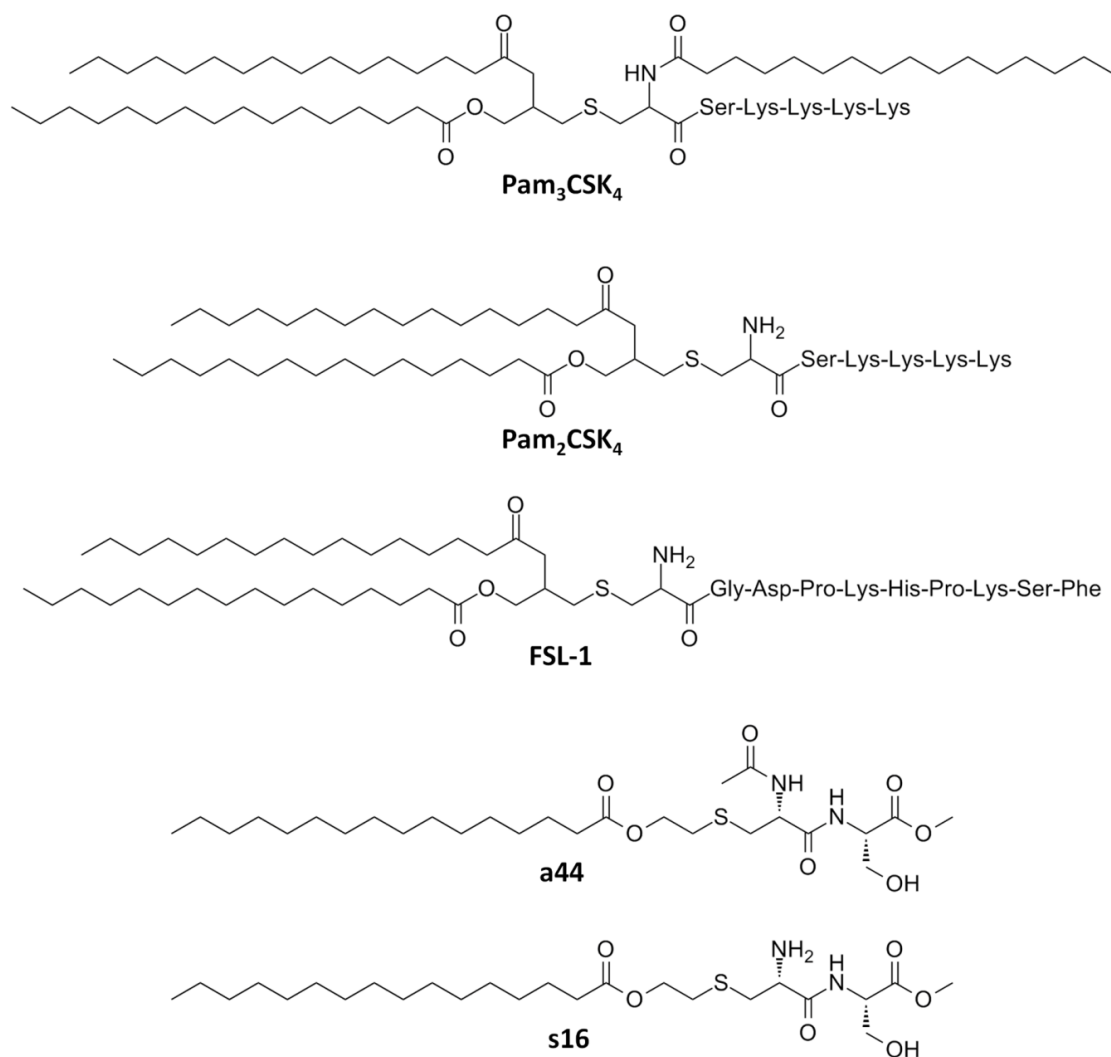
Figure 5: Synthetic lipopeptides as TLR2 agonists.
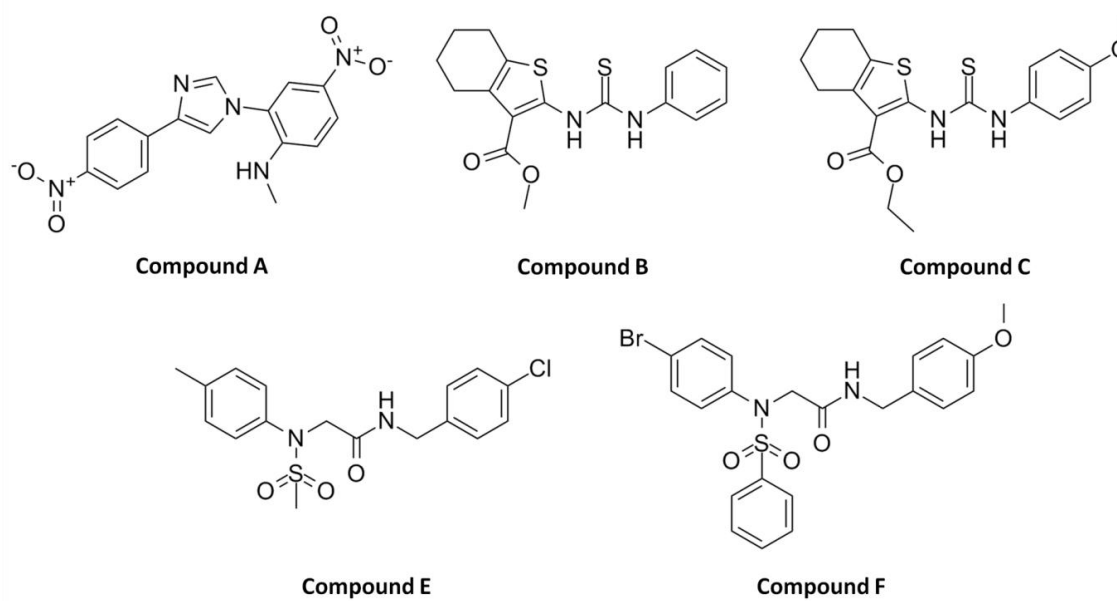
Figure 6: Small molecule agonists of TLR2.

1.1.5.2 TLR2 antagonists

A common strategy employed to develop TLR2 antagonists has been the synthesis of lipopeptide mimetics with alternative chemical scaffolds. Following this strategy, Spyvee *et al.* [68] identified a phospholipid (**s1**, Figure 7) with antagonistic activity on TLR2 (IC_{50} 3.07 μM) and TLR4 (IC_{50} 0.35 μM) by high-throughput screening. In a next step the authors performed SAR studies on the compounds that did not lead to any substantial changes in activity or specificity. In another study, Seyberth and coworkers [69] designed lipolanthionine peptides as TLR2 inhibitors. A strong relationship between the inhibitory potency and the fatty acid chain length was observed in the performed SAR studies. The most potent compound identified was a myristic acid derivative (**s17**, Figure 7).

s17

s1
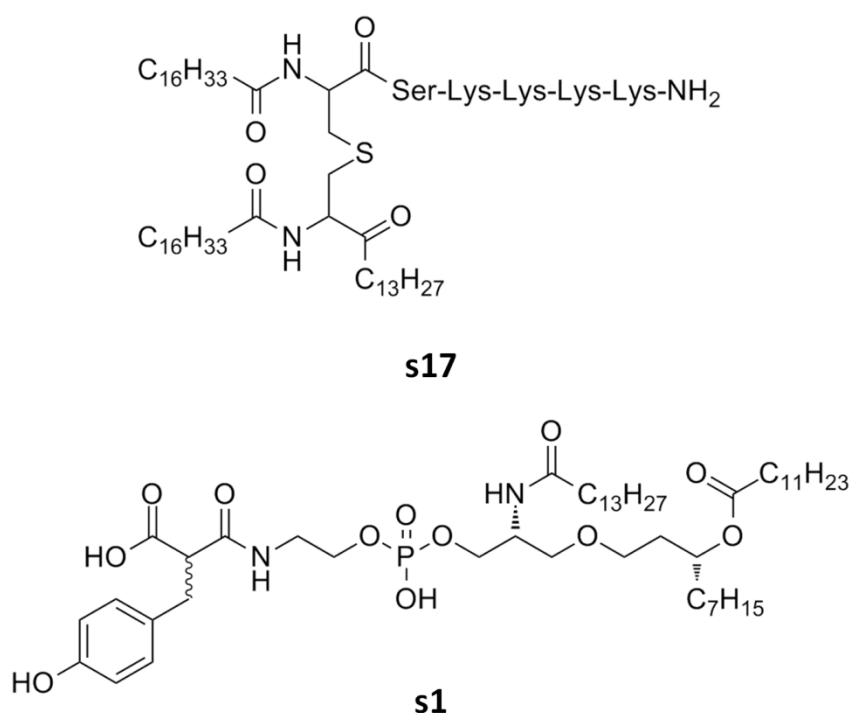
Figure 7: Lipid TLR2 antagonists

When we started our study, no small molecules with clear evidence for TLR2 antagonism were reported. Only a few compounds showing inhibitory activity on TLR2 signaling without information on the actual target protein were known, these compounds are shown in Figure 8. We distinguish between TLR2 antagonism, by which we mean binding to the receptor and impeding the signaling cascade, and TLR2 signaling inhibition, which is the case when a compound acts on any target in the signaling cascade activated by TLR2. In 2010, RSCL-0409 a synthetic gluco-disaccharide ([4-*O*-chloroacetyl-2,3-di-*O*-acetyl-6-*O*-

levulinoyl-β-D-glucopyranosyl]-(1-3)-1-*O*-(p-methoxyphenyl)-2-deoxy-2-N-trichloroacetyl-4,6-*O*-benzylide-ne-α-D-glucopyranoside) was described by Kalluri *et al*. [70] to inhibit cytokine production induced by TLR2 as well as TLR4 ligands. A study suggesting potential inhibitory activity of tricyclic molecules like amitriptyline on TLR2 and TLR4 activity was also published in 2010 by Hutchinson *et al*. [71]. Zhou and coworkers [59] performed a high-throughput study and identified E567, a compound able to inhibit TLR2 signaling activated by Lymphocytic choriomeningitis virus stimulation. A last compound described as TLR2 signaling inhibitor with unspecified target protein was Sparstolonin B (SsnB), a molecule extracted from a Chinese herb published in 2011 [72]. These compounds are shown in Figure 8.
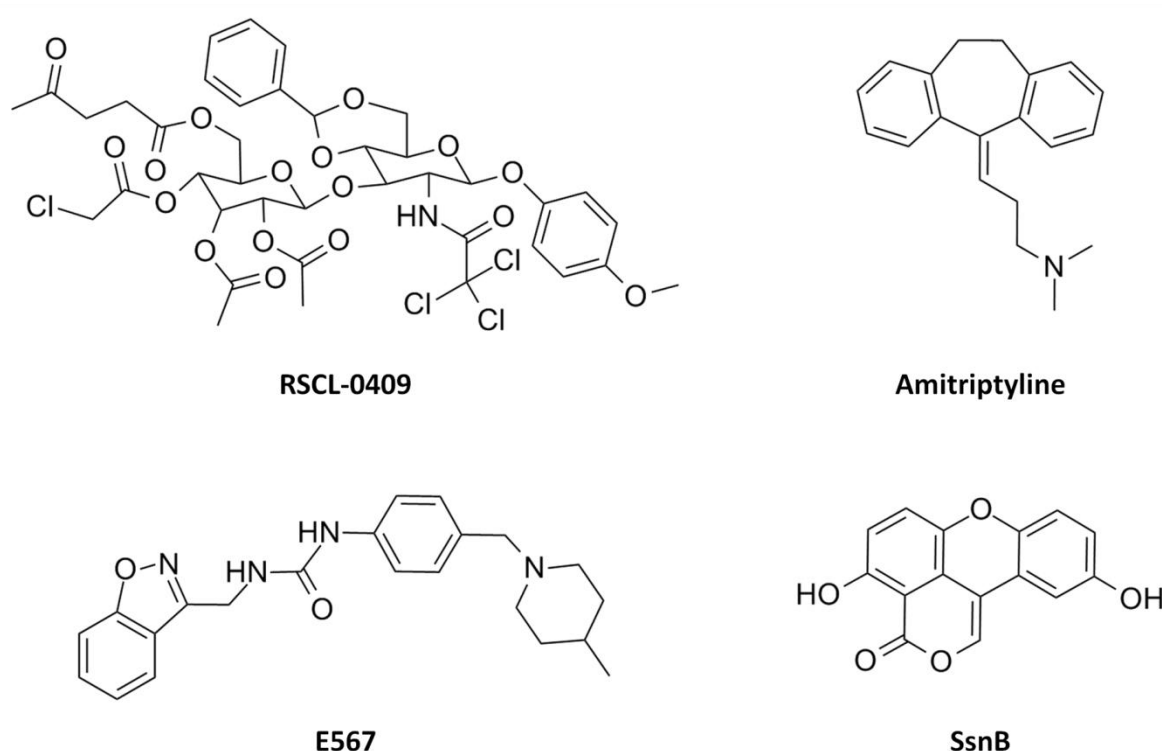
Figure 8: Compounds related to TLR2 signaling inhibition with unspecified target protein.

In a recent study published by Cheng *et al*. [60] in 2012, a high-throughput screening campaign leading to the selection of TLR2/1 small molecule antagonists and further lead optimization was presented. Out of nine initial high-throughput hits the most potent compound NCI35676 (IC_{50} 2.45 ± 0.25) was selected for further SAR analysis, leading to CU-CPT22 the most potent derivative (IC_{50} 0.58 ± 0.09) and 26 other derivatives with TLR2/1 antagonistic activity (Figure 9). This study will be discussed in more detail in chapter

4.4 with regards to the binding mode of the antagonists suggested by the authors. In chapter 4.5 the newly discovered TLR2 antagonists will be integrated into our model of TLR2 antagonism. Finally, the performance of high-throughput screening compared to virtual screening will be discussed based on this work in chapter 5.1.2.
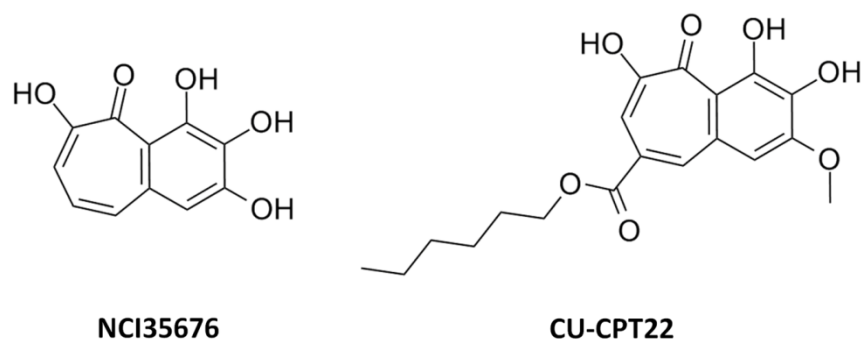
Figure 9: Recently published TLR2/1 antagonists NCI35676 and CU-CPT22 [60].

2 Aim & objectives

Toll-like receptors (TLRs) represent the first barrier in innate immune response through the recognition of pathogens. In collaboration with TLR1 and TLR6, TLR2 is mainly responsible for the detection of gram positive bacteria and fungi. However, its contribution to the onset and progression of several inflammatory processes and autoimmune diseases has been broadly reported. Hence, drugs targeting TLR2 have been advocated promising therapeutic tools for the treatment of these conditions. In this research project, we aimed at identifying new small molecule TLR2 antagonists and determining and explaining their mode of action through molecular modeling and virtual screening. For this purpose, we proceeded according to the following steps:

1. Structure-based virtual screening using 3D pharmacophore models derived from molecular interaction fields (MIFs). To generate these models, MIFs of a putative binding site for small molecule TLR2 antagonists were calculated and carefully analyzed, leading to the identification of those chemical features that are essential for binding.

2. Ligand-based virtual screening through a two-stepped shape- and feature-based search with putatively specific TLR2 modulators as query structures.

3. Integration of the experimental results of the biological validation and characterization of the compounds (performed by our cooperation partner Dr. Santos-Sierra at the Medicinal University of Innsbruck, Austria).

4. Docking studies to explain and carefully analyze the binding mode of the identified TLR2 antagonists and agonists.

5. Analysis of TLR2 antagonists published during the course of this study by other research groups with regards to their binding mode by molecular docking.

6. Generation of a refined, integrative 3D pharmacophore model collection that reflects the current knowledge on small molecule TLR2 antagonists, including the data generated in this study and the one published during its realization. This model should be able to predict whether and how a small molecule binds to TLR2.

To reach these goals, a variety of molecular modeling and virtual screening techniques were employed. An overview on the applied computational methods is given in the next chapter.

3   Computational methods

3.1   Molecular modeling and virtual screening in drug discovery

Computational techniques like molecular modeling and virtual screening have proven to be of great use in drug development and render different stages of the drug discovery process more efficient, goal-oriented and less risky [73]. This is especially true for lead discovery and optimization. A great variety of methods and tools have been developed for the generation of computational models that describe the interactions between macromolecules and ligands. These can be categorized into structure- and ligand-based methods depending on the employed experimental data. While in structure-based methods the investigations and predictions are centered on experimentally determined 3D structures of macromolecules, ligand-based methods attempt to identify crucial physicochemical and structural properties of active compounds and identify analogs purely based on ligand information. An overview on key computational approaches employed for drug discovery is given in Figure 10.

For structure-based design, the starting information is the crystal structure of the macromolecule. The main source for such information is the Protein Data Bank (PDB) [74]. The most frequently used structure-based 3D methods are 3D pharmacophores and protein-ligand docking. In the ideal setting, the ligand-binding site is already known, which is the case when the compound is co-crystallized with the macromolecule. Alternatively, cavity detection algorithms can be utilized to predict putative binding sites. Molecular interaction fields (MIFs) can then be calculated to define interaction hot spots with the macromolecular surface by using molecular probes that mimic enthalpic interactions like hydrogen bonding or Coulomb interactions, on the one hand, or entropic effects like hydrophobic contacts, on the other hand.

When no 3D structure of the studied target protein is available, ligand-based methods are employed. These methods require the knowledge of a set of known bioactive molecules, which can be used to search for novel potential analogs. Ligand-based approaches can be categorized into 2D and 3D methods. 2D methods rely on the calculation and comparison of scalar molecular properties to identify compounds with similar calculated molecular attributes (descriptors). Several methods for selecting and weighing of descriptors exist, for example

neural networks or linear correlation of measured biological activity (quantitative structure-activity relationships, QSAR). On the other hand, 3D methods search for novel ligands starting from known molecules by 3D alignment in terms of molecular volume (shape-based screening) or in terms of chemical functions (3D pharmacophores) [75].

In the following chapters, an overview on the key structure- and ligand-based modeling techniques utilized in this study will be given. First, cavity detection tools will be discussed (chapter 3.1.1). Then the concept of MIFs (chapter 3.1.2) will be introduced, followed by docking (chapter 3.1.3). Next, the two key virtual screening techniques employed in this work, 3D pharmacophore modeling and shape- and feature-similarity based screening, will be presented (chapters 3.1.4 and 3.1.5). Finally, different metrics used to assess the performance of virtual screening methods will be introduced (chapter 3.1.6). In doing so, we will focus on the software tools and methods employed in the course of this work.
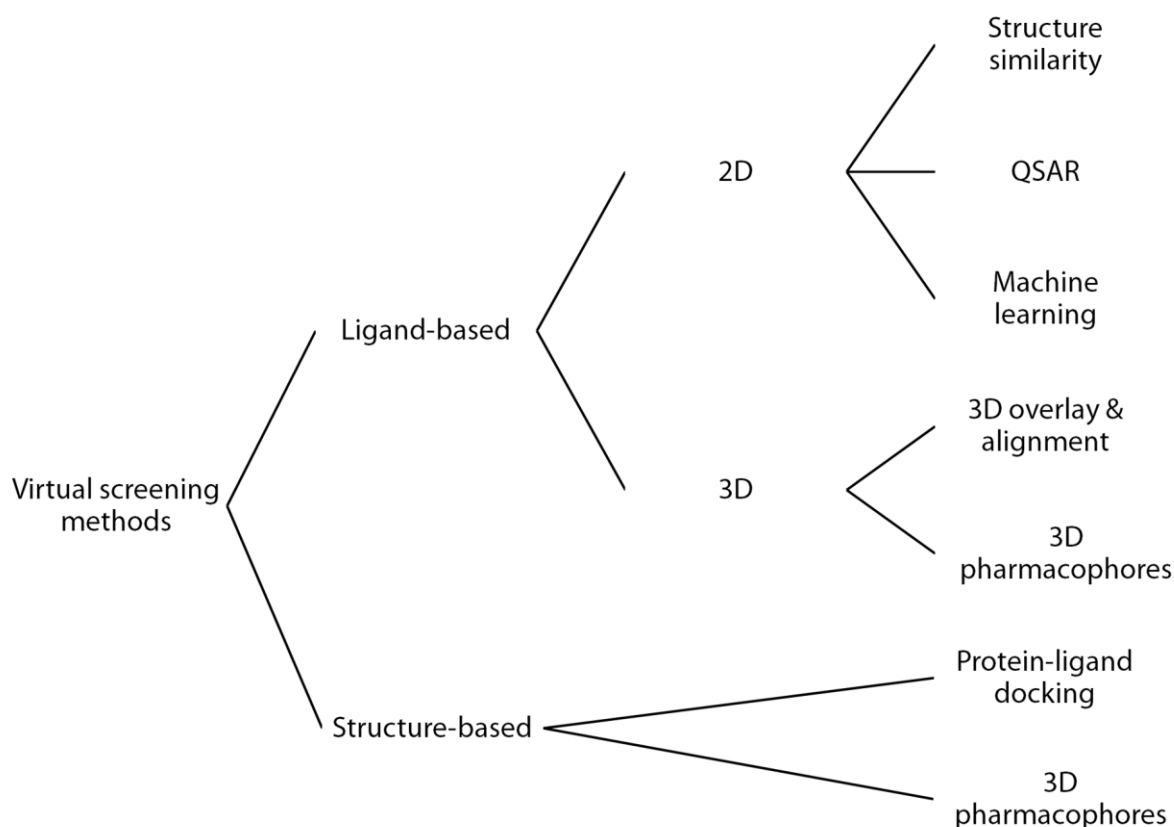
Figure 10: Overview on the main computational approaches in virtual screening.

3.1.1 Cavity detection

The first step in the structure-based design of small molecule ligands for target proteins with unknown ligand binding site is the definition of pockets suitable for small molecule binding. Several *in silico* methods for cavity detection based on the protein crystal structure have been developed. In general, they can be classified into geometry- and energy-based methods [76, 77]. Binding of small molecules usually requires the presence of a concave binding cleft on the surface of the target protein. Geometry-based prediction algorithms are based on the detection of such deepenings from a steric point of view. In energy-based methods energetically favorable locations for ligand interaction are determined [76, 77]. The principles on which the algorithms are based vary as reviewed by Laurie and Jackson [77]. Many employed method for geometry-based cavity detection are either based on a 3D grid surrounding the protein or on a definition of the molecular surface. In the grid based methods, the interface of the macromolecule surface is found by defining the grid points that do not coincide with the protein atoms. In surface based methods the protein boundaries are usually defined through a "solvent" probe that rolls across the surface [77]. Energy-based methods are based on the calculation of the interaction energy between a probe molecule and the protein at given points of its surface [77]. In general, cavity detection adds a certain degree of uncertainty to a structure-based modeling approach and therefore renders this strategy more challenging.

In our study two cavity detection algorithms were employed. The first was SiteFinder [78], a geometry-based approach. Here the binding sites are calculated though the definition of α-spheres, these are defined as spheres that contact only four atoms on their boundaries and do not contain atoms. The collection of spheres is filtered, removing those that represent inaccessible areas of the protein and areas with too extensive solvent contact. In the next step, the α-spheres are divided into "hydrophilic" and "hydrophobic" spheres and are then clustered into potential binding sites. The second binding site prediction algorithm used was Q-SiteFinder [79], an energy-based method. The non-bonded interaction energy of a simple methyl probe is calculated for grid points around the protein according to the GRID force field [80]. Probes with favorable binding energy are then retained and clustered into putative binding sites. These are then ranked according to their total interaction energy, which is the sum of the energies of the single grid points belonging to a cluster.

3.1.2 Molecular interactions fields (MIFs)

In order to design molecules that bind into a specific protein cavity, not only geometric information has to be considered. The chemical features of the atoms of the binding site directly influence the energetic characteristics of the protein-ligand interaction. Molecular interaction fields (MIFs) were first implemented 1985 by Goodford in the GRID program [80], and can be used to identify, analyze and visualize regions with favorable interaction potential. This technology is now further developed by the group of Cruciani and has led to several new methods [81-83]. The program calculates the interaction between a macromolecular system, the "target", and a "probe" that is a small molecule or molecular fragment. The probes are selected to imitate the chemical characteristics of specific molecular fragments. GRID points are established throughout and around the protein and the potential energy E_{xyz} of the probe at each point is calculated according to the empirical energy functions:

$$E_{xyz} = \sum E_{lj} + \sum E_{el} + \sum E_{hb}$$ Equation 1

where E_{lj} is the Lennard-Jones potential (i.e. steric contacts), E_{el} is the electrostatic term (including charge interactions) and E_{hb} is a heuristic hydrogen bonding function. The results can then be used to map binding sites, visualize energy contours surfaces and identify key interactions on the protein surface. Furthermore they can be employed to derive 3D pharmacophores for structure-based virtual screening [84].

The implementation of interaction potential maps within the software Molecular Operating Environment (MOE) was used in this study [85]. It is based on a similar idea as the aforementioned GRID force-field [80] with the three-term interaction energy defined in Equation 1. Several probe types are available e.g.: "N1" – amide NH group, "O" – carbonyl oxygen atoms. In addition a "DRY" – probe that calculates the hydrophobic energies is included as well, with modified interaction energy:

$$E_{xyz} = \sum E_{lj} + S - \sum E_{hb}$$ Equation 2

where S is the entropy set to -0.848 J/K this representing the favorable interactions of

hydrophobic particles in aqueous environment. The hydrogen bond function E_{hb} is subtracted in order to integrate the cost of breaking the hydrogen bonding network in the protein's hydration shell oriented towards the solvent.

3.1.3 Docking & scoring

Molecular docking is an important tool in structure-based computer aided drug design. Its aim is to predict the correct binding mode of a ligand to a protein with known 3D structure. Docking generally is described as a multi-step process. First, docking programs generate binding poses in the protein binding site. Second, the affinity of the ligand binding to the protein based on the given pose is estimated based on various scoring functions [86].

Several docking tools are available, which differ from each other in terms of ligand placement and scoring algorithms. Different strategies for pose generation have been implemented. Stochastic methods induce random changes to single ligands or populations of ligands and then assay the poses on basis of predefined probability functions [86]. Like this, GOLD [87] and AutoDock [88] employ genetic algorithms while LigandFit [89] is based on a Monte Carlo optimization. Other algorithms are based on ligand fragmentation and incremental reconstruction approaches (e.g. FlexX [90]), molecular shape (e.g. DOCK [91], FRED [92]) or on three-dimensional molecular similarity (e.g. Surflex [93]). While the treatment of ligand flexibility in the generation of docking poses is advanced, the protein is mostly treated as rigid due to the high computational costs of allowing the degrees of freedom necessary to take into account protein flexibility [86]. GOLD allows for a limited degree of side chain flexibility.

Scoring functions can generally be subdivided into three categories: force-field-based (e.g. GoldScore [94], DOCK[95]), empirical (e.g. FlexX [90]) and knowledge-based (e.g. DrugScore [96]) scoring functions [86, 97]. Force-field based scoring functions aim at calculating the approximate atomic interaction energies for the system. In empirical scoring functions, parameter coefficients are fitted to experimental data as binding affinities of known crystallized ligand-protein complexes. Knowledge-based scoring functions are designed to reproduce experimental structures more than binding affinities. Here parameters are derived from the statistical analysis of interaction distances between specific atom pairs in crystallized protein-ligand structures [86, 97, 98].

Whereas pose generation works reliably to sample possible ligand geometries, the performance of scoring functions heavily varies depending on the target protein [99], mainly due to the challenging task of predicting the entropic component of ligand binding. Regardless, docking remains an invaluable tool in drug design [100].

In this study, GOLD [87] was employed for docking experiments. As mentioned before, the pose optimization in this software is based on a genetic algorithm. An evolutionary process of data structures called chromosomes is mimicked in order to identify the ideal binding poses. The utilized GoldScore [94] is a force-field based scoring function that contains terms for H-bonding energy, van der Waals energy, ligand torsion strain and metal interactions.

3.1.4 3D pharmacophore modeling and virtual screening

The term pharmacophore was officially defined by the IUPAC in 1998 [101] as: *A pharmacophore is the ensemble of steric and electronic features that is necessary to ensure the optimal supramolecular interactions with a specific biological target structures and to trigger (or to block) its biological response.* In other words, ligand binding to a macromolecule can be described by a 3-dimensional pattern of chemical features (such as hydrogen bonds, charges, aromatic rings and lipophilic areas). These features describe both enthalpic (hydrogen bonds, charge interactions) and entropic (hydrophobic contacts) contributions to ligand binding. These do not represent real molecules or functional groups but are an abstract concept of the molecular interaction capacities of a given group of ligands towards a specific target structure [102]. As an example, primary amines, as guanidines or amidino moieties can all be described as a positive ionizable feature, which will therefore identify all ligands that show this specific chemical functionality. Since 3D pharmacophores implicitly encode the similarity of bioisosteric groups, they are especially helpful for scaffold hopping, a main challenge in drug development [103]. Due to their easy interpretability and modification pharmacophore modeling has become an established method in drug design. Several software packages for pharmacophore modeling are available, which include: Phase [104, 105], Catalyst [106], MOE [85, 107] and LigandScout [108, 109] among others. In this study pharmacophore modeling was performed using LigandScout, since it represents the first tool that allows for automated structure-based pharmacophore creation and offers the best geometric screening accuracy with high computational efficiency and reasonable virtual

screening speed.

3D pharmacophores can be derived in a ligand- as well as in a structure-based manner. Structure-based pharmacophores are derived from protein-ligand complexes, focusing on the key interactions between a bound compound and the macromolecule. In LigandScout [108, 109], starting from a PDB crystal structure, the correct hybridization states are assigned and a pharmacophore model automatically generated. This occurs based on a set of rules that detect and classify protein-ligand interactions based on chemical characteristics and geometry. In ligand-based pharmacophore generation, key chemical features of a set of active ligands are arranged in 3D space taking into consideration their conformational flexibility. The quality of the alignment is crucial for the accuracy of the pharmacophore model and its prediction power [109]. The chemical features shared by all ligands are assumed to be essential for ligand binding. Like this, a putative binding mode of a set of active compounds can be inferred without a protein crystal structure. Ligand-based pharmacophore modeling is only possible when all ligands bind to the same binding site and share the same binding mode.

Due to their simplicity and abstract nature, 3D pharmacophores can effectively be used as virtual screening filters to search large compound libraries [110, 111]. In LigandScout [108, 109], an efficient pattern-matching algorithm was implemented, which allows for improved geometric accuracy. In the ideal case, the hit compounds resulting from the database search will differ from the compounds utilized for model generation in terms of chemical and molecular features, leading to the identification of novel lead scaffolds [110].

