UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE FARMACIA
DEPARTAMENTO DE QUÍMICA ORGÁNICA Y FARMACEUTICA

TESIS DOCTORAL

Computational studies of Toll-like receptor 4

Estudios computacionales del receptor Toll-like 4

MEMORIA PARA OPTAR AL GRADO DE DOCTORA

PRESENTADA POR

Lucía Pérez Regidor

DIRECTORA

Sonsoles Martín Santamaría

Madrid, 2018

FACULTAD DE FARMACIA

DEPARTAMENTO DE QUÍMICA ORGÁNICA Y FARMACÉUTICA

DOCTORADO EN QUÍMICA MÉDICA

"ESTUDIOS COMPUTACIONALES DEL RECEPTOR TOLL-LIKE 4"

"COMPUTATIONAL STUDIES OF TOLL-LIKE RECEPTOR 4"

Memoria para optar al grado de Doctor presentada por

LUCÍA PÉREZ REGIDOR

Madrid, Julio 2017

Esta Tesis ha sido realizada en el Centro de Investigaciones Biológicas (CIB-CSIC), Madrid

DIRECTORA: Sonsoles Martín Santamaría

FACULTY OF PHARMACY

DEPARTMENT OF ORGANIC CHEMISTRY AND PHARMACEUTICAL

PhD Medicinal Chemistry

"COMPUTATIONAL STUDIES OF TOLL-LIKE RECEPTOR 4"

PhD Dissertation presented by

LUCÍA PÉREZ REGIDOR

Madrid, July 2017

This Thesis has been realized in Centro de Investigaciones Biológicas (CSIC), Madrid

SUPERVISOR: Sonsoles Martín Santamaría

FACULTAD DE FARMACIA

DEPARTAMENTO DE QUÍMICA ORGÁNICA Y FARMACÉUTICA

DOCTORADO EN QUÍMICA MÉDICA

"ESTUDIOS COMPUTACIONALES DEL RECEPTOR TOLL-LIKE 4"

Memoria para optar al grado de Doctor presentada por

LUCÍA PÉREZ REGIDOR

Madrid, Julio 2017

Esta Tesis ha sido realizada en el Centro de Investigaciones Biológicas (CSIC), Madrid

DIRECTORA: Sonsoles Martín Santamaría

To my parents and sister,

Acknowledgements

I would like to express my gratitude to all the people who have collaborated in this work, especially my supervisor, Dra. Sonsoles Martín-Santamaría, for giving me the opportunity to be part of her research group, but above all, for the guidance, monitoring and supervision, as well as for the motivation and for trusting me over the years.

I would like also to thank Prof. Nurisso, Prof. Peri and Prof. Walker for give me the opportunity to work in their laboratories, giving me a wonderful time and teaching me during my secondments.

Also I would like to thank, Dra. Marta Gutierrez, who introduce me to the research. Without her, I would probably not have decide to do a PhD.

I would like to thank the COST Action and the Spanish Ministry of Economy and Competiveness for the funds over these years (Refs. COST BM1003, CTQ 2011-22724, BES-2012-053653 and CTQ2014-57141-R).

In addition, I would like to thank my colleagues and friends from my laboratory: Alessandra, Javi, Joan, Jean-Marc and Laura, by their welcome, friendship and good advices.

Also my best friends: Chus, Patri, Laura, Rocio, that they have behaved with me like a sister, and they have helped me and accompanied me in many ways, motivated me with their joy, and their way of living life. Also my good friends from university: Fatima, Carmen, Maria, Angela and Nidia, thank you for trusting and supporting me all the time.

Of course also wanted to extend my gratitude and also to dedicate this thesis to my family, especially to my parents: mum, that has unfortunately not come to witness the moment, but I know that she is watching me wherever she is, and dad, and last but not least, my sister. Without their support, love and motivation this would not have been possible.

These final words I would like to dedicate especially to Ricardo, who has made my life better and happier in every way, supporting me and helping me in many difficult moments, but giving me the happiest and most wonderful moments of my life.

CONTENTS

List of abbreviations:

AD = Alzheimer's Disease

ALS = Amyotrophic Lateral Sclerosis

AMBER = Assisted Model Building with Energy Refinement

AnxA2 = Annexin A2

CD14 = Cluster of Differentiation 14

CHARMM = Chemistry at Harvard Molecular Mechanics

CHOL = Cholesterol

CMC = Critical Micellar Concentration

DAMPs = Damage-Associated Molecular Patterns

DFT = Density Functional Theory

DIC = diisopropylcarbodiimide

DIPEA = diisopropylethylamine

DOPC = 1,2-dipalmitoyl-sn-glycero-3- phosphocholine

DOPE = 1,2-Dioleoyl-sn-glycero-3-phosphoethanolamine

DPPE = 1,2-dipalmitoyl-sn-glycero-3-phosphoethanolamine

CNT = Carbon Nanotubes

DAMP = Damage-Associated Molecular Patterns

DFT = Density Functional Theory

DPPC = 1,2-dipalmitoyl-sn-glycero-3-phosphocholine

ED = Extracellular Domain

ESP = Electrostatic Potential

E55 = E5564 or Eritoran

FA = Fatty Acid

FBS = Fetal Bovine Serum

FF = Force Field

FLAP = Fingerprints for Ligands And Proteins

GAFF = General AMBER Force Field

GM = Gangliosides

HOBt = N-hydroxybenzotriazole

HTVS = High Throughput Virtual Screening

IAANS = Iodoacetylaminonaphtyl sulfate

ID = Intracellular Domain

IEF-3 = Interferon Regulatory Factor 3

ISM = Implicit Solvent Model

KBP = Knowledge-Based Pharmacophore

LBP = LPS-binding Protein

logP = logarithm of the Partition Coefficient

clogP = Calculated logP

LA = Lauroyl

LJ = Lennard-Jones

LOS = Lipooligosaccharide

LPB = Lipopolysaccharide Binding Protein

LPS = Lipopolysaccharide

LRR = Leucine-rich Repeat

Mal or TIRAP = Adaptor Protein Necessary to Recruit MyD88 to TLR2 and TLR4

MD = Molecular Dynamics

MD-2 = Myeloid Differentiation Factor 2

MIFs = Molecular Interaction Fields

MM = Molecular Mechanics

MMFF = Merck Molecular Force Field

MM/GBSA = Molecular Mechanics/Generalized Born Surface Area

MM/PBSA = Molecular Mechanics/Poisson-Boltzmann Surface Area

MPLA = Monophosphoryl Lipid A

MY = Myristoyl

MyD88 = Myeloid Differentiation Primary Response Gene (88)

NMA = Normal Mode Analysis

NMR = Nuclear Magnetic Resonance

NO = Nitric Oxide

NPM = N-pyrene maleimide

OL = Oleoyl

OPLS = Optimized Potentials for Liquid Simulations

PA = Palmitoyl

PAINS = Pan Assays Interference Compounds

PAMPs = Pathogen-Associated Molecular Patterns

PDB = Protein Data Bank

PC = Phosphatidylcholine

PMB = p-methoxybenzyl

pNPP = paranitrophenylphosphate

POPC = 1palmitoyl-2-oleoyl-sn-glycero-3-phosphocoline

POPE = 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine

PPI = Protein-Protein Interaction

PRRs = Pattern-Recognition Receptors

PS = Phosphatidylserine

PTX = Paclitaxel

QM = Quantum Mechanical

RMSD = Root Mean Square Deviation

RMSF = Root Mean Square Fluctuation

RsLA = *Rhodobacter sphaeroides* Lipid A

SA = Surface Area

sAP = secreted Alkaline Phosphatase

SAR = Structure-Activity Relationship

SASA = Solvent Accessible Surface Area

SFN = Sulforaphane

siRNA = small-interfering RNA

SM = Sphingomyelin

SP = Standard Precision

SP-A = Surfactant Protein A

ssRNA = single-stranded RNA

ST = Stearoyl

STD = Saturation Transfer Difference

ST2 or IL1RL1 = Interleukin 1 Receptor-Like 1

ST2L = Membrane-Bound Receptor form of ST2

TD = Transmembrane Domain

TIR = Toll-like/Interleukin-1 Receptor Domain

TLC = Thin-Layer Chromatography

TLRs = Toll-like Receptors

TLR2 = Toll-like receptor 2

TLR3 = Toll-like receptor 3

TLR4 = Toll-like receptor 4

TLR7 = Toll-like receptor 7

TLR8 = Toll-like receptor 8

TPSA = Topological Polar Surface Area

TRAM = TRIF-Related Adaptor Molecule

TRIF = TIR-domain-containing adapter-including interferon-α

VDW = Van der Waals

VINA = Vina is not AutoDock

VS = Virtual Screening

WT = Wild-Type

XN = Xanthohumol

XP = Extra Precision

ABSTRACT

COMPUTATIONAL STUDIES OF TOLL-LIKE RECEPTOR 4

This Thesis is focused on the molecular modeling and computational study of the molecular recognition processes involving Pattern Recognition Receptors (PRRs), in particular, Toll-like receptors (TLRs). TLRs are the main actors in innate immunity and are specialized in the recognition of pathogen associated molecular patterns (PAMPs).

In particular, TLR4 is located in the plasma membrane where, together with the MD-2 protein, it binds to lipopolysaccharides, membrane constituents of Gram-negative bacteria, forming a heterodimeric complex. TLR4 agonists can be used as adjuvants in vaccine development and in cancer immunotherapy. TLR4 antagonists have also been studied for their promising application in septic shock, chronic inflammation and autoimmunity. However, the mechanism at atomic level for such activation/inactivation process remains unknown. Our research has been focused on the study of the mechanism of the TLR4/MD-2 system by means of computational approaches.

In order to carry out our research objectives, we use a combination of several computational tools: geometry optimization, charges calculations, docking, virtual screening, and molecular dynamics simulations of protein complexes and membranes.

The main objective of this Thesis is to elucidate the ligand-protein interactions of TLR4 at atomic detail through computational techniques. Computational methodologies will be applied to the study of the molecular mechanisms involved in the TLRs functionality, and in the recognition of PAMPs. Ligand-protein docking and virtual screening will be used as a source of new compounds able to modulate the TLRs behavior with possible therapeutic applications, and also as biological probes.

CHAPTER 3: Theoretical binding modes will be predicted for reported modulators of the TLR4/MD-2 system with agonist and antagonist activity. In particular, we will focus our work in synthetic LPS analogues and small molecules with a non LPS-like structure. We will undertake a computational study of some representative compounds to unveil some of these patterns of interactions.

CHAPTER 4: The cationic glycolipid IAXO-102, a potent TLR4 antagonist, will be used as scaffold to design new potential TLR4 modulators and fluorescent labels for the TLR4 receptor complex. These compounds will be synthetized by collaborators. Our modelling studies will led us to the proposal of 3D models for the interaction with CD14 and TLR4/MD-2 accounting for their binding properties and also for their antagonistic activity.

CHAPTER 5: Virtual screening strategies from commercial, *in-house* and generic drugs libraries, followed by biological assays, will allow us to identify new chemical entities for the development of novel TLR4 modulators with a LPS non-related structure.

CHAPTER 6: The computational study of the full TLR4/MD-2 system will be performed, simulating the TLR4/MD-2 complex in the membrane environment. Several models of the monomer TLR4/MD-2 system inserted in different membranes will be built and simulated. These different models will be useful for the final building of the complete TLR4/MD-2 dimer and will provide us insights into the mechanism of TLR4 agonism/antagonism. The analysis of the molecular dynamics simulations will lead us to understand the key ligand-receptor and protein-protein interactions implicated in the molecular recognition events and the dimerization process.

Our work has led to the following conclusions:

Theoretical binding modes have been predicted for reported modulators of the TLR4/MD-2 system, with agonist and antagonist activity. In particular, we focused our work in synthetic glycolipids and non LPS-like molecules. For all these TLR4 modulators, it is clear that, despite their different chemical structure, they must share a common pattern of interactions with TLR4. We have undertaken a computational study of some representative compounds to unveil some of these patterns of interactions.

The cationic glycolipid IAXO-102, a potent TLR4 antagonist targeting both MD-2 and CD14 co-receptors, has been used as scaffold to design new potential TLR4 modulators and fluorescent labels for the TLR4 receptor complex (membrane TLR4/MD-2 dimer and CD14). Our modelling studies have led to the proposal of 3D models for the interaction with CD14 and TLR4/MD-2 accounting for their binding properties and also for their antagonistic activity.

To propose new chemical scaffolds for the development of new ligands able to modulate TLR4 functions, we have performed virtual screening. Virtual screening strategies from commercial and *in-house* libraries, followed by biological assays, have allowed us to identify new chemical entities for the development of novel TLR4 modulators with a LPS-non-related structure. So far, we have identified seven novel compounds with a promising TLR4 antagonist activity.

The computational study of the full TLR4/MD-2 heterodimer was performed, simulating the full complex inserted in the membrane environment. The analysis of the molecular dynamics simulations led us to understand the key interactions implicated in the dimerization process at atomic level.

Summarizing, molecular modelling approaches have been used to elucidate the molecular recognition mechanisms of TLR4/MD-2 modulation, with focus on the agonist/antagonist conformational changes of the TLR4/MD-2 system, and to provide some hints for the design of novel binders, hopefully with therapeutic potential. We also have collaborated with experimental groups to synthesize the designed compounds and to perform biological assays. The study of the binding mode of several reported TLR4/MD-2 modulators has been undertaken and the simulation of the TLR4/MD-2 system in different membranes environments, by means of a combination of docking calculations, molecular dynamics simulations and virtual screening protocols have been performed.

RESUMEN

ESTUDIOS COMPUTACIONALES DEL RECEPTOR TOLL-LIKE 4

En esta Tesis Doctoral se han empleado técnicas de modelado molecular y se ha llevado a cabo el estudio computacional de los procesos de reconocimiento molecular que implican Receptores de Reconocimiento de Patrones (PRRs), en particular, los receptores Toll-like (TLRs). Los TLRs son los principales actores en la inmunidad innata y se especializan en el reconocimiento de patrones moleculares asociados a patógenos (PAMPs).

En particular, el receptor TLR4 se localiza en la membrana plasmática donde, junto con la proteína MD-2, se une a lipopolisacáridos, constituyentes de membrana de bacterias Gram-negativas, que forman un complejo heterodimérico. Los agonistas de TLR4 pueden ser útiles como coadyuvantes en el desarrollo de la vacuna y en la inmunoterapia contra el cáncer. Los antagonistas de TLR4 también han sido estudiados por su prometedora aplicación en choque séptico, inflamación crónica y autoinmunidad. Sin embargo, el mecanismo a nivel atómico para tal proceso de activación/inactivación sigue siendo desconocido. Nuestra investigación se ha centrado en el estudio del mecanismo del sistema TLR4/MD-2 mediante métodos computacionales.

Con el fin de llevar a cabo nuestros objetivos de investigación, hemos utilizado una combinación de varias herramientas computacionales: optimización de la geometría, cálculos de carga, docking, cribado virtual y simulaciones de dinámica molecular de complejos y membranas de proteínas.

El objetivo principal de esta Tesis es elucidar las interacciones ligando-proteína del receptor TLR4 a nivel atómico a través de técnicas computacionales. Metodologías

computacionales se aplicarán para el estudio de los mecanismos moleculares involucrados en la funcionalidad de los receptores Toll-like, y en el reconocimiento de los PAMPs. Técnicas de acoplamiento ligando-proteína y cribado virtual serán utilizadas, dando lugar a una fuente de nuevos compuestos capaces de modular el comportamiento de los TLRs con posibles aplicaciones terapéuticas, y también como sondas biológicas.

CAPÍTULO 3: Para los moduladores descritos del sistema TLR4/MD-2 se predecirán modos de unión teórica. En particular, nuestro trabajo se centrará en análogos de LPS. Realizaremos un estudio computacional de algunos compuestos representativos para desvelar algunos de estos patrones de interacciones.

CAPÍTULO 4: El glicolípido catiónico IAXO-102, un potente antagonista de TLR4, se utilizará como esqueleto para diseñar nuevos moduladores de TLR4 y marcadores fluorescentes para el complejo TLR4. Estos compuestos serán sintetizados por otros colaboradores. Nuestros estudios de modelado nos permitirán diseñar nuevos compuestos y proponer modelos de interacción tanto para el CD14 como para el complejo TLR4/MD-2.

CAPÍTULO 5: Las estrategias de cribado virtual de diferentes bibliotecas de compuestos, seguidas por ensayos biológicos, nos permitirán identificar nuevas entidades químicas para el desarrollo de nuevos moduladores de TLR4 con una estructura diferente al LPS.

CAPÍTULO 6: Se llevará a cabo el estudio computacional del sistema TLR4/MD-2, simulando el complejo TLR4/MD-2 con diferentes modelos de membrana. Varios modelos del monómero TLR4/MD-2 insertado en diferentes membranas serán construidos y se llevará a cabo la simulación de dinámica molecular. Estos diferentes modelos serán útiles para la construcción final del dímero TLR4/MD-2 completo y nos proporcionarán información sobre el mecanismo agonista/antagonista del receptor

TLR4. El análisis de las simulaciones de dinámica molecular nos permitirá comprender las interacciones clave ligando-receptor y proteína-proteína implicadas en los eventos de reconocimiento molecular y el proceso de dimerización.

Nuestro trabajo ha llevado a las siguientes conclusiones:

Se han predicho modos de unión teórica para los moduladores descritos del sistema TLR4/MD-2. En particular, hemos centrado nuestro trabajo en glicolípidos sintéticos y moléculas que no tienen estructura de tipo LPS. Para todos estos moduladores de TLR4, está claro que, a pesar de su diferente estructura química, deben compartir un patrón común de interacciones con el receptor TLR4. Hemos llevado a cabo un estudio computacional de algunos compuestos representativos para revelar algunos de estos patrones de interacciones.

El glicolípido catiónico IAXO-102, un potente antagonista de TLR4, se ha utilizado como esqueleto para diseñar nuevos moduladores de TLR4 y marcadores fluorescentes para el complejo TLR4/MD-2. Nuestros estudios de modelización nos han permitido diseñar nuevos compuestos y nos han permitido proponer modelos de interacción tanto para el CD14 como para el complejo TLR4/MD-2. Todos estos compuestos han presentado una actividad antagonista para el complejo del TLR4/MD-2.

El estudio computacional del heterodímero TLR4/MD-2 completo se realizó simulando el complejo completo insertado en la membrana. El análisis de las simulaciones de dinámica molecular nos llevó a entender las interacciones claves implicadas en el proceso de dimerización a nivel atómico.

Resumiendo, técnicas de modelado molecular se han utilizado para elucidar los mecanismos de reconocimiento molecular de la modulación del receptor TLR4/MD-2,

centrándonos en los cambios conformacionales agonistas/antagonistas del sistema TLR4/MD-2, que nos ha permitido el diseño de nuevos ligandos, con cierto potencial terapéutico. También colaboramos con grupos experimentales para sintetizar los compuestos diseñados y realizar ensayos biológicos. Se ha realizado el estudio del modo de unión de varios moduladores TLR4/MD-2 y se ha llevado a cabo la simulación del sistema TLR4/MD-2 en diferentes entornos de membranas, mediante una combinación de cálculos de docking, simulaciones de dinámica molecular y protocolos de cribado virtual.

CHAPTER 1

Introduction

1.   Introduction

1.1   Understanding Toll Like Receptors

All living organisms constantly face attack from environmental microorganisms and need to cope with perpetual invasions into the body. The vertebrate immune response can be classified into innate and acquired immunity, being innate immunity the first line of defence against pathogens. Toll-like receptors (TLRs) have a fundamental role in early innate immunity, and are responsible for initiating and propagating inflammation. Evidences indicate a role for TLRs in immune and inflammatory diseases,[1] and increasingly in cancer.[2] Its relevance was highlighted in the award of the 2011 Nobel Prize in Medicine[3] to Beutler and Hoffmann for their studies on the activation of the innate immune system. Therefore, TLRs have emerged as novel targets for drug design. TLR agonists are currently under development for the treatment of cancer, allergies, and viral infections, and also as adjuvants as part of potent vaccines to be used in prevention or treatment of cancer and infectious diseases. As inappropriate TLR stimulation leads to inflammation and autoimmunity, significant efforts have also been directed towards the development of compounds as TLR antagonists.[4]

TLRs trigger two mechanisms of the immune response, the innate and the adaptive immunity, that work together to combat infection in mammals. The adaptive response generates antibody-secreting B cells and cytotoxic T cells that are specific and efficient at targeting pathogen. One disadvantage of this mechanism is that require more time to be developed than the innate response.[5] Since TLRs are the first responders to danger signals, they are pivotal in the research for fighting infectious and inflammatory diseases. New strategies for modulating the immune response depend on the understanding of the TLRs cell biology: structure, cell localization, signal transduction pathways and expression patterns.

In 1989, Janeway proposed that cells use pattern recognition to identify pathogens. Receptors recognize and subsequently bind to structural shapes or patterns called pathogen-associated molecular patterns (PAMPs) which are present in

entire groups of pathogens,[6] but not in the host. According to Janeway's theory, receptors cannot precisely recognize a particular microbe, but they can identify it as a foreign entity. Ten years after Janeway´s proposal, the first human pattern-recognition receptors (PRRs) were identified. Using the amino acid sequence of the Toll gene from the fruit fly,[7] related sequences were searched in the Human Genome Project database, finally leading to the identification of TLRs.[8]

Since some TLR binders are originated from the host, these new ligands are hypothesized to act as damage signals (damage-associated molecular patterns, or DAMPs) to alert the body of cell and tissue injury, this is evident in cases of necrosis, ischemic injury, etc.[9] For example, blocking various TLRs (such as TLR2 and TLR4) with antagonists may be useful in these circumstances to prevent an overactive immune response.[10-11] There is also evidence that TLRs contribute to the development of atherosclerosis and Alzheimer's disease through sensing of damage signals in the form of oxidized lipoproteins.[12]

Given their therapeutic potential, there is considerable interest in pharmaceuticals that modulate TLR activation. TLR antagonists hold great clinical promise for the treatment of numerous inflammatory conditions and are under investigation for the treatment of viral infections, redirecting allergic helper T cell responses and as anticancer therapeutics. Some TLR agonists have also proven to be safe and efficacious as vaccine adjuvants in humans and are currently used in Europe.[13-16]

The human TLR family comprises of 10 to 12 type I transmembrane glycoproteins with a single transmembrane domain, a conserved cytoplasmic Toll-like/interleukin-1 receptor signaling domain,[17-20] and an extracellular antigen recognition domain comprising of 19–25 tandem leucine-rich repeat (LRR) modules.[21] The LRR modules have 20~30 amino acid residues with conserved ''LxxLxLxxN'' motifs.[22-23]

TLRs generally function as heterodimers. Many ligands with distinct PAMPs exist, so the ten human TLRs are able to recognize more than ten different PAMPs. In

fact, the list of known TLR binders keeps growing. Heterodimer formation also increases binders diversity. For example, TLR2 associates with TLR1 and TLR6, and the association with proteins outside the TLR family also increase diversity, for example, TLR4 recognizes LPS in association with the accessory proteins MD-2 and CD14. Upon binding of the ligands to the extracellular domains of TLRs, rearrangement of the receptor complex is promoted, thus triggering the recruitment of specific adaptor proteins to the intracellular TIR domains.[24] In particular, MyD88 is a universal adapter protein used by almost all TLRs (except TLR3) to activate the transcription factor NF-κβ. Mal (also known as TIRAP) is another adaptor protein necessary to recruit MyD88 to TLR2 and TLR4. TLR expression is particularly significant in different types of white blood cells: mast cells, macrophages, and dendritic cells. The innate immune response is initiated by mast cells and macrophages, whereas the adaptive immune response is primarily initiated by dendritic cells.[25]

TLRs 1, 2, 4, 5, and 6 are located primarily in the plasma membrane, where they recognize components of microbial cell walls and membranes unique to pathogens (Figure **1.1**). The best characterized ligands are bacterial, examples include: lipopolysaccharide (LPS) and lipoteichoic acid from the cell wall, lipoproteins from the cell membrane, and flagellin, a structural component of bacterial flagella. TLRs 3, 7, 8, and 9 are situated in the membranes of endosomes and lysosomes, these TLRs bind to microbial nucleic acids, such as DNA from most organisms, and double and single stranded RNA from RNA viruses. Since these TLRs cannot distinguish self-nucleic acids (those of the host cell) on structural differences alone, and recognition of foreign nucleic acids (those of the pathogen) largely depends on the location in the cell.[26]
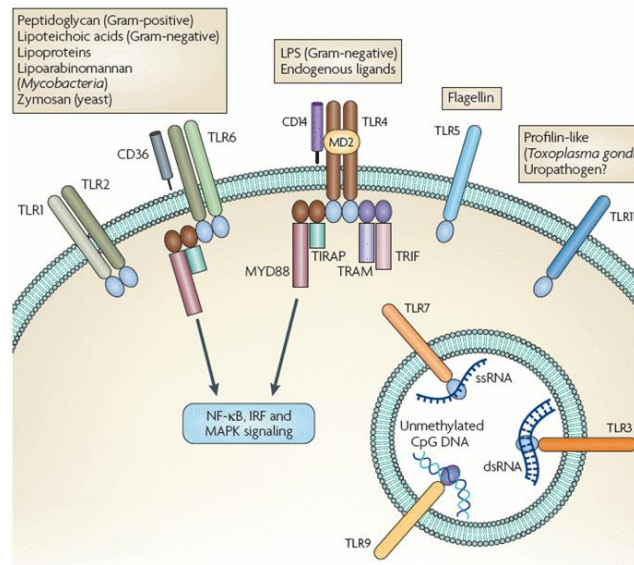
Figure 1.1. Human Toll Like Receptor family.[2]

1.1.1 Relevance of TLR4 as Therapeutic Target

TLR4 and the Inmune Response

TLR4 was the first TLR identified,[8] and was characterized as the receptor for LPS as it generates an innate immune response upon LPS stimulation.[27-29] LPS, glycolipids produced by Gram-negative bacteria, are composed of an oligosaccharide core and a highly variable O antigen polysaccharide component, along with a hydrophobic lipid A segment containing multiple lipid acyl tails and a phosphorylated glucosamine disaccharide headgroup (Figure **1.2**). TLR4 was found to require to be associated to an additional protein, MD-2 (myeloid differentiation factor 2), to be activated by LPS, and it was discovered that mice lacking MD-2 do not respond to LPS. There have been identified a number of MD-2 polymorphisms that modify LPS binding and/or activation.[30] MD-2 is very flexible, and allosterically transmits this conformational plasticity, in a ligand dependent manner, to a phenylalanine residue (Phe126) in the cavity mouth previously implicated in TLR4 activation. The assumption is that Phe126 is the "molecular switch" in endotoxic signalling.

LPS interaction with TLR4/MD-2 involves at least two other proteins, these are lipopolysaccharide binding protein (LPB) and cluster of differentiation 14 (CD14). LPS

first binds to LBP in serum and it is then transferred to CD14 (Figure **1.2**). The major role for CD14 is to enhance the sensitivity of the TLR4/MD-2 signaling complex, causing the binding affinity toward LPS to drop to picomolar concentrations. It was identified that mice without CD14 are resistant to endotoxic shock. Unlike the rest of the TLR family, TLR4 does not recognize the lipid in isolation, but when bound to MD-2.[31]

TLR4 activation has been associated with certain autoimmune diseases, noninfectious inflammatory disorders, and neuropathic pain, as well as metabolic syndrome in multiple tissues and cardiovascular diseases, suggesting a wide range of possible clinical settings for the application of TLR4 antagonists.[32-33] However, agonists of TLR4 can be useful as adjuvants in vaccine development and in cancer immunotherapy.[29]

Targeting TLR is actually a sparkling field for translational cancer research. The expression of TLRs mediating innate immune response on tumor cells, influence in the proliferation and migration of these cells. The activation of TLRs may play opposite role, antitumor or protumor. Better understanding the mechanism of TLRs in cancer biology will contribute to expand the opportunities for pharmacological intervention and discovery new strategies and new candidates for drug against cancer.[34]
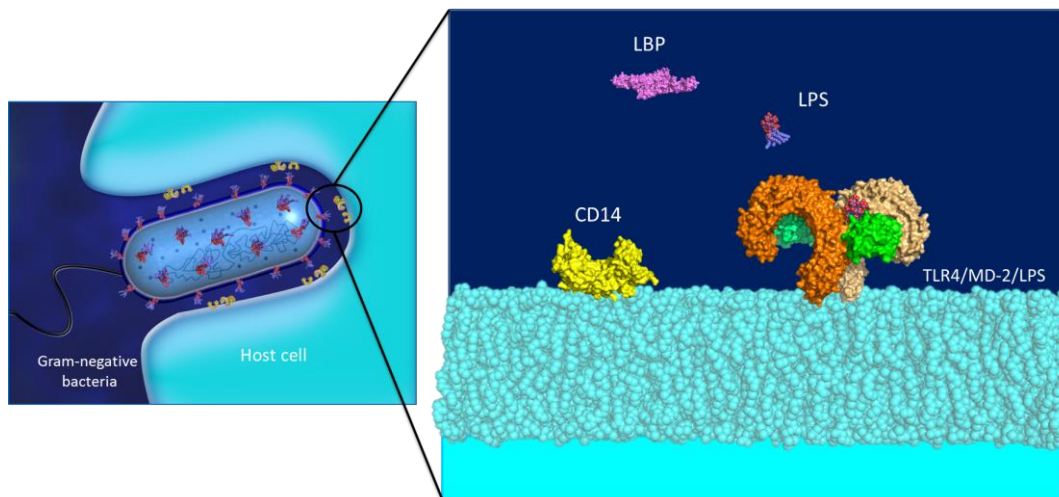

Figure 1.2. Left: Bacteria engulped by a macrophage during infection. Right: Detail of the LPS recognition by TLR4, in concert with the accessory proteins LBP (LPS-binding protein), and CD14 (who transfers LPS from bacteria membranes or aggregates in serum to MD-2).

Therapeutic Applications of TLR4 Agonists

TLR agonists have shown to improve currently applied anticancer vaccination protocols,[35-36] and are also the focal point for new vaccine development as non-infectious subunit vaccines.[16, 36] For example, the natural product monophosphoryl lipid A (MPLA), a detoxified component of LPS from Salmonella Minnesota which contains the lipid A (Figure **1.3**) moiety that binds to TLR4/MD-2, is incorporated into several vaccines,[16, 37] including vaccines for Hepatitis B (Fendrix™),[38] and cervical cancer (Cervarix™)[39-40] and in the immunotherapy for melanoma.[41]

Also synthetic TLR4 agonists have been designed and assayed. Compound E6020[42] have shown good adjuvant activity with antitumoral trastuzumab,[43] or enhancing vaccine efficacy (Figure **1.3**).[44] Lipid A mimetics, such as the aminoalkyl glucosaminide phosphates, have been developed as TLR4 stimulants[45] with good adjuvant activity,[46] including the potent vaccine adjuvant RC-529,[47] and the bioisoster CRX-547, which has reduced toxicity in comparison to RC-529.[48] Small molecules pyrimido[5,4- b]indoles have shown to stimulate TLR4 and could potentially be used as adjuvants or immune modulators,[49] synthetic analogues of natural product Euodenine A have exhibited potent and selective agonism towards TLR4,[50] and synthetic peptides to mimic the TLR4/LPS interaction have also been reported.[51]
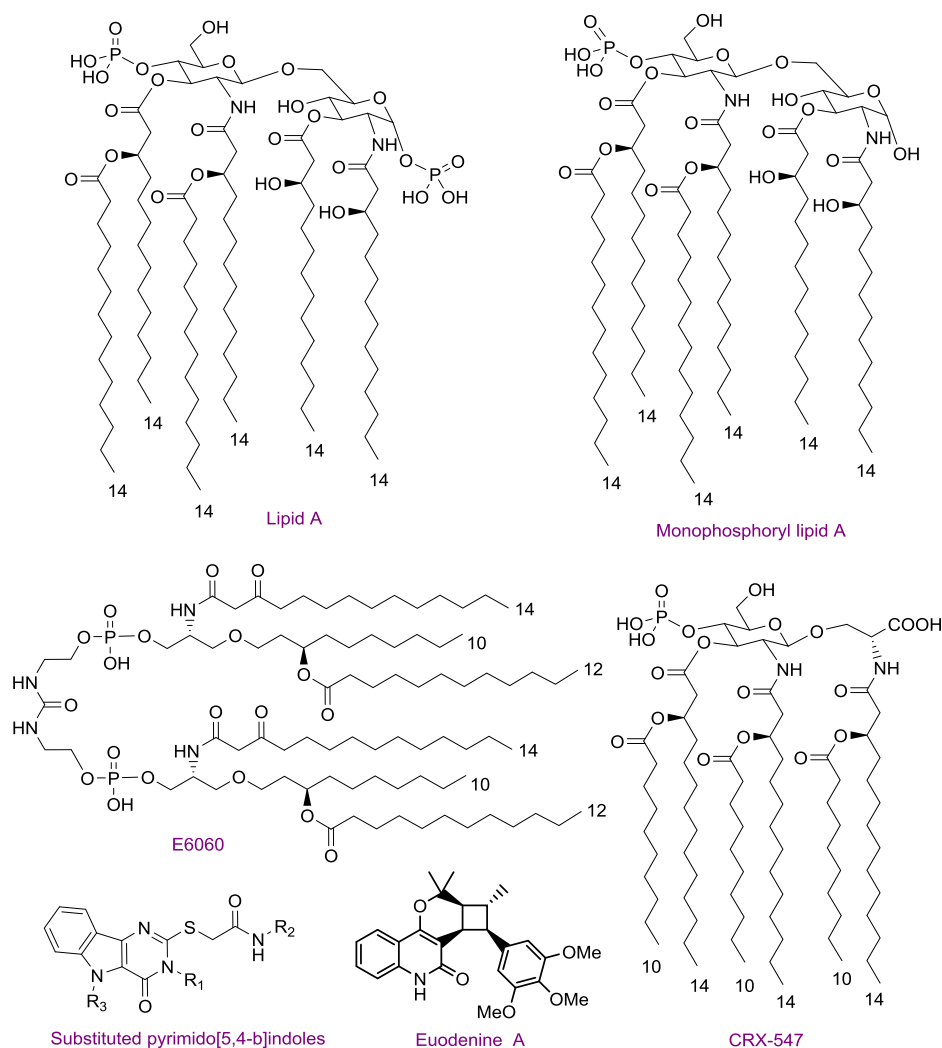
Figure 1.3. Chemical structures of selected TLR4 agonists.

Therapeutic Applications of TLR4 Antagonists

The design of LPS mimetics with TLR4 antagonist activity is an emerging strategy for the treatment of sepsis,[52] combined with the challenge of obtaining good drug-like properties. Lipid IVa is an underacylated lipid A analogue with intriguing properties, being antagonist in human TLR4 but agonist in mouse (see Table **1.1** for references). The tetraacylated synthetic compound eritoran (or E5564, Figure **1.4**) reached phase III in clinical trials, but failed to demonstrate sufficient efficacy in late stage human trials, although it has recently shown promising activity in preventing influenza induced acute lung injury, through a TLR4 antagonism mechanism.[53]

9

Simplified derivatives without the phosphate group have also been reported,[54-56] exploring the presence of a cation. Of merit, new glycolipids and benzylammonium lipids (for example, IAXO-102 and IAXO-103, Figure **1.4**) are the first family derived from a monosaccharide core with effective TLR4 antagonist activity.[55] Other synthetic lipid A analogues include, for example, compound D1,[57] and one lipid X mimetic (Figure **1.4**),[58] exhibiting a TLR4 antagonist mechanism by blocking the interaction of LPSs with both CD14 and MD-2 proteins.
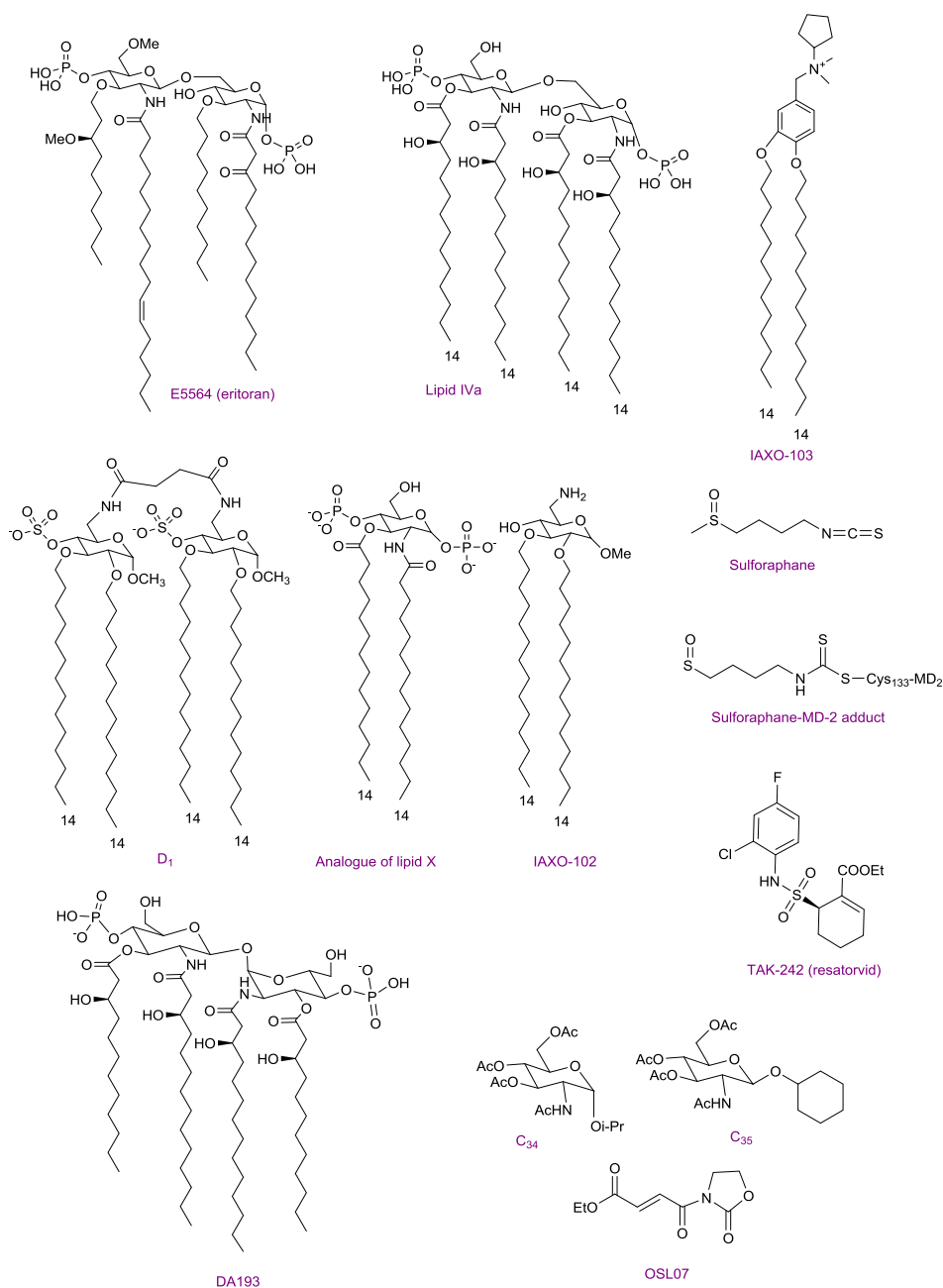

Figure 1.4. Chemical structures of selected TLR4 antagonists.

Several small non LPS-like molecules with TLR4 antagonist activity have also been developed (Figure **1.4**), such as ethyl 4-oxo-4-(oxazolidin-3-yl)-butenoate derivatives (OSL07), benzothiazole-based inhibitors, ethyl phenyl-sulfamoyl-cyclohexene-carboxylate derivatives (TAK-242 or resatorvid), and β-amino alcohol derivatives.[59-62] However, no successful progress was shown when reaching clinical phases (for example, in the case of compound OSL07). Examples of non-lipid TLR4 antagonists based on dendrimer architecture can also be found in the recent literature, showing that the presence of lipidic chains is not an absolute requirement for an MD-2 antagonist, and thus opening interesting opportunities for immunity modulation.[63-64]

1.1.2   Reported X-Ray Structures of TLR4: Key Interactions

Scientists have elucidated the three-dimensional structure of TLRs by means of X-ray crystallography. Presently, the 3D structure is available for TLRs 1, 2, 3, 4, and 6, as hetero/homo-dimers, and in complex with some ligands (agonists and antagonists) and/or co-receptors.[65-66] The analysis of the ligand-receptor interactions at atomic detail gives understanding to the signaling processes which gives rise to the design of molecules with properties required in TLR modulators (agonist/antagonist properties).[67-68] Still, the molecular features that drive the recognition processes have yet to be unraveled.

Regarding TLR4, to the best of our knowledge, 15 X-ray crystallographic structures have been reported containing any of the partners forming the TLR4/MD-2/ligand complex (Table **1.1**). Deposited at the Protein Data Bank (PDB), we can find: 1) TLR4 protein, either the nearly complete TLR4 chain (PDB-ID: 2Z63) or either fragments of the chain (PDB-ID: 2Z62, and 2Z66); 2) MD-2 protein bound to different ligands (PDB-ID: 2E56, and 2E59); 3) the TLR4/MD-2 complex not forming the dimeric (active) symmetric complex, either bound to antagonist ligands (PDB-ID: 3ULA, and 2Z65), or without any ligand (PDB-ID: 5IJB, 2Z64); and 4) the complete multimeric complex of the active TLR4 (PDB-ID: 5IJB, 5IJD, 3FXI, 3VQ1, 3VQ2, and 4G8A) composed by two

TLR4/MD-2/ligand units interacting in a symmetric manner, forming the extracellular complex which finally triggers the TLR4 signaling pathway through the interaction of each intracellular TIR domain belonging to the TLR4.[67] Table **1.1** also contains information about the relevant protein CD14.

Table 1.1. X-Ray crystallographic structures of TLR4 and CD14 deposited at the Protein Data Bank. [a]The structures are noted as multimer when two TLR4/MD-2 heterodimers are found in the structure. [b]Not natural complexes. [c]*E. coli* LPS. [d]Re-chemotype of *E. coli* LPS. [e]Human TLR4 polymorphism D299G and T399I. [f]Human TLR4 decoy. [g]Three units of myristic acid.

PDB-ID	Organism	Proteins	Ligand	Structure[a]	MD-2 Conformation	Resolution (Å)
5IJD[69]	Mouse-inshore hagfish hybrid	TLR4/MD-2	Re-LPS[d]	Multimer	Agonist	2.7
5IJC[69]	Mouse-inshore hagfish hybrid	TLR4/MD-2	Neoseptin-3	Multimer	Agonist	2.57
5IJB[69]	Mouse-inshore hagfish hybrid	TLR4 fragment/MD-2	None	Heterodimer	Agonist	2.91
3FXI[67]	Human	TLR4/MD-2	LPS[c]	Multimer	Agonist	3.10
3VQ1[70]	Mouse	TLR4/MD-2	Lipid IVa	Multimer	Agonist	2.70
3VQ2[70]	Mouse	TLR4/MD-2	Re-LPS[d]	Multimer	Agonist	2.48
4G8A[71]	Human	TLR4/MD-2[e]	Re-LPS[d]	Multimer	Agonist	2.40
2Z64[72]	Mouse	TLR4/MD-2	None	Heterodimer	Antagonist	2.84
3ULA[73]	Human-inshore hagfish hybrid[f]	TLR4 fragment/MD-2	E55	Heterodimer[b]	Antagonist	3.60
2Z65[72]	Human-inshore hagfish hybrid	TLR4 fragment/MD-2	E55	Heterodimer[b]	Antagonist	2.70
2Z63[72]	Human-inshore hagfish hybrid	TLR4	None	Monomer	N/A	2.00
2Z62[72]	Human-inshore hagfish hybrid	TLR4 fragment	None	Monomer	N/A	1.70

2Z66[72]	Human-inshore hagfish hybrid	TLR4 fragment	None	Tetramer[b]	N/A	1.90
2E56[74]	Human	MD-2	Myristic acid[g]	Monomer	Antagonist	2.00
2E59[74]	Human	MD-2	Lipid IVa	Monomer	Antagonist	2.21
1WWL[75]	Mouse	CD14	---	---	---	2.50
4GLP[76]	Human	CD14	---	Amino-terminal pocket	---	4.00

Structure of TLR4

From the structural point of view TLR4 belongs to the superfamily of the LRR proteins, which are characterized for having a typical horse-shoe-like conformation, which contains several parallel β-strands in its concave surface and loops in its convex surface (Figure **1.7**).[67] Differences between the β-strands in the concave surface allow dividing this region in three different domains. First, the N-terminal domain, which includes modules from 1 to 6 (Figure **1.7**), and has no sequence homology with the typical LRR modules, preventing this region to be highly hydrophobic. Second, the C-terminal domain is formed by modules 13 to 22 (Figure **1.5**). And third, the central domain, which is built by typical LRR modules, except one variable residue which has different residue lengths over the set of β-strands, deeply related with the horse-shoe-like shape of this kind of proteins. The surfaces of the N-terminal and central domains provide charge complementarity in order to bind to its co-receptor MD-2, leading to a stable 1:1 heterodimer.[68] This heterodimer dimerizes when bound to the bacterial LPS, leading to the multimeric complex, and triggering the immune response. When bound to an antagonist ligand, the heterodimer is not capable of dimerization.
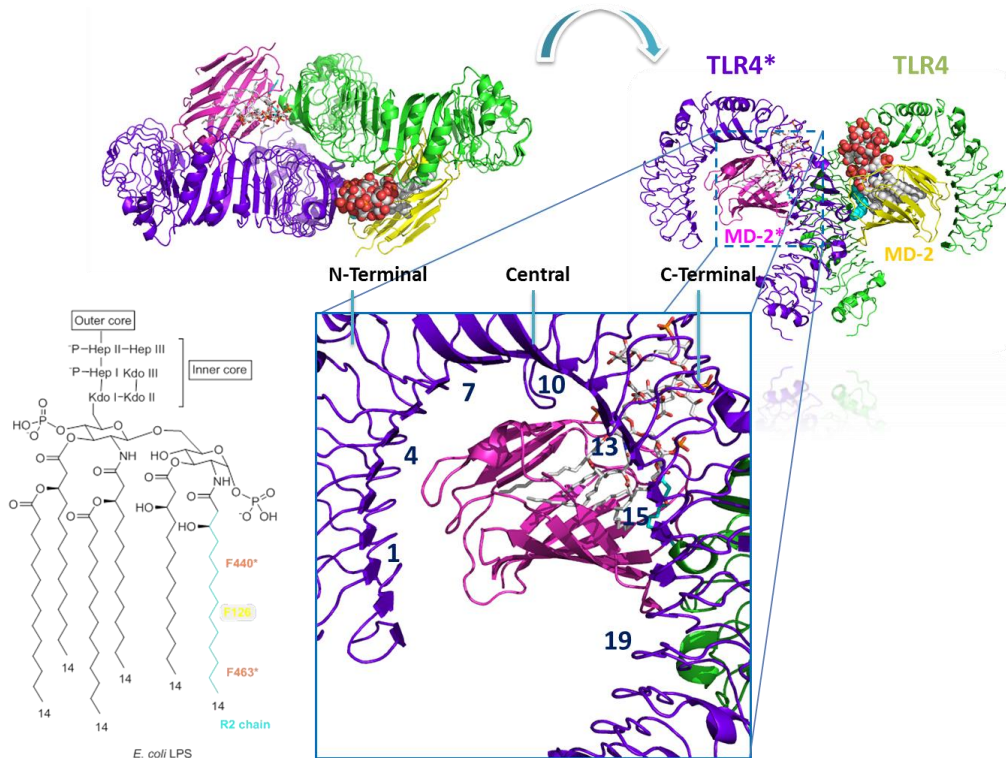
Figure 1.5. Multimer complex composed of two copies of heterodimers of the TLR4/MD-2/LPS complex arranged symmetrically (PDB-ID: 3FXI). Highlight in the box shows the N-terminal, central, C-terminal domains of the TLR4.

Structure of MD-2

Before describing any structural attributes of these crystallographic structures, it is worth pointing out that the main structural differences observed between the activated and inactivated (agonist and antagonist) conformations of the TLR4/MD-2 complex are mainly related to the twist of a small loop of MD-2, comprised of Phe126 and Leu124 (Figure **1.6**).[74] The twist of this small loop acts as an ON/OFF switch, allowing the binding of the agonist (ON) or the antagonist (OFF). The formation of the TLR4/MD-2/agonist complex builds a hydrophobic region and a dimerization interface. This promotes the coupling of the second TLR4/MD-2/ligand partner (referred to here as TLR4*/MD-2*). Binding of antagonists leads to changes into the characteristics of this hydrophobic interface precluding formation of the activated TLR4/MD-2/ligand multimer.

The structure of the MD-2 is characterized by two antiparallel β-sheets, containing three and six β-strands respectively. These two β-sheets adopt a β-cup-like fold, forming a large internal hydrophobic pocket of 1000 Å^2, open to host the large lipophilic fatty acid (FA) chains from the LPS.[74] This pocket is completely built by hydrophobic residues on the interior, and positively charged residues surrounding the entrance, allowing the binding of the polar groups from the lipid IVa (PDB-ID: 2E59). The formation of the disulfide bond between Cys25 and Cys51, and the hydrogen bond between Tyr34 and Tyr36, are crucial for the stability of the MD-2 structure.
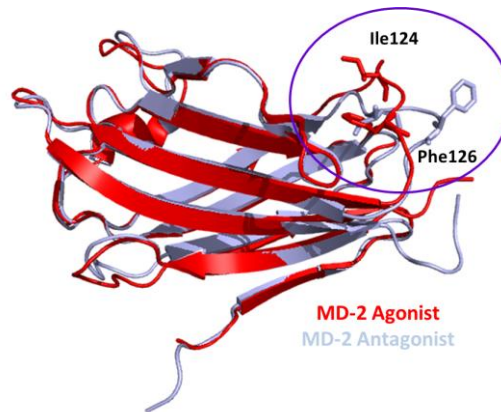
Figure 1.6. Superimposition of agonist (PDB-ID: 3FXI) and antagonist (PDB-ID: 2E59) conformations of the human MD-2. Purple circles show main conformational change which mainly involves the loop containing Ile124 and Phe126.

Structure of the TLR4/MD-2 Complex

The TLR4/MD-2 complex interacts through a narrow and long interface, noted as the primary contact interface[67] which is formed before the binding of LPS, and is divided by the A and B patches belonging to the N-terminal and central domains of TLR4 respectively (Figure **1.9**). On one hand, the A patch in TLR4 is characterized by being highly evolutionarily conserved, and negatively charged, which allows the interaction with the positively charged Arg68 and Lys109 residues of MD-2. On the other hand, the B patch is built by a poorly conserved area, positively charged, and interacts with the negatively charged residues in the F β-strand of MD-2 (Figure **1.7**), patterns can be found in both, the apo and the eritoran-bound TLR4/MD-2 complexes.

Additionally, in the TLR4/MD-2/eritoran complex, it can be observed that eritoran does not interact directly with TLR4, while the FA chains are placed deeply inside the hydrophobic pocket, and the two phosphate groups interact with positively charged residues at the entrance of the pocket, not observing direct polar interactions of the disaccharide with MD-2.[72]
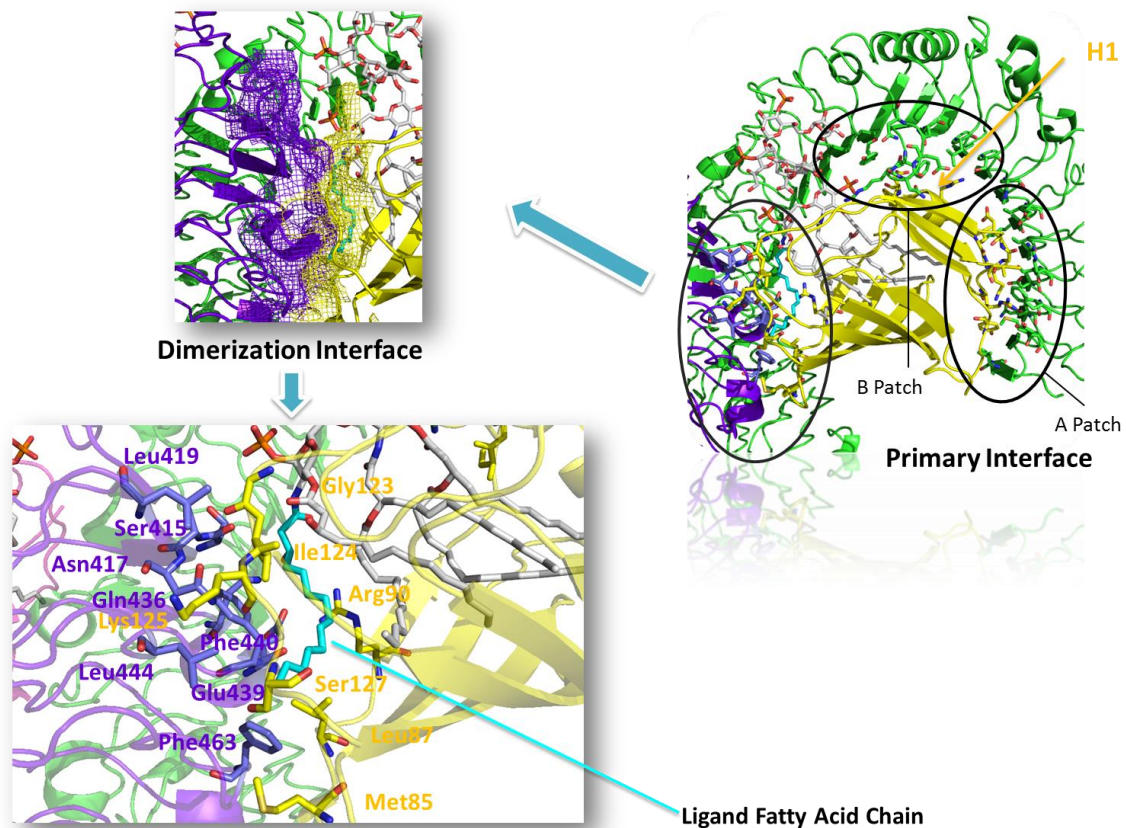

Figure 1.7. Two views of the dimerization and the primary interfaces of TLR4/MD-2/LPS complex (PDB-ID: 3FXI).

Structure of the Activated TLR4/MD-2/Agonist Multimeric Complex

Analysis of the structures of the activated multimeric complex of TLR4/MD-2/agonist, composed of two copies of the heterodimer TLR4/MD-2/ligand arranged symmetrically, reveals that each copy of the heterodimer interacts through the so-called dimerization interface.[67] This involves the LRR modules 15-17 of the C-terminal

domain (Figures **1.5, 1.7** and **1.8**), and the Phe126 and Leu87 loops that connect the G-H and the E-F β-strands of MD-2 (Figure **1.8**). Some interesting interactions that can be highlighted are the CH-π interactions between Phe463 (TLR4*) and Leu87 side chains, and polar interactions between Glu439 (TLR4*) and Arg90 (MD-2). Interestingly, the Phe126 loop of MD-2 is the one related with the main conformational change responsible with the agonist/antagonist behaviour (ON/OFF switch). One of the FA chains from the LPS does not penetrate into the MD-2 pocket, meaning this is partially exposed to the outer, and thus completing the hydrophobic surface of MD-2 that will interact with TLR4*. In particular, this FA chain interacts with the hydrophobic side chains located at the G β-strands of the MD-2, which could be responsible for triggering the conformational change of MD-2, which in turn, promotes the dimerization, and the final activation of the TLR4 signaling pathway. In addition to this particular (and key) interaction, other hydrophobic interactions contribute to the final assembly of the full multimeric complex.

MD-2 can only be found in the antagonist conformation, either bound to antagonists (MD-2/lipid IVa, PDB-ID: 2E59, and TLR4/MD-2/Eritoran, PDB-ID: 2Z65), or without any ligand (TLR4/MD-2, PDB-ID: 2Z64). This observation may suggest that the conformational change of the Phe126 loop which leads to the dimerization, and to the activation of the immune response, is promoted upon agonist binding. The MD-2 conformational change could be explained by the induced fit paradigm rather than by the conformational selection from the MD-2 conformational landscape. This has also been suggested by reported NMR studies on hexaacylated endotoxin bound to wild-type and F126A mutant MD-2, which indicate that re-orientation of the aromatic side chain of Phe126 is induced by binding of hexaacylated endotoxin, preceding interaction with TLR4.[77] Multimeric complexes are available with natural agonist binder LPS (PDB-ID: 3FXI, human TLR4/MD-2), and with lipid IVa (PDB-ID: 3VQ1, mouse TLR4/MD-2). In this way, lipid IVa acts as an "accidental" agonist or antagonist depending on the species dual binder, able to activate the TLR4 immune response in mice, but to block the TLR4 system in humans. As an agonist, it is able to bridge between the two phosphate binding sites of the two TLR4/MD-2 units. As a "flipped"

antagonist (the glucosamine backbone is in opposite orientation than LPS) it is buried more deeply into the cavity of MD-2 (4-5 Å) and thus can only connect to one phosphate binding site. Eritoran is not able to bridge between the two TLR4/MD-2 subunits of any of the analyzed species.
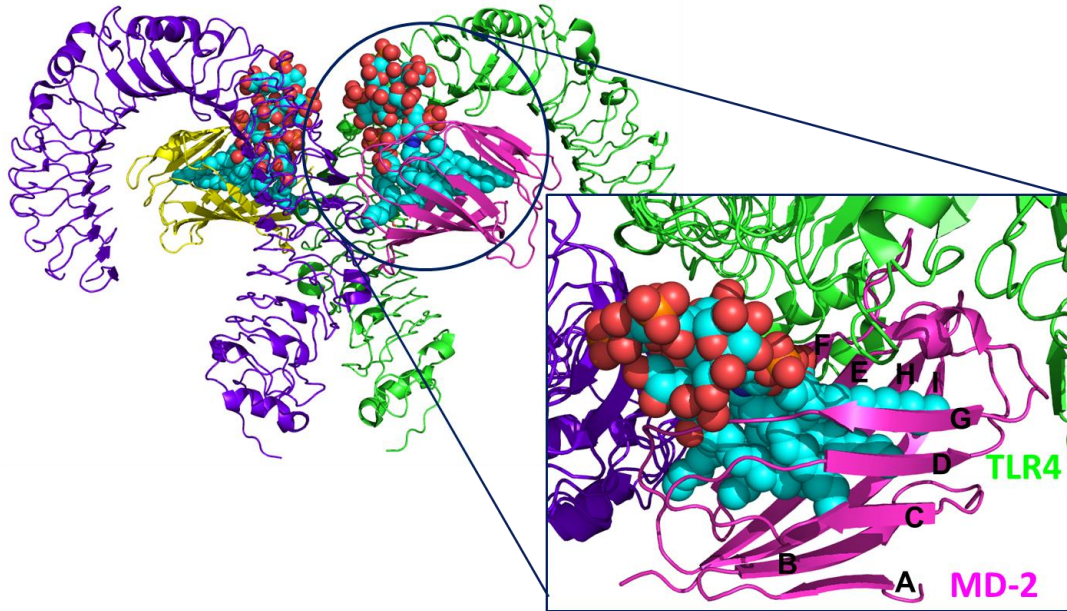

Figure 1.8. Multimer complex composed of two copies of heterodimers of the TLR4/MD-2/LPS complex organized symmetrically (PDB-ID: 3FXI).Box at the bottom shows the standard naming for each β-strand of MD-2.

Structure of CD14

CD14 acts as a co-receptor for the detection of LPS, and can bind LPS only in the presence of the lipopolysaccharide-binding protein (LBP). Human (*h*CD14) and mouse (*m*CD14) CD14 are characterized by a bent solenoid typical of leucine-rich repeat (LRR) proteins, with a large hydrophobic pocket found on the amino-terminal side (PDB-ID: 4GLP for *h*CD14, and 1WWL for *m*CD14). Similarly to MD-2, CD14 is also characterized by having a wide lipophilic pocket, but with fewer number of polar residues at the rim, and it is capable of recognizing other microbial and cellular molecular determinants, in addition to LPS, such as lipopeptides which are the PAMP recognized by TLR2.
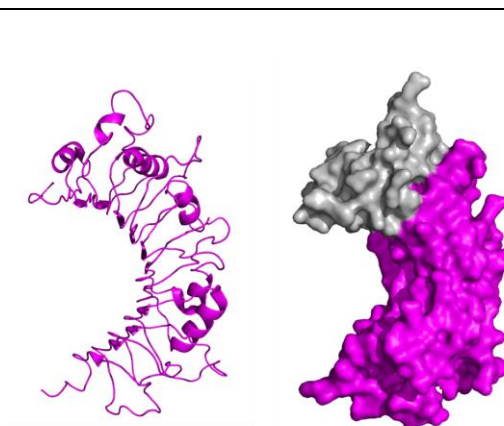
Figure 1.9. 3D structure of the amino-terminal pocket of human CD14 (PDB-ID: 4GLP). On the left, *h*CD14 in cartoon style. On the right, view in surface; highlight in grey the similar binding pocket in comparison with MD-2 pocket.

1.2 Reported Molecular Modeling Studies of the TLR4/MD-2 System

As a result of the increasing development in computing technology,[78] research to date has allowed the exploration of the dynamics of these high size TLR4/MD-2 system by means of MD simulations, and also binding properties of reported TLR4/MD-2 modulators. A recent review by our group has summarized these studies. We here give a brief overview, as follows.[79]

1.2.1 Computational Studies of the TLR4/MD-2 Ectodomain

There are several reported MD simulation studies focused on the extracellular domain of the TLR4/MD-2 complex, more specifically on its ability to recognize lipid A and lipid IVa. Garate *et al.*[80] reported several MD simulations of different TLR4/MD-2/ligand complexes, concluding in to highlight the hydrophobicity of the MD-2 pocket and its ability to close promptly in an aqueous environment, due to the flexibility of the helix connecting MD-2 with TLR4 (helix H1). MD-2 was observed to fluctuate less due to the presence of TLR4, reducing the number of degrees of freedom. Another interesting conclusion is the key role that charged phosphates play in the early recognition of lipids with the corresponding impact on the formation of heterotetramers. Plasticity of MD-2 has also been observed by DeMarco *et al.* after several MD simulations performed in

complex with variably-acylated lipid A molecules from *Escherichia coli* and *Neisseria meningitides*, concluding that the level of acylation of these ligands greatly influences the final architecture of the dimerization interface,[81] and the final agonist/antagonist conformation of the TLR4/MD-2 system. This agonist/antagonist transition has also been studied by MD simulation by Paramo *et al.*[82] finding that the opening of the Phe126 switch disrupts the arrangement of nearby side chains from Leu87, Val82 and Met85 of MD-2, in agreement with NMR studies.[77]

The TLR4 ectodomain and its dimerization mechanism have also been subjected to computational studies. MD simulations of the TLR4/MD-2/TLR4*/MD-2* comlex by de Aguiar *et al.* [83] revealed pronounced conformation and structure alterations in the N- and C-terminal domains of the TLR4 ectodomains, while MD-2 underwent structural rearrangements and interacted with TLR4 and partner TLR4*, reinforcing the stabilizing role of MD-2 for the TLR4 complexation. In a related work, Anwar et al. performed interesting computational studies of the TLR4 signaling mechanism by studying the species-specific behavior of TLR4/MD-2 in the recognition of *Rhodobacter sphaeroides* lipid A (RsLA): human, murine, horse and hamster TLR4/MD-2 systems.[84] The data suggested a relationship between the flexibility of two loops (the on/off switch Phe126 loop of MD-2, from residues 123–129, and the MD-2 loop containing residues 81–89, which are the residues interacting with the partner TLR4*), and the agonist/antagonist activity of the ligand, thus providing a plausible explanation for the species-specific behavior of RsLA regarding TLR4 activation.
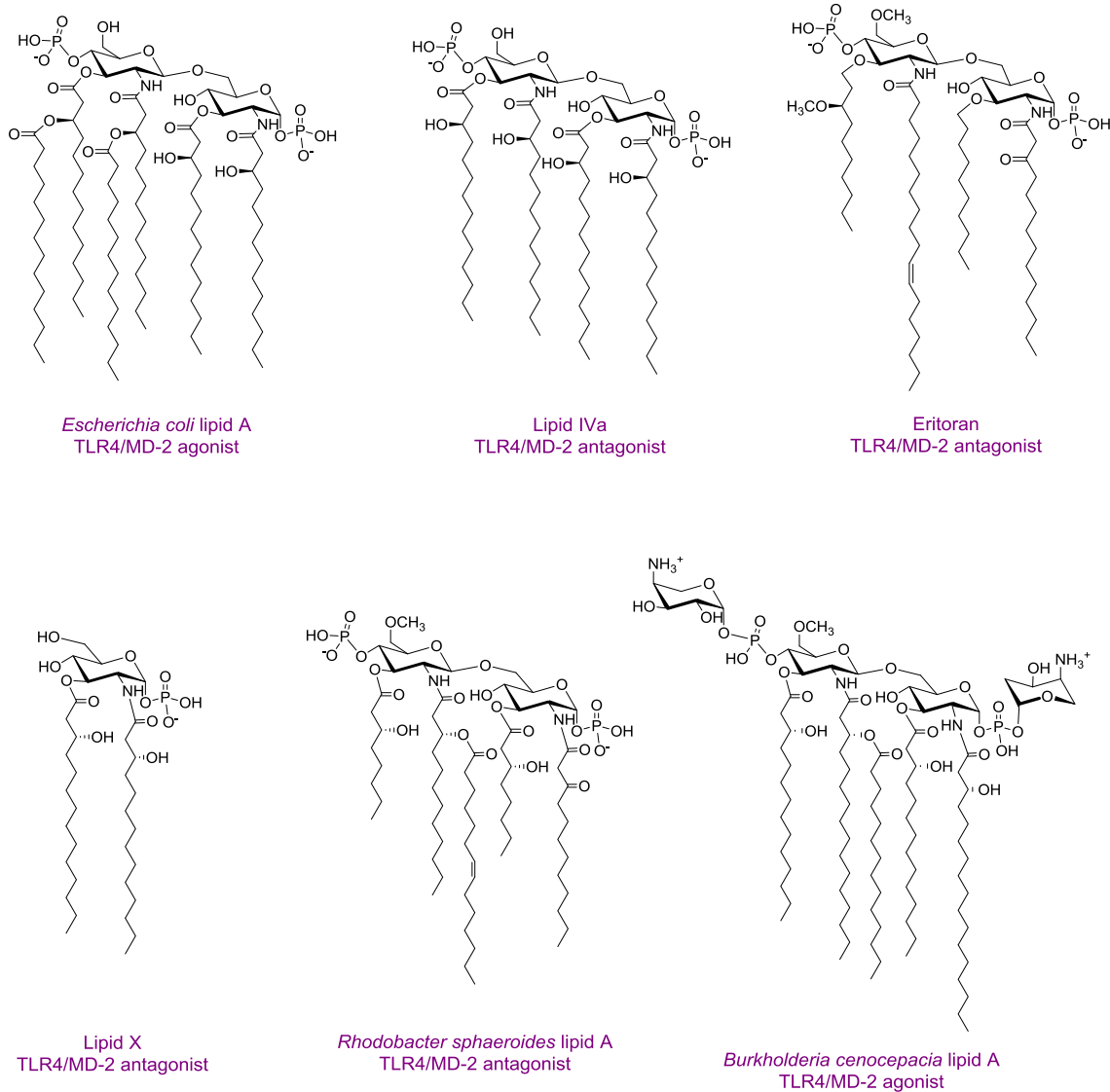
Escherichia coli lipid A
TLR4/MD-2 agonist

Lipid IVa
TLR4/MD-2 antagonist

Eritoran
TLR4/MD-2 antagonist

Lipid X
TLR4/MD-2 antagonist

Rhodobacter sphaeroides lipid A
TLR4/MD-2 antagonist

Burkholderia cenocepacia lipid A
TLR4/MD-2 agonist

Figure 1.10. Lipid A and synthetic lipid A analogues with activity as TLR4 modulators. Activity is referred to *h*TMR4/MD-2.

Computational Strategies to Study Mutant TLR4 and MD-2 Proteins

Herein, we report three studies that used in silico-mutated TLR4/MD-2 systems to serve different purposes; namely, to transform the mouse complex interface into the human one by mutating the residues laying there into their human counterparts, to estimate the influence of the mutations on the binding affinity of a ligand and to

evaluate the impact of mutations on the structural shape and plasticity of the MD-2 binding pocket.

In a 2009 study, Slivka et al.[85] used Rosetta software[86] to compare the binding energy of a truncated MD-2 with the original one. MD-2 was truncated (termed MD-2-I) to keep only the residues identified as playing a major role in maintaining the TLR4/MD-2 heterodimer stability. The docking experiment was performed targeting both a partial human TLR4 retrieved from the Protein Data Bank (PDB-ID: 2Z65) and a full-length TLR4 humanized model built by mutating the residues at the TLR4/MD-2 heterodimer interface in the mouse crystal structure (PDB-ID: 2Z64) into their human counterparts (TLR4: F160L, G234N, K263R, D264N, T290A; MD-2: H96R, H98R). In the first case, the affinity of MD-2-I was found higher than the one of the full-length MD-2. When docked against the human TLR4 model, MD-2-I exhibited a lower affinity than the full length MD-2. Altogether, these results indicate that MD-2-I is theoretically able to bind TLR4 and might even compete with the full-length MD-2. This was confirmed by cell assay experiments showing that the addition of compound MD-2-I abolishes their responsiveness to LPS stimulation. Flow cytometry analyses on cells (HEK293 cell line, transfected with all proteins involved in the TLR4 activation pathway) incubated with LPS covalently linked to fluorescein isothiocyanate (LPS-FITC), suggest that MD-2-I impedes TLR4/MD-2 dimerization. The SEAP assay shows that MD-2-I also alters downstream signaling.

A 4-aminoarabinose-containing lipid A from the opportunistic bacterium *Burkholderia cenocepacia* (Figure **1.10**) and its aminoarabinose-deficient equivalent were docked to *h*MD-2 and *m*MD-2.[87] These docked models were used to build a full dimer complex in order to perform 0.1-ps MD simulations. In both the human and the murine systems, the wild-type (WT) LPS obtained a better predicted free energy of binding than the aminoarabinose-deprived one, with both the AutoDock[88] and AutoDock VINA[89] docking programs. An energy analysis was conducted to estimate the per residue contribution to the total ligand binding energy for both WT and mutated TLR4/MD-2/TLR4*/MD-2* complexes (D294A, R322A, S415A* and S416A*) using the MM-ISMSA method.[90] This study permitted the identification of the mutated residues

as major contributors to the total binding energy of *B. cenocepacia* LPS and suggested that the ammonium groups of Ara4N stabilize the complex by providing additional anchorage interactions.

Recently, the critical role of residue 135 of MD-2, located deeply inside the hydrophobic pocket, was reported by Vasl et al.[91]. *h*MD-2 has the ability to bind LPS in the absence of TLR4, while *m*MD-2 is responsive to LPS only when engaged in a complex with TLR4. Site-directed mutagenesis was used on *h*MD-2, to mutate Val135 to its murine alanine counterpart. This single point mutation led to a mutant V135A *h*MD-2 lacking the ability of binding LPS. A series of MD simulations of the WT MD-2 and the V135A mutant MD-2 in solution and in complex with TLR4 was performed to study the conformational changes. In the case of the WT *h*MD-2, the authors reported an abrupt decreased of the SASA and volume in the first nanoseconds of the simulation, describing it as a hydrophobic collapse. This phenomenon was not observed in the V135A systems, suggesting that Val135 is primordial to confer plasticity to MD-2. This tendency was confirmed by another simulation of MD-2 in complex with three myristic acids (as observed in some crystal structures). The V135A mutant *h*MD-2 needed a much longer simulation time to adapt its shape to the three myristic acids than the wild-type. The authors concluded that this loss of plasticity could incapacitate *h*MD-2 for binding LPS.[91]

1.2.2 Computational Studies on the Intracellular Domain of TLR4

The intracellular domain of the TLR4 transmembrane protein contains a Toll/interleukin-1 receptor homology domain, which is a common feature of all adaptors involved in the initiation of TLR4 signaling, mediating protein-protein interactions between the TLR4 and the signal transduction components. TLR4 has two distinguished signaling pathways involving primarily four TIR-domain-containing adaptors. In the first pathway, the MyD88 adapter-like (Mal) acts as a "sorting" adaptor by recruiting the myeloid differentiation primary response gene 88 (MyD88), the "signaling" adaptor, to the plasma membrane. In the second pathway, the TRIF-

related adaptor molecule (TRAM) plays the role of "sorting" adaptor, which recruits the TIR-domain-containing adapter-inducing interferon-β (TRIF), the "signaling" adaptor, to the membrane to initiate the signal. As a major component of theses adaptors, the TIR domain is believed to play a central role in the recruitment processes.[24, 92]

The crystal structures of human TLR1 (PDB-ID: 1FYV) and TLR2 (PDB-ID: 1FYW) revealed the structural basis of the TIR domain[93] followed by the crystal structure of TLR10 TIR domain (PDB-ID: 2J67)[94] and the solution structure of MyD88 TIR domain resolved by NMR (PDB-ID: 2JS7 and 2Z5V).[95] Prior to that release, two homology models of the TIR domain of MyD88 were reported. Both were built based on the TLR2 TIR domain crystal structure (PDB-ID: 1FYW) resolved by X-ray crystallography.[96-97] In 2012, the crystal structure of Mal was also resolved by X-ray crystallography (PDB-ID: 3UB2).[98]

The lack of structural information for the TIR domain of TLR4 has driven the creation of models to clarify the recruitment of adaptors from a structural perspective. Dunne et al.[99] built monomer models of TLR4, Mal and MyD88 using comparative modeling and loop refining techniques. They noted differences in the electrostatic surface potentials suggesting that adaptor binding is driven by electrostatic complementarity. This point was also emphasized in a study by Kubarenko et al.[100] in which they compared the surface charges of TIR domains of the crystal structure of *h*TLR2 and of the models of *h*TLR3 and *h*TLR4 and noted that the surface charge distribution of the BB loop and the αC-helix (Figure **1.11**) present similarities in TLR2 and TLR4 and differ between TLR3 and TLR4. The authors considered that these findings could explain why TLR2 and TLR4 recruit MyD88, whereas TLR3 does not. In the computational study by Gong et al.[101], it was highlighted that, whereas the BB-loop is highly conserved among TIR-domains, the APBS electrostatic surfaces differ. The authors hypothesized that this finding might explain the specificity and selectivity of adaptors recruitment. An experimental study showed that a single point mutation in the TIR domain of murine TLR4 (P712H) renders the system hyporesponsive to LPS

stimulation. The authors noted that their data do not suggest a direct role for this residue.
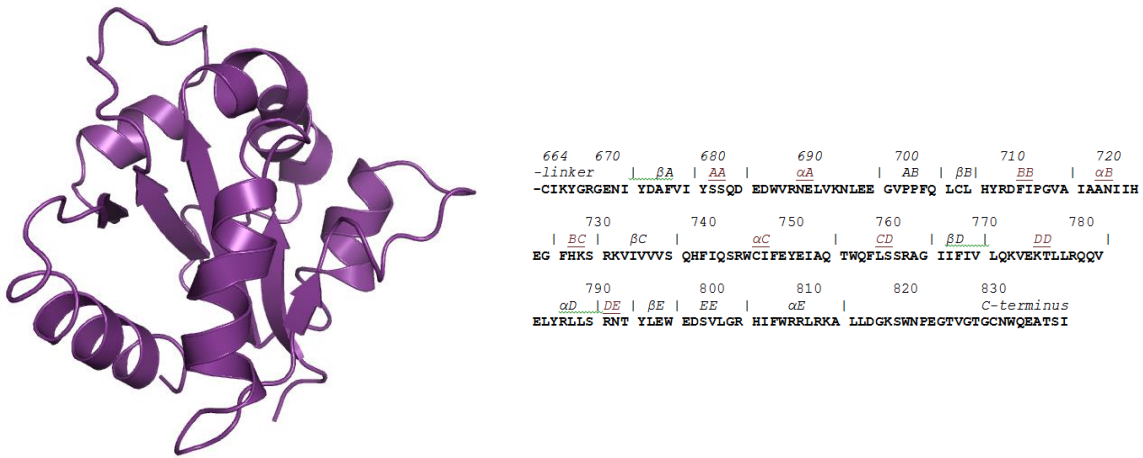

Figure 1.11. Intracellular TIR domain of TLR4. Left: 3D representation of homology model. Right: FASTA sequence.

Dunne et al.[99] used a docking procedure based on hydrophobicity and geometry. Their results suggest that Mal and MyD88 bind at two distinct binding sites (non-overlapping): the DD- and DE-loops of Mal forming interactions with the BB-loop and αC helix of TLR4-TIR domain and the AA- and DD-loop of MyD88 with the CD-loop of TLR4 (Figure **1.11**). The biological relevance of this binding mode was later questioned, as it was discovered that TLR4 activation required homodimerization. In line with that, in 2007, Miguel et al.[102] reported the first 3D model of the dimer of the TIR domain of TLR4; a dimer composed of two identical subunits, arranged in a two-fold axis of symmetry (Figure **1.12a**). Despite the observation that some loops are differently oriented, the overall monomeric fold and the secondary structure of each subunit are very similar to the monomer model reviewed above.[99] This dimer model outlines significant interactions between the BB-loops of each monomer. A flat, but slightly curved surface was observed and attributed to the side facing the membrane. The authors also reported a docking study of TRAM and Mal with the TLR4 dimeric model in which the two adaptors bind at either sides of the dimer interface formed by the union of the two TLR4-TIR domains, which are identical due to the symmetry. The

residues of the adaptors found at the TLR4 interface are mostly located on the BB-loop suggesting that the BB-loop of all three TIR-containing structures is of critical importance for binding specificity and selectivity.

Gong et al.[101] performed a docking study based on the geometry, hydrophobicity and electrostatic complementarity of the molecular surface reporting a dimeric model different from the model described above (Figure **1.12b**). In another study, Basith et al.,[103] used in silico approaches (homology modeling, protein-protein docking and MD simulations) to investigate the inhibitory effect of ST2L toward TLR4 activation. ST2 (IL-33r) is a member of the Toll-like/IL-1 receptor superfamily known to negatively regulate MyD88-dependant signaling pathway. The authors reported a TLR4-TLR4 homodimer model[102] (Figure **1.12a**), and their docking study also gave a similar binding mode for Mal (at each side of the dimer). Their results indicate that MyD88 is recruited by Mal, and that ST2L prevents the recruitment of MyD88 by binding at the Mal interface. Thus, according to these results, ST2L successfully competes with MyD88 to bind at the Mal interface.
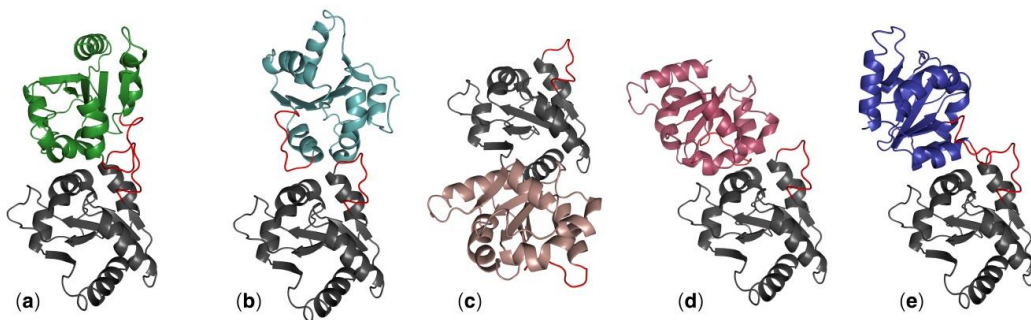
Figure 1.12. Representation of the different ways the dimer is proposed by published computational strategies to be assembled in the literature by computational strategies. (**a**) First reported by Miguel et al.[102] ; (**b**) reported by Gong et al.[101], (**c**–**e**) reported by Guven-Maiorov et al.[104]. The monomer has been built by homology modeling, and the secondary structure representation has been altered to resemble the other models. The dimers have been assembled manually, fitting as best as possible the schemes present in each paper, to provide an overview of the variety of binding poses reported so far. The dimmers shown do not have the pretention of being as precise as those shown in the original papers and should be considered schematic.

26

In a later study, Bovijn et al.[105] reported a homology model constructed based on the crystal structure of the dimeric TLR10 TIR domain. This model is also in agreement with the first model reported by Miguel et al.[102]. The authors proposed that Mal and TRAM adaptors are competing for binding an extended site formed by the reunion of two TLR4 intracellular domains. An experimental mutation study showed that all mutations that impaired Mal binding also impaired TRAM binding, strengthening the idea that Mal and TRAM bind to the same molecular surface. They define the TLR4/TLR4* dimer interface as binding site II, composed of residues from the BB-loop, DD-loop and αC (Figure **1.11**). Then, they describe that the binding site for TRAM and Mal is formed by the reunion of two site I (as defined in the study: residues from αA αB BB and BC), which is in disagreement with the binding site proposed by Miguel et al.[102]. The authors thus argue that their model is supported by experimental data and residue conservation analysis. The binding site III is defined as being located at the opposite direction of binding site I and might be implicated in the interferon regulatory factor 3 (IRF-3) activation.

Singh et al.[106] studied the importance of the highly conserved β-sheets among TLRs' TIR domain and revealed their primordial implications in the communication network. MD simulations of 100 ns of models based on sequence similarity were performed.. They reported interactions between the backbone atoms of the first β-sheet with the BB-loop and the third β-sheet. The authors identified four interacting hubs mainly constituted of hydrophobic residues. Among them, three are in the β-sheets just before the BB-loop, the αC helix and the DD-loops, stressing their role in TIR/TIR interaction. This hypothesis was further supported by analyzing the mutations known to completely abrogate signaling. They show that mutantsIFI767-769AAA and L815A disturb the interacting network, thus explaining the impaired TIR domain homodimerization capacity. In a very recent paper by Guven-Maiorov et al.,[104] the authors used computational techniques to describe the architecture of the signalosome of TLR4. They built three models of the intracellular part of the TLR4 protein (Figure **1.12c–e**). These three dimer models are all unprecedented despite that the secondary structure of the monomer is in great agreement with all of the

published models. Furthermore, the authors used two of their models (Figure **1.12c,d**) to propose different binding modes for Mal.

1.2.3 Binding Mode of TLR4 Modulators

Several computational studies have been performed in order to clarify the binding mode of TLR4/MD-2 agonist and antagonist ligands. The unveiling of the molecular recognition process at atomic detail is one of the major challenges in TLR4/MD-2 modulation. Molecular modeling, docking studies and MD simulations have already provided relevant contributions about the ligand/receptor interactions with promising impact for rational drug design.[79]

Synthetic LPS mimetics

Inspired by the LPS structure, different ligands have been designed and synthesized. One of the first compounds to enter into clinical trials was Eritoran, a synthetic lipid A mimetic, potent TLR4 antagonist, which reached phase III clinical trials as an antisepsis agent, but failed since the study did not meet its primary endpoint of reduction in 28-day all-cause mortality in patients with severe sepsis.[107] Eritoran is a tetraacylated lipid A, the structural analogue of lipid A from RsLA, antagonist of human TLR4 and agonist of TLR4 from mouse and horse. In order to analyze the species-dependent activity of Eritoran, Scior et al.[108] built homology models by means of the SCWRL4 program.[109] Afterwards, docking of Eritoran was performed with AutoDock in order to determine the characteristics of the agonist/antagonist binding in the TLR4 structures from different species: human, mouse and horse. Some key amino acids were identified as relevant in species-specific binding: Lys58 (that corresponds to Asn in mouse and to Glu in horse), Lys388 (which is a Ser in mouse and a Lys in horse) and Gln436 (which is an Arg in mouse and a Gln in horse). The different pattern of interactions that are presented by these different residues impairs the TLR4-TLR4* bridging role of the ligand, thus preventing the effective dimerization and the agonist activity.

Modifications of the chemical structure of the lipid A scaffold have served as a starting point for the design of novel TLR4 modulators. One modification reported by Cighetti et al.[58] was the diphosphorylation of the scaffold of lipid X (Figure **1.10**), a biosynthetic precursor of lipid A, leading to Compound **1** (Figure **1.13**), which has been found to be an antagonist in both human and mouse TLR4/MD-2. This compound was also shown to stimulate CD14 internalization in bone-marrow-derived murine macrophages, thus demonstrating targeting of also CD14 in a TLR4-independent manner. In order to propose 3D models for the ligand recognition processes, computational studies were undertaken on both CD14 and MD-2 proteins. Docking calculations in MD-2 with AutoDock and AutoDock VINA,[89] followed by MD simulations of the resulting complexes with the Impact program,[110] led to the identification of two possible binding poses: the most stable one (in terms of predicted binding energy) allocated both FA chains inside the MD-2 binding pocket, mimicking the lipid IVa binding to MD-2 in the crystal structure (PDB-ID: 2E59). The MD-2/Compound **1** complex is stabilized by hydrophobic interactions between its two FA chains and aliphatic and aromatic residues from the MD-2 pocket together with polar interactions at the rim of MD-2, involving mainly the phosphate groups and side chains from Ser118 and Arg96 residues and, in some cases, interactions between the amide CO or ester CO groups from Compound **3** and the Ser120 OH group. This result was in agreement with NMR experiments performed by the authors that clearly showed FA chain-protein interactions. In a few cases, calculations predicted a second docked binding pose for Compound **1** presenting only one FA chain inside the MD-2 hydrophobic pocket, while the second FA chain was lying over Ile124. Interestingly, in the agonist conformation, this residue has moved towards the inside of MD-2 and Phe126 occupies its place. This synchronism allows the agonist/antagonist switch. In the bound/unbound equilibrium, this alternative binding pose could co-exist with the first and most stable one. Cighetti et al.[58] also combined docking and MD simulations to propose a binding mode for Compound **1** with CD14. CD14 is also characterized by having a wide lipophilic pocket, but with fewer polar residues at the rim. Compound **1** was predicted to bind with the saccharide moiety and the phosphate groups at the entrance of the CD14 hydrophobic cavity and with the FA chains inside the pocket, in

agreement with the CD14 binding properties observed experimentally. In addition to its properties to prevent TLR4 signaling, Compound **1** has also been proposed as a promising hit as TLR4 modulator because of its favorable solubility properties and for its lack of toxicity according to the MTT tests.
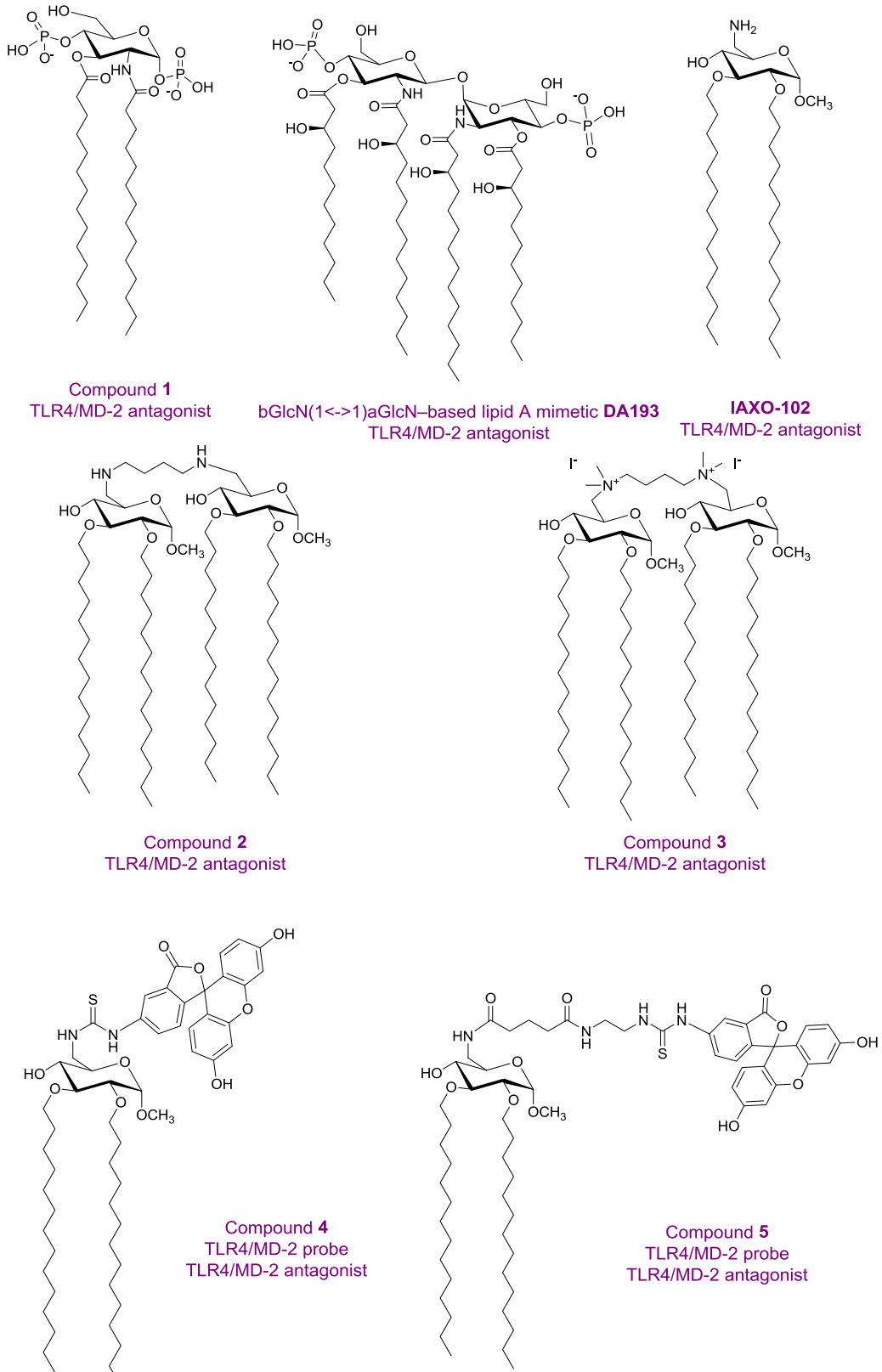
Figure 1.13. Synthetic LPS mimetics studied by computational approaches.

