UNIVERSITÁ DEGLI STUDI DI MILANO-BICOCCA
Facoltà di Scienze Matematiche, Fisiche e Naturali
Dipartimento di Biotecnologie e Bioscienze
Dottorato in Biotecnologie Industriali – XXVI Ciclo

Design, synthesis and biological characterization of new small-molecule TLR4 modulators

Tutor: Prof. Francesco Peri

Roberto Cighetti
Matr. 073690

Anno Accademico 2012-2013

Al mio caro nonno,
al mio caro zio

E certamente diverrà come un albero
piantato presso ruscelli d'acqua,
Che dà il suo proprio frutto nella sua stagione
E il cui fogliame non appassisce,
E ogni cosa che fa riuscirà.

Salmo 1:3

Summary

Abstract

Activation of Toll-like receptor 4 (TLR4) and subsequent intracellular signaling in response to minute amounts of circulating endotoxins (Gram-negative bacterial lipopolysaccharides, LPS), results in the rapid triggering of proinflammatory processes necessary for optimal host immune responses to invading Gram-negative bacteria in mammalians.

TLR4 does not bind directly to LPS, and TLR4 activation by endotoxin is a complex event, involving the participation of other LPS-binding proteins — namely LBP, CD14, and MD-2 — and ending with the formation of the activated $(TLR4{\cdot}MD\text{-}2{\cdot}LPS)_2$ complex. In particular, CD14 was the first identified Pattern Recognition Receptor (PRR) that binds directly to LPS, and chaperones the formation of the $(TLR4{\cdot}MD\text{-}2{\cdot}LPS)_2$ complex. CD14 is also required for endotoxin-induced TLR4 endocytosis and relocalization of the entire LPS receptor complex into the endosome, where a second signaling pathway initiates.

Excessively potent and deregulated TLR4 activation and signaling causes severe syndromes such as septic shock, associated with a high mortality (50–70%), and organ-specific damages. Efficient and selective TLR4 antagonists with chemical structures simpler than that of lipid A are therefore required for the development of potential new drugs with a wide array of medical and pharmacological applications (from sepsis to CNS pathologies).

In this thesis, the synthesis of three monosaccharide LPS mimetics based on a glucosamine core linked to two fatty acid chains and bearing one or two phosphate groups is described. Two of them, each with one phosphate group,

resulted quite active in inhibiting LPS-induced TLR4 signaling and cytokine production in HEK-Blue™ cells and murine macrophages. The compound with two phosphate groups, however, was found to be more active in efficiently inhibiting TLR4 signal in both cell types.

The direct interaction between this compound and the MD-2 receptor was investigated by NMR spectroscopy and molecular modeling/docking analysis. This compound also interacts directly with the CD14 receptor, stimulating its internalization by endocytosis. The dual targeting of MD-2 and CD14, accompanied by good solubility in water and lack of toxicity, suggests the use of monosaccharide 3 as a lead compound for the development of drugs directed against TLR4-related syndromes.

A derivatization of this compound was also started in order to couple it to a fluorescent moiety or to a Gadolinium-chelating agent, for microscopy and MRI studies respectively.

Riassunto

L'attivazione del *Toll-Like Receptor* 4 (TLR4) e il conseguente *signaling* intracellulare in risposta a esigue quantità di endotossine (lipopolisaccaridi di batteri Gram-negativi, LPS) presenti nel circolo sanguigno porta al rapido innesco di processi pro-infiammatori necessari per una risposta immunitaria ottimale dell'organismo ospite contro l'invasione di batteri Gram-negativi.

Il TLR4 non lega però direttamente i LPS e la sua attivazione da parte dell'endotossina è un processo complesso che prevede la partecipazione di altre proteine che legano i LPS (LBP, CD14 e MD-2) e porta infine alla formazione del complesso attivato $(TLR4{\cdot}MD\text{-}2{\cdot}LPS)_2$. In particolare, CD14 è stato il primo recettore che lega LPS ad essere scoperto e sembra che favorisca la formazione del complesso $(TLR4{\cdot}MD\text{-}2{\cdot}LPS)_2$. CD14 è anche necessario per l'endocitosi di TLR4 indotta da LPS e la localizzazione dell'intero complesso recettoriale per l'endotossina negli endosomi, da dove parte un secondo *pathway* del segnale.

Un'attivazione di TLR4 eccessivamente potente e sregolata causa però gravi sindromi come lo shock settico, associato a un'alta mortalità (50-70%) e a danni agli organi anche molto gravi. Sono perciò necessari degli antagonisti (inibitori) di TLR4 efficaci e selettivi con una struttura chimica più semplice di quella del LPS, per lo sviluppo di potenziali nuovi farmaci con un'ampia gamma di applicazioni mediche e farmacologiche (dalla sepsi alle patologie del sistema nervoso).

In questa tesi è descritta la sintesi di monosaccaridi mimetici del LPS basati su un *core* di glucosammina legata a due catene di acidi grassi che portano uno o

due gruppi fosfato. Di questi, i due composti con un solo fosfato hanno mostrato una certa attività come inibitori dell'attivazione di TLR4 e della produzione di citochine in cellule HEK-Blue® e in macrofagi murini. Il composto con due gruppi fosfato è risultato invece molto più efficace nell'inibire il segnale di TLR4 in entrambi i modelli cellulari.

L'interazione diretta tra questo composto e il recettore MD-2 è stata investigata mediante spettroscopia NMR e analisi bioinformatiche (*molecular modelling/docking*). Questo composto interagisce direttamente anche con il recettore CD14, inducendone l'internalizzazione per endocitosi.

Il *targeting* sia di CD14 che di MD-2, unito a una buona solubilità in acqua e all'assenza di tossicità, rendono il monosaccaride difosfato presentato in questa tesi un possibile *lead compound* per lo sviluppo di farmaci contro le sindromi collegate al TLR4.

E' stata inoltre avviata una funzionalizzazione di questo composto in modo da coniugarlo a una molecola fluorescente o a un chelante del gadolinio per studi, rispettivamente, di microscopia o *imaging* mediante risonanza magnetica.

Introduction

Gram stain and bacterial classification

Bacteria are the most common and widespread human pathogens. For classification purposes they can be divided into two groups, depending on a microbiology stain protocol, known as the Gram staining (2). Gram-negative bacteria are those bacteria that do not retain crystal violet dye after the staining. In a Gram stain test, a counterstain (commonly safranin) is added after the crystal violet, coloring all Gram-negative bacteria with a red or pink color. On the other hand, Gram-positive bacteria will retain the crystal violet dye when washed in a decolorizing solution (**Figure 1**). This test is useful in classifying two distinct types of bacteria based on their differences, that reside mainly in the structure of their cell envelope.

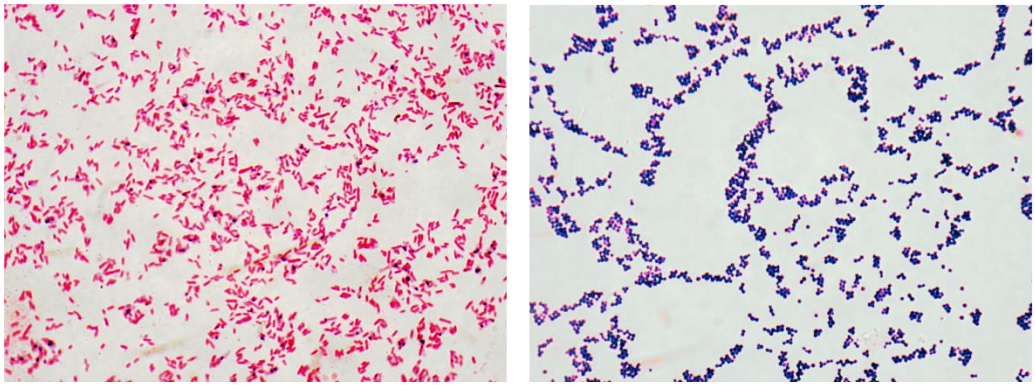


Figure 1. Gram staining of *Escherichia coli* (Gram-negative, pink, on the left) and *Staphylococcus sp.* (Gram-positive, violet-blue, on the right) (3)

Gram-positive bacteria have a cell envelope that consists of a single membrane and a thick layer of peptidoglycan (or murein, a complex polysaccharide) known as cell wall. Gram-negative bacteria instead have a cell wall that is made up of two membranes (the inner and the outer membrane)

that enclose a region known as the periplasmic space (or periplasm), in which lies a thin peptidoglycan layer, which is linked with the overlaying outer membrane. The cell wall of Gram-positive bacteria tends to be 2 to 8 times as thick as the Gram-negative wall.

When thin sections of Gram-negative bacteria are viewed in the transmission electron microscope (TEM), the membranes can be distinguished and they appear as electron-dense dark stripes. There are two parallel thickly stained lines separated by an almost transparent region. The dark regions are the charged head groups of molecules called phospholipids (**Figure 2**).

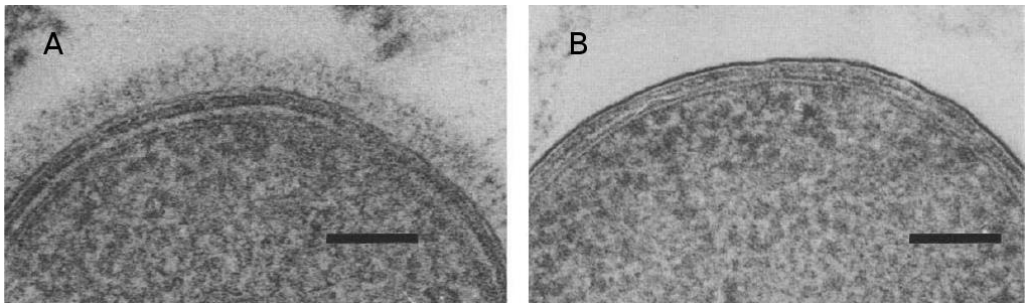


Figure 2. *Pseudomonas aeruginosa* cells. Scale markers indicate 100 nm. Figure A shows serotype PAO1, which has a longer LPS O-chain, visible as the fringe above the outer membrane, while Figure B shows serotype PFV, which has a shortest O-antigen.

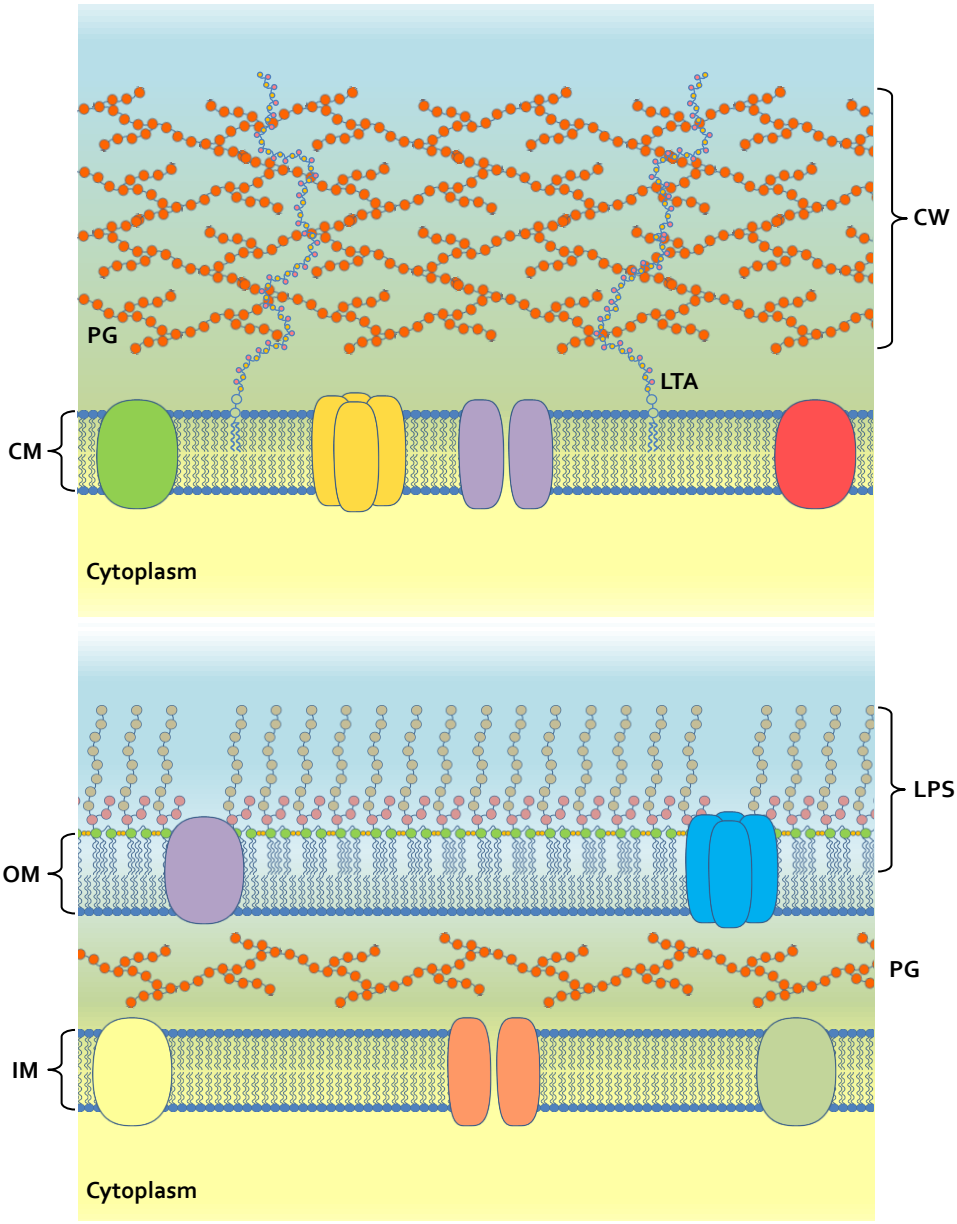


Figure 3. Scheme of the bacterial envelope of Gram-positive (up) or Gram-negative bacteria (down). In the upper figure , the cell membrane (CM), the cell wall (CW), peptidoglycan (PG) and lipoteichoic acid (LTA) are shown. The bottom figure shows the inner membrane (IM) and the outer membrane (OM), the thin PG layer into the periplasm and the location of LPS, component of the outer leaflet of the OM. Self work.

Phospholipids make up the bulk of bacterial membranes. In Gram-positive bacteria and in the inner membrane of Gram-negative bacteria the phospholipids are arranged fairly evenly on either "leaflet" of the membrane. In contrast, the outer membrane of Gram-negative bacteria is asymmetric, with the majority of the phospholipids located at the inner leaflet of the membrane, while the outer leaflet is mainly constituted by a glycolipid termed lipopolysaccharide (LPS), beside some phospholipids and proteins (**Figure 3**).

The asymmetrical arrangement of the Gram-negative outer membrane confers various functions to the bacterium. Proteins allow the diffusion of compounds across the outer membrane, as long as they can fit into the pore that runs through the center of the protein. In addition, other proteins function to specifically transport compounds to the inside of the bacterium in an energy-dependent manner. The LPS component of the outer membrane is capable of various chemical arrangements that can influence the bacterium's ability to elude host immune defenses. When it's released from the bacterium, LPS is also referred to as endotoxin, as it can be toxic to mammals, including humans.

A two-membrane system allows the existence of the periplasm. This was once thought to be just functionless empty space. Now, however, the periplasm is known to have very important functions in the survival and operation of the bacterium. The region acts as a buffer between the external environment and the interior of the bacterium. Specialized transport proteins and enzymes are located exclusively in this region, like environment-sensing proteins that determine the response of a bacterium to environmental signals, as occurs during chemotaxis.

The lipopolysaccharide

The lipopolysaccharide (LPS), synthesized by most wild-type (wt) Gram-negative bacteria (S-form LPS), is the major constituent of the outer leaflet of their outer membrane. This macromolecule consists of three regions: the O-polysaccharide chain, which is made up of repeating oligosaccharide units, the core oligosaccharide and the lipid A, which is responsible for the endotoxic activity of the entire molecule. (6; 7; 8) (**Figure 4**).

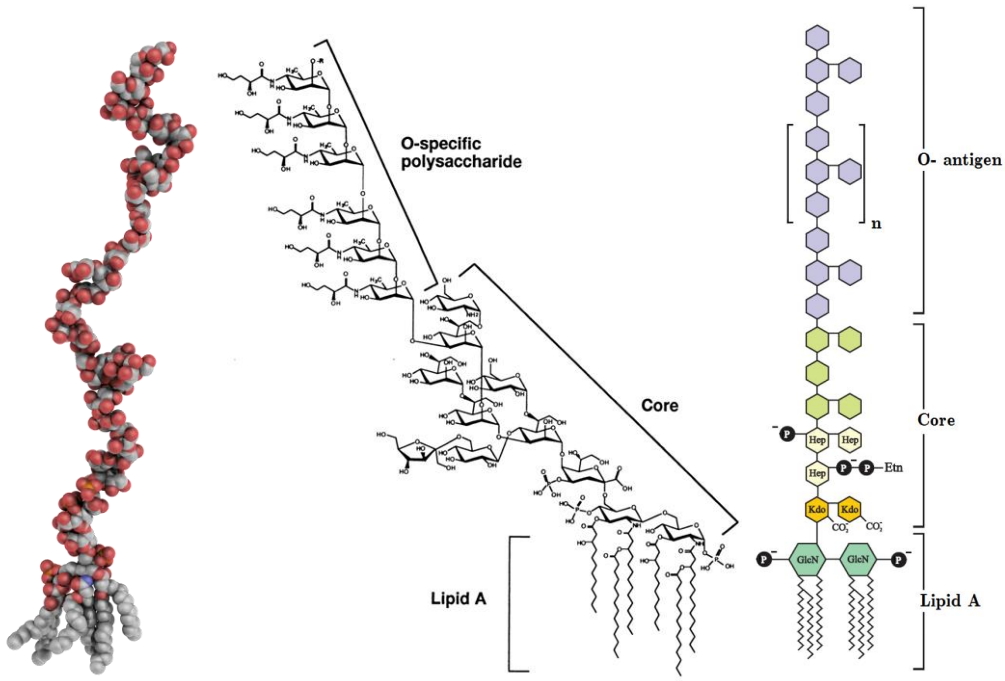


Figure 4. General structure of bacterial LPS. The O-specific antigen is variable among species. Lipid A is a conserved structure, although some modifications of its structure have been reported. A pictorial sphere model of LPS is also shown. Self work from PDB ID: 1QFG. All the molecular images are obtained with: The PyMOL Molecular Graphics System, Version 1.3, Schrödinger, LLC.

Some mutant bacteria that are defective of some enzymes involved in the synthesis of LPS form colonies described as "rough". On the other hand, the wild-type bacteria form colonies that appear "smooth" because of the large amount of sugars contained in their membranes. The different LPS types that were extracted from the colonies and chemically characterized, were then named after the appearance of the colonies: the LPS with the longest polysaccharide O-antigen was named S-LPS and the truncated version with little or no O-antigen sugars become the R-form LPS.

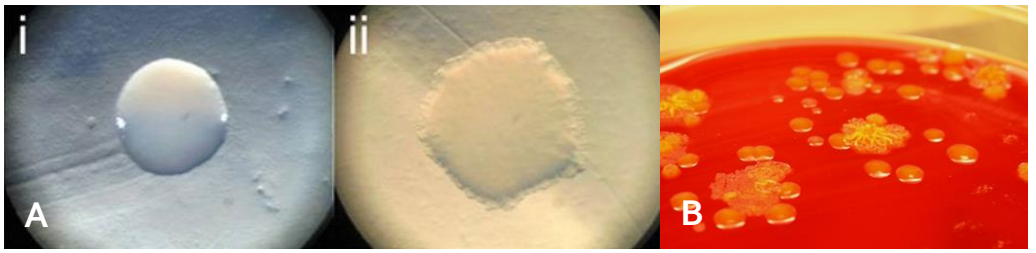


Figure 5. A) Micrograph of a smooth (i) and a rough (ii) colony of *Escherichia coli*. The rough strain carries a mutation in a gene involved in LPS biosynthesis. (6) B) Appearance of smooth and rough colonies of *Pseudomonas luteola* in a blood culture. (7)

It was later discovered that not only the O-antigen was subjected to mutations, but also the core-oligosaccharide may be present in different degrees of completion, depending on the class ("semi-rough", SR, or "rough", Ra to Re) to which the mutant belongs (7; 8) (**Figure 6A**).

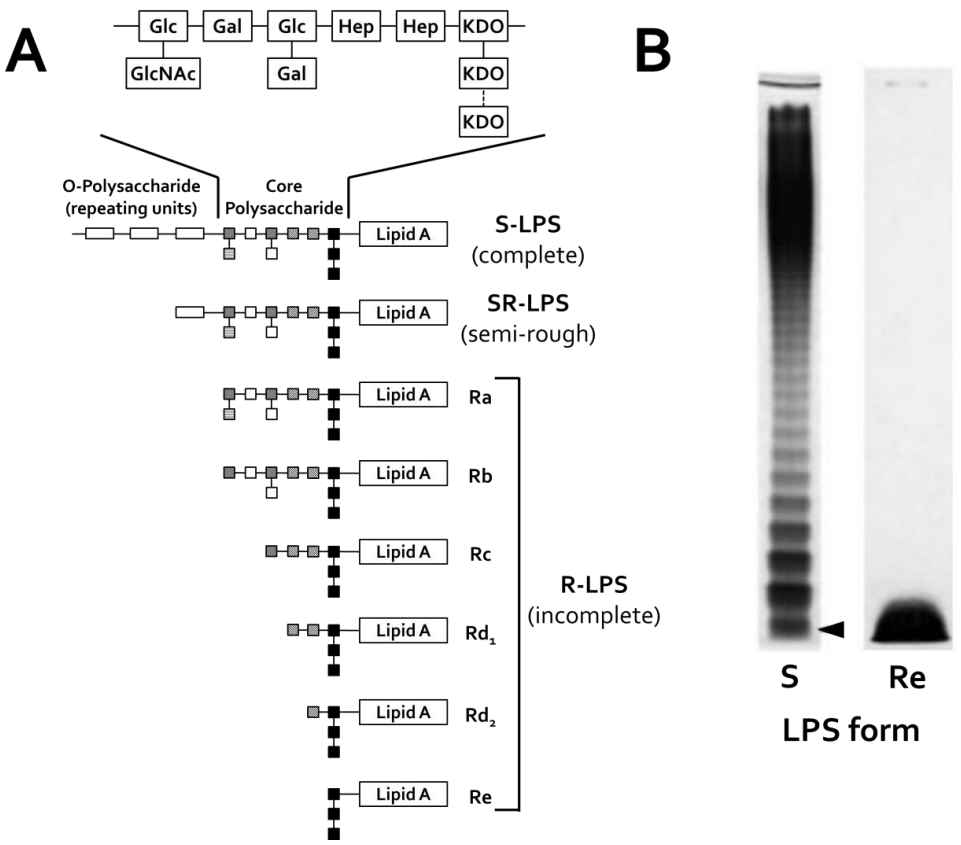


Figure 6. Schematic representation of the different LPS chemotypes. (A) Depending on the completeness of the LPS molecule, different R-form LPS, SR-form LPS, and S-form LPS can be distinguished. Abbreviations: KDO, 2-keto-3-desoxyoctonate; Hep, heptose; Glc, glucose; Gal, galactose; GlcNac, *N*-acetyl-glucosamine. (B) LPS from *S. abortus equi* and *S. minnesota* R595 was extracted, purified and size-separated by SDS-PAGE. The arrowhead depicts the R-form contained in the S-form preparation. Adapted from (9).

Notably, LPS from wt bacteria are always highly heterogeneous mixtures of S-form LPS molecules containing 1 to over 50 repeating oligosaccharide units and contain ubiquitously a varying proportion of R-form molecules lacking the O-specific chain (**Figure 6B**). Many clinically relevant Gram-negative bacteria

synthesize this type of LPS. Further, Gram-negative wt bacteria, such as *Chlamydia* and *Neisseria* synthesize LPS in which the number of sugar residues is highly reduced, and thus resemble R-form LPS, at least in their physicochemical properties. LPS are amphipathic molecules whose hydrophobicity decreases with increasing length of the sugar part (7; 8). Based upon these differences, S- and R-form LPS show marked differences in the kinetics of their blood clearance and cellular uptake as well as in the ability to induce different immune responses.

The immune system

The immune system is a remarkably versatile defense system that protect animals from invading pathogenic microorganisms and cancer. It is able to generate an enormous variety of cells and molecules capable of specifically recognizing and eliminating a limitless variety of foreign invaders. These cells and molecules act together in a dynamic network whose complexity rivals that of the nervous system.

The immune system includes innate and adaptive components

Immunity—the state of protection from infectious disease —has both a less specific and more specific component. The less specific component, innate immunity, provides the first line of defense against infection. Most components of innate immunity are present before the onset of infection and constitute a set of disease-resistance mechanisms that are not specific to a particular pathogen but that include cellular and molecular components that recognize classes of molecules peculiar to frequently encountered pathogens. Phagocytic cells, such as macrophages and neutrophils, barriers such as skin, and a variety of antimicrobial compounds synthesized by the host all play important roles in innate immunity. In contrast to the broad reactivity of the innate immune system, which is uniform in all members of a species, the specific component, adaptive immunity, does not come into play until there is an antigenic challenge to the organism. Adaptive immunity responds to the challenge with a high degree of specificity as well as the remarkable property of "memory". Typically, there is an adaptive immune response against an antigen within five or six days after the initial exposure to that antigen.

Exposure to the same antigen some time in the future results in a memory response: the immune response to the second challenge occurs more quickly than the first, is stronger, and is often more effective in neutralizing and clearing the pathogen. The major agents of adaptive immunity are lymphocytes and the antibodies and other molecules they produce. Because adaptive immune responses require some time to marshal, innate immunity provides the first line of defense during the critical period just after the host's exposure to a pathogen. In general, most of the microorganisms encountered by a healthy individual are readily cleared within a few days by defense mechanisms of the innate immune system before they activate the adaptive immune system. In this thesis only the innate immunity system will be discussed.

Innate Immunity

The defensive barriers of Innate Immunity can be divided into four types: anatomic, physiologic, phagocytic, and inflammatory.

The skin and the mucosal surfaces: anatomic protective barriers against infection

Physical and anatomic barriers that tend to prevent the entry of pathogens are an organism's first line of defense against infection. The skin and the surface of mucous membranes are included in this category because they are effective barriers to the entry of most microorganisms. The pH of the skin is maintained between 3 and 5; this pH inhibits the growth of most microorganisms. Breaks in the skin resulting from scratches, wounds, or abrasion are obvious routes of infection. The skin may also be penetrated by biting insects (e.g., mosquitoes, mites, ticks, fleas, and sandflies); if these harbor pathogenic organisms, they can introduce the pathogen into the body as they feed (malaria, bubonic plague and Lyme disease, for example).

The conjunctivae and the alimentary, respiratory, and urogenital tracts are lined by mucous membranes, not by the dry, protective skin that covers the exterior of the body. Although many pathogens enter the body by binding to and penetrating mucous membranes, a number of nonspecific defense mechanisms tend to prevent this entry. For example, saliva, tears, and mucous secretions act to wash away or entrap potential invaders and also contain antibacterial or antiviral substances. In addition, nonpathogenic organisms tend to colonize the epithelial cells of mucosal surfaces. These

normal flora generally outcompete pathogens for attachment sites on the epithelial cell surface and for necessary nutrients. Some organisms can escape these defense mechanisms and thus are able to invade the body through mucous membranes. Adherence of bacteria to mucous membranes is due to interactions between hairlike protrusions on a bacterium, called fimbriæ or pili, and certain glycoproteins or glycolipids that are expressed only by epithelial cells of the mucous membrane of particular tissues. For this reason, some tissues are susceptible to bacterial invasion, whereas others are not.

Physiologic barriers to infection

The physiologic barriers that contribute to innate immunity include temperature, pH, and various soluble and cell-associated molecules. Many species are not susceptible to certain diseases simply because their normal body temperature inhibits growth of the pathogens. Gastric acidity is an innate physiologic barrier to infection because very few ingested microorganisms can survive the low pH of the stomach contents. A variety of soluble factors contribute to innate immunity, among them the soluble proteins lysozyme, interferon, and complement. Lysozyme, a hydrolytic enzyme found in mucous secretions and in tears, is able to cleave the peptidoglycan layer of the bacterial cell wall. Interferon comprises a group of proteins typically produced by virus-infected cells. Among the many functions of the interferons is the ability to bind to nearby cells and induce a generalized antiviral state. Complement is a group of serum proteins that circulate in an inactive state. A variety of specific and nonspecific immunologic mechanisms can convert the inactive forms of complement proteins into an active state with the ability to damage the membranes of pathogenic organisms, either

destroying the pathogens or facilitating their clearance. Complement may function as an effector system that is triggered by binding of antibodies to certain cell surfaces, or it may be activated by reactions between complement molecules and certain components of microbial cell walls. Reactions between complement molecules or fragments of complement molecules and cellular receptors trigger activation of cells of the innate or adaptive immune systems. Many of the molecules involved in innate immunity have the property of pattern recognition, the ability to recognize a given class of molecules. Because there are certain types of molecules that are unique to microbes and never found in multicellular organisms, the ability to immediately recognize and combat invaders displaying such molecules is a strong feature of innate immunity. Molecules with pattern recognition ability may be soluble, like lysozyme and the complement components described above, or they may be cell-associated receptors.

Cells that ingest pathogens: a phagocytic barrier to infection

Another important innate defense mechanism is the ingestion of extracellular particulate material by phagocytosis. Phagocytosis is one type of endocytosis, the general term for the uptake by a cell of material from its environment. Most phagocytosis is conducted by specialized cells, such as blood monocytes, neutrophils, and tissue macrophages.

Most cell types are capable of other forms of endocytosis, such as receptor-mediated endocytosis, in which extracellular molecules are internalized after binding by specific cellular receptors, and pinocytosis, the process by which cells take up fluid from the surrounding medium along with any molecules contained in it. Ligation of many of the cell-surface receptors that recognize pathogens leads to phagocytosis of the pathogen, followed by its death inside the phagocyte. (1)

Phagocytosis is an active process, in which the bound pathogen is first surrounded by the phagocyte membrane and then internalized in a membrane-bounded vesicle known as a phagosome, which is then acidified (pH= ~3.5–4.0). In addition to being phagocytic, macrophages and neutrophils have granules, called lysosomes, that contain enzymes, proteins, and peptides that can mediate an intracellular antimicrobial response. The phagosome fuses with one or more lysosomes to generate a phagolysosome in which the lysosomal contents are released to destroy the pathogen. Upon phagocytosis, macrophages and neutrophils also produce a variety of other toxic products that help kill the engulfed microorganism (Reactive Oxygen Species, Nitric Oxides, antimicrobial peptides and enzymes). (**Figure 7**)

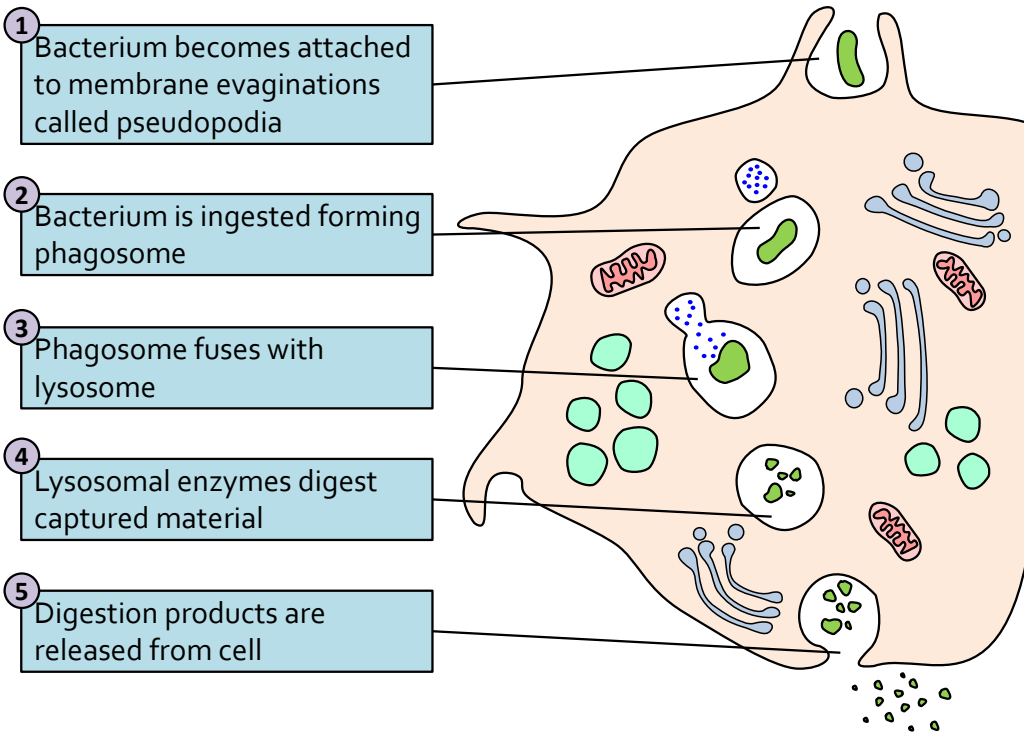


Figure 7. The phagocytosis process. Self work.