3.1.5 Shape-based virtual screening

Shape-based virtual screening is an increasingly popular, purely ligand-based screening method [75]. The basic idea of shape-based screening is that a complementary image of the binding site can be generated, from the 3D-shape of the ligand [112]. One key advantage of shape-based virtual screening is that it makes the search for analogs of a single compound possible. Furthermore, even though it would seem necessary to know the bioactive conformation of a compound for a successful search for analogs it has been shown that the performance of the method is not influenced by this factor [112].

One of the most frequently used algorithms for shape-based search is ROCS (Rapid Overlay of Chemical Structures) [113]. The volume overlay of the compounds is assessed based on Gaussian spheres, which makes the fast calculation of the atom overlap mathematically efficient. The quantification of the atom overlap occurs through calculating the Tanimoto coefficient [114] between the Gaussian spheres. This encodes the similarity between the common features and those in at least one of the molecules. In order to include information on the chemical functionality of the molecules, a rudimentary pharmacophore scoring was later added to improve the similarity representation. This so-called "color-score" quantifies the similarity between the chemical characteristics of the query structure and the library compounds. Optimization of both shape overlay and chemical functionality substantially enhances the precision of the screening method [112].

3.1.6 Analyzing virtual screening performance

The general aim of virtual screening is the retrieval of a significantly higher amount of active compounds from a database than a random selection. There are several measures that can be used to assess the performance of a virtual screening protocol and will be introduced next. If a virtual screening protocol selects n molecules from a database with N entries, the hit list is composed of active compounds (true positive, *TP*) and inactive compounds (false positive, *FP*). The active molecules that are not retrieved are the false negative (*FN*), the inactives that are correctly not retrieved are the true negatives (*TN*) (Figure 11).
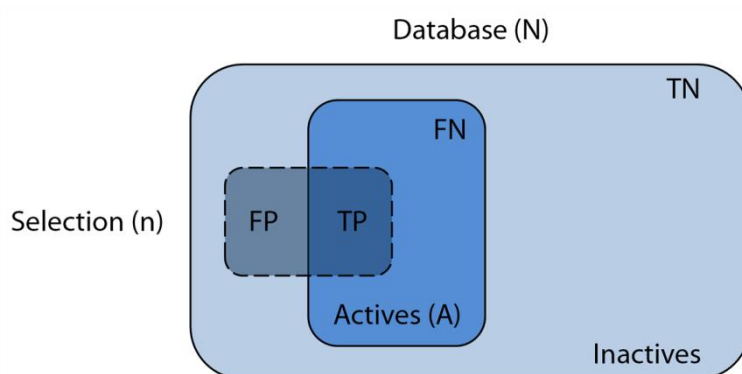
Figure 11: Selection of *n* molecules form a database with *N* entries [115].

Several metrics exist to assess the quality of the selection. One of them is the sensitivity (*Se*, true positive rate, Equation 3) describing the ratio of retrieved actives towards the total

number of actives in the database [111, 115].

$$Se = \frac{TP}{TP + FN}$$ **Equation 3**

Specificity (*Sp*, false positive rate, Equation 4) is the amount of true negative compounds divided by the total number of negative compounds in the database [111, 115].

$$Sp = \frac{TN}{TN + FP}$$ **Equation 4**

Another metric is the Yield of Actives (*Ya*, Equation 5) that is the measure for the amount of truly active compounds in comparison the number of selected molecules *n* [111, 115].

$$Ya = \frac{TP}{n}$$ **Equation 5**

The enrichment factor (*EF*) is a descriptor that quantifies the improvement of the hit rate by a virtual screening protocol compared to random selection [111, 115].

$$EF = \frac{TP/n}{A/N}$$ **Equation 6**

Receiver operating characteristic (ROC) curves are also used to analyze the performance of virtual screening methods [116]. The ROC curve describes *Se* as a function of (1-*Sp*) for any possible change of *n*. In the ideal case all actives are found before the inactives, which causes the plot to rise to an *Se* of 1 at 1-*Sp* of 0, afterwards only inactives can be found and the plot continues as an horizontal line until the upper right corner. In the case of random selection the curve trends towards the diagonal line. The distribution achieved by a virtual screening workflow with overlapping selection of actives and decoys, which is able to rank actives higher than decoys, is located between those two curves (Figure 12) [115, 116].

24
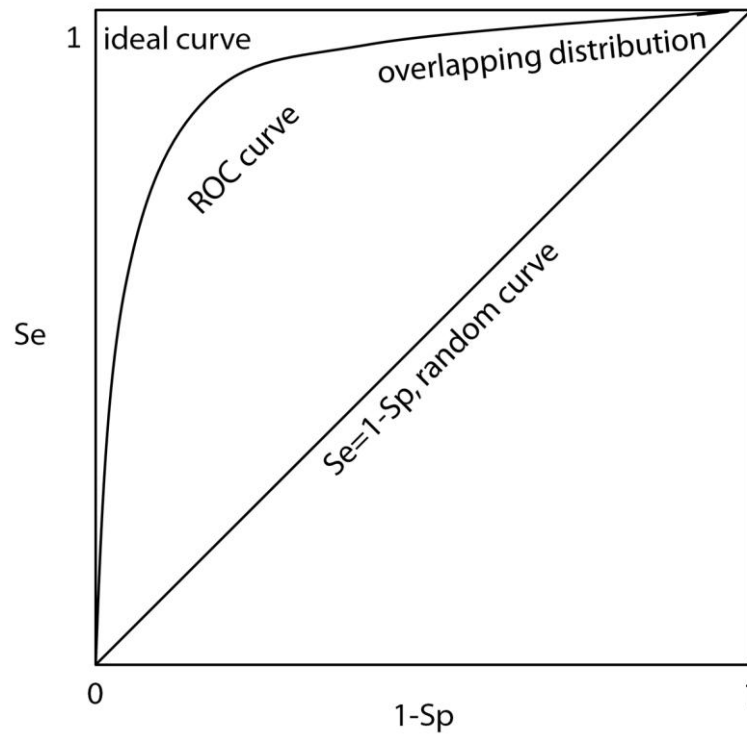
Figure 12: ROC curves for an ideal, random and an overlapping distribution of actives and decoys [115].

A scalar metric to evaluate ROC curves is to calculate the ROC functions integral, the *area under the curve* (*AUC*, Equation 7). It is calculated by measuring the sum of the area of all rectangles formed by *Se* and 1-*Sp* for a given threshold (Equation 6) [115].

$$AUC = \sum_i \left[(Se_{i+1})(Sp_{i+1} - Sp_i) \right]$$ **Equation 7**

4 Results

4.1 Virtual screening for TLR2 antagonists

Our screening campaign for small molecule antagonists of TLR2 was performed in a combined ligand- and structure-based approach. In the structure-based part of virtual screening, 3D pharmacophores of the TLR2 binding site were derived from the TLR2 crystal structure and used for screening. In the ligand-based part of the study, a two-stepped shape- and feature-based screening was performed with compounds that had been reported to modulate TLR2 activity. Virtual screening hits were selected for biological validation, which was performed by our cooperation partner Dr. Sandra-Santos Sierra at the Medicinal University of Innsbruck. The compounds were tested in a cell-based NF-κB reporter assay with subsequent transfection assays. In the following chapters the model development, the virtual screening with the experimental validation and finally the refinement of the model will be described and discussed for the structure- (chapter 4.1.1) and ligand-based screening (chapter 4.1.2).

4.1.1 Structure-based virtual screening

As no drug-like molecule antagonists of TLR2 were known when our study was initiated, we decided to perform a structure-based analysis starting with the crystal-structure of human TLR2 co-crystallized with TLR1 and the lipopeptide ligand Pam_3CSK_4 (PDB code: 2Z7X) [20]. For this purpose, we focused on the extracellular part of TLR2 even though a crystal structure of the intracellular domain was available [28]. This was done for several reasons: (i) the only known ligand binding site of TLR2 is in the extracellular domain, (ii) more information on the signaling mechanism of this domain is available and (iii) by targeting the extracellular domain of the receptor, limitations caused by cell permeability can be avoided. First, a thorough structural analysis of the receptor was performed in order to identify the most suitable pocket for small molecule binding. In the next step, we calculated MIFs of the identified binding pocket, in order to elucidate key interactions necessary for ligand binding. These were then used to define 3D pharmacophores, which were used to screen libraries of commercially available compounds. Next, a selection of virtual hits was chosen for biological validation. Finally, the generated model was refined integrating the obtained results.

4.1.1.1 Structural analysis and cavity detection

In 2007, the X-ray structure of the heterodimer formed by the extracellular domains of human TLR2 and TLR1 co-crystallized with the lipopeptide ligand Pam_3CSK_4 was published by Jin *et al.* [20]. Two years later, the structure of murine TLR2 dimerized with TLR6 and the diacylated lipopeptide Pam_2CSK_4 was published [21]. With the elucidation of these two structures, the basis for structure-based design of TLR2 ligands was provided.

By sequence homology, TLR2 belongs to the TLR1/TLR2/TLR6/TLR10 subfamily of TLRs, whose extracellular domains are characterized by three sub-domains: the N-terminal, central and C-terminal region [25]. The opening of the ligand binding sites of TLR2 and TLR1 are situated at the border of the C-terminal and central domain of the proteins [20]. This is a very unusual location for the ligand binding sites of TLRs, as these are usually located on the concave surface of the receptors [34, 35, 117, 118].

TLR2 and TLR1 as well as TLR2 and TLR6 form an m-shaped dimer in the ligand bound form (Figure 13). In the crystal structure of the TLR2/1 dimer, the two ester-bound lipid chains of the agonist Pam_3CSK_4 are located in the TLR2 binding site; the amide-bound lipid chain in the TLR1 cavity. The peptide head-group of the ligand reaches out of the protein and interacts with surface residues of both receptors through a network of hydrogen bonds [20]. In the TLR2/6 dimer, the orientation of the two ester-bound lipid chains and the peptide head-group are similar to their position in the TLR2/1 dimer. In contrast, the lipid binding cavity present in TLR1 is shortened in TLR6 by the bulky side chain of Phe343 that is not present in TLR1. This explains why TLR2/6 specifically recognizes diacylated lipopeptides like Pam_2CSK_4 as they lack the amide-bound chain [21]. The structures of the lipopeptide agonists and their binding modes are shown in Figure 14.
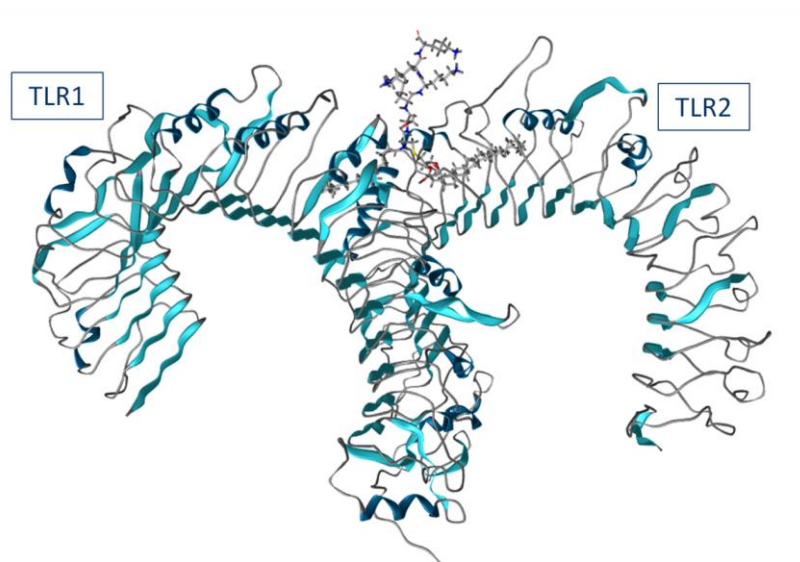
Figure 13: TLR2/1 heterodimer with co-crystallized ligand Pam$_3$CSK$_4$. TLR1 is depicted as backbone ribbon in the front, TLR2 in the back of the picture. The ligand is shown in stick representation.
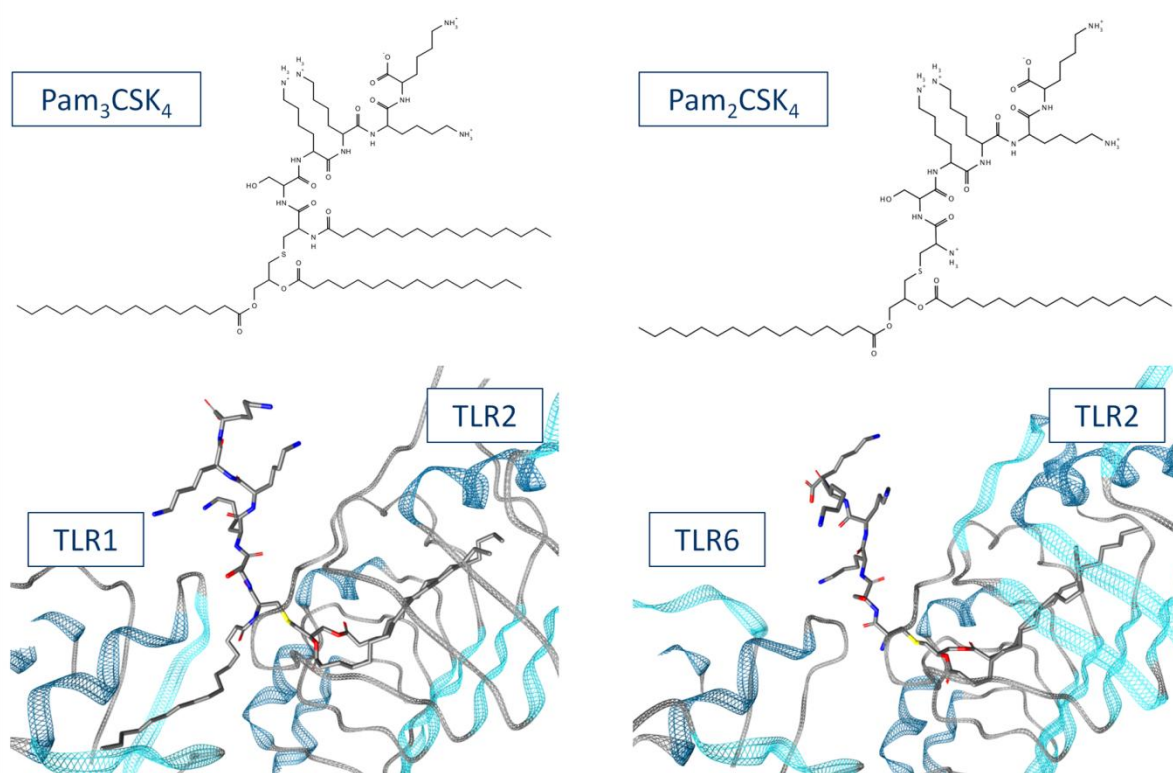
Figure 14: Lipopeptide ligands of TLR2 and its co-receptors TLR1 and TLR6. TLR2 recognizes tri- and diacylated lipopeptides in cooperation with TLR1 and TLR6, respectively. The chemical structures (top) and the binding modes (bottom) of the lipopeptides are shown. TLR2/1 and the specific agonist Pam$_3$CSK$_4$ are shown on the left, TLR2/6 and Pam$_2$CSK$_4$ on the right.

For the dimerization of TLR1 with TLR2, the binding of the triacylated ligand Pam_3CSK_4 is indispensable, as it bridges the two proteins through its three lipid chains. Besides these lipid chain interactions, the dimer is held together by hydrophobic, hydrogen-bonding and ionic interactions. The dimerization interface is characterized by a small hydrophobic core surrounded by hydrogen-bonding and ionic interactions [20]. In the case of TLR6, the stabilization of the dimer through the lipopeptide is missing; instead, stronger protein-protein interactions are formed between the monomers. The hydrophobic portion of the interaction surface area is increased by approximately 80 % compared to the TLR2/1 dimer and contains additional hydrophobic and ionic interactions [21].

Given the fact that our study aimed at finding small molecules that antagonize TLR2 response, we decided to focus our structure-based study on TLR2, because we hypothesized that compounds binding here would act on TLR1 as well as TLR6 specific signaling. As no binding site for small molecules acting as TLR2 antagonists was known, the first step for the rational design of such compounds was the identification and analysis of a potential binding site. The only binding area known for TLR2 is the one for the endogenous lipopeptide ligands, which are much larger than the small molecules we aspired to develop. The binding site search was performed using the crystal structure of human TLR2 (PDB code: 2Z7X [34]) co-crystallized with the lipopeptide (Pam_3CSK_4) and TLR1. TLR1 was removed and the remaining TLR2 monomer analyzed with two cavity detection programs: Q-SiteFinder an energy-based binding-site prediction tool [79], and Site Finder [85], a geometrical pocket detection algorithm based on α-spheres. Only pockets for which the results of both programs coincided were taken into account. The analysis was focused on areas of the protein that were part of the dimerization interface with TLR1 or part of the lipopeptide binding site. This was based on the assumption that small molecules could inhibit TLR2 signaling by impeding its dimerization, this being possible by binding to a part of the protein-protein interface or to the lipopeptide binding site. Figure 15A shows the TLR2 monomer with marked cavities. The cavity in the interface is small and very shallow and is only situated at the border of the interface, what makes it less suitable for ligand binding. Because of this, the further analysis focuses on the lipopeptide binding site.

After having discarded the other cavities, the endogenous ligand binding site was analyzed in more detail. Because the pocket is very big, its volume comprising approx. 1200 A^2 [35], it

29

was subdivided into smaller cavities more appropriated for small molecule binding (shown in Figure 15). The binding pocket is roughly trapezoidal and comprises four minor cavities, as shown in Figure 15B. The first sub-pocket (P1) is located close to the aperture of the lipopeptide binding site and is where Pam_3CSK_4 forms the only hydrogen bonds to the inside of the cavity. Pockets P2 and P3 are situated deeper in the binding site and are smaller than P1. Pocket P4 is at the opposite end of the cavity and is the biggest of the four cavities. As it is far away from the opening of the binding site, we rationalized that it would be poorly accessible for small molecule ligands. The cavities P1, P2 and P3 are closer to the aperture, making them more suitable for small molecule binding. Therefore, we focused our further analysis on this region of the lipopeptide binding site.
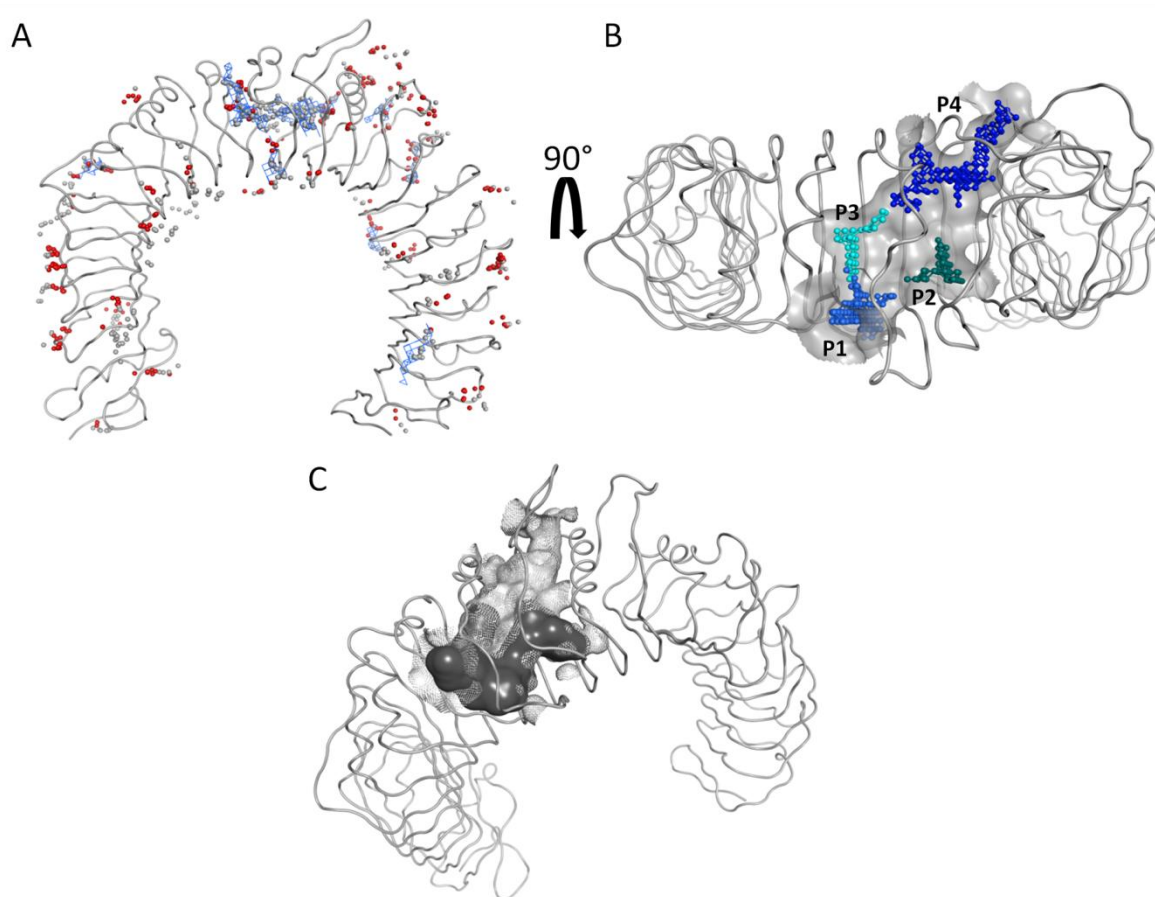
Figure 15: TLR2 monomer cavity detection. A) TLR2 with potential binding pockets calculated with Q-SiteFinder (blue lines) and Site Finder (red and grey spheres). B) Selected binding site with four sub-pockets P1-P4. C) Lipopeptide binding site volume, the front region is shown in dark grey, the back in light grey.

4.1.1.2 Deriving 3D pharmacophores for the putative binding site

In order to elucidate key interactions necessary for ligand binding to TLR2, we calculated molecular interaction fields (MIFs) [80] of the cavity using MOE 2010.10. [85]. These feature-based MIFs were subsequently used to determine favorable interaction hotspots for different chemical probes, which could then be transferred into a 3D pharmacophore. Three main probe types were used: "N1+" (cationic H-bond donor), "O-" (anionic H-bond acceptor) and "DRY" (hydrophobic probe). The resulting mapping of the binding site is shown in Figure 16. In the main part of the pocket hydrophobic contacts are favored (DRY probe shown in yellow), which is caused by the hydrophobic side chains of leucine rich repeats 9~12 that point into the pocket. The side chains of hydrophilic residues point towards the outside of the pocket, which explains why H-bonds can only be formed through backbone interactions. The MIF analysis shows that H-bond donors are mostly energetically favored to interact with the carbonyl-oxygens of Ser346 and Leu350. The amide nitrogens of Leu350 and Phe349 are suitable for H-bond formation with H-bond acceptors, which is the case with the carbonyl oxygen of Pam$_3$CSK$_4$ in the crystal structure.
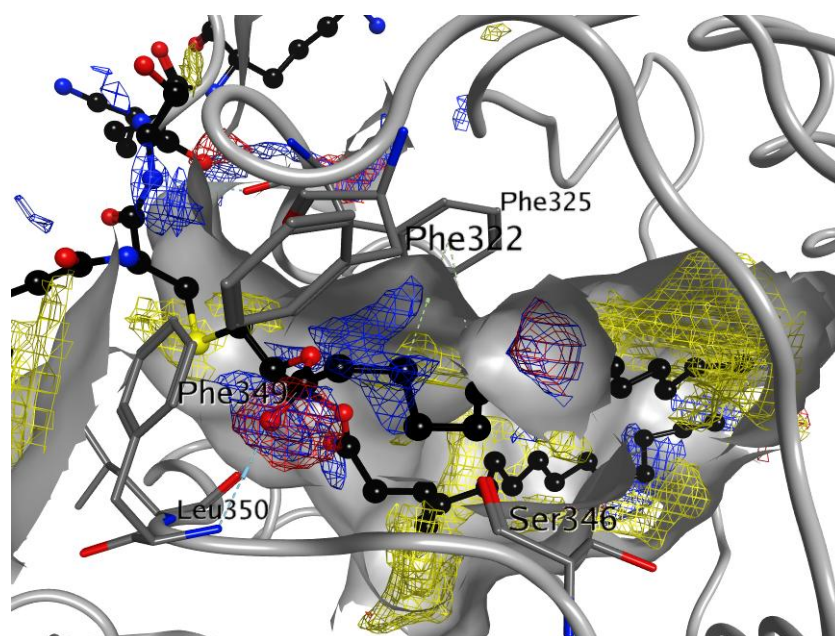
Figure 16: TLR2 ligand binding site with calculated molecular interaction fields. The binding pocket and the protein residues are depicted in grey. The molecular interaction fields are shown as grids. Zones that favor hydrophobic interactions are drawn in yellow; interaction hotspots for hydrogen bond acceptors in red and those for hydrogen donors in blue.

Next, using LigandScout 3.1. [108, 109, 119], two 3D pharmacophore models describing the previously identified key interactions for ligand binding and the geometric shape of the cavity were defined. The first (mod_1, Figure 17 left), comprised four features: An H-bond donor (HBD) with the carbonyl oxygen of Leu350, an H-bond acceptor (HBA) with the amine nitrogen of Phe349 and two hydrophobic areas. The first (HYD1) was located close to Phe325, Ile319 and Val348 and the second (HYD2) was situated deeper inside the pocket, in the area of Ile314, Phe284, Phe295, Leu266 and Leu328. For the second model (mod_2, Figure 17 right) two similar hydrophobic areas were defined. In addition, the 3D pharmacophore model contained two H-bond donor interactions one with the carbonyl oxygens of Ser346 and another optional donor with Leu350. An H-bond acceptor feature with the amide nitrogen of Phe349 was also included. In order to discard compounds that would not fit the binding pocket sterically, exclusion volumes spheres were defined on the protein side in both models to better describe the steric constraints formed by the pocket.
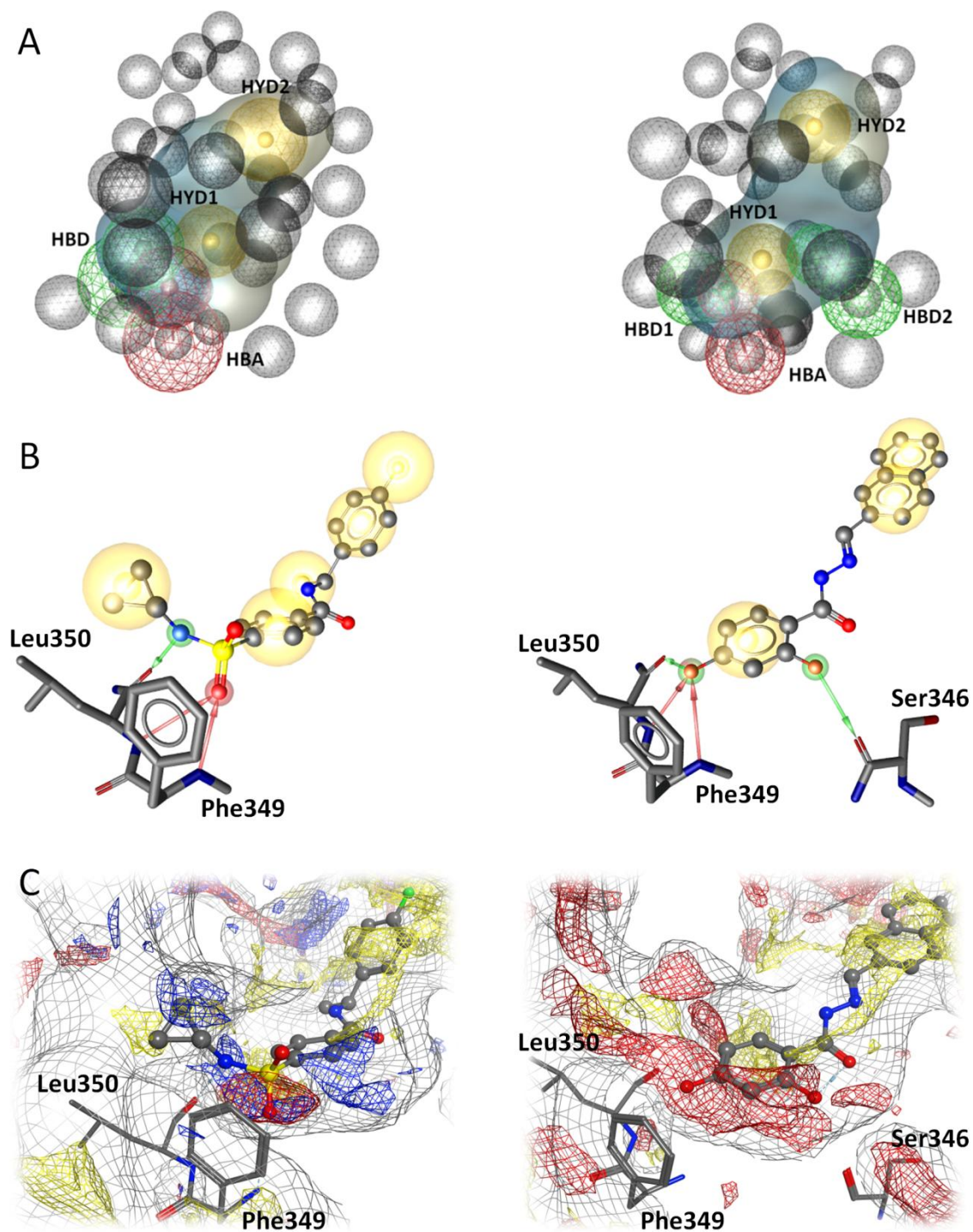
Figure 17: 3D pharmacophore models developed for virtual screening. The following is shown for mod_1 (left) and mod_2 (right): A) 3D pharmacophore with exclusion volume spheres. B) Virtual screening hit with highlighted interactions to the TLR2 binding site. C) Virtual screening hit in the TLR2 binding site with calculated MIFs (red – H-bond acceptor, blue – H-bond donor, yellow – hydrophobic)

4.1.1.3 Virtual screening protocol

In the next step the previously defined models were used for the virtual screening of a library of 2,831,238 commercially available compounds from different vendors. Virtual screening was performed using LigandScout 3.1. and led to 7426 hits for mod_1 and 4435 for mod_2. Subsequently, the screening hits were ranked by geometric pharmacophore fit. The most promising structures (150 compounds with the highest score) were docked into the TLR2 binding site using GOLD Suite, v5.1. [120] and GoldScore [94]. Afterwards, the resulting binding poses were minimized in the binding site using the LigandScout implementation of the most recent version of the Merck Molecular Force Field (MMFF94) [121-127]. Next, the binding poses were visually inspected putting special emphasis in the fulfillment of the previously defined interaction patterns based on the MIFs. Through this procedure five virtual hits per pharmacophore were selected for experimental validation.

Biological testing was performed in a cell-based NF-κB reporter assay in cooperation with Dr. Sandra Santos-Sierra at the Medicinal University of Innsbruck. The cell line HEK293-TLR2 was stimulated with the TLR2/1 and TLR2/6 agonists Pam_3CSK_4 and Pam_2CSK_4, respectively and the relative inhibition with and without compound was measured. Compounds that decreased NF-κB activity by more than 1/3 at a concentration of 200 µM were selected for further investigation. In addition to the NF-κB reporter plasmid a Renilla control plasmid was transfected into the cells to monitor potential toxicity of the compounds. A drawback of the employed assay is that the activity of the compounds is not directly assessed at the level of TLR2 but in form of the NF-κB activation. As described in chapter 1.1.3, TLR2 activation leads to the activation of the aforementioned transcription factor. Nevertheless, this takes place through a signaling cascade involving several proteins like adaptor molecules and kinases. By measuring NF-κB activity compounds acting on other parts of the signaling cascade cannot be distinguished from compounds acting directly on the receptor. To overcome this ambiguity, further transfection and inhibition assays with the downstream proteins were performed. When the cells are transfected with a downstream protein, the NF-κB activity increases. If, after transfection, the inhibitor is added and a decreased activity can be measured, this implies that the compound is acting on the transfected protein or further downstream. Through sequential transfection and inhibition assays, compound targets can be identified and true TLR2 antagonists distinguished from compounds with inhibitory activity on the signaling pathway. What cannot be determined through this method is whether

compounds possibly act on the receptor as well as on other downstream proteins.