31

Another strategy to mimic lipid A was the design of tetraacylated lipid A mimetics based on the βGlcN (1↔1) αGlcN scaffold analogue by substituting the β(1→6) with βα(1↔1) glycosidic linkage in order to confer rigidity to the molecule.[111] In particular, Compound **DA193** (Figure **1.13**) resulted in being a dose-dependent antagonist in human and mouse, according to assays performed in HEK293 cells transiently transfected with membrane CD14 (*m*CD14)/*h*MD-2TLR4, HEK293 cells transfected with *h*MD-2TLR4 only and assays on human macrophage-like cell line (THP-1). In order to propose an atomistic understanding of the interactions between the ligand and the receptor, MD simulations of 11 ns were performed starting from two possible binding orientations of the ligand into the MD-2 protein: one with the α-GlcN ring facing the Phe126 loop and the second one with the β-GlcN facing the Phe126 loop with an energy difference similar to that found for orientations of lipid A in the binding site of *h*MD-2. Dissociation constants, calculated from MD simulations of the MD-2/DA193 complex, estimated a binding to MD-2 20-fold stronger than lipid A and three-fold more than lipid IVa. It was concluded that the conformational rigidity of the βα(1↔1) diglucosamine backbone of these tetraacylated lipid A mimetics ensures strong binding to MD-2, in two possible binding poses, unlike the native lipid A structures.

The commercial TLR4 antagonist IAXO-102 (Figure **1.13**)[112] has also served as inspiration for the rational design of TLR4 modulators and probes. The cationic glycolipid IAXO-102, a potent TLR4 antagonist targeting both MD-2 and CD14 co-receptors, has been used as scaffold to design new potential TLR4 modulators and fluorescent labels for the TLR4 receptor complex (membrane TLR4/MD-2 dimer and CD14). The primary amino group of IAXO-102, not involved in direct interaction with MD-2 and CD14 receptors, has been exploited to covalently attach a fluorescein (Compound **4** and **5**) or to link two molecules of IAXO-102 through diamine and diammonium spacers, obtaining 'dimeric' Compound **2** and **3** (Figure **1.13**). The structure-based rational design of compounds **2**-**4** was guided by the optimization of MD-2 and CD14 binding. Compounds **4** and **5** inhibited TLR4 activation, in a concentration-dependent manner, and signaling in HEK-Blue TLR4 cells. The fluorescent labeling of murine macrophages

by compound **4** was inhibited by LPS and was also abrogated when cell surface proteins were digested by trypsin, thus suggesting an interaction of fluorescent probe **4** with membrane proteins of the TLR4 receptor system (See Chapter 4).

Computational studies of natural LPSs

Rhodobacter sphaeroides lipid A (RsLA, Figure **1.10**)[84] has five acyl chains, with one unsaturated and two shorter chains than *Escherichia coli* lipid A. It is an antagonist in human and mouse, but an agonist in horse, although, intriguingly, the horse TLR4/MD-2 sequence is more closely related to the human sequence than to the mouse one. To clarify the species-specific response, a computational-aided study of the three 3D structures was undertaken. A homology model was built for horse and hamster TLR4/MD-2, with human and murine X-ray crystallographic structures as templates (PDB-ID: 3FXI and 2Z64) by means of MODELLER.[113] The role of Arg385 had been proven in horse TLR4 complex activation by lipid IVa[114] through polar interactions between the guanidinium moiety and the phosphate group of lipid IVa. In fact, in other species, this residue is substituted by glycine in human and hamster and by an alanine in murine. The docked structure with AutoDock VINA of the horse TLR4/MD-2/RsLa complex closely resembled the pose of lipid IVa in murine crystal structure of TLR4/MD-2. On the contrary, the docked binding pose found in the hamster MD-2 was similar to the lipid IVa pose in the crystal structure from chicken (PDB-ID: 3MU3) and human. The difference between the species was mainly attributed to the different characteristic of each protein. By docking studies on *h*MD-2 with AutoDock, it has been observed that the longest chain of RsLPS could be accommodated in MD-2 by folding the chains itself as has been observed with the Eritoran fatty acid chains. The polar head (diglucosamine) is always exposed to the solvent.

Molecular modeling by Irvine et al. has also showed that the different human/horse TLR4 responses towards RsLA is related to two different amino acids, Gly384 and Ser441, in human TLR4 (Arg385 and Pro442 in horse).[115] Residue Arg385 in horse TLR4, although located around a 9 Å distance from the docked RsLA, could

establish a long-range electrostatic interaction with a phosphate group of RsLA, while the Pro442 is situated near the dimerization interface with TLR4* and interacts with an FA chain of RsLA by van der Waals interactions. This hypothesis was confirmed by experimental assays with transfected HEK293 cells with G384R/S441P hTLR4 with eqMD-2 and R385G/P442S eqTLR4 with hMD-2. It was observed that the R385G/P442S mutations in horse caused a complete loss of activity, and in human, the double mutant G384R/P441S TLR4 was unable to activate the signaling event. Since the double mutation did not revert the activity, other residues must be required. The docking of RsLA in human TLR4/MD-2 shares some similarity with the Eritoran crystal structure, such as the folding of the longest acyl chain and the polar interaction with charged residues of MD-2. RsLPS can adopt two orientations depending on the position of 1-PO$_4$ (oriented towards primary TLR4 in the case of horse and towards partner TLR4* in the case of human). This fact leads to different contacts between acyl chains of RsLPS and the hydrophobic pocket of MD-2. Moreover, superimposition of docked RsLA with X-ray crystallography poses of lipid A and lipid IVa showed that RsLA and lipid A acyl chains occupy more volume than lipid Iva, an, more importantly, the R2 chain of RsLA and lipid A protrudes from MD-2 and establishes interactions with the partner TLR4 in contrast to the R2 chain of lipid IVa, which is folded into the MD-2 pocket.

As mentioned above, the severe pathogen *B. cenocepacia* LPS has been reported by Di Lorenzo et al. to strongly activate human TLR4/MD-2, despite the fact that its lipid A has only five acyl chains.[90] The Ara4N residues in lipid A have been shown to contribute to TLR4-lipid A interactions, and experiments in a mouse model of LPS-induced endotoxic shock confirmed the proinflammatory potential of *B. cenocepacia* penta-acylated lipid A. A combination of docking calculations and MD simulations, together with experimental mutagenesis of the TLR4/MD-2 interacting surfaces, suggested that the longer acyl chains allow reaching deeper regions inside the MD-2 pocket, thus compensating the absence of one FA chain and, at the same time, allowing the exposure of the fifth FA chain on the surface of MD-2. This enables interactions with partner TLR4* and promotes its dimerization. The replacement of Val82 by Phe enhanced the inflammatory response, and it was related to the changes of van der Waals

interactions into stronger CH–π interactions with the FA chain, longer than the corresponding one on *E. coli* LPS. Interestingly, the presence of the positively-charged ammonium groups in the Ara4N seems to favor the electrostatic interactions and, consequently, the binding, whereas uncharged amino acids are critical for responses to *Bordetella pertussis* lipid A, for example.[116] As described in Section 1.2.1, this model for the TLR4/MD-2/LPS$_{BC}$ complex was used to generate a computational mutant TLR4/MD-2/LPS$_{BC}$ complex (D294A, R322A, S415A* and S416A*), which was submitted to MD simulations and energy analysis for quantification of the per residue contributions to the final binding energy.[87] Altogether, these results provided a molecular model for the activation of the human TLR4/MD-2 complex by penta-acylated lipid A, which sheds some light onto the comprehension of the molecular recognition of LPS by TLR4/MD-2.

Computational Studies of Non LPS-Like TLR4 Modulators

The species-specific discrimination of TLR4 ligands by MD-2 is exemplified by taxanes, in particular paclitaxel (PTX; Figure **1.14**), a proinflammatory murine TLR4/MD-2 ligand, which activates the subsequent inflammatory cytokine response.[117-119] Zimmer et al. demonstrated with different experiments that the activation of TLR4 by PTX requires the *m*MD-2 protein, being independent from TLR4 species.[120] This requirement is due to the electrostatic potential surfaces, hydrophobicity, binding pocket size and the conformational gating of the 123–130 amino acids loop. *h*MD-2 and mMD-2 have a very large cavity volume that in principle allows lipid IVa, PTX and Eritoran to fit inside. The computational study identified the key PTX/protein interactions responsible of the differences in the mouse/human TLR4/MD-2 binding mode and of the subsequent different agonist/antagonist behaviour. In the best predicted MD-2/PTX binding poses, the benzamido group of PTX is very close to Phe126, suggesting that a π-stacking interaction may exist between both aromatic groups. Also, the Lys125 side chain establishes a hydrophobic contact with the phenyl ring, and the phenyl group of PTX establishes a cation-π interaction with the Lys122 side chain, which is the only different amino acid in the MD-2 species-conserved sequence Phe119–Gly123. In

*h*MD-2, the multiple interactions attract the Gly123–Lys130 loop to form a concave surface facing the docked PTX. The same loop in the mouse protein is oriented in the reverse direction. The presence of a Glu122 instead of the Lys122 in *m*MD-2 leads to a completely different binding pose, possibly due to the absence of the cation-π interaction.[120] Other work by Resmana et al.[119] proposed a similar binding mode for paclitaxel and the analogue docetaxel, on the basis of docking performed with AutoDock in *h*MD-2 (PDB-ID: 2E59). Also in this case, the most favorable docked binding poses of both taxanes oriented the benzoyl group towards the nearby region formed by Ile61, Phe76, Leu78, Phe119 and Phe151 of *h*MD-2.
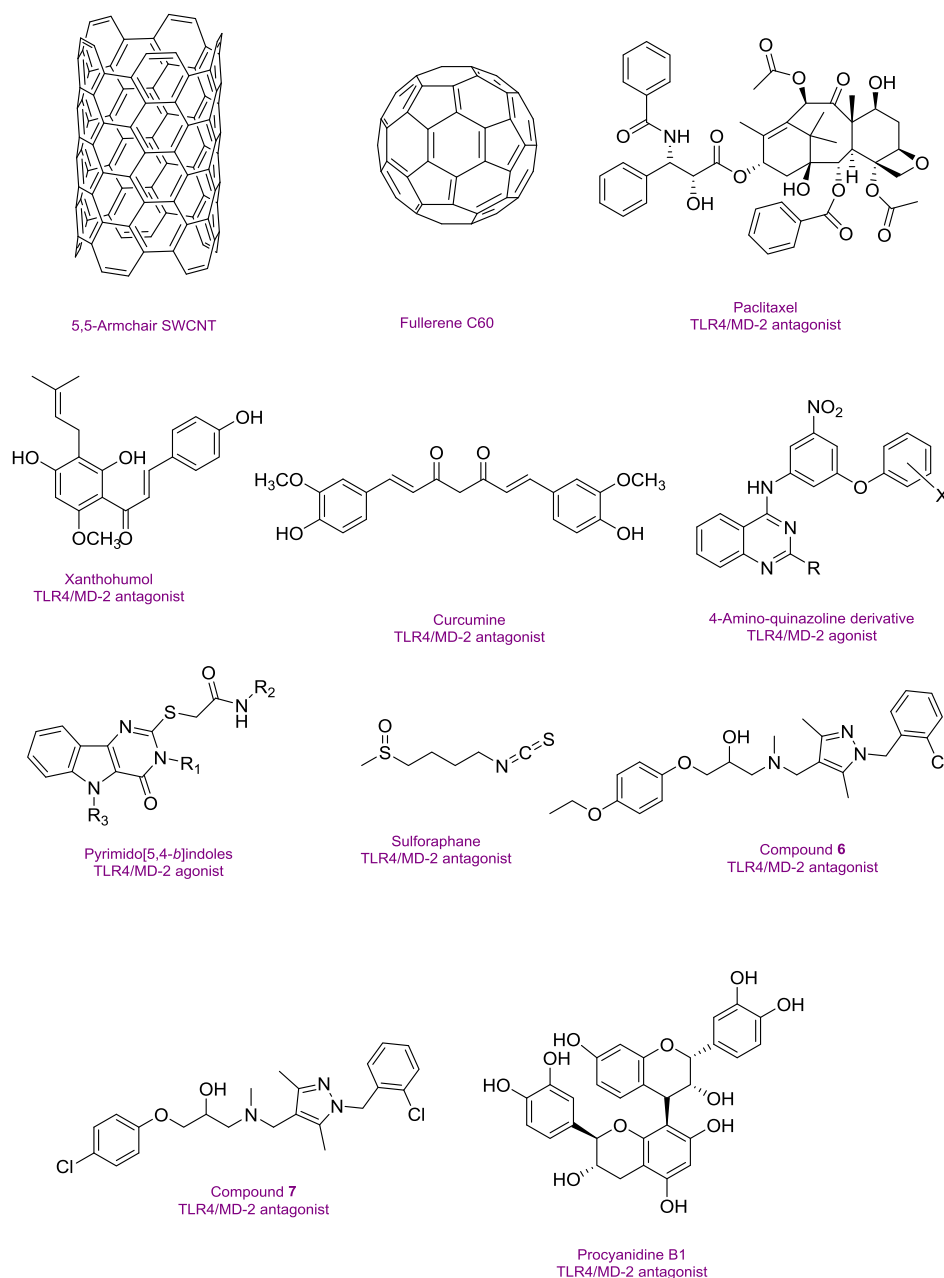
Figure 1.14. Non-LPS-like TLR4/MD-2 modulators studied by computational approaches.

A docked binding mode for a prenylated chalcone-type (xanthohumol; Figure **1.14**) into the antagonist conformation of *h*MD-2 (PDB-ID: 2E59) has been proposed by Fu et al.[121]. The results highlighted the importance of the H-bonds between the OH groups present in the xanthohumol and residues Tyr102 and Arg90. Moreover, another H-bond between the OH of the phenolic group and Glu92 was identified from the docking studies, but this interaction was rapidly broken during the subsequent MD

simulation (50 ns), leading to a final MD-2/xanthohumol complex stabilized by the above-mentioned interactions. An analogue behavior was found for curcumin (Figure **1.14**) after docking with AutoDock also in the same crystal structure of hMD-2 (PDB-ID: 2E59). The hMD-2/curcumin complex resulting from the docking was subjected to MD simulations leading to a stable complex with equivalent interactions with Tyr102 and Arg90. Accordingly, experimental studies with MD-2 mutants (MD-2R90A/Y102A) have pointed to a direct binding of curcumin to MD-2 in the same binding site as LPS. This ligand would occupy a large part of the hydrophobic pocket and form H-bonds with residues Arg90 and Tyr102. Analogously, the H-bond with Gly92 was broken during the simulation. In addition, MD simulations have revealed that the presence of the ligand stabilizes the complex. In particular, MD simulations of the apo-state and bound state of MD-2 have shown that, in the case of the apo-state, MD-2 suffers an important conformational change, reducing the volume of the cavity entrance, in agreement with other similar MD simulations performed on the apo MD-2, whereas the bound MD-2[82, 91] shows good stability.[122]

Cell-based high throughput screening (HTS) allowed the identification of novel chemical entities as potent NFκB activators as selective TLR4 ligands: substituted pyramid[5-4-b]indole derivatives[49] and 4-amino-quinazolines.[123] From the former family, one hit compound was selected (Figure **1.14**; R_1 = phenyl, R_2 = cyclohexyl, R_3 = H). A series of pyrimido[5,4-b]indole rings with carboxamides substituted with various alkyl, cycloalkyl, aromatic and heteroaromatic groups was synthesized and biologically tested in order to establish the SAR. One of the most active compounds (Figure **1.14**; R_1 = phenyl, R_2 = 3,3-dimethylbutyl, R_3 = H) was docked in the mouse TLR4/MD-2 system. The ligand was predicted to bind within the LPS-binding pocket. forming H-bonds and multiple hydrophobic interactions. This computational study supported that active compounds appeared to bind primarily to MD-2 in the TLR4/MD-2 complex.

From the second HTS, one 4-amino-quinazoline (Figure **1.14**; R = COOEt, X = H) was identified with selective agonist activity for human TLR4/MD-2 rather than mouse.[123]. Moreover, the results from the computational study underlined the importance of the Lys122, which happens to be a glutamic acid in mouse. This could

produce an electrostatic repulsion effect with the nitro group, thus justifying the decreased activity in mTLR4/MD-2.[123] Several analogues were synthesized to establish the basis for SAR, confirming the relevant role of the nitro group for the TLR binding and guiding further optimization of the lead compound.

It was shown by liquid chromatography-mass spectrometry analysis that the compound termed sulforaphane (SFN; Figure **1.14**) forms a covalent bond with the residue Cys133 of *h*MD-2. Covalent docking methods were applied in an attempt to explain the propensity of SFN to impair LPS engagement with the MD-2 hydrophobic pocket. The authors proposed a model in which SFN, once covalently linked to Cys133, occupies the same position as the R3" lipid chain of LPS (PDB-ID: 3FXI) and XA2 lipid chain of lipid IVa (PDB-ID: 2E59). This model suggests that SFN sterically prevents other LPS/lipid A from approaching or settling inside the pocket.[124] The same mechanism was reported for the caffeic acid phenethyl ester compound, using only experimental methods.[125]

A series of compounds built by functionalizing pyrazole rings was reported by Bevan et al.[126] to inhibit TLR4 activation. Experimental studies indicated that two compounds (Compounds 6 and 7; Figure **1.14**) were the lead inhibitors. The results indicate that both compounds independently bind at the surface of TLR4 where a protruding loop of MD-2 is normally found in the crystal structure. These predicted binding modes suggest that these compounds compete with MD-2 for binding TLR4, thus preventing or impairing the formation of the TLR4/MD-2 complex, resulting in a TLR4 able to carry out its innate immunity role.

Polyphenol procyanidin B1 (Figure **1.14**) has been shown to be able to regulate innate and adaptive immunity by, inter alia, impairing LPS-induced inflammatory responses in human monocytes.[127-129] In order to explain its mode of action at atomic level, the authors undertook experimental and docking studies.[130] They noted a high degree of similarity in terms of the interactions found in the predicted binding pose with the TLR4/MD-2 system when compared to the interactions established by LPS with TLR4/MD-2 in the crystal structure (PDB-ID: 3FXI).

Computational Studies of Fullerenes, Nanotubes and Dendrimers as TLR4 Ligands

Large molecules have also been subjected to computational studies to unveil their mechanism of action as TLR4/MD-2 binders. Among them, we found interesting examples, such as fullerenes, carbon nanotubes (CNT) and dendrimers. Several studies have indicated a strong impact of carbon nanostructures on the immune system by inducing pro-inflammatory activity through their recognition as pathogens by the TLRs.[131-132] Turabekova et al. have undertaken a theoretical study to analyze 5,5-armchair SWCNT and C_{60} fullerene (Figure **1.14**) interactions with the available X-ray structures of TLRs homo- and hetero-dimer extracellular domains.[133] The authors have searched possible binding sites able to host such nanostructures by identifying the most favorable pockets in terms of hydrophilicity/hydrophobicity and size. In the case of TLR4, the MD-2 binding pocket was detected as the possible binding site. The nanostructures were docked in the environment of the hydrophobic pocket where it interacts with aromatic residues (Phe and Tyr side chains) through π–π interactions and with aliphatic residues through CH–π and lipophilic interactions (Leu, Ile, Ala, Val and Pro).A pair of Lys residues from the rim were found to be accessible for establishing π-cation bonding.

On the other hand, Barata et al.[64] have shown that partially glycosylated polyamidoamine (PAMAM) dendrimer inhibits TLR4/MD-2/LPS-induced inflammation. Molecular modeling studies indicate that the hydrophilic surface bind to the entrance of MD-2 cavity.. Crucially, dendrimer glucosamine interferes with the electrostatic binding between LPS and polar residues Ser118, Tyr102 and Lys91 of MD-2. It was also determined that the bioactivity was due to their surface properties, such as the electrostatic and polar surface, their flexibility and their density.

Computational Studies of Proteins as TLR4 Modulators

Peptide-related molecules have also been explored as putative TLR4 modulators by computational strategies aiming to shed light onto their mechanisms of interaction. Among them, S100A8 is a small protein expressed in neutrophils and platelets, among

other cells, and it can be recognized by TLR4, as part of damage-associated molecular patterns,[134] thus activating TLR4-mediated immune response. To study how TLR4 recognizes S100A8, a rigid body docking was performed.[135] Human S100A8 crystal structure (PDB-ID: 1MR8) was docked on mouse TLR4/MD-2 crystal structure (PDB-ID: 2Z64) using ZDOCK,[136] followed by a clustering/re-ranking method. Fifty-four thousand structures were initially generated and ranked, taking into account desolvation and electrostatic energy and shape complementarity. The top five models were examined as possible complex structures. In all of these models, C-terminal residues of S100A8 are located on the interface with TLR4/MD-2, in agreement with experimental data, suggesting that the *C*-terminal region plays a crucial role in TLR4/MD-2/S100A8 recognition.

Another protein, annexin A2 (AnxA2), has been demonstrated to activate human macrophages through TLR4-mediated signaling. Annexins are calcium-dependent proteins that are involved in cell motility, endocytosis and ion channel formation, among others cellular processes.[137] Recently, experimental data suggested that AnxA2 binds to the TRIF/TRAM/TLR4 internalized complex, although the mechanism remains unclear. Protein-protein docking showed how this complex is formed.[138] Since there is no crystal structure available for neither TRIF, nor the TRAM protein, 3D structures were constructed with the SWISS-MODEL homology modeling server.[139] The TRIF model was docked on TRAM using ZDOCK program. The best predicted TRIF/TRAF complex was then docked on mouse TLR4 (PDB-ID: 3VQ1), and finally, human AnxA2 (PDB–ID: 4HRE) was docked on this complex. The results showed that the complex is formed through both electrostatic and hydrophobic interactions..

A similar approach was used with another protein, the surfactant protein A (SP-A).[140] This protein downregulates inflammation, binds to TLR4 and stops cytokine release. A protein-protein docking of SP-A trimer on the TLR4/MD-2 complex was performed using GRAMM-X.[141] Among the 100 poses predicted by the server and taking into account experimental data suggesting that SP-A mainly binds to MD-2, only three poses were kept. To identify the interacting residues, binding hotspots were predicted using shape specificity and biochemical contact features. Twelve residues of SP-A were

found to interact with the TLR4/MD-2 complex. Using this information, a 20-residue peptide (SPA4) containing the interacting residues of SP-A was synthesized, and it was showed to bind to TLR4 and suppress an inflammatory response.

1.2.4   Virtual Screening in Toll Like Receptors

In the context of drug discovery, virtual screening (VS) techniques have already proved to make hit identification more goal-oriented, allowing the access to a huge number of chemically diverse binders (from public and commercial databases) with a relatively low-cost in terms of time and materials. This computational approach has been subjected to extensive attention and revision over the years, from the early perspective of being an emerging method,[142] until the current time where new challenges are faced.[9, 143-147] We could say that TLRs are not standard receptors which could be approached following classical strategies in drug design. The complexity of the system and the characteristics of their complexation with the PAMPs make them especially difficult to tackle following classical procedures in drug design and discovery. This is why TLRs constitute a special case study in this context. We herein report successful cases of VS approaches that have led to TLR modulators either with agonist or antagonist activity.

Virtual Screening Studies in TLR2

TLR2 heterodimerization either with TLR1 or TLR6 mediates specific ligand recognition of bacterial lipopetides.[148] The X-ray crystallographic structures of both extracellular heterodimers have been resolved assisted by homology modeling in the past few years in complex with the triacylated[65] and diacylated[149] synthetic lipopeptides Pam3CSK4 and Pam2CSK4, respectively. The crystal structure of the TLR2/TLR1 heterodimer[65] with the triacylated lipoprotein revealed that the two ester-linked lipid chains are inserted into the large TLR2 pocket in extended conformation, and the remaining amide-bound lipid chain is inserted into a narrow channel present in TLR1. The binding site is mainly composed of hydrophobic residues from Leucine-rich

repeat (LRR) modules 9-12 in both receptors. The peptidic head establishes contacts with polar groups from Phe349 from TLR2, and Gly313 and Gln316 from TLR1. Interestingly, in the case of the (mouse) TLR2-TLR6 heterodimer co-crystalized with the diacylated lipopeptide, the TLR2-lipid interaction and strong PPIs seem to be the prime force for heterodimerization and signalling since TLR6 channel is shortened by the presence of the bulky side chains from Phe343 and Phe365. A H-bond between the Phe319 (TLR6) backbone and the first peptide bond of the lipopeptide is herein detected.

Regarding the application of VS tool for the finding of novel TLR2 modulators, Zhong *et al.*[150] report the identification of a natural product-like inhibitor of TLR2/TLR1 heterodimerization (code ZINC12899676, Table **1.2**) following a structure-based VS strategy, through the docking of a collection of natural products and natural product-like compounds from ZINC database (> 90 000 compounds) to a TLR2/1 ectodomain model based on the TLR2/TLR1/Pam3CSK4 crystal structure (PDB-ID: 2Z7X). Flexible ligand docking was performed using the virtual library screening module in the ICM-Pro program[151] at the TLR2/TLR1 heterodimeric interface. The 17 best ranked solutions according to the Full ICM Score compounds were selected for biological testing. Among these 17 compounds, compound ZINC12899676 (Table **1.2**) could decrease the secretion of pro-inflammatory cytokines TNF-α and IL-6 in RAW 264.7 macrophages stimulated with the most studied TLR2/TLR1 agonist, Pam3CSK4. It showed that could reduce the secretion of TNF-α by 44% over the concentration range of 0.25 to 4 mM, with an IC$_{50}$ value of *ca.* 6.1 mM, and the secretion of IL-6 by 56% on the concentration range of 0.25 to 2 mM, with an IC$_{50}$ value of *ca.* 1.9 mM, displaying similar potency to the only other TLR2/TLR1 small molecule antagonist reported to date (CU-CPT22)[152-153] with no cytotoxic activity being detected. Compound ZINC12899676 also demonstrated its ability to reduce the phagocytic activity of RAW 264.7 cells.

The mechanism of the antagonist activity exhibited by compound ZINC12899676 is proposed to be by displacement of the synthetic lipopeptide Pam3CSK4, as shown by docking studies where two key H-bonds were identified:.. Complementary biological and biophysical tests corroborated this possible mechanism
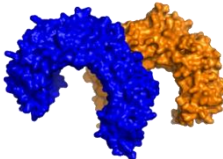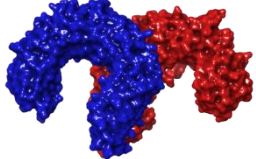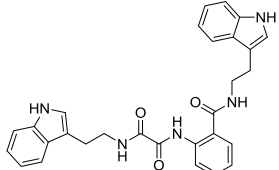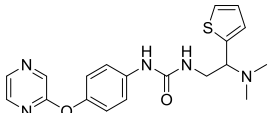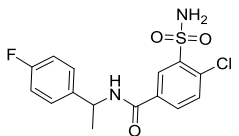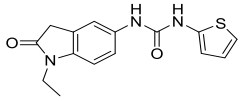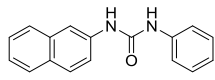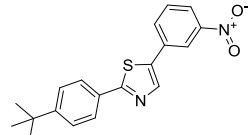
of action. A fluorescence polarization assay demonstrated the ability of ZINC12899676 to disrupt Pam3CSK4-mediated TLR1/TLR2 heterodimerization in a dose-dependent manner, with an IC_{50} value of *ca.* 7.2 mM. An immunoprecipitation assay was used to confirm the inhibitory effect on lipoprotein-induced TLR1/TLR2 heterodimerization exhibiting similar potency to reference compound CU-CPT22. Compound ZINC12899676 could attenuate NF-kB-luciferase reporter assay in RAW 264.7 cells with greater potency than CU-CPT22, and in HEK293T cells transfected with pZERO-TLR1, pCMV-Flag-TLR2 and pNF-kB-Luc, and was able to downregulate IκBα and IKKα/β phosphorylation and IκBα expression *in cellulo*.

Other interesting results in this field are the work reported by Murgueitio *et al.*.[154] The authors report the analysis of TLR2 monomer to predict and locate ligand binding. A subsequent structure-based strategy was followed by centering the VS on the lipopeptide binding site sub-pockets P1-P3. A 3D-pharmacophore model was then constructed using LigandScout[155] revealing a hotspot for H-bond acceptors. Two hydrophobic areas, defined as HYD1 (Ile319, Phe325, and Val348) and HYD2 (Leu266, Phe284, Phe295, Ile314, and Leu328) were also characterized. This model was validated and used to screen a library of more than 2 800 000 commercially available compounds from different vendors (ASINEX, Life Chemicals, Maybridge, ChemBridge, ENAMINE HTS Collection, and SPECS) with the help of LigandScout. 150 compounds with the highest pharmacophore fit score were docked into the TLR2 binding pocket using GOLD[156-158] and, after careful visual inspection, five of them were selected for biological testing on a NF-kB reporter assay in the cell line HEK293-TLR2. Compound with code MolPort-001-796-266 (Table **1.3**) exhibited antagonistic activity, and the IC_{50} value was measured in human monocytes obtaining μM values. Its presumed binding mode was studied by means of docking techniques into the TLR2 binding site, displaying key H-bond and hydrophobic interactions with residues located deep inside the TLR2 pocket..

A ligand-based strategy was followed using a shape- and feature-based similarity screening assisted by ROCS and using three reported small-molecule TLR2 signaling modulators[159] and E567[160] against a NCI compound library of 260 071

compounds. Five hundred hits arose from the VS and, after visual inspection, 39 were selected for biological testing. Out of them, 4 exhibited antagonist activity (hit rate: 10%): compounds ZINC16769362 and ZINC398557 (that were identified from compound B as query structure) and compounds ZINC1758666 and ZINC585632 (from E567 as query) (Table **1.2**).

Table 1.2. 2D Chemical structure of TLR2 modulators identified by VS techniques and mentioned in this review. The database codes are provided. [a] MolPort is a supplier of chemicals included in several VS databases (www.molport.com).

TLR2/TLR1	TLR2/TLR6
3D structure from PDB-ID: 2Z7X	3D structure from PDB-ID: 3A79
ZINC: ZINC12899676[150]	ENAMINE: Z416323354[154]
TLR2-TLR1 heterodimerization inhibitor	MolPort[a]: MolPort-009-315-475
	TLR2/1 & TLR2/6 inhibitor
MolPort[a]: Molport-001-796-266[154]	MolPort[a]: MolPort-009-737-181[154]
TLR2/1 & TLR2/6 inhibitor	TLR2/1 & TLR2/6 inhibitor with a decrease of cell viability
ZINC: ZINC1676936[154]	MolPort[a]: MolPort-002-914-354[154]
NCI: Plated 2007 44661	TLR2/1 & TLR2/6 inhibitor

TLR2/1 & TLR2/6 inhibitor
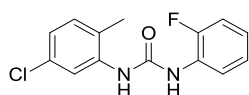
ZINC: ZINC398557[154]

NCI: Plated 2007: 205636

MolPort[a]: MolPort-001-835-401

C29

TLR2 TIR domain inhibitor[161]

TLR2/1 & TLR2/6 inhibitor

ZINC: ZINC1758666[154]

NCI: Plated 2007: 17379

TLR2/1 & TLR2/6 inhibitor
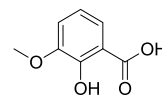
C29L (*o*-vanillin)[161]

TLR2 TIR domain inhibitor
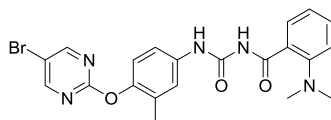
ZINC: ZINC585632[154] TLR2/1 & TLR2/6 inhibitor

The same procedure was repeated using compounds A, B, ZINC16769362, and ZINC585632 as queries, this time against the collection of more than 2 800 000 commercially available compounds used in the structure-based approach. From this procedure, 22 compounds were selected for biological testing and three of them displayed antagonistic activity (Z416323354, MolPort-009-737-181, and MolPort-002-914-354, Table **1.2**). Compounds were also tested for TLR2-specificity and toxicity and the decrease of pro-inflammatory cytokine TNF-α was evaluated in human monocytes.

Additional computational docking studies of ZINC16769362, which showed the lowest IC_{50}, were carried out showing that the ligand is embedded into a narrow sub-pocket at the end of the binding site thus interfering with lipopeptide binding. . The substitution pattern of the phenyl moiety was shown to be crucial for activity, since compounds with other pattern of substitution were inactive, as well as the presence of aromatic rings, as compounds with aliphatic rings were inactive. Overall, these results

shown to be very promising for the identification of several novel TLR2 antagonist with activity in the µM range by using virtual screening techniques.

Mistry *et al.*[161] have also reported the identification of two novel small molecule inhibitors of TLR2 signaling by targeting a pocket within the so-called BB loop of the TLR2 TIR domain. The TIR domain, located on the cytosolic face of all TLRs and adaptor proteins[24] (in TLR2, MyD88 and TIRAP) has been proven to be key for signaling through the mediation of certain homotypic and heterotypic PPIs[162] that triggers downstream signaling cascades and ends in the production of pro-inflammatory cytokines and chemokines.[163] The crystal structures of human TLR2 and TLR1, as well as the P681H mutant of the TLR2 TIR domain[93] revealed that the BB loop connects strand β-B and helix α-B and sticks out of the structure. The P681H mutation in the BB loop has shown to preclude the recruitment MyD88 and therefore TLR2 signaling.

A pocket within this BB loop of *h*TLR2, formed by 10 residues (Tyr647, Cys673, Asp678, Phe679, Ile680, Lys683, Asp687, Asp688, Asp691, and Ser692) neighboring the highly conserved Pro681 and Gly682 pair, was selected as the target for searching new TLR2 modulators. Flexible ligand docking of a collection of commercially available small molecules and FDA-approved compounds (> 1 million compounds) was performed using the DOCK algorithm[164] based on the anchor-and-grow search method.[165] First, a primary docking was performed where each rotatable bond was minimized while created without reminimizing the other bonds, with a minimization of the complete molecule once it was built. The most favored conformation of each molecule in terms of interaction energy was conserved. This resulted in the selection of 50 000 compounds that were subjected to a secondary docking step with an additional simultaneous minimization step of all rotatable bonds against the crystal structure (PDB-ID: 1FYW) and three additional conformations obtained from MD simulations of the protein.[166]

The top 1 000 compounds that exhibited the most favorable interaction energies, taking into account every protein conformation, led to the selection of 149 compounds and 20 FDA-approved drugs attending to chemical diversity and physicochemical properties for biological testing in HEK293T-TLR2 transfectants.

Among them, compound C29 (Table **1.2**) was able to inhibit/disrupt/block both TLR2-TLR1 and TLR2-TLR6 signaling induced by synthetic and bacterial agonist in human cell lines. Pam3CSK4- and Pam2CSK4-induced IL-8 mRNA was decreased by compound C29 in stably transfected HEK-*h*TLR2 in a dose dependent manner, as well as IL-1β gene expression in the human monocytic cell line THP-1. Other effects were not exhibited in other TLR agonist- of TNF-α induced signaling nor cytotoxic effects. This behaviour was also observed when HEK-*h*TLR2 and THP-1 cells were stimulated with heat-killed or live Gram-positive and Gram-negative bacteria.

Notwithstanding, C29 only showed activity on TLR2/1 signaling pathway, disrupting only P3C-and *Staphylococcus aureus* lipoteichoic acid-induced IL-1β mRNA in murine macrophages. C29L (o-vanillin), a derivative from the imine cleavage of C29 in alkaline conditions (NaOH, 65 μM), displayed similar activity and potency in NF-κB luciferase reporter assay in HEK293T cells and has the advantage of a better water solubility. It was shown to be active both *in vitro* and *in vivo*. In this work, Mistry *et al*. also performed an Alanine scanning mutagenesis of every residue within the BB loop using Y647A as a control mutation as it has been reported to play no role in TLR2 signaling.[167] All 10 BB loop pocket mutants resulted crucial for TR2/1 signaling but not for TLR2/6 signaling, were mutations C673A, I680A, K683A, and S692A were found to not be needed for TL2/6 signaling.
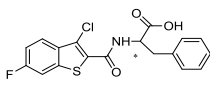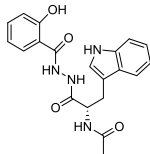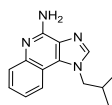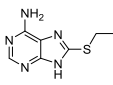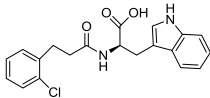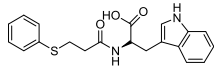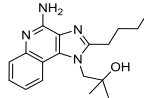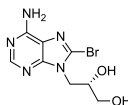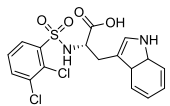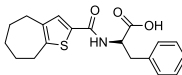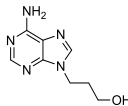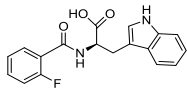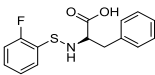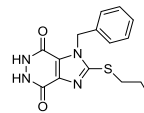
Virtual Screening Studies in TLR3

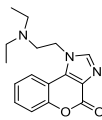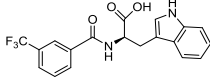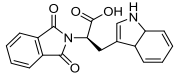Toll-like receptor 3 (TLR3) is located at the membrane of the endoplasmic reticulum, endosomes, multivesicular bodies, and lysosomes. TLR3 forms a large horseshoe shape that contacts with a neighboring horseshoe, yielding a dimer of two horseshoes. The overall horseshoe-shaped structure of the ectodomain TLR3 is formed by 23 repeating LRRs, ligand-binding domain that is composed of leucine-rich repeats (LRRs).[168] Some X-ray crystallographic structures are available from mouse (PDB-ID: 3CIG and 3CIY) and from human (PDB-ID: 2AOZ and 1ZIW). TLR3 recognizes specifically dsRNA, and the activation of the receptor induces the secretion of type I interferons

and pro-inflammatory cytokines, like a TNF-α, IL-1 and IL-6, triggering immune cell activation and recruitment of the adaptor molecule TRIF *via* TIR domain interaction.[169] In contrast to other TLR ligands, dsRNA signaling occurs *via* MyD88-independent pathways.[170] It has also been reported to recognize synthetic analogue polyinosinic-polycytidylic acid, Poly(I:C).[171] Therefore, the TLR3/dsRNA complex constitutes an important target in multiples infectious diseases and cancer, as it has been shown to be implicated in several infection models like a herpes simplex encephalitis,[172] West Nile disease, phlebovirus, vaccinia and Influenza A.[173-176] It has also been reported that double-stranded DNA from necrotic cells during inflammation or viral infection activates the signal of TLR3.[177]

Cheng *et al.* have reported the development of small-molecule probes that exhibited activity as competitive inhibitors of dsRNA binding to TLR3.[178] The authors performed a VS in the dsRNA binding domain of TLR3 using the ENAMINE drug database. The docking protocol was performed into the dsRNA binding domain of mouse TLR3 (PDB-ID: 3CIY) with Glide program. A HTVS protocol was employed for the first docking and ranking, followed by SP protocol for the top 10 000 compounds. The resultant top 5 000 compounds were subsequently docked using the more accurate and computationally intensive XP mode of Glide. First top-ranked 100 compounds were selected and re-ranked by predicted binding energy. The authors finally selected nine hits compounds for evaluation by cell assay of TLR3 activation (ENAMINE codes are: T5528092, T5631009, T5630975, T0519-9149, T5626448, T5643856, T5260630, T55994342, T0505-4844, Table **1.3**).

Table 1.3. 2D Chemical structure of TLR3 and TLR7 modulators identified by VS techniques and mentioned in this review. The database codes are provided.

TLR3		TLR7	
3D structure from PDB-ID: 3CIY		No X-ray crystallographic structure available -	
(R) Compound 4a[178] TLR3 inhibitor	ENAMINE: T5528092[178] TLR3 inhibitor	Query 1 (Imiquimod)[179]	ZINC: ZINC1667204[179] TLR7 inhibitor
T5631009[178] TLR3 inhibitor	ENAMINE: T5630975[178] TLR3 inhibitor	Query 2[179]	ZINC: ZINC39698[179] TLR7 inhibitor
T0519-9149[178] TLR3 inhibitor	ENAMINE: T5626448[178] TLR3 inhibitor	ZINC: ZINC12382420[179] TLR7 inhibitor	ZINC: ZINC36416[179] TLR7 inhibitor
ENAMINE: T5643856[178] TLR3 inhibitor	ENAMINE: T5260630[178] TLR3 inhibitor	ZINC: ZINC4756232[179] TLR7 inhibitor	
ENAMINE: T55994342[178] TLR3 inhibitor	ENAMINE: T0505-4844[178] TLR3 inhibitor	ZINC: ZINC8686004[179] TLR7 inhibitor	

Most of these nine hits resulted to share a structural motif: the chemical structure of a D-amino acid conjugated with an aromatic substituent, thus yielding a new pharmacophore for the TLR3 binding site. To select the best ranked compounds, they took into account different benchmarks: a) predicted binding energy and spatial complementarity; b) reasonable chemical structures found in the dsRNA-binding site of TLR3; c) existence of at least one H-bond between the ligand and one of the dsRNA-recognizing residues on the TLR3 surface (*e.g.* His539, Asn541, and Ser571); d) protonation state and tautomeric form of the ligand had to be acceptable.

A dsRNA, Poly(I:C) was employed to selectively activate TLR3 signaling, resulting in the activation of nitric oxide (NO) synthase and the production of NO in RAW 264.7 macrophage cells.[180] They monitored the NO level as an indicator of Poly(I:C)-induced TLR3 activation to evaluate the inhibitory activity. Hit compounds T5626448 and T5260630, both derivatives of D-Phe, were identified with IC_{50} values of 154 µM and 145 µM respectively. Different analogues were synthesized and SAR analysis was performed. Finally, only one compound, a T5626448 derivative (compound 4a in Table **1.3**), was identified as a very potent dose dependent TLR3 antagonist, with a low µM IC_{50} value (3.44 ± 0.41 µM). However, in the case of T5260630 analogues, not significant improvement in the activity was observed, so they only focused on the T5626448 derivative family.

Compound 4a was also tested against homologous TLRs: TLR1/2, TLR2/6, TLR3, TLR4 and TLR7 using TLR specific ligands, but only TLR3 inhibition was observed. Other different biological assays were performed, finding that compound 4a did not affect cytochrome P450 CYP3A4, CYP2D6, and CYP2C19 isoforms. Tests on RAW 264.7 macrophages were also carried out showing low toxicity, and kinase profiling showed that 4a demonstrates negligible inhibition activity against a panel of 12 representative kinases. Biophysical tests were also carried out, with a negative control, to demonstrate that 4a binds to TLR3. Fluorescence anisotropy assay demonstrated that this compound competes with dsRNA for binding to TLR3 with a K_i value of 2.96 µM. By an ELISA assay, 4a was also demonstrated to inhibit the downstream signaling transduction mediated by the formation of the TLR3/ds RNA complex, showing that

this compound almost completely abolishes the TL3-mediated inflammation response at its IC_{90} concentration (27 μM). Finally the inhibitory effects of TNF-α by compound 4a at 10 μM were also tested with a result of 60% inhibition, agreeing with the results observed in the NO synthase assay.

Virtual Screening Studies in TLR4

In the search of novel TLR4 modulators, Yin *et al.* have applied a computational methodology to the identification of small drug-like inhibitors of TLR4/MD-2 PPIs.[181] The authors have developed a novel *in silico* screening methodology incorporating molecular mechanics (MM) and implicit solvent methods[181] to evaluate binding free energies, in order to improve affinity prediction accuracy without reducing screening speed. The ENAMINE database collection was screened against the TLR4/MD-2 complex of the crystal structure of the human TLR4 TV3 hybrid-MD-2-Eritoran complex (PDB-ID: 2Z65). The library was clustered to ensure the least possible computational work, while keeping as much of the full chemical diversity of the available library as possible. A combination of Jarvis-Patrick and Li algorithms[182-183] was used; as well as the Tanimoto similarity calculation[184-187] with Daylight fingerprints in order to measure the distance between the molecules. About 86 000 clusters were isolated. Then, the compounds representing the cluster centroids were taken, and an additional filter that matched the molecular volume to the biding site was applied.

Fast molecular docking for the generation of binding poses and subsequent MD simulations were performed to rank the ligand poses according to their binding affinities. The hits were profiled against a library of 500 representative human proteins as a selectivity filter in order to remove the non-specific inhibitors. Finally, as a proof of concept, the compounds were screened against both TLR4 and MD-2 to validate the strategy.[181, 188] Two compounds, T5342126 and T6071187 (Table **1.4**) were identified as potential TLR4- and MD-2-specific antagonists, respectively, completely abolishing LPS-induced activation of signaling. Their biological activity and selectivity were tested *in vitro* using Akt1 and nitric oxide in RAW264.7 cells.
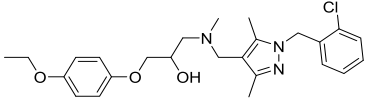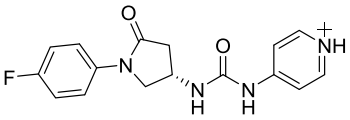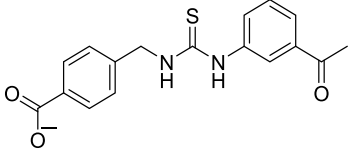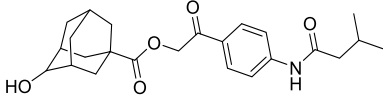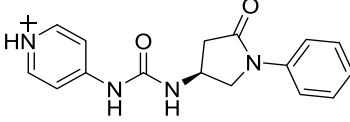
In another study, Gobec *et al.*[188] performed parallel ligand-based and structure-based virtual screenings in order to identify novel TLR4 antagonists targeting the TLR4/MD-2 interface using the crystal structure of the human TLR4 TV3 hybrid-MD-2-Eritoran complex (PDB-ID: 2Z65). For both ligand-based and structure-based virtual screening, they used the ZINC drug-like subset (~ 11.3 million drug-like compounds) from the ZINC database.[189]
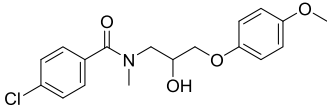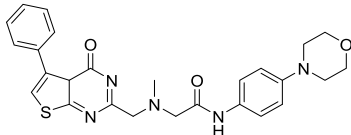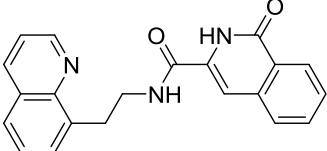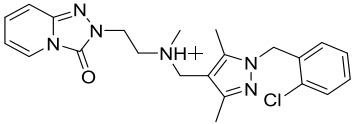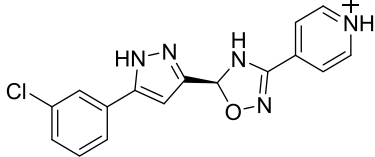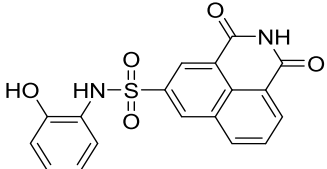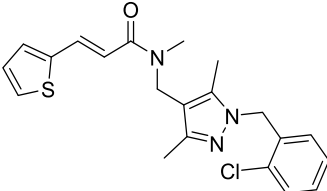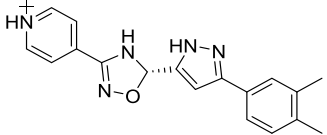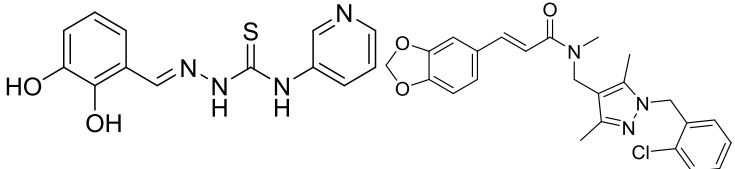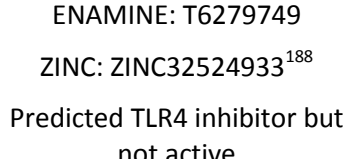
Regarding the ligand-based virtual screening, they used the OMEGA software[190] on the compound T5342126 (Table **1.4**),[188] a known TLR4 antagonist, to generate 5 query conformers. ROCS software was then used to compare the database to all query conformers. The single best overlay hits were ranked according to the TanimotoCombo scoring function,[190] considering similarities in the molecular shape and color of atom types. Thereby 5 compounds were identified (ZINC51408124, ZINC464832, 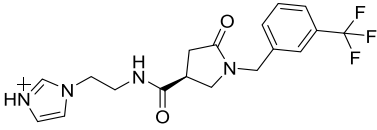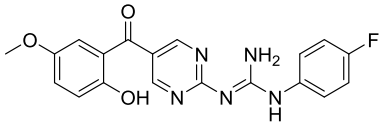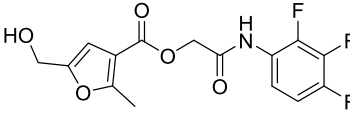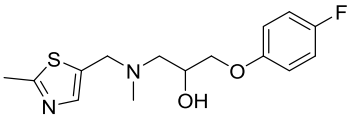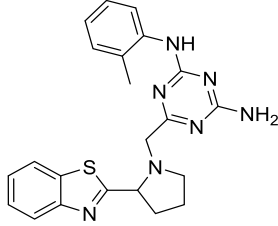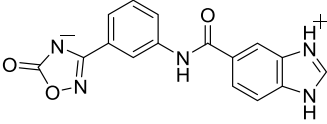ZINC26905159, ZINC32525142 and ZINC32524933, Table **1.4**) and evaluated *in vitro*. Unfortunately, these compounds were either not water soluble, or not active, or presented cytotoxicity on HEK293 cells.

For the structure-based virtual screening, before the docking process, they performed an enriching procedure, using ROCS software between the database and T5342126, the query molecule, in order to reduce the number of compound and to enrich it. Two sets of 25 000 compounds each, were created: set 1 with the highest shape similarities to T5342126, using ShapeTanimoto algorithm, and set 2 with both the highest shape and color (atom type) similarities to T5342126, using the TanimotoCombo algorithm.[191] Both sets were merged and the duplicates were removed, leading to a total of 49 600 unique compounds left. The docking procedure was carried out using FlexX program and the active site was defined as an area of TLR4 within 8 Å around the interacting MD-2 loop (Gly97-Leu108). LeadIT-implemented pharmacophore constraints were performed then in order to keep only the compounds that can form interactions with at least one of the polar amino-acid residues such as Ser183 and Asp209 of TLR4, and Arg106 of MD-2. At the end of the implementation, 25 750 compounds had been kept, and the docking procedure was performed. The compounds have been finally ranked according to their best scoring conformation

using LeadIT score and 40 were selected and assessed *in vitro*. After the first *in vitro* assay, only 14 compounds were sufficiently water-soluble, up to 500 μM, and completely non-cytotoxic at 100 μM. Those received further biological evaluation using HEK-Blue^TM *h*TLR4 cells, and 3 compounds with promising antagonistic activities were discovered: ZINC25778142, ZINC49563556 and ZINC3415865.

Table 1.4. 2D Chemical structure of TLR4 modulators identified by VS techniques and mentioned in this review. The database codes are provided.

ENAMINE: T5342126[192]	ZINC: ZINC04272679[193]	ZINC: ZINC00611718[193]
TLR4 inhibitor	Predicted TLR4 inhibitor	Predicted TLR4 inhibitor
ENAMINE: T6071187[192]	ZINC: ZINC04272561[193]	ZINC: ZINC48141941[193]
MD-2 inhibitor	Predicted TLR4 inhibitor	Predicted TLR4 inhibitor
ENAMINE: T5339238 ZINC: ZINC25778142[188]	ZINC: ZINC09535665[193]	ENAMINE: T6969316 ZINC: ZINC51408124[188]
TLR4 inhibitor	Predicted TLR4 inhibitor	TLR4 activity not determined (solubility problems)

ENAMINE: T5458371

ZINC: ZINC49563556[188]

TLR4 inhibitor

ZINC: ZINC70039563[193]

Predicted TLR4 inhibitor

ZINC: ZINC464832[188]

TLR4 activity not determined
(cytotoxicity on HEK293 cells)

ENAMINE: T5315798

ZINC: ZINC3415865[188]

TLR4 inhibitor

ZINC: ZINC29450369[193]

Predicted TLR4 inhibitor

ENAMINE: T6417643

ZINC: ZINC26905159[188]

Predicted TLR4 inhibitor but
not active

ZINC: ZINC64951618[193]

Predicted TLR4 inhibitor

ZINC: ZINC41124663[193]

Predicted TLR4 inhibitor

ENAMINE: T6280209

ZINC: ZINC32525142[188]

Predicted TLR4 inhibitor but
not active

ZINC: ZINC64951738[193]

Predicted TLR4 inhibitor

ZINC: ZINC08687988[193]

Predicted TLR4 inhibitor

ENAMINE: T6279749

ZINC: ZINC32524933[188]

Predicted TLR4 inhibitor but
not active
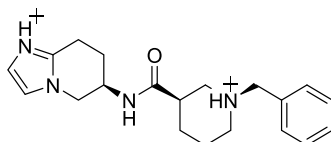
ZINC: ZINC72278680[193]

Predicted TLR4 inhibitor

In other work, Sowdhamini *et al.*[193] used homology modeling, docking, and virtual screening techniques, in combination with known experimental data, molecular mechanics calculation to identify novel and potential small molecule inhibitors of TRAM-mediated TLR4 signaling. For this purpose, they identified TLR10 TIR dimer as the best model to build the TIR domain of TLR4 as a dimer. Then, they modeled the C-terminal region of the A46 poxviral protein containing the VIPER motif, using the crystal structure of A52 poxviral protein (PDB-ID: 2VVW) as a template. This motif is capable to bind the TIR domain of different adaptor proteins. After having obtained the two models, they performed a two-phase docking for creating reliable models of the complex between the TRAM TIR homology model and the VIPER peptide segment. A virtual screening was then, performed onto the complex. They used the lead-like and drug-like subsets of the ZINC database, totaling 32 million of compounds. The ligands 2D structures were converted into their 3D structure including all possible stereoisomers, tautomers, and ionization states under a pH range of 6-8, the hydrogens were added and the structures were optimized and minimized in LigPrep. The library was preliminary screened based on ADMET properties and reactive functional groups, using Qikprop and Lipinski's rule of five.[194] The amino acid residues constituting the BB loop (110-122) and alphaC helix (141-154) of the TLR4 TIR domain were selected for generating the receptor grid.

Glide was used for the docking by concatenating the three protocols: HTVS, SP and XP. The top 10% compounds, based on the Glide score, obtained from the HTVS step were retained for the subsequent step. These were re-docked using the SP module. The XP module was used to perform a more extensive docking of the top 10% compounds carried forward from the SP step. Final ranking of the compounds was

based on their Glide XP Score. The compounds having similar scaffold were then clustered using CANVAS[195-196] resulting in a pool of 265 chemically diverse structures. These selected compounds were submitted to induced fit docking within the Maestro suite[197-199] to restrict the flexibility only into the binding site. For this purpose, they used the Glide SP protocol to generate 2 000 poses for each molecule within the binding site. Finally, they inspected the top 20 receptor-ligand poses for each ligand to see if potential interaction between the binding site residues and the ligand atoms were maintained or disrupted upon incorporating flexibility to the residues, and the ligands with more interactions conserved throughout most of the poses were selected. Binding free energy calculations were performed on the top two poses generated during induced-fit docking of each compound. These complexes between the TRAM TIR homology model and each ligand were ranked according to this analysis and a final structural analysis of the ligand/receptor interactions was performed, shortlisting 12 molecules (Table **1.4**). Interestingly, compound ZINC08687988 remained firmly bound in the pocket even after incorporating a considerable degree of conformational flexibility during the MD simulations carried out in the complexes. To date, no further biological testing has been performed yet.

Virtual Screening Studies in TLR7

Toll-like receptor 7 (TLR7) is intracellularly located at the membranes of endosomes, endoplasmic reticulum, multivesicular bodies, and lysosomes.[200] Its function is related to defense against viral infection by recognizing single-stranded RNA (ssRNA) and small-interfering RNA (siRNA) from viruses,[201-202] including human immunodeficiency virus, influenza, and vesicular stomatitis virus.[203] The host can also utilize TLR7 to detect RNA released into endolysosomes by phagosomal bacteria. Several synthetic ligands have also been reported to modulate TLR7, such as imidazoquinoline derivatives (resiquimod and imiquimod), and guanine analogues.[204] Also, TLR7 recognizes guanosine- and uridine-rich ssRNA, and synthetic polyuridines act as potent ligands.[202] The development of new antagonist modulators could have

important applications for the treatment of autoimmune disorders, like rheumatoid arthritis, Sjogren's syndrome, and systemic lupus erythematosus.[14]

Since no X-ray crystallographic structure of TLR7 is available to date, in order to identify TLR7 modulators, Gobec *et al*. undertook a ligand-based VS.[179] ROCS was employed to carry out the screening protocol from which six compounds with three novel chemical scaffolds were discovered. The authors employed ZINC database and OMEGA software to prepare the compound library. With the help of ROCS, two query compounds were identified as TLR7 binders: query 1 (imiquimod) and query 2 (1-(4-amino-2-butyl-1*H*-imidazo[4,5-*c*]quinolin-1-yl)-2-methylpropan-2-ol) (Table **1.3**). Imiquimod (query 1) is a TLR7 agonist currently used for topical treatment of genital warts caused by human papillomavirus, actinic keratosis, and superficial basal cell carcinoma,[205] and query 2 compound was developed in the last years in a systematic SAR exploration study as the most potent imidazoquinoline with TLR7 agonist activity.[206]

From queries 1 and 2, the authors performed parallel VS studies. The results were ranked taking into account the TanimotoCombo score, and the best results from both VS were finally merged. The best 25 ranked compounds were selected and submitted to biological assays, only considering soluble and available compounds. Cytotoxicity tests were performed with HEK-Blue[TM] *h*TLR7 determined using a propidium-iodide based staining method and none of the compounds showed cytotoxicity at 250 µM. In the subsequent step, the soluble compounds were assayed for TLR7 agonist activity at 250 and 500 µM using the reporter assay but none of the compounds showed any notable agonist activity. Finally, to evaluate the antagonist activity, the compounds were tested using HEK239 cell line co-transfected with *h*TLR7 gene using imiquimod as a control. Six compounds were identified as antagonists at the µM scale containing three novel chemical scaffolds: chromeno[3,4-*d*]imidazole-4-one, 1*H*-imidazo[4,5-*d*]pyridazine-4,7-dione, and 6-amino-9*H*-purine (ZINC codes 12382420, 1667204, 39698, 36416, 4756232, and 8686004, Table **1.3**). The authors also propose a simple and straightforward synthesis of derivatives from the chromeno[3,4*d*]imidazole-4-one scaffold which showed promising TLR7 antagonistic activities.

Virtual Screening Studies in TLR8

Toll-like receptor 8 (TLR8) is an endosomal membrane receptor that recognizes single stranded RNA (ssRNA) from viruses. TLR8[207] is expressed in monocytes and myeloid dendritic cells.[208-209] TLR8 signaling pathways are mediated by MyD88; this adaptor protein activates NF-κB, IRF-7, and p38 MAPK, resulting in the induction of pro-inflammatory cytokines such as TNF-α, IL-6, IL-1β, IL-12, and antiviral type I interferons. Therefore, TLR8 is a promising target in the development of vaccine adjuvants and anticancer agents.[210] The 3D structure is well-known and six X-Ray crystallographic structures of human TLR8 in complex with six agonists are available (PDB-ID: 3W3J, 3W3K, 3W3N, 3WN4, 4Q8Z, and 4QC0). [211-213]TLR8 consists of an extracellular domain with a horseshoe-shape containing 26 LRR modules, being the ssRNA binding site very large and flexible. Ligand binding induces reorganization of the pre-organized TLR8 dimer finally enabling downstream signaling processes.[211]

To overcome the difficulty of targeting a flexible binding site, Pei *et al.*[195-196] have performed an enrichment assessment of multiple virtual screening methods, and developed a combined strategy to improve the performance of virtual screening for TLR8 agonists. First, they have created a knowledge-based pharmacophore (KBP) by merging structure-based pharmacophore and previous SAR analysis including furo[2,3-*c*]pyridines, furo[2,3-*c*]quinoles, thiazolo[4,5-*c*]quinolones, 3-*R*-quinolone-2-amine, and C7-methoxycarbonyl-imidazoquines. The combination of the KBP screening with ROCS search was used to improve the efficiency of the virtual screening process. The authors prepared a benchmarking data set merging 13 known active compounds,[214-217] 15 known inactive compounds, and decoys from ZINC database.[218] So finally they had 13 actives and 1302 decoys. The benchmarking data set was generated from their recently developed MUBD-Decoymaker protocol.[219]

The six TLR8 crystal structures were used to generate SB pharmacophores and shape-based 3D similarity search queries by means of LigandScout software. Eight pharmacophore models were derived with similar backbones in agreement with reported SAR for TLR8 agonists: three hydrophobic centroids, two aromatic rings, one H-bond donor, and one H-bond acceptor. The eight KBPs were used to screen the

benchmarking data set in order to select the most robust KBP. The authors selected the so-called "Phar1" as the priority KBP and reserved it for the subsequent antagonist verification.

In order to perform an antagonist verification data set to test the agonist/antagonist selectivity of their selected KBP, twenty reported antagonists were used.[206] For the shape-based 3D similarity search, the authors performed ROCS queries through the alignment of the six ligands from the six crystal structures, using TanimotoShape and TanimotoCombo scores. Among the resulting queries, the so-called "Query4" was taken into further analysis because of its excellent performance. As an additional step in the protocol, a comparative study was performed with four docking programs: AutoDock VINA, GOLD, Surflex-Dock and Glide. Cross-docking runs were performed with 20 cognate ligands and five dimer TLR8 complexes. Average and median RMSD values were statistically analysed to determine which program and which crystal structure best matched VS. Taken together, GOLD was identified as the most suitable docking program in conjunction with PDB-ID: 3W3J for the VS evaluation of the protocol.

Finally, the selected pharmacophore "Phar1" was combined with the ROCS "Query4" in different ways to get to the best performance as VS strategy for TLR8 agonists. Final docking with GOLD and PDB-ID: 3W3J, led to the screening of seven compounds, being three of them known active ligands as TLR8 agonists. The authors conclude that this "Phar1_Q4_Gold" strategy was proved to be a promising practice for the identification of novel TLR8 agonists. Indeed, this computational effort can be of help for the design of efficient VS strategies in other TLRs.

1.3 Objectives

The main objective of this Thesis is to deepen into the elucidation of the molecular recognition processes involving TLR4/MD-2 at atomic detail by means of computational techniques. Computational methodologies, such as MD simulations, protein-protein docking and membrane simulations will be applied to the study of the molecular mechanisms involved in the TLRs functionality, and in the recognition of PAMPs, such as natural lipopolysaccharides (LPS), synthetic glycolipids and non-LPS modulators. This knowledge will be used to help in the design and identification of new modulators.

These studies will be carried out by addressing the following specific objectives:

- CHAPTER 3: Reported modulators of the TLR4/MD-2 system with agonist and antagonist activity will be study, more concrete, the theoretical binding modes will be predicted. In particular, we will focus our work in synthetic glycolipids lipid A analogues and non LPS-like molecules. For all these reported TLR4/MD-2 modulators, there is not binding mode proposed. It is clear that, although these molecules have a different chemical structure, they must share a common pattern of interactions when bound to TLR4/MD-2. We will undertake a computational study of some representative compounds to unveil some of these patterns of interactions.

- CHAPTER 4: The cationic glycolipid IAXO-102, a potent TLR4 antagonist targeting both MD-2 and CD14 co-receptors, will be use as scaffold to design new potential TLR4 modulators and fluorescent labels for the TLR4 receptor complex (membrane TLR4/MD-2 dimer and CD14). Our modelling studies will led us to the proposal of 3D models for the interaction with CD14 and TLR4/MD-2 accounting for their binding properties and also for their antagonistic activity.

- CHAPTER 5: Virtual screening strategies from commercial and in-house libraries, followed by biological assays, will be used to the identification of new chemical entities with activity in the TLR4 complex, useful for the development of novel TLR4 modulators with a non LPS-related structure.

- CHAPTER 6: The computational building of the full structure of the TLR4/MD-2 heterodimer will be addressed, simulating the TLR4/MD-2 complex in the membrane environment. We are very much interested in the study of the dynamics of the full TLR4 at atomic level. The analysis of the molecular dynamics simulations will led us to understand the key interactions implicated in the molecular recognition events and in the dimerization process of this complex.

Bibliography:

1.   Monie, T. P.; Bryant, C. E.; Gay, N. J., Activating immunity: lessons from the TLRs and NLRs. *Trends Biochem. Sci.* **2009,** *34* (11), 553-561.

2.   Rakoff-Nahoum, S.; Medzhitov, R., Toll-like receptors and cancer. *Nat. Rev. Cancer* **2009,** *9*, 57-63.

3.   www.nobelprize.org/nobel_prizes/medicine/laureates/2011/   The   Official Website of the Nobel Prize.

4.   Hennessy, E. J.; Parker, A. E.; O'Neill, A. J., Targeting of Toll-like receptors: Emerging therapeutics? *Nat. Rev. Drug Discov.* **2010,** *9*, 293–307.

5.   Akira, S.; Takeda, K., Toll-like receptor signalling. *Nat. Rev. Immunol.* **2004,** *4*, 499-511.

6.   Janeway, C. A., Approaching the asymptote? Evolution and revolution in immunology. *Cold Spring Harb. Symp. Quant. Biol.* **1989,** *54*, 1–13.

7.   Lemaitre, B.; Nicolas, E.; Michaut, L.; Reichhart, J. M.; Hoffmann, J. A., The dorsoventral regulatory gene cassette spätzle/Toll/cactus controls the potent antifungal response in Drosophila adults. *Cell* **1996,** *86*.

8.   Medzhitov, R.; Preston-Hurlburt, P.; Janeway, C. A., A human homologue of the Drosophila Toll protein signals activation of adaptive immunity. *Nature* **1997,** *388*, 394–397.

9.   Schneider, G., Virtual screening: an endless staircase? *Nat. Rev. Drug Discov.* **2010,** *9* (4), 273-6.

10.   Gill, R.; Tsung, A.; Billiar, T., Linking oxidative stress to inflammation: Toll-like receptors. *Free Radic. Biol. Med.* **2010,** *48*, 1121–1132.

11.   Yang, H.; Hreggvidsdottir, H. S.; Palmblad, K.; Wang, H.; Ochani, M.; Li, J.; Lu, B.; Chavan, S.; Rosas-Ballina, M.; Al-Abed, Y.; Akira, S.; Bierhaus, A.; Erlandsson-Harris, H.; Andersson, U.; Tracey, K. J., A critical cysteine is required for HMGB1 binding to Toll-like receptor 4 and activation of macrophage cytokine release. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, 11942–11947.

12.   Stewart, C. R.; Stuart, L. M.; Wilkinson, K.; M., v. G. J.; Deng, J.; Halle, A.; Rayner, K. J.; Boyer, L.; Zhong, R.; Frazier, W. A.; Lacy-Hulbert, A.; El Khoury, J.; Golenbock, D. T.; Moore, K. J., CD36 ligands promote sterile inflammation through assembly of a Toll-like receptor 4 and 6 heterodimer. *Nat. Immunol.* **2010,** *11*, 155–161.

13.   O'Neill, L. A.; Bryant, C. E.; Doyle, S. L., Therapeutic targeting of Toll-like receptors for infectious and inflammatory diseases and cancer. *Rev. Pharmacol.* **2009,** *61*, 177–197.

14.   Kanzler, H.; Barrat, F. J.; Hessel, E. M.; Coffman, R. L., Therapeutic targeting of innate immunity with Toll-like receptor agonists and antagonists. *Nat. Med.* **2007,** *13*, 552–559.

15.   Makkouk, A.; Abdelnoor, A. M., The potential use of toll-like receptor (TLR) agonists   and   antagonists   as   prophylactic   and/or   therapeutic   agents. *Immunopharmacol. Immunotoxicol.* **2009,** *31* (3), 331-338.

16.   Casella, C. R.; Mitchell, T. C., Putting endotoxin to work for us: monophosphoryl lipid A as a safe and effective vaccine adjuvant. *Cell. Mol. Life Sci.* **2008,** *65*, 3231–3240.

17.   West, A. P.; Koblansky, A. A.; Ghosh, S., Recognition and signaling by toll-like receptors. *Annu. Rev. Cell Dev. Biol.* **2006,** *22*, 409–437.

18.	Kaisho, T.; Akira, S., Toll-like receptor function and signaling. *J. Allergy Clin.* **2006,** *117*, 979-987.

19.	Iwasaki, A.; Medzhitov, R., Toll-like receptor control of the adaptive immune responses. *Nat. Immunol.* **2004,** *5*, 987-995.

20.	Gay, N. J.; Gangloff, M., Structure and function of Toll receptors and their ligands. *Annu. Rev. Biochem.* **2007,** *76*, 141-165.

21.	Matsushima, N.; Tanaka, T.; Enkhbayar, P.; Mikami, T.; Taga, M.; Yamada, K.; Kuroki, Y., Comparative sequence analysis of leucine-rich repeats (LRRs) within vertebrate toll-like receptors. *BMC Genom.* **2007,** *8*, 124.

22.	Kajava, A. V., Structural diversity of leucine-rich repeat proteins. *J. Mol. Biol.* **1998,** *277* (3), 519-527.

23.	Kobe, B.; Kajava, A. V., The leucine-rich repeat as a protein recognition motif. *Curr. Opin. Struct. Biol.* **2001,** *11*, 725–732.

24.	O'Neill, L. A.; Bowie, A. G., The family of five: TIR-domain-containing adaptors in Toll-like receptor signalling. *Nat. Rev. Immunol.* **2007,** *7*, 353–364.

25.	Iwasaki, A.; Medzhitov, R., Regulation of adaptive immunity by the innate immune system. *Science* **2010,** *327*, 291–295.

26.	Botos, I.; Segal, D. M.; Davies, D. R., The Structural Biology of Toll-like Receptors. *Structure* **2011,** *19* (4), 447-459.

27.	Poltorak, A.; He, X.; Smirnova, I.; Liu, M. Y.; Van Huffel, C.; Du, X.; Birdwell, D.; Alejos, E.; Silva, M.; Galanos, C.; Freudenberg, M.; Ricciardi-Castagnoli, P.; Layton, B.; Beutler, B., Defective LPS signaling in C3H/HeJ and C57BL/10ScCr mice: Mutations in Tlr4 gene. *Science* **1998,** *282*, 2085-2088.

28.	Hold, G. L.; Bryant, C. E., The Molecular Basis of Lipid A and Toll-Like Receptor 4 Interactions. In *Bacterial Lipopolysaccharides: Structure, Chemical Synthesis, Biogenesis and Interaction with Host Cells*, Knirel, Y. A.; Valvano, M. A., Eds. Springer Vienna: Vienna, 2011; pp 371-387.

29.	Peri, F.; Piazza, M.; Calabrese, V.; Cighetti, R., Modulation of Lipopolysaccharide Signalling Through TLR4 Agonists and Antagonists. In *Bacterial Lipopolysaccharides: Structure, Chemical Synthesis, Biogenesis and Interaction with Host Cells*, Knirel, Y. A.; Valvano, M. A., Eds. Springer Vienna: Vienna, 2011; pp 389-416.

30.	Mancek-Keber, M.; Gradisar, H.; Iñigo Pestaña, M.; Martinez de Tejada, G.; J., J. R., Free thiol group of MD-2 as the target for inhibition of the lipopolysaccharide-induced cell activation. *J. Biol. Chem.* **2009,** *284*, 19493-500.

31.	Gioannini, T.-L.; Teghanemt, A.; Zhang, D.; Coussens, N. P.; Dockstader, W.; Ramaswamy, S.; Weiss, J. P., Isolation of an endotoxin–MD-2 complex that produces Toll-like receptor 4-dependent cell activation at picomolar concentrations. *Proc. Natl. Acad. Sci. U.S.A.* **2004,** *101* (12), 4186-4191.

32.	Li, M.; Zhou, Y.; Feng, G.; Su, S. B., The critical role of Toll-like receptor signaling pathways in the induction and progression of autoimmune diseases. *Curr. Mol. Med.* **2009,** *9* (3), 365-74.

33.	Jialal, I.; Kaur, H.; Devaraj, S., Toll-like receptor status in obesity and metabolic syndrome: a translational perspective. *J. Clin. Endocrinol. Metab.* **2014,** *99* (1), 39-48.

34.	Shi, M.; Chen, X.; Ye, K.; Yao, Y.; Li, Y., Application potential of toll-like receptors in cancer immunotherapy: Systematic review. *Medicine* **2016,** *95* (25), e3951.

35.   Baxevanis, C. N.; Voutsas, I. F.; Tsitsilonis, O. E., Toll-like receptor agonists: current status and future perspective on their utility as adjuvants in improving antic ancer vaccination strategies. *Immunotherapy* **2013,** *5* (5), 497-511.

36.   Garçon, N.; Chomez, P.; Van Mechelen, M., GlaxoSmithKline Adjuvant Systems in vaccines: concepts, achievements and perspectives. *Expert Rev. Vaccines* **2007,** *6* (5), 723-39.

37.   Casella, C. R.; Mitchell, T. C., Inefficient TLR4/MD-2 heterotetramerization by monophosphoryl lipid A. *PloS one* **2013,** *8* (4).

38.   Thoelen, S.; De Clercq, N.; Tornieporth, N., A prophylactic hepatitis B vaccine with a novel adjuvant system. *Vaccine* **2001,** *19* (17-19), 2400-3.

39.   Paavonen, J.; Jenkins, D.; Bosch, F. X.; Naud, P.; Salmerón, J.; Wheeler, C. M.; Chow, S. N.; Apter, D. L.; Kitchener, H. C.; Castellsague, X.; de Carvalho, N. S.; Skinner, S. R.; Harper, D. M.; Hedrick, J. A.; Jaisamrarn, U.; Limson, G. A.; Dionne, M.; Quint, W.; Spiessens, B.; Peeters, P.; Struyf, F.; Wieting, S. L.; Lehtinen, M. O.; Dubin, G.; group, H. P. s., Efficacy of a prophylactic adjuvanted bivalent L1 virus-like-particle vaccine against infection with human papillomavirus types 16 and 18 in young women: an interim analysis of a phase III double-blind, randomised controlled trial. *Lancet* **2007,** *369* (9580), 2161-70.

40.   Harper, D. M.; Franco, E. L.; Wheeler, C.; Ferris, D. G.; Jenkins, D.; Schuind, A.; Zahaf, T.; Innis, B.; Naud, P.; De Carvalho, N. S.; Roteli-Martins, C. M.; Teixeira, J.; Blatter, M. M.; Korn, A. P.; Quint, W.; Dubin, G.; Group, G. H. V. S., Efficacy of a bivalent L1 virus-like particle vaccine in prevention of infection with human papillomavirus types 16 and 18 in young women: a randomised controlled trial. *Lancet* **2004,** *364* (9447), 1757-65.

41.   Mitchell, M. S.; Kan-Mitchell, J.; Kempf, R. A.; Harel, W.; Shau, H. Y.; Lind, S., Active specific immunotherapy for melanoma: phase I trial of allogeneic lysates and a novel adjuvant. *Cancer Res.* **1988,** *48* (20), 5883-93.

42.   Przetak, M.; Chow, J.; Cheng, H.; Rose, J.; Hawkins, L. D.; Ishizaka, S. T., Novel synthetic LPS receptor agonists boost systemic and mucosal antibody responses in mice. *Vaccine* **2003,** *21* (9-10), 961-70.

43.   Wang, S.; Astsaturov, I. A.; Bingham, C. A.; McCarthy, K. M.; von Mehren, M.; Xu, W.; Alpaugh, R. K.; Tang, Y.; Littlefield, B. A.; Hawkins, L. D., Effective antibody therapy induces host-protective antitumor immunity that is augmented by TLR4 agonist treatment. *Cancer Immunol. Immunother.* **2012,** *61* (1), 49-61.

44.   Morefield, G. L.; Hawkins, L. D.; Ishizaka, S. T.; Kissner, T. L.; Ulrich, R. G., Synthetic Toll-like receptor 4 agonist enhances vaccine efficacy in an experimental model of toxic shock syndrome. *Clin. Vaccine Immunol.* **2007,** *14* (11), 1499-1504.

45.   Stöver, A. G.; Correia, J. D. S.; Evans, J. T.; Cluff, C. W.; Elliott, M. W.; Jeffery, E. W.; Johnson, D. A.; Lacy, M. J.; Baldridge, J. R.; Probst, P., Structure-activity relationship of synthetic toll-like receptor 4 agonists. *J. Biol. Chem.* **2004,** *279* (6), 4440-4449.

46.   Johnson, D. A., Synthetic TLR4-active glycolipids as vaccine adjuvants and stand-alone immunotherapeutics. *Curr. Top. Med. Chem.* **2008,** *8* (2), 64-79.

47.   Johnson, D. A.; Gregory Sowell, C.; Johnson, C. L.; Livesay, M. T.; Keegan, D. S.; Rhodes, M. J.; Terry Ulrich, J.; Ward, J. R.; Cantrell, J. L.; Brookshire, V. G., Synthesis and biological evaluation of a new class of vaccine adjuvants: aminoalkyl glucosaminide 4-phosphates (AGPs). *Bioorg. Med. Chem. Lett.* **1999,** *9* (15), 2273-2278.

48. Bowen, W. S.; Minns, L. A.; Johnson, D. A.; Mitchell, T. C.; Hutton, M. M.; Evans, J. T., Selective TRIF-dependent signaling by a synthetic toll-like receptor 4 agonist. *Sci. Signal.* **2012,** *5* (211), ra13.

49. Chan, M.; Hayashi, T.; Mathewson, R. D.; Nour, A.; Hayashi, Y.; Yao, S.; Tawatao, R. I.; Crain, B.; Tsigelny, I. F.; Kouznetsova, V. L., Identification of substituted pyrimido [5, 4-b] indoles as selective Toll-like receptor 4 ligands. *J. Med. Chem.* **2013,** *56* (11), 4206-4223.

50. Neve, J. E.; Wijesekera, H. P.; Duffy, S.; Jenkins, I. D.; Ripper, J. A.; Teague, S. J.; Campitelli, M.; Garavelas, A.; Nikolakopoulos, G.; Le, P. V., Euodenine A: A Small-Molecule Agonist of Human TLR4. *J. Med. Chem.* **2014,** *57* (4), 1252-1275.

51. Shanmugam, A.; Rajoria, S.; George, A. L.; Mittelman, A.; Suriano, R.; Tiwari, R. K., Synthetic toll like receptor-4 (TLR-4) agonist peptides as a novel class of adjuvants. *PloS one* **2012,** *7* (2), e30839.

52. Wittebole, X.; Castanares-Zapatero, D.; Laterre, P.-F., Toll-like receptor 4 modulation as a strategy to treat sepsis. *Mediators Inflamm.* **2010,** *2010*.

53. Shirey, K. A.; Lai, W.; Scott, A. J.; Lipsky, M.; Mistry, P.; Pletneva, L. M.; Karp, C. L.; McAlees, J.; Gioannini, T. L.; Weiss, J., The TLR4 antagonist Eritoran protects mice from lethal influenza infection. *Nature* **2013,** *497* (7450), 498-502.

54. Peri, F.; Marinzi, C.; Barath, M.; Granucci, F.; Urbano, M.; Nicotra, F., Synthesis and biological evaluation of novel lipid A antagonists. *Bioorg. Med. Chem.* **2006,** *14* (1), 190-9.

55. Piazza, M.; Rossini, C.; Della Fiorentina, S.; Pozzi, C.; Comelli, F.; Bettoni, I.; Fusi, P.; Costa, B.; Peri, F., Glycolipids and benzylammonium lipids as novel antisepsis agents: synthesis and biological characterization. *J. Med. Chem.* **2009,** *52* (4), 1209-13.

56. Peri, F.; Granucci, F.; Costa, B.; Zanoni, I.; Marinzi, C.; Nicotra, F., Inhibition of lipid a stimulated activation of human dendritic cells and macrophages by amino and hydroxylamino monosaccharides. *Angew. Chem. Int. Ed. Engl.* **2007,** *46* (18), 3308-12.

57. Piazza, M.; Calabrese, V.; Damore, G.; Cighetti, R.; Gioannini, T.; Weiss, J.; Peri, F., A Synthetic Lipid A Mimetic Modulates Human TLR4 Activity. *ChemMedChem* **2012,** *7* (2), 213-7.

58. Cighetti, R.; Ciaramelli, C.; Sestito, S. E.; Zanoni, I.; Kubik, Ł.; Ardá-Freire, A.; Calabrese, V.; Granucci, F.; Jerala, R.; Martín-Santamaría, S., Modulation of CD14 and TLR4· MD-2 Activities by a Synthetic Lipid A Mimetic. *Chembiochem : a European journal of chemical biology* **2014,** *15* (2), 250-258.

59. Park, S.-J.; Kang, S. H.; Kang, Y. K.; Eom, Y.-B.; Koh, K. O.; Kim, D. Y.; Youn, H.-S., Inhibition of homodimerization of toll-like receptor 4 by 4-oxo-4-(2-oxo-oxazolidin-3-yl)-but-2-enoic acid ethyl ester. *Int. Immunopharmacol.* **2011,** *11* (1), 19-22.

60. Jin, G. H.; Li, H.; An, S.; Ryu, J.-H.; Jeon, R., Design, synthesis and activity of benzothiazole-based inhibitors of NO production in LPS-activated macrophages. *Bioorg. Med. Chem. Lett.* **2010,** *20* (21), 6199-6202.

61. Kawamoto, T.; Ii, M.; Kitazaki, T.; Iizawa, Y.; Kimura, H., TAK-242 selectively suppresses Toll-like receptor 4-signaling mediated by the intracellular domain. *Eur. J. Pharmacol.* **2008,** *584* (1), 40-48.

62. Chavez, S. A.; Martinko, A. J.; Lau, C.; Pham, M. N.; Cheng, K.; Bevan, D. E.; Mollnes, T. E.; Yin, H., Development of β-Amino Alcohol Derivatives That Inhibit Toll-

like Receptor 4 Mediated Inflammatory Response as Potential Antiseptics. *J. Med. Chem.* **2011,** *54* (13), 4659–4669.

63.   Barata, T. S.; Teo, I.; Brocchini, S.; Zloh, M.; Shaunak, S., Partially glycosylated dendrimers block MD-2 and prevent TLR4-MD-2-LPS complex mediated cytokine responses. *PLoS Comput. Biol.* **2011,** *7* (6), e1002095.

64.   Barata, T.; Teo, I.; Lalwani, S.; Simanek, E.; Zloh, M.; Shaunak, S., Computational design principles for bioactive dendrimer based constructs as antagonists of the TLR4-MD-2-LPS complex. *Biomaterials* **2011,** *32* (33), 8702-8711.

65.   Jin, M. S.; Kim, S. E.; Heo, J. Y.; Lee, M. E.; Kim, H. M.; Paik, S. G.; Lee, H.; Lee, J. O., Crystal structure of the TLR1-TLR2 heterodimer induced by binding of a tri-acylated lipopeptide. *Cell* **2007,** *130*, 1071–1082.

66.   Liu, L.; Botos, I.; Wang, Y.; Leonard, J. N.; Shiloach, J.; Segal, D. M.; Davies, D. R., Structural basis of toll-like receptor 3 signaling with double-stranded RNA. *Science* **2008,** *320*, 379–381.

67.   Park, B. S.; Song, D. H.; Kim, H. M.; Choi, B. S.; Lee, H.; Lee, J. O., The structural basis of lipopolysaccharide recognition by the TLR4-MD-2 complex. *Nature* **2009,** *458*, 1191–1195.

68.   Park, B. S.; Lee, J.-O., Recognition of lipopolysaccharide pattern by TLR4 complexes. *Exp. Mol. Med.* **2013,** *45* (12), e66.

69.   Wang, Y.; Su, L.; Morin, M. D.; Jones, B. T.; Whitby, L. R.; Surakattula, M. M.; Huang, H.; Shi, H.; Choi, J. H.; Wang, K. W.; Moresco, E. M.; Berger, M.; Zhan, X.; Zhang, H.; Boger, D. L.; Beutler, B., TLR4/MD-2 activation by a synthetic agonist with no similarity to LPS. *Proc. Natl. Acad. Sci. U.S.A.* **2016,** *113* (7), E884-93.

70.   Ohto, U.; Fukase, K.; Miyake, K.; Shimizu, T., Structural basis of species-specific endotoxin sensing by innate immune receptor TLR4/MD-2. *Proc. Natl. Acad. Sci. U.S.A.* **2012,** *109* (19), 7421-6.

71.   Ohto, U.; Yamakawa, N.; Akashi-Takamura, S.; Miyake, K.; Shimizu, T., Structural analyses of human Toll-like receptor 4 polymorphisms D299G and T399I. *J. Biol. Chem.* **2012,** *287* (48), 40611-7.

72.   Kim, H. M.; Park, B. S.; Kim, J. I.; Kim, S. E.; Lee, J.; Oh, S. C.; Enkhbayar, P.; Matsushima, N.; Lee, H.; Yoo, O. J.; Lee, J. O., Crystal structure of the TLR4-MD-2 complex with bound endotoxin antagonist Eritoran. *Cell* **2007,** *130* (5), 906-17.

73.   Han, J.; Kim, H. J.; Lee, S. C.; Hong, S.; Park, K.; Jeon, Y. H.; Kim, D.; Cheong, H. K.; Kim, H. S., Structure-based rational design of a Toll-like receptor 4 (TLR4) decoy receptor with high binding affinity for a target protein. *PLoS one* **2012,** *7* (2), e30929.

74.   Ohto, U.; Fukase, K.; Miyake, K.; Satow, Y., Crystal structures of human MD-2 and its complex with antiendotoxic lipid IVa. *Science* **2007,** *316*, 1632-4.

75.   Kim, J. I.; Lee, C. J.; Jin, M. S.; Lee, C. H.; Paik, S. G.; Lee, H.; Lee, J. O., Crystal structure of CD14 and its implications for lipopolysaccharide signaling. *J. Biol. Chem.* **2005,** *280* (12), 11347-51.

76.   Kelley, S. L.; Lukk, T.; Nair, S. K.; Tapping, R. I., The crystal structure of human soluble CD14 reveals a bent solenoid with a hydrophobic amino-terminal pocket. *J. Immunol.* **2013,** *190* (3), 1304-11.