The most important of these antibacterial molecules are hydrogen peroxide (H_2O_2), the superoxide anion (O_2^-), and nitric oxide (NO), which are directly toxic to bacteria. They are generated by lysosomal NADPH oxidases and other enzymes in a process known as the respiratory burst, as it is accompanied by a transient increase in oxygen consumption. Neutrophils are short-lived cells, dying soon after they have accomplished a round of phagocytosis. Macrophages, on the other hand, are long-lived and continue to generate new lysosomes. Macrophages can make this response immediately on encountering an infecting microorganism and this can be sufficient to prevent an infection from becoming established.

A key feature that distinguishes pathogenic from nonpathogenic micro-organisms is their ability to overcome innate immune defenses. Pathogens display a variety of strategies to avoid being immediately destroyed by macrophages. Many extracellular pathogenic bacteria are coated with a thick polysaccharide capsule that is not recognized by any phagocyte receptor. Other pathogens, for example mycobacteria, can grow inside macrophage phagosomes by inhibiting fusion with a lysosome. Without such devices, a microorganism must enter the body in sufficient numbers to simply overwhelm the immediate innate host defenses and establish a focus of infection.

Another important effect of the interaction between pathogens and tissue macrophages is activation of macrophages to release cytokines and other mediators that set up a state of inflammation in the tissue and bring neutrophils and plasma proteins to the site of infection.

Receptors that signal the presence of pathogens and induce cytokines also have another important role. This is to induce the expression of so-called co-stimulatory molecules on both macrophages and dendritic cells, another type of phagocytic cell present in tissues, thus enabling these cells to initiate an adaptive immune response. The cytokines released by macrophages make an important contribution both to local inflammation and to other induced but nonadaptive responses that occur in the first few days of a new infection. (10)

Inflammation: a complex sequence of events that stimulates immune responses

Tissue damage caused by a wound or by an invading pathogenic microorganism induces a complex sequence of events collectively known as the inflammatory response. As described above, a molecular component of a microbe, such as LPS, may trigger an inflammatory response via interaction with cell surface receptors. The end result of inflammation may be the marshalling of a specific immune response to the invasion or clearance of the invader by components of the innate immune system.

Inflammation plays three essential roles in combating infection. The first is to deliver additional effector molecules and cells to sites of infection to augment the killing of invading microorganisms by the front-line macrophages. The second is to provide a physical barrier preventing the spread of infection, and the third is to promote the repair of injured tissue. Inflammation at the site of infection is initiated by the response of macrophages to pathogens.

Many of the classic features of the inflammatory response were described as early as 1600 BC, in Egyptian papyrus writings. In the first century AD, the Roman physician Celsus described the "four cardinal signs of inflammation" as *rubor* (redness), *tumor* (swelling), *calor* (heat), and *dolor* (pain). In the second century AD, another physician, Galen, added a fifth sign: *functio laesa* (loss of function). The cardinal signs of inflammation reflect the three major events of an inflammatory response:

1. Vasodilation—an increase in the diameter of blood vessels—of nearby capillaries occurs, while on the other hand the vessels that carry blood

away from the affected area constrict, resulting in an engorgement of the capillary network. The engorged capillaries are responsible for tissue redness (erythema) and an increase in tissue temperature. It also causes a reduction in the velocity of blood flow, especially along the surfaces of small blood vessels.

2. An increase in capillary permeability facilitates an influx of fluid and cells from the engorged capillaries into the tissue. The endothelial cells lining the blood vessel are activated to express adhesion molecules that promote the binding of circulating leukocytes. The combination of slowed blood flow and induced adhesion molecules allows leukocytes to attach to the endothelium and migrate into the tissues. The fluid that accumulates (exudate) has a much higher protein content than fluid normally released from the vasculature. Accumulation of exudate contributes to tissue swelling (edema), and pain as well as the accumulation of plasma proteins that aid in host defense. All these changes are initiated by the cytokines produced by activated macrophages. These changes are induced by a variety of inflammatory mediators released as a consequence of the recognition of pathogens. These include the lipid mediators of inflammation prostaglandins, leukotrienes, and platelet-activating factor (PAF) which are rapidly produced by macrophages through enzymatic pathways that degrade membrane phospholipids. Their actions are followed by those of the cytokines and chemokines (chemoattractant cytokines) that are synthesized and secreted by macrophages in response to pathogens. The cytokine tumor necrosis factor-α (TNF-α), for example, is a potent activator of endothelial cells.

3. Influx of phagocytes into the tissues is facilitated by the increased permeability of the capillaries. Once inflammation has begun, the first cells attracted to the site of infection are generally neutrophils. They are followed by monocytes, which differentiate into more tissue macrophages. In the later stages of inflammation, other leukocytes such as eosinophils and lymphocytes also enter the infected site. The emigration of phagocytes is a multistep process that includes adherence of the cells to the endothelial wall of the blood vessels (margination), followed by their emigration between the capillary-endothelial cells into the tissue (diapedesis or extravasation), and, finally, their migration through the tissue to the site of the invasion (chemotaxis). As phagocytic cells accumulate at the site and begin to phagocytose bacteria, they release lytic enzymes, which can damage nearby healthy cells. The accumulation of dead cells, digested material, and fluid forms a substance called pus.

The events in the inflammatory response are initiated by a complex series of events involving a variety of chemical mediators whose interactions are only partly understood. Some of these mediators are derived from invading microorganisms, some are released from damaged cells in response to tissue injury, some are generated by several plasma enzyme systems, and some are products of various white blood cells participating in the inflammatory response.

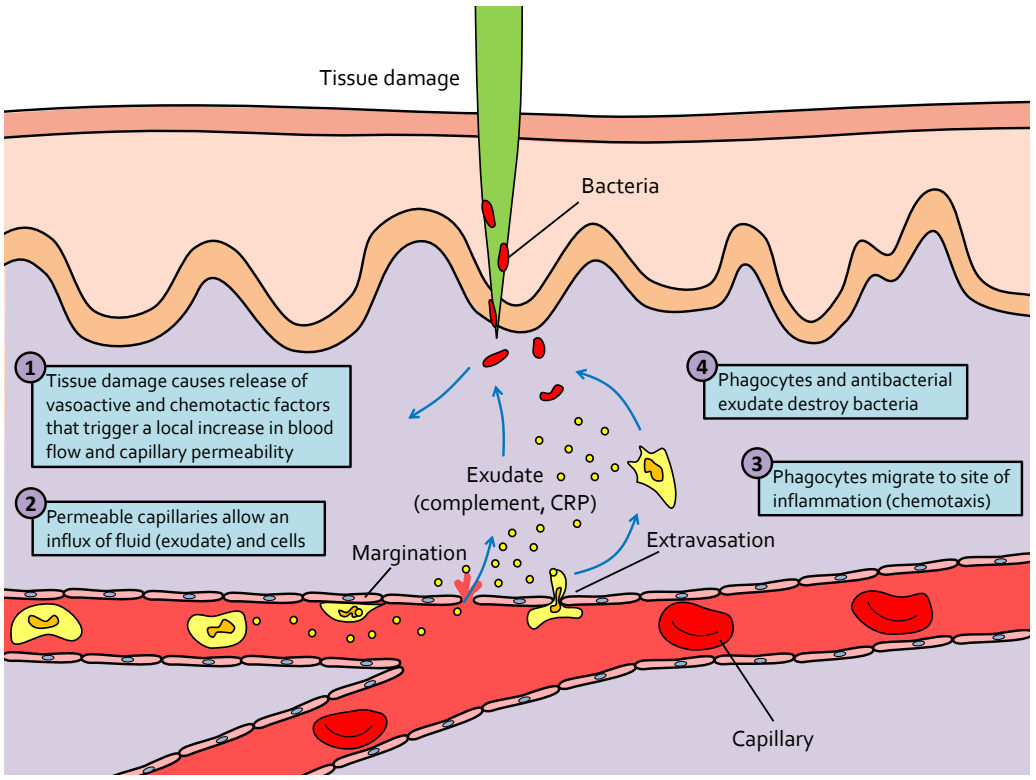


Figure 8. Major events in the inflammatory response. A bacterial infection causes tissue damage with release of various vasoactive and chemotactic factors. These factors induce increased blood flow to the area, increased capillary permeability, and an influx of white blood cells, including phagocytes and lymphocytes, from the blood into the tissues. The serum proteins contained in the exudate have antibacterial properties, and the phagocytes begin to engulf the bacteria, as illustrated in **Figure 7**. Self work.

Among the chemical mediators released in response to tissue damage are various serum proteins called acute-phase proteins. The concentrations of these proteins increase dramatically in tissue-damaging infections. C-reactive protein is a major acute-phase protein produced by the liver in response to tissue damage (**Figure 9**). It binds to the C-polysaccharide cell-wall component found on a variety of bacteria and fungi (hence the name). This

binding activates the complement system, resulting in increased clearance of the pathogen either by complement-mediated lysis or by a complement-mediated increase in phagocytosis.

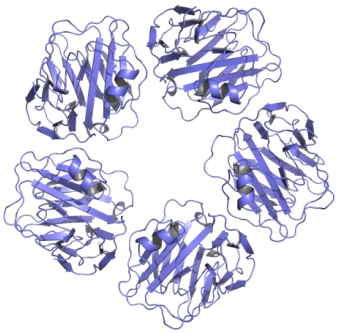


Figure 9. C-reactive protein.

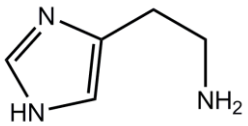


Figure 10. Histamine. It derives from the decarboxylation of the amino acid histidine, catalyzed by the enzyme L-histidine decarboxylase.

One of the principal mediators of the inflammatory response is histamine (**Figure 10**), a chemical released by a variety of cells in response to tissue injury. Histamine binds to receptors on nearby capillaries and venules, causing vasodilation and increased permeability. Another important group of inflammatory mediators, small peptides called kinins, are normally present in blood plasma in an inactive form. Tissue injury activates these peptides, which then cause vasodilation and increased permeability of capillaries. A particular kinin, called bradykinin, also stimulates pain receptors in the skin. This effect probably serves a protective role, because pain normally causes an individual to protect the injured area. Vasodilation and the increase in capillary permeability in an injured tissue also enable enzymes of the blood-clotting system to enter the tissue. These enzymes activate an enzyme cascade that results in the deposition of insoluble strands of fibrin, which is the main component of a blood clot. The fibrin strands wall off the injured area from the rest of the body and serve to prevent the spread of infection. Once the inflammatory response has subsided and most of the debris has been cleared

away by phagocytic cells, tissue repair and regeneration of new tissue begins. Capillaries grow into the fibrin of a blood clot. New connective tissue cells, called fibroblasts, replace the fibrin as the clot dissolves. As fibroblasts and capillaries accumulate, scar tissue forms. (1)

Pattern-Recognition Receptors

The receptors of adaptive and innate immunity differ. Antibodies (Ab) and T-cell receptors (TCR), the receptors of adaptive immunity, recognize details of molecular structure and can discriminate with extreme specificity between antigens featuring only slight structural differences. The receptors of innate immunity recognize broad structural motifs that are highly conserved within microbial species but are generally absent from the host. Because they recognize particular overall molecular patterns, such receptors are called pattern-recognition receptors (PRRs). Patterns recognized by this type of receptor include combinations of sugars, certain proteins, particular lipid-bearing molecules, and some nucleic acid motifs. These molecules are generally termed PAMPs (Pathogen-Associated Molecular Patterns), although others (11) (12) prefer the acronym MAMPs (Microbe-Associated Molecular Patterns, on the grounds that most microbes, not only pathogens, express the molecules detected). Typically, the ability of pattern-recognition receptors to distinguish between self and nonself is perfect because the molecular pattern targeted by the receptor is produced only by the pathogen and never by the host. This contrasts sharply with the occasional recognition of self antigens by receptors of adaptive immunity, which can lead to autoimmune disorders. Like Ab*s* and TCR*s*, pattern-recognition receptors are proteins. However, the genes that encode PRRs are present in the germline of the organism. In contrast, the genes that encode the enormous diversity of Ab*s* and TCR*s* are not present in the germline, but they are generated by an extraordinary process of genetic recombination. Many different pattern-

recognition receptors have been identified and several examples appear in **Table 1**.

Characteristic	Innate immunity	Adaptive immunity
Specificity	Specific for conserved molecular patterns or types	Specific for details of antigen structure
Self/nonself discrimination	Perfect. Never recognizes self.	Excellent: but imperfect. Occasional reaction with self antigens.
Receptors of the adaptive immune system		
Receptor (location)	**Target (source)**	**Effect of recognition**
Antibody (B-cell membrane, blood, tissue fluids)	Specific components of pathogen	Labeling of pathogen for destruction and removal
T-cell receptor (T-cell membrane)	Proteins or certain lipids of pathogen	Induction of pathogen-specific humoral and cellmediated immunity
Receptors of the innate immune system		
Complement (bloodstream, tissue fluids)	Microbial cell-wall components	Complement activation, opsonization
Mannose-binding lectin (MBL) (bloodstream, tissue fluids)	Mannose-containing microbial carbohydrates (cell walls)	Complement activation, opsonization
C-reactive protein (CRP) (bloodstream, tissue fluids)	Phosphatidylcholine (microbial membranes)	Complement activation, opsonization
LPS-binding protein (LBP) (bloodstream, tissue fluids)	Bacterial lipopolysaccharide (LPS)	Delivery to cell-membrane LPS receptor
TLRs (cell membrane and internal compartments)	Viral, bacterial or Fungi components (See **Table 2**)	Attracts phagocytes, activates macrophages, dendritic cells. Induces secretion of several cytokines; induces production of interferon, an antiviral cytokine
Scavenger receptors (many) (cell membrane)	Many targets; gram-positive and gram-negative bacteria, apoptotic host cells	Induces phagocytosis or endocytosis

Table 1. A comparison between the innate immune system receptors and those of the adaptive immune system (top). Some of the PRR*s* known (bottom). For TLRs ligands and localization details, see next Chapter and **Table 2.**

Some are present in the bloodstream and tissue fluids as soluble circulating proteins and others are on the membrane of cells such as macrophages, neutrophils, and dendritic cells. Mannose-binding lectin (MBL) and C-reactive protein (CRP) are soluble pattern receptors that bind to microbial surfaces and promote their opsonization. Both of these receptors also have the ability to activate the complement system when they are bound to the surface of microbes, thereby making the invader a likely target of complement-mediated lysis. Yet another soluble receptor of the innate immune system, lipopolysaccharide-binding protein (LBP), is an important part of the system that recognizes and signals a response to LPS. Pattern-recognition receptors found on the cell membrane include scavenger receptors and the toll-like receptors. Scavenger receptors are present on macrophages and many types of dendritic cells, and are involved in the binding and internalization of Gram-positive and Gram-negative bacteria, as well as the phagocytosis of apoptotic host cells. The exact roles and mechanisms of action of the many types of scavenger receptors known to date are under active investigation. The toll-like receptors (TLR*s*) are important in recognizing many microbial patterns. Typically, signals transduced through the TLRs cause transcriptional activation and the synthesis and secretion of cytokines, which promote inflammatory responses that bring macrophages and neutrophils to sites of inflammation. TLR signaling can also result in the recruitment and activation of macrophages, NK cells, and dendritic cells, key agents in the presentation of antigen to T cells. The links to T cells and cytokine release shows the intimate relationship between innate and adaptive responses. Ten human TLR functions have been determined to date. **Table 2** summarizes the receptors of adaptive immunity and lists many pattern-recognition receptors of innate

immunity. The microbial targets and physiological sites of some TLR*s* are shown in **Figure 12**. Next chapter will discuss the features and functions of TLR*s* more in detail.

Toll-Like Receptors

Toll-like receptors (TLR*s*) are a class of proteins that play a key role in the innate immune system. They are non-catalytic receptors that recognize structurally conserved molecules derived from microbes. Once these microbes have breached physical barriers such as the skin or intestinal tract mucosa, they are recognized by TLR*s* which activate immune cell responses.

TLRs are so named because of their similarity to the protein coded by the *Toll* gene identified in *Drosophila melanogaster*.[1]

TLR*s* Diversity

TLRs are a type of pattern recognition receptor (PRR) and recognize molecules that are broadly shared by pathogens but distinguishable from host molecules, collectively referred to as PAMPs. TLRs together with the Interleukin-1 (IL-1) receptors form a receptor superfamily, known as the "Interleukin-1 Receptor/Toll-Like Receptor Superfamily"; all members of this family have in common a so-called TIR (Toll-IL-1 receptor) domain. (**Figure 11)**

Three subgroups of TIR domains exist. Proteins with subgroup 1 TIR domains are receptors for interleukins that are produced by macrophages, monocytes and dendritic cells and all have extracellular Immunoglobulin (Ig) domains. Proteins with subgroup 2 TIR domains are classical TLRs, and bind directly or indirectly to molecules of microbial origin. A third subgroup of proteins

[1] The origin of Toll name is quite bizarre. The gene in question, when mutated, makes the Drosophila flies look unusual, or 'weird'. The researchers were so surprised that they spontaneously shouted out in German "*Das ist ja toll!*" which translates as "*That´s amazing!*".

containing TIR domains consists of adaptor proteins that are exclusively cytosolic and mediate signaling from proteins of subgroups 1 and 2.

Toll-like receptors are membrane-spanning proteins that share a common structural element in their extracellular region, repeating segments of 24 to 29 amino acids containing the sequence xLxxLxLxx (x is any amino acid and L is leucine). These structural motifs are called leucine-rich repeats (LRR*s*; **Figure 11**). All TLR*s* contain several LRRs, and a subset of the LRRs make up the extracellular ligand-binding region of the TLR. As shown in **Figure 11**, TIR domains have three regions, highly conserved among all members of the TIR family, called boxes 1, 2, and 3, that serve as binding sites for intracellular proteins participating in the signaling pathways mediated by TLRs.

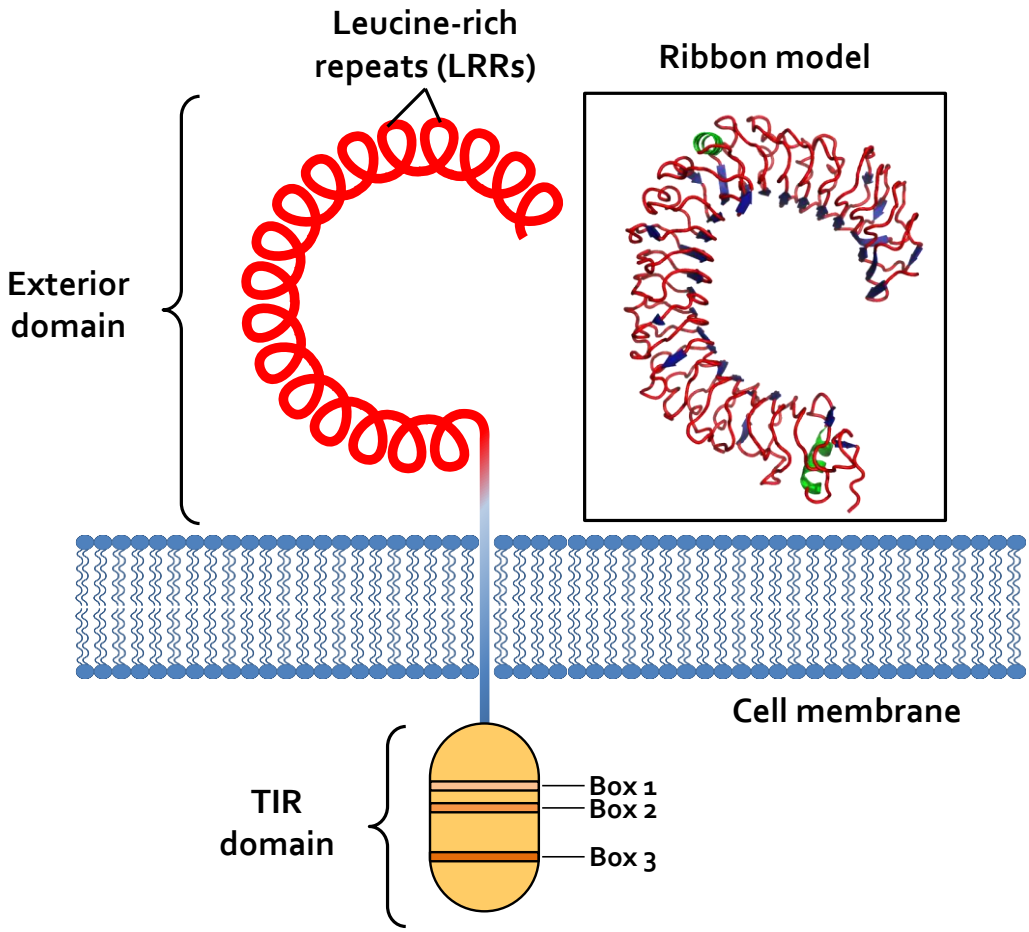


Figure 11. Structure of a Toll-like receptor (TLR). Toll-like receptors have an exterior region that contains many leucine-rich repeats (LRRs), a membrane-spanning domain, and an interior domain called the TIR domain. The ligand-binding site of the TLR is found among the LRRs. The TIR domain interacts with the TIR domains of other members of the TLR signal-transduction pathway; three highly conserved sequences of amino acids called box 1, 2, and 3 are essential for this interaction and are characteristic features of TIR domains.

TLRs are present in vertebrates, as well as in invertebrates. Some genes which have a similar sequence to TLRs have been found also in plants and seem to

be required for host defence against infection. TLRs thus appear to be one of the most widespread and ubiquitous components of the immune system.

It has been estimated that most mammalian species have between ten and fifteen types of Toll-like receptors. Thirteen TLRs (named simply TLR1 to TLR13) have been identified in humans and mice together, and equivalent forms of many of these have been found in other mammalian species (21; 22; 23). Other mammals may express TLRs which are not found in humans and vice versa. Other non-mammalian species may have TLRs distinct from mammals. This may complicate the process of using experimental animals as models of human innate immunity.

Ligands

These receptors recognize molecules that are constantly associated with threats (i.e. pathogen or cell stress) and are highly specific to these threats (i.e. cannot be mistaken for self molecules). Pathogen-associated molecules that induce a TLR-mediated response are usually critical to the pathogen's function and cannot be eliminated or changed through mutation. These features in pathogens include bacterial cell-surface LPS, lipoproteins, lipopeptides and lipoarabinomannan; proteins such as flagellin from bacterial flagella; double-stranded RNA of viruses or the unmethylated CpG islands of bacterial and viral DNA; and certain other RNA and DNA (**Figure 12**).

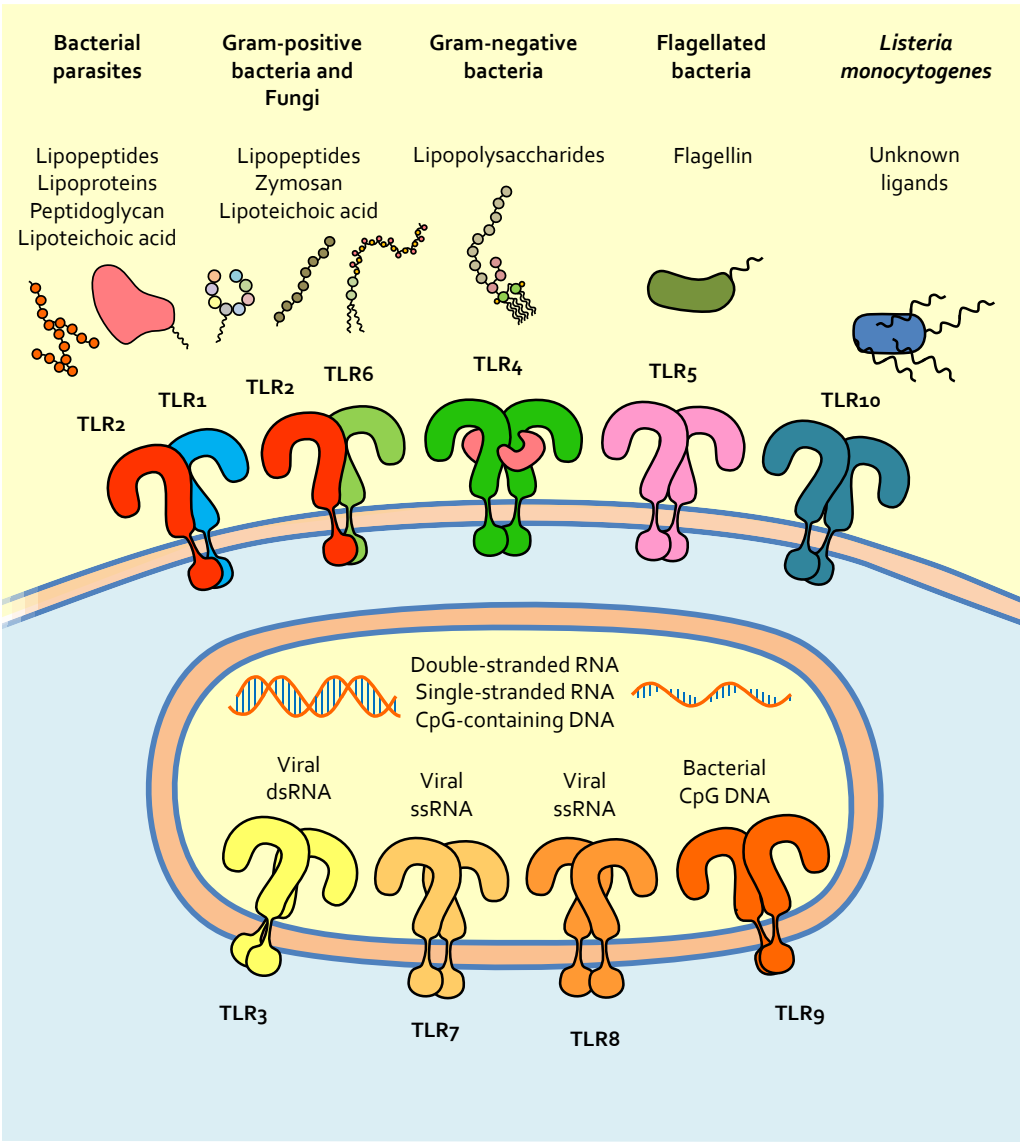


Figure 12. Location and targets of some TLRs. Many pattern-recognition receptors are extracellular and target microbes or microbial components in the bloodstream and tissue fluids, causing their lysis or marking them for removal by phagocytes. Other PRRs, like TLRs, are present either on the cell membrane or in intracellular compartments and bind to a broad variety of microbes or microbial products. Engagement of these receptors triggers signaling pathways that promote inflammation.

Endogenous ligands

It has been proposed that endogenous activators of TLRs might participate in autoimmune diseases. TLRs have been suspected of binding to host molecules including fibrinogen (involved in blood clotting), heat shock proteins (HSPs) and host DNA.

Signaling

Though most TLRs appear to function as homodimers, TLR2 forms heterodimers with TLR1 or TLR6, each dimer having a different ligand specificity. TLR10 seems to form both homodimers and heterodimers with either TLR1 or TLR2. TLRs may also depend on other co-receptors for full ligand sensitivity, such as in the case of TLR4's recognition of LPS, which requires MD-2. CD14 and LPS Binding Protein (LBP) are known to facilitate the presentation of LPS to MD-2.

When activated, TLRs recruit adapter molecules within the cytoplasm of cells in order to propagate a signal. Four adapter molecules are known to be involved in signaling. These proteins are known as MyD88, TIRAP (also called Mal), TRIF and TRAM (19; 20; 21). The adapters activate other molecules within the cell, including certain protein kinases (IRAK1, IRAK4, TBK1, and IKKi) that amplify the signal, and ultimately lead to the induction or suppression of genes that orchestrate the inflammatory response. In all, thousands of genes are activated by TLR signaling, and collectively, the TLRs constitute one of the most pleiotropic yet tightly regulated gateways for gene modulation.

Summary of known mammalian TLRs

TLRs are activated by different ligands they bind, which, in turn are located on different types of organisms or structures. They also have different adapters to respond to activation and are located sometimes at the cell surface and sometimes to internal cell compartments. Furthermore, they are expressed by different types of leucocytes or other cell types. **Table 2** presents a resume of this differences between various TLRs. (22) Next chapter contains more detailed information on TLR4 and the associated co-receptors.

Receptor	Ligand(s)	Ligand location	Adapter(s)	Location	Cell types
TLR1	multiple triacyl lipopeptides	Bacteria	MyD88, MAL	cell surface	Monocytes, macrophages a subset of dendritic cells, B lymphocytes
TLR2	multiple glycolipids	Bacteria	MyD88, MAL	cell surface	Monocytes, macrophages, myeloid dendritic cells [24], mast cells
	multiple lipopeptides	Bacteria			
	multiple lipoproteins	Bacteria			
	lipoteichoic acid	Bacteria			
	HSP70	Host cells			
	zymosan (β-glucan)	Fungi			
TLR3	dsRNA, poly I:C	Viruses	TRIF	cell compartment	Dendritic cells B lymphocytes
TLR4	LPS	Gram-negative bacteria	MyD88, MAL, TRIF, TRAM	cell surface	Monocytes, macrophages, myeloid dendritic cells [24], mast cells, intestinal epithelium
	several heat shock proteins	Bacteria and host cells			
	fibrinogen	host cells			
	heparan sulfate fragments	host cells			
	hyaluronic acid fragments	host cells			
TLR5	flagellin	Bacteria	MyD88	cell surface	Monocytes, macrophages, a subset of dendritic cells, intestinal epithelium
TLR6	Multiple diacyl lipopeptides	Mycoplasma	MyD88/MAL	cell surface	Monocytes, macrophages, mast cells, B lymphocytes
TLR7	imidazoquinoline	small synthetic compounds	MyD88	cell compartment	Monocytes, macrophages, plasmacytoid dendritic cells, B lymphocytes
	loxoribine				
	bropirimine				
	single-stranded RNA				
TLR8	small synthetic compounds, ssRNA		MyD88	cell compartment	Monocytes, macrophages, a subset of dendritic cells, mast cells
TLR9	unmethylated CpG Oligodeoxynucleotide	Bacteria	MyD88	cell compartment	Monocytes, macrophages, plasmacytoid dendritic cells, B lymphocytes
TLR10	unknown	*Listeria monocytogenes*	MyD88	cell surface	Monocytes, macrophages B lymphocytes

Table 2. Known human TLRs, their localization and their ligands. (23)

Toll-Like Receptor 4 and its ligands

TLR4 is a cell surface, single-spanning transmembrane protein containing 799 aminoacid residues with an extracellular leucine-rich repeat (LRR) domain and an intracellular TIR domain, like other TLRs (**Figure 13**).