Two out of five molecules selected with mod_1 and four out of five of those selected with mod_2 showed inhibitory activity against TLR2 signaling. Nevertheless, through the transfection assays, some of these compounds were shown to act downstream of the receptor protein, leaving one compound selected by mod_1 (compound **75**) with direct antagonistic activity on TLR2. This corresponds to a hit rate of 20 % for mod_1 and of 10 % in general for the structure-based screening. The results are summarized in Table 1.

Table 1: Results of the biological validation of the 3D pharmacophore-based screening

Model	Selected compounds	TLR2 signaling inhibitors	TLR2 antagonists	Hit Rate
mod_1	5	2	1	20 %
mod_2	5	4	0	0 %
Total	10	6	1	10 %

4.1.1.4 Novel compounds identified by structure-based virtual screening

Compound **75**, the identified TLR2 antagonist is shown in Figure 18. Remarkably, the sulfamoylbenzamide scaffold has not been reported as TLR2 ligand previously. The chloro-substitution seems to be essential for activity as other tested compounds that lack the chlorine substituent do not shown activity in the performed assay (compound **74**, Appendix 1). The fluorine also seems to be important for activity. Compound **73** that differs from **75** by a methyl- instead of a fluorine substituent and an ethyl-instead of a methyl group and was experimentally discarded as a TLR2 antagonist, as it was showing activity downstream of the receptor. However, TLR2 activity could still occur with this compound. A potential structural explanation of this fact is given in the next chapter.
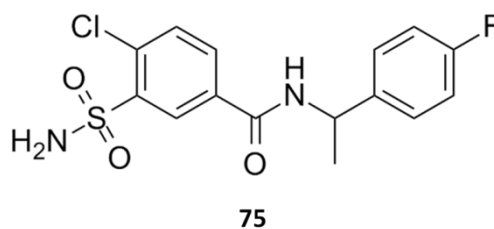

75

Figure 18: TLR2 antagonist identified by structure-based virtual screening

4.1.1.5 Model refinement and evaluation

In a post validation step, the information gained through the biological testing was integrated into the model in order to describe the requirements for binding in more detail. For this purpose, the identified TLR2 antagonist **75** and the inactive virtual hits were re-docked into the TLR2 binding site and the interaction pattern necessary for binding analyzed (Figure 19). Compound **75** had been selected through virtual screening with mod_1. This model contained an H-bond donor interaction with the carbonyl oxygen of the Leu350 backbone and an H-bond acceptor interaction with the amide nitrogen of the Phe349 backbone. Both features could be confirmed through the binding pose of **75**. The two hydrophobic contact regions HYD1 and HYD2 were validated as well. In addition to those, the chlorine substituent of the sulfamoylbenzamide moiety forms hydrophobic contacts with Phe349, Val348, Phe322, Phe325 and Ile319 in the front part of the binding pocket. Compounds with the same scaffold but lacking this substituent are not active, which shows that this interaction is essential for TLR2 antagonism. Likewise, the hydrophobic contacts formed by the fluorine substituent of the benzyl ring and the deepening formed by Ile314, Phe295, Leu289, Phe284 and Leu266 seems to be necessary for activity as well. One compound with a methyl-group instead of the fluorine (**73**, Appendix 1) was tested as well and found to be inactive.

Virtual screening with the second 3D pharmacophore mod_2 did not result in any biologically active TLR2 antagonists. To investigate possible reasons for this result, the tested compounds were re-docked as well. All compounds fulfilled the defined interactions with Ser346, Leu350 and Phe349 and form the hydrophobic contact areas. Nevertheless, the interactions with the cavity deepenings described for **75** did not take place. This compound geometrically fits the binding site in a much better way than the virtual hits obtained with mod_2, which could be a reason for lacking activity on the receptor. Alternatively, the interactions of the sulfonamide with Leu350 and Phe349 could be more favorable for ligand binding than the H-bonds formed by the hydroxyl-groups and Leu350 and Ser346.
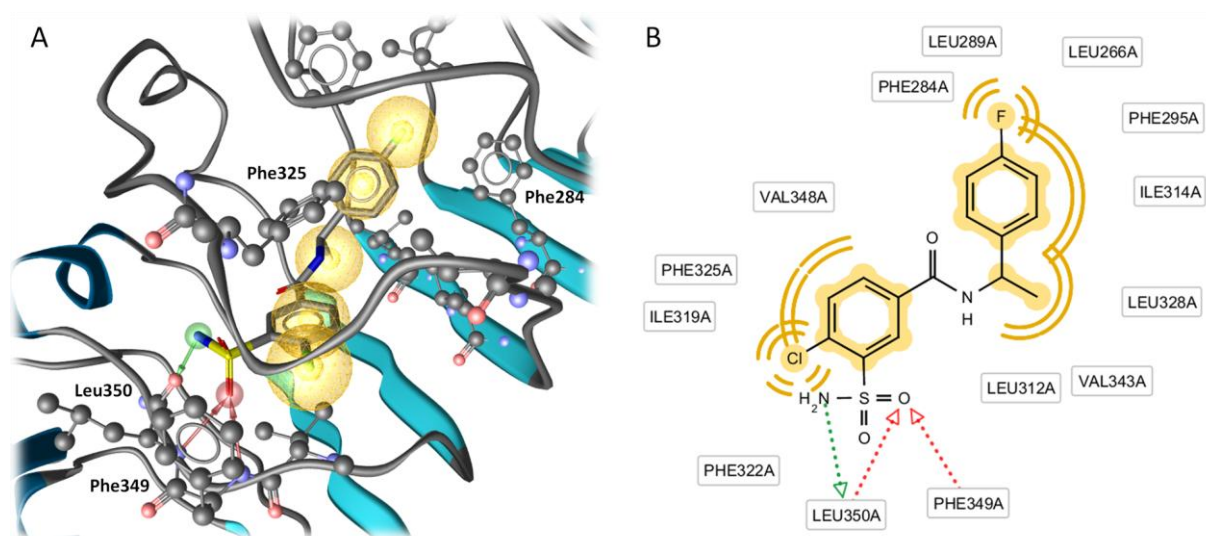
Figure 19: Predicted binding mode of TLR2 antagonist 75. A) 3D graphic of the predicted binding mode. Pharmacophoric features are represented in the following color code: hydrophobic contacts – yellow spheres, H-bond donors – green arrows, H-bond acceptors – red arrows. B) 2D representation of the interactions formed between the compound and the TLR2 binding pocket.

4.1.2 Ligand-based virtual screening

In parallel to the screening with 3D pharmacophores, a ligand-based virtual screening was performed in order to identify TLR2 antagonists. For this purpose, the first step was a thorough literature research for drug-like compounds known to bind or modulate TLR2. In this case also antagonists were taken into account. Then, a two-stepped shape- and feature-based search was carried out. This screening method was selected because it allows identifying analogues of single compounds without further knowledge on their binding mode. In the first screening the compound collection provided by the National Cancer Institute (NCI) [128] was screened with known TLR2 modulators extracted from the literature. In the second screening cycle the search with known ligands was extended with new query structures (two hit molecules from the first screening and one new TLR2 signaling inhibitor) and by screening a bigger library of commercially available compounds. Furthermore, a 3D database search using experimentally validated hits from the first screening cycle was performed. In the next chapters the query structures, the virtual screening and its biological results will be presented and discussed.

4.1.2.1   Virtual screening protocol I: Screening of the NCI library

The first screening was performed with three query structures extracted from literature shown in Figure 20. The first two structures (Compound A and Compound B) had been identified by high-throughput screening by Guan *et al*. [58] and were described to have agonistic activity on TLR2 (chapter 1.1.5.1). These molecules were used to search for antagonist for TLR2, as the only lipids or carbohydrates with TLR2 inhibitory activity were known and are not suitable as query structures for drug-like small molecules. By analogizing molecules that were known to bind TLR2, we expected to find other binders that eventually would have the opposite activity due to small changes in their chemical structure. The third molecule used for the screening was E567, a compound that had also been identified through high-throughput screening and described as an inhibitor of TLR2 signaling with unknown target protein (chapter 1.1.5.2) [59]. Because no other information on small molecule TLR2 binders was available when we started the study, this approach appeared to be the most promising starting point to identify TLR2 antagonists. The program ROCS (Rapid Overlay of Chemical Structures, described in chapter 3.1.5) [129] was used for this screening step. The software compares the molecular volume based on Gaussian spheres and a rudimentary pharmacophore-like similarity scheme. The inclusion of chemical information in the overlay makes it more precise than purely shape-based screening. Another advantage of the selected software is that the knowledge of the bioactive conformation does not necessarily influence the achieved enrichment of the screening [112].
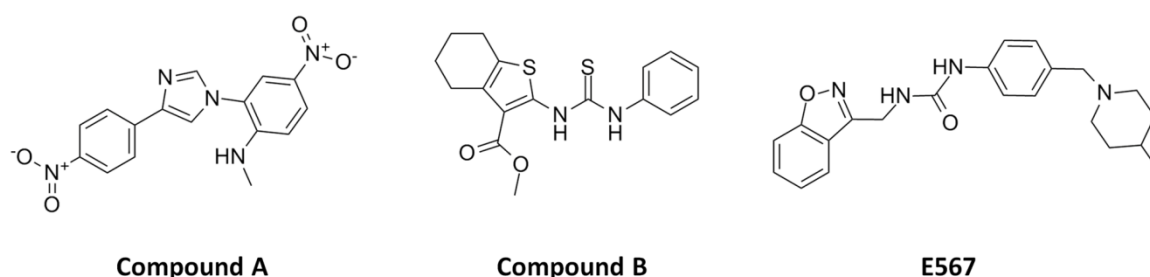
Figure 20: Query structures used for the first shape- and feature-based screening.

As mentioned above, in the first shape- and feature-based screening the compound collection of the NCI was screened [128]. All 260.071 structures contained in the database were transferred into a multi-conformational structure database using the software tool OMEGA [130]. Standard 3D structures of the query molecules were generated using CORINA [131]

38

and used for shape and feature based similarity search with ROCS [129]. The 500 best hits according to the Tanimoto Combo Score (described in chapter 3.1.5) were saved for every query structure. Careful visual inspection regarding similarity, drug-likeness and availability resulted in a list of 39 compounds which were biologically tested. 6, 29 and 4 analogs were selected for compound A, B and E567, respectively. The main reason for the fact that most of the tested compounds were analogs of B was the commercial availability.

All selected virtual hits were tested in the NF-κB reporter assay described above. Compounds acting on targets downstream of TLR2 were detected by transfection assays and discarded, as described for the structure-based screening. Through this procedure eight compounds with antagonistic activity on TLR2 could be identified. In a next step, the purity of the hit compounds was tested by a previously developed HPLC procedure. Unfortunately, three compounds had to be discarded because they did not meet the necessary purity criteria. Finally, five novel small molecule antagonists of TLR2 could be confirmed. Three compounds were analogs of compound B, whereas the two remaining hits were derived from E567. Overall, five out of 39 tested compounds could be confirmed as TLR2 antagonists, corresponding to a hit rate of 13 %. An overview on the results is given in Table 2.

Table 2: Overview on the results of the first shape- and feature-based screening.

Query structure	Selected compounds	TLR2 signaling inhibitors	TLR2 antagonists	Hit rate
Compound A	4	2	0	0 %
Compound B	29	6	3	10 %
E567	6	2	2	33 %
Total	39	10	5	13 %

4.1.2.2 Novel compounds identified by ligand-based virtual screening I

The compounds identified by the first ligand-based virtual screening are shown in Figure 21. Compounds **5**, **22** and **27** share a phenyl-urea scaffold and were identified with compound B as a query structure. Their scaffold had not been reported at the time of their discovery. However, in their high-throughput screening study Cheng *et al.* [60] identified one compound that shares this scaffold. This will be further discussed in chapter 4.5 dealing with the integration of the data of this study into the TLR2 antagonism model. Other tested compounds

with a phenylthiourea scaffold (**2** and **4**) or with aliphatic ring-systems (**12**) are inactive. Three further compounds that shared the phenyl-3-(2,4-dimethoxyphenyl)urea scaffold of **27** with a different substitution pattern on the halogenated ring (**18** – 3-chloro, **19** – 2-chloro and **21** 2-carboxyl-substituted) were tested as well and were all found to be inactive. The other two discovered antagonists **40** (*N*-(3-(1*H*-indol-1-yl)propyl)-4-methoxyaniline) and **44** (*N*-((4-(5-bromopyrimidin-2-yl)oxy)-3-methylpehnyl)carbamoyl)-2-(dimethylamino)benz-amide) were analogs of E567.
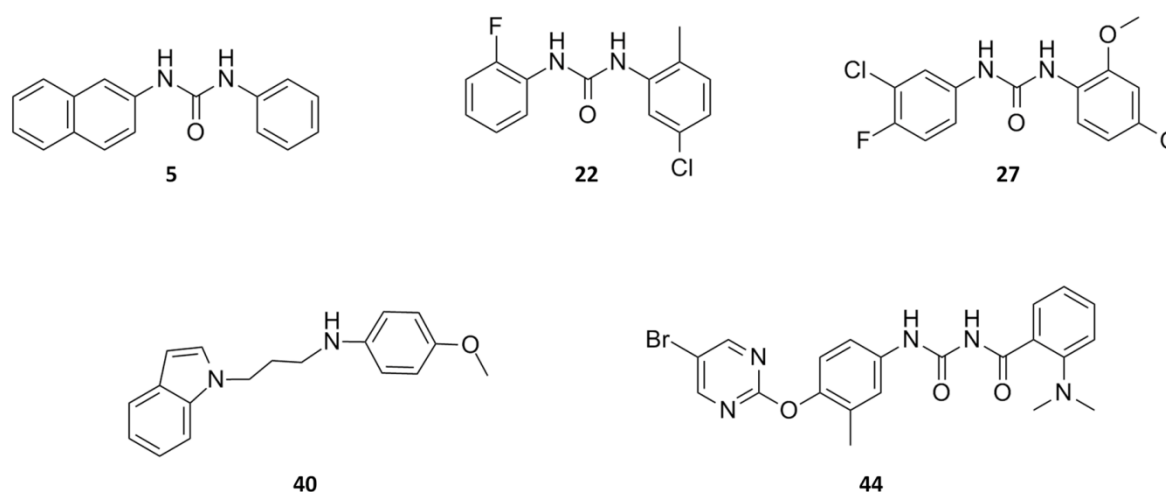
Figure 21: Chemical structures of the validated TLR2 antagonists.

4.1.2.3 Virtual screening protocol II: Virtual screening of a commercial compound library

In the second part of the ligand-based feature- and shape-based screening the search was extended by including new query molecules, as well as by screening a bigger collection of commercially available compounds. The screened database comprised 2,831,238 structures from different vendors. The NCI compound library was excluded of this part of the screening because of the difficulties in terms of purity and availability encountered in the first screening. Sparstolonin B (SsnB, Figure 22) was included as a query structure as it had been published in the meantime. This compound had been isolated from the Chinese herb (*Spaganium stoloniferum*), and then showed to have inhibitory activity on TLR2 and TLR4 signaling [72]. The screening with compounds A and B was repeated against the commercial library. In contrast, E567 was excluded as a query structure as experimental results of the transfection assays showed that it could potentially act on MyD88 and not on TLR2. In addition to the

TLR2 modulators described in literature, two of the validated TLR2 antagonists (**5** and **44**) from the first screening were included as query structures, too. These compounds were selected because of their structural divergence from each other and because they seemed to be the most promising compounds in terms of activity considering the preliminary biological results available at that point of the study.

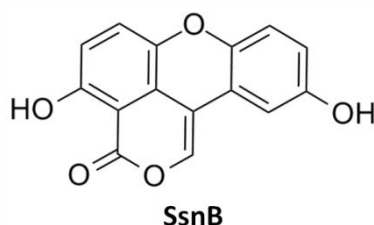**Figure 22: Chemical structure of the TLR2 signaling inhibitor Sparstolonin B (SsnB).**

The screening was performed using ROCS [129], as described before for the first search (chapter 4.1.2.1). For the query performed with the previously in-house identified TLR2 antagonists an additional docking step was included into the virtual screening protocol. The hits from the shape- and feature-based screening were docked into the TLR2 binding site in order to assess their binding mode. Docking was performed using GOLD Suite v5.1. [120] and GoldScore [94]. The poses were inspected, paying special attention to the presence of the key interactions defined in the structure-based screening (chapter 4.1.1.2). Through this procedure, four analogs of compound B, six of compound A, four of SsnB and six of each **1** and **5** were selected for experimental validation.

The 26 selected candidates were tested biologically in the aforementioned NF-κB reporter assay with subsequent transfection experiments. The purity of the active compounds was checked by HPLC. The results of the biological validation are summarized in Table 3. In total six TLR2 signaling inhibitors were identified. Of these, four could be confirmed to act directly on the receptor through the transfection assays. Two of the confirmed TLR2 antagonists were analogs of compound A, one of compound **5** and the last one of **44**. Neither the search with compound B nor with SsnB retrieved any TLR2 antagonists. In contrast, one of the analogs of compound B (**61**) and three (**47 - 49**) of the molecules derived from SsnB had agonistic activity on TLR2 (Figure 24). This is not unexpected in the case of the compound B analog as the query structure was actually described to be a TLR2 agonist [58].

On the contrary, SsnB was described to be a TLR2 signaling inhibitor [72]. This clearly shows how small changes in the chemical features of a compound can lead to reversed activity. The mode of action and potential binding mode of the compounds with agonistic activity will be described in more detail in chapter 4.3.3.

Overall, a hit rate of 15 % could be achieved in the second shape- and feature-based screening. This is an improvement compared to the first screening in which 13 % of the tested compounds showed activity. The screening with hit molecules from the first screening as query structures (**5** and **44**) resulted in one active out of six tested compounds (hit rate 17 %) in both cases. In contrast, the search with known modulators of TLR2 only led to two antagonists out of six tested molecules, in the case of compound A, and no hits for the other two query structures (hit rate 11 %). Interestingly, four agonists of TLR2 could be identified with these two structures.

Table 3: Overview on the results of the second shape- and feature-based screening.

Query structure	Selected compounds	TLR2 signaling inhibitors	TLR2 agonists	TLR2 antagonists	Hit rate
Compound A	6	3	0	2	33 %
Compound B	4	0	1	0	0 %
SsnB	4	0	3	0	0 %
5	6	1	0	1	17 %
44	6	2	0	1	17 %
Total	26	6	4	4	15 %

4.1.2.4 Novel compounds identified by ligand-based virtual screening II

The chemical structures of the TLR2 antagonists identified in the second ligand-based virtual screening are depicted in Figure 23. Compound **77** was derived from the previous hit **5** and was found to inhibit TLR2 activity. This shows that the naphthalene can be replaced by an ethyloxoindolinyl and the phenyl by a thiophen, if the urea core is conserved. Analogs of **5** with di-substituted the phenyl-ring (**57** and **69,** Appendix 1) were found to be inactive. Possible structural reasons for this will be explained in more detail in chapter 4.3.1. **66** (1-(2-(dimethylamino)-2-(thiophen-2-yl)ethyl)-3-(4-(pyrazin-2-yloxy)phenyl)urea)   was   derived

from the previous hit **44**. The two remaining compounds are analogs of compound A, the agonists extracted from literature. The three-ring system is conserved in both structures as well as the nitrophenyl-moiety. A more detailed discussion on potential activity cliffs between agonists and antagonists of this class will be given in chapter 4.1.2.5.

Interestingly, in the second shape- and feature-based screening not only antagonists of TLR2 but also agonists could be identified. The chemical structures of these compounds are shown in Figure 24. Compound **61** was derived from compound B, the other three compounds (**47 - 49**) from SsnB. Main structural features of the compounds in comparison to their query structures will be discussed in the next chapter.
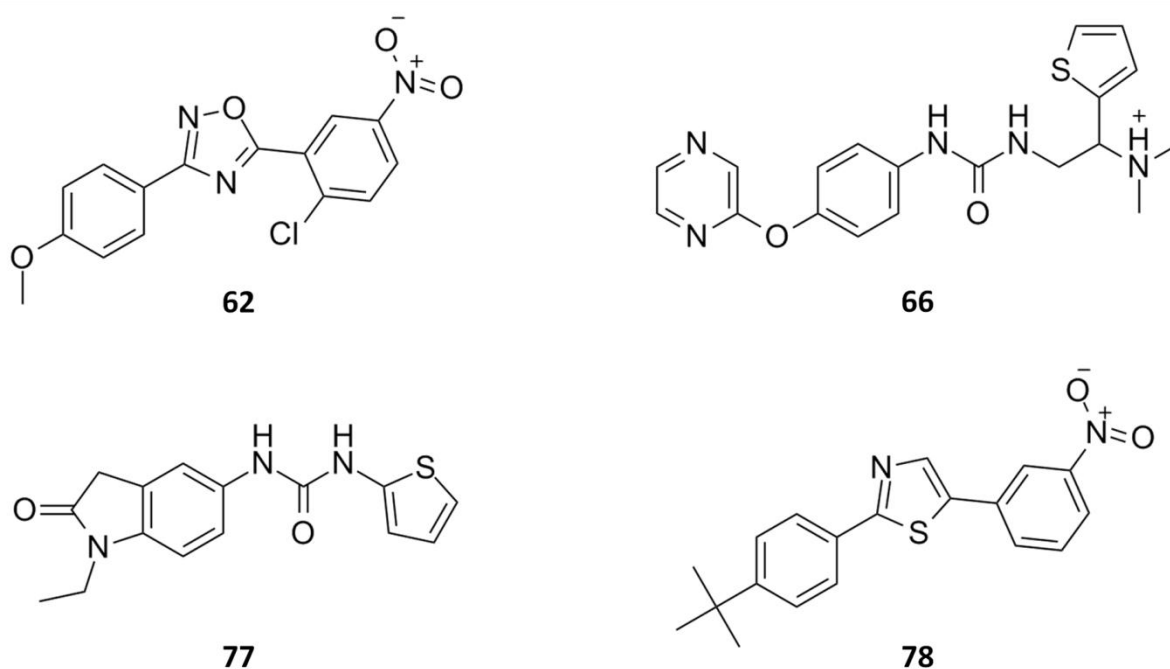

Figure 23: Chemical structures of the TLR2 antagonists identified in the second shape- and feature based screening.
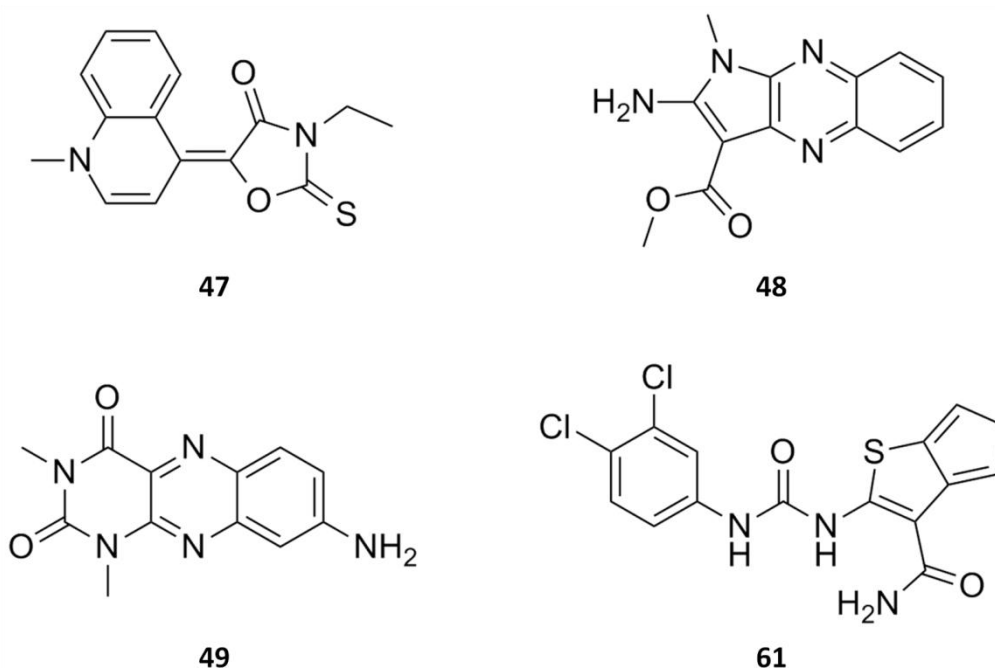
Figure 24: TLR2 agonists discovered in the virtual screening study for TLR2 antagonists.

4.1.2.5 Agonist query structures as templates for antagonist identification

As described before, in the last chapters our shape- and feature-based screening with TLR2 modulators led to the discovery of agonists as well of antagonists. The activity of the hit compounds sometimes was contrary to the effect described for the query structure. This motivated us to analyze the chemical features of structurally related query structures and hit molecules in order to elucidate which of them are crucial for agonistic or antagonistic activity.

In terms of pharmacophoric features compound A (N-methyl-4-nitro-2-(4-(4(nitrophenyl)-1H-imidazol-1-yl)aniline), the first agonist used as a query structure, is characterized by three aromatic rings, four strong H-bond acceptors represented by the oxygens of the nitro-groups, one H-bond acceptor in the imidazol ring and one H-bond donor that is caused by the amine nitrogen of the aniline (Figure 25A). The shape- and feature-based search performed with this molecule led to two hits that showed antagonistic activity on TLR2 (compounds **62** and **78**, Figure 25B and C) and no compounds with agonistic activity. Both antagonists also contain the three aromatic rings as well as the H-bond acceptor in the five-membered ring. In contrast, one of the nitro-groups as well as the amine group of the aniline are not present in those molecules. This causes the loss of two H-bond acceptors besides the H-bond donor and the addition of an hydrophobic contact area on the phenyl-ring that is not substituted with the

nitro-group. This could be the reason for the changed activity. The ability to form hydrophobic contacts with this phenyl-ring seems to be especially important for antagonistic activity, as observed with **10** that has a methoxy-substituent and is much less active than **9** that is substituted with a trimethyl-group.
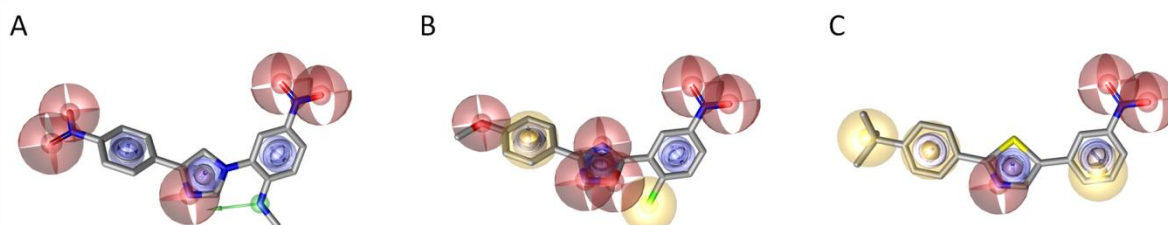

Figure 25: Pharmacophoric features of Compound A and its analogs with antagonistic activity. A) Compound A with pharmacophoric features, B) Antagonist **62** with pharmacophoric features, C) Compound **78** with pharmacophoric features. Chemical features are color coded: H-bond acceptor – red sphere, H-bond donor – green arrow, aromatic ring – blue circle, hydrophobic contact – yellow spheres.

Interestingly, the search with compound B (methyl 2-(3-phenylthioureido)-4,5,6,7-tetrahydrobenzo[*b*]thiophene-3-carboxylate) as a query structure led to one molecule with agonistic and three molecules with antagonistic activity (Figure 26). The query structure contains a phenylthioureido-scaffold which is replaced by phenylurea in the screening hits. In terms of pharmacophoric features the two H-bond donors formed by the two nitrogens and the H-bond acceptor represented by the sulfur in compound A as well as the carbonyl-oxygen in the other compounds, seem to be necessary for both agonistic and antagonistic activity. The two ring systems are also necessary for activity. In the case of the antagonists, only compounds with highly hydrophobic six-membered rings showed activity. The aromaticity is also crucial, as compounds with no or aliphatic rings were found to be inactive (**13**, **31**). For the agonists, the substitution of the thiophene with a strong H-bond acceptor (the carboxylate- and carboxamide-oxygen in compound B and **14**, respectively) seems to be necessary for the receptor activation.
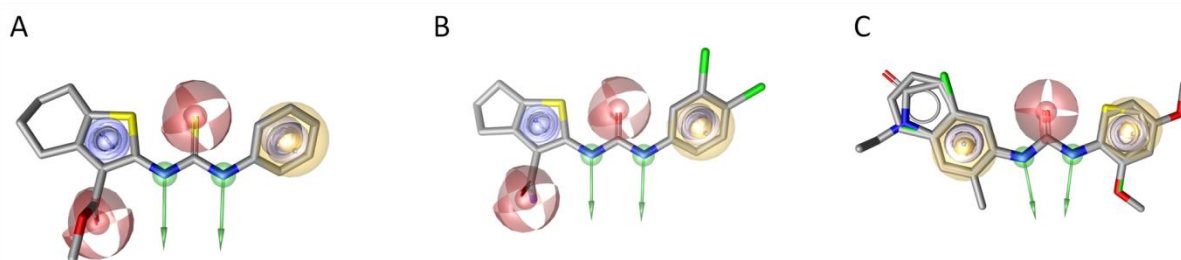
Figure 26: Pharmacophoric features of TLR2 agonists and antagonists derived from compound B. A) Compound B with pharmacophoric features B) Agonist **61** identified by similarity search with compound B C) Antagonists identified with Compound B as a query structure and shared-feature pharmacophore. Chemical features are color coded: H-bond acceptor – red sphere, H-bond donor – green arrow, aromatic ring – blue circle, hydrophobic contact – yellow spheres.

As mentioned before, in the second shape- and feature-based screening we included SsnB an inhibitor of TLR2 signaling as query structure. Interestingly, neither TLR2 antagonists nor signaling inhibitors could be found with this structure. In contrast, three molecules with agonistic activity were identified. In Figure 27A, SsnB is shown with highlighted pharmacophoric features. The main features of the 4,10-dihydroxybenzo[*kl*]xanthen-3(*2H*)-one are the three aromatic rings, the H-bond acceptor caused by the carbonyl-oxygen, the two hydroxyl-substituents on the xanthene scaffold and the hydrophobic contact areas on the same. Speaking in terms of pharmacophoric features all discovered TLR2 agonists lack the two hydroxyl-groups as well as the four-ring structure. That could possibly be the reason for their diverging activity. Nevertheless, it is important to keep in mind that SsnB was published to be an inhibitor of TLR2 and TLR4 signaling what means that its target protein could also be another protein apart from TLR2. This makes the reliable determination of the features necessary for a certain activity difficult to determine with the available information.
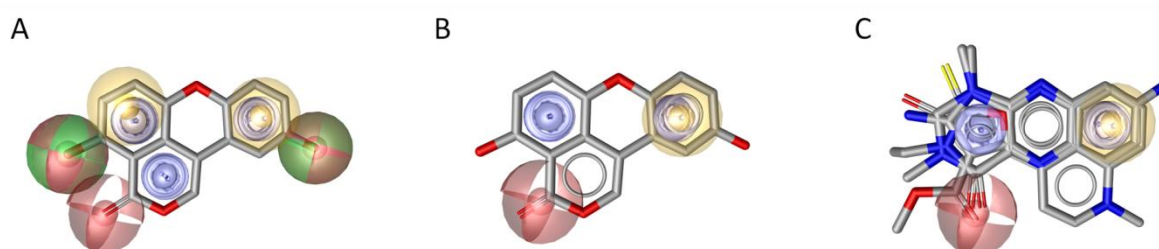

Figure 27: Comparison of the pharmacophoric features of TLR2 signaling inhibitor SsnB and its analogs with agonistic activity. A) SsnB with marked pharmacophoric features. B) SsnB with chemical features that are conserved in the agonists. C) Agonists with shared feature pharmacophore.

