

UNIVERSITY OF NAPLES FEDERICO II



PhD in Chemical Sciences
XXX Cycle

Structural characterization and semi-synthetic
modification of bacterial glycolipids

PhD Student: Marcello Ziaco

Tutor:
Prof. M. M. Corsaro

Supervisor:
Dr. A. Guaragna

Index

ABBREVIATIONS	V
SUMMARY	VIII

SECTION I: Introduction

<u>Chapter 1: Gram-negative bacteria</u>	2
1.1 The Kingdom of Life	2
1.2 Bacterial cell envelope	3
1.3 Lipopolysaccharides (LPSs)	5
1.3.1 The Lipid A moiety: structure and activity	8
1.3.2 The core oligosaccharide: structure and activity	12
1.3.3 The O-polysaccharide moiety: structure and activity	13
1.4 LPS and Elicitation of Host Immune response	14
1.4.1 Innate immunity and Adaptive Immunity	14
1.4.2 LPS and Innate Immunity	19
References	22
<u>Chapter 2: Structural characterization of LPS and LOS</u>	24
2.1 Extraction and Purification of LPS and LOS	24
2.2. Structural determination of the saccharide moieties	25
2.2.1 Chemical analyses of the saccharide moieties	28
2.2.2 NMR Spectroscopy analysis of the saccharide moieties	30
2.2.3 Mass spectrometry analysis of the saccharide moieties	33
2.3 Structural determination of the lipid A moiety	35
2.3.1 Chemical analyses of the lipid A moiety	35
2.3.2 Mass spectrometry analysis of the lipid A moiety	36

References	38
------------------	----

SECTION II: Lipid As isolated from Extremophiles

<u>Chapter 3: Extremophiles, an overview</u>	45
3.1 Life-Forms in inhospitable environments	45
3.2 Psychrophiles	47
3.3 Molecular and physiological adaptation	49
3.4 Industrial and biomedical applications	53
References	55
<u>Chapter 4: <i>Colwellia psychrerythraea</i> strain 34H</u>	57
4.1 Lipopolysaccharides and lipid A structures	58
4.1.1 Isolation and compositional analysis of lipid A	59
4.1.2 Mass spectrometry analysis of lipid A	60
4.1.3 NMR spectroscopy of lipid A	69
4.1.4 Evaluation of biological activity	72
4.2 Conclusions	73
4.3 Experimental section	75
References	80
<u>Chapter 5: <i>Psychrobacter arcticus</i> strain 273-4</u>	82
5.1 Lipopolysaccharides and lipid A structures	82
5.1.1 Isolation and compositional analysis of lipid A	84
5.1.2 Mass spectrometry analysis of lipid A	85
5.1.3 Evaluation of biological activity	90
5.2 Conclusions	91
5.3 Experimental section	91
References	94

SECTION III: Semisynthetic derivatives from *Escherichia coli* lipid A

<u>Chapter 6: Multigram-scale extraction and purification of lipid A from <i>E. coli</i></u>	97
6.1 Isolation, purification and compositional analysis of	

lipid A from <i>E. coli</i> K4 fermentation	99
6.2 Selective de- <i>O</i> -phosphorylation of lipid A	102
6.3 Structural characterization and biological assays	103
6.4 Conclusions	106
6.5 Experimental section	107
References	112
<u>Chapter 7: Functionalization on the C-6 of GlcN II of <i>E. coli</i></u>	
<u>monophosphoryl lipid A</u>	114
7.1 Site-selective chemical reactions	115
7.2 Insertion of a clickable moiety	119
7.3 Semisynthetic “self-adjuvant” anticancer vaccine candidates	133
7.3.1 <i>E. coli</i> MPLA - Tn antigen conjugate	137
7.3.2 <i>E. coli</i> MPLA - TF antigen conjugate	141
7.4 Biological assays	146
7.5 Multivalent Tn antigens conjugates - MPLA	152
7.6 Conclusions	155
7.7 Experimental section	156
References	170
<u>Chapter 8: Lipid pattern modification of <i>E. coli</i> monophosphoryl</u>	
<u>lipid A</u>	173
8.1 Selective deacylation reactions	173
8.2 Selective re-acylation with shorter chains	177
8.3 Biological assays	185
8.4 Conclusions	187
8.5 Experimental section	187
References	191
<u>Chapter 9: Fluorescent derivatives of <i>E. coli</i> lipid A</u>	
9.1 Flippases	193
9.2 Fluorescent mono/diphosphoryl lipid As and ABC transporters	194
9.3 Imaging on HEK and macrophages cells	201

9.4 Conclusions	205
9.5 Experimental section	205
References	207
Appendix	
<u>PhD Course Activity Summary</u>	210
Attended Courses	210
Attended Seminars	210
Publications	212
Attended congresses/workshops/summer schools/contribution	213
Other Activities	214

Portions of this work have been adapted from the following articles that were co-written by the author:

- A. Casillo, M. Ziaco, B. Lindner et al., Unusual Lipid A from a cold adapted bacterium: detailed structural characterization. *ChemBioChem* **2017**, *18*, 1 - 11.
- D. D'Alonzo, M. Cipolletti, M. Ziaco et al., A Semisynthetic Approach to New Immunoadjuvant Candidates: Site-Selective Chemical Manipulation of *Escherichia coli* Monophosphoryl Lipid A. *Chem. Eur. J.* **2016**, *22*, 1-12.
- M. Ziaco, S. Górska, S. Traboni et al., Development of clickable monophosphoryl lipid As towards semi-synthetic conjugates with tumor-associated carbohydrate antigens. *J. Med. Chem.* **2017**, *60*, 9757-9768.

ABBREVIATIONS

AFGPs = Antifreeze glycoproteins

AFP = Antifreeze proteins

AMG = Acetylated methyl glycoside

APC = Antigen presenting cells

BMDMs = Bone marrow-derived macrophages

CDMT = 2-chloro-4,6-dimethoxy-1,3,5-triazine

COSY = Correlation spectroscopy

CS = Chondroitin sulfate

CSPs = Cold-shock proteins

DCC = Dicyclohexylcarbodiimide

DIPEA = *N,N*-diisopropylethylamine

DIPC = *N,N'*-diisopropylcarbodiimide

DMAP = 4-(*N,N*-dimethylamino)pyridine

DMF = *N,N*-dimethylformamide

DMSO = Dimethyl sulfoxide

DO = Dissolved-oxygen

DPAP = 2,2-dimethoxy-2-phenylacetophenone

DPLA = Diphosphoryl lipid A

EI = Electron ionisation

ELISA = Enzyme-linked immunosorbent assay

EPSs = Exopolysaccharides

ESI = Electro Spray Ionisation

ESI FT-ICR MS = Fourier transform ion cyclotron resonance mass
spectrometry

GalNAc = *N*-Acetyl-galactosamine

GlcN = Glucosamine
GlcNAc = *N*-Acetyl-glucosamine
HEK = Human Embryonic Kidney
HMBC = Heteronuclear multiple bond correlation
HOBt = 1-hydroxybenzotriazole
HSQC-DEPT = Heteronuclear Single Quantum Coherence-
Distorsionless Enhancement by Polarization Transfer
IBX = 2-iodoxybenzoic acid
Kdo = 3- deoxy-D-manno octulosonic acid
LBP = LPS- binding protein
LOS = Lipooligosaccharide
LPS = Lipopolysaccharide
MALDI = Matrix Assisted Laser Desorption Ionisation
MHC = Major histocompatibility complex
MPLA = Monophosphoryl lipid A
NBD = 4-Chloro-7-nitrobenzofurazan
NOESY = Nuclear Overhauser enhancement spectroscopy
NMM = 4-methylmorpholine
OM = Outer membrane
PAMPs = Pathogen Associated Molecular Patterns
PMN = Polymorphonuclear neutrophilic leukocytes
PRRs = Pattern Recognition Receptors
PyBOP[®] = (benzotriazol-1-yloxy)tripyrrolidinophosphonium
hexafluorophosphate
RAFT = Regioselectively Addressable Functionalised Template
SEC = Size exclusion chromatography
TACA = Tumor-associated carbohydrate antigen

THP-1 = Human monocytic cell line

TLRs = Toll-Like Receptors

TOCSY = Total correlation spectroscopy

TBAB = Tetrabutylammonium bromide

TBDMS = tert-butyldimethylsilyl

TBDPS = tert-butyldiphenylsilyl

TBTU = *O*-(benzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium
tetrafluoroborate

TEMPO = 2,2,6,6-tetramethylpiperidine-1-oxyl

TMSOTf = Trimethylsilyl trifluoromethanesulfonate

SUMMARY

The present PhD thesis aims to discover new natural lipid A structures and to obtain semi-synthetic derivatives thereof to be tested as candidates for the pharmacological use as vaccine adjuvants.

The term lipid A indicates a family of glycolipids representing the inflammation-inducing moiety of lipopolysaccharides (LPSs) and lipooligosaccharides (LOSs), that are amphiphilic molecules embedded in the Gram negative bacterial outer membrane, of which they are the major constituents and very often the major virulence factor.

The general structure of LPSs is characterized by three distinct portions:

- The **lipid A**, generally composed of a β -D-GlcN-(1 \rightarrow 6)-D-GlcN disaccharide (GlcN=2-amino-2-deoxy-glucopyranose) with a variable number (from three to seven) of acyl chains linked to the positions 2 and 3 of both glucosamine residues through amide or ester bonds, respectively. The structure is generally completed by one or two phosphate groups, linked at position O-4 of the non-reducing unit (GlcN II) and/or at position O-1 of the pseudo-reducing residue (GlcN I) with an α -configuration.¹
- The **core oligosaccharide**, distinguishable in an inner core and an outer core. This oligosaccharide is linked to the lipid A and is well conserved among different strains of the same bacteria and consists of uncommon sugar residues such as heptoses (L-glycero-D-manno-heptose and, less commonly, D-glycero-D-manno-heptose) and Kdo (3- deoxy-D-manno octulosonic acid).
- The **O-chain** polysaccharide built up of oligosaccharide repeating units. The O-chain polysaccharide is the most variable portion of the LPS also within bacteria belonging to the same genus; it consists of up to 50 identical repeating units composed of two to eight different glycosyl residues (heteroglycans) or, in some bacteria, of identical sugars (homoglycans). This moiety can be absent and in that case LPSs are named LOSs.

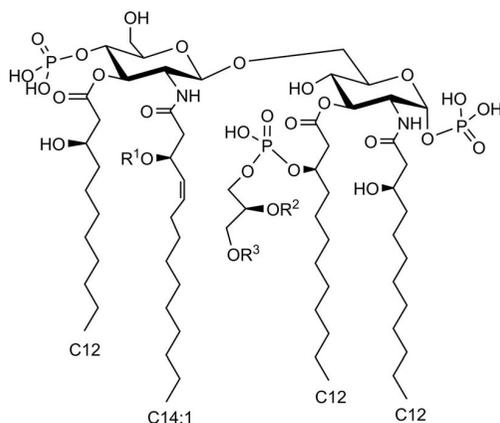
Furthermore, LPS elicits host immune response through the Toll-like receptors-4 (TLR-4) and Myeloid differentiation factor-2 (MD-2) receptorial complex expressed on macrophages and neutrophils. Toll-

like receptors (TLRs) play an important role in the recognition of molecular signals of microbial invasion (Microbial Associated Molecular Patterns, MAMPs) from virus, bacteria, protozoa and fungi. In particular, TLR-4 ligands, such as lipid A, trigger innate immunity response in the host through chemokine and cytokine signalling cascades.² This response is highly modulated by several structural features of lipid A,³ such as the number of phosphate groups and acyl chains as well as the length and the distribution of the latter.⁴ Indeed, hexa-acylated diphosphoryl lipid As possess strong inflammatory activity, that may be associated to a number of severe pathophysiological symptoms (fever, diarrhoea, blood pressure drop, septic shock), whereas underacylated lipid As can block or inhibit TLR-4, thus acting as TLR-4-antagonists. The deletion of the phosphate group at the anomeric position of GlcN I can furnish the 1-*O*-diphosphorylated form of lipid A, known as monophosphoryl lipid A (MPLA), was revealed to have significantly reduced endotoxicity but retain the strong immunostimulatory activity. MPLA has been proved to be clinically safe as a vaccine adjuvant to treat infectious and neoplastic diseases. The development of new TLR-4-stimulators to be used as immuno-adjuvants has attracted much effort in the last years. To this aim, the research is focused on the obtainment of new example of lipid A structures, both through isolation from wild-type or engineered bacterial extracts^{1,5} and by synthetic⁶ and semisynthetic approaches.⁷

In the former frame, extremophiles may be a promising source of non-toxic LPSs and lipid As. In particular, cold environments, due to the still low contamination by pathogenic microorganisms, may represent an untapped reservoir of immunomodulating molecules. The outer membrane of Gram-negative bacteria is mainly constituted by LPSs, and it was demonstrated that structural changes could be present in these macromolecules isolated from cold-adapted bacteria.

In particular, the analysis of *C. psychrerythraea* strain 34H genome, evidences the features that allow the live at freezing temperatures and for this reason this bacterium is considered as a model for the study of cold adapted life-style. The structures of the lipid A moiety from *C. psychrerythraea* strain 34H and its partially deacylated derivative were completely characterized in this Ph.D. work by high-resolution mass spectrometry (HR-MS), NMR spectroscopy, and chemical analysis. The high heterogeneity of the lipid A structure, showed by the fatty acids

analysis, was confirmed by the complexity of MS and MS/MS spectra. These experiments indicated a complex acylation pattern ranging from tetra- to hepta-acylated glycoforms. Furthermore, this lipid A displayed a structure that is quite new among known LPSs. An unusual structure with a 3-hydroxy unsaturated tetradecenoic acid as a component of the primary acylation pattern was identified. In addition, the presence of a partially acylated phosphoglycerol moiety on the secondary acylation site at the 3-position of the reducing 2-amino-2-deoxyglucopyranose unit caused tremendous natural heterogeneity in such lipid A structure.



[M-H]	R ¹	R ² + R ³
1780.021	C10:0	H + C10:0
1808.058	C12:0	H + C10:0
1890.128	C12:0	H + C16:1
2019.047	C10:0	C10:0 + C16:0
2046.289	C12:0	C10:0 + C16:0
2073.290	C12:0	C12:0 + C16:1
2100.325	C12:0	C14:0 + C16:1

Structure of *C. psychrerythraea* 34H lipid A.

The double bond is arbitrarily shown in the *cis* configuration.

The peculiar structural features showed by *C. psychrerythraea* 34H lipid A could be related to the cold-adaptation mechanisms adopted by this bacterium.⁸ Following the same strategy, during this Ph.D. work the lipid A moiety derived from the LPS of *Psychrobacter arcticus* 273-4 grown at 4°C, a Gram-negative bacterium isolated from Kolyma region of Siberian permafrost core⁹, was completely elucidated. *P. arcticus* 273-4 is considered a psychro-tolerant microorganism, capable to grow at both low, even subzero temperatures, and even more than 20°C. The lipid A

and its partially deacylated derivatives were completely characterized by chemical analysis, electrospray ionization Fourier transform ion cyclotron (ESI FT-ICR) and Orbitrap mass spectrometry. As the chemical analysis and the mass spectrometry showed, the lipid A of *P. arcticus* is characterized by the presence of shorter acyl chains than lipid A from *C. psychrerythraea* 34H; decanoic acid (C10:0) as secondary acyl groups and mainly 3-hydroxy-dodecanoic acid [C12:0(3-OH)], 3-hydroxy-tetradecanoic acid [C14:0(3-OH)] and minor amount of 3-hydroxy-undecadecanoic acid [C11:0(3-OH)], 3-hydroxy-tridecanoic acid [C13:0(3-OH)] as primary acyl groups. Therefore, the lipid A moiety from both the studied psychrophilic bacteria display several structural features that could be related to the adaptation mechanisms and in particular an increased incorporation of unsaturated, short, and/or branched fatty acids to maintain membrane fluidity as part of a set of physiological actions necessary to survival in cold conditions.¹⁰

Preliminary biological activity assays were performed to characterize the effect of *C. psychrerythraea* 34H and *P. arcticus* 273-4 lipid A on TNF secretion in the presence of a serial dilution of *Escherichia coli* O111:B4 LPS. The elucidation of new lipid A structures is fundamental for the comprehension and the enlargement of the number of structure-activity relationships necessary for the design of new immunomodulatory molecules. Actually, these components are crucial for the development of modern formulations of recombinant or synthetic vaccines, in order to obtain strong and long-lasting responses also in elderly and immunocompromised patients.¹¹

These molecules are usually obtained by purification of wild-type or engineered bacterial extracts without very good results for the purity,¹ as well as through total chemical synthesis, that requires lengthy procedures with a high number of synthetic steps and very difficult to apply on a large-scale production.

An emerging approach is the use of semi-synthetic strategies that combine the advantages of a fermentative access to complex natural products with the development of site-selective chemical reactions modifying these structures.

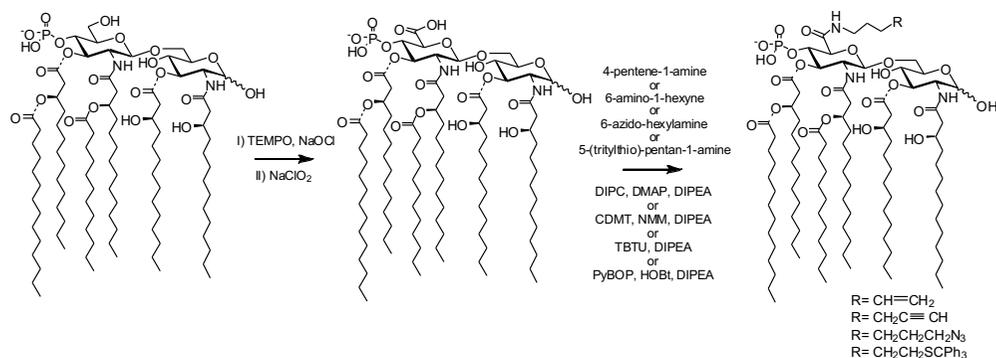
In particular, Prof. Corsaro and co-workers developed a semi-synthetic approach relying upon the use of lipid A in gram quantities from fed-batch fermentation of *E. coli*, after de-*O*-phosphorylation at *O*-1 position

of GlcN I, as a scaffold for the obtainment of several other lipid As and derivatives thereof through few, tailored chemical modifications.⁷

These site-selective reactions were developed in this Ph.D. work, with a special focus on the functionalization of the primary hydroxyl group at position C-6 of GlcN II and on the modification of the lipid pattern.

The former was attractive not only for the expected higher reactivity of the primary alcohol with respect to the secondary hydroxyls, but also because it could be biologically well tolerated since lipid As and the core part of LPSs are linked through a glycosidic bond between this position and the Kdo or, much more rarely, 2-octulosonic acid (Ko). In order to insert a chemical handle that would be easily and orthogonally derivatizable with respect to the other functionalities present on lipid A structure, the installation of a carboxylic acid function at position 6 of GlcN II was performed. This derivatization could not only serve as a suitable moiety for further functionalization, but also mimics the Kdo structure (lipid A-Kdo structures have been demonstrated to induce stronger immune responses than their parent lipid As).^{6, 12}

In order to insert a clickable moiety on such oxidized lipid A structure, this derivative was treated with diverse amines carrying a suitable functionalization at their ω -position (double or triple bond, azide, S-trityl as protected thiol). The first aim was to find the best conditions for the amide formation reaction using a model amine such as 4-penten-1-amine. The coupling was tested under several typical peptide coupling conditions. After having detected PyBOP[®] and TBTU as the best coupling reagent under appropriate conditions, these were employed to install the other clickable functionalities indicated above.



Insertion of a clickable moiety on *E. coli* monophosphoryl lipid A.

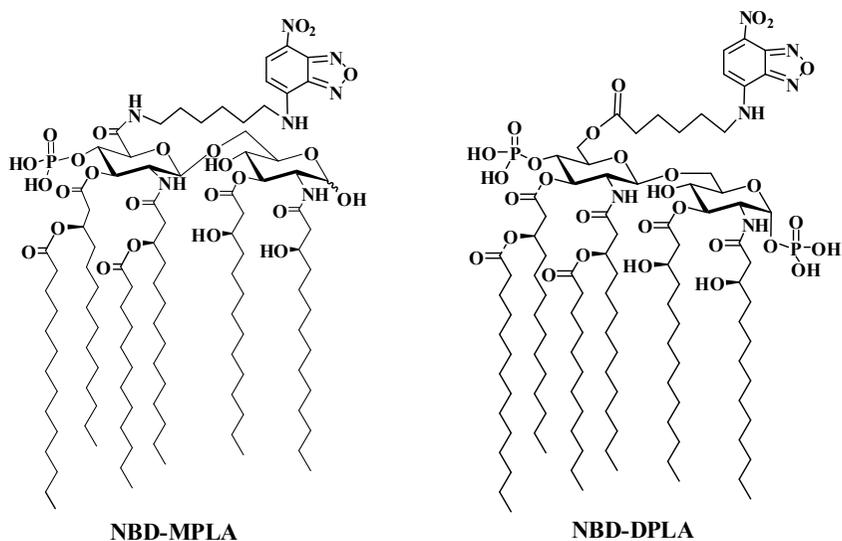
The production of these clickable lipid A derivatives opens a straightforward access to their conjugation with other interesting biomolecules, such as tumor-associated carbohydrate antigens (TACAs). In order to demonstrate the feasibility of this approach, the thiol-equipped derivatives of Tn and TF antigens were synthesized and then conjugated with a clickable lipid A derivative carrying a double bond moiety, through a UV-mediated thiol-ene reaction, to give two potential self-adjuvant anticancer vaccine candidates. A preliminary evaluation of the immunological activity of these lipid A-TACA antigen conjugates, as well as of other semi-synthesized lipid A derivatives, was performed.¹⁶ It was shown that most of them ameliorated production of proinflammatory cytokines IL-6, TNF- α , and IFN- γ compared to control. The tested derivatives also induced chemokine CXCL8 (IL-8) and immunoregulatory cytokine IL-10 thus highlighting the promising immunobiological applications. Additionally, the described compounds exhibited very low toxicity which could facilitate their use in vaccines and give them an advantage over the currently used adjuvants.

Following this strategy, a collaboration with Prof. Renaudet (Univ. Grenoble Alpes, CNRS, France) has been started, in order to develop the first Regioselectively Addressable Functionalised Template (RAFT) scaffold-MPLA conjugates as potent multivalent immunogenic vaccine candidates.