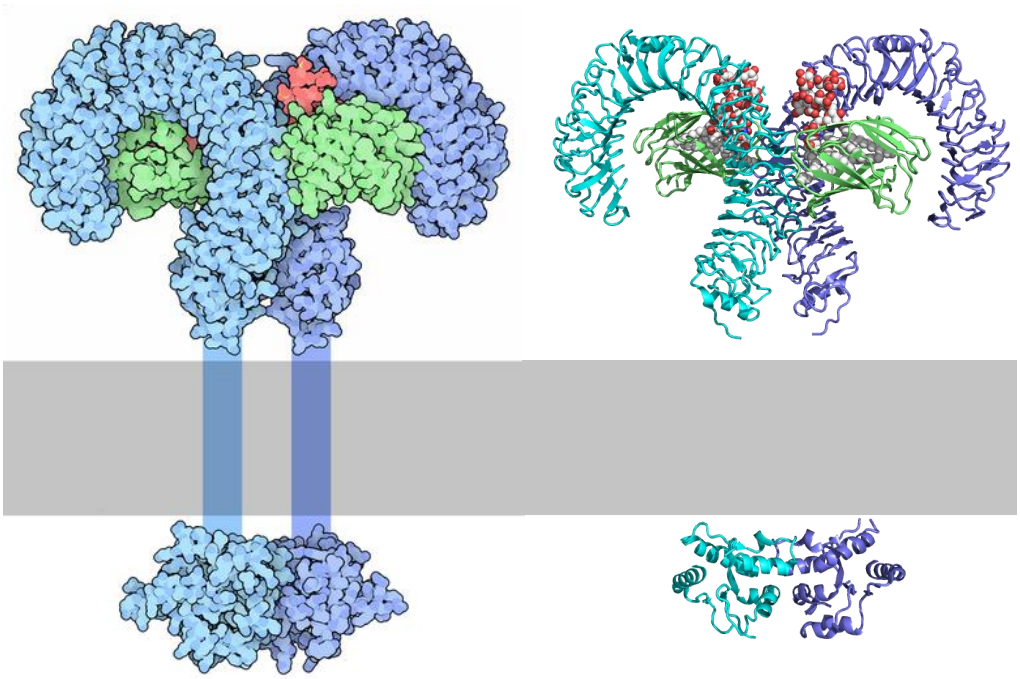


Figure 13. Crystal structure of human $(TLR4{\cdot}MD\text{-}2{\cdot}LPS)_2$ complex. The receptor complex has been co-crystallized with *E. coli* LPS Ra. On the left, a sphere representation of the extracellular and the cytoplasmic domains (David Goodsell & RCSB Protein Data Bank), on the left a cartoon representation of the receptors, with R-LPS shown as spheres (self work from PDB ID: 3FXI).

Among TLRs, TLR4 is the receptor that selectively recognize bacterial lipopolysaccharide (LPS), thus providing an effective response against Gram-negative bacteria that eluded the physical and anatomic barriers that

represent the organism's first line of defense against infection (27; 28; 29). It activates the host defense effector system and rapidly triggers pro-inflammatory processes. This aim is achieved by the coordinate and sequential action of four endotoxin-binding proteins: the LPS binding protein (LBP), the cluster differentiation antigen 14 (CD14), the myeloid differentiation protein (MD-2) and TLR4 itself. This sequential process starts with the binding of LBP to LPS aggregates, in form of micelles or membrane blebs (spontaneously released by Gram-negative bacteria), and ends up with the formation of the activated $(TLR4{\cdot}MD\text{-}2{\cdot}LPS)_2$ complex that has a pivotal role in initiating the inflammatory cascade. LBP interacts with endotoxin-rich bacterial membranes or endotoxin aggregates, catalyzing extraction and transfer of LPS monomers to CD14 that in turn transfers LPS monomers to MD-2 and to $(TLR4{\cdot}MD\text{-}2{\cdot}LPS)_2$ heterotetramer. (**Figure 14**)

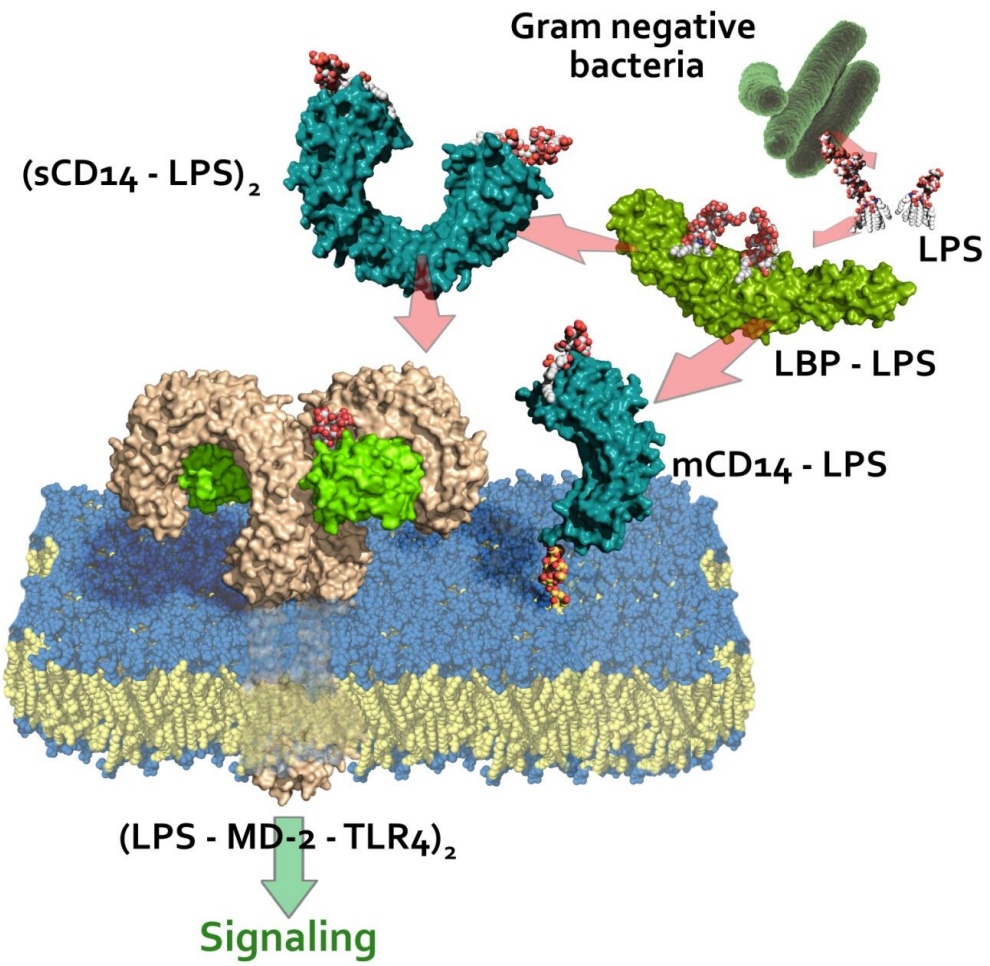


Figure 14. TLR4–associated proteins that are involved in LPS sensing. Accordingly to the currently accepted mechanism, the first protein involved in the process is the LBP, than LPS is bound by soluble or GPI-anchored CD14 (indicated as sCD14 and mCD14 respectively, in the figure) and finally it enters the hydrophobic pocket present in protein MD-2, that associate with TLR4 inducing its dimerization. Self work from PDB ID: 3FXI (for TLR4·MD-2 complex), 1O77 (for the TIR domain of TLR4), 2OBD and 1BP1 (for LBP), 1WWL (for CD14), 1QFG (for LPS).

Receptor dimerization leads to the recruitment of adapter proteins to the intracellular TIR domain of TLR4, initiating the intracellular signal cascade

that culminates in translocation of transcription factors to the nucleus and the biosynthesis of cytokines.

Variations in the aggregation state and the 3D form of endotoxin aggregates may directly influence the kinetics and potency of TLR4 activation and signaling.

Before discussing TLR4-associated receptors and some of the molecules that stimulate (agonists) or inhibit (antagonists) TLR4 signaling, it will be briefly resumed the crucial role of TLR4 in immune and inflammatory responses, since it is involved in three functions:

- **Sensory function.** LPS recognition by TLR4 mediates rapid production of cytokines and informs innate immunity cells that Gram-negative bacteria are present inside the organism.
- **Effector function.** the response to infection through TLR4 initiates rapid recruitment of inflammatory cells to the site of infection and activate them to induce an arsenal of antibacterial functions. (30)
- **Regulation of the adaptive responses.** the TLR4-mediated innate response to a pathogen can be decisive in determining both the nature and the magnitude of the adaptive response. (31)

TLR4-associated receptors: LPS-binding protein (LBP)

Human LBP is a 58-60 kDa serum glycoprotein, with 44% sequence identity with human bactericidal/permeability-increasing protein (BPI). The crystal structure of murine BPI has been determined, which allows the construction of a reliable tertiary structural model of LBP. The tertiary structure of BPI

contains two barrel domains arranged in a banana-like shape which are connected by a proline-rich linker. Each domain is composed of an antiparallel β-stranded layer twisted around the long α-helix. In the pocket between the helix and the beta-strand, a phospholipid molecule has been cocrystallized in each domain (40). The tip of the N-terminal domain of LBP contains a cluster of cationic residues, which are essential for the LPS binding and signaling. The phospholipid-binding site of LBP seems a suitable site for binding of the lipid A moiety. However, the size of the binding pocket is by far too small to accommodate the acyl chains of the typical lipid A. Therefore, the location of the lipid A-binding site on LBP is not quite clear, since the distance between the cationic and hydrophobic site is also too large to allow the interaction of the lipid A molecule with both sites.

LPS response is enhanced by the serum LBP that disaggregates LPS micelles with the cooperation of soluble CD14 (sCD14); serum albumin also plays a role in facilitating the extraction of LPS monomers from the aggregates (27). LBP seems to have a concentration-dependent dual role: low concentrations of LBP enhance the LPS-induced activation (42; 43), while the acute-phase rise in LBP concentration inhibit LPS-induced cellular stimulation (30; 31; 32).

TLR4 associated receptors: MD-2

The Myeloid Differentiation factor 2 or MD-2 (also know as Lymphocyte antigen 96, LY96) is a 160 aa protein (18.4 kDa) with a β-cup fold structure composed of two antiparallel β-sheets forming a large hydrophobic pocket for ligand binding (33; 34). MD-2 binds to the ectodomain of TLR4 and is essential for LPS signaling (35). The monomeric form of MD-2 binds to a LPS monomer (36; 37; 38; 39).

The TLR4·MD-2 heterodimer has complex ligand specificity and can be activated by structurally diverse molecules that will be presented in this section. Minor changes in synthetic derivatives can dramatically change their activity and induce a switch between agonist and antagonist activity. The majority of synthetic TLR4 agonists and antagonists are MD-2 ligands, so that MD-2 is considered the principal target for the pharmaceutical intervention on innate response to LPS (39). The very recent determination of the crystal structure of $(TLR4{\cdot}MD\text{-}2{\cdot}LPS)_2$ complex (55), together with crystallographic data of MD-2 bound to TLR4 antagonists lipid IVa (49) and Eritoran (E5564) (10), has revealed some fundamental structural aspects of the TLR4 dimerization process and the molecular basis of TLR4 agonism and antagonism already discussed in the previous paragraph.

In the crystal structures of MD-2 bound to the two antagonists lipid IVa (PDB ID: 2E59) and Eritoran (PDB ID: 2Z65), the four lipid chains of the antagonists completely fill the available space in the MD-2 binding cavity (**Figure 15**, A and B). *E. coli* LPS has two more lipid chains than these antagonists, so it has been proposed that global structural changes may take place in the MD-2 structure to accommodate the extra lipid chains. Unexpectedly, the crystal structure of the $(TLR4{\cdot}MD\text{-}2{\cdot}LPS)_2$ complex (PDB ID: 3FXI) has shown that the size of the MD-2 pocket is unchanged and that additional space for lipid binding is generated by displacing the phosphorylated glucosamine disaccharide moiety by about 6 Å towards the exterior of the lipophilic pocket (**Figure 15** C) (55).

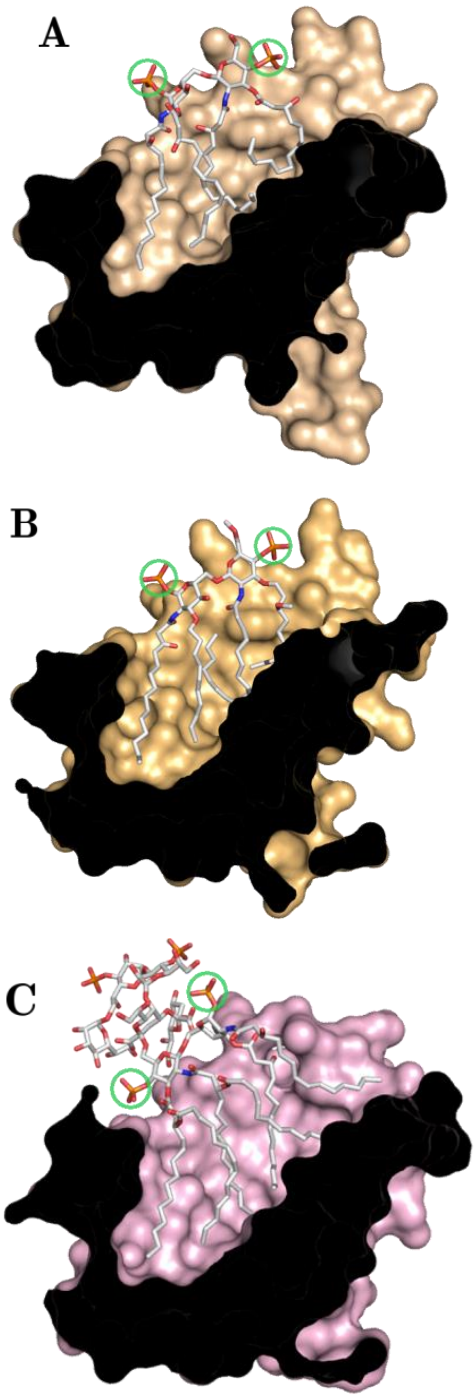


Figure 15. Crystal structures of MD-2 bound to: A) Lipid IVa (PDB ID: 2E59); B) Eritoran (PDB ID: 2Z65); C) LPS Ra (PDB ID: 3FXI). Self work. For comments, see text.

This shift of the diglucosamine backbone repositions the phosphate groups (in particular, that in C-1 position protrudes from MD-2) such that they can interact with positively charged residues of the two TLR4 molecules, thus promoting dimerization and formation of the activated $(TLR4{\cdot}MD\text{-}2{\cdot}LPS)_2$ complex.

Interestingly, the glucosamine backbones of the antagonists are not only translated but also rotated by 180° respect to the LPS agonist, thus interchanging the two phosphate groups (55). More research is needed to establish whether all LPS derivatives with four lipid chains cause a similar rotation of the phosphorylated-glucosamine backbone.

The overall folding of TLR4 and MD-2 is not disturbed by LPS binding and dimerization.

In the $(TLR4{\cdot}MD\text{-}2{\cdot}LPS)_2$ structure the lipid chains of lipid A are completely buried inside the MD-2 cavity, but the chain on C-2 is partially exposed on the MD-2 surface generating an hydrophobic interface for the interaction with the second TLR4 of the complex. The esters and amide groups connecting the lipids to the glucosamine backbone or to other lipid chains are exposed on the surface of MD-2. They interact with hydrophilic side chains on the MD-2 surface and on the surface of the two TLR4 molecules. The two phosphate groups of the lipid A bind to the TLR4-MD-2 complex by interacting with positively charged residues in the two TLR4 subunits and in MD-2, and by hydrogen bonding to MD-2 Ser118 (55).

TLR4-associated receptors: CD14

CD14 is expressed on the surface of monocytes and granulocytes as a glycosylphosphatidylinositol (GPI)-anchored glycoprotein or in soluble form in the serum. The monomeric subunit of CD14 has an curved structure with a concave surface formed by a large β-sheet with the repetition of LRR motifs. The concave surface contains both helices and loops in no regular pattern, so it is rugged, with several grooves and pockets that seems to be crucial for ligand binding. The main characteristic of its structure is the N-terminal hydrophobic pocket (**Figure 16**).

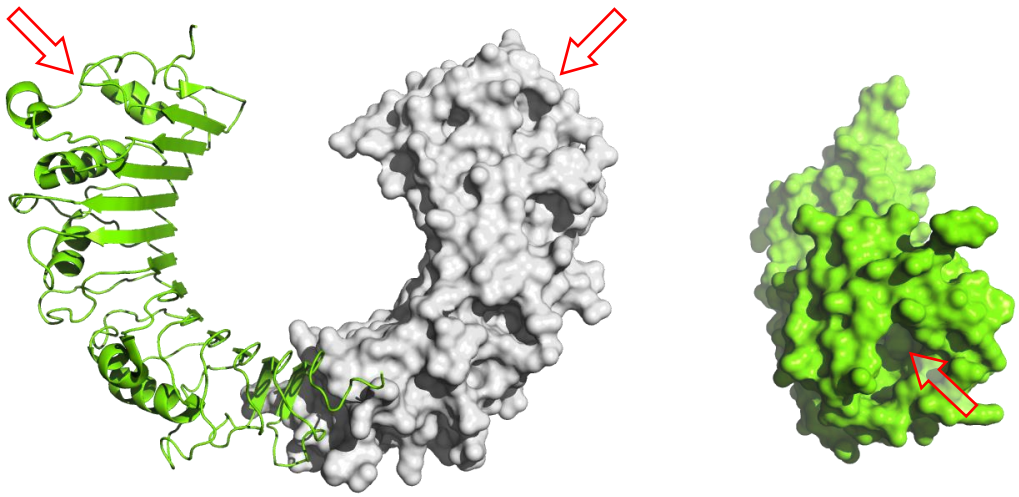


Figure 16. Crystallographic structure of CD14, presented as a dimer (side view, one chain shown as green cartoons, the other as white surface) and as a monomer (top view, green surface). The position of the N-terminal hydrophobic pocket is indicated by red arrows. Self work from PDB ID: 1WWL.

CD14 was the first endotoxin receptor identified, until TLR4 was characterized as a LPS receptor in 1998 by Beutler and co-workers. Thus, a lot of scientific work before 1998 was focused on CD14/LPS interaction, and CD14 biological

role was hence overestimated. A discrete number of studies, however, pointed out that something was missing and led to the hypothesis that at least one other LPS receptor was present on innate immunity cells, unknown at that time. After 1998 it was soon clear to the scientific world that TLR4 is the principal endotoxin receptor and that it controls the major signal pathway following LPS exposure. The scientific interest thus dramatically moved towards TLR4 functions and signaling, giving it such an importance that CD14 and its role as a LPS receptor have been somehow neglected from that time on.

After TLR4 discovery, the main function of CD14 has long been considered that of concentrating LPS to increase the sensitivity of TLR4 complex to endotoxin. CD14 in fact results essential for the activation of TLR4·MD-2 complex when LPS concentration is low and when S-LPS is present, while R-LPS and highly concentrated LPS can activate TLR4 also in the absence of CD14.

Moreover, it has also been recently shown that CD14 controls a signal pathway on its own, dependent on PLCγ2, Syk, ITAM-containing proteins and Ca^{2+} signalling, resulting in NFAT transcription factor activation and production of IL-2, involved in cell life-cycle and edema formation. (42) This signal pathway involves CD14 internalization from the cell membrane to internal compartments. CD14 also has a fundamental role in TLR4·MD-2 internalization after LPS binding. This was demonstrated using CD14-defective cells treated with LPS, in which no TLR4 internalization was observed, while in wt cells CD14 and TLR4 colocalize inside the cells after exposure to endotoxin. This process is CD14-dependent and involves the same proteins as the NFAT-activating pathway mentioned before. (43)

The fact that R-LPS (and not S-LPS) can activate TLR4·MD-2 complex in absence of CD14 could reflect a different intracellular activation level and signaling, because R-LPS binding to TLR4·MD-2 activates a MyD88-dependent pathway while S-LPS, by binding to CD14, activates a MyD88-independent pathway (44).

R-form LPS induces higher levels of TNF-α than S-form LPS *in vivo*

Huber *et al.* analyzed the in vivo TNF-α responses of wt and CD14-deficient mice to Re- and S-form LPS after intravenous injection (9). R-form LPS induced a strong dose-dependent TNF-α response in wt and a lower one in $CD14^{-/-}$ mice. In contrast, S-form LPS induced no detectable TNF-α response in $CD14^{-/-}$ mice and its TNF-α-inducing activity in wt mice was lower compared to R-form LPS. This demonstrates that *in vivo* the activation capacity of R-form is superior to that of S-form LPS and this supports the results previously obtained by the same research group in the *in vitro* studies.

The message of the study of Huber *et al*. is that the S-form LPS, which has long been considered as the classical form of endotoxically active LPS, activates a narrower spectrum of TLR4/MD-2-expressing cells than R-form LPS and with a lower potency, both *in vitro* and *in vivo*. They showed that S-form LPS is practically devoid of stimulatory activity for mast cells (MC), while R-form LPS is a potent activator. This difference is based on a differential requirement for CD14 in the activation of cells by the two types of LPS. While the S-form requires CD14, R-form LPS can activate cells, regardless of the presence or absence of this LPS-binding protein. This explains why R-form LPS in contrast to S-form strongly stimulates MC, which lack CD14. Moreover, R-form LPS induces higher TNF-α responses than S-form *in vivo*,

demonstrating that the different activation capacities of the two LPS chemotypes are also present *in vivo*. Soluble CD14 (sCD14) provides help during activation of CD14-negative cells by LPS. Interestingly, in the case of murine cells high concentrations of sCD14 are required that are above the normal plasma levels. Such concentrations were found in *P. acnes*-treated or *S. typhimurium*-infected mice. This suggests that the up-regulation of sCD14 in the course of an inflammatory response ensures the contribution of MC and other CD14-negative LPS target cells to the antibacterial defense. This mechanism, however, might provoke also an enhanced risk of endotoxic shock, or potentiate acute allergic reactions. These results also provide an explanation for why R-form LPS and free lipid A induce strongly oxidative burst in human granulocytes, while S-form LPS is totally inactive in this respect (9). In human granulocytes, CD14 occurs intracellularly and is only sporadically expressed on the cell surface. Thus, in normal granulocytes, CD14, which is essential for the activity of S-form LPS, is not readily available on the cell surface. (45)

A varying part of the LPS isolated from all S-form wt bacteria is of the R-form type. The R-form fraction isolated from S-form LPS is a strong, CD14-independent activator of MC and other cell types. Therefore, the low activity of S-form LPS preparations for MC and other cell types devoid of CD14 (9), could be due to the portion of R-form LPS they contain. The presence of CD14 enhances the response of cells to both LPS chemotypes. A crystallographic study by Kim *et al*. (10) proposed that a large hydrophobic pocket found on the N-terminal side of the CD14 structure is responsible for the binding of the lipid portion of LPS. The CD14-independent activation by R-form LPS suggests that it is either capable of binding directly to the extracellular portion

of TLR4·MD-2 or that it integrates into the cell membrane and subsequently binds to the receptor complex. The difference in physicochemical properties between R- and S-form LPS may form the basis for the differences in LBP and CD14 requirement. It is conceivable that the lack of the long polysaccharide chains, which increases the hydrophilicity of R-form LPS, allows a better incorporation and mobility of the LPS in the mammalian cell membrane, thus providing a better access to the signaling receptor. This would also explain why the highly hydrophobic lipid A, which is completely devoid of core sugar constituents, is at least as powerful an activator as R-form LPS. Alternatively, the observed difference between R- and S-form LPS in the requirement for LPS-binding proteins might be related to possible differences in the structure of their lipid A moiety. The information on the structure and biological activity of lipid A so far has been obtained from studies on the lipid A of R-form LPS. Lipid A from pure genuine S-form LPS is not available yet. MC are found in almost all connective tissues and thus frequently encounter gram-negative bacteria during infection. Since there is ample evidence that MC, TLR and infectious agents play a crucial role in allergic, inflammatory and chronic disorders (9), it is likely that MC-derived cytokines such as TNF-α and IL-6 induced by R-form LPS contribute to the development of these pathologies. In contrast, the cytokine responses of MC to R-form LPS under physiological conditions are very likely beneficial.

In summary, it was demonstrated that R- and S-LPS chemotypes differ in their capacity to activate immune cells. Under physiological conditions, the S-form is limited in its activating capacity to cells that express mCD14, while the R-form has the potential to stimulate all cells that express TLR4·MD-2.

Since precise information about the molecular weight of LPS are not available, especially because S-LPS appears to be an heterogeneous mixture of molecules with different O-antigen lengths, it's not possible to speak about LPS concentration in terms of molarity (unless it is referred to very definite R-LPS molecules of known molecular structure). Thus, LPS concentrations are typically indicated as w/V, for example mg/ml. Since S-LPS molecules are far bigger than R-LPS or Lipid A, a comparison between the activity of S-LPS and R-LPS, as the one indicated in the aforementioned paper, could be misleading. In fact, since the molecular mechanism of TLR4 have been elucidated and pointed out that the active portion of the molecule is the Lipid A moiety, we can consider that in 1 mg of S-LPS and 1 mg of R-LPS the concentration of Lipid A structures is significantly different. Most of the weight of S-LPS is in fact represented by the very long sugar chains that are not required for innate immunity activation. A cautious approach to comparisons between the activities of R-LPS and S-LPS is therefore recommended.

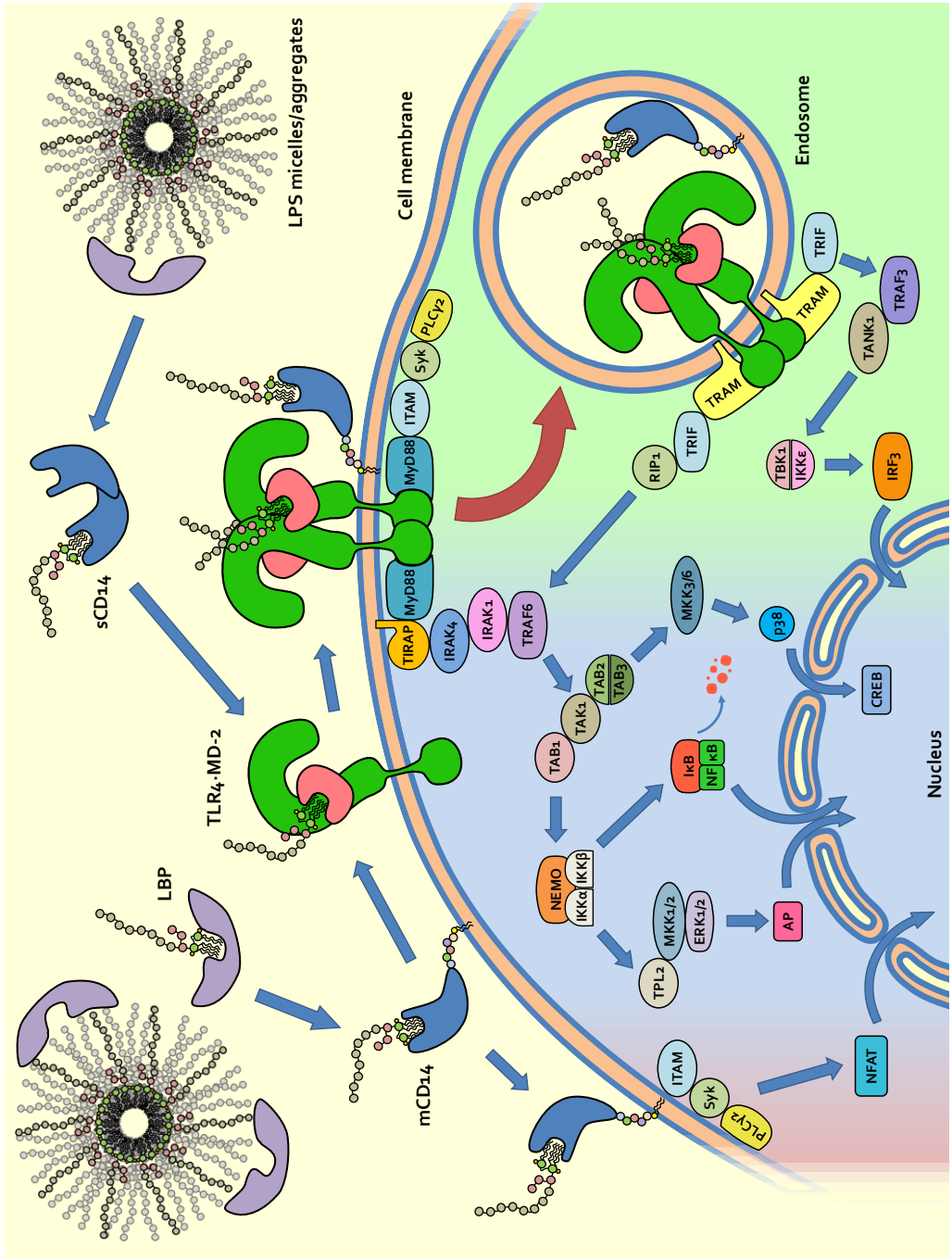


Figure 17. The complex signaling pathway of TLR4. Shades indicate the three separate pathways. Red shade: CD14 signal pathway; Blue shade: MyD88-depemdent pathway; Green shade: MyD88-independent pathway. Self work.

Dysregulated immune responses: Septic Shock

Septic shock is an example of how inappropriate functioning of host defenses can lead to disastrous consequences. Septic shock is a form of shock is generally caused by bacterial infection. Other causes of shock include massive crush injuries and burns.

In people with septic shock, vascular resistance and blood pressure drop despite normal-to-high cardiac output. The heart rate increases as the body tries to compensate for decreasing blood pressure. Another sign that is frequently associated with septic shock is disseminated intravascular coagulation (DIC), which can be seen as blackish or reddish skin lesions (petechiae). Septic shock is a serious condition because severely reduced levels of blood flow deprive essential organs of oxygen and nutrients. The consequent failure of organs, such as the kidneys, heart, brain, and lungs, is the cause of death in septic-shock patients.

Septicemia (bacteria infecting and growing in the bloodstream) occurs in more than 500,000 patients per year in the United States alone, millions worldwide each year, making sepsis one of the most common and life-threatening diseases all over the world (**Figure 18**).

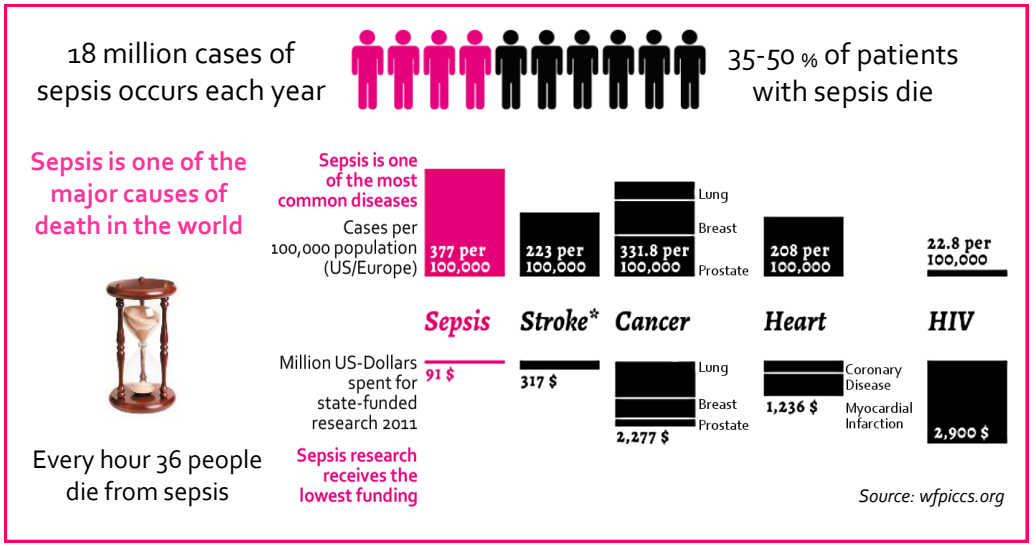


Figure 18. Sepsis is one of the most common diseases and one of the most life-threatening all over the world. Despite this, very little money is invested in research on sepsis. Adapted from (42).

A quarter of the patients with septicemia die in the hospital. Even those who survive may have long-term aftereffects, such as stroke or permanent damage to the lungs or other organs. Moreover, those who survive an episode of septic shock have a significantly greater risk of dying during the next 5-year period than people with the same underlying conditions but no previous episode of shock. Septic shock has serious implications and represent a huge expense for hospital administrations and insurances.