4.2 Biological characterization of antagonists

In the next step, the biological activity of the TLR2 antagonists discovered through our virtual screening was characterized in more detail by our cooperation partner Sandra Santos-Sierra of the Medicinal University of Innsbruck (Austria). The conception of the experiments was partially performed in mutual effort. In the following sections the results of the performed assays will be presented. First, the relative inhibition of NF-κB signaling was assessed (chapter 4.2.1), then the selectivity of the compounds towards TLR3 and TLR4 was measured (chapter 4.2.2) and finally, the impact of the antagonists on TNF-α production was determined (chapter 4.2.3).

4.2.1 Relative inhibition in NF-κB luciferase reporter assays

As during the virtual screening the NF-κB reporter assays had only been carried through at high concentration to identify bioactive TLR2 antagonists it was repeated in a concentration dependent manner. The toxicity of the compounds was also assessed in the same assay; none of the compounds showed significant toxic effects.

The resulting relative inhibitions for TLR2/1 and TLR2/6 signaling at 15 µM are depicted in Figure 28. Compound **44**, one of the molecules selected as query structures for the second screening, showed the highest relative inhibitions of TLR2/1 50.26 % and TLR2/6 83.82 %. In contrast, compound **5** that showed the best results in the preliminary screening (at a concentration of 200 µM antagonist) and was therefore selected as a query structure for the second screening only showed low relative inhibitions of 16.46 % for TLR2/1 and 11.40 % for TLR2/6.
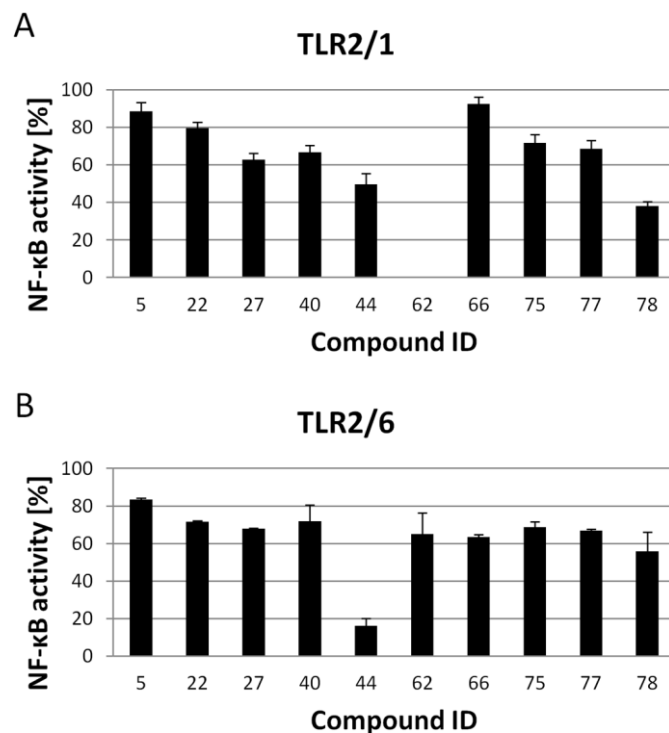
Figure 28: Relative inhibition of TLR2/1 and TLR2/6 measured in an NF-κB luciferase reporter gene assay. A) TLR2/1 signaling was activated by Pam_3CSK_4 (200 ng/mL) stimulation and B) TLR2/6 signaling by Pam_2CSK_4 (50 ng/mL). Values represent the mean ± standard deviation.

4.2.2 Compound selectivity towards TLR3 and TLR4

TLRs are a family of highly homologous proteins. Thus, the selectivity of the TLR2 antagonists against other TLRs (TLR3 and TLR4) was also assessed. For this purpose, a NF-κB reporter gene assay in RAW264.7 macrophages was performed. The cells were stimulated with PolyI:C, the specific TLR3 agonist, and LPS, the specific TLR4 agonist, and the relative inhibition of the antagonists measured at a concentration of 50 µM. The compounds did neither show significant antagonistic activity towards TLR3 nor TLR4. This is a further proof of the fact that the compounds act on the receptor itself and not on proteins further downstream of the signaling cascade, as these are shared by TLR2 and the other tested TLRs. Furthermore, specificity towards other proteins is a good indicator for the exclusion of other unspecific inhibition mechanisms. The results of the selectivity assays are shown in Figure 29.

Figure 29: Activity of the identified TLR2 antagonists towards TLR3 and TLR4. NF-κB reporter gene assays in RAW264.7 cells were performed. The TLR3 activation was r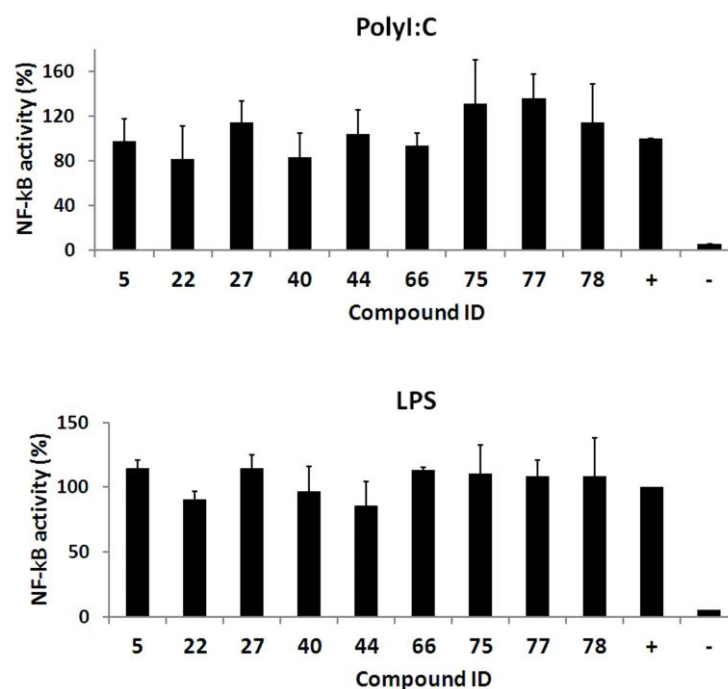ealized with 10 µg/mL PolyI:C, the TLR4 induction with 20 ng/mL LPS. No significant inhibitory effect on the tested receptors could be observed. Values represent the mean ± standard deviation.

4.2.3 Determination of the inhibitory activity on TNF-α production

The final step of the biological characterization of the TLR2 antagonists was to determine whether the observed inhibition of NF-κB signaling also resulted in a reduced production of inflammatory cytokines. The production of TNF-α in human monocytes was measured by ELISA after TLR2/1 and TLR2/6 stimulation with and without antagonist. Finally, IC_{50} values for each compound were determined. The obtained values are summarized in Table 4. Three compounds (**5**, **22** and **27**) showed activity in the low micromolar range against TLR2/1, TLR2/6 and TLR2/1 respectively. Compound **40**, **77** and **75** had IC_{50} around 10 µM against TLR2/1, TLR2/6 and TLR2/6, respectively.

In the case of compounds **5**, **27** and **40** the determined IC_{50} values suggest selectivity for TLR2/1 signaling. The obtained values for **22** seem to show a preference for TLR2/6. This is surprising, as no significant selectivity could be observed in the NF-κB reporter assays described above. This divergence could be due to the different assay sensitivity that also leads to differing concentration dependence of the inhibitory signal. Assuming that measurement of

the TNF-α concentration is the most reliable signal, there still can be several reasons for the observed differences in TLR2/1 and TLR2/6 signaling inhibition. On the one hand, the antagonists could actually bind to the TLR1 or TLR6 lipopeptide binding site in addition to binding to the TLR2 cavity. This could lead to an enhanced activity towards one of the receptors. A more detailed analysis of this hypothesize is performed in chapter 4.3.1. On the other hand, other assay related, potential causes for the divergence in activity against TLR2/1 and TLR2/6 are possible and will be discussed in chapter 5.1.1.

Table 4: IC_{50} for TLR2/1 and TLR2/6 antagonism. The IC_{50} values were determined by the relative inhibition of TNF-α production in human monocytes. Values represent the mean ± standard deviation.

Compound ID	TLR2/1 IC_{50} [µM]	TLR2/6 IC_{50} [µM]
5	3.33 ± 1.14	16.69 ± 1.04
22	30.66 ± 1.65	4.49 ± 1.47
27	5.34 ± 1.72	>200
40	16.72 ± 1.52	48.18 ± 1.34
44	26.16 ± 1.11	30.72 ± 1.59
62	~100	>200
66	39.95 ± 1.29	12.15 ± 1.25
75	28.01 ± 1.23	10.91 ± 1.38
77	18.5 ± 1.28	11.60 ± 1.16
78	23.74 ± 1.41	11.13 ± 1.22

4.3 Analysis of experimental results and refined binding models

In order to analyze the binding mode of the TLR2 modulators identified by ligand-based screening, docking studies were performed. The identified antagonists (**5**, **22**, **27**, **40**, **44**, **62**, **66, 77** and **78**) were docked into the front part of the TLR2 lipopeptide binding site. Compounds **5** and **27** showed a possible TLR1 selectivity in the human monocytes assay. As mentioned before we hypothesized that this could be due to an additional binding site in the co-receptors. Potential binding modes for explaining this finding were investigated, as well. For the TLR2 agonists identified in the course of this study, a binding hypothesis was developed and docking performed. The obtained results and surmised binding modes will be summarized in the following sections.

4.3.1 Docking studies with discovered TLR2 antagonists

The potential binding mode of the discovered TLR2 antagonists was analyzed through docking into the TLR2 binding site. For this purpose, in a first study, the binding pocket was defined to be the front part of the pocket that had been also selected for the structure-based virtual screening (chapter 4.1.1.1). The surmised poses for **62**, **66**, **77** and **78** are shown in Figure 30. All four compounds are embedded in the front part of the pocket. **62**, **77** and **78** form H-bonds to the backbone nitrogens of Leu350 and Phe349 through the carbonyl-oxygen in the case of **77** and an oxygen of the nitro-group in the cases of **62** and **78**. The comparison of **62** and **78**, which both contain a similar scaffold with three aromatic rings and a nitro-group, suggests that the hydrophobic contacts formed by the tri-methyl substituent of **78** are important for activity, as **62** that only is substituted with a methoxy-group is a much weaker antagonists. **66** does not form the H-bonds to Phe349 and Leu350. In contrast, the dimethylamino-substituent is possibly positively charged and interacts with the π-system of the Phe284 side chain.
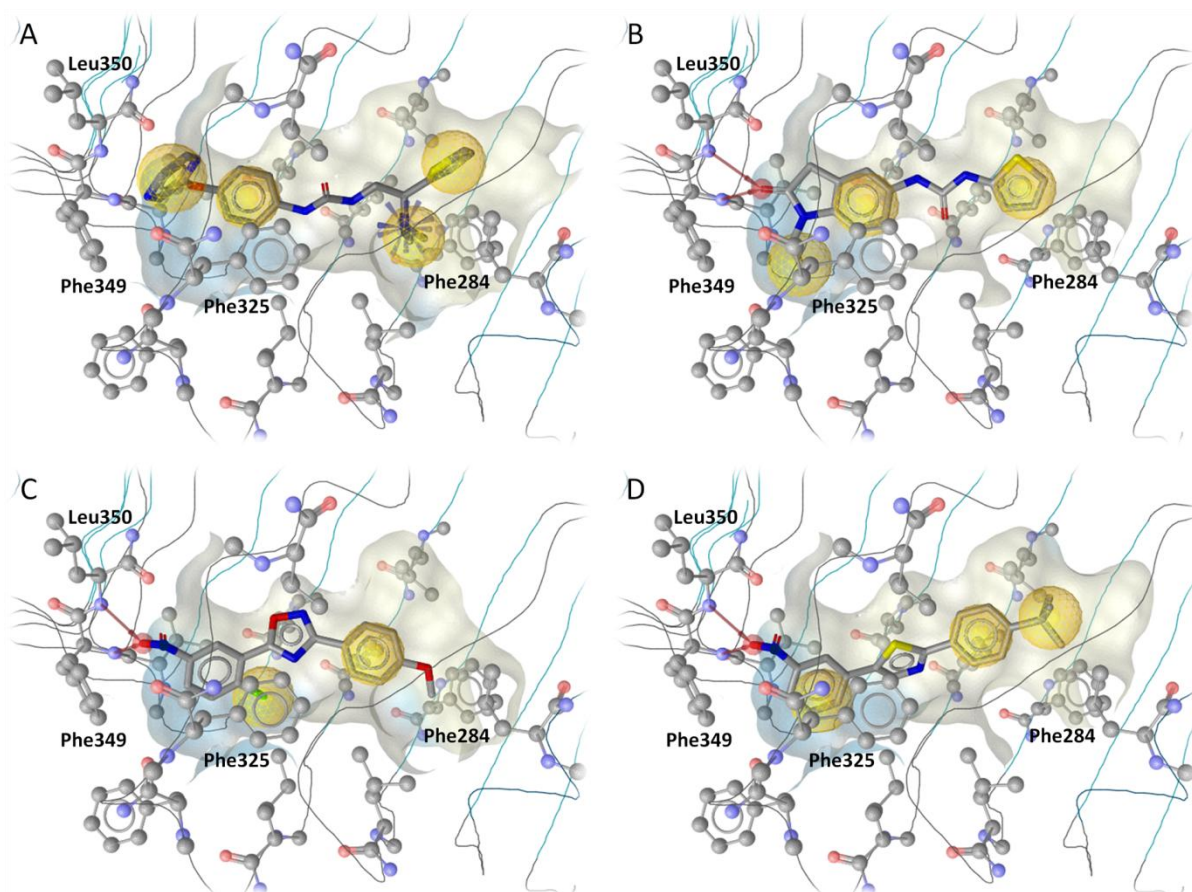
Figure 30: Docking poses of TLR2 antagonists identified by ligand-based virtual screening. The ligands (A: **66**, B: **77**, C: **62**, D: **78**) are shown in stick mode, interacting protein residues in ball and stick. Chemical features are color coded: H-bond acceptor – red arrow, hydrophobic contact – yellow spheres, positive ionizable area – blue star.

The docking study into the front part of the TLR2 binding site did not result in satisfying binding poses for the compounds with phenylurea scaffold (**5**, **22** and **27**). Nevertheless, we were especially interested in elucidating their potential binding mode due to their low IC$_{50}$ and their similar scaffolds. Thus, we extended the screening to the back part of the pocket. We docked the compounds into sub-pocket P4 that is the sub-pocket located further away from the aperture of the lipopeptide binding site (as described in chapter 4.1.1.1). An overview on the two binding regions in the TLR2 ligand binding site is shown in Figure 31. Ligands binding the back site of the pocket also interfere with Pam$_3$CSK$_4$ binding and could so antagonize TLR2 activity.

In the identified binding poses, the compounds are embedded into a narrow sub-pocket and form H-bonds with the carbonyl-oxygens of the backbone of Asp305 and Pro306, which

possibly further stabilize binding (Figure 32). In the case of both compounds, the hydrogen bonds are formed by the two amine nitrogens and the backbone. In addition to these interactions, the fluorine substituent of **27** forms a hydrogen-bond like electrostatic interaction with the backbone nitrogen of Ile276 [132]. This fluorine interaction seems to be necessary for activity, as other tested compounds lacking this substituent showed no inhibitory activity towards TLR2. In the case of **5**, the substitution pattern of the phenyl-ring is important as well. Compounds with substituents in a 2 and 5 or 2 and 4 positions are inactive, as they do not fit the narrow cavity anymore.
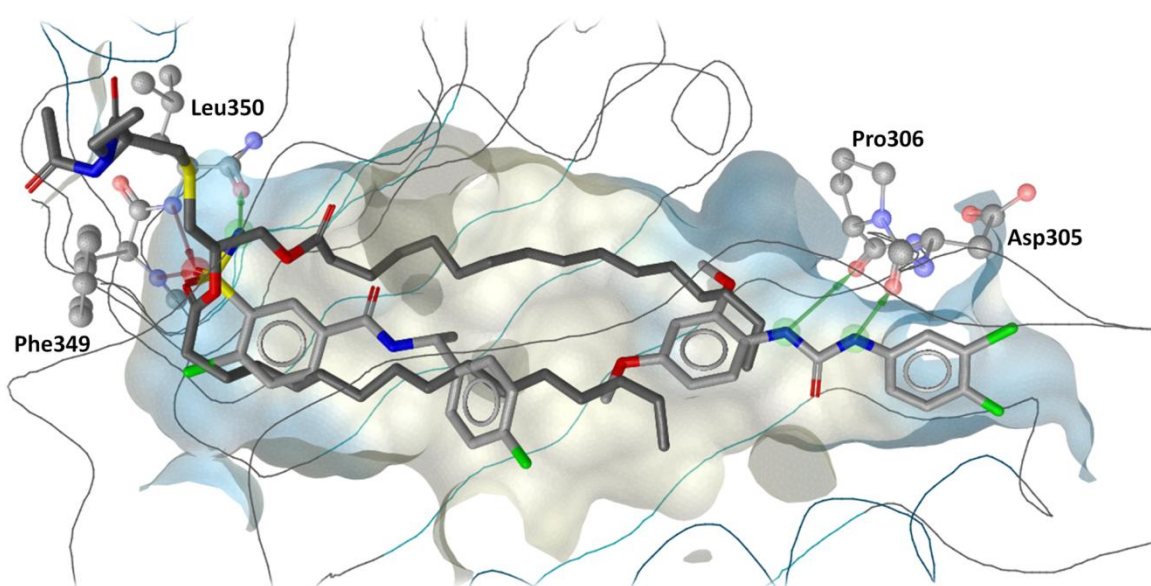
Figure 31: Overview on the TLR2 front and back ligand binding sites. The protein backbone is shown as lines, the binding surface is color encoded (yellow – hydrophobic, blue – lipophilic). The lipid part of Pam$_3$CSK$_4$ is shown in dark grey for orientation, the peptide moieties are hidden for more clarity. **75** docked into the front of the pocket and **27** into the back are depicted as an example. Protein residues forming key H-bonds are shown as well as the interactions with the ligands. Interactions are color-coded: hydrophobic – yellow spheres; H-bond donor – green arrow; H-bond acceptor – red arrow.

Figure 32: Predicted binding mode of compounds **5** (A) and **27** (B) in the TLR2 binding site. On the left the observed interactions between the protein and the ligand are shown in 2D, on the right the 3D image of the docked antagonist is depicted. Interactions are color-coded: hydrophobic – yellow spheres; H-bond donor – green arrow; H-bond acceptor – red arrow.

4.3.2 Potential mechanism for co-receptor selectivity
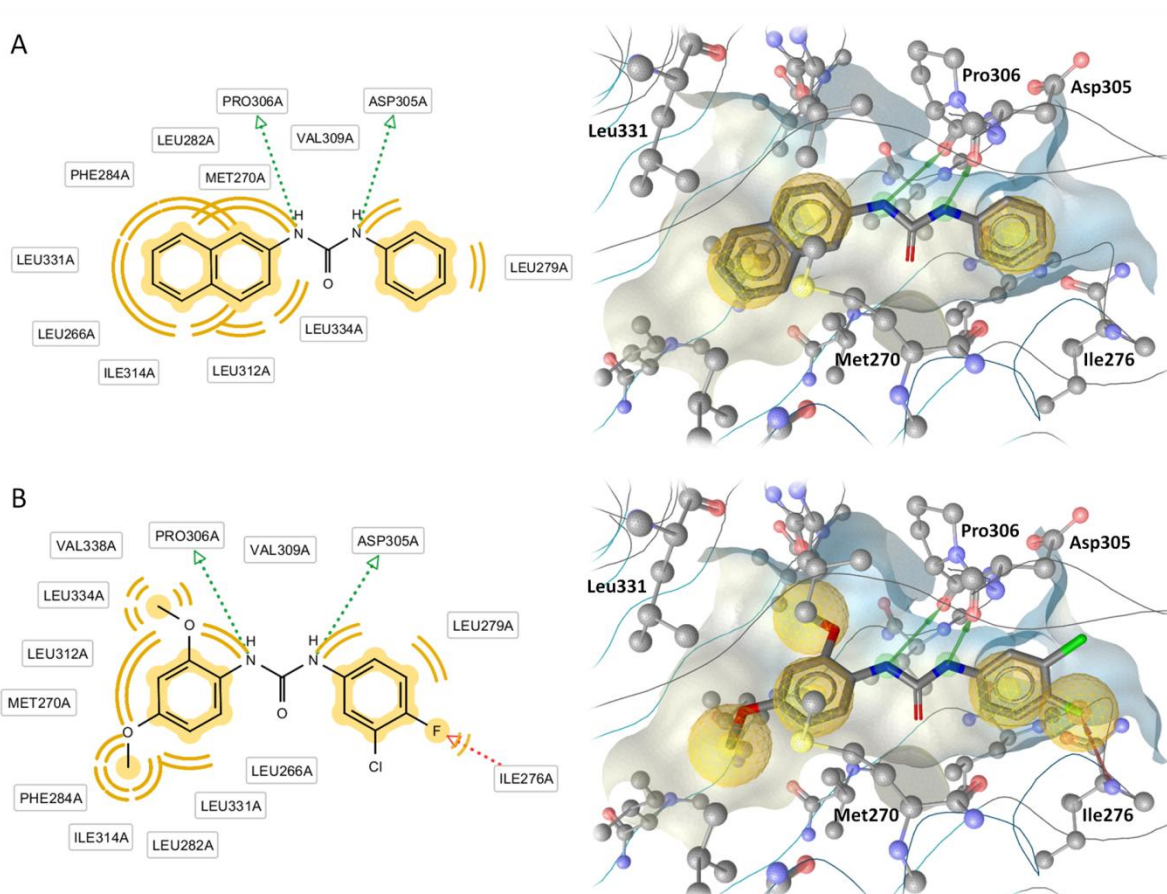
As described above, two of the identified antagonists (**5** and **27**) showed selectivity towards the inhibition of TLR2/1 signaling in the ELISA assays with human monocytes. Based on this finding, we hypothesized that a possible explanation for the divergence in activities could be that these molecules bind into the TLR1 binding site in addition to binding to the one in TLR2 and thus inhibit TLR2/1 in a higher extent. This binding site is much smaller in TLR6 as it is blocked by a phenyl-residue [21], which could be an explanation for the enhanced activity towards TLR1. Therefore, we docked the two selective molecules into the TLR1 ligand binding site. The docking was performed using the aforementioned crystal structure of TLR2 dimerized with TLR1 and Pam$_3$CSK$_4$ (PDB: 2Z7X [20]) after removal of the lipopeptide ligand and the TLR2 monomer. In the case of TLR1, the whole binding site was taken into

54

account for docking, as it is much smaller than in TLR2. The selective ligands fitted the TLR1 binding site well, this being a possible explanation for their selectivity. The ligands are positioned deep inside the pocket and nearly fill it out completely, forming hydrophobic contacts with the highly lipophilic inner part of the pocket. In the case of **5** and **27**, the binding pose is stabilized by an H-bond between the carbonyl-oxygen of the ligand and the hydroxyl-group of Tyr320.
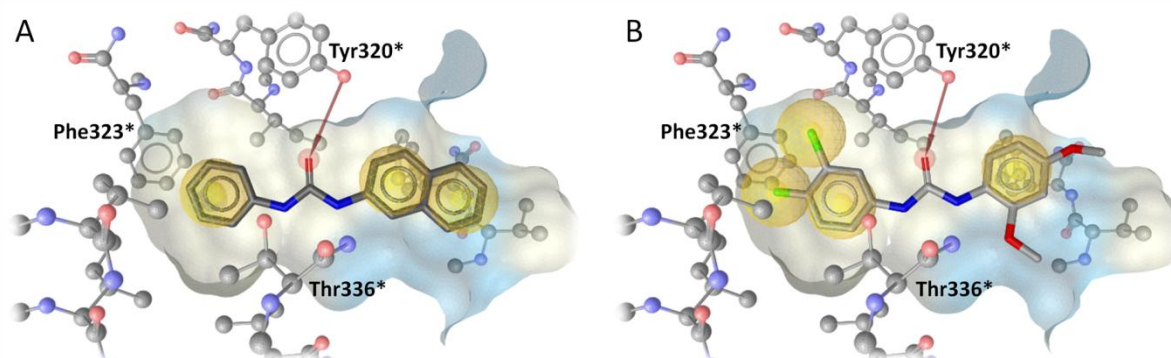

Figure 33: Docking poses of TLR2/1 specific antagonists. Predicted binding poses of the two most potent TLR2/1 specific antagonists (A: **5**, B: **27**) in the TLR1 ligand binding site. The 3D representation of the ligand binding poses is shown with the binding pocket of TLR1. Important interacting residues are shown as ball and stick, further residues are omitted for better clarity. Chemical features are color coded: H-bond acceptor – red arrow, hydrophobic contact – yellow spheres.

4.3.3 Binding model for small molecule TLR2 agonists

In the course of our study, four compounds with agonistic activity towards TLR2 signaling were identified. The compounds stimulated NF-κB activity in HEK-TLR2 cells weakly. Their activity was found to be much stronger when applied together with the lipopeptide agonists (Pam$_3$CSK$_4$ or Pam$_2$CSK$_4$), which could be a sign for a synergistic mechanism. We were interested in determining a possible binding mode of the compounds that would explain such a behavior. For this purpose a docking study was initiated with the TLR2/1 dimer. The fact that the activity of the compounds seemed to be synergistic to the lipopeptide ligands suggests that they bind to a different region of the protein than the endogenous ligands. Until recently, no structural information on how small molecules could activate TLRs was available. Recently, the first crystal structure of a TLR (TLR8) co-crystallized with a small molecule

agonist was released [36] and made it possible to understand how small molecules activate this receptor. The compounds bind to the interface of the two TLR8 monomers and causing them to rearrange in a way that the activated dimer is formed as described in chapter 1.1.2. In this form, the compounds are located in a small aperture between the two monomers and interact with key residues on both of them. We hypothesized that TLR2 agonists could possibly act in a similar manner, binding to the interface between TLR2 and TLR1 or TLR6 and thus supporting the dimerization and signaling. This might also explain why the activity is stronger in cooperation with the lipopeptides.

For our docking study, we used the crystal structure of TLR2 co-crystallized with TLR1 and Pam$_3$CSK$_4$ (PDB code: 2Z7X [20]) that we had previously used for docking. We performed a cavity detection of the whole dimer using Q-SiteFinder [79]. The main binding site found was the lipopeptide binding site. However, as the activity of the agonists seemed to be synergistic we hypothesized that they had to bind to an alternative area of the protein. For this reason, we were interested in finding binding sites adjacent to the dimerization interface, as we figured that the activation mechanism could be similar to the one of TLR8 activation with small molecules. Two main cavities at the interface were detected (b1 and b2, Figure 34A). After the addition of the missing hydrogens to the protein, we performed docking into both cavities. Only the docking into the front cavity (b1, Figure 34B) led to satisfactory results as the other cavity is too small and shallow to fit the molecules appropriately.

In the surmised binding poses, the compounds are embedded between the two receptor monomers and interact with both TLRs holding them together. As an example the binding poses of **47** and **61** are shown in Figure 34. In the case of **47** the H-bonds are formed between the carbonyl-oxygen of the thioxopyrrolidinone ring and the backbone nitrogen of Asp362*[1] and the sulfur and Lys347. The methylquinoline forms hydrophobic contacts to Arg316, His318 and Thr363*. In the case of **61** the binding takes place through the formation of H-bonds to His318, Glu386*, Arg316 and Cys287, as well as hydrophobic contacts to Lys385*, Gly383* Leu371, Thr361* and Lys347.

It is important to keep in mind that structural rearrangement after binding of the small molecules is likely to occur, leading to binding poses different from the ones achieved with

---

[1] In this chapter, residues belonging to TLR1 will be distinguished from those pertaining to TLR2 by a "*".

the available crystal structure. In order to draw reliable conclusions about the binding mode of the compounds a more detailed biological characterization would be necessary. Mutation studies of the receptors involved in binding or ideally co-crystallization could help to confirm the actual mechanism of TLR2 antagonism by small molecules.
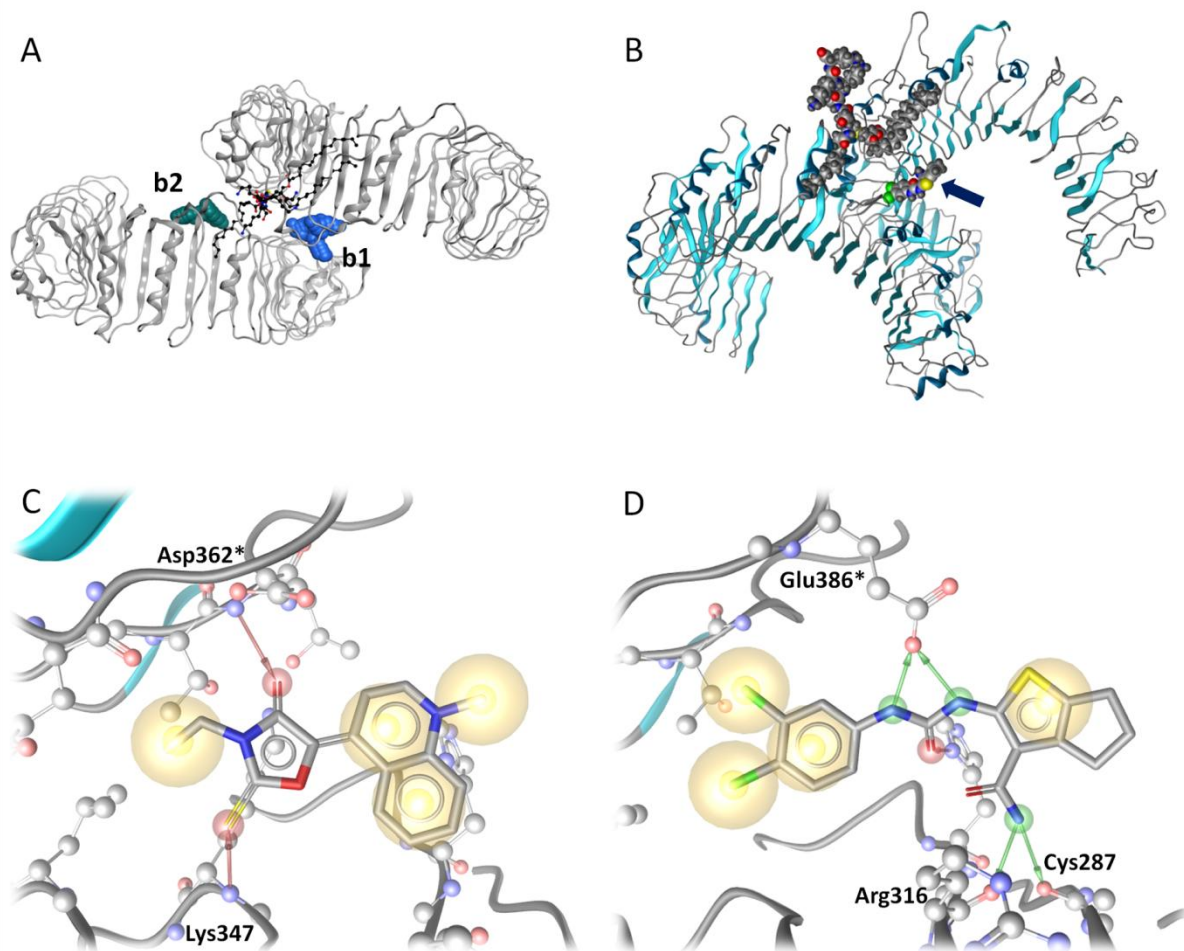

Figure 34: Depiction of the hypothesized mechanism of small molecule TLR2 agonists. A) b1 and b2, the two binding sites detected at the protein-protein interface. B) Selected binding site b1. C) Docking pose of **47** and D) of **61**.