Moreover, we finely tuned the conditions suitable for the semisynthesis of other new lipid A derivatives, by selective modification of the lipid pattern (selective deacylation and re-acylation with shorter chains). In this way, a small collection of new lipid A derivatives was rapidly prepared, avoiding lengthy total synthetic approaches. Notably, potential access to several other modified lipid A structures could be opened by this strategy, because the reactions developed on *E. coli* lipid A scaffold should also be applicable on lipid A substrates from different bacterial sources. Some of the new compounds were assayed for preliminary in vitro immunological tests, showing somewhat reduced inflammatory activity and a cytokine production profile similar to the already used immunomodulant MPL[®] *Salmonella* Minnesota lipid A, thus highlighting the promising immunobiological applications.

Finally, in the last part of this PhD work, two different fluorescent lipid A derivatives were semi-synthesized, with the aim to improve the

understanding of the potential role of lipid A in transmembrane protein stability, in mediating drug export and in the mechanisms of ABC transporters.



Semi-synthetic fluorescent mono- (left) and diphosphoryl (right) lipid A derivatives.

In collaboration with Prof. Tampé (Institute of Biochemistry, at the Goethe-University in Frankfurt, Germany) these two derivatives were exploited in several biological assays in order to develop some *in vitro* fluorescence-based lipid A flippase assays, focused on the ABC transporters mechanisms comprehension (data still under investigation).¹⁷ Finally, these two fluorescent derivatives, in collaboration with Prof. Huser (Department of Physics, University of Bielefeld, Germany), were subjected to a preliminary cell imaging test aimed to observe their staining with two different cell lines on a deconvolution microscope (experiments still in progress).

References

1. Molinaro, A. et al., *Chem. Eur. J.* **2015**, *21* (2), 500–519.
2. Medzithov, R. *Nature* **2008**, *454*, 428-435.
3. Park, B.S. et al., *Nature* **2009**, *458*, 1191-1196.
4. (a) Li, Y. et al., *Mar. Drugs* **2013**, *11*, 3197-3208; (b) Rietschel, E.T. et al., *FASEB J.* **1994**, *8*, 217-225.

5. (a) Needhama, B.D. et al., *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 1464-1469; (b) Carillo, S. et al., *Chem. Eur. J.* **2011**, *17*, 7053-7060.
6. Gao J., Guo Z., *Med. Res. Rev.* **2017**, (in press)
doi: 10.1002/med.21447.
7. (a) D'Alonzo, D. et al., *Chem. Eur. J.* **2016**, *22*, 11053-11063; (b) Pieretti, G. et al., *Appl. Microbiol. Biotechnol.* **2014**, *98*, 7781-7790.
8. Casillo A. et al., *ChemBioChem* **2017**, *18*, 1 – 11.
9. Vishnivetskaya, T. et al., *Extremophiles* **2000**, *4*, 165–173.
10. Carty S.M. et al., *J. Biol. Chem.* **1999**, *274*, 9677 – 9685.
11. (a) Kwissa, M. et al., *Expert Rev. Vaccines* **2007**, *6*, 673 – 684; (b) Marciani, D.J. *Drug Discovery Today* **2003**, *8*, 934 – 943.
12. (a) Zughailer, S.M. et al., *Infect Immun.* **2004**, *72*, 371-380; (b) Yoshizaki, H. et al., *Angew. Chem. Int. Ed.* **2001**, *40*, 1475-1480.
13. Saito, Y. et al., *Tetrahedron* **2008**, *64*, 3578–3588.
14. Lee, J. W. et al., *Tetrahedron Lett.* **2001**, *42*, 2709–2711.
15. Jagadish, B. et al., *J. Med. Chem.* **2012**, *55*, 10378–10386.
16. Ziaco M. et al., *J. Med. Chem.* **2017**, *60*, 9757-9768.
17. Bechara, C. et al., *Nat. Chemistry* **2015**, *7*, 255 – 262.

SECTION I

Introduction

Chapter 1: Gram-negative bacteria

1.1 The Kingdom of Life

The three-domain system, which classifies organisms on Earth into three distinct kingdoms - Archea, Bacteria and Eukarya - was designed by the American microbiologist Carl Woese in the 1990. Otherwise, an alternative biological classification system currently used is known as “the six-kingdoms system of life” and include Archaeobacteria, Eubacteria, Protista (comprising one-celled organisms), Fungi, Plantae and Animalia. While Archaeobacteria and Eubacteria constitute the Archaea and Bacteria domains respectively (both representing the Prokarya kingdom), Protista, Fungi, Plantae and Animalia together form the Eukarya kingdom of life. Most prokaryotic cells have diameters in the range of 1-10 μm , much smaller than most eukaryotic cells (typically 10-100 μm). Prokaryotes have a cellular organization different from that of eukaryotes; these latter have a membrane-enclosed nucleus and numerous membrane-enclosed organelles, whereas prokaryotic cells lack these structural features. Typically, Archea domain includes all the extremophiles prokaryotic organisms that live and thrive in inhospitable, from a human viewpoint, environments such as alkaline and acidic waters, boiling hot springs, high pressure waters, ultra-saline brines, without excluding a combination of aforesaid chemical and physical extremes which is typical of polyextremophilic microbes (Mesbah and Wiegel, 2008). Bacteria comprise a wide group of microorganisms characterized by a bewildering assortment of size, shapes and arrangements reflecting the diverse environments in which they grow and reproduce (Pommerville, 2010). Despite the high structural variability,

Bacteria can be divided on the basis of their basic cell morphologies: a bacterial cell with a rod shape is called “bacillus” whereas spherical and curved cells are called “cocci” and “spirilli” respectively (Raven and Johnson, 2001). Some rod-shaped and spherical bacteria form colonies adhering end-to-end after their cellular division forming chains, whereas spirilla bacteria generally do not form associations with other cells. However, early systems for classifying bacteria relied on differential stains such as the colorimetric assay developed by the Danish physician Hans Gram, in use to date, that allows the distinction of bacteria in Gram-positive and Gram-negative on the basis of the different response to the test. The differences between Gram-positive and Gram-negative bacteria are related to diversities in the structure and chemical composition of their cell wall: Gram-positive bacteria have the thicker cell envelope, nearly uniformly dense layers, and stain the purple color of the crystal-violet typically used in the Gram stain, whereas Gram-negative bacteria contain just a single or two layers and lose the purple-colored dye (Staley et al., 2007).

1.2 Bacterial cell envelope

The complex system of layers that surrounds bacterial cells is referred collectively as the cell envelope. It consists of a cytoplasmic membrane (or plasma membrane), cell wall, and for some bacteria, capsule and glycocalyx from inner to outer surface. Both Gram-positive and Gram-negative bacteria present the cytoplasmic membrane, namely a phospholipid bilayer surrounding the cytoplasm and representing a physical semi-permeable barrier that regulates the transport of nutrients and metabolic products in and out the cell. This membrane is, in turn, enclosed by the cell wall, an important structure that maintains the

cellular shape and protects the cell from swelling and rupturing (Raven and Johnson, 2001). The bacterial cell wall differs markedly from archaeal and eukaryotic cells walls in containing peptidoglycan (or murein) which is a network of saccharide molecules connected by polypeptide cross-links. In detail, it is composed of β -(1 \rightarrow 4)-linked disaccharide chains, alternating *N*-acetylglucosamine and *N*-acetylmuramic acid residues in turn cross linked by short peptide stems composed of alternating L- and D-amino acids (Figure 1.1). It is possible to classify peptidoglycans on the basis of the third amino acid residue of the stem peptide; indeed, in Gram-negative bacteria this residue is the unusual *meso*-diaminopimelic acid (*m*-DAP), while in Gram-positive bacteria is commonly a lysine (Figure 1.1).

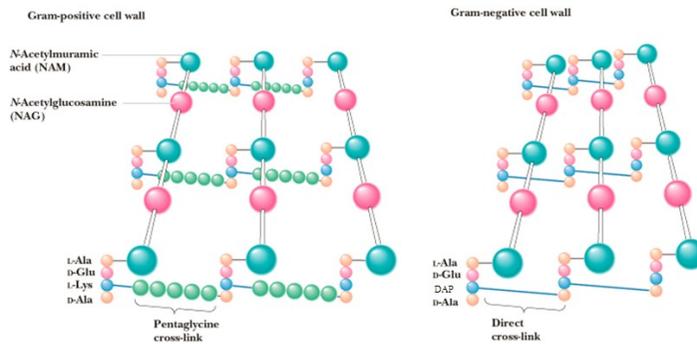


Figure 1.1 Peptidoglycan structure in Gram-positive (left) and Gram-negative (right) bacteria.

Furthermore, in Gram-positive bacteria the external portion of the cell is composed by 95% of rigid peptidoglycan cell wall that may be one of the reason why they retain the crystal-violet of Gram stain (Figure 1.2); on the contrary, in Gram-negative cells there is a thin layer of peptidoglycan, almost 5-10% of the cell wall, surrounded by an additional asymmetric phospholipid bilayer, termed outer membrane (OM) (Figure 1.2). This

external membrane is a unique feature of Gram-negative bacteria and represents the first barrier that protects bacterial cell from environmental stress factors, thus it is essential for bacterial survival. The outer membrane is separated by a gel-like compartment, termed periplasm, from the cytoplasmic membrane and shows to have a structure similar to that of the plasma membrane, from which nonetheless differs for important physical and chemical properties. The inner portion of the OM is mainly composed of phospholipids, while the external leaflet is occupied for the 75% by glycolipids known as lipopolysaccharides (LPS) (Alexander and Rietschel, 2001). Furthermore, the outer membrane is also constituted by a large number of proteins, among which there is the Braun's lipoprotein (BLP or murein lipoprotein), that acts as a bridge between the outer membrane and the peptidoglycan layer.

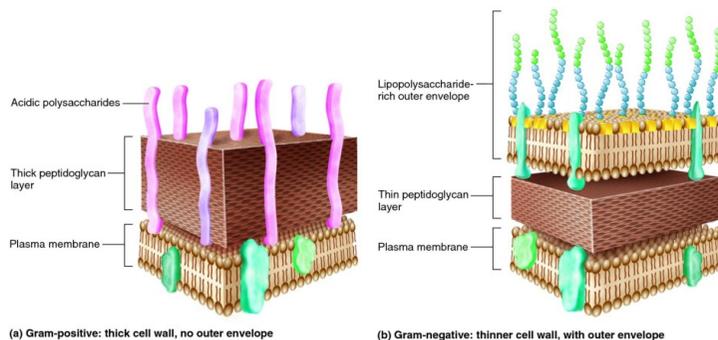


Figure 1.2 Cell-wall structure of Gram-positive (left) and Gram-negative bacteria (right).

1.3 Lipopolysaccharides (LPSs)

Lipopolysaccharides are heat-stable amphiphilic molecules indispensable for viability and survival of Gram-negative bacteria, as they heavily contribute to the structural integrity of the OM and to the protection of the bacterial cell envelope (Holst et al., 1996).

Being intercalated into the outer membrane, they form hydrophobic interactions with the phospholipid bilayer composing the inner layer, giving rise to the typical asymmetry of Gram-negative bacteria OM (Holst et al., 1996). The highly ordered structure and low fluidity of the LPS monolayer, stabilized by electrostatic interactions of divalent cations (as Ca^{2+} and Mg^{2+}) with negatively charged groups present on LPS molecules, are responsible for the increase of permeability to hydrophobic and higher molecular weight hydrophilic compounds, but also for the bacteria resistance to external stress factors. Indeed, most of the commonly used antibiotics directed against Gram-negative bacteria, such as polymyxin B, are able to destabilize abovementioned ionic interactions leading to the disruption of membrane integrity (Cardoso et al., 2007). In addition, since they are exposed toward the external environment, LPS molecules participate in crucial mechanisms of host-bacterium interactions like colonization, virulence in the case of pathogen and opportunistic bacteria, adhesion and symbiosis (Silipo et al., 2010). Among all these activities, LPS has been shown to be the most potent immunostimulant molecule playing a key role in the pathogenesis of Gram-negative infections and triggering the immune system in a wide range of eukaryotic organisms ranging from insects to humans (Alexander and Rietschel, 2001). LPSs belonging to different bacterial species possess different structures, but also LPS of an individual bacterial strain is not a single molecule, possessing a specific chemical structure, but rather a blend of various molecules characterized by an intrinsic size and structural heterogeneity (Raetz, 1990; Raetz and Whitfield, 2002). Furthermore, bacteria of the same species producing diverse LPS molecules under different growth conditions were also found. Despite the high structural heterogeneity, all LPSs broadly

comply with a common basic architecture composed of well-conserved domains in which three distinct regions encoded by different gene clusters, can be distinguished (Figure 1.3). Indeed, they consist of a polysaccharide (also known as O-antigen, O-side chain or O-specific polysaccharide) characterized by a highly variable structure which constitutes the chemical basis for the serological classification of bacterial strains. O-antigen is covalently linked to an oligosaccharide part (core), in turn, linked to a glycolipid portion (lipid A) that is the most conserved region of the entire macromolecule, even among different species belonging to the same genus (Holst et al., 1996).

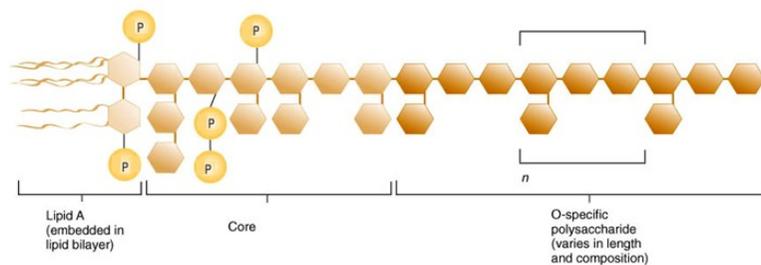


Figure 1.3 Schematized structure of Gram-negative LPSs.

The lipid A moiety is embedded in the outer leaflet of the OM whereas the sugar chain is oriented outwards (Raetz and Whitfield, 2002). The O-polysaccharide chain is not ubiquitous, as it can be absent or partly truncated in some Gram-negative strains. Bacterial colonies can be catalogued by gross colony morphology as rough or smooth on the basis of the occurrence of the O-side chain, being absent in the former and present in the latter (Lüderitz et al., 1966). The terminology currently used to designate the different LPS types, namely with or without the O-polysaccharide portion, is S- LPS or R-LPS (or lipooligosaccharide, LOS), respectively (Raetz and Whitfield, 2002).

1.3.1 The lipid A moiety: structure and activity

Lipid A represents the hydrophobic and endotoxic principle of the LPS molecule since it acts as a potent stimulator of the host innate immune system (see paragraph 1.4.2). It possesses a highly conservative structure consisting of a β -(1 \rightarrow 6)-linked D-glucosamine disaccharide backbone substituted with a number of amide- and ester-linked 3-hydroxy fatty acids at the positions 2 and 3 (Figure 1.4), respectively. The acyl groups that are directly linked to the sugar backbone are defined primary and some of them are further acylated at the hydroxy groups by secondary acyl chains. Furthermore, the sugar backbone is generally α -phosphorylated at position O-1 of the reducing glucosamine (GlcN I) and at position O-4' of the non-reducing glucosamine (GlcN II) (Figure 1.4) (Zahringer et al., 1999). The firstly characterized lipid As were from *Escherichia coli* (Figure 1.4) and *Salmonella enterica* LPSs in 1983. *E. coli* lipid A is built up of the following sugar backbone [P \rightarrow 4- β -D-GlcN-(1 \rightarrow 6)- α -D-GlcN-1 \rightarrow P] acylated at position 2 and 3 of both GlcNs by four C14:0 (3-OH) (Figure 1.4). The primary fatty acids located on the GlcNII were both esterified at their hydroxyl group by two secondary acyl chains: one C14:0 (3-OH) was esterified by a C12:0 while the other by a C14:0 (Figure 1.4).

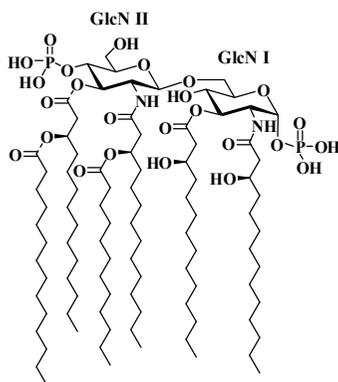


Figure 1.4 Lipid A from *Escherichia coli*.

Despite their general architecture, lipid As also present a micro-heterogeneity due to the presence of subtle chemical differences depending on a wide number of factors including bacterial adaptation, incomplete biosynthesis, changing of environment, presence of external stimuli and chemical modifications resulting from the procedures used for lipid A extraction from bacterial cells. Micro-heterogeneity has been observed in the acylation (number, type and distribution of acyl chains) and phosphorylation patterns, and, less commonly, also in the disaccharide backbone as in the case of bacterial species as *Aquifex pyrophilus* (Hellerqvist et al., 1971), *Brucella abortus* (Hellerqvist et al., 1969), whose GlcN residues may be replaced with 2,3-diamino-2,3-dideoxy-D-glucopyranose (GlcN3N) residues. Moreover, phosphate groups can be derivatized with further phosphate groups, producing a pyrophosphate, but also by other polar substituents such as 4-amino-4-deoxy-L-arabinopyranose (4-amino-arabinose, Ara4N) and 2-amino-ethanol group (EtN), or by acid residues such as galacturonic acid (GalA); phosphate groups can be absent, as in the case of *Bdellovibrio bacteriovorus* lipid A, characterised by the replacement of phosphate groups with two mannose residues generating a totally neutral lipid A. Concerning the acylation pattern, fatty acids can be attached to the disaccharide backbone either symmetrically (3+3, e.g. *Neisseria meningitidis*) or asymmetrically (4+2, e.g. *Escherichia coli*). Finally, lipid A fatty acids present less frequently further peculiar structural features such as a methyl branch, different functional and hydroxyl groups, length of the chains up to 28 carbon atoms, odd numbered carbon chains and unsaturation. Even the subtlest variation in chemical structure of lipid A can affect the LPS molecule bioactivity, since this latter is strictly structure-related; indeed, the lipid A intrinsic conformation is

responsible for its agonistic and antagonistic activity on innate immune system (Brandenburg et al., 1993; Rietschel et al., 1994; Seydel et al., 2000; Fukuoka et al., 2001; Oikawa et al., 2004).

A plethora of studies aimed to the elucidation of the LPS structure-activity relationship have highlighted that the hexa- acylated bisphosphorylated lipid A with an asymmetric 4:2 fatty acid distribution, found in the majority of enterobacteria, as the aforesaid *E. coli* lipid A, is considered to have the highest immunostimulatory capacity in mammalian cells. In contrast, the hypoacylated synthetic precursor of *E. coli* lipid A, the tetra- acylated lipid A (Lipid IVa), showed to have weak agonistic effects for some species of mammals and is well known to possess a strong antagonistic activity on human cells (Golenbock et al., 1991).

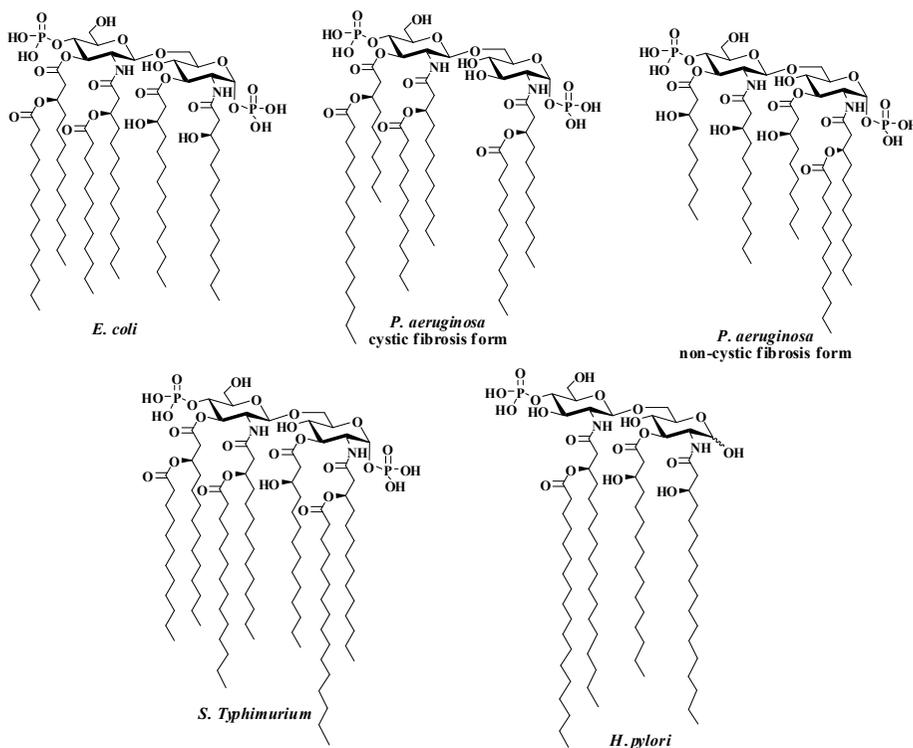


Figure 1.5 Lipid A chemical structure diversity in Gram-negative bacteria.

These different biological effects are correlated to two interconnected structural parameters: the molecular shape of the lipid A and the tilt angle between the di-glucosamine backbone and fatty acid chains that is the inclination of lipid A hydrophilic moiety respect to the hydrophobic portion (Netea et al., 2002). In details, the molecular shape possessed by lipid A influences its ability to be recognized by host immune system receptors, indeed it has been shown that at 37°C, in aqueous solution and in physiological conditions, the most agonistic lipid A has a truncated cone form that drives to a hexagonal supramolecular structure, while lipid As presenting an antagonistic activity assume a cylindrical shape leading to a lamellar structure (Netea et al., 2002). Regarding the tilt angle, it has been found that the most active form has a conical structure with a tilt angle $>50^\circ$, while the antagonist structures present smaller values of tilt angles (Seydel et al., 2000): species with a tilt angle $<25^\circ$, such as lipid IVA and the penta-acylated and symmetrically (3+3) hexaacylated lipid As, act as antagonists; species with an angle between 25 - 50° , as monophosphorylated lipid A, have a lower bioactivity (Figure 1.6) (Seydel et al., 2000).

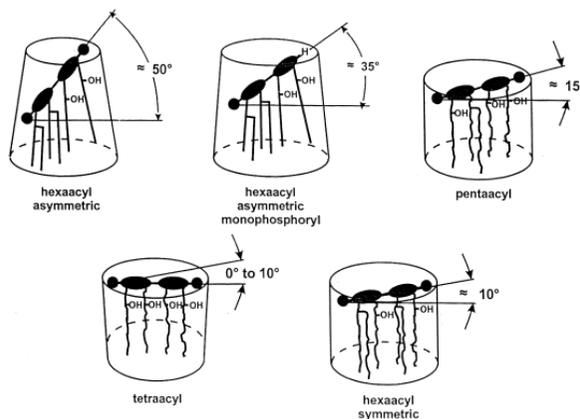


Figure 1.6 Tilt angle of lipid A depends from the acylation pattern.