77.   Yu, L.; Phillips, R. L.; Zhang, D. S.; Teghanemt, A.; Weiss, J. P.; Gioannini, T. L., NMR studies of hexaacylated endotoxin bound to wild-type and F126A mutant MD-2 and MD-2· TLR4 ectodomain complexes. *J. Biol. Chem.* **2012,** *287* (20), 16346-16355.

78.   Lindert, S.; Kekenes-Huskey, P. M.; McCammon, J. A., Long-Timescale Molecular Dynamics Simulations Elucidate the Dynamics and Kinetics of Exposure of the Hydrophobic Patch in Troponin C. *Biophys. J.* **2012,** *103* (8), 1784-1789.

79.   Billod, J. M.; Lacetera, A.; Guzman-Caldentey, J.; Martin-Santamaria, S., Computational Approaches to Toll-Like Receptor 4 Modulation. *Molecules (Basel, Switzerland)* **2016,** *21* (8).

80.   Garate, J. A.; Oostenbrink, C., Lipid A from lipopolysaccharide recognition: structure, dynamics and cooperativity by molecular dynamics simulations. *Proteins* **2013,** *81* (4), 658-74.

81.   DeMarco, M. L.; Woods, R. J., From agonist to antagonist: structure and dynamics of innate immune glycoprotein MD-2 upon recognition of variably acylated bacterial endotoxins. *Mol. Immunol.* **2011,** *49* (1-2), 124-33.

82.   Paramo, T.; Piggot, T. J.; Bryant, C. E.; Bond, P. J., The structural basis for endotoxin-induced allosteric regulation of the Toll-like receptor 4 (TLR4) innate immune receptor. *J. Biol. Chem.* **2013,** *288* (51), 36215-25.

83.   de Aguiar, C.; Costa, M. G.; Verli, H., Dynamics on human Toll-like receptor 4 complexation to MD-2: the coreceptor stabilizing function. *Proteins* **2015,** *83* (2), 373-82.

84.   Anwar, M. A.; Panneerselvam, S.; Shah, M.; Choi, S., Insights into the species-specific TLR4 signaling mechanism in response to Rhodobacter sphaeroides lipid A detection. *Sci. Rep.* **2015,** *5*, 7657.

85.   Slivka, P. F.; Shridhar, M.; Lee, G. I.; Sammond, D. W.; Hutchinson, M. R.; Martinko, A. J.; Buchanan, M. M.; Sholar, P. W.; Kearney, J. J.; Harrison, J. A.; Watkins, L. R.; Yin, H., A peptide antagonist of the TLR4-MD2 interaction. *Chembiochem : a European journal of chemical biology* **2009,** *10* (4), 645-9.

86.   Rohl, C. A.; Strauss, C. E.; Misura, K. M.; Baker, D., Protein structure prediction using Rosetta. *Methods Enzymol.* **2004,** *383*, 66-93.

87.   Di Lorenzo, F.; Kubik, L.; Oblak, A.; Lore, N. I.; Cigana, C.; Lanzetta, R.; Parrilli, M.; Hamad, M. A.; De Soyza, A.; Silipo, A.; Jerala, R.; Bragonzi, A.; Valvano, M. A.; Martin-Santamaria, S.; Molinaro, A., Activation of Human Toll-like Receptor 4 (TLR4).Myeloid Differentiation Factor 2 (MD-2) by Hypoacylated Lipopolysaccharide from a Clinical Isolate of Burkholderia cenocepacia. *J. Biol. Chem.* **2015,** *290* (35), 21305-19.

88.   Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J., AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility. *J. Comput. Chem.* **2009,** *30* (16), 2785-2791.

89.   Trott, O.; Olson, A. J., AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. *J. Comput. Chem.* **2010,** *31* (2), 455-461.

90.   Klett, J.; Nunez-Salgado, A.; Dos Santos, H. G.; Cortes-Cabrera, A.; Perona, A.; Gil-Redondo, R.; Abia, D.; Gago, F.; Morreale, A., MM-ISMSA: An Ultrafast and Accurate Scoring Function for Protein-Protein Docking. *J. Chem. Theory Comput.* **2012,** *8* (9), 3395-408.

91.   Vasl, J.; Oblak, A.; Peternelj, T. T.; Klett, J.; Martin-Santamaria, S.; Gioannini, T. L.; Weiss, J. P.; Jerala, R., Molecular Basis of the Functional Differences between

Soluble Human Versus Murine MD-2: Role of Val135 in Transfer of Lipopolysaccharide from CD14 to MD-2. *J. Immunol.* **2016,** *196* (5), 2309-18.

92.   Narayanan, K. B.; Park, H. H., Toll/interleukin-1 receptor (TIR) domain-mediated cellular signaling pathways. *Apoptosis* **2015,** *20* (2), 196-209.

93.   Xu, Y.; Tao, X.; Shen, B.; Horng, T.; Medzhitov, R.; Manley, J. L.; Tong, L., Structural basis for signal transduction by the Toll/interleukin-1 receptor domains. *Nature* **2000,** *408* (6808), 111-5.

94.   Hasan, U.; Chaffois, C.; Gaillard, C.; Saulnier, V.; Merck, E.; Tancredi, S.; Guiet, C.; Briere, F.; Vlach, J.; Lebecque, S.; Trinchieri, G.; Bates, E. E., Human TLR10 is a functional receptor, expressed by B cells and plasmacytoid dendritic cells, which activates gene transcription through MyD88. *J. Immunol.* **2005,** *174* (5), 2942-50.

95.   Ohnishi, H.; Tochio, H.; Kato, Z.; Orii, K. E.; Li, A.; Kimura, T.; Hiroaki, H.; Kondo, N.; Shirakawa, M., Structural basis for the multiple interactions of the MyD88 TIR domain in TLR4 signaling. *Proc. Natl. Acad. Sci. U.S.A.* **2009,** *106* (25), 10260-5.

96.   Loiarro, M.; Sette, C.; Gallo, G.; Ciacci, A.; Fanto, N.; Mastroianni, D.; Carminati, P.; Ruggiero, V., Peptide-mediated interference of TIR domain dimerization in MyD88 inhibits interleukin-1-dependent activation of NF-{kappa}B. *J. Biol. Chem.* **2005,** *280* (16), 15809-14.

97.   Jiang, Z.; Georgel, P.; Li, C.; Choe, J.; Crozat, K.; Rutschmann, S.; Du, X.; Bigby, T.; Mudd, S.; Sovath, S.; Wilson, I. A.; Olson, A.; Beutler, B., Details of Toll-like receptor:adapter interaction revealed by germ-line mutagenesis. *Proc. Natl. Acad. Sci. U.S.A.* **2006,** *103* (29), 10961-6.

98.   Lin, Z.; Lu, J.; Zhou, W.; Shen, Y., Structural Insights into TIR Domain Specificity of the Bridging Adaptor Mal in TLR4 Signaling. *PloS one* **2012,** *7* (4), e34202.

99.   Dunne, A.; Ejdeback, M.; Ludidi, P. L.; O'Neill, L. A.; Gay, N. J., Structural complementarity of Toll/interleukin-1 receptor domains in Toll-like receptors and the adaptors Mal and MyD88. *J. Biol. Chem.* **2003,** *278* (42), 41443-51.

100.   Kubarenko, A.; Frank, M.; Weber, A. N., Structure-function relationships of Toll-like receptor domains through homology modelling and molecular dynamics. *Biochem. Soc. Trans.* **2007,** *35* (Pt 6), 1515-8.

101.   Gong, J.; Wei, T.; Stark, R. W.; Jamitzky, F.; Heckl, W. M.; Anders, H. J.; Lech, M.; Rossle, S. C., Inhibition of Toll-like receptors TLR4 and 7 signaling pathways by SIGIRR: a computational approach. *J. Struct. Biol.* **2010,** *169* (3), 323-30.

102.   Nunez Miguel, R.; Wong, J.; Westoll, J. F.; Brooks, H. J.; O'Neill, L. A.; Gay, N. J.; Bryant, C. E.; Monie, T. P., A dimer of the Toll-like receptor 4 cytoplasmic domain provides a specific scaffold for the recruitment of signalling adaptor proteins. *PloS one* **2007,** *2* (8), e788.

103.   Basith, S.; Manavalan, B.; Govindaraj, R. G.; Choi, S., In silico approach to inhibition of signaling pathways of Toll-like receptors 2 and 4 by ST2L. *PloS one* **2011,** *6* (8), e23989.

104.   Guven-Maiorov, E.; Keskin, O.; Gursoy, A.; VanWaes, C.; Chen, Z.; Tsai, C. J.; Nussinov, R., The Architecture of the TIR Domain Signalosome in the Toll-like Receptor-4 Signaling Pathway. *Sci. Rep.* **2015,** *5*, 13128.

105.   Bovijn, C.; Ulrichts, P.; De Smet, A. S.; Catteeuw, D.; Beyaert, R.; Tavernier, J.; Peelman, F., Identification of interaction sites for dimerization and adapter

recruitment in Toll/interleukin-1 receptor (TIR) domain of Toll-like receptor 4. *J. Biol. Chem.* **2012,** *287* (6), 4088-98.

106.   Singh, S.; Pandey, K.; Rathore, Y. S.; Sagar, A.; Pattnaik, U. B.; Ashish, A communication network within the cytoplasmic domain of toll-like receptors has remained conserved during evolution. *J. Biomol. Struct. Dyn.* **2014,** *32* (5), 694-700.

107.   Barochia, A.; Solomon, S.; Cui, X.; Natanson, C.; Eichacker, P. Q., Eritoran tetrasodium (E5564) treatment for sepsis: review of preclinical and clinical studies. *Expert Opin. Drug Metab. Toxicol.* **2011,** *7* (4), 479-94.

108.   Scior, T.; Lozano-Aponte, J.; Figueroa-Vazquez, V.; Yunes-Rojas, J. A.; Zahringer, U.; Alexander, C., Three-dimensional mapping of differential amino acids of human, murine, canine and equine TLR4/MD-2 receptor complexes conferring endotoxic activation by lipid A, antagonism by Eritoran and species-dependent activities of Lipid IVA in the mammalian LPS sensor system. *Comput. Struct. Biotechnol. J.* **2013,** *7*, e201305003.

109.   Krivov, G. G.; Shapovalov, M. V.; Dunbrack, R. L., Jr., Improved prediction of protein side-chain conformations with SCWRL4. *Proteins* **2009,** *77* (4), 778-95.

110.   http://www.schrodinger.com, G. I. s. A. o.

111.   Garate, J. A.; Stockl, J.; Fernandez-Alonso Mdel, C.; Artner, D.; Haegman, M.; Oostenbrink, C.; Jimenez-Barbero, J.; Beyaert, R.; Heine, H.; Kosma, P.; Zamyatina, A., Anti-endotoxic activity and structural basis for human MD-2.TLR4 antagonism of tetraacylated lipid A mimetics based on betaGlcN(1<-->1)alphaGlcN scaffold. *Innate Immun.* **2015,** *21* (5), 490-503.

112.   Piazza, M.; Calabrese, V.; Baruffa, C.; Gioannini, T.; Weiss, J.; Peri, F., The cationic amphiphile 3,4-bis(tetradecyloxy)benzylamine inhibits LPS signaling by competing with endotoxin for CD14 binding. *Biochem. Pharmacol.* **2010,** *80* (12), 2050-6.

113.   Webb, B.; Sali, A., Comparative Protein Structure Modeling Using MODELLER. *Curr. Protoc. Protein Sci.* **2016,** *86*, 2.9.1-2.9.37.

114.   Walsh, C.; Gangloff, M.; Monie, T.; Smyth, T.; Wei, B.; McKinley, T. J.; Maskell, D.; Gay, N.; Bryant, C., Elucidation of the MD-2/TLR4 interface required for signaling by lipid IVa. *J. Immunol.* **2008,** *181* (2), 1245-54.

115.   Irvine, K. L.; Gangloff, M.; Walsh, C. M.; Spring, D. R.; Gay, N. J.; Bryant, C. E., Identification of key residues that confer Rhodobacter sphaeroides LPS activity at horse TLR4/MD-2. *PloS one* **2014,** *9* (5), e98776.

116.   Maeshima, N.; Evans-Atkinson, T.; Hajjar, A. M.; Fernandez, R. C., Bordetella pertussis Lipid A Recognition by Toll-like Receptor 4 and MD-2 Is Dependent on Distinct Charged and Uncharged Interfaces. *J. Biol. Chem.* **2015,** *290* (21), 13440-53.

117.   Kawasaki, K.; Akashi, S.; Shimazu, R.; Yoshida, T.; Miyake, K.; Nishijima, M., Mouse toll-like receptor 4. MD-2 complex mediates lipopolysaccharide-mimetic signal transduction by Taxol. *J. Biol. Chem.* **2000,** *275*, 2251-2254.

118.   Manthey, C. L.; Qureshi, N.; Stutz, P. L.; Vogel, S. N., Lipopolysaccharide antagonists block taxol-induced signaling in murine macrophages. *J. Exp. Med.* **1993,** *178* (2), 695-702.

119.   Resmana, N.; Gradisar, H.; Vasl, H.; Keber, M. M.; Pristovsek, P.; Jerala, R., Taxanes inhibit human TLR4 signaling by binding to MD-2. *FEBS Lett.* **2008,** *582*, 3929–3934.

120.   Zimmer, S. M.; Liu, J.; Clayton, J. L.; Stephens, D. S.; Snyder, J. P., Paclitaxel binding to human and murine MD-2. *J. Biol. Chem.* **2008,** *283* (41), 27916-26.

121.   Fu, W.; Chen, L.; Wang, Z.; Zhao, C.; Chen, G.; Liu, X.; Dai, Y.; Cai, Y.; Li, C.; Zhou, J.; Liang, G., Determination of the binding mode for anti-inflammatory natural product xanthohumol with myeloid differentiation protein 2. *Drug Des. Devel. Ther.* **2016,** *10*, 455-63.

122.   Wang, Z.; Chen, G.; Chen, L.; Liu, X.; Fu, W.; Zhang, Y.; Li, C.; Liang, G.; Cai, Y., Insights into the binding mode of curcumin to MD-2: studies from molecular docking, molecular dynamics simulations and experimental assessments. *Mol. Biosyst.* **2015,** *11* (7), 1933-8.

123.   Nour, A.; Hayashi, T.; Chan, M.; Yao, S.; Tawatao, R. I.; Crain, B.; Tsigelny, I. F.; Kouznetsova, V. L.; Ahmadiiveli, A.; Messer, K.; Pu, M.; Corr, M.; Carson, D. A.; Cottam, H. B., Discovery of substituted 4-aminoquinazolines as selective Toll-like receptor 4 ligands. *Bioorg. Med. Chem. Lett.* **2014,** *24* (21), 4931-8.

124.   Koo, J. E.; Park, Z. Y.; Kim, N. D.; Lee, J. Y., Sulforaphane inhibits the engagement of LPS with TLR4/MD2 complex by preferential binding to Cys133 in MD2. *Biochem. Biophys. Res. Commun.* **2013,** *434* (3), 600-5.

125.   Kim, S. Y.; Koo, J. E.; Seo, Y. J.; Tyagi, N.; Jeong, E.; Choi, J.; Lim, K. M.; Park, Z. Y.; Lee, J. Y., Suppression of Toll-like receptor 4 activation by caffeic acid phenethyl ester is mediated by interference of LPS binding to MD2. *Br. J. Pharmacol.* **2013,** *168* (8), 1933-45.

126.   Bevan, D. E.; Martinko, A. J.; Loram, L. C.; Stahl, J. A.; Taylor, F. R.; Joshee, S.; Watkins, L. R.; Yin, H., Selection, Preparation, and Evaluation of Small- Molecule Inhibitors of Toll-Like Receptor 4. *ACS Med. Chem. Lett.* **2010,** *1* (5), 194-198.

127.   Byun, E.-B.; Sung, N.-Y.; Byun, E.-H.; Song, D.-S.; Kim, J.-K.; Park, J.-H.; Song, B.-S.; Park, S.-H.; Lee, J.-W.; Byun, M.-W.; Kim, J.-H., The procyanidin trimer C1 inhibits LPS-induced MAPK and NF-κB signaling through TLR4 in macrophages. *Int. Immunopharmacol.* **2013,** *15* (2), 450-456.

128.   Terra, X.; Palozza, P.; Fernandez-Larrea, J.; Ardevol, A.; Blade, C.; Pujadas, G.; Salvado, J.; Arola, L.; Blay, M. T., Procyanidin dimer B1 and trimer C1 impair inflammatory response signalling in human monocytes. *Free Radic. Res.* **2011,** *45* (5), 611-9.

129.   Jung, M.; Triebel, S.; Anke, T.; Richling, E.; Erkel, G., Influence of apple polyphenols on inflammatory gene expression. *Mol. Nutr. Food Res.* **2009,** *53* (10), 1263-80.

130.   Xing, J.; Li, R.; Li, N.; Zhang, J.; Li, Y.; Gong, P.; Gao, D.; Liu, H.; Zhang, Y., Anti-inflammatory effect of procyanidin B1 on LPS-treated THP1 cells via interaction with the TLR4-MD-2 heterodimer and p38 MAPK and NF-kappaB signaling. *Mol. Cell Biochem.* **2015,** *407* (1-2), 89-95.

131.   Dumortier, H., When carbon nanotubes encounter the immune system: desirable and undesirable effects. *Adv. Drug Deliv. Rev.* **2013,** *65* (15), 2120-6.

132.   Dong, J.; Ma, Q., Advances in mechanisms and signaling pathways of carbon nanotube toxicity. *Nanotoxicology* **2015,** *9* (5), 658-76.

133.   Turabekova, M.; Rasulev, B.; Theodore, M.; Jackman, J.; Leszczynska, D.; Leszczynski, J., Immunotoxity of nanoparticles: a computational study suggests that

CNTs and C60 fullerenes might be recognized as pathogens by Toll-like receptors. *Nanoscale* **2014,** *6* (7), 3488-95.

134.   Hiratsuka, S.; Watanabe, A.; Sakurai, Y.; Akashi-Takamura, S.; Ishibashi, S.; Miyake, K.; Shibuya, M.; Akira, S.; Aburatani, H.; Maru, Y., The S100A8-serum amyloid A3-TLR4 paracrine cascade establishes a pre-metastatic phase. *Nat. Cell Biol.* **2008,** *10* (11), 1349-55.

135.   Deguchi, A.; Tomita, T.; Ohto, U.; Takemura, K.; Kitao, A.; Akashi-Takamura, S.; Miyake, K.; Maru, Y., Eritoran inhibits S100A8-mediated TLR4/MD-2 activation and tumor growth by changing the immune microenvironment. *Oncogene* **2016,** *35* (11), 1445-56.

136.   Chen, R.; Li, L.; Weng, Z., ZDOCK: an initial-stage protein-docking algorithm. *Proteins* **2003,** *52* (1), 80-7.

137.   Gerke, V.; Creutz, C. E.; Moss, S. E., Annexins: linking Ca2+ signalling to membrane dynamics. *Nat. Rev. Mol. Cell Biol.* **2005,** *6* (6), 449-61.

138.   Zhang, S.; Yu, M.; Guo, Q.; Li, R.; Li, G.; Tan, S.; Li, X.; Wei, Y.; Wu, M., Annexin A2 binds to endosomes and negatively regulates TLR4-triggered inflammatory responses via the TRAM-TRIF pathway. *Sci. Rep.* **2015,** *5*, 15859.

139.   Bordoli, L.; Kiefer, F.; Arnold, K.; Benkert, P.; Battey, J.; Schwede, T., Protein structure homology modeling using SWISS-MODEL workspace. *Nat. Protoc.* **2009,** *4* (1), 1-13.

140.   Awasthi, S.; Anbanandam, A.; Rodgers, K. K., Structure of a TLR4-interacting SPA4 peptide. *RSC advances* **2015,** *5* (35), 27431-27438.

141.   Tovchigrechko, A.; Vakser, I. A., GRAMM-X public web server for protein-protein docking. *Nucleic Acids Res.* **2006,** *34* (Web Server issue), W310-4.

142.   Lyne, P. D., Structure-based virtual screening: an overview. *Drug Discov. Today* **2002,** *7* (20), 1047-55.

143.   Haga, J. H.; Ichikawa, K.; Date, S., Virtual Screening Techniques and Current Computational Infrastructures. *Curr. Pharm. Des.* **2016,** *22* (23), 3576-84.

144.   Zhu, T.; Cao, S.; Su, P. C.; Patel, R.; Shah, D.; Chokshi, H. B.; Szukala, R.; Johnson, M. E.; Hevener, K. E., Hit identification and optimization in virtual screening: practical recommendations based on a critical literature analysis. *J. Med. Chem.* **2013,** *56* (17), 6560-72.

145.   Lionta, E.; Spyrou, G.; Vassilatis, D. K.; Cournia, Z., Structure-based virtual screening for drug discovery: principles, applications and recent advances. *Curr. Top. Med. Chem.* **2014,** *14* (16), 1923-38.

146.   Yan, X.; Liao, C.; Liu, Z.; Hagler, A. T.; Gu, Q.; Xu, J., Chemical Structure Similarity Search for Ligand-based Virtual Screening: Methods and Computational Resources. *Curr. Drug Targets* **2016,** *17* (14), 1580-1585.

147.   Sheng, C.; Dong, G.; Miao, Z.; Zhang, W.; Wang, W., State-of-the-art strategies for targeting protein-protein interactions by small-molecule inhibitors. *Chem. Soc. Rev.* **2015,** *44* (22), 8238-59.

148.   Takeuchi, O.; Kawai, T.; Muhlradt, P. F.; Morr, M.; Radolf, J. D.; Zychlinsky, A.; Takeda, K.; Akira, S., Discrimination of bacterial lipoproteins by Toll-like receptor 6. *Int. Immunol.* **2001,** *13* (7), 933-40.

149. Kang, J. Y.; Nan, X.; Jin, M. S.; Youn, S. J.; Ryu, Y. H.; Mah, S.; Han, S. H.; Lee, H.; Paik, S. G.; Lee, J. O., Recognition of lipopeptide patterns by Toll-like receptor 2-Toll-like receptor 6 heterodimer. *Immunity* **2009,** *31*, 873-84.

150. Zhong, Z.; Liu, L. J.; Dong, Z. Q.; Lu, L.; Wang, M.; Leung, C. H.; Ma, D. L.; Wang, Y., Structure-based discovery of an immunomodulatory inhibitor of TLR1-TLR2 heterodimerization from a natural product-like database. *Chem. Commun. (Camb.)* **2015,** *51* (56), 11178-81.

151. Abagyan, R.; Totrov, M.; Kuznetsov, D., ICM—A new method for protein modeling and design: Applications to docking and structure prediction from the distorted native conformation. *J. Comput. Chem.* **1994,** *15* (5), 488-506.

152. Cuevas, C. D.; Ross, S. R., Toll-like receptor 2-mediated innate immune responses against Junin virus in mice lead to antiviral adaptive immune responses during systemic infection and do not affect viral replication in the brain. *J. Virol.* **2014,** *88* (14), 7703-14.

153. Cheng, K.; Wang, X.; Zhang, S. Y.; Yin, H., Discovery of small molecule inhibitors of the TLR1-TLR2 complex. *Angew. Chem. Int. Ed.* **2012,** *51* (49), 12246-12249.

154. Murgueitio, M. S.; Henneke, P.; Glossmann, H.; Santos-Sierra, S.; Wolber, G., Prospective virtual screening in a sparse data scenario: design of small-molecule TLR2 antagonists. *ChemMedChem* **2014,** *9* (4), 813-22.

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

156. Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R., Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* **1997,** *267* (3), 727-48.

157. Verdonk, M. L.; Cole, J. C.; Hartshorn, M. J.; Murray, C. W.; Taylor, R. D., Improved protein-ligand docking using GOLD. *Proteins* **2003,** *52* (4), 609-23.

158. Jones, G.; Willett, P.; Glen, R. C., Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation. *J. Mol. Biol.* **1995,** *245* (1), 43-53.

159. Guan, Y.; Omueti-Ayoade, K.; Mutha, S. K.; Hergenrother, P. J.; Tapping, R. I., Identification of Novel Synthetic Toll-like Receptor 2 Agonists by High Throughput Screening. *J. Biol. Chem.* **2010,** *285*, 23755-62.

160. Zhou, S.; Cerny, A. M.; Bowen, G.; Chan, M.; Knipe, D. M.; Kurt-Jones, E. A.; Finberg, R. W., Discovery of a novel TLR2 signaling inhibitor with anti-viral activity. *Antiviral Res.* **2010,** *87* (3), 295-306.

161. Mistry, P.; Laird, M. H. W.; Schwarz, R. S.; Greene, S.; Dyson, T.; Snyder, G. A.; Xiao, T. S.; Chauhan, J.; Fletcher, S.; Toshchakov, V. Y.; MacKerell, A. D.; Vogel, S. N., Inhibition of TLR2 signaling by small molecule inhibitors targeting a pocket within the TLR2 TIR domain. *Proc. Natl. Acad. Sci. U.S.A.* **2015,** *112* (17), 5455-60.

162. Kopp, E. B.; Medzhitov, R., The Toll-receptor family and control of innate immunity. *Curr. Opin. Immunol.* **1999,** *11* (1), 13-8.

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

164. Kuntz, I. D.; Blaney, J. M.; Oatley, S. J.; Langridge, R.; Ferrin, T. E., A geometric approach to macromolecule-ligand interactions. *J. Mol. Biol.* **1982,** *161* (2), 269-88.

165. Kuntz, I. D., Structure-based strategies for drug design and discovery. *Science* **1992,** *257* (5073), 1078-82.

166.   Brooks, B. R.; Brooks, C. L.; MacKerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M., CHARMM: The Biomolecular Simulation Program. *J. Comput. Chem.* **2009,** *30* (10), 1545-1614.

167.   Gautam, J. K.; Ashish; Comeau, L. D.; Krueger, J. K.; Smith, M. F., Jr., Structural and functional evidence for the role of the TLR2 DD loop in TLR1/TLR2 heterodimerization and signaling. *J. Biol. Chem.* **2006,** *281* (40), 30132-42.

168.   Bell, J. K.; Botos, I.; Hall, P. R.; Askins, J.; Shiloach, J.; Segal, D. M.; Davies, D. R., The molecular structure of the Toll-like receptor 3 ligand-binding domain. *Proc. Natl. Acad. Sci. U.S.A.* **2005,** *102* (31), 10976-10980.

169.   Barton, G. M.; Medzhitov, R., Toll-like receptor signaling pathways. *Science* **2003,** *300* (5625), 1524-5.

170.   Assmann, T. S.; Brondani, L. d. A.; Bouças, A. P.; Canani, L. H.; Crispim, D., Toll-like receptor 3 (TLR3) and the development of type 1 diabetes mellitus. *Arch. Endocrinol. Metab.* **2015,** *59*, 4-12.

171.   Amarante, M. K.; Watanabe, M. A., Toll-like receptor 3: involvement with exogenous and endogenous RNA. *Int. Rev. Immunol.* **2010,** *29* (6), 557-73.

172.   Zhang, S. Y.; Jouanguy, E.; Ugolini, S.; Smahi, A.; Elain, G.; Romero, P.; Segal, D.; Sancho-Shimizu, V.; Lorenzo, L.; Puel, A.; Picard, C.; Chapgier, A.; Plancoulaine, S.; Titeux, M.; Cognet, C.; von Bernuth, H.; Ku, C. L.; Casrouge, A.; Zhang, X. X.; Barreiro, L.; Leonard, J.; Hamilton, C.; Lebon, P.; Heron, B.; Vallee, L.; Quintana-Murci, L.; Hovnanian, A.; Rozenberg, F.; Vivier, E.; Geissmann, F.; Tardieu, M.; Abel, L.; Casanova, J. L., TLR3 deficiency in patients with herpes simplex encephalitis. *Science* **2007,** *317* (5844), 1522-7.

173.   Wang, T.; Town, T.; Alexopoulou, L.; Anderson, J. F.; Fikrig, E.; Flavell, R. A., Toll-like receptor 3 mediates West Nile virus entry into the brain causing lethal encephalitis. *Nat. Med.* **2004,** *10* (12), 1366-73.

174.   Gowen, B. B.; Hoopes, J. D.; Wong, M. H.; Jung, K. H.; Isakson, K. C.; Alexopoulou, L.; Flavell, R. A.; Sidwell, R. W., TLR3 deletion limits mortality and disease severity due to Phlebovirus infection. *J. Immunol.* **2006,** *177* (9), 6301-7.

175.   Hutchens, M.; Luker, K. E.; Sottile, P.; Sonstein, J.; Lukacs, N. W.; Nunez, G.; Curtis, J. L.; Luker, G. D., TLR3 increases disease morbidity and mortality from vaccinia infection. *J. Immunol.* **2008,** *180* (1), 483-91.

176.   Pashine, A.; Valiante, N. M.; Ulmer, J. B., Targeting the innate immune response with improved vaccine adjuvants. *Nat. Med.* **2005,** *11* (4 Suppl), S63-8.

177.   Oshiumi, H.; Matsumoto, M.; Funami, K.; Akazawa, T.; Seya, T., TICAM-1, an adaptor molecule that participates in Toll-like receptor 3-mediated interferon-beta induction. *Nat. Immunol.* **2003,** *4* (2), 161-7.

178.   Cheng, K.; Wang, X.; Yin, H., Small-Molecule Inhibitors of the TLR3/dsRNA Complex. *J. Am. Chem. Soc.* **2011,** *133* (11), 3764–3767.

179.   Švajger, U.; Horvat, Ž.; Knez, D.; Rožman, P.; Turk, S.; Gobec, S., New antagonists of toll-like receptor 7 discovered through 3D ligand-based virtual screening. *Med. Chem. Res.* **2014,** *24* (1), 362-371.

180.   Heitmeier, M. R.; Scarim, A. L.; Corbett, J. A., Double-stranded RNA-induced inducible nitric-oxide synthase expression and interleukin-1 release by murine macrophages requires NF-kappaB activation. *J. Biol. Chem.* **1998,** *273* (24), 15301-7.

181.   Case, D. A.; Darden, T. A.; Cheatham, T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K. M.; Roberts, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Swails, J.; Götz, A. W.; Kolossváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wolf, R. M.; Liu, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Cai, Q.; Ye, X.; Wang, J.; Hsieh, M.-J.; Cui, G.; Roe, D. R.; Mathews, D. H.; Seetin, M. G.; Salomon-Ferrer, R.; Sagui, C.; Babin, V.; Luchko, T.; Gusarov, S.; Kovalenko, A.; Kollman, P. A. *AMBER 12*, University of California, San Francisco., 2012.

182.   Jarvis, R. A.; Patrick, E. A., Clustering Using a Similarity Measure Based on Shared Near Neighbors. *IEEE Trans. Comput.* **1973,** *22* (11), 1025-1034.

183.   Willett, J., *Similarity and Clustering in Chemical Information Systems*. John Wiley & Sons, Inc. New York, NY, USA: 1987; p 266.

184.   Willett, P., Similarity-based virtual screening using 2D fingerprints. *Drug Discov. Today* **2006,** *11* (23–24), 1046-1053.

185.   Todeschini, R.; Consonni, V.; Xiang, H.; Holliday, J.; Buscema, M.; Willett, P., Similarity Coefficients for Binary Chemoinformatics Data: Overview and Extended Comparison Using Simulated and Real Data Sets. *J. Chem. Inf. Model.* **2012,** *52* (11), 2884-2901.

186.   Salim, N.; Holliday, J.; Willett, P., Combination of Fingerprint-Based Similarity Coefficients Using Data Fusion. *J. Chem. Inf. Model.* **2003,** *43* (2), 435-442.

187.   Bajusz, D.; Rácz, A.; Héberger, K., Why is Tanimoto index an appropriate choice for fingerprint-based similarity calculations? *J. Cheminform.* **2015,** *7* (1), 1-13.

188.   Švajger, U.; Brus, B.; Turk, S.; Sova, M.; Hodnik, V.; Anderluh, G.; Gobec, S., Novel toll-like receptor 4 (TLR4) antagonists identified by structure- and ligand-based virtual screening. *Eur. J. Med. Chem.* **2013,** *70*, 393-9.

189.   Irwin, J. J.; Sterling, T.; Mysinger, M. M.; Bolstad, E. S.; Coleman, R. G., ZINC: a free tool to discover chemistry for biology. *J. Chem. Inf. Model.* **2012,** *52* (7), 1757-68.

190.   Hawkins, P. C.; Skillman, A. G.; Warren, G. L.; Ellingson, B. A.; Stahl, M. T., Conformer generation with OMEGA: algorithm and validation using high quality structures from the Protein Databank and Cambridge Structural Database. *J. Chem. Inf. Model.* **2010,** *50* (4), 572-84.

191.   Fontaine, F.; Bolton, E.; Borodina, Y.; Bryant, S. H., Fast 3D shape screening of large chemical databases through alignment-recycling. *Chem. Cent. J.* **2007,** *1* (1), 1-14.

192.   Joce, C.; Stahl, J. A.; Shridhar, M.; Hutchinson, M. R.; Watkins, L. R.; Fedichev, P. O.; Yin, H., Application of a novel in silico high-throughput screen to identify selective inhibitors for protein-protein interactions. *Bioorg. Med. Chem. Lett.* **2010,** *20* (18), 5411-3.

193.   Mahita, J.; Harini, K.; Rao Pichika, M.; Sowdhamini, R., An in silico approach towards the identification of novel inhibitors of the TLR-4 signaling pathway. *J. Biomol. Struct. Dyn.* **2015**, 1-18.

194.   Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J., Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* **2001,** *46* (1-3), 3-26.

195.   Duan, J.; Dixon, S. L.; Lowrie, J. F.; Sherman, W., Analysis and comparison of 2D fingerprints: Insights into database screening performance using eight fingerprint methods. *J. Mol. Graph. Model.* **2010,** *29* (2), 157-170.

196.   Sastry, M.; Lowrie, J. F.; Dixon, S. L.; Sherman, W., Large-Scale Systematic Analysis of 2D Fingerprint Methods and Parameters to Improve Virtual Screening Enrichments. *J. Chem. Inf. Model.* **2010,** *50* (5), 771-784.

197.   Farid, R.; Day, T.; Friesner, R. A.; Pearlstein, R. A., New insights about HERG blockade obtained from protein modeling, potential energy mapping, and docking studies. *Bioorg. Med. Chem.* **2006,** *14* (9), 3160-3173.

198.   Sherman, W.; Day, T.; Jacobson, M. P.; Friesner, R. A.; Farid, R., Novel Procedure for Modeling Ligand/Receptor Induced Fit Effects. *J. Med. Chem.* **2006,** *49* (2), 534-553.

199.   Sherman, W.; Beard, H. S.; Farid, R., Use of an Induced Fit Receptor Structure in Virtual Screening. *Chem. Biol. Drug Des.* **2006,** *67* (1), 83-84.

200.   Hacker, H.; Mischak, H.; Miethke, T.; Liptay, S.; Schmid, R.; Sparwasser, T.; Heeg, K.; Lipford, G. B.; Wagner, H., CpG-DNA-specific activation of antigen-presenting cells requires stress kinase activity and is preceded by non-specific endocytosis and endosomal maturation. *EMBO J.* **1998,** *17* (21), 6230-40.

201.   Hornung, V.; Guenthner-Biller, M.; Bourquin, C.; Ablasser, A.; Schlee, M.; Uematsu, S.; Noronha, A.; Manoharan, M.; Akira, S.; de Fougerolles, A.; Endres, S.; Hartmann, G., Sequence-specific potent induction of IFN-alpha by short interfering RNA in plasmacytoid dendritic cells through TLR7. *Nat. Med.* **2005,** *11* (3), 263-70.

202.   Blasius, A. L.; Beutler, B., Intracellular Toll-like Receptors. *Immunity* **2010,** *32* (3), 305-315.

203.   Eberle, F.; Sirin, M.; Binder, M.; Dalpke, A. H., Bacterial RNA is recognized by different sets of immunoreceptors. *Eur. J. Immunol.* **2009,** *39* (9), 2537-47.

204.   Diebold, S. S.; Kaisho, T.; Hemmi, H.; Akira, S.; Reis e Sousa, C., Innate antiviral responses by means of TLR7-mediated recognition of single-stranded RNA. *Science* **2004,** *303* (5663), 1529-31.

205.   Miller, R. L.; Gerster, J. F.; Owens, M. L.; Slade, H. B.; Tomai, M. A., Imiquimod applied topically: a novel immune response modifier and new class of drug. *Int. J. Immunopharmacol.* **1999,** *21* (1), 1-14.

206.   Shukla, N. M.; Malladi, S. S.; Day, V.; David, S. A., Preliminary evaluation of a 3H imidazoquinoline library as dual TLR7/TLR8 antagonists. *Bioorg. Med. Chem.* **2011,** *19* (12), 3801-11.

207.   Ohto, U.; Tanji, H.; Shimizu, T., Structure and function of toll-like receptor 8. *Microb. Infec.* **2014,** *16* (4), 273-282.

208.   Hornung, V.; Rothenfusser, S.; Britsch, S.; Krug, A.; Jahrsdorfer, B.; Giese, T.; Endres, S.; Hartmann, G., Quantitative expression of toll-like receptor 1-10 mRNA in cellular subsets of human peripheral blood mononuclear cells and sensitivity to CpG oligodeoxynucleotides. *J. Immunol.* **2002,** *168* (9), 4531-7.

209.   Marques, J. T.; Williams, B. R. G., Activation of the mammalian immune system by siRNAs. *Nat. Biotech.* **2005,** *23* (11), 1399-1405.

210.   Philbin, V. J.; Dowling, D. J.; Gallington, L. C.; Cortes, G.; Tan, Z.; Suter, E. E.; Chi, K. W.; Shuckett, A.; Stoler-Barak, L.; Tomai, M.; Miller, R. L.; Mansfield, K.; Levy, O., Imidazoquinoline Toll-like receptor 8 agonists activate human newborn monocytes and

dendritic cells through adenosine-refractory and caspase-1-dependent pathways. *J. Allergy Clin. Immunol.* **2012,** *130* (1), 195-204.e9.

211.   Tanji, H.; Ohto, U.; Shibata, T.; Miyake, K.; Shimizu, T., Structural reorganization of the Toll-like receptor 8 dimer induced by agonistic ligands. *Science* **2013,** *339* (6126), 1426-9.

212.   Yoo, E.; Salunke, D. B.; Sil, D.; Guo, X.; Salyer, A. C.; Hermanson, A. R.; Kumar, M.; Malladi, S. S.; Balakrishna, R.; Thompson, W. H.; Tanji, H.; Ohto, U.; Shimizu, T.; David, S. A., Determinants of activity at human Toll-like receptors 7 and 8: quantitative structure-activity relationship (QSAR) of diverse heterocyclic scaffolds. *J. Med. Chem.* **2014,** *57* (19), 7955-70.

213.   Kokatla, H. P.; Sil, D.; Tanji, H.; Ohto, U.; Malladi, S. S.; Fox, L. M.; Shimizu, T.; David, S. A., Structure-based design of novel human Toll-like receptor 8 agonists. *ChemMedChem* **2014,** *9* (4), 719-23.

214.   Kokatla, H. P.; Yoo, E.; Salunke, D. B.; Sil, D.; Ng, C. F.; Balakrishna, R.; Malladi, S. S.; Fox, L. M.; David, S. A., Toll-like receptor-8 agonistic activities in C2, C4, and C8 modified thiazolo[4,5-c]quinolines. *Org. Biomol. Chem.* **2013,** *11* (7), 1179-98.

215.   Kokatla, H. P.; Sil, D.; Malladi, S. S.; Balakrishna, R.; Hermanson, A. R.; Fox, L. M.; Wang, X.; Dixit, A.; David, S. A., Exquisite Selectivity for Human Toll-Like Receptor 8 in Substituted Furo[2,3-c]quinolines. *J. Med. Chem.* **2013,** *56* (17), 6871-6885.

216.   Salunke, D. B.; Yoo, E.; Shukla, N. M.; Balakrishna, R.; Malladi, S. S.; Serafin, K. J.; Day, V. W.; Wang, X.; David, S. A., Structure–Activity Relationships in Human Toll-like Receptor 8-Active 2,3-Diamino-furo[2,3-c]pyridines. *J. Med. Chem.* **2012,** *55* (18), 8137-8151.

217.   Schiaffo, C. E.; Shi, C.; Xiong, Z.; Olin, M.; Ohlfest, J. R.; Aldrich, C. C.; Ferguson, D. M., Structure–Activity Relationship Analysis of Imidazoquinolines with Toll-like Receptors 7 and 8 Selectivity and Enhanced Cytokine Induction. *J. Med. Chem.* **2014,** *57* (2), 339-347.

218.   http://zinc15.docking.org.

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

CHAPTER 2

Computational Techniques

2.   Computational Techniques In Drug Design And In Molecular Recognition Studies

The principal aim of computational techniques in drug design is to predict whether a given molecule or ligand will bind to a biological target and, if so, how strongly. This allows the computational design of novel molecules with enhanced (predicted) affinity, and guides the synthesis in the optimization protocol. When studying molecular recognition events, it also is important the prediction of binding poses between small molecules and macromolecules, and/or interactions between macromolecules (e.g., dimerization events). Docking techniques predict the possible binding poses at a concrete binding site of the target, and provides a first estimation of the binding energy. The estimation of the strength of the intermolecular interaction between the ligand and its biological target is most often performed by molecular mechanics (MM) and molecular dynamics (MD). These methods are also used to predict the conformation of the small molecules and to model conformational changes in the target that may occur when the small molecule binds to it. Semi-empirical, *ab initio* quantum mechanics methods and density functional theory (DFT) are often used to provide optimized parameters for the molecular mechanics calculations and also to estimate the electronic properties (electrostatic potential, polarizability, etc.) of the drug candidate that will influence binding affinity.

In this Thesis, diverse computational techniques have been applied to decipher some of the basis at the atomic level regarding the mechanism of functioning and the ligand recognition processes involving the TLR4/MD-2 system.

2.1   Docking

Molecular docking has become an important tool for drug discovery. Basically, the aim of molecular docking is to give a prediction of the ligand-receptor complex structure using computation methods. Docking can be achieved through two interrelated steps: first, by sampling conformations of the ligand in the active site of the protein; then, ranking these conformations via a scoring function. Ideally, sampling

algorithms should be able to reproduce the experimental binding mode, and the scoring function should also rank the best bound pose among all the generated conformations.

The interaction between protein and ligand has been thought as key-lock system. This key-lock model is easy to reproduce with a rigid docking where both ligand and protein are treated as rigid. The global vision of this model evolved with the advent of the induced-fit theory by Koshland, in which a mutual conformation change of both ligand and protein occurs at the moment of the interaction.[1] However, these two theories are now better substituted by the concept of conformational ensembles. The real three-dimensional structure will be represented by an ensemble of structures which interconvert through dynamic processes. Thus, the study of dynamics and its impact in conformation and molecular recognition is of paramount importance within this field. Karush defines this as a configurational adaptability in which the best-fitting configuration for a biomolecule after interacting with a ligand would become selected from the whole structural ensemble.[2] Furthermore, Weber suggested that the binding with a ligand shift the conformational equilibrium of the receptor in favor of those conformers in the dynamic ensemble that are most complementary to the ligand.[3]

The most popular method treats only the ligand as flexible, although receptor flexibility is becoming incorporated in several docking programs. The flexibility of a protein can be included into the binding site before the docking or afterward it is possible to generate different ensembles structures from MD simulation or from Normal Mode Analysis (NMA). Some programs use different strategy in order to include the flexibility in the receptor. For example, AutoDock4 program allows the selection of some bonds of the side chain of several residues as rotatable bonds when performing the docking of the ligand. Another strategy is to perform a conformational search and a refinement of some side chains of selected residues or in the region of the binding site while docking the ligand (*e.g.* Glide program).[4-5] Obviously this techniques have an higher computational cost than considering the receptor as rigid.[6]

In structure-based drug design, the scoring function is one of the most important components.[7] It is important to development of an energy scoring function that can rapidly and accurately describe the interaction between protein and ligand.

There are three important applications of scoring functions in molecular docking:

1.   The first of these is the determination of the binding mode and site of a ligand on a protein. Given a protein target, molecular docking generates hundreds of thousands of putative ligand binding orientations/conformations at the active site around the protein. A scoring function is used to rank these ligand orientations/conformations by evaluating the binding tightness of each of the putative complexes.

2.   The second application of a scoring function, which is related to the first application, is to predict the absolute binding affinity between protein and ligand. This is particularly important in lead optimization. Lead optimization refers to the process to improve the tightness of binding for low-affinity hits or lead compounds that have been identified. During this process, an accurate scoring function can greatly increase the optimization efficiency and save costs by computationally predicting the binding affinities between the protein and modified ligands before the much more expensive step of ligand synthesis and experimental testing.

3.   The third application, perhaps the most important one in structure-based drug design, is to identify the potential drug hits/leads for a given protein target by searching a large ligand database, i.e. virtual database screening. A reliable scoring function should be able to rank known binders most highly according to their binding scores during database screening. Given the expensive cost of experimental screening and sometimes unavailability of high-throughput assays, virtual database screening has played an increasingly important role in drug discovery.

All of these three applications, ligand binding mode identification, binding affinity prediction, and virtual database screening, are related to each other.

The different categories of scoring function could be:

-       Force field scoring function:

Force field (FF) scoring functions[8] are developed based on physical atomic interactions including van der Waals (VDW) interactions, electrostatic interactions, and bond stretching/bending/torsional forces. Force field functions and parameters are usually derived from both experimental data and ab initio quantum mechanical calculations according to the principles of physics. Despite its lucid physical meaning, a major challenge in the force field scoring functions is how to treat the solvent in ligand binding. The scoring function is composed of two energy components of Lennard-Jones VDW and an electrostatic term (Equation [2.1]):

$$E = \sum_i \sum_j \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} + \frac{q_i q_j}{\varepsilon(r_{ij}) r_{ij}} \right) \qquad \textbf{[2.1]}$$

where r_{ij} represents the distance between an atom i of the protein and an atom j of the ligand. Aij and Bij are the van der Walls parameters, and q_i and q_j are the atomic charges of the atom i and j respectively. Here, the effect of solvent is implicitly considered by introducing a simple distance-dependent dielectric constant $\varepsilon(r_{ij})$ n the Coulombic term.

- Empirical scoring function:

Empirical scoring function[9-10] estimates the binding affinity of a complex on the basis of a set of weighted energy terms. Compared to the force field scoring functions, the empirical scoring functions are much faster in binding score calculations due to their simple energy terms. Glide Score is one of the examples of empirical scoring function (Equation [2.2]):

$$\Delta G = \sum_i W_i \Delta G_i \qquad \textbf{[2.2]}$$

ΔG_i represents different energy terms such as VDW energy, electrostatics, hydrogen bond, desolvation, entropy, hydrophobicity, etc. ΔG_i stands for the individual empirical energy terms and the corresponding coefficients W_i.

- Knowledge-based scoring function:

Knowledge-based scoring function[11-12] employs energy potentials that are derived from the structural information embedded in experimentally determined atomic structures:

$$\omega(r) = -k_B T ln \left[\frac{\rho(r)}{\rho^*(r)} \right]$$

[2.3]

In this Equation [2.3] the scoring function is dependent on the density of the protein-ligand atom pair at the distance *r* ($\rho(r)$) in the training set, and the pair density in a reference state ($\rho^*(r)$) where there are no interatomic interactions at the absolute temperature T. k_B is the Boltzmann constant.

In this thesis, three different docking protocols were used: AutoDock4,[13] VINA[14] and Glide.[5, 15]

2.1.1 AutoDock4

AutoDock4[13] combines an empirical free energy force field with a Lamarckian Genetic Algorithm. The force field evaluates binding in two steps. The ligand and protein start in an unbound conformation. In the first step, the intramolecular energetics are estimated for the transition from these unbound states to the conformation of the ligand and protein in the bound state. The second step then evaluates the intermolecular energetics of combining the ligand and protein in their bound conformation.

The force field includes six pair-wise evaluations *(V)* and an estimate of the conformational entropy lost upon binding ΔS_{conf} (Equation [2.4])

$$\Delta G = (V_{bound}^{L-L} - V_{unbound}^{L-L}) + (V_{bound}^{P-P} - V_{unbound}^{P-P}) \\ + (V_{bound}^{P-L} - V_{unbound}^{P-L} + \Delta S_{conf})$$

[2.4]

where L refers to the "ligand" and P refers to the "protein" in a ligand-protein docking calculation.

Each of the pair-wise energetic terms includes evaluations for dispersion/repulsion, hydrogen bonding, electrostatics, and desolvation (Equation [2.5]):

$$V = W_{vdW} \sum_{i\,j} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}} \right) + W_{hbound} \sum_{i\,j} E(t) \left(\frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}} \right)$$
$$+ W_{elec} \sum_{i\,j} \left(\frac{q_i q_j}{\varepsilon(r_{ij}) r_{ij}} \right) + W_{sol} \sum_{i\,j} \left(S_i V_j + S_j V_i \right) e^{\left(-\frac{r_{ij}^2}{2\sigma^2} \right)}$$
[2.5]

The weighting constants W have been optimized to calibrate the empirical free energy based on a set of experimentally determined binding constants. The first term is a typical 6/12 potential for dispersion/repulsion interactions. The parameters are based on the Amber force field. The second term is a directional H-bond term based on a 10/12 potential. The parameters C and D are assigned to give a maximal well depth of 5 kcal/mol at 1.9Å for hydrogen bonds with oxygen and nitrogen, and a well depth of 1 kcal/mol at 2.5Å for hydrogen bonds with sulfur. The function E(t) provides directionality based on the angle t from ideal H-bonding geometry. The third term is a screened Coulomb potential for electrostatics. The final term is a desolvation potential based on the volume of atoms (V) that surround a given atom and shelter it from solvent, weighted by a solvation parameter (S) and an exponential term with distance-weighting factor $\sigma=3.5$Å.[16]

In the final ranking and clustering-based scoring methods also the loss of entropy (ΔS_{conf}) associated to the binding is considered. This contribution is strictly related to the number of rotatable bonds (N_{tors}) in the ligand (Equation [2.6]).

$$\Delta S_{conf} = W_{conf} N_{tors}$$
[2.6]

2.1.2 VINA

AutoDock VINA (Vina is not AutoDock) is an open-source molecular docking program.[14] It has no graphical interface but it is compatible with MGLTools.[17] However,

although MGLTools needs other files, such as AutoDock and AutoGrid parameter files (GPF, DPF) and grid map files, VINA does not need them. All it requires is the 3D structures of the molecules to be docked and the specification of the search space including the binding site. One limitation in VINA is that the maximum number of predicted binding poses is limited to 20 per ligand.

AutoDock VINA uses a hybrid scoring function (Equation [2.7]), where the summation is over all of the pairs of atoms that can move relative to each other, normally excluding 1–4 interactions. Each atom is assigned a type t_i, and a symmetric set of interaction functions $f_{t_i t_{ij}(r_{ij})}$ of the interatomic distance r_{ij} should be defined.

$$c = \sum_{i<j} f_{t_i t_{ij}(r_{ij})}$$ [2.7]

where i and j represent two atoms and t_i and t_j their types and r_{ij} the interatomic distance, excluding the 1-4 interaction. A symmetric set of interaction functions $f_{t_i t_j}$ should be defined. Basically, this sum corresponds to the sum of intermolecular and intramolecular contributions. Finally, the optimization algorithm used is the Iterated Local Search global optimizer.[18]

It is inspired by X-score[19] and tuned using the PDBbing[20-21] and extracting empirical information from both the conformational preferences of the receptor-ligand complexes and the experimental affinity measurements. It is both an empirical and a knowledge-based function. Regarding the optimization algorithm, the Iterated Local Search global optimizer is used, and to treat ligand flexibility and optimization, VINA uses a stochastic method with the Iterated Local Search global optimizer.[18, 22]

2.1.3   GLIDE

Glide is a commercial docking program provided by Schrödinger.[5, 15, 23] It uses a hierarchical series of filters to search for possible locations of the ligand in the active-site region of the receptor. It has a systematic method to treat ligand flexibility, with

an exhaustive search algorithm. The predicting binding affinity and the ranking of the binding pose is performed using GlideScore function based on ChemScore function (Equation [2.8]).

$$\Delta G_{bind} = C_o + C_{lipo} \sum f(r_{lr})$$
$$+ C_{hbond} \sum g(\Delta r)h(\Delta \alpha) \qquad \text{[2.8]}$$
$$+ C_{metal} \sum f(r_{lm}) + C_{rotb}H_{rotb}$$

The Glide protocol is intuitive and relies on 4 steps: the ligands and protein preparation, the receptor grid generation, and the docking process. Before launching the docking step, Glide has to generate a grid that represents the shape and the properties of the receptor, using several different sets of fields that provide progressively more accurate scoring of ligand poses. The grid permits to dock only the relevant region of the receptor, thus saving time calculations.

Regarding the last point, the full docking VS workflow includes 3 docking stages: HTVS, SP (Standard Precision) and XP (eXtra Precision). The first stage performs High Throughput Virtual Screening (HTVS) docking. It is intended for rapid screening of very large number of ligands and has much more restricted conformational sampling than SP docking. The second stage performs SP docking. It is appropriate for screening ligands of unknown quality in large numbers. The third stage is the XP docking and scoring. It is a more powerful and discriminating procedure using an implementation of a modified and expanded version of the ChemScore scoring function, called GlideScore[5] (Equation [2.9]) and categorized as an empirical scoring function. Glide can be used to perform virtual screening, accurate binding mode precision and furthermore, Glide exhibits excellent docking accuracy and high enrichment across a diverse range of receptor types.

$$\Delta G_{bind} = C_{lipo-lipo} \sum f(r_{lr})$$

$$+ C_{hbond-neut-neut} \sum g(\Delta r)h(\Delta \alpha)$$

$$+ C_{hbond-neut-charged} \sum g(\Delta r)h(\Delta \alpha)$$

$$+ C_{hbond-charged-charged} \sum g(\Delta r)h(\Delta \alpha) \qquad \textbf{[2.9]}$$

$$+ C_{max-metal-ion} \sum f(r_{lm}) + C_{rotb}H_{rotb}$$

$$+ C_{polar-phob}V_{polar-phob} + C_{coul}E_{coul} + C_{vdW}E_{vdW}$$

$$+ solvation\ terms$$

The lipophilic-lipophilic term and H-bonding term is same as defined in ChemScore. But the H-bonding term is separated into different weights that depend on whether the donor and acceptor are neutral and the other is charged, or both are charged. The metal-ligand interaction term also usesthe same functional form as in ChemScore but varies in three principal ways: a) this term considers only itneractions with anionic acceptor atoms; b) it ocunts just the single best interaction when two or more metal ligations are found; and c) the net charge on th metal ion is assessed in the unligated apo–protein. If the net carge is positive, the preferene for anionic ligand is incorporated into the scoring functio. Else, if net charge is neutral, the prefernce is suppressed. The seventh term in the fucntion rewards situations in chich a polar but non-H-bonding atom is found in a hydrophobic region.

The major components are the contributions from the Coulomb and vdW interactions energies between the ligand and the receptor (Equation [2.10]). Another major component is hte introduction of a solvation model and incoporation of the solvation effects. Glide docks explicit waters into the active site for each energetically competitive ligand pose and utilize empirical scoring terms that assess the exposure of various groups to the explicit waters.

$$XPGlideScore = E_{coul} + E_{vdW} + E_{bind} + E_{penaltyd} \qquad \textbf{[2.10]}$$

$$E_{bind} = E_{hyd\_enclosure} + E_{hb\_nn\_motif} + E_{hb\_cc\_motif} + E_{PI} + E_{hb\_pair} + E_{phobic\_pair}$$

[2.11]

where, $E_{hyd\_enclosure}$ represents hydrophobic enclosure; $E_{hb\_nn\_motif}$ represents the special neutral-neutral H-bond motifs; $E_{hb\_cc\_motif}$ represents special charged H-bond motifs, E_{PI} pi-stacking and pi-cation interactions; and E_{phobic_pair} hydrophobic atom-atom pair energy term. E_{hb_pair} term is same as defined in ChemScore scoring function.

$$E_{penalty} = E_{desolv} + E_{ligand\_strain}$$

[2.12]

where E_{desolv} represents desolvation penalties, and E_{ligand_strain} contac penalties (penalizing strain energy).

2.2 Molecular Mechanics and Molecular Dynamics Simulations

Molecular mechanics (MM) uses classical mechanics to model molecular systems. The potential energy of all systems in molecular mechanics is calculated using force fields. Molecular mechanics can be used to study molecule systems ranging in size and complexity from small to large biological systems or material assemblies with many thousands to millions of atoms.

The force field refers to the functional form and parameter sets used to calculate the potential energy of a system of atoms. All-atom force fields provide parameters for every type of atom in a system. Force field methods (also known as molecular mechanics; MM) ignore the electronic motions and calculate the energy of a system as a function of the nuclear positions only. The first assumption of these molecular mechanics is the Born-Oppenheimer approximation; MM is based upon a rather simple model of the interactions within a system with contributions from processes such as the stretching of bonds, the opening and closing angles and the rotations about single bonds.

The basic functional form of potential energy in MM includes bondend terms for interactions, and nonbonded terms that describe the long-range electrostaic and van der Waals forces, as in Equation [2.13].

$$V(R)_{total} = V(R)_{bonded} + V(R)_{non-bonded} \qquad \textbf{[2.13]}$$

The potential energy function (Equation [2.13]) could be divided into two main terms: bonded potential energy (Equation [2.14]) and non-bonded (Equation [2.15]):

$$V(R)_{bonded} = \sum_{bonds} K_b \, (b - b_o)^2 + \sum_{angles} K_\theta \, (\theta - \theta_o)^2$$
$$+ \sum_{dihedrals} K_\phi [1 + \cos(n\phi - \gamma)] \qquad \textbf{[2.14]}$$

In the case the of the bond energy term, K_b represents the force constant, while b and b_o correspond to the current bond leght and the equilibrated bond lenght, respectively.

In the case the of the angle-bending energy term, K_θ is the force constant for the system composed by the three atoms, and the θ and θ_o are the current angle and the angle at equilibrium, respectively. And finally in the case of the dihedral energy term, K_ϕ is the force constant, ϕ is the actual angle, and γ is the phse shift determining the position of the minima.

$$V(R)_{non-bonded} = \sum_{\substack{nonbonded \\ atom\ pairs}} \left(\varepsilon_{ij} \left[\left(\frac{R_{\min ij}}{r_{ij}} \right)^{12} - \left(\frac{R_{\min ij}}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j}{\varepsilon_D r_{ij}} \right) \qquad \textbf{[2.15]}$$

The non-bonded interactions consist of Lennard-Jones repulsion and dispersion $\varepsilon_{ij} \left[\left(\frac{R_{\min ij}}{r_{ij}} \right)^{12} - \left(\frac{R_{\min ij}}{r_{ij}} \right)^{6} \right]$ as well as Coulomb electrostatics $\frac{q_i q_j}{\varepsilon_D r_{ij}}$ (Figure **2.1**). The 12-6 Lennard-Jones (LJ) potential is used to describe the repulsion between two atoms (*i*

and *j*) in which the overlap of the electron clouds of both atoms induces dipoles generating an attractive component.
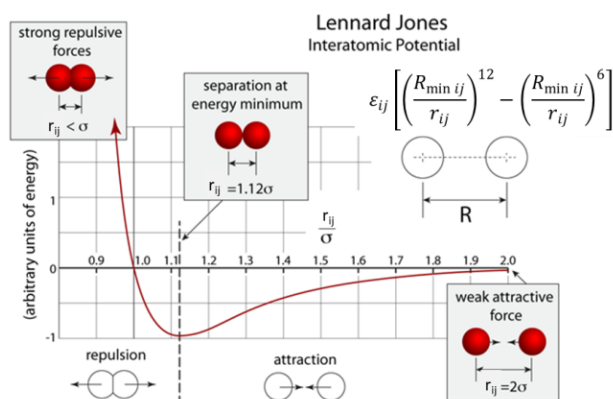
Figure 2.1. Lennard-Jones potential for two atoms (extracted and modified from http://atomsinmotion.com/book/chapter5/md).

The electrostatic term is calculated using Coulomb's law for the system, this electrostatic contribution due to the partial charges of each atom (q_i and q_j) in distance-dependent manner (r_{ij}), taking into account the permittivity of the solvent (ε_D). To study biological systems the most accurate approach would use QM calculations, but they are very time consuming (Figure **2.2**).
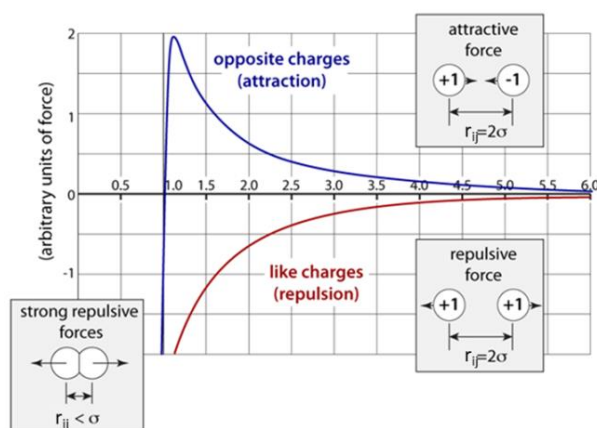
Figure 2.2. Representation of Coulombic electrostatic term (extracted and modified from http://atomsinmotion.com/book/chapter5/md)

The various contributions ara schematically represented in Figure **2.3**.
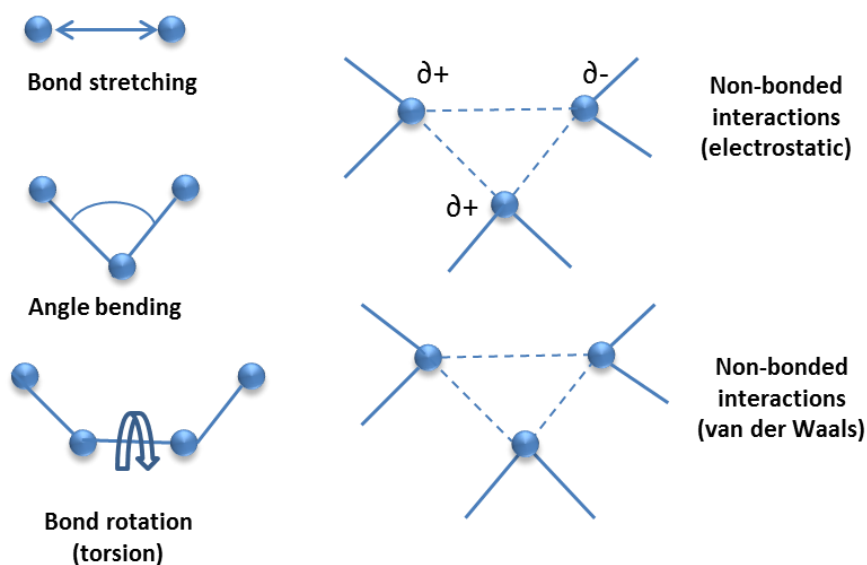
Figure 2.3. Schematic representation of the key contributions to a molecular mechanics force field: bond stretching, angle bending, torsional terms and non-bonded interactions.

Bond stretching: The most elementary approach is to use a Hooke's law in which the energy varies with the square of the displacement from the reference bond leght b_o.

Angle bending: The angle-bending energy term is also frequently described using a Hooke's law or harmonic potential. The contribution of each angle is characterized by a force constant and a reference value. Rather less is required to distort an angle away from equilibrium than to stretch or compress a bond, and the force constant are proportionately smaller.

Torsional terms: Most of the variation in structure and relative energies is due to the complex interplay between the torsional and non-bonded contributions. The exitence of barriers to rotation about chemical bonds is fundamental to understanding the structural properties of molecules and conformational analysis. Many force fields are used for modelling flexible molecules where the major changes in conformation

are due to rotations about bonds; in order to simulate this it is essential that the force field properly represents the energy profiles of such changes.

Non-bonded interactions: Independent molecules and atoms interact thourght non-bonded foreces, which also play an importatn role in determining the structure of individual molecular species. This non bonden interactions do not depend upon a specific bonding relationship between atoms. They are "through-space" interactions, and are usually modelled as a function of some inverse power of the distance. The non-bonden terms in a force fiel are usually considered in two groups, one comprisin electrostatic interactions ans the other van der Waals interactions.

Molecular Dynamics Simulation

MD relies on Newton's second law of motion, described in this equation:

$$\frac{d^2 x_i}{dt^2} = \frac{F_{xi}}{m_i}$$ **[2.16]**

where x_i is the position in a Cartesian representation of the particle i, F_{xi} is the force acting on such a particle with a mass of m_i, and t is the time. To solve this Equation [2.16], a Taylor series expansion is used as a propagator to simulate the system's time evolution. The position of the particle in step (n+1) is calculated using molecular coordinates and velocity from (n) and (n-1). Iterations continue until sufficient time steps (defined by the user) have been produced; the information (coordinationates and velocity) from all the steps constitute what is called MD trajectory. A classical MD simulation protocol is composed of four important steps:

1.   The structural data coming either from an experimental NMR/X-ray or homology modelling are prepared. Missing hydrogen atoms were added and protonation state of ionisable groups was computed by using Maestro Protein Preparation Wizard. Atom types and charges were assigned according to AMBER ff10 force field.

2. Minimization phase with the aim of removing the bad contacts between the solute and solvent, and a heating phase for increasing the atom velocities through gradual scaling until the system reaches the proper temperature. Initial velocities are generally calculated using the standard temperature-dependent Maxwell-Boltzmann distribution.

3. The equilibration phase which consists of allowing the system to relax, where energy, temperature, volume, pressure, and Root Men Square Deviation (RMSD) are monitore to ensure the stability of the system.

4. And the final step, production phase, where the trajectories are collected for further analysis.

Molecular dynamics simulation algorithms have been implemented in a number of simulation software packages: AMBER (Assisted Model building with Energy Refinment), CHARMM (Chemistry at Harvard Molecular Mechanics), Desmond, ESPResSo, GROMACS, GROMOS, LAMMPS, Martini, NAMD and OPLS (Optimized Potentials for Liquid Simulations). Based on classic mechanics and given appropriate initial conditions, 'in theory' we could calculate exactly how a system of interacting particles evolves over time, deriving any desired property of the system through statistical mechanics.

Accurate force fields are required for molecular mechanics. AMBER (Assisted Model building with Energy Refinment) is the most common force field used to describe proteins, developed in 1995 by Peter Kollman's group. The Merck Molecular Force Field (MMFF) and the General AMBER Force Field (GAFF) have been developed for the descripcion of small, drug-like molecules. Also lipids can be described in silico through force fields such as Lipid 14 force field. The AMBER force field was used for both ligand and protein, with the gaff parameter used for the ligand and the ff10 parameter used for the protein complex. MD simulations were run using the sander module of AMBER 14.

2.2.1 GAFF

GAFF[24] is a general force field for organic molecules, which has parameters for most organic and pharmaceutical molecules. This force field uses 33 basic atom types and 22 special atom types to cover almost all the chemical space composed of H, C, N, O, S, P, F, Cl, Br, and I. For the basic atom types, all the bond length, bond angle, and torsional angle parameters are available or can be calculated with empirical rules. Special atom types were introduced to describe certain chemical environments accurately, such as conjugated single and double bonds. The charge method used in GAFF is HF/6-31G* RESP charge. The van der Waals parameters of GAFF are as same as those used by the traditional AMBER force field.

2.2.2 GLYCAM06

GLYCAM06[25] is the most general biomolecular force filed for carbohydrates. GLYCAM development followed the general approach employed in biomolecular force fields of defining a single dihedral angle term for each molecular-class-specific linkage. QM calculations were employed to compute properties that are difficult or impossible to access experimentally, such as, bond and valence angle deformation force constants, dihedral angle rotational barriers, and electrostatic properties. The generality of the parameters is exemplified by the utilization of a common set of terms for α and β carbohydrate anomers. The feature to have just one C atom type facilitates the simulation of ring-flipping, having equilibrium between conformers with axial and equatorial substituents at the anomeric centre. This force field can be transferable to all carbohydrate ring conformations and sizes, also can be readily extendible to carbohydrate derivatives and other biomolecules, including monosaccharides and complex oligosaccharides, be rigorously assessed in terms of the relative accuracy of its component terms, and avoid the use of 1–4 electrostatic or non-bonded scaling factors. GLYCAM turns the 1-4 non-bonded interactions off to correctly reproduce the rotation of the ω-angle. In a study of the ω-angle rotation (O5—C5—C6—O6) in monosaccharides, we observed that O6 may interact with either O4 (in a 1–5

relationship) or O5 (in a 1–4 relationship) and the use of 1–4 scaling therefore unbalanced these interactions leading to an inability to correctly predict rotamer populations.

2.2.3   AMBER ffSB14

Ff14SB improve the accuracy of protein side chain and backbone parameters from ff99SB.[26] ff99SB uses the functional form and many of the parameters derived in ff94[27] and ff99,[28] which are associated with the Amber software.[29] Some characteristic features of ff94 include fixed partial charges on atom centers, explicit use of all hydrogen atoms, no specific functional form for hydrogen bonding, and dihedral parameters fit to relative quantum-mechanical (QM) energies of alternate rotamers of small molecules. In particular, the protein ϕ/ψ dihedrals have specific rotational parameters that affect relative energies of alternate backbone conformations. The "backbone" dihedral parameters (that can alter the secondary structure) are very important component of protein force fields. In AMBER, each dihedral profile is defined by a set of four atoms. The set of atoms used to define ϕ and ψ for glycine is as expected, following ϕ and ψ along the main chain (ϕ= C-N-Cα-C, ψ= N-Cα-C-N).[30]

2.2.4   LIPID 14

Lipid bilayers were set up and molecular dynamics run with Amber and the Lipid14 force field. Amber 14 includes Lipid14[31] which is a modular lipid AMBER force field, allowing the simulation of a number of lipids via the combination of different head and tail groups, a modular lipid force for tensionless lipid phospholipid simulations. Lipid14 includes the modular charge derivation framework developed in Lipid11[32] as well as a reparameterization of key van der Waals and dihedral angles as performed in GAFFlipid.[33] Hydrocarbon parameters have been refined, resulting in good reproduction of thermodynamic and dynamic properties for a number of simple carbon chains. To reproduce the experimental density and heat of vaporization of alkanes covering a range of chain lengths, LJ and torsion parameters were modified.

Given that both the torsion and LJ parameters affected the simulated density and heat of vaporization, these parameters were altered simultaneously, with the $CH_2-CH_2-CH_2-CH_2$ torsion being fitted to ab initio data. These modifications were tested analyzing different parameters of a selection of hydrocarbon chains. In the case of the lipid partial charge, standard AMBER RESP protocol was used to generate partial charges from quantum mechanical (QM) optimized structures, using six different orientations of a single conformation. A greater number of conformations were used per residue, with the partial charges calculated as an average over all conformations. The head and tail group starting structures were extracted from previous in-house bilayer simulations. This allows one to obtain Boltzmann weighted charges, introducing a temperature dependence. The electrostatic potential (ESP) was calculated directly from the conformations extracted from a bilayer simulation, with no QM optimization performed on those structures. Charges were derived using the standard AMBER RESP protocol.[34]

The modular nature of the force field allows for many combinations of lipid head groups and tail groups as well as rapid parameterization of further lipid types. Several van der Waals and dihedral angle parameters have been refined to fit experimental data and quantum energies as well as a new partial charge derivation for the head groups and tail groups. The force field was validated on six principle lipid bilayer types. The lipid bilayer structural features compare favorably with experimental measures such as area per lipid, bilayer thickness, NMR order parameters, scattering data, and lipid lateral diffusion. The interaction of other species, such as small molecules or proteins, with lipid membranes can be studied in AMBER using the Lipid14 force field.

2.3   Free energy of binding: MM-PBSA/MM-GBSA and MM-ISMSA

The Molecular Mechanics/Poisson-Boltzmann Surface Area (MM/PBSA) and the Molecular Mechanics/Generalized Born Surface Area (MM/GBSA) methods calculate

binding free energies for macromolecules by combining molecular mechanics calculations and continuum solvation models.[35]

The Gibbs free-energy formula is applied to calculate ΔG (equation [2.17]):

$$\Delta G = \Delta H - T\Delta S \qquad\qquad \textbf{[2.17]}$$

where ΔG is the variation of the free energy in a system, ΔH the variation of the enthalpy, ΔS the variation of the entropy and T the tempearture of the system.

The AMBER software[36] used to poduce MD simulations allow us to calculate the absolute free energies of binding by Molecular Mechanics Generalized Born (MM-GBSA) and Molecular Mechanics Poisson-Boltzmann Surface Area (MM-PBSA) approaches. The $\Delta G_{bind,solv}$ is calculated as the difference between ΔG of the solvated complex (ΔG_{solv}^{C}) and the sum of the ΔG of the solvated receptor(ΔG_{solv}^{R})and ligand (ΔG_{solv}^{L}), taking also in consideration the ΔG associated to the complex *in vacuum* (equation [2.18]).

$$\Delta G_{bind,solv} = \Delta G_{bind,vac} + \Delta G_{solv}^{C} - (\Delta G_{solv}^{R} + \Delta G_{solv}^{L}) \qquad \textbf{[2.18]}$$

The energy of the binding is calculated *in vacuum* (equation [2.19]), using classical methods and taking into account theenrgy differences associated to bonds, angles and dihedrals upon binding and the electrostaics and van der Waals terms (ΔE_{MM}) and also the conformatioanl entropy change ($T\Delta S$) computed by normal mode analysis (NMA) on a set of conformational snapshot from MD simulation.

$$\Delta G_{bind,vac} = \Delta E_{MM} - T\Delta S \qquad\qquad \textbf{[2.19]}$$

The solvent free energy (ΔG_{solv}) (equation [2.20]) has two contributions, the electrostatic (polar contribution) and non-electrostatic (non-polar) contributions. The electrostatic contribution ($\Delta G_{PB/GB}$) of the solvation to the free energy is calculated by solving either the Generalized Born (GB) or the Poisson-Boltzmann (PB) equation.

The non-electrostatic contribution (ΔG_{SA}) is assessed by solvent accessible surface area (SASA).

$$\Delta G_{solv} = \Delta G_{PB/GB} + \Delta G_{SA} \qquad\qquad \textbf{[2.20]}$$

Another an ultrafast and accurate scoring function for protein–protein docking is MM-ISMSA.[37] It includes a molecular mechanics (MM) part based on a 12–6 Lennard-Jones potential; an electrostatic component based on an implicit solvent model (ISM) with individual desolvation penalties for each partner in the protein–protein complex plus a hydrogen bonding term; and a surface area (SA) contribution to account for the loss of water contacts upon protein–protein complex formation.

2.4   Virtual Screening Strategies in the Search of Novel TLR4 Modulators

It is imperative to find new chemical entities as TLR modulators with drug-like properties in order to facilitate their development as drugs. In the context of drug discovery, virtual screening techniques have already proved to make hit identification more goal-oriented, allowing the access to a huge number of chemically diverse binders (from public and commercial databases) with a relatively low-cost in terms of time and materials. This computational approach has been subjected to extensive attention and revision over the years, from the early perspective of being an emerging method,[38] until the current time where new challenges are faced.[39-44] TLRs are not standard receptors which could be approached following classical strategies in drug design. The complexity of the system and the characteristics of their complexation with the PAMPs make them especially difficult to tackle following classical procedures in drug design and discovery. This is why TLRs constitute a special case study, in this context.

2.4.1 Virtual Screening Protocol

General strategies for a VS protocol include several steps that are summarized in Figure **2.3.** The availability of the 3D coordinates of the target is mandatory, either from X-ray crystallography, NMR or homology modeling. Prior knowledge about the ligand binding site may help in the identification of proper binders although, in some approaches, the search for novel binding pockets can be an additional interesting and challenging element in the drug discovery process.
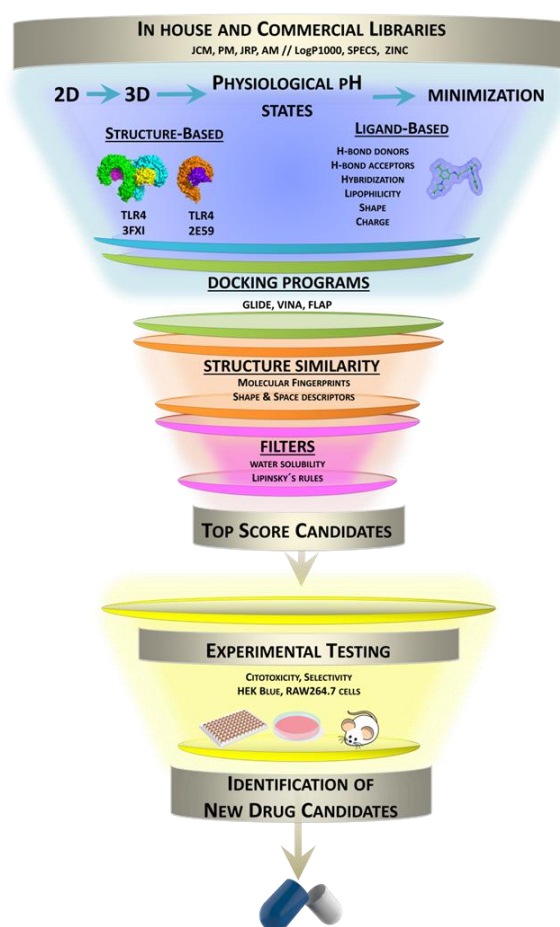
Figure 2.3. Summary of the VS protocols applied for the search of novel TLR modulators: access to databases and preparation/filtering of small-molecules; docking calculations; selection of candidates; experimental testing, and final identification of drug candidates.

2.4.2 Database Processing and Inclusion of Decoys

Database processing constitutes a fundamental step in VS approaches. It is crucial to generate the proper chemical library, with the adequate geometries, ionization states, conformations, etc. Furthermore, it is very important to discard any molecule that will not be a good candidate in the further steps of the VS study in relation to the particular system on hand. A good database processing will assure a rigorous and well-conducted virtual screening, as well as it will avoid computational cost and identification of unsuitable drug candidates.

To prepare the databases and inclusion of decoys are some different software:

- LigPrep,[45] a software created by Schrödinger LLC, is a collection of tools designed to prepare high quality, all-atom 3D structures for large numbers of drug-like molecules, starting from 2D or 3D structures. LigPrep starts by converting the input structure files to Maestro[46] format. The LigPrep process consists of a series of steps that perform conversions, apply corrections to the structures, generate variations on the structures, eliminate unwanted structures, and optimize the geometry. LigPrep produces a single low-energy 3D structure with defined chiralities for each processed input structure and it can also produce a number of structures from each starting geometry with varying ionization states, tautomeric forms, stereoisomers, and ring conformations. Additionally, LigPrep offers the option to eliminate molecules from the collection to be screened using various criteria including molecular weight or quantity and types of functional groups composing the molecule.

- AutoDockTools[47] is the graphical interface implemented within the Python Molecular Viewer to make AutoGrid and AutoDock (both are required to be used as docking program) widely accessible tools.[13, 47-48] It facilitates the formatting of input molecule files, with a set of methods that guide the user through protonation, calculation of charges, and specification of rotatable bonds in the ligand and the protein. As a brief outline of the preparation process, the ligand is loaded to the graphical interface, and ADT prepares it for AutoDock docking program. Polar hydrogens are added, charges are calculated, and nonpolar hydrogens are merged

with the heavier atoms to which they are attached. If the ligand file presents no charges, ADT will compute Gasteiger charges. Then, AutoDock atom types are assigned to each atom. Regarding the ligand preparation for the virtual screening, it is important to treat flexibility of the ligands. For this purpose, ligand flexibility is assigned in several steps. First, a root atom is chosen, which will act as the fixed position during coordinate transformation in the docking simulation. To find the optimal atom, the number of atoms in each branch is evaluated, and the root atom that minimizes the size of the largest branch is chosen. However, the ligand flexibility can be limited. As a limitation, each step in ADT has to be launched manually, one by one, as well as the preparation of each ligand. However, it is possible, with simple scripts, to do it automatically.

2.4.3 Docking Tools for VS

Molecular docking is a well-established method to investigate how a ligand interacts with its receptor. It integrates an automated computer algorithm that determines how a compound may bind in the active site of a target (binding mode and ligand/receptor interactions) and that tries to predict how tightly it binds (prediction of the binding energy), revealing the electrostatic and steric complementarity between the protein and the ligand.[49-51]

Nowadays, most of the docking programs are characterized by (i) the specific method to treat ligand flexibility,[52] which can be divided into three categories: systematic methods (incremental construction and conformational search); random or stochastic methods (Monte Carlo, Genetic Algorithms and Tabu search); and simulation methods[53] (molecular dynamics and energy minimization); (ii) the scoring function,[52] classified into three categories: ii.a) force field-based scoring functions,[54] where a classic force field is employed to compute the noncovalent ligand-target interactions, such as van der Waals and electrostatic energies (they are often augmented by a GB/SA or PB/SA term in order to account for the solvation effect); ii.b) empirical scoring function,[10] where the overall binding free energy is calculated by adding the contributions from several energetic terms, including hydrogen bond (H-

bond) interaction and hydrophobic interaction (the weighting factors of all terms are calibrated from a set of known complexes with experimentally determined structures and binding affinities); and ii.c) knowledge-based scoring functions,[11-12] where the ligand/target interactions are computed as a sum of distance-dependent statistical potentials between the ligand and the target (only the structural information of ligand-target complexes is needed, which is being accumulated rapidly due to structural biology advances).

1.   FLAP

The new molecular modeling tool FLAP (Fingerprints for Ligands and Proteins) provides a common reference framework for comparing molecules, using GRID Molecular Interaction Fields (MIFs). The MIFs describe the type, strength and direction of the molecular interactions between two biological partners. These MIFs are then condensed into discrete pharmacophoric points representing favorable and unfavorable interactions using a weighted energy-based and space coverage function. Using these discrete points, all four-point quadruplets are generated, and the resulting pharmacophore quadruplet fingerprint describes the target of interest. In addition to the fingerprints, the GRID MIFs are retained. The targets are then aligned by matching quadruplets in Cartesian space and a field similarity computed using the pre-calculated MIFs. Hence, the fingerprints are used to find matching pharmacophoric regions and the entire fields are used to score the match. Ligands and/or proteins can be the targets of interest.

FLAP consists of a graphical user interface, and several command-line programs which execute the various tasks. FLAP is based on four probes only, H, O, N1, and DRY, which respectively characterize the shape, hydrogen-bond acceptor, hydrogen-bond donor, and hydrophobic interactions.

We used this program to perform structure-based and ligand- based virtual screening. The aim of running structure-based VS is that the 3D structure of a receptor binding site is known, and can be used as a target for FLAP. When performing

alignments, the MIFs of a drug candidate are compared with the receptor MIFs. These similarities are used to rank the potential interactions of drug candidates with the receptor. Notice that the structure-based enrichments are in general lower than the ligand-based enrichments. This is most likely due to the fact that a structure-based search uses a larger cavity definition (there is more noise). Moreover, a ligand-based approach is based on the fact that active compounds should be chemically very similar, biasing the search results to favor a particular chemical species. The structure-based approach has the advantage that ligands binding different subpockets can be identified, and should not be as closely tied to the chemotype of a ligand template.

Virtual screening employing a Ligand-Based Approach is the best way to select potential new candidate drugs from a library of compounds with known three-dimensional structure but unknown activity against the biological target (the so-called "decoys"). Having access to a series of compounds with known activity on a specific biotarget, FLAP is able to align the molecules from the database (decoys) to one specific and chosen active compound (template). It then computes the GRID based - MIFs similarity) between decoys and the template assigning a score that can be used to rank the most similar compounds The assumption is that the higher the similarity with the template, the higher the probability of similar mechanism of action at the receptor site. Once scores are produced for each molecule of the dataset, it can be necessary to evaluate how well the known active compounds are recognized by FLAP through the use of an Enrichment Plot or a ROC curve.

2. GLIDE

Glide is a commercial docking program provided by Schrödinger[5, 15, 23] and designed to dock only the relevant region of the receptor, thus saving time calculations. The full docking VS workflow includes 3 docking stages: HTVS, SP (Standard Precision) and XP (Extra Precision). The first stage performs High Throughput Virtual Screening (HTVS) docking. It is intended for rapid screening of very large number of ligands and has much more restricted conformational sampling than SP

2. COMPUTATIONAL TECHNIQUES

docking. The second stage performs SP docking. It is appropriate for screening ligands of unknown quality in large numbers. The third stage is the XP docking and scoring. It is a more powerful and discriminating procedure using an implementation of a modified and expanded version of the ChemScore scoring function, called GlideScore[5] and categorized as an empirical scoring function. Glide can be used to perform virtual screening, accurate binding mode precision and furthermore, Glide exhibits excellent docking accuracy and high enrichment across a diverse range of receptor types.

2.4.4 Pan Assay Interference Compounds, PAINS

A computational tool for identifying new lead compounds for the validated drug targets in a virtual screening is high-throughput screening. However, it has also introduced a large number of molecules which interfere drug screening. These compounds, they are known by the name of Pan Assay Interference Compounds, PAINS, which interfere with the progress of drug screening in various ways, such as interfering with a biochemical assay, modifying the protein or aggregate-based inhibitors. PAINS may result from the presence of functional groups that can rect with biological molecules independent of a specific molecular recognition event. The physical properties of some small molecules coul cause them the falsely score as a hits in assays. So it is of vital significance to remove them. New leads are commonly identified using in-house discovery programs, competitor monitoring, and public sources of information, such as the current literature or databases of active molecules (ChemBl, PubChem, ChemSpider, UniChem, etc).[55-57] A number of substructural features which can help to identify compounds that appear as frequent hitters (promiscuous compounds) in many biochemical highthroughput screens.The compounds identified by such substructural features are not recognized by filters commonly used to identify reactive compounds. Even though these substructural features were identified using only one assay detection technology, such compounds have been reported to be active from many different assays. In fact, these compounds are increasingly prevalent in the literature as potential starting points for further exploration, whereas they may not be.[58] They have identified rhodanines, phenolic

Mannich bases, hydroxyphenylhydrazones, alkylidene barbiturates, alkylidene heterocycles, 1,2,3-aralkylpyrroles, activated benzofurazans, 2-amino-3-carbonylthiophenes, catechols and quinones as PAINS. The most common causes of PAINS activity are metal chelation, chemical aggregation, redox activity, compound fluorescence, cysteine oxidation or promiscuous binding. Many PAINS have multiple functionalities, causing different types of interference and resulting in in vitro and in vivo activity. From the pharmaceutical industry GSK has reported filters that their medicinal chemists use to reject molecules containing undesired functional groups.[59-60] Also Abbott reported their ALARM NMR tool and subsequent procedures to remove or flag potentially thiol reactive compounds.[56, 61] And Eli-Lilly reported on the filters they use at the front end of their open-innovation platform in order to discriminate the molecules they are interested in testing from unwanted ones.[62] In the academia, University of Dundee published a series of SMARTs filters to remove unwanted groups.[63] Sean Ekins et al. published a study toward the phenotypic screening of *Mycobacterium tuberculosis* and subsequent findings regarding unwanted groups.[64] And also Jonathan Baell from the Monash University published his report on pan assays interference compounds (PAINS) a list of structural features of frequent hitters from six different and independent assays. Computational filters exist to remove known PAINS from chemical libraries and an experienced medicinal chemist will be quickly able to identify a PAINS-type structure.

There are two servers to remove these PAINS: PAINS-remover, http://cbligand.org/PAINS/ and http://zinc15.docking.org/patterns/home/.