4.4 TLR2 antagonists with benzotropolone scaffold

As mentioned in chapter 1.1.5.2, Cheng *et al.* [60] recently published a study describing the selection of TLR2/1 antagonists by high-throughput screening and further lead optimization. In a first step, they screened the *NCI diversity library* (1363 compounds) through a phenotypic assay that used the inhibition of nitric oxide (NO) production in mouse macrophages as read-out. Nine initial hits with 70 % inhibition of TLR2 signaling (here NO production) at 3 μM compound concentration could be identified. Out of those, one compound (NCI35676, Figure 35) was selected for further characterization and optimization. Concentration dependence of the inhibitory activity was shown and an IC_{50} of 2.45 ± 0.25 μM defined. The selectivity-profile towards other TLR receptors was measured as well, at a concentration of 3 μM. It shows high specificity towards TLR2/1 signaling. It is noteworthy to state that no significant inhibition of TLR2/6 (FSL-1 used as agonist) signaling was reported. Only a slight inhibition of TLR3 signaling could be measured.

In an additional synthesis step, the SAR around the benzotropolone scaffold was explored leading to 26 further compounds that were also characterized with regards to their antagonistic activity on TLR2/1 signaling. The most potent hit found was CU-CPT22 (Figure 35) with an IC_{50} of 0.58 ± 0.09 μM. The higher activity of CU-CPT22 appeared to be due to the addition of a six-carbon aliphatic chain at the R^6 position. Selectivity against other TLR receptors was measured as well, interestingly at an inhibitor concentration of 0.5 μM. While the inhibition towards TLR3 completely disappears, a notable inhibition of TLR4 signaling (around 15 %) is detected. A possible explanation for this could be that CU-CPT22 resembles subunits of LPS, the specific TLR4 ligand through, because of its hydroxylated aromatic ring system with hexyl substitution. The antagonist could thus bind to TLR4 or its co-receptors MD-2 or CD-14. Several TLR4 antagonists with similar structures (hydroxylated ring system and lipid chain) are known [133].

In a last modeling step, Cheng *et al*. performed flexible docking into the TLR2/1 dimer in order to elucidate the binding mode of CU-CPT22. In the binding pose they propose (Figure 36), the ring-system of the antagonists is located in the TLR2 cavity and a hydrogen bond is formed between the oxygen of the carbonyl-group and the amine nitrogen of Phe349. According to their analysis, the hexyl chain at position R^6 bridges over the phenyl-residue of Phe349 and into the TLR1 lipid binding site, where it forms hydrophobic contacts. Cheng *et*

al. also state that these interactions of the chain with TLR1 are the reason for the selectivity of the antagonist, as the lipid binding site of TLR1 is not present in TLR6.
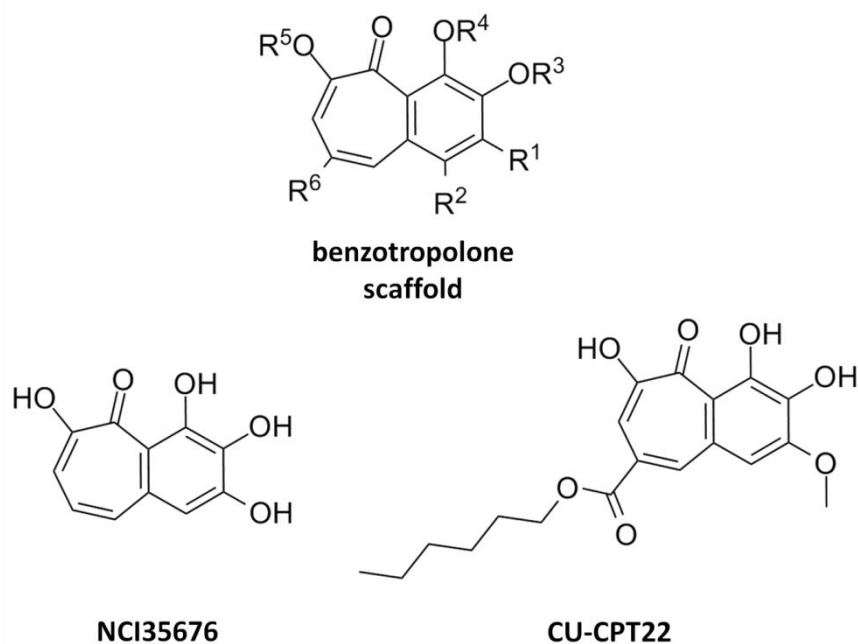
Figure 35: Chemical structures of the initial screening hit NCI35676 and CU-CPT22
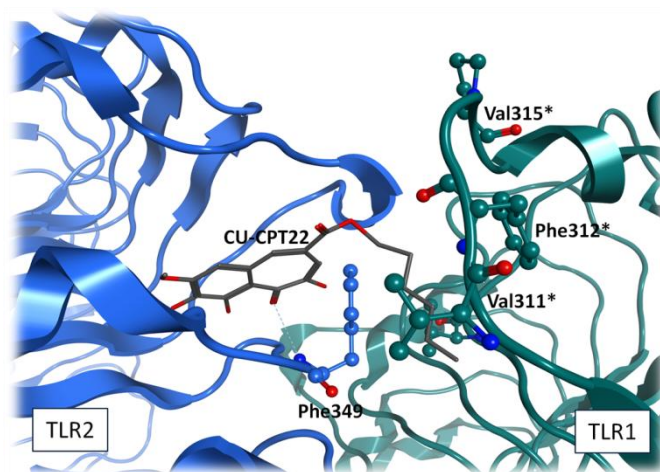
Figure 36. Reproduced binding pose of CU-CPT22 into the TLR2/1 dimer (PDB: 2Z7X) proposed by Cheng *et al*. [60]

The rationale Cheng *et al*. employ to select the binding pose is based on the stated selectivity of CU-CPT22 towards TLR2/; however, this argumentation has several flaws. First, it is unlikely that the interaction of the hexyl chain with TLR1 is the actual reason for selectivity, as the starting molecule (NCI35676) without the 6-carbon substituent is highly selective. The

59

relative selectivity is actually less pronounced in the case of CU-CPT22 than in the case of NCI35676. Several other compounds tested also lack the 6-carbon chain as well and thus could not interact with TLR1 (**c1 – c14**) [2]. Others, like compound **c24**, have a substituent (tolyl) that is too bulky and rigid to fulfill the suggested binding mode and so would have to be much less active, which is not the case according to the published activity values. The SAR described for the 26 compounds is not explained by the proposed binding mode.

Furthermore, the crystal-structure used for the docking study is the activated dimer. As described in chapter 1.1.2., TLRs exist as pre-arranged dimers on the cell membranes, which re-arrange and finally dimerize only after ligand binding [20]. The recent crystallization of the TLR8 homodimer with and without small molecule agonists is further proof for this mechanism [36]. So, if CU-CPT22 impedes the lipopeptide ligand to bind to TLR2 it is unlikely that the dimer will exist in the activated form that was used for docking.

To understand how the measured selectivity occurs, it is important to analyze the experimental methodology in detail. In the performed phenotypic assay, the murine macrophages were stimulated with agonists specific for the different TLRs in the presence of the tested antagonist and the relative inhibition in comparison to stimulated cells without antagonists was assessed. In the case of TLR2/1 the lipopeptide agonist Pam_3CSK_4 was used and in the case of TLR2/6, FSL-1. The concentrations applied in both cases vary significantly in the case of Pam_3CSK_4, 200 ng/mL (66 nM) and in the case of FSL-1, 10 ng/mL (6 nM). This is due to the fact that FSL-1 is a much stronger agonists than Pam_3CSK_4. Hence, the relative inhibitions measured are difficult to compare, as the differences could also be caused by the different agonistic potencies of the agonists used to stimulate TLR2. Furthermore, the peptide part of Pam_3CSK_4 and FSL-1 differ from each other and could have led to a different binding mode and thus different effects of the antagonists.

With this in mind, we set up a docking study with the compounds identified in the mentioned publication in order to suggest a more relevant binding mode. We first tried to reproduce the binding pose described in the TLR2/1 dimer and then performed a docking study into the TLR2 monomer.

---

[2] In this text a "c" will be added to the compound-ids of the structures described by Cheng *et al*. in order to better distinguish them.

4.4.1 Unraveling the mechanism of TLR2 antagonism

As a first step in order to elucidate the binding mode of the compounds with benzotropolone scaffold, we performed docking studies to reproduce the binding mode proposed by Cheng *et al*. In addition to CU-CPT22, we selected two additional compounds that represented key structural features of the SAR (Figure 37). NCI35676 was selected, as it was the original hit. Compounds **c24** was docked as well because it is the only compound with a rigid scaffold (tolyl substituent instead of 6-carbon alkyl chain) and comparable activity.
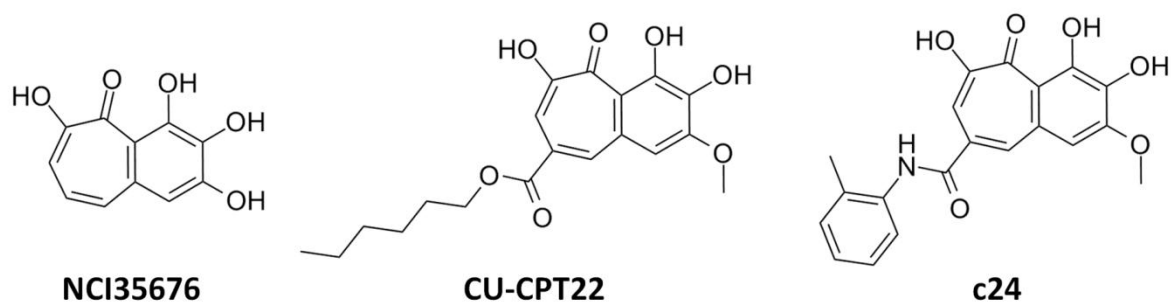

Figure 37: Compounds docked to elucidate the binding mode of the TLR2 antagonists with benzotropolone scaffold.

The crystal structure of the TLR2/1 dimer with co-crystallized lipopeptide (PDB-code: 2Z7X) was used for docking after removal of all ligands. In order to avoid bias, the whole ligand binding pockets of TLR2 and TLR1 were considered as part of the binding site [119]. When no docking constrains were set, a binding pose similar to the one proposed by Cheng *et al*. (hydrophobic contacts with Val311*, Phe312*, Pro315* and Val339* and H-bond between Phe349 of TLR2 and the carbonyl-oxygen of the ligand) [60] could not be generated. NCI35676 can obviously not fulfill the described interactions, as it lacks the lipophilic substituent. For CU-CPT22 and compound **c24**, the docking was repeated setting docking constrains (distance to TLR1 channel and H-bond with Phe349). For CU-CPT22 a similar pose to the published one could be generated. Nevertheless, the docking pose showed a highly strained and unfavorable conformational geometry and the H-bond with Phe349 was formed by one of the hydroxy-oxygens, not the carbonyl-oxygen, as suggested by the image of their publication. As mentioned above before, compound **c24** has a too rigid scaffold to bind in the proposed mode; it is not possible to generate a binding pose with hydrophobic contacts to the TLR1 ligand binding site and H-bonds to Phe349 or Leu350.

After having discarded the proposed binding mode we re-analyzed the binding in order to elucidate the actual binding mode of the described compounds. In most of the poses we generated in the docking study into the TLR2/1 dimer cavity without constrains, the compounds were embedded solely into the TLR2 binding site with no contact to TLR1. This is in accordance with our mechanism of antagonism developed in the structure-based part of our screening (chapter 4.1.1), where we surmise that binding to the TLR2 binding site is sufficient to inhibit its activity. For this reason, we decided to perform a more focused docking study in the TLR2 binding site to explore the SAR of the compound series published by Cheng *et al*. As the IC_{50} values indicated in the paper were determined in a murine cell system, we decided to use the crystal structure of murine TLR2 dimerized with TLR6 (PDB code: 3A79) instead of the human structure used by the authors, as small differences in the binding site could possibly lead to different activities of the compounds. The docking was performed defining the whole cavity as the binding pocket after removal of the ligand, the TLR6 monomer, and protonation of the protein. The resulting binding poses were rescored and ranked by the "interaction feature count", which counts the number of all occuring pharmacophoric interactions (H-bonds, hydrophobic contacts, π-stacking etc.) between the ligand conformation and the protein. Furthermore, we focused on poses in which our three test compounds formed interactions to the same residue.

Through this procedure we defined a binding mode that fully explained the SAR described by Cheng *et al*., as shown for the test compounds in Figure 38. The ring system of the compounds is located in the front pocket of the TLR2 ligand binding site. In the case of NCI35676, the following interactions were formed: the hydroxyl-group at position R^5 forms an H-bond with Ser346, the carbonyl-oxygen interacts with Phe349, the hydroxyl-group at position R^3 with Leu350 and the hydroxyl-group at position R^1 forms an H-bond with Tyr326. Hydrophobic contacts take place between the aromatic ring-system and Val248, Phe325, Ile319 and Leu328. The reason for the stepwise decrease in activity when hydroxyl-groups are methylated (**c4**, **c5** and **c6**) becomes clear with this binding mode, as the H-bonds formed are necessary for binding and each of them contributes to binding affinity. The insertion of a carboxyl-group at position R^6 (**c10**) decreases the activity of the compounds as it places the hydrophilic carboxyl-group close to highly hydrophobic residues (Ile319, Val348). The esterification enhances the lipophilicity of the group and thus increases the binding affinity (**c11**). The introduction of long carbon chains at the R^6 position cause an increase in binding

affinity as they permit the formation of hydrophobic contacts and make the compounds more similar to the endogenous lipopeptide ligands. The introduction of the tolyl substituent (**c24**) causes no decrease of activity as the binding site is big enough to fit the compound and hydrophobic contacts can be formed by this group as well. The slight decrease of activity of **c19**, the compound with the longest carbon chain (tetradecane), can also be explained by this binding mode, as the chain is too long for the binding pocket and does not fit as well as the compounds with shorter chains. In the case of CU-CPT22, the chain length is long enough to increase the binding potency but short enough to fit the pocket easily. The hydroxyl-groups form the H-bonds to Ser346, Phe349 and Leu350 as is the case for the original hit compound NCI35767. This is the reason why this compound shows the lowest IC_{50} value.
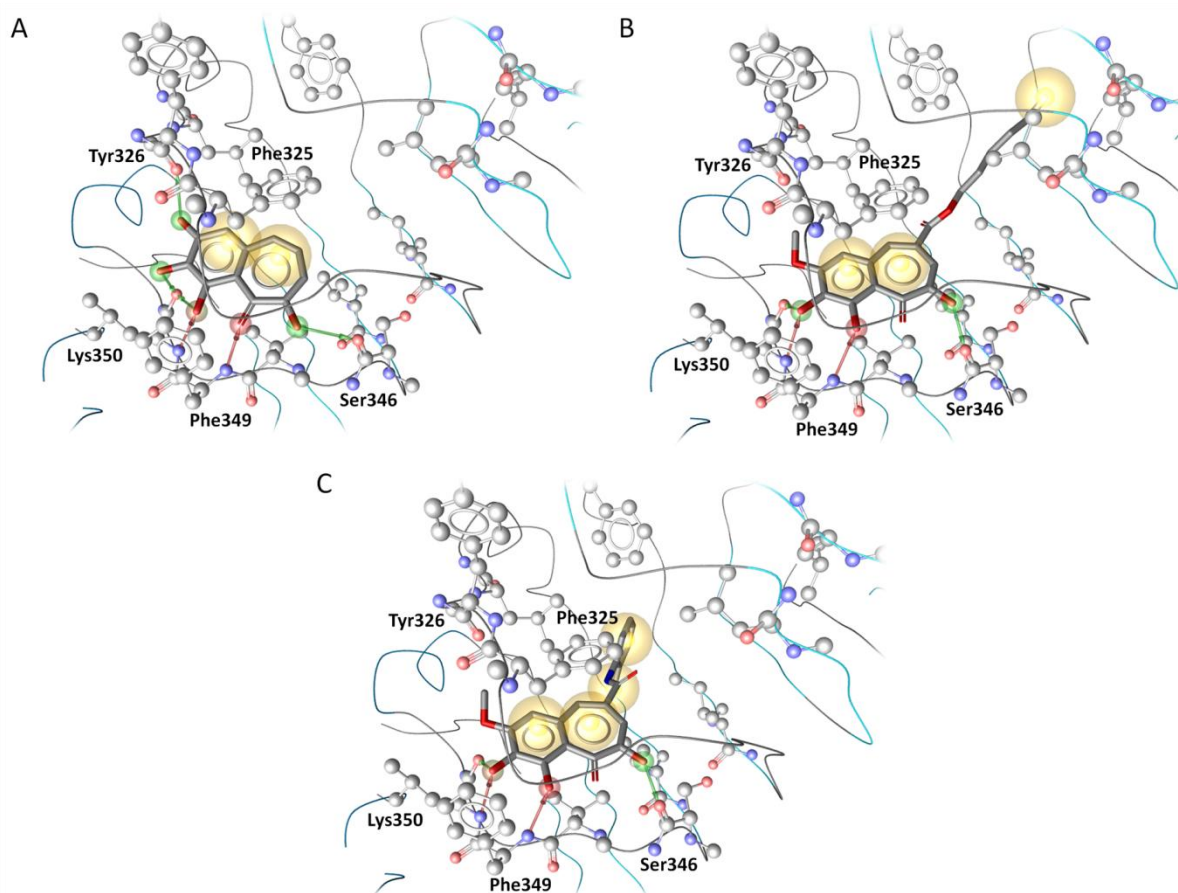

Figure 38: Predicted binding mode of compounds A) NCI35676, B) CU-CPT22 and C) c24 in the TLR2 binding site. The ligands are shown in stick mode, interacting protein residues in ball and stick. Chemical features are color coded: hydrogen-bond acceptor – red arrow, H-bond donor – green arrow, hydrophobic contact – yellow spheres.
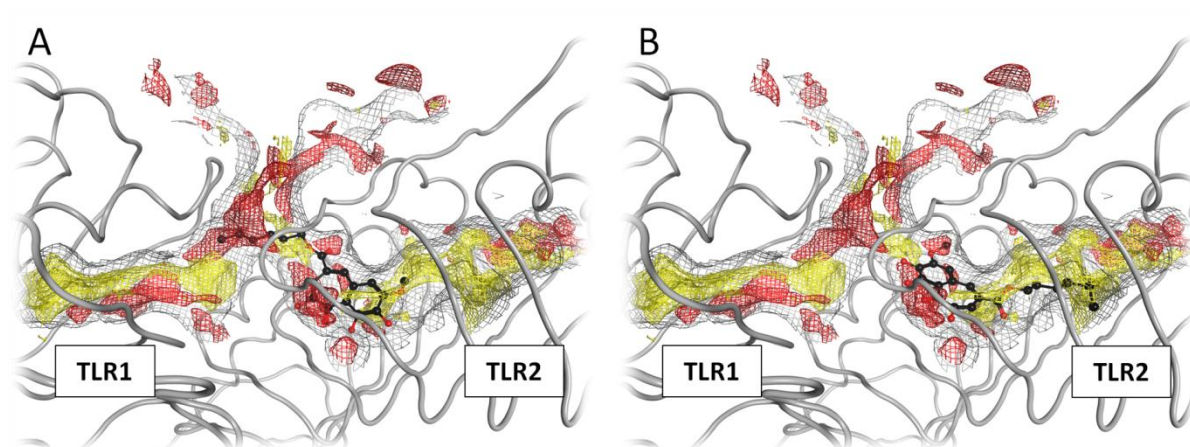
Figure 39: Predicted binding mode for CU-CPT22 in the TLR2/1 binding site with calculated MIFs. A) Reproduced docking pose predicted by Cheng *et al*. B) Docking pose resulting from our study. The MIFs shown were calculated with a hydrophobic probe (yellow) and a hydroxyl-group probe (red). The protein backbone is shown in grey, the ligand in black.

In order to compare the binding mode we propose to the one published by Cheng *et al*. we calculated MIFs of the TLR2/1 binding site and assessed their overlap of the docked compounds (Figure 39). The interaction potentials for a hydroxyl-group and for hydrophobic contacts were calculated. It is clearly visible that the binding pose we propose the hydroxyl-groups lay in the area more favorable for their interaction as well as the alkyl chain. In contrast in the published pose, the alkyl chain lays in an area more favorable for hydroxyl-groups.

We included a further docking validation step using the human TLR2 crystal structure (PDB code: 2Z7X) in order to check whether the binding into the human homolog functions in a similar manner. Even though murine and human TLR2 are highly homologous proteins, small changes in the binding pocket could lead to changes in activity. The main difference between the front region of the murine and human TLR2 binding site is Phe355, which is replaced by Leu355 in the human receptor. Our analysis showed that the binding mode in the murine protein was very similar to the one in the human receptor, though in the murine structure the ligands were additionally stabilized through the phenyl residue.

In conclusion, we performed a docking study to find out the binding mode of TLR2 antagonists with benzotropolone scaffold and proposed an alternative binding mode that fully explains the described SAR of this chemical series; however, our binding mode does not

explain the selectivity towards TLR1 observed by the authors. We believe that this is caused by the employed TLR2/6 antagonist FSL-1, which is much more potent than the TLR2/1 agonist Pam_3CSK_4.

4.5   Integrative model of TLR2 antagonism through small molecules

During the course of this study, ten small molecule TLR2 antagonists were identified and validated. In the aforementioned publication by Cheng *et al*. [60] that recently was published, nine primary hits from the *NCI-2 diversity library* were discovered as TLR2 antagonists and 26 derivatives of one of the compounds were described and IC_{50}s determined. Due to this, we decided to integrate the newly available information into a refined TLR2 antagonism model. In the next chapters, an overview on the available scaffolds and binding modes of small molecule TLR2 antagonists is given (chapter 4.5.1). Then, the generation of a collection of 3D pharmacophores representing this data is described, highlighting key features of the obtained models (chapter 4.5.2).

4.5.1   Chemical scaffolds and binding modes

Apart from the 26 benzotropolone derivatives synthesized and characterized by Cheng *et al*. [60], the authors published nine further compounds that were identified as TLR2 antagonists through high-throughput screening. An overview on these TLR2 antagonists and the molecules discovered in our screening campaign is given in Figure 40. As a first step to integrate the new data, we structurally compared the newly published compounds to our screening hits. One of the antagonists (NCI522131) contains a phenylurea scaffold similar to the one of compounds **5**, **22** and **27**, discovered in our work. This further validates the TLR2 antagonistic activity of compounds with such a scaffold. The recently published compound with phenylurea scaffold (NCI522131) is especially similar to compound **27**. The fluorine substituent on the halogenated phenyl-ring is replaced by chlorine and the methoxy-group at position 2 is replaced by a hydroxyl, whereas the other methoxy moiety is missing. A similar compound with phenylthiourea scaffolds (NCI201634) was described, as well. The remaining compounds do not show any noteworthy similarity to the antagonists discovered in our screening, nor are there other reoccurring scaffolds in this set of molecules.
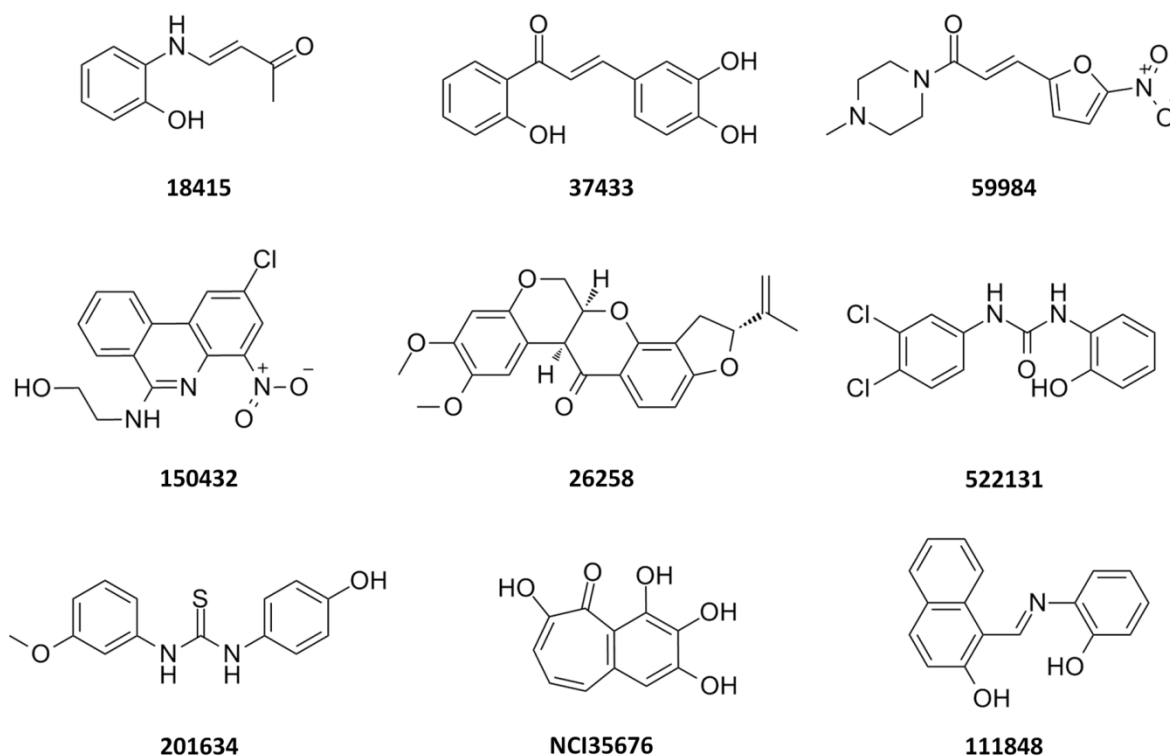
Figure 40: TLR2 antagonsits discovered by high-throughput screening by Cheng _et al._ [60]

The binding mode of the benzotropolone derivatives was discussed in chapter 4.4.1. The compounds bind to the front part of the TLR2 ligand binding site, forming H-bonds with Leu350, Phe349 and Ser346. According to the docking quality categories described in chapter 5.2 this binding mode is highly reliable, as it explains the SAR of the chemical class well. Furthermore, binding of compound CU-CPT22 to TLR2 was confirmed experimentally in the publication [60]. Our binding hypothesis for **75** is similar, lying in the front region of the TLR2 cavity as well with H-bonds to Phe349 and Leu350 (chapter 4.1.1.5).

In contrast, the predicted binding site for the compounds with phenylurea scaffold is located at the back part of the binding pocket (chapter 4.3.1). As NCI522131 shares this scaffold, we performed a docking study with this compound to verify whether it could bind in a similar manner as the compounds discovered in our study. The docking was performed as described for the other antagonists with this scaffold in chapter 4.3.1. The identified binding pose is shown in Figure 41. The binding pose is in agreement with the previously identified binding mode for this group of compounds. An H-bond is formed between the nitrogen of the phenylurea scaffold and the carbonyl-oxygen of Pro306. In addition, an H-bond between the hydroxylgroup of the ligand and Leu334 is formed. This further validates the binding mode

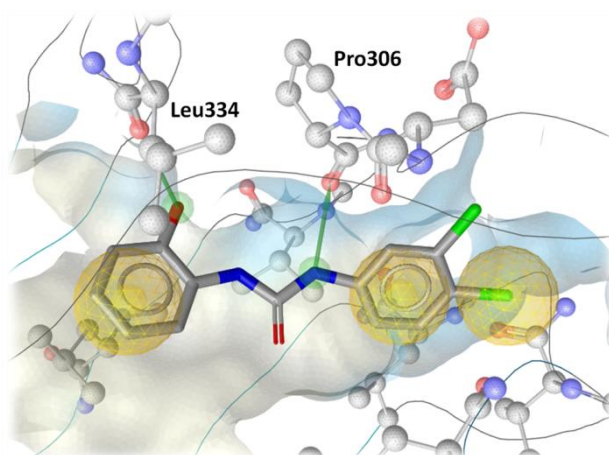hypothesized for TLR2 antagonists with a phenylurea scaffold.

Figure 41: Binding hypothesis for NCI522131. Important interacting protein residues are shown in ball-and-stick style. Interactions are color-coded: hydrophobic – yellow spheres; H-bond donor – green arrow.

4.5.2 Essential chemical features of the integrative 3D pharmacophore model

The first step to generate an integrative model of small molecule TLR2 antagonists was to create a validation dataset. The antagonists identified in our study and the compounds published by Cheng *et al*. [60] were merged into a set of actives comprising 45 molecules. Then, all compounds tested in our study that could not be confirmed as TLR2 antagonists including TLR2 downstream inhibitors and agonists were defined as inactives, which led to a list of 65 structures. Using an in-house KNIME workflow for decoy generation, myDecoyFinder, a set of 1303 decoys was generated starting from the 45 active molecules. The workflow is an adaption of the decoy selection for the Database of useful Decoys (DUD) [134] and the software DecoyFinder [135]. The underlying principle for decoy selection is that decoys have to be similar to a specific ligand by physical properties (cLogP, molecular weight, number of H-bonding groups) but different in ligand topology [134].

We decided to generate structure-based 3D models only for the benzotropolones and compound **75**. To generate the structure-based pharmacophore of the benzotropolones, **c24** was selected for pharmacophore generation as it has a rigid scaffold that makes the determination of pharmacophoric features more accurate. A 3D pharmacophore model was generated using the automatic pharmacophore generation protocol in LigandScout 3.1. [119].

This model comprised a total of six features: an H-bond donor towards the carbonyl oxygen of Ser346, an H-bond donor to the carbonyl oxygen of Leu350, two H-bond acceptors with the amine nitrogens of Leu350 and Phe349 respectively, a π-interaction with Phe325 and a hydrophobic contact region deeper in the pocket close to Phe295, Leu328, Leu331 and Ile314. In a next step, the H-bond donor feature towards the carbonyl oxygen of Leu350 was set as optional, as we considered it less relevant than the other features. The hydrophobic region was set optional so that the model would also recognize compounds with short or without alkyl substituents like for example the initial hit molecule NCI35676. Compound **c24** in the TLR2 binding site and the 3D pharmacophore generated (mod_c24) are shown in Figure 42.