The L-glycero-D-manno-heptose residues are often decorated with charged substituents like phosphate, pyrophosphate, Ara4N or uronic acids, as found in bacteria belonging to *Pseudomonas* genus and in many other species. It is speculated that all the substituents bearing a positively charged free amino group, as Ara4N residues, might play a role in pathogenesis since they reduce the negatively charged surface on the outer membrane rendering it positively charged or in an isoelectric state that, in turn, confers resistance to antibiotic compounds and antimicrobial peptides (Raetz and Whitfield, 2002; Hamad et al., 2012). The outer core region is the most exposed portion, often branched, and characterised by a higher structural variability than the inner core region. It is typically characterised by common hexoses such as glucose, galactose, *N*-acetyl glucosamine and *N*-acetyl galactosamine and it may also contain residues as 6-deoxy-L-mannose (L-Rha) and *N*-acetyl-2,6-dideoxy-D-glucosamine (D- QuiNAc).

1.3.3 The O-polysaccharide moiety: structure and activity

The O-chain polysaccharide is the most variable portion of the LPS also within bacteria belonging to the same genus; it consists of up to 50 identical repeating oligosaccharide units composed of two to eight different glycosyl residues (heteroglycans) or, in some bacteria of identical sugars (homoglycans). A single bacterium produces LPSs with O-chains characterised by a wide range of lengths as a result of incomplete synthesis of the polysaccharide chain (Raetz and Whitfield, 2002); this different degree of polymerization is responsible for the “ladder-like” pattern, showed by SDS-PAGE (Kittelberger and Hilbink, 1993) typical of a smooth LPS. The high structural variability of the O-polysaccharide is ascribable to the large number of sugar residues (in

both pyranose and furanose rings, anomeric and absolute configurations) that can build up the repeating units as well as to the glycosidic sequence and to the presence of non-carbohydrate substituents such as phosphates, amino acids, sulphates, acetyl or formamide groups, often present in a non-stoichiometric fashion (Adinolfi et al., 1996). The function of these substituents is frequently unknown although it can be speculated that bacteria can modify their LPSs to mask themselves to the host immune system.

1.4 LPS and elicitation of host immune response

1.4.1 Innate immunity and adaptive immunity

From the moment of birth and throughout our lifetime, humans are continuously exposed to a multitude of potential pathogen bacteria, through inhalation, direct contact and ingestion. The ability to avoid most of these infections depends on the adaptive immune system which remembers the previous encounters with specific pathogens and destroys them when they attack again (Alberts et al., 2007). Adaptive immune responses, however, are slow to develop on the first exposure to a new pathogen, since it is required the activation and expansion of specific cells that take approximately seven days prior to give an effective response. For this reason, the first days of exposure to a new pathogen see the intervention of the innate immune system (Alberts et al., 2007). Indeed, innate immune system is the first line of defence against invading organisms while the adaptive immune system acts as a second line of defence and also affords protection against re-exposure to the same pathogen; moreover, the innate immune responses are not specific to a particular pathogen in the way that the adaptive immune responses are.

The strategy of recognition typical of the innate immune system relies on the detection of constitutive and conserved microbial molecular targets, that are common to many pathogens but are absent in the host; these latter are known as PAMPs (Pathogen Associated Molecular Patterns) and are recognised by specific receptors, the PRRs (Pattern Recognition Receptors) present on the surface of some phagocytic cells that rapidly destroy them. It is worth to note that in vertebrates, the skin and other epithelial surfaces, including those lining lung and gut, provide a physical barrier between the inside of the body and the outside world, thus can be considered part of the innate immune system (Alberts et al., 2007). The microorganisms that pass this first barrier are recognised by the sentinels of innate immunity which are phagocytes, or white blood cells, or leukocytes, that are distinguishable in: neutrophils, polymorphonuclear neutrophilic leukocytes (PMN), macrophages and monocytes (Figure 1.8).

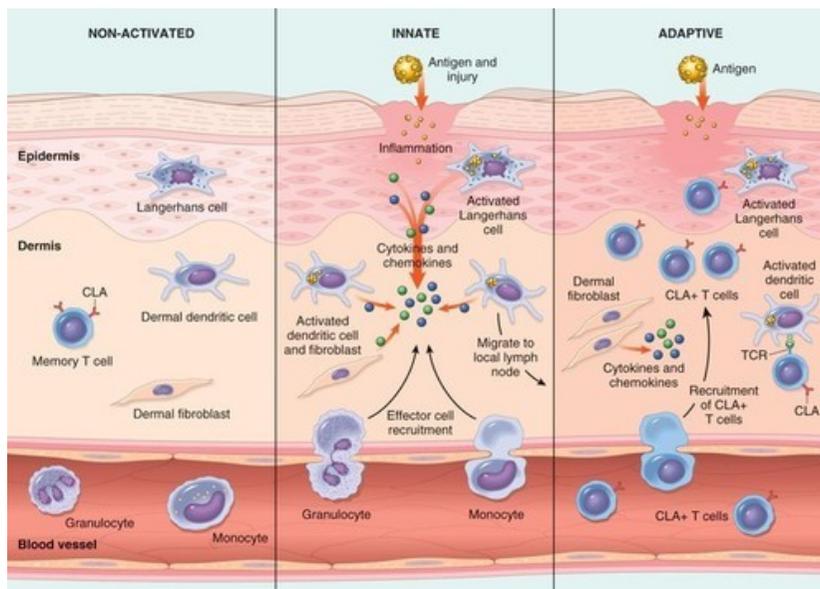


Figure 1.8 Overview of the innate and adaptive immune responses.

Macrophages and neutrophils show a variety of cell surface receptors, which make them able to recognise different types of microorganisms; the attack of a ligand to these receptors induces actin polymerization that allows the plasma membrane of phagocytes to surround the pathogens and to incorporate them forming a phagosome which becomes acidified (Figure 1.9). In addition to being phagocytic, macrophages and neutrophils have granules, called lysosomes, that contain lysozyme and hydrolases; the phagosome fuses with one or more lysosomes to generate a phagolysosome in which the lysosomal contents are released to degrade bacterial cell wall (Figure 1.9). After degradation, phagolysosomes exhibit the antigenic determinants on their cell surface with the consequent triggering of the mechanism of specific immune response (Figure 1.9) (Janeway et al., 2001).

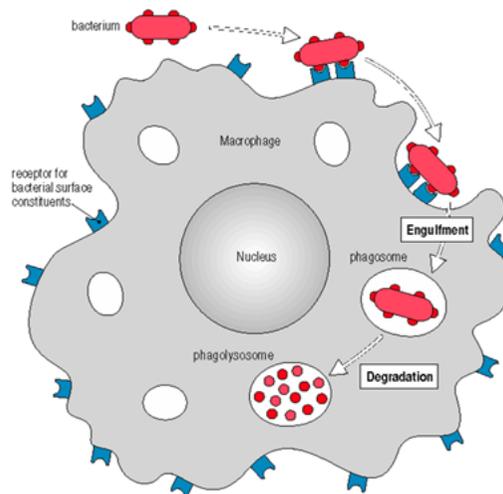


Figure 1.9 Phagocytosis mechanism as starting point of the innate immunity.

Compared to the innate immunity, specific immunity (or acquired or adaptive immunity) represents a more sophisticated system of defence. A peculiarity of this second line of defence is the ability to discriminate between exogenous and autologous antigens (Janeway et al., 2001).

Therefore, acquired immunity strengthens the role of the innate immune response, amplifying, focusing and directing this response against foreign agents. It is possible to schematically distinguish two types of specific immunity: humoral and cell-mediated immunity (Figure 1.10). The main components of acquired immunity are two other types of white blood cells that are called lymphocytes, which are divided into B lymphocytes (expression of humoral immunity and developed in the bone marrow) and T lymphocytes (expression of cell-mediated immunity and developed in the thymus). The acquired immune response consists in a first step of antigen-processing by macrophages and dendritic cells (termed “antigen presenting cells” or APC) and a second step of presentation to T and B lymphocytes which ensure the recognition by specific receptors (TCR and BCR, T and B cell receptors) expressed on their cell membrane. The binding of the antigen to these receptors determines the activation, proliferation and differentiation into effector cells by a process of clonal selection. In details, in the humoral immune response, activated B cells (or plasma cells) secrete antibodies (or immunoglobulins) that circulate in the bloodstream and permeate other body fluids, where they bind specifically to the foreign antigen that stimulated their production, causing its destruction (Figure 1.10). In cell-mediated immune response, some T cells, termed cytotoxic, attack the leukocytes which present the antigenic determinants complexed to proteins of the major histocompatibility complex of class I (MHC I), and destroy them (Figure 1.10).

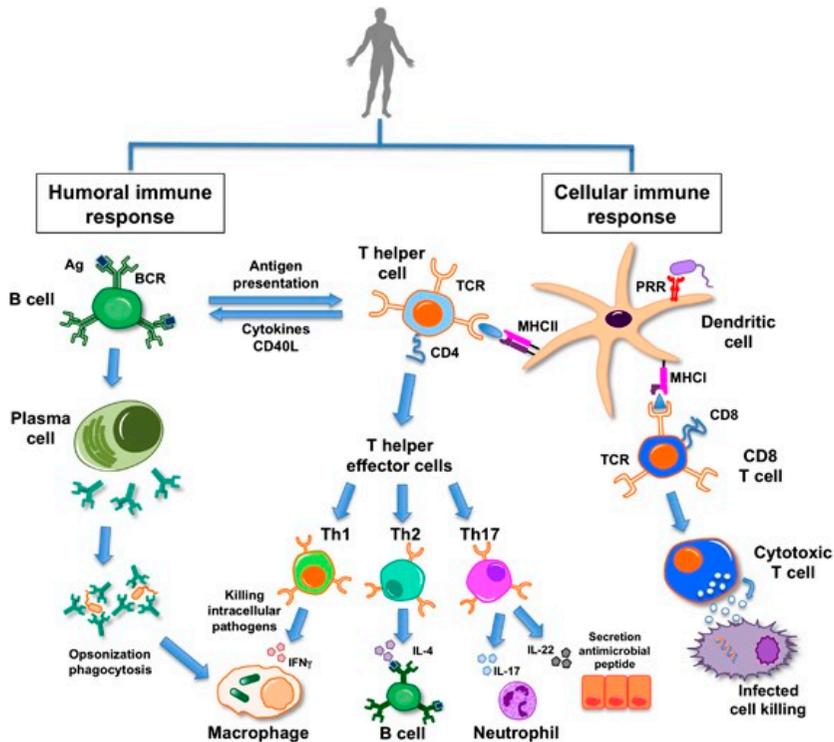


Figure 1.10 The humoral and cell-mediated immune responses.

Other T cells, the helper T immature cells, recognise the complex formed between the antigen and MHC class II proteins, and produce cytokines, known as macrophage colony-stimulating factor and gamma interferon, which promote the activity of macrophages. Moreover, the helper T cells secrete interleukin-2, which stimulates the proliferation of abovementioned cytotoxic T cells (Alberts et al., 2007). Following expulsion of the antigen, the immune response gradually comes to an end. Some cells, produced during clonal expansion, undergo a phase of differentiation into memory cells, which are capable, in case of a possible new encounter with the same antigen, to respond more quickly and with higher efficiency (Alberts et al., 2007).

1.4.2 LPS and innate immunity

As described in the above paragraph 1.4.1, innate immune response is activated after recognition of bacterial conservative molecular targets termed PAMPs, by specific innate immune receptors, termed PRRs (Janeway, 1989). PAMPs have three common features that make them ideal targets for innate immune recognition (Medzhitov, 2001):

- are produced only by microbes and not by host
- are invariant between microorganisms of a given class
- are essential for microbial survival

In case of Gram-negative bacteria, lipopolysaccharides respond to all aforesaid three features, being recognised as PAMPs by a component of a family of PRRs known as Toll-Like Receptors (TLRs). Upon recognition of the real endotoxic principle of LPS, the lipid A moiety, the host innate immune system is immediately activated. The core oligosaccharide or, in smooth bacterial colonies, the O-chain, are the antigenic determinants recognised by the adaptative immunity. TLRs comprise a family of ten transmembrane receptors in humans, which are characterised by an extracellular leucine-rich repeat (LRR) domain and an intracellular Toll/IL-1 receptor (TIR) domain (Hashimoto et al., 1988; Medzhitov et al., 1997). Among all TLRs, TLR4 is the receptor designated to recognise the LPS macromolecule, but its detection requires several accessory molecules. The first host protein involved in LPS recognition is a serum protein, the LBP (LPS- binding protein) that forms a high affinity complex with lipid A released from bacterial lysis or replication, or still bound to the outer membrane of the intact microbial cell (Pålsson-McDermott and O'Neill, 2004). The N-terminal domain of

LBP is responsible for lipid A binding while the C-terminal domain is specific for further recognition by CD14 (Figure 1.11). This latter can be found in two forms, soluble (secreted into serum) and membrane bound (as glycosphosphoinositol-linked protein on the surface of macrophages) (Haziot et al., 1996), and has a crucial role in the enhancement of host response to endotoxins since it facilitates the recognition of LPS by the final receptor complex TLR4/MD-2. It was demonstrated that CD14-deficient mice have a profound defect in responsiveness to LPS, thus showing the pivotal role of CD14 in LPS-recognition process (Haziot et al., 1996). Once CD14 binds to the LPS, it has been shown that CD14 can chaperone LPS from LBP to TLR4 at the cell surface. This latter is linked, through its extracellular domain, to a small glycoprotein, MD-2 (myeloid differentiation factor 2), which turns out to be absolutely necessary for the proper functioning of TLR4 receptor; indeed, the same MD-2 protein is able, by itself, to recognise lipid A with a good affinity, and also to discriminate different lipid A types. Therefore, TLR4 might be considered as a "pseudo-receptor", since, in the absence of MD-2, does not recognise its ligand. The TLR4/MD-2 complex has been shown to bind LPS with higher affinity than soluble MD-2 (Akashi et al., 2003). However, the binding affinity of MD-2 or the MD-2/LPS complex for TLR4 is the same (Visintin et al., 2005). TLR4 with the MD-2 protein forms an heterodimer that recognizes a common structural motif in different LPS molecules. The binding of LPS to TLR4/MD-2 receptorial complex can activate two different immune response pathways (Figure 1.11) (Park et al., 2009).

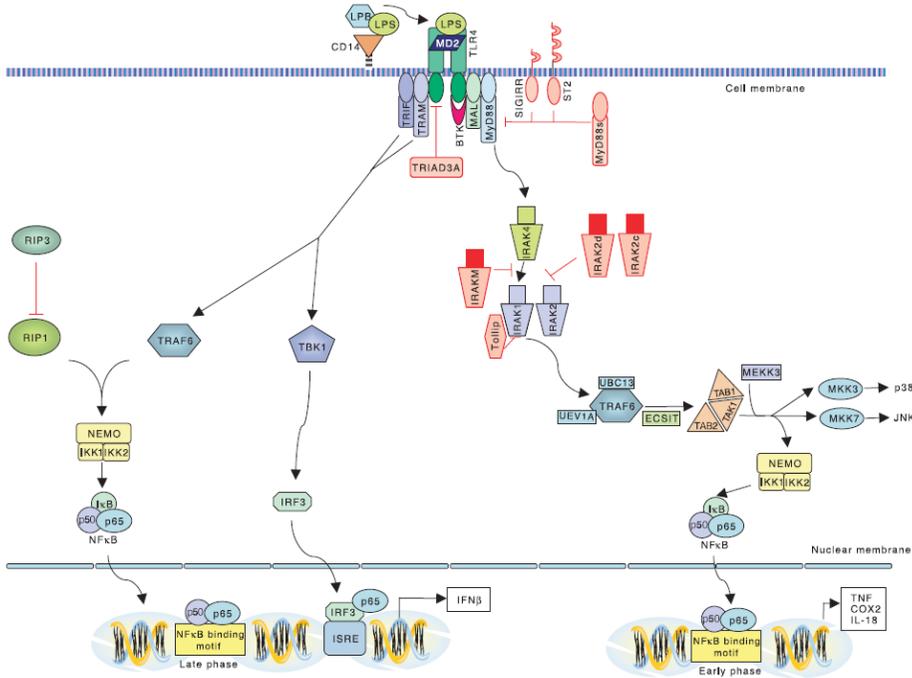


Figure 1.11 Schematized model of LPS signaling.

In the first case, it results in activation of the transcription factor NF- κ B, which regulates the expression of genes encoding for inflammatory cytokines; whereas the second pathway results in activation of MAP-kinases which regulate the transcription of genes involved in increasing the stability of particular regions of the mRNA. In both cases, the final result is the amplification of the transduction signal with the consequent massive production of inflammatory proteins thus eliciting the inflammatory process (Figure 1.11) (Park et al., 2009). If the inflammatory response is amplified and uncontrolled, due to the high toxicity of the LPS from the infecting pathogen, there may occur a fulminating septic shock syndrome.

References

- Adinolfi M. et al., *Carbohydr. Res.*, **1996**, *284*, 119-133.
- Akashi S. et al., *J. Exp. Med.*, **2003**, *198*, 1035-1042.
- Alberts B. et al., *Molecular Biology of the Cell*, 4th edition (Ed.: Garland Science), New York, chapter 24, **2007**.
- Alexander C., Rietschel E.T. *J. Endotoxin Res.*, **2001**, *7*, 167-202.
- Brandenburg K. et al., *Eur. J. Biochem.*, **1993**, *218*, 555-563.
- Cardoso L.S., et al., *Microb. Cell Fact.*, **2007**, *6*, 1.
- De Soyza A. et al., *Innate Immun.*, **2008**, *14*, 127-144.
- Fukuoka S. et al., *Biochim. Biophys. Acta.*, **2001**, *1510*, 185-197.
- Golenbock D.T. et al., *J. Biol. Chem.*, **1991**, *266*, 19490-19498.
- Hamad M.A. et al., *Mol. Microbiol.*, **2012**, *5*, 962-974.
- Hashimoto C. et al., *Cell*, **1988**, *52*, 269-279.
- Haziot A. et al., *Immunity*, **1996**, *4*, 407-414.
- Hellerqvist C.G. et al., *Acta Chem. Scand.*, **1971**, *25*, 955-961.
- Hellerqvist C.G. et al., *Carbohydr. Res.*, **1969**, *9*, 237-241.
- Holst O. et al., *FEMS Immunol. Med. Microbiol.*, **1996**, *16*, 83-104.
- Janeway C.A. Jr. *Cold Spring Harb. Symp. Quant. Biol.*, **1989**, *54*, 1-13.
- Janeway C.A. Jr. et al., *Immunobiology*, 5th edition (Ed.: Garland Science), New York, **2001**.
- Kittelberger R., Hilbink F., *J. Biochem. Biophys. Methods*, **1993**, *26*, 81-86.
- Lüderitz O. et al., *Ann. NY Acad. Sci.*, **1966**, *133*, 347-349.
- Medzhitov R., *Nat. Rev. Immunol.*, **2001**, *1*, 135-45.
- Medzhitov R., Preston-Hurlburt P., Janeway C.A. Jr., *Nature*,

1997, 388, 394- 397.

- Mesbah N.M., Wiegel J., *Ann. NY Acad. Sci.*, **2008**, *1125*, 44-57.
- Netea M.G. et al., *Trends Immunol.*, **2002**, *23*, 135-139.
- Oikawa M. et al., *Org. Biomol. Chem.*, **2004**, *2*, 3557- 3565.
- Pålsson-McDermott E.M., O'Neill L.A., *Immunology*, **2004**, *113*,153-162.
- Park B.S. et al., *Nature*, **2009**, *458*, 1191-5.
- Pommerville J.C., Alcamo's fundamental of microbiology, 9th Edition, (Eds.: Jones & Bartlett Publishers), Burlington, Chapter 4, **2010**.
- Raetz C.R., *Annu. Rev. Biochem.*, **1990**, *59*, 129-70.
- Raetz C.R., Whitfield C., *Annu. Rev. Biochem.*, **2002**, *71*, 635-700.
- Raven P.H., Johnson G.B., *Biology* 9th Edition, (Ed.: McGraw Hill), Chapter 34, **2011**.
- Rietschel E.T. et al., *FASEB J.*, **1994**, *8*, 217-225.
- Seydel U. et al., *Eur. J. Biochem.*, **2000**, *267*, 3032-3039.
- Silipo A. et al., Prokaryotic cell wall compounds structure and biochemistry (Eds. Konig H., Herald C., Varma A.) Springer-Verlag, Berlin, Germany, 133-154, **2010**.
- Staley J.T. et al., *Microbial Life* 2nd Edition, chapter 4, **2007**.
- Visintin A. et al., *J. Immunol.*, **2005**, *175*, 6465-6472.
- Zahringer U., Lindner B., Rietschel E.T., Endotoxin in Health and Disease (Eds.: H. Brade, S. M. Opal, S. N. Vogel, D. C. Morrison), Marcel Dekker, New York, Basel, 93- 114, **1999**.

Chapter 2: Structural characterization of LPS and LOS

2.1 Extraction and purification of LPS and LOS

The first approach aimed to the structural elucidation of LPS/LOS is represented by their extraction from intact microbial cells. This is conventionally achieved through two complementary procedures, that lead to the selective isolation of R-LPS (LOS) and S-LPS, exploiting their different amphiphilic character; indeed, the latter has higher hydrophilicity, due to the presence of the polysaccharide moiety of the O-chain, whereas the former, which lacks such portion, has a more pronounced hydrophobic character. Thus, these different physico-chemical properties allow to perform a selective extraction of the LOS and LPS, with a high degree of purity. Lyophilized bacterial cells are firstly washed with acetone, water and ethanol to remove cell contaminants. After this initial step of purification, the procedure of extraction can take place. Typically, rough-type LPSs are extracted with a phenol-chloroform-petroleum ether procedure (PCP) (Galanos et al., 1969), while smooth-type LPSs are extracted with the hot phenol-water protocol (Westphal and Jann, 1965). The PCP extraction consists in the treatment of dried bacterial cells with phenol/chloroform/light petroleum ether in proportions 5:8:2 v/v/v; LOS is precipitated from pure phenol adding drops of water. After this extraction, cells undergo a second treatment with 1:1 v/v 90% phenol/water at 68°C. The two phases are then extensively dialysed and liophilised. As previously mentioned, typically the presence of the long O-chain moiety increases the hydrophilic nature of endotoxins, thus LPS molecules are extracted in

the water phase even though several factors, as the occurrence within the repeating unit of hydrophobic groups as well as charged monosaccharides or the length of the polysaccharide chains, may modulate LPS solubility in water. Thus, it can be detected either in the water or in the phenol extract.