Bibliography:

1.   Koshland, D. E., The Key–Lock Theory and the Induced Fit Theory. *Angew. Chem. Int. Ed.* **1995,** *33* (23-24), 2375-2378.
2.   Karush, F., Heterogeneity of the Binding Sites of Bovine Serum Albumin1. *J. Am. Chem. Soc.* **1950,** *72* (6), 2705-2713.
3.   Weber, G., Ligand binding and internal equilibria in proteins. *Biochemistry* **1972,** *11* (5), 864-78.
4.   Morris, G. M.; Goodsell, D. S.; Huey, R.; Olson, A. J., Distributed automated docking of flexible ligands to proteins: parallel applications of AutoDock 2.4. *J. Comput. Aided Mol. Des.* **1996,** *10* (4), 293-304.
5.   Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S., Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem.* **2004,** *47* (7), 1739-49.
6.   Lexa, K. W.; Carlson, H. A., Protein flexibility in docking and surface mapping. *Q. Rev. Biophys.* **2012,** *45* (3), 301-43.
7.   Huang, S. Y.; Grinter, S. Z.; Zou, X., Scoring functions and their evaluation methods for protein-ligand docking: recent advances and future directions. *Physical chemistry chemical physics : PCCP* **2010,** *12* (40), 12899-908.
8.   Huang, N.; Kalyanaraman, C.; Irwin, J. J.; Jacobson, M. P., Physics-based scoring of protein-ligand complexes: enrichment of known inhibitors in large-scale virtual screening. *Journal of chemical information and modeling* **2006,** *46* (1), 243-53.
9.   Bohm, H. J., The development of a simple empirical scoring function to estimate the binding constant for a protein-ligand complex of known three-dimensional structure. *Journal of computer-aided molecular design* **1994,** *8* (3), 243-56.
10.   Eldridge, M. D.; Murray, C. W.; Auton, T. R.; Paolini, G. V.; Mee, R. P., Empirical scoring functions: I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. *Journal of computer-aided molecular design* **1997,** *11* (5), 425-45.
11.   Verkhivker, G.; Appelt, K.; Freer, S. T.; Villafranca, J. E., Empirical free energy calculations of ligand-protein crystallographic complexes. I. Knowledge-based ligand-protein interaction potentials applied to the prediction of human immunodeficiency virus 1 protease binding affinity. *Protein engineering* **1995,** *8* (7), 677-91.
12.   Gohlke, H.; Hendlich, M.; Klebe, G., Knowledge-based scoring function to predict protein-ligand interactions. *Journal of molecular biology* **2000,** *295* (2), 337-56.
13.   Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J., AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility. *J. Comput. Chem.* **2009,** *30* (16), 2785-2791.
14.   Trott, O.; Olson, A. J., AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. *J. Comput. Chem.* **2010,** *31* (2), 455-461.
15.   Halgren, T. A.; Murphy, R. B.; Friesner, R. A.; Beard, H. S.; Frye, L. L.; Pollard, W. T.; Banks, J. L., Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *J. Med. Chem.* **2004,** *47* (7), 1750-9.

16.   Stouten, P. F. W.; Frömmel, C.; Nakamura, H.; Sander, C., An Effective Solvation Term Based on Atomic Occupancies for Use in Protein Simulations. *Mol. Simul.* **1993,** *10* (2-6), 97-120.

17.   Janeway, C. A., Approaching the asymptote? Evolution and revolution in immunology. *Cold Spring Harb. Symp. Quant. Biol.* **1989,** *54*, 1–13.

18.   Baxter, J., Local Optima Avoidance in Depot Location. *J. Oper. Res. Soc.* **1981,** *32* (9), 815-819.

19.   Wang, R.; Lai, L.; Wang, S., Further development and validation of empirical scoring functions for structure-based binding affinity prediction. *Journal of computer-aided molecular design* **2002,** *16* (1), 11-26.

20.   Wang, R.; Fang, X.; Lu, Y.; Wang, S., The PDBbind database: collection of binding affinities for protein-ligand complexes with known three-dimensional structures. *J. Med. Chem.* **2004,** *47* (12), 2977-80.

21.   Wang, R.; Fang, X.; Lu, Y.; Yang, C. Y.; Wang, S., The PDBbind database: methodologies and updates. *J. Med. Chem.* **2005,** *48* (12), 4111-9.

22.   Blum, C.; Blesa Aguilera, M. J.; Roli, A.; Sampels, M., *Hybrid Metaheuristics: An Emerging Approach to Optimization*. Springer Publishing Company, Incorporated: 2008; p 290.

23.   Friesner, R. A.; Murphy, R. B.; Repasky, M. P.; Frye, L. L.; Greenwood, J. R.; Halgren, T. A.; Sanschagrin, P. C.; Mainz, D. T., Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *J. Med. Chem.* **2006,** *49* (21), 6177-96.

24.   Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A., Development and testing of a general amber force field. *J. Comput. Chem.* **2004,** *25* (9), 1157-74.

25.   Kirschner, K. N.; Yongye, A. B.; Tschampel, S. M.; Gonzalez-Outeirino, J.; Daniels, C. R.; Foley, B. L.; Woods, R. J., GLYCAM06: a generalizable biomolecular force field. Carbohydrates. *J. Comput. Chem.* **2008,** *29* (4), 622-55.

26.   Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C., ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* **2015,** *11* (8), 3696-713.

27.   Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A., A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* **1995,** *117* (19), 5179-5197.

28.   Wang, J.; Cieplak, P.; Kollman, P. A., How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J. Comput. Chem.* **2000,** *21* (12), 1049-1074.

29.   Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M., Jr.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J., The Amber biomolecular simulation programs. *J. Comput. Chem.* **2005,** *26* (16), 1668-88.

30.   Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C., Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins* **2006,** *65* (3), 712-25.

31.   Dickson, C. J.; Madej, B. D.; Skjevik, Å. A.; Betz, R. M.; Teigen, K.; Gould, I. R.; Walker, R. C., Lipid14: The Amber Lipid Force Field. *J. Chem. Theory Comput.* **2014,** *10* (2), 865-879.

32.   Skjevik, A. A.; Madej, B. D.; Walker, R. C.; Teigen, K., LIPID11: a modular framework for lipid simulations using amber. *J. Phys. Chem. B.* **2012,** *116* (36), 11124-36.

33.   Dickson, C. J.; Rosso, L.; Betz, R. M.; Walker, R. C.; Gould, I. R., GAFFlipid: a General Amber Force Field for the accurate molecular dynamics simulation of phospholipid. *Soft. Matter.* **2012,** *8* (37), 9617-9627.

34.   Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A., A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J. Phys. Chem.* **1993,** *97* (40), 10269-10280.

35.   Hou, T.; Wang, J.; Li, Y.; Wang, W., Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. *Journal of chemical information and modeling* **2011,** *51* (1), 69-82.

36.   Case, D. A.; Darden, T. A.; Cheatham, T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K. M.; Roberts, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Swails, J.; Götz, A. W.; Kolossváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wolf, R. M.; Liu, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Cai, Q.; Ye, X.; Wang, J.; Hsieh, M.-J.; Cui, G.; Roe, D. R.; Mathews, D. H.; Seetin, M. G.; Salomon-Ferrer, R.; Sagui, C.; Babin, V.; Luchko, T.; Gusarov, S.; Kovalenko, A.; Kollman, P. A. *AMBER 12*, University of California, San Francisco., 2012.

37.   Klett, J.; Nunez-Salgado, A.; Dos Santos, H. G.; Cortes-Cabrera, A.; Perona, A.; Gil-Redondo, R.; Abia, D.; Gago, F.; Morreale, A., MM-ISMSA: An Ultrafast and Accurate Scoring Function for Protein-Protein Docking. *J. Chem. Theory Comput.* **2012,** *8* (9), 3395-408.

38.   Lyne, P. D., Structure-based virtual screening: an overview. *Drug Discov. Today* **2002,** *7* (20), 1047-55.

39.   Schneider, G., Virtual screening: an endless staircase? *Nat. Rev. Drug Discov.* **2010,** *9* (4), 273-6.

40.   Haga, J. H.; Ichikawa, K.; Date, S., Virtual Screening Techniques and Current Computational Infrastructures. *Curr. Pharm. Des.* **2016,** *22* (23), 3576-84.

41.   Zhu, T.; Cao, S.; Su, P. C.; Patel, R.; Shah, D.; Chokshi, H. B.; Szukala, R.; Johnson, M. E.; Hevener, K. E., Hit identification and optimization in virtual screening: practical recommendations based on a critical literature analysis. *J. Med. Chem.* **2013,** *56* (17), 6560-72.

42.   Lionta, E.; Spyrou, G.; Vassilatis, D. K.; Cournia, Z., Structure-based virtual screening for drug discovery: principles, applications and recent advances. *Curr. Top. Med. Chem.* **2014,** *14* (16), 1923-38.

43.   Yan, X.; Liao, C.; Liu, Z.; Hagler, A. T.; Gu, Q.; Xu, J., Chemical Structure Similarity Search for Ligand-based Virtual Screening: Methods and Computational Resources. *Curr. Drug Targets* **2016,** *17* (14), 1580-1585.

44.   Sheng, C.; Dong, G.; Miao, Z.; Zhang, W.; Wang, W., State-of-the-art strategies for targeting protein-protein interactions by small-molecule inhibitors. *Chem. Soc. Rev.* **2015,** *44* (22), 8238-59.

45.   Hennessy, E. J.; Parker, A. E.; O'Neill, A. J., Targeting of Toll-like receptors: Emerging therapeutics? *Nat. Rev. Drug Discov.* **2010,** *9*, 293–307.

46.   Akira, S.; Takeda, K., Toll-like receptor signalling. *Nat. Rev. Immunol.* **2004,** *4*, 499-511.

47.   Takeuchi, O.; Kawai, T.; Muhlradt, P. F.; Morr, M.; Radolf, J. D.; Zychlinsky, A.; Takeda, K.; Akira, S., Discrimination of bacterial lipoproteins by Toll-like receptor 6. *Int. Immunol.* **2001,** *13* (7), 933-40.

48.   Morris, G. M.; Huey, R.; Olson, A. J., Using AutoDock for ligand-receptor docking. *Curr. Protoc. Bioinformatics* **2008,** *Chapter 8*, Unit 8.14.

49.   Sousa, S. F.; Ribeiro, A. J.; Coimbra, J. T.; Neves, R. P.; Martins, S. A.; Moorthy, N. S.; Fernandes, P. A.; Ramos, M. J., Protein-ligand docking in the new millennium--a retrospective of 10 years in the field. *Curr. Med. Chem.* **2013,** *20* (18), 2296-314.

50.   Moroni, E.; Paladino, A.; Colombo, G., The Dynamics of Drug Discovery. *Curr. Top. Med. Chem.* **2015,** *15* (20), 2043-55.

51.   Spyrakis, F.; Cavasotto, C. N., Open challenges in structure-based virtual screening: Receptor modeling, target flexibility consideration and active site water molecules description. *Arch. Biochem. Biophys.* **2015,** *583*, 105-19.

52.   Halperin, I.; Ma, B.; Wolfson, H.; Nussinov, R., Principles of docking: An overview of search algorithms and a guide to scoring functions. *Proteins* **2002,** *47* (4), 409-443.

53.   Sliwoski, G.; Kothiwale, S.; Meiler, J.; Lowe, E. W., Computational Methods in Drug Discovery. *Pharmacol. Rev.* **2014,** *66* (1), 334-395.

54.   Genheden, S.; Ryde, U., The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert. Opin. Drug. Discov.* **2015,** *10* (5), 449-61.

55.   Xie, X. Q., Exploiting PubChem for Virtual Screening. *Expert Opin. Drug Discov.* **2010,** *5* (12), 1205-1220.

56.   Huth, J. R.; Song, D.; Mendoza, R. R.; Black-Schaefer, C. L.; Mack, J. C.; Dorwin, S. A.; Ladror, U. S.; Severin, J. M.; Walter, K. A.; Bartley, D. M.; Hajduk, P. J., Toxicological evaluation of thiol-reactive compounds identified using a la assay to detect reactive molecules by nuclear magnetic resonance. *Chem. Res. Toxicol.* **2007,** *20* (12), 1752-9.

57.   http://depth-first.com/articles/2011/10/12/sixty-four-free-chemistry-databases/ Sixty-Four Free Chemistry Databases.

58.   Baell, J. B.; Holloway, G. A., New Substructure Filters for Removal of Pan Assay Interference Compounds (PAINS) from Screening Libraries and for Their Exclusion in Bioassays. *J. Med. Chem.* **2010,** *53* (7), 2719-2740.

59.   Nadin, A.; Hattotuwagama, C.; Churcher, I., Lead-oriented synthesis: a new opportunity for synthetic chemistry. *Angew. Chem. Int. Ed. Engl.* **2012,** *51* (5), 1114-22.

60.   Hann, M.; Hudson, B.; Lewell, X.; Lifely, R.; Miller, L.; Ramsden, N., Strategic Pooling of Compounds for High-Throughput Screening. *J. Chem. Inform. Comput. Sci.* **1999,** *39* (5), 897-902.

61.   Huth, J. R.; Mendoza, R.; Olejniczak, E. T.; Johnson, R. W.; Cothron, D. A.; Liu, Y.; Lerner, C. G.; Chen, J.; Hajduk, P. J., ALARM NMR: A Rapid and Robust Experimental Method To Detect Reactive False Positives in Biochemical Screens. *J. Am. Chem. Soc.* **2005,** *127* (1), 217-224.

62.   Bruns, R. F.; Watson, I. A., Rules for Identifying Potentially Reactive or Promiscuous Compounds. *J. Med. Chem.* **2012,** *55* (22), 9763-9772.

63.   Brenk, R.; Schipani, A.; James, D.; Krasowski, A.; Gilbert, I. H.; Frearson, J.; Wyatt, P. G., Lessons learnt from assembling screening libraries for drug discovery for neglected diseases. *ChemMedChem* **2008,** *3* (3), 435-44.

64.   Ekins, S.; Kaneko, T.; Lipinski, C. A.; Bradford, J.; Dole, K.; Spektor, A.; Gregory, K.; Blondeau, D.; Ernst, S.; Yang, J.; Goncharoff, N.; Hohman, M. M.; Bunin, B. A., Analysis and hit filtering of a very large library of compounds screened against Mycobacterium tuberculosis. *Mol. Biosyst.* **2010,** *6* (11), 2316-24.

CHAPTER 3

Computational Studies of Reported

TLR4/MD-2 Ligands

3.1   Introduction

As mentioned in Chapter 1, LPSs are large glycolipids consisting of a highly conserved lipid moiety, known as lipid A, and a polysaccharide composed of an oligosaccharidic core (outer and inner part) and a *O*-specific polysaccharide. Recently, structural information on TLR4 has become available, providing insights on the binding at the atomic level.[1] In the X-ray crystallographic structure of TLR4/MD-2 in complex with potent agonist *Escherichia coli* LPS (PDB- ID: 3FXI), MD-2 is able to accommodate up to five fatty acid (FA) chains of the *E. coli* lipid A into its large hydrophobic cavity. The sixth chain protrudes from the MD-2 hydrophobic pocket, thus completing the dimerization interface. The phosphate groups of the LPS are anchored to the polar MD-2 rim, and the polysaccharide moiety establishes a network of polar interactions with TLR4. Therefore, all of the structural components of the LPS molecule play a fundamental role in the TLR4/MD-2 recognition and in the binding.[2]

Lipid A is composed of FA chains of different lengths attached to a 1,4-β-diphosphorylated diglucosamine backbone (Figure **3.1**). The agonistic activity of lipid A has been mainly assigned to the number (established as six), length and chemical structure of the attached FA chains, as well as variability in the level of phosphorylation and the number and types of substituted groups found attached to the phosphate residues. Recent findings have pointed to the need of revisiting this paradigm since some examples of LPS bearing a penta-acylated lipid A, together with positively-charged residues decorating the lipid A, have been reported with immunostimulatory ability. Therefore, these data suggest that subtle changes in lipid A structure may profoundly impact the innate immune response from the host.
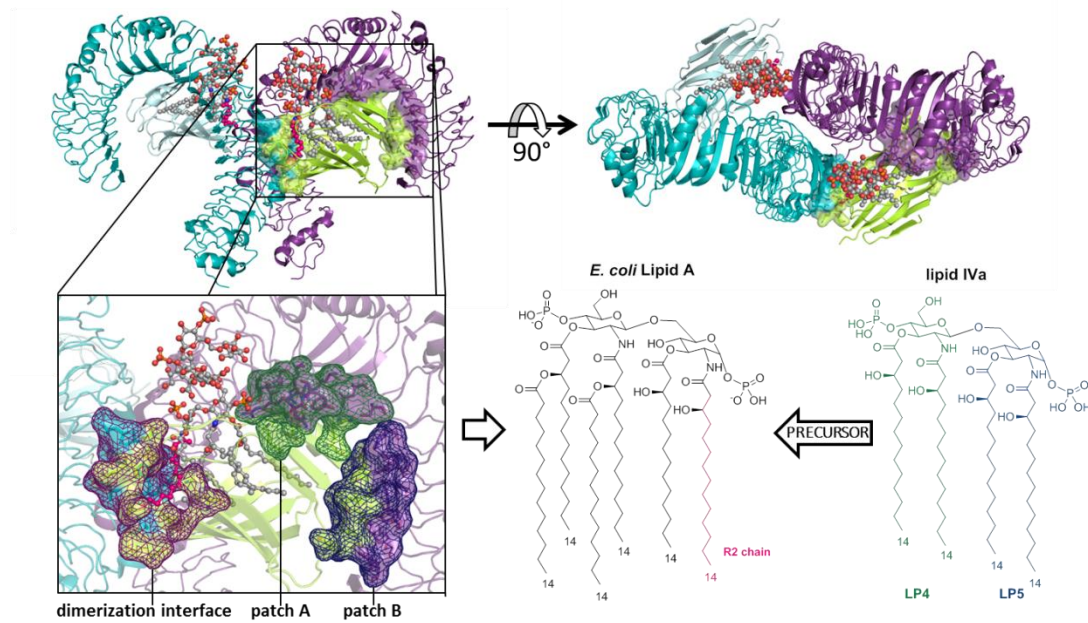
Figure 3.1. Showing the contact interface of the dimerized TLR4/MD-2 complex bound to LPS. (PDB-ID 3FXI). At the right, primary contact interface between the MD-2 in purple and the TLR4 in green, which includes two different contact regions, the Patch A and Patch B. At the left the dimerization interface, remarking the channel created in the surroundings of Phe126.

In a previous work performed by members of our group, the study was focused in the binding of lipid IVa,[3] which is a synthetic analogue of lipid A, that is the most outer region of the LPS used by the gram-negative bacteria to anchor the host outer membrane during the invasion process. The main interest of studying the binding of lipid IVa is the difference in the mechanism action of the molecule across different species; in particular, lipid IVa is an agonist of the immune response in murine and an antagonist in human. Lipid IVa bound conformation differs significantly when looking at the mouse and human crystal structures by having opposite orientation in the direction of the sugar moiety and phosphate groups. Interestingly there are certain substitutions in the MD-2 sequence that may stabilize these two different conformations of the lipid IVa. The strongest interaction found out from our energy analysis is the polar interaction between the Lys122 from MD-2 that interacts with and the O5 from the sugars and the oxygen from the glycosidic linkage. Lys122 from the MD-2 is substituted by Glu122 in

mouse that interacts with the hydroxyls from the sugar moieties. This is the main difference found in the protein sequence that facilitates two different binding modes of lipid IVa in terms of polar interactions; other substitutions were found, but only affecting to hydrophobic residues which kept their hydrophobic nature. It is remarkable that, in human, there is an interaction between Lys122 from the MD-2 and the oxygen atoms from the sugar core of lipid IVa exactly in the same way as the Ly360 from the TLR4 is interacting with the lipid IVa. Another mayor difference we found that the antagonist human conformation interaction pattern lacks of important interactions with the TLR4. In the mouse agonist system we observed the interaction with the O5 from the sugar moieties and oxygen atom from the glycosidic linkage with Lys360 from the TLR4, this interaction could contribute to stabilize the agonist conformation in terms of the global motion previously described for the maintaining the TLR4/MD-2 complex.

Small molecules pyrimido[5,4- b]indoles have shown to stimulate TLR4 and could potentially be used as adjuvants or immune modulators.[4] Synthetic analogues of natural product Euodenine A have exhibited potent and selective agonistic activity towards TLR4.[5] It has also been reported that synthetic peptides mimic the TLR4/LPS interactions.[6] On the other hand, several small non LPS molecules with TLR4 antagonist activity have also been developed, such as ethyl 4-oxo-4-(oxazolidin-3-yl)-butenoate derivatives (OSL07), benzothiazole-based inhibitors, ethyl phenyl-sulfamoylcyclohexenecarboxylate derivatives (TAK-242 or resatorvid), and β-amino alcohol derivatives.[7-10] However, no successful progress was shown in clinical phases (for example, in the case of compound OSL07).

Due to the wide range of possible therapeutic usages of TLR4 modulators, several approaches for their design have been exploited, leading to four general classes of therapeutics: LPS mimetics, small molecule modulators, peptides, and monoclonal antibodies.[11] Compared with all the designed TLR4 therapeutic agents, small molecules show interesting advantages given that LPS analogues show important solubility and toxicity limitations,[12] and peptides have poor pharmacokinetics.[13] Therefore, small molecules that do not structurally resemble to LPS represent a different promising and interesting alternative. A deep and extensive study of the

molecular mechanisms by which these TLR4 ligands modulate TLR4 activity emerges as an interesting strategy to understand TLR4 pharmacodynamics, which is essential for drug design and development processes. Also computational chemistry methodologies appear as necessary, powerful and time-saving tools for the prediction of ligand binding modes with the TLR4/MD-2 systems.

There is experimental evidence confirming that most of the TLR4 small molecule modulators exert their pharmacological activity through binding to MD-2 inner pocket. Among them we found the TLR4 agonists such as paclitaxel,[14] some opioids (remifentanil),[15] and TLR4 antagonists like xanthohumol (XN),[16] sulforaphane (SFN),[17] N-pyrene maleimide (NPM), dalcetrapib (JTT705), iodoacetylaminonaphtyl sulfate (IAANS),[18] curcumin,[19] cinnamamides,[20] and certain opioids such as naxolone.[15] However, TLR4 small molecule modulators show varied physicochemical properties, suggesting that not all of them may bind to MD-2 pocket, which is extremely hydrophobic. Additionally, MD-2 binding ligands may not be exerting its function by binding uniquely to MD-2 pocket, modulating the receptor by binding to other unknown binding sites.[21]

The presence of drug binding sites other than MD-2 has been further confirmed by experimental studies which proof that the TLR4 agonist as DiC14-amidine acts as a protein-protein interaction (PPI) stabilizing agent, binding to the TLR4/TLR4* interface of the activated system.[22] There is also experimental evidence that this dimerization interface is targeted by nickel ions, which activate the receptor complex.[23] Furthermore, although there is no experimental evidence for drug binding, Gobec and coworkers and Yin and coworkers[4] reported series of TLR4 modulating agents which presumably act as PPI inhibitors by disrupting the TLR4/MD-2 interface in the TLR4/MD-2 system.[24] Other compounds, such as 4-aminoquinazolines and pyrimido-[5-4,b]-indoles show TLR4 agonist activity through binding to the TLR4/MD-2* interface of the TLR4/MD-2/TLR4*/MD-2 system.[4, 25] The design and development of PPI stabilizing agents and PPI inhibitors has appeared as a relatively new approach for the achievement of drugs with different pharmacodynamic profile.[26]

Many of the ligands that were found to modulate TLR4 activity do not present neither experimental nor computational studies of their binding mode: this was the case of the heme group,[27] bryostatin-1 (bryo-1)[28] and certain polyoxygenated cholesterol ester hydroperoxides (BEP-CE),[29] as these compounds were proven to activate TLR4. Additionally, there are several compounds which are able to inhibit the dimerization of the TLR4/MD-2 system whose binding is unknown, such as glycyrrhizin, and isoliquiritigenin.[30] Furthermore, although eudonenines require MD-2 to activate TLR4, their binding modes have not been yet fully defined.[5] Finally, the binding mode of certain tryciclics such as imipramine, desipramine and amitriptyline is only supported by computational studies thus there is no experimental evidence proving their binding to MD-2.[31]

Our work aimed to propose a binding mode for some reported TLR4/MD-2 binders, and to understand, at atomic level, the main ligand-macromolecule interactions taking place for the agonist/antagonist behaviour comparing with lipid A and lipid IVa.

In particular, we focused our work in simplified LPS analogues, among them (Figure **3.2**): **P01** and **P03**[12] are small molecules with two fatty acid chains containing an ammonium group; the tetraacylated synthetic compound **Eritoran**[32] reached phase III in clinical trials, but failed to demonstrate sufficient efficacy in late stage human trials, although it has recently shown promising activity in preventing influenza-induced acute lung injury, through a TLR4 antagonism mechanism; agonist compounds **ONO-4007**[33] (a LPS-like compound with TLR4 agonist activity through the induction of TNF-α production in tumor cells, with no further clinical development due to the limited water solubility), antagonist compound **D1**[33] (a synthetic lipid A analogue, which inhibits endotoxin activation of TLR4 by precluding interaction of the endotoxin with both CD14 and TLR4/MD-2) and the antagonists compounds **A1** and **A2**.

Non LPS molecules have also been developed among them (Figure **3.3**): **euodenine A,**[5] a natural product isolated from the leaves of *Euodia asteridula*, with no LPS structure; its synthetic analogues have exhibited potent and selective agonism

towards the TLR4, **Paclitaxel**,[34] and **pyrimido[5,4-b]indoles**[4] have shown to stimulate TLR4 and could potentially be used as adjuvants or immune modulators.

For all these TLR4 modulators there are not binding mode proposed. It is clear that although these molecules have a different chemical structure, they must share a common pattern of interactions with TLR4. We have undertaken a computational study of some representative compounds to unveil some of these patterns of interactions. Some of them have engender an antagonistic response and other agonist response in the TLR4 complex. These molecules were docked comfortably into the hydrophobic pocket MD-2, thus creating a stable complex between the MD-2 and TLR4 in molecular dynamics. There is something common among them that do have this behaviour and we are studying by molecular modelling technique.

It is interesting to note that a number of small molecules have been reported as TLR4 modulators. Although it is plausible these molecules bind MD-2, we have to consider that they might be targeting other sites/pockets (due to its reduced volume and size in comparison with the "native" ligand LPS). Moved by this idea, we undertook a computational analysis of the different accessible pockets of the TLR4/MD-2 system to identify/discover new/alternative binding sites where these small molecules could be binding.

For this, a combined approach involving binding site prediction and docking studies was implemented to determine primary and secondary drug binding sites for known TLR4 ligands, and evaluate the results to deduce the underlying molecular mechanism that reflects ligand pharmacodynamics. Our objectives here were, the identification of possible binding pockets in different conformations of the TLR4/MD-2 system, and the proposal of the binding modes for known TLR4 small molecule modulators. Different new primary drug binding sites and secondary sites were found in the TLR4/MD-2 systems, especially in protein-protein interfaces. Additionally, regarding the specific targeted docking studies, new binding modes for the TLR4 modulation through different ligands were proposed.

LPS-like Molecules
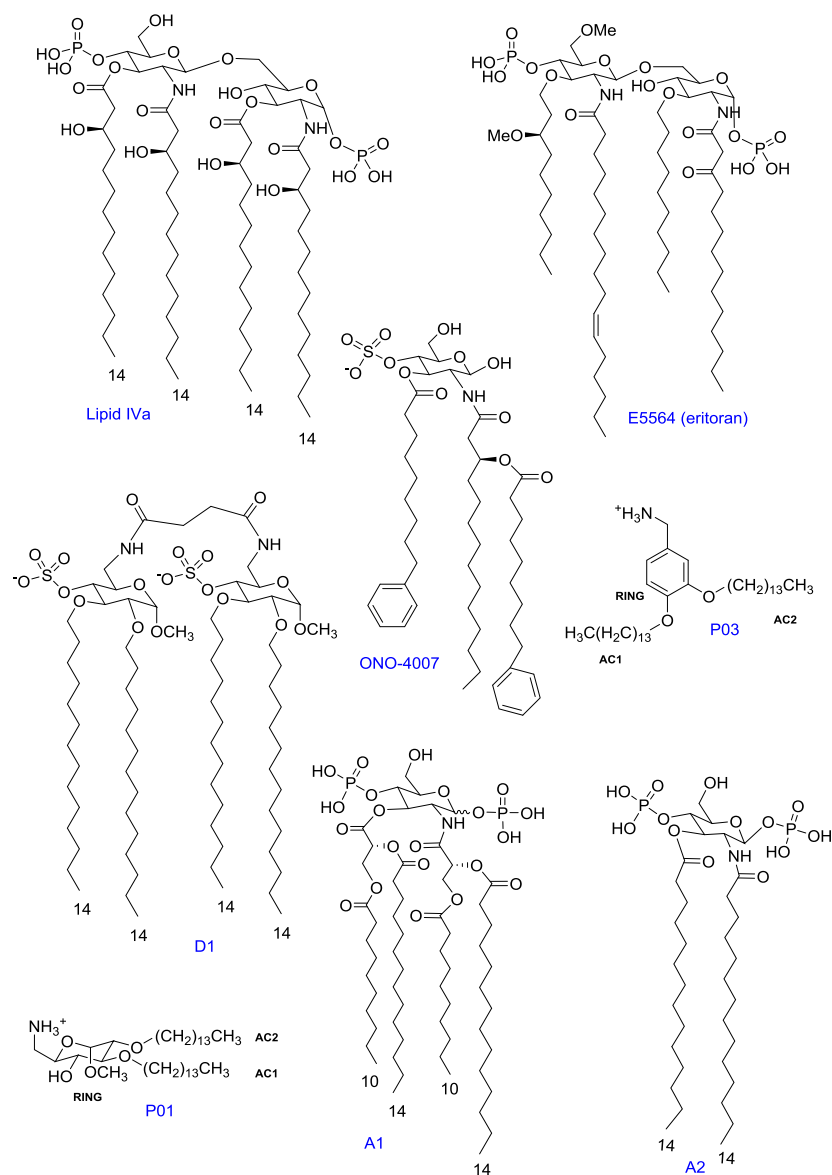
Figure 3.2. Chemical structures LPS-like molecules.

Non LPS-like Molecules
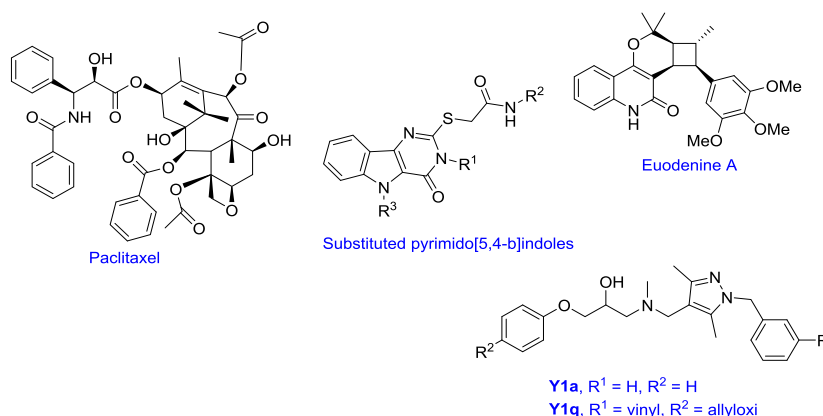
Figure 3.3. Chemical structures non LPS-like molecules

3.2 Results and Discussion

3.2.1 Interaction in TLR4/MD-2 Modulators with LPS-like Structure

To evaluate putative TLR4/MD-2 binding properties of these compounds, we undertook docking studies in the *h*TLR4/MD-2 system. The X-ray structure PDB-ID: 3FXI was used for the agonist conformation. Since the X-ray structure of the *h*TLR4/MD-2 complex is not available in the complex with an antagonist, we used a hybrid in house model by superimposing MD-2 structure from PDB-ID: 2E59 in one of the MD-2 subunits of the TLR4/MD-2 complex (PDB-ID: 3FXI, chain C) (for more details, see Chapter 3). AutoDock4 and VINA programs were used to carry out the docking calculations leading to similar predicted binding poses in all cases, without main differencies in the observed ligand-TLR4/MD-2 interactions.

As in the X-ray crystallographic structures of LPS and **lipid IVa** in both conformation, agonist and antagonist, the predicted binding poses for all the LPS-like molecules showed a general tendency to bury their fatty acid (FA) chains deeply buried into the hydrophobic pocket of MD-2, establishing van der Waals and CH-π interactions with the side chains of the lipophilic residues of the MD-2 pocket, and polar interactions between the sugars and the polar residues of MD-2 rim. The

resulting docked complexes of TLR4/MD-2 with **P01**, **P03**, **eritoran**, **D1**, and **ONO-4007** were submitted to MD simulations for a deeper analysis of the ligand-receptor interactions.

3.2.1.1   LPS-like Antagonist Ligands

Studies into de binding pose of synthetic LPS-like compounds were performed for the following ligands: **P01**, **P03**, **D1**, **A1** and **A2**, together with endotoxin, **eritoran** and **lipid IVa**[35] as reference compounds.

For all the studied LPS-like compounds, docking calculations led to similar results in both MD-2 and TLR4/MD-2 systems. The FA chains were allocated inside the hydrophobic MD-2 pocket, and sugar moieties were found to be accommodated in the MD-2 rim, by establishing polar interactions. In particular, the studied LPS-like antagonists were found to establish many hydrophobic interactions in common to those observed for **lipid IVa** [35] inside MD-2: Ile32, Ile44, Ile46, Val48, Ile52, Leu61, Ile63, Tyr65, Leu71, Leu74, Phe76, Leu78, Ile80, Val82, Ile94, Phe104, Val113, Ile117, Phe119, Phe121, Ile124, Val135, Leu146, Phe147, Leu149, Phe151 and Ile153. Also common polar interactions were found in the docked poses for most of the LPS-like antagonists, in particular with Tyr65, Arg90, Glu92, Tyr102, Ser118, Ser120, Lys122, Tyr131, and Cys133. Additionally, **D1** and **A2** were found to interact with Arg264 from TLR4.

Regarding **P01** and **P03,**[12] our results showed that most of the docking models displayed a very rational binding mode with the FA chains toward the pocket (Figure **3.3**), but they were not able to provide an unique binding mode. This result suggests that the ligand may be fluctuating in the hydrophobic pocket while being bound to the MD-2, and an adaptation of the MD-2 hydrophobic pocket could take place, as shown by our calculations of the volume of the hydrophobic pocket.

For both **P01** and **P03**, two docking calculations were done respectively, first in the isolated MD-2 and second in the TLR4/MD-2 complex. Similar results were

obtained in terms of the general orientation of the molecule as lower energy binding modes tended to bury their FA chains inside the MD-2 pocket and the sugar moiety located at the rim of the MD-2 establishing polar interactions. From the all docking obtained, the poses with lowest binding energy and having the consistent orientation of the FA chains towards the pocket were chosen for MD simulation studies.
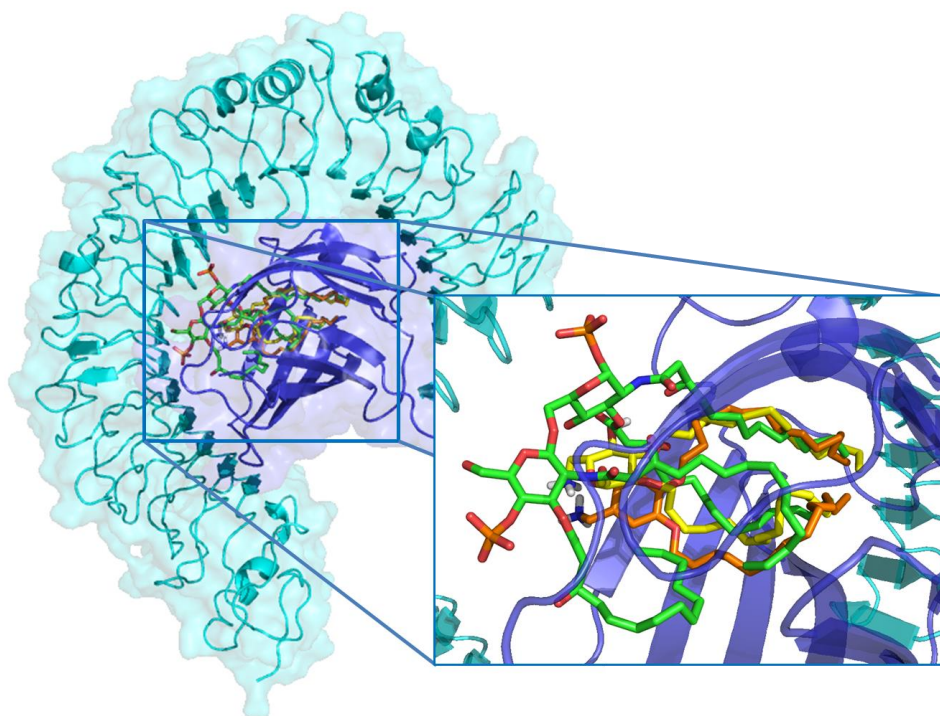
Figure 3.4. P01 (yellow) and **P03** (orange) docking binding pose, superimposed with **lipid IVa** (green) in TLR4(cyan)/MD-2(dark blue) system.

The MD simulations were performed from the resulting docked poses in the TLR4/MD-2 complex in antagonist conformation with both ligands **P01** and **P03**. In all cases, from the RMSD analysis of the MD simulations (Annex Figure **3.1**), it is observed that the receptor stays stable throughout the whole of the simulation during the run time of 50 ns. In contrast, the ligand had some significant movements that have been identified as different regions of stability, corresponding to different conformations of the ligand during the MD simulation. So, we have obtained minimized average structures and studied in detail for each conformation their significant interactions. As shown in the RMSD graph (Annex Figure **3.1**) different polar interactions stabilize

different conformations of the ligand, allow us to suggest different possible docking poses of these ligands in the TLR4/MD-2 complex, with the hypothesis that there is a period of stabilization, then, there will be polar bonds holding at that specific time. Therefore, we studied the hydrophobic and polar interactions occurring in the parts of the simulation where the ligand was stable identifying interactions present in all the simulations of **P01** and **P03**. Looking at hydrophobic contacts, we observed many common interactions to all the docking models and average structures from these simulations as Ile46, Ile52, Leu61, Ile63, Leu74, Phe76, Leu78, Ile80, Ile94, Phe104, Ile117, and Val135. In the case of polar interactions only few of them were found in the majority of the binding modes, the interaction with Tyr65 is present in all the solutions but at the end of the MD simulation of the TLR4/MD-2/**P01** simulation is weaker; or the interaction with Glu92, that is present again in all the proposed binding modes but is only missing at the beginning of the MD simulation of the TLR4/MD-2/**P01** and at the end of the simulation of the TLR4/MD-2/**P03**. Some other remarkable polar interactions stabilizing certain stable binding modes, i.e. interactions observed during some moments of the MD simulation are: the interaction with Arg90 which is formed along the MD simulation of the TLR4/MD-2/**P01** or the interaction Tyr102, which is found as a very strong interaction in all the binding modes from the docking models proposed by AutoDock, but which is lost along all the MD simulations; the interaction with Ser118 which is observed in some of the binding modes of **P01** but not in the ones from **P03**; opposite to the interaction with Ser120, which can be seen in some of the binding modes of **P03** but is more rarely seen in the **P01** ones; or Lys122 or Gly123 present in only few of the **P01** and **P03** complexes. For example, along the MD simulation of the MD-2 protein with **P01** and **P03**, according to the RMSD graph of the ligand there is a period of stabilization then there will be polar bonds holding at that specific time, for example with Tyr102 and in other period of time with Glu92. These results show a difference when comparing with other compounds, for example of **eritoran** which shows a significant stability in its proposed binding mode and during the MD simulation as it will be shown below, due to the big size of the ligand which completely fills the hydrophobic pocket.

In particular for **P01**, we also computed the RMSD of the different parts of the **P01** molecule, as the two acyl chains (AC-1 and AC-2) and the sugar and aromatic ring respectively (sugar scaffold) (Annex Figure **3.2**). In general, one of the acyl chains, AC-1, is more stable along the MD simulation but the other acyl chain, AC-2, and the sugar scaffold move more along the MD. These results mean that across the MD simulation the ligand migrates from its original docking pose to a more central position within the MD-2 pocket. The MD simulation implemented on the **P01** molecule but with the binding pose selected from the docking (Figure **3.5**) performed with the full TLR4/MD-2 complex showed no key areas of stability across the whole simulation. Comparing the observed polar interactions to the previous simulation, we found that Glu92 is still present during most of the simulation, but new polar interactions were observed as Arg90. Again to understand the main reason of the motion of the ligand we computed the RMSD of the different parts of the **P01** molecule, and the contacts over the different stable regions of the MD, which showed that in this case AC-1 and the sugar scaffold were fluctuating along the whole simulation but the other AC-2 was stable along the MD.
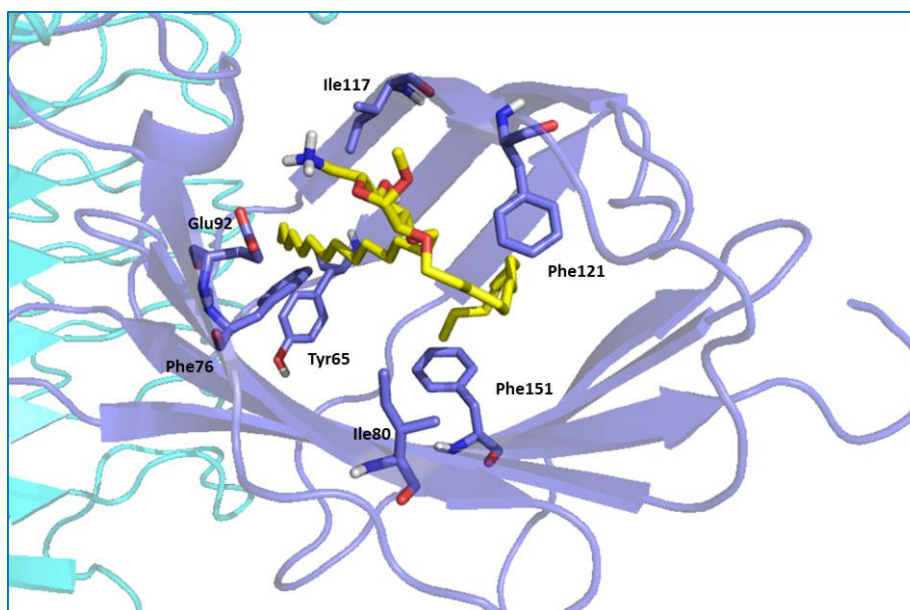
Figure 3.5. Detail of interactions of average structure of **P01** during the MD simulation with residues of MD-2 pocket.

In the case of the docking studies for **P03** using MD-2 and TLR4/MD-2 systems, we obtained several predicted binding poses. One of them was further studied by means of MD simulation. The MD simulation of the docked **P03**/MD-2 showed that the ligand was fairly stable with peaks and troughs varying from 9Å to 12Å. The polar interactions found were Glu92, Ser120 and Gly123. When observing the RMSD of the **P03** ligand we observe that during all the MD simulation again only one of the FA chains was stable. The final simulation that was performed for the small antagonist molecules was the **P03** from the docked full complex being run in and MD simulation with the full complex. The complex like all other simulations stayed stable throughout the whole simulation. The ligand on the other hand shows some varying regions of stability. Again in this simulation we observe that one of the FA chain remains stable along the MD simulation, but also the sugar scaffold. The other FA chain moves away from its initial position, reaching other conformation in which it is stable (Figure **3.6**).
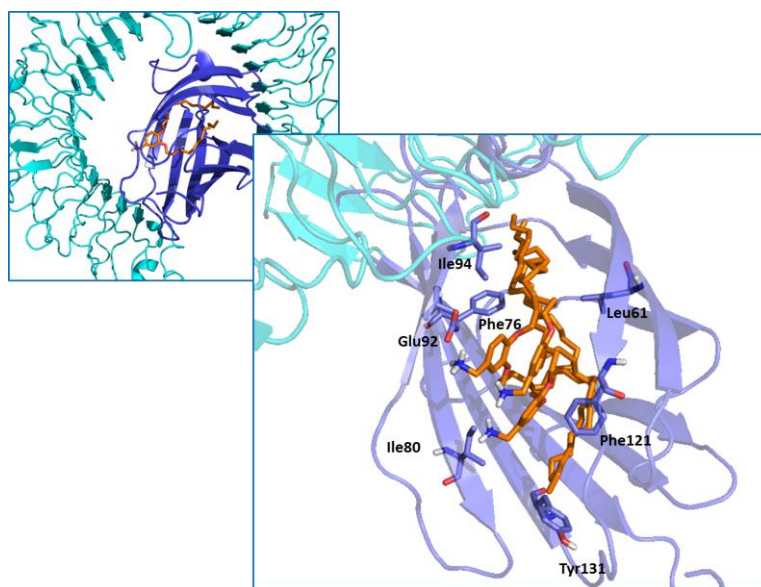
Figure 3.6. P03 (orange) binding pose on the top, and highlight three binding poses around the MD simulation with the interactions with MD-2 pocket.

So one thing to take into consideration from this result is that **P01** along with **P03** are small molecules, if converted into dimers there will not be nearly as much space available for the ligand to move around the MD-2 pocket throughout the MD

simulation. Further studies were performed to unravel these questions regarding ligand-receptor interactions with relevance for future design of novel binders (for more details see Chapter 4).[36]

In the case of **eritoran**, a crystallographic structure in complex with a hybrid Insore Hagfish/Human TLR4/MD-2 complex can be found at the PDB (PDB-ID: 2Z65). We used this initial X-ray crystallographic structure to build a *h*TLR4/MD-2 in antagonist conformation using the TLR4/MD-2 model from the antagonist MD-2 (PDB-ID: 2E59) and the agonist TLR4 (PDB-ID: 3FXI), and the **lipid IVa** after superimposing the MD-2 subunit found the hybrid model (PDB-ID: 2Z65) to the MD-2 of our model. This model was submitted to a 50 ns MD simulation under explicit water which was found to be stable during the simulation time (Annex Figure **3.3**).

The RMSD values of the ligand observed during the whole simulation allow us to confirm that this binding mode is stable, and that the ligand is placed filling the hydrophobic pocket. Further analysis of single residue interactions, we observed some stable polar interactions between **eritoran** and Lys122, Arg90 and Tyr102, also observed in the TLR4/MD-2/**lipid IVa** model, but not in many **P01** or **P03** binding modes. Only the interaction with Leu54 and Thr115 disappear along he MD simulation, but is possible to observe new interactions with Val24, Val48, Lys58, Tyr65, Ser118, Lys132, Leu146 and Val152 (Figure **3.7** and Annex Figure **3.4**).
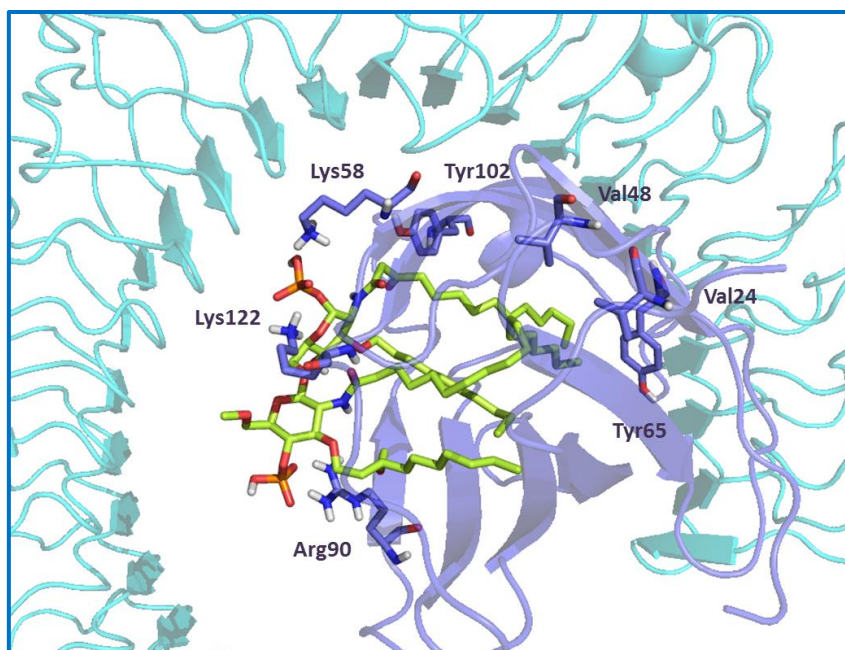
Figure 3.7. Principal interactions of **eritoran** (pistachio green color, average structure during the MD simulation) with TLR4/MD-2 system.

Compound **D1** from the point of view of the chemical structure is very different from other ligands. The succinic diamide linker with the two negatively charged groups from the sulfate mimics the distance between the phosphates of **lipid IVa**. So, the sulfate role could represent an alternative anchorage group to design new ligands. In the case of **D1**, the FA chains establishes hydrophobic interactions in the MD-2 pocket between the alkyl side chains from residues Ile44, Ile46, Leu61, Ile63, Tyr65, Leu71, Leu74, Phe76, Leu78, Ile80, Val93, Ile94, Tyr102, Val113, Ile117, Phe119, Ile124, Tyr131, Cys133, Val135, Phe147, and CH-π interactions with Phe104 and Phe151 side chains. Additionally, polar interactions are established with the residues at the MD-2 rim between one of the phosphate groups and Glu92, Arg96, Asp100 and Lys122 and some interactions with TLR4 between Asn339 and Arg264 side chains (Figure **3.8**).
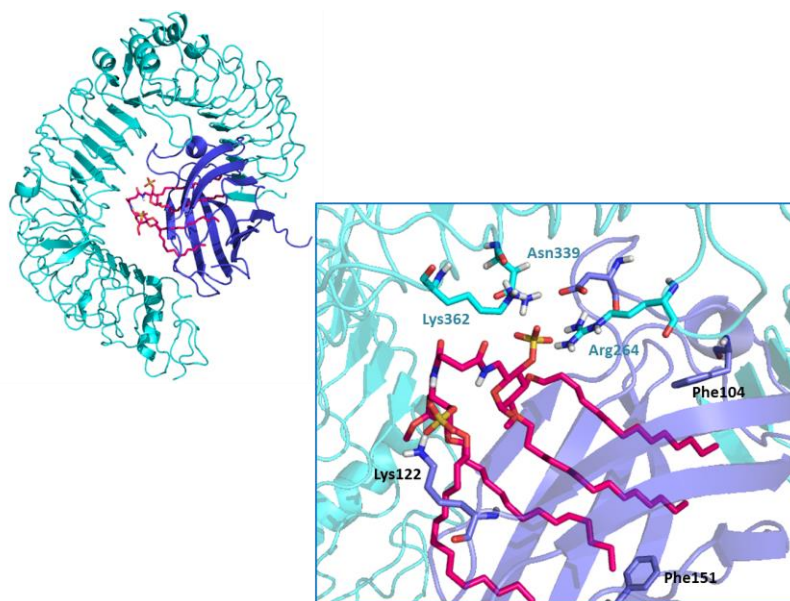
Figure 3.8. Principal interactions of **D1** (magent) with TLR4/MD-2 system.

Overall, from docking calculations and MD simulations, we can say that the predicted binding pose is analogue to the antagonist binding pose found for **lipid IVa** in the crystallographic structure (PDB-ID: 2E59), with the FA chains deeply buried into the lipophilic pocket of MD-2 and with polar interactions with the polar residues from the MD-2 rim. Stability of such predicted complex was tested by means of MD simulations. Along the MD simulation (50 ns) is possible to observe that the receptor is stable and the ligand migrates from its original docking pose to a more central position within the MD-2 pocket, where two FA chains (the two FA chains in the deep of the pocket) remain stable from the beginning to the end of the simulation. The other two fatty acid chains are moving from the beginning position and remain stable around MD since 20 ns. Due to the change in the ligand position, **D1** goes toward the innermost part of the MD-2 pocket. Along the MD simulation the interactions with TLR4 (with Arg264, Asn339, Lys341 and Lys362) are lost. It stabilizes in the period from 20 ns to the end (Figure **3.9**). The interactions were analyzed monitoring the distances between the residues from TLR4/MD-2 complex which interact with **D1** along the MD simulation, being possible to observe the loss of interactions with TLR4 (with Arg264, Asn339, Lys341 and Lys 362), also others interactions were lost, among them: Val82, Leu87,

130

Val93, Arg96, Asp100 but some interactions were established with Val24, Ile32, Val48, Ile52, Lys58, Gly59, Pro67, Thr115, Ser118, Ser120 and Leu149, due to a slight change in the initial position of the ligand.
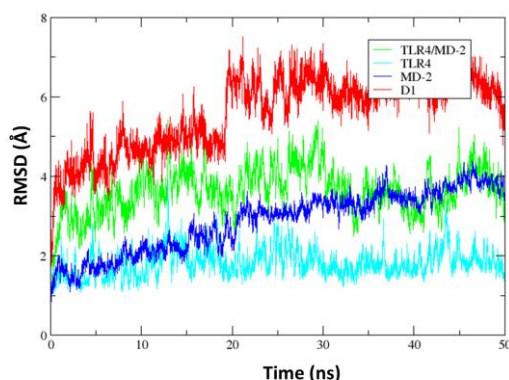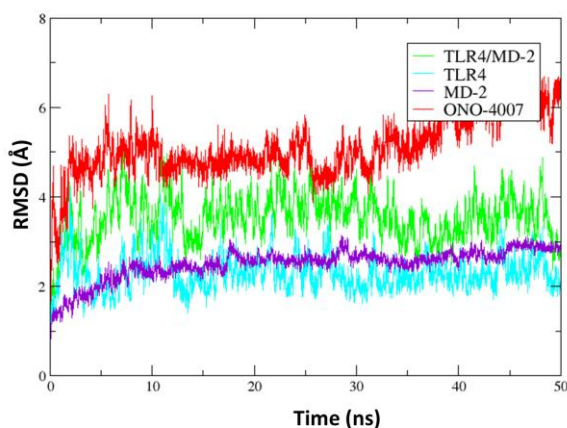

Figure 3.9. MD simulation of the complex TLR4/MD-2/**D1**(50 ns): RMSD (Å) is represented for the α-Carbons for TLR4/MD-2 (green), TLR4 (cyan), and MD-2 (dark blue), and the heavy atoms for the ligand **D1** (red).

For selected binding pose, free energy of binding was calculated by means of the MM-GBSA method computed with Primea, obtaining: -115 kcal mol^{-1}. This calculation of the free energy of binding by MM-GBSA approach can be appropriate to overcome the possible underestimation of binding energy for big hydrophobic ligands.

Comparing with **lipid IVa**, as representative TLR4 antagonist, in the multimeric complex of TLR4/MD-2/antagonist from the X-ray structure, some interesting interactions are identified (Figure **3.10**).
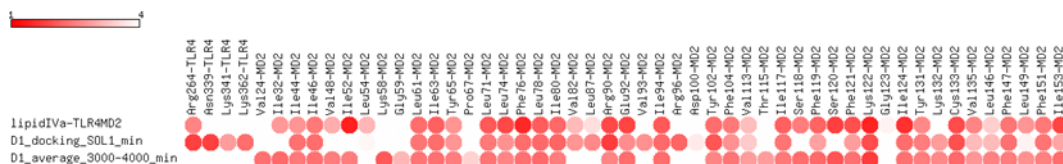
Figure 3.10. Comparison the contacts between **D1** (from the docking and average structures of the MD simulation) and **lipid IVa** with the residues from the TLR4/MD-2 antagonist system.

Compounds **A1** and **A2** are simplification to a single sugar of the **lipid A** chemical structure. **A1** presents four fatty acid chains, and **A2** with only two. To evaluate MD-2 binding properties of **A1** and **A2**, we performed docking studies in the hybrid model TLR4/MD-2 *in-house* system, since the X-ray structure of the TLR4/MD-2 complex is not available in the antagonist conformation. Most of the best docked solutions correspond to binding poses for these compounds with FA chains deeply confined inside the MD-2 hydrophobic pocket, in a way equivalent to **lipid IVa**.

In the case of the compound **A1**, the two branched FA chains establishes hydrophobic interactions in the MD-2 pocket between the alkyl side chains from residues Val24, Ile32, Ile44, Ile46, Val48, Ile52, Leu54, Leu61, Ile63, Leu71, Leu74, Leu78, Ile80, Val82, Ile94, Phe104, Phe119, Phe121, Ile117, Ile124, Tyr131, Cys133, Val135, Phe147, Leu149, Ile153, and CH-π interactions with Phe76, Tyr102 and Phe151 side chains. Additionally, polar interactions are established with the residues at the MD-2 rim, one phosphate group participate in hydrogen bond with Ser120 and the second phosphate group with Lys122; also establishes polar contact with Arg90 and Ser120, and also one of the phosphates from the ligand interact with Lys122 (Figure **3.11**).
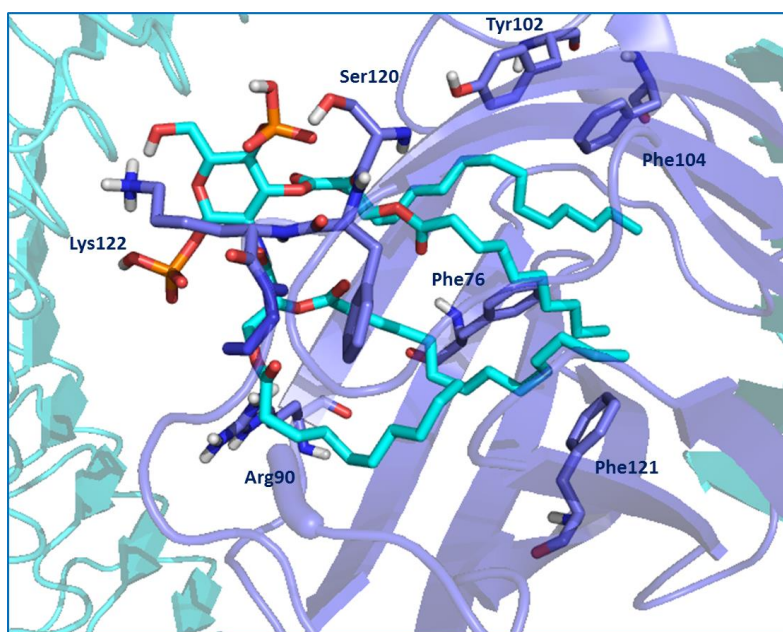
Figure 3.11. Interactions details of **A1** with TLR4/MD-2 system.

In the case of the compound **A2**, the three best binding poses are shown presenting both FA chains inside the pocket, with the polar phosphate groups and the sugar placed at the rim of MD-2, similar than lipid **IVa** (Annex Figure **3.5**).

In all the cases, the two FA chains established hydrophobic contacts with some residues and CH-π interactions with Phe151 in a similar way to **lipid IVa**. In one case, polar interaction was also identified in these binding poses, even with Arg264, presenting in TLR4. Also we observed in the most cases, that the phosphates groups participate in hydrogen bonds with Ser118 and Lys122 from the hydrophilic rim, adopting three preferred binding poses, all of them similar to the **lipid IVa**. But the interactions that all the solutions share each other were with Ile32, Ile46, Val48, Ile52, Leu61, Ile63, Phe119, Phe121, Phe151 and Ile153. Comparing with **lipid IVa** all of them establish interactions with Phe119, Ser120, Phe121, Phe151 and Ile153 (Figure **3.12**).
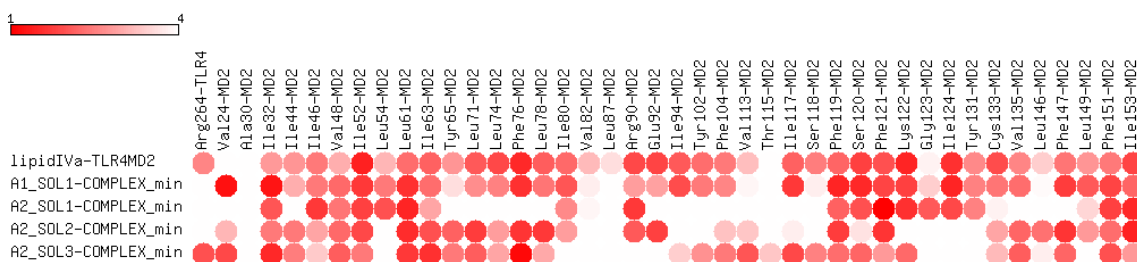
Figure 3.12. Comparison the contacts between **A1** and **A2** (from the docking structures) and **lipid IVa** with the residues from the TLR4/MD-2 antagonist system.

Here is showed a perspective of the TLR4 antagonist conformation with **lipid IVa** and the superimposition with different antagonist LPS-like ligands (Figure **3.13**):
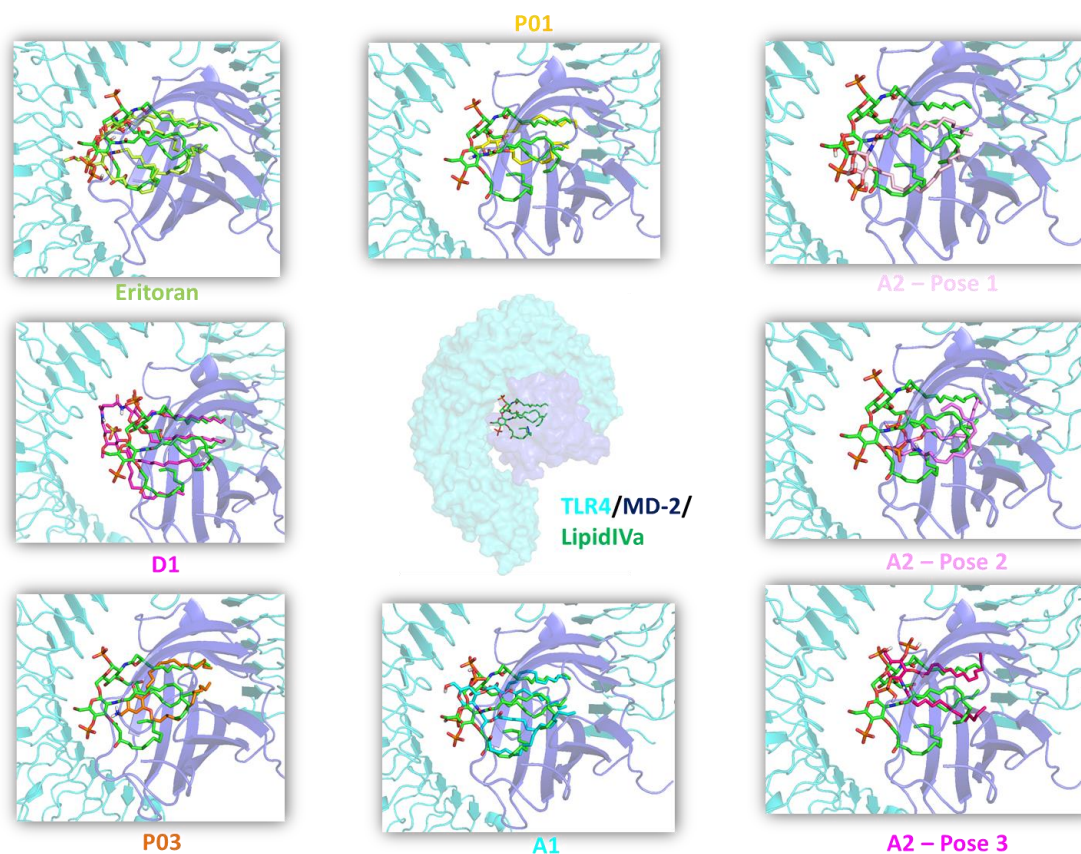
Figure 3.13. TLR4 antagonist conformation with **lipid IVa** (green) superimposition with: **eritoran**, **D1**, **P01**, **P03**, **A1**, **A2**.

3.2.1.2 LPS-like Agonist Ligands

Regarding **ONO-4007**, which has an agonist behavior, AutoDock and AutoDock4 Vina docked binding poses were equivalent to that observed for **lipid A** in the TLR4/MD-2 complex (PDB-ID: 3FXI), and which can be considered as the agonist-like binding mode. We selected the best binding pose to perform 50 ns of MD simulations (Figure **3.14**). At the beginning of the MD simulation, the ligand showed some drastic movements. Visibly it can be seen that the ligand almost instantly began to move away from the position that it originally sat in when the docking was performed, because the chains move towards the deep of the pocket but the receptor was very stable around the MD simulation (Figure **3.14**). We also identified an alternative binding pose with

one phenylalkyl chain occupying the channel and the other two chains inside the MD-2 pocket (Figure **3.16**).
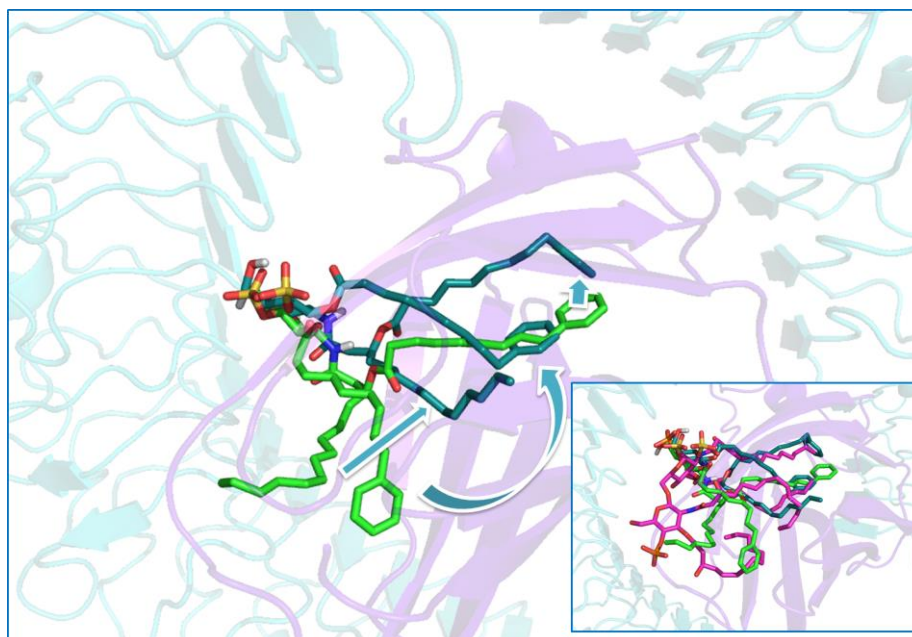
Figure 3.14. Change the position of ligand **ONO-4007** at the beginning of the MD simulation in green and during the MD simulation in dark green. At the bottom the superimposition with **lipid A** (pink).

In the docking pose, the two phenylalkyl chains buried inside the MD-2 pocket (with hydrophobic interactions with residues Phe119, Phe126, Ile52, Tyr65), and the alkyl chain placed in the channel, stabilizing the "on" conformation of Phe126. The FA chains establish hydrophobic interactions with Ile44, Ile46, Ile52, Leu54, Leu61, Ile63, Leu71, Phe76, Leu78, Ile80, Val82, Ile117, Lys122, Gly123, Ile124, Cys133, Val135, Leu149 and Phe151, CH-π interactions with Tyr65, Phe119, Phe121, Tyr131 and also π-π interactions with Phe147. We also identified polar interactions between **ONO-4007** and Arg90, Ser118 and Ser120 from MD-2 and Arg264, Asn339 and Lys362 from TLR4. After the re-colocation of the ligand new interactions were established, and new interactions, the interactions with Ile44, Ile46, Ile52, Leu61, Ile63, Tyr65, Leu71, Phe76, Leu78, Arg90, Glu92, Ile117 Phe119, Ser120, Cys133, Val135, Phe147, Leu149 and Phe151 were maintained. New hydrophobic interactions with Val24, Ile32, Val48, Lys58, Leu74, Ile94, Tyr102, Phe104, Val113, Asn114, Thr115, Ile117, and Leu146 were

observed. The contact with Arg264, Asn339 and Lys362 from TLR4, Leu54, Ile80, Val82, Phe121, Lys122, Gly123, Ile124, Phe126, Tyr131, and Lys132 MD-2 were lost.
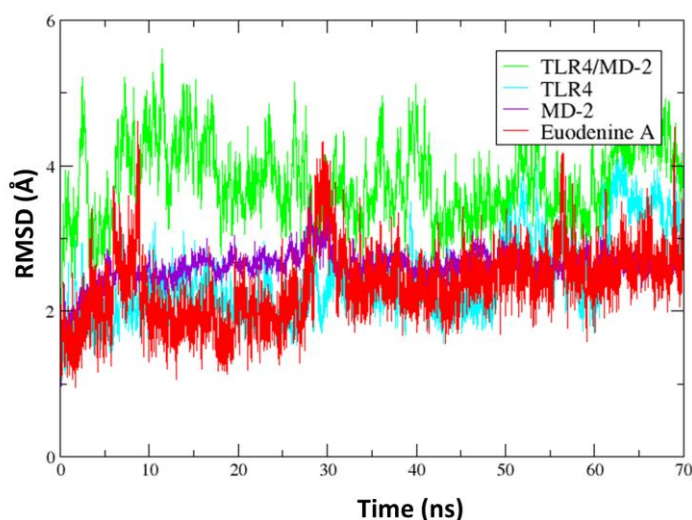

Figure 3.15. MD simulation of the complex TLR4/MD-2/**ONO-4007** (50 ns): RMSD (Å) is represented for the α-Carbons for TLR4/MD-2 (green), TLR4 (cyan), and MD-2 (purple), and the heavy atoms for the ligand **ONO-4007** (red).

For selected binding pose, free energy of binding was calculated by means of the MM-GBSA and MM-ISMSA methods, obtaining: -63 kcal mol^{-1} and -79 kcal mol^{-1} respectively.

Docked binding poses of the agonist ligands revealed common interactions with **lipid A**, some of them are Val24, Ile32, Ile44, Ile46, Val48, Ile52, Lys58, Leu61, Ile63, Tyr65, Phe76, Leu78, Arg90, Glu92, Ile94, Tyr102, Phe104, Val113, Thr115, Ile117, Ser118, Phe119, Ser120, Cys133, Val135, Phe147, Leu149 and Phe151 (Figure **3.16**).
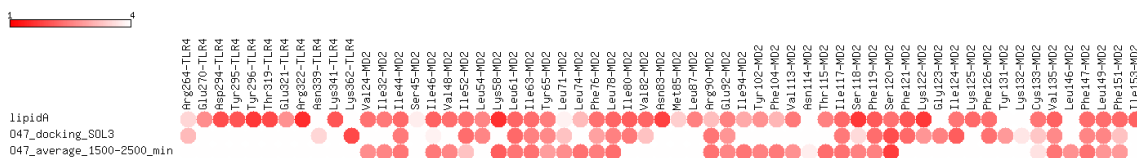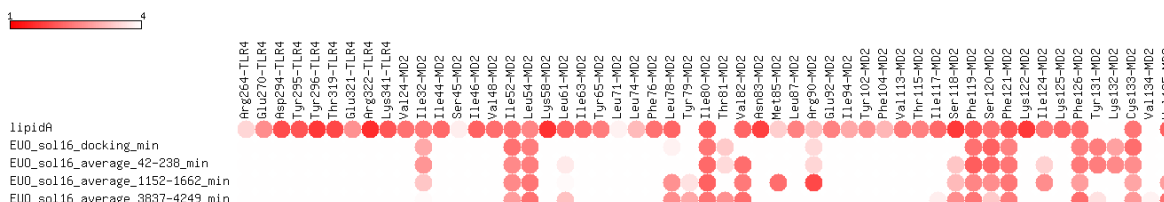

Figure 3.16. Comparison the contacts between **ONO-4007** (from the docking and average structures from MD) and **lipid A** with the residues from the TLR4/MD-2 agonist system.

Also in the case of the binding pose with one phenylalkyl chain occupying the channel and the other two chains inside the MD-2 pocket, was possible to observe from the RMSD along the MD simulation (30 ns) that almost instantly **ONO-4007** moved away from the original position, and keeping stable the complex for around 25 ns. After that time the phenylalkyl FA chain went outside of the pocket and the others FA chains occupied the channel. The complex stayed stable throughout the whole of the simulation and although it moves around a little no drastic changes were observed during the run time of 30 ns (Figure **3.18**).
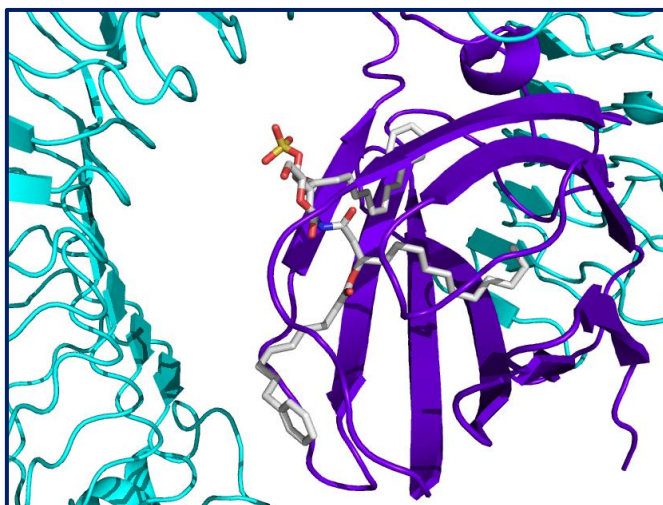
Figure 3.17. Docking pose of **ONO-4007** (grey) with TLR4/MD-2 system, but with one phenylalkyl chain occupying the channel.
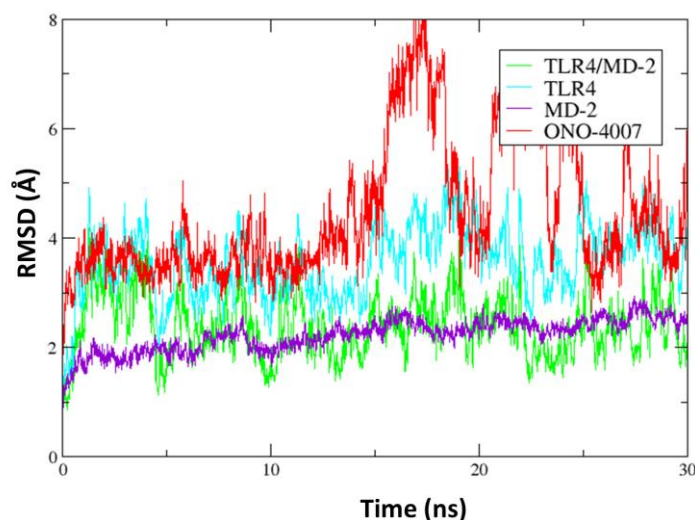
Figure 3.18. MD simulation of the complex TLR4/MD-2/**ONO-4007** (30 ns): RMSD (Å)
is represented for the α-Carbons for TLR4/MD-2 (green), TLR4 (cyan), and MD-2
(purple), and the heavy atoms for the ligand **ONO-4007** (red).

3.2.2 Interactions in TLR4/MD-2 Modulators With Non LPS-like Structure

Scarce small molecules have been reported to modulate TLR4/MD-2 system.
Because these ligands do not have a LPS-like chemical structure, they are very
interesting binders to be studied. Some molecules have already shown agonist activity
such as **euodenine A** and others with antagonist activity such as ethyl 4-oxo-4-
(oxazolidin-3-yl)-butenoate derivatives (OSL07), benzothiazole-based inhibitors, ethyl
phenyl-sulfamoylcyclohexenecarboxylate derivatives (TAK-242 or resatorvid),
paclitaxel and β-amino alcohol derivatives. To evaluate putative TLR4/MD-2 binding
properties of these compounds, we also undertook docking studies in the agonist and
antagonist *h*TLR4/MD-2 system.

3.2.2.1 Non LPS-like Antagonist Ligands

Non-LPS-like antagonist ligands displayed common hydrophobic and polar
interactions with **lipid IVa**: Ile52, Leu61, Phe76, Leu78, Ile80, Glu92, Ile94, Ile117,

Phe119, Ser120, Phe121, Cys133, Val135, Phe151 and Ile153.

Paclitaxel is a compound isolated from the bark of *Taxus brevifolia,* has been discovered as a cancer chemotherapeutic agent that bound to β-tubulin and prevented mitosis through microtubule overstabilization.[37] Also **paclitaxel** has been shown implication in the induction of apoptosis of cancer cells via the TLR4 innate immune pathway.

Docking of **paclitaxel** on the TLR4/MD-2 complex in antagonist conformation (using as complex the same model as in all antagonist models) was computed. Several low energy docking solutions were found in which the ligand was showing three different possible binding modes, one displaying an extended conformation of the ligand (typical in non-polar environments) and two similar to the two previously described paclitaxel conformations, T-Taxol conformation (tubuline-bound conformation) and polar conformation.[38-39]

From the lowest binding energy docking models, a total of 14 models (2 in extended conformation, 6 in polar conformation and 6 in T-Taxol conformation) were selected and further studied were performed by means of MD simulation. Initially, 2 ns of MD simulation was carried out. From these initial MD run, only two of the docking poses shown stability, hence, longer MD simulation of 50 ns was performed to provide reliable conformations for the TLR4/MD-2/**paclitaxel** complex (Annex Figure **3.6**).

In particular TLR4/MD-2/**paclitaxel** extended model shown mainly hydrophobic interactions due to the high amount of hydrophobic moieties of **paclitaxel**, as Val24, Ile46, Val48, Leu61, Ile63, Phe76, Leu78, Ile80, Ile92, Phe104, Phe119, Phe121, Val135, Phe147 among others, but also some polar interactions with key residues observed in the TLR4/MD-2/**lipid-IVa** model as Arg90, Glu92, Ser118 and Ser120 (Figure **3.19**).
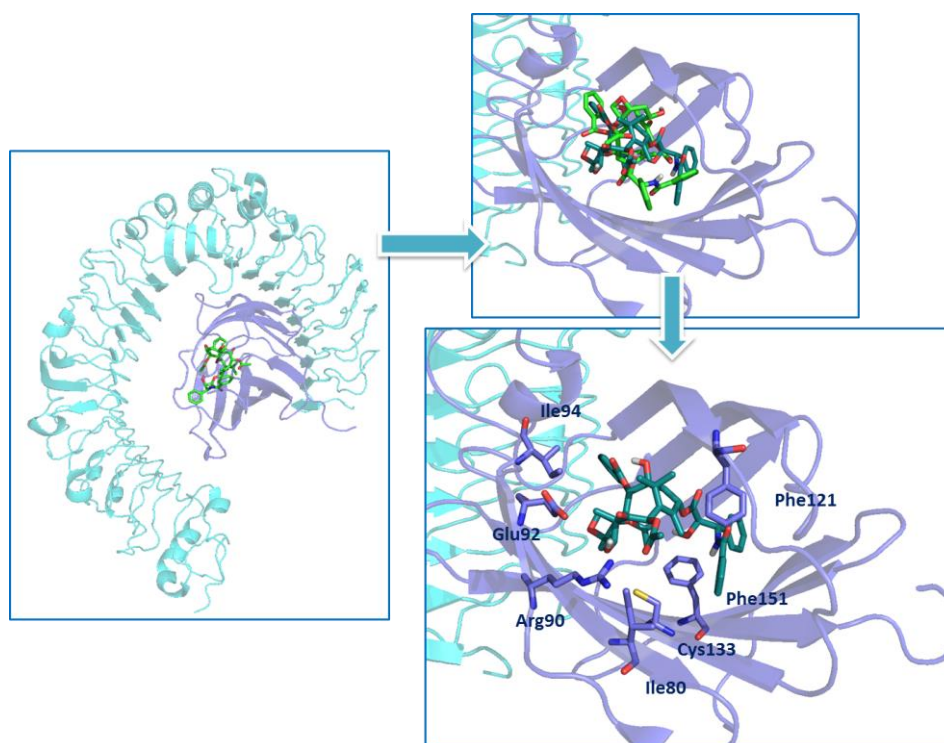
Figure 3.19. On the left, docking pose of **paclitaxel**; on the top right, during the MD simulation change the position establishing in MD-2 pocket, and on the bottom right, in dark green the new binding pose showing in sticks the interactions with MD-2 residues.

Activation of **paclitaxel** is quite specific as its analogue docetaxel does not have immunostimulatory activity (Figure **3.20**).[37, 40]

Paclitaxel   **Docetaxel**
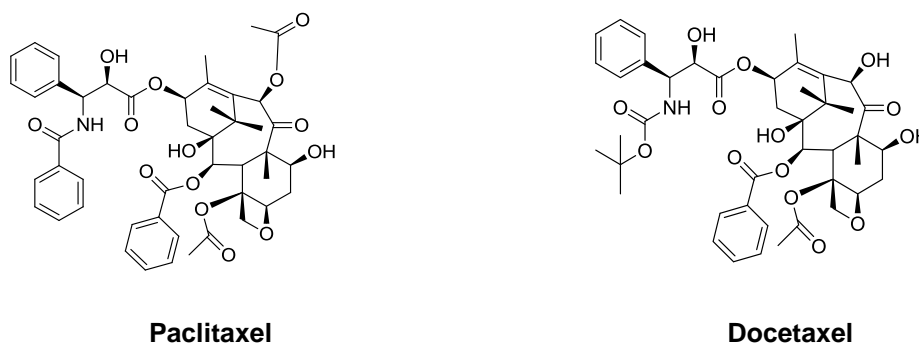
Figure 3.20. Chemical structures of **paclitaxel** and **docetaxel**.

3.2.2.2   Non LPS-like Agonist Ligands

Euodenine A is one of the few small molecules reported with TLR4/MD-2 agonist behavior. **Euodenine A**, a natural product, was identified as an agonist activity of the human TLR4 receptor, this ligand does not have a LPS-like chemical structure, and it is a very interesting binder to be studied. Other small molecules have been reported to modulate TLR4/MD-2 system but with antagonist behavior. To provide a model for understanding the mechanism, docking calculations with **euodenine A** were undertaken into agonist (PDB-ID: 3FXI) and antagonist (PDB-ID: 2E59) *h*TLR4/MD-2 proteins. We used the hybrid *in-house* model. Docking studies led to binding poses into the region inside the MD-2 pocket defined by Phe119, Phe121 and Phe151.
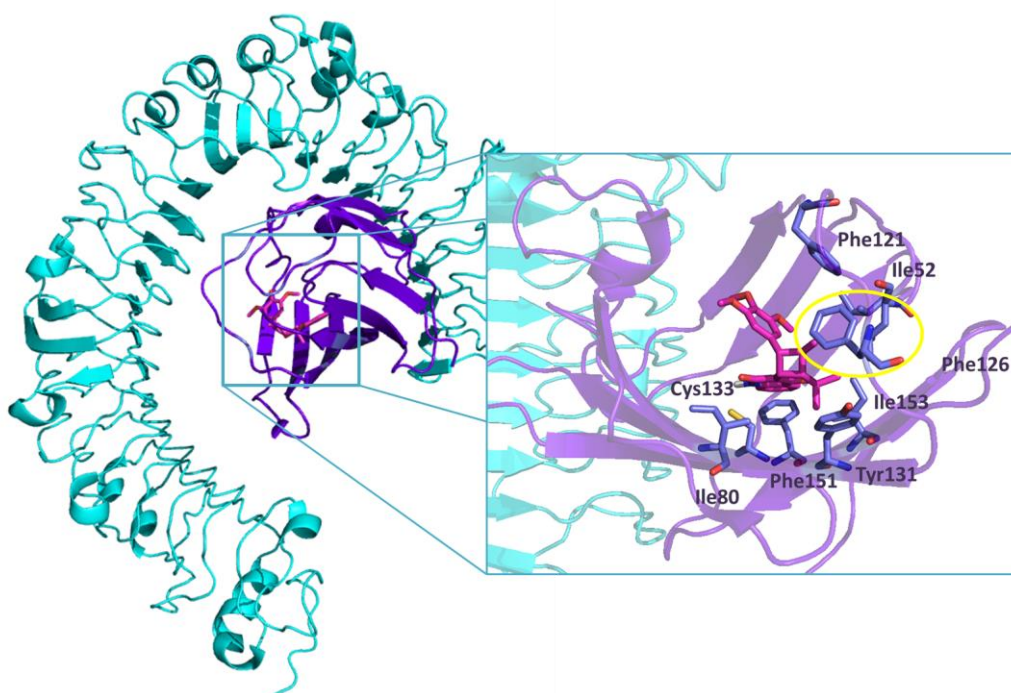
Figure 3.21. On the left, binding pose of **euodenine A** in TLR4/MD-2 system, and highlight interactions at atomic detail for an average structure during the MD simulation and highlight the Phe126 which is the agonist/antagonist conformation switch.

From the analysis of the docking poses hydrophobic interactions with Ile32, Ile52, Leu54, Leu78, Ile80, Thr81, Phe119, Cys133 and Ile153 were detected, CH-π interactions were observed with Phe126, Tyr131 and Phe151, also π-π interactions

with Phe121. Polar interactions with Lys132 and Ser120 were also established (Figure **3.21**).

Analyzing the MD simulation from the monomer around 70 ns, it can be observed that the receptor was stable throughout the whole of the simulation and although it moved a little, no drastic changes during the run time of 70 ns were registered (Figure **3.22**).

Figure 3.22. MD simulation of the complex TLR4/MD-2/**euodenine-A** (70 ns): RMSD (Å) is represented for the α-Carbons for TLR4/MD-2 (green), TLR4 (cyan), and MD-2 (purple), and the heavy atoms for the ligand **eudenine A** (red).

For selected binding pose, free energy of binding was calculated by means of the MM-GBSA and MM-ISMSA methods, obtaining: -38 kcal mol^{-1} and -48 kcal mol^{-1} respectively.

The analysis of the water molecules around the MD-2 was performed and it was observed that there was always only one water molecule inside the MD-2 pocket; in the starting geometry of the MD simulation no water molecules were present into the pocket. Along the MD simulation a water molecule sneaked into the pocket establishing interaction with the NH group from **euodenine A** amide group. It is

curious because was not always the same molecule and in other cases was not possible to observe this. The water molecule was surrounded by Leu35, Tyr36, Cys67, Val68 and Val69.

Comparing with **lipid A**, was possible to observe some common interactions (Figure **3.23**).

Figure 3.23. Comparison the contacts between **euodenine A** (from the docking and average structures from MD simulation) and **lipid A** with the residues from the TLR4/MD-2 agonist system.

A *in-house* dimer model TLR4/MD-2/**euodenine A** was built (Figure **3.24**), starting from an average structure from the MD simulation from the monomer, and we were performed the MD simulation of the monomer. Also the dimer was submitted to 50 ns of MD simulation. The RMSD of the dimer shown high stability for the receptor along the MD, as well as for the two **euodenine A** (Figure **3.25**).
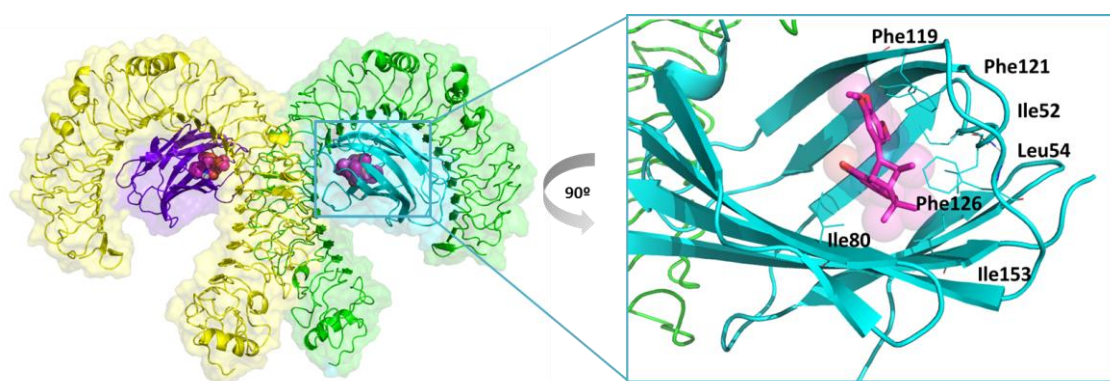
Figure 3.24. TLR4/MD-2 agonist *in-house* model, with two molecules of **euodenine A** inside the MD-2 pocket. Highlight the major LPS mimic interactions.
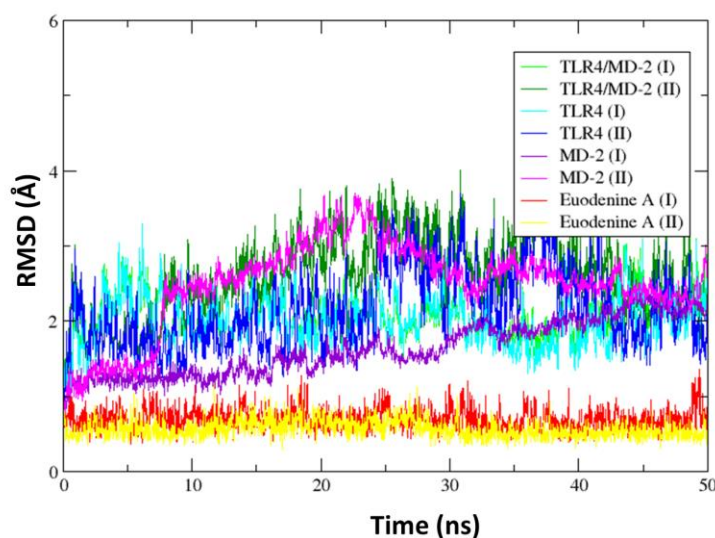
Figure 3.25. MD simulation of the full dimer complex TLR4/MD-2/**eudenine-A** (50 ns): RMSD (Å) is represented for the α-Carbons for TLR4/MD-2 (green), TLR4'/MD-2' (dark green), TLR4 (cyan), TLR4' (dark blue), and MD-2 (purple), MD-2' (magenta), and the heavy atoms for the ligand **eudenine A** (red) and **euodenine A**' (yellow).