The term "sepsis" has now been defined more precisely in an effort to aid diagnosis of the various stages of disease. The first stage of shock, called systemic inflammatory response syndrome (SIRS), is characterized by a temperature over 38 °C or under 36 °C, a higher-than-normal heart rate, a higher-than-normal respiratory rate, and an unnaturally high or unnaturally

low neutrophil count. The second stage, termed "sepsis", is SIRS with a culture-documented infection (i.e. laboratory results showing the presence of bacteria in the bloodstream). The third stage is "severe sepsis", characterized by organ dysfunction and very low blood pressure. The fourth stage is "septic shock", which is characterized by low blood pressure despite fluid administration.

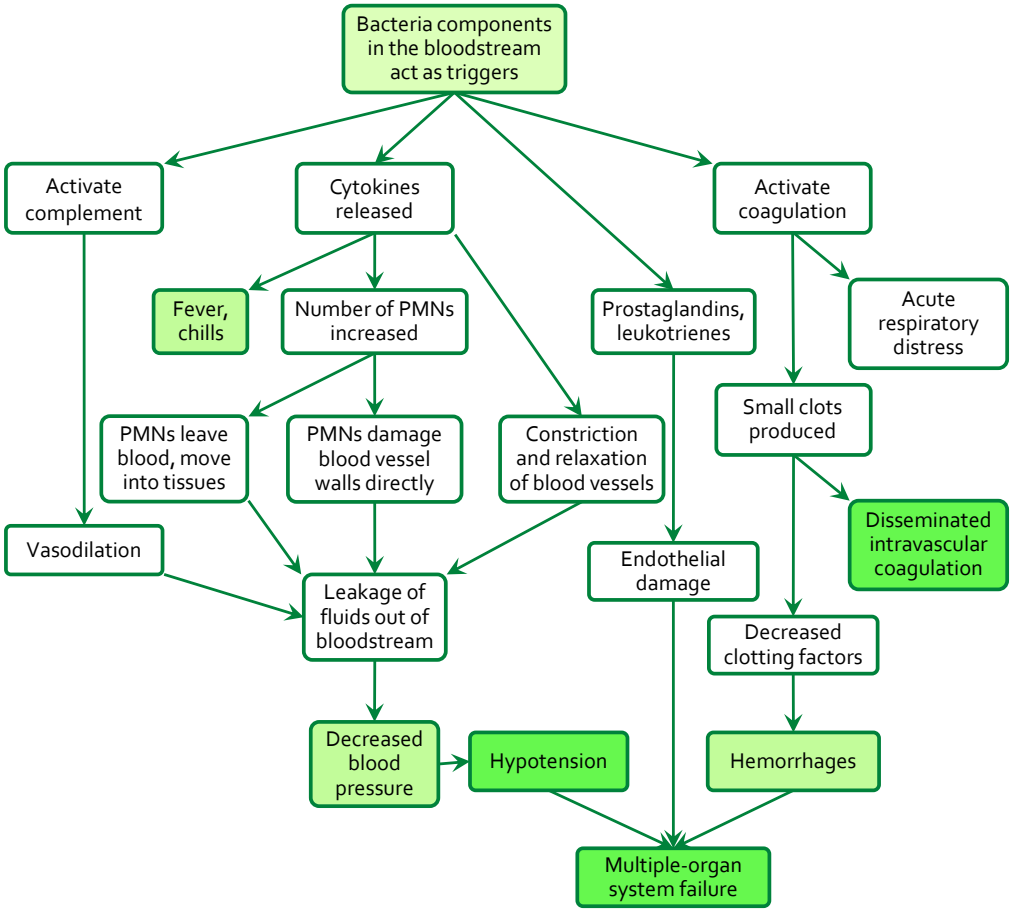


Figure 19. Sources of hypotension, DIC, and internal hemorrhages seen in cases of septic shock. Adapted from (41)

How does a bacterial infection of the bloodstream produce such a serious condition? The steps involved in septic shock are illustrated in **Figure 19**.

The immune system has the purpose not only to destroy the invading microorganisms but also to confine the inflammatory response to certain areas of the body. Septic shock is an example of what happens when an inflammatory response is triggered throughout the body. Shock occurs when bacteria or their products reach high enough levels in the bloodstream to trigger complement activation, cytokine release, and the coagulation cascade in many parts of the body. The effects of this are illustrated in **Figure 19**. High levels of cytokines, especially TNF-α, IL-1, IL-6, IL-8, and IFN-γ, cause increased levels of polymorphonuclear leukocytes (PMNs or granulocytes) in the blood and encourage these PMNs to leave the blood vessels throughout the body. This leads to massive leakage of fluids into surrounding tissue. PMNs and macrophages activated by IFN-γ also damage blood vessels directly, resulting in loss of fluid from blood vessels. Activation of complement throughout the body further increases the transmigration of phagocytes. The vasodilating action of C3a, C5a, leukotrienes, and prostaglandins contributes to leakage of fluids from blood vessels and further reduces the ability of blood vessels to maintain blood pressure. Some cytokines cause inappropriate constriction and relaxation of blood vessels, an activity that undermines the ability of the circulatory system to maintain normal blood flow and normal blood pressure. Widespread triggering of the coagulation system produces the clots that can plug capillaries (manifested as DIC). More seriously, it depletes the blood of essential clotting factors, so that damage to endothelial cells caused by phagocytes and cytokines leads to hemorrhages in many parts of the body. Hemorrhages not only contribute to hypotension, but also

damage vital organs. Once septic shock enters the phase where organs start to fail, it is extremely difficult to treat successfully, and the death rate exceeds 70%.

Treatment is most likely to be effective if it is begun early in the infection. Diagnosis of septic shock in its early stages is not straightforward, however, because the early symptoms of shock (fever, hypotension, and tachycardia) are nonspecific. Also, the transition from the early stages to multiple-organ failure can occur with frightening rapidity. Hundreds of thousands of cases of septic shock occur in the United States each year. Many of these occur in patients hospitalized for some other condition than infectious disease. Accordingly, a massive effort has been made to develop new techniques for treating septic-shock patients more successfully, especially in cases where the disease has reached the point that treatment with antibacterial agents is no longer sufficient to avert disaster. Early efforts to combat septic shock centered on administration of glucocorticoids, which downregulate cytokine production. Physicians have long believed that administration of steroids or ibuprofen would help shock victims because these compounds mitigate various aspects of the inflammatory response, but there is now general agreement that glucocorticoid treatment is not effective in treating most types of shock. Clinical trials have now shown no significant effects from either of these therapies. More recently, attention has focused on the cytokines, such as TNF-α, which seem to play such a central role in the pathology of shock. Antibodies or other compounds that bind and inactivate cytokines have been tested for efficacy in clinical trials. The outcome of early clinical trials has been disappointing, but newer anti-cytokine agents now being tested appear to be more promising. Nonetheless, it is clear that this

type of therapy will never be as effective as catching septic shock in its very early stages. Antibiotic therapy administered early enough can stop the shock process, but the right antibiotic must be chosen. This means more laboratory tests, which is in conflict with the aim of health management organizations that want to save money by minimizing the use of tests. Microbiological testing tends to be expensive because it requires skilled technicians and is less automated than other types of clinical tests. Although it is true that bacterial identification and antibiotic susceptibility tests cost money, septic shock costs even more. If the causative bacteria are resistant to most antibiotics, the cost can rise by as much as 100-fold. Moreover, patients who have developed shock and survived may have experienced strokes and permanent damage to vital organs. The effect of shock, in terms of risk of untimely death, lasts years beyond the actual shock experience. Then there are the lawsuits. Patients who through neglect or misdiagnosis develop strokes or other long-term damage are starting to sue hospitals. Consumer advocates are also suing to obtain information about infection rates in hospitals, especially those leading to serious conditions like septic shock. The rate of postsurgical infections is higher than most people realize, especially in large big-city hospitals. Increasing antibiotic resistance is only making things worse. Not surprisingly, hospitals guard their incidence of hospital-acquired infection figures in strict secrecy. One impediment to early diagnosis of shock has already been mentioned: the nonspecific nature of the signs and symptoms. Another impediment to early diagnosis is the fact that so many different types of bacteria can cause septic shock. A definitive microbial diagnosis cannot be made in about a third of patients with clinical signs of sepsis. Bacteria are the microorganisms most frequently implicated in septic shock (approximately

80% of cases), but many different species of Gram-positive and Gram-negative bacteria can cause shock. Because no single antibiotic is effective against all of these bacterial pathogens, it is important to determine the species of bacterium causing the infection. Current research efforts are focused on defining a spectrum or profile of molecular biomarkers that are indicative of different sepsis causes. The bad news so far is that rapid detection methods like PCR may not be effective because concentrations of bacteria below the PCR detection level can cause sepsis. (41)

Currently, over 30 pharmaceutical products have been in the development stage to treat this condition, yet only few have reached the market (44). Many of these target specific inflammatory mediators have been unsuccessful because of the complex nature of sepsis. For the treatment of sepsis, there are a few products that are being investigated in clinical studies via blocking different mechanisms of the body's innate immune system. Eli Lilly's Xigris® was one of the few drugs currently available on the market to treat sepsis. Xigris® is a recombinant human activated protein C that has anti-inflammatory, anti-thrombotic and pro-fibrinolytic properties to block the coagulation cascade which plays a critical role in the development of organ failure due to sepsis (45). In addition, simvastatin and atorvastatin had also shown to have some non-specific anti-inflammatory effects contributing to their clinical benefits in treating sepsis (50). However, statins are currently not been considered as a treatment for sepsis. To find a more specific target, scientists have identified TLR4 as one of the candidates in blocking the innate immune system. Only two TLR4 antagonists, E5564 and TAK-242, have made far into the clinical phase, but eventually failed because they were found lacking in efficacy. (46) (47)

TLR4-active compounds

The past 15 years of scientific research on TLR4 led to the discovery of an incredible number of TLR4 ligands and interactors. Many of them come from medicinal chemistry, as they are synthesized with the purpose of being TLR4 agonists or antagonists. Most of them are disaccharide Lipid A mimetics. But a surprisingly vast number of molecules that are able to bind TLR4 or its co-receptors are of natural origin, mostly coming from plants. **Figure 20** sums up (some of) the known TLR4 ligands.

The most obvious and important use of TLR4 ligands is to inhibit the LPS-triggered TLR4 activation, that is the key event in causing sepsis and septic shock, which originate from a dysregulated TLR4 activation with a subsequent life-threatening cytokine overproduction, as seen in previous section. Thus, molecules with activity as TLR4 antagonists are interesting hit or lead compounds for anti-sepsis drug development.

A second reason of interest for the discovery and characterization of TLR4 ligands is the development of new agonists that could be employed as adjuvants, i.e. as molecules that elicit a controlled immune response, which is required for a proper vaccine function. Many efforts have been made to synthesize new Lipid A mimetics deprived of the toxic activity that can induce the response pathway following TLR4 binding.

TLR4 is also involved in sterile inflammation, that is, an inflammation not caused by viruses and microbes (32; 33). Sterile inflammation is often the product of tissue injury, and it is significant that main endogenous ligands for TLR4 are molecules either released by damaged or dying cells or triggered by

chemical or physical tissue injury (such as heat shock proteins, extracellular matrix degradation products, HMGB-1, β-defensin, surfactant protein A, minimally modified LDL). Some of those endogenous ligands are also shown in **Figure 20**. Since TLR4 is somehow involved in the onset of neuropathic pain, very recent studies were focused on the use of TLR4 antagonists for neuropathic and chronic pain relief. (34)

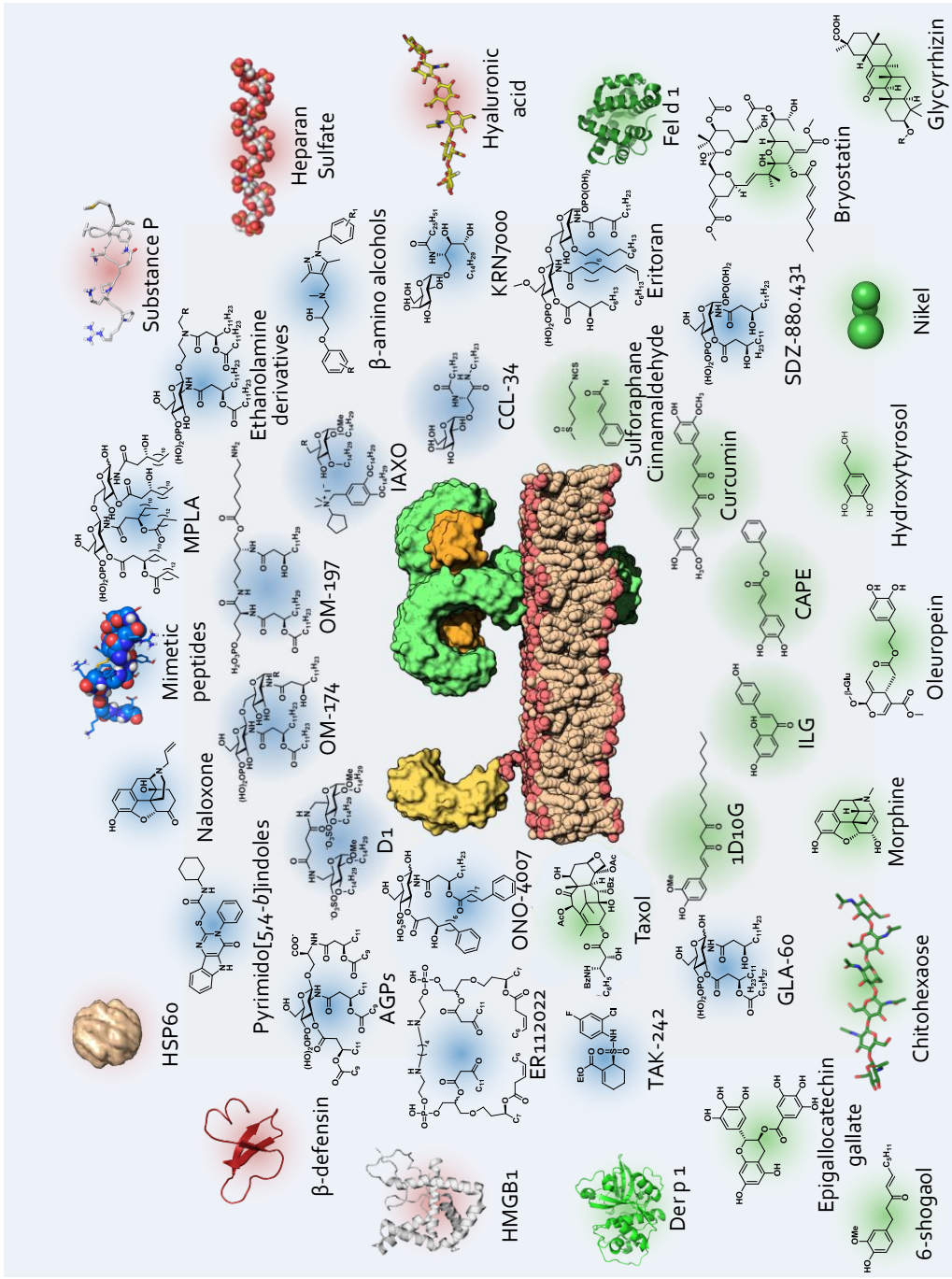
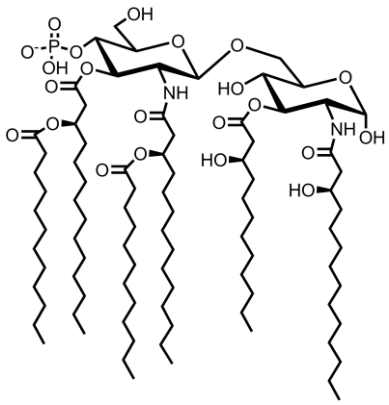


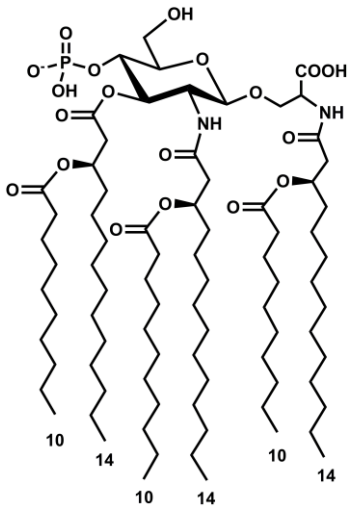
Figure 20. Some of known TLR4 interactors. Blue shade: synthetic ligands. Green shade: natural ligands. Red shade: endogenous ligands. The list is not exhaustive. Self work.

In next section a short list of compounds active on the TLR4·MD-2 complex is presented.



Monophosphoryl lipid A (MPLA)

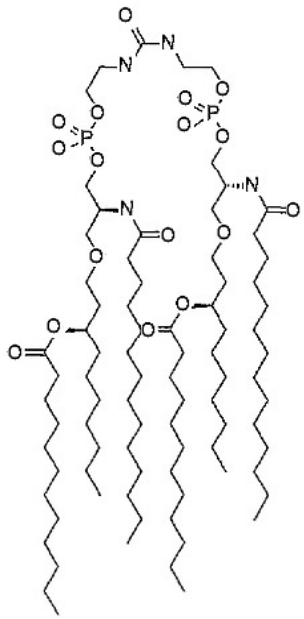
MPLA was initially obtained by chemical modification of Lipid A, as a mixture of the two differently acylated variants, the hexa- and pentaacyl monophosphates. The two forms were equivalent with respect to adjuvant activity, but the former is up to 20-times more active than the latter in inducing nitric oxide synthesis in murine macrophages (61). These observations parallel studies on lipid A variants showing that hexaacylation is a prerequisite to the full expression of endotoxic activities and that underacylated lipid As tend to be much less active and may even possess antagonist activity. Chemically synthesized, pure MPLA is the first TLR4 agonist to be approved for use as an adjuvant in human vaccines (*e.g.* FENDrix® vaccine for hepatitis B virus) in Europe. The effect of fatty acid structure on endotoxic (agonist) activity in the MPL series was investigated systematically by Johnson *et al.*, by synthesizing a series of molecules that differ only in the length of secondary lipid chains, from C_4 to C_{12} (57). Secondary fatty chains length had an unpredictable and profound effect on the agonist activity of MPLA. Short chains derivatives (C_4 or C_6) were inactive in stimulating TLR4-dependent cytokine production, whereas the C_{10} homolog exhibited the highest level of cytokine induction and the greatest pyrogenicity (63).



Aminoalkyl glucosaminide phosphates (AGP*s*)

In AGP*s* the reducing glucosamine residue has been replaced by an acylated amino acid or another acylated function. These compounds generally retain significant activity as TLR4 modulators but have a simplified structure with a reduced number of stereogenic carbons, and can be therefore obtained by a simpler synthesis than lipid A. Many homologues have been synthesized, by changing the length of the alkyl chains and a systematic SAR study has been conducted on these molecules (64).

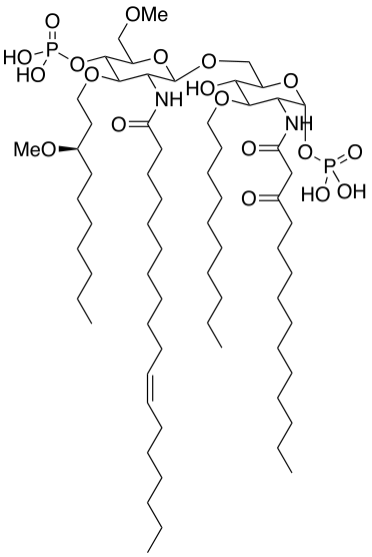
Lipid A mimetics with linear scaffold



Compound ER112022 (65), ER803022 and other similar compounds were developed at the Eisai Research Center (Boston) as acyclic lipid A analogues. All these compounds containing six symmetrical lipid chains and two phosphate groups attached to linear linkers with different chemical structures (66). Compound ER112022 activated NF-κB in HEK293 cells transfected with TLR4 and MD-2 in a TLR4-dependent manner; CD14 was not essential to the activity but enhanced the sensitivity of the response (65).

E5531 and E5564

Compound E5531 is an analogue of the lipid A from *Rhodobacter capsulatus* (a non-toxic lipid A and a LPS antagonist), first developed in Eisai Laboratories (Boston). Although E5531 demonstrated potent inhibition of LPS toxicity when added to blood *in vitro* and *in vivo*, this activity decreased by time because of its interaction with plasma lipoproteins.



A second-generation LPS antagonist, Eritoran (E5564) was then developed by Eisai, derived from the structure of the weakly agonistic endotoxin found in *R. sphaeroides*. E5564 is a potent *in vitro* antagonist of endotoxin that directly binds to the hydrophobic pocket of MD-2 (10), competitively inhibits lipid A binding to the same site and thereby prevents TLR4 dimerization and intracellular signaling, as indicated by crystallographic data (PDB ID: 2Z65 and 3ULA).

On January 2011, after a Phase III clinical trial, based on the preliminary findings from the ACCESS (*A Controlled Comparison of Eritoran and Placebo in Patients with Severe Sepsis*) trial, Eisai Inc. announced that the company will not submit marketing authorization applications to regulatory authorities in the United States, the European Union and Japan, as previously planned. The decision was based on the fact that the study did not meet its primary endpoint of reduction in 28-day all-cause mortality in patients with severe sepsis. (58)

Other synthetic and natural compounds with a chemical structure unrelated to that of lipid A are active on TLR4-LPS signaling by binding MD-2. Because their profound structural difference with lipid A, these molecules act as antagonists rather than agonists, since their binding to MD-2 does not result in the formation of the activated $(TLR4{\cdot}MD\text{-}2)_2$ complex required for signaling. Among these, it will be cited:

- The fluorescent probe bis-ANS, 4,4'-bis(1-anilinonaphthalene-8-sulfonate)
- The anti-rheumatic auranofin (*Ridaura*®)
- The CEPT inhibitor JTT-705
- The natural mitotic inhibitor Paclitaxel (*a.k.a.* Taxol)
- The natural polyphenol Curcumin found in turmeric
- The synthetic compound TAK-242, that binds to the intracellular portion of TLR4, thus blocking its signaling cascade

Our research group synthesized TLR4-active glycosylamino- and benzylammonium-lipids, that showed activity in inhibiting LPS- and lipid A-induced cytokine production in innate immunity cells such as macrophages or dendritic cells (59; 60). They were also employed in very recent (still unpublished) studies on aortic aneurism, skin inflammatory diseases and Lateral Amyotrophic Sclerosis. (58) (50) (57)

Purpose of the work

The purpose of the current thesis project is to synthesize new chemical entities that are active as antagonists on TLR4 receptor complex.

In order to reach this objective, new monosaccharide molecules were designed, based on the structure of Lipid A, the endotoxic portion of LPS, that binds to TLR4-associated receptors. The total synthesis of LPS (described in (64) and (65) and of LPS analogues with a disaccharidic scaffold (like Eritoran or others) is very complicated, because it implies the orthogonal protection of the many hydroxyl groups present on precursor sugars and a glycosylation step which generally has low yields or is hardly reproducible in different laboratories. Thus we decided to simplify the structure of LPS, by designing monosaccharides that mimic the reducing or the non-reducing moiety of LPS. With the goal of "keeping it simple" at first we decided not to insert branched chains in the structure, which synthesis requires a more complicated pathway. The acyl chains of LPS, composed by either branched or not (R)-3-hydroxymyristic acid, are therefore substituted in our analogues by a simple, linear chain of myristic acid ($C_{13}H_{27}COOH$).

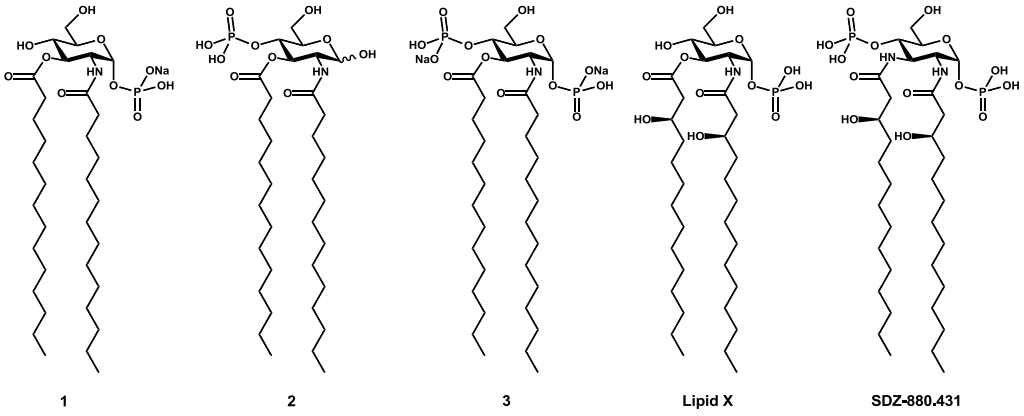


Figure 21. Chemical structure of synthetic monosaccharides **1** - **3**, of Lipid X and diphosphate Lipid X analogue SDZ-880.431.

Although many Lipid A mimetics were developed in recent years, the medical need for antisepsis agents is not decreased. Safer and more active compounds are still required in order to understand and combat the life-threatening sepsis and septic shock syndromes. Since full Lipid A mimetics are very complicated molecules and their synthesis involves a lot of chemical reactions, we decided to design and synthesize very simple monosaccharide mimetics of Lipid X, one of the precursors of Lipid A synthesis in bacteria which proved to be a good antagonist of LPS in experimental models. (66)

Other Lipid X monosaccharide mimetics were synthesized and tested in the past 25 years, (GLA compounds, (67) (68)) with a vast (yet not exhaustive) combination of different chain types (branched or not), lengths (2-18 carbon atoms), bonds (ester, amide, ether, amine, …), phosphate positions (1, 4, 6, …). Quite surprisingly, most of those compounds, despite being thoroughly tested for agonism on cells and/or whole blood, were not tested for the antagonism against LPS. This reveals that probably at that time the research groups were more interested in the development of new vaccine adjuvants rather than antisepsis agents. TLR4 was still unknown at that time, so the scientific research before year 2000 didn't take it in account for LPS responses. Also, some molecular structures with particular combinations of acyl chains and phosphate groups were not developed.

So, we decided to synthesize the most simple Lipid X mimetics with not-branched, not-hydroxylated C-14 chains and to investigate the activity of three different phosphate positions (1, 4, and both 1 and 4).

In this thesis, the synthesis and the biological characterization of monosaccharides **1**-**3** is presented (**Figure 21**). Compound **1** corresponds to a

Lipid X mimetic with phosphate in the anomeric (C-1) position, while **2** has a phosphate ester in C-4 position and **3** is phosphorylated on both C-1 and C-4 positions. Extensive structure-activity studies are available on lipid X mimetics formed by a GlcNac monosaccharide with a C-4 phosphate and acylated in C-2 and C-3 positions with different linear and branched FA chains. (68) (67) While compounds with two C_{14} FA acyl chains in C-3 and a linear C_{14} chain in C-2 have TLR4 agonist activity in human and mouse macrophages, compounds with different acylation patterns (included compound **1**, named GLA-26, with two linear acyl chains) acted as agonist in murine macrophages and antagonist in human monocytes. (67) (69) Compound **2** with a phosphate in C-1 has been described (compound 880.244, (70)) as a non-active or very weak TLR4 modulator. (70) (71) (72) Among all lipid X synthetic analogues, N^2,N^3-diacyl-glucosamine-1,4-bisphosphate (also known as 3-aza Lipid X 4-phosphate or SDZ-880.431, **Figure 21**) showed the most potent antagonis activity on both murine macrophages and human monocytes (**Figure 22**). (66) Compound **3** closely resembles to diphosphoryl lipid X but its structure is further simplified by the presence of a glucosamine scaffold (instead of 2,3-diamino-2,3-dideoxyglucose, which synthesis is more complicated) and the removal of C-3 hydroxyls on FA chains.

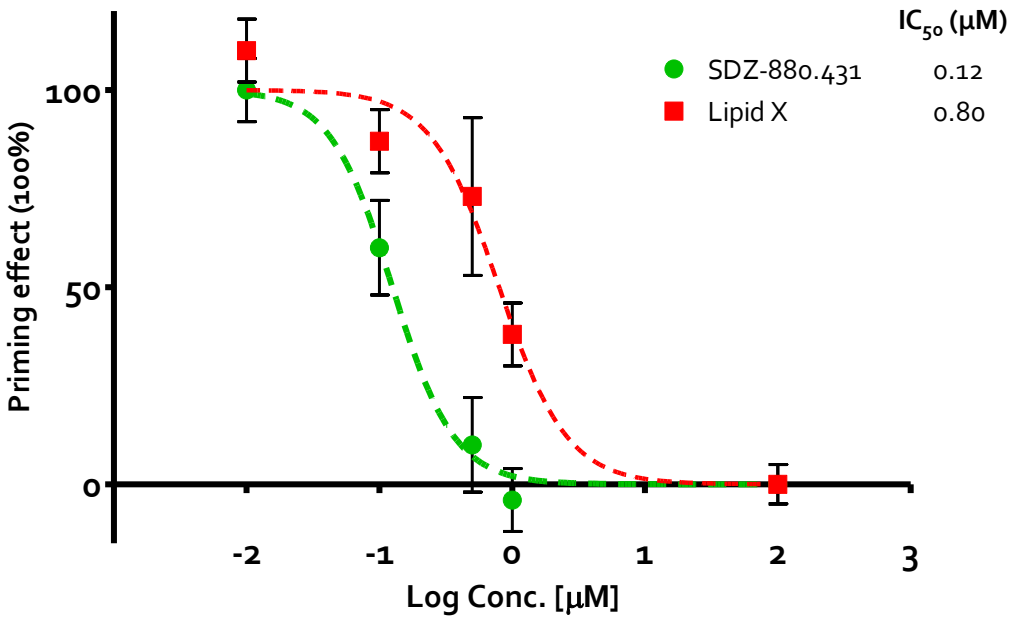


Figure 21. Priming effect of Lipid X and synthetic mimetic SDZ-880.431 on neutrophils in presence of *Salmonella Minnesota* Re595 LPS (10 ng/ml). Normalized results are shown as the mean ± SE, (*N* = 3). 0 % = pyrogen-free buffer, 100% = LPS 10 ng/ml. Data derived from (66).

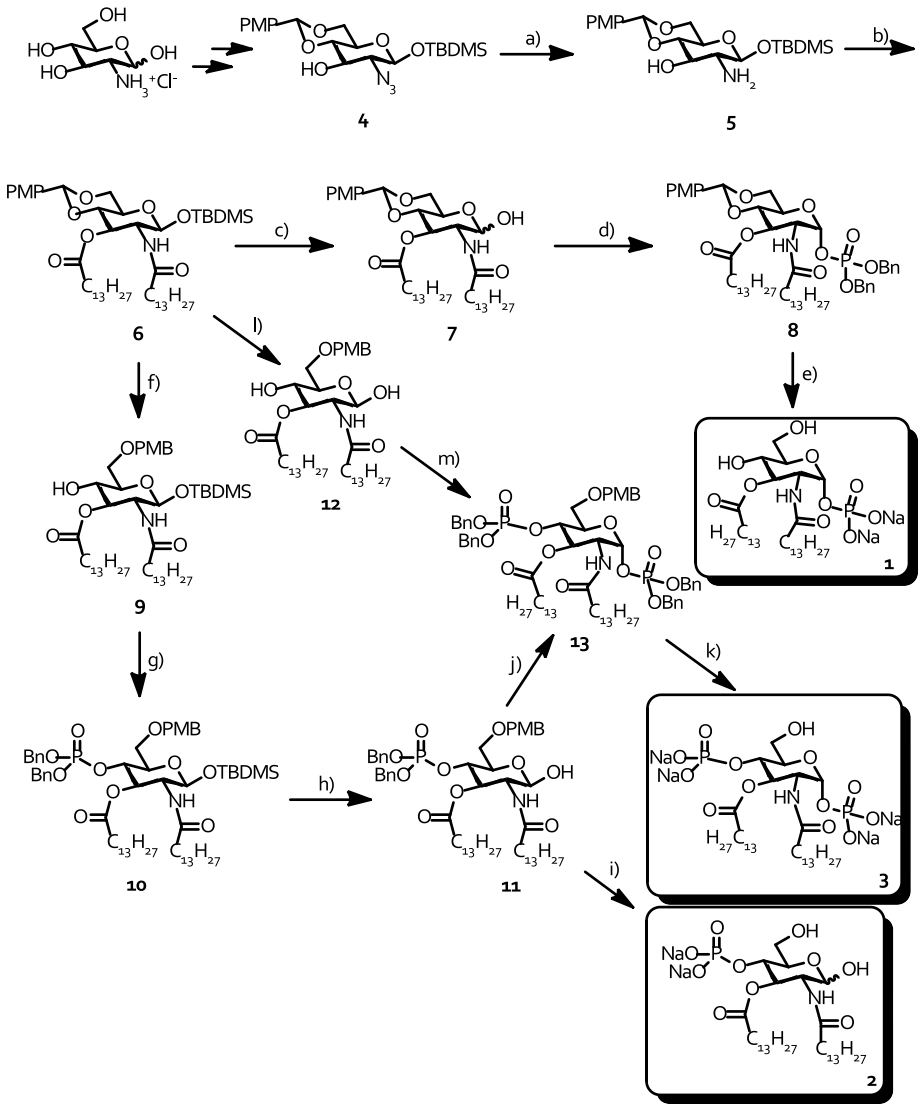
Furthermore, for the most promising compound, a functionalization has been designed in order to convert it into a fluorescent derivative or into a paramagnetic probe for magnetic resonance imaging (MRI).