The model generated with compound **75** (mod_75) is shown in Figure 42 as well. It was derived from the docking pose generated and discussed in chapter 4.1.1.5. It comprises the two H-bond acceptor interactions with the nitrogens of Leu350 and Phe349, the H-bond donor with the carbonyl-oxygen of Leu350, π-interaction with Phe325 and three hydrophobic areas, representing the interactions formed by the chlorine, the methyl-group and the fluorine.

For both models, exclusion volumes were added on residues of the binding site in order to define the volume of potential ligands. The feature tolerance and the amount and size of the exclusion volumes were further optimized by iteratively screening the set of actives, inactives and decoys and modifying the pharmacophore features until the highest possible specificity was reached.
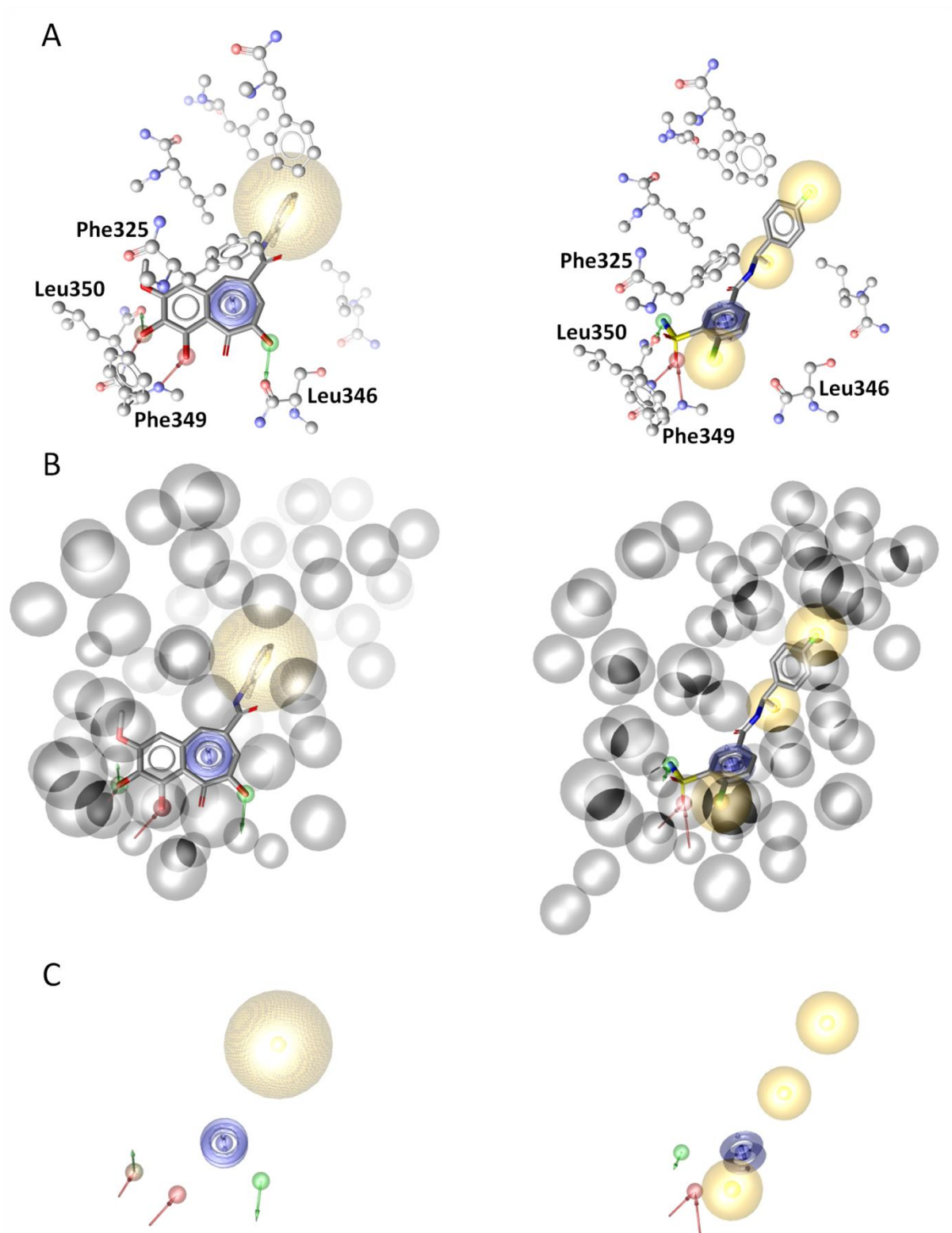
Figure 42: Binding mode of c24 (left) and compounds 75 (right) derived structure-based pharmacophores. A) Binding mode of the compounds in the TLR2 binding site. B) Ligand with 3D pharmacophore with exclusion volumes. C) 3D Pharmacophore (left - mod_c24, right – mod_75)

To analyze the remaining 16 compounds, a ligand-based approach was followed. First, the compounds were clustered with respect to their pharmacophoric features in LigandScout 3.1. This resulted in a total of nine clusters. One of them was integrated by the compounds with phenylurea scaffold **5**, **22**, **27** and NCI522131. A merged feature pharmacophore was generated with these molecules using the automatic pharmacophore generation protocol in LigandScout (Figure 43). A total of ten models were generated. One model was selected for further optimization, due to its specificity and the quality of the compound alignment (mod_pu). It comprised two π-interaction features for the two phenyl rings, two hydrophobic contact areas at the rings and one further for a hydrophobic substituent, one H-bond acceptor for the carbonyl-oxygen and one directed H-bond donor for the two nitrogens of the urea moieties as mandatory core features. In addition, two optional hydrophobic contact areas for potential substituents of the phenyl-moieties and one optional H-bond acceptor were included. Additionally, exclusion volume spheres were generated to sterically constrain the 3D environment around the best alignment of the ligands.

Another cluster was formed by compound **44** and its derivative **66**. A pharmacophore (mod_44_66) was generated for this cluster as is shown in Figure 43. Two H-bond acceptors and one π-interaction were included to describe the pyrimidine or pyrazin ring, respectively. The methylphenyl or phenyl groups are represented by a hydrophobic contact area and a π-interaction. Furthermore, the pharmacophore includes an H-bond acceptor (carbonyl-oxygen), two H-bond donors (amine nitrogens of the urea- or carbamoyl-moeity), hydrophobic contact area (thiophen or benzamide) and an optional H-bond donor (dimethylamin-group of compound **66**).
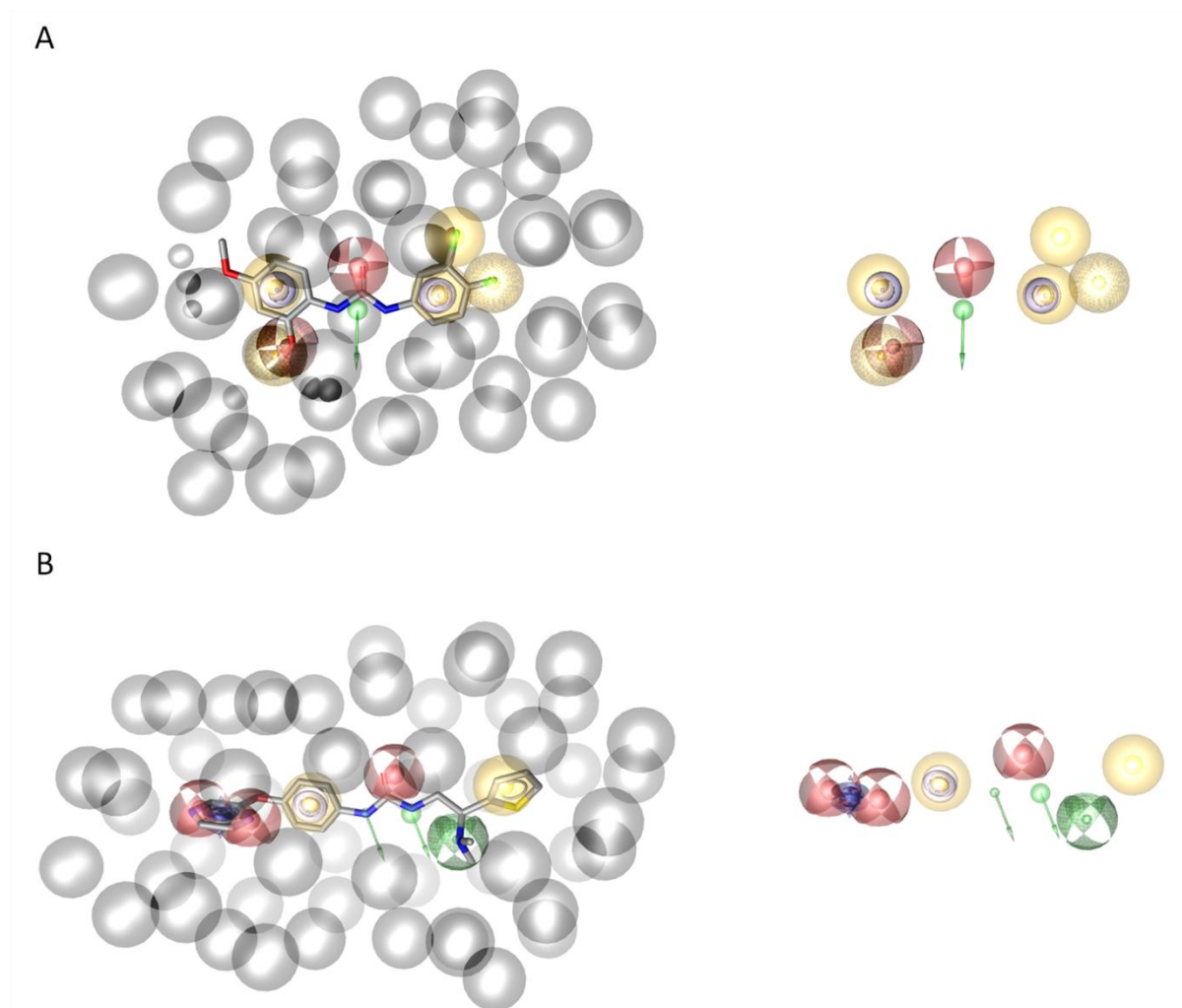
Figure 43: Ligand-based pharmacophore for A) the compounds with phenylurea scaffold and B) compounds 44 and 66. The models with exclusion volumes and aligned ligand are shown on the left, and the 3D pharmacophores alone on the right.

For the two models, the feature tolerance and the size and amount of exclusion volumes were optimized through iterative screening of the validation data set as described above for the structure-based models. With the remaining 14 compounds, no satisfactory pharmacophores in terms of selectivity could be generated. This was due to the fact that the molecules are chemically too distant to each other and do not contain enough common features. Besides, the compounds retrieved from the high-throughput screening have not been validated to act on TLR2 but only to inhibit its signaling pathway in a phenotypic assay.

The four generated 3D pharmacophores were used to screen the previously generated active (45 compounds), inactive (65 compounds) and decoy (1303 compounds) sets in order to

assess the specificity and selectivity of the model collection. The pharmacophore collection retrieved a total of 31 hits. 24 actives, 3 inactives and 4 decoys were found through this screening. An overview on the number of actives, inactives and decoys found for each model is given in Table 5. Our pharmacophore collection retrieves 54.55 % of the actives and 0.51 % of the decoys (including inactives). The specificity of the model is reflected by the AUC_{100} value of 0.77. pAUC values of 1.0 are retrieved at 1 %, 5 % and 10 %. At 1 %, 5 % and 10 % the model collection achieves enrichment factors of 25.2, 24.8 and 24.8 respectively (Figure 44).

Table 5: Overview on the hits found with the generated 3D pharmacophores.

Pharmacophore	Actives	Inactives	Decoys	Total hits
mod_c24	17	0	4	21
mod_75	1	1	0	2
mod_pu	5	2	0	7
mod_44_66	2	0	0	2
Total	24[a]	3	4	31

[a]Compound 44 is found both by mod_44_66 and mod_pu, because of this, the total amount of actives found is 24, even though the sum of the retrieved compounds for the single models is 25
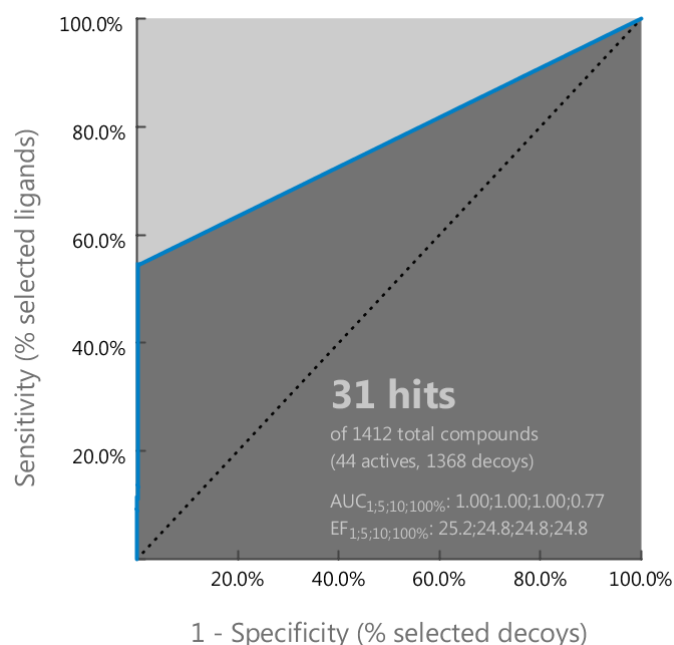

Figure 44: ROCS plot for the integrative 3D pharmacophore collection for TLR2 antagonists. AUC and EF are displayed. The screening was performed against the validation data set.

5 Discussion

5.1 Virtual screening

The main goal of this research project was the discovery of novel small molecule TLR2 antagonists through virtual screening. TLR2 has only recently begun to be studied as a potential drug target. Hence, the available information for drug design at the start of the study was scarce. At that point, two small molecule agonists of TLR2 and one signaling inhibitor with unspecified target protein were the only known small molecule modulators. Even though, crystal structures of TLR2 dimerized with TLR1 or TLR6 and the respective lipopeptides were available, no information on a small molecule binding site had been reported. In order to gain maximal information from the available data for virtual screening, a combined strategy workflow was implemented. It combined solely structure-based virtual screening through MIF derived 3D pharmacophores, and ligand-based virtual screening through shape- and feature-based search. In chapters 4.1 to 4.2, the different steps of the workflow and its results were described. In the following chapter 5.1.1, our virtual screening campaign will be discussed in terms of yields and activities. Then, our study will be compared to the aforementioned work by Cheng *et al.* [60], focusing on the high-throughput part of the publication (chapter 5.1.2).

5.1.1 Virtual screening evaluation: Yields and activities

As mentioned before, the search for small molecule TLR2 antagonists was performed through a workflow that integrated a structure- and a two-stepped ligand-based screening with subsequent biological characterization of the identified antagonists. An overview on the screening workflow is given in Figure 45. For the structure-based study two 3D pharmacophores (mod_1 and mod_2) of potential TLR2 ligands were derived from previously calculated MIFs of the putative small molecule binding site, which had previously been defined by a careful cavity detection procedure. The 3D pharmacophores were then used to screen a commercial compound collection containing about 3 million structures. Five virtual hits per model were selected for biological testing. The performed TLR2 specific NF-κB reporter assays led to one biologically validated TLR2 antagonist, which corresponds to a hit rate of 10 %. This is a highly satisfactory result for a prospective structure-based screening without a known or ideally co-crystallized small molecule antagonist and validates the defined

putative binding site and binding mode.

In the first step of the ligand-based screening, two compounds that had been published as TLR2 agonists [58] and one compound described as a TLR2 signaling inhibitor in literature [59] were used as query structures in a shape- and feature-based screening. This method was selected, as it allows searching for analogs of a single compound in a highly efficient manner (chapter 3.1.5). The query was performed against the compound collection provided by the National Cancer Institute (NCI) [128] which comprises ~200.000 small molecules. Five out of 39 biologically tested compounds could be confirmed as TLR2 antagonists (hit rate 13 %). In the second ligand-based screening, two of the query structures employed before, the two most promising hits from the first search and an additional TLR2 signaling inhibitor from literature [72] were used as query structures to select 26 compounds out of the aforementioned commercial compound collection with approximately 3 million compounds. Four of the chosen candidates for biological testing resulted to be active in a biological assay, corresponding to a hit rate of 15 % for the second part of the ligand-based search. Besides, four TLR2 agonists could be identified. In total, 75 compounds were selected for testing in a biological assay; out of these, ten could be confirmed as TLR2 antagonists and 22 further compounds as TLR2 signaling inhibitors with unspecified target. This results in an overall hit rate (for TLR2 antagonists) of 13 % for TLR2. Three of the identified TLR2 antagonists show a phenylurea scaffold whereas the remaining compounds are highly diverse in terms of their chemical structure. The diversity is remarkable and boosts the value of the identified hit molecules, as they can be independently used as lead structures to be further optimized in terms of binding affinity and pharmacokinetic as well as pharmacodynamic properties.
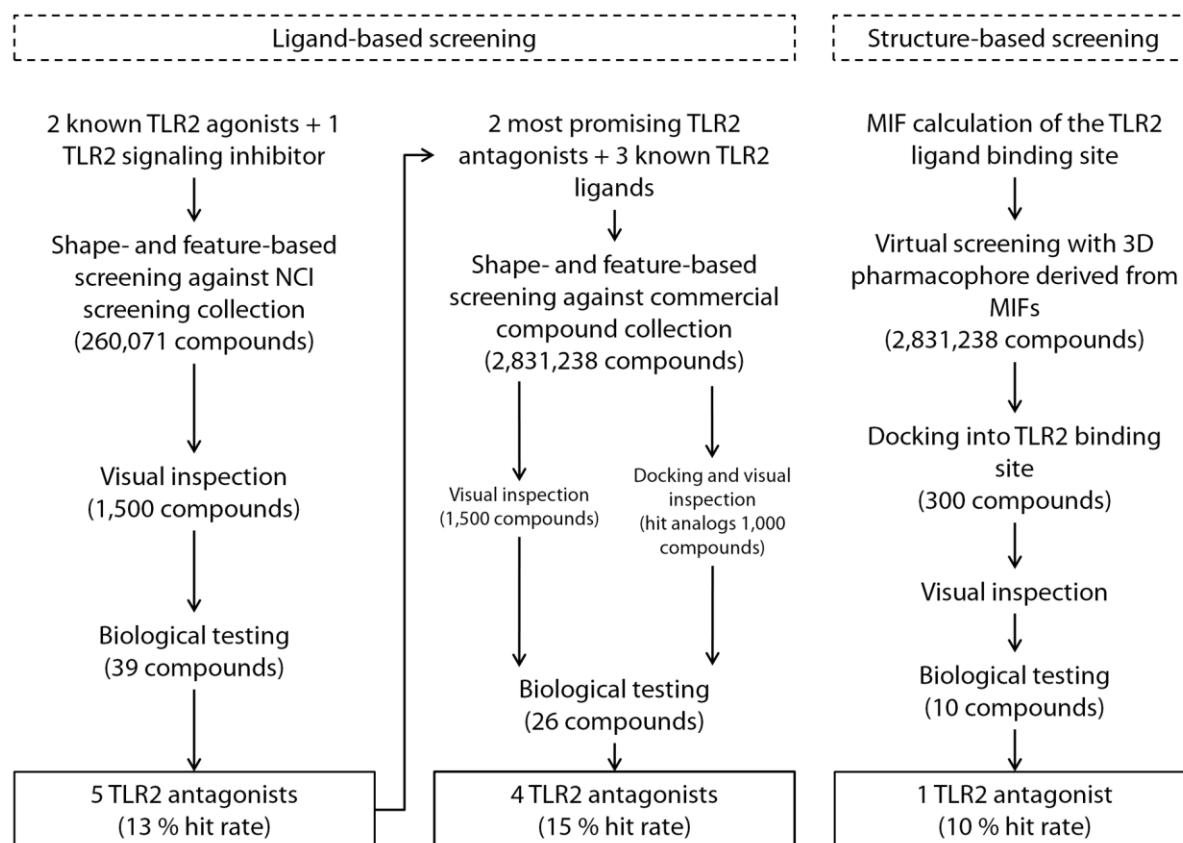
Figure 45: Overview of the virtual screening workflow.

The biological characterization of the identified TLR2 antagonists was performed through two main assays. First, relative inhibition was assessed in a NF-κB reporter assay. Then, ELISA assays with human monocytes were performed to determine the amount of released TNF-α production. It is noteworthy that the inhibitory values obtained in both assays highly diverge from each other. This can be attributed to the differences between the two employed cell systems. HEK293-TLR2 cells used in the NF-κB reporter assays are different from the human monocytes, the cells employed in the second assay. Thus, the incubation with the compounds could have a different impact on their TLR2 response. Furthermore, the expression levels of the TLR2 and its co-receptors can be different in both cellular systems, leading to a different signal.

In the biological characterization, three of the discovered antagonists (**5**, **22**, and **27**) showed IC_{50} values in the low micromolar range (<5 μM) against TLR2/1, TLR2/6 and TLR2/1 signaling respectively. Three other compounds showed values around 10 μM (**40**, **75** and **77**). In general, the activity range of the compounds is quite promising considering their low

molecular weight and the fact that they are primary hits, leaving space for further chemical optimization.

The determined IC_{50} values of some compounds (**5** and **27**) seem to reflect certain selectivity towards one of the two co-receptors. In chapter 4.3.2, a potential structural explanation for this specificity was developed and will be discussed in chapter 5.2. Nevertheless, the observed preference for one of the two receptors could also be caused by the experimental method. This way, the divergence between the TLR2/1 and the TLR2/6 inhibition could be caused by the different activating potencies of the lipopeptide TLR2 agonists employed for stimulation of the monocytes in the assay. Pam_3CSK_4 is applied in a concentration of 200 ng/mL (132 nM) and Pam_2CSK_4 at 50 ng/mL (16 nM) to obtain a similar response in terms TNF-α production. This makes it difficult to compare the inhibitory values obtained. Another reason for the different levels of inhibition could be the aforementioned varying expression level of the co-receptors (TLR1 or TLR6) in the cellular systems [136]. In order to clarify the preference of certain compounds for one of the co-receptors, a more detailed biological validation through additional assays, especially binding assays, will be necessary.

5.1.2 Virtual screening evaluation: Comparison to a recent experimental high-throughput screening

As mentioned before, a recent study describing the selection of TLR2/1 antagonists by high-throughput screening and further lead optimization was published by Cheng and co-workers [60]. A phenotypic assay that used the inhibition of nitric oxide (NO) production in mouse macrophages as a read-out was used to screen the *NCI-2 diversity library* that comprises 1363 small molecules. Through this assay, nine initial biologically active TLR2 antagonists could be discovered. Given the fact that this was the first successful screening described for TLR2 antagonists we compared it to our study in terms of activity, yields and chemical scaffolds. The conclusions we draw from this analysis will be elucidated next. The analysis of the binding mode of the chemically optimized compound will be discussed in chapter 5.3.

The nine initial hits of the study of Cheng *et al*. were described to reduce TLR2 signaling in terms of NO-production in murine macrophages at least 70 % at a compound concentration of 3 µM. The IC_{50} of our compounds determined measuring the TNF-α production in human

monocytes range from $3.33 \pm 1.14 \, \mu M$ to $48.18 \pm 1.34 \, \mu M$. This corresponds to a similar potency range, considering the fact that inhibitory constants derived from different assays can vary to a wide extent. Furthermore, the measurement of NO production instead of TNF-α production seems to be more sensitive, as the most active compound presented in the study CU-CPT22 was described to show an IC_{50} value of $0.58 \pm 0.09 \, \mu M$ but to only inhibit TNF-α production by 60 % at a concentration of $8 \, \mu M$ [60]. Nonetheless, to be able to compare the potency of all recently discovered TLR2 antagonists in a reliable manner, measurements under the same assay conditions would have to be performed.

As the high-throughput screening published by Cheng and coworkers was performed against the *NCI-2 diversity collection*, which is a subset of the NCI collection we screened in our first ligand-based virtual screening, we were interested in finding out whether we eventually had identified the same or similar TLR2 antagonists. None of the nine hits presented in the study was found through our screening. Still, two of the compounds (NCI201634 and NCI522131) are very similar to compounds we identified (**5**, **22** and **27**) as they contain a phenylthioureido- and phenylureido-scaffold respectively as previously mentioned in chapter 4.5.1. This is a further proof of the efficiency and accuracy of our screening and testing method and a validation of this chemical scaffold for the design of small molecule TLR2 antagonists.

Another interesting aspect that can be compared between the two studies are the achieved hit rates. An overview on the obtained values is given in Table 6. In their high-throughput screening campaign, the authors tested the whole *NCI-2 diversity library* comprising 1363 compounds in a phenotypic assay. This resulted in nine initial hits, hence a hit rate of 0.7 %. Yet, we performed a combined virtual screening approach against around three million compounds in order to select 75 compounds for biological testing. Of these, ten could be confirmed as biologically active TLR2 antagonists, this corresponds to a hit rate of 13 %. This yield is ~19 times higher than the one achieved by Cheng and coworkers and clearly shows that virtual screening makes lead discovery more goal-oriented and efficient in terms of time and materials.

Table 6: Comparison of hit rates of experimental high-throughput screening and virtual screening

Screening Technology	High-throughput screening by Cheng *et al*. [60]	Virtual screening
Biologically tested compounds	1363	75
Active hits	9	10
Hit rate	0.7 %	13 %

5.2 Evaluation of the modeled binding modes derived by docking

In chapters 4.3.1 to 4.3.3, potential binding modes for the discovered TLR2 modulators were derived by docking. As no binding assays or mutational studies were performed, the binding poses could not be validated experimentally. They were derived by careful analysis of the protein and compound structure and the available information on the mechanism of TLR2 signaling. However, a different modulator mechanism as for example allosterism or binding to the intracellular domain cannot be entirely ruled out. In this chapter, we attempt to discuss and classify the reliability of the hypothesized binding modes in order to give a better understanding of their quality and potential usage for further modeling studies or design of experiments.

At the start of our study, we hypothesized that small molecules with antagonistic activity on TLR2 would have to bind into the TLR2 lipopeptide binding site. In this cavity, we identified four smaller sub-pockets of which the three closer to the cavity opening were merged into the "front" binding region (chapter 4.1.1.1). The structure-based screening was performed with focus on this segment and effectively led to one TLR2 ligand with antagonistic activity, which validates the hypothesis. The compound was re-docked into the TLR2 binding site leading to a binding mode that explained its activity and the reason for the inactivity of other virtual hits found with the same model. This, and the fact that the compound binding pattern of the compound was directly derived from this part of the structure, makes its binding pose quite reliable.

For some of the discovered antagonist, plausible binding poses in this region of the pocket could also be identified and in some cases (**62** and **78**) explain the observed activity

differences in structurally related compounds, making their binding poses more reliable. Nevertheless, for other compounds, which are less similar to other tested molecules (**40**, **44**, **66** and **77**), this was not possible and makes it more difficult to validate the hypothesized binding poses.

In the case of the compounds with a phenylurea-scaffold, no plausible binding mode, which could explain the observed SAR could be found in the front region of the pocket. Hence, the docking studies were extended to the back part of the binding site. There, a sub-pocket into which the compounds fitted well and which explained the SAR of the molecules could be identified.

With the same compounds, a further docking study into the TLR1 binding sites was performed in order to elucidate an observed preference for one of the two co-receptor sub-types. However, with regard to eventual compound selectivity towards TLR2/1 signaling it is important to keep the method, through which the inhibitory constants are determined, and its limitations in mind. The antagonistic activity is measured in a cell-based assay in which TLR2 signaling is triggered by the lipopeptide agonists Pam_3CSK_4 and Pam_2CSK_4, which show different potencies. Furthermore, the expression levels of the receptors in the cells are not known. These factors can influence the measured activities and thus lead to unreliable, deceiving selectivity results. More extensive biological characterization of the antagonists will be necessary to confirm the ligand preferences and their action mode. This increases the uncertainty of the binding poses.

The same is true for the binding mode described for the discovered TLR2 agonists. Even though the rationale for the binding mode was derived from the recently released crystal structure of TLR8 and a small molecule agonist, the binding mode could be very different and structural re-arrangement after ligand binding could occur. Defining interacting residues by mutational studies would be very helpful to confirm the binding pose and activation mechanism.

In order to rank the reliability of the achieved binding modes, we classified them into categories from A to D, ranging from A – good reliability to D – very low reliability. An overview on the classification of the different docking poses is given in Table 7. The binding

mode of the compound identified by structure-based screening was the only one given a score of A. The binding poses of the molecules with phenyl-urea scaffold in TLR2 were classified into category B, the docking solutions of the remaining antagonists in the TLR2 binding site into category C. Finally, the binding modes of compounds with agonistic activity and of the potentially selective antagonists were rated with D. This gives an overview on the credibility of the binding poses and is especially useful for eventual further modeling and compound optimization steps in order to give the structural information the right emphasis. Like this, binding information of category A or B could be employed for structure-based screening, though in the later case the combination with ligand-based data would be recommendable. In both categories C and D more emphasis should be put on ligand-based screening employing the surmised binding modes as a selection criterion to identify molecules that could validate the hypothesis. Nevertheless, all binding modes should be checked by further, more in-depth experimental characterization through measurements of the binding affinity or mutational experiments.

Table 7: Classification of the binding modes into reliability categories and potential usage.

Reliability Category	Equivalent	Docking poses	Potential usage
A	good reliability	**75**	Structure-based modeling
B	moderate reliability	**5**, **22** and **27**	Combined ligand- and structure-based modeling
C	low reliability	**40**, **44**, **62**, **66**, **77** and **78**	Combined ligand- and structure-based modeling
D	very low reliability	**47 – 49**, **61** agonists, **5**, **22** and **27** co-receptor selective	Hypothesis generation

5.3 Binding mode of TLR2 antagonists with benzotropolone scaffold

In chapter 4.4.1 the description of our extensive docking study with the newly identified TLR2 antagonists with benzotropolone scaffold by Cheng *et al.* [60] was described. First, docking experiments to reproduce the binding mode proposed by the authors of the study were performed, but did not lead to any satisfying binding pose. This was in accordance with our expectations, because as described in chapter 4.4 the rationale employed for pose selection by the authors of the study is partially inconsistent. Due to this fact, a refined docking study was performed in order to elucidate a potential binding mode of the compounds that would better explain the SAR of the chemical series.

The first and crucial step in docking is the selection of the crystal structure. As the activities of the benzotropolones were measured in murine cells, we decided to use the murine crystal structure of TLR2/6 and not the human structure employed in our previous re-docking experiment. As we hypothesized that binding to TLR2 alone would be sufficient for inhibitory activity, we only used the TLR2 monomer for docking. In order to avoid bias in the docking experiment, the whole lipopeptide binding site was defined as ligand binding site. Through a procedure that ranked the generated docking poses by the number of pharmacophoric interactions with the protein and the accordance of the binding mode of different benzotropolone derivatives, a binding mode could be generated that fully explained the SAR published by Cheng *et al*. Furthermore, the interactions formed between the benzotropolone derivatives and TLR2 are in good accordance with the key interactions identified in chapter 3.1.2 through the analysis of MIFs. Like this, the compounds form H-bonds to Leu350, Phe349 and Ser346. Speaking in terms of the docking evaluation categories introduced in chapter 5.2, this gives the binding hypothesis "good reliability" (reliability category A).