In terms of the protein-protein interaction, the distance between TLR4 and TLR4' were monitored around dimer MD, analyzing the polar interactions and comparing them with the X-ray crystallographic structure (PDB-ID: 3FXI). The interactions analyzed were (TLR4-TLR4'): Asn365(TLR4)-Ser386(TLR4'), Ser386(TLR4)-Ser386(TLR4'), Asn433(TLR4)-Asn433(TLR4') and Gln507(TLR4)-Gln507(TLR4'). All of them remained stable around the MD except for the case of Gln507 located in the lower part of TLR4 (Annex Figure **3.7**).

Also the docking of the **euodenine A** in the TLR4/MD-2 antagonist conformation was performed. Two binding poses were chosen in which the ligand occupied a more central position in the MD-2 pocket. Analyzing the MD simulation (70 ns) it was observed that the ligand moved away from the original central pose, not getting a stable position throughout the MD simulation.

3.2.3 Small Molecules

In this study, a combined approach involving binding site prediction and docking studies was implemented to determine primary and secondary drug binding sites for known TLR4 ligands, and evaluate the results to deduce the underlying molecular mechanism that reflects ligand pharmacodynamics. Therefore, two main objectives were aimed for this work: the identification of possible binding pockets in different conformations of the TLR4/MD-2 system, and the proposal of the binding modes for known TLR4 small molecule modulators. Different new primary drug binding sites and secondary sites were found in the TLR4/MD-2 system, especially in protein-protein interfaces. Additionally, regarding the specifically targeted docking studies, new binding modes for the TLR4 modulation through different ligands are proposed.

Binding Site Prediction Studies

SiteMap. As a result of the structural differences between the receptor structures used in this study, certain binding sites were identified in every receptor conformation, while others were identified in only two or even one protein structure. Identified binding sites and their scoring values can be found in Annex Table **3.1**.

MD-2 pocket was reported as a single binding site in both TLR4/MD-2 systems (Figure **3.26a**). In the dimer conformation, MD-2 pocket was identified as part of two binding sites, named A and Q, which also display a solvent exposed surface corresponding to part of the TLR4/MD-2 and TLR4/MD-2 interfaces, respectively. Another binding site was predicted in the surface of MD-2, named M.

The TLR4/MD-2 interface was identified as several binding sites (Figure **3.26b**). Interestingly, the binding region where the polar interactions of LPS occur was identified as two different binding sites, one comprising the solvent exposed region of binding site A and another binding site, named as L. Additionally, another binding site was identified in all the receptor structures which was described in the TLR4/MD-2/TLR4*/MD-2* system as three different components: B, which comprised most of the surface of this dimerization interface for the binding site; X, a region between the

LRRs of TLR4 and MD-2, and a small, narrow pocket located in the TLR4 LRRs, named as W.

The TLR4/TLR4* dimerization interface was reported as five different binding sites (Figure **3.26c**). Two binding sites encompassed the vast majority of this interface: J and C comprised the lower half of the interface, while the upper half was composed by binding sites P, U and I, being the latter two also reported in both agonist and antagonist TLR4/MD-2 systems. Three different binding regions were identified in the TLR4/MD-2* dimerization interface (Figure **3.26d**): the solvent exposed part of Q, M and a highly solvent exposed binding site, named as R. Other binding sites were reported in the TLR4 LRRs (Figure **3.26e**). A binding region was also detected in the TLR4 central domain, comprising binding sites F, G, H and O. Near the TLR4 N-terminal domain, a series of binding sites were reported (D, K, N, S and Y). Due to their location, binding sites E (transmembrane end) and T (inner cavity) were not considered for further studies. Additionally, another binding site was only identified in one monomer of the TLR4/MD-2/TLR-4*/MD-2* system; named as V, the identification of this binding site could be an artefact secondary to the removal of glycosylations in the protein preparation process.
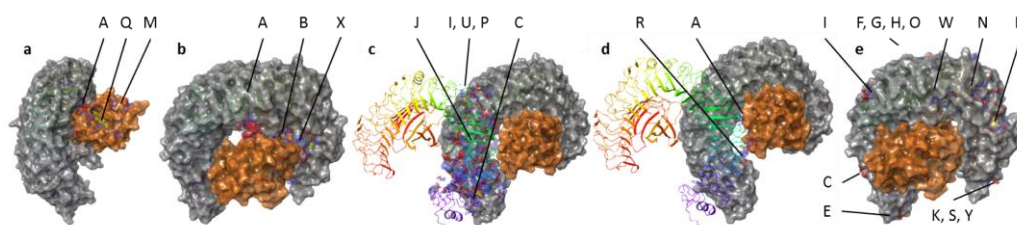
Figures 3.26a-26e. From left to right: binding sites at MD-2, TLR4/MD-2 interface, TLR4/TLR4* interface, TLR4*/MD-2 interface, and TLR LRRs. Acceptor and donor maps are shown in red and blue, respectively; hydrophobic maps are represented in yellow. MD-2 surface is colored in orange; TLR4 surface is colored in grey.

CASTp. Probe radius variation resulted in jobs with different characteristics). A probe radius of value of 1.7 Å was defined as it combined analysis simplicity with minimum loss of SiteMap binding site identification (Figures **3.27a** and **3.27b**). With

the exceptions of binding sites that were not recognized (L, S), partly recognized in each (V) or a determined structure (M and F), all SiteMap reported binding sites were totally recognized by CASTp. These results were consistent with SiteMap output, given that binding site F was only reported in the antagonist conformation, binding sites J, P, U and R are located in homodimerization interfaces, and M is absent in the antagonist conformation.
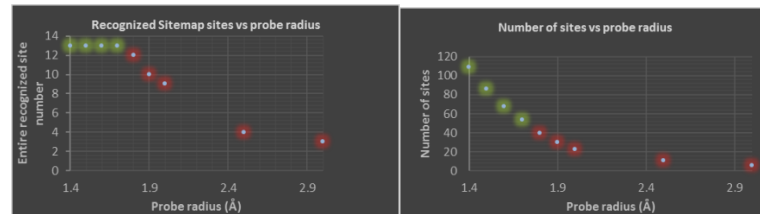
Figure 3.27a. Graphs showing the effect of probe radius variation in the number of recognized SiteMap binding sites (left) and the total number of identified sites (right). CASTp jobs wherein there is a loss of identification of SiteMap binding sites are depicted in red; adequate CASTp jobs are represented in green.
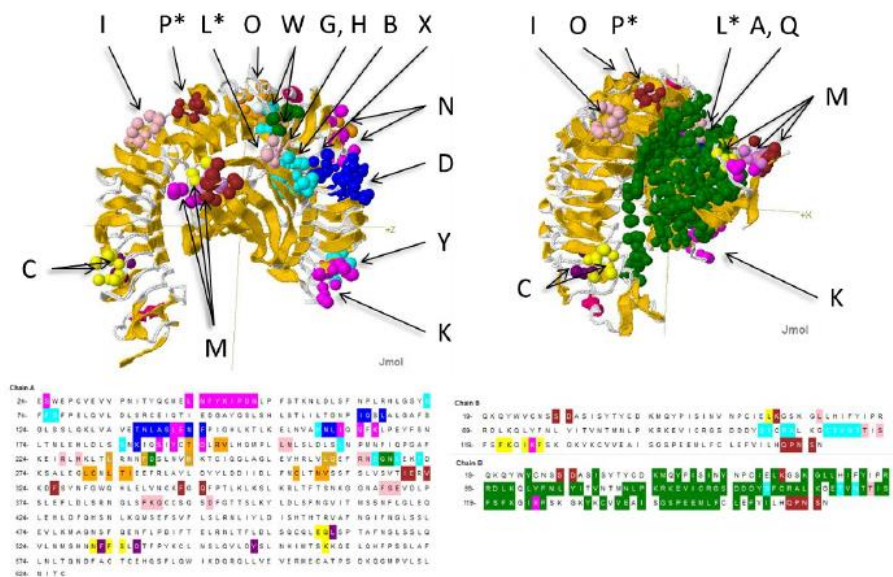
Figures 3.27b. Results of pocket identification by CASTp for the TLR4/MD-2 system in agonist conformation. Left: predicted pockets excluding MD-2 internal pocket; Right: MD-2 inner binding site (right) is reported as a huge pocket, even comprising parts of TLR-4. In this pocket, sites A and Q are encompassed; certainly all SiteMap binding sites are identified by CASTp. Residues forming pockets are represented by spheres and FASTA code letters of different color (green: A, Q, part of W; magenta: pocket K, parts of N and M; yellow: parts of C and M; cyan: Y,

147

B and part of W pocket; purple: part of C; dark blue: D and X; orange: O and N; brown: I and part of P). Regarding the protein structure, β-sheets and α-helices are colored in orange and red, respectively; loops are depicted in white. *: these pockets were only partially recognized by CASTp.

Binding Site Rating and Evaluation

Predicted binding sites were rated regarding their scoring values and their identification by SiteMap and CASTp in all the receptor structures. As expected, A and Q shown the highest values for SiteScore, supporting the evidence that MD-2 pocket is the primary binding site for the TLR4/MD-2 complex. Although binding site M shown low scoring values, its consideration as a binding site could be of potential usefulness due to its proximity to the external region of Q binding site. As a result of the high scoring values of binding sites that comprise the TLR4/MD-2 interface, A, L, B and W appear to be potential drug binding sites. Moreover, the presence of pharmacological activity in this binding region is supported by previous studies over certain TLR4 antagonists. Regarding the TLR4/TLR4* interface, C, J and P were reported as highly scored binding sites. Due to their location, these binding sites are potential candidates for the binding of PPI modulators.[59, 83] Regarding the TLR4/MD-2* interface, the notable scoring values of M and Q was not a particular surprise given that Phe126, located in this region, plays a crucial role in the receptor activation where several TLR4 modulators have been reported to bind.[58, 60]

Docking Studies

Receptor-Based. The average value for all minimum binding energies performed in docking calculations was -7.17 kcal mol^{-1}, with a standard deviation value of 1.45, which was indicative of the fact that there was a wide distribution of binding energy values between different docking calculations.

When averaging all the minimum binding energies of all dockings carried out in a determined grid box, a differential distribution of energies was observed between

different grid boxes for the same receptor conformation, suggesting that ligands displayed more favourable binding energies for specific regions in the TLR4/MD-2 complex. Particularly, AL (-7.72 kcal mol^{-1} for the average of the three receptor structures) and AQ (-7.68 kcal mol^{-1}) grid boxes, comprising MD-2 pocket, showed the most favourable binding energy results followed by DN (TLR4 LRRs, -7.63 kcal mol^{-1}), C (TLR4 LRRs -7.59 kcal mol^{-1}), B1W (TLR4/MD-2 interface, -7.49 kcal mol^{-1}) and B2 (TLR4/MD-2 interface, -7.32 kcal mol^{-1}). KS, J, FGHO, MQR, IUP and V grid boxes showed less favourable results (-7.29, -7.19, -7.11, -6.67, -6.46 and -5.90 kcal mol^{-1}, respectively).

Subsequently, when averaging the minimum binding energies for a determined grid box in a determined conformation, a differential distribution for these energies was observed between different receptor conformations for a specific grid box. These results suggest that certain receptor conformations allowed ligands to establish more favourable conformations, especially in the case of grid boxes which encompass parts of the TLR4*/TLR4* and TLR4*/MD-2 systems, suggesting the presence of ligand poses which display interactions in the TLR4/MD-2/TLR-4*/MD-2* systems with both TLR4/MD-2 dimers. For example, J grid box showed an energy value of -7.69 kcal mol^{-1} in the TLR4/MD-2/TLR-4*/MD-2* system, while less favourable values were found in the agonist and antagonist TLR4/MD-2 systems (-6.91 and -6.94 kcal mol^{-1}, respectively). Thus, J, C, IUP and MQR grid boxes were potentially identified as possible binding sites for PPI modulating agents. Interestingly, a considerable decrease in binding energy was found for B1W grid box (-7.63, -7.47, -7.35 kcal mol^{-1} for antagonist, agonist TLR4/MD-2 systems and TLR4/MD-2/TLR4*/MD-2* system, respectively), which may be reflected by the conformational change occurring in the TLR4/MD-2 interface between agonist and antagonist conformations. When considering the standard deviation related with the different minimum energy values calculated for each ligand in each grid box of each structure substantial differences was found in the calculated values, suggesting that there are regions of the TLR4 which are more likely to bind specific ligands than others. Finally, a correlation between certain energy values for a determined grid box of a determined conformation and

SiteScore values can be found, which confirms the potential of the predictive score obtained by SiteMap.

Results. Binding site mapping has been performed over all the surface of the monomer and dimer TLR4/MD-2 systems in three different structures. The binding sites reported by Sitemap have been successfully validated by CASTp. The scoring of these binding sites generally correlates with docking results, being MD-2 pocket the primary binding site for TLR4 small molecule modulators. Binding site prediction also defined the existence of other well scored sites, mainly comprising the TLR4/TLR4*, TLR4/MD-2* and TLR4/MD-2 interfaces. As some molecules have been determined to bind to these regions of the receptor, these scoring values are consistent with experimental evidence. Additionally, regarding docking results, binding site C, which is a new highly scored pocket in the C-terminal domain of TLR4, appears to bind molecules with low binding energies. Some of these TLR4 modulators could exert their pharmacological activity by binding to this site, either stabilizing PPI interactions or by other unknown mechanisms. Furthermore, binding site D, a pocket present in the N-terminal domain of TLR4, appears to be an important binding site for certain TLR4 modulators. The potential of our binding site prediction studies is established, as these two novel binding sites have never been reported in previous studies.
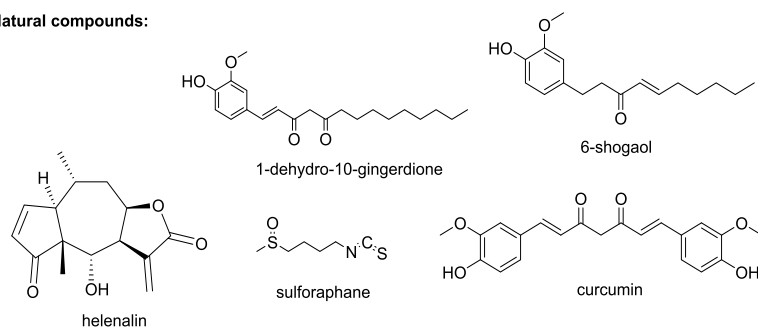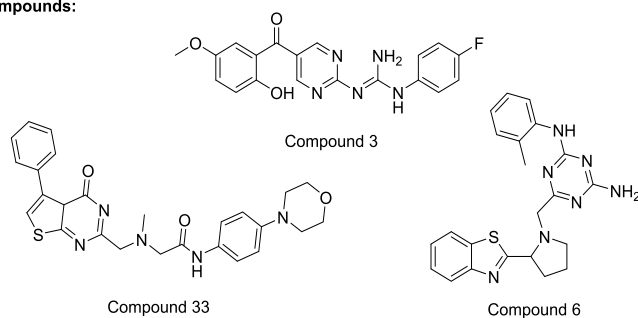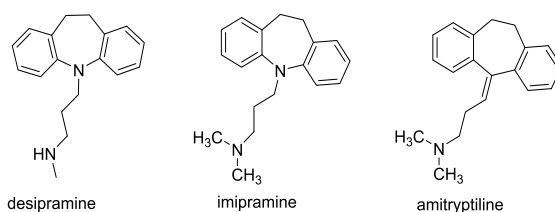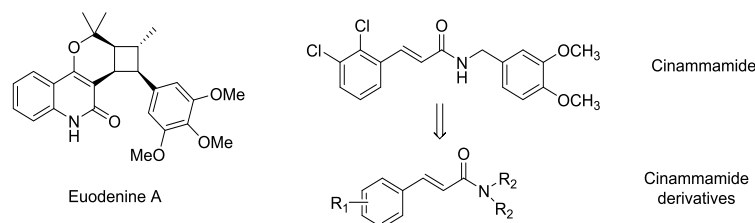
Figure 3.28. Some chemical structure, given names and families of TLR4 ligands.

Additionally, evaluation of the binding energies from different TLR4 modulators to the different predicted binding sites has permitted to propose their preferential binding to certain areas of the TLR4/MD-2 complex. Thus, the binding site of many natural and natural-like products with TLR4 agonist/antagonist activity has been proposed: bryostatin-1,[28] 1-Dehydro-10-gingerdione,[41] polyoxygenated cholesterol ester hydroperoxides,[29] 6-shogaol,[41] curcumin[19] and euodenins[5] appear to bind MD-2

151

selectively, while hemin likely binds both TLR4/TLR4* interface and MD-2 pocket (Figure **3.28**). Furthermore, Isoliquiritigenin[30] displayed very good energy values for regions comprising both TLR4/TLR4* and TLR4/MD-2 interfaces. However, the hypothesis that ILG forms covalent bonds with TLR4/MD-2 cannot be rejected. Additionally, three binding modes for GL have been proposed: this compound may be binding to either two regions of the TLR4/TLR4* interface or MD-2 pocket.

On the other hand, the predicted binding site of other TLR4 ligands has been in agreement with the experimental evidences shown in previous studies. Although all of them presented low binding energies for MD-2 pocket, several differences could be found in certain compound series. Although tricyclic compounds[31] and cinnamamides are thought to bind MD-2 pocket, favourable binding energies were found for binding site B, located in the TLR4/MD-2 interface (Figure **3.20**). Certain opioids[15] showed remarkable favourable binding energies for binding sites C and D. Additionally, the TLR4/TLR4* interface could be targeted by paclitaxel. Therefore, a more detailed study of these compounds is necessary in order to assess mechanisms that may unravel their binding modes. Other synthetic compounds, which are thought to bind the TLR4/MD-2 interface, displayed very good binding energies for site B. Nevertheless, they also showed favourable binding energies for MD-2 pocket. Thus, their binding mode should be reconsidered.

Other compounds (sulforaphane,[17] helenalin, acrolein, cinnamaldehyde[42-43]), showed low selectivity, as similar predicted binding energies for all the binding sites were observed. This could be mainly due to their small sized chemical structure that leads to predicted binding to a large variety of regions in the receptor surface. However, a mechanism through establishment of covalent bonds with the receptor cannot be excluded, which would require additional computational studies to assess this possibility. Last but not least, this study has permitted to consider that the feasibility of MD-2 binding assays in which 4,4'-dianilino-1,1'-binaphthyl-5,5'-disulfonic acid[41] participates may be compromised if TLR4/MD-2 systems were used, as it has been found that site C could be a secondary binding site for this compound. It can be concluded that this combined computational approach has permitted to identify new

receptor binding sites in the TLR4/MD-2 complex and predict plausible binding poses to them. Ultimately, this may help in understanding the principles that govern TLR4 activity and thus the development of new TLR4/MD-2 targeting modulators.

3.2.4 Overview of Agonist and Antagonist Interaction Patterns

Agonists ligands (**lipid A**, **euodenine A** and **ONO-4007**) always bind to the MD-2 region delimited for Ile32, Ile52, Leu54, Leu61, Leu78, Ile80, Val82, Arg90, Ser118, Phe119, Ser120, Phe121, Ile124, Phe126, Cys133 and Phe151 (Figure **3.29**).
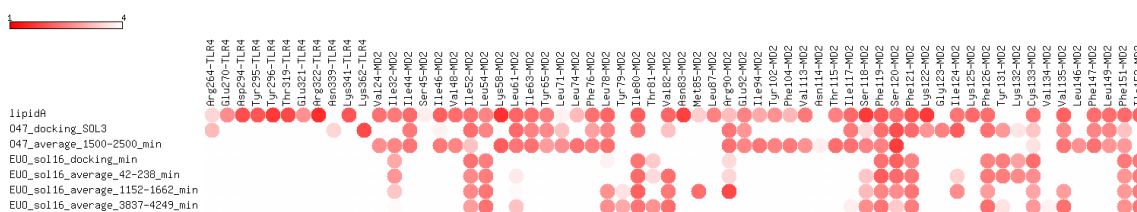
Figure 3.29. Comparison the contacts between the agonist ligands (from the docking and average structures from MD) and lipid A with the residues from the TLR4/MD-2 agonist system.

Antagonists ligands (**lipid IVa**, **eritoran**, **P01**, **P03**, **paclitaxel**, **D1**, **A1**, **A2**) share the following interactions with MD-2 pocket: Ile32, Ile44, Ile46, Val48, Ile52, Leu61, Ile63, Phe76, Leu78, Ile80, Arg90, Glu92, Ile94, Tyr102, Phe104, Ile117, Phe119, Ser120, Phe121, Cys133, Val135, Phe147, Leu149, Phe151 and Ile153 (Figure **3.30**).
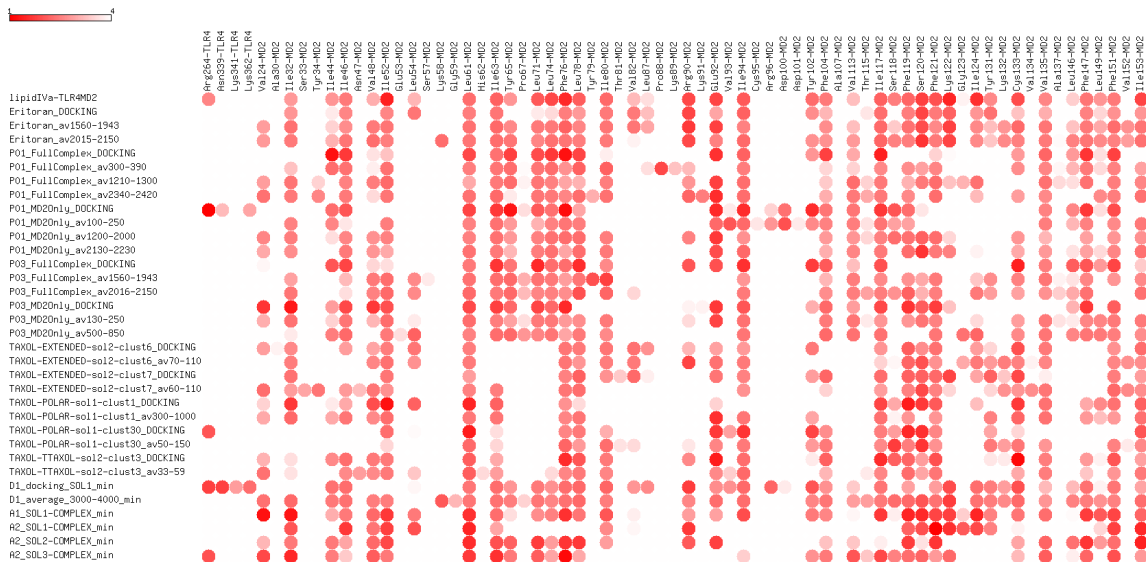
Figure 3.30. Comparison the contacts between the antagonists ligands (from the docking and average structures from MD) and **lipid IVa** with the residues from the TLR4/MD-2 antagonist system.

In the case of LPS-like ligands (**ONO-4007** and **lipid A**, among the agonists, and **lipid IVa**, **D1** and **eritoran**, among the antagonists), there is like a wall in the MD-2 pocket where one FA chain is driven towards this region cover by Ile63 and Tyr65 from one of the β-sheet, Val113, Thr115 and Ile117 from another β-sheet and some amino acids residues from the bottom of the pocket like Leu71, Leu74, Ile94, Tyr102 and Phe104. There are three residues at the bottom of the MD-2 pocket (Val135, Phe147 and Leu149), where one of the fatty acid chain from LPS-ligand always interacts. And also around Phe126, LPS-like ligands interact with Phe121 and Ile124 (Figure **3.31**).

154
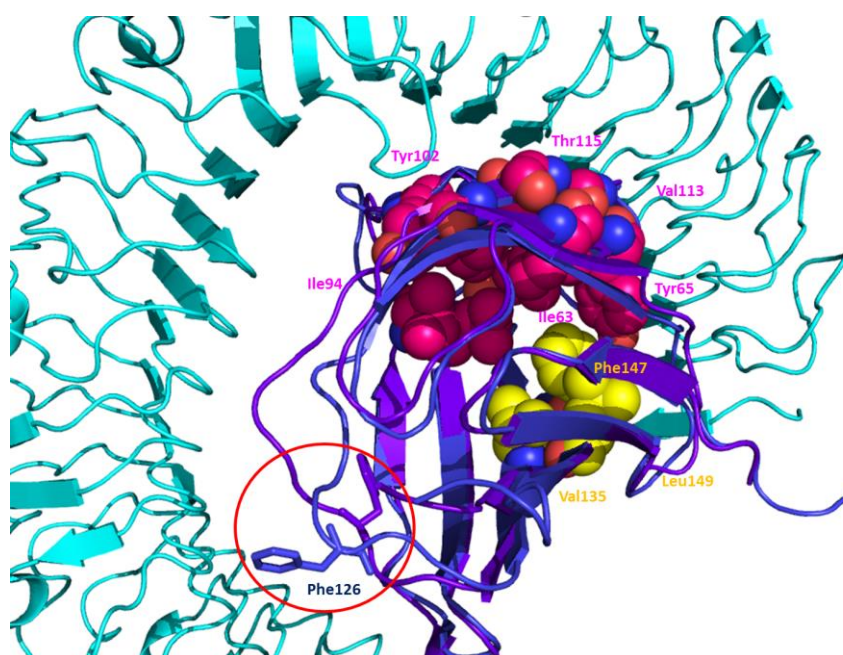
Figure 3.31. Hightlighted residues where LPS-like ligands interact. The pink spheres represent like a wall where one FA chain always goes towards this direction in the pocket. In yellow spheres the deep residues where another FA chain always interacts. Circle red highlights the principal difference between MD-2 agonist/antagonist conformations (Phe126).

According to our docking calculations, **Paclitaxel** (non-LPS ligand) does not bind into the pocket depth, this molecule doesn't establish interactions with the bottom of the MD-2 pocket where are the amino acids residues: Tyr65, Pro67, Leu71, Leu74, Val113 and Thr115.

Euodenine A (agonist) and **paclitaxel** (antagonist) share common interactions with MD-2 pocket, both interact with Ile52, Leu78, Ile80, Val82, Arg90, Ser118, Phe119, Ser120, Phe121, Cys133, Phe151 and Ile 153, both remain at the rim of MD-2 and they do not bind in the pocket background. The binding mode of these molecules is similar (Figure **3.32** and **3.33**). These interactions can be also observed in **lipid A** and **lipid IVa**.
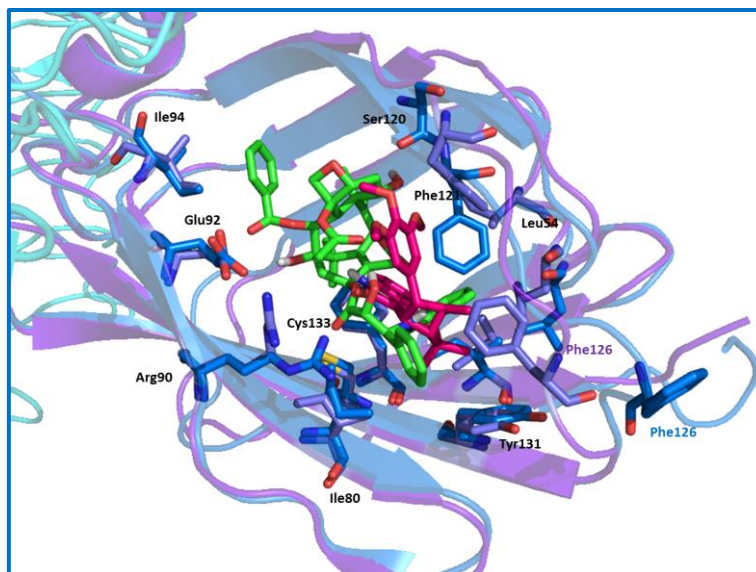
Figure 3.32. Superimposition of the binding pose of **euodenine A** (pink) and **paclitaxel** (green) in the TLR4/MD-2 complex. MD-2 agonist in purple and MD-2 antagonist in dark blue.
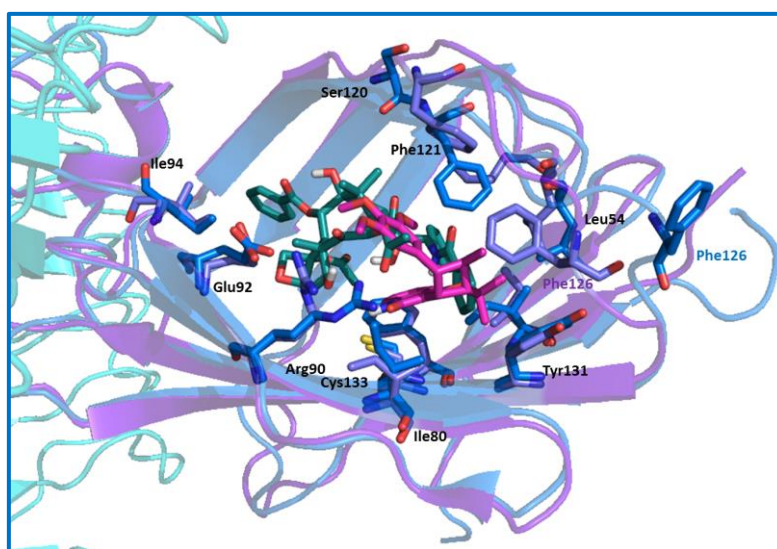
Figure 3.33. Superimposition of the average structure obtained from MD simulation of **euodenine A** (light pink) and **paclitaxel** (dark green). MD-2 agonist conformation is shown in purple and MD-2 antagonist in dark blue.

The plasticity of the TLR4/MD-2 complex was analyzed calculating the solvent accesible volume on hydrophobic pockets of MD-2 by using CASTp server which identifies and computes the molecular area and volume for cavities and pockets of a

given protein.[44-45]For this, the binding pose of the docking and the average structures from the MD simulations were used for all the ligands; the ligand was removed and the plasticity of the MD-2 pocket was analyzed. The data corresponding to the hydrophobic pockets of MD-2 proteins were selected. In the case of the smallest compounds and without LPS-like structure (**euodenine A** and **paclitaxel**) is possible to observe that the MD-2 pocket acquires a more closed conformation and the pocket gets a bit smaller in comparison with the LPS-like structure and more bigger ligands (**ONO-4007**, **D1**, **eritoran**, **P01**, **P03**, **A1** and **A2**). It is possible to confirm that MD-2 is very plastic, adapting a conformation in a ligand dependent manner (Annex Table **3.2**).

RMSD of the α-carbons of the loop from Leu78 to Lys89 and from Phe121 to Gln129 (loops implicated in the dimerization interface) from all the ligands was calculated. It was possible to observe that MD-2 is very flexible around the MD simulation (Figure **3.34** and **3.35**). We can conclude that MD-2 is very flexible but differences between agonist and antagonist conformation are very difficult to be detected.
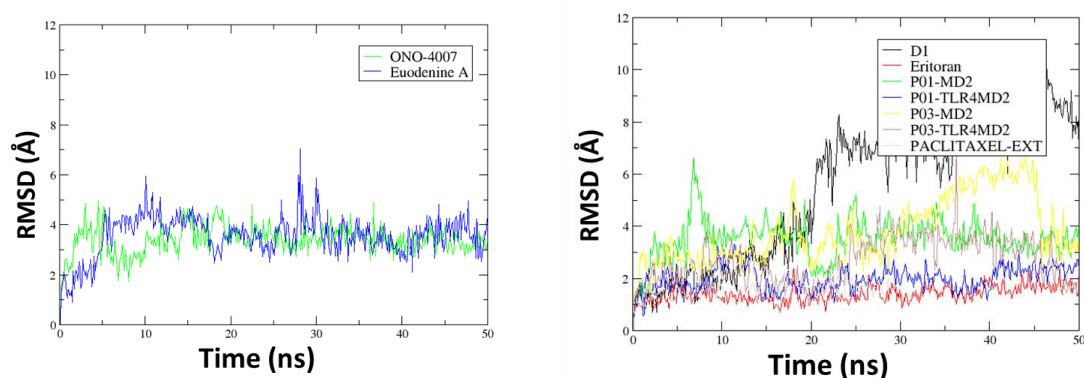
Figure 3.34. RMSD from Leu78 to Lys89 of the α-Carbons from MD-2 loops implicated in the dimerization interface from the agonist (left) and antagonist (right) ligands.
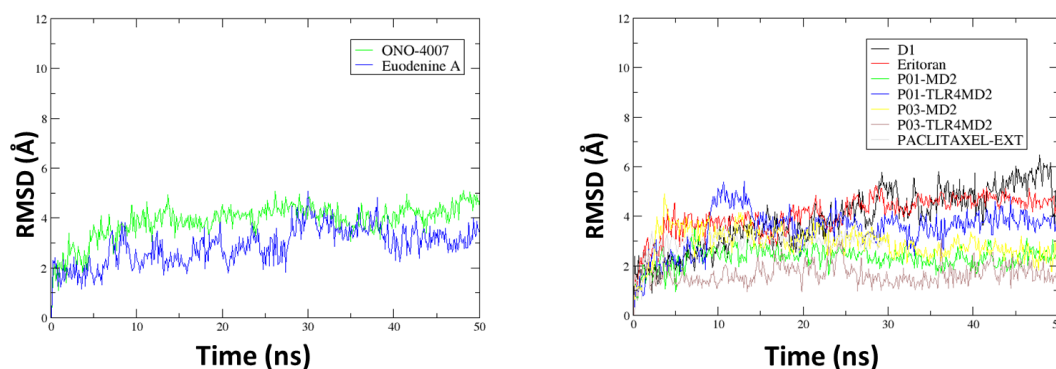
Figure 3.35. RMSD from Phe121 to Gln129 α-Carbons from MD-2 loops implicated in the dimerization interface from the agonist (left) and antagonist (right) ligands.

Regarding the plasticity of de MD-2 pocket implicated in the dimerization interface distances between Met85-Lys125, Leu87-Ile124 and Val82-Phe126 from MD-2 pocket were analyzed in agonist and antagonist ligands (Figure **3.36**). The distance in this region in the edge of the MD-2 pocket remains stable along the MD in all the cases, not being possible to see any significant differences between agonists and antagonists (Annex Figure **3.8**). Except in the case of **D1**, is could be possible to observe some drastic movements since 20 ns, which is the time where **D1** adopt a new position going towards the deep of the pocket. It is possible to think that it could be due to the size of **D1**, which has a bigger size than the other ligands.
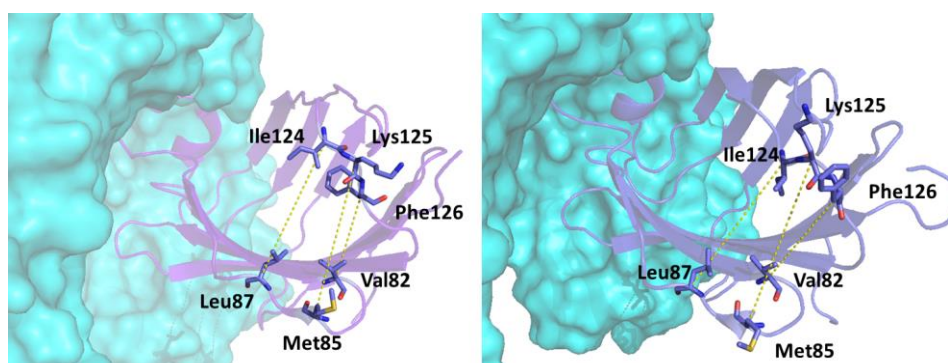
Figure 3.36. Distances between Met85-Lys125, Leu87-Ile124 and Val82-Phe126 from MD-2 pocket were analyzed in MD simulation of agonist and antagonist ligands. Agonist conformation on the left and antagonist conformation on the right.

The polar interactions between TLR4 and MD-2 were monitored around the MD from agonists and antagonist ligands comparing with the TLR4/MD-2 monomer structure (PDB-ID: 3FXI and 2E59). For the agonist conformation the distance interactions monitored were (TLR4/MD-2): Ser317-Asp101, Arg264-Asp101, Arg289-Ser98, Arg289-Asp99, Arg234-Asp100, Glu266-Ser103, Ser183-Arg106, Ser184-Arg106, Glu135-Thr112, Arg87-Thr112 and Glu42-Arg68. For the antagonist conformation the distance interactions monitored were (TLR4/MD-2): Asn339-Asp100, Ser317-Asp100, Arg234-Asp99, Arg264-Asp100, Glu266-Ser103, Asn265-Ser103, Glu135-Thr112 and Arg87-Thr112 (Annex Figure **3.9a** and **3.9b**). It was not possible to observe any changes in plasticity of MD-2 pocket. The interactions remained stable throughout the MD simulation (Annex Figure **3.9a** and **3.9b**).

The number of water molecules around Phe126 and Tyr131 were detected in the agonist and the antagonist conformation, being able to see that Phe126 in the antagonist conformation is exposure to the solvent having around 26 water molecules, while in the agonist conformation, has around 9 molecules of water in a 4 Å radius. In the case of Tyr131 is not possible to observe many changes between agonist and antagonist conformation (Annex Figure **3.10**).

The electrostatic potential from MD-2 pocket was calculated for all the ligands using APBS tool,[46] a molecular solvation based in the Poisson-Boltzmann approximation. It was not possible to observe significant differences between agonist and antagonist conformations (Annex Figure **3.11**).

The angle formed for Asp530(TLR4), Phe272(TLR4) and Pro50(MD-2) was calculated for all the ligands around the MD simulation (Figure **3.37**), it could be said that there were no abrupt changes throughout the MD simulations.
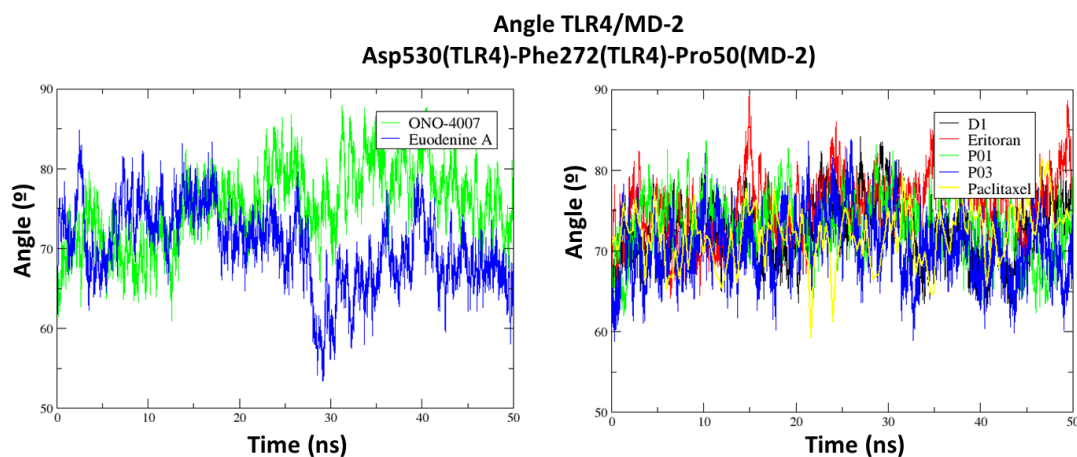
Figure 3.37. Angle measurement in the TLR4/MD-2 system between Asp530(TLR4)-Phe272(TLR4)-Pro50(MD-2). Agonists ligands on the left, antagonists ligands on the right.

3.3 Conclusions

Theoretical binding modes have been predicted for reported modulators of the TLR4/MD-2 system, with agonist and antagonist activity. In particular, we focused our work in simplified LPS analogues and non-LPS molecules have also been developed. For all these TLR4 modulators there are not binding mode proposed. It is clear that although these molecules have a different chemical structure, they must share a common pattern of interactions with TLR4. We have undertaken a computational study of some representative compounds to unveil some of these patterns of interactions. Some of them have engendered an antagonistic response and other agonist response in the TLR4 complex.

3.4 Materials and Methods

3.4.1 Macromolecule Preparation

In the case of the agonist conformation of the TLR4/MD-2 monomer, 3D coordinates from TLR4/MD-2 heterodimer were obtained from the Protein Data Bank,[47] (PDB-ID: 3FXI) and then chains A and C were extracted and considered as TLR4/MD-2 monomer in agonist conformation. In the case of the antagonist

conformation, since the full crystallographic structure of the TLR4/MD-2 complex is not available, a *in-house* model was used. This model was built using the human MD-2 protein in antagonist conformation (PDB-ID: 2E59) superimposed onto the MD-2 subunit of the agonist full complex (PDB-ID: 3FXI chain C) through PyMOL. Then, coordinates from the TLR4 chain of the 3FXI adjacent to the superimposed MD-2 (PDB-ID: 3FXI chain A) and the superimposed MD-2 in antagonist conformation were retained, forming the TLR4/MD-2 monomer in antagonist conformation. Finally, both agonist and antagonist structures were subjected to 10.000 cycles of steepest descent energy minimization under the OPLS_2005 force field via Maestro.[48]

3.4.2 Ligand Preparation

We built the full 3D structure of the agonists and antagonists ligands, with the help of ChemDraw® and Corina.[49] Optimization of the 3D geometries was performed using Macromodel from Maestro Suite, using steepest descent minimization under the MMFFs force field until a convergence in energy of 0.5 kcal/mol was reached.

3.4.3 Docking Studies

Starting geometries of the ligands were docked in the agonist or antagonist conformation of the TLR4/MD-2 complex respectively, by using AutoDock4[50] in the case of all of the ligands, Vina[51] for some molecules (specifically **ONO 4007**, **D1** and **euodenine A**) and also Glide[52-55] (in the case of **euodenine A** to compare the results obtained with AutoDock4 and Vina).

For the docking studies performed with AutoDock4, the TLR4/MD-2 system was prepared with the help of AutodockTools by assigning Kollman charges and setting the grid as follows. For the TLR4/MD-2 in agonist conformation, the grid point spacing was set at 0.375 Å, the center of the grid box were in the midst of Leu78, Val135, Cys133 and Ile80 residues from MD-2 pocket and number of grid points in x, y, z was 88, 96, 86. For the TLR4/MD-2 in antagonist conformation, the grid point spacing was set at

0.375 Å, the center of the grid box was close to Glu92 and number of grid points in x, y, z was 84, 96, 74.

Ligand was also prepared with the help of AutodockTools, retaining the charges obtained from the optimization process.

All docking calculations with AutoDock4 were performed using the Lamarckian genetic algorithm,[56] number of individuals in population 150, maximum number of energy evaluations 2500000-5000000, maximum number of generations 27000, number of top individuals to survive to next generation 1, rate of gene mutation 0.02, rate of crossover 0.8, window size 10, Alpha parameter of Cauchy distribution 0.0, Beta parameter Cauchy distribution 1.0, run 200. Disposition of the ligand was defined as random changes in the torsion angles, location and overall orientation of the molecule.

In the docking studies with Vina the following boxes were used: the TLR4/MD-2 system in agonist conformation, the grid point spacing was set at 1 Å, the center of the grid box were in the midst of Leu78, Val135, Cys133 and Ile80, and number of grid points in x, y, z was 33, 36, 33. For the antagonist conformation, the grid point spacing was set at 1 Å, the center of the grid box was close to Glu92 residue from MD-2 pocket, and number of grid points in x, y, z was 32, 36, 28. For all of them, the maximum number of binding modes to generate was 50, and the exhaustiveness of the global search (roughly proportional to time) was increased to 20.

And docking calculations with Glide were performed using the grid points in x, y, z 58, 47, and 45, and the lengths of the inner box for agonist protein conformation were: 10 Å, 10 Å, 10 Å (x, y, z) and for the outer box: 30 Å, 30 Å, 30 Å (x, y, z). Epik state penalties for docking were used, and the non-polar part of the ligand potential were soften by scaling the van der Waals radii of ligand atoms with small partial charges. To do so, the scaling was 0.8, and the partial charge cutoff was 0.15. Standard Precision (SP) was performed and 200 docking poses per ligand was set. The Dock flexibility method was used for SP docking allowing us to penalize non-planar conformation for amide bonds. A post-docking minimization was also performed, as well as constraints for the docking stage.

3.4.4   Molecular Dynamic Simulations

Molecular dynamic simulation was carried out for the some of the complexes obtained from the docking studies, i.e. **eritoran**, **P01**, **P03**, **paclitaxel**, **euodenine A**, **ONO-4007** and **D1**. The AMBER force field was used for both ligand and protein, with the ff10 parameter used for the TLR4/MD-2complex, and the gaff parameter were applied for the ligand with the help of antechamber. Then counter ions were added to the systems as to neutralize it (9 Na$^+$ atoms), and then they were solvated by using TIP3P waters in a cubic box with a 10 Å distance around the TLR4/MD-2 complex. Minimization was performed using Sander and MD simulations were run using the pmemd, which are distributed within the AMBER 12 package.[57] A 1 fs integration step and the shake algorithm on every hydrogen-containing bond.[58] The smooth particle mesh Ewald method[59] was used to represent the electrostatic attractions in the system while each simulation was under periodic boundary conditions, and the grid spacing was 1 Å. Initial annealing of the system occurred steadily and lightly from 100 K to 300 K over 25 ps. The temperature was then kept constant at 300 K during 50 ps with progressive energy minimizations and also a solute restraint. The solute restraints were gradually released, which was closely followed by a 20 ps heating period which went from 100 K to 300 K, once completed the restraints were removed. Each of the simulations lasted 50 ns. The systems then advanced in an isothermal-isobaric ensemble.

3.4.5   Average Structures

Average structures were extracted from molecular dynamics simulations of the ligands with ptraj of AmberTools 13. All the average structures were minimized with 5000 steps of steepest descent minimization with position restrain (force constant of 10 kcal mol^{-1} Å^{-2}) for all nonhydrogen atoms, plus 5000 steps of steepest descent minimization with no restrains.

3.4.6 Plasticity of the TLR4/MD-2 Complex

Solvent accesible volume was calculated on hydrophobic pockets of MD-2 by using CASTp server.[45] CASTp server identifies and computes the molecular area and volume for cavities and pockets of a given protein. The data corresponding to the hydrophobic pockets of MD-2 proteins were selected.

3.4.7 Binding Site Prediction Studies

Two different programs based in different algorithms were used to predict and score potential drug binding sites: SiteMap[60] and CASTp.[44]

SiteMap. Binding sites were predicted and scored for the mol2 files prepared receptor structures. SiteMap default parameters for the identification of site points were used. Requiring at least 15 site points per reported site, a maximum value of 30 sites per run was determined. Calculations were performed in a standard grid, using a restrictive definition of hydrophobicity. Maps were cropped at 4 Å from nearest site point. Results were visualized using Maestro. Site point groups were saved in xyz format.

CASTp. On a first approach, a probe radius input parameter was fixed. For this purpose, different jobs using the TLR4/MD-2 system and varying the probe radius were submitted to CASTp server. The correspondent calculations were retrieved, loaded and visualized using Pymol. SiteMap results were contrasted with CASTp results using the xyz exports. SiteMap sites that were either located next to the membrane domain or forming cavities inside TLR4 were not used for this evaluation. For each job, the number of pockets which were entirely recognized by CASTp was plotted against the number of reported pockets. Once a determined probe radius was fixed, two different jobs for the TLR4/MD-2/TLR4*/MD-2* and antagonist TLR4/MD-2 systems were submitted to CASTp server. Results were again visualized and compared with SiteMap results. Two different parameters were used to rank binding sites: detection of the binding site by CASTp in the different receptor structures and SiteScore.

3.5   Annex III

a)
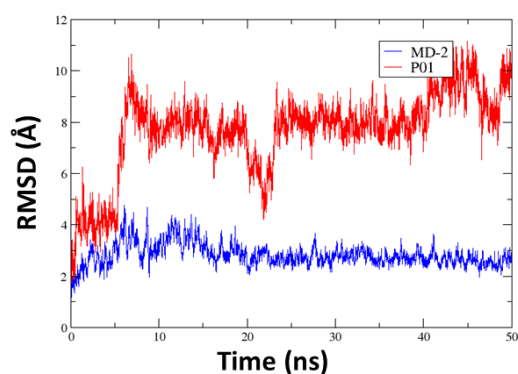
b)
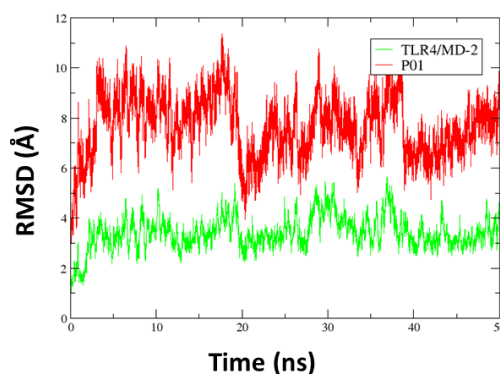
c)
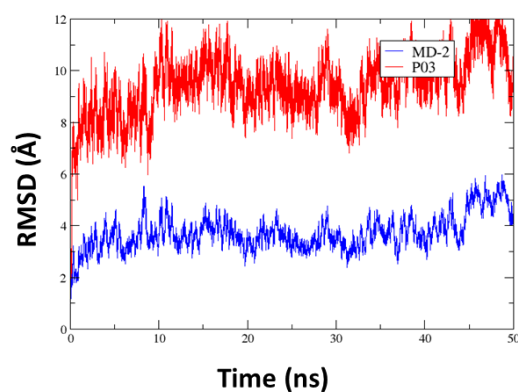
d)
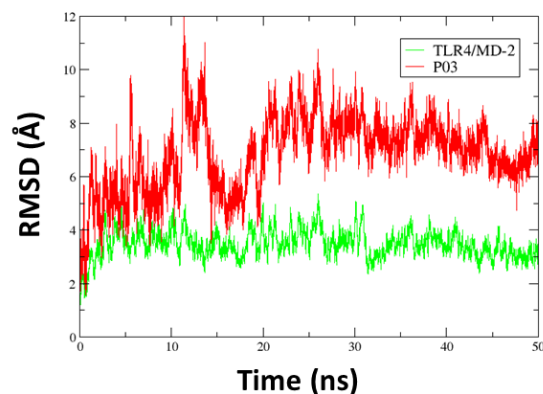
Annex Figure 3.1. a) MD simulation of MD-2/**P01** (50 ns): RMSD (Å) is represented for the α-Carbons for MD-2 (dark blue) and the heavy atoms for the ligand **P01** (red), b) MD simulation of TLR4/MD-2/**P01** (50 ns): RMSD (Å) is represented for the α-Carbons for TLR4/MD-2 (green) and the heavy atoms for the ligand **P01** (red), c) MD simulation of MD-2/**P03** (50 ns): RMSD (Å) is represented for the α-Carbons for MD-2 (dark blue) and the heavy atoms for the ligand **P03** (red), d) MD simulation of TLR4/MD-2/**P03** (50 ns): RMSD (Å) is represented for the α-Carbons for TLR4/MD-2 (green) and the heavy atoms for the ligand **P03** (red).

a)

b)

c)

d)
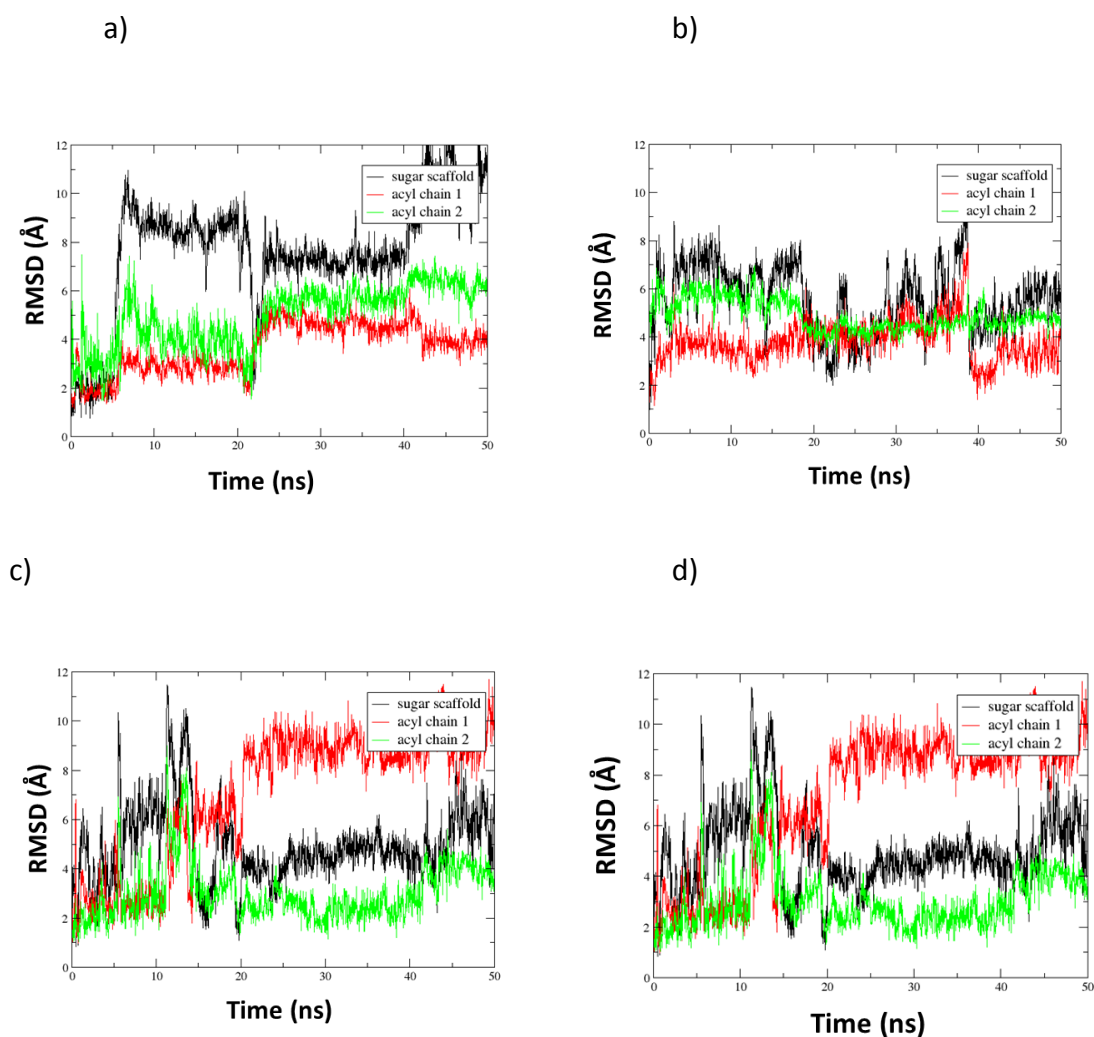
Annex Figure 3.2. RMSD of the positions of the atoms belonging to the P01 ligand. Black acyl chain 1, red acyl chain 2 and green sugar moiety. a) P01 docking in MD-2 and simulation in full complex. b) P01 docking in TLR4/MD-2 and simulation in "MD-2 only". c) P03 docking in TLR4/MD-2 and simulation in full complex. d) P03 docking in TLR4/MD-2 and simulation in "MD-2 only".
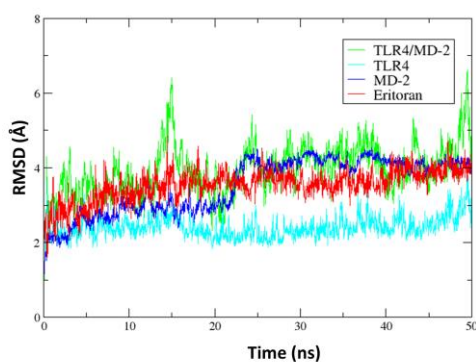
Annex Figure 3.3. MD simulation of the complex TLR4/MD-2/**Eritoran** (50 ns): RMSD (Å) is represented for the α-Carbons for TLR4/MD-2 (green), TLR4 (cyan), and MD-2 (dark blue), and the heavy atoms for the ligand **Eritoran** (red).
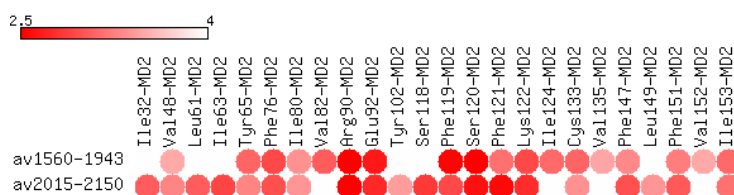
Annex Figure 3.4. Contacts between **Eritoran** ligand and the residues from the TLR4/MD-2 present in two different average structures of the MD simulation.
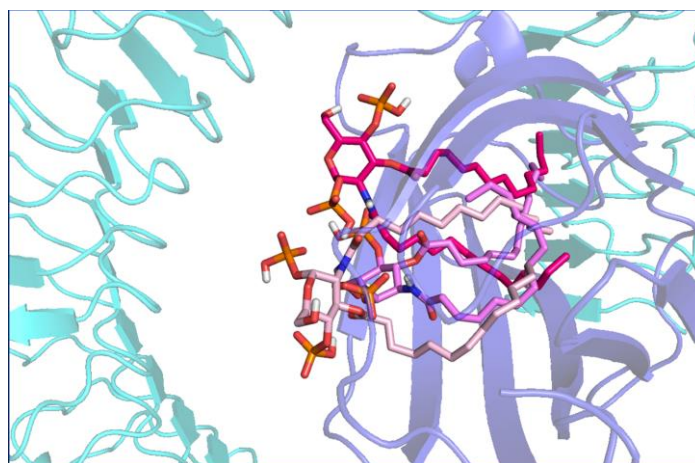
Annex Figure 3.5. Three binding poses (orange, magenta and pink) of compound **A2** in TLR4/MD-2 system.
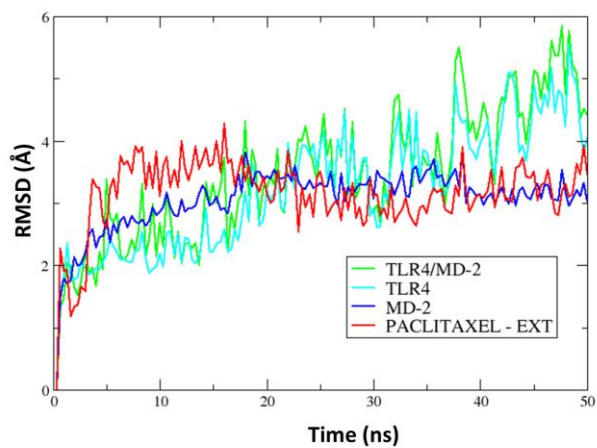
Annex Figure 3.6. MD simulation of the complex TLR4/MD-2/**paclitaxel** (50 ns): RMSD (Å) is represented for the α-Carbons for TLR4/MD-2 (green), TLR4 (cyan), and MD-2 (dark blue), and the heavy atoms for the ligand **paclitaxel** (red).
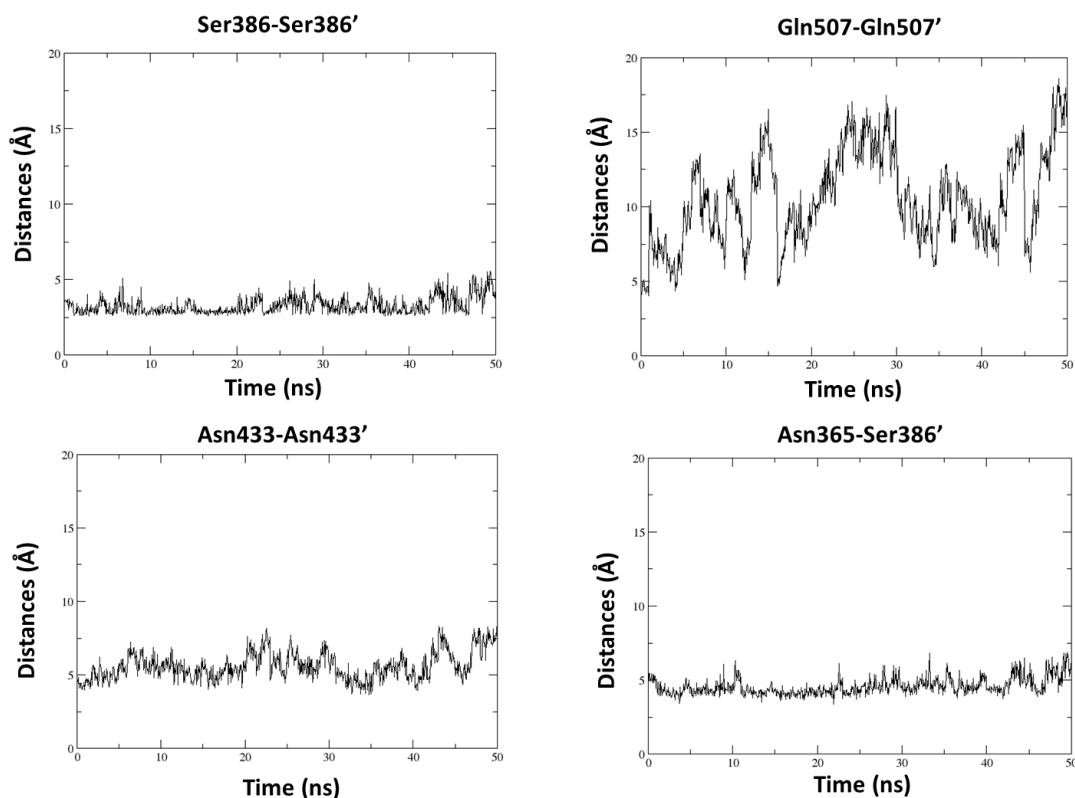
Annex Figure 3.7. Distances between TLR4-TLR4' residues around dimer MD simulation of the **euodenine A** in the full complex.

Annex Table 3.1. Identified SiteMap binding sites in all the receptor structures, with their correspondent SiteScore and Dscore values, and their location in the TLR4/MD-2 complex.

Binding site name	dimerization state	Conformation	SiteScore	Dscore	Location
A, Q	homodimer	agonist	1,228	1,34	MD-2 pocket
A, Q	heterodimer	antagonist	1,205	1,302	MD-2 pocket
A, Q	heterodimer	agonist	1,147	1,229	MD-2 pocket
A	homodimer	agonist	1,126	1,2	MD-2 pocket
Q	homodimer	agonist	1,073	1,118	MD-2 pocket

B, W, X	homodimer	agonist	1,023	1,048	TLR4/MD-2 heterodimerization interface and TLR4
B, X	heterodimer	antagonist	1,018	1,035	TLR4/MD-2 heterodimerization interface and TLR4
B, X	homodimer	agonist	1,012	1,039	TLR4/MD-2 heterodimerization interface and TLR4
B, W, X	heterodimer	agonist	1,01	1,041	TLR4/MD-2 heterodimerization interface and TLR4
C, C	homodimer	agonist	1,009	0,959	TLR4*
T	homodimer	agonist	0,983	0,968	TLR4 (internal cavity)
J	homodimer	agonist	0,981	0,937	TLR4/TLR4 homodimerization interface
L	homodimer	agonist	0,971	0,842	TLR4/MD-2 heterodimerization interface
P, I	homodimer	agonist	0,969	0,897	TLR4/TLR4 homodimerization interface
F, G, H	homodimer	agonist	0,922	0,933	TLR4

F, G, H, O	homodimer	agonist	0,91	0,954	TLR4
C	heterodimer	agonist	0,886	0,892	TLR4
E	homodimer	agonist	0,841	0,84	TLR4 (near transmembrane domain)
C	heterodimer	antagonist	0,838	0,84	TLR4
G, H	heterodimer	agonist	0,838	0,875	TLR4
E	homodimer	agonist	0,835	0,845	TLR4 (near transmembrane domain)
D	heterodimer	antagonist	0,835	0,844	TLR4
W	homodimer	agonist	0,827	0,817	TLR4
L	homodimer	agonist	0,786	0,543	TLR4/MD-2 heterodimerizatio n interface
W	homodimer	agonist	0,78	0,741	TLR-4
U	homodimer	agonist	0,778	0,747	TLR4/TLR4 homodimerization interface
V	homodimer	agonist	0,777	0,695	TLR-4
E	heterodimer	agonist	0,77	0,775	TLR4 (near transmembrane domain)
D	homodimer	agonist	0,762	0,744	TLR4
E	heterodimer	antagonist	0,747	0,735	TLR4 (near transmembrane domain)
D	heterodimer	agonist	0,738	0,718	TLR4

D	homodimer	agonist	0,705	0,674	TLR4
L	heterodimer	agonist	0,703	0,474	TLR4/MD-2 heterodimerizatio n interface
M	homodimer	agonist	0,686	0,595	MD-2 (external binding site)*
M	homodimer	agonist	0,676	0,58	MD-2 (external binding site)*
F, G	heterodimer	antagonist	0,668	0,659	TLR4
I	heterodimer	antagonist	0,649	0,366	TLR4
I	heterodimer	agonist	0,64	0,476	TLR4
G, H	heterodimer	antagonist	0,627	0,586	TLR4
M	heterodimer	agonist	0,623	0,51	MD-2 (external binding site)
O	homodimer	agonist	0,62	0,516	TLR4
X	homodimer	agonist	0,614	0,566	TLR4
R	homodimer	agonist	0,58	0,524	TLR4/MD-2 homodimerization interface
K, Y	heterodimer	antagonist	0,577	0,504	TLR4
N	heterodimer	agonist	0,552	0,471	TLR4
S, Y	homodimer	agonist	0,547	0,491	TLR4
O	heterodimer	agonist	0,543	0,455	TLR4

Annex Table 3.2. Calculated SASA and volume of MD-2 pockets by CASTp for the binding poses and average structure during the MD simulations of the different ligands.

LIGAND	Nº ATOM	DOCKING/AVERAGE	SASA (Å^2)	Volume (Å^3)
Euodenine A	61	docking:	617.650	602.904
		average 42-238:	635.431	495.813
		average 1152-1662:	393.377	309.073
		average 3837-4249:	334.191	261.028
ONO-4007	142	docking:	667.810	715.539
		average 1500-2500:	739.338	837.885
D1	238	docking:	781.697	1007.779
		average 3000-4000	819.913	781.971
Eritoran	213	docking:	613.745	670.638
		average 1560-1943:	771.575	988.573
		average 2015-2150:	591.068	659.075
Paclitaxel	113	EXT-sol2-clust6_docking	635.865	681.513
		EXT-sol2-clust6_average	591.347	525.086
		EXT-sol2-clust7_docking	645.843	721.360
		EXT-sol2-clust7_average	622.278	638.440
		POLAR_sol1_clust30_docking	660.009	690.237
		POLAR_sol1_clust30_average	557.977	470.107
		POLAR_sol1_clust1_docking	660.004	690.231
		POLAR_sol1_clust1_av300-1000	653.755	717.826
		T-TAXOL_docking	660.016	690.214
		T-TAXOL_average	830.141	1005.880
P01	113	Docking	660.009	690.237

		FullComplex_av300-390	605.486	497.741
		FullComplex_av1210-1300	498.415	379.840
		FullComplex_av2340-2420	459.895	318.623
		docking	539.230	527.605
		MD2only_av100-250	831.267	860.676
		MD2only_av1200-2000	518.032	358.442
		MD2only_av2130-2230	668.878	575.479
P03	104	docking	660.009	690.237
		FullComplex_av1560-1943	572.964	439.545
		FullComplex_av2016-2150	708.288	595.026
		docking	539.221	527.594
		MD2only_av130-250	733.237	536.505
		MD2only_av500-850	678.205	647.867
A1	193	docking	699.866	798.737
	115	docking_SOL1	699.866	798.737
		docking_SOL2	699.866	798.737
		docking_SOL3	699.866	798.737

Leu87-Ile124

Met85-Lys125
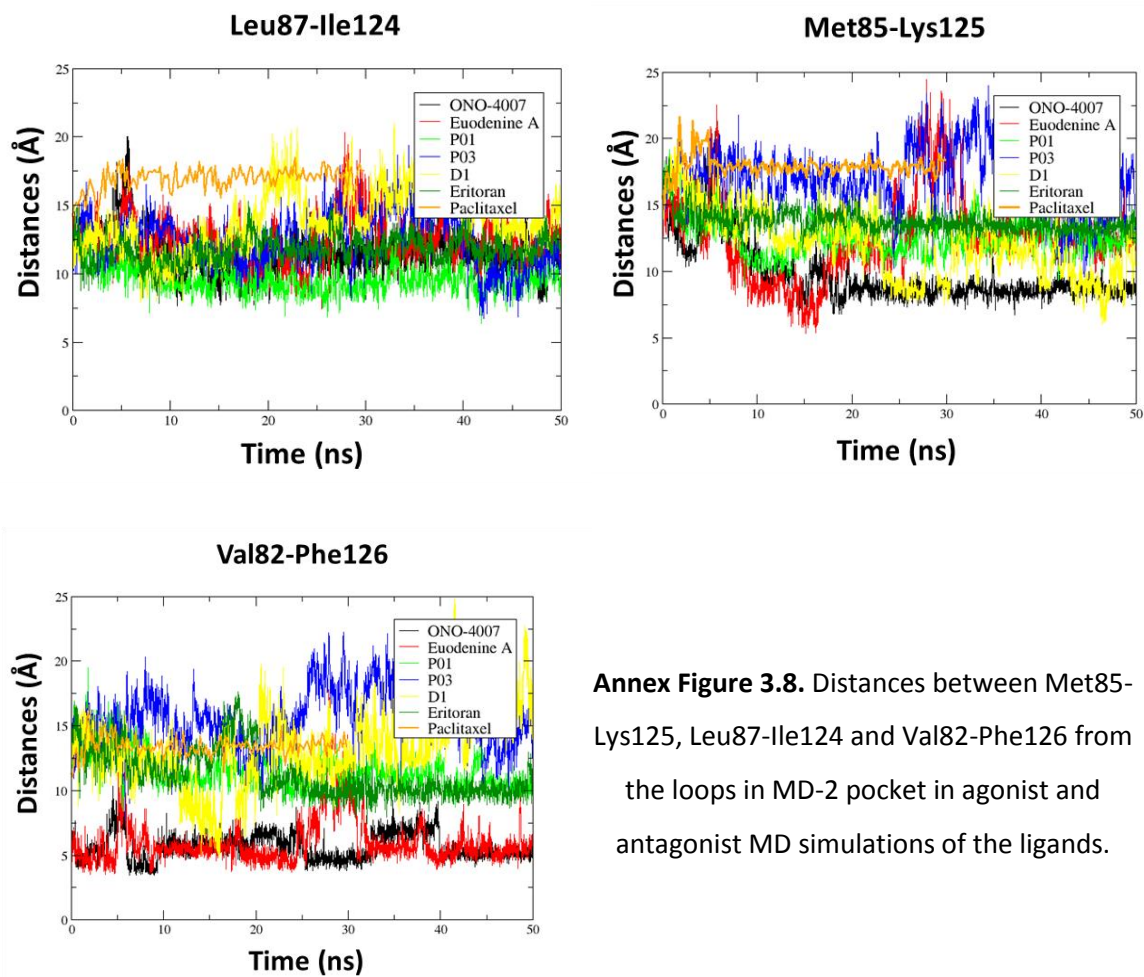
Val82-Phe126

Annex Figure 3.8. Distances between Met85-Lys125, Leu87-Ile124 and Val82-Phe126 from the loops in MD-2 pocket in agonist and antagonist MD simulations of the ligands.

Ligand: **Euodenine A**
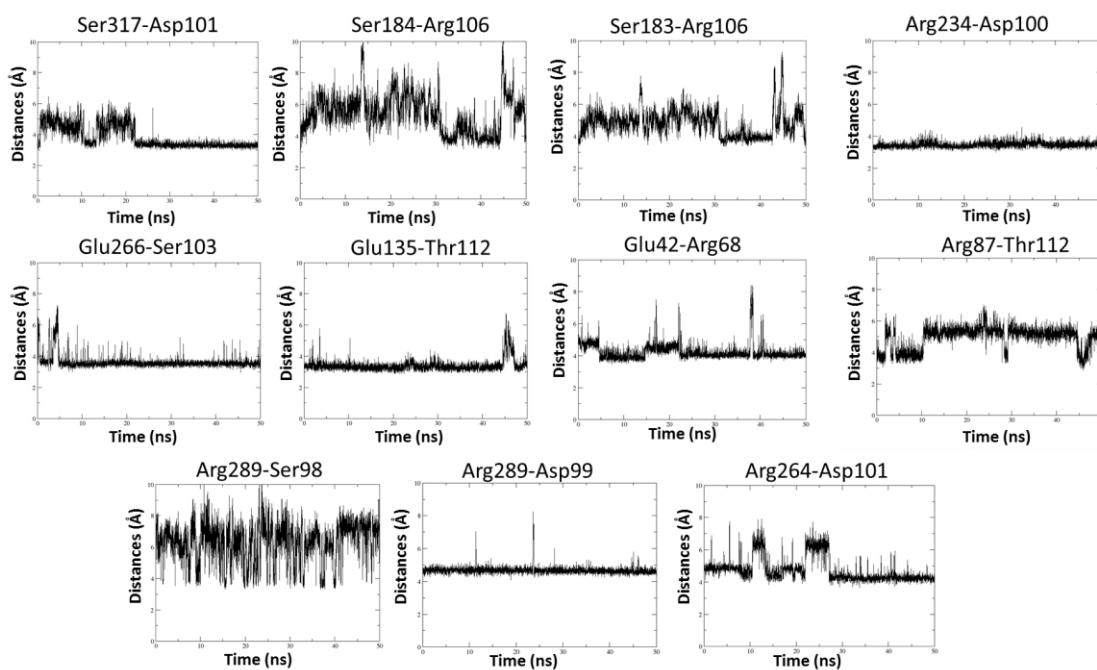
Ligand: **ONO-4007**
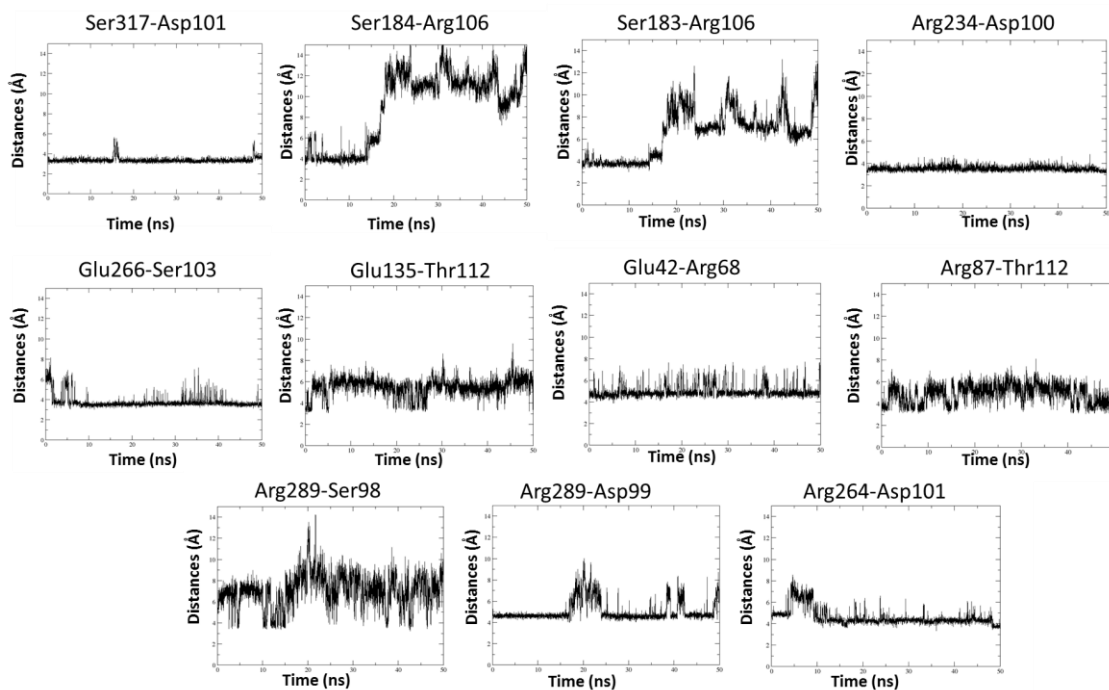
Annex Figure 3.9a. Polar interactions monitorized between TLR4 and MD-2 residues from **Euodenine A** and **ONO-4007** MD simulation.

176
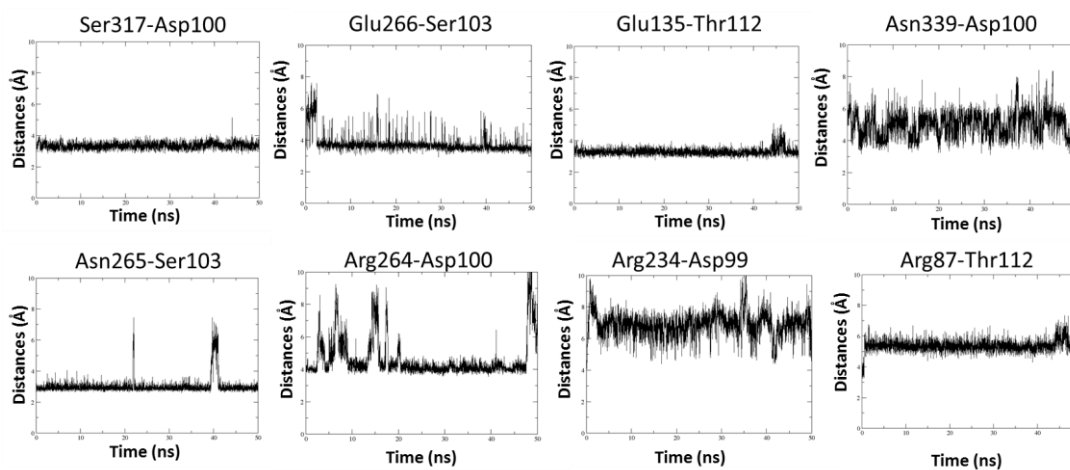
Ligand: **D1**

Ligand: **Eritoran**
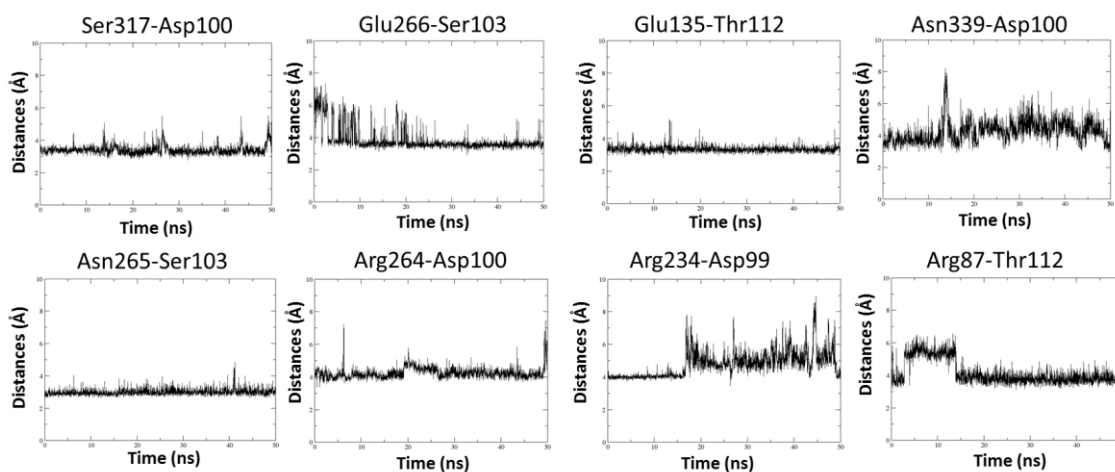
Ligand: **P01**
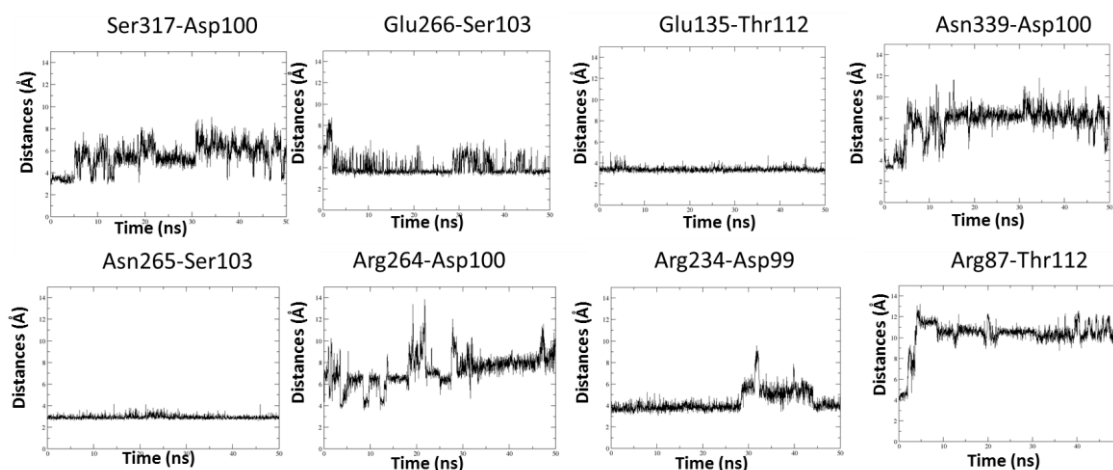
Annex Figure 3.9b. Polar interactions monitorized between TLR4 and MD-2 residues from **D1, Eritoran** and **P01** MD simulation.

ANTAGONISTS	Time (ns)	Average of the water molecules around Phe126	Average of the water molecules around Tyr131
P01_FullComplex	50	24.724	6.544
P01_MD2Only	50	23.248	10.208
P03_FullComplex	50	26.062	8.432
P03_MD2Only	50	26.906	10.038
D1	50	26.102	8.56
Eritoran	50	18.5	6.586
TAXOL-EXT-clust6	30	26.7733	7.59333
TAXOL-EXT-clust7	30	28.4333	14.1933
TAXOL-POLAR-clust1	50	26.552	13.804
TAXOL-POLAR-clust30	30	25.4933	6.36
TAXOL-TTAXOL	30	28.2133	9.08667

AGONISTS	Time (ns)	Average of the water molecules around Phe126	Average of the water molecules around Tyr131
Euodenine A	70	9.84286	10.0271
ONO 4007	50	8.954	7.994

Annex Figure 3.10. Evaluation of the number of water molecules around Phe126 and Tyr131 in a radius of 4 Å from the center of mass of the both residues. The average of the water molecules was taken from the MD simulation starting from the docked binding poses.

APBS from MD-2 pocket:

Ligand **Euodenine A**

docking   average 42-238   average 1152-1662   average 3837-4249

Ligand **ONO-4007**
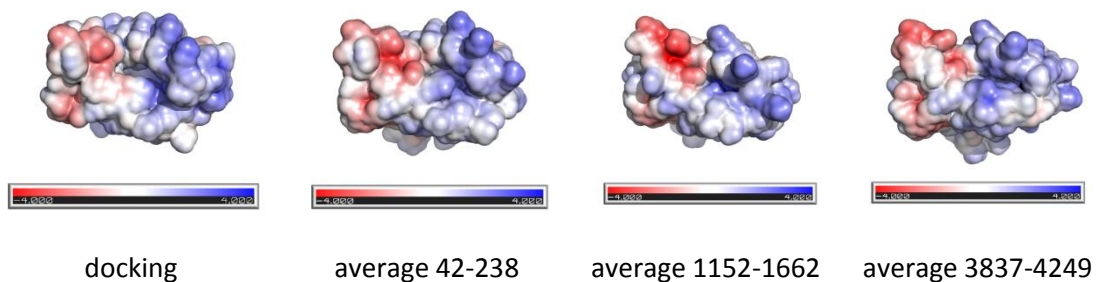
docking   average 1500-2500

Ligand **Eritoran**
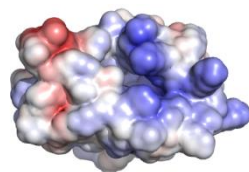
docking | average 1560-1943 | average 2015-2150
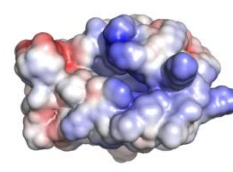
Ligand **D1**
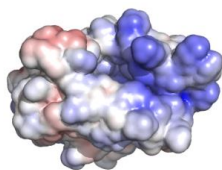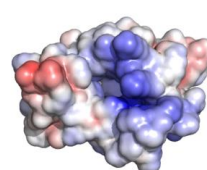
docking | average

Ligand **P01**
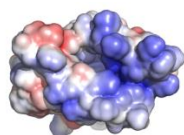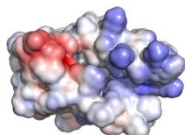
docking | average 100-250 | average 1200-2000 | average 2130-2230
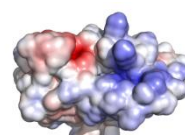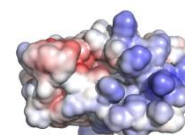
Ligand **P03**
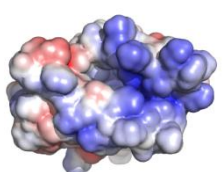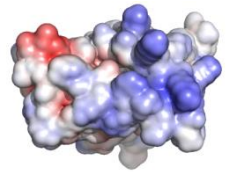
docking | average 130-250 | average 500-850
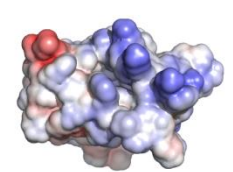
Ligand **A1** | Ligand **A2**

docking   docking

Annex Figure 3.11. Electrostatic potential of MD-2 in Pymol plotted on the solvent accessible surface. The surface colors are clamped at red (-) or blue (+).

Bibliography:

1.   Klett, J.; Reeves, J.; Oberhauser, N.; Perez-Regidor, L.; Martin-Santamaria, S., Modulation of toll-like receptor 4. Insights from x-ray crystallography and molecular modeling. *Curr. Top Med. Chem.* **2014,** *14* (23), 2672-83.

2.   Meng, J.; Lien, E.; Golenbock, D. T., MD-2-mediated ionic interactions between lipid A and TLR4 are essential for receptor activation. *J. Biol. Chem.* **2010,** *285* (12), 8695-702.

3.   Fujimoto, Y.; Adachi, Y.; Akamatsu, M.; Fukase, Y.; Kataoka, M.; Suda, Y.; Fukase, K.; Kusumoto, S., Synthesis of lipid A and its analogues for investigation of the structural basis for their bioactivity. *J. Endotoxin Res.* **2005,** *11* (6), 341-7.

4.   Chan, M.; Hayashi, T.; Mathewson, R. D.; Nour, A.; Hayashi, Y.; Yao, S.; Tawatao, R. I.; Crain, B.; Tsigelny, I. F.; Kouznetsova, V. L., Identification of substituted pyrimido [5, 4-b] indoles as selective Toll-like receptor 4 ligands. *J. Med. Chem.* **2013,** *56* (11), 4206-4223.

5.   Neve, J. E.; Wijesekera, H. P.; Duffy, S.; Jenkins, I. D.; Ripper, J. A.; Teague, S. J.; Campitelli, M.; Garavelas, A.; Nikolakopoulos, G.; Le, P. V., Euodenine A: A Small-Molecule Agonist of Human TLR4. *J. Med. Chem.* **2014,** *57* (4), 1252-1275.

6.   Shanmugam, A.; Rajoria, S.; George, A. L.; Mittelman, A.; Suriano, R.; Tiwari, R. K., Synthetic toll like receptor-4 (TLR-4) agonist peptides as a novel class of adjuvants. *PloS One* **2012,** *7* (2), e30839.

7.   Park, S.-J.; Kang, S. H.; Kang, Y. K.; Eom, Y.-B.; Koh, K. O.; Kim, D. Y.; Youn, H.-S., Inhibition of homodimerization of toll-like receptor 4 by 4-oxo-4-(2-oxo-oxazolidin-3-yl)-but-2-enoic acid ethyl ester. *Int. Immunopharmacol.* **2011,** *11* (1), 19-22.