After the extraction procedure, enzymatic treatments to remove proteins and nucleic acids are executed. The typical protocol consists in LPS/LOS purification with nucleases (DNAse and RNAse) and protease followed by dialysis in order to remove digested material. The screening for detection of LPS is realised by SDS polyacrylamide electrophoresis gel (SDS-PAGE) followed by silver nitrate staining which allows to define the typology of the extracted material and its purification degree (Kittelberger and Hilbink, 1993). The presence of LPS is testified by the observation of the typical “ladder-like” pattern of the electrophoretic migration, due to the presence of molecules differing in their structure for the number of repeating units composing the O-polysaccharide moiety. Conversely, LOS appears as a dark band that quickly migrates at the bottom of the gel, denoting the lower molecular weight of LOS lacking O-chain domain.

2.2 Structural determination of the saccharide moities

Structural investigation of LPS/LOS is a difficult task since their amphiphilic nature determines the tendency to form micelles with low solubility in both aqueous and apolar organic solvents. Several strategies have been developed to overcome this obstacle, including the usage of solvent mixtures and chemical modifications or degradation of the molecules to improve dissolution in simple solvent systems. The

common protocol consists in the analysis of the lipid portion and saccharide moiety separately by using several hydrolysis techniques, as aqueous 1% acetic acid or acetate buffer. The above mild conditions are sufficient to cleave selectively the acid-labile glycosidic linkage between the Kdo residue and the non-reducing glucosamine of the lipid A, thus releasing the oligo-/polysaccharide moiety in the water solution.

This linkage is very labile due to:

- I. the absence of a whichever electron withdrawing group at position adjacent to anomeric (*C*-3 in this case since anomeric is *C*-2) that favors the formation of the reaction intermediate oxonium ion;
- II. the passage from chair to half-chair conformation during the formation of the oxonium ion fastened by the presence of non-substituted carbons;
- III. the presence of a hydroxyl group in axial configuration at *C*-5 that contributes to a steric energy release in the formation of the intermediate ion.

The insoluble lipid A moiety can then be recovered by centrifugation and several washes with distilled water, whereas the supernatant is constituted by O-antigen still linked to core region. This does not complicate the structural investigation of the O-antigen since, in terms of monosaccharide residues number, the contribute of the core fraction is negligible respect to the O-chain. A disadvantage of the above described methodology, based on an acid treatment, is that it produces a Kdo reducing unit with a structural microheterogeneity (α and β anomers of pyranose and furanose rings, condensed or anhydro forms) that can

render the study of the core oligosaccharide of LOS by NMR spectroscopy particularly difficult. Thus, typically, the procedure to determine the primary structure of the core oligosaccharide portion considers the de-lipidation of the intact LOS by means of alkaline treatment in two steps: first, a mild hydrazinolysis in anhydrous conditions to remove ester-linked fatty acids from lipid A; second, a strong alkaline hydrolysis is performed to obtain the cleavage of amide-linked acyl chains too (Holst, 2000). The obtained product is represented by a homogenous phosphorylated oligosaccharide, comprehensive of the lipid A disaccharide backbone, that can be easily studied through chemical and spectroscopical investigations. Obviously, base-labile substituents of glycidic or non-glycidic nature are lost during this procedure. Therefore, it is possible to use the acid approach as a complement to the alkaline degradation, in order to evaluate the occurrence of base-labile groups lost during the alkaline treatment (i.e. pyrophosphate groups).

The primary structure determination of the saccharide portions so far obtained currently exploits state-of-art NMR experiments and soft-ionisation mass spectrometry techniques, together with compositional and linkage chemical analyses that allow the determination of:

- the quali-quantitative composition of sugar residues
- the absolute configuration of each monosaccharide
- the cyclisation ring sizes
- the glycosylation points of each monosaccharide
- the anomeric configuration of the glycosidic linkages
- the monosaccharide sequences
- the location of non-carbohydrates substituents

Additional information on the nature and the sequence of monosaccharides can be obtained from ad hoc selective cleavages of the polysaccharide such as partial acid hydrolysis, acetolysis, Smith degradation, β -elimination and solvolysis.

2.2.1 Chemical analyses of the saccharide moieties

Chemical analyses are useful to obtain important preliminary information regarding primary structure of poly/oligosaccharides. The procedure that provides the usage of Gas-Chromatography coupled with Mass-Spectrometry (GC-MS), can only be realized after conversion of the monosaccharides into apolar and volatile derivatives. There is a plethora of protocols that can be employed to identify the monosaccharide type as well as the glycosylation positions, thus the appropriate choice of the derivatisation method can highlight specific features of the sugar residues within the native LPS/LOS structure. Typically, the qualitative analysis by GC-MS is achieved by treating the oligo- /polysaccharide with MeOH/HCl leading to solvolysis of the molecule and to the formation of the *O*-methyl glycoside for each monosaccharide. Subsequent acetylation with acetic anhydride in pyridine produces the per-acetylated *O*-methyl glycosides (AMG), that can be injected and analysed by GC-MS. Comparison of the retention times from the GC analysis and the fragmentation pattern from the MS spectra with opportune standards allows to the identification of the monosaccharide residues. Quantification analysis can then be obtained by using an internal standard, that usually is per-acetylated inositol. Since the AMG protocol provides a solvolysis in acid conditions, several isomers for each monosaccharide may form (i.e. pyranose and furanose,

both as α and β anomers) resulting in the occurrence of many peaks in the corresponding chromatogram. Although not severely influencing the identification of the monosaccharides, this can lead to their misquantification. Therefore, an alternative and complementary approach that can be used leads to the formation of acetylated alditol derivatives. In this case, after strong acid hydrolysis with trifluoroacetic Acid (TFA), the carbonyl moiety of the monosaccharide residues is reduced with NaBH_4 , thus providing a single molecule from each monosaccharide. The alditols are then acetylated and finally analysed via gas-chromatographic techniques.

A further procedure of chemical analysis that delivers the distinction between enantiomers provides that the solvolysis is executed with an enantiomerically pure alcohol as 2-(+)-octanol or 2-(+)-butanol. Thus, the absolute configuration of each monosaccharide residues can be determined (Leontein and Lönngren, 1978). Indeed, after acetylation and injection to GC-MS, the comparison of the retention time of the peracetylated 2-(+)-octyl glycosides with the ones of a mixture of peracetylated 2-(\pm)-octyl glycosides of standard monoses in D or L configuration allows the assignment of the monosaccharide configuration.

The determination of ring size and glycosylation site of the monosaccharides is a further information obtainable through GC-MS analysis. The procedure consists in an extensively *O*-methylation of the oligo/polysaccharide with CH_3I in strongly alkaline conditions. Then, the permethylated oligo-polysaccharide is hydrolysed in acid conditions and reduced with a marked boro hydride (NaBD_4). The alditols so obtained have free hydroxyl groups only at the positions previously involved in glycosidic linkages and cyclisation, and can be then acetylated. These

partially methylated acetylated alditols can be analysed by GC-MS and the fragments observed in the MS spectra are diagnostic for specific substitution patterns of acetyl and methoxyl groups, since molecule fragmentations preferably occur in correspondence of a methoxyl group, better sustaining the released positive charge because of resonance effects. To obtain the desired structure information it is necessary to consider that the position of the acetyl groups in the fragments accounts for the attachment point or for the position of cyclisation of the pyranose or furanose ring whereas the methyl groups are located on positions not involved in linkages. The reduction of the carbonilic function with sodium borodeuteride discriminates between the fragments originated from the reduced position (even masses) and those originated from the last position (odd masses) (Hakomori, 1964). Information obtained from these chemical analyses helps and confirms the interpretation of subsequent NMR and MS experiments.

2.2.2 NMR spectroscopy analysis of the saccharide moieties

Nuclear Magnetic Resonance (NMR) represents the most useful and decisive technique in the field of structural investigation of carbohydrates. This is due to the possibility of analysing molecules in solution in a native state, owing to the good solubility observed for oligo- and polysaccharides in aqueous solutions. During the structural study of saccharides, nuclei usually observed are ^1H , ^{13}C and ^{31}P , the last being necessary to detect phosphate groups and unusual phosphorous-containing substituents.

Analysing a typical monodimensional ^1H -NMR spectrum of an oligo-/polysaccharide, it is possible to distinguish three different regions:

- The region between 5.5 and 4.6 ppm typically related to the anomeric protons signals
- The region between 4.6 and 2.6 ppm where the ring proton signals are located
- The region between 2.5 and 1.0 ppm that is typical of the deoxy position signals

Furthermore, from a ^{13}C NMR spectrum it is possible to obtain information about:

- 95-105 ppm: anomeric carbon atoms involved in a glycosidic linkage
- 80-60 ppm: oxymethylene or carbinolic carbon atoms
- 60-45 ppm: carbon atoms linked to nitrogen
- ~ 30 ppm: aliphatic methylene carbons of deoxy-sugar residues
- 25-15 ppm: methylene carbon atoms of 6-deoxy-sugar residues and of acetyl substituents.

The region relative to the ring proton signals is considerably narrow, and usually the identification of all chemical shifts is not possible on the basis of the sole one dimensional NMR analysis. Thus, a combination of homo- and hetero-nuclear 2D- NMR experiments (such as DQF-COSY, TOCSY, NOESY, $^1\text{H},^{31}\text{P}$ - HSQC, $^1\text{H},^{13}\text{C}$ -HSQC, $^1\text{H},^{13}\text{C}$ -HMBC) are required in order to assign all the spin systems, to determine the location and the nature of non-carbohydrate substituents and to characterise the monosaccharide sequence. In detail, TOCSY and DQF-COSY spectra allow the correct identification and assignment of all ring protons, and on the basis of these data, the assignment of all ^{13}C resonances usually

follows from the analysis of the $^1\text{H}, ^{13}\text{C}$ -HSQC spectrum; the intra-residue pattern of dipolar correlations in the 2D NOESY and ROESY spectra gives the proof for specific configurations; as example, NOE correlations between H-1/H-3 and H-5 are univocally diagnostic for a β -*gluco*-configuration, on the other hand a NOE connectivity between H-1 and only H-2 is indicative of a α -*gluco*-configuration. In general, proton orientations on the rings can be deduced from the $^3J_{\text{H,H}}$ coupling constant values that follow the Karplus law and assume high values ($8 \div 10$ Hz) in case of *trans*-diaxial orientation of vicinal protons and considerably lower values (< 4 Hz) in case of equatorial/axial or equatorial/equatorial orientation. These values can be attained from DQF-COSY spectrum. Anomeric proton and carbon chemical shifts are the first suggestion for the determination of the anomeric configuration of each monomer, since usually α -configured proton signals appear at lower fields respect to corresponding β -anomers, whereas the opposite situation occurs for carbon chemical shifts. Anomeric configurations are then confirmed by the $^3J_{\text{H1,H2}}$ and by $^1J_{\text{C1,H2}}$ values. In detail, in monosaccharides with the H-2 in axial orientation (as glucose, galactose), a $^3J_{\text{H1,H2}}$ around 8 Hz is indicative of a β - orientation, whereas below 3 Hz of an α -configuration. Sugars with the H-2 in equatorial orientation (as mannose) show in both anomeric configuration a $^3J_{\text{H1,H2}}$ value below 3 Hz. Moreover, a $^1J_{\text{C1,H1}}$ below 165 indicates a β -anomer, above 170 Hz a α -anomer. The downfield shift of ring proton signals (acylation shift) is useful to determine *O*- and *N*-acylation sites. *O*- phosphorylated protons undergo a similar downfield shift. In the $^1\text{H}, ^{13}\text{C}$ -HSQC spectrum, the down-field shift, or glycosylation shift, of carbon resonances ($2 \div 10$ ppm), allows to locate the positions of glycosidation and these may find confirmation in information previously obtained by means of chemical analyses.

Information regarding the cyclisation ring is obtained by the observation in the $^1\text{H},^{13}\text{C}$ -HMBC spectrum of intra-residual long range correlations, i.e. H-1/C-5 and C-1/H-5 for pyranosidic rings or H-1/C-4 and C-1/H-4 for furanosidic rings. These latter are also identified by a low field displacement of all the ring resonances for the monosaccharide. Information concerning the sequence of the monosaccharides within the oligosaccharide (or the repeating unit of the polysaccharide) are then deduced from the observation of the inter-residual dipolar correlations in 2D NOESY and ROESY spectra (Figure 2.1) and from the existence of scalar long range correlations, in the $^1\text{H},^{13}\text{C}$ -HMBC spectrum, among the ^1H and ^{13}C nuclei involved in glycosidic linkages. Eventually, $^1\text{H},^{31}\text{P}$ -HSQC allows the localisation of the phosphorylation sites within the saccharide structure.

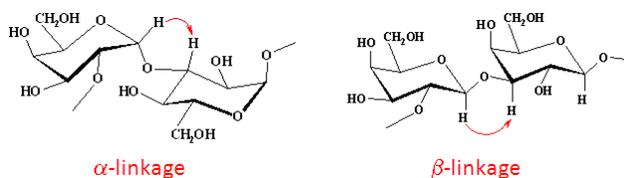


Figure 2.1 Inter-residue NOE correlations between disaccharides

2.2.3 Mass spectrometry analysis of the saccharide moieties

Mass spectrometry represents to date an essential technique in the structural characterisation of lipopolysaccharides and lipooligosaccharides especially with the development of soft ionisation techniques like MALDI (Matrix Assisted Laser Desorption Ionisation), ESI (ElectroSpray Ionisation), and the fragment analysis in MS/MS experiments. Apart from the Electron ionisation (EI) normally used in GC-MS analyses, mass spectrometric techniques allowing the definition

of the molecular weight of the intact molecule are particularly useful, since they can provide additional information to the 1D and 2D NMR analyses that sometimes can be hard to construe. Usually, for carbohydrates, MALDI and ESI MS spectra are performed in negative ion mode, due to the presence of a great number of potentially ionising groups (carboxyl or phosphate functions), nevertheless spectra in positive ion mode are also performed

MALDI-TOF analysis, in particular, gives an ideal profile of the sample by sketching all the molecules composing it, thus giving a picture of the eventual presence of different glycoforms differing for the presence of one or more glycidic or non-glycidic residues. This methodology consists in a soft ionisation of the sample that is suspended in a specific matrix assisting the analyte in ionisation phenomenon. The matrix has important roles among which the protection of the sample from the laser energy assuring a soft ionisation with formation of single charged ions. Recently it has been pioneered the birth of new MALDI MS approaches through development of new methods of preparation and experimental study of the intact sample yielding a “picture” of the whole endotoxin, without any loss of structural information that can arise when the MS analysis is executed on degraded derivatives provided for the standard structural elucidation protocols (Sturiale et al., 2005). These procedures provide good spectra of intact molecular ions of lipopolysaccharides comprehensive of the heterogeneity inferred by the lipid A presence, showing high sensitivity also at relatively high molecular mass. The usage of high energy laser source allows to cleave the labile linkage between the Kdo and the glucosamine backbone lipid A, in a similar way to the acetate buffer treatment (see paragraph 2.2). In case of LOS, in the deriving spectra there are three regions of signals originated by the lipid

A, the core oligosaccharide and the intact LOS. In LPS, due to the dispersion of molecular weight of the O-polysaccharide, information on the size of the repeating unit can be deduced.

2.3 Structural determination of the lipid A moiety

The lipid fraction obtained by precipitation, after mild acid treatment of LPS (or LOS) consists of a mixture of intrinsically heterogeneous lipid A molecules. The structural approach makes an extensive use of mass spectrometry analysis by MALDI-TOF and ESI MS techniques, as well as NMR spectroscopy and chemical analyses both on the native and selectively degraded lipid A fractions. The lipid A structure characterisation requires the determination of:

- the saccharide backbone
- the amide and the ester linked fatty acids
- the distribution of the acyl chains on the sugar backbone
- the phosphorylation sites
- the sites of heterogeneity relative to the acylation pattern (number, type and distribution of the acyl chains) or the non-stoichiometric presence of phosphate substituents.

The determination of the nature, ring size, attachment point and absolute configuration of the sugar components of lipid A is performed as already described (see paragraph 2.2.1).

2.3.1 Chemical analyses of the lipid A moiety

Fatty acids determination is usually achieved through GC-MS analysis of their methyl ester derivatives. Lipid A is first treated with aqueous HCl,

then neutralised with NaOH, and acyl chains are then extracted in CHCl₃, methylated with diazomethane and analysed by GC-MS. *O*-linked acyl chains can be extracted from the entire molecule through an alkaline treatment with NaOH, followed by methylation with diazomethane and analysis by GC-MS. The choice of cleavage conditions allows the selective detection of the linkage type of the fatty acids (Rietschel, 1976).

2.3.2 Mass spectrometry analysis of the lipid A moiety

Mass spectrometry is the most common approach used in lipid A structural determination. MALDI-TOF and ESI MS data allow gaining insights into the number of lipid A species present in the fraction, the presence of polar heads and the distribution of acyl residues on each GlcN units of the disaccharide backbone (Que et al., 2000). Novel general approaches gaining structural insights on the native and selectively degraded lipid A fractions have been developed. In particular, combining MALDI MS or ESI MS and selective chemical degradations it is possible to understand the acylation profile of the molecule. Firstly, the amide linked fatty acids are identified by MS analysis of completely de-*O*-acylated lipid A with anhydrous hydrazine (Holst, 2000). An enormous advantage in lipid A structural analysis is the treatment with aqueous NH₄OH (Silipo et al., 2002). This procedure selectively cleaves the ester-linked acyloxyacyl groups, leaving untouched the amide-linked acyloxyacyl groups. Thus, MS analysis of the ammonium hydroxide treated lipid A allows the detection of the secondary fatty acid distribution, which is one of most demanding issues in the structural determination of the LPS/LOS lipidic moiety (Sforza et al., 2004; Sturiale et al., 2005). MALDI spectra can be executed in positive and

negative mode. The use of the former is expedient to differentiate the fatty acid allocation between the two GlcNs of lipid A in case of asymmetric distribution; this is possible since during the execution of positive ion MALDI MS spectra a selective ionisation after fragmentation of the GlcN backbone occurs, and the only oxonium ion that is visible is relative to the non-reducing GlcN (Domon and Costello, 1988). On the other hand, both negative and positive mass spectra furnish indication of the presence of non-stoichiometric phosphorylation sites. The intact or fully de-*O*-deacylated lipid A are also used to study the location of the phosphorylation sites.

References

- Domon B., Costello C.E., *Biochemistry*, **1988**, *27*, 1534-1543.
- Galanos C., Luderitz O., Westphal O., *Eur. J. Biochem.*, **1969**, *9*, 245-249.
- Hakomori S., *J. Biochem.*, **1964**, *55*, 205-208.
- Holst O., *Methods Mol. Biol.*, **2000**, *145*, 345 -353.
- Kittelberger R., Hilbink F., *J. Biochem. Biophys. Methods*, **1993**, *26*, 81-86.
- Leontein K., Lönnngren J., *Methods Carbohydr. Chem.*, **1978**, *62*, 359-362.
- Que N.L.S. et al., *J. Biol. Chem.*, **2000**, *275*, 28006- 28016.
- Rietschel E.T., *Eur. J. Biochem.*, **1976**, *64*, 423-428.
- Sforza S. et al., *J. Mass Spectrom.*, **2004**, *39*, 378-83.
- Silipo A. et al., *J. Lipid Res.*, **2002**, *43*, 2188 - 2195.
- Sturiale L. et al., *Rapid Commun. in Mass Spectrom.*, **2005**, *19*, 1829- 1834.
- Westphal O., Jann K., *Methods Carbohydr. Chem.*, **1965**, *5*, 83-91.

SECTION II

Lipid As isolated from Extremophiles

Chapter 3: Extremophiles, an overview

3.1 Life-Forms in inhospitable environments

Belonging to both two dominia of Archaea and Bacteria, and to either Gram-positive and Gram-negative family, extremophiles are life-forms capable to live and thrive in inhospitable, from a human viewpoint, environments such as alkaline and acidic waters, boiling hot springs, high pressure waters, ultra saline brines, without excluding a combination of aforesaid chemical and physical extremes which is typical of polyextremophilic microbes (Mesbah and Wiegel, 2008). On the basis of the physical or chemical parameters, these microorganisms can be distinguished in: acidophiles and alkalophiles that live at extreme pH values; thermophiles that are able to live at extremely high temperatures, ranging from 40°C up to 100°C; psychrophiles able to survive in frozen habitats, i.e. glaciers or polar seas; halophiles that live in extreme salt concentrations from 3M to saturation of NaCl; barophiles that require elevated pressure for proliferation and xerophiles that can be isolated from extremely anhydrous environments. Several studies of extremophiles have helped redraw the three-domain system of life. Indeed, the typical dogma held that living organisms could be grouped into two domains: Bacteria, with a simple structural architecture, and Eukarya, whose cells are more complex. The discovery of microorganisms in environments that were assumed to be sterile supported to the once heretical proposal that yet a third group, the Archaea, exists. Many microbiologists believe that ancestors of today's archeal species might represent a type of organism that first inhabited Earth when it was a young and hot planet. These unique features led the American microbiologist Carl Woese to propose these microorganisms

be lumped together and called Archeobacteria, as already described in Section I, Chapter 1. Anatomically, archaeans differ from bacteria since their cell membrane is composed by isoprenoid and isoprenyl glycerol ether-linked lipids, with fully saturated acyl chains (Gulik et al., 1985). The presence of these lipids and the production of exopolysaccharides allowing the formation of biofilms, are responsible for the extraordinary stability shown by extremophiles. The notable resistance offered by these components to hydrolysis at acidic and alkaline pHs and at high temperature, has been proposed as the feature that allowed the development of archaeal extremophile strains. Recently, researchers have focused on DNA and completely sequenced a number of bacterial and archaeal genomes; noticeably, these studies have demonstrated that some genes of archaea are similar to bacterial genes and others to eukaryotic ones. On the other hand, a large fraction of archaeal genes appear to be unique. These unshared genes establish archaea's separate identity. Owing to their unique features, extremophiles have attracted much interest since they showed to possess a high biotechnological potential with a number of important industrial applications comprising bioremediation and biomining processes (Taylor et al., 2012). Of particular interest are the enzymes that help extremophiles to function in hostile circumstances. The commonly used enzymes degrade or stop working when exposed to heat or other extreme conditions, and so manufacturers that rely on them must take specific steps to protect the molecules during reactions or storage.