Results and Discussion

Synthesis of lipid X mimetics **1**-**3**

Compounds **1**-**3**, previously shown in **Figure 21**, were synthesized using a divergent strategy, starting from the common intermediate **4**, (73) obtained from commercial D-glucosamine hydrochloride (**Scheme 1**). Compound **4** has been treated with PPh_3 in THF/H_2O to transform the azido group into amine (compound **5**), then acylated in C-2 and C-3 positions using myristic acid in the presence of the condensing agent 1-Ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC), thus affording **6**. Monosaccharide **6** was deprotected on anomeric C-1 position using tetrabutylammonium fluoride (TBAF) and AcOH to obtain **7**, which was phosphorylated using tetrabenzyldiphospate in the presence of lithium bis(trimethylsilyl)amide affording the α-anomer **8** exclusively. Catalytic hydrogenation with Pd/C allowed to remove simultaneously *p*-methoxybenzylidene and benzyl groups on phosphate, affording compound **1**. Alternatively, the benzylidene protecting group of compound **6** was reductively opened by treatment with $NaCNBH_3$ in dry THF to obtain compound **9** protected as *p*-methoxybenzyl (PMB) ether on C-4 position. Compound **9** was phosphitilated using dibenzyl *N,N*-diisopropylphosphoramidite and imidazolium triflate, then oxidized to phosphate with *m*-chloroperbenzoic acid to obtain **10**. Compound **10** was deprotected on C-1 using TBAF and AcOH affording **11**, then catalytic hydrogenation with Pd/C allowed to obtain compound **2**. To have access to compound **3**, two synthetic routes were employed. The first one involved the phosphorylation of the anomeric position of compound **11** with $LiN(TMS)_2$ and tetrabenzyldiphosphate followed by the hydrogenation on Pd/C. An alternative route was serendipitously discovered, leading to the synthesis of **3**

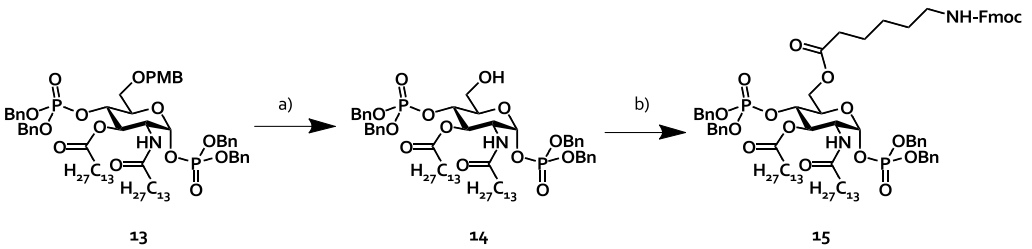
through less synthetic steps. That implied the treatment of compound **6** with $NaCNBH_3$ with regioselective opening of the p-methoxybenzylidene ring and formation of C-6 PMB ether. This reaction was quenched by adding an acidic solution (HCl 1M in dioxane at pH1 for 1h), that promoted the removal of the anomeric TBDMS protective group, thus obtaining compound **12** in one reaction step. The double phosphorylation on C-1 and C-4 positions using phosphoramidite plus oxidation method, followed by catalytic hydrogenation afforded compound **3**. Compounds **1**, **2** and **3** were recovered as triethylammonium salts after hydrogenation, then treated with Amberlite IR 120 Na^+ exchange resin to exchange the counterion. The final compounds for biological characterization were used as sodium salts.



Scheme 1. a) PPh_3, THF/H_2O, 60 °C, 87%; b) $C_{13}H_{27}COOH$, EDC, DMAP, DCM, 97%; c) TBAF, AcOH, THF -15 °C to RT, 76%; d) tetrabenzyl diphosphate, $LiN(TMS)_2$, THF, -78 to -20 °C, 43%; e) H_2, Pd/C, AcOH, MeOH/DCM, quant.; f) $NaCNBH_3$, 4Å MS, THF, 84%; g) $(BnO)_2PNi\text{-}Pr_2$, Imidaz. Trifl., DCM, then *m*CPBA, 0 °C, 42%; h) TBAF, AcOH, THF, -15 °C to RT, 57%; i) H_2, Pd/C, AcOH, MeOH/DCM, quant.; j) tetrabenzyl diphosphate, $LiN(TMS)_2$, THF, -78 to -20 °C, 71%; k) H_2, Pd/C, AcOH, MeOH/DCM, quant.; l) $NaCNBH_3$, 4Å MS, THF, then HCl in dioxane to pH=1.5, 61%; m) $(BnO)_2PNi\text{-}Pr_2$, Imidaz. Trifl., DCM, then *m*CPBA, 0 °C, 79.5%.

Synthesis of C-6 functionalized derivative of Compound **3**.

In order to couple Compound **3** to a fluorescent or paramagnetic probe (for microscopy and MRI studies, respectively) a functionalization has been designed that involves C-6 *p*-methoxybenzyl ether removal from Compound **13** with TFA 5% in DCM for 30 minutes to give intermediate **14,** which is subsequently coupled to *N*-Fmoc 6-aminohexanoic acid with EDC and DMAP to give the Fmoc-C_6-derivative **15** (**Scheme 2**). The intermediate compound **14** resulted unstable in solution at room temperature, since TLC and NMR analysis revealed that a phosphate migration occurs between C-4 and C-6 positions of the sugar, giving the 1,6-bisphospate molecule. These reactions were performed by Davide Zaccarin and the products are about to be fully characterized (NMR and MS).



13 **14** **15**

Scheme 2. a) TFA 5% in CH_2Cl_2, 0.5 h, RT, 86%; b) *N*-Fmoc 6-aminohexanoic acid, EDC, DMAP, dry CH_2Cl_2, 47.5%.

Further work on compound **15** will involve Fmoc removal from the terminal amine of the C_6 linker and coupling to fluorescein or to a synthetic Gadolinium-chelating moiety developed in the research group of Dr. Philippe Maurin in Lyon, France. This synthesis is planned for the first months of 2014.

Compound **15** will also be hydrogenated to obtain the free phosphate groups and its biological activity will be tested as soon as possible in order to understand if the functionalization of the molecule on C-6 position with a linker and a bulky group (Fmoc) could affect the LPS antagonism. This could give give a good indication of the possibility to functionalize the groups not directly involved in binding to endotoxin receptors without impairing the biological activity of the molecule.

Inhibition of LPS-Induced, TLR4-dependent NF-κB activation in HEK-Blue™ cells

The ability of molecules **1**, **2** and **3** to interfere with LPS-triggered TLR4 activation in HEK-Blue™ cells was investigated. HEK-Blue™ cells are HEK293 cells stably transfected with human TLR4, MD-2, and CD14 genes. In addition, HEK-Blue™ cells stably express a secreted Alkaline Phosphatase (sAP) produced upon activation of NF-κB. LPS binding activates TLR4 and NF-κB leading to sAP secretion, which is detected by an alkaline phosphatase substrate in cell culture media. When supplied alone, compounds **1**-**3** were unable to stimulate TLR4-dependent sAP production in a range of concentrations between 0 and 25 μM, thus confirming the lack of any agonist activity for the three monosaccharides on human TLR4. Cells were then pre-treated with increasing concentrations of synthetic monosaccharides (from 0 to 25 μM) and then stimulated with *E. coli* O55:B5 LPS. In this concentration range compounds **1** and **2** were inactive or vey weakly active in inhibiting LPS-stimulated TLR4 signaling, while **3** was active. The experiment was run for molecule **3** at two LPS concentrations, 10 ng/mL and 1μg/mL, by administering LPS 30 min after the pretreatment with compound **3**. At both these LPS concentrations the % activation of HEK cells reached similar values, thus indicating that the lower concentration (10 ng/mL) also saturated the signal corresponding to TLR4-dependent NF-κB stimulation (the 100% of vertical scale, **Figure 23**).

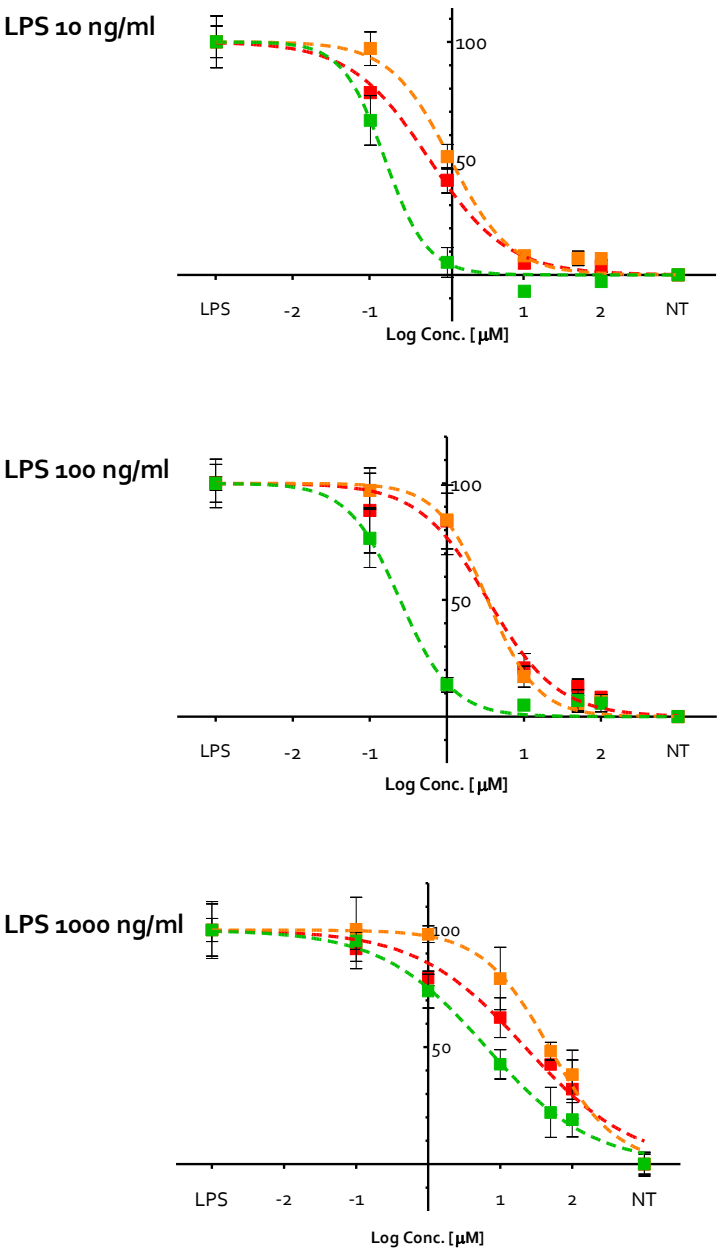


Figure 23. HEK-Blue™ cells assay of Compound **1** (red), Compound **2** (orange) and Compound **3** (green) in presence of three different LPS concentrations (10, 100 and 1000 ng/ml). Compounds **1-3** inhibit in a dose-dependent way the LPS-triggered TLR4 activation (normalized A_{405} n=3 independent experiments. 0%=DMEM, 100%=LPS).

	LPS 10 ng/ml	LPS 100 ng/ml	LPS 1000 ng/ml
Cpd. 1	0.57	3.42	22.11
Cpd. 2	1.08	3.37	47.7
Cpd. 3	0.15	0.25	6.36

Table 3. IC_{50} values for Compounds **1**-**3**, in presence of different LPS concentrations on human receptors expressed in HEK-Blue™ cells.

Compound **3** turned out to be more active as antagonist at low LPS concentration (10 ng/mL), with a calculated IC_{50} of 0.46 μM, while when LPS was more concentrated, the IC_{50} raised at 3.42 μM (**Figure 23**). As a negative control a HEK-293 cell line (*InvivoGen*) transfected with the same plasmids as HEK-Blue™ but without TLR4, MD-2 and CD14 genes was used and no effect was observed (data not shown).

Toxicity assay.

The toxicity of all compounds was assessed by MTT test and all compounds showed no or very limited toxicity up to the highest concentration tested (50 μM, **Figure 24**).

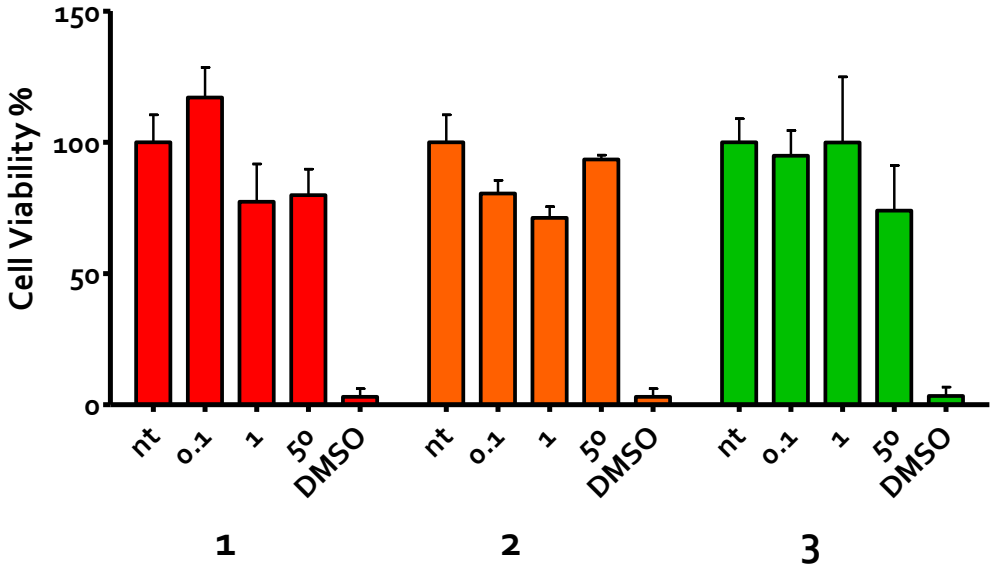


Figure 24. Cell viability assay (MTT). Normalized cell viability in presence of increasing concentrations of compounds **1**, **2** or **3** (μM) is indicated as % values (100%=DMEM, 0%=DMSO).

Inhibition of LPS-Induced TLR4 activation in murine macrophages

The ability of molecules **1**-**3** to modulate the LPS-stimulated TLR4 pathway was further investigated in bone marrow-derived murine macrophages (BMMΦ). TNF-α production was taken as read out of TLR4-MyD88 pathway activation. BMMΦ from wild type and $CD14^{-/-}$ mice were treated with increasing concentrations (0-25 μM) of compounds **1**-**3** in DMEM+BSA 0.03% in the presence or absence of LPS. Two LPS concentrations were tested: 10 ng/ml and 1μg/ml, and LPS was administered 30 min after the pretreatment with synthetic compounds. The LPS-induced TNF-α production after one night incubation was assessed by ELISA test (**Figure 25**). As expected, high concentration of smooth LPS activated TLR4 signaling also in the absence of CD14, (74) while at low LPS concentration the signal was absent in CD14-defective cells. Compound **1** was inactive in both cell types, molecules **2** showed a weak antagonist effect on wt macrophages at low LPS concentration. Compound **3** showed a dose-depended LPS antagonistic activity, in wt cells at low LPS concentration, and in $CD14^{-/-}$ at high LPS concentration (**Figure 25**). The antagonist activity on both wt and CD14-defective cells, suggests that molecule **3** competes with LPS for the interaction with both CD14 and MD-2·TLR4 complex.

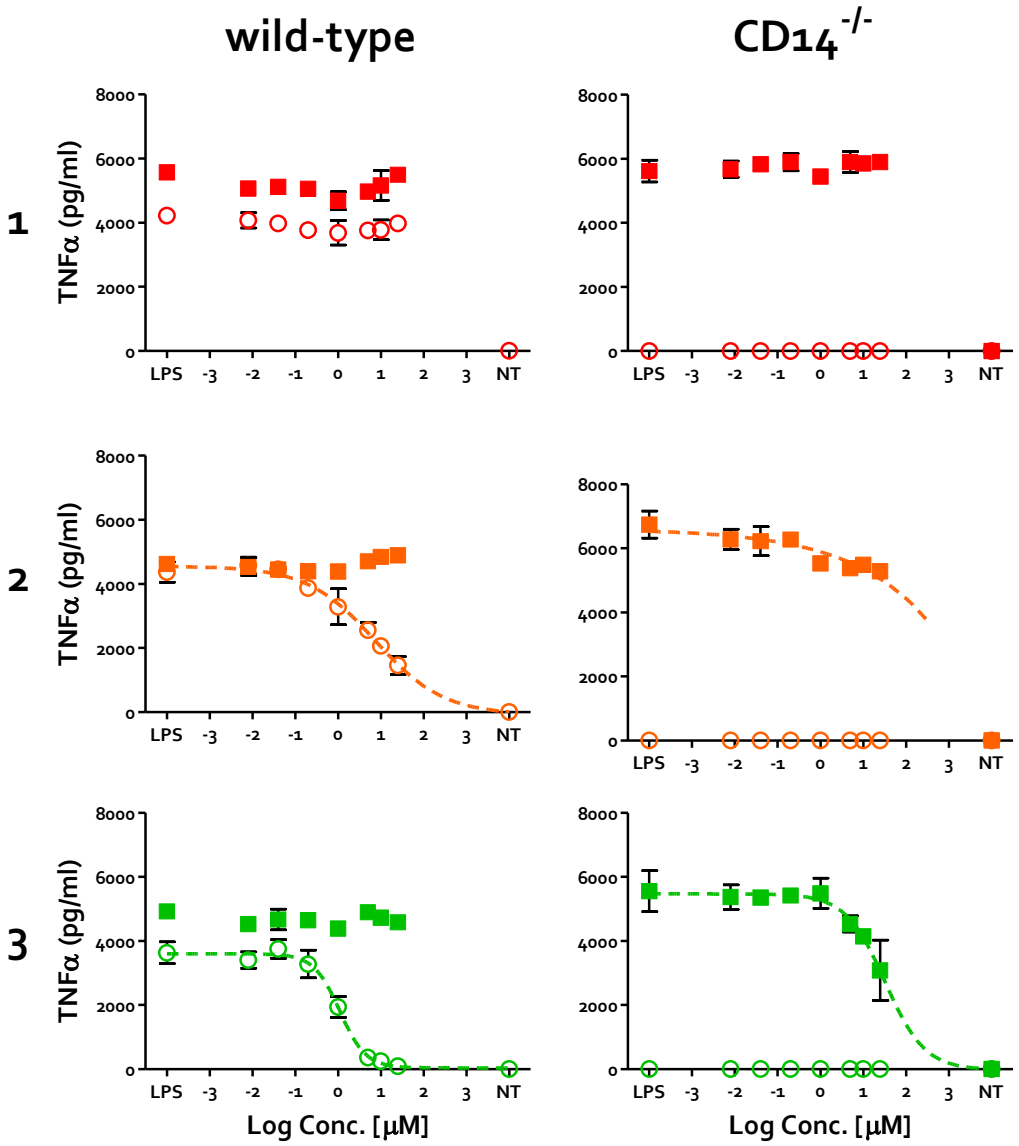


Figure 25. Effect of compounds **1** - **3** on LPS-induced TNF-α production by BM-derived macrophages. WT or CD14-/- BM-macrophages were pre-incubated for 30 min with synthetic compounds and then treated with high LPS concentration (1000 ng/mL, black squares) or low LPS (10 ng/mL, open circles). Readout was the TNF-α production after one night incubation.

Compound **3** selectively stimulates endocytosis of CD14 (and not of MD-2·TLR4 complex)

Since CD14 favors the activation of the TLR4-MyD88 pathway at low LPS doses, and compounds **2** and **3** are more active as inhibitors at low LPS concentrations, we evaluated whether CD14 could be directly targeted by compounds **2** and **3**.

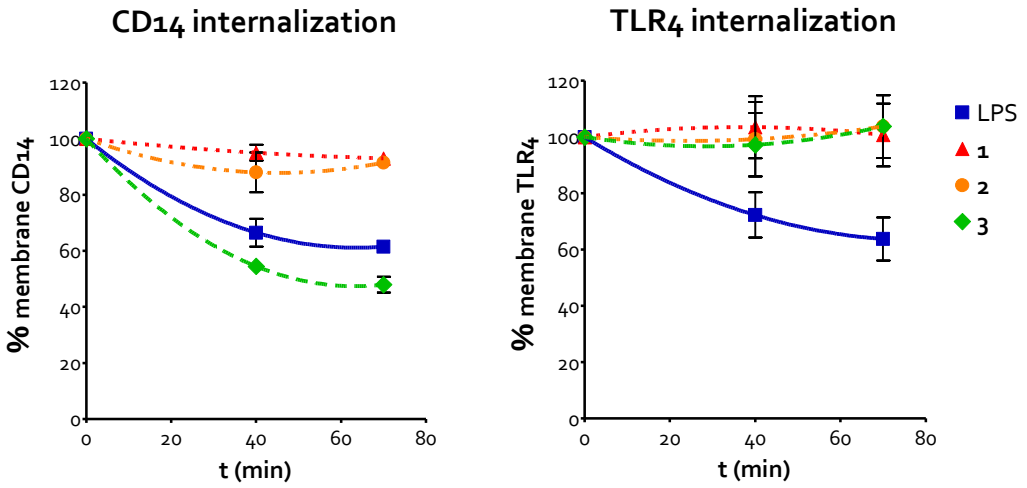


Figure 26. CD14 and TLR4 internalization induced by compounds **1**, **2** and **3**. BMMΦ were treated with LPS (1 μg/mL) or with synthetic compounds and incubated at 37 °C for the time indicated. Flow cytometry was then used to examine receptor endocytosis by determining the surface levels of the proteins indicated. Panels represent the mean fluorescence intensity (MFI) of specific receptor staining at each time point.

We analyzed the capacity of molecules **1**-**3** to induce CD14 and MD-2.TLR4 complex internalization on BMMΦ. As a matter of fact, after LPS or lipid A binding, CD14 is efficiently internalized together with the entire LPS receptor

complex. (43) BMMΦ were incubated with compounds **1**-**3** at a concentration of 10 μM, and the amount of CD14 remaining at the cell surface analyzed by flow cytometry over-time. Molecule **3**, showing the best antagonistic activity, was also capable of efficiently inducing CD14 internalization, while molecules **1** and **2** did not show any effect on CD14 surface expression (**Figure 26**). Interestingly MD-2.TLR4 complex was not internalized following molecules **1**-**3** exposure (**Figure 26**). These results suggest that molecule **3**, an antagonist of the TLR4 signal pathway, can directly interact with CD14 causing its internalization.

Further Biological characterization

The activity of diphosphate monosaccharide **3** was further characterized by Dr. Stefania Sestito in cooperation with the research group of Prof. Roman Jerala at the National Institute of Chemistry in Ljubljana (Slovenia).

Tests were performed on HEK-293 cells transiently transfected with TLR4, MD-2 genes and the NF-κB reporter luciferase gene. This generates endotoxin-responsive HEK cells very similar to the commercial HEK-Blue-TLR4 cell line, with a different reporter gene for NF-κB activation (a luciferase enzyme that is used to oxidize luciferin in presence of O_2 and ATP to oxiluciferin, with subsequently production of light which is quantified through a luminometer). LPS for all the experiments was used at a concentration of 50 ng/ml.

Human VS murine receptors

HEK-293 cells were transfected with human TLR4 gene (hTLR4), human MD-2 gene (hMD-2) and/or the corresponding murine genes (mTLR4 and mMD-2), in various combinations:

- hTLR4/hMD-2
- mTLR4/mMD-2
- hTLR4/mMD-2

Transfection was performed 4 hours before the activity assay.

The activation of cells transfected with the human receptors was compared to that of cells transfected with murine receptors, following a pre-treatment with Compound **3** and the LPS stimulus. Results are reported in **Figure 27**.

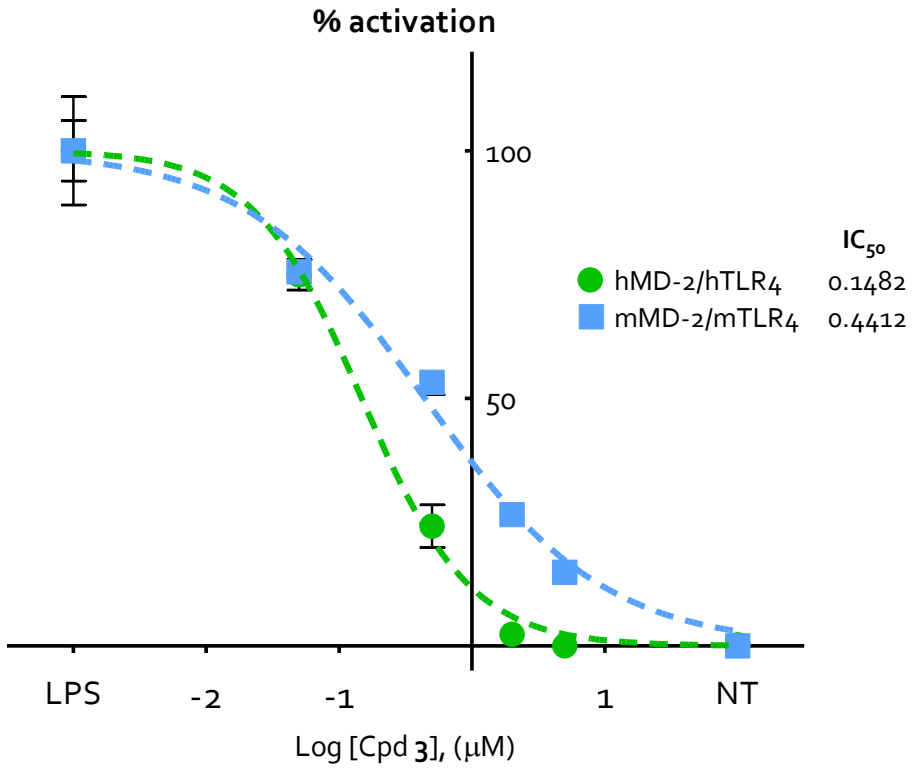


Figure 27. Normalized activation of HEK-293 cells expressing hTLR4 and hMD-2 receptors (green line) or mTLR4 and mMD-2 receptors (blue line), following exposure to LPS (50 ng/ml). Cells were pretreated with Compound **3** at 0.05–5 µM concentrations. Values represent normalized means ± St.Dev. (0% = NT, 100% = LPS) and are representative of 3 independent experiments.

It's noteworthy that the IC_{50} obtained in Slovenia for HEK-293 cells transfected with the human TLR4 and MD-2 receptors is the same obtained in Italy for the same experiment on HEK-Blue™ with the same receptors but a

different reporter. Also, the activity on murine receptors resulted lower than that on human receptors. This is coherent with the observations on murine BMMΦ.

Pre-treatment VS post-treatment

Compound **3** was tested on cells both in pre-treatment (60' before LPS administration) or post-treatment (60' and 180' after LPS administration), in order to individuate a possible activity of Compound **3** after cells were exposed to LPS, which will represent a positive hint for a future use of Compound **3** as an anti-sepsis agent. Results are reported in **Figure 28**.

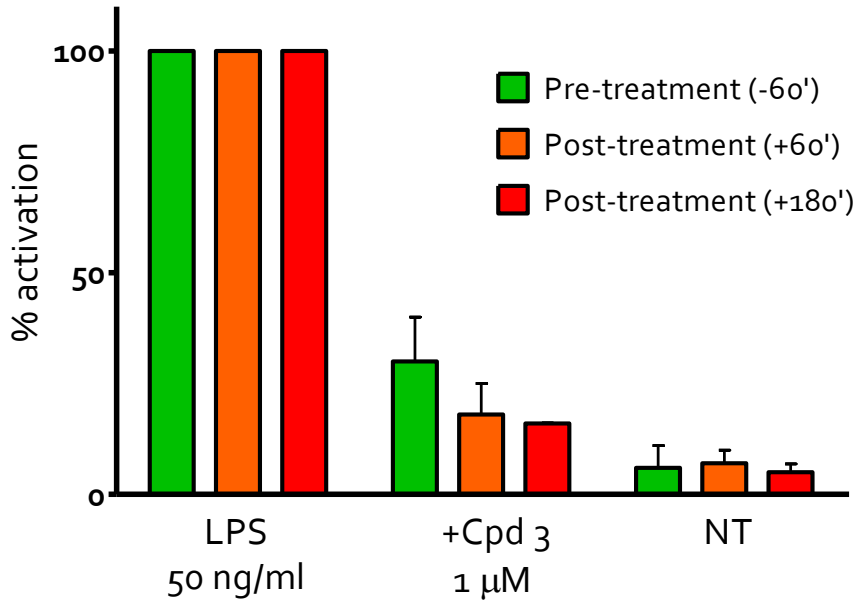


Figure 28. A comparison between NF-κB activation following exposure to LPS (50 ng/ml) in HEK-293 cells expressing hTLR4 and hMD-2 receptors pre-treated (60' before) or post-treated (60' and 180' after) with Compound **3**. Values represent normalized means ± St.Dev. (0% = 0, 100% = LPS). NT= vehicle.

When comparing the efficacy of Compound **3** before (-60') and after (+60', +180') LPS administration to cells, no significative differences were observed. This is a very interesting result for two different reasons. First, the possibility to reduce NF-κB activation also 1 and 3 hours after LPS exposure is a good indication that Compound **3** could be used in the treatment of sepsis, in which LPS is yet widespread in the organism (generally the blood concentration of endotoxin in septic patients is far lower than those used in the experimental models). A second reason of interest is based on the observation of very fast responses to LPS in cells. In particular, after a few minutes of exposure to LPS, endotoxin-sensitive cells ($TLR4^+$, $MD\text{-}2^+$, $CD14^+$) immediately start signaling through MyD88-dependent pathway that ultimately leads to NF-κB activation and translocation to the nucleus. Also, as early as 30' after exposure to endotoxins, CD14 induces its internalization along with TLR4 and MD-2·LPS complexes into endosomes, from where a TRIF/TRAM-mediated signal pathway leads to a "late" NF-κB activation. Thus, it seems very unlikely that any antagonism performed outside the cell at any time after the activation of "early" NF-κB signal pathway and, especially, after TLR4·MD-2 and CD14 internalization from the membrane could possibly reduce or block the synthesis of gene products controlled by NF-κB (as the reporter luciferase gene).

At the moment, we have no reasonable explanation for such a behavior.

MD-2 mutants

Compound **3** was also tested on two single-mutants of MD-2 which are used to investigate the structural determinants of binding between the endotoxin receptors and its ligands (**Figure 29**). In one of these mutants, a Lys residue

involved in LPS phosphate binding is replaced by a negatively charged amino acid (K122E). In the other mutant a Val residue located in the hydrophobic pocket is changed in a bulky Phe residue, thus reducing the volume of the pocket (V82F).

Results obtained for Compound **3** on these mutants are reported in **Figure 30**.