5.4 Integrative model for TLR2 small molecule antagonists

At the start of our study, no small molecule antagonists of TLR2 were known. On the one hand, our structure-based modeling studies were based on the crystal structure of the receptor; on the other hand, our ligand-based screening was performed starting from two small molecule agonists of TLR2 and two TLR2 signaling inhibitors with unspecified target protein. The fact that we could successfully identify ten TLR2 antagonists and Cheng *et al*. identified

a total of 35 compounds with the same biological effect during the time of our study represents a significant increase in available data. As a consequence, we decided to refine our TLR2 antagonism model integrating the new information.

First, the available data was integrated into a set of 45 active and 65 inactive molecules. The actives were TLR2 antagonists discovered in our screening or by Cheng *et al.*; the inactives included all compounds biologically tested in our screening that did not show TLR2 antagonism including TLR2 agonists and TLR2 signaling inhibitors. Furthermore we generated a set of 1303 decoys from the ZINC database. These three datasets are extremely valuable for model optimization and validation. In an optimal virtual screening experiment, the model would be able to select all TLR2 antagonists and none of the decoys or inactives.

As the available TLR2 antagonists are highly diverse and belong to different chemotypes and are likely to possess a different interaction mode to the receptor, we decided to generate a collection of 3D pharmacophores instead of one single model. Two structure-based models were generated for **75** and the benzotropolone derivatives. In agreement with the reliability considerations presented in chapter 5.2, only the pharmacophores for these compounds were generated based on the binding information from the previous docking studies. Two additional ligand-based pharmacophores were generated representing the compounds with phenylurea scaffold and compound **44** and **66**. The remaining compounds were found to be to distant from each other in terms of pharmacophoric features, so that the generation of meaningful models for these compounds was not possible, as a different binding mode had to be assumed taking into account that they had been characterized as TLR2 signaling inhibitors and not antagonists.

In order to assess the selectivity and sensitivity of the model collection, the validation dataset of actives (45), inactives (65) and decoys (1303) was screened. Ideally, the validation of the model would have occurred on a different dataset than the model optimization. This is not possible in the case of TLR2 antagonists due to the small amount of available data. 54.55 % of the actives and only 0.51 % of the decoys were retrieved by the model. This shows a good sensitivity (the model finds a high amount of actives) and a very high specificity (only very few false positives were found). We were still interested in inspecting the remaining actives that were not found by the model. Basically, three categories of compounds could be

identified: i) compounds with very different pharmacophoric features that could not be clustered for model generation, ii) benzotropolone derivatives with methoxy-substituents that were in the set of actives even though they were much less active than the other compounds of this series, and iii) benzotropolone derivatives with very long aliphatic chains that would interfere with the exclusion volumes. The specificity of the model is reflected by the AUC_{100} value of 0.77. The EF at 1 %, 5 % and 10 % the model collection retrieves enrichment factors of 25.2, 24.8 and 24.8 respectively. This clearly shows that the model is able to discriminate active from decoy molecules very well and would be a useful tool for future virtual screening.

6 Conclusion & outlook

TLR2 is a critical signaling molecule in various severe clinical conditions such as sepsis and rheumatoid arthritis, and has been proposed as a promising drug-target for the treatment of these diseases. In this study, our aim was the identification of small molecule TLR2 antagonists. As no small molecule TLR2 antagonists were known at the start of our project, and only limited information on the TLR2 signaling mechanism was available, we decided to develop a combined ligand- and structure-based virtual screening workflow.

In this study, we successfully developed and applied a virtual screening workflow for the identification of TLR2 antagonists. Our work combines structure- and ligand-based screening: In the structure-based part, molecular interaction fields of a previously identified putative binding site were calculated, which led to the development of two 3D pharmacophores. These were then used for virtual screening and gave rise to the discovery of one new TLR2 antagonist. This result validates the hypothesized binding pocket and our 3D pharmacophore. In the ligand-based part of our work a two stepped shape- and feature-based screening was performed leading to the discovery of several promising antagonists of TLR2 with previously unknown scaffolds. In total, 10 out of 75 tested compounds showed antagonistic activity (13 % hit rate). In addition, four TLR2 agonists were identified. These compounds can now be further optimized and used for lead development or as tool compounds to elucidate TLR2 signaling.

In a further effort to understand the mechanism of TLR2 modulation by small molecules, agonists and antagonists were structurally analyzed to discover chemical features that cause the reversed activity. Furthermore, docking studies were performed with the identified modulators and with a series of TLR2 antagonists published during the course of this work in order to predict their binding mode to the receptor.

In a final step, we integrated the TLR2 antagonists identified in our study, the newly released molecules and the gained knowledge on the structural features necessary for TLR2 antagonism. This was performed through the analysis of the molecular structures and the generation of a 3D pharmacophore collection, reflecting the current knowledge on TLR2 antagonism through small molecules. The model was validated and refined and can now be

used for further virtual screening and the discovery and optimization of further promising TLR2 antagonist lead structures.

7   Experimental section

7.1   Virtual screening

7.1.1   Screening libraries

For the first ligand-based virtual screening the compound library provided by the National Cancer Institute [128] comprising 260,071 structures was utilized. The structure-based as well as the second ligand-based screening were performed against an in-house library of commercially available compounds from different vendors: Asinex [137-139], Enamine HTS Collection [140], Specs [141], Life Chemicals [142], Chembridge [67] and Maybridge [143]. In total, the library comprised 2,831,238 molecules. An overview of the amount of molecules from each vendor is given in Table 8.

Table 8: Overview of the extended compound collection of commercially available structures.

Database	Compounds
Asinex	436,012
LifeChemicals	344,693
Maybridge	55,810
Chembridge	503,802
Enamine HTS Collection	1,286,559
Specs	204,362
Total	2,831,238

In order to prepare the libraries for virtual screening the protonation states of the compounds were calculated and small molecular fragments and salts removed using MOE 2010.10 [85]. For pharmacophore-based screening proprietary database files that serve as an input for virtual screening in LigandScout 3.1. [119] were generated using the command-line tool idbgen. A maximum of 25 conformations per molecule were generated using OMEGA with FAST parameters [144]. For the ligand-based virtual screening conformations of the compounds were calculated using Omega [130], Openeye's conformer model generator and then stored in OEBinary v2-file format. This serves as input for ROCS calculations.

7.1.2   Protein preparation for structure-based virtual screening

For structure-based modeling the crystal structure of the TLR2/1 dimer with Pam_3CSK_4 (PDB code: 2Z7X [20]) from the Protein Data Bank [145] was used. The TLR1 monomer and all ligands apart from Pam_3CSK_4 and water molecules were removed. For docking, the lipopeptide was removed as well. The protonation state of the protein was calculated with the "Protonate 3D" application included in MOE 2010.10 suite [85].

7.1.3   Cavity detection

A binding site analysis of the TLR2 monomer was performed after removing TLR1. The two cavity detection programs Q-SiteFinder [79] and Site Finder, MOE 2010.10 [85] were employed. Only pockets where both algorithms coincided were taken into account for further steps. The search was concentrated on the lipopeptide binding site and on residues that form part of the protein dimerization interface. Two main cavities were found; one is the lipopeptide binding site, the other one is on the dimerization interface. The binding site on the dimerization interface was discarded as it was small and shallow. The big lipopeptide binding site was sub-divided into four sub-pockets; further analysis was focused on the area located close to the aperture of the binding site.

7.1.4   Molecular interaction fields

Molecular interaction fields (MIFs) of the TLR2 binding site were calculated in order to identify potential interaction hotspots. The analysis was performed using MOE 2010.10. The interaction potential for a probe group and the protein were calculated and energetically favorable areas were determined and displayed graphically. The following probes were selected to calculate the molecular interaction fields: "DRY" – hydrophobic moieties, "N1+" – H-bond donors with a positive charge and "O-" – acceptors with a single negative charge. Through manual adjustment of the interaction potentials a precise mapping of the binding site was provided, relating areas with favorable interaction potential to specific protein residues.

7.1.5   3D pharmacophore-based virtual screening

The program LigandScout 3.1. [119] was used to generate 3D pharmacophores. The

calculated molecular interaction fields were used to define key interactions necessary for ligand binding, as no crystal structure of TLR2 with co-crystallized small molecule binder was available. 3D pharmacophores defining the binding hypothesis derived from the mapping of the binding site were generated and refined by optimizing the size of the tolerance spheres and by the addition of exclusion volumes to define the binding site volume. Retrospective optimization and validation of the models based on a set of known active molecules was not possible, as no TLR2 antagonists had been published until then. The models were evaluated qualitatively and optimized by screening different small compound libraries and checking how well the virtual hits fitted the TLR2 binding site after minimization. Finally, the two 3D pharmacophores (mod_1 and mod_2) discussed in chapter 4.1.1.2 were selected for virtual screening.

Virtual screening was performed in LigandScout 3.1. against the collection of commercially available compounds described in chapter 7.1.1 using standard settings. To assess whether virtual hits actually fulfilled the previously defined interaction pattern with TLR2 a docking step was included. The 150 structures with the best pharmacophores fit scores that were found to be the most promising were docked into the TLR2 binding site, after calculating their 3D structures using Corina [131]. Ten docking poses per molecule were generated using GOLD Suite v5.1. [120] with standard parameters and GoldScore [94] as a scoring function. The docking poses were then ranked by the docking score and minimized in the binding site under an MMFF94 force field in LigandScout 3.1. In a last step compounds that fulfilled the interaction pattern that had been defined and fitted the binding site well were selected through visual inspection. For this purpose, the pharmacophoric interactions were visualized in LigandScout 3.1. and the distances as well as the angles of the interactions were evaluated to find compounds with a good interaction pattern.

7.1.6 Shape- and feature-based screening

The shape- and feature-based screening was performed using the Openeye software ROCS v3.1.2. [129] described in chapter 3.1.5. Default settings were used for screening. Previously, a biologically relevant conformation of the query structure had been calculated with Corina. 500 hits per query structure were retrieved. Out of these, candidates for biological validation were selected based on their similarity to the query structure and variability between each

other. The first shape- and feature-based screening was performed against the NCI database using three molecules form literature as query structures. The second screening was performed against the compound library of commercially available molecules described in chapter 7.1.1, two hit molecules from the previous screening and three molecules from literature were used as query structures. For the selected compounds an additional docking step was included into the second screening, using the two hit molecules from the first search as query structure. The 500 best compounds from the shape- and feature-based screening were docked into the TLR2 binding site using GOLD Suite v5.1 and GoldScore. The docking poses were inspected with regard to the key interactions defined in the structure-based part of the work and compounds that resembled the query structures and fitted the binding site well were selected for biological testing.

7.1.7 NF-κB Luciferase reporter gene assays

The NF-κB Luciferase reporter gene assays were performed by our cooperation partner at the Medicinal University of Innsbruck. HEK293-TLR2 cells were seeded into 96-well tissue culture plates (4×10^4 cells/well) in DMEM plus 10 % FBS and 0.5 % Ciprofloxacin (Sigma Aldrich). After 24 hours, the cells were transiently transfected with 25 ng of Elam.luc-pCDNA (reporter plasmid for NF-κB activity) and 25 ng of Renilla (NF-κB independent reporter plasmid) using Trans-IT LT1 transfection reagent (Mirus) in a ratio 3:1 (reagent:DNA). After overnight incubation, the cells were incubated for 1 h with the indicated compounds at 200 µM (initial screening) or with increasing concentrations (dose-response curves). Control wells without compounds were incubated with DMSO as the stock solutions of the compounds had been prepared in this solvent at a concentration of 10 mM. The stimulation was started by the addition of Pam_2CSK_4 (50 ng/mL) or Pam_3CSK_4 (200 ng/mL) (Invivogen). 5 hours later, the cells were lysed in reporter passive lysis buffer (Promega). As a last step, a luminometer plate-reader (Victor X, Perkin Elmer) was utilized to measure gene activity. The luciferase activity was normalized to the Renilla values and the relative inhibition determined.

7.1.8 Human-TNF-α-ELISA in human monocytes

A further step in the biological characterization of the identified TLR2 antagonists were TNF-α-ELISAs in human monocytes, which were also performed by our cooperation partner Dr.

Santos-Sierra at the Medicinal University of Innsbruck. PBMC's (peripheral blood mononuclear cells) were isolated from blood taken from healthy donors using Histopaque-1077 (Sigma Aldrich) according to the manufacturer's protocol. In the following step, 10^5 cells/well were seeded in a 96-well plate with DMEM plus 10 % FBS and 0.5 % Ciprofloxacin. After overnight incubation, the adherent cells were pre-stimulated with the different compounds or as a control with DMSO for one hour. Subsequently, stimulation with Pam_2CSK_4 (50 ng/mL), Pam_3CSK_4 (200 ng/mL) or medium took place for 18 h. TNF-α levels were measured by ELISA (Biolegend) according to the supplier's protocol in the collected supernatants. The levels of TNF-α obtained after stimulation with Pam_2CSK_4 or Pam_3CSK_4 were 1000-1250 pg/ml and the amount detected in unstimulated samples was lower than 15 pg/ml. The calculation of the IC_{50} values was done with Graphpad Prism software using a non-linear regression fit of dose-response curves with a Hill slope of -1.

7.1.9 Purity checks

A purity check of the active compounds was performed in order to assess their quality. HPLC separation was chosen as analysis method. 100 µL of a 10 mM solution of the compounds in DMSO was dissolved with 900 µL of MeOH (HPLC grade) first. Then 100 µL of this solution were further dissolved with 400 µL of MeOH and used for measurements. The runs were performed for 15 min under isocratic conditions with 80 % Acetonitril and 20 % H_2O. A vial with only DMSO was used as a control. Compounds with more than one HPLC peak were defined as impure.

7.2 Docking studies with discovered TLR2 antagonists

7.2.1 General protein preparation and docking settings

All docking experiments were performed using the software GOLD Suite v5.1 [120]. Default settings were applied and the GoldScore [94] used as scoring function. Ligand conformations were generated using Corina previous to docking [131]. The protein structures were prepared in MOE 2010.10 [85] by removing all ligands and water molecules and assigning the right protonation state to the residues of the macromolecule. The docking poses were visually analyzed in LigandScout 3.1. after previous MMFF94 [146] energy minimization of the ligand in the protein binding site. Therefore, the pharmacophoric interactions of the

compound with the macromolecule were visualized in order to assess the quality of the binding pose.

7.2.2 Docking studies with biologically validated TLR2 antagonists

In order to elucidate the binding mode of the discovered TLR2 antagonists docking studies were performed. The crystal structure of TLR2 with TLR1 and the lipopeptide ligand Pam$_3$CSK$_4$ (PDB code: 2Z7X [20]) was used for the docking studies, after removal of TLR1. All ligands were docked into the TLR2 binding site. The interactions between the molecule and the cavity residues were examined. Poses containing the key interactions that had been determined through the molecular interaction fields were prioritized as well as those that were similar to the binding mode of the co-crystallized ligand.

7.2.3 Docking with compounds with potential co-receptor selectivity

Some of the TLR2 antagonists discovered through our virtual screening show a much higher inhibitory activity towards one of the co-receptors. We hypothesized that these compounds could bind to the co-receptor in addition to binding to TLR2 and thus performed additional docking studies into TLR1 for TLR1 specific antagonists.

For docking into TLR1, the crystal structure of TLR2 with TLR1 and the lipopeptide ligand Pam$_3$CSK$_4$ (PDB code: 2Z7X [20]) was used. TLR2, all ligands and water molecules were removed. Cavity detection was performed as described for TLR2 using the two algorithms Q-Sitefinder [79] and Site Finder, MOE 2010.10. The TLR1 lipopeptide binding site was found to be the most suitable for ligand binding. The whole cavity was defined as binding site for docking, as it is smaller than the TLR2 pocket.

7.2.4 Docking of discovered small molecule TLR2 agonists

For the docking of the discovered small molecule agonists the complete dimer of TLR2 with TLR1 and Pam$_3$CSK$_4$ was used (PDB code: 2Z7X). Cavity detection was performed as described above. It led to two cavities at the interface of the protein; one of them was discarded as it was too shallow to fit the identified agonists. The other pocket was used for docking. All identified TLR2 agonists were docked into this pocket.

7.3   TLR2 antagonists with benzotropolone scaffold

7.3.1   Generation of ligand conformations

Before the actual docking was performed, 3D conformations of the compounds had to be generated. The ring geometries generated with CORINA [131] and through minimization in an MMFF94 [146] force field diverged from each other. The first conformation showed a planar ring, whereas the second was characterized by a bent ring system with unfavorable angles for sp2 hybridized carbon atoms. Therefore, we decided to perform a more thorough optimization of the ring conformation by quantum mechanics using Gaussian09 [147]. The resulting conformation was characterized by a fully planar ring system and was used for all docking studies. The input file was generated using antechamber. In the optimized structure the ring system is planar; hence the conformation generated by Corina for the compounds was used for docking experiments.

7.3.2   Docking of TLR2 antagonists with benzotropolone scaffold

The compounds were first docked into the dimer of TLR2 with TLR1 (PDB code: 2Z7X) in order to reproduce the binding mode proposed in the original publication of the compounds [60], after removal of the lipopeptide, all other ligands and water molecules. To avoid bias the whole dimer lipopeptide binding site (in TLR2 and TLR1) was defined as binding site for docking. In order to reproduce the binding pose constraints had to be defined (distance to TLR1 channel and H-bond with Phe349).

In the second part of the docking study to elucidate the binding mode of the compounds, the murine structure of TLR2/6 with Pam_2CSK_4 was utilized (PDB code: 3A79 [21]). The TLR6 monomer was removed and the macromolecule protonated. 20 docking poses per molecule were generated using GOLD Suite v5.1 [120] and the GoldScore [94]. The docking poses were visualized and further analyzed using LigandScout 3.1. [119]. The "interaction feature count", which represents the number of all occurring chemical features like H-bond acceptor, -donor, hydrophobic contact areas was calculated and used to rank the binding poses of the docked molecules. Docking poses in which the interactions formed by the compounds selected for docking coincided were preferred. In addition MIFs of the binding site with a phenolic-hydroxyl- ("OH") and a hydrophobic ("DRY") probe were calculated and the

overlay of the molecular structures and the MIFs assessed.

7.4   Integrative model of TLR2 antagonism through small molecules

7.4.1   Validation datasets

In order to validate and refine the generated model a set of actives, inactives and decoys was generated. The compounds discovered by Cheng *et al*. [60] and the nine TLR2 antagonists identified in our screening were integrated into a list of 45 actives. The compounds tested during our screening that did not display TLR2 antagonism (including TLR2 agonists and downstream inhibitors) were defined as the inactive set that comprised 65 compounds. Finally a set of 1303 decoys was generated using an in house KNIME workflow (MyDecoyFinder implemented by Dr. Susanne Dupré) that selects decoys from the ZINC DB based on the following similarity criteria:

- Molecular weight ± 25 Da
- Mannhold LogP ± 1.0
- Number of H-bond donors ± 1
- Number of H-bond acceptors ± 2
- Number of rotable bonds ± 1
- Tanimoto coefficient < 0.75

The datasets were generated in smiles format. Subsequently, proprietary database files that serve as an input for virtual screening in LigandScout 3.1 [119] were generated using the command-line tool idbgen. Standard parameters (OMEGA FAST, maximally 25 conformations per molecule) were employed for database generation.

7.4.2   Generation of structure-based 3D pharmacophores

Structure-based 3D pharmacophores of compound **75** and **c24** were generated based on the selected docking pose for each compound using LigandScout 3.1. [119]. Pharmacophoric features that were considered less important were removed and exclusion volumes placed on the pocket residues in order to define the volume of potential ligands. Through iterative screening of the active, inactive and decoy set the feature tolerances were adjusted so that a maximal amount of actives and a minimal amount of inactives or decoys were found.

7.4.3 Generation of ligand-based 3D pharmacophores

Ligand-based pharmacophores were generated for the remaining compounds. For this purpose, the compounds were first clustered by pharmacophoric features using LigandScout 3.1 with a maximum of 4 conformations and a cluster distance of 0.45. Merged feature pharmacophores were generated allowing two omitted features and selecting pharmacophore fit and atom overlap as scoring function. Ten models per cluster were generated and the one that distinguished actives from decoys and inactives best was selected for further optimization. This was performed by adjusting the feature tolerances and adding exclusion volumes in an iterative procedure of screening the validation dataset and model modification.

7.4.4 Model validation

The 3D pharmacophore collection was validated by screening the set of actives, inactives and decoys in order to evaluate its prediction power in the LigandScout 3.1 [119] screening mode. In order to assess the early recovery of actives the EF at 1 %, 5 % and 10 % was calculated. The screening results were represented in a ROC plot and the AUCs at 1 %, 5 %, 10 % and 100 % calculated.

8   Summary

Toll-like receptors (TLRs) play a pivotal role in the onset of innate immunity against invading microbial pathogens through the recognition of "pathogen associated molecular patterns" (PAMPs) [10]. Recently, they have been identified as potential drug-targets due to their role in the advent of various severe clinical conditions such as cancer, rheumatoid arthritis and pathogen sepsis [39]. Accordingly, the main goal of this work was to discover small molecule TLR2 antagonists by virtual screening and molecular modeling.

For this purpose, we developed and performed a combined structure- and ligand-based virtual screening workflow. In the structure-based part of the study, we first identified a putative small molecule binding site of which we then calculated molecular interaction fields (MIFs). These were used to identify key interactions necessary for ligand binding that were later transferred into 3D pharmacophores. Subsequently, the generated models were employed to perform virtual screening of a commercial library comprising about 3 million compounds. Through consecutive docking and visual inspection the generated virtual hits were prioritized and compounds selected for biological validation. This led to one out of ten biologically active TLR2 antagonists.

In the ligand-based part of the study, a two-stepped shape- and feature-based search was performed. The first screening was performed using two known TLR2 agonists and one TLR2 signaling inhibitor with unknown target protein as query structures to search the compound collection provided by the NCI. In the second search, the two TLR2 agonists were employed as query structures. Furthermore, one TLR2 signaling inhibitor and the two most promising hits from the first virtual screening were included as query structures. Virtual screening steps were followed by experimental validation of the selected compounds. In total, ten TLR2 antagonists were identified by ligand-based virtual screening.

Overall, 75 compounds were selected through virtual screening and biologically tested in an NF-κB reporter assay. In total, ten compounds showed antagonistic activity in this assay and were further characterized with regard to their effect on cytokine production in human monocytes. All compounds showed IC_{50} values in the micromolar range, three of them in the low micromolar range ($<5~\mu M$). In addition, four TLR2 agonists could be identified. In a next

step, the binding modes of the discovered TLR2 modulators were elucidated by docking studies. All discovered compounds were structurally analyzed to elucidate potential activity cliffs, necessary for agonistic or antagonistic activity.

Furthermore, we analyzed a series of TLR2 antagonists that were published during the course of this study. The binding mode of a series of benzotropolone derivatives could consistently be determined by docking. In the last part of our study, we integrated these compounds and the antagonists discovered in our study into a 3D pharmacophore model collection including all currently available information on TLR2 antagonism. The models were optimized and validated and can now be used to discover further TLR2 antagonists through virtual screening.

9 Zusammenfassung

Toll-like Rezeptoren (TLRs) spielen eine entscheidende Rolle bei der Aktivierung der angeborenen Immunantwort gegenüber mikrobiellen Krankheitserregern. Dies geschieht durch die Erkennung sogenannter PAMPs (*pathogen associated molecular patterns*), dies sind molekulare Strukturen, die ausschließlich in Pathogenen vorkommen [10]. Aufgrund ihrer Rolle bei der Enstehung verschiedener Krankheiten wie Krebs, rheumatoider Arthritis und Sepsis sind TLRs kürzlich als Target für die Arzneistoffentwicklung etabliert worden [47]. Dementsprechend war das Hauptziel dieser Arbeit die Entdeckung von TLR2 Antagonisten durch virtuelles Screening und Molecular Modelling.

Zu diesem Zweck wurde ein kombiniertes struktur- und liganden-basiertes virtuelles Screening durchgeführt. Im strukturbasierten Teil der Studie wurde als erstes eine putative Bindestelle für kleine Moleküle identifiziert, von der anschließend *molecular interaction fields* (MIFs) berechnet wurden. Hierdurch konnten für die Bindung von Liganden notwendige Interaktionen erkannt und mittels 3D Pharmakophoren modelliert werden. Im anschließenden Schritt wurde ein virtuelles Screening einer Datenbank mit ca. 3 Millionen kommerziell verfügbaren Verbindungen durchgeführt. Virtuelle Hits wurden durch Protein-Liganden-Docking und visuelle Inspektion priorisiert und zehn Moleküle für die biologische Testung ausgewählt. Dies führte zu einem biologisch aktiven TLR2 Antagonisten mit bisher unbekannter chemischer Struktur.

Im ligandenbasierten Teil der Studie wurde ein zweistufiges virtuelles Screening durch sterische Überlagerung durchgeführt. Im ersten Teil wurden zwei bekannte TLR2 Agonisten und ein Inhibitor der von TLR2 induzierten Signalkaskade als Suchstrukturen verwendet. Im zweiten Teil wurden neben den beiden bekannten TLR2 Agonisten ein weiterer Inhibitor der Signalkaskade und die zwei vielversprechendsten Hits aus dem ersten Teil benutzt. Die durch virtuelles Screening ausgewählten Strukturen wurden biologisch validiert. Insgesamt führte der ligandenbasierte Teil der Arbeit zur Entdeckung zehn bisher unbekannter, biologisch aktiver TLR2 Antagonisten.

Insgesamt wurden 75 Verbindungen durch virtuelles Screening ausgewählt und biologisch in einem NF-κB-Reporter-Assay getestet. Zehn davon zeigten antagonistische Aktivität und

wurden daraufhin in Hinblick auf ihre Wirkung auf die Produktion von Zytokinen in humanen Monozyten untersucht und ihre IC_{50}-Werte bestimmt. Alle Verbindungen zeigten Aktivität im mikromolaren Bereich, drei unter 5 µM. Darüber hinaus konnten vier TLR2 Agonisten entdeckt werden. Im folgenden Schritt wurde der Bindungsmodus der identifizierten TLR2 Modulatoren durch Dockingstudien aufgeklärt. Außerdem wurden die Verbindungen strukturell analysiert, um so die chemischen Merkmale zu identifizieren, die Agonisten von Antagonisten unterscheiden.

Desweiteren analysierten wir eine Reihe von TLR2 Antagonisten, die im Verlauf dieser Arbeit anderweitig veröffentlicht wurden. Der Bindungsmodus einer Reihe von Benzotropolon-Derivanten wurde durch molekulares Docking bestimmt. Im letzten Teil unserer Studie wurde eine Sammlung von 3D Pharmakophoren generiert, die die gesamten derzeit verfügbaren Informationen über TLR2 Antagonisten integriert. Die Modelle wurden optimiert und validiert und können nun dazu verwendet werden, durch virtuelles Screening weitere TLR2 Antagonisten zu entdecken.

10 References

1. Anderson, K.V., L. Bokla, and C. Nusslein-Volhard, *Establishment of dorsal-ventral polarity in the Drosophila embryo: the induction of polarity by the Toll gene product.* Cell, 1985. **42**(3): p. 791-8.

2. Anderson, K.V., G. Jurgens, and C. Nusslein-Volhard, *Establishment of dorsal-ventral polarity in the Drosophila embryo: genetic studies on the role of the Toll gene product.* Cell, 1985. **42**(3): p. 779-89.

3. Siegmund-Schultze, N., *Toll-like Rezeptoren: Neue Zielstruktur für immunstimulierende Medikamente.* Dtsch Arztebl, 2007. **104**(16): p. A 1072 - A 1073.

4. Lemaitre, B., et al., *The dorsoventral regulatory gene cassette spatzle/Toll/cactus controls the potent antifungal response in Drosophila adults.* Cell, 1996. **86**(6): p. 973-83.

5. Kawai, T. and S. Akira, *The role of pattern-recognition receptors in innate immunity: update on Toll-like receptors.* Nat Immunol, 2010. **11**(5): p. 373-84.

6. Uematsu, S. and S. Akira, *The role of Toll-like receptors in immune disorders.* Expert Opin Biol Ther, 2006. **6**(3): p. 203-14.

7. Henneke, P., et al., *Lipoproteins are critical TLR2 activating toxins in group B streptococcal sepsis.* J Immunol, 2008. **180**(9): p. 6149-58.

8. Rakoff-Nahoum, S. and R. Medzhitov, *Toll-like receptors and cancer.* Nat Rev Cancer, 2009. **9**(1): p. 57-63.

9. Palsson-McDermott, E.M. and L.A. O'Neill, *The potential of targeting Toll-like receptor 2 in autoimmune and inflammatory diseases.* Ir J Med Sci, 2007. **176**(4): p. 253-60.

10. Akira, S. and K. Takeda, *Toll-like receptor signalling.* Nat Rev Immunol, 2004. **4**(7): p. 499-511.

11. Janeway, C.A., Jr. and R. Medzhitov, *Innate immune recognition.* Annu Rev Immunol, 2002. **20**: p. 197-216.

12. Kawai, T. and S. Akira, *Toll-like receptors and their crosstalk with other innate receptors in infection and immunity.* Immunity, 2011. **34**(5): p. 637-50.

13. Matzinger, P., *Tolerance, danger, and the extended family.* Annu Rev Immunol, 1994. **12**: p. 991-1045.

14. Kono, H. and K.L. Rock, *How dying cells alert the immune system to danger.* Nat Rev Immunol, 2008. **8**(4): p. 279-89.

15. Bianchi, M.E., *DAMPs, PAMPs and alarmins: all we need to know about danger.* J Leukoc Biol, 2007. **81**(1): p. 1-5.

16. Piccinini, A.M. and K.S. Midwood, *DAMPening inflammation by modulating TLR signalling.* Mediators Inflamm, 2010. **2010**.

17. Aliprantis, A.O., et al., *Cell activation and apoptosis by bacterial lipoproteins through toll-like receptor-2.* Science, 1999. **285**(5428): p. 736-9.

18. Schwandner, R., et al., *Peptidoglycan- and lipoteichoic acid-induced cell activation is mediated by toll-like receptor 2.* J Biol Chem, 1999. **274**(25): p. 17406-9.

19. Underhill, D.M., et al., *The Toll-like receptor 2 is recruited to macrophage phagosomes and discriminates between pathogens.* Nature, 1999. **401**(6755): p. 811-5.

20. Jin, M.S., et al., *Crystal structure of the TLR1-TLR2 heterodimer induced by binding of a tri-acylated lipopeptide.* Cell, 2007. **130**(6): p. 1071-82.

21. Kang, J.Y., et al., *Recognition of lipopeptide patterns by Toll-like receptor 2-Toll-like*

receptor 6 heterodimer. Immunity, 2009. **31**(6): p. 873-84.

22. Guan, Y., et al., *Human TLRs 10 and 1 share common mechanisms of innate immune sensing but not signaling.* J Immunol, 2010. **184**(9): p. 5094-103.

23. Asea, A., et al., *Novel signal transduction pathway utilized by extracellular HSP70: role of toll-like receptor (TLR) 2 and TLR4.* J Biol Chem, 2002. **277**(17): p. 15028-34.

24. Scheibner, K.A., et al., *Hyaluronan fragments act as an endogenous danger signal by engaging TLR2.* J Immunol, 2006. **177**(2): p. 1272-81.

25. Botos, I., D.M. Segal, and D.R. Davies, *The structural biology of Toll-like receptors.* Structure, 2011. **19**(4): p. 447-59.

26. Manavalan, B., S. Basith, and S. Choi, *Similar Structures but Different Roles - An Updated Perspective on TLR Structures.* Front Physiol, 2011. **2**: p. 41.

27. Bell, J.K., et al., *Leucine-rich repeats and pathogen recognition in Toll-like receptors.* Trends Immunol, 2003. **24**(10): p. 528-33.