Since enzymes from extremophiles - termed "extremozymes" - remain active when other enzymes would fail, they can potentially remove the need for those steps, thereby increasing efficiency and reducing costs

(Madigan and Mairs, 1997). Furthermore, they can form the basis of entirely new enzyme-based processes (Madigan and Mairs, 1997).

3.2 Psychrophiles

Much of life on Earth has evolved to colonize low-temperature environments. In fact, at temperatures permanently below 5°C, the cold biosphere represents by far the largest fraction of the global biosphere (Siddiqui et al., 2013). Permanently cold environments cover about one-fifth of the surface of the Earth including terrestrial ice sheets of Antarctica and Greenland, mountain glaciers and sea ice in Antarctic and Arctic Ocean. Sea ice covers a larger area of the Earth than glaciers, approaching 10% of the ocean's surface, but with an average thickness of only 2-3 meter, the global volume of sea ice is much lower than that of the glacial environment. In these environments, the temperature plays a critical role in the selection and survival of a variety of organisms such as bacteria, archaea, yeasts, algae, insects, fishes and plants. These cold-adapted organisms are termed psychrophiles. The term psychrophile (from Greek *psychros* meaning cold and *philes* meaning loving) was first used by Schmidt-Nielsen in 1902 (Morita 1975). For many years, cold-adapted organisms have been classified into two partially overlapping groups: psychrophiles, having minimum, optimum and maximum growth temperature at or below 0°C, 15°C and 20°C, respectively, and psychrotolerant, with growth maximum temperature above 25°C, but with the capacity to grow at very low temperatures (Morita, 1975). Recently, among the additional definition proposed, such as "moderate psychrophiles" (Helmke et al., 2004), and "psychro-active" (Laucks et al., 2005), the terms "eurypsychrophiles" and "stenopsychrophiles" have been favored (Feller et al., 2003; Bakermans et al., 2004; Cavicchioli, 2006).

Stenopsychrophile, formerly "true psychrophile", refers to a microorganism with a restricted growth-temperature range; in fact, they are most frequently isolated from thermally stable cold marine environments. Eurypsychrophile, formerly psychrotolerant, describes a microorganism that prefers permanently cold environment, but can also tolerate a wide range of temperatures reaching up into the mesophilic range; these microbes dominate, in fact, cold terrestrial ecosystems (Helmke et al., 2004). The genomic differences in cold adaptation between stenopsychrophiles and eurypsychrophiles have yet to be thoroughly investigated (Bakermans et al., 2004). However, this distinction is ignored at present and all the cold-tolerant bacteria could be dubbed as psychrophiles (Chattopadhyay et al., 2014).

Cold adaptation of microorganisms is the result of intrinsic genome-wide changes that facilitate the growth at low temperatures. Comparative studies of cold-adapted microorganisms are beginning to reveal which adaptations are common to all the psychrophiles and which are specific to the particular environment each psychrophile inhabits. Cold marine environments are distinctly different from cold terrestrial ones like permafrost; marine environments tend to have high thermal stability as well as stable solute concentrations, while terrestrial environments do not. Consequently, genomic analysis of microorganisms isolated from marine environments likely reflects adaptations to stable low temperatures, while the ones from terrestrial environments may reveal unique mechanisms of cold adaptation related to heterogeneous environmental chemistry due to the extreme temperatures fluctuations.

Ecological limiting factors, such as nutrient and water availability, salinity, pressure, UV irradiation and temperature, are all characteristics of cold environments. Psychrophilic microorganisms in glacial ice are restricted

to small amounts of unfrozen water inside the permafrost soil or the ice, and to brine channels. In some terrestrial habitats, these stresses dictate that psychrophilic communities develop niches, not only in brine channels, but also frequently associated with lithic habitats, such as cryptoendolithic and chasmolithic environments (De Maayer et al., 2014).

3.3 Molecular and physiological adaptation

Cold temperatures impose severe physicochemical constraints on cellular function by negatively influencing cell integrity, water viscosity, solute diffusion rate, membrane fluidity and macromolecular interactions (De Maayer et al., 2014). The ability of an organism to survive and grow in cold conditions is therefore dependent on a various number of adaptive strategies (Figure 3.1). The analysis of numerous psychrophilic genomes and metagenomes has indicated the presence of a large number of features contributing to genome plasticity, such as plasmids, transposable and other mobile genetic elements. One of the most obvious disadvantages of life at very low temperatures is the low rates of catalysis: psychrophilic enzymes must therefore be suitably adapted to maintain adequate catalytic rates for cellular function. These enzymes are generally characterized by higher flexibility and lower thermostability than their mesophilic counterparts. One of the strategies is to reduce the content of amino acids such as arginine and proline; the last one is involved in the formation of multiple hydrogen bonds and salt bridges, and reduces the conformational flexibility. Following a downshift of temperature from 37°C to 10°C in mesophilic *E. coli*, the synthesis of most of the cellular proteins is repressed for a lag period of 4-5 h. Among these proteins, the most prominently up-regulated genes are

those encoding cold-shock proteins (CSPs), a family of small, single-stranded nucleic acid binding proteins that regulate a variety of cellular processes including transcription, translation, protein folding and membrane fluidity (De Maayer et al., 2014). These stress proteins have been found to occur both in mesophilic and psychrophilic bacteria. A second class of proteins, called cold acclimation proteins (Caps), is found to be constitutively expressed in psychrophilic bacteria while in mesophiles are expressed only in response to cold exposure. The CSPs are believed to facilitate transcription and translation at low temperature in the mesophilic bacteria. However, the exact role of CSPs in cold adaptation of psychrophiles is yet to be elucidated (Chattopadhyay 2006).

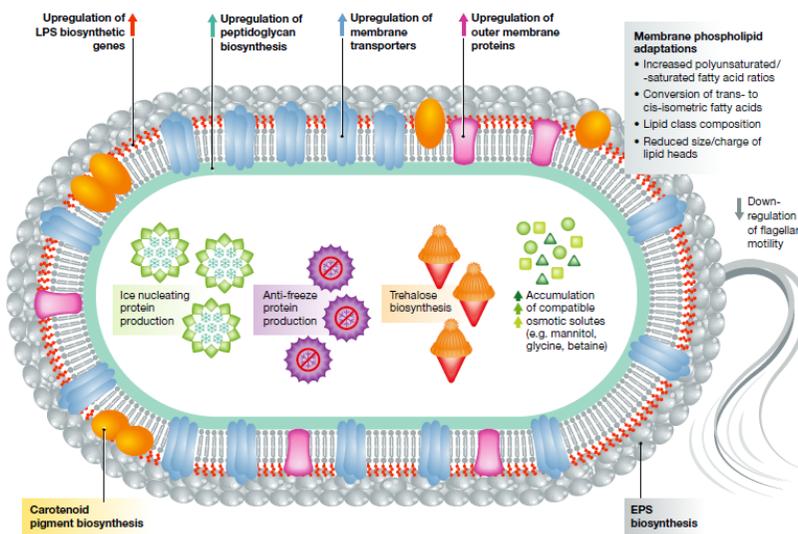


Figure 3.1: Common physiological adaptations in a psychrophilic prokaryote.

Furthermore, the increased solubility of oxygen at low temperatures poses an increased risk from reactive oxygen species (ROS), and different approaches seem to be involved in protecting against free radical damage. *C. psychrerythraea* contains three copies of catalase

genes as well as two different superoxide dismutase (SOD) genes, one of which is a nickel-containing SOD never reported before in proteobacteria (Méthé et al., 2005).

It has long been known that one of the most significant impacts of low temperature is on membrane fluidity, essential for its structural integrity and for the cellular functionality. In fact, recent transcriptome analyses have shown that exposure to cold temperatures induces a rapid up-regulation of genes involved in membrane biogenesis, such as fatty acids, LPS, peptidoglycan and outer membrane proteins biosynthesis (De Maayer et al., 2014). The lipid composition governs the physical properties of the membranes and hence it is not surprising that this varies with the thermal habitat of the microorganism. In general, a reduction in temperature leads to a higher content of unsaturated, polyunsaturated and methyl-branched fatty acids, and/or a shorter acyl-chain length. This altered composition is thought to have a key role in increasing membrane fluidity by introducing steric constraints that change the packing order and reduce the number of interactions in the membrane. Lipopolysaccharides (LPSs) are the major constituent of the bacteria outer membrane that forms a protective barrier around the cell, and their biosynthesis is increased by lowering the temperature. Some of the genes encoding for enzymes involved in the LPS biosynthesis, as lipopolysaccharide biosynthesis polymerase and lipid A biosynthesis acyltransferase (LpxP) are induced (Gao et al., 2006). It is believed that the induction of LpxP is related to the function of enhancing the outer membrane fluidity after cold shock with the acylation of lipid A with palmitoleate instead of laurate (Carty et al., 1999). Carotenoid pigments represent another class of membrane fluidity modulators. Both polar and non-polar pigments are produced by Antarctic bacteria and have

been postulated to buffer membrane fluidity and assist maintaining homeoviscosity during temperature fluctuations (De Maayer et al., 2014). One of the main consequences of lowering temperature is the formation and the propagation of cytoplasmic ice crystals, resulting in the destruction of membranes, gas bubbles formation and organelles disruption, determining a total cellular damage (Fuller 2004). Psychrophilic microorganisms display the expression of compatible solutes biosynthesis - glycine, betaine, sucros, threalose and mannitol- during growth at low temperature. Compatible solutes act globally and not specifically, allowing the microorganisms to occupy broader thermal niches and to react quickly to changes in temperature (Chin et al., 2010). In particular, the threalose disaccharide may prevent the denaturation and aggregations of proteins, scavenge free radicals and stabilize cellular membranes. To contrast low temperatures, some species of fungi, bacteria, plants and insects produce antifreeze proteins (AFP) and glycoproteins (AFGPs), also known as thermal hysteresis proteins. These possess the ability to inhibit the formation of ice and can lower the freezing temperature of a solution non-colligatively, without affecting its melting temperature.

Moreover, it is known that some bacteria produce ice-nucleating agents (INAs) that serve as templates for ice crystallization and provide resistance to desiccation (Margesin et al., 2007). Sea ice through its seasonal lifetime is recently known to hold high concentrations of sugar-based exopolymers in its liquid brine inclusions. These exopolysaccharides (EPSs) are understood to serve as natural cryoprotectants not only against potential ice-crystal damage, but also by further depressing the freezing point of the liquid water remaining available within the ice matrix (Marx et al., 2009). The ability of a

microorganism to surround itself with a highly hydrated shell helps to buffer the cell against the osmotic stress of high salt concentrations, and may provide it with protection against desiccation, and predation by protozoans (Kumar et al 2007). Therefore, EPSs are essential in the aggregate formation, in the mechanism of adhesion to surfaces, in the uptake of nutrient and in the formation of biofilm (Poli et al., 2010). EPS is a common component of biofilm and its production is an important feature of the mature biofilm (Fux et al., 2003). In addition, EPSs form the matrix that embeds the bacteria, where additional free bacteria can be entrapped (Costerton et al.,1995; Bianciotto et al., 2001). The mechanisms of adaptation to cold stress have received considerable attention in the last few decades, particularly for the biotechnological potentiality of these organisms and their biomolecules.

3.4 Industrial and biomedical applications

The scientific interest in cold-adapted molecules has increased substantially in recent years, as evidenced by the growing number of recently isolated examples. The necessity of all components of a microbial cell to adapt to the cold implies that a broad range of cellular products is available for biotechnological applications. The vast majority of studies related to biotechnological applications have focused on cold-adapted enzymes that possess high catalytic activity at low temperature and low thermostability at elevated temperatures. Cold-adapted enzymes are of great interest because their employment could avoid heating and consequently save energy. This is very important, especially in food industry since this could prevent heat-sensitive substrates from chemical degradation. In addition to their enzymes, cold-adapted microorganisms have a range of molecules of interest in terms of investigation, patenting

and commercial products. For examples, polyunsaturated fatty acids could be used as dietary supplements for aquaculture, livestock and human diets; pharmaceutical and cosmetic industries are more interested in physical properties of EPS, like high viscosity, gelling capacity or high resistance in a wide range of temperature and pH (Freitas et al., 2011). From the biomedical point of view, bacterial LPSs isolated from cold environments may also be a source of potential immunostimulant molecules. During these years, an ever-growing attention has been devoted to the research of new lipid As, obtained either from engineered bacterial extracts (Needhama et al. 2013) or synthetic (Gao et al., 2017; Maiti et al. 2010; Shimoyama et al. 2011) and semisynthetic approaches (Pieretti et al. 2014; D'Alonzo et al. 2016), that could find application as possible adjuvants. In this context, extremophiles may be a promising source of non-toxic LPSs and lipid As derived from these. In particular, cold environments, due to the still low contamination by pathogenic microorganisms, may represent an untapped reservoir of immunomodulating molecules. Indeed, lipid As isolated from several psychrophiles have been reported to antagonize *E. coli* LPS toxicity, while others have proved able to act as weak agonists. Both types are interesting, due to the possibility of exploiting them for the treatment of inflammatory conditions and as immunoadjuvants, respectively.

References

- Bakermans C. and Neelson, K. H., *J. Bacteriol.* **2004**, *186*, 2340-2345.
- Bianciotto V. et al., *Eur.J Histochem.* **2001**, *45*, 39-49.
- Carty S.M., et al., *J. Biol. Chem.* **1999**, *274*, 9677-9685.
- Cavicchioli R., *Nat Rev Microbiol.* **2006**, *4*, 331-343.
- Chattopadhyay K. et al., *Current Biotechnol.* **2014**, *3*, 100-116.
- Chattopadhyay K., *J. Biosci.* **2006**, *31*, 157-165.
- Chin J.P. et al., *PNAS* **2010**, *107*, 7835-7840.
- Costerton J.W. et al., *Annu. Rev. Microbiol.* **1995**, *49*, 711-745.
- D'Alonzo D. et al., *Chem Eur J.* **2016**, *22*, 11053-11063.
- De Maayer M. et al., *EMBO Rep.* **2014**, *15*, 508-517.
- Feller G. and Gerday C. *Nat. Rev. Microbiol.* **2003**, *1*, 200-208.
- Freitas F. et al., *Trends in Microbiol.* **2011**, *29*, 388-398.
- Fuller B.J., *CryoLetters* **2004**, *25*, 375-388.
- Fux C.A. et al., *Expert Rev. Anti-Infect. Ther.* **2003**, *1*, 667-683.
- Gao H. et al., *J. Bacteriol.* **2006**, *188*, 4560-4569.
- Gao J., Guo Z. *Med. Res. Rev.* **2017**, (*in press*),
doi: 10.1002/med.21447.
- Gulik A. et al., *J. Mol. Biol.*, **1985**, *182*, 131-49.
- Helmke E. and Weyland H. *Cell. Mol. Biol.* **2004**, *50*, 553-561.
- Kumar A.S. et al., *Bull. Environ. Contam. Toxicol.* **2007**, *79*,
217-221.
- Laucks M.L. et al., *App. Spectrosc.* **2005**, *59*, 1222-1228.
- Madigan M.T., Mairs B.L., *Scientific American*, **1997**, *276*, 82-87.
- Maiti K.K. et al., *Eur. J. Org. Chem.* **2010**, 80-91.

- Margesin R. et al., *Naturwissenschaften*, **2007**, *94*, 77-99.
- Marx JG, et al., *Can. J. Microbiol.* **2009**, *55*, 63-72.
- Mesbah N.M., Wiegel J., *Ann. NY Acad. Sci.*, **2008**, *1125*, 44-57.
- Methé B. A. et al., *Proc. Natl. Acad. Sci.* **2005**, *102*, 10913-10918.
- Morita R.Y Psychrophilic bacteria. *Bacteriol. Rev.* **1975**, *39*, 144-167.
- Pieretti G. et al., *Appl Microbiol Biotechnol.* **2014**, *98*, 7781-7791.
- Poli A. et al., *Mar. Drugs* **2010**, *8*, 1779-1802.
- Shimoyama A. et al., *Chem. Eur. J.* **2011**, *17*, 14464-14474.
- Siddiqui K.S. et al., *Annu. Rev. Earth Planet. Sci.* **2013**, *41*, 87-115.
- Taylor M.P. et al., *Extremophiles: Microbiology and Biotechnology* (Ed.: R. P. Anitori), Caister Academic Press, Norfolk (UK), 1-24, **2012**.

Chapter 4: *Colwellia psychrerythraea* strain 34H

The family *Colwelliaceae* is a part of the order of *Alteromonadales*, located within class *Gammaproteobacteria*. It contains twelve psychrophilic or psychrotolerant species (Deming et al., 1988), two of which are also barophilic (Collins et al., 2010). The species of family *Colwelliaceae* are principally Gram-negative, rod- to curved rod-shaped cells and facultative anaerobic. Deming et al. (1988) first described the genus *Colwellia* for a psychrophilic microorganism isolated from Flounder eggs collected in Norway. This name was given in honor of the american microbiologist Professor Rita Colwell (Bowman 2014). *Colwellia* species are ubiquitous present in cold and polar marine ecosystem, including ice brine channels, marine biofilms and aquaculture system. Nowadays, only the genome of the sea ice bacterium *Colwellia psychrerythraea* strain 34H had determined. Analysis of the genome suggests that a collection of synergistic changes in the overall genome content and amino acid composition supports its psychrophilic lifestyle (Methe et al., 2005). *Colwellia psychrerythraea* strain 34H, isolated from subzero Arctic marine sediments (Huston et al., 2004), grows at temperatures as low as -12°C, with optimal and maximal growth temperatures of 8-9°C and 19°C, respectively (Huston et al., 2004). The analysis of *C. psychrerythraea* genome evidences the features that allow living at freezing temperatures and for this reason this bacterium is considered a model to study cold adapted life-style.

4.1 Lipopolysaccharides and lipid A structures

The cells extracted by the PCP method revealed a rough-LPS (LOS), as already found for other psychrophilic microorganisms (Corsaro et al., 2004; Corsaro et al., 2008; Carillo et al., 2011). The structure, obtained by chemical analysis, NMR spectroscopy and ESI mass spectrometry, is reported below (Figure 4.1) (Carillo et al., 2013):

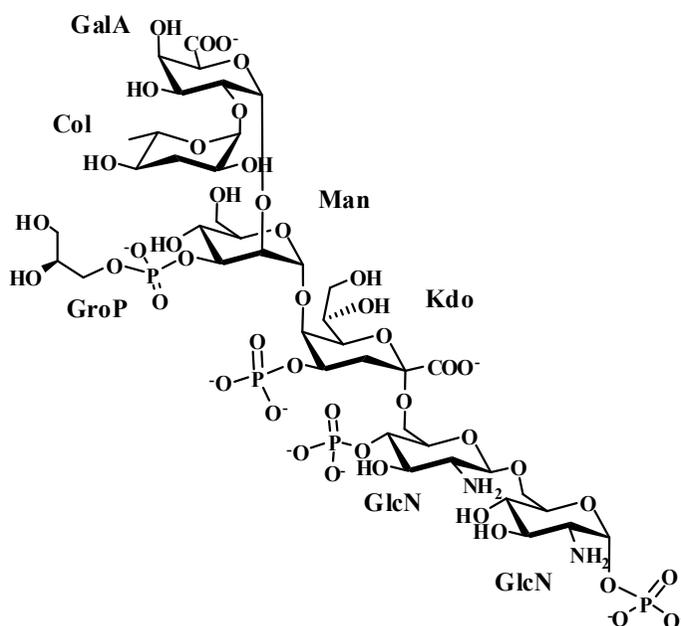


Figure 4.1: Core structure of LOS from *C. psychrerythraea* 34H.

Only few sugar units, some of which are acidic monosaccharides, constitute the core region. This structural feature, together with the presence of several phosphate groups, gives to the LOS a high charge density. Furthermore, it is worth noting the lack of heptose residues in the inner core of LOS; in their place, an α -mannose residue is linked to the Kdo. This structural feature is commonly found in *Rhizobiaceae* family, but has never been found in extremophiles. Finally, colitose and

phosphoglycerol, already found in some O-polysaccharides chains (Silipo et al., 2005; Kenyon et al., 2011), are present (Figure 4.1).

4.1.1 Isolation and compositional analysis of lipid A

Extraction of *Colwellia psychrerythraea* 34H dried cells, grown at 4°C, by the phenol/chloroform/light petroleum (Galanos et al. 1969) yielded about 1.1% of LOS (Carillo et al., 2013). The LOS sample was recovered as precipitate after washing in chloroform/methanol (1:2, v/v) to a concentration of 5 mg mL⁻¹ to remove phospholipids as supernatant. The lipid A fraction was obtained by 1% acetic acid hydrolysis of LOS suspension. The lipid A was recovered as pellet after centrifugation of the crude reaction, washed twice with water and freeze-dried. Sugar analysis revealed the presence of D-glucosamine as unique monosaccharide. Fatty acid methyl esters, obtained after methanolysis of the lipid A and extraction with hexane, were recognised in the GC-MS chromatogram by comparison of their retention times with standards and from their EI-MS spectra. GC-MS analysis revealed the presence of 3-hydroxy-dodecanoic [C12:0(3-OH)] and 3-hydroxy-tetradecenoic acids [C14:1(3-OH)], as hydroxylated species. The presence of an unsaturated hydroxylated species was quite uncommon. The analysis also revealed the occurrence of decanoic (C10:0), dodecanoic(C12:0), dodecenoic (C12:1), tetradecanoic (C14:0), tetradecenoic (C14:1), and hexadecenoic (C16:1) acids. In addition, signals attributable to hexadecanoic (C16:0) and octadecanoic (C18:0) acids suggested the presence in the LOS sample of membrane phospholipids. Finally, less abundant signals were attributed to pentadecanoic (C15:0) and pentadecenoic (C15:1) acids.

4.1.2 Mass spectrometry analysis of lipid A

The characterization of *Colwellia psychrerythraea* 34H lipid A, started and partially reported in the PhD thesis of Dr. Casillo, was continued and completed during this PhD work. Indeed, as already reported, the lipid A, was analyzed by ESI FT-ICR mass spectrometry in the negative-ion mode. The obtained mass spectrum was complex and showed five signal clusters (Figure 4.2, Table 4.1), corresponding to a number of glycoforms differing for their acylation and phosphorylation degree.