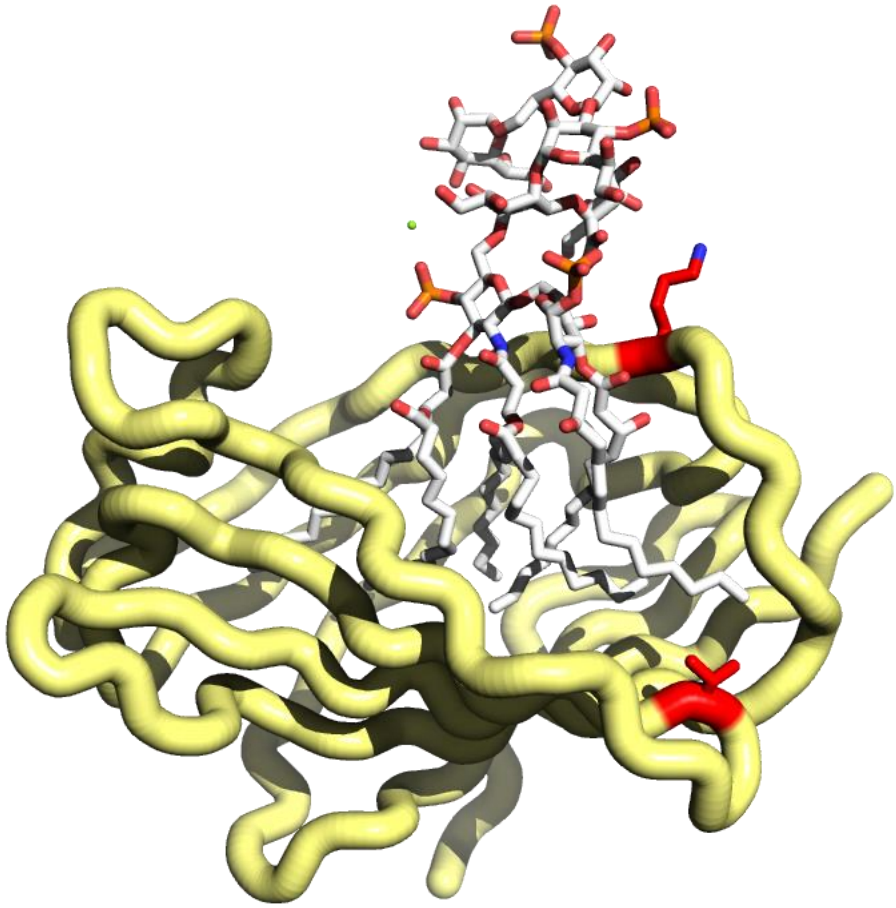


Figure 29. MD-2 crystallographic structure, in complex with R-LPS. The side chains of the two mutated residues are shown as red sticks (V82 and K122). Self work from PDB ID: 3FXI.

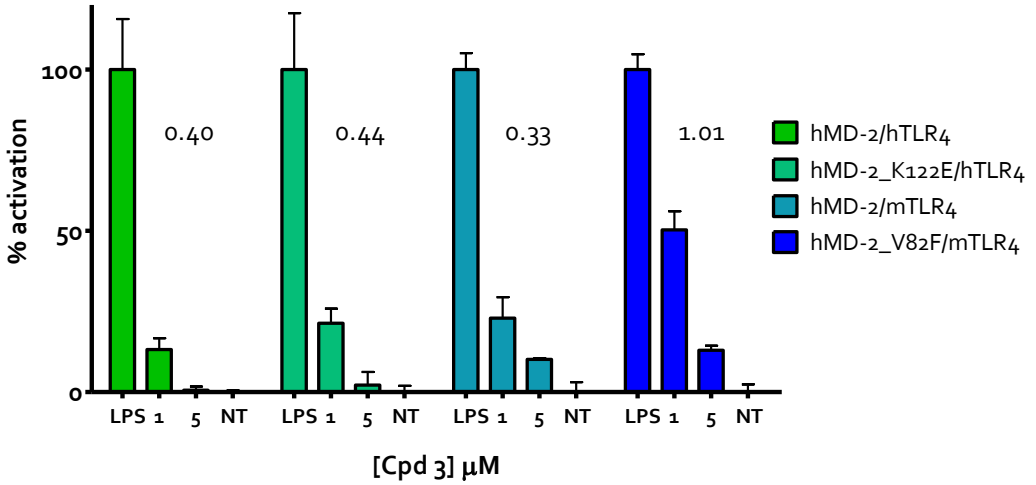


Figure 30. A comparison between NF-κB activation following exposure to LPS (50 ng/ml) in cells expressing different sets of endotoxin receptors pre-treated (60' before LPS) with Compound **3**. Values represent normalized means ± St.Dev. (0% = NT, 100% = LPS). Mutants are presented along with their wild type counterpart. Compound **3** was tested at 1 and 5 μM concentrations.

Other experiments were performed on a cell line of mature human monocytes (Mono Mac 6, Braunschweig, Germany), monitoring IL-6 production in presence of two concentrations of molecule **3**. Results are shown in **Figure 31**.

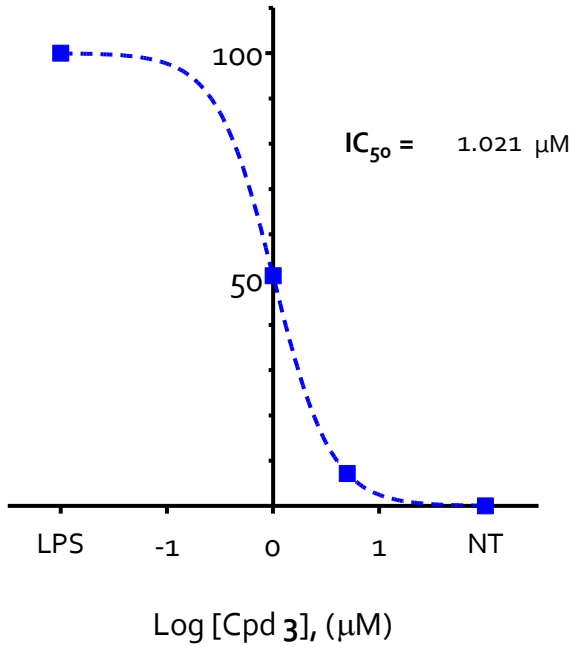


Figure 31. Normalized IL-6 production (pg/ml) following exposure to LPS (50 ng/ml) in Mono Mac 6 cells pre-treated (60' before LPS) with Compound **3** at 1 and 5 µM concentrations. Values represent normalized means ± St.Dev. (0% = NT, 100% = LPS) and are representative of 6 independent experiments.

NMR binding experiments: interaction between synthetic molecule **3** and MD-2

The binding of monosaccharide **3** to MD-2 receptor in solution was investigated by means of NMR techniques by Dr. Carlotta Ciaramelli. Because the tendency of compound **3** to form stable aggregates at the concentration required for NMR measurements (150-300 μM), it turned out to be impossible to detect ligand-protein interaction by Saturation Transfer Difference (STD) experiments. However, it was possible to record good quality ^{1}H-NMR spectra of **3**, of MD-2 and of a **3**:MD-2 mixture (ratio 5:1, at a 150 μM ligand concentration, **Figure 32**).

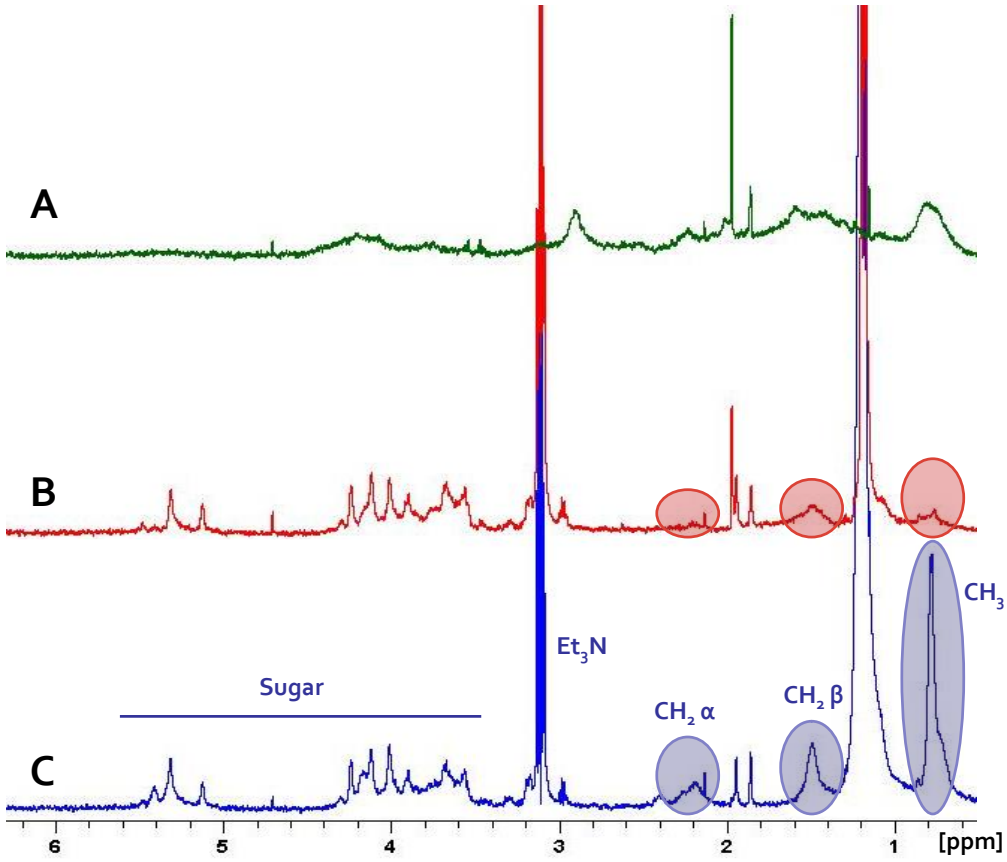


Figure 32. A) ^{1}H-NMR of 30 μM MD-2 protein in deuterated acetate buffer at pH=5, 298K, 120 scans; B) ^{1}H-NMR of 30 μM MD-2 protein and 150 μM of compound **3** in deuterated acetate buffer at pH=5, 298K, 120 scans; C) ^{1}H-NMR of 150 μM **3** in deuterated acetate buffer at pH=5, 298K, 120 scans.

Selective attenuation of the signals corresponding to the fatty acid C_{14} chain protons of **3** was observed upon addition of MD-2 to the monosaccharide sample solution (**Figure 32**). Although broadening of all the resonance signals was detected, the decrease in signal intensities was significantly higher for those protons belonging to the FA aliphatic chains, in particular for the signals of the ω–methyl groups and for those of the contiguous CH_2 moieties,

followed by the rest of the chain protons. In contrast, the intensity of the signals corresponding to the protons on the sugar ring proved to be practically unaltered. The experimentally observed reduction in intensity, due to specific line broadening of these signals, arises from the change in the transverse relaxation time of these signals. This dramatic change likely arises from the existence of interaction between **3** and MD-2, precisely focused on the aliphatic side chain region. In fact, the FA chains-MD-2 interactions outlined here are reminiscent of those previously observed for positively charged amphiphiles. (63) These data suggest the existence of a major interaction of both FA chains of the sugar ligand with the hydrophobic binding cavity of MD-2, as also confirmed by the docking calculations (see next section). The exchange process between the free and bound states provides the basis for the increased relaxation rate and the corresponding observed increase of line width. Additional features of the interaction were also investigated by DOSY (diffusion ordered spectroscopy) NMR. Firstly, the aggregation properties of **3** were evaluated by use of DOSY. The DOSY of the free ligand showed a strikingly small diffusion coefficient, corresponding to a high molecular weight species in solution. This evidence can be correlated with the observation that compound **3** forms relatively large aggregates in aqueous solution (**Figure 33**).

This was also confirmed by the determination of the critical micelle concentrations (CMC) for compounds **1**-**3** through pyrene fluorescence measurements, by Dr. Julio Rodriguez-Lavado (University of Seville, Spain). (75) The CMC values for compounds **1**, **2** and **3** are 13, 28, and 9 μM, respectively.

The DOSY spectrum of the **3**/MD-2 mixture (at 5:1 molar ratio) was also recorded. As stated above, the signals of the ligand aliphatic chains were no longer visible in the NMR spectrum, thus strongly suggesting the interaction of this part of the molecule with MD-2. In addition, the observation of a higher diffusion coefficient for the ligand molecule indicated that the existing aggregate for isolated **3** is indeed disrupted.

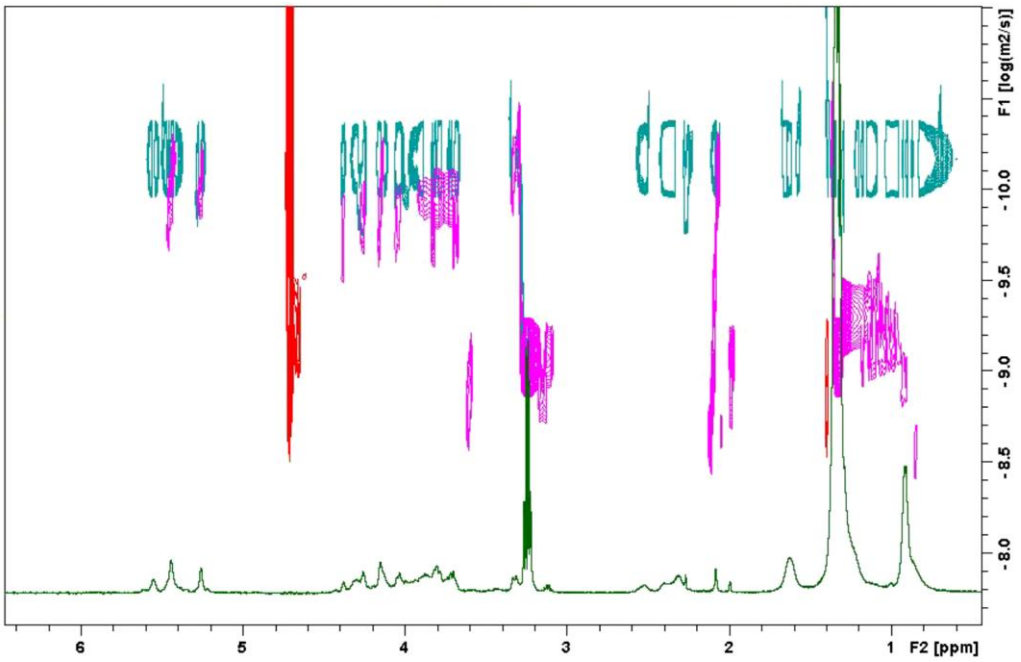


Figure 33. Blue: DOSY spectrum of compound **3** (300 μM) in deuterated acetate buffer at pH=5. Magenta: DOSY spectrum of **3**:MD-2 mixture at 5:1 molar ratio (330 μM compound **3** and 60 μM MD-2) in deuterated acetate buffer at pH = 5. Green: ^{1}H spectrum of compound **3** (300 μM) in deuterated acetate buffer at pH = 5.

Molecular modeling studies and docking of monosaccharide **3** with CD14 and MD-2

Docking studies were performed with the AutoDock (75) and AutoDock Vina (76) programs, by the protocol described in the Experimental section. Firstly, the use of these computational programs was validated by docking the natural antagonist lipid IVa, with employment of the X-ray crystallographic structure of the human MD-2 protein in complexation with lipid IVa (PDB ID: 2E59) as the starting geometry. (77) Both the AutoDock and the Vina programs satisfactorily reproduced the crystallographic binding mode, showing the four FA acyl chains inside the lipophilic pocket, and the phosphorylated glucosamine moieties located at the entrance to the cavity (data not shown). Once the docking procedures were validated, the same docking protocol was applied to compound **3**. Reasonable binding poses were predicted in both proteins (CD14 and MD-2), thus pointing to this compound being a suitable binder. The docked theoretical MD-2 binding energies for compounds **1** and **2** were significantly higher than that for compound **3**, by at least 8 $kJmol^{-1}$ (data not shown), thus pointing to a more favorable binding process for **3**. Analysis of the AutoDock and AutoDock Vina docked binding poses in MD-2 showed the ability of compound **3** to bind in two different fashions, with close predicted binding energies. Most of the best docked solutions corresponded to binding poses with the two FA chains deeply confined inside the MD-2 pocket, similarly to what had been seen with lipid IVa. One of the FA chains establishes hydrophobic contacts and CH–π interactions with Leu74, Phe76, Phe104, and Ile117, in a similar way to the equivalent FA chain present in lipid IVa (**Figure 34**). The second FA chain is directed into the region delimited by Ile52, Leu54, Phe121, Ile124, Tyr131, and Ile153, a subpocket also occupied by a FA chain in the complex with lipid IVa. In few cases, results from docking showed a second binding mode, with one

FA chain extending towards Val82 and placed over Ile124 (**Figure 35**). Interestingly, this Ile124 moves "up" in the agonist conformation, and its position is occupied by Phe126, thus acting as an agonist/antagonist switch. Polar interactions were also identified in some of the docked binding poses. One phosphate group participates in hydrogen bonds—with Ser118, for instance—and is always located in the vicinity of Lys58 and/or Lys122, similarly to one of the lipid IVa phosphates. The second phosphate group is found in the vicinity of positively charged side chains, such as that from Arg96 or, in other docked solutions, exposed to the outside. In addition, in some of the docking results either the amide CO group or an ester CO group from compound **3** was found to establish a hydrogen bond with the Ser120 CO group. These predicted binding poses are in agreement with the NMR experiments and provide a 3D model for the interaction of the FA chains with MD-2 protein, as well as for putative polar interactions involving the phosphate groups. Docking calculations for compound **3** into CD14 were also carried out. In this case, the CD14 protein, like the MD-2 protein, also possesses a highly lipophilic wide pocket, but with fewer charged residues in the opening portion. Compound **3** showed binding poses presenting both FA chains inside the pocket, with the polar phosphate groups and sugar placed at the entrance of the cavity (**Figure 36**), thus supporting the experimental evidence of CD14 binding properties for this compound.

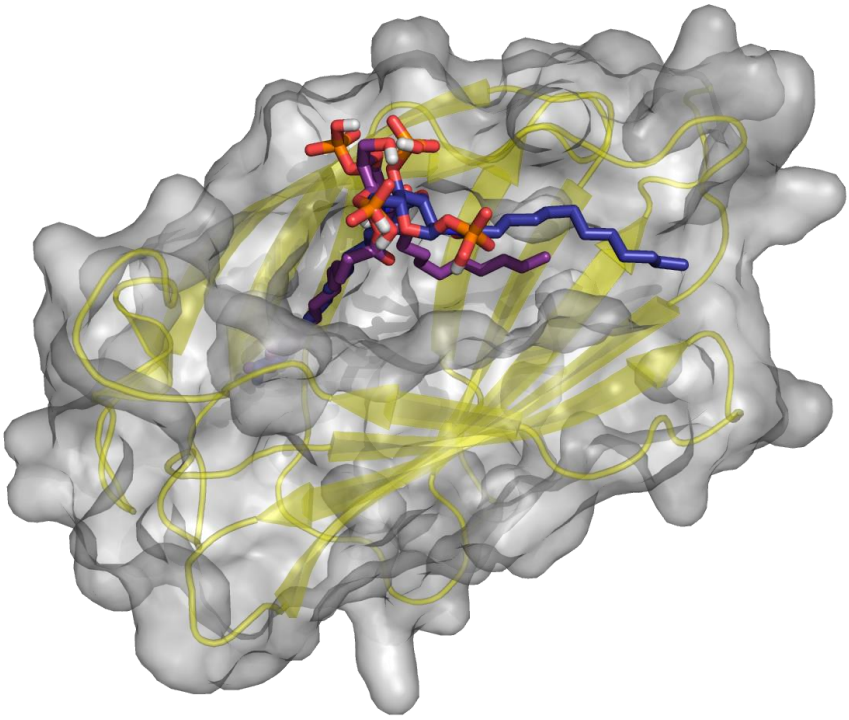


Figure 34. Superposition of the AutoDock binding poses of compound **3** characterized for having two (violet) or one (blue) FA chain oriented inside the lipophilic MD-2 pocket.

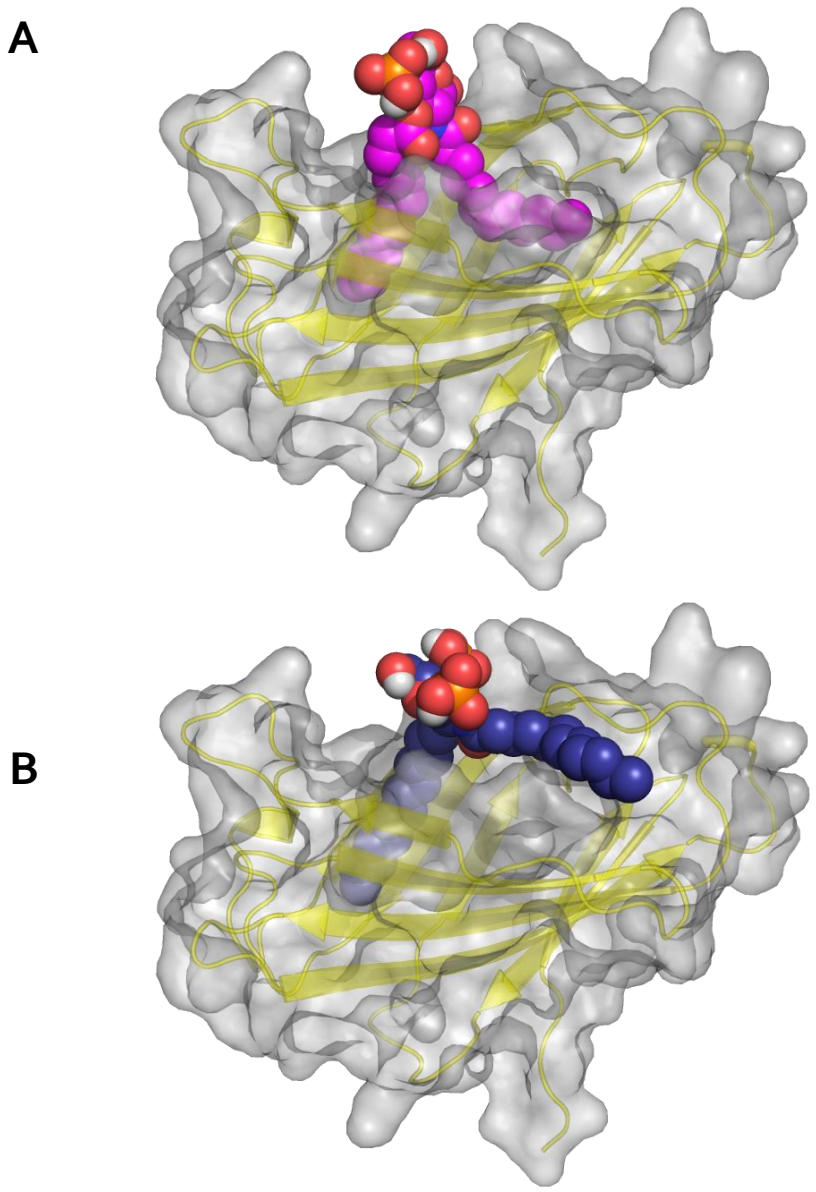


Figure 35. AutoDock binding poses of compound **3** characterized for having two (A) or one (B) FA chains oriented inside the lipophilic MD-2 pocket. In figure B, one of the FA chains is placed in an "hydrophobic tunnel" pointing outward. This is a position similar to that assumed by one acyl chain of LPS (both in human and murine receptors) and Lipid IV a (in murine MD-2 only) that are responsible for TLR4 dimerization.

Polar interactions were also identified in some of the predicted binding poses (not in all of them). In the docked poses with one of the alkyl chain inside the channel, AutoDock calculations led to the observation of two hydrogen bonds involving one of the phosphate groups (oxygen atom and the NH group from Lys122, and phosphate OH group with Ser120 CO). The second phosphate is in close vicinity to OH group from Tyr102, establishing an electrostatic interaction. In the case of the Vina docked poses, the hydrogen bond between one phosphate group and the Lys122 NH group is also identified (second phosphate group is exposed to the outside). For the second predicted binding mode, characterized by both FA chains inside the lipophilic pocket, AutoDock docked poses did not exhibit remarkable polar interactions, while Vina led to few poses with one phosphate group in the vicinity of the amide group from Gln62 side chain, indicating a possible hydrogen bond. These predicted binding poses are in agreement with the NMR experiments and provide a 3D model for the interaction of the FA chains with MD-2 protein, as well as putative polar interactions involving the phosphate groups. In the case of CD14, this protein also owns a highly lipophilic wide pocket as MD-2 protein, but with fewer charged residues in the opening portion. Compound **3** showed binding poses presenting both FA chains inside the pocket, and the polar phosphate groups and sugar placed at the outer region.

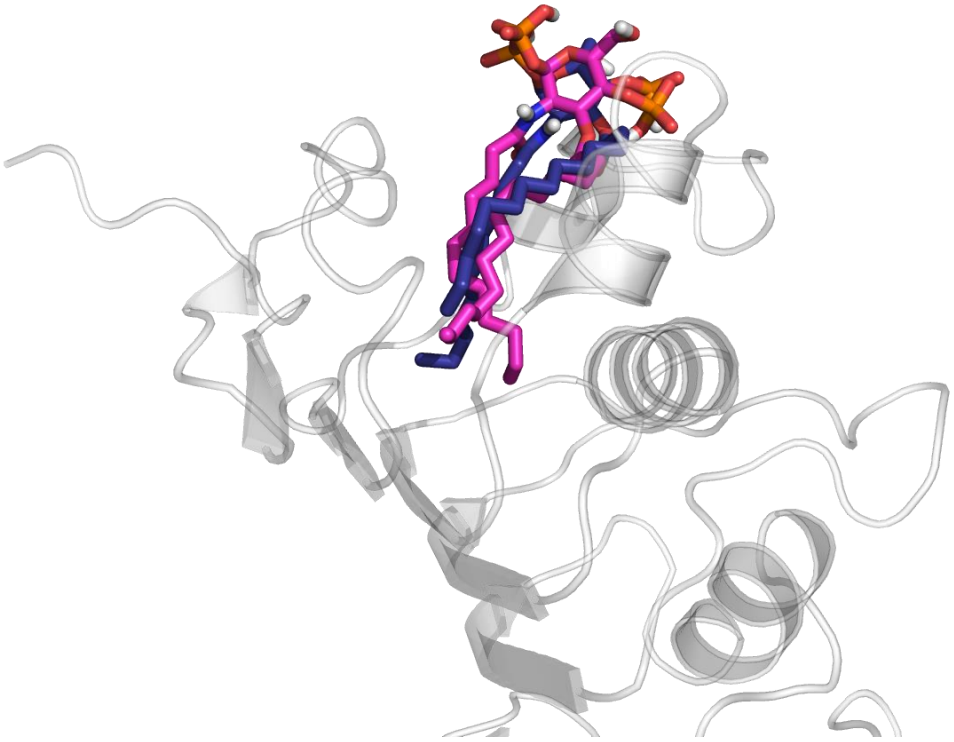


Figure 36. Superimposition of docked binding solutions of compound **3** in CD14 from AutoDock 4.2 (depicted in blue) and AutoDock Vina (depicted in pink).

Conclusions

With the aim to obtain TLR4 antagonists with a chemical structure simpler than lipid A, mono- and diphosphate lipid X analogues were synthesized (compounds **1**-**3**). Monosaccharides **1** and **2** with one phosphate group linked respectively to anomeric (C-1) and C-4 positions, showed some biological activity in HEK cells expressing endotoxin receptors TLR4, MD-2 and CD14, and compound **2** was weakly active as antagonist in murine macrophages. Monosaccharide **3**, a diphosphate lipid X analogue lacking the C-3 hydroxyls on FA acyl chains, showed to be the most active compound in inhibiting in a dose-dependent way the LPS-stimulated NF-κB activation in HEK-Blue™ cells and LPS-induced TNF-α production in macrophages (IC_{50} in the high nM range). This results indicates that the phosphate number and position could influence the biological activity of the molecules. The presence or absence of the (R)-3-hydroxyl groups on the fatty acid chains of Lipid X mimetics seems to be not relevant for activity, as neither is the presence of an ester or an amide linkage in C-3 position. This can be easily derived from a comparison between the IC_{50} of compound SZD-880.431, previously reported in literature (66) as 0.12 μM against 10 ng/ml LPS, and the IC_{50} of Compound **3**, that is 0.15 μM against the same concentration of LPS. This is a clear indication that the structural differences between the two compound are negligible when it comes to the biological activity as endotoxin antagonists. Kinetic parameters and experiments at different LPS concentrations indicate a competitive antagonism mechanism for compound **3** towards LPS for the binding to endotoxin-sensitive cells. Monosaccharide **3** was much more active in inhibiting cytokine production at low (10 ng/mL) rather than at high (1 μg/mL) LPS concentrations. This could be due mainly to the increase in agonist concentration that results in a right-shift of the inhibition curve, as expected

in competitive antagonism model. It is known that in presence of low S-LPS concentrations, the CD14-catalyzed extraction of endotoxin monomers from aggregates in solution and their presentation to MD-2·TLR4 complex is essential to TLR4 activation and signaling. (74) On the other hand, it has been reported that high concentrations of S-LPS can bind directly to the MD-2·TLR4 complex and activate the TLR4 cascade without need of the CD14 chaperone. The higher activity of compound **3** in the experimental conditions in which CD14 contributes to TLR4 activation, is a first indication then compound **3** probably interferes in both LPS/MD-2·TLR4 and LPS/CD14 interactions. This behavior recalls that of a first generation of positively charged monosaccharides developed in our laboratory that blocked *in vitro* and *in vivo* TLR4 activation by mainly displacing endotoxin from CD14 interaction. (63) The ability of molecule **3** to stimulate CD14 internalization (but not TLR4·MD-2 complex) provide further evidence of a direct interaction between the synthetic molecule and CD14. The interaction with MD-2 was characterized by NMR studies, allowing the identification of the FA acyl chain moieties as the part of compound **3** that directly interacts with MD-2. Once again in analogy with what observed with synthetic cationic amphiphiles, (63) the FA acyl chains of the molecule show to interact with MD-2 more strongly than the sugar part.

Monosaccharide **3** is a promising lead for the development of drugs targeting TLR4 activation in a variety of medical settings. (79) (80) (81) (82) (83) This non-toxic TLR4 antagonist shows a high water solubility (completely soluble up to 1 mM concentration) in contrast with lipid A and lipid A mimetics previously developed by us and other groups that suffer from poor solubility in aqueous media. (83) This is probably due to the high phosphate/hydrophobic

chains ratio, which favors hydrophilicity and water-solubility. Thus, it is quite strange that in more than 20 years of medicinal chemistry research on CD14/TLR4-modulating molecules and Lipid A analogues, the diphosphate monosaccharide scaffold has been exploited in so little extent. Water-solubility is an important prerequisite for good bioavailability and favorable pharmacokinetic (distribution) properties. According to the data presented in this thesis, compound **3** interacts efficiently with both MD-2 and CD14 receptors. Multiple targeting could explain and contribute to molecule's efficacy in blocking TLR4 signal.

In particular, MD-2 binding could result in an antagonism since Compound **3** is probably unable to induce the dimerization of TLR4·MD-2 complex and the consequent activation of the signal pathways, while it keeps the hydrophobic pocket occupied and not available for LPS binding. Compound **3** is not able to induce receptor dimerization probably because it is too small compared to Lipid A and/or it has only two FA chains, which are not sufficient to create a hydrophobic dimerization interface between a MD-2 molecule (associated to a TLR4 receptor) and another TLR4.

CD14 binding, according to our current hypothesis, results instead in some kind of receptor activation (a sort of CD14 "agonism"), inducing its internalization and a decrease in the number of CD14 molecules available on the cell surface for LPS binding and presentation to TLR4·MD-2. The precise molecular mechanism through which this GPI-anchored protein triggers a signal pathway activation is still unknown, although some of the intracellular proteins associated to CD14 agonism have been described. (42)

CD14 has been described as one of the macrophage receptors involved in apoptotic cell recognition, although the mechanism of its action and the binding partners are not fully understood. It would appear, however, that human monocyte-derived macrophages, which are known to clear apoptotic cells via a CD14 component (67) are signalled by apoptotic cells to activate anti-inflammatory pathways that are dominant over pro-inflammatory signals, including LPS. (68) It might be predicted, therefore, that anti-inflammatory CD14-generated signals are dominant over pro-inflammatory CD14-generated signals in macrophages. (69) This is another possible explanation of the CD14-mediated activity of Compound **3** in inhibiting LPS response in cells. Further work will be performed in order to identify the intracellular mechanisms associated to CD14 activation by Compound **3** (**Figure 37**).