8.   Jin, G. H.; Li, H.; An, S.; Ryu, J.-H.; Jeon, R., Design, synthesis and activity of benzothiazole-based inhibitors of NO production in LPS-activated macrophages. *Bioorg. Med. Chem. Lett.* **2010,** *20* (21), 6199-6202.

9.   Kawamoto, T.; Ii, M.; Kitazaki, T.; Iizawa, Y.; Kimura, H., TAK-242 selectively suppresses Toll-like receptor 4-signaling mediated by the intracellular domain. *Eur. J. Pharmacol.* **2008,** *584* (1), 40-48.

10.   Chavez, S. A.; Martinko, A. J.; Lau, C.; Pham, M. N.; Cheng, K.; Bevan, D. E.; Mollnes, T. E.; Yin, H., Development of β-Amino Alcohol Derivatives That Inhibit Toll-like Receptor 4 Mediated Inflammatory Response as Potential Antiseptics. *J. Med. Chem.* **2011,** *54* (13), 4659–4669.

11.   Hennessy, E. J.; Parker, A. E.; O'Neill, L. A., Targeting Toll-like receptors: emerging therapeutics? *Nat. Rev. Drug Discov.* **2010,** *9* (4), 293-307.

12.   Peri, F.; Piazza, M.; Calabrese, V.; Damore, G.; Cighetti, R., Exploring the LPS/TLR4 signal pathway with small molecules. *Biochem. Soc. Trans.* **2010,** *38* (5), 1390-5.

13.   Renukuntla, J.; Vadlapudi, A. D.; Patel, A.; Boddu, S. H.; Mitra, A. K., Approaches for enhancing oral bioavailability of peptides and proteins. *Int. J. Pharm.* **2013,** *447* (1-2), 75-93.

14.   Resman, N.; Gradisar, H.; Vasl, J.; Keber, M. M.; Pristovsek, P.; Jerala, R., Taxanes inhibit human TLR4 signaling by binding to MD-2. *FEBS Lett.* **2008,** *582* (28), 3929-34.

15. Hutchinson, M. R.; Northcutt, A. L.; Hiranita, T.; Wang, X.; Lewis, S. S.; Thomas, J.; van Steeg, K.; Kopajtic, T. A.; Loram, L. C.; Sfregola, C.; Galer, E.; Miles, N. E.; Bland, S. T.; Amat, J.; Rozeske, R. R.; Maslanik, T.; Chapman, T. R.; Strand, K. A.; Fleshner, M.; Bachtell, R. K.; Somogyi, A. A.; Yin, H.; Katz, J. L.; Rice, K. C.; Maier, S. F.; Watkins, L. R., Opioid activation of toll-like receptor 4 contributes to drug reinforcement. *J. Neurosci.* **2012,** *32* (33), 11187-200.

16. Fu, W.; Chen, L.; Wang, Z.; Zhao, C.; Chen, G.; Liu, X.; Dai, Y.; Cai, Y.; Li, C.; Zhou, J.; Liang, G., Determination of the binding mode for anti-inflammatory natural product xanthohumol with myeloid differentiation protein 2. *Drug Des. Devel. Ther.* **2016,** *10*, 455-63.

17. Koo, J. E.; Park, Z. Y.; Kim, N. D.; Lee, J. Y., Sulforaphane inhibits the engagement of LPS with TLR4/MD2 complex by preferential binding to Cys133 in MD2. *Biochem. Biophys. Res. Commun.* **2013,** *434* (3), 600-5.

18. Mancek-Keber, M.; Gradisar, H.; Inigo Pestana, M.; Martinez de Tejada, G.; Jerala, R., Free thiol group of MD-2 as the target for inhibition of the lipopolysaccharide-induced cell activation. *J. Biol. Chem.* **2009,** *284* (29), 19493-500.

19. Gradisar, H.; Keber, M. M.; Pristovsek, P.; Jerala, R., MD-2 as the target of curcumin in the inhibition of response to LPS. *J. Leukoc. Biol.* **2007,** *82* (4), 968-74.

20. Chen, G.; Zhang, Y.; Liu, X.; Fang, Q.; Wang, Z.; Fu, L.; Liu, Z.; Wang, Y.; Zhao, Y.; Li, X.; Liang, G., Discovery of a New Inhibitor of Myeloid Differentiation 2 from Cinnamamide Derivatives with Anti-Inflammatory Activity in Sepsis and Acute Lung Injury. *J. Med. Chem.* **2016,** *59* (6), 2436-51.

21. Huang, J.; Xu, Z.; Wang, D.; Ogata, C. M.; Palczewski, K.; Lee, X.; Young, N. M., Characterization of the secondary binding sites of Maclura pomifera agglutinin by glycan array and crystallographic analyses. *Glycobiology* **2010,** *20* (12), 1643-53.

22. Lonez, C.; Irvine, K. L.; Pizzuto, M.; Schmidt, B. I.; Gay, N. J.; Ruysschaert, J. M.; Gangloff, M.; Bryant, C. E., Critical residues involved in Toll-like receptor 4 activation by cationic lipid nanocarriers are not located at the lipopolysaccharide-binding interface. *Cell Mol. Life Sci.* **2015,** *72* (20), 3971-82.

23. Zoroddu, M. A.; Peana, M.; Medici, S.; Potocki, S.; Kozlowski, H., Ni(II) binding to the 429-460 peptide fragment from human Toll like receptor (hTLR4): a crucial role for nickel-induced contact allergy? *Dalton Trans.* **2014,** *43* (7), 2764-71.

24. Svajger, U.; Brus, B.; Turk, S.; Sova, M.; Hodnik, V.; Anderluh, G.; Gobec, S., Novel toll-like receptor 4 (TLR4) antagonists identified by structure- and ligand-based virtual screening. *Eur. J. Med. Chem.* **2013,** *70*, 393-9.

25. Nour, A.; Hayashi, T.; Chan, M.; Yao, S.; Tawatao, R. I.; Crain, B.; Tsigelny, I. F.; Kouznetsova, V. L.; Ahmadiiveli, A.; Messer, K.; Pu, M.; Corr, M.; Carson, D. A.; Cottam, H. B., Discovery of substituted 4-aminoquinazolines as selective Toll-like receptor 4 ligands. *Bioorg. Med. Chem. Lett.* **2014,** *24* (21), 4931-8.

26. Scott, D. E.; Bayly, A. R.; Abell, C.; Skidmore, J., Small molecules, big targets: drug discovery faces the protein-protein interaction challenge. *Nat. Rev. Drug Discov.* **2016,** *15* (8), 533-50.

27. Piazza, M.; Damore, G.; Costa, B.; Gioannini, T. L.; Weiss, J. P.; Peri, F., Hemin and a metabolic derivative coprohemin modulate the TLR4 pathway differently through different molecular targets. *Innate Immun.* **2011,** *17* (3), 293-301.

28.   Ariza, M. E.; Ramakrishnan, R.; Singh, N. P.; Chauhan, A.; Nagarkatti, P. S.; Nagarkatti, M., Bryostatin-1, a naturally occurring antineoplastic agent, acts as a Toll-like receptor 4 (TLR-4) ligand and induces unique cytokines and chemokines in dendritic cells. *J. Biol. Chem.* **2011,** *286* (1), 24-34.

29.   Choi, S. H.; Yin, H.; Ravandi, A.; Armando, A.; Dumlao, D.; Kim, J.; Almazan, F.; Taylor, A. M.; McNamara, C. A.; Tsimikas, S.; Dennis, E. A.; Witztum, J. L.; Miller, Y. I., Polyoxygenated cholesterol ester hydroperoxide activates TLR4 and SYK dependent signaling in macrophages. *PLoS One* **2013,** *8* (12), e83145.

30.   Honda, H.; Nagai, Y.; Matsunaga, T.; Saitoh, S.; Akashi-Takamura, S.; Hayashi, H.; Fujii, I.; Miyake, K.; Muraguchi, A.; Takatsu, K., Glycyrrhizin and isoliquiritigenin suppress the LPS sensor toll-like receptor 4/MD-2 complex signaling in a different manner. *J. Leukoc. Biol.* **2012,** *91* (6), 967-76.

31.   Hutchinson, M. R.; Loram, L. C.; Zhang, Y.; Shridhar, M.; Rezvani, N.; Berkelhammer, D.; Phipps, S.; Foster, P. S.; Landgraf, K.; Falke, J. J.; Rice, K. C.; Maier, S. F.; Yin, H.; Watkins, L. R., Evidence that tricyclic small molecules may possess toll-like receptor and myeloid differentiation protein 2 activity. *Neuroscience* **2010,** *168* (2), 551-63.

32.   Shirey, K. A.; Lai, W.; Scott, A. J.; Lipsky, M.; Mistry, P.; Pletneva, L. M.; Karp, C. L.; McAlees, J.; Gioannini, T. L.; Weiss, J., The TLR4 antagonist Eritoran protects mice from lethal influenza infection. *Nature* **2013,** *497* (7450), 498-502.

33.   Piazza, M.; Calabrese, V.; Damore, G.; Cighetti, R.; Gioannini, T.; Weiss, J.; Peri, F., A Synthetic Lipid A Mimetic Modulates Human TLR4 Activity. *Chem. Med. Chem.* **2012,** *7* (2), 213-7.

34.   Kawasaki, K.; Akashi, S.; Shimazu, R.; Yoshida, T.; Miyake, K.; Nishijima, M., Mouse toll-like receptor 4.MD-2 complex mediates lipopolysaccharide-mimetic signal transduction by Taxol. *J. Biol. Chem.* **2000,** *275* (4), 2251-4.

35.   Ohto, U.; Fukase, K.; Miyake, K.; Satow, Y., Crystal structures of human MD-2 and its complex with antiendotoxic lipid IVa. *Science* **2007,** *316*, 1632-4.

36.   Ciaramelli, C.; Calabrese, V.; Sestito, S. E.; Pérez-Regidor, L.; Klett, J.; Oblak, A.; Jerala, R.; Piazza, M.; Martín-Santamaría, S.; Peri, F., Glycolipid-based TLR4 Modulators and Fluorescent Probes: Rational Design, Synthesis, and Biological Properties. *Chem. Biol. Drug Des.* **2016**.

37.   Fitzpatrick, F. A.; Wheeler, R., The immunopharmacology of paclitaxel (Taxol), docetaxel (Taxotere), and related agents. *Int. Immunopharmacol.* **2003,** *3* (13-14), 1699-714.

38.   Alcaraz, A. A.; Mehta, A. K.; Johnson, S. A.; Snyder, J. P., The T-Taxol Conformation. *J. Med. Chem.* **2006,** *49* (8), 2478-2488.

39.   Zimmer, S. M.; Liu, J.; Clayton, J. L.; Stephens, D. S.; Snyder, J. P., Paclitaxel binding to human and murine MD-2. *J. Biol. Chem.* **2008,** *283* (41), 27916-26.

40.   Kirikae, T.; Ojima, I.; Kirikae, F.; Ma, Z.; Kuduk, S. D.; Slater, J. C.; Takeuchi, C. S.; Bounaud, P. Y.; Nakano, M., Structural requirements of taxoids for nitric oxide and tumor necrosis factor production by murine macrophages. *Biochem. Biophys. Res. Commun.* **1996,** *227* (1), 227-35.

41.   Park, S. H.; Kyeong, M. S.; Hwang, Y.; Ryu, S. Y.; Han, S. B.; Kim, Y., Inhibition of LPS binding to MD-2 co-receptor for suppressing TLR4-mediated expression of

inflammatory cytokine by 1-dehydro-10-gingerdione from dietary ginger. *Biochem. Biophys. Res. Commun.* **2012,** *419* (4), 735-40.

42.   Youn, H. S.; Lee, J. K.; Choi, Y. J.; Saitoh, S. I.; Miyake, K.; Hwang, D. H.; Lee, J. Y., Cinnamaldehyde suppresses toll-like receptor 4 activation mediated through the inhibition of receptor oligomerization. *Biochem. Pharmacol.* **2008,** *75* (2), 494-502.

43.   Zhao, L.; Lee, J. Y.; Hwang, D. H., Inhibition of pattern recognition receptor-mediated inflammation by bioactive phytochemicals: a review of recent research. *Nutr. Rev.* **2011,** *69* (6), 10.1111/j.1753-4887.2011.00394.x.

44.   Binkowski, T. A.; Naghibzadeh, S.; Liang, J., CASTp: Computed Atlas of Surface Topography of proteins. *Nucleic Acids Res.* **2003,** *31* (13), 3352-5.

45.   Dundas, J.; Ouyang, Z.; Tseng, J.; Binkowski, A.; Turpaz, Y.; Liang, J., CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. *Nucleic Acids Res.* **2006,** *34* (Web Server issue), W116-8.

46.   Baker, N. A.; Sept, D.; Joseph, S.; Holst, M. J.; McCammon, J. A., Electrostatics of nanosystems: application to microtubules and the ribosome. *Proceedings of the National Academy of Sciences of the United States of America* **2001,** *98* (18), 10037-41.

47.   http://www.rcsb.org/pdb/, **(accessed on 7th June 2017)**.

48.   Schrodinger a) Suite 2012: Schrodinger Suite 2012 Protein Preparation Wizard; b) Epik version 2.3, S., LLC, New York, NY, 2012; c) Impact version5.8, Schrodinger, LLC, New York, NY, 2012; d) Prime version 3.1, Schrodinger, LLC, New York, NY, 2012.

49.   https://www.molecular-networks.com/online_demos/corina_demo.

50.   Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J., AutoDock treats the ligand as a flexible unit and the protein as a rigid unit.(a) J. *Comput. Chem.* **1998,** *19*, 1639.

51.   Trott, O.; Olson, A. J., AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **2010,** *31* (2), 455-461.

52.   Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S., Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem.* **2004,** *47* (7), 1739-49.

53.   Halgren, T. A.; Murphy, R. B.; Friesner, R. A.; Beard, H. S.; Frye, L. L.; Pollard, W. T.; Banks, J. L., Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *J. Med. Chem.* **2004,** *47* (7), 1750-9.

54.   Friesner, R. A.; Murphy, R. B.; Repasky, M. P.; Frye, L. L.; Greenwood, J. R.; Halgren, T. A.; Sanschagrin, P. C.; Mainz, D. T., Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *J. Med. Chem.* **2006,** *49* (21), 6177-96.

55.   http://www.schrodinger.com/Glide.

56.   Klett, J.; Núñez-Salgado, A.; Dos Santos, H. G.; Cortés-Cabrera, Á.; Perona, A.; Gil-Redondo, R.; Abia, D.; Gago, F.; Morreale, A., MM-ISMSA: An Ultrafast and Accurate Scoring Function for Protein–Protein Docking. *J. Chem. Theory Comput.* **2012,** *8* (9), 3395-3408.

57.   Case, D. A.; Darden, T. A.; Cheatham, T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K. M.; Roberts, B.; Hayik, S.; Roitberg, A.;

Seabra, G.; Swails, J.; Götz, A. W.; Kolossváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wolf, R. M.; Liu, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Cai, Q.; Ye, X.; Wang, J.; Hsieh, M.-J.; Cui, G.; Roe, D. R.; Mathews, D. H.; Seetin, M. G.; Salomon-Ferrer, R.; Sagui, C.; Babin, V.; Luchko, T.; Gusarov, S.; Kovalenko, A.; Kollman, P. A. *AMBER 12*, University of California, San Francisco., 2012.

58.   Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J., *J. Comput. Phys.* **1977,** *23*, 237-34.

59.   York, D. M.; Darden, T. A.; Pedersen, L. G., The effect of long-range electrostatic interactions in simulations of macromolecular crystals: a comparison of the Ewald and truncated list methods. *J. Chem. Phys.* **1993,** *99* (10), 8345-8348.

60.   Halgren, T., New method for fast and accurate binding-site identification and analysis. *Chem. Biol. Drug Des.* **2007,** *69* (2), 146-8.

CHAPTER 4

Glycolipid-Based TLR4 Modulators and Fluorescent Probes:

Rational Design, Synthesis and Biological Properties

4.1 Introduction

As it has been mentioned in the Introduction, the induction of inflammatory responses by endotoxin is achieved by the coordinated and sequential action of four principal endotoxin-binding proteins: the lipopolysaccharide-binding protein (LBP), the cluster differentiation antigen CD14, the myeloid differentiation protein (MD-2), and TLR4.[1] CD14-dependent or independent TLR4 activation by endogenous factors (danger or damage-associated molecular patterns, DAMPs) such as heat-shock proteins, fibronectin, and oxidized phospholipids has been recently related to a wide array of inflammatory disorders, including neuroinflammation and neurological diseases, such as amyotrophic lateral sclerosis (ALS),[2] neuropathic pain,[3] and Alzheimer's disease (AD).[4] Consensus is growing that TLR-directed compounds will provide in the near future new specific drugs against a wide array of diseases still lacking specific pharmacological treatment,[5] and the EU is now strongly committed to support academic and industrial research focused on TLR modulation by small molecules and antibodies.

Small molecules able to interact with membrane CD14 and TLR4/MD-2 dimer are not only new hit for drug development, but also potential templates to develop selective chemical probes allowing in vitro and in vivo imaging of innate immune receptors and ligand/receptor co-localization studies. The portion of LPS that binds to the CD14 and MD-2 receptors is lipid A. Even though lipid A chemical structure varies in different bacterial species, it is generally composed by a hydrophilic domain formed by a glucosamine disaccharide bearing two phosphate groups and a hydrophobic domain formed by linear and branched fatty acid lipid chains attached to the disaccharide core through ester or amide bonds.[6-7] The lipid A structure can be mimicked by synthetic glycolipids bearing anionic phosphates. Synthetic TLR4 activators (agonists)[8] and inhibitors (antagonists)[9-10] have been developed with a variety of clinical and pharmacological applications.[11]

Although the negative charges on phosphate are important for the interaction of lipid A and its synthetic analogs with the CD14 and MD-2

receptors,[12] it has been recently found that also cationic lipids with positively charged head groups are active in modulating TLR4 activity as agonists or antagonists.[13] While the mechanism of action of other cationic lipids seems to be mainly based on interaction with LPS and stabilization or solubilization of insoluble LPS aggregates, specific interaction with MD-2 and CD14 co-receptors was clearly demonstrated for cationic glycolipids developed by our group.[14-16] These compounds are active in inhibiting TLR4-dependent cytokine production in cells and in blocking TLR4 activation by LPS[15] and endogenous[17] stimulation in animal models. Size-exclusion chromatography revealed that incubation of soluble CD14 (sCD14) with these compounds inhibited the transfer of lipooligosaccharides (LOS) from CD14 to TLR4/MD-2.[14] Evaluation of transfer of LOS from monomeric sCD14 to His6-tagged CD14 or MD-2 by co-capture to metal chelating resin clearly showed that the cationic lipids derived from D-glucose or benzylamine inhibit the transfer of LOS from sCD14 to CD14-His6, but not the transfer of LOS from sCD14 to MD-2. Finally, saturation transfer difference (STD) NMR data demonstrated direct binding of the cationic lipids to CD14, through acyl chains mainly.[14] Altogether, these data suggest therefore that the cationic lipid tails insert into the hydrophobic pocket of CD14 and compete with LPS or LOS lipid chains. In fact, the wideness of MD-2 pocket leaves space to accommodate extra units of LPS-like ligands with two fatty acid chains such as IAXO-102 ligand.

With this starting background, we undertook the computer-assisted, structure-based rational design, the synthesis, and a preliminary evaluation of biological activity of compounds **4.1-4.4** (Figure **4.1**). Labeling properties of fluorescent compound **4.1** were studied in murine macrophages. The synthesis and biological evaluation of the compounds were performed at the laboratory of Prof. Peri at the University Milano-Bicocca. All the experimental details can be found elsewhere.[18]

Compounds **4.1** and **4.2** are derived from the glycolipid IAXO-102 previously developed by Francesco Peri's group from Milano (Italy) and now commercially available (Adipogen, www.adipogen.com). Compound **4.1** has a

fluorescein unit directly linked to the glucose C-6, while fluorescein is linked to the sugar through a glutaryl-diaminoethyl-thiourea linker in compound **4.2** (Figure **4.1**). Compounds **4.3** and **4.4** are composed by two glycolipid units with the same structure of IAXO-102 connected through, respectively, C4 diamino and di-ammonium linkers (Figure **4.1**).
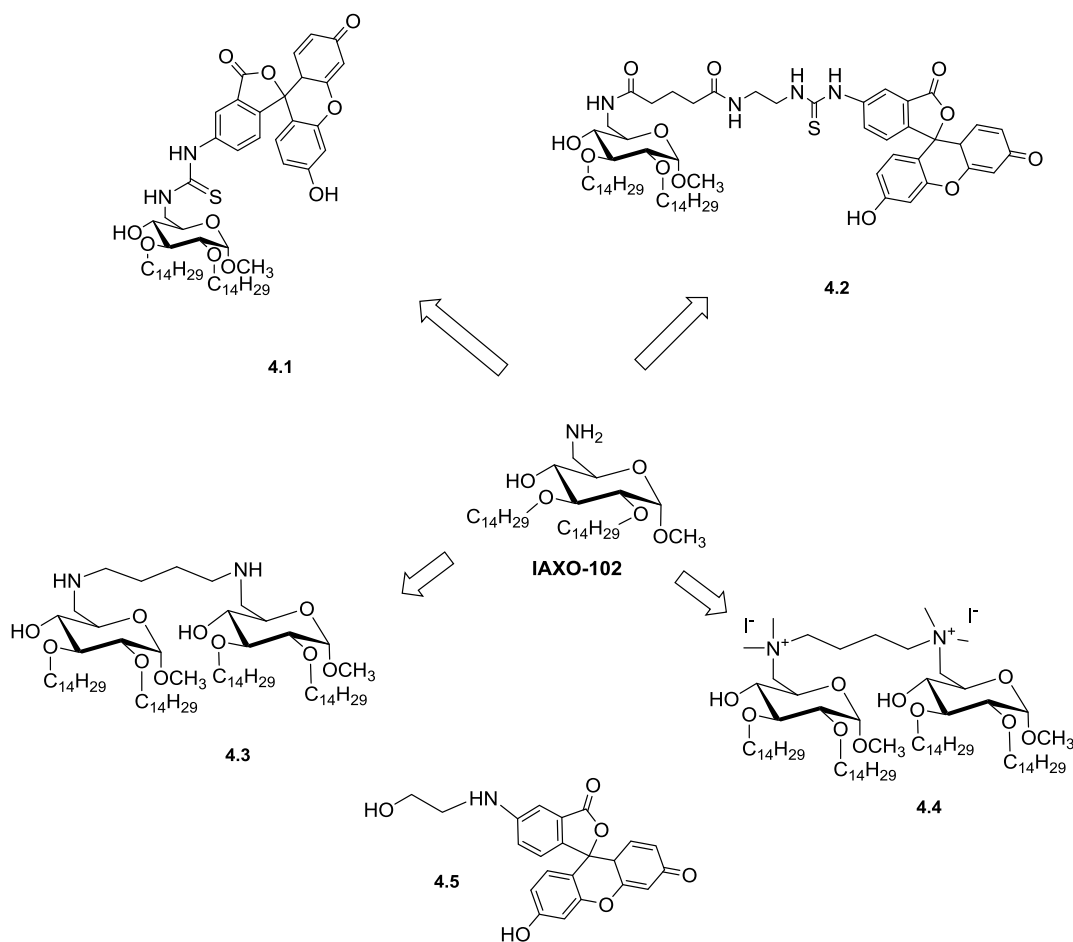
Figure 4.1. Glycolipid derivatives of IAXO-102: fluorescent probes **4.1** and **4.2**, dimeric derivatives **4.3** and **4.4**, and control compound **4.5**.

The use of permanently charged ammonium groups in compound **4.4**, instead of pH-sensitive amines of **4.3**, is aimed at improving water solubility of these molecules and at the same time optimizing charge interactions with CD14 and MD-2 receptors. Compound **4.5**, a water-soluble fluorescein derivative, was

also synthesized to be used as a negative control for fluorescence labeling studies on cells.

4.2   Results

Structures of CD14 and MD-2 Binding Pockets

Structurally, both human and mouse CD14 (*h*CD14 and *m*CD14) are characterized by a bent solenoid typical of leucine-rich repeat (LRR) proteins, with a large hydrophobic pocket found on the amino-terminal side (PDB-ID: 4GLP for *h*CD14, and 1WWL for *m*CD14). In the TLR4/MD-2 complex, MD-2 protein is responsible for LPS binding and it is characterized by a wide lipophilic pocket that hosts the fatty acid chains from LPS. Our calculations showed that both *h*CD14 and *h*MD-2 pockets share a similar topology in terms of solvent accessible surface area (SASA) and volume (Figure **4.2** and Table **4.1**). However, CD14 possesses fewer polar residues in the rim and it is capable of recognizing other microbial and cellular molecular determinants, in addition to LPS.[9] Therefore, despite being very similar in lipophilicity, SASA, and volume, the pockets differ in the polarity of the rim, allowing MD-2 to be more selective than CD14 in the recognition of LPS.
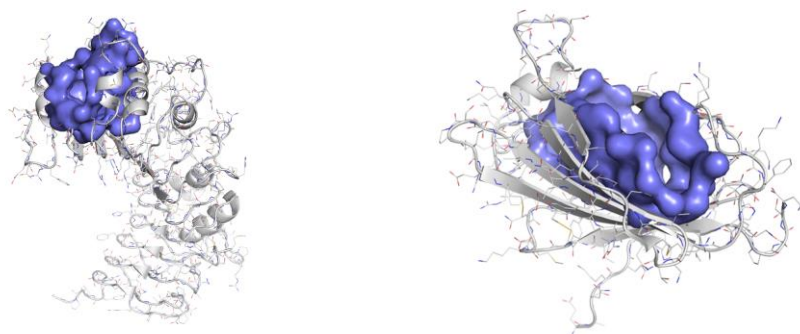
Figure 4.2. Binding pockets identified by CASTp26 in *h*CD14 (left, PDB-ID: 4GLP) and *h*MD-2 (right, PDB-ID: 2E59).

Rational Design of IAXO-102-derived Fluorescent Probes 4.1 and 4.2

Based on IAXO-102 structure (Figure **4.1**), we aimed to develop small-molecule fluorescent probes for endotoxin receptors, able to bind to CD14 and TLR4/MD-2 complex. Then, we designed compounds **4.1** and **4.2** (Figure **4.1**) where IAXO-102 is chemically linked to fluorescein through a thiourea 0-atoms linker (compound **4.1**) or a glutaryl-diaminoethyl 9-atoms linker (compound **4.2**). The C-6 position of the glucose moiety in IAXO-102 has been selected as attachment point for fluorescein because it is apart from lipid chains that directly interact with *h*CD14 and MD-2 receptors, according to NMR binding studies on IAXO-102.[14]

To estimate its CD14 binding properties, designed compounds **4.1** and **4.2** were submitted to docking studies using the X-ray crystal structure of *h*CD14 (PDB-ID: 4GLP), focusing on the amino-terminal pocket, that presumably binds acylated ligands including LPS. As the 3D structure of CD14-bound ligands are not available, and given that the crystallographic structure for CD14 (PDB-ID: 4GLP) adopts a closed conformation of the pocket, thus preventing an efficient exploration by docking, we first undertook a normal mode analysis (NMA) of the protein to represent motions and conformational changes. Finally, docking calculations were performed in three different conformations of *h*CD14 obtained by NMA (see Experimental Section), thus approaching a flexible protein docking protocol. Calculations predicted binding poses for compounds **4.1** and **4.2**, with a general tendency to bury their fatty acid chains inside the hydrophobic pocket, with the sugar pyranose ring remaining toward the external portion (Table **4.2**). Only, 2% of the docking results (performed in the three different structures from NMA) predicted fluorescein moiety inserted inside the CD14 pocket with one or two FA chain outside the pocket, and always with unfavorable theoretical binding energy. These results are in agreement with the binding modes proposed for other TLR4 modulators interacting with *h*CD14.[19] Our calculations have predicted that the thiourea linker and the fluorescein moiety establish polar interactions with the hydrophilic rim, without adopting any preferred binding pose. This

heterogeneity can be observed in the docked poses shown in Figure **4.3**. These results would not be incompatible with the fluorescein tag dangling from the CD14 into the solvent. Similar docking results were obtained for *m*CD14 (Figure **4.4**). Selected docked poses from the most populated clusters were submitted to minimization and analysis of the free energy of binding using the MM-GBSA approach (see Experimental Section).
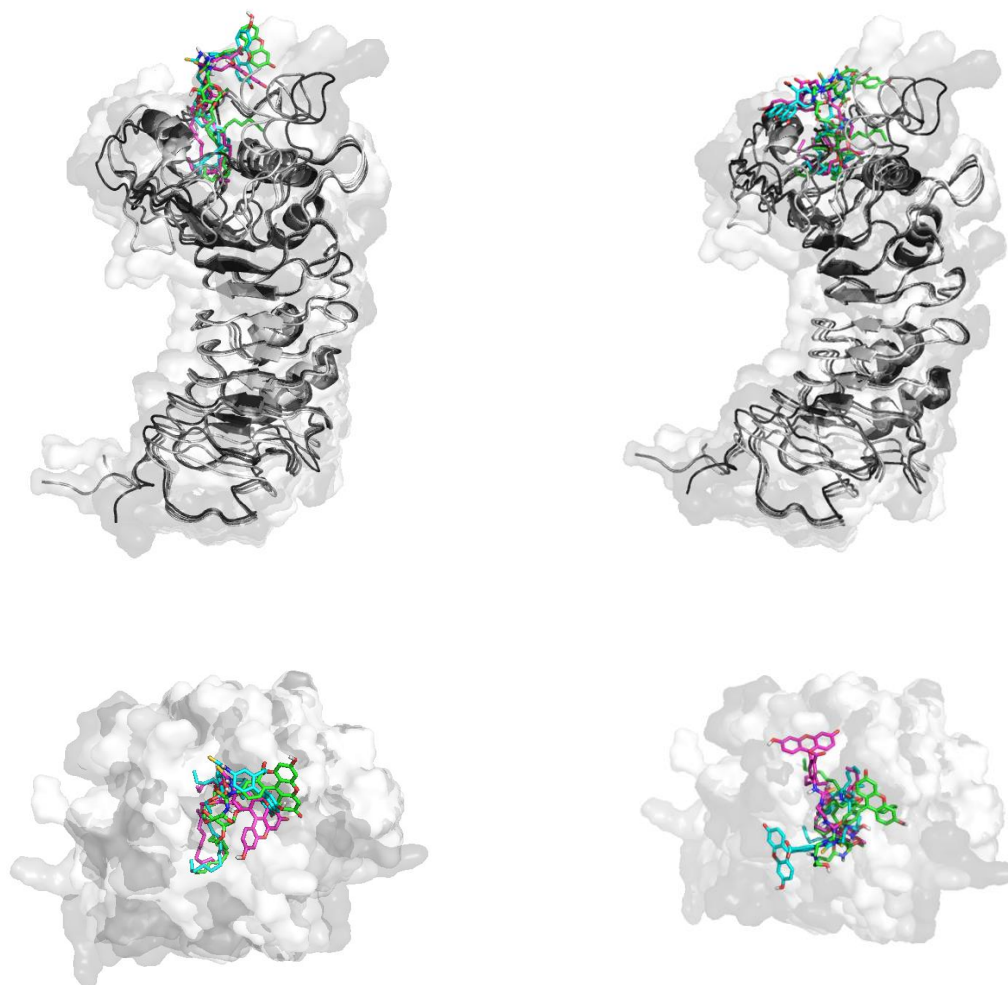
Figure 4.3. Rational design of compounds **4.1** and **4.2**. Docked poses for compound **4.1** (left, lateral, and top views) and compound **4.2** (right, lateral, and top views) binding to *h*CD14 (PDB-ID: 4GLP). Fluorescein remains at the *h*CD14 rim and the IAXO-102 scaffold binds into the *h*CD14 pocket. Fatty acid chains are buried inside the hydrophobic pocket. Three different conformations of *h*CD14 from NMA are superimposed.
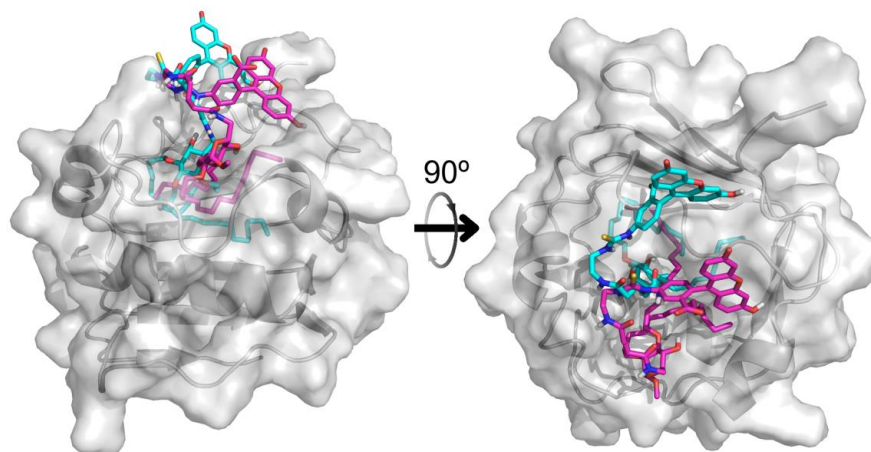
Figure 4.4. Docked poses for compound **4.2** binding to *m*CD14 (PDB-ID: 1WWL). Fluorescein remains at the *m*CD14 rim and the IAXO-102 scaffold binds into the *m*CD14 pocket (on the left). Fatty acid chains are buried inside the hydrophobic pocket (on the right).

To evaluate putative MD-2 binding properties of compounds **4.1** and **4.2**, we also undertook docking studies in the *h*TLR4/MD-2 system. As the X-ray structure of the *h*TLR4/MD-2 is not available in complex with an antagonist, we used a hybrid model built by us from PDB-IDs 3FXI and 2E59 (see Experimental Section). Similarly to CD14, the calculations predicted binding poses with FA chains buried inside the MD-2 hydrophobic pocket (Table **4.2**). These results are in agreement with the antagonist binding pose found for lipid IVa, a precursor of lipid A, in the crystallographic structure (PDB-ID: 2E59). In the case of compounds **4.1** and **4.2**, the thiourea linker and the fluorescein moiety resulted outside MD-2, adopting different orientations (Figure **4.5**) and establishing transient polar interactions with MD-2 residues Arg90, Glu92, Lys122, and Tyr102. In the case of compound **4.2**, the longer linker allows the fluorescein moiety to reach the TLR4 region delimited by Asn361, Lys362, Gly363, and Arg264. In a dynamic environment, it is likely the fluorescein tag may fluctuate into the solvent.
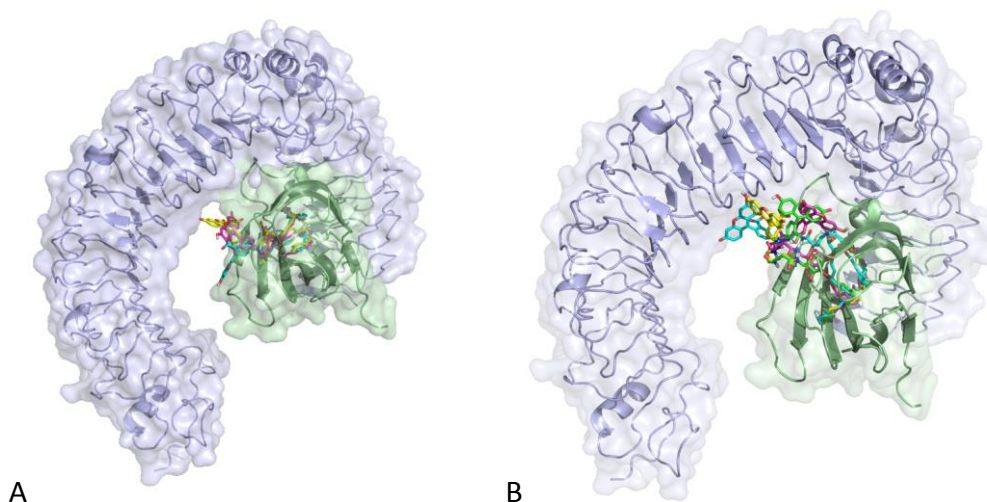
Figure 4.5. Rational design of compounds **4.1** and **4.2**. Docked poses for compound **4.1** (A) and compound **4.2** (B) binding to TLR4/MD-2 (hybrid model from PDB-ID 2E59) showing three preferred binding poses.

Rational Design of IAXO-102 derivatives 4.3 and 4.4

Previous computational and NMR studies performed by our group showed that LPS-like compounds bearing two FA chains are able to bind MD-2, inserting both chains inside the pocket, by adopting different biding poses.[19] In fact, the wideness of MD-2 pocket leaves room to accommodate the ligand while leaving enough empty space to host a second ligand molecule. Duplication of the structure of IAXO-102 would still allow both ammonium groups to remain at the rim of MD-2, while allowing the four FA chains to go inside the pocket, following a dimer-based design strategy already used for TLR4 ligands.[20-21] Thus, different dimeric structures were designed and docked to CD14 and TLR4/MD-2 (data not shown), finally leading to the selection of molecules **4.3** and **4.4**.

Predicted binding poses of compounds **4.3** and **4.4** on the *h*TLR4/MD-2 receptor complex showed a general tendency to bury their fatty acid chains inside the hydrophobic pocket of MD-2, similarly to lipid A and lipid IVa, and to place the sugar moieties in the rim of MD-2, establishing polar interactions. Selected docked poses from the most populated clusters were submitted to analysis of the free energy of binding using the MM-GBSA approach (Table **4.2**). The computed energy

values were reasonably good when compared with the antagonist lipid IVa, whose binding energy was also computed as reference value (Table **4.2**).

Analysis of the docked poses in the *h*TLR4/MD-2 system for both compounds **4.3** and **4.4** showed that the FA chains establish lipophilic interactions with the alkyl and aromatic residues building the inside of MD-2 pocket, mainly Ile44, Ile46, Ile52, Leu61, Ile63, Tyr65, Leu78, Ile94, Ile117, Phe121, Ile124, Val135, Phe147, and Ile153. In addition, CH–π interactions were observed for both compounds between the ligand aliphatic chains and Phe76, Phe104, and Phe151 side chains. No CH–π interactions were observed involving the sugar CH groups. For the analysis of the lipophilic interactions, distances of FA chains of compounds **4.3** and **4.4** with residues alkyl chains were measured (distance to compound **4.3** / distance to compound **4.4**): Ile44 (3.2 Å / 3.4 Å), Ile46 (3.6 Å / 3.5 Å), Ile52 (2.9 Å / 3.5 Å), Leu61 (4 Å / 3.6 Å), Ile63 (3.2 Å / 3.7 Å), Tyr65 (3.4 Å / 3.6 Å), Leu78 (3.7 Å / 3.6 Å), Ile94 (3.4 Å / 3.1 Å), Ile117 (3.3 Å / 3.4 Å), Phe121 (3.4 Å / 3.2 Å), Ile124 (3.3 Å / 3.6 Å), Val135 (3.3 Å / 3.3 Å), Phe147 (3.6 Å / 3.4 Å) and Ile153 (3.2 Å / 3.5 Å). Distance for CH-π interactions were also measured with Phe76 (3.2 Å / 3.1 Å), Phe104 (3.6 Å / 3.3 Å) and Phe151 (4.1 Å / 4.6 Å).

Both compounds also showed a similar pattern of polar interactions. Several hydrogen bonds can be identified between hydroxyl groups from both glucose moieties (glucoses A and B in Figure **4.6**). Distances in the docked TLR4/MD-2 ligand complex are (compound **4.3**/compound **4.4**): Arg90 side chain and O3 from glucose A (2.6 Å/1.9 Å), Glu92 side chain and O4 from glucose A (2.2 Å /2.2 Å), and Arg96 side chain and O5 from glucose B (3.8 Å /3.1 Å). The ammonium group from glucose B is in the close vicinity of Glu92 side chain (distance of 4.5 Å /4.7 Å), establishing a favorable electrostatic interaction for both compounds. In the case of compound **4.3**, the presence of polar hydrogen atoms leads us to suggest the possible formation of a hydrogen bond. Overall, the reasonably good predicted binding properties prompted us to synthesize and test both compounds.[18]
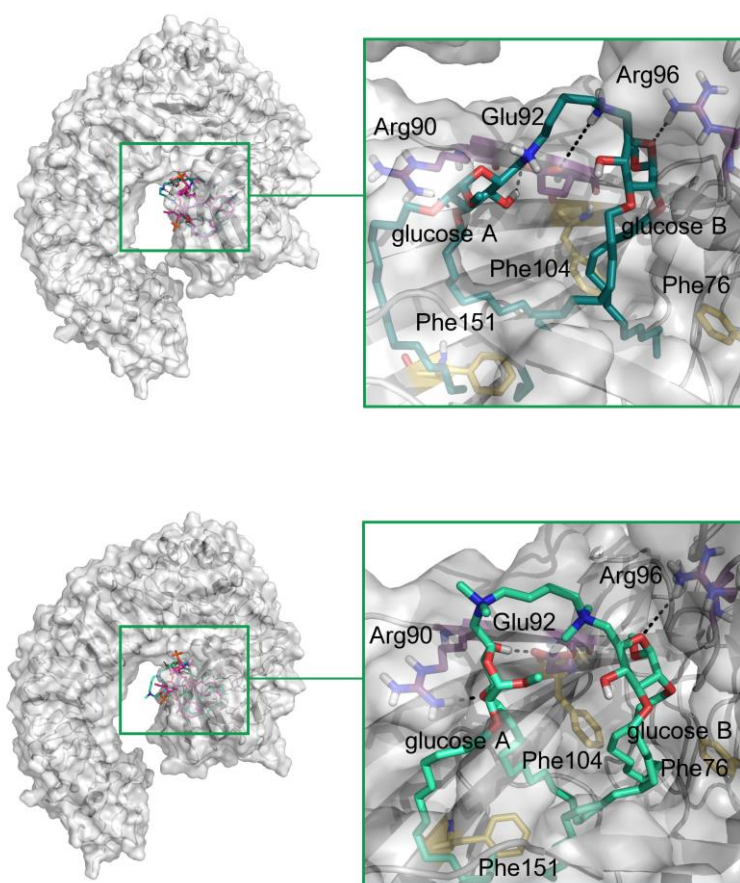
Figure 4.6. Rational design of compounds **4.3** (top) and **4.4** (bottom). Docked poses for compounds **4.3** and **4.4** binding to TLR4/MD-2. Superimposed lipid IVa (magenta) is shown as reference in the global views (left, top, and bottom). Details of some ligand–receptor interactions are shown (right, top, and bottom).

4.3 Discussion

The TLR4 antagonist IAXO-102 has proven to serve as inspiration for rational design of new glycolipid-based TLR4 inhibitors: compounds **4.1** and **4.2** are fluorescein labeled analogs of the parent molecule; compounds **4.3** and **4.4** are homodimers based on another dimeric version of TLR4 antagonist already described by our group.[20] Molecular modeling studies have assisted in the structure based design of new molecules and in the rationalization of the putative binding modes for the three molecules **4.1**-**4.4**, showing that the primary amino group at of IAXO-102 is not directly involved in the binding to CD14 and TLR4/MD-2 receptors, while FA chains are inserted in the lipophilic pockets of the receptors. The primary amino group has been covalently attached to a fluorescent tag through a long short 0-atom or a long 9-atoms linker (molecules **4.1** and **4.2**) or used to link two glucose molecules through di-amine and di-ammonium spacers

(molecules **4.2** and **4.3**). All compounds proved to inhibit in a concentration-dependent manner TLR4 activation and signaling in HEK-Blue[TM] cells expressing *h*TLR4, compounds **4.1** and **4.2** being more active than **4.3** and **4.4**. In preliminary confocal microscopy experiments,[18] compound **4.1** was able to fluorescently label the surface of murine macrophages, and labeling was abrogated when cells were pre-treated with LPS or with trypsin. This would suggest a selective interaction of compound **4.1** with LPS-binding proteins CD14 and MD-2 of the TLR4 receptor system. The reference fluorescein ethanolamine did not bind to cells in the same experimental conditions, thus suggesting again that compound **4.1** interact with LPS-binding proteins on the macrophage surface.

We still do not completely exclude some non-specific interactions of compounds **4.1** and **4.2** with cell membrane, including insertion into membrane bilayer, and we are further investigating this property for these and other compounds of the IAXO series. We are also investigating why the presence of the 9-atoms linker makes compound **4.2** much less efficient than **4.1** in fluorescence labeling of cells.

Homodimers **4.3** and **4.4** with increased hydrophobic part inhibited LPS-stimulated TLR4 signal in cells very weakly.[18] Both these compounds have very poor solubility in aqueous solutions with tendency to form aggregates. Experimentally was found that critical micellar concentration (CMC) for compound **4.3** and **4.4** are, respectively, 18.1 and 29.7 µM using an established fluorescence technique based on pyrene.[22] Both compounds present quite low CMC values, if compared to other cationic glycolipids active as TLR4 modulators.[23] The low value of CMC could account for the weak TLR4 activity of these compounds.

In the concentration range used for biological characterization (10–200 µM), these molecules form aggregates, so that activity on TLR4 receptor system is due to very few residual monomers able to interact with CD14 and MD-2 receptors. While being active in cells, fluorescent compounds **4.1** and **4.2** failed to inhibit LPS-induced cytokine production in vivo. In fact, calculated logP values for compounds **4.1** and **4.2** (9.89 and 9.78, respectively)

(http://www.molinspiration.com) indicate a high predicted lipophilicity, accounting for poor solubility and unfavorable distribution properties that could explain the lack of in vivo activity. Given that the physico-chemical properties can be tuned, the very low toxicity on cells of compounds **4.1** and **4.2** is however an important prerequisite for drug development. Another interesting result from this study is the possibility to chemically label with fluorophores TLR4 antagonists preserving their TLR4 activity. Overall, these findings suggest that molecules **4.1** and **4.2** could be promising hits for the development of TLR4/MD-2 modulators and probes.

4.4   Materials and Methods

Building of Ligands

3D co-ordinates of the compounds **4.1-4.4** were built in Corina[24] from the SMILES code. Optimization of the ligands was performed with MMFFs force field using MacroModel.[25] Two fragments from compounds **4.2** were built and optimized with MMFFs force field: the glycolipid moiety F1 and the fluorescein moiety plus the linker F2 (Figure **4.7**). Fluorescein moiety was modeled according to the protonation state at experimental conditions. 3D co-ordinates of human CD14 protein (PDB-ID: 4GLP) and mouse CD14 (PDB-ID: 1WWL) were refined and optimized under the Protein Preparation Wizard module of Maestro,[25] using AMBER force field. In the case of human CD14, three geometries resulting from the normal mode analysis (NMA) were considered for docking purposes as described below. For the building of the human TLR4/MD-2 system, we used the (hetero) monomer from crystallographic TLR4/MD-2/LPS heterodimer (PDB-ID: 3FXI, agonist conformation), by replacement of the MD-2 protein by the human MD-2 in antagonist conformation (from PDB-ID: 2E59) by superimposition of the Cα trace. Ligand was deleted, missing hydrogens were added, and protonation state of ionizable groups was computed using the Protein Preparation Wizard module of Maestro. This structure was submitted to 100000 steps of steepest

descent minimization with MacroModel,[25] optimized with AMBER force field, and finally used as TLR4/MD-2 macromolecule for the docking calculations.
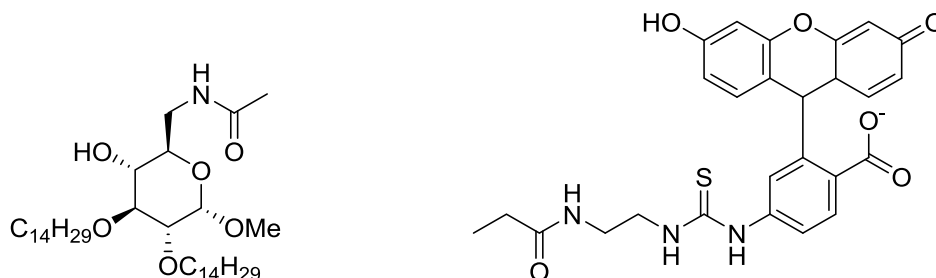
Figure 4.7. Preparation of the macromolecules: fragments F1 and F2 of molecule **4.2.**

SASA and Volume Calculations

Solvent accessible surface area (SASA) and solvent accessible volume were calculated on hydrophobic pockets of CD14 (PDB-ID: 4GLP) and MD-2 (PDB-ID: 2E59) using CASTp server.[26] CASTp server identifies and computes the molecular area and volume for cavities and pockets of a given protein. In our case, we selected the data corresponding to the hydrophobic pockets of both CD14 and MD-2 proteins (Table **4.1**).

Table 4.1. Calculated SASA and volume of
*h*CD14 (PDB-ID: 4GLP) and *h*MD-2 (PDB-ID:
2E59) pockets by CASTp.

Molecule	SASA (Å^2)	Volume (Å^3)
*h*CD14	623	607
*h*MD-2	591	621

Normal Mode Analysis

To compute the low frequency normal modes of CD14, the elastic network model was used with help of the Webinterface ElNemo.[27] We focused our attention on the sequence ranging from Glu26 to Leu135 that corresponds to the binding pocket, and normal mode 4 was found to account for the opening/closing of the pocket. From the trajectory corresponding to this normal mode, we selected three structures to be used for docking purposes: the one corresponding to the crystallographic structure (CD14-a), the one corresponding to the last structure of the displacement in one of the directions of the normal mode (CD14-b), and the last located structure in the opposite direction (CD14-c). The three structures were submitted to 100 000 steps of steepest descent minimization with MacroModel[25] and optimized with AMBER force field, before being used for docking calculations.

Docking Calculations

Docking calculations were performed by means of AUTODOCK 4.2..[28] Analysis was performed with the help of AUTODOCKTools. For CD14 as macromolecule, only the sequence ranging from Ala3 to Leu130 was considered for docking purposes and three geometries were considered: CD14-a, CD14-b, and CD14-c (as reported above). The grid point spacing was set at 0.375 Å, with a number of grid points in xyz of 78, 72, 84 for hCD14, and 66, 72, 88 for mCD14.

Due to the big size of compound **4.2**, and subsequently the high number of degrees of freedom, when performing docking on the hTLR4/MD-2 receptor, the docking was performed in two steps. First, fragment F1 (Figure **4.7**) was docked into the hTLR4/MD-2 model using a box with 48, 76, 56 grid points in xyz, and with a spacing set at 0.375 Å. Best docked pose (with the FA chains inside the pocket and favorable binding energy) was selected for the second step. Second, fragment F2 (Figure **4.2**) was docked starting from the macromolecule containing the best docked solution for F1, using a new box for the TLR4/MD-2/F1 system, with a grid

point spacing set at 0.375 Å , and a number of grid points in xyz of 38, 38, 36. From the best docked solutions for F2, inside the TLR4/MD-2/F1 system, a full geometry for compound **4.2** was built considering appropriate relative fragments orientations, allowing a reasonable F1-F2 connection. Thus, resulting geometry for compound **4.2** was minimized (10 000 cycles of steepest descent minimization with AMBER force field as implemented in Maestro) and subsequently docked into TLR4/MD-2 system. The grid point spacing was set at 0.375 Å, and the number of grid points in xyz was 84, 96, 74. For all cases, the docking protocol was as follows. All allowed torsional bonds were considered rotatable, and Lamarckian algorithm was used (number of individuals in population 150, run 200). Results are shown in Table **4.2**. For compounds **4.1**, **4.3**, and **4.4**, the docking was performed directly into TLR4/MD-2 system. The grid point spacing of the box was set at 0.375 Å, and the number of grid points in xyz was 84, 96, 74.

Table 4.2. Results from the docking calculations of compounds **4.1**, **4.2, 4.3** and **4.4** in macromolecules CD14 and TLR4/MD-2 by means of AutoDock. Binding energy from MM-GBSA analysis of selected docking poses is provided.

Compound	Macromolecule	Selected docked solution	MM-GBSA binding energy (kcal mol^{-1})
4.1	*h*TLR4/MD-2	pose 1	-123.25
4.1	*h*CD14-a	pose 1	-74.40
		pose 2	-120.20
		pose 3	-110.93
	*h*CD14-b	pose 1	-94.97
		pose 2	-107.51

		pose 3	-96.33
	*h*CD14-c	pose 1	-95.04
		pose 2	-63.33
		pose 3	-51.37
4.1	*m*CD14	pose 1	-134.80
4.2	*h*TLR4/MD-2	pose 1	-112.1
		pose 2	-82.2
		pose 3	-129.5
4.2	*h*CD14-a	pose 1	-24.7
	*h*CD14-b	pose 1	-115.1
	*h*CD14-c	pose 1	-102.3
		pose 2	-90.6
4.2	*m*CD14	pose 1	-106.0
		pose 2	-98.4
4.3	*h*TLR4/MD-2	pose 1	-142.1
		pose 2	-142.0
		pose 3	-146.9
4.4	*h*TLR4/MD-2	pose 1	-136.2
		pose 2	-153.4
		pose 3	-130.1
Lipid IVa	*h*TLR4/MD-2	pose 1	-201.9

MM-GBSA Calculations

For selected binding poses, free energy of binding was calculated by means of the MM-GBSA method computed with Prime. In our hands, AUTODOCK scoring function has proved to be an efficient tool to predict binding for LPS and LPS-like ligands. Additional calculation of the free energy of binding by MM-GBSA approach can be appropriate to overcome the possible underestimation of binding energy for big hydrophobic ligands (Table **4.2**).

4.5   Annex IV

Experimental Section

Synthesis of the compounds

The compounds were synthesized at the Prof. Peri´s laboratory following the route shown in Scheme **4.1.** Details can be found elsewhere.[18]
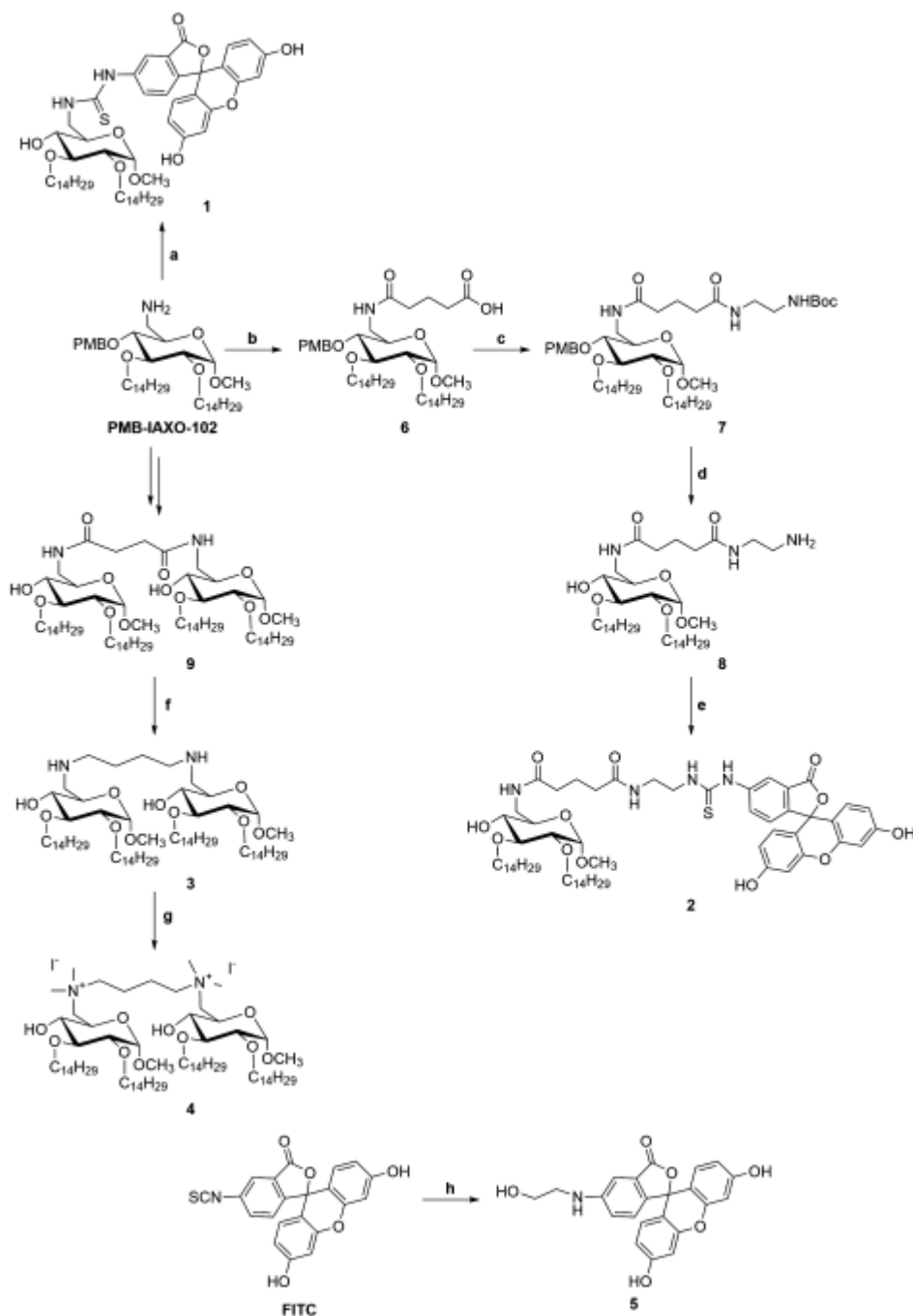
Scheme 4.1. Reactions and conditions: (a) TFA, CH$_2$Cl$_2$, 25 °C, 1 h, 98%; FITC, CH$_2$Cl$_2$, 25 °C, 3.5 h, 92%; (b) Glutaric anhydride, dry pyridine, 25 °C, 2 h, 87%; (c) N-Boc-ethylenediamine, DIC, HOBt, DIPEA, dry DMF, 40 °C, 40 h, 67%; (d) TFA, CH$_2$Cl$_2$, 25 °C, 1.5 h, 73%; (e) FITC, CH$_2$Cl$_2$, 25 °C, 2.5 h, 50%; (f) LiAlH$_4$, dry THF, dry CH$_2$Cl$_2$, 50 °C, 4 h, 78%; (g) CH3I, Na$_2$CO$_3$, dry DMF, 40 °C, 24 h, 73%; (h) NH$_2$(CH$_2$)$_2$OH, THF, 40 °C, 10 h, 76%.

Biological Characterization

Biological characterization was performed at the Prof. Peri´s lab. Complete details can be found elsewhere.[18]

HEK-Blue TLR4 Assay

HEK-Blue-TLR4 cells (InvivoGen, Toulouse, France) and parental cell line HEK-Blue Null 2 (InvivoGen) were cultured according to manufacturer's instructions. Briefly, cells were cultured in DMEM high glucose medium supplemented with 10% fetal bovine serum (FBS), 2 mM glutamine, 19 penicillin/ streptomycin, 19 Normocin (InvivoGen), and 19 HEK-Blue Selection (InvivoGen). Cells were detached by the use of a cell scraper, and the cell concentration was estimated using Trypan Blue (Sigma-Aldrich, St Louis, MO, USA). The cells were diluted in DMEM high glucose medium supplemented as described before and seeded in multiwall plate at a density of 2×10^4 cells/well in 200 μl. After 14 h incubation (37 °C, 5% CO_2, 95% humidity), supernatants were removed, cell monolayers were washed with warm PBS without Ca^{2+} and Mg^{2+} and treated with increasing concentrations of synthetic compounds dissolved in DMSO-ethanol (1:1) and diluted in DMEM. After 30 min, the cells were stimulated with 100 ng/mL LPS from *Escherichia coli* (*E. coli*) O55:B5 (Sigma-Aldrich) and incubated 14 h at 37 °C, 5% CO_2, and 95% humidity. As a control, the cells were treated with or without LPS (100 ng/mL) alone. Then, the supernatants were collected, and 50 lL of each sample was added to 100 μl PBS, pH 8, and 0.84 mM paranitrophenylphosphate (pNPP) for a final concentration of 0.8 mM pNPP. Plates were incubated for 2-4 h in the dark at 25 °C, and then, the plate reading was assessed using a spectrophotometer at 405 nm (LT 4000, Labtech). The results were normalized with positive control (LPS alone) and expressed as the mean of percentage ± SD of at least three independent experiments.

Activity on HEK-Blue^{TM} Cells

The ability of molecules **4.1**-**4.4** to interfere with LPS-stimulated TLR4 activation in HEK-Blue *h*TLR4 cell model was investigated. HEK293 cell line is stably transfected with human TLR4, MD-2, and CD14 genes. In addition, HEK-Blue^{TM} cells stably express a secreted alkaline phosphatase (sAP) produced upon activation of NF-kB. LPS binding activates TLR4 and NF-kB, leading to sAP secretion, which is detected by an alkaline phosphatase substrate in cell culture media. In this assay, cells were pretreated with increasing concentrations of compounds **4.1**-**4.4** and then stimulated with *E. coli* LPS (100 ng/mL, Figure **4.8**).

Fluorescent probes **4.1** and **4.2** inhibited TLR4 activation in a dose-dependent way. Compounds **4.1** and **4.2** had IC50 of the same order of magnitude (about 25 and 10 µM, respectively). Compounds **4.3** and **4.4** too induced an inhibition of the TLR4 pathway, but with lower potencies (IC50 262 and 173 µM, respectively). Reference compound **4.5** (fluorescein-ethanol) turned out to be totally inactive in inhibiting TLR4 activation in HEK-Blue cells.
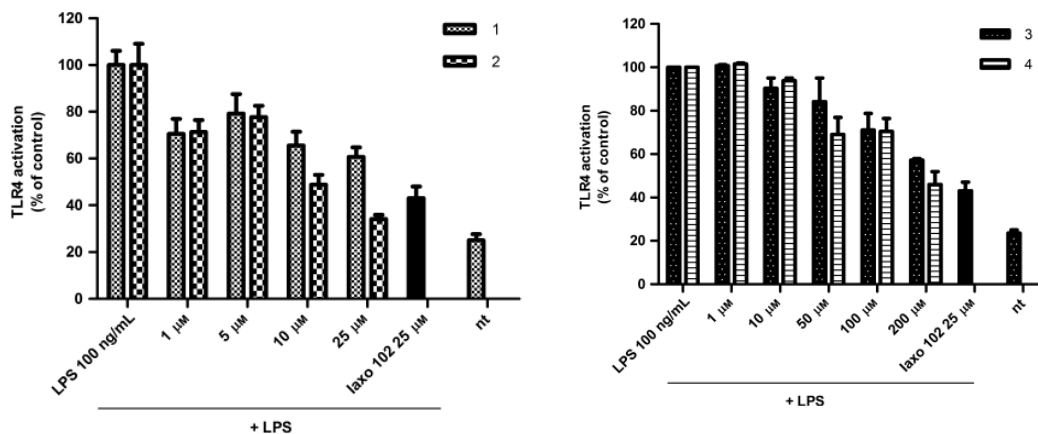
Figure 4.8. Dose-dependent inhibition by compounds **4.1**, **4.2** (left) and **4.3**, **4.4** (right) of LPS-stimulated TLR4 activation. HEK-Blue cells were treated with increasing concentrations of synthetic molecules and then stimulated with LPS. TLR4 activation is monitored as sAP production. The TLR4 inhibition of compound IAXO-102 at 25 µM concentration has been inserted as a reference. nt = HEK-Blue cells treated with 0.25% DMSO–ethanol in DMEM. The results are normalized to activation by LPS alone and

expressed as the mean of percentage ±SD of three independent experiments.

As a negative control, compounds **4.1**-**4.4** were tested in Null cell line (InvivoGen), transfected with the same plasmids as HEK-Blue cells but without TLR4, MD-2, and CD14 genes, and no effect was observed. All compounds were tested by MTT assay and showed no or very low toxicity in the concentration range used for biological characterization (Figure **4.9**).
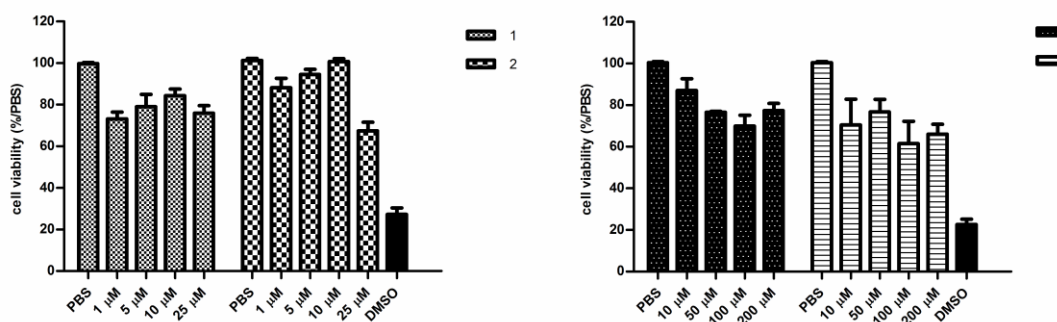
Figure 4.9. Cell viability assay (MTT assay) on compounds **4.1**-**4.4**.

Labeling of Murine Macrophages by Fluorescent Probe 4.1

It was hence explored if compound **4.1** can be used as a fluorescent probe in cells naturally expressing TLR4, MD-2, and CD14. RAW264.7 murine macrophages were treated with 10 µM compound **4.1** and with fluorescein-ethanol **4.5** as a negative control. Confocal microscopy images of the two slides showed that only the fluorescent probe **4.1** bound the cells (Figure **4.10A**), while the negative control **4.5** was completely removed during the washing steps with cold PBS. Selectivity of compound **4.1** for membrane bound TLR4/MD-2 and CD14 endotoxin receptors was tested with a competition assay toward the natural ligand (LPS). Cells were treated with LPS (*E. coli* 055:B5, 100 ng/mL), then incubated with compound **4.1** (10 µM), and analyzed in confocal microscopy. As expected, no fluorescence was detected suggesting that LPS and compound **4.1** compete for the binding with *h*CD14 and TLR4/MD-2 complex.
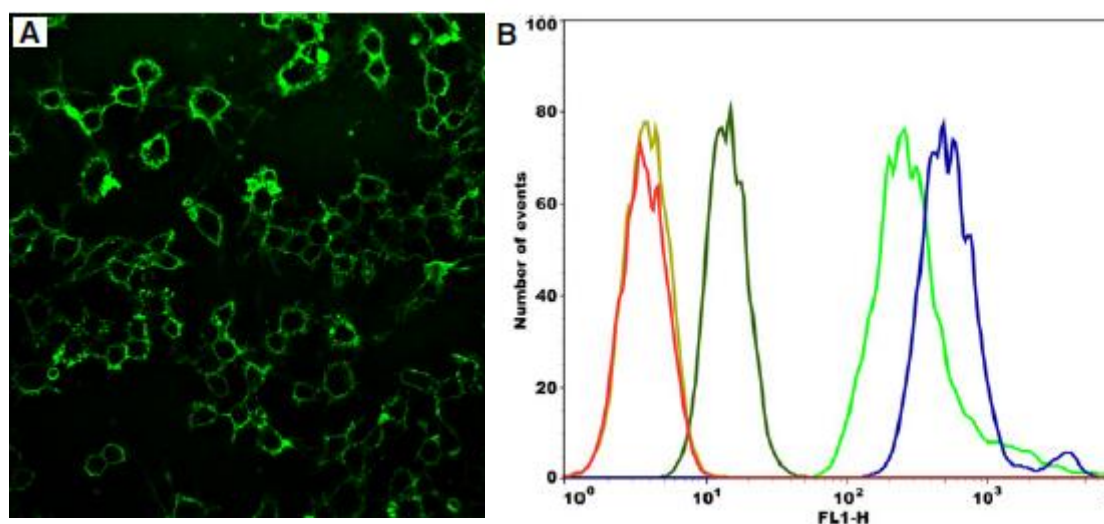
Figure 4.10. Labeling of membrane TLR4 receptor complex on RAW264.7 cells. (A) Confocal micrographs show fluorescence when cells were treated with 10 μM compound **4.1**. (B) FACS analysis of RAW264.7 cells. Red: no compound, blue: compound **4.1**, green: control compound **4.5**, mustard: trypsin + compound **4.1**; light green: LPS-FITC.

To investigate which part of the observed fluorescence was caused by non-specific interaction of the lipophilic part of molecule 1 with cell membrane, RAW264.7 cells were pretreated with trypsin to hydrolyze the extracellular membrane protein components, included endotoxin receptors. Cells were treated at 37 °C for 10 min with 200 μL trypsin and then were incubated with compound **4.1**. Very little fluorescence was observed on cells surfaces. The same samples described in Figure **4.10A** were analyzed with a cytofluorimetric analysis (Figure **4.10B**) to confirm the selectivity of compound **4.1** interaction with the TLR4 receptor system on the whole cellular population. RAW264.7 cells population treated with compound **4.1** (10 μM) shows a remarkable shift toward higher fluorescence content, similar to that observed when cells were treated with fluorescent LPS, while, when exposed to control compound **4.5** (10 μM) or pretreated with trypsin before the addition of **4.1**, limited or no increase in fluorescence was observed.

Compound **4.2** showed a similar behavior than **4.1** as fluorescent probe, but with a very low efficiency in cell labeling.

In Vivo Cytokines Production

As fluorescent compounds **4.1** and **4.2** were the most active in inhibiting TLR4 activation and signaling in HEK-cells, their inhibitory potential was tested in vivo in C57/Bl6 mice, estimating the production of IL-6 and TNF-α cytokines (Figure **4.11**). Unfortunately, the activity found in human cells was not observed in mice, as both compounds did not show significant inhibition of LPS-stimulated cytokine production.
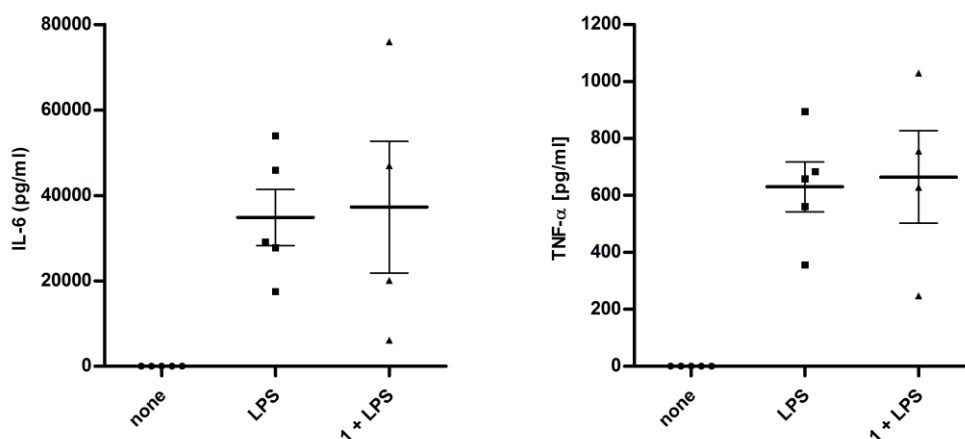
Figure 4.11. In vivo activity of compounds **4.1** and **4.2**. C57/Bl6 mice were injected ip with synthetic molecules (2 x 10^{-7} mol/mouse), followed 1 h later by ip injection of LPS (1 x 10^{-9} mol/mouse). Three hours later, sera were collected and TNF-α and IL-6 concentrations were determined by ELISA assay.

Bibliography:

1. Jerala, R., Structural biology of the LPS recognition. *Int. J. Med. Microbiol.* **2007,** *297* (5), 353-63.

2. Casula, M.; Iyer, A. M.; Spliet, W. G.; Anink, J. J.; Steentjes, K.; Sta, M.; Troost, D.; Aronica, E., Toll-like receptor signaling in amyotrophic lateral sclerosis spinal cord tissue. *Neuroscience* **2011,** *179*, 233-43.

3. Cao, L.; Tanga, F. Y.; Deleo, J. A., The contributing role of CD14 in toll-like receptor 4 dependent neuropathic pain. *Neuroscience* **2009,** *158* (2), 896-903.

4. Rivest, S., Regulation of innate immune responses in the brain. *Nat. Rev. Immunol.* **2009,** *9* (6), 429-439.

5. Peri, F.; Calabrese, V., Toll-like receptor 4 (TLR4) modulation by synthetic and natural compounds: an update. *J. Med. Chem.* **2014,** *57* (9), 3612-22.

6. Rietschel, E. T.; Kirikae, T.; Schade, F. U.; Mamat, U.; Schmidt, G.; Loppnow, H.; Ulmer, A. J.; Zahringer, U.; Seydel, U.; Di Padova, F.; et al., Bacterial endotoxin: molecular relationships of structure to activity and function. *FASEB J.* **1994,** *8* (2), 217-25.

7. Seydel, U.; Schromm, A. B.; Blunck, R.; Brandenburg, K., Chemical structure, molecular conformation, and bioactivity of endotoxins. *Chem. Immunol.* **2000,** *74*, 5-24.

8. Ulrich, J. T.; Myers, K. R., Monophosphoryl lipid A as an adjuvant. Past experiences and new directions. *Pharm. Biotechnol.* **1995,** *6*, 495-524.

9. Kim, H. M.; Park, B. S.; Kim, J. I.; Kim, S. E.; Lee, J.; Oh, S. C.; Enkhbayar, P.; Matsushima, N.; Lee, H.; Yoo, O. J.; Lee, J. O., Crystal structure of the TLR4-MD-2 complex with bound endotoxin antagonist Eritoran. *Cell* **2007,** *130* (5), 906-17.

10. Barochia, A.; Solomon, S.; Cui, X.; Natanson, C.; Eichacker, P. Q., Eritoran tetrasodium (E5564) treatment for sepsis: review of preclinical and clinical studies. *Expert Opin. Drug Metab. Toxicol.* **2011,** *7* (4), 479-94.

11. Molinaro, A.; Holst, O.; Di Lorenzo, F.; Callaghan, M.; Nurisso, A.; D'Errico, G.; Zamyatina, A.; Peri, F.; Berisio, R.; Jerala, R.; Jimenez-Barbero, J.; Silipo, A.; Martin-Santamaria, S., Chemistry of lipid A: at the heart of innate immunity. *Chemistry (Weinheim an der Bergstrasse, Germany)* **2015,** *21* (2), 500-19.

12. Park, B. S.; Song, D. H.; Kim, H. M.; Choi, B. S.; Lee, H.; Lee, J. O., The structural basis of lipopolysaccharide recognition by the TLR4-MD-2 complex. *Nature* **2009,** *458*, 1191–1195.

13. Lonez, C.; Vandenbranden, M.; Ruysschaert, J.-M., Cationic lipids activate intracellular signaling pathways. *Adv. Drug Deliv. Rev.* **2012,** *64* (15), 1749-1758.

14. Piazza, M.; Yu, L.; Teghanemt, A.; Gioannini, T.; Weiss, J.; Peri, F., Evidence of a specific interaction between new synthetic antisepsis agents and CD14. *Biochemistry* **2009,** *48* (51), 12337-44.

15. Piazza, M.; Rossini, C.; Della Fiorentina, S.; Pozzi, C.; Comelli, F.; Bettoni, I.; Fusi, P.; Costa, B.; Peri, F., Glycolipids and benzylammonium lipids as novel antisepsis agents: synthesis and biological characterization. *J. Med. Chem.* **2009,** *52* (4), 1209-13.

16. Piazza, M.; Calabrese, V.; Baruffa, C.; Gioannini, T.; Weiss, J.; Peri, F., The cationic amphiphile 3,4-bis(tetradecyloxy)benzylamine inhibits LPS signaling by

competing with endotoxin for CD14 binding. *Biochem. Pharmacol.* **2010,** *80* (12), 2050-6.

17.   Bettoni, I.; Comelli, F.; Rossini, C.; Granucci, F.; Giagnoni, G.; Peri, F.; Costa, B., Glial TLR4 receptor as new target to treat neuropathic pain: efficacy of a new receptor antagonist in a model of peripheral nerve injury in mice. *Glia* **2008,** *56* (12), 1312-9.

18.   Ciaramelli, C.; Calabrese, V.; Sestito, S. E.; Perez-Regidor, L.; Klett, J.; Oblak, A.; Jerala, R.; Piazza, M.; Martin-Santamaria, S.; Peri, F., Glycolipid-based TLR4 Modulators and Fluorescent Probes: Rational Design, Synthesis, and Biological Properties. *Chem. Biol. Drug. Des.* **2016,** *88* (2), 217-29.

19.   Cighetti, R.; Ciaramelli, C.; Sestito, S. E.; Zanoni, I.; Kubik, Ł.; Ardá-Freire, A.; Calabrese, V.; Granucci, F.; Jerala, R.; Martín-Santamaría, S., Modulation of CD14 and TLR4· MD-2 Activities by a Synthetic Lipid A Mimetic. *ChemBioChem* **2014,** *15* (2), 250-258.

20.   Piazza, M.; Calabrese, V.; Damore, G.; Cighetti, R.; Gioannini, T.; Weiss, J.; Peri, F., A Synthetic Lipid A Mimetic Modulates Human TLR4 Activity. *ChemMedChem* **2012,** *7* (2), 213-7.

21.   Calabrese, V.; Cighetti, R.; Peri, F., Molecular simplification of lipid A structure: TLR4-modulating cationic and anionic amphiphiles. *Mol. Immunol.* **2015,** *63* (2), 153-61.

22.   Hofmann, A. M.; Wurm, F.; Hühn, E.; Nawroth, T.; Langguth, P.; Frey, H., Hyperbranched Polyglycerol-Based Lipids via Oxyanionic Polymerization: Toward Multifunctional Stealth Liposomes. *Biomacromolecules* **2010,** *11* (3), 568-574.

23.   Rodriguez Lavado, J.; Sestito, S. E.; Cighetti, R.; Aguilar Moncayo, E. M.; Oblak, A.; Lainscek, D.; Jimenez Blanco, J. L.; Garcia Fernandez, J. M.; Ortiz Mellet, C.; Jerala, R.; Calabrese, V.; Peri, F., Trehalose- and glucose-derived glycoamphiphiles: small-molecule and nanoparticle Toll-like receptor 4 (TLR4) modulators. *J. Med. Chem.* **2014,** *57* (21), 9105-23.

24.   Sadowski, J.; Gasteiger, J.; Klebe, G., Comparison of Automatic Three-Dimensional Model Builders Using 639 X-ray Structures. *J. Chem. Inf. Comput. Sci.* **1994,** *34* (4), 1000-1008.

25.   Maestro, v., Schrödinger, LLC, New York, NY, 2012.

26.   Binkowski, T. A.; Naghibzadeh, S.; Liang, J., CASTp: Computed Atlas of Surface Topography of proteins. *Nucleic Acids Res.* **2003,** *31* (13), 3352-3355.

27.   http://www.sciences.univ-nantes.fr/elnemo/.

28.   Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J., AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility. *J. Comput. Chem.* **2009,** *30* (16), 2785-2791.

CHAPTER 5

Computational Approaches Towards the Discovery of Toll-Like Receptor 4 Modulators Using Virtual Screening

5.1   Introduction

Identification of drug-like molecules with potential therapeutic applications for the treatment of TLR-related diseases has attracted considerable interest due to their clinical potential. TLR modulators have the potential to be used with different biomedical applications, especially in the field of infection,[1] inflammation[2] and autoimmune diseases,[3] and also in cancer[4-5] and in central nervous system (CNS) disorders such as Alzheimer's disease.[6] However, just few candidates are currently under clinical development due to the difficulty to find molecules with appropriate physic-chemical properties and low toxicity.[7] Therefore, it is imperative to find new chemical entities, and not necessary with LPS-like structure, as TLR modulators with drug-like properties in order to facilitate their development as drugs. There are some small molecules, described in Chapter 3, that exemplify this possibility. For example, some pyrimido[5,4-b]indoles that have shown to stimulate TLR4 and could potentially be used as adjuvants or immune modulators;[8] synthetic analogues of natural product euodenine A have exhibited potent and selective agonist towards TLR4;[9] and synthetic peptides to mimic the TLR4/LPS interaction have also been reported.[10] Also several small non LPS-like molecules with TLR4 antagonist activity have been developed, such as ethyl 4-oxo-4-(oxazolidin-3-yl)-butenoate derivatives (OSL07),[11] benzothiazole-based inhibitors,[12] ethyl phenyl-sulfamoylcyclohexenecarboxylate derivatives (TAK-242 or resatorvid),[13] and β-amino alcohol derivatives.[14]

In the context of drug discovery, virtual screening (VS) techniques have already proved to make hit identification more goal-oriented, allowing the access to a huge number of chemically diverse binders (from public and commercial databases) with a relatively low-cost in terms of time and materials. This computational approach has been subjected to extensive attention and revision over the years, from the early perspective of being an emerging method,[15] until the current time where new challenges are faced.[16-21] We could say that TLRs are not standard receptors which could be approached following classical strategies in drug design. The complexity of the system and the characteristics of their complexation with the pathogen associated molecular patterns (PAMPs) make them especially difficult to tackle following classical

procedures in drug design and discovery. This is why TLRs constitute a special case study in this context. These VS approaches constitute a current strategy in drug design for the identification of novel chemical entities with a given binding ability.[22]

Specifically, on the field of TLR4 research, VS studies have been recently reported leading to novel ligands with drug-like properties, trying to overcome the solubility problems associated with LPS mimetics (see Chapter 1). Among these works, Joce *et al*.[23] have developed a novel *in silico* screening methodology including molecular mechanics and implicit solvent methods to incorporate the evaluation of binding free energies and have screened the Enamine database collection.[24] The resulting clusters were filtered by selecting the representative compounds that were submitted to fast molecular docking for the generation of binding poses and subsequent MD simulations to rank the ligand poses according to their predicted binding affinities. Final filtering led to the identification of compounds T5342126 and T6071187 (Figure **5.1**) as small drug-like inhibitors of the TLR4/MD-2 protein-protein interactions. Their biological activity and selectivity were tested in vitro, and their TLR4/MD-2 antagonist activity was confirmed. In another study, Švajger *et al*.[25] performed parallel ligand-based and structure-based virtual screenings in order to identify novel TLR4 antagonists targeting the TLR4/MD-2 interface, by using the ZINC drug-like subset (~11.3 million drug-like compounds) from the ZINC database.[26] The identified ligands after ligand-based VS resulted in being either insoluble in water, or inactive, or presented cytotoxicity on HEK293 cells. However, the structure-based VS identified 40 putative TLR4/MD-2 ligands that were assessed *in vitro*. After the first assays, only 14 compounds were sufficiently water-soluble and completely non-cytotoxic at 100 μM. These compounds received further biological evaluation, and finally, three compounds with promising antagonistic activities were discovered: ZINC25778142, ZINC49563556 and ZINC3415865 (Figure **5.1**).
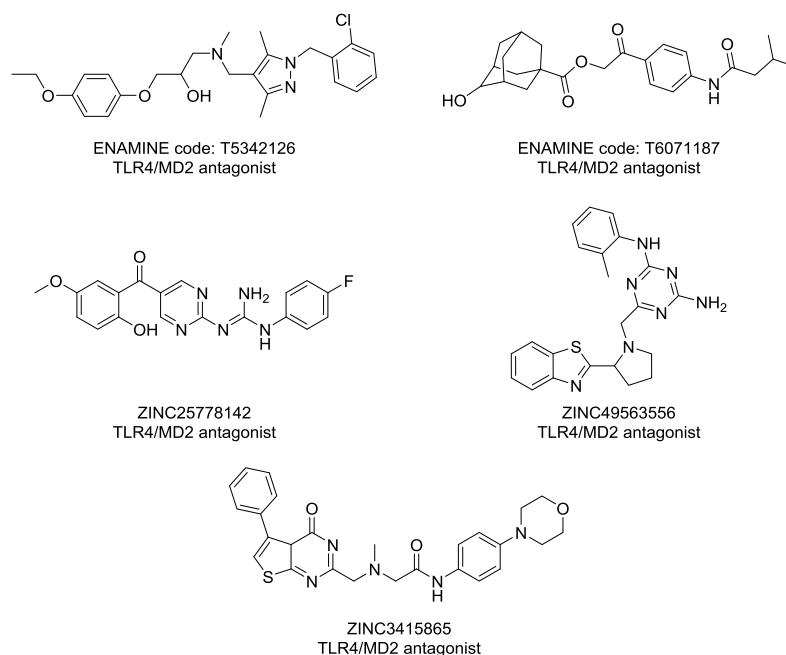
Figure 5.1. Novel TLR4/MD-2 modulators identified by VS approaches.

On the other hand, despite the huge effort of spent time and money on research and development, the number of new drugs brought to market drastically decreases each year.[27] Significant investments by pharmaceutical companies for optimizing the drug discovery pipeline have been undertaken, and new techniques such as structure-based drug design, combinatorial chemistry and high throughput screening (HTS) techniques have emerged. However, the impact of these innovations has not been as important as it was expected both in short and long term.[28] Drug repositioning (also known as drug repurposing, drug redirecting, or drug reprofiling) is a process of discovering new uses outside the scope of the original medical indication for existing drugs. Before 2004, no traces of this process have been found in the literature,[29] but it has gained an increasing attention within the international drug development community over the last few years, and represents a new promising direction.[30-35]

Different terms are used to describe drug repositioning, but all mean a way to find new indications for existing drugs or potential drug candidates, including those in clinical development where mechanism-of-action is relevant to multiple diseases: drugs that have failed to demonstrate efficacy for a particular indication during Phase

II or Phase III trials but with no major safety concerns; drugs that have been discontinued for commercial reasons; marketed drugs for which patents are close to expiry; and drugs candidates from academic institutions and public sector laboratories that have not been fully pursued yet, are also took into account. In this way, drug repositioning represents unique translational opportunities, and is believed to offer great benefits over the *de novo* drug discovery, reducing the development risks and timeline to potentially 3-12 years,[35] substantially increasing the probability of success to brought drugs into market due to existing knowledge about the drugs, and providing relatively inexpensive solutions as therapies for rare and neglected diseases[36-39] that frequently offer limited potential revenue to pharmaceutical companies.

Before the advancements in computational modeling that have led to rational drug repurposing, successful repurposing examples as sildenafil (Viagra®), acetyl salicylic acid (Aspirin®),[40] and thalidomide[41-42] have been due to serendipity. However, recent research has shown that bioinformatics-based approaches have the potential to offer insights into the complex relationships among drugs, targets and diseases for successful repositioning. Given the availability of X-ray crystallographic structures of a number of proteins and identified functional binding sites, and also the advent of molecular docking for the prediction of the free energy of binding of a ligand and its positioning within a defined binding pocket, "computational drug repositioning" is a promising and efficient tool for discovering new uses from existing drugs and holds the great potential for precision medicine in the age of big data.

In this work, we aimed to identify novel TLR4 modulators with non LPS-like structure by means of computational virtual screening. We have followed a virtual screening protocol, and used different commercial and in-house databases. We also present the application of computer-aided drug repositioning in the search of novel TLR4 modulators.

5.2   Results and Discussion

Considering the uncertainty (not yet known) in the binding modes of the reported small molecules with TLR4 activity, and tacking into account that there are plausible pockets of TLR4/MD-2 as binding sites (See Chapter 5), the VS can be considered an effective approach to identify new molecular entities as putative TLR4 binders.

Receptors

Presently, there are several available 3D structures of TLR4, as hetero/homo-dimers, and in complex with some ligands (agonists and antagonists) and/or co-receptors.[43] In the case of the agonist conformation of the *h*TLR4/MD-2 monomer complex, 3D coordinates from TLR4/MD-2 heterodimer were obtained from the PDB (PDB-ID: 3FXI).[44] In the case of the antagonist conformation, since the full crystallographic structure of the *h*TLR4/MD-2 complex is not available, a model built by us was used. This model was built using the human MD-2 protein in antagonist conformation (PDB-ID: 2E59)[45] superimposed onto the MD-2 subunit of the agonist full complex (PDB-ID: 3FXI chain C) through PyMOL (see Chapter 2). Also in order to consider different antagonist conformations of TLR4, we used PDB-ID: 2E56 (only in the case of SPECS and Log P 1000 databases).

Databases

Database processing constitutes a fundamental step in VS approaches. It is crucial to generate the proper chemical library, with the adequate geometries, ionization states, conformations, etc. Furthermore, it is very important to discard any molecule that will not be a good candidate in the further steps of the VS study in relation to the particular system on hand. A good database processing will assure a rigorous and well-conducted virtual screening, as well as it will avoid computational cost and identification of unsuitable drug candidates.