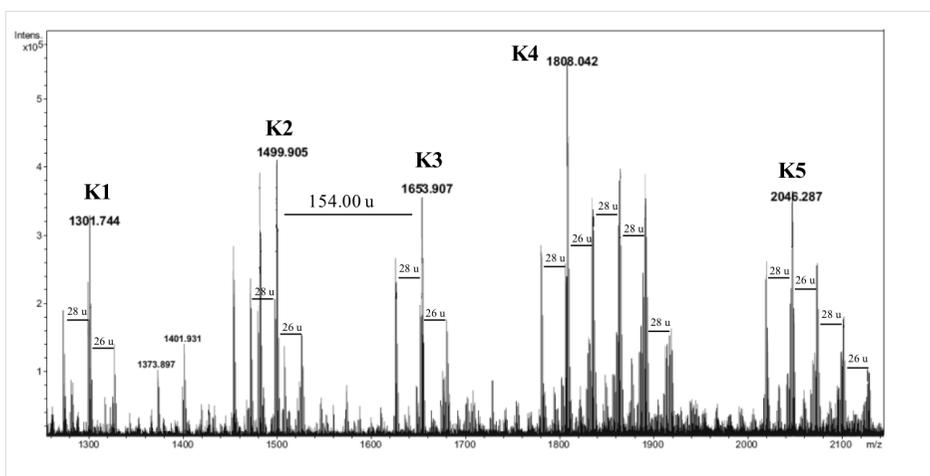


Figure 4.2: Negative ions ESI FT-ICR mass spectrum of the lipid A from *Colwellia psychrerythraea* 34H.

Table 4.1: Composition of the main species present in the negative ions ESI FT-ICR mass spectrum of the lipid A from *Colwellia psychrerythraea* 34H.

Species	[M-H] ⁻ calculated	[M-H] ⁻ observed	Composition
K1	1301.751	1301.744	(GlcN) ₂ P ₂ [C12:0(3OH)] ₂ [C14:1(3OH)][C12:0]
K2	1499.916	1499.905	(GlcN) ₂ P ₂ [C12:0(3OH)] ₃ [C14:1(3OH)][C12:0]
K3	1653.919	1653.907	(GlcN) ₂ P ₂ [C12:0(3OH)] ₃ [C14:1(3OH)]GroP-[C12:0]
K4	1808.058	1808.042	(GlcN) ₂ P ₂ [C12:0(3OH)] ₃ [C14:1(3OH)]GroP-[C12:0][C10:0]
K5	2046.290	2046.287	(GlcN) ₂ P ₂ [C12:0(3OH)] ₃ [C14:1(3OH)]GroP-[C12:0][C10:0][C16:0]

In particular, they ranged from tetra-acylated to hepta-acylated glycoforms, of which the most abundant are named **K1-K5**. For the species **K1** with $[M-H]^-$ at 1301.744 m/z , with composition $\text{GlcN}_2\text{P}_2[\text{C12:0(3-OH)}]_2[\text{C14:1(3-OH)}][\text{C12:0}]$ (m/z calculated: 1301.751 Da), revealed a tetra-acylated lipid A with only three hydroxylated fatty acids. The species **K2** at 1499.905 m/z (m/z calculated: 1499.916 Da), was identified as penta-acylated lipid A with an additional C12:0(3-OH), suggesting an asymmetric distribution of hydroxylated primary fatty acids. The species at 154.002 Da higher, indicated as **K3**, differed from **K2** in the uncommon presence of a glycerol phosphate residue (GroP). Although the presence of GroP linked to saccharides belonging to the O-chain of a LPS is not unusual, (Wang et al. 2008; Jachymek et al. 1999) the linkage of GroP to lipid A has only been reported in lipid A from *Vibrio fischeri* (Phillips et al. 2011), which shares high phylogenetic proximity with *Colwellia* (Methe et al. 2005).

Most-abundant species **K4** at m/z 1808.042 was identified as a hexa-acylated species carrying an additional C10:0 ($\Delta m=154.135$ Da). Species **K5** at m/z 2046.287 displayed an additional C16:0 ($\Delta m=238.248$ Da). All the species comprise heterogeneity in unsaturation owing to substitution of C12:0 with C12:1 ($\Delta m=-2$ Da). In addition, heterogeneity due to the presence of C14:1 ($\Delta m=+26$ Da) or C10:0 (-28 Da) instead of a C12:0 fatty acid was displayed. Finally, further heterogeneity was observed for **K4** and **K5** in the length of the fatty acids and in the presence of an uneven number of C atoms, which is in agreement with the compositional analysis by GC-MS (C15 :0 and C15 :1). Lipid A was then treated with NH_4OH (Silipo et al. 2002), which hydrolyzed only the acyl and acyloxacyl esters and left the acyl and acyloxacyl amides intact.

Lipid A NH₂OH was analyzed by negative-ion ESI FT-ICR MS, and the spectrum displayed four signals clusters centered at species **M1-M4** (Figure 4.3, Table 4.2).

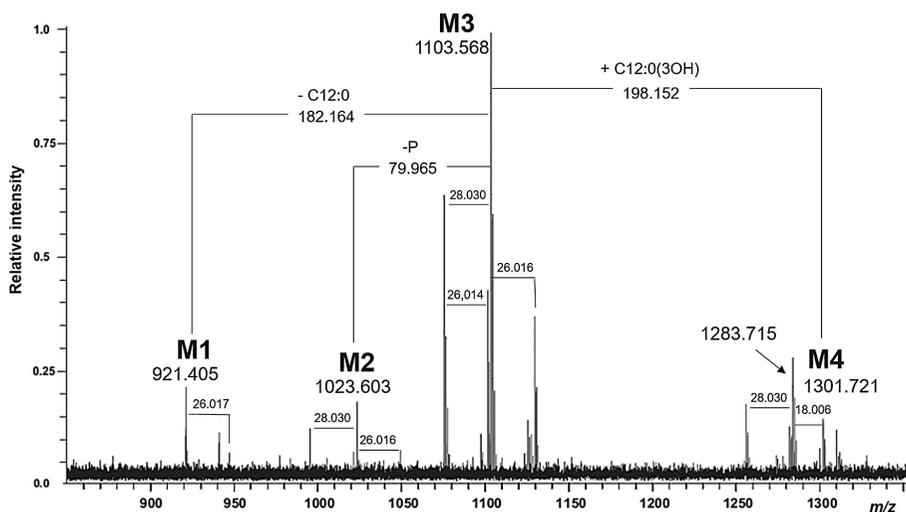


Figure 4.3: Negative ions mass spectrum of the lipid A_{NH₂OH} from *Colwellia psycherythraea* 34H.

Species	[M-H] ⁻ calculated	[M-H] ⁻ observed	Composition
M1	921.419	921.405	GlcN ₂ P ₂ [C12:0(3OH)][C14:1(3OH)]
M2	1023.622	1023.603	GlcN ₂ P[C12:0(3OH)][C14:1(3OH)][C12:0]
M3	1103.589	1103.568	GlcN ₂ P ₂ [C12:0(3OH)][C14:1(3OH)][C12:0]
M4	1301.754	1301.721	GlcN ₂ P ₂ [C12:0(3OH)] ₂ [C14:1(3OH)][C12:0]

Table 4.2: Composition of the main species of the lipid A_{NH₂OH} observed in the ESI FT-ICR mass spectrum.

Less abundant signals, appearing at higher ($\Delta m = +26$ Da) and lower ($\Delta m = -28$ Da) m/z values relative to the m/z values of species **M1-M4** and **M2-M4**, respectively, were attributed to the presence of C14:1 or C10:0,

respectively, instead of C12:0. In addition, all the main species, except for **M1**, presented a shift in mass of $\Delta m = -2$ Da; this suggested secondary acylation by dodecenoic acid instead of dodecanoic acid. All the species clearly indicated that C12:0(3-OH) and C14:1(3-OH) were linked as amides to either GlcN I or GlcN II.

To obtain more detailed information on the distribution of fatty acids and the linkage of GroP on the disaccharidic backbone, lipid A was studied in the positive-ion mode with a Q-Exactive Plus Orbitrap (Thermo, Bremen). The mass spectrum showed intensive adduct ions $[M+Et_3N+H]^+$ of corresponding negative-ion species **K1-K5** (Figure 4.4), however, with slightly different relative intensities.

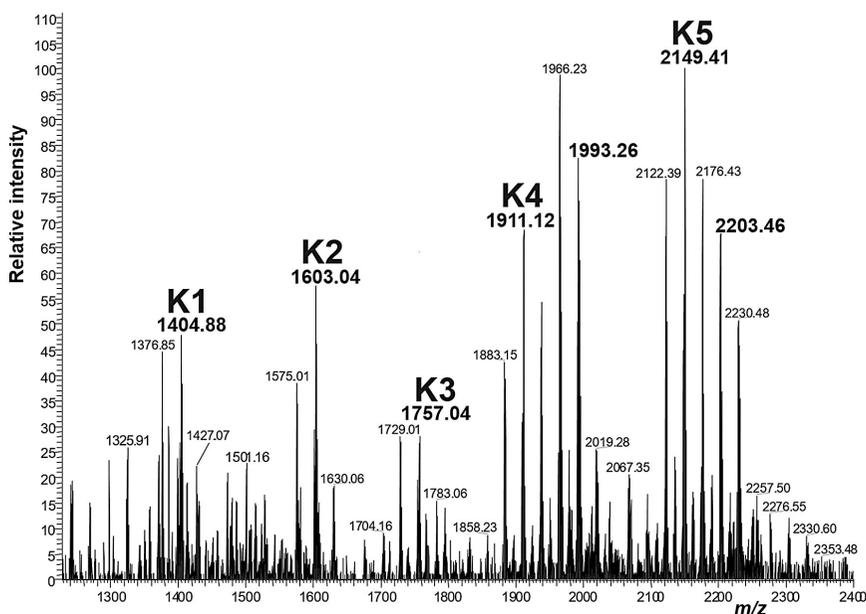


Figure 4.4: Positive-ion Q-Exactive Plus mass spectrum of lipid A from *C. psychrerythraea* 34 H. The mass numbers given in bold indicate the $[M+Et_3N+H]^+$ ions that were selected for MS/MS.

In general, the MS/MS spectra showed the expected loss of Et_3N ($\Delta m = 101$ Da) and the cleavage of the phosphate linked to GlcN I

($\Delta m=97.89$ Da). Further fragmentation led to the appearance of both the **B1** and **Y1** fragments (Domon et al. 1988), together with fragments resulting from the loss of fatty acids mainly as free acids instead of ketenes.

The **B1** fragment ions at m/z 846.55 and 818.52, with compositions of GlcNP[C14:1(3OH)][C12:0(3OH)][C12:0] and GlcNP[C14:1(3OH)][C12:0(3OH)][C10:0], respectively, were found in the MS/MS spectra (Figures 4.5-4.7), which definitively established the substitution on GlcN II. Indeed, whereas the analysis of the lipid A NH₂OH product had previously indicated C14:1(3-OH) as an acyl amide at the 2-position, the presence of C12:0(3-OH) at the 3-position was deduced from the **B1** fragment. Taking together the results of the analysis of the lipid A NH₂OH and **B1** fragment ions, the occurrence of C12:0 as an acyloxacyl of C14:1(3-OH) was inferred. Moreover, the fragmentation spectra of the main species showed the presence of signals corresponding to the elimination of C12:0 as a free acid ($\Delta m=200.18$ Da) from **B1**, which thus confirmed that C14:1(3-OH) was acylated by the same fatty acid in all of the main species (Figures 4.5-4.7). Starting from penta-acylated species **K2**, the MS/MS spectrum of the corresponding signal at m/z 1603.04 (Figure 4.5) displayed the **B1** fragment ion at m/z 846.55 and a signal at m/z 1403.94 corresponding to the simultaneous elimination of a phosphate group and water (**B2** fragment). Less-intense fragments were observed at m/z 1203.76 and 1187.94 owing to the loss of C12:0 or C12:0(3OH), respectively (Figure 4.5).

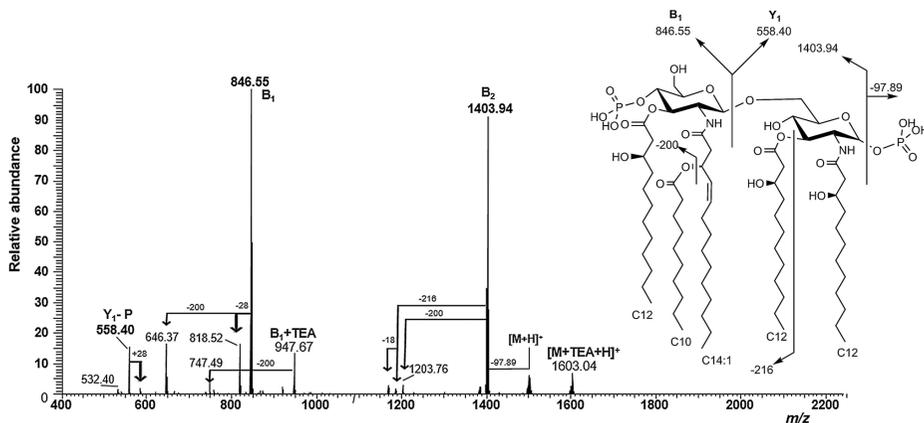


Figure 4.5: MS/MS spectrum of m/z 1603.04. TEA = triethylamine.

This finding indicated that both fatty acids were linked as esters (Kussak et al. 2002). The spectrum also contained the signal at m/z 558.40, which could be attributed to the **Y1** fragment lacking a phosphate group, and this definitely confirmed the structure of the penta-acylated species reported in the inset of Figure 4.5. To establish the position of the GroP moiety, the $[K3+Et_3N+H]^+$ adduct ion at m/z 1757.04 was selected for MS/MS analysis (Figure 4.6), and the obtained spectrum was clearly related to that of the penta-acylated form. In fact, whereas the **B1** fragment remained the same, the **Y1** ion shifted in mass by $\Delta m = +154.0$ Da, which thus indicated the localization of the phosphoglycerol on GlcN I. In addition, the simultaneous loss of $\Delta m = 97.89$ Da (phosphate group and water) and $\Delta m = 388.19$ Da (moiety composed of the 3-hydroxydodecanoic acid and phosphoglycerol group) from the $[K3+H]^+$ parent ion at m/z 1557.94 (**B2** fragment) suggested that the GroP moiety was linked to C12:0(3OH), which was in turn linked to the 3- position of GlcN I.

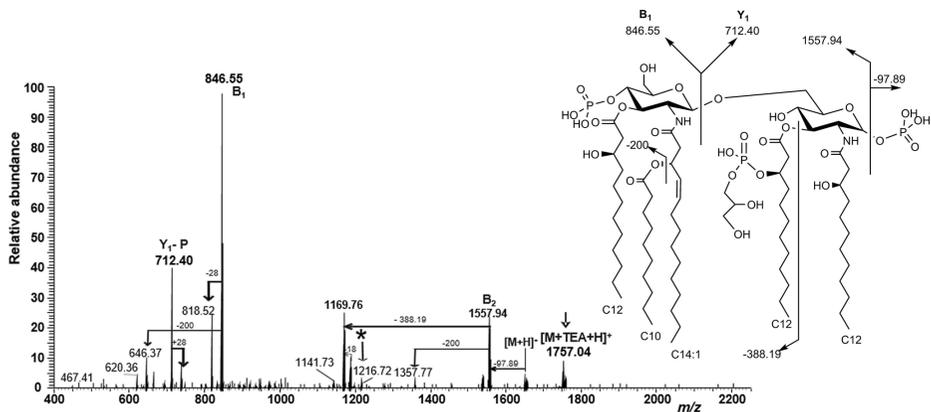


Figure 4.6: MS/MS spectrum of m/z 1757.04. TEA = triethylamine.

If selected for MS/MS analysis, the hexa-acylated lipid A species at m/z 1911.12 gave the fragmentation reported in Figure 4.7. In this spectrum, the **Y1** ion fragment at m/z 866.54 indicated that the proximal GlcN moiety had the additional decanoic acid linked as an acyloxacyl ester. Furthermore, the signal at m/z 1187.77 arose from the elimination of $524.31 + 97.89$ Da from the $[K_4+H]^+$ parent ion; this suggested that the glycerol moiety was substituted by decanoic acid.

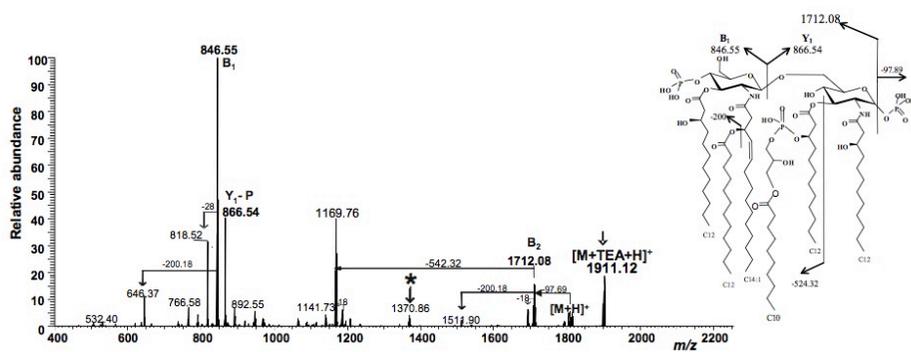


Figure 4.7: MS/MS spectrum of m/z 1911.12. TEA = triethylamine.

The last fatty acid was assigned to the *sn*-1 position thus to form a lysophosphatidic acid moiety, as demonstrated by NMR spectroscopy (see paragraph 4.1.3). Upon selecting higher-molecular-weight species for

MS/MS analysis, the spectra obtained revealed the additional substitution of the phosphoglycerol moiety. Indeed, the fragmentation spectrum of the hepta-acylated species $[\mathbf{K5}+\text{Et}_3\text{N}+\text{H}]^+$ at m/z 2149.25 (Figure 4.8), containing an additional hexadecanoic acid with respect to **K4**, gave a fragment ion due to the loss of 780.55 Da (m/z 1169.76).

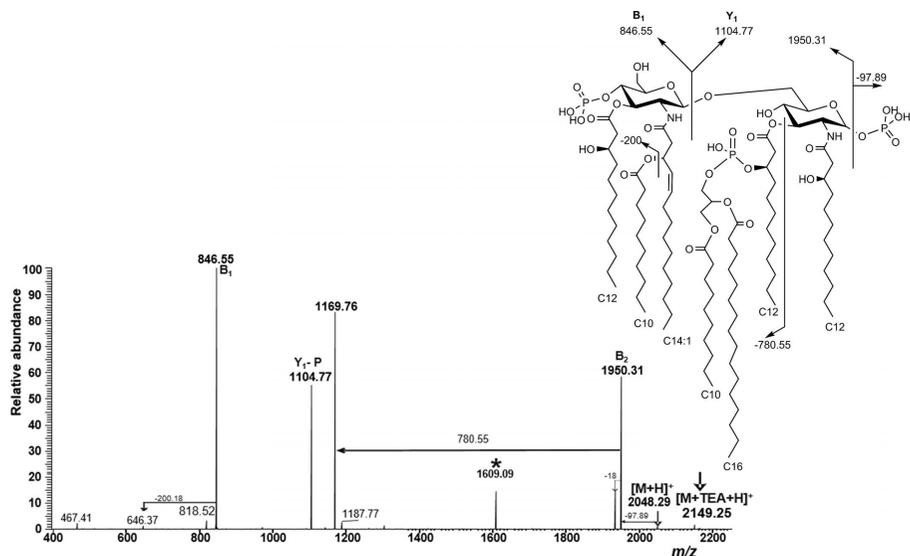


Figure 4.8: MS/MS spectrum of m/z 2149.25. TEA = triethylamine.

This fragment is consistent with the mass predicted for the loss of the entire substituent at the *O*-3 position of GlcN I as free acid plus a water molecule. The same GroP substitution was revealed in the MS/MS spectrum of the species at 28 Da lower than the mass of **K5** (Figure 4.9), for which the **B1** fragment ion indicated a decanoic acid as the secondary fatty acid.

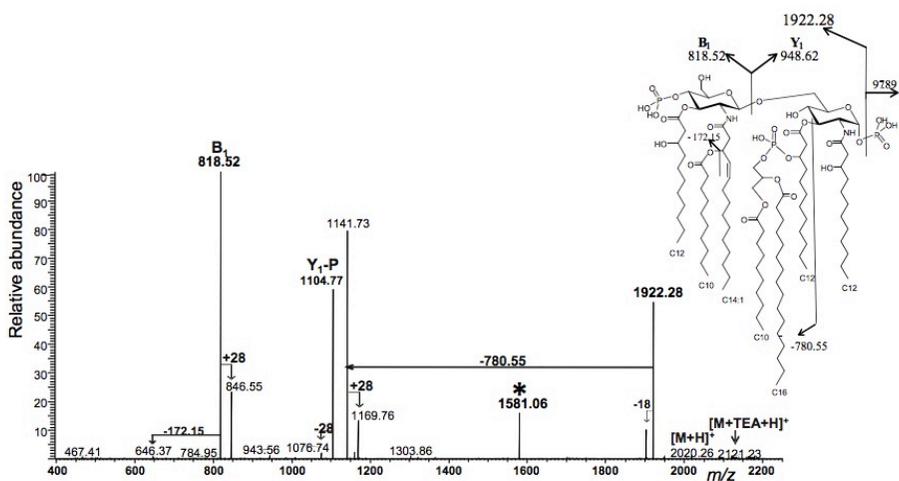


Figure 4.9: MS/MS spectrum of m/z 2121.23. TEA = triethylamine.

Other hepta-acylated species were analyzed by inducing fragmentation and revealed a different substitution on the GroP moiety (Figure 4.10); this indicated heterogeneity in the acylation of this moiety.