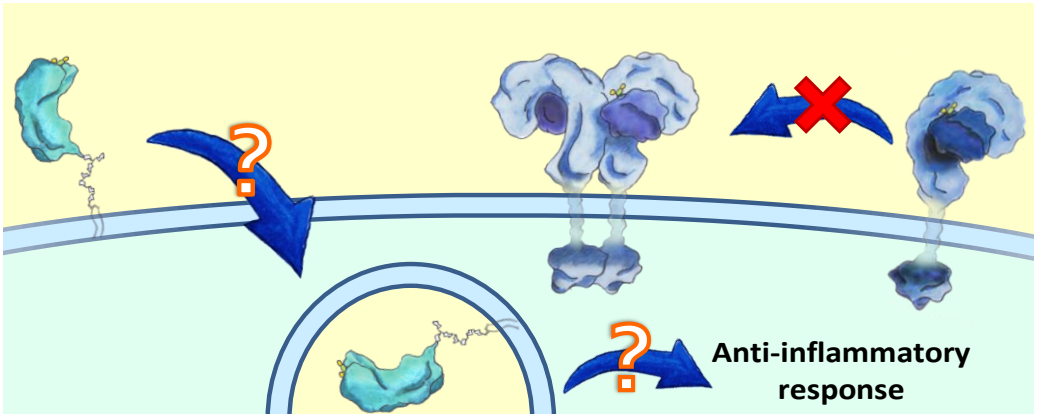


Figure 37. Proposed model for Compound **3** activity on CD14 and MD-2 receptors. For description, see text. Self work.

Moreover, a derivatization of Compound **3** (the most active, i.e. that with the lowest IC_{50} value against LPS) in C-6 position has been planned and it is currently in progress. In particular, the esterification of the hydroxyl group in

C-6 position with a *N*-Fmoc protected amino acid moiety (*N*-Fmoc 6-aminohexanoic acid) led to the incorporation of a C_6 linker with a easily-functionalizable terminal amine. This linker is located in the region of the molecule that (based on molecular docking simulations and NMR data) is supposed to interact less with the endotoxin receptors, thus being the most prone to modifications that are expected not to interfere with molecule binding to the target proteins. This could lead, after a few synthetic steps, to the production of a fluorescent derivative of Compound **3** that could be useful for imaging of the innate immunity endotoxin receptors through fluorescence, confocal or high-resolution microscopy, and in other studies (e.g. flow cytometry). Also, the introduction on C-6 position of a paramagnetic probe (a Gadolinium-chelating agent) for MRI is planned to start in the first months of 2014, starting from the same common intermediate as the fluorescent derivatives.

Experimental section

Chemistry

General procedures. All the reagents were commercially available and used without further purification unless indicated otherwise. All solvents were anhydrous grade unless indicated otherwise. When dry conditions were required, the reactions were carried out in oven-dried glassware under a slight pressure of argon. Reaction were magnetically stirred and monitored by thin-layer chromatography (TLC) on silica gel. TLC was performed on Silica Gel 60 F254 plates (Merck) with UV detection, or using a molibdate developing solution [5% aq H_2SO_4 with 4% $(NH_4)_6Mo_4O_7.4H_2O$ and 0.2% $Ce(SO_4)_2$] followed by heating at 120°C. Flash column chromatography was performed on silica gel 230–400 mesh (Merck). The petroleum ether used as eluent in chromatography has boiling range of 40–60°C. ^1H and ^{13}C-NMR spectra were recorded on a Varian 400 MHz Mercury instrument at 300 K. Chemical shifts are reported in ppm downfield from TMS as internal standard. Mass spectra were recorded on ESI-MS triple quadrupole (model API2000 QTrap™, Applied Biosystems).

Phosphoryl 2-deoxy-3-O-tetradecanoyl-2-tetradecanoylamino-α-D-glucopyranoside (1). Compound **8** (0.05 g, 0.05 mmol) was dissolved in dry CH_2Cl_2/MeOH 1:2 (6 mL), then Pd on activated charcoal (catalytic amount) and AcOH were added. The reaction was stirred at r.t. under H_2 atmosphere overnight, monitoring the disappearance of starting material by TLC (Toluene/AcOEt 7:3). Triethylamine (100 µL) was then added to the reaction mixture, and the suspension was filtered on a syringe filter in order to remove Pd/C catalyst, and washed with CH_2Cl_2. The product was then passed on a Amberlite IR 120 Na^+ exchange resin in order to remove triethylamine and

give compound **1** as sodium salt (0.04 g, quantitative). ^{1}H NMR (400 MHz, CD_3OD + 3% D_2O, 25 °C, TMS): δ = 5.44 (dd, ^{3}J(H,P)= 6.8 Hz, ^{3}J(H,H)= 3.4 Hz, 1H; H-1), 5.20 (dd, ^{3}J(H,H)= 10.6, 9.4 Hz, 1H; H-3), 4.15 (dt, ^{3}J(H,H)= 5.7, 3.2, 1H; H-2), 3.95 (m, 1H; H-5), 3.85 (dd, ^{3}J(H,H)= 11.8, 2.0 Hz, 1H; H-6a), 3.72 (dd, ^{3}J(H,H)= 11.8, 5.2 Hz, 1H; H-6b), 3.63–3.54 (m, 1H; H-4), 2.42–2.27 (m, 2H; CH_2α chains), 2.27–2.14 (m, 2H; CH_2α chains), 1.56 (m, 4H; CH_2β chains), 1.37–1.20 (m, 40H; CH_2 bulk), 0.89 (t, ^{3}J(H,H)= 6.8 Hz, 6H; CH_3). ^{13}C NMR (101 MHz, CD_3OD + 3% D_2O, 25 °C, TMS): δ = 175.23, 174.18, 93.77, 73.33, 72.91, 68.06, 60.60, 51.83, 35.80, 33.81, 33.31, 31.67, 29.47, 29.42, 29.38, 29.29, 29.19, 29.14, 29.08, 28.95, 25.69, 25.31, 24.60, 24.56, 22.35, 13.16, 7.88. ESI MS calc. 679.44, found 678.6, 339.4.

2-deoxy-4-O-phosphoryl-3-O-tetradecanoyl-2-tetradecanoylamino-α,β-D-glucopyranose (2). Compound **11** (0.14 g, 0.15 mmol) was dissolved in dry CH_2Cl_2/MeOH 1:2 (6 mL), then Pd on activated charcoal was added in catalytic amount. The reaction was stirred at r.t. under H_2 overnight, monitoring the disappearance of starting material by TLC (ETP/AcOEt 7:3). Triethylamine (100 μL) was added to the reaction mixture, and the suspension was filtered on a syringe filter in order to remove Pd/C catalyst. The pure triethylammonium salt was then passed on a Amberlite IR 120 Na^+ exchange resin in order to remove triethylamine and give compound **2** as sodium salt (mixture of α- and β- anomers, 5:1), in quantitative yield. ^{1}H NMR (400 MHz, CD_3OD+2.5%$CDCl_3$, 25 °C, TMS): δ = 7.66 (d, ^{3}J(H,H)= 9.6 Hz, 1H; NH), 5.34 (t, ^{3}J(H,H)= 9.8 Hz, 1H; H-3α), 5.14 (t, ^{3}J(H,H)= 9.7 Hz, 1H; H-3β), 5.05 (d, ^{3}J(H,H)= 3.6 Hz, 1H; H-1α), 4.68 (d, ^{3}J(H,H)= 8.4 Hz, 1H; H-1β), 4.26–4.10 (m, 3H; H-2α, H-4α, H-4β), 3.98-3.90 (m, 3H; H-5α, H-6aα, H-6aβ), 3.85–3.75 (m, 2H; H-2b; H-6bβ), 3.68 (m, 1H; H-6bα), 3.40 (m, 1H; H-5β), 2.44-2.11 (m, 8H; CH_2α

chains), 1.56 (m, 8H; $CH_2\beta$ chains), 1.45–1.17 (m, 80 H; CH_2 bulk), 0.89 (t, ^3J(H,H)=6.7 Hz, 12H; CH_3). ^{13}C NMR (101 MHz, CD_3OD+2.5%$CDCl_3$, 25 °C, TMS): δ = 78.97, 178.15, 99.68, 95.74, 75.93, 75.44, 75.16, 64.65, 59.48, 56.55, 50.39, 46.21, 40.31, 39.87, 37.99, 35.90, 33.69, 33.63, 33.59, 33.55, 33.52, 33.33, 33.25, 29.98, 28.69, 26.55, 17.27, 11.93. ESI MS calc. 679.44, found 678.52.

Phosphoryl 2-deoxy-4-O-phosphoryl-3-O-tetradecanoyl-2-tetradecanoylamino-α-D-glucopyranoside (3) Compound **13** (0.08 g, 0.06 mmol) was dissolved in dry CH_2Cl_2/MeOH 1:2 (5 mL), then Pd on activated charcoal was added in catalytic amount. The reaction was stirred at r.t. under H_2 atmosphere overnight (TLC: AcOEt). Triethylamine (150 µL) was then added to the reaction mixture and the suspension was filtered on a syringe filter. The triethylammonium salt was then passed on a Amberlite IR 120 Na^+ exchange resin in order to remove triethylamine and give compound **3** as sodium salt in quantitative yield (61 mg). ^{1}H NMR (400 MHz, CD_3OD, 25 °C, TMS): δ =5.45 (dd, ^3J(H,P)= 6.4, ^3J(H,H)= 3.5 Hz, 1H; H-1), 5.37 (t, ^3J(H,H)= 9.9 Hz, 1H; H-3), 4.31 – 4.18 (m, 2H; H-2, H-4), 4.02 – 3.95 (m, 2H; H-5, H-6a), 3.72 (d, ^2J(H,H)= 12.6 Hz, 1H; H-6b), 2.46 – 2.14 (m, 4H; $CH_2\alpha$ chains), 1.56 (m, 4H; $CH_2\beta$ chains), 1.39 – 1.20 (m, 40H, CH_2 chains), 0.89 (t, ^3J(H,H)= 6.6 Hz, 6H; CH_3 chains). ^{13}C NMR (101 MHz, CD_3OD, 25 °C, TMS): δ = 174.60, 173.29, 94.11, 72.22, 71.89, 70.86, 60.22, 53.41, 52.16, 46.14, 35.87, 33.77, 31.72, 29.51, 29.45, 29.42, 29.37, 29.35, 29.32, 29.28, 29.15, 29.09, 25.71, 24.87, 24.49, 22.37, 13.09, 7.74. ESI MS calc. 759.41, found 758.5, 378.7.

***tert*-butyldimethylsilyl 2-amino-2-deoxy-4,6-*O*-(4-methoxybenzylidene)-β-D-glucopyranoside (5) .** Compound **4** (1 g, 2.28 mmol) was dissolved in a mixture of THF (25 mL) and H_2O (1.65 mL, 91.42 mmol), then PPh_3 (1.8 g, 6.86

mmol) was added and the reaction was stirred at 60° C for 2 h, monitoring the disappearance of starting material by TLC (CH_2Cl_2/MeOH 9.8:0.2). The solvents were evaporated *in vacuo* and the crude product was purified with flash chromatography (CH_2Cl_2/MeOH from 9.9:0.1 to 9:1) to give compound **5** (0.84 g, 87%). ^{1}H NMR (400 MHz, $CDCl_3$, 25 °C, TMS): δ = 7.40 (d, ^3J(H,H)= 8.7 Hz, 2H; PMP), 6.86 (d, ^3J(H,H)= 8.8 Hz, 2H; PMP), 5.45 (s, 1H; C*H*-PMP), 4.49 (d, ^3J(H,H)= 7.6 Hz, 1H; H-1), 4.22 (dd, ^3J(H,H)= 10.4, ^2J(H,H)= 4.9 Hz, 1H; H-6a), 3.77 (m, 4H; OCH_3), 3.73 (t, ^2J(H,H)= 10.4 Hz, 1H; H-6b), 3.57–3.45 (m, 2H; H-3, H-4), 3.45–3.32 (m, 1H; H-5), 2.71 (t, ^3J(H,H)= 8.4 Hz, 1H; H-2), 0.90 (s, 9H; Si-*t*Bu), 0.11 (s, 6H; $(CH_3)_2$-Si). ^{13}C NMR (101 MHz, $CDCl_3$, 25 °C, TMS): δ = 160.16, 129.72, 127.64, 113.66, 101.84, 99.36, 81.43, 72.81, 68.72, 66.59, 59.80, 55.29, 25.72, 17.96, -4.04, -5.16.

tert-butyldimethylsilyl 2-deoxy-4,6-*O*-(4-methoxybenzylidene)-3-*O*-tetradecanoyl-2-tetradecanoylamino-β-D-glucopyranoside (**6**). Myristic acid (2.70 g, 11.82 mmol) was dissolved in dry CH_2Cl_2 under Ar atmosphere and EDC (3.40 g, 17.73 mmol) was added. After 30 min compound **5** (0.84 g, 1.97 mmol) and DMAP (0.36 g, 2.95 mmol) were added and the mixture was stirred at room temperature monitoring the disappearance of starting material by TLC (ETP/AcOEt 9:1). The reaction mixture was diluted with CH_2Cl_2 and washed with saturated $NaHCO_3$ solution and brine. The organic layer was dried over anhydrous Na_2SO_4, filtered and the solvents were evaporated *in vacuo*. The crude product was purified with flash chromatography (ETP/AcOEt 9:1) to give compound **6** (1.6 g, 97% yield). ^{1}H NMR (400 MHz, $CDCl_3$, 25 °C, TMS): δ = 7.30 (d, ^3J(H,H)= 8.5 Hz, 2H; PMP), 6.80 (d, ^3J(H,H)= 8.5 Hz, 2H; PMP), 5.53 (d, ^3J(H,H)= 9.6 Hz, 1H; H-1), 5.40 (s, 1H; C*H*-PMP), 5.12 (t, ^3J(H,H)= 9.9 Hz, 1H; H-3), 4.65 (d, ^3J(H,H)= 8.0 Hz, 1H;

NH), 4.19 (dd, ^{3}J(H,H)=10.5, 4.9 Hz, 1H; H-6a), 4.02 (m, 1H; H-2), 3.76–3.59 (m, 5H; OCH_3, H-4, H-6b), 3.41 (m, 1H; H-5), 2.25 (m, 2H; CH_2α chains), 2.13–1.92 (m, 2H; CH_2α chains), 1.54–1.45 (m, 4H; CH_2β chains), 1.07 (m, 40H; CH_2 chains), 0.83–0.80 (m, 15H; CH_3 chains, Si-*t*Bu), 0.03 (s, 3H; $(CH_3)_2$-Si), 0.00 (s, 3H; $(CH_3)_2$-Si). ^{13}C NMR (101 MHz, $CDCl_3$, 25 °C, TMS): δ =174.57, 172.89, 160.26, 129.76, 127.63, 113.74, 101.51, 97.60, 78.89, 71.93, 68.84, 66.83, 56.49, 55.48, 37.19, 34.51, 32.17, 29.93, 29.77, 29.72, 29.68, 29.66, 29.62, 29.57, 29.29, 25.83, 25.76, 25.27, 22.94, 18.07, 14.38, -3.87, -5.09.

2-deoxy-4,6-*O*-(4-methoxybenzylidene)-3-*O*-tetradecanoyl-2-tetradecanoylamino-α,β-D-glucopyranose (7). Compound **6** (1.48 g, 1.78 mmol) was dissolved in dry THF (40 mL) under Ar atmosphere, cooled to -15 °C and a solution of TBAF (0.62 g, 1.96 mmol) and AcOH (128 µl, 2.24 mmol) in THF (20 mL) was added. The reaction was stirred at -15° C for 1 h, then at 0 °C for 1 h and at r.t. until complete consumption of starting material (TLC: ETP/AcOEt 6:4). The solution was diluted with water and extracted with CH_2Cl_2. The organic layer was dried over anhydrous Na_2SO_4, filtered and the solvents were evaporated *in vacuo*. The crude product was purified with flash chromatography (ETP/AcOEt 8:2 to 6:4) to give compound **7** (0.98 g, 76%). ^{1}H NMR (400 MHz, $CDCl_3$, 25 °C, TMS): δ = 7.36 (d, ^{3}J(H,H)= 8.6 Hz, 2H; PMP), 6.86 (d, ^{3}J(H,H)= 8.6 Hz, 2H; PMP), 5.91 (d, ^{3}J(H,H)= 9.1 Hz, 1H; H-1), 5.49 (s, 1H; C*H*-PMP), 5.38 (t, ^{3}J(H,H)= 10.0 Hz, 1H; H-3), 5.27 (s, 1H; NH), 4.36–4.20 (m, 2H; H-2, H-6a), 4.12 (m, 1H; H-5), 3.81–3.66 (m, 5H; OCH_3, H-4, H-6b), 3.11 (s, 1H; OH), 2.31 (m, 2H; CH_2α chains), 2.14 (m, 2H; CH_2α chains), 1.58 (m, 4H; CH_2β chains), 1.25 (m, 40H; CH_2 chains), 0.88 (t, ^{3}J(H,H)= 6.7 Hz, 6H; CH_3 chains). ^{13}C NMR (101 MHz, $CDCl_3$, 25 °C, TMS): δ =177.11, 175.69, 174.33, 173.47, 160.02, 129.50, 129.24, 127.43, 113.49, 101.48, 92.45, 79.13, 69.50,

68.86, 62.81, 58.09, 55.31, 55.20, 52.80, 36.73, 34.29, 31.93, 29.72, 29.68, 29.55, 29.49, 29.45, 29.38, 29.33, 29.05, 25.64, 25.00, 22.69, 14.13.

Dibenzylphosphoryl 2-deoxy-4,6-*O*-(4-methoxybenzylidene)-3-*O*-tetradecanoyl-2-tetradecanoylamino-α-D-glucopyranoside (8). Compound **7** (0.08 g, 0.11) and tetrabenzyl diphosphate (0.24 g, 0.45 mmol) were dissolved in dry THF (10 mL) under Ar athmosphere, cooled to -78 °C, then $LiN(TMS)_2$ 1 M in THF (88 µl, 0.45 mmol) was added dropwise. The reaction was stirred at -78 °C for 1 h, then at -20 °C for 1 h (TLC: toluene/AcOEt 7:3). The reaction mixture was then quenched with a saturated solution of $NaHCO_3$ and extracted with AcOEt. The organic layer was washed with brine, dried over anhydrous Na_2SO_4, filtered and the solvents were evaporated *in vacuo*. The crude product was purified with flash chromatography (CH_2Cl_2/MeOH 8:2), to give compound **8** (0.047 g, 43%). ^{1}H NMR (400 MHz, $CDCl_3$, 25 °C, TMS): δ = 7.47 – 7.22 (m, 12H; Bn, PMP), 6.88 (d, ^3J(H,H)= 8.5 Hz, 2H; PMP), 5.69 (m, 1H; H-1), 5.63 (d, ^3J(H,H)= 8.9 Hz, 1H; NH), 5.45 (s, 1H; C*H*PMP), 5.25 (t, ^3J(H,H)= 10.1 Hz, 1H; H-3), 5.14 – 4.81 (m, 4H; CH_2PMB,), 4.37 (m, 1H; H-2), 4.07–4.02 (m, 1H; H-6a), 3.93 (m, 1H; H-5), 3.80 (s, 3H; OMe), 3.70 (m, 2H; H-4, H-6b), 2.27 (m, 2H; CH_2α chains), 1.94-1.81 (m, 2H; CH_2α chains), 1.54 (m, 2H; CH_2β chains), 1.43(m, 2H; CH_2β chains), 1.25 (m, 40H; CH_2 chains), 0.88 (t, ^3J(H,H)= 6.8 Hz, 6H; CH_3 chains). ^{13}C NMR (101 MHz, $CDCl_3$, 25 °C, TMS): δ = 175.38, 174.42, 159.20, 137.70, 137.81, 129.84, 128.60, 128.78, 128.41, 128.40, 128.18, 128.01, 113.56, 102.06, 94.92, 78.49, 71.31, 70.98, 67.93, 64.38, 56.04, 54.23, 36.73, 34.29, 31.93, 29.72, 29.66, 29.55, 29.52, 29.45, 29.38, 29.33, 29.05, 25.64, 25.00, 22.80, 14.15.

tert-butyldimethylsilyl 2-deoxy-6-*O*-(4-methoxybenzyl)-3-*O*-tetradecanoyl-2-tetradecanoyl-amino-β-D-glucopyranoside (9).

Compound **6** (1.02 g, 1.22 mmol) was dissolved in dry THF (50 mL) under Ar atmosphere and 4Å molecular sieves (3.7 g) and $NaCNBH_3$ (1.15 g, 18.38 mmol) were added. The reaction was stirred at r.t. for 3h, monitored with TLC (ETP/AcOEt 8:2). The reaction mixture was then cooled in an ice bath and HCl (4M in dioxane) was added dropwise to reach pH=2. The solution was then neutralized by adding saturated $NaHCO_3$ solution, diluted with CH_2Cl_2 and extracted, washing the organic solution with brine. The organic layer was dried over anhydrous Na_2SO_4, filtered and the solvents were evaporated *in vacuo*. The crude product was purified with flash chromatography (ETP/AcOEt 8.5:1.5) to give compound **9** (0.85 g, 84%). ^{1}H NMR (400 MHz, $CDCl_3$, 25 °C, TMS): δ = 7.23 (d, ^3J(H,H)= 8.6 Hz, 2H; PMB), 6.88 (t, J = 8.7 Hz, 2H: PMB), 5.60 (d, ^3J(H,H)= 9.3 Hz, 1H; H-1), 5.01 (dd, ^3J(H,H)= 10.7, 9.2 Hz, 1H; H-3), 4.67 (d, ^3J(H,H)= 8.0 Hz, 1H; NH), 4.56–4.41 (m, 2H; CH_2PMB), 3.93 (dd, ^3J(H,H)= 18.8, 9.5 Hz, 1H; H-2), 3.80 (s, 3H; OMe), 3.77-3.66 (m, 3H; H-4, H-5, H-6a), 3.52 (m, 1H; H-6b), 3.07 (d, ^3J(H,H)= 2.7 Hz, 1H; OH), 2.35–2.30 (m, 2H; CH_2α chains), 2.12–2.05 (m, 2H; CH_2α chains), 1.56 (m, 4H; CH_2β chains), 1.24 (m, 40H; CH_2 chains), 0.95–0.80 (m, 15H; Si-*t*Bu, CH_3 chains), 0.09 (s, 3H; $(CH_3)_2$Si), 0.06 (s, 3H; $(CH_3)_2$Si). ^{13}C NMR (101 MHz, $CDCl_3$, 25 °C, TMS): δ = 174.87, 172.72, 159.32, 129.34, 113.84, 96.61, 73.66, 73.46, 71.30, 70.61, 55.22, 36.99, 34.29, 31.92, 29.71, 29.66, 29.51, 29.43, 29.40, 29.37, 29.32, 29.13, 25.67, 25.56, 24.96, 22.69, 17.89, 17.88, 14.12, -4.07.

***tert*-butyldimethylsilyl 2-deoxy-4-*O*-dibenzylphosphoryl-6-*O*-(4-methoxybenzyl)-3-*O*-tetradecanoyl-2-tetradecanoylamino-β-D-**

glucopyranoside (10). Compound **9** (0.83 g, 0.99 mmol) was dissolved in dry CH_2Cl_2 (20 mL) under Ar atmosphere, then imidazolium tosylate (0.43 g, 1.78 mmol) and dibenzyl *N*,*N*-diisopropyl phosphoramidite (524 µl, 1.59 mmol) were added and the reaction was stirred at r.t. for 2 h (TLC ETP/AcOEt 8.5:1.5). The solution was then cooled in an ice bath and *m*CPBA (0.34 g, 1.99 mmol) was added. The reaction was stirred at 0 °C for 1 h, then the mixture was diluted with CH_2Cl_2, washed with a saturated solution of $NaHCO_3$ and brine. The organic layer was dried over anhydrous Na_2SO_4, filtered and the solvents were evaporated *in vacuo*. The crude product was purified with flash chromatography (ETP/AcOEt 9.5:0.5) followed by trituration in ETP, to give compound **10** (0.46 g, 42%).^{1}H NMR (400 MHz, $CDCl_3$, 25 °C, TMS): δ = 7.27–7.08 (m, 12H; Bn, PMB), 6.73 (d, ^3J(H,H)= 8.6 Hz, 2H; PMB), 5.23 (d, ^3J(H,H)= 9.3 Hz, 1H; H-1), 5.14 (t, ^3J(H,H)= 10.7, 9.2 Hz, 1H; H-3), 4.87–4.77 (m, 4H; C*H*$_2$Ph), 4.66 (d, ^3J(H,H)= 7.9 Hz, 1H; NH), 4.40–4.30 (m, 3H; H-4, C*H*$_2$PMB), 3.81 (dd, ^3J(H,H)= 19.0, 9.2 Hz, 1H; H-2), 3.67 (m, 4H; OMe, H-6a), 3.53 (m, 2H; H-5, H-6b), 2.06 (m, 2H; CH_2α chains), 1.97 (m, 2H; CH_2α chains), 1.44 (m, 2H; CH_2β chains), 1.33 (d m, 2H; CH_2β chains), 1.16 (m, 40H; CH_2 chains), 0.89–0.67 (m, 15H; Si-*t*Bu, CH_3 chains), 0.03 (s, 3H; $(CH_3)_2$-Si), 0.00 (s, 3H; $(CH_3)_2$-Si). ^{13}C NMR (101 MHz, $CDCl_3$, 25 °C, TMS): δ =174.18, 172.57, 159.02, 135.54, 130.23, 129.14, 128.54, 127.92, 127.89, 113.63, 96.36, 74.33, 74.08, 73.03, 72.63, 69.50, 31.93, 31.90, 29.67, 29.65, 29.52, 29.38, 29.32, 29.15, 25.59, 24.66, 22.70, 17.88, 14.13.

2-deoxy-4-*O*-dibenzylphosphoryl-6-*O*-(4-methoxybenzyl)-3-*O*-tetradecanoyl-2-tetradecanoylamino-β-D-glucopyranose (11). Compound **10** (0.63 g, 0.58 mmol) was dissolved in dry THF (10 mL) under Ar atmosphere, cooled to -15 °C, then a solution of TBAF (0.25 g, 0.79 mmol) and AcOH (50

µl, 0.9 mmol) in dry THF (3 mL) was added and the reaction was stirred at 0° C for 2 h, then at r.t. for 2 h (TLC ETP/AcOEt 7:3). The solution was diluted with water and extracted with CH_2Cl_2. The organic layer was dried over anhydrous Na_2SO_4, filtered and the solvents were evaporated *in vacuo*. The crude product was crystallized from CH_2Cl_2 to give compound **11** (0.32 g, 57%).^{1}H NMR (400 MHz, $CDCl_3$, 25 °C, TMS): δ =7.25 (m, 12H; Bn, PMB), 6.80 (d, ^3J(H,H)= 8.4 Hz, 2H; PMB), 5.74 (d, ^3J(H,H)= 9.0 Hz, 1H; H-1), 5.37 (t, ^3J(H,H)= 10.0 Hz, 1H; H-3), 5.22 (bs, 1H; NH), 4.86 (m, 4H; CH_2Ph), 4.48–4.37 (m, 3H; CH_2PMB, H-4), 4.23 (m, 1H; H-2), 4.15 (m, 1H; H-5), 3.70 (m, 4H; OCH_3, H-6a), 3.59 (dd, ^2J(H,H)= 10.5, ^3J(H,H)= 6.3 Hz, 1H; H-6b), 3.45 (s, 1H; OH), 2.09 (m, 4H; CH_2α chains), 1.53 (m, 2H; CH_2β chains), 1.40 (m, 2H; CH_2β chains), 1.24 (s, 40H; CH_2 chains), 0.88 (t, ^3J(H,H)= 6.5 Hz, 6H; CH_3 chains). ^{13}C NMR (101 MHz, $CDCl_3$, 25 °C, TMS): δ =174.68, 173.21, 159.40, 135.72, 129.94, 129.70, 128.81, 128.19, 128.14, 113.93, 91.74, 74.30, 73.29, 70.99, 69.81, 68.57, 68.46, 55.36, 52.28, 36.93, 34.22, 32.15, 29.90, 29.75, 29.60, 29.36, 25.79, 24.82, 22.92, 14.36.

2-deoxy-6-*O*-(4-methoxybenzyl)-3-*O*-tetradecanoyl-2-tetradecanoylamino-β-D-glucopyranose (12). Compound **6** (1.03 g, 1.23 mmol) was dissolved in dry THF (50 mL) under Ar atmosphere and 4Å molecular sieves (3 g) and $NaCNBH_3$ (1.16 g, 18.45 mmol) were added. The mixture was stirred at rt for 3 h, then HCl (1 M in dioxane) was added dropwise until pH=1.5, monitoring the consumption of starting material by TLC (ETP/AcOEt 6:4). The solution was neutralized using $NaHCO_3$ then extracted with CH_2Cl_2, the organic phase was dried over anhydrous Na_2SO_4, filtered and the solvent was evaporated *in vacuo*. The crude product was purified with flash chromatography (ETP/AcOEt/MeOH 7:3:0.05) affording compound **12**

(0.54 g, 61%). ^{1}H NMR (400 MHz, $(CD_3)_2CO$, 25 °C, TMS): δ = 7.27 (d, ^{3}J(H,H)= 8.5 Hz, 2H; PMB), 6.90 (d, ^{3}J(H,H)= 8.6 Hz, 2H; PMB), 6.53 (d, ^{3}J(H,H)= 9.2 Hz, 1H; H-1), 5.88 (d, ^{3}J(H,H)= 3.0 Hz, 1H; OH), 5.18 (dd, ^{3}J(H,H)=10.8, 9.3 Hz, 1H; H-3), 5.08 (t, ^{3}J(H,H)= 3.9 Hz, 1H; NH), 4.56 – 4.40 (m, 3H; CH_2PMB, OH), 4.08 (m, 1H; H-2), 4.02 (m, 1H; H-5), 3.79 (s, 3H; OCH_3), 3.71 (m, 2H; H-6a, H-6b), 3.62 (m, 1H; H-4), 2.27 (m, 2H; CH_2α chains), 2.18 (m, 2H; CH_2α chains), 1.56 (m, 4H; CH_2β chains), 1.29 (s, 40H; CH_2 chains), 0.88 (t, ^{3}J(H,H)=6.5 Hz, 6H; CH_3 chains). ^{13}C NMR (101 MHz, $CDCl_3$, 25 °C, TMS): δ = 175.38, 174.51, 159.15, 132.58, 129.10, 113.46, 94.501, 74.18, 73.72, 72.15, 71.43, 69.15, 56.04, 55.19, 50.40, 46.21, 40.31, 39.88, 37.99, 35.92, 33.71, 33.66, 33.59, 33.52, 33.51, 33.33, 33.26, 29.98, 28.68, 26.55, 17.27, 11.94.