28. Xu, Y., et al., *Structural basis for signal transduction by the Toll/interleukin-1 receptor domains.* Nature, 2000. **408**(6808): p. 111-5.

29. O'Neill, L.A. and A.G. Bowie, *The family of five: TIR-domain-containing adaptors in Toll-like receptor signalling.* Nat Rev Immunol, 2007. **7**(5): p. 353-64.

30. Slack, J.L., et al., *Identification of two major sites in the type I interleukin-1 receptor cytoplasmic region responsible for coupling to pro-inflammatory signaling pathways.* J Biol Chem, 2000. **275**(7): p. 4670-8.

31. Nyman, T., et al., *The crystal structure of the human toll-like receptor 10 cytoplasmic domain reveals a putative signaling dimer.* J Biol Chem, 2008. **283**(18): p. 11861-5.

32. Khan, J.A., et al., *Crystal structure of the Toll/interleukin-1 receptor domain of human IL-1RAPL.* J Biol Chem, 2004. **279**(30): p. 31664-70.

33. Bella, J., et al., *The leucine-rich repeat structure.* Cell Mol Life Sci, 2008. **65**(15): p. 2307-33.

34. Liu, L., et al., *Structural basis of toll-like receptor 3 signaling with double-stranded RNA.* Science, 2008. **320**(5874): p. 379-81.

35. Park, B.S., et al., *The structural basis of lipopolysaccharide recognition by the TLR4-MD-2 complex.* Nature, 2009. **458**(7242): p. 1191-5.

36. Tanji, H., et al., *Structural reorganization of the Toll-like receptor 8 dimer induced by agonistic ligands.* Science, 2013. **339**(6126): p. 1426-9.

37. Kang, J.Y., et al., *Recognition of Lipopeptide Patterns by Toll-like Receptor 2-Toll-like Receptor 6 Heterodimer.* Immunity, 2009.

38. Akira, S., S. Uematsu, and O. Takeuchi, *Pathogen recognition and innate immunity.* Cell, 2006. **124**(4): p. 783-801.

39. Barbalat, R., et al., *Toll-like receptor 2 on inflammatory monocytes induces type I interferon in response to viral but not bacterial ligands.* Nat Immunol, 2009. **10**(11): p. 1200-7.

40. Rakoff-Nahoum, S. and R. Medzhitov, *Toll-like receptors and cancer.* Nature Reviews Cancer, 2009. **9**(1): p. 57-63.

41. Sriskandan, S. and D.M. Altmann, *The immunology of sepsis.* Journal of Pathology, 2008. **214**(2): p. 211-223.

42. Henneke, P., et al., *Lipoproteins are critical TLR2 activating toxins in group B streptococcal sepsis.* J Immunol, 2008. **180**(9): p. 6149-6158.

43. Rice, T.W., et al., *A randomized, double-blind, placebo-controlled trial of TAK-242 for the treatment of severe sepsis.* Critical Care Medicine, 2010. **38**(8): p. 1685-1694.

44. Tidswell, M., et al., *Phase 2 trial of eritoran tetrasodium (E5564), a Toll-like receptor 4 antagonist, in patients with severe sepsis.* Critical Care Medicine, 2010. **38**(1): p. 72-

83.

45. Opal, S.M., et al., *Effect of Eritoran, an Antagonist of MD2-TLR4, on Mortality in Patients With Severe Sepsis The ACCESS Randomized Trial.* Jama-Journal of the American Medical Association, 2013. **309**(11): p. 1154-1162.

46. Palsson-McDermott, E.M. and L.A.J. O'Neill, *The potential of targeting Toll-like receptor 2 in autoimmune and inflammatory diseases.* Ir J Med Sci, 2007. **176**(4): p. 253-260.

47. Connolly, D.J. and L.A. O'Neill, *New developments in Toll-like receptor targeted therapeutics.* Curr Opin Pharmacol, 2012. **12**(4): p. 510-8.

48. Seibl, R., et al., *Expression and regulation of Toll-like receptor 2 in rheumatoid arthritis synovium.* Am J Pathol, 2003. **162**(4): p. 1221-7.

49. Ultaigh, S.N., et al., *Blockade of Toll-like receptor 2 prevents spontaneous cytokine release from rheumatoid arthritis ex vivo synovial explant cultures.* Arthritis Res Ther, 2011. **13**(1): p. R33.

50. Lee, M.S., et al., *Role of TLR2 in the pathogenesis of autoimmune diabetes and its therapeutic implication.* Diabetes Metab Res Rev, 2011. **27**(8): p. 797-801.

51. Kim, D.H., et al., *Inhibition of autoimmune diabetes by TLR2 tolerance.* J Immunol, 2011. **187**(10): p. 5211-20.

52. Arslan, F., et al., *Myocardial ischemia/reperfusion injury is mediated by leukocytic toll-like receptor-2 and reduced by systemic administration of a novel anti-toll-like receptor-2 antibody.* Circulation, 2010. **121**(1): p. 80-90.

53. Arslan, F., et al., *Treatment with OPN-305, a humanized anti-Toll-Like receptor-2 antibody, reduces myocardial ischemia/reperfusion injury in pigs.* Circ Cardiovasc Interv, 2012. **5**(2): p. 279-87.

54. Garay, R.P., et al., *Cancer relapse under chemotherapy: Why TLR2/4 receptor agonists can help.* European Journal of Pharmacology, 2007. **563**(1-3): p. 1-17.

55. Morales, A., D. Eidinger, and A.W. Bruce, *Intracavitary Bacillus Calmette-Guerin in the treatment of superficial bladder tumors.* J Urol, 1976. **116**(2): p. 180-3.

56. Murata, M., *Activation of Toll-like receptor 2 by a novel preparation of cell wall skeleton from Mycobacterium bovis BCG Tokyo (SMP-105) sufficiently enhances immune responses against tumors.* Cancer Sci, 2008. **99**(7): p. 1435-40.

57. Wang, X., C. Smith, and H. Yin, *Targeting Toll-like receptors with small molecule agents.* Chem Soc Rev, 2013. **42**(12): p. 4859-66.

58. Guan, Y., et al., *Identification of novel synthetic toll-like receptor 2 agonists by high throughput screening.* J Biol Chem, 2010. **285**(31): p. 23755-62.

59. Zhou, S., et al., *Discovery of a novel TLR2 signaling inhibitor with anti-viral activity.* Antiviral Res, 2010. **87**(3): p. 295-306.

60. Cheng, K., et al., *Discovery of small-molecule inhibitors of the TLR1/TLR2 complex.* Angew Chem Int Ed Engl, 2012. **51**(49): p. 12246-9.

61. Buwitt-Beckmann, U., et al., *Lipopeptide structure determines TLR2 dependent cell activation level.* FEBS J, 2005. **272**(24): p. 6354-64.

62. Buwitt-Beckmann, U., et al., *TLR1- and TLR6-independent recognition of bacterial lipopeptides.* J Biol Chem, 2006. **281**(14): p. 9049-57.

63. Spohn, R., et al., *Synthetic lipopeptide adjuvants and Toll-like receptor 2--structure-activity relationships.* Vaccine, 2004. **22**(19): p. 2494-9.

64. Okusawa, T., et al., *Relationship between structures and biological activities of mycoplasmal diacylated lipopeptides and their recognition by toll-like receptors 2 and 6.* Infect Immun, 2004. **72**(3): p. 1657-65.

65. Agnihotri, G., et al., *Structure-activity relationships in toll-like receptor 2-agonists*

leading to simplified monoacyl lipopeptides. J Med Chem, 2011. **54**(23): p. 8148-60.

66. Salunke, D.B., et al., *Structure-Activity Relationships in Human Toll-like Receptor 2-Specific Monoacyl Lipopeptides.* J Med Chem, 2012. **55**(7): p. 3353-3363.

67. *Chembridge Screening Library*, 2013, Chembridge Corporation: San Diego, CA, USA.

68. Spyvee, M.R., et al., *Toll-like receptor 2 antagonists. Part 1: preliminary SAR investigation of novel synthetic phospholipids.* Bioorg Med Chem Lett, 2005. **15**(24): p. 5494-8.

69. Seyberth, T., et al., *Lipolanthionine peptides act as inhibitors of TLR2-mediated IL-8 secretion. Synthesis and structure-activity relationships.* J Med Chem, 2006. **49**(5): p. 1754-65.

70. Kalluri, M.D., et al., *Novel synthetic gluco-disaccharide RSCL-0409--a lipopolysaccharide-induced Toll-like receptor-mediated signalling antagonist.* FEBS J, 2010. **277**(7): p. 1639-52.

71. Hutchinson, M.R., et al., *Evidence that tricyclic small molecules may possess toll-like receptor and myeloid differentiation protein 2 activity.* Neuroscience, 2010. **168**(2): p. 551-63.

72. Liang, Q., et al., *Characterization of sparstolonin B, a Chinese herb-derived compound, as a selective Toll-like receptor antagonist with potent anti-inflammatory properties.* J Biol Chem, 2011. **286**(30): p. 26470-9.

73. McInnes, C., *Virtual screening strategies in drug discovery.* Curr Opin Chem Biol, 2007. **11**(5): p. 494-502.

74. Berman, H.M., et al., *The Protein Data Bank.* Acta Crystallogr D Biol Crystallogr, 2002. **58**(Pt 6 No 1): p. 899-907.

75. Murgueitio, M.S., et al., *In silico virtual screening approaches for anti-viral drug discovery.* Drug Discovery Today, 2012. **9**(3): p. e219-e225.

76. Leis, S., S. Schneider, and M. Zacharias, *In silico prediction of binding sites on proteins.* Curr Med Chem, 2010. **17**(15): p. 1550-62.

77. Laurie, A.T. and R.M. Jackson, *Methods for the prediction of protein-ligand binding sites for structure-based drug design and virtual ligand screening.* Curr Protein Pept Sci, 2006. **7**(5): p. 395-406.

78. Labute, P. and M. Santavy. *Locating Binding Sites in Protein Structures.* [cited 2013 17.Jun.]; Available from: http://www.chemcomp.com/journal/sitefind.htm.

79. Laurie, A.T. and R.M. Jackson, *Q-SiteFinder: an energy-based method for the prediction of protein-ligand binding sites.* Bioinformatics, 2005. **21**(9): p. 1908-16.

80. Goodford, P.J., *A computational procedure for determining energetically favorable binding sites on biologically important macromolecules.* J Med Chem, 1985. **28**(7): p. 849-57.

81. Cross, S., et al., *GRID-based three-dimensional pharmacophores I: FLAPpharm, a novel approach for pharmacophore elucidation.* J Chem Inf Model, 2012. **52**(10): p. 2587-98.

82. Cross, S., et al., *GRID-based three-dimensional pharmacophores II: PharmBench, a benchmark data set for evaluating pharmacophore elucidation methods.* J Chem Inf Model, 2012. **52**(10): p. 2599-608.

83. Broccatelli, F., et al., *A novel approach for predicting P-glycoprotein (ABCB1) inhibition using molecular interaction fields.* J Med Chem, 2011. **54**(6): p. 1740-51.

84. Cross, S. and G. Cruciani, *Molecular fields in drug discovery: getting old or reaching maturity?* Drug Discovery Today, 2010. **15**(1-2): p. 23-32.

85. *MOE, 2010.10*, 2010, Chemical Computing Group: Montreal, Quebec, Canada.

86. Kitchen, D.B., et al., *Docking and scoring in virtual screening for drug discovery:*

methods and applications. Nat Rev Drug Discov, 2004. **3**(11): p. 935-49.

87. Jones, G., et al., *Development and validation of a genetic algorithm for flexible docking.* Journal of Molecular Biology, 1997. **267**(3): p. 727-48.

88. Morris, G.M., et al., *Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function.* Journal of Computational Chemistry, 1998. **19**(14): p. 1639-1662.

89. Venkatachalam, C.M., et al., *LigandFit: a novel method for the shape-directed rapid docking of ligands to protein active sites.* Journal of Molecular Graphics & Modelling, 2003. **21**(4): p. 289-307.

90. Rarey, M., et al., *A fast flexible docking method using an incremental construction algorithm.* Journal of Molecular Biology, 1996. **261**(3): p. 470-489.

91. Ewing, T.J.A., et al., *DOCK 4.0: Search strategies for automated molecular docking of flexible molecule databases.* J Comput Aided Mol Des, 2001. **15**(5): p. 411-428.

92. McGann, M.R., et al., *Gaussian docking functions.* Biopolymers, 2003. **68**(1): p. 76-90.

93. Jain, A.N., *Surflex: Fully automatic flexible molecular docking using a molecular similarity-based search engine.* J Med Chem, 2003. **46**(4): p. 499-511.

94. Verdonk, M.L., et al., *Improved protein-ligand docking using GOLD.* Proteins, 2003. **52**(4): p. 609-23.

95. Meng, E.C., B.K. Shoichet, and i.D. Kuntz, *Automated Docking with Grid-Based Energy Evaluation.* J. Comput. Chem. , 1992. **13**(4): p. 505-524.

96. Gohlke, H., M. Hendlich, and G. Klebe, *Knowledge-based scoring function to predict protein-ligand interactions.* Journal of Molecular Biology, 2000. **295**(2): p. 337-56.

97. Huang, S.Y., S.Z. Grinter, and X. Zou, *Scoring functions and their evaluation methods for protein-ligand docking: recent advances and future directions.* Phys Chem Chem Phys, 2010. **12**(40): p. 12899-908.

98. Huang, N., et al., *Physics-based scoring of protein-ligand complexes: enrichment of known inhibitors in large-scale virtual screening.* J Chem Inf Model, 2006. **46**(1): p. 243-53.

99. Warren, G.L., et al., *A critical assessment of docking programs and scoring functions.* J Med Chem, 2006. **49**(20): p. 5912-5931.

100. Kirchmair, J., et al., *Development of anti-viral agents using molecular modeling and virtual screening techniques.* Infect Disord Drug Targets, 2011. **11**(1): p. 64-93.

101. Wermuth, C.G., et al., *Glossary of terms used in medicinal chemistry (IUPAC Recommendations 1997).* Annu. Rep. Med. Chem., 1998. **33**: p. 385-395.

102. Wermuth, C.G., *Pharmacophores: Historical Perspective and Viewpoint from a Medicinal Chemist,* in *Pharmacophores and Pharmacophore Searches,* T. Langer and R.D. Hoffmann, Editors. 2006, WILEY-VCH Verlag GmbH & Co.KGaA: Weinheim.

103. McGaughey, G.B., et al., *Comparison of topological, shape, and docking methods in virtual screening.* J Chem Inf Model, 2007. **47**(4): p. 1504-19.

104. Dixon, S.L., A.M. Smondyrev, and S.N. Rao, *PHASE: a novel approach to pharmacophore modeling and 3D database searching.* Chem Biol Drug Des, 2006. **67**(5): p. 370-2.

105. Dixon, S.L., et al., *PHASE: a new engine for pharmacophore perception, 3D QSAR model development, and 3D database screening: 1. Methodology and preliminary results.* J Comput Aided Mol Des, 2006. **20**(10-11): p. 647-71.

106. Guner, O., O. Clement, and Y. Kurogi, *Pharmacophore modeling and three dimensional database searching for drug design using catalyst: recent advances.* Curr Med Chem, 2004. **11**(22): p. 2991-3005.

107. Labute, P., et al., *Flexible alignment of small molecules.* J Med Chem, 2001. **44**(10): p. 1483-90.

108. Wolber, G. and T. Langer, *LigandScout: 3-D pharmacophores derived from protein-bound ligands and their use as virtual screening filters.* J Chem Inf Model, 2005. **45**(1): p. 160-9.

109. Wolber, G., A.A. Dornhofer, and T. Langer, *Efficient overlay of small organic molecules using 3D pharmacophores.* J Comput Aided Mol Des, 2006. **20**(12): p. 773-88.

110. Wolber, G. and T. Langer, *Pharmacophore definition and 3D searches.* Drug Discov Today, 2004. **1**(3): p. 203-207.

111. Seidel, T., et al., *Strategies for 3D pharmacophore-based virtual screening.* Drug Discov Today, 2010. **7**(4): p. e221-e228.

112. Kirchmair, J., et al., *How to optimize shape-based virtual screening: choosing the right query and including chemical information.* J Chem Inf Model, 2009. **49**(3): p. 678-92.

113. Grant, J.A., M.A. Gallardo, and B.T. Pickup, *A fast method of molecular shape comparison: A simple application of a Gaussian description of molecular shape.* Journal of Computational Chemistry, 1996. **17**(14): p. 1653-1666.

114. Willett, P., J.M. Barnard, and G.M. Downs, *Chemical similarity searching.* Journal of Chemical Information and Computer Sciences, 1998. **38**(6): p. 983-996.

115. Kirchmair, J., et al., *Evaluation of the performance of 3D virtual screening protocols: RMSD comparisons, enrichment assessments, and decoy selection--what can we learn from earlier mistakes?* J Comput Aided Mol Des, 2008. **22**(3-4): p. 213-28.

116. Triballeau, N., et al., *Virtual screening workflow development guided by the "receiver operating characteristic" curve approach. Application to high-throughput docking on metabotropic glutamate receptor subtype 4.* J Med Chem, 2005. **48**(7): p. 2534-47.

117. Kim, H.M., et al., *Crystal structure of the TLR4-MD-2 complex with bound endotoxin antagonist Eritoran.* Cell, 2007. **130**(5): p. 906-17.

118. Yoon, S.I., et al., *Structural basis of TLR5-flagellin recognition and signaling.* Science, 2012. **335**(6070): p. 859-64.

119. *LigandScout3.1.*, 2013, Inte:ligand: Vienna, Austria.

120. *GOLD Suite, v5.1*, 2011, Cambridge Crystallographic Data Centre: Cambridge, UK.

121. Halgren, T.A., *Merck molecular force field .1. Basis, form, scope, parameterization, and performance of MMFF94.* Journal of Computational Chemistry, 1996. **17**(5-6): p. 490-519.

122. Halgren, T.A., *Merck molecular force field .2. MMFF94 van der Waals and electrostatic parameters for intermolecular interactions.* Journal of Computational Chemistry, 1996. **17**(5-6): p. 520-552.

123. Halgren, T.A., *Merck molecular force field .3. Molecular geometries and vibrational frequencies for MMFF94.* Journal of Computational Chemistry, 1996. **17**(5-6): p. 553-586.

124. Halgren, T.A. and R.B. Nachbar, *Merck molecular force field .4. Conformational energies and geometries for MMFF94.* Journal of Computational Chemistry, 1996. **17**(5-6): p. 587-615.

125. Halgren, T.A., *Merck molecular force field .5. Extension of MMFF94 using experimental data, additional computational data, and empirical rules.* Journal of Computational Chemistry, 1996. **17**(5-6): p. 616-641.

126. Halgren, T.A., *MMFF VI. MMFF94s option for energy minimization studies.* Journal of Computational Chemistry, 1999. **20**(7): p. 720-729.

127. Halgren, T.A., *MMFF VII. Characterization of MMFF94, MMFF94s, and other*

widely available force fields for conformational energies and for intermolecular-interaction energies and geometries. Journal of Computational Chemistry, 1999. **20**(7): p. 730-748.

128. *NCI Open Database*, 2010, Developmental Therapeutics Program, Division of Cancer Treatment and Diagnosis, National Cancer Institute.

129. *ROCS, v3.1.2*, 2011, OpenEye Scientific Software: Santa Fe, New mexico, USA.

130. *OMEGA, v2.4.6.*, 2007, OpenEye Scientific Software: Santa Fe, New Mexico, USA.

131. *CORINA, 3.4.*, 2006, Molecular Networks GmbH: Erlangen, Germany.

132. Razgulin, A.V. and S. Mecozzi, *Binding properties of aromatic carbon-bound fluorine.* J Med Chem, 2006. **49**(26): p. 7902-6.

133. Peri, F. and M. Piazza, *Therapeutic targeting of innate immunity with Toll-like receptor 4 (TLR4) antagonists.* Biotechnol Adv, 2012. **30**(1): p. 251-60.

134. Huang, N., B.K. Shoichet, and J.J. Irwin, *Benchmarking sets for molecular docking.* J Med Chem, 2006. **49**(23): p. 6789-801.

135. Cereto-Massague, A., et al., *DecoyFinder: an easy-to-use python GUI application for building target-specific decoy sets.* Bioinformatics, 2012. **28**(12): p. 1661-2.

136. Zarember, K.A. and P.J. Godowski, *Tissue expression of human Toll-like receptors and differential regulation of Toll-like receptor mRNAs in leukocytes in response to microbes, their products, and cytokines.* J Immunol, 2002. **168**(2): p. 554-61.

137. *Asinex Platinum Collection*, 2013, Asinex Ltd: Moscow, Russia.

138. *Asinex Gold Collection*, 2013, Asinex Ltd.: Moscow, Russia.

139. *Asinex Synergy Library*, 2013, Asinex Ltd.: Moscow, Russia.

140. *Enamine HTS Collection* 2013, Enamine Ltd.: Kiev, Ukraine.

141. *Specs Screening Compounds*, 2013, Specs: Delft, The Netherlands.

142. *Life Chemical Compound Collection for HTS*, 2013, Life Chemical Inc.: Niagara-on-the-Lake, Ontario, Canada.

143. *Maybridge Screening Collection*, 2013, Thermo Fisher Scientific: Waltham, MA, USA.

144. Kirchmair, J., et al., *Comparative performance assessment of the conformational model generators omega and catalyst: a large-scale survey on the retrieval of protein-bound ligand conformations.* J Chem Inf Model, 2006. **46**(4): p. 1848-61.

145. *Protein Data Bank*, 2011, Research Collaboratory for Structural Biology: USA.

146. Halgren, T.A., *Characterization of MMFF94, MMFF94s, and other widely available force fields for conformational energies and for intermolecular-interaction energies and geometries.* Abstracts of Papers of the American Chemical Society, 1998. **216**: p. U702-U702.

147. *Gaussian09*, 2013, Gaussian, Inc.: Wallingford, Connecticut, USA.

Curriculum vitae

Der Lebenslauf ist in der Online-Version aus Gründen des Datenschutzes nicht enthalten.

Publications

Peer reviewed articles

[1]   von Eichborn J, Murgueitio MS, Dunkel M, Koerner S, Bourne PE, Preissner R.: PROMISCUOUS: a database for network based drug-repositioning, *Nucleic Acids Res.* 2011 jan; 39(Dababase issue):D 1060-6.

[2]   Ahmed J, Meinel T, Dunkel M, Murgueitio MS, Adams R, Blasse C, Eckert A, Preissner S, Preissner R, CancerResource: a comprehensive database of cancer-relevant proteins and compound interactions supported by experimental knowledge, *Nucleic Acids Res.* 2011 Jan;39(Database issue):D960-7.

[3]   Woelke AS, Murgueitio MS, Preissner R, Theoretical Modeling Techniques and Their Impact on Tumor Immunology, *Clin Dev Immunol*. 2010;2010:271794.

[4]   Woelke AL, von Eichborn J, Murgueitio MS, Worth CL, Castiglione F, Preissner R. Development of Immune Specific Interaction Potentials and Their Application in the Multi-Agent System VaccImm, PloS ONE 6:e23257 (2011)

[5]   Murgueitio MS, Bermudez M, Mortier J, Wolber G, In silico virtual screening approaches for anti-viral drug discovery. Drug Discovery Today: Technologies, 2012. **9**(3): p. e219-e225.

[6]   Murgueitio MS, Henneke, P, Glossmann, H, Santos-Sierra, S, Wolber, G. Prospective virtual screening in a sparse data scenario: Design of small molecule TLR2 antagonists. *Submitted.*

Conference Proceedings

[7]   Meinel T, Müller MS; Ahmed J, Yildirimman R, Dunkel M, Herwig R, Preissner R, SOAP/WSDL-based Web Services for Biomedicine: Demonstrating the Technique with the CancerResource, XII Mediterranean Conference on Medical and Biological Engeneering and Computing 2010 (IFMBE Proceedings), Chalkidiki, Greece. Berlin Heidelberg: Springer. 835-838.

[8]   Poster: von Eichborn J, Murgueitio MS, Preissner R, PROMISCUOUS: A resource to explore drug-target relationships, 25th German Conference in Bioinformatics 2010 Braunschweig, Germany

[9]   Oral Presentation: Murgueitio MS, Santos-Sierra S, Wolber G, Targeting the first

barrier in immune response by designing TLR antagonists: a shape based approach, Second Summer School on Medicinal Chemistry in Drug Discovery: The Pharma Perspective 2011 Madrid, Spain

[10]   Poster: Murgueitio MS, Santos-Sierra S, Wolber G, Discovery of novel TLR modulators by molecular modeling and virtual screening, 7th German Conference on Chemoinformatics 2011 Goslar, Germany

[11]   Oral Presentation: Murgueitio MS, Santos-Sierra S, Wolber G, Modulation of the first barrier in immune response by the design of TLR antagonists using computational methods, 243rd ACS National Meeting and Exposition 2012 San Diego, CA, USA

[12]   Poster: Murgueitio MS, Santos-Sierra S, Wolber G, Virtual screening for Toll-like receptor antagonists: a combined approach, 19th EuroQSAR 2012 Vienna, Austria

[13]   Oral Presentation: Murgueitio MS, Santos-Sierra S, Wolber G, M Design of novel TLR antagonists by molecular modeling, 245rd ACS National Meeting and Exposition 2013 New Orleans, LA, USA

[14]   Poster: Murgueitio MS, Santos-Sierra S, Wolber G, Insights into TLR2 antagonism by small molecules through molecular modeling, XXII National Meeting on Medicinal Chemistry 2013 Rome, Italy

Appendix

Appendix 1: Inactive compounds virtual screening
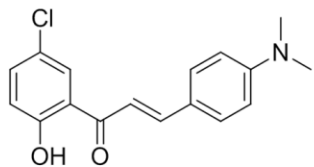
1
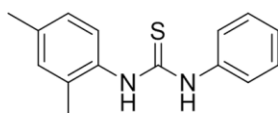
2
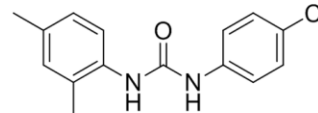
3
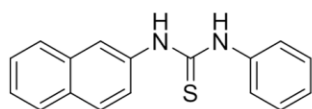
4
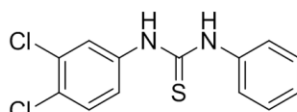
6
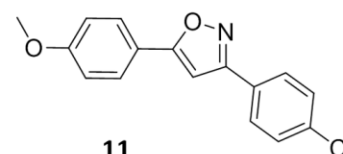
11
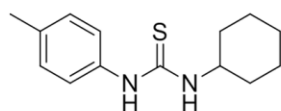
12
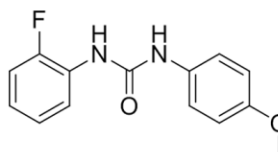
13
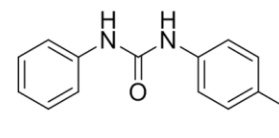
14
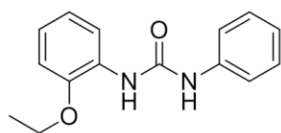
15
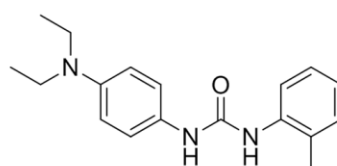
16
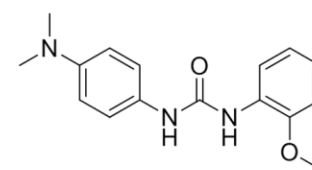
17
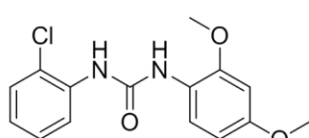
18
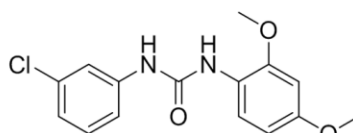
19

20

21

23

24

25

26

27

109
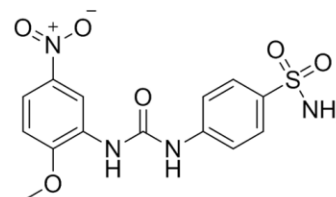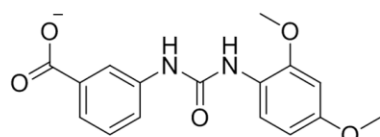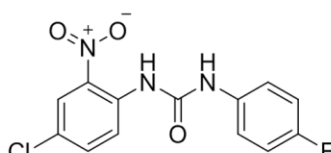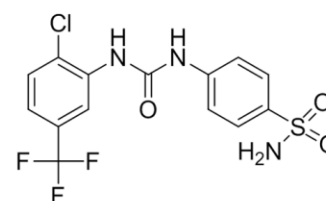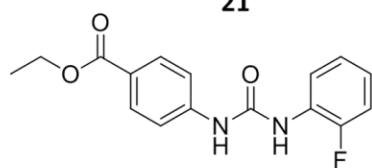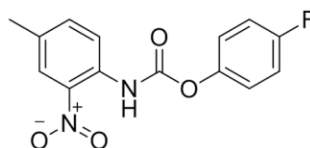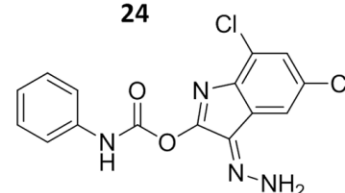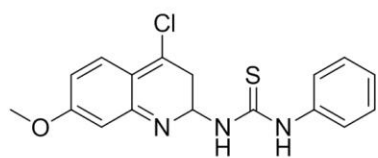
29
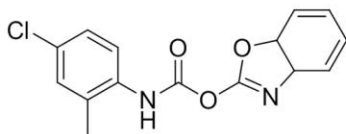
30
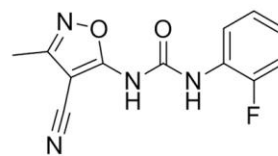
32
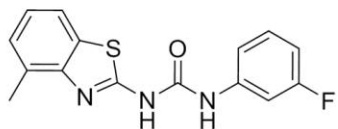
33
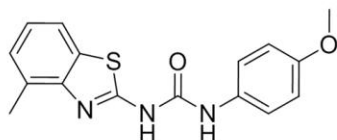
34
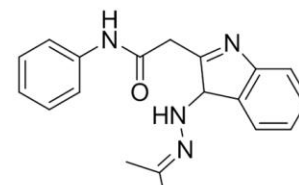
35
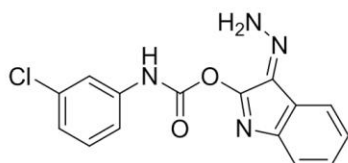
36
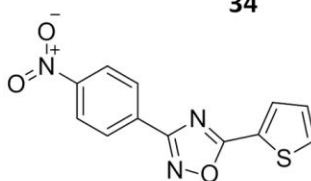
37
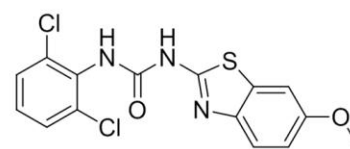
38
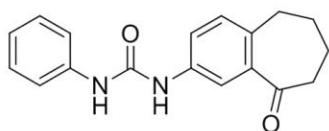
39
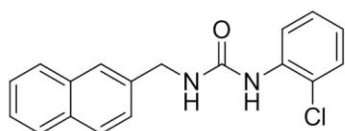
41

42

43

46

54

55

56

57

61

63

64

110

65

67

68

69

70

71

72

73

74

76

79

80

81

82

83

84

85

86

87

88

111