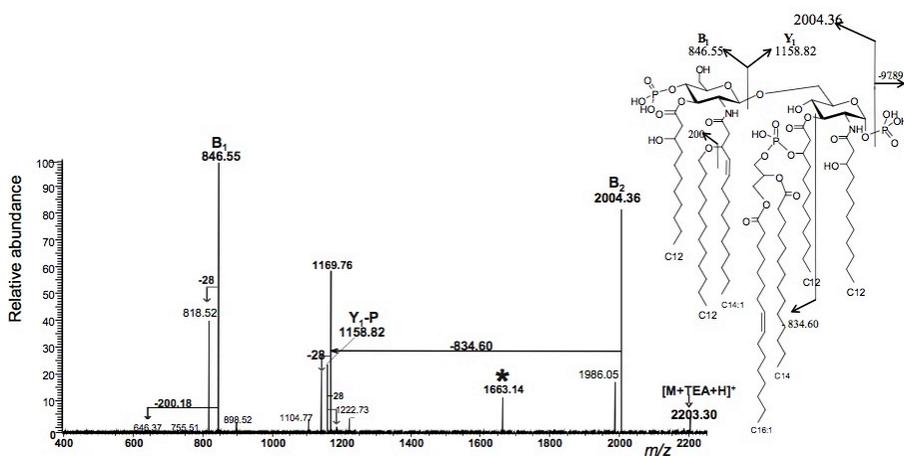
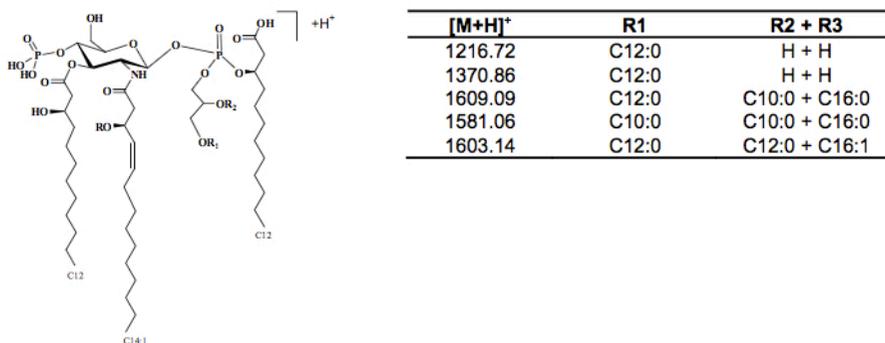


Figure 4.10: MS/MS spectrum of m/z 2203.30. TEA = triethylamine.

Finally, all the signals marked by an asterisk cannot be straightforwardly explained by the normal fragmentation pathways known for lipid A structures. These fragment ions are observed only from lipid A species

carrying the GroP and might be due to gas-phase rearrangements during CID. The m/z values observed for these ions in the reported MS/MS spectra (Figures 4.6-4.10) are in excellent agreement with the suggested structures. However, further experiments have to be performed to confirm this suggestion (Scheme 4.1).



Scheme 4.1: Suggested structures for the fragment ions marked by an asterisk in the reported MS/MS spectra (Figures 4.6-4.10).

4.1.3 NMR spectroscopy of lipid A

In addition to the data obtained from the mass spectrometry experiments, the NMR spectroscopy was here used to prove some details of the lipid A structure. To confirm the presence of a phosphoglycerol residue covalently linked to lipid A, 2D NMR spectroscopy experiments were performed (¹H,¹H COSY, ¹H,¹H TOCSY, ¹H,¹³C HSQC-DEPT, ¹H,¹³C HMBC, and ¹H,³¹P HMBC). Two signals were detected in the proton anomeric region of the HSQC experiment at $\delta=5.02$ and 4.57 ppm, whereas the carbon signals resonated at $\delta=91.0$ and 98.0 ppm, respectively (Table 4.3).

	Sugar portion ^[a]						Lipid portion ^[a]						Glycerol moiety ^[a]
	1	2	3	4	5	6	α	β	γ	δ	$\omega-1$	ω	
A							C12:0(3-OH)			1.18			3.87, 3.80
GlcN I (α)	5.02	4.05	5.07	3.87	n.d.	4.06, 4.08	C14:1(3-OH)	2.50, 2.39	5.12	29.2	1.18		63.1
	91.0	52.0	73.2	68.4		69.7		40.3	67.3	5.26	29.2		5.22
										130.0			128.9-CH ₂ -
												0.78-0.75	(-12.7) ^[b]
												13.5	5.09
													69.8
B	4.57	4.17	4.95	4.07	3.29	3.76, 3.82	R1-R3	2.11-2.25	1.48	1.12-1.24			-CH(OH)-
GlcN II (β)	98.0	50.0	72.6	68.8	73.8	58.9	C10:0-C16:0	34.1	24.6	22.3-31.6			4.28, 4.04
													62.0
													-CH ₂ -

[a] ¹H (600 MHz, 298 K) and ¹³C (150 MHz, 298 K) chemical shifts (in plain and in italics, respectively) expressed in δ relative to residual CHCl₃ (¹H: δ = 7.26 ppm; ¹³C: δ = 77.0 ppm) in CDCl₃/CD₃OD (4:1, v/v); n.d.: not defined. [b] ³¹P (160 MHz, 298 K) chemical shift expressed in δ relative to 85% phosphoric acid (external standard).

Table 4.3: NMR spectroscopy data of *C. psychrerythraea* 34H lipid A.

These two crosspeaks are diagnostic of the α configuration of GlcN I and the β configuration of GlcN II, respectively. All the proton and carbon chemical shifts of the GlcN disaccharide were assigned and reported (Table 4.3, Figure 4.11).

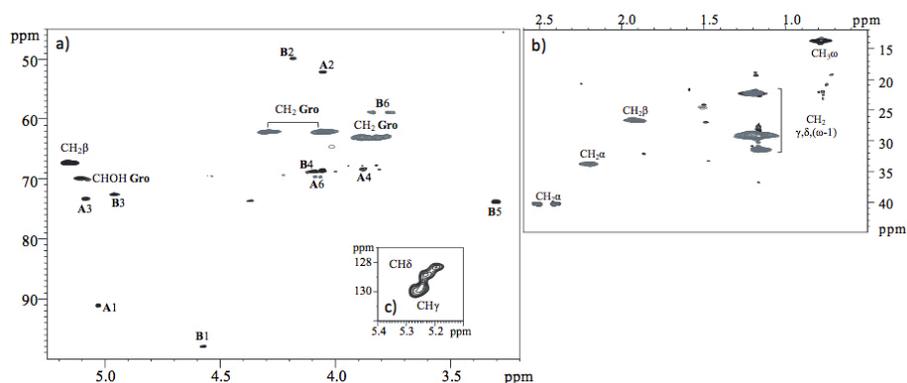


Figure 4.11: ¹H, ¹³C HSQC-DEPT NMR spectrum in 4:1 v/v CDCl₃-CD₃OD of the lipid A from *C. psychrerythraea* 34H grown at 4 °C showing the (a) anomeric and carbinolic region; (b) methylene and methyl groups region; and (c) double-bonds region.

In addition, it was possible to assign the position of the double linkage of C14 :1(3-OH). In fact, the carbon signal at δ =67.3 ppm, assigned to the β carbon atom of the 3-hydroxy fatty acids, was correlated in the HMBC experiment with both diastereotopic α -CH₂ protons at δ =2.39/2.50 ppm and with a CH olefinic proton at δ =5.26 ppm, which thus indicated that

the double linkage was located between carbon atoms C4 and C5 (Figure 4.12). The NMR spectroscopy data also revealed the presence of a GroP moiety linked to lipid A. In the HSQC-DEPT experiment, the crosspeak at $\delta = 5.09/69.8$ ppm was attributed to the CHOH group of the lysophosphatidic acid moiety. In particular, one of the signals of the CH₂ group of the Gro residue occurring at $\delta = 3.87/63.1$ ppm in the HSQC-DEPT experiment was correlated in the ¹H, ³¹P HMBC experiment with a ³¹P signal at $\delta = -12.7$ ppm, the chemical shift of which indicated the phosphodiester linkage (Helander et al. 1994) (Table 4.3). In addition, one of the signals of the CH₂ group of the Gro residue ($\delta = 4.04$ ppm) correlated with a carbonyl signal at $\delta = 173.3$ ppm in the ¹H, ¹³C HMBC experiment (Figure 4.12), which in turn correlated with the α -CH₂ protons of the acyl ester at $\delta = 2.21$ ppm.

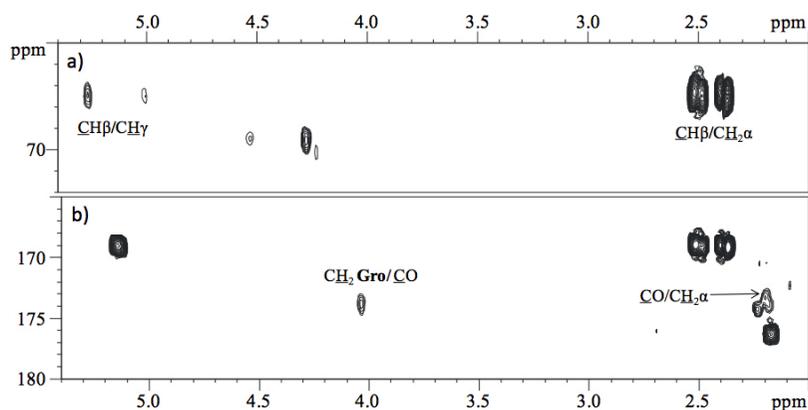
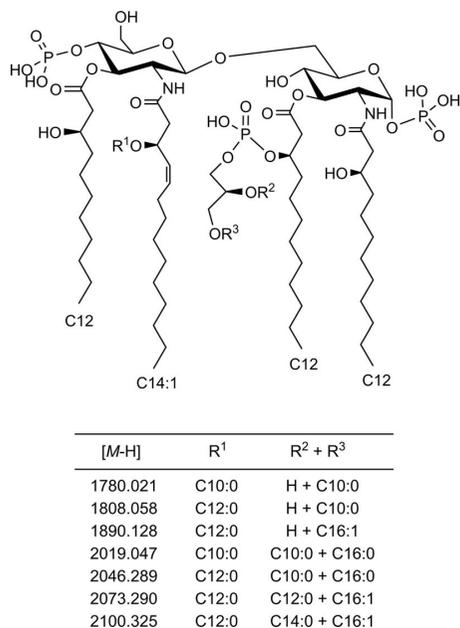


Figure 4.12: ¹H, ¹³C HMBC NMR spectrum in 4:1 v/v CDCl₃-CD₃OD of the lipid A from *C. psychrerythraea* 34H grown at 4 °C showing the (a) carbinolic region; and (b) carboxylic region.

These last correlations indicated that in the most-abundant species, the CH₂OH position of the Gro moiety was always acylated. All together, these results established the four fatty acids at the primary acylation sites

on lipid A of *C. psychrerythraea* 34H and also showed microheterogeneity at the secondary acylation site at the 2'-position and in the GroP moiety (Scheme 4.2).



Scheme 4.2: Structure of *C. psychrerythraea* 34H lipid A.

The double bond is arbitrarily shown in the *cis* configuration.

4.1.4 Evaluation of biological activity

Biological activity assays were performed to characterize the effect of *C. psychrerythraea* lipid A on TNF secretion in the presence of a serial dilution of *Escherichia coli* O111:B4 LPS.

Therefore, we used the human monocytic cell line THP-1, which we differentiated into mature macrophages a 24h-long pre-treatment with 200 mM phorbol 12-myristate 13-acetate (PMA). Macrophage-like THP1 cells express MD-2 and CD14, which are necessary for the transfer of monomeric LPS to the TLR4 cell surface receptor and further downstream signalling (Park et al. 2009). Cells were

simultaneously stimulated with different concentrations of *C. psychrerythraea* lipid A and *E. coli* O111:B4 LPS or were further left untreated.

After 24h, medium was taken, and the amount of secreted TNF was measured by ELISA. As seen in Figure 4.13, *C. psychrerythraea* lipid A did not show a clear effect on LPS-induced TNF production in THP-1 cells. To conclude, *C. psychrerythraea* lipid A has no agonistic or antagonistic effect on LPS-induced TNF production on human macrophages.

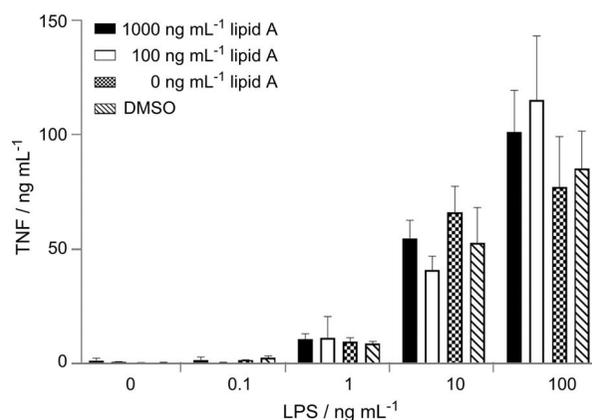


Figure 4.13: Effect of *C. psychrerythraea* lipid A on LPS-induced TNF production in THP-1 cells. Cells were incubated with the indicated amounts of *C. psychrerythraea* lipid A in the presence or absence of different concentrations of *E. coli* O111:B4 LPS, as indicated.

4.2 Conclusions

C. psychrerythraea 34H represents a good model to understand how cold-adapted bacteria contrast harsh environmental conditions. The structure of *C. psychrerythraea* 34H lipid A was established, the characterization of which may help to understand the structure-function relationships (Casillo et al. 2017). The results of chemical, MS, and

NMR spectroscopy analyses revealed a structure that is quite unusual for lipid As. The primary acylation pattern was found to be asymmetric, and reducing GlcN was found to be substituted at positions 2 and 3 by two 3-hydroxydodecanoic acids, whereas the distal GlcN was found to carry a 3-hydroxydodecanoic acid and 3-hydroxy-tetradecenoic acid at positions 3' and 2', respectively. To our knowledge, a 3-hydroxy unsaturated tetradecenoic acid together with a phosphoglycerol moiety on a secondary acylation site at the 3-position of the reducing GlcN are found only in *Vibrio fischeri* lipid A (Phillips et al. 2011). In addition, the level of micro-heterogeneity in the *Colwellia* structure is high owing to the presence of unsaturated secondary fatty acids and to a wide range of lengths for the saturated chains. Finally, another site of heterogeneity is constituted by the phosphoglycerol moiety, which evokes in turn acylated at positions *sn*-1 and *sn*-2 to some extent, to produce the hexa-acylated and hepta-acylated glycoforms, respectively.

The lipid A moiety displays several structural features that could be related to the adaptation mechanisms adopted by *C. psychrerythraea* 34H. The presence of unsaturated fatty acids in phospholipids and lipid As from cold-adapted bacteria and mesophiles grown at a temperature lower than their optimal one is well documented (Carty et al. 1999). In this context, the structure of *C. psychrerythraea* 34H lipid A here reported is in agreement with the genomic data of the microorganism, the coding sequences of which indicate both unsaturated and polyunsaturated fatty acid biosynthesis (Hashimoto et al. 2015). In contrast, a comparison with lipid A structures from other *Colwellia* species revealed substantial differences, the most important of which was the absolute absence of the phosphoglycerol moiety in both the *C. piezophila* and *C. hornerae* species (Sweet et al. 2015). In addition, no

unsaturated fatty acids have so far been reported for the structures from these two *Colwellia* species grown whether at 4 or 15°C.

More intriguing is the simultaneous presence of the phosphoglycerol moiety on both the core (Carillo et al. 2013) and lipid A moieties of *C. psychrerythraea* 34H LPS. One can assume that these unusual structural features of the LPS molecule have an effect on the outer membrane of *C. psychrerythraea* 34H cells. In particular, the high salt concentration to which *C. psychrerythraea* 34H cells are exposed in brine channels forces the cells to be in osmotic balance with the environment, and the phosphoglycerol moiety could act as an osmolyte (Thomas et al. 2002). Biological activity assays performed in the human macrophage cell line THP-1 did not show an agonistic or antagonistic effect of *C. psychrerythraea* lipid A on tumor necrosis factor production in these cells at any concentration used.

To summarize, a new lipid A structure from *C. psychrerythraea* 34H was characterized; it enlarges the number of structure-activity relationships data available for the design of new immunomodulatory molecules (D'Alonzo et al. 2016; Marzabadi et al. 2017; Choma et al. 2017).

4.3 Experimental section

4.3.1 Isolation of the LPS

C. psychrerythraea 34H was isolated from Arctic marine sediments (Méthé et al. 2005) and was grown at 4°C as previously reported (Carillo et al. 2011). Dried bacteria cells (4.8 g) were extracted by the PCP method (Fuller, 2004), which gave 52 mg of LPS (yield 1.1 % of dried cells). The LPS sample was recovered as precipitate after washing in chloroform/methanol (1:2, v/v) to a concentration of 5 mg mL⁻¹ to remove phospholipids as supernatant.

4.3.2 Mild acid hydrolysis of the LPS

The LPS (23 mg) was incubated with 1% aqueous CH_3COOH (3mL) for 3h at 100°C . The sample was then centrifuged at $10000g$ for 30min, and two fractions were recovered: lipid A (7.5 mg) as pellet and the saccharide portion (15.5 mg) as surnatant.

4.3.3 Chemical analysis

The lipid A fraction (0.5 mg) was analyzed as acetylated methyl glycoside derivatives (AMG). Briefly, $\text{HCl}/\text{CH}_3\text{OH}$ (1mL, 1.25M) was added to the sample, and the reaction was performed at 80°C for 20 h. The crude mixture was extracted with hexane (2x). The hexane layer, containing the fatty acids as methyl ester derivatives, was analyzed by GC-MS. The analyses were performed with an Agilent Technologies gas chromatograph 6850A equipped with a mass-selective detector 5973 N and a Zebron ZB-5 capillary column (Phenomenex, $30\text{m}\times 0.25\text{mm}$ i.d., flow rate 1 mL min^{-1} , He as carrier gas). The following temperature program was used for lipid analysis: 110°C for 3 min, 140 to 280°C at $10^\circ\text{C min}^{-1}$.

4.3.4 Lipid A purification

Lipid A (5 mg) was purified through a short Celite column ($\approx 5\text{cm}^3$). Gel-filtration chromatography was performed with a Sephadex LH-20 ($1\times 50\text{cm}$, flow rate 48 mL h^{-1} , eluent $\text{CHCl}_3/\text{MeOH}$ 3:2, v/v) to obtain 1.6 mg of pure product.

4.3.5 NH_4OH hydrolysis of lipid A

Lipid A (0.5 mg) was incubated with concentrated aqueous NH_4OH (200 μL), as reported (Silipo et al. 2002). The sample was dried and analyzed by ESI FT-ICR mass spectrometry.

4.3.6 NMR spectroscopy

1D and 2D NMR spectra were recorded with a Bruker DRX-600 (^1H : 600 MHz, ^{13}C : 150 MHz) instrument equipped with a cryoprobe in $\text{CDCl}_3/\text{CD}_3\text{OD}$ (4 :1, v/v; residual CHCl_3 as internal standard, ^1H : $\delta = 7.26$ ppm; ^{13}C : $\delta = 77.0$ ppm) at 298 K. All two-dimensional homo- and heteronuclear experiments (COSY, TOCSY, HSQC-DEPT, and HMBC) were performed by using standard pulse sequences available in the Bruker software. The mixing time for TOCSY was 100 ms. ^{31}P NMR spectra were recorded with a Bruker Avance-DRX 400 (^{31}P : 160 MHz) instrument, and chemical shifts are relative to an external standard (phosphoric acid).

4.3.7 ESI FTICR mass spectrometry

Electrospray ionization Fourier-transform ion cyclotron resonance mass spectrometry was performed by using an APEX Qe (Bruker Daltonics) equipped with a 7 Tesla actively shielded magnet and a dual ESI/MALDI ion source. For negative-ion mode, samples (10 ng mL^{-1}) were dissolved in propan-2-ol/water/ Et_3N (50:50:0.001, v/v/v) and were sprayed at a flow rate of 2 mL min^{-1} . The capillary entrance voltage was set to 3.8 kV, and the dry gas temperature was set to 200°C .

4.3.8 Q-Exactive Plus Orbitrap

For MS/MS analyses in the positive-ion mode, the lipid A samples were analyzed with a Q-Exactive Plus (Thermo, Bremen, Germany) by using the H-ESI source with 4.0 kV, a sheath gas flow rate of 5 au, and a transfer capillary temperature of 250°C . The sample was dissolved in propan-2-ol/water/30mm ammonium acetate (50:50:4 v/v/v) adjusted with acetic acid to pH 4.5. For enhanced fragmentation analysis,

triethylamine was added as described (Kondakova et al. 2005), and measurements were performed in direct infusion mode by using a syringe pump at a flowrate of 5 mL min⁻¹. The S-lens RF (Radio Frequency) level was set to 100 and a source fragmentation of 50 eV was applied. The complete isotopic cluster of a precursor was selected by using 10 Da isolation window. The resolution was set to 280000 full width at half maximum (FWHM) defined at m/z 200. For MS/MS, normalized collision energy (NCE) was set to either 13 or 14 NCE. For further in-depth MS/MS analysis of selected molecular species, nano-ESI experiments were performed by using the Nanomate Triversa (Advion, Ithaca, US). The spray voltage was set to 1.1 kV, and the additional backpressure was set to 7.6 kPa. The source fragmentation was set to 100 eV, and the precursor selection was performed with unit resolution specifically to isolate the monoisotopic signal.

4.3.9 Reagents and cell cultures

The human myelomonocytic THP1 cell line was obtained from ATCC (Manassas, USA). Phorbol 12-myristate 13-acetate (PMA) was purchased from Sigma-Aldrich. *E. coli* O111:B4 LPS was purchased from InvivoGen (Invitrogen). Lipid A was reconstituted in DMSO to provide 1 mg mL⁻¹ stock solutions. Further dilutions were made in cell medium [RPMI1640 (Life Technologies)] supplemented with 10% fetal calf serum (FCS), l-glutamine (2 mM), sodium pyruvate (0.4 mM), and 2-mercaptoethanol (4 μM) so that the final amount of DMSO in the cell culture did not exceed 0.01%.

4.3.10 Biological activity assay in human macrophage-like cell line THP1

THP-1 cells were grown in RPMI1640 cell-culture medium supplemented with 10% FCS, l-glutamine (2mM), sodium pyruvate (0.4

mM), and 2-mercaptoethanol (4 μ M). Cells were seeded in a 96-well plate at 10^5 cells well⁻¹ in complete medium (140 μ L) and were stimulated with PMA (200 nM) for 24 h to induce differentiation into macrophage-like cells. The next day, the cells were washed with complete culture medium to discard the cells that did not adhere, refreshed with complete medium (180 μ L), and left for 1 h to recover. Cells were simultaneously stimulated with *C. psychrerythraea* lipid A in the presence or absence of *E. coli* O111:B4 LPS at the indicated concentrations. The *C. psychrerythraea* lipid A and *E. coli* O111:B4 LPS were added as solutions in complete medium (10 μ L) to reach a 200 μ L total volume of the well after stimulation. The cells were incubated overnight (18–24 h) at 37°C and 5% CO₂, and the supernatants were analyzed for TNF production by ELISA (BD Biosciences). Data are representative of three independent experiments. Error bars indicate standard error of the mean of triplicate samples.