Different commercial, public and in-house databases have been used: Log P 1000, SPECS and ZINC as commercial databases, and as in-house databases, a diversity collection of compounds from laboratories of Prof. Péter Mátyus (PM) from Semmelweis University (Budapest), Prof. Jose Carlos Menéndez[46] (JCM) from Complutense University of Madrid, Prof. J. R. Pedro (JRP),[47-60] and Prof. A. Marco (AM)[61-66] from the University of Valencia.

Commercial databases:

ZINC15 (ZINC Is Not Commercial 2015)[26, 67] is a public access database and tool set, developed to enable ready access to compounds for virtual screening, ligand discovery, pharmacophore screens, benchmarking, and force field development. Nowadays ZINC15 database contains over 120 million purchasable compounds. For the purpose of this work, we were only interested in the approved compounds which represented, that time, a total of 2.459 structures categorized under the substance subset called **WORLD** that is standing for approved drugs in major jurisdictions, including the FDA. Being a computational drug repositioning study, the compounds present in the ZINC15 database were filtered by clinically approved drugs. Thus 2 459 from the WORLD subset over 100 million compounds in total were kept for the repurposing study. These compounds were submitted to a preparation process and the number of compounds increased from 2 459 to 2 949.

Log P 1000 dataset[68] a small diverse subset of the ZINC database.[26, 69] The subset was obtained by a similarity search based on 128 molecular VolSurf+ descriptors[70] covering biologically relevant properties such as shape, surface, volume, molecular weight, polar surface area, hydrogen bonding capacity, lipophilicity and solubility. This was followed by an additional elimination of permanently charged compounds and compounds with a molecular weight lower than 150 Da, resulting in a diverse set of drug-like compounds.

SPECS dataset is a database of commercially available drug-like compounds.[71-72] Due to the large number of compounds (almost 300.000) and computational

limitations, a reduction of the final screening set was necessary. The MOE software[73] was used to perform a cluster analysis on which a diverse subset was created. Therefore the fingerprint of each molecule was calculated in form of a bit-packed version of the molecular access system (MACCS) structural keys (BIT_MACCS),[74] encoding 166 unique features. The Tanimoto coefficient was used as a measure of similarity between fingerprints.[75-77] A similarity of 85% was used for the cluster search. This resulted in a reduced diverse subset of SPECS comprising 23.774 compounds that were used for screening.

In-house databases:

We selected in-house collections with a wide range of chemical structures from different collaborators expert in different types of chemistry. It is important to mention that these chemical libraries are available to perform the biological assay, in the case of these compounds give very good results in the VS studies.

First, we used the diversity collection of heterocyclic compounds, based on their interesting structural characteristics, from Prof. Mátyus from Semmelweis University with around 1964 molecules, a second in-house dataset with quinoline, quinazoline and acridine structures from Prof. Menéndez from Complutense University of Madrid, with 68 compounds, and a third in-house collection of 25 and 85 compounds from Prof. Pedro and Prof. Marco (Universidad de Valencia) respectively, including pyrroles, indoles, naptholes, heterocyclic derivatives from Prof. Pedro and analogues of natural products colchicine and pironetin from Prof. Marco library. All of them had the available samples to be tested in case there were successfully screened.

On the other hand, given that paclitaxel had shown antagonistic activity in human TLR4, while agonistic activity in mouse TLR4, showing the species-specific ligand recognition by MD-2, we were prompted to include tubuline binders in our VS approach. Prof. Marco from the University of Valencia is a well-recognized synthetic chemist specialized in the synthesis of natural products analogues, being analogues to tubuline binders among them. The binding of paclitaxel to TLR4 had been

demonstrated, however not the induction of the cytokine response. Based on these reported results, we included other tubulin binders and related compounds as putative TLR4 ligands in order to discover novel TLR4 modulators. We chose a family of compounds analogue to natural products colchicine and pironetin (Figure **5.2**). Regarding their antitumoral activity, and their ability to bind to tubulin components and microtubules, paclitaxel is a tubulin-interacting drug that stabilizes microtubules, while colchicine causes disruption of microtubules, and pironetin derivates bind to α-tubulin, inhibiting tubulin assembly. These opposite effects are due to the different tubulin sites with which they interact. We also included compounds derived from stilbene, like resveratrol since they are studied for their antimitotic properties and their antitumor activity, all of them from Prof. Marco´s laboratory from the University of Valencia.
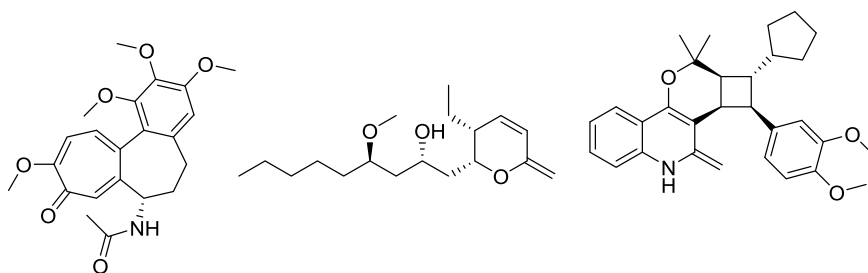
Figure 5.2. Colchicine, pironetin and euodenine A structures.

Filtering

In order to prepare the databases for the VS, different tautomers were considered according physiological pH leading to the corresponding increase in the total number of screened compounds (see Materials and Methods). Finally, in this study, a database composed by around five hundred thousand compounds was built, including known binders (data from the literature) and decoys.

We have considered the following filters:

1- Lipophilicity of the molecules: a maximum logP of 6 were considered, taking into account that the natural LPS and reported synthetic glycolipids have a logP very high: 29.14 ± 0.83, 14.35±0.73 and 13.53±0.47 for lipid IVa, P01 and ONO-

4007 respectively. This limit is a reasonable margin above the value of 5 according to Lipinski´s rules (oral bioavailability).

2- Molecular weight (MW): we considered a wide range between 300 and 700 Da given the MW of glycolipids targeting TLR4, with a reasonable margin above the value of 500 according to Lipinski´s rules.

3- pH: only possible tautomers at physiological pH were considered within a range of 7± 0.5.

4- Prediction of favourable binding from at least two docking programs and in two different conformations of TLR4.

Protocols - Docking programs for virtual screening (SBVS and LBVS)

Molecular docking screening was performed against the different databases based on both, the agonist conformation of *h*TLR4/MD-2 complex from PDB-ID: 3FXI, and our modeled antagonist conformation of *h*TLR4/MD-2 complex. Ligand Based (LBVS) and Structure Based (SBVS) VS were carried following the protocols showed in the Figure **5.3**.
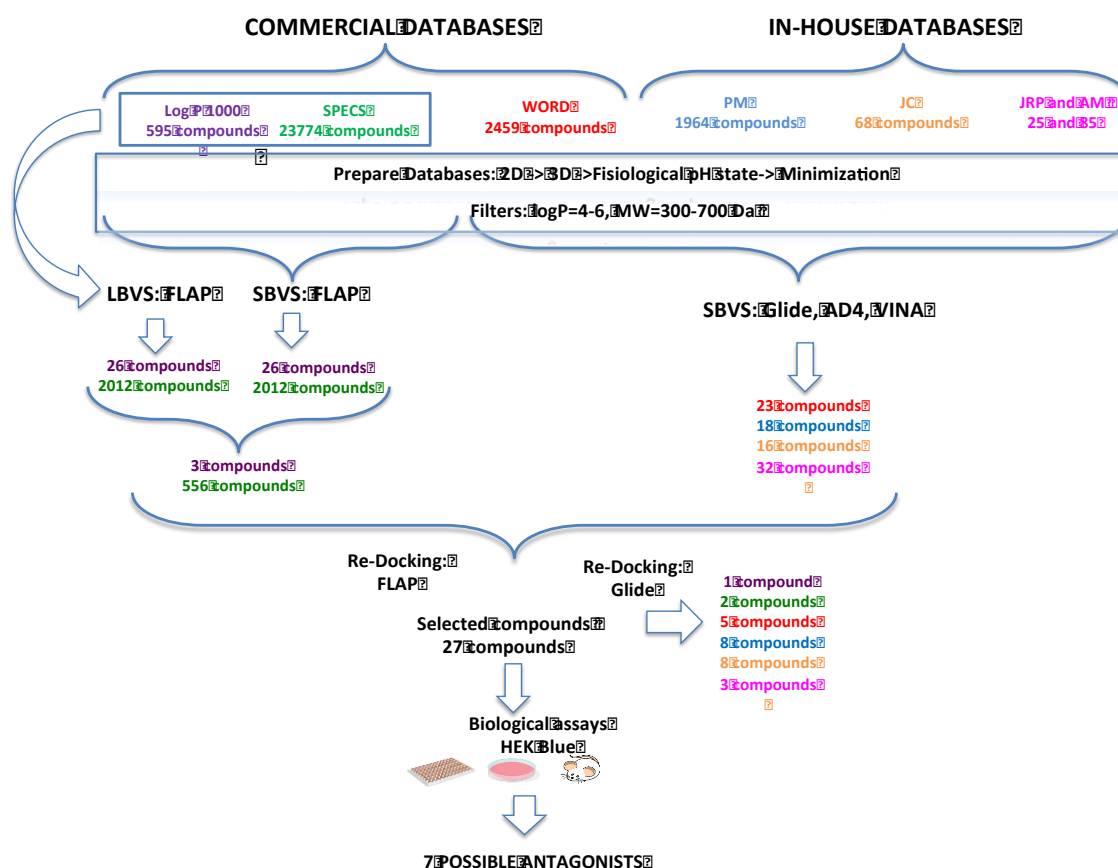
Figure 5.3. Flow chart of VS protocol.

Structure-Based Virtual Screening (SBVS) with FLAP

The investigated compounds from Log P 1000, SPECS datasets are drug-like molecules and not glycolipids like the known binders lipid A, lipid IVa or eritoran. The lipid chains in those glycolipids contain a vast number of free bonds which would increase the docking time exponentially. An additional complication would probably arise from the circumstance that most scoring functions are calibrated on drug-like molecules and would likely have difficulties evaluating the interactions correctly, especially in the entropic term, due to the many degrees of freedom and the large hydrophobic surface.[78]

The literature defines three categories of compounds that may inhibit MD-2.[79] The first class consists of inhibitors that compete with LPS for the binding in the hydrophobic pocket but without being able to trigger the final dimerization; Paclitaxel

is one example for this class of inhibitors.[80-81] Molecules of the second category bind covalently to the residue Cys133; compound JTT705 is one example.[82] The final class of inhibitors does not enter the hydrophobic pocket completely but binds in the opening region of the cavity and prevent LPS from entering the pocket; representatives of this category are compounds JSH, curcumin, xanthohumol and isoxanthohumol (Figure 5.**4**).[83-85] The majority of the side chains of the residues form the MD-2 pocket are hydrophobics (Leu, Ile, Phe and Val), but the rim of the cavity, on the other hand, contains almost no hydrophobic residues. The surface contains many positively and negatively charged amino acids which are important for the interaction between TLR4 and MD-2.

FLAP's SBVS method was used to perform target based VS on the TLR4/MD-2 receptor with SPECS and Log P 1000 databases. As a benchmark, the method was applied initially to the set of known active compounds. The result is shown in Table **5.1** with the ligands ranked according to their *Glob-Sum* score. This score is a global similarity score calculated by summing the four single contributions: shape (H), hydrogen-bond acceptor (N1), hydrophobic (DRY), and hydrogen-donor acceptor (O) descriptors; *Glob-Sum* is the global sum of all four energy values. Note that the scores of the single contributions are derived from the individual best conformation for this type of score, which might be different, while the *Glob-Sum* score comes from the one conformation for which the sum of scores is maximal.

Table 5.1. Known antagonists of MD-2, ranked descending by *Glob-Sum* score obtained from SBVS.

Antagonist	*Glob-Sum*	Antagonist	Glob-Sum
Paclitaxel	3.245	6-shogaol	2.498
JSH	2.714	Isoxanthohumol	2.465
Curcumin	2.695	Isoquiritigenine	2.239
1D10G	2.669	Cinnamaldehyde	2.179

CAPE	2.621	C34	2.136
Xanthohumol	2.611	OSL7	1.799
JTT705	2.513	Sulforaphane	1.707

Since the *Glob-Sum* score does not reflect any experimental binding affinity, the results of the known ligands allow having an idea at which value a screened ligand can be considered as a potential hit. The highest score was obtained by paclitaxel (*Glob-Sum*=3.245) which was then used as a cutoff value for the screened unknown ligands. From Log P 1000 and SPECS libraries, 26 and 2012 compounds were obtained, respectively, having a score equal to or higher than 3.245. The highest contribution to the global score is given by the hydrophobic score which can easily be explained by the high hydrophobicity of the target pocket and the screening model that is obtained from it.

Structure-Based Virtual Screening (SBVS) with Glide, Autodock and VINA

WORD database from ZINC and in-house databases (PM, JCM, JRP and AM) were docked into both agonist and antagonist protein conformations, using three docking programs, Glide, AutoDock and VINA, to avoid the limitation of one scoring function. The receptor grid was set up in order to fully contain the *E. coli* LPS, allowing small molecules to interact with the entire MD-2 pocket, as well as its rim and its entrance (see Materials and Methods). During the docking process, all the ligands were kept to facilitate visual inspections, comparisons and selections between the three docking programs. 50 poses per ligand were generated with AutoDock, 20 poses per ligand with VINA (which is the maximum for the program), and only one pose per ligand was generated with Glide, using HTVS, SP and XP protocols in order to also facilitate the comparisons, choosing Glide as the main docking software. Either with Glide, AutoDock or VINA, the scoring results for all the compounds were consistent and correlated to each other. However, the correlation between AutoDock and VINA is stronger than between Glide and AutoDock or VINA. For the docking program

validation analysis, either with, Glide, AutoDock or VINA, the scoring results for all the compounds were consistent and correlated to each other. However, the correlation between AutoDock and VINA is stronger than between Glide and AutoDock or VINA. The docked compounds, as well as all their corresponding predicted binding poses, were visually analyzed to detect any computational errors. The docking scores and the delta docking scores, defined by the average score of all the poses from one ligand for each docking program were analyzed. Among all the compounds, according to each scoring function, only the top 25% from each docking program, has been kept for the next analysis step. Among the databases, 23, 18, 16 and 32 compounds were selected respectively, being ranked at the top 25% for at least two docking programs at the same time and for one or both conformations, prioritizing a correlation with Glide, were kept for visual cluster rank analysis.
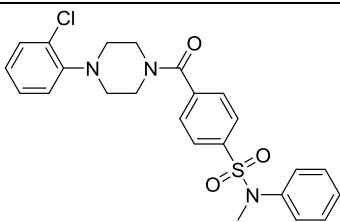
Molecular docking using Glide

The molecules were subjected to a grid-based ligand docking with energetics (Glide, Schrodinger, version 6.9)[86-88] using the Virtual Screening Workflow protocol (See Materials and Methods). Regarding the docking step parameters, Epik state penalties for docking were used, and the non-polar part of the ligand potential were soften by scaling the van der Waals radii of ligand atoms with small partial charges. The full workflow includes three docking stages, each step differing from the preceding step in the amount of time taken to dock each molecule and the scoring system used to evaluate each pose. The first stage performs HTVS (High Throughput Virtual Screening) docking. The ligands that are retained are then passed to the next stage, which performs SP (Standard Precision) docking. The survivors of this stage are passed onto the third stage, which performs XP (eXtra Precision) docking, a more powerful and discriminating procedure. The Dock flexibility method was used for HTVS, SP and XP dockings allowing us to penalize non-planar conformation for amide bonds. A post-docking minimization was also performed, as well as constraints for the docking stages. One pose per compound state was generated and 100 % of the best compounds that passed the HTVS, SP and XP docking have been kept. For HTVS and SP

docking, all states have been retained, but only the best scoring state for the XP docking.

Molecular docking using AutoDock and VINA.

Docking was also performed independently with both VINA[89] and AutoDock.[90] In AutoDock the Lamarckian evolutionary algorithm was chosen and all parameters were kept default except for the number of genetic algorithm (GA) runs which was set to 50 to sample more docked poses. VINA (Vina Is Not AutoDock) uses an Iterated Local Search global optimizer[91-92] based on Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm which approximates Newton's method and the number of docking poses was set to 20, which is the maximum for the program. TLR4/MD-2 receptors were kept rigid and the ligands were set partially flexible (i.e. maximum of 32 dihedral angles) for AutoDock and totally flexible for VINA.

Docking program validation analysis

Either with Glide, AutoDock or VINA, the scoring results for all the compounds were consistent and correlated to each other. However, the correlation between AutoDock and VINA is stronger than between Glide and AutoDock or VINA.

Molecular docking data analysis: score and cluster ranks

The all the docked compounds, as well as all their corresponding predicted binding poses, were visually analyzed to detect any computational errors. The docking scores and the delta docking scores, defined by the average score of all the poses from one ligand for each docking program were analyzed. Among all the compounds, according to each scoring function, only the top 10 % in the case of WORD and PM databases, and 20% from JCM, JRP and AM databases from each docking program, that is to say 89 compounds has been kept for the next analysis step, re-docking with Autodock and Glide.

Among them, 89 compounds being ranked at the top for at least two docking programs at the same time and for one or both conformations, prioritizing a correlation with Glide, were kept for visual cluster rank analysis.

LBVS identified potential hit candidates to inhibit TLR4/MD-2

In the case of the commercial databases, we have performed LBVS. It was the first step of the search for TLR4/MD-2 inhibitors. Even though several inhibitors of the TLR4/MD-2 complex acting on MD-2 were found in the literature, only a minority shows promising characteristics to become an available drug. Eritoran for example showed promising results in phase I and II clinical trials, but in phase II failed in showing better properties than existing treatments for sepsis.[93]

The FLAP LBVS method uses the common reference framework to align a set of candidate molecules to the template binder, to find the optimal overlap according to the GRID MIFs. The similarity between the fields is quantified by the Tanimoto coefficient. In the output table the user can see the individual scores obtained by the single MIF contributions (*Glob-Prod*), as well as a global score representing the sum (*Glob-Sum*). for each compound.

In this work, for the LBVS with FLAP, a set of known active antagonists of MD-2 was built based on a literature search (Figure **5.4**). The two datasets Log P 1000 and SPECS were screened on each known active separately and ranked by their obtained *Glob-Sum* scores.
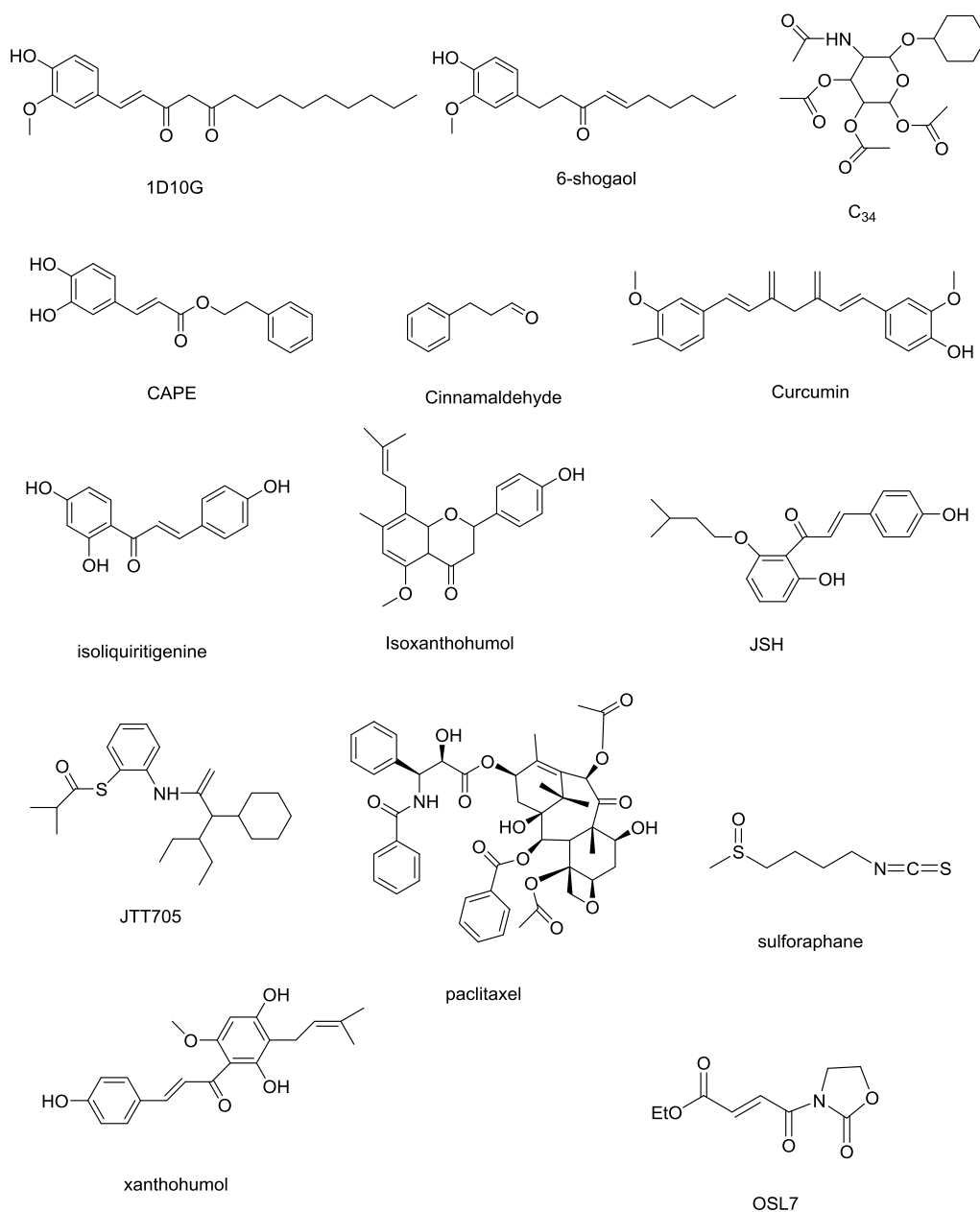
Figure 5.4. Known antagonists of the MD-2 reported in the literature.

LBVS was here performed individually by using the 14 known ligands as templates for the screening. The best ranked results are shown in Table **5.2**. The 2D representations of the compounds of Log P 1000 and SPECS can be found in Annex Figure **5.1** and **5.2**, respectively. The similarity between template and test molecule of the single contributions is a value between 0 (no similarity) and 1 (high similarity). The four single contributions are shape (H), hydrogen-bond acceptor (O), hydrogen-bond

donor (N1) and hydrophobic (DRY) potential. The global score value that was used as result for the screening analysis is the *Glob-Sum* which is the global sum of all four energy values.[94] Note that the scores of the single contributions are derived from the individual best conformation for this type of score, which might be different, while the *Glob-Sum* score comes from the one conformation for which the sum of scores is maximal.

Table 5.2. Best ranked compounds of Log P 1000 (blue) and SPECS (white) set for each known ligand (black).

Template	Compound	Glob-Sum	H	N1	DRY	O
6-shogaol	152	1.326	**0.663**	0.508	0.224	0.239
	481	1.742	0.598	0.271	0.254	**0.702**
Xanthohumol	568	1.269	**0.699**	0.283	0.340	0.124
	19907	1.912	0.703	0.368	0.508	0.359
Paclitaxel	383	0.847	**0.505**	0.175	0.134	0.337
	20513	1.022	**0.565**	0.171	0.105	0.321
1D10G	368	1.152	**0.579**	0.203	0.181	0.310
	20700	1.857	**0.654**	0.306	0.260	0.734
JSH	492	1.165	**0.598**	0.359	0.229	0.144
	21315	1.421	**0.515**	0.371	0.304	0.329
Isoliquiritigenine	42	1.181	**0.637**	0.364	0.195	0.010
	120	1.706	**0.750**	0.431	0.343	0.294
Isoxanthohumaol	138	1.054	**0.638**	0.234	0.308	0.010
	28	1.493	**0.634**	0.430	0.304	0.305

CAPE	575	**0.625**	0.242	0.181	0.243
	22298	0.587	0.230	0.159	**0.700**
Curcumin	548	**0.631**	0.242	0.204	0.010
	23010	0.519	0.264	0.173	**0.623**
Sulforaphane	46	**0.650**	0.361	0.166	0.000
	3203	**0.684**	0.296	0.000	0.000
Cinnamaldehyde	40	**0.684**	0.581	0.383	0.000
	23599	0.580	**0.673**	0.273	0.000
OSL7	35	**0.648**	0.295	0.128	0.000
	1171	**0.702**	0.445	0.191	0.000
C34	187	**0.506**	0.205	0.137	0.512
	10959	0.560	0.216	0.102	**0.903**
JTT705	439	**0.539**	0.391	0.188	0.010
	14650	**0.592**	0.347	0.188	0.000

Table **5.2** shows that for each of the known actives the best scoring SPECS compound scored higher than the best scoring one from the Log P 1000 database. This could be explained by the sole fact that the SPECS set contain a much higher number of compounds than Log P 1000. Consequently the probability is higher to find a good scoring compound.

Regarding the single contributions of the four similarities, the shape similarity (H) seems to have the highest impact on the global score in most of the cases. In four cases (6-shogaol, CAPE, Curcumin, C34) the hydrogen-bond acceptor and in one case the hydrogen-bond donor (N1) similarity made the biggest contribution to the global score. All five compounds are from the SPECS set. The reason why the influence of hydrophobic (DRY) similarity is comparatively low might be the relatively small size of

the compounds. While strong hydrogen-bond similarities can be derived from single donor or acceptor atoms, the hydrophobic potential needs larger apolar surfaces to show a strong impact.

RE-docking with FLAP and GLIDE

To better understand the interactions of the potential inhibitors retrieved by LBVS and SBVS with TLR4/MD-2, a molecular re-docking approach was carried out. In order to narrow down the number of compounds to dock, only molecules were selected which obtained a good score in the LB and the SBVS approaches.

In SBVS, in total, 2038 compounds (26 from Log P 1000 and 2012 from SPECS) obtained a score higher than the cutoff value of 3.245. Since an analogous cutoff value was not available for the LBVS approach, the same number of compounds was chosen here, i.e. the top ranked 26 and 2012 compounds from Log P 1000 and SPECS, respectively. Of the Log P 1000 set 3 common compounds were found in the top ranks of both LBVS and SBVS, while SPECS shared 556 top-ranked compounds. This total number of 559 compounds still seemed large, considering the time-consuming FLAP docking program. For this reason only the top 100 highest scoring ligands were taken for the docking. This selection procedure was found to be in agreement with examples from the literature.[95-98]

In the case of SBVS with Glide, AD4 and VINA:

Among them, 23 compounds from ZINC database, 18 from PM, 16 from JCM and 32 from JRP and AM, being ranked at the top 25% for at least two docking programs at the same time and for one or both conformations, prioritizing a correlation with Glide, were kept for visual cluster rank analysis.

Visually analyzing each most probable cluster for each molecule in the three docking programs, **5** compounds from WORD database have proven to outperform all the others: compounds **146**, **157**, **177**, **208** and **212; 8** compounds from **PM (PM1097, PM1811, PM1779, PM567, PM1090, PM810, PM1758** and **PM1200), 8** from **JCM (MS14, MS20, MS21, MS32, MS35, MS40, MS45, MS49) and 3** from **JRC** and **AM**
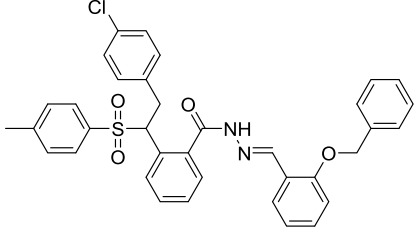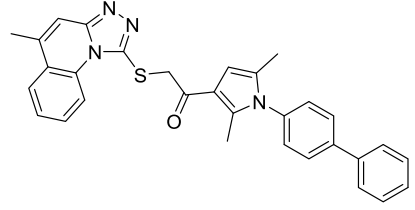
(JRP07, JRp07p and **JRP10).** This analysis was based on the ligand/receptor interactions identified by visual inspection for each compound. Each of the most probable clusters in AutoDock and VINA were found to be relatively similar to the best pose calculated by Glide. For a deeper pose interactions analysis with Glide, a re-docking process was performed. The results obtained by the docking of the known compounds allowed setting a cutoff value. For the screening scores obtained by docking, only the ones with a score equal to or higher than the cutoff value were retained for further biological analysis.

Discussion : Identification of key residues that interact with the screened ligands

Three ligands of the combined Log P 1000 and the SPECS datasets that were obtained by docking had an S-score equal to or higher than the threshold of 1.074, obtained by the best-scoring known inhibitor sulforaphane. The compounds, their 2D description and the respective scores are listed in Table **5.3**. The highest scoring compound is ID-5382 from the Log P 1000 set, with an S-score of 1.231. The two compounds of the SPECS set AG-690/11203225 and AF-399/15128553 obtained a score of 1.114 and 1.074 respectively.

Table 5.3. 2D description and the respective scores from ID-5382, AG-690/11203225 and AF-399/15128553.

Compound	S-score	Structure
ID-5382	1.231	

AG-690/11203225	1.114	
AF-399/15128553	1.074	

Compound ID-5382 is in close contact with the hydrophobic residues Ile46, Leu61, Leu78, Phe121, Ile124, Val135 and Phe151. Compound AG-690/11203225 interacts with the residues Ile52, Phe76, Leu78, Ile80, Val82, Glu92, Phe121, Ile124, Val135 and Ile153. Finally, the hydrophobic partners of AF-399/15128553 are Ile46, Leu61, Ile80, Val82, Leu87, Phe121, Ile124, Tyr131 and Phe151.

The two amino acids are able to build salt bridges with the compounds ID-5382 and AG-690/11203225 due to their sulfonyl group. This would explain why the polar score is significantly higher for these two ligands than for compound AF-399/15128553 which possesses no sulfonyl groups. The latter one only forms hydrogen-bonds between Arg90 and Lys122 and its nitrogen located in the pentacycle. Interactions with Cys133, as the ones reported in the literature[99] could not be observed. As already discussed, is the simulation of covalent bonds not possible in FLAP docking.

From J. R. Pedro and A. Sanz-Marco databases the most of the compounds, stablishing stacking interactions with Phe76 and the CH-π interactions observed are with the side chain of Cys133, Phe151, Phe104 and Leu61. Other interactions observed are hydrophobic with the residues Val24, Ile32, Ile44, Val48, Ile52, Leu78, Ile80, Ile94, Ile117, Phe119, Val135 and Ile153.

From P. Matyus databases, the majority of the compounds stablish π-π with Phe104 and Phe151, also CH-π interactions with Phe76 and Phe121. Other interactions observed are hydrophobic with Ile32, Ile52, Leu61, Ile117, Val135, leu149 and Ile153.

And from J. C. Menendez databases the principal interactions observed are: π-π interactions with Phe76 and Phe151, and also hydrophobix interactions with Ile32, Ile52, Leu61, Leu63, Ile94 and Val135.

In all the cases the docked ligand is located at the entry of the hydrophobic cavity of TLR4/MD-2 in a similar pose. The principle interactions are hydrophobic and polar ones. All the compounds show polar interactions with Arg90 and Lys122. The hydrophobic interactions, however, are more wide spread and not with the same set of amino acids for all the compounds. Arg90 is assumed to participate in interactions with sulforaphane, JTT705, isoxanthohumol, isoliquiritigenin, CAPE and JSH. Lys122 interacts with OSL07 and cinnamaldehyde. Some of the identified side chains are also participating in the interaction with known ligands. Ile80 for example interacts with xanthohumol, JSH, OSL07, cinnamaldehyde and 6-shogaol. The side chains Phe121 and Ile124 interact with all known ligands. Hydrophobic interactions with the ligand are basically with aromatic cycles.

Regarding drug repurposing results, the analysis revealed 5 compounds outperforming the remaining ones: compounds **56**, **146**, **177**, **179** and **208**. Surprisingly, compared to the previous analysis done only with the best Glide pose, compounds **157** and **212** did not show good results in the last analysis. Indeed, having a more wide number of poses in Glide permitted to see, for these two compounds, that the first pose was not part of the most probable cluster, or any cluster at all, for both conformations. Moreover, it has been shown that the most probable clusters for these two compounds were ranked in a low energy position, and with a medium total percentage of interaction against the main residues. Compound **208**, previously revealed in the first analysis, having a good cluster position, was reported to have medium total percentage of interaction against the main residues. Compound **56** revealed having similar problem as compounds **157** and **212**. However, the most probable cluster was ranked in a good position. Compounds **56**, **157**, **208** and **212** were kept as a query for future structure similarity search.

In ascending order of potential prediction, compounds **146**, **177** and **179** outperformed all the compounds. Compounds **146** and **177**, already revealed by the

first cluster analysis, have shown having in each pose, interactions with almost all the main residues. Moreover, in about 50% of the poses, they were able to make 2 hydrogen bonds at the same time, and in about 70% of the poses, able to make 2 salt bridge interactions simultaneously. Regarding compound **179**, it was predicted having the highest affinity potential with all the main residues. It interacts with all the main residues with high affinity, making in 80% of the poses, up to 3 hydrogen bonding and a salt bridge in 50% of the poses. Compounds **146**, **177** and **179** were also kept as queries for future structure similarity search.

Regarding compound **146**, it is known as Diphenoxylate. It is a meperidine congener used as an antidiarrheal, usually in combination with atropine. At high doses, it acts like morphine. Its unesterified metabolite difenoxin has similar properties and is used similarly. It has little or no analgesic activity. According to DrugBank (www.drugbank.ca), it is categorized as: analgesics, opioid, antidiarrheals, antiperistaltic agents, alimentary tract and metabolism, antidiarrheals, intestinal anti-Inflammatory/anti-infective agents, and antipropulsives. Because TLR pathways can be related to inflammatory and microbial pathologies, it can be conceivable that Diphenoxylate could have a certain affinity for TLR4. It has also been shown that Diphenoxylate can regulate NF-κ-B,[100] a protein present downstream in the TLR pathway. Moreover, some studies have proven the binding between morphine and TLR4,[101-103] that could suggest also a conceivable effect of Diphenoxylate to TLR4.

Compound **177** is known as Ono-Rs 411 or Pranlukast. It is a cysteinyl leukotriene receptor-1 antagonist. It antagonizes or reduces bronchospasm caused, principally in asthmatics, by an allergic reaction to accidentally or inadvertently encountered allergens. It is classified as: anti-asthmatic agents, respiratory system, drugs for obstructive airway diseases, leukotriene receptor antagonists, cytochrome P-450 CYP2C9 inhibitors, cytochrome P-450 CYP2C9 inducers, and CYP3A4 inhibitors. Besides, some studies have shown that Pranlukast can inhibit NF-κ-B activation,[104-105] a protein present downstream in the TLR activation pathway. It has also been shown that it indirectly induces cytoplasmic membrane depolarization of Gram-negative

bacteria, promoting *E. coli* outer membrane detachment,[106] which some are recognized by TLR4.

Compound **179** is known as Vemurafenib, a V600 mutant BRAF enzyme inhibitor for the treatment of late-stage melanoma.[107] Vemurafenib inhibits the active form of the kinase,[108-109] firmly anchoring itself in the ATP-binding site. By inhibiting only the active form of the kinase, it selectively inhibits the proliferation of cells with unregulated BRAF, normally those that cause cancer. It is classified as: antineoplastic agents, protein kinase inhibitors, antineoplastic and immunomodulating agents, cytochrome P-450 CYP1A2 inhibitors, cytochrome P-450 CYP1A2 inducers, CYP2D6 inducers, CYP2D6 inducers (strong), and CYP3A4 inhibitors. Up to date, it has been shown that TLR4 and its signaling pathway promote the migration of human melanoma cells,[110-111] but no studies showing an effect from Vemurafenib to TLR4 have been done yet.

All hit structures show a very common scaffold and binding pattern: two hydrophobic moieties separated by a polar linker. The larger hydrophobic part occupies the hydrophobic MD-2 cavity, while the smaller one is placed in the same hydrophobic side region where also one of the lipid A alkyl chains is located in the bound X-ray structure. Key interactions are those stablished with residues Arg90, capable of making salt bridges and hydrogen bonds, Phe121, able to make strong hydrophobic interactions, and situated closely to Phe126, and Tyr131, also able to make hydrogen bonds. These interactions were common for all the compounds, and conferred them a strong predicted binding energy. The polar linker seems to be interacting with two of the positively charged amino acids Arg90 and Lys122 at the entry region of the pocket which have already been described in the literature to interact with known active compounds. The literature reports a covalent interaction between Cys133 of MD-2 and some of the known actives. This observation could be reproduced with some of the known actives. The identified screening hits, however, represent an interesting scaffold for a new class of possible inhibitors for the TLR4/MD-2 complex.

Activity on HEK293 Cells Transfected with *h*TLR4/MD-2.

Then biological testing has been performed in HEK-Blue *h*TLR4 cells to assess the potential of the compounds as agonist or antagonist molecules. Indeed, with molecular docking, it is only possible to predict the binding and the affinity of a molecule for a biological or chemical entity, but it is not possible, or it is very complex and not accurate, to predict the activity of a compound without doing biological assays. Further, if the biological testing reveals to show an activity for the compounds, several strategies will be followed starting by screening wider libraries using structure similarity search, creating pharmacophores, then a new docking protocol using the brand-new compounds.

The ability of molecules to interfere with LPS-triggered TLR4 activation in HEK-Blue *h*TLR4 cells model was investigated. This HEK293 cell line is stably transfected with human TLR4, MD-2, and CD14 genes. In addition, Hek-Blue^{TM} cells stably express a secreted Alkaline Phosphatase (SEAP) produced upon activation of NF-kβ. LPS binding activates TLR4 and NF-kB leading to SEAP secretion, which is detected by an alkaline phosphatase substrate in cell culture media (Figure **5.5**).
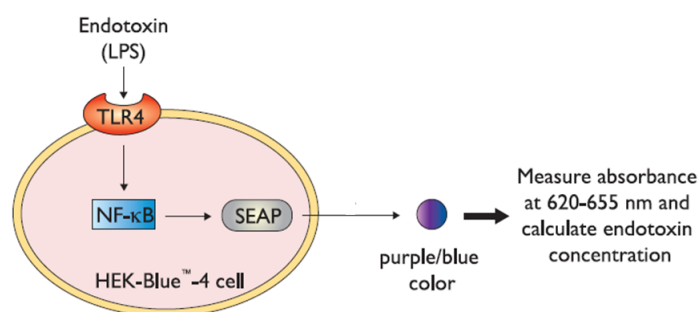
Figure 5.5. Cell-based colorimetric assay for the detection of biological active endotoxin.

In this assay, HEK293 cells transfected with human CD14 and TLR4/MD-2 were treated with increasing concentrations of synthetic molecules and then stimulated with LPS (LPS, 100 ng/ml). TLR4 activation is monitored as SEAP production. The results are normalized to activation by LPS alone and expressed as the mean of percentage ±

SD of three independent experiments. Compounds B (**ID-5382**), F (**MS21**), H (**MS32**), I (**MS35**), X (**PM1090**) and Z (**PM1200**) inhibited TLR4 activation in a dose-dependent way (Figure **5.6**). As a negative control, compounds were tested in Null cell line (InvivoGen), transfected with the same plasmids as HEK-Blue but without TLR4, MD-2, and CD14 genes, and no effect was observed. The toxicities of all compounds are being assaying by MTT assay and no inhibitory effects on cell viability has being observed in the concentration range used for biological characterization.
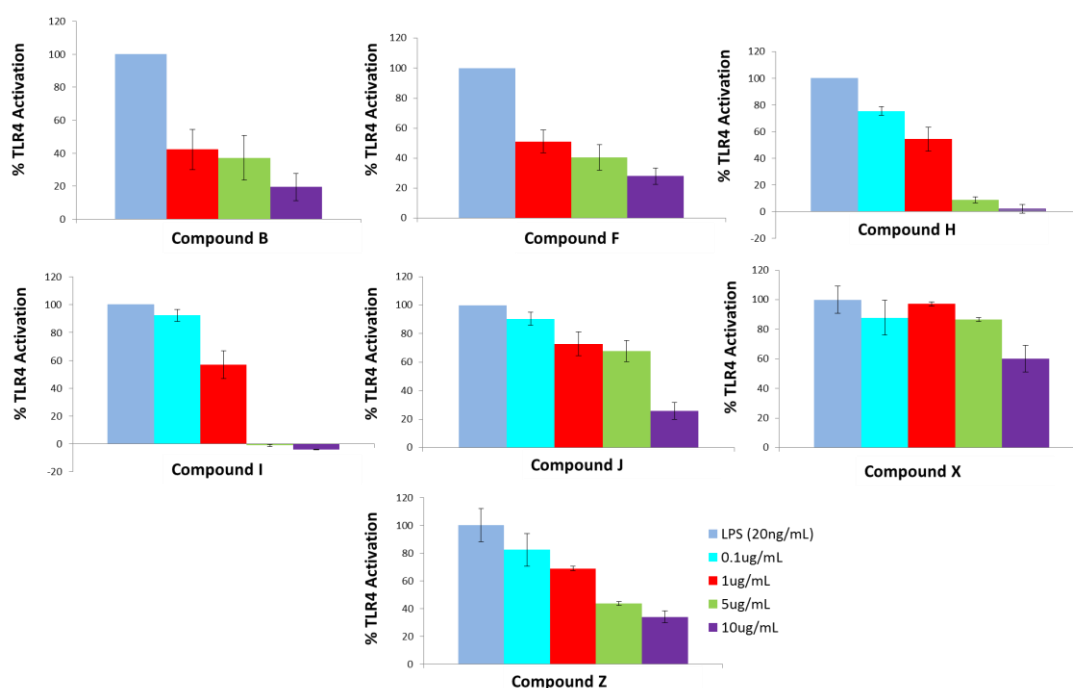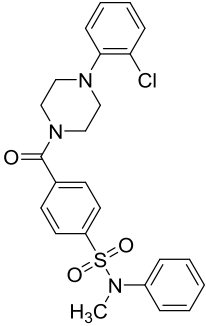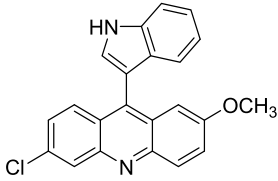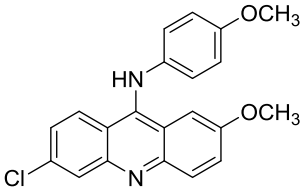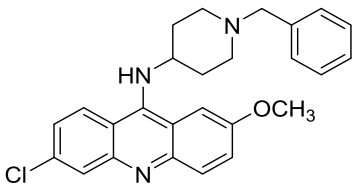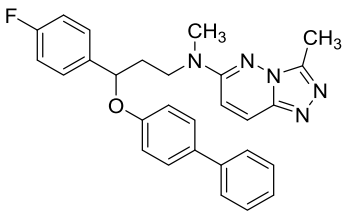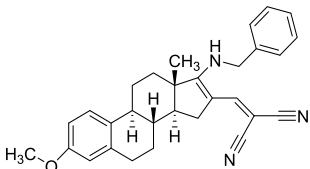
Figure 5.6. Results are expressed in % of TLR4 activation. Positive control (LPS 20ng/mL) represents 100% of activation. Dose-dependent inhbiton of LPS-stimulated TLR4 activation by compounds.

We finally obtained the following possible antagonsit compounds:

Compound	Structure	LogP ChemSketch	LogP Molinspiration
ID-5382 (B)		5.3 (exp)	
MS21 (F)		5.89+/- 0.40	6.331
MS32 (H)		4.20+/- 0.83	6.434
MS35 (I)		4.68+/-0.84	6.728
PM1090 (X)		5.70+/- 0.89	6.063
PM1200 (Z)		6.29+/- 0.45	6.116

5.3   Conclusions

In this work, we have applied virtual screening and computational repositioning strategies for the finding of novel TLR4 modulators. The computational protocol has made use of different conformations of TLR4/MD-2 system, and a deep ligand/receptor analysis, including ligand-based and structure-based virtual screening, leading to a robust approach for the final identification of 7 possible antagonist compounds: Compounds B (**ID-5382**), F (**MS21**), H (**MS32**), I (**MS35**), X (**PM1090**) and Z (**PM1200**) inhibited TLR4 activation in a dose-dependent way as putative of TLR modulators. The identified screening hits, however, represent interesting scaffolds for a new class of possible inhibitors for the TLR4/MD-2 complex.

5.4   Materials and Methods

Computational Methods

Library Preparation

Importation. All the databases were saved as a SD File and imported in Maestro software (Schrodinger, version 10.4),[112] which is an all-purpose molecular modeling environment. During the importation process, the chirality and the atom type of each compound has been checked.

Ligand Preparation using LigPrep. LigPrep (Shrodinger, version 3.6)[113] is a program specialized in preparing all-atom 3D structure of drug like molecules, was used for many purposes: to refine the geometry of the ligands imported from the databases; to generate accurate, energy minimized 3D molecular structures; to expand tautomeric, ring conformation, and stereoisomers in order to produce broad chemical and structural diversity from each input structure and to predict protonation states. The 3D structures were minimized using OPLS 2005;[114] to generate ionization states, Epik[115-117] was used, in order to simulate the physiological pH. In many cases, the compounds contain water molecules or ions, these extra molecules were removed with Desalt option. The generating tautomer options were also used in order to generate up to 8 tautomers per input structure. Regarding the setting stereoisomer options, the choice

of retaining the specified chiralities to keep this information from the input file and fixed these chiralities for the entire calculation has been made. The number of stereoisomers generated was limited up to 32 per ligand. From a 2D structure, it is not immediately obvious which ring conformations give the lowest energy or are preferred for binding to an active site. Therefore it was decided to generate one low energy ring conformation per ligand with LigPrep. The final output was in Maestro format to keep the total information calculated for all the compounds. For the virtual screening, the compounds were selected according to their molecular weight and their lipophilicity, between 300 Da and 700 Da, and between 4 and 6 respectively, using the property calculation tool from the Maestro software.

Protein Preparation

In the case of the agonist conformation of the TLR4/MD-2 monomer, 3D coordinates from TLR4/MD-2 heterodimer were obtained from the PDB (PDB-ID: 3FXI).[44] By contrast, in the case of the antagonist conformation, since the full crystallographic structure of the TLR4/MD-2 complex is not available, a model built by us was used. This model was built using the human MD-2 protein in antagonist conformation (PDB-ID: 2E59)[45] superimposed onto the MD-2 subunit of the agonist full complex (PDB-ID: 3FXI chain C) through PyMOL. Then, coordinates from the TLR4 chain of the 3FXI adjacent to the superimposed MD-2 (PDB-ID: 3FXI chain A) and the superimposed MD-2 in antagonist conformation were retained, forming the TLR4/MD-2 monomer in antagonist conformation. Finally, both agonist and antagonist the structures were subjected to 10.000 cycles of steepest descent energy minimization under the Amber force field via Maestro (see Chapter 3). Also PDB-ID: 2E56 were used to consider different antagonist conformation of MD-2.

Receptor Grid Preparation

Glide. For preparing the receptor grids for the two protein conformations, Glide software (Schrodinger, version 6.9) was used.[86-88] All the parameters from the software

were kept at their default values. We only determined where the scoring grids will be positioned and their sizes. The coordinates of the box were set up to fully contain *E. coli* LPS. Glide software uses two "boxes" that can be parametrized to organize the calculation: the inner box, which can be monitored in the advanced panel, and where the ligand center is allowed to move within that box during the site point search; and the outer box, which is the box within all the ligand atoms must be contained. Its size is function of the inner box, and the inner box has to be included within the outer box. For the inner box, the center was set up at residue serine 120 and the lengths of the boxes for both protein conformations were the following ones: 33 Å in X, 40 Å in Y and 35 Å in Z. For the outer box, 10 Å has been chosen, that is to say 10 Å bigger than the inner box (43 Å in X, 50 Å in Y and 45 Å in Z).

AutoDock, FLAP and VINA. As the receptor grids were already set up with Glide, the same grids have been chosen for the softwares. Glide coordinates were kept for VINA, but were converted in AutoDock coordinates using scaling calculation tool. In the case of FLAP, The pockets of MD-2 were identified and defined by FLAP's pocket search algorithm.

Docking

Structure Based Virtual Screening (SBVS) with FLAP

The FLAP software explicitly distinguishes between the so called *SBVS* method and docking.[118-119] While in FLAP docking is primarily used for pose prediction and a more precise quantification of binding energies, SBVS is a tool for large-scale virtual screenings. Even though docking is often used as a structure based virtual screening technique,[120] the term *SBVS* will hereafter refer only to FLAP's correspondent screening program.

The SBVS program first creates MIFs of the receptor's binding site. During screening, the MIFs of the ligand are compared with those of the binding site. Time consuming calculations describing each atom-atom interaction are not needed here.

One downside of this method is that there is no energetic penalty for atom clashing with the target. In some scenarios however, this might even be an advantage, since it overcomes the rigidity of the target to some extent.

The SBVS in FLAP was performed on the 3D structure of the human co-receptor MD-2. The structure was obtained from the PDB (PDB code: 2E56)[45] and the MOE software was used to prepare the protein by removing water molecules, adding hydrogens and missing atoms and side chains.[73] The optimized structure was loaded into FLAP and the *Search for pockets* function was used to define the binding area. The results are then treated in analogy to the LBVS approach.

SBVS with Glide

The molecules were subjected to a grid-based ligand docking with energetics (Glide, Schrodinger, version 6.9)[86-88] using the Virtual Screening Workflow protocol. It is designed to run an entire sequence of jobs for screening large collections of compounds against one or more targets. However, as the compounds and the grids had already been prepared, in this case, only the docking steps of the program have been used. The compound files and the receptor grid files were imported into the Virtual Screening Workflow program. Regarding the docking step parameters, Epik state penalties for docking were used, and the non-polar part of the ligand potential were soften by scaling the Van der Waals radii of ligand atoms with small partial charges. To do so, the scaling factor was 0.80, and the partial charge cutoff was 0.15. The full workflow includes three docking stages, each step differing from the preceding step in the amount of time taken to dock each molecule and the scoring system used to evaluate each pose. The first stage performs HTVS (High Throughput Virtual Screening) docking. The ligands that are retained are then passed to the next stage, which performs SP (Standard Precision) docking. The survivors of this stage are passed onto the third stage, which performs XP (eXtra Precision) docking, a more powerful and discriminating procedure.

The Dock flexibility method was used for HTVS, SP and XP dockings allowing us to penalize non-planar conformation for amide bonds. A post-docking minimization was also performed, as well as constraints for the docking stages. One pose per compound state was generated and 100 % of the best compounds that passed the HTVS, SP and XP docking have been kept. For HTVS and SP docking, all states have been retained, but only the best scoring state for the XP docking.

Ligand re-docking using Glide

The shortlisted molecules were submitted to a re-docking procedure using Glide. All the parameters were kept as mentioned in the docking paragraph using Glide, except for the docking poses which were set to 50 per molecule.

Molecular re-docking using FLAP

The FLAP software implements a fragmentation-based docking algorithm, called FLAPdock, which works as follows. MIFs are calculated for the target binding site, in a similar manner to the SBVS approach but with more points to describe the site in more detail (Reference manual for FLAP 2.0, © 2014 Molecular Discovery Ltd). A set of ligand conformations is generated using a stochastic search and a customized implementation of the MM3 force field[121] with a cutoff of 30 kcal mol^{-1} to remove high energy and duplicate conformations respectively. The ligands are then split into fragments with only 1-3 rotatable bonds. For each fragment conformation, GRID MIFs are calculated. The first fragment is docked into the binding site and the best scoring solutions, according to the global S-Score, are retained for the next iteration. In the next step, the next fragment, is attached to the first one and scored in the same way. The S-Score is a scoring function that includes terms from the GRID MIF similarities (Hydrogen-bonding and hydrophobic interactions as well as shape matching), Lennard-Jones and electrostatic interactions. It was validated, amongst other targets, on those of the Astex and DUD datasets.[119, 122-123] In each iteration the best scoring solutions are kept and filtered by RMS clustering. Once the reconstruction of the ligand has finished,

the final pose can be optionally optimized by minimization and the final score is recalculated. The benefit of FLAPdock towards the SBVS method lies in the more detailed chemical interactions that are considered for docking and the respect of steric clashes that are not regarded in the SBVS method. In order to obtain score reference values, a set of known MD-2 inhibitors was docked, followed by the docking of compounds from the Log P 1000 and the SPECS dataset.

Biological characterization

HEK-Blue TLR4 assay. HEK-Blue-TLR4 cells (InvivoGen, Toulouse, France) and parental cell line HEK-Blue Null 2 (InvivoGen) were HEK-Blue cells were used to test the agonist or antagonist effect of different compounds. This cell line expresses TLR4, MD-2 and CD14 and do not express any other TLR. The activation of TLR4 leads to the expression of SEAP, a protease that enzimatically hydrolyze a molecule present in the media. The amount of hydrolized molecule can be mesured using colorimetric methods. These cells were cultured according to manufacturer's instructions. Briefly, cells were cultured in DMEM high glucose medium supplemented with 10% fetal bovine serum (FBS), 1% glutamine, 1% penicillin/ streptomycin, 1X Normocin (InvivoGen). Experiments were performed when 70-80% of confluence was reached. Cells were detached by the use of PBS, tapping the flask and the cell concentration was estimated using. Four different compound concentrations were used: 0.1, 1, 5 and 10 µg/mL. 20 µL of compound dilution were added in a 96-well plate, in triplicate (3 wells for each concentration), seeded in multiwall plate at a density of 2 x 10^4 cells/well in 200 µl. LPS was used as positive control (20ng/mL final concentration) and PBS 1x was used as negative control. Cells were detached using 4mL of PBS and 140.000 cell/mL solution was prepared using Detections Media. 180 µL of this solution were added into each well (25.000 cells/well). After a 30 min incubation, 20 µL of LPS solution were added in each well (final LPS concentration: 20ng/mL) (LPS was diluted in PBS as well). Plates were incubated for 16 h in the dark at 37 °C, 95% of humidity and 5% of CO_2 and then, the plate reading was assessed using a spectrophotometer at 620 nm. The

results were normalized with positive control (LPS alone) and expressed as the mean of percentage ± SD of at least three independent experiments.

5.5   Annex V
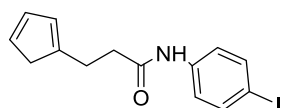
Compound 152
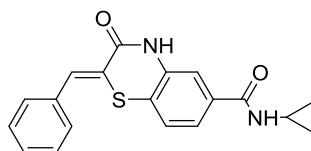
Compound 568
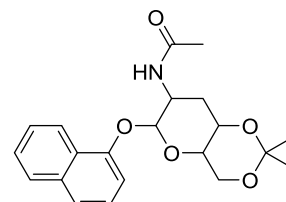
Compound 383
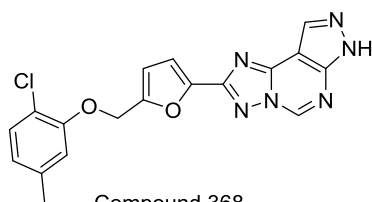
Compound 368
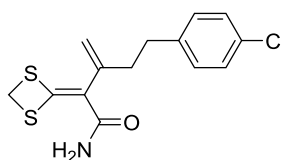
Compound 492
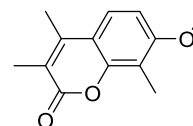
Compound 42
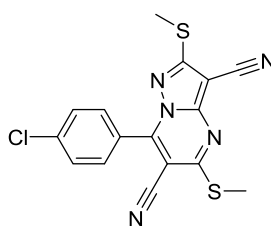
Compound 138
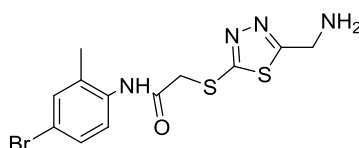
Compound 575
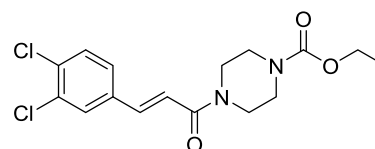
Compound 548
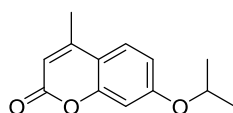
Compound 46
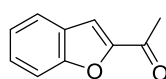
Compound 40
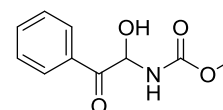
Compound 35
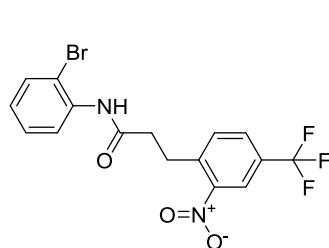
Compound 187
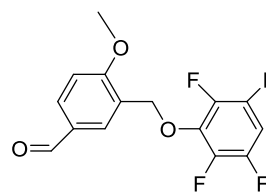
Compound 439

Annex Figure 5.1. Top scoring compounds obtained by LBVS on th Log P 1000 database
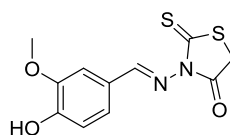
Compound 481
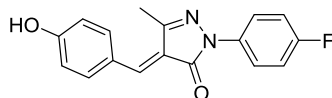
Compound 19907
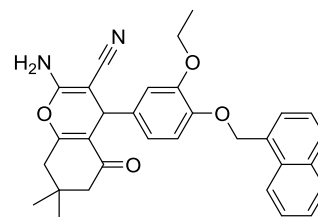
Compound 20513
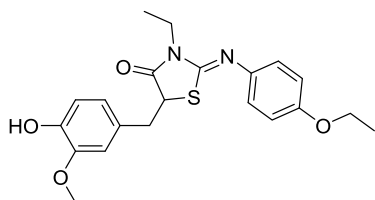
Compound 20700
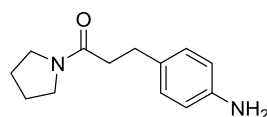
Compound 21315
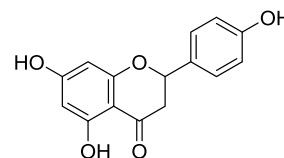
Compound 120
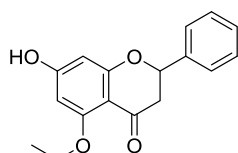
Compound 28
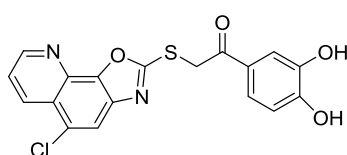
Compound 22298
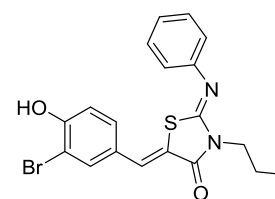
Compound 23010
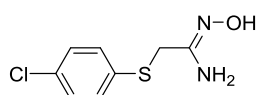
Compound 3203
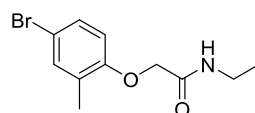
Compound 23599
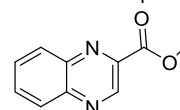
Compound 1171
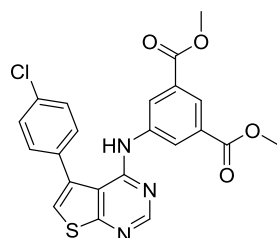
Compound 10959
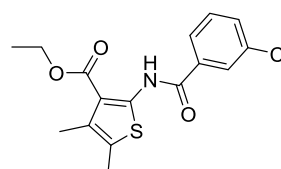
Compound 14650

Annex Figure 5.2. Top scoring compounds obtained by LBVS on the SPECS database.

Table 5.1. 2D Chemical structure of predicted TLR4 modulators identified by computational drug repurposing, and kept for future structure similarity search.
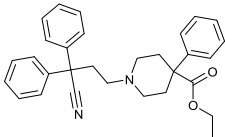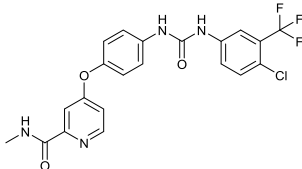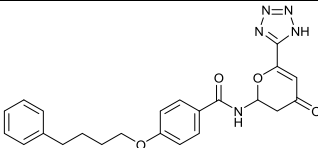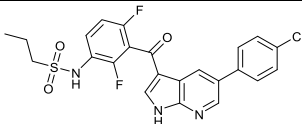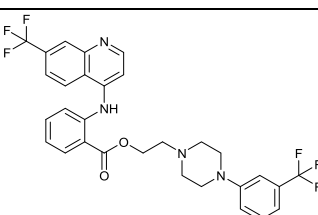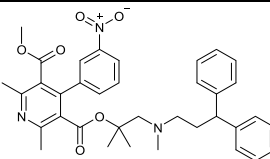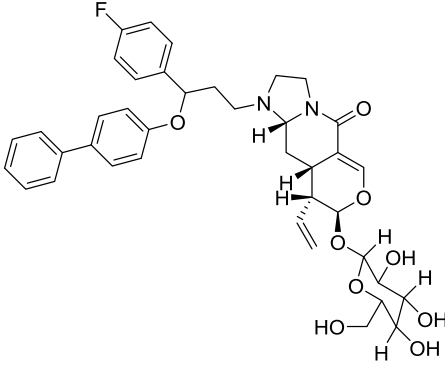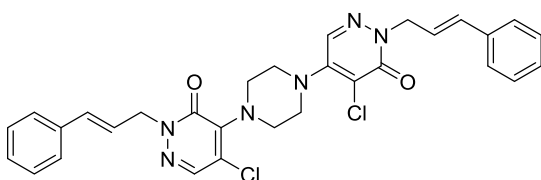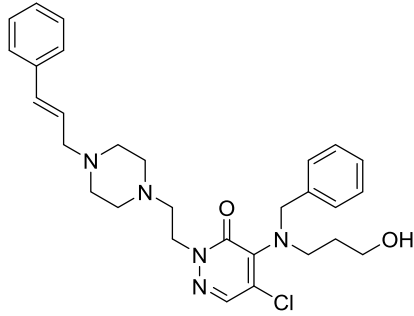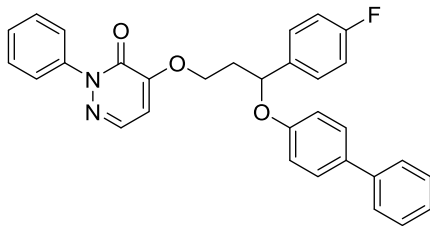
Compound number	ZINC ID	Usual/Commercial name	2D structure
146	3830716	Diphenoxylate	
157	1493878	Sorafenib	
177	15919406	Ono-Rs 411	
179	52509366	Zelboraf	
208	53073961	Antrafenine	
212	19685790	Lercanidipine	

Table 5.2. 2D Chemical structure from PM databases obtained from SBVS.

Number	Compound	Structure
1	PM1097_p_R/1097	
2	PM1811	
3	PM1779	
4	PM567S ó R	
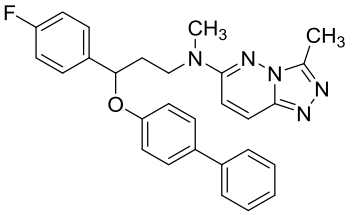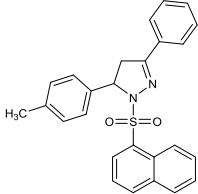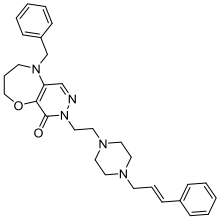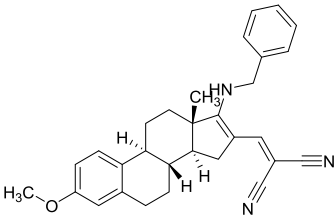
5	PM1090	
6	PM810	
7	PM1758	
8	PM1200	

Table 5.2. 2D Chemical structure from JCM

databases obtained from SBVS.

Number	Compound	Structure
1	MS_35/35p	
2	MS_29	
3	MS_40	
4	MS_34	
5	MS_31	
6	MS_22	
7	MS_45	
8	MS_21	
9	MS_32	

10	MS_26	
11	MS_14	
12	MS_49	
13	MS_37	
14	MS_46	
15	MS_20	

Final compounds:
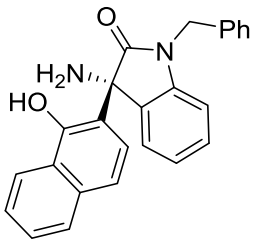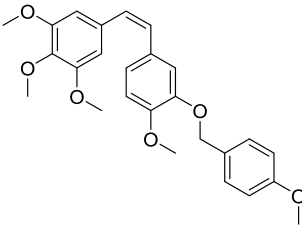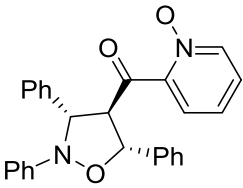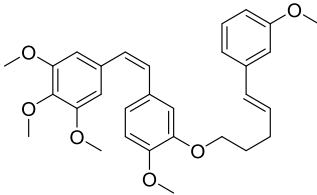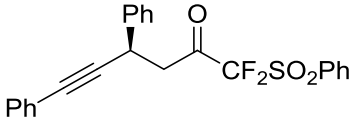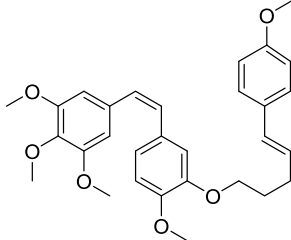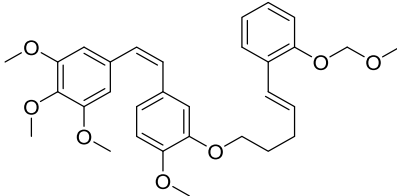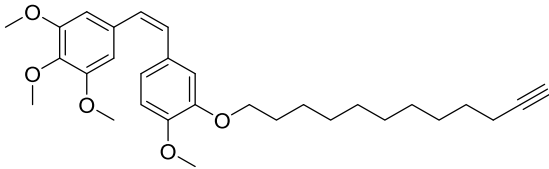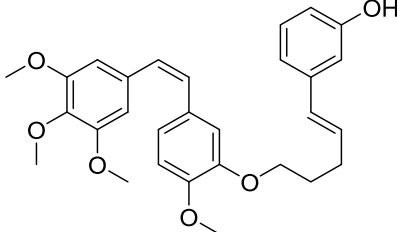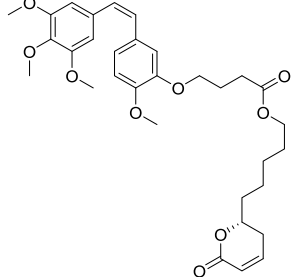
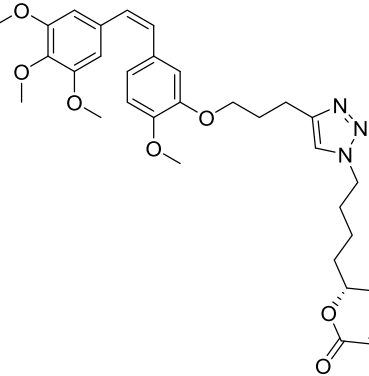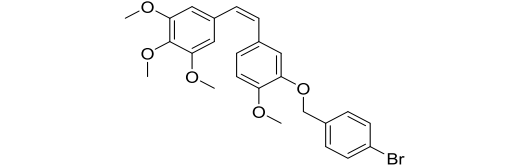Quinoline family: MS-**14**, MS-**20**
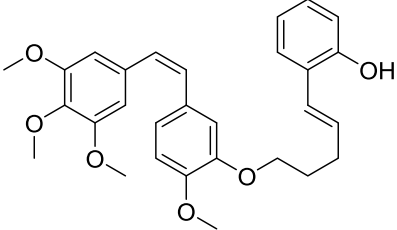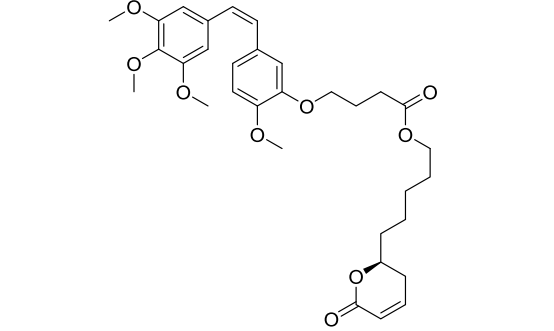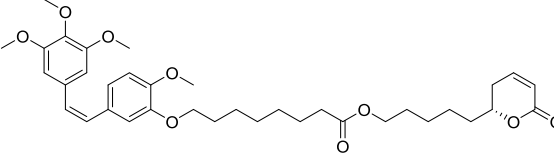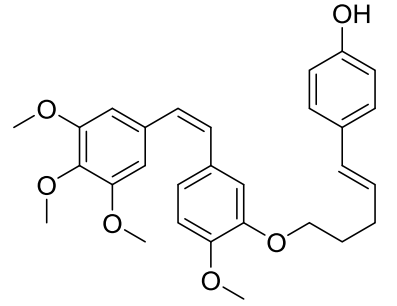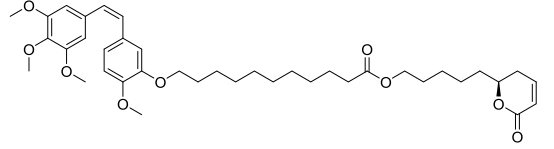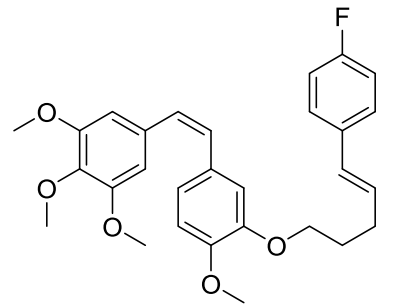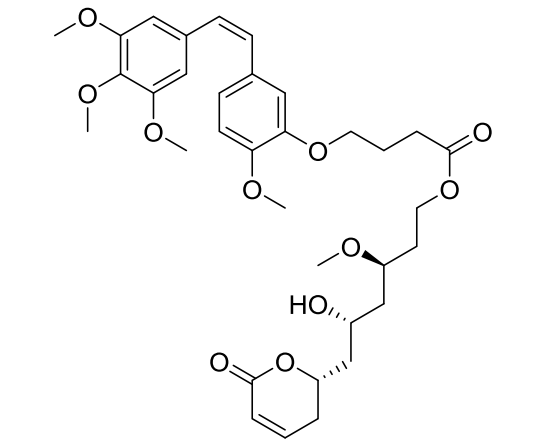
Quinazoline family: MS-**40**, MS-**45**, MS-**49**
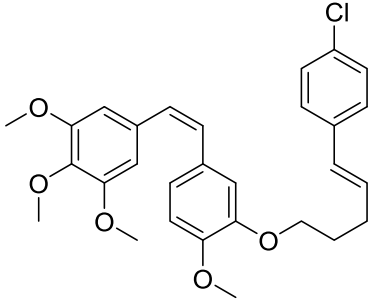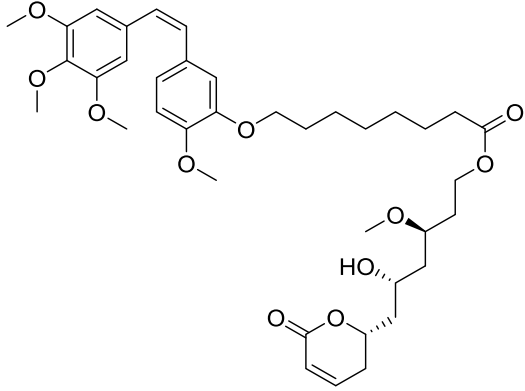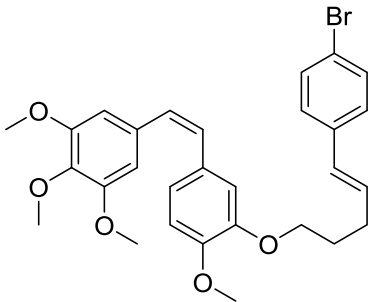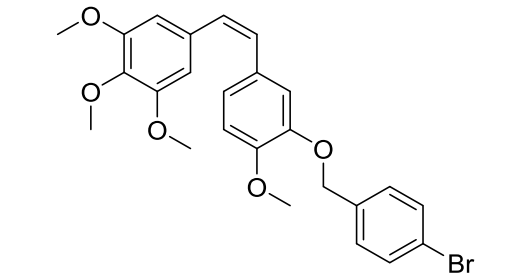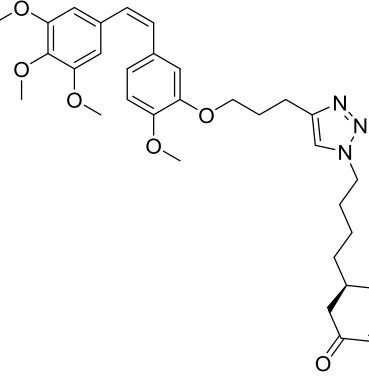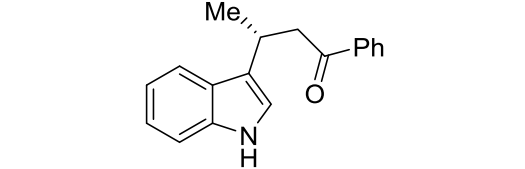
Acridine family: MS-**32** and MS-**21**
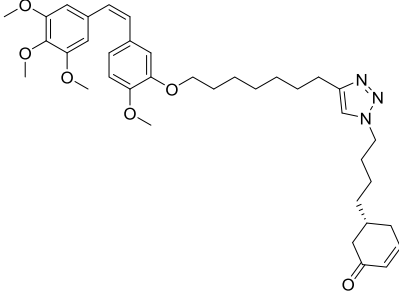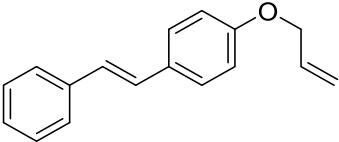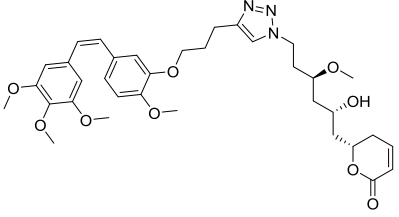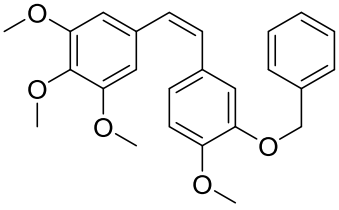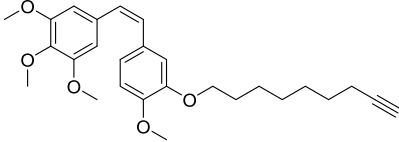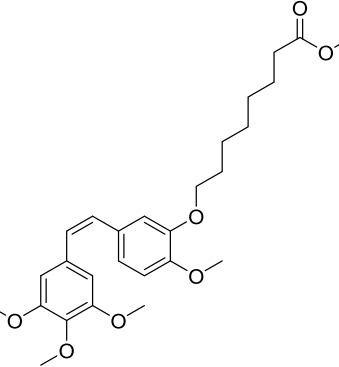
MS-**35**, MS-**29**=MS-**22**, MS-**26**, MS-**31**, MS-**34**, MS-**37**

Table 5.3. 2D Chemical structure from J. R. Pedro and A. Sanz-Marco databases obtained from SBVS.

Name	Structure	Name	Structure
JRP07		**AM15**	
JRP10		**AM18**	
JRP18		**AM19**	
AM20 **momo**		**AM54**	
AM21		**AM57**	

AM20		AM58	
		AM59	
AM22		AM62	
AM23		AM65	

258

AM24		AM66	
AM25		AM72	
AM40		AM16	
AM41		JRP01	

AM42		AM08	
AM48		AM14	
AM53		AM71	

Bibliography:

1.   Mifsud, E. J.; Tan, A. C.; Jackson, D. C., TLR Agonists as Modulators of the Innate Immune Response and Their Potential as Agents Against Infectious Disease. *Front. Immunol.* **2014,** *5*, 79.

2.   Joosten, L. A.; Abdollahi-Roodsaz, S.; Dinarello, C. A.; O'Neill, L.; Netea, M. G., Toll-like receptors and chronic inflammation in rheumatic diseases: new developments. *Nat. Rev. Rheumatol.* **2016,** *12* (6), 344-57.

3.   Gooshe, M.; Aleyasin, A. R.; Abdolghaffari, A. H.; Rezaei, N., Toll like receptors: a new hope on the horizon to treat multiple sclerosis. *Expert Rev. Clin. Immunol.* **2014,** *10* (10), 1277-9.

4.   Nelson, M. H.; Diven, M. A.; Huff, L. W.; Paulos, C. M., Harnessing the Microbiome to Enhance Cancer Immunotherapy. *J. Immunol. Res.* **2015,** *2015*, 368736.

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

6.   Gambuzza, M. E.; Sofo, V.; Salmeri, F. M.; Soraci, L.; Marino, S.; Bramanti, P., Toll-like receptors in Alzheimer's disease: a therapeutic perspective. *CNS Neurol. Disord. Drug Targets* **2014,** *13* (9), 1542-58.

7.   Veber, D. F.; Johnson, S. R.; Cheng, H. Y.; Smith, B. R.; Ward, K. W.; Kopple, K. D., Molecular properties that influence the oral bioavailability of drug candidates. *J. Med. Chem.* **2002,** *45* (12), 2615-23.

8.   Chan, M.; Hayashi, T.; Mathewson, R. D.; Nour, A.; Hayashi, Y.; Yao, S.; Tawatao, R. I.; Crain, B.; Tsigelny, I. F.; Kouznetsova, V. L.; Messer, K.; Pu, M.; Corr, M.; Carson, D. A.; Cottam, H. B., Identification of substituted pyrimido[5,4-b]indoles as selective Toll-like receptor 4 ligands. *J. Med. Chem.* **2013,** *56* (11), 4206-23.

9.   Neve, J. E.; Wijesekera, H. P.; Duffy, S.; Jenkins, I. D.; Ripper, J. A.; Teague, S. J.; Campitelli, M.; Garavelas, A.; Nikolakopoulos, G.; Le, P. V.; de, A. L. P.; Pham, N. B.; Shelton, P.; Fraser, N.; Carroll, A. R.; Avery, V. M.; McCrae, C.; Williams, N.; Quinn, R. J., Euodenine A: a small-molecule agonist of human TLR4. *J. Med. Chem.* **2014,** *57* (4), 1252-75.

10.   Shanmugam, A.; Rajoria, S.; George, A. L.; Mittelman, A.; Suriano, R.; Tiwari, R. K., Synthetic Toll like receptor-4 (TLR-4) agonist peptides as a novel class of adjuvants. *PLoS One* **2012,** *7* (2), e30839.

11.   Park, S. J.; Kang, S. H.; Kang, Y. K.; Eom, Y. B.; Koh, K. O.; Kim, D. Y.; Youn, H. S., Inhibition of homodimerization of toll-like receptor 4 by 4-oxo-4-(2-oxo-oxazolidin-3-yl)-but-2-enoic acid ethyl ester. *Int. Immunopharmacol.* **2011,** *11* (1), 19-22.

12.   Jin, G. H.; Li, H.; An, S.; Ryu, J. H.; Jeon, R., Design, synthesis and activity of benzothiazole-based inhibitors of NO production in LPS-activated macrophages. *Bioorg. Med. Chem. Lett.* **2010,** *20* (21), 6199-202.

13.   Kawamoto, T.; Ii, M.; Kitazaki, T.; Iizawa, Y.; Kimura, H., TAK-242 selectively suppresses Toll-like receptor 4-signaling mediated by the intracellular domain. *Eur. J. Pharmacol.* **2008,** *584* (1), 40-8.

14.   Chavez, S. A.; Martinko, A. J.; Lau, C.; Pham, M. N.; Cheng, K.; Bevan, D. E.; Mollnes, T. E.; Yin, H., Development of beta-amino alcohol derivatives that inhibit Toll-like receptor 4 mediated inflammatory response as potential antiseptics. *J. Med. Chem.* **2011,** *54* (13), 4659-69.

15.   Lyne, P. D., Structure-based virtual screening: an overview. *Drug Discov. Today* **2002,** *7* (20), 1047-1055.

16.   Schneider, G., Virtual screening: an endless staircase? *Nat. Rev. Drug Discov.* **2010,** *9* (4), 273-6.

17.   Haga, J. H.; Ichikawa, K.; Date, S., Virtual Screening Techniques and Current Computational Infrastructures. *Curr. Pharm. Des.* **2016,** *22* (23), 3576-84.

18.   Zhu, T.; Cao, S.; Su, P. C.; Patel, R.; Shah, D.; Chokshi, H. B.; Szukala, R.; Johnson, M. E.; Hevener, K. E., Hit identification and optimization in virtual screening: practical recommendations based on a critical literature analysis. *J. Med. Chem.* **2013,** *56* (17), 6560-72.

19.   Lionta, E.; Spyrou, G.; Vassilatis, D. K.; Cournia, Z., Structure-based virtual screening for drug discovery: principles, applications and recent advances. *Curr. Top. Med. Chem.* **2014,** *14* (16), 1923-38.

20.   Yan, X.; Liao, C.; Liu, Z.; Hagler, A. T.; Gu, Q.; Xu, J., Chemical Structure Similarity Search for Ligand-based Virtual Screening: Methods and Computational Resources. *Curr. Drug Targets* **2016,** *17* (14), 1580-1585.

21.   Sheng, C.; Dong, G.; Miao, Z.; Zhang, W.; Wang, W., State-of-the-art strategies for targeting protein-protein interactions by small-molecule inhibitors. *Chem. Soc. Rev.* **2015,** *44* (22), 8238-59.

22.   Perez-Regidor, L.; Zarioh, M.; Ortega, L.; Martin-Santamaria, S., Virtual Screening Approaches towards the Discovery of Toll-Like Receptor Modulators. *Int. J. Mol. Sci.* **2016,** *17* (9).

23.   Joce, C.; Stahl, J. A.; Shridhar, M.; Hutchinson, M. R.; Watkins, L. R.; Fedichev, P. O.; Yin, H., Application of a novel in silico high-throughput screen to identify selective inhibitors for protein-protein interactions. *Bioorg. Med. Chem. Lett.* **2010,** *20* (18), 5411-3.

24.   http://www.enamine.net, **(accessed on 6th June 2017)**.

25.   Svajger, U.; Brus, B.; Turk, S.; Sova, M.; Hodnik, V.; Anderluh, G.; Gobec, S., Novel toll-like receptor 4 (TLR4) antagonists identified by structure- and ligand-based virtual screening. *Eur. J. Med. Chem.* **2013,** *70*, 393-9.

26.   Irwin, J. J.; Sterling, T.; Mysinger, M. M.; Bolstad, E. S.; Coleman, R. G., ZINC: a free tool to discover chemistry for biology. *J. Chem. Inf. Model.* **2012,** *52* (7), 1757-1768.

27.   Chong, C. R.; Sullivan, D. J., Jr., New uses for old drugs. *Nature* **2007,** *448* (7154), 645-6.

28.   Shibayama, S.; Tanikawa, K.; Fujimoto, R.; Kimura, H., Effect of mergers and acquisitions on drug discovery: perspective from a case study of a Japanese pharmaceutical company. *Drug Discov. Today* **2008,** *13* (1-2), 86-93.

29.   Langedijk, J.; Mantel-Teeuwisse, A. K.; Slijkerman, D. S.; Schutjens, M. H., Drug repositioning and repurposing: terminology and definitions in literature. *Drug Discov. Today* **2015,** *20* (8), 1027-34.

30.   Allarakhia, M., Open-source approaches for the repurposing of existing or failed candidate drugs: learning from and applying the lessons across diseases. *Drug Des. Dev. Ther.* **2013,** *7*, 753-766.

31.   Allison, M., NCATS launches drug repurposing program. *Nat. Biotechnol.* **2012,** *30* (7), 571-572.

32.	Marusina, K.; Welsch, D. J.; Rose, L.; Brock, D.; Bahr, N., The CTSA Pharmaceutical Assets Portal–a public–private partnership model for drug repositioning. *Drug Discov. Today* **2012,** *8* (3), 77-83.

33.	Murteira, S.; Ghezaiel, Z.; Karray, S.; Lamure, M., Drug reformulations and repositioning in pharmaceutical industry and its impact on market access: reassessment of nomenclature. *J. Mark Access Health Policy* **2013,** *1*.

34.	Novac, N., Challenges and opportunities of drug repositioning. *Trends Pharmacol. Sci.* **2013,** *34* (5), 267-272.

35.	Ashburn, T. T.; Thor, K. B., Drug repositioning: identifying and developing new uses for existing drugs. *Nat. Rev.* **2004,** *3* (8), 673-683.

36.	Ekins, S.; Williams, A. J.; Krasowski, M. D.; Freundlich, J. S., In silico repositioning of approved drugs for rare and neglected diseases. *Drug Discov. Today* **2011,** *16* (7), 298-310.

37.	Muthyala, R., Orphan/rare drug discovery through drug repositioning. *Drug Discov.Today* **2012,** *8* (3), 71-76.

38.	Sardana, D.; Zhu, C.; Zhang, M.; Gudivada, R. C.; Yang, L.; Jegga, A. G., Drug repositioning for orphan diseases. *Brief. Bioinform.* **2011,** *12* (4), 346-356.

39.	Aubé, J., ACS Med. Chem. Lett. *ACS medicinal chemistry letters* **2012,** *3* (6), 442-444.

40.	Fagan, S. C., Drug Repurposing for Drug Development in Stroke. *Pharmacother.* **2010,** *30* (7P2), 51S-54S.

41.	Singhal , S.; Mehta , J.; Desikan , R.; Ayers , D.; Roberson , P.; Eddlemon , P.; Munshi , N.; Anaissie , E.; Wilson , C.; Dhodapkar , M.; Zeldis , J.; Siegel , D.; Crowley , J.; Barlogie , B., Antitumor Activity of Thalidomide in Refractory Multiple Myeloma. *N. Engl. J. Med.* **1999,** *341* (21), 1565-1571.

42.	Yang, T.-J.; Yang, T.-S.; Liang, H.-M., Thalidomide and congenital abnormalities. *The Lancet* **1963,** *281* (7280), 552-553.

43.	Klett, J.; Reeves, J.; Oberhauser, N.; Perez-Regidor, L.; Martin-Santamaria, S., Modulation of toll-like receptor 4. Insights from x-ray crystallography and molecular modeling. *Curr. Top Med. Chem.* **2014,** *14* (23), 2672-83.

44.	Park, B. S.; Song, D. H.; Kim, H. M.; Choi, B. S.; Lee, H.; Lee, J. O., The structural basis of lipopolysaccharide recognition by the TLR4–MD-2 complex. *Nature* **2009,** *458* (7242), 1191-1195.

45.	Ohto, U.; Fukase, K.; Miyake, K.; Satow, Y., Crystal structures of human MD-2 and its complex with antiendotoxic lipid IVa. *Science* **2007,** *316* (5831), 1632-1634.

46.	Staderini, M.; Cabezas, N.; Bolognesi, M. L.; Menendez, J. C.; Carlos, J., *Solvent- and chromatography-free amination of π-deficient nitrogen heterocycles under microwave irradiation. A fast, efficient and green route to 9-aminoacridines, 4-aminoquinolines and 4-aminoquinazolines and its application to the synthesis of the drugs amsacrine and bistacrine*. Elsevier: Kidlington, ROYAUME-UNI, 2013; Vol. 69.

47.	Sanz-Marco, A.; Garcia-Ortiz, A.; Blay, G.; Pedro, J. R., Catalytic asymmetric conjugate addition of terminal alkynes to beta-trifluoromethyl alpha,beta-enones. *Chem. Commun.* **2014,** *50* (18), 2275-8.

48.	Blay, G.; Hernandez-Olmos, V.; Pedro, J. R., The construction of quaternary stereocenters by the Henry reaction: circumventing the usual reactivity of substituted glyoxals. *Chemistry (Weinheim an der Bergstrasse, Germany)* **2011,** *17* (13), 3768-73.

49.   Barroso, S.; Blay, G.; Al-Midfa, L.; Munoz, M. C.; Pedro, J. R., Copper(II)-bis(oxazoline) catalyzed asymmetric Diels-Alder reaction with alpha'-arylsulfonyl enones as dienophiles. *J. Org. Chem.* **2008,** *73* (16), 6389-92.

50.   Holmquist, M.; Blay, G.; Munoz, M. C.; Pedro, J. R., Enantioselective addition of nitromethane to 2-acylpyridine N-oxides. Expanding the generation of quaternary stereocenters with the Henry reaction. *Org. Lett.* **2014,** *16* (4), 1204-7.

51.   Blay, G.; Hernandez-Olmos, V.; Pedro, J. R., Enantioselective Henry addition of methyl 4-nitrobutyrate to aldehydes. Chiral building blocks for 2-pyrrolidinones and other derivatives. *Org. Lett.* **2010,** *12* (13), 3058-61.