Dibenzylphosphoryl 2-deoxy-4-*O*-dibenzylphosphoryl-6-*O*-(4-methoxybenzyl)-3-*O*-tetradecanoyl-2-tetradecanoylamino-α-D-glucopyranoside (13). Compound **12** (0.30 g, 0.41 mmol) was dissolved in dry CH_2Cl_2 (20 mL) under Ar atmosphere, then imidazolium triflate (0.48 g, 1.87 mmol) and dibenzyl *N*,*N*-diisopropyl phosphoramidite (547 µl, 1.66 mmol) were added and the reaction was stirred at r.t. for 1.5 h (TLC: ETP/AcOEt 7.5:2.5). The solution was then cooled in an ice bath and *m*CPBA (0.50 g, 2.90 mmol) was added. The reaction was stirred at 0 °C for 2 h, then the mixture was diluted with CH_2Cl_2, washed with a saturated solution of $NaHCO_3$ and brine. The organic layer was dried over anhydrous Na_2SO_4, filtered and the solvents were evaporated *in vacuo*. The crude product was purified with flash chromatography (ETP/AcOEt 7:3) affording compound **13** (0.19 g, 79,5%). ^{1}H NMR (400 MHz, $CDCl_3$, 25 °C, TMS): δ = 7.40 – 7.22 (m, 20H; PMB), 7.17 (d, ^{3}J(H,H)= 8.6 Hz, 2H; PMB), 6.76 (d, ^{3}J(H,H)= 8.6 Hz, 2H; PMB), 5.69 (dd, ^{3}J(H,H)= 6.0, 3.3 Hz, 1H; H-1), 5.59 (d, ^{3}J(H,H)= 9.1 Hz, 1H; NH), 5.28 (dd,

^{3}J(H,H)= 10.8, 9.3 Hz, 1H; H-3), 5.07 – 4.85 (m, 8H; CH_2Bn), 4.61 (dd, ^{3}J(H,H)= 18.8, 9.2 Hz, 1H; H-4), 4.35 (m, 3H; CH_2PMB, H-2), 4.00 (d, ^{3}J(H,H)= 9.7 Hz, 1H; H-5), 3.71 (s, 3H; OCH_3), 3.56 (s, 2H; H-6a, H-6b), 2.13 (t, ^{3}J(H,H)= 7.7 Hz, 2H; CH_2α chains), 1.92 – 1.76 (m, 2H; CH_2α chains), 1.40 (m, 4H; CH_2β chains), 1.35 – 1.07 (m, 40H; CH_2 chains), 0.88 (t, ^{3}J(H,H)= 6.6 Hz, 6H; CH_3 chains). ^{13}C NMR (101 MHz, $CDCl_3$, 25 °C, TMS): δ = 174.29, 173.05, 159.07, 135.81, 135.74, 135.52, 135.44, 135.37, 135.33, 135.26, 129.79, 129.34, 128.75, 128.72, 128.66, 128.64, 128.59, 128.55, 128.51, 128.04, 127.99, 127.94, 127.91, 113.55, 96.19, 73.04, 72.83, 71.73, 71.67, 70.14, 69.80, 69.74, 69.69, 69.64, 69.59, 69.55, 69.50, 69.30, 69.24, 69.19, 67.30, 55.12, 51.70, 36.29, 33.94, 31.94, 29.73, 29.69, 29.53, 29.39, 29.36, 29.35, 29.27, 29.13, 25.39, 24.60, 22.71, 14.16.

Biology

HEK-Blue™ assay.

HEK-Blue™ hTLR4 cells (*InvivoGen*) were cultured according to the manufacturer's instructions. Cells were cultured in Dulbecco's modified Eagle's medium (DMEM) high-glucose medium supplemented with fetal bovine serum (FBS, 10 %), glutamine (2 mm), 1× Normocin (*InvivoGen*), and 1× HEK-Blue™ Selection (*InvivoGen*). Cells were detached by use of a cell scraper, and the cell concentration was estimated by use of a cell counter. Cells were then diluted in DMEM high-glucose medium supplemented with FBS (10 %), glutamine (2 mm), 1× Normocin (*InvivoGen*), and cell suspension (200 mL, 20000 cells) was added to each well. After overnight incubation (37 °C, 5% CO_2, 95% humidity), cells reached 80% confluency. Supernatant was removed, and cell monolayers were washed with warm PBS without Ca^{2+} and

Mg^{2+}, preincubated for 30 min in 190 mL DMEM+0.03% BSA, supplemented with compounds **1**, **2**, or **3** (different concentrations in different wells, from 0 to 50 mm). LPS (*E. coli* O55:B5 strain, *Sigma–Aldrich*) at a final concentration of 10, 100, or 1000 $ngmL^{-1}$ was then added as stimulus (10 mL per well), and cells were incubated overnight under the same conditions as above. After the incubation, supernatants were collected. Each sample (50 mL) was added to a pNPP solution in PBS (0.8 mm, 100 mL). Plates were incubated in the dark at room temperature and then analyzed spectrophotometrically (absorbance at 405 nm).

Murine-bone-marrow-derived macrophages (BMMFs) assay.

Murine BMMFs were obtained from wild-type and $CD14^{-/-}$ mice and cultured by the published procedure. (88) Cells were washed with PBS and detached by treatment with EDTA in PBS (2 mm). The cellular suspension was then centrifuged at 188*g* for 5 min, and the cellular pellet was resuspended in DMEM high-glucose+FBS 10% and plated in a 96-well plate (2×10^5 cells per well). After overnight incubation (37 °C, 5% CO_2, 95% humidity), supernatant was removed, and cell monolayers were washed with warm PBS without Ca^{2+} and Mg^{2+}, preincubated for 30 min in 190 mL DMEM+0.03% BSA, and supplemented with compounds **1**, **2**, or **3** (different concentrations in different wells, from 0 to 50 mm). LPS (*E. coli* O55:B5 strain, *Sigma–Aldrich*) at a final concentration of 10, 100, or 1000 $ngmL^{-1}$ was then added as stimulus (10 mL per well), and cells were incubated overnight under the same conditions as above. After the incubation, supernatants were collected, and TNF concentration was detected through an ELISA.

MTT cell viability assay.

100 µL of hTLR4-HEK-BlueTM cells were seeded in DMEM without red phenol into 96-well microplates (approximately 1×10^4 cells per well). After incubation (overnight, 37°C, 5% CO_2), the cells were treated with 50 µL of 3× compound **1**, **2** or **3** (final contentrations: 0-50 µM) and the plates were incubated overnight. Then, 10 µL of MTT solution (5 mg/mL in PBS) were added in each well and after 3 h (37 °C, 5% CO_2) the supernatant was removed (taking care not to remove the crystals) and the formazan crystals were dissolved by DMSO (100 µL). Formazan concentration in the wells was determined measuring the absorbance at 570 nm, subtracting the absorbance at 630 nm. 50 µl DMSO and 50 µl PBS instead of the compounds were used as controls.

Flow cytometry analysis.

BMMFs were washed twice with PBS and detached by treatment with EDTA in PBS (2 mm). The cellular suspension was then centrifuged at 188*g* for 5 min, and the cellular pellet was resuspended in DMEM high glucose+FBS 10% (100 mL) and incubated at 37 °C, 5% CO2, and 95% humidity for 40 or 70 min with compound **1**, **2**, or **3** (10 mm) or LPS (1 $mgmL^{-1}$). Cellular suspensions were then centrifuged at 188*g*, 4 °C, for 5 min, and the cellular pellet was resuspended with ice-cold PBS. The receptor-specific fluorescent antibodies were then added (aCD14: FITC rat anti-mouse CD14, clone SA2–8, *eBIOSCIENCE*; aTLR4: PE rat anti-mouse TLR4, clone SA15–21, *Biolegend*), at 1 $mgmL^{-1}$, for 20 min in ice in the dark. The cells were then washed twice with ice-cold PBS, and fluorescence was analyzed with a FaCSCalibur flow cytometer (*BD Biosciences*).

Production and isolation of MD-2.

Recombinant MD-2 was produced in *E. coli* as described in literature (89) with use of solubilization of inclusion bodies in Gdn·HCl (6 M) followed by purification and refolding on a C8 reversed-phase column. Eluted fractions were lyophilized and dialyzed against Milli-Q water. Biological activity of each batch of MD-2 has been assessed for the ability to support LPS induced TLR4 activation in HEK293 cells transfected with TLR4.

Determination of CMCs through pyrene fluorescence measurements. (75)

For every sample, independently of the synthetic compound's concentration, a final pyrene concentration of 0.6 mm is desired. A mother solution of pyrene in THF (1 mM) was prepared in a 25 mL volumetric flask, and diluted into a solution (237 mm, 6 mL of the concentrated solution, followed by addition of 19 mL of THF, to a final volume of 25 mL). An aliquot of this solution (252 mL) is then diluted with Milli-Q water to a final volume of 100 mL. Aqueous solutions of each compound (0.6 mM in pyrene) were prepared at concentrations from 1 mM to 0.05 mM, by two-fold serial dilutions. The 1 mM solutions were prepared by adding pyrene solution (0.6 mM) to each compound and sonicating for 1 h. The serial dilutions were incubated for 1 h at 37.0 °C. For fluorescence measurements, the solution (1.5 mL) was placed in a conventional 1 cm quartz cuvette; fluorescence spectra were recorded with a Varian Cary Eclipse spectrofluorophotometer at (37±0.1) °C, with use of 5 mm excitation and emission slits. The onset of micelle formation can be observed in a shift of the fluorescence excitation spectra of the samples at an emission wavelength of 372 nm. In the concentration range of aqueous micellar solutions, a shift of the excitation band in the 335 nm region toward higher

wavelengths confirms the incorporation of pyrene in the hydrophobic interiors of micelles. The ratio of the fluorescence intensities at 339 and 335 nm was used to quantify the shift of the broad excitation band. The critical micelle concentrations were determined from the crossover point in the low concentration range.

NMR binding experiments.

NMR experiments were performed with a 600 MHz DRX spectrometer (*Bruker*) equipped with a cryoprobe. ^{1}H NMR spectra were recorded at 298 K, by acquisition of 120 scans. DOSY spectra were recorded at 311 K with the stebpgls19 Bruker pulse sequence by acquisition of 160 scans, with a diffusion time of 300 ms, a gradient length of 1.9 ms, and a gradient ramp from 2% to 95% in 32 linear steps. Protein samples were prepared by diluting the stock solution of MD-2 (0.11 mM in deuterated acetate buffer at pH 5) with the same buffer. Ligand samples were prepared by dissolving solid compound **3** in deuterated acetate buffer at pH 5. For ^{1}H NMR experiments the final concentrations reached for the analysed samples were 30 mM MD-2 and 150 mM compound **3**, whereas for DOSY experiments final concentrations of 60 mM MD-2 and 300 mM compound **3** were needed.

Molecular modelling.

3D coordinates of compound **3** were built with the aid of Maestro (version 9.3, *Schrödinger*, LLC, New York, NY, 2012). Phosphate groups were considered to be monoprotonated. Molecular mechanics optimization (UFF force field), semiempirical calculations (AM1), and DFT (B3LYP/6–31G*) were subsequently applied by use of Gaussian03. (90) Final geometry was

submitted to MD simulations with implicit water and MM3* as force field, by use of Schrödinger Maestro 9.3 Impact 5.8, [Maestro, v. 9.3, *Schrödinger*, LLC, New York, NY, 2012. Suite 2012: Impact v. 5.8, *Schrödinger*, LLC, New York, NY, 2012] and the MM3* force field, dielectric constant 80.0, number of MD steps 100, and time step of 0.001 ps. 3D coordinates of human MD-2 protein (from PDB ID: 2E59), and CD14 protein (from PDB ID: 1WWL) were refined and minimized by use of the Protein Preparation Wizard module of Maestro and the Amber force field. (91) (92) (93) (94) In the case of CD14, only the sequence from Ala3 to Leu130 was considered for docking purposes. Compound **3** was docked into both MD-2 and CD14 proteins with the aid of AutoDock 4.2., (76) and separately with the aid of AutoDock Vina 1.1.2. (77) Predicted binding energies ranged from –2 to –6 kcalmol^{-1} in the AutoDock results and from –6 to –9 kcal mol^{-1} in the AutoDock Vina results. For MD-2 the Autogrid grid point spacing was set at 0.375 Å, center coordinates of the grid box were –0.379, 17.201, 16.216 (*x*, *y*, *z*), and number of grid points in *xyz* was 58, 92, 82. For CD14 the Autogrid grid point spacing was set at 0.375 Å, center coordinates of the grid box were 13.500, 51.000, 56.500 (*x*, *y*, *z*), leading to 66, 72, 88 (*x*, *y*, *z*) grid points. All allowed torsional bonds were considered rotatable. Docking calculations with AutoDock were performed by use of Genetic Algorithm (number of runs 250, number of individuals in population 150). Docking calculations with AutoDock Vina were also performed. Coordinates and dimensions of grid boxes, starting geometries, and general methodology were the same as for AutoDock. 3D structures of the docked complexes were optimized by performing MD simulations with Impact (implicit water, and AMBER* force field).

References

1. *Use of the gram stain in microbiology.* **Beveridge, T.J.** 2001, Biotech. Histochem., Vol. 76, p. 111–118.

2. **CSK Himachal Pradesh Agricultural University, Palampur (India).** *www.hillagric.ernet.in/edu/covas/vph/students/practicalslides.htm.* [Online] [Riportato: 28 October 2010.]

3. **Galanos C., Freudenberg M. A., Lüderitz O., Rietschel E.T. and Westphal O.** Chemical, physicochemical and biological properties of bacterial lipopolysaccharides. [book auth.] Cohen E. *Biomedical applications of the horseshoe crab (Limulidae).* New York : Liss A.R., 1979, pp. 321-332.

4. *Bacterial lipopolysaccharides and innate immunity.* **Alexander C. and Rietschel E.T.** 2001, J. Endotoxin Res., Vol. 7, p. 167-202.

5. **Lüderitz O., Freudenberg M.A., Galanos C., Lehmann V., Rietschel E.T. and Shaw D.H.** Lipopolysaccharides of gram-negative bacteria. [book auth.] Bronner F. and Kleinzeller A. *Current topics in membranes & transport: Membrane lipids of Prokaryotes.* New York : Academic Press, 1982, pp. 79-150.

6. *O-antigen modulates infection-induced pain states.* **Charles N. Rudick et al.** 8, 2012, PLoS One, Vol. 7, p. e41273.

7. *New Associations With Pseudomonas luteola Bacteremia: A Veteran With A History Of Tick Bites And A Trauma Patient With Pneumonia.* **Forest W. Arnold et al.** 2, 2005, The Internet Journal of Infectious Diseases, Vol. 4.

8. *R-form LPS, the master key to the activation of TLR4/MD-2-positive cells.* **Huber M., Kalis C., Keck S., Jiang Z., Georgel P., Du X., Shamel L., Sovath S., Mudd S., Beutler B., Galanos C. and Freudenberg M.A.** 2006, Eur. J. Immunol., Vol. 36, p. 701-711.

9. **Goldsby R.A., Kindt T.J., Kuby J., Osborne B.A.** *Immunology, 5th edition.* USA : WH Freeman, 2002. pp. 1-9,24,69-72.

10. **Janeway C., Travers P., Walport M., Shlomchik M.** *Immunobiology, 5th edition - The Immune System in Health and Disease.* New York : Garland Science, 2001.

11. *Are innate immune signaling pathways in plants and animals conserved?* **F.M., Ausubel.** 2005, Nature Immunology, Vol. 6, p. 973-979.

12. *Intestinal epithelial barrier and mucosal immunity - Innate and acquired plasticity of the intestinal immune system.* **Didierlaurent A., Simonet M. and Sirard J.C.** 12, 2006, Cellular and Molecular Life Sciences, Vol. 62, p. 1285-1287.

13. **Du X et al.** 3, 2000, Eur Cytokine Netw., Vol. 11, p. 362-71.

14. **RJ, Chuang TH and Ulevitch.** 3, 2000, Eur Cytokine Netw, Vol. 11, p. 372-8.

15. **Tabeta K et al.** 10, 2004, Proc Natl Acad Sci U.S.A., Vol. 101, p. 3516-21.

16. *TLR2 is constitutively expressed within the kidney and participates in ischemic renal injury through both MyD88-dependent and -independent pathways.* **A.A. Shigeoka et al.** 10, 2007, J. Immunol., Vol. 178, p. 6252–8.

17. *TRAM is specifically involved in the Toll-like receptor 4-mediated MyD88-independent signaling pathway.* **M. Yamamoto et al.** 11, 2007, Nat. Immunol., Vol. 4, p. 1144–50.

18. *Essential role for TIRAP in activation of the signaling cascade shared by TLR2 and TLR4.* **M. Yamamoto et al.** 6913, 2002, Nature, Vol. 420, p. 324–9.

19. en.wikipedia.org/wiki/Toll-like_receptor. *en.wikipedia.org.* [Online] [Riportato: 15 October 2010.]

20. *The instructive role of dendritic cells on T-cell responses.* **F. Sallusto and A. Lanzavecchia.** Suppl 3, 2002, Arthritis Res., Vol. 4, p. S127–32.

21. *Defective LPS signaling in C3H/HeJ and C57BL/10ScCr mice: mutations in Tlr4 gene.* **Poltorak A. et al.** 1998, Science, Vol. 282, p. 2085-2088.

22. *Identification of Toll-like receptor 4 (Tlr4) as the sole conduit for LPS signal transduction: genetic and evolutionary studies.* **Beutler B., Du X., Poltorak A.** 2001, J Endotoxin Res, Vol. 7, p. 277-280.

23. *TLR4 as the mammalian endotoxin sensor.* **B., Beutler.** 109-120, 2002, Curr Top Microbiol Immunol, Vol. 270.

24. *Toll-like receptors stimulate human neutrophil function.* **Hayashi F., Means T., Luster A.** 2003, Blood, Vol. 102, p. 2660-2669.

25. *Toll-like receptor control of the adaptive immune responses.* **Iwasaki A. and Medzhitov R.** 2004, Nat Immunol , Vol. 5, pp. 987-995.

26. **Beamer et al.** 1997, Science, Vol. 276, p. 1861–1864.

27. **Gioannini T et al.** 2002, J Biol Chem, Vol. 277, p. 47818-47825.

28. **Corradin et al.** 1992, J Leukoc Biol, Vol. 52, p. 363-8.

29. **Dentener et al.** 1993, J Immunol, Vol. 151, p. 4258-4265.

30. **Lamping et al.** 1998 , J Clin Invest , Vol. 101, p. 2065-2071.

31. **Zweigner et al.** 2001, Blood, Vol. 98, p. 3800-3808.

32. **Hamann et al.** 2005, Infect Immun , Vol. 73, p. 193-200.

33. *Cell.* **Kim et al.** 2007, Vol. 130, p. 906-917.

34. **Ohto et al.** 2007, Science, Vol. 316, p. 1632-1634.

35. **Shimazu et al.** 1999, J Exp Med , Vol. 189, p. 1777-1782.

36. **Re and Strominger.** 2002 , J Biol Chem , Vol. 277, p. 23427-23432.

37. **Gioannini T et al.** 2004 , Proc Natl Acad Sci USA , Vol. 101, p. 4186-4191.

38. **Gioannini et al.** 2005, J Endotoxin Res, Vol. 11, p. 117-123.

39. **Visintin et al.** 2005, J Immunol , Vol. 175, p. 6465-6472.

40. **Park et al. .** 2009, Nature, Vol. 458, p. 1191-1195.

41. *Crystal structure of CD14 and its implications for lipopolysaccharide signaling.* **Kim J.I. et al.** 2005, J. Biol. Chem., Vol. 280, p. 11347–11351.

42. *Role of CD14 in host protection against infections and in metabolism regulation.* **Granucci, I. Zanoni and F.** 32, 2013, Front. Cell. Infect. Microbiol., Vol. 3.

43. **I. Zanoni, R. Ostuni, L.R. Marek, S. Barresi, R. Barbalat, G.M. Barton, F. Granucci, and J.C. Kagan.** 4, 2011, Cell, Vol. 147, p. 868-880.

44. **Jiang et al.** 2005, Nat Immunol, Vol. 6, p. 565-570.

45. *Expression patterns of the lipopolysaccharide receptor CD14, and the FCgamma receptors CD16 and CD64 on polymorphonuclear neutrophils: data from patients with severe bacterial infections and lipopolysaccharide-exposed cells.* **Wagner, C.,**

Deppisch, R., Denefleh, B., Hug, F., Andrassy, K. and Hansch, G. M. 2003, Shock, Vol. 19, p. 5-12.

46. http://www.wfpiccs.org/projects/sepsis-initiative/. *wfpiccs.org.* [Online]

47. **Abigail A. Salyers, Brenda A. Wilson, Dixie D. Whitt, Malcolm E. Winkler.** *Molecular Pathogenesis – A molecular approach.* 3rd edition. Washington, DC : Amer Society for Microbiology, 2011. p. 47-50.

48. *The Pathophysiology and Treatment of Sepsis.* **I.E., Hotchkiss R.S. and Karl.** 2003, N. Engl. J. Med., Vol. 348, p. 138.

49. *Drotrecogin alfa (activated recombinant human activated protein C) for the treatment of severe sepsis.* **G., Bernard.** 2003, Crit. Care Med., Vol. 31, p. 85–93.

50. *Statins decrease toll-like receptor 4 expression and downstream signaling in human CD14 monocytes.* **Methe H. et al.** 2005, Arterioscler. Thromb. Vasc. Biol., Vol. 25, p. 1439.

51. www.drugs.com/ clinical_trials/phase-iii-study-eritoran-does-not-meet-primary-endpoint-11082.html. *drugs.com.* [Online]

52. **Michael J. Barratt, Donald E. Frail.** *Drug Repositioning: Bringing New Life to Shelved Assets and Existing Drugs.* s.l. : John Wiley & Sons, 2012. p. 16.

53. *Therapeutic targeting of innate immunity with Toll-like receptor agonists and antagonists.* **Kanzler H., Barrat F., Hessel E., Coffman R.** 2007, Nat Med, Vol. 13, p. 552-559.

54. *Toll-like receptors in neurodegeneration.* **Okun E., Griffioen K., Lathia J., Tang S., Mattson M., Arumugam T.** 2009, Brain Res Rev, Vol. 59, p. 278-292.

55. *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.* **I. Bettoni et al.** 2008, Glia, Vol. 56, p. 1312-1319.

56. **Myers, Ulrich and.** 1995, Pharm Biotechnol , Vol. 6, p. 495-524.

57. **Johnson et al.** 1999, J Med Chem, Vol. 42, p. 4640-4649.

58. **Johnson.** 2008 , Curr Top Med Chem, Vol. 8, p. 64-79.

59. **Stöver et al.** 2004, J. Biol. Chem. , Vol. 279, p. 4440-4449.

60. **Lien et al.** 2001, J Biol Chem , Vol. 276, p. 1873-1880.

61. **Brandenburg et al.** 2004 , Biochemistry, Vol. 43, p. 4039-4046.

62. *Amino and Hydroxylamino Monosaccharides Inhibit Lipid A Stimulated Activation of Human Dendritic Cells and Macrophages.* **Peri F., Granucci F., Costa B., Zanoni I., Marinzi C. and Nicotra F.** 2007, Angew. Chem. Int. Ed., Vol. 46, pp. 3308–3312.

63. **Piazza et al.** 2009, Biochemistry, Vol. 48, p. 12337-1234.

64. *Total synthesis of Escherichia coli lipid A.* **Masahiro Imoto, Hiroyuki Yoshimura, Nobuki Sakaguchi, Shoichi Kusumoto, Tetsuo Shiba.** 12, 1985, Tetrahedron Letters, Vol. 26, p. 1545-1548.

65. *Total synthesis of Escherichia coli lipid A, the endotoxically active principle of cell-surface lipopolysaccharide.* **Masahiro Imoto, Hiroyuki Yoshimura, Tetsuo Shimamoto, Nobuki Sakaguchi, Shoichi Kusumoto, Tetsuo Shiba.** 6, 1987, Bull. Chem. Soc. Jpn., Vol. 60, p. 2205-2214.

66. **R. L. Danner, A. L. Van Dervort, M. E. Doerfler, P. Stuetz, J. E. Parrillo.** 1990, Pharm. Res., Vol. 7, p. 260-263.

67. **K. Funatogawa, M. Matsuura, M. Nakano, M. Kiso, A. Hasegawa.** 1998, Infect. Immun., Vol. 66, p. 5792-5798.

68. **A. L. Van Dervort, M. E. Doerfler, P. Stuetz, R. L. Danner.** 1992, J. Immunol., Vol. 149, p. 359-366.

69. **M. Matsuura, M. Kiso, A. Hasegawa.** 1999, Infect. Immun., Vol. 67, p. 6286-6292.

70. **L. Brade, K. Brandenburg, H. M. Kuhn, S. Kusumoto, I. Macher, E. T. Rietschel, H. Brade.** 1987, Infect. Immun., Vol. 55, p. 2636-2644.

71. **H. M. Kuhn, L. Brade, B. J. Appelmelk, S. Kusumoto, E. T. Rietschel, H. Brade.** 1992, Infect. Immun., Vol. 60, p. 2201-2210.

72. **R. Tamai, Y. Asai, M. Hashimoto, K. Fukase, S. Kusumoto, H. Ishida, M. Kiso, T. Ogawa.** 2003, Immunology, Vol. 110, p. 66-72.

73. **J. Chen, Y. Zhou, C. Chen, W. Xu, B. Yu.** 2008, Carbohydr. Res., Vol. 343, p. 2853-2862.

74. **M. Triantafilou, K. Triantafilou, N. Fernandez.** 2000, Eur. J. Biochem, Vol. 267, p. 2218-2226.

75. **A. M. Hofmann, F. Wurm, H. Frey.** 2011, Macromolecules, Vol. 44, p. 4648-4657.

76. **D. S. Goodsell, G. M. Morris, A. J. Olson.** 1996, J. Mol. Recognit., Vol. 9, p. 1-5.

77. **O. Trott, A. J. Olson.** 2010, J. Comput. Chem., Vol. 31, p. 455-461.

78. **U. Ohto, K. Fukase, K. Miyake, Y. Satow.** 2007, Science, Vol. 316, p. 1632-1634.

79. *Modulation of lipopolysaccharide-induced macrophage gene expression by Rhodobacter sphaeroides lipid A and SDZ 880.431.* **C L Manthey, P Y Perera, N Qureshi, P L Stütz, T A Hamilton, and S N Vogel.** 8, 1993, Infect. Immun., Vol. 61, p. 3518-3526.

80. **Rangel-Frausto, M. S.** 2005, Arch. Med. Res., Vol. 36, p. 672-681.

81. **M. Casula, A. M. Iyer, W. G. Spliet, J. J. Anink, K. Steentjes, M. Sta, D. Troost, E. Aronica.** 2011, Neuroscience, Vol. 179, p. 233-243.

82. **L. Cao, F. Tanga, J. Deleo.** 2009, Neuroscience, Vol. 158, p. 896-903.

83. **Rivest, S.** 2009, Nat. Rev. Immunol., Vol. 9, p. 429-439.

84. **F. Peri, M. Piazza.** 2012, Biotechnol. Adv., Vol. 30, p. 251-260.

85. *Recognition of apoptotic cells by human macrophages: inhibition by a monocyte/macrophage-specific monoclonal antibody.* **Gregory, Pauline K. Flora and Christopher D.** 1994, European Journal of Immunology, Vol. 11, p. 2625–2632.

86. *Macrophages that have ingested apoptotic cells in vitro inhibit proinflammatory cytokine production through autocrine/paracrine mechanisms involving TGF-beta, PGE2, and PAF.* **V A Fadok, D L Bratton, A Konowal, P W Freed, J Y Westcott, P M Henson.** 4, 1998, J. Clin. Invest., Vol. 101, p. 890-898.

87. **Gregory, Christopher D.** Non-Inflammatory/Anti-Inflammatory CD14 Responses: CD14 in Apoptosis. [a cura di] Robert S. Jack. *CD14 in the inflammatory response.* s.l. : Karger Medical and Scientific Publishers, 2000, p. 135.

88. **F. Granucci, C. Vizzardelli, N. Pavelka, S. Feau, M. Persico, E. Virzi, M. Rescigno, G. Moro, P. Ricciardi-Castagnoli.** 2001, Nat. Immunol., Vol. 2, p. 882-888.

89. **M. Mancek-Keber, R. Jerala.** 2006, FASEB J., Vol. 20, p. 1836-1842.

90. Gaussian 03, Gaussian, Inc., Wallingford, CT, USA, 2003. [Online]

91. Suite 2012: Schrödinger Suite 2012 Protein Preparation Wizard. [Online]

92. Epik version 2.3, Schrödinger, LLC, New York, NY, 2012. [Online]

93. Impact version 5.8, Schrödinger, LLC, New York, NY, 2012. [Online]

94. **Prime version 3.1, Schrödinger, LLC, New York, NY, 2012.** [Online]

Publications

Cighetti R, Ciaramelli C, Zanoni I, Calabrese V, Granucci F, Martin-Santamaria S, Jimenez-Barbero J, Peri F., Modulation of CD14 and TLR4/MD-2 activities by a synthetic lipid A mimetic, *ChemBioChem*, 2014, **15** (2), 250-258

Peri F, Calabrese V, Piazza M, Cighetti R, Synthetic molecules and functionalized nanoparticles targeting the LPS-TLR4 signaling: A new generation of immunotherapeutics. *Pure and Applied Chemistry*, 2012, **84**, 97-106

Piazza M, Calabrese V, Damore G, Cighetti R, Gioannini T, Weiss J and Peri F, A Synthetic Lipid A Mimetic Modulates Human TLR4 Activity. *ChemMedChem*, 2012, **7**, 213-217

Peri F, Piazza M, Calabrese V, Cighetti R, Modulation of LPS signaling through TLR4 agonists and antagonists, in: Yuriy A. Knirel & Miguel A. Valvano., *Bacterial lipopolysaccharide - Structure, chemical synthesis, biogenesis and interaction with host cells*, Ed. Springer-Verlag, 2011

Peri F, Piazza M, Calabrese V, Damore G, Cighetti R, Exploring the LPS/TLR4 signal pathway with small molecules. *Biochemical Society Transactions*, 2010, **38**, 1390-1395

Communications

BtBs Day, University of Milano-Bicocca, Milan, December 5, 2013. Oral presentation "Small molecules. From design to biological characterization" and poster presentation "New molecular tools to investigate LPS recognition".

24th Joint Glycobiology Meeting, Wittenberg (Germany), November 24-26, 2013. Poster presentation and flash oral presentation "New molecular tools to investigate LPS recognition". Winner of poster prize.

32nd annual ESRA congress 2013, Glasgow, September 4-7, 2013 . P. De Negri, T. Tirri, F. Peri, R. Cighetti, G. D'Orazio, "Stability evaluation of Ziconotide-Morphine Hydrochloride and Ziconotide-Baclofen admixtures under experimental conditions" presented as e-poster by Dr. Pasquale de Negri

15th International Congress of Immunology, Milan, August 22-27, 2013. "New molecular tools to investigate LPS recognition", accepted as poster presentation.

15th ICI Satellite Meeting "Endotoxin, TLR4 signaling and beyond", Cinisello Balsamo (Milan), August 21, 2013. "New molecular tools to investigate LPS recognition", accepted as poster presentation.

European Young Investigators Workshop "Deciphering the Glycome – from synthesis to applications", Potsdam (Berlin), March 17-20, 2013. Oral presentation of my research topic at the "Glycan Forum", March 20, 2013

COST BM1003 Winter School "Microbial infection from the chemistry perspective: the bottom-up approach", Madrid, January 23-25, 2013, oral presentation of my research topic

Workshop "*Glycosynthesis*", Prague (Czech Republic), April, 18th, 2012, organized by European Network "*Carbohydrate multivalent systems as tools to study pathogen interactions with DC-SIGN*" (CARMUSYS)

XVI School of Pure and Applied Biophysics – Multimodal Methods for Cell Imaging and Tracking, Venice (Italy), January 30th – February 3rd, 2012, Poster: *"A fluorescent small-molecule probe for CD14 detection"*

Workshop "*Human applications and clinical trials design*", Toledo (Spain), June 28-29, 2011, organized by Eurpean Network "*Carbohydrate multivalent systems as tools to study pathogen interactions with DC-SIGN*" (CARMUSYS)

Oral presentation at the event "Science & Art", May 4-8, 2011, Modern Art Gallery of Villa Reale, Milano (Italy), organized by Prof. Francesco Peri in collaboration with Comune di Milano

Workshop "*Glycoscience: tools and molecules*", Milano (Italy), December 28-29, 2010, organized by Eurpean Network "*Carbohydrate multivalent systems as tools to study pathogen interactions with DC-SIGN*" (CARMUSYS)

EFMC-ISMC 2010, International Symposium on Medicinal Chemistry, September 5-9, 2010, Bruxelles, Belgium, Poster: "New sugar-derived compounds targeting selectively CD14 and MD-2 receptors as anti sepsis and anti neuropathic pain agents", also published in: *Drugs of the future*, 2010, 35 (supplement A), p. 184

Workshop - Structural glycoscience: methods and instrumentation, June 29 – July 1, 2010, ESRF, Grenoble, France