52.   Blay, G.; Incerti, C.; Muñoz, M. C.; Pedro, J. R., Enantioselective LaIII-pyBOX-Catalyzed Nitro-Michael Addition to (E)-2-Azachalcones. *Eur. J. Org. Chem.* **2013,** *2013* (9), 1696-1705.

53.   Blay, G.; Fernandez, I.; Munoz, M. C.; Pedro, J. R.; Recuenco, A.; Vila, C., Enantioselective synthesis of tertiary alcohols through a zirconium-catalyzed Friedel-Crafts alkylation of pyrroles with alpha-ketoesters. *J. Org. Chem.* **2011,** *76* (15), 6286-94.

54.   Blay, G.; Fernandez, I.; Monleon, A.; Pedro, J. R.; Vila, C., Enantioselective zirconium-catalyzed Friedel-Crafts alkylation of pyrrole with trifluoromethyl ketones. *Org. Lett.* **2009,** *11* (2), 441-4.

55.   Barroso, S.; Blay, G.; Muñoz, M. C.; Pedro, J. R., Highly Enantio- and Diastereoselective Inverse Electron Demand Hetero-Diels–Alder Reaction using 2-Alkenoylpyridine N-Oxides as Oxo-Heterodienes. *Adv. Synth. Catal.* **2009,** *351* (1-2), 107-111.

56.   Sanz-Marco, A.; Blay, G.; Munoz, M. C.; Pedro, J. R., Highly enantioselective copper(I)-catalyzed conjugate addition of 1,3-diynes to alpha,beta-unsaturated trifluoromethyl ketones. *Chem. Commun.* **2015,** *51* (43), 8958-61.

57.   Blay, G.; Fernandez, I.; Pedro, J. R.; Vila, C., Highly enantioselective Friedel-Crafts alkylations of indoles with simple enones catalyzed by zirconium(IV)-BINOL complexes. *Org. Lett.* **2007,** *9* (13), 2601-4.

58.   Barroso, S.; Blay, G.; Munoz, M. C.; Pedro, J. R., Highly enantioselective nitrone cycloadditions with 2-alkenoyl pyridine N-oxides catalyzed by Cu(II)-BOX complexes. *Org. Lett.* **2011,** *13* (3), 402-5.

59.   Blay, G.; Fernandez, I.; Munoz, M. C.; Pedro, J. R.; Vila, C., Synthesis of functionalized indoles with a trifluoromethyl-substituted stereogenic tertiary carbon atom through an enantioselective Friedel-Crafts alkylation with beta-trifluoromethyl-alpha,beta-enones. *Chemistry (Weinheim an der Bergstrasse, Germany)* **2010,** *16* (30), 9117-22.

60.   Blay, G.; Fernández, I.; Monleón, A.; Muñoz, M. C.; Pedro, J. R.; Vila, C., Synthesis of Functionalized Indoles with an α-Stereogenic Ketone Moiety Through an Enantioselective Friedel–Crafts Alkylation with (E)-1,4-Diaryl-2-butene-1,4-diones. *Adv. Synth. Catal.* **2009,** *351* (14-15), 2433-2440.

61.   Vilanova, C.; Diaz-Oltra, S.; Murga, J.; Falomir, E.; Carda, M.; Redondo-Horcajo, M.; Diaz, J. F.; Barasoain, I.; Marco, J. A., Design and synthesis of pironetin analogue/colchicine hybrids and study of their cytotoxic activity and mechanisms of interaction with tubulin. *J. Med. Chem.* **2014,** *57* (24), 10391-403.

62.   Vilanova, C.; Torijano-Gutierrez, S.; Diaz-Oltra, S.; Murga, J.; Falomir, E.; Carda, M.; Alberto Marco, J., Design and synthesis of pironetin analogue/combretastatin A-4 hybrids containing a 1,2,3-triazole ring and evaluation of their cytotoxic activity. *Eur. J. Med. Chem.* **2014,** *87*, 125-30.

63.   Marco, J. A.; Garcia-Pla, J.; Carda, M.; Murga, J.; Falomir, E.; Trigili, C.; Notararigo, S.; Diaz, J. F.; Barasoain, I., Design and synthesis of pironetin analogues with simplified structure and study of their interactions with microtubules. *Eur. J. Med. Chem.* **2011,** *46* (5), 1630-7.

64.   Marti-Centelles, R.; Cejudo-Marin, R.; Falomir, E.; Murga, J.; Carda, M.; Marco, J. A., Inhibition of VEGF expression in cancer cells and endothelial cell differentiation by synthetic stilbene derivatives. *Bioorg. Med. Chem.* **2013,** *21* (11), 3010-5.

65.   Panos, J.; Diaz-Oltra, S.; Sanchez-Peris, M.; Garcia-Pla, J.; Murga, J.; Falomir, E.; Carda, M.; Redondo-Horcajo, M.; Diaz, J. F.; Barasoain, I.; Marco, J. A., Synthesis and biological evaluation of truncated alpha-tubulin-binding pironetin analogues lacking alkyl pendants in the side chain or the dihydropyrone ring. *Org. Biomol. Chem.* **2013,** *11* (35), 5809-26.

66.   Torijano-Gutierrez, S.; Diaz-Oltra, S.; Falomir, E.; Murga, J.; Carda, M.; Marco, J. A., Synthesis of combretastatin A-4 O-alkyl derivatives and evaluation of their cytotoxic, antiangiogenic and antitelomerase activity. *Bioorg. Med. Chem.* **2013,** *21* (23), 7267-74.

67.   Sterling, T.; Irwin, J. J., ZINC 15-Ligand Discovery for Everyone. **2015**.

68.   Martel, S.; Gillerat, F.; Carosati, E.; Maiarelli, D.; Tetko, I. V.; Mannhold, R.; Carrupt, P. A., Large, chemically diverse dataset of logP measurements for benchmarking studies. *Eur. J. Pharm. Sci.* **2013,** *48* (1-2), 21-9.

69.   Irwin, J. J.; Shoichet, B. K., ZINC--a free database of commercially available compounds for virtual screening. *J. Chem. Inf. Model.* **2005,** *45* (1), 177-82.

70.   Cruciani, G.; Pastor, M.; Guba, W., VolSurf: a new tool for the pharmacokinetic optimization of lead compounds. *Eur. J. Pharm. Sci.* **2000,** *11, Supplement 2*, S29-S39.

71.   Cummins, D. J.; Andrews, C. W.; Bentley, J. A.; Cory, M., Molecular diversity in chemical databases: comparison of medicinal chemistry knowledge bases and databases of commercially available compounds. *J. Chem. Inf. Comput. Sci.* **1996,** *36* (4), 750-63.

72.   http://www.specs.net/, **(accessed on 6th June 2017)**.

73.   *MOE 2011.10* **Chemical Computing Group Inc.: Montreal, 2011**.

74.   Wild, D. J.; Blankley, C. J., Comparison of 2D fingerprint types and hierarchy level selection methods for structural grouping using Ward's clustering. *J. Chem. Inf. Comput. Sci.* **2000,** *40* (1), 155-62.

75.   Tanimoto, T. T., *An Elementary Mathematical Theory of Classification and Prediction*. International Business Machines Corporation: 1958.

76.   Willett, P.; Barnard, J. M.; Downs, G. M., Chemical Similarity Searching. *J. Chem. Inf. Comput. Sci.* **1998,** *38* (6), 983-996.

77.   Lipkus, A. H., A proof of the triangle inequality for the Tanimoto distance. *J. Math. Chem.* **1999,** *26* (1), 263-265.

78.   Scior, T.; Lozano-Aponte, J.; Figueroa-Vazquez, V.; Yunes-Rojas, J. A.; Zähringer, U.; Alexander, C., Three-dimensional mapping of differential amino acids of human, murine, canine and equine TLR4/MD-2 receptor complexes conferring endotoxic

activation by lipid A, antagonism by Eritoran and species-dependent activities of Lipid IVA in the mammalian LPS sensor system. *Comput. Struct. Biotechnol. J.* **2013,** *7*, e201305003.

79.   Park, S. H.; Kim, N. D.; Jung, J. K.; Lee, C. K.; Han, S. B.; Kim, Y., Myeloid differentiation 2 as a therapeutic target of inflammatory disorders. *Pharmacol. Ther.* **2012,** *133* (3), 291-8.

80.   Resman, N.; Gradisar, H.; Vasl, J.; Keber, M. M.; Pristovsek, P.; Jerala, R., Taxanes inhibit human TLR4 signaling by binding to MD-2. *FEBS Lett.* **2008,** *582* (28), 3929-34.

81.   Kawasaki, K.; Akashi, S.; Shimazu, R.; Yoshida, T.; Miyake, K.; Nishijima, M., Mouse toll-like receptor 4.MD-2 complex mediates lipopolysaccharide-mimetic signal transduction by Taxol. *J. Biol. Chem.* **2000,** *275* (4), 2251-4.

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

83.   Gradisar, H.; Keber, M. M.; Pristovsek, P.; Jerala, R., MD-2 as the target of curcumin in the inhibition of response to LPS. *J. Leukoc. Biol.* **2007,** *82* (4), 968-74.

84.   Roh, E.; Lee, H. S.; Kwak, J. A.; Hong, J. T.; Nam, S. Y.; Jung, S. H.; Lee, J. Y.; Kim, N. D.; Han, S. B.; Kim, Y., MD-2 as the target of nonlipid chalcone in the inhibition of endotoxin LPS-induced TLR4 activity. *J. Infect. Dis.* **2011,** *203* (7), 1012-20.

85.   Peluso, M. R.; Miranda, C. L.; Hobbs, D. J.; Proteau, R. R.; Stevens, J. F., Xanthohumol and related prenylated flavonoids inhibit inflammatory cytokine production in LPS-activated THP-1 monocytes: structure-activity relationships and in silico binding to myeloid differentiation protein-2 (MD-2). *Planta Med.* **2010,** *76* (14), 1536-43.

86.   Small-Molecule Drug Discovery Suite 2015-4: Glide, v., Schrödinger, LLC, New York, NY, 2015.

87.   Friesner, R. A.; Murphy, R. B.; Repasky, M. P.; Frye, L. L.; Greenwood, J. R.; Halgren, T. A.; Sanschagrin, P. C.; Mainz, D. T., Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *J. Med. Chem.* **2006,** *49* (21), 6177-6196.

88.   Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S., Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem.* **2004,** *47* (7), 1739-49.

89.   Trott, O.; Olson, A. J., AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **2010,** *31* (2), 455-461.

90.   Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J., AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.* **2009,** *30* (16), 2785-2791.

91.   Baxter, J., Local optima avoidance in depot location. *J. Oper. Res. Soc.* **1981**, 815-819.

92.   Blum, C.; Roli, A., Hybrid metaheuristics: an introduction. In *Hybrid Metaheuristics*, Springer: 2008; pp 1-30.

93.   Opal, S. M.; Laterre, P. F.; Francois, B.; LaRosa, S. P.; Angus, D. C.; Mira, J. P.; Wittebole, X.; Dugernier, T.; Perrotin, D.; Tidswell, M.; Jauregui, L.; Krell, K.; Pachl, J.;

Takahashi, T.; Peckelsen, C.; Cordasco, E.; Chang, C. S.; Oeyen, S.; Aikawa, N.; Maruyama, T.; Schein, R.; Kalil, A. C.; Van Nuffelen, M.; Lynn, M.; Rossignol, D. P.; Gogate, J.; Roberts, M. B.; Wheeler, J. L.; Vincent, J. L., Effect of eritoran, an antagonist of MD2-TLR4, on mortality in patients with severe sepsis: the ACCESS randomized trial. *Jama* **2013,** *309* (11), 1154-62.

94.   Sirci, F.; Istyastono, E. P.; Vischer, H. F.; Kooistra, A. J.; Nijmeijer, S.; Kuijer, M.; Wijtmans, M.; Mannhold, R.; Leurs, R.; de Esch, I. J.; de Graaf, C., Virtual fragment screening: discovery of histamine H3 receptor ligands using ligand-based and protein-based molecular fingerprints. *J. Chem. Inf. Model.* **2012,** *52* (12), 3308-24.

95.   Yadav, M. K.; Pandey, S. K.; Swati, D., Drug target prioritization in Plasmodium falciparum through metabolic network analysis, and inhibitor designing using virtual screening and docking approach. *J. Bioinform. Comput. Biol.* **2013,** *11* (4), 1350003.

96.   Scior, T.; Bender, A.; Tresadern, G.; Medina-Franco, J. L.; Martinez-Mayorga, K.; Langer, T.; Cuanalo-Contreras, K.; Agrafiotis, D. K., Recognizing pitfalls in virtual screening: a critical review. *J. Chem. Inf. Model.* **2012,** *52* (4), 867-81.

97.   Collins, J. C.; Armstrong, A.; Chapman, K. L.; Cordingley, H. C.; Jaxa-Chamiec, A. A.; Judd, K. E.; Mann, D. J.; Scott, K. A.; Tralau-Stewart, C. J.; Low, C. M. R., Prospective use of molecular field points in ligand-based virtual screening: efficient identification of new reversible Cdc25 inhibitors. *MedChemComm* **2013,** *4* (8), 1148-1155.

98.   Liu, C.; He, G.; Jiang, Q.; Han, B.; Peng, C., Novel Hybrid Virtual Screening Protocol Based on Molecular Docking and Structure-Based Pharmacophore for Discovery of Methionyl-tRNA Synthetase Inhibitors as Antibacterial Agents. *Int. J. Mol. Sci.* **2013,** *14* (7), 14225-14239.

99.   Koo, J. E.; Park, Z. Y.; Kim, N. D.; Lee, J. Y., Sulforaphane inhibits the engagement of LPS with TLR4/MD2 complex by preferential binding to Cys133 in MD2. *Biochem. Biophys. Res. Commun.* **2013,** *434* (3), 600-5.

100.   Wishart, D. S.; Knox, C.; Guo, A. C.; Cheng, D.; Shrivastava, S.; Tzur, D.; Gautam, B.; Hassanali, M., DrugBank: a knowledgebase for drugs, drug actions and drug targets. *Nucleic Acids Res.* **2008,** *36* (Database issue), D901-6.

101.   Lewis, S. S.; Hutchinson, M. R.; Rezvani, N.; Loram, L. C.; Zhang, Y.; Maier, S. F.; Rice, K. C.; Watkins, L. R., Evidence that intrathecal morphine-3-glucuronide may cause pain enhancement via toll-like receptor 4/MD-2 and interleukin-1β. *Neuroscience* **2010,** *165* (2), 569-583.

102.   Hutchinson, M. R.; Zhang, Y.; Shridhar, M.; Evans, J. H.; Buchanan, M. M.; Zhao, T. X.; Slivka, P. F.; Coats, B. D.; Rezvani, N.; Wieseler, J.; Hughes, T. S.; Landgraf, K. E.; Chan, S.; Fong, S.; Phipps, S.; Falke, J. J.; Leinwand, L. A.; Maier, S. F.; Yin, H.; Rice, K. C.; Watkins, L. R., Evidence that opioids may have toll-like receptor 4 and MD-2 effects. *Brain Behav. Immun.* **2010,** *24* (1), 83-95.

103.   Hutchinson, M. R.; Northcutt, A. L.; Hiranita, T.; Wang, X.; Lewis, S.; Thomas, J.; van Steeg, K.; Kopajtic, T. A.; Loram, L.; Sfregola, C.; Galer, E.; Miles, N. E.; Bland, S. T.; Amat, J.; Rozeske, R. R.; Maslanik, T.; Chapman, T.; Strand, K.; Fleshner, M.; Bachtell, R. K.; Somogyi, A. A.; Yin, H.; Katz, J. L.; Rice, K. C.; Maier, S. F.; Watkins, L. R., Opioid activation of Toll-Like receptor 4 contributes to drug reinforcement. *J. Neurosci.* **2012,** *32* (33), 11187-11200.

104.	Ishinaga, H.; Takeuchi, K.; Kishioka, C.; Suzuki, S.; Basbaum, C.; Majima, Y., Pranlukast inhibits NF-kappaB activation and MUC2 gene expression in cultured human epithelial cells. *Pharmacology* **2005,** *73* (2), 89-96.

105.	Woszczek, G.; Chen, L.-Y.; Alsaaty, S.; Nagineni, S.; Shelhamer, J. H., Concentration dependent non-CysLT(1) receptor mediated inhibitory activity of leukotriene receptor antagonists. *J. Immunol.* **2010,** *184* (4), 2219-2225.

106.	Ishioka, S.; Hozawa, S.; Haruta, Y.; Hiyama, K.; Maeda, A.; Yamakido, M., Pranlukast, a cysteinyl leukotriene antagonist, reduces serum eosinophil cationic protein levels in patients with asthma. *Hiroshima J. Med. Sci.* **1999,** *48* (4), 105-110.

107.	Bollag, G.; Tsai, J.; Zhang, J.; Zhang, C.; Ibrahim, P.; Nolop, K.; Hirth, P., Vemurafenib: the first drug approved for BRAF-mutant cancer. *Nat. Rev. Drug Discov.* **2012,** *11* (11), 873-886.

108.	Tsai, J.; Lee, J. T.; Wang, W.; Zhang, J.; Cho, H.; Mamo, S.; Bremer, R.; Gillette, S.; Kong, J.; Haass, N. K., Discovery of a selective inhibitor of oncogenic B-Raf kinase with potent antimelanoma activity. *Proc. Natl. Acad. Sci. U.S.A.* **2008,** *105* (8), 3041-3046.

109.	Bollag, G.; Hirth, P.; Tsai, J.; Zhang, J.; Ibrahim, P. N.; Cho, H.; Spevak, W.; Zhang, C.; Zhang, Y.; Habets, G., Clinical efficacy of a RAF inhibitor needs broad target blockade in BRAF-mutant melanoma. *Nature* **2010,** *467* (7315), 596-599.

110.	Goto, Y.; Arigami, T.; Kitago, M.; Nguyen, S. L.; Narita, N.; Ferrone, S.; Morton, D. L.; Irie, R. F.; Hoon, D. S., Activation of Toll-like receptors 2, 3, and 4 on human melanoma cells induces inflammatory factors. *Molecular cancer therapeutics* **2008,** *7* (11), 3642-3653.

111.	Takazawa, Y.; Kiniwa, Y.; Ogawa, E.; Uchiyama, A.; Ashida, A.; Uhara, H.; Goto, Y.; Okuyama, R., Toll-like receptor 4 signaling promotes the migration of human melanoma cells. *The Tohoku journal of experimental medicine* **2014,** *234* (1), 57-65.

112.	Schrödinger Release 2015-4: Maestro, v., Schrödinger, LLC, New York, NY, 2015.

113.	Schrödinger Release 2015-4: LigPrep, v., Schrödinger, LLC, New York, NY, 2015.

114.	Banks, J. L.; Beard, H. S.; Cao, Y.; Cho, A. E.; Damm, W.; Farid, R.; Felts, A. K.; Halgren, T. A.; Mainz, D. T.; Maple, J. R., Integrated modeling program, applied chemical theory (IMPACT). *J. Med. Chem.* **2005,** *26* (16), 1752-1780.

115.	Schrödinger Release 2015-4: Epik, v., Schrödinger, LLC, New York, NY, 2015.

116.	Greenwood, J. R.; Calkins, D.; Sullivan, A. P.; Shelley, J. C., Towards the comprehensive, rapid, and accurate prediction of the favorable tautomeric states of drug-like molecules in aqueous solution. *J. Comput. Aided Mol. Des.* **2010,** *24* (6-7), 591-604.

117.	Shelley, J. C.; Cholleti, A.; Frye, L. L.; Greenwood, J. R.; Timlin, M. R.; Uchimaya, M., Epik: a software program for pK a prediction and protonation state generation for drug-like molecules. *J. Comput. Aided Mol. Des.* **2007,** *21* (12), 681-691.

118.	Baroni, M.; Cruciani, G.; Sciabola, S.; Perruccio, F.; Mason, J. S., A common reference framework for analyzing/comparing proteins and ligands. Fingerprints for Ligands and Proteins (FLAP): theory and application. *J. Chem. Inf. Model.* **2007,** *47* (2), 279-94.

119.	Cross, S.; Baroni, M.; Goracci, L.; Cruciani, G., GRID-based three-dimensional pharmacophores I: FLAPpharm, a novel approach for pharmacophore elucidation. *J. Chem. Inf. Model.* **2012,** *52* (10), 2587-98.

120.   Lyne, P. D., Structure-based virtual screening: an overview. *Drug Discov. Today* **2002,** *7* (20), 1047-55.

121.   Allinger, N. L.; Yuh, Y. H.; Lii, J. H., Molecular mechanics. The MM3 force field for hydrocarbons. 1. *J. Am. Chem. Soc.* **1989,** *111* (23), 8551-8566.

122.   Verdonk, M. L.; Berdini, V.; Hartshorn, M. J.; Mooij, W. T.; Murray, C. W.; Taylor, R. D.; Watson, P., Virtual screening using protein-ligand docking: avoiding artificial enrichment. *J. Chem. Inf. Comput. Sci.* **2004,** *44* (3), 793-806.

123.   Huang, N.; Shoichet, B. K.; Irwin, J. J., Benchmarking Sets for Molecular Docking. *J. Med. Chem.* **2006,** *49* (23), 6789-6801.

CHAPTER 6

Computational Approaches to the Structure of Toll-Like Receptor 4 and its Membrane Interactions

6.1   Introduction

Cell membranes, also known as the plasma membrane or cytoplasmic membrane, consists in a lipid bilayer with ions, channels and proteins embedded, and separate the cell interior from the outside environment (Figure **6.1**).[1] The membrane has important implications for many cellular processes, e.g., protein trafficking and aggregation, membrane fusion, and signal transduction. Lipids are the main components of lipid bilayers and play an important role in many cell signaling and physiological processes. Changes in expression levels of individual lipid species have been implicated in many diseases including: cancers, diabetes, Alzheimer's disease, HIV entry, and aterosclerosis.[2-3]
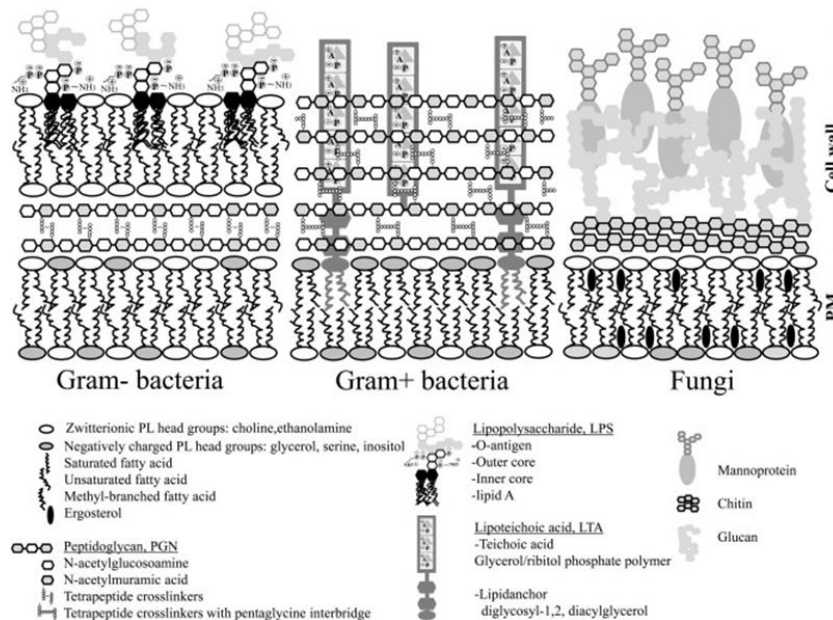
Figure 6.1. Cell envelopes of various microbial families.[1]

A typical plasma membrane requires a lipid organization,[4] which is formed by hundreds of different lipids. All the lipid molecules in cell membranes are amphipathic, being phopholipids the most abundant ones, together with glycolipids and sterols.[5] The fatty chains (FA) in phospholipids and glycolipids may be saturated or unsaturated and usually contain an even number of carbon atoms, typically between 16 and 20, where the 16- and 18-carbon FAs the most common ones. The polar head groups are

exposed to water and nonpolar lipid tails groups go inside the membrane.[6-8] The length and the degree of unsaturation of FA chains have an effect on membrane fluidity, preventing the FA from packing together as tightly, thus decreasing the melting temperature (increasing the fluidity) of the membrane. In particular, mammalian membrane is composed by phosphatidylcholine (PC), sphingomyelin (SM), and gangliosides (GM) in the outer leaflet and phosphatidylcholines (PE), phosphatidylserine (PS), and other charged lipids in the inner leaflet; also eukaryotic plasma membrane contains approximately 20-50% sterols (Figure **6.2**).[6, 9]
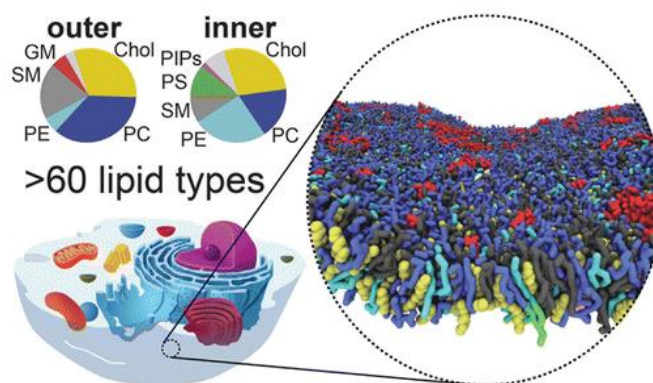
Figure 6.2. Lipid organization of the plasma membrane extracted from Marrink *et al.*[4]

The importance of study of receptors in a membrane environment may be used to explore in detail the interactions of membrane proteins and specific lipids, yielding predictions of lipid binding sites in good agreement with available structural data. Molecular dynamics (MD) simulation approaches provide important tools which allow us to simulate both individual membrane proteins and more complex membrane systems. Thus, MD simulations have become a valuable addition to the range of experimental structural and biophysical techniques for studying membrane proteins and their interactions with lipids.[10-12]

For all of this, the computational study of the membranes at atomic details has become an essential tool for elucidate the structural and dynamic organization of cellular membranes and to understand the different mechanims where the membranes are implicated, such as: mechanims of diffusion through ions or

membranes, mechanisms of signalling and also the active transport throught transporters.

TLR4 together with MD-2, is one of the main receptors involved in innate immunity. TLR4/MD-2 modulation constitutes a challenging and sparkling area of research with high potential for the development of novel drugs. The molecular modelling approaches have been used to elucidate the molecular recognition mechanisms of TLR4/MD-2 modulation, with focus on the agonist/antagonist conformational changes of the TLR4/MD-2 system, and to provide some hints for the design of novel binders, hopefully with therapeutic potential. And the analysis of the MD simulation of the dimer TLR4/MD-2 complex in the membrane environment could help us to identify the key ligand-receptor and protein-protein interactions governing the molecular recognition events and the dimerization process. Our previously calculated complexes of TLR4 with reported agonists will be also calculated accounting for the validation of the proposed binding modes.

6.2   Results and Discussion

Several models of the TLR4/MD-2 system inserted in different membranes have been built and simulated. These different models will be useful for the final building of the complete TLR4/MD-2 dimer and will provide us insights into the mechanism of TLR4 agonism/antagonism. The MD simulations of the full complex will help us to identify the key ligand-receptor and protein-protein interactions governing the molecular recognition events and the dimerization process at atomic level.

We started carrying out the building and simulation of different models of membranes. With these models, we inserted the monomer TLR4/MD-2 with three molecules of myristic acid. X-ray structure is available for the extracellular domain (ED) in complex with lipid A. The intracellular (ID) and transmembrane domains (TD) were modelled by homology modelling by members of our group and were used for this study. The building and simulation of the full dimer complex of TLR4/MD-2 system in the bilayer membrane is in progress, especially the build and simulation of the full

dimer complex of TLR4/MD-2 system in the bilayer membrane in complex with agonists ONO-4007 (LPS-like molecule)[13] and Euodenine A (non-LPS-like molecule).[14] The binding poses have been already modelled by means of docking techniques plus MD simulations (see Chapter 3). The proposed binding modes will be assessed and compared with the model obtained for lipid A.

Membrane models:

We started this approach with the building of different simple <u>symmetric membrane models</u> (Table **6.1**). It is often important to understand the dynamics of the bilayer itself before proceeding with protein systems.

The lipid bilayers models are: POPC (1palmitoyl-2-oleoyl-sn-glycero-3-phosphocoline), POPE(1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine)-POPC [1:1], CHOL (Cholesterol)-POPC [1:1], DPPC(1,2-dipalmitoyl-sn-glycero-3-phosphocholine)-POPC [1:1], DPPC-POPE [1:1] (Annex Table **6.1**, structure of the lipid and sterols).

Also reviewing the literature and given that the composition of the mammalian plasma membranes are composed by different type of lipids and sterols, we decided to approach the complexicity of real cell membranes and build a more complex <u>asymmetric membrane model</u> and which resembled more the composition of the mammalian plasma membrane, with following composition: The <u>outer leaflet</u> is composed by 35% cholesterol and 65% lipids (60% DPPC, DPPE (0 insaturation), 20% POPE, POPC (1 insaturation), 20% DOPE, DOPC (>1 insaturation)), and the <u>inner leaflet</u> by 30% cholesterol and 70% lipids (50% DPPC, DPPE, 20% POPE, POPC and 30% DOPE, DOPC) (Figure **6.3** and **6.4**).

OUTER	INNER
65% lipids:	70% lipids:
0 insaturation:	**0 insaturation:**
60% { DPPC (30%) / DPPE (30%)	50% { DPPC (25%) / DPPE (25%)
1 insaturation:	**1 insaturation:**
20% { POPE (10%) / POPC (10%)	20% { POPE (10%) / POPC (10%)
> 1 insaturation:	**> 1 insaturation:**
20% { DOPE (10%) / DOPC (10%)	30% { DOPE (15%) / DOPC (15%)
35 % Cholesterol	**30%** Cholesterol

Figure 6.3. Composition of the asymmetric membrane model built by us.

OUTSIDE: 263 molecules
35 % CHL = 92 molecules − spheres purple
65% lipids = 171 molecules
   51 molecules **DPPC** (0 insaturation)
   52 molecules **DPPE** (0 insaturation)
   17 moelcules **POPE** (1 insaturation)
   17 molecules **POPC** (1 insaturation)
   17 molecules **DOPE** (>1 insaturation)
   17 moelcules **DOPC** (>1 insaturation)
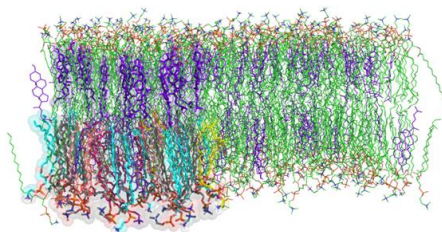
INSIDE= 250 molecules
30% CHL = 75 molecules
70 % lipids = 175 molecules
   44 molecules **DPPC**
   44 molecules **DPPE**
   17 molecules **POPE**
   18 molecules **POPC**
   26 molecules **DOPE**
   26 moelcules **DOPC**
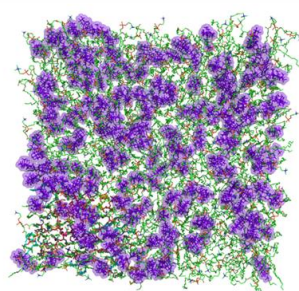
Figure 6.4. Our asymmetric membrane model built by us, with the lipids and sterols represented in different colors.

Regarding TLR4/MD-2 system, an X-ray structure is available for the extracellular domain in complex with lipid A (PDB-ID: 3FXI). The intracellular and transmembrane domains have been modelled by members of our group by homology modelling and were used for this study (Ref. Uniprot O00206).

The MD simulations of these models were carried out, and with the area per molecule and the electron density were analyzed (see Figures **6.5** and **6.6**).
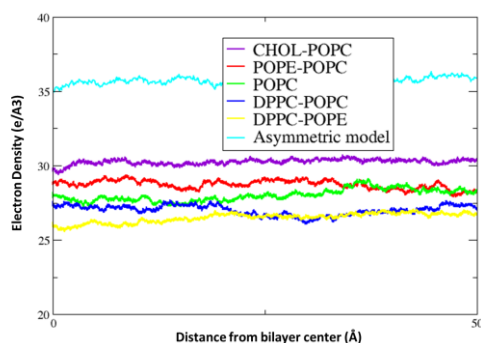
Area per molecule:

Figure 6.5. Area per molecule for different models of membranes during MD simulation (50 ns).
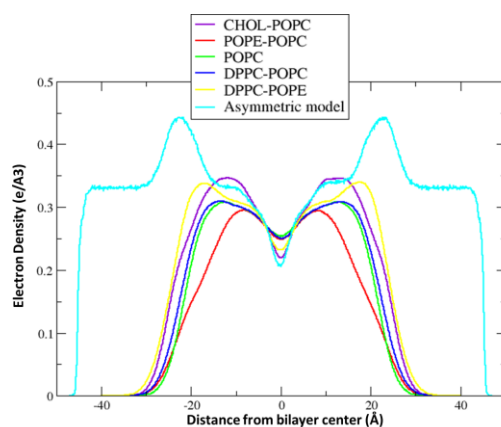
Figure 6.6. Electron density profile for the different models of membranes during MD simulations (50 ns).

In order to undertand the behaviour of the TLR4/MD-2 system in the membrane enviorment, we started to build different models from the most simple until more complex with the different models of membranes as follows:

1- The building and simulation of the transmembrane and intracellular domain of TLR4 in the different bilayer membranes environment complex (100 ns for POPC, POPE-POPC and CHOL-POPC, DPPC-POPE and DPPC-POPC) was carried out, with the corresponding analysis of different parameters.

2- The building and simulation of the full monomer system (extracellular, transmembrane and intracellular domains) of the complete complex of agonist TLR4/MD-2 system with three molecules of myristic acid in the different bilayer membranes environment complex (100 ns for POPC, POPE-POPC and CHOL-POPC, DPPC-POPE and DPPC-POPC) was carried out, with the corresponding analysis of different parameters (Figure **6.7**).

3- The building and simulation of the full dimer system of the complete monomer complex of agonist TLR4/MD-2 system in the different bilayer membranes environment in our asymmetric complex bilayer model is in progress with the corresponding analysis of different parameters.
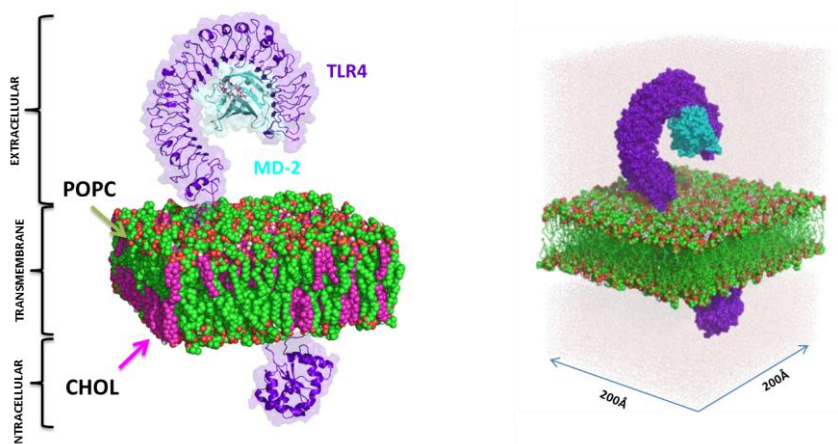

Figure 6.7. Example of the monomer TLR4/MD-2 complex with three molecules of myristic acid in the CHOL-POPC bilayer membrane.

ANALYSIS:

The analysis was performed using the cpptraj module of AmberTools15.[15] After production several lipid bilayer parameters, including area per lipid, electron density profiles, RMSD, RMSF, distances and angles were analyzed. The results for the transmembrane + intracellular domain in the membrane environment are very similar compared with the full monomer system (extracellular + transmembrane + intracellular domains of TLR4).

Area per molecule: The area per molecule is a common experimental structural parameter used in the validation of lipid bilayer simulations. The area is determined from the specified cross-section of the box.

Very little is known about the experimental data of membranes. For the case POPC membrane, the experimental data indicate that area per lipid is between 64.3 and 68.3 Å^2.[16-17] Our models indicate that the area per lipid is around 65 Å^2, so satisfactory are in agreement with the experimental value. In the case of POPE, the experimental data indicates that the area per lipid is around 59-60 Å^2.[18] The area per molecule of the bilayer systems are seen to be stable during the simulation time and shows lower fluctuations in all the cases. In the case if CHOL-POPC the area per molecule is less, because the sterols occupy less area in relation to lipids (Figure **6.8**).

TLR4 (ED+TD+ID)/MD-2 with 3 myristic acids
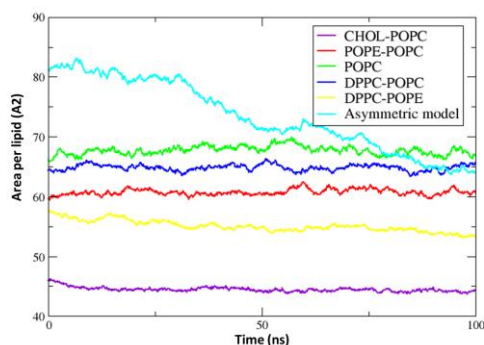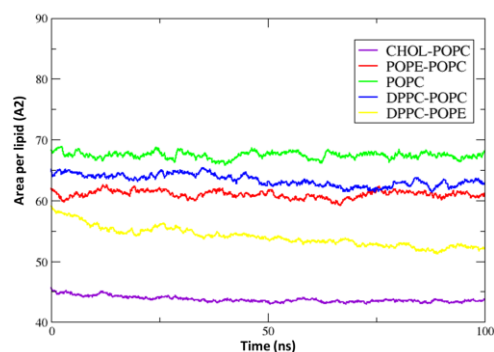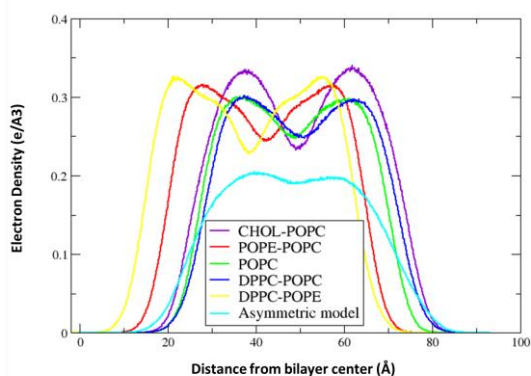
TLR4 (TD+ID)

Figure 6.8. Area per molecule of the TLR4/MD-2 system with three molecules of myristic acids on the left and TLR4 (TD+ID) on the right, for the different bilayer systems.

Electron density profile: The electron density profile provides a time-averaged measurement of the density of electrons through the lipid bilayer. The total electron density profile for the lipid bilayers was carried out. It is possible to compute the thickness of the lipid bilayer using the peak-to-peak distance.

Regarding the experimental data, the thickness for different models of membranes is: 37 Å for POPC, [17] and 39.5 Å for POPE.[18] The results obtained for the transmembrane and intracellular domains are very similar comparing with the full system; in all the cases we have obtained a thickness around 40 Å (Figure **6.9**).

TLR4 (ED+TD+ID)/MD-2 with 3 myristic acids
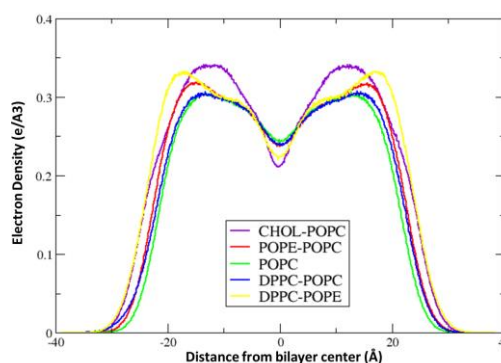
TLR4 (TD+ID)

Figure 6.9. Electron density profile of the bilayer systems.

RMSD: The root mean-square deviation (RMSD) of the backbone of the protein (TLR4 and MD-2) and the heavy atoms of the ligands (three molecules of myristic acid) were calculated around the molecular dynamic (MD) simulations (100ns). In the case of the TLR4, we have monitorized also the extracellular (ED), transmembrane (TD) and intracellular (ID) domain.

From studying the RSMD of the simulations in all the cases, it could be observed that TLR4 (full system), at the beginning had some significant movements but acquire certain stability since 40 ns of the MD simulation, except in the case of CHOL-POPC membrane which has drastic movements throughout all the MD simulation (Figure **6.10**) but in the case of the RMSD separate into fragments (ED, TD and ID), it could be possible to observe stability in all the models for the TD, but in the case of the ED and ID all the models are stable, except in the case of DPPC-POPC model which has some drastic movements throughout the whole of the simulation (Figure **6.11**).
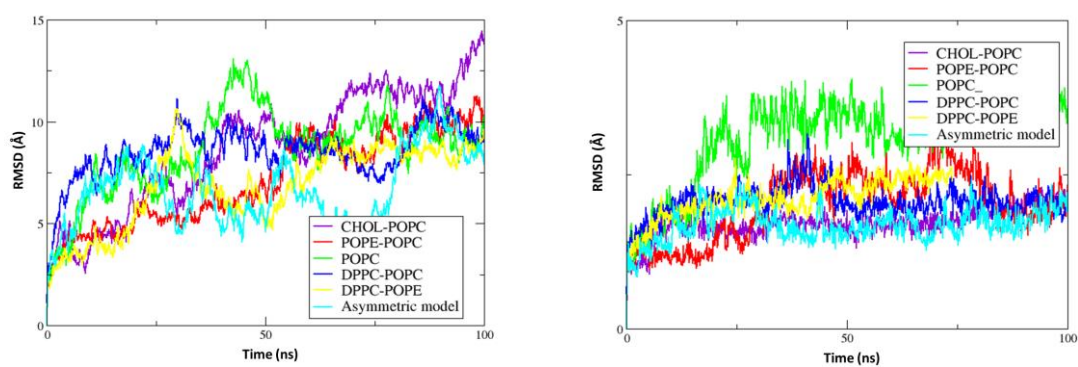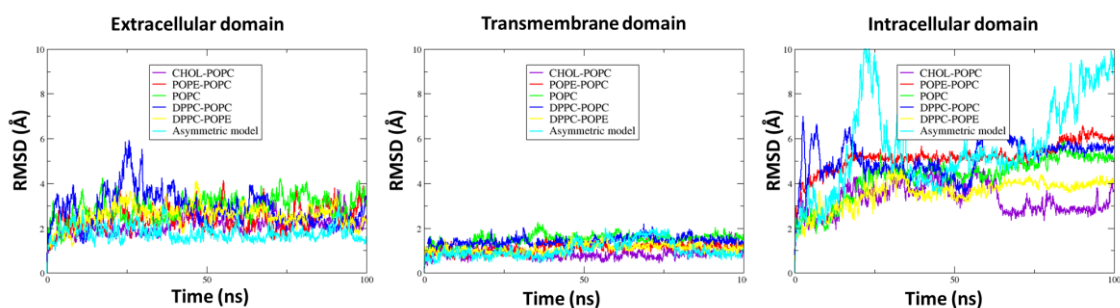
Figure 6.10. RMSD of the backbone of the TLR4/MD-2 monomer (EC, TM and IC domains) in complex with three myristic acids (left) and the MD-2 protein (right).

TLR4 (ED+TD+ID)/MD-2 with 3 myristic acids
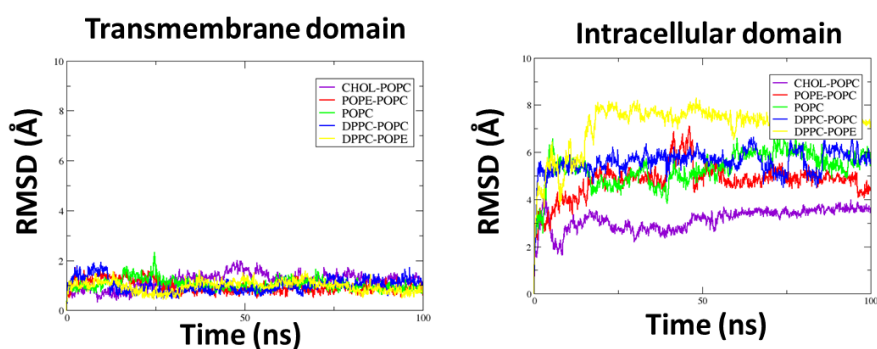

TLR4 (TD+ID)

Figure 6.11. RMSD of the backbone of the extracellular, transmembrane and intracellular TLR4 domains for the different models of membranes.

In the case of the MD-2 and the three molecules of myristic acid (Figure **6.12**), MD-2 remains stable throughout the whole of the simulation; regarding the ligands, the three molecules of myrisitc acid in the case of POPC membrane change the positions since 30 ns, but from here remains stable until the end of the simulation. In the case of POPE-POPC, the myristic acids 146 and 147 change the position from 45 ns. In the DPPC-POPC model, the ligands change at the beginning but remain stable during the whole of the simulation. For the others models, the molecules of myrist acid remain stable during the whole of the simulation.
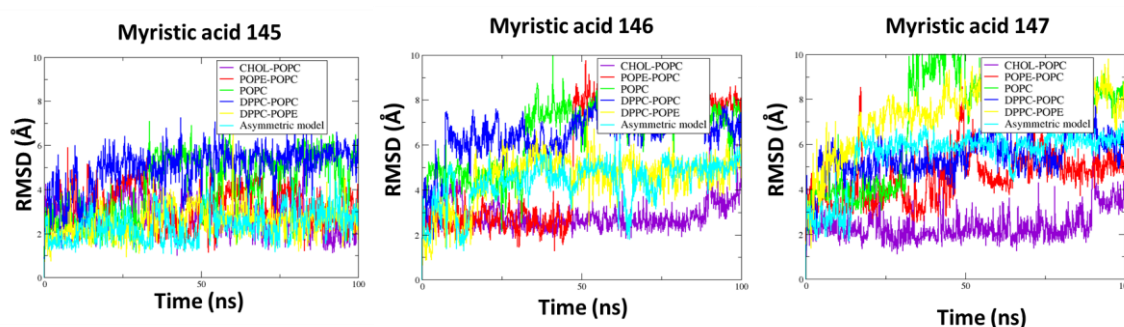

Figure 6.12. RMSD of the heavy atoms of the three molecules (145, 146 and 147) of myristic acid for the different membranes.

RMSF: We have measure the root mean-square fluctuation (RMSF), a measure of the average atomic mobility of backbone atoms (N, Cα and C atoms) during the MD simulations, in order to evaluate internal fluctuation in the different models (figure **6.13**).

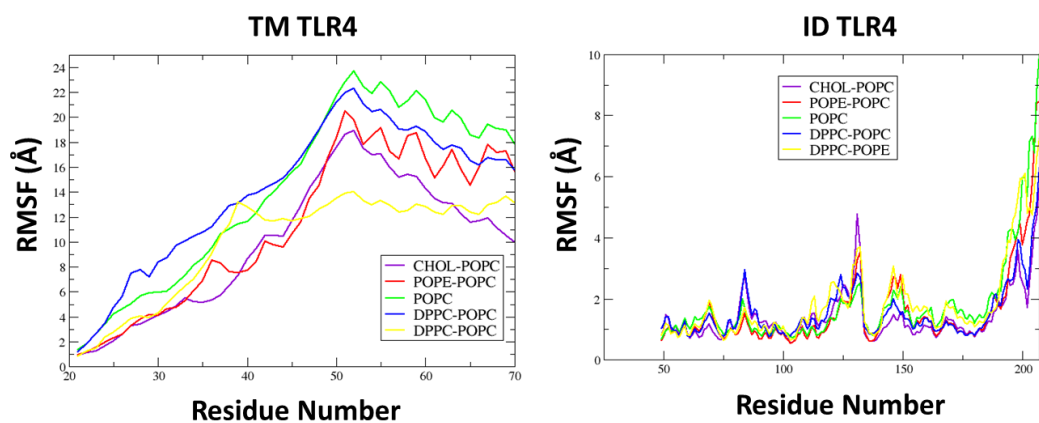## TLR4 (ED+TD+ID)/MD-2 with 3 myristic acids
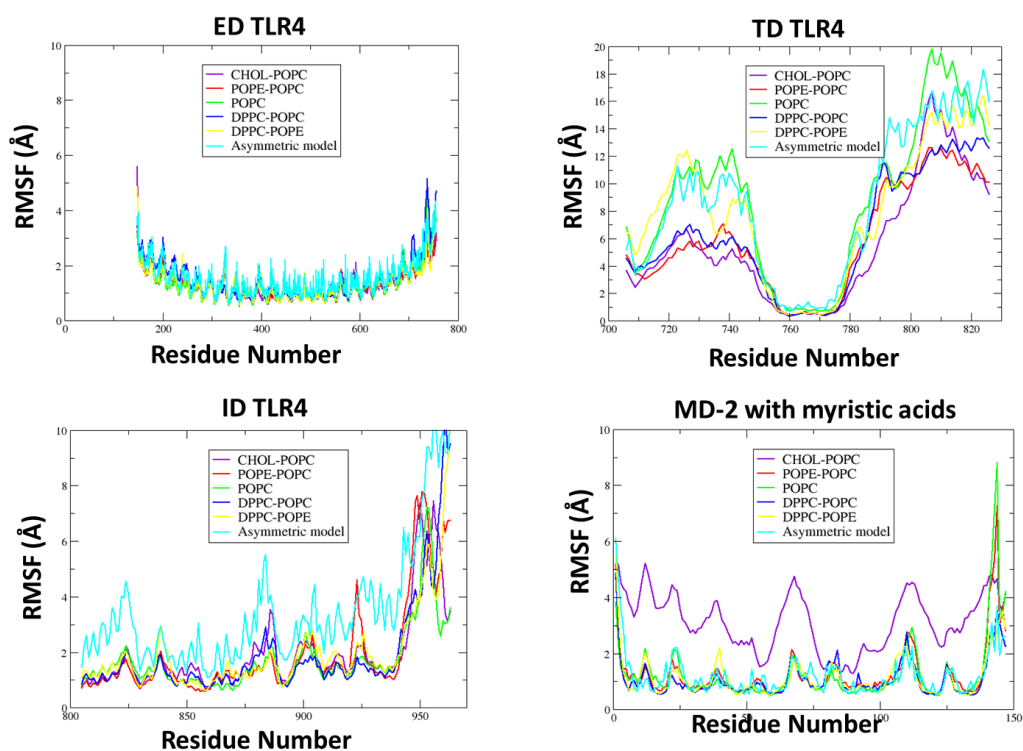
TLR4 (TD+ID)

Figure 6.13. RMSF of the backbone of different parts of TLR4 (ED, TD and ID) and MD-2 with myristic acids (top) and TLR4 (TD and ID) (bottom) for the different models of membranes.

In the case of the MD-2 pocket, the graphics show the average RMSD of the amino acid residues. It is possible to observe that the mayor fluctuations correspond to the residues in the loops implicate in the dimerization interface.

For the ED, the pics of fluctuation correspond to the amino acids residues in the loops, the mayor fluctuations are due to the first residue which is the initial part of TLR4, and the final residue correspond to threonine, which is in the end of the TLR4 extracellular domain, and very close to the membrane.

Distances: The lenght for the TD of TLR4 was measured from Lys631 to Lys653. The distance is stabley around the MD simulations both cases (TD +ID domain of TLR4 adn TLR4 complete in the models of membranes), the distance is around 35 Å (Figure **6.14**).

TLR4 (ED+TD+ID)/MD-2 with 3 myristic acids
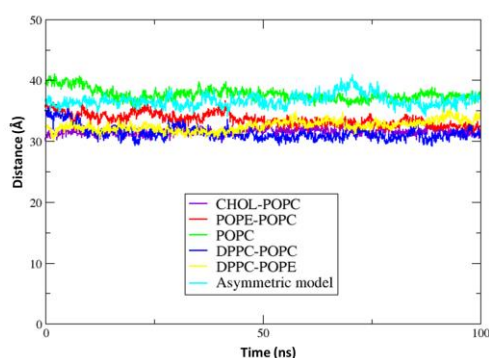
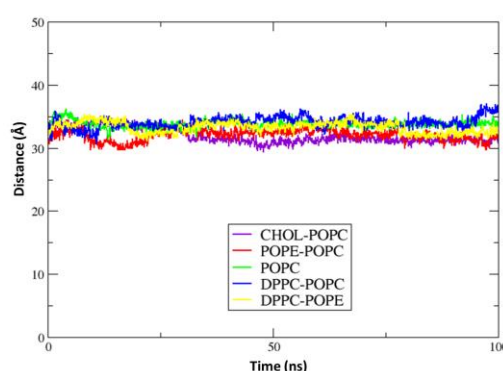**TLR4 (TD+ID)**

Figure 6.14. Distances for the TM domain from Lys631 to Lys653 of TLR4 in the different models of membranes.

Also the distance from Lys653 to the middle of the ID of TLR4 was measured, and it was possible to observe that the ID approaches to the membrane in all the models, establishing interactions with the head gropus of the inner leaflet from the membrane (Figure **6.15**).

Regarding the ID of TLR4 from <u>CHOL-POPC membrane</u> (Figure **6.15** and **6.16**), this domain aprroach to the membrane establishing polar interactions the residues Gly663, Tyr667, Arg689, Arg669, Glu698, Gly699 and Gln704 eith the polar head groups of POPC inner leaflet, also CH-π interaction between the Phe656 and Pro714. The portion of linker since Tyr653 to Cys664 is inserted in the membrane.
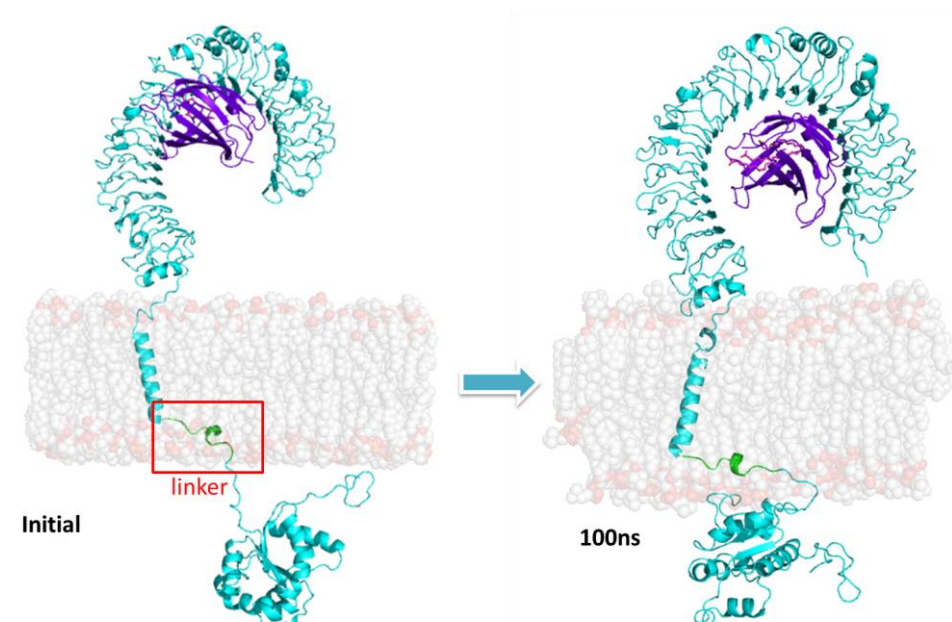

Figure 6.15. MD simulation of TLR4/MD-2 with three molecules of myristic acid in CHOL-POPC membrane. The ID is aprroach towards the inner leaflet of the membrane, establishing polar interactions with the head groups of the inner leaflet.
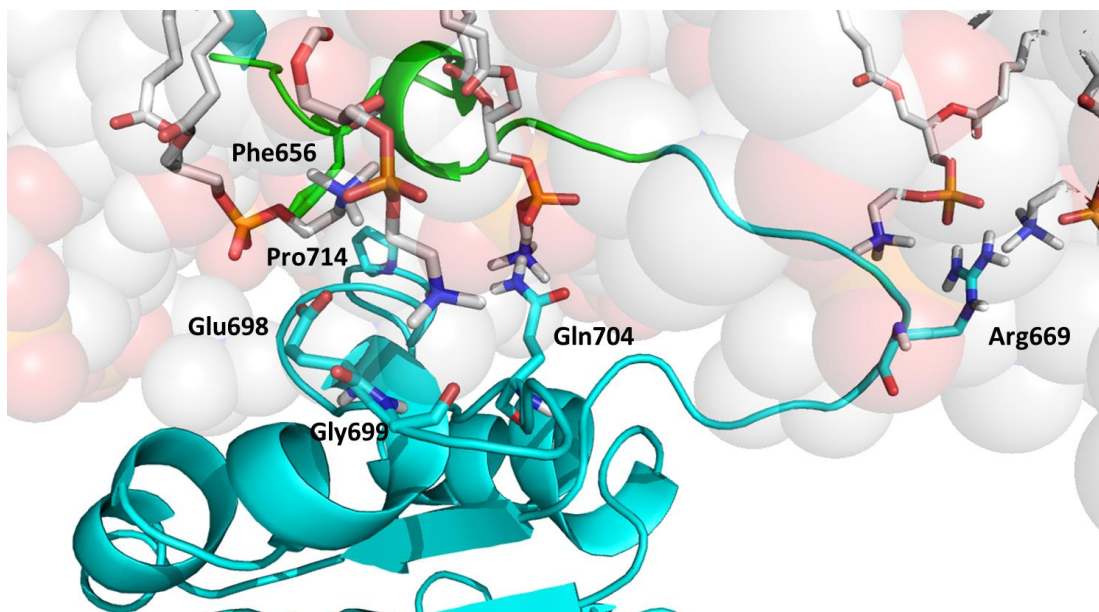
Figure 6.16. Polar interactions of the ID of TLR4 with the head groups of POPC from the inner leaflet of the CHOL-POPC membrane.

Regarding the model with <u>POPC membrane</u>, also de portion of the linker since Tyr653 to Gly668 is inserted in the membrane and is possible to observe polar interaction from the residues Gly663, Lys666, Tyr667, Arg669, Glu698, Gly828, Thr829 and Trp833 of TLR4 with the head groups of the inner leaflet of the membrane.

Regarding the model with <u>POPE-POPC membrane</u>, the linker is also inserted in the mebrane from Tyr653 to Cys664 and the residues from TLR4: Gly663, Lys666, Arg669 and Gln704 establish polar interactions with the head groups of the lipids.

Regarding the model with <u>DPPC-POPC</u>, the residues of TLR4 Gly663, Lys666, Tyr667 and Arg669 and establish polar interactions with the head groups of the membrane. Also the linker from Tyr653 to Cys664 is inserted in the membrane.

Regarding the model with <u>DPPC-POPE</u> Lys666, Arg669, Glu671, Arg763, Asp817 and Lys819 from TLR4 establish polar interactions with the head groups of the membrane, also the linker since Tyr653 to Cys664 is inserted in the membrane.

For the models built with only TD and ID of TLR4 in the different membranes enviorment was possible observe the same interactions from the ID of TLR4 with the

head groups of the inner leaflet of the membranes and the linker is also inserted in the membrane.

It could be said that the ID approach to the membrane, due to contains residues in this part that interact with the heads groups of the inner leaflet.

Secondary structure – helicity of the transmembrane and intrecellular domain of TLR4: DSSP method of Kabsch and Sander[19] were calculated for TD and ID of TLR4 , which assigns secondary structure types for residues based on backbone amide (N-H) and carbonyl (C=O) atom positions. We have measured the evolution of secondary structure over 100 ns. In the case of the ID for all the models, we observed stable structural features with minor variations in turn or loop regions around the MD simulation. And the secondary structure of the transmembrane region was stable, which was a α-helix around residues Ile633 and Val651 (Annex Figure **6.1** and Annex Figure **6.2**).

Angles: In order to see the torsion and movement of the TLR4 during the MS simulation toward the membrane, we measured two angles, angle 123 between Cys390-Cy391(1), Cys585- Cys627 (2) and Lys653 (3), and angle 234 between Cys585-Cys627 (2), Lys653 (3) and Cys735-Cys736 (4) (Figure **6.17**). We observed in both models, TD+ID domain of TLR4 and TLR4 complete in the different models of membranes, and we can conclude that it remains more or less stable throughout the simulation, not be possible to observe great changes at different angles (Figure **6.18**).
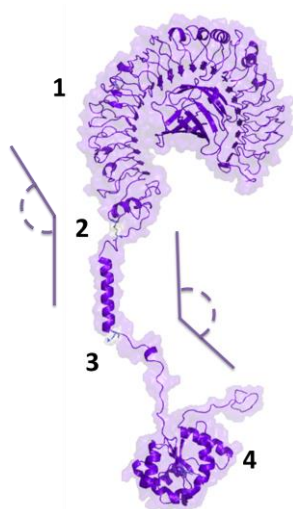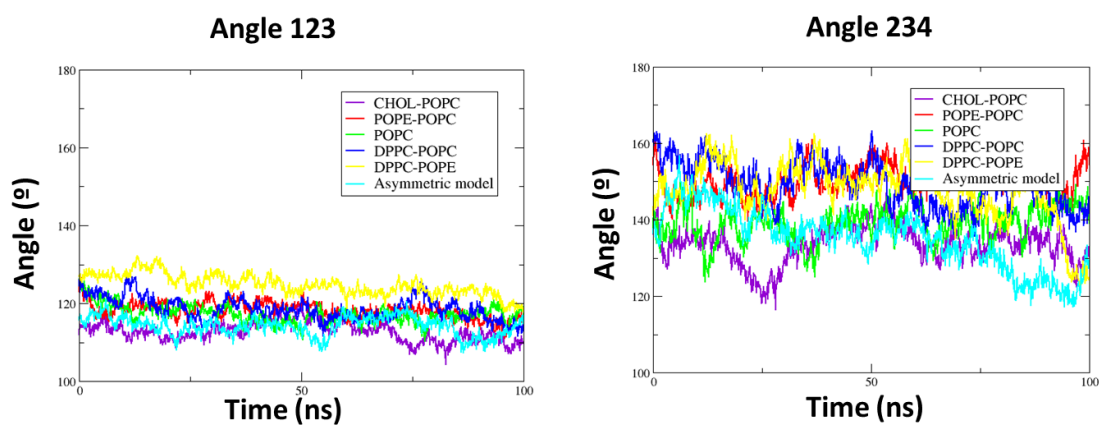
Figure 6.17. Angles measured in the TLR4 during the MD simulation of the different models of membranes.

TLR4 (ED+TD+ID)/MD-2 with 3 myristic acids

Angle 123

Angle 234
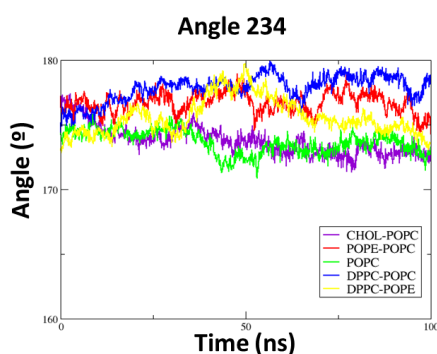
TLR4 (TD+ID)

Angle 234

Figure 6.18. Angle 134 between Cys390-Cys391 (1), Cys585-Cys627 (2), Lys653 (3) and Cys735-Cys736 (4) from TLR4.

6.3   Conclusions

We have presented results from MD simulation of bilayers with different kind of lipids and also including cholesterol in order to see the behaviour of the TLR4/MD-2 system in different membrane enviorments. The analysis of the MD simulation led us to understand better the key ligand-receptor and protein-protein interactions implicated in the molecular recognition events and in the dimerization process.

Firslty, we started this approach with the building of different symmetric models, and the analysis were carried out. The area per lipid is different depending if the composition include saturation or insaturation lipids and also depending if their composition include or not cholesterol. Regarding the membranes with cholesterol or insaturated lipids the area per lipid is less than the others membranes with saturated lipids. But very little is know about the experimental data of membranes.In relation to the thickness of the models, is approximatley the same for all the models, 40 Å in all the cases. An asymmetric membrane model including cholesterol has been built trying to represent the complexity of the mamalian membranes. We have perfomed the MD simulation, and we are going to use this model to insert now the TLR4/MD-2 system.

The TLR4/MD-2 monomeric complex has been simulated inserted in all the membrane models (in the asymmetric model , the MD simualtion is still running), and was possible to observe that the membranes with DPPC (saturated lipid) are more instable comparing with the membranes with unsaturated lipids or choleterol. We are trying to undertand this fact and to figure out what is happen, because the analysis of the membranes with DPPC-POPC and DPPC-POPE give worse results in comparation with the models POPC, CHOL-POPC, POPE-POPC and asymmetric models.

We are working now in the full dimer structure (ED, TD and ID) with the different models of membranes, in order to have more insights, to understand better the full TLR4 dimer.

6.4   Materials and Methods

Recently has been developed a new generation of all-atom phospholipid force fields for MD simulations of membrane bilayers (Amber Lipid Force Field).[20]

Lipid bilayers were set up and molecular dynamics run with Amber and the Lipid14 force field. Amber 14 includes Lipid14 (Table **6.1**) a modular lipid force for tensionless lipid phospholipid simulations. Lipid14 includes the modular charge derivation framework developed in Lipid11 as well as a reparameterization of key van der Waals and dihedral angles as performed in GAFFlipid. [21 21 20 21]

Table **6.1**. Lipid14 residues name.

Lipid 14 Residues		
	Description	**LIPID14 Residue Name**
Acyl Chain	Lauroyl (12:0)	LA
	Myristoyl (14:0)	MY
	Palmitoyl (16:0)	PA
	Stearoyl (18:0)	ST
Head Group	Oleoyl (18:1 n-9)	OL
	Phosphatidylcholine	PC

The lipid bilayers structures were built using the CHARMM Membrane Builder GUI,[22] an internet based solution to generating lipid bilayer structures as well as membrane-bound protein structures. The membranes were created with a rectangular box, 22.5 water layers thick on the top and bottom of the system and 200 length of XY (Table **6.2**). All systems are modeled using suitable AMBER parameters, and converted to Lipid14 PDB format using the charmmlipid2amber.x script.[23] Formatted structure files were loaded into the program Leap, and parameters and topology were assigned. Glycerophospholipid parameters from Lipid14 were used for the lipids.

Table 6.2. Composition of the different models of membranes.

Membrane Model	Lipid Type	Upper Leaflet Number	Lower Leaflet Number	Average Area (Å²)	Area Upper (Å)	Area Inner (Å)
CHOL-POPC	CHOL	370	370	40071.00	200.18	200.18
	POPC	370	370			
POPE-POPC	POPE	315	315	40036.50	200.09	200.09
	POPC	315	315			
POPC	POPC	586	586	40023.80	200.06	200.06
DPPC-POPC	DPPC	305	305	40046.50	200.12	200.12
	POPC	305	305			
DPPC-POPE	DPPC	329	329	40072.20	200.18	200.18
	POPE	329	329			
ASYMMETRIC MODEL	CHOL DPPC DPPE POPE POPC DOPE DOPC	263	250	14221	119.25	119.25

Regarding TLR4/MD-2 system, an X-ray structure is available for the extracellular domain in complex with lipid A (PDB-ID: 3FXI). The intracellular and transmembrane domains have been modelled by our Computational Chemical Biology

lab (at CIB-CSIC) by homology modelling and were used for this study (Ref. Uniprot O00206).

The following protocol was used for molecular dynamics of lipid bilayer, lipid bilayer with the intracellular and transmembrane TLR4 domains, and finally the full system TLR4/MD-2: a minimization phase, following two steps of heat, holding to equilibrate periodic box dimensions and finally production with constant pressure.
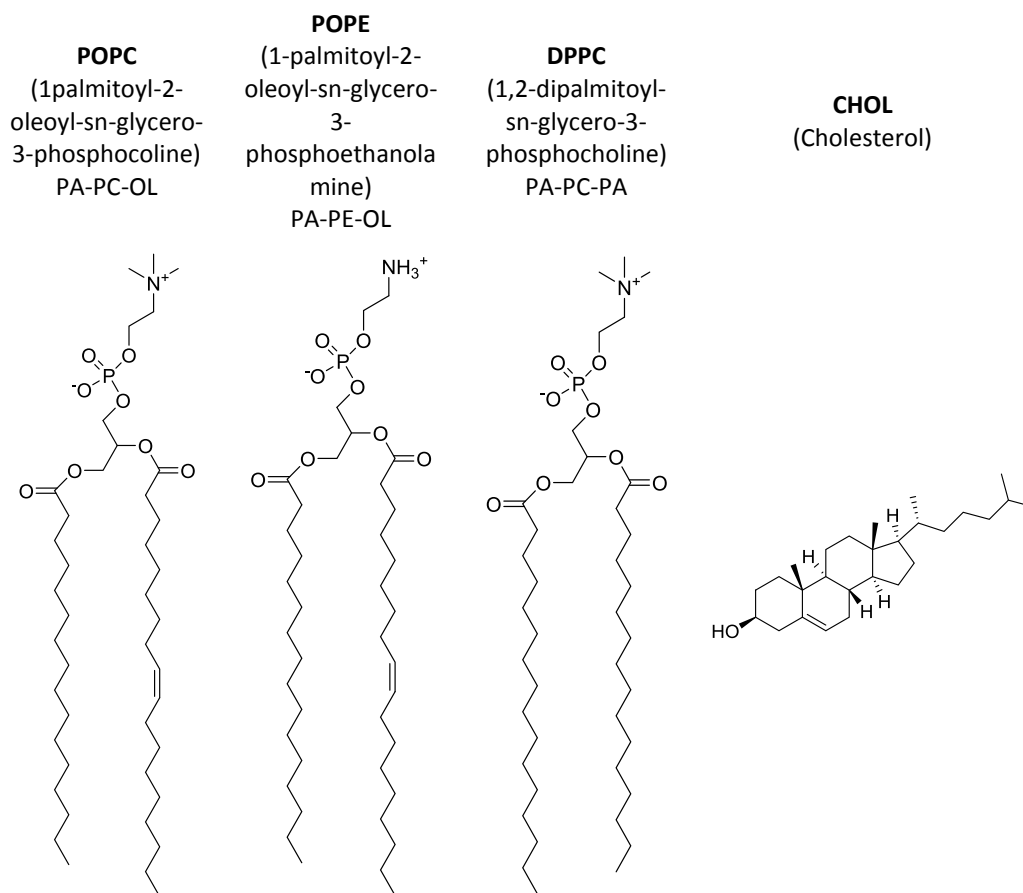
The full system was minimized for 10000 steps, of which the first 5000 steps used the steepest descent method and the remaining steps used the conjugate gradient method. The system was then heated from 0 K to 100 K using Langevin dynamics[24] for 5 ps at constant volume, with weak restraints on the lipid (force constant 10 kcal mol^{-1} Å^{-2}). The systems were heated through two sequential runs to 303 K while keeping the lipid fixed. First the systems were heated to 100 K and then slowly to the production temperature. The Langevin thermostat was used for the initial heating. The second phase of heating slowly increased the temperature to the desired production temperature. This time an anisotropic Berendsen weak-coupling barostat was used to also equilibrate the pressure in addition to the use of the Langevin thermostat to equilibrate the temperature.

To hold, in order to equilibrate the system's periodic boundary condition dimensions, it was necessary to run 5ns MD with a barostat. The system's dimensions and density must equilibrate before proceeding with production MD. Because the periodic boundary condition box dimensions are changing, it was necessary to increase the "skinnb" value and to restart the MD simulation after 500ns. This avoided most "skinnb" errors. If these cells change size, due to the box changing size, by too much then it will cause the code to halt with an error related to skinnb. Once the system was equibrated the box size fluctuations were small and so this was not an issue during production.During the production (100 ns), the temperature was controlled using the Langevin thermostat while pressure was controlled using the anisotropic Berendsen barostat.

The analysis was performed using the cpptraj module of AmberTools15.[15]

6.5 Annex VI

Annex Table 6.1. Different type of lipids and sterols.

| **POPC**
(1palmitoyl-2-oleoyl-sn-glycero-3-phosphocoline)
PA-PC-OL | **POPE**
(1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine)
PA-PE-OL | **DPPC**
(1,2-dipalmitoyl-sn-glycero-3-phosphocholine)
PA-PC-PA | **CHOL**
(Cholesterol) |

DPPE
(1,2-dipalmitoyl-sn-glycero-
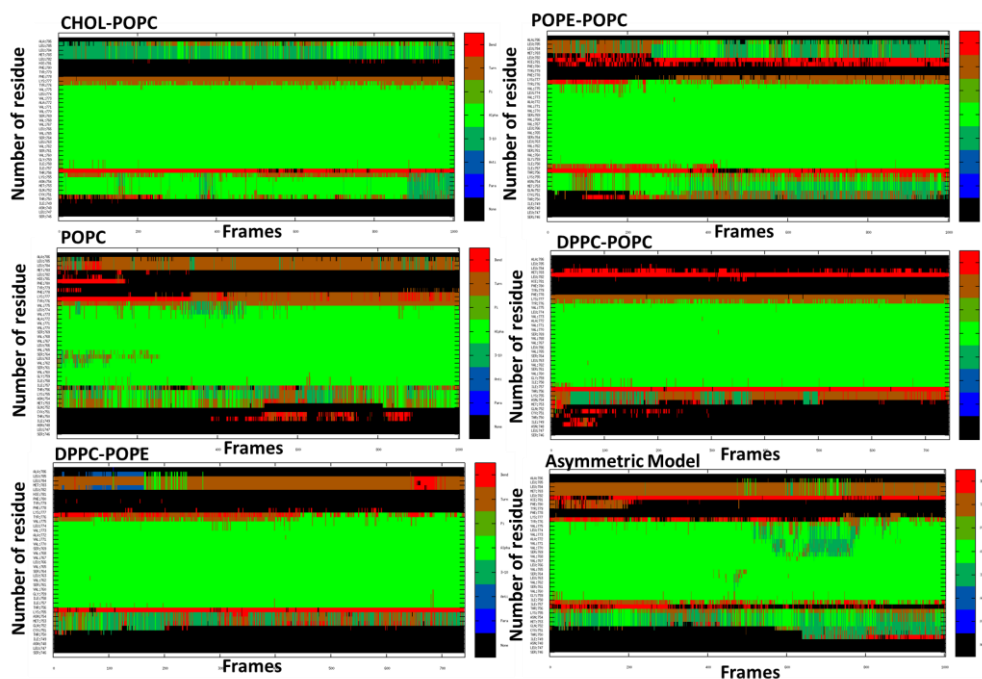3-phosphoethanolamine)
PA-PE-PA
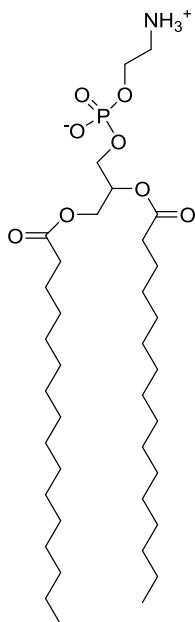
DOPE
(1,2-Dioleoyl-sn-glycero-3-
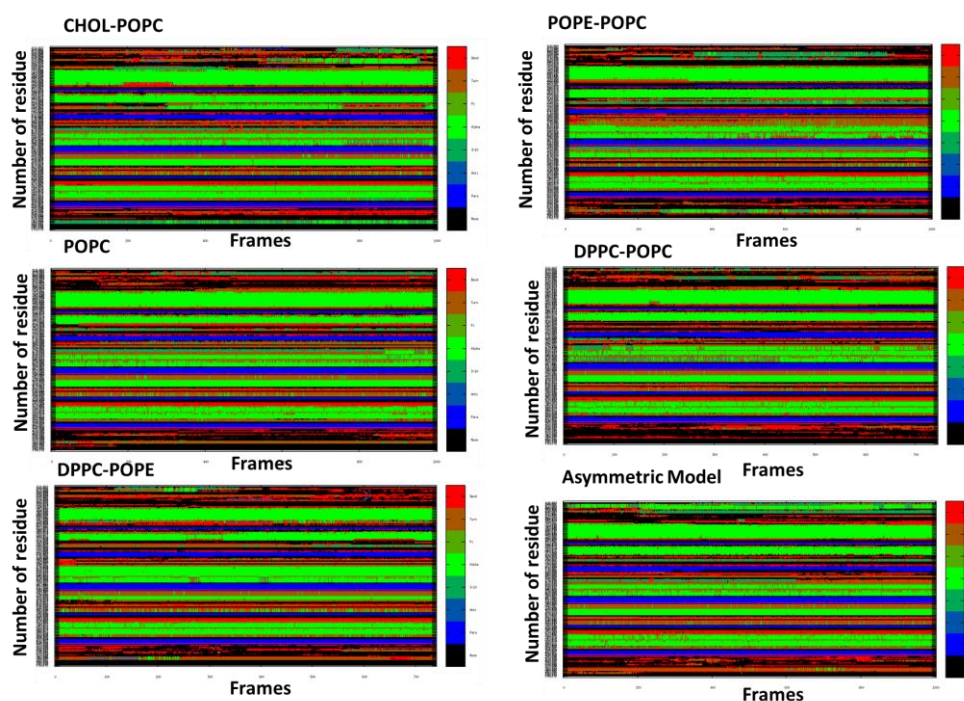phosphoethanolamine)
OL-PE-OL
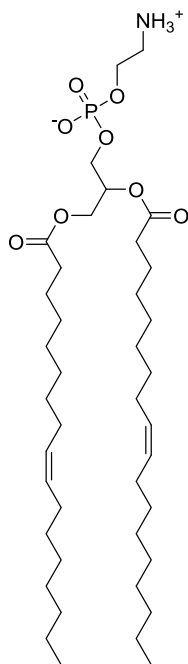
DOPC
(1,2-dipalmitoyl-sn-glycero-
3- phosphocholine)
OL-PC-OL
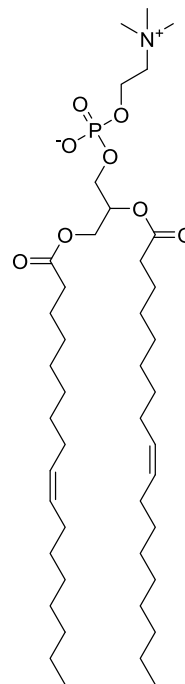
Annex Figure 6.1. The helicity of the transmembrane domain of TLR4 for all the membranes models.

Annex Figure 6.2. The helicity of the intracellular domain of TLR4 for all the membranes models.

Bibliography:

1.   Malanovic, N.; Lohner, K., Gram-positive bacterial cell envelopes: The impact on the activity of antimicrobial peptides. *Biochim. Biophys. Acta* **2016,** *1858* (5), 936-46.

2.   van Meer, G., Cellular lipidomics. *EMBO J.* **2005,** *24* (18), 3159-65.

3.   Holthuis, J. C.; Menon, A. K., Lipid landscapes and pipelines in membrane homeostasis. *Nature* **2014,** *510* (7503), 48-57.

4.   Ingolfsson, H. I.; Melo, M. N.; van Eerden, F. J.; Arnarez, C.; Lopez, C. A.; Wassenaar, T. A.; Periole, X.; de Vries, A. H.; Tieleman, D. P.; Marrink, S. J., Lipid organization of the plasma membrane. *J. Am. Chem. Soc.* **2014,** *136* (41), 14554-9.

5.   Phillips, R.; Ursell, T.; Wiggins, P.; Sens, P., Emerging roles for lipids in shaping membrane-protein function. *Nature* **2009,** *459* (7245), 379-385.

6.   van Meer, G.; Voelker, D. R.; Feigenson, G. W., Membrane lipids: where they are and how they behave. *Nat. Rev. Mol. Cell Biol.* **2008,** *9* (2), 112-24.

7.   Sampaio, J. L.; Gerl, M. J.; Klose, C.; Ejsing, C. S.; Beug, H.; Simons, K.; Shevchenko, A., Membrane lipidome of an epithelial cell line. *Proc. Natl. Acad. Sci. U. S. A.* **2011,** *108* (5), 1903-7.

8.   Klose, C.; Surma, M. A.; Simons, K., Organellar lipidomics--background and perspectives. *Curr. Opin. Cell. Biol.* **2013,** *25* (4), 406-13.

9.   Mouritsen, O. G.; Zuckermann, M. J., What's so special about cholesterol? *Lipids* **2004,** *39* (11), 1101-13.

10.   Stansfeld, P. J.; Sansom, M. S., Molecular simulation approaches to membrane proteins. *Structure (London, England : 1993)* **2011,** *19* (11), 1562-72.

11.   Laganowsky, A.; Reading, E.; Allison, T. M.; Ulmschneider, M. B.; Degiacomi, M. T.; Baldwin, A. J.; Robinson, C. V., Membrane proteins bind lipids selectively to modulate their structure and function. *Nature* **2014,** *510* (7503), 172-5.

12.   Chavent, M.; Duncan, A. L.; Sansom, M. S., Molecular dynamics simulations of membrane proteins and their interactions: from nanoscale to mesoscale. *Curr. Opin. Struct. Biol.* **2016,** *40*, 8-16.

13.   Piazza, M.; Calabrese, V.; Damore, G.; Cighetti, R.; Gioannini, T.; Weiss, J.; Peri, F., A synthetic lipid A mimetic modulates human TLR4 activity. *ChemMedChem* **2012,** *7* (2), 213-7.

14.   Neve, J. E.; Wijesekera, H. P.; Duffy, S.; Jenkins, I. D.; Ripper, J. A.; Teague, S. J.; Campitelli, M.; Garavelas, A.; Nikolakopoulos, G.; Le, P. V.; de, A. L. P.; Pham, N. B.; Shelton, P.; Fraser, N.; Carroll, A. R.; Avery, V. M.; McCrae, C.; Williams, N.; Quinn, R. J., Euodenine A: a small-molecule agonist of human TLR4. *J. Med. Chem.* **2014,** *57* (4), 1252-75.

15.   Roe, D. R.; Cheatham, T. E., PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* **2013,** *9* (7), 3084-3095.

16.   Kucerka, N.; Nieh, M. P.; Katsaras, J., Fluid phase lipid areas and bilayer thicknesses of commonly used phosphatidylcholines as a function of temperature. *Biochim. Biophys. Acta* **2011,** *1808* (11), 2761-71.

17.   Kucerka, N.; Tristram-Nagle, S.; Nagle, J. F., Structure of fully hydrated fluid phase lipid bilayers with monounsaturated chains. *J. Membr. Biol.* **2005,** *208* (3), 193-202.

18.   Rappolt, M.; Hickel, A.; Bringezu, F.; Lohner, K., Mechanism of the Lamellar/Inverse Hexagonal Phase Transition Examined by High Resolution X-Ray Diffraction. *Biophys. J.* **2003,** *84* (5), 3111-3122.

19.   Kabsch, W.; Sander, C., Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **1983,** *22* (12), 2577-637.

20.   Dickson, C. J.; Madej, B. D.; Skjevik, Å. A.; Betz, R. M.; Teigen, K.; Gould, I. R.; Walker, R. C., Lipid14: The Amber Lipid Force Field. *J. Chem. Theory Comput.* **2014,** *10* (2), 865-879.

21.   Dickson, C. J.; Rosso, L.; Betz, R. M.; Walker, R. C.; Gould, I. R., GAFFlipid: a General Amber Force Field for the accurate molecular dynamics simulation of phospholipid. *Soft Matter* **2012,** *8* (37), 9617-9627.

22.   Jo, S.; Lim, J. B.; Klauda, J. B.; Im, W., CHARMM-GUI Membrane Builder for mixed bilayers and its application to yeast membranes. *Biophys. J.* **2009,** *97* (1), 50-8.

23.   Case, D. A.; Darden, T. A.; Cheatham, T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K. M.; Roberts, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Swails, J.; Götz, A. W.; Kolossváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wolf, R. M.; Liu, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Cai, Q.; Ye, X.; Wang, J.; Hsieh, M.-J.; Cui, G.; Roe, D. R.; Mathews, D. H.; Seetin, M. G.; Salomon-Ferrer, R.; Sagui, C.; Babin, V.; Luchko, T.; Gusarov, S.; Kovalenko, A.; Kollman, P. A. *AMBER 12*, University of California, San Francisco., 2012.

24.   Pastor, R. W.; Brooks, B. R.; Szabo, A., An analysis of the accuracy of Langevin and molecular dynamics algorithms. *Mol. Phys.* **1988,** *65* (6), 1409-1419.

CHAPTER 7

Conclusions

CONCLUSIONS

In this Thesis, we have contributed to the elucidation of the molecular recognition events involving TLR4/MD-2 and some modulators at atomic level. This work has been carried out by means of computational methodologies, such as MD simulations, ligand-protein and protein-protein docking, virtual screening, and membrane simulations. Ligand-protein docking and virtual screening has been employed as a source of new compounds able to modulate the TLR4 behavior with possible therapeutic applications. We have combined the computational work with experimental biological assays, and a fruitful collaboration with Prof. Peri´s group for the synthesis of compounds. Overall, we have helped to provide new insights for the understanding of the molecular recognition events underlying the biological functions of the TLR4.

Theoretical binding modes have been predicted for reported modulators of the TLR4/MD-2 system, with agonist and antagonist activity. In particular, we focused our work in synthetic glycolipids and non LPS-like molecules. For all these TLR4 modulators, it is clear that, despite their different chemical structure, they must share a common pattern of interactions with TLR4. We have undertaken a computational study of some representative compounds to unveil some of these patterns of interactions.

The cationic glycolipid IAXO-102, a potent TLR4 antagonist targeting both MD-2 and CD14 co-receptors, has been used as scaffold to design new potential TLR4 modulators and fluorescent labels for the TLR4 receptor complex (membrane TLR4/MD-2 dimer and CD14). Our modelling studies have led to the proposal of 3D models for the interaction with CD14 and TLR4/MD-2 accounting for their binding properties and also for their antagonistic activity.

To propose new chemical scaffolds for the development of new ligands able to modulate TLR4 functions, we have performed virtual screening. Virtual screening strategies from commercial and in-house libraries, followed by biological assays, have allowed us to identify new chemical entities for the development of novel TLR4

modulators with a LPS-non-related structure. So far, we have identified seven novel compounds with a promising TLR4 antagonist activity.

The computational study of the full TLR4/MD-2 heterodimer was performed, simulating the full complex inserted in the membrane environment. The analysis of the molecular dynamics simulations led us to understand the key interactions implicated in the dimerization process at atomic level.

CONCLUSIONES

En esta Tesis, hemos contribuido a la elucidación de los eventos de reconocimiento molecular del complejo TLR4/MD-2 y algunos moduladores a nivel atómico. Este trabajo se ha realizado por medio de metodologías computacionales, como simulaciones de dinámica molecular, acoplamiento ligando-proteína, cribado virtual y simulaciones de la membrana. Técnicas de acoplamiento ligando-proteína y cribado virtual han sido utilizadas, dando lugar a una fuente de nuevos compuestos capaces de modular el comportamiento de los TLRs con posibles aplicaciones terapéuticas, y también como sondas biológicas.

Hemos combinado el trabajo computacional con ensayos biológicos experimentales y hemos llevado a cabo una fructífera colaboración con el grupo del Prof. Peri para la síntesis de compuestos. En general, hemos ayudado a proporcionar nuevas ideas para la comprensión de los eventos de reconocimiento molecular subyacente a las funciones biológicas de los receptores TLR4.

Se han predicho modos de unión teórica para los moduladores descritos del sistema TLR4/MD-2. En particular, hemos centrado nuestro trabajo en glicolípidos sintéticos y moléculas que no tienen estructura de tipo LPS. Para todos estos moduladores de TLR4, está claro que, a pesar de su diferente estructura química, deben compartir un patrón común de interacciones con el receptor TLR4. Hemos emprendido un estudio computacional de algunos compuestos representativos para revelar algunos de estos patrones de interacciones.

El glicolípido catiónico IAXO-102, un potente antagonista de TLR4, se ha utilizado como esqueleto para diseñar nuevos moduladores de TLR4 y marcadores fluorescentes para el complejo TLR4. Nuestros estudios de modelización nos han permitido diseñar nuevos compuestos y a proponer modelos de interacción tanto para el CD14 como para el complejo TLR4/MD-2. Todos estos compuestos han presentado una actividad antagonista para el complejo del TLR4.

El estudio computacional del heterodímero TLR4/MD-2 completo se realizó simulando el complejo completo insertado en la membrana. El análisis de las

simulaciones de dinámica molecular nos llevó a entender las interacciones claves implicadas en el proceso de dimerización a nivel atómico.