References

- Bowman J.P. The family *Colwelliaceae*. In The prokaryotes-*Gammaproteobacteria*. 2014 (Ed.s) Rosenberg et al., Chapter 10.
- Carillo S. et al., *Chemistry* 2011, 17, 7053-7060.
- Carillo S. et al., *Eur J Org Chem*. 2013, 18, 3771-3779.
- Casillo A. et al., *ChemBioChem* 2017, 18, 1 - 11.
- Carty S.M. et al., *J. Biol. Chem.* 1999, 274, 9677 - 9685.
- Choma A. et al., *Biochim. Biophys. Acta Mol. Cell. Biol. Lipids* 2017, 1862, 196-209.
- Collins R.E. et al., *Environ. Microbiol.* 2010, 12, 1828-1841.
- Corsaro M.M. et al., *Chemistry*, 2008, 14, 9368-9376.
- Corsaro M.M. et al., *J Bacteriol.* 2004, 186, 29-34.
- D'Alonzo D. et al., *Chem. Eur. J.* 2016, 22, 11053-11063.
- Deming J.W. et al., *Syst Appl Microbiol.* 1988, 10, 152-160.
- Domon B., Costello C.E., *Glycoconjugate J.* 1988, 5, 397 - 409.
- Fuller B. J. *Cryo-Letters*, 2004, 25, 375 - 388.
- Galanos C. et al., *Eur. J. Biochem.* 1969, 9, 245-249.
- Hashimoto M. et al., *J. Basic Microbiol.* 2015, 55, 838-845.
- Helander I.M., Kilpelainen I., Vaara M., *Mol. Microbiol.* 1994, 11, 481- 487.
- Huston A.L. et al., *Appl Environ Microbiol* 2004, 70, 3321-3328.
- Jachymek W. et al., *Eur. J. Biochem.* 1999, 266, 53-61.
- Kenyon J.J. et al., *Glycobiology*, 2011, 21, 1131-1139.
- Kussak A., Weintraub A., *Anal. Biochem.* 2002, 307, 131-137.
- Marzabadi C.H., Franck R.W., *Chem. Eur. J.* 2017, 23, 1728 - 1742.

- Methé, B. A. et al., *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 10913-10918.
- Park B.S. et al., *Nature* **2009**, *458*, 1191-1195.
- Phillips N.J. et al., *J. Biol. Chem.* **2011**, *286*, 21203 - 21219.
- Silipo A. et al., *J. Lipid Res.* **2002**, *43*, 2188-2195.
- Silipo A. et al., *Carb. Research.* **2005**, *340*, 2693-2697.
- Sweet C.R. et al., *Mar. Drugs* **2015**, *13*, 4701 - 4720.
- Thomas D.N., Dieckmann G.S., *Science* **2002**, *295*, 641 - 644.
- Wang Z., Liu X., Li J., Altman E., *Carbohydr. Res.* **2008**, *343*, 483-488.

Chapter 5: *Psychrobacter arcticus* strain 273-4

Until now, only LPSs from marine Arctic (Corsaro et al. 2008; Carillo et al. 2013) and Antarctic (Carillo et al. 2011; Corsaro et al. 2001) Gram-negative microorganisms have been characterized, but very little is known about isolates from permafrost. It has been shown that viable bacteria are abundant in Siberian permafrost (Vishnivetskaya et al. 2000; Carillo et al. 2011), and the most frequently isolated from the Kolyma permafrost of northeast Siberia include *Arthrobacter*, *Exiguobacterium*, *Flavobacterium*, *Sphingomonas*, and *Psychrobacter* (Rodrigues et al. 2009; Rivkina et al. 2000; Vishnivetskaya et al. 2000). *Psychrobacter* is considered an indicator genus for permafrost and other polar environments (Shi et al. 1997), suggesting that many of its members are adapted to low temperatures and have evolved molecular-level changes that aid survival at low temperatures.

Psychrobacter arcticus 273-4 is a Gram-negative bacterium isolated from a 20,000-to-30,000-year-old continuously frozen permafrost horizon in the Kolyma region in Siberia that was not exposed to temperatures higher than 4 °C during isolation (Rivkina et al. 2000).

5.1 Lipopolysaccharides and lipid A structures

As already reported in a previous work (Casillo et al. 2015), the structural characterization of the carbohydrate backbone of the LOS of *Psychrobacter arcticus* 273-4 grown at 4 °C was already performed. The structure showed a particular inner core region, with a residue of glucose linked to the Kdo in place of a *manno*-heptose. This structural feature has been found only in *Colwellia psychrerythrae* 34H, which showed a

mannose residue linked to the Kdo (see paragraph 4.1).

Generally, the oligo- and polysaccharides produced by marine bacteria are distinguished by the acidic character (Muldoon et al. 2003), and by the occurrence of unusual sugars (Kenne et al. 1983), non-sugar substituents (Komandrova et al. 2010; Hanniffy et al. 1999; Nazareno et al. 2003; Carillo et al. 2013) or highly phosphorylated structures (Corsaro et al. 2004). Although *P. arcticus* 273-4 was isolated from Arctic permafrost it displays similar characteristics of cold-adapted marine isolates, due to the presence of the unusual residue of *N*-acetylmuramic acid (NAM), commonly encountered as a component of bacterial cell-wall peptidoglycan, already found in the O-specific polysaccharide of *Yersinia ruckerii* (Beynon et al. 1994) and *Proteus penneri* (Zych et al. 1998), but found for the first time in an oligosaccharide core (Casillo et al. 2015). It is well known that cold-adapted microorganisms are able to modify the fluidity of the cellular membrane in response to a lowering of temperature by producing a higher content of unsaturated, polyunsaturated and methyl-branched fatty acids in phospholipids (Chintalapati et al. 2004; Driessen et al. 1996). Instead, how bacteria modify the glycolipid portion of LPS structures in response to the cold stress is still poorly understood. Even though only few LPS structures from cold-adapted bacteria have been characterized (Corsaro et al. 2008; Carillo et al. 2013; Corsaro et al. 2001; Carillo et al. 2011), their attractive feature is the production of rough lipopolysaccharides. Moreover, it is worth noting that *Psychrobacter arcticus* 273-4, a permafrost isolate, shares this feature with marine isolates. Thus, the LOS isolated from *P. arcticus* 273-4 grown at 4°C was hydrolyzed under mild acidic conditions to cleave the unstable Kdo glycosidic linkage between the lipid A and the saccharidic region. After centrifugation, the supernatant containing the

core oligosaccharidic portion of the LOS was separated from a precipitate constituted by the lipid A. After further purifications steps, the characterization of lipid A portion, has been a topic of this thesis. The lipid A from *P. arcticus* grown at 25°C, was already reported (Korneev et al 2014). During this thesis, we therefore speculated and evaluated the influence of the temperature on the lipid A structure.

5.1.1 Isolation and compositional analysis of lipid A

The extraction of *P.arcticus* 273-4 dried cells by hot phenol-water method gave LPS in about 3.1% yield (Casillo et al. 2015). The LPS was subjected to 5% acetic acid hydrolysis to obtain lipid A fraction. The lipid A was recovered as pellet after centrifugation of the crude reaction, and washed twice with water and then freeze-dried. The fatty acid methyl ester derivatives were obtained after methanolysis of lipid A sample and extraction with hexane, and analyzed by gas chromatography mass spectrometry (GC-MS) (Figure 5.1).

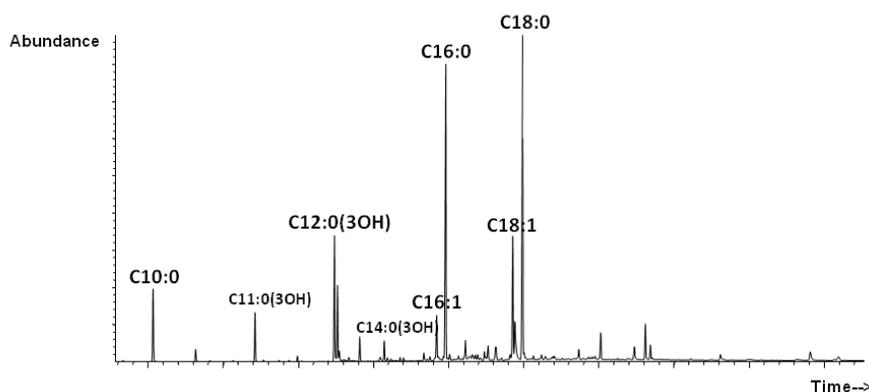


Figure 5.1: Chemical composition of lipid a portion.

The analysis showed the presence of mainly 3-hydroxy-dodecanoic acid [C12:0(3-OH)], 3-hydroxy-tetradecanoic acid [C14:0(3-OH)], decanoic

acid (C10:0), and a minor amount of 3-hydroxy-undecadecanoic acid [C11:0(3-OH)], and 3-hydroxy-tridecanoic acid [C13:0(3-OH)]. In addition, signals attributable to longer chains (C16:0; C:16:1; C18:0; C18:1) suggested the presence of membrane phospholipids. Phospholipids contaminants were removed from LPS by extraction with a chloroform-methanol mixture, and a very low contamination due to phospholipids was revealed by GC-MS. The presence of phosphate groups was deduced from NMR spectra of delipidated core oligosaccharide and confirmed by mass spectrometry (Casillo et al. 2015).

5.1.2 Mass spectrometry analysis of Lipid A

The charge-deconvoluted ESI FT-ICR mass spectrum obtained in negative ion mode (Figure 5.2) revealed for the lipid A from *P. arcticus*, mainly three signal clusters (**M1-M3**) representing tetra-, penta-, and hexa-acylated lipid A species (Table 5.1).

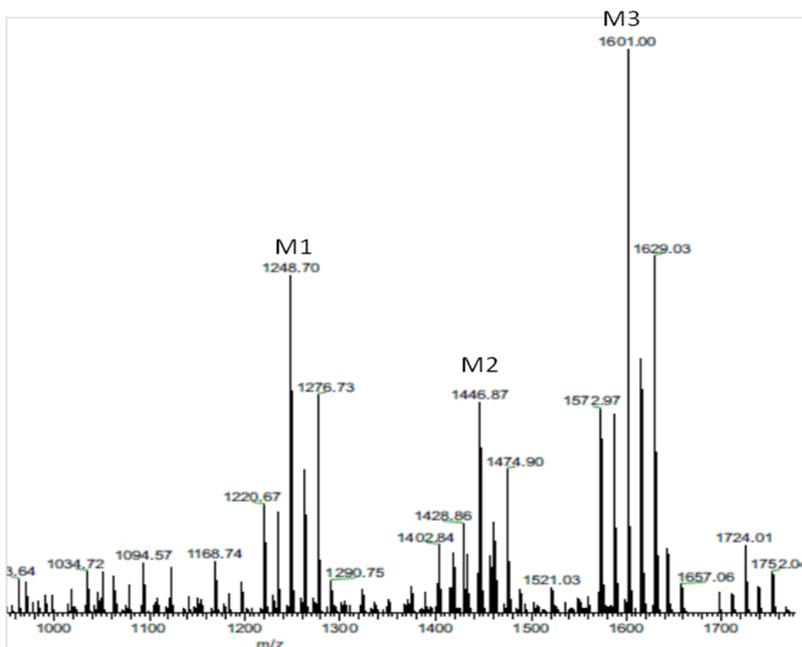


Figure 5.2: Charge-deconvoluted negative ions ESI FT-ICR mass spectrum of lipid A from *P. arcticus* 273-4.

Table 5.1. Composition of the main species present in the charge-deconvoluted negative ions ESI FT-ICR mass spectrum of the lipid A from *P. arcticus* 273-4.

Species	[M-H] ⁻ Calculated	[M-H] ⁻ Observed	Composition
M1	1248.715	1248.70	GlcN ₂ P ₂ [C12:0(3-OH)] ₃ (C10:0)
M2	1446.880	1446.87	GlcN ₂ P ₂ [C12:0(3-OH)] ₁ (C10:0)
M3	1601.019	1601.00	GlcN ₂ P ₂ [C12:0(3-OH)] ₁ (C10:0) ₂
M4	1724.016	1724.01	GlcN ₃ P ₂ PEtN[C12:0(3-OH)] ₁ (C10:0) ₂

Chemical analysis and mass spectra results, allowed to assign to the tetra-acylated specie **M1** (calculated molecular mass 1248.715 Da) the following composition: GlcN₂P₂[C12:0(3-OH)]₃(C10:0). The specie **M2** at 1446.87 *m/z* (calculated molecular mass 1446.880 Da), was identified

as penta-acylated lipid A with an additional C12:0(3-OH). Finally, species at 154.13 u higher, indicated as **M3**, differs from the **M2** for the additional presence of a decanoic residue. This species at 1601.00 *m/z* (calculated molecular mass 1601.019 Da), was attributed to the hexa-acylated lipid A, with the following composition: GlcN₂P₂[C12:0(3-OH)]₄(C10:0)₂. Moreover, **M4** signal cluster displays an additional phosphoethanolamine respect to **M3**. All the species display heterogeneity caused by variation of the fatty acids chain length (± 14.01 Da, ± 28.03 Da), with the replacement of one or more C12:0(3-OH) chains with C14:0(3-OH) or C11:0(3-OH), together with the lack of a phosphate group (79.96 Da).

To understand the position of the secondary fatty acids, the lipid A from *P.arcticus* was treated with conc. NH₄OH (Silipo et al. 2002), a procedure that allows to hydrolyse acyl and acyloxacyl esters, leaving intact acyl and acyloxacyl amides.

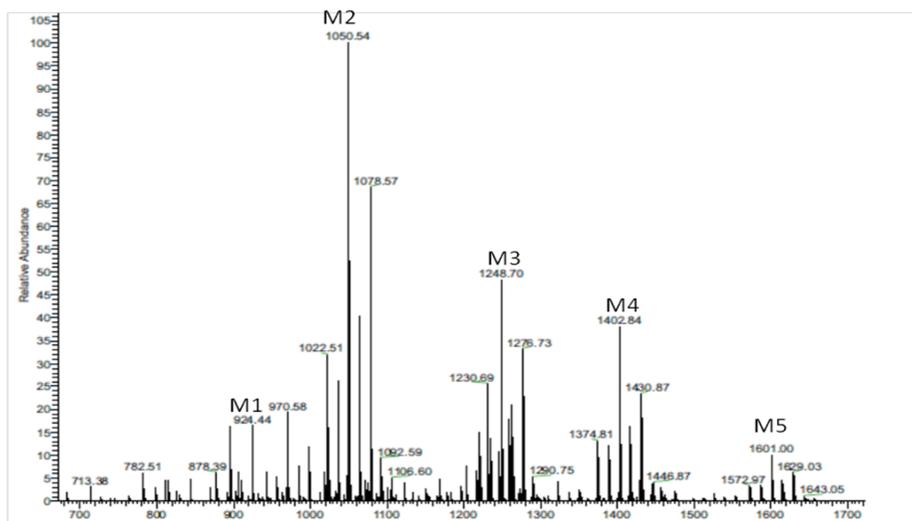


Figure 5.3: Charge-deconvoluted negative ions ESI FT-ICR mass spectrum of the Lipid A_{NH₄OH} from *Psychrobacter arcticus* 273-4.

Table 5.2. Composition of the main species of the *Psychrobacter arcticus* 273-4 Lipid $A_{\text{NH}_2\text{OH}}$ observed in the ESI FT-ICR mass spectrum.

Species	[M-H] ⁻ calculated	[M-H] ⁻ observed	Composition
M1	896.411	846.41	GlcN ₂ P ₂ [C12:0(3OH)] ₂
M2	1050.550	1050.54	GlcN ₂ P[C12:0(3OH)] ₂ [C10:0]
M3	1248.715	1248.70	GlcN ₂ P ₂ [C12:0(3OH)] ₃ [C10:0]
M4	1402.854	1402.84	GlcN ₂ P ₂ [C12:0(3OH)] ₃ [C10:0] ₂
M5	1601.019	1601.00	GlcN ₂ P ₂ [C12:0(3OH)] ₄ [C10:0] ₂

The charge deconvoluted negative ions ESI FT-ICR mass spectrum of the lipid $A_{\text{NH}_2\text{OH}}$ (Figure 5.3; Table 5.2) revealed the presence of five main signals. The most abundant signal **M2** with measured mass of 1050.54 u corresponded to a molecular species the composition of which was identified as GlcN₂P₂[C12:0(3-OH)]₂(C10:0) (calculated molecular mass 1050.550 Da), thus indicating that the decanoic acid was linked as acyloxyamide. In order to obtain the exact distribution of fatty acids on each glucosamine residue, the lipid A was analyzed by MS and MS/MS in the positive ion mode, performed with a Q-Exactive Hybrid Quadrupole Orbitrap.

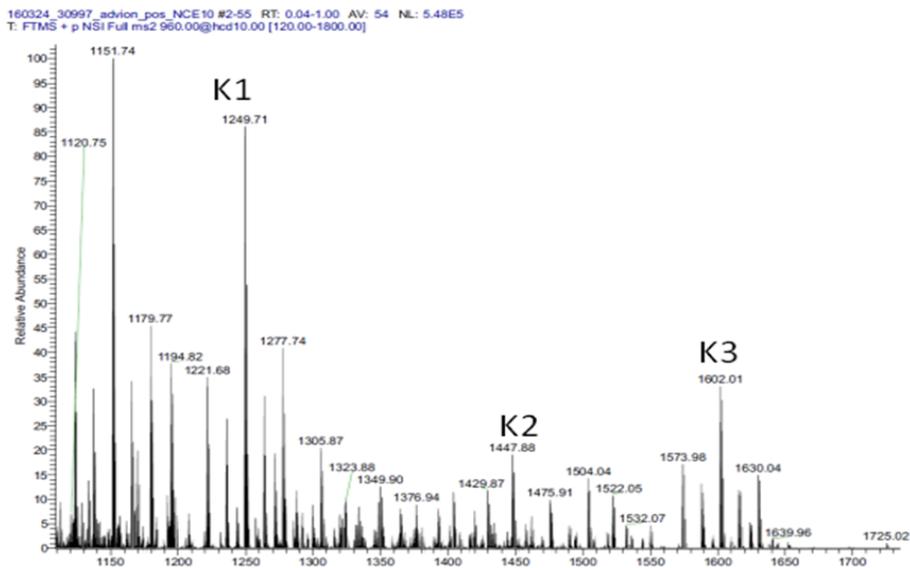


Figure 5.4: Positive-ion Q-Exactive Hybrid mass spectrum of lipid A from *Psychrobacter arcticus* 273-4.

The positive ions MS spectrum showed the presence of three species **K1-K3** (Figure 5.4), as found in negative ion mode, even if a different relative intensity among the signals was detected.

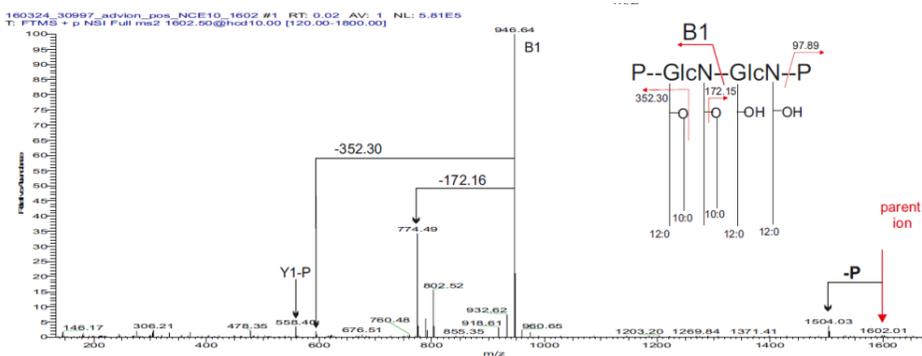


Figure 5.5: MS/MS spectrum of m/z 1602.01.

The glycoform at m/z 1602.01, corresponding to the hexacyl-diphosphorylated specie, was selected for the fragmentation. The MS/MS spectrum showed the presence of a **B** fragment ion at m/z 946.64 Da, corresponding to $\text{GlcNP}[\text{C12:0}(3\text{-OH})_2(\text{C10:0})_2]$, thus

indicating that both the secondary fatty acids were linked to the distal non-reducing end glucosamine (Figure 5.5).

5.1.3 Evaluation of biological activity

Preliminary biological activity assays were performed to characterize the effect of *Psychrobacter arcticus* 273-4 lipid A on TNF secretion in the presence of a serial dilution of *Escherichia coli* O111:B4 LPS, as described for *C. psychrerythraea* in paragraph 4.1.4.

As shown in Figure 5.6, *P. arcticus* lipid A seemed to be quite agonist on LPS-induced TNF production in THP-1 cells and displays only a slight agonistic effect at the highest concentration. However, further experiments to evaluate these properties in details are needed.

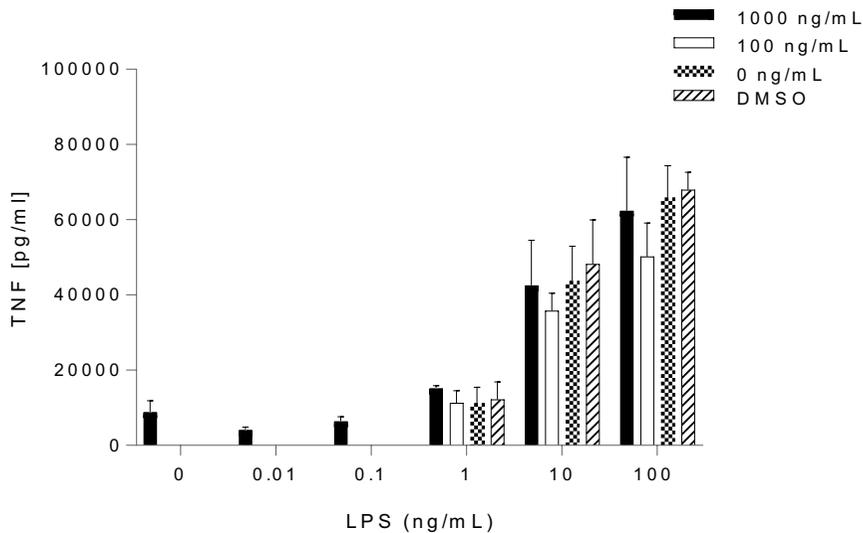


Figure 5.6: Effect of *P. arcticus* lipid A on LPS-induced TNF production in THP-1 cells. Cells were incubated with the indicated amounts of *P. arcticus* lipid A in the presence or absence of different concentrations of *E. coli* O111:B4 LPS, as indicated.

5.2 Conclusions

In addition to the complete structure of the sugar backbone of the LPS from the permafrost isolate *Psychrobacter arcticus* 273-4, that was already reported (Casillo et al 2015), here we investigated the structure of the lipid A portion. The results of chemical and mass spectrometry analyses revealed a symmetric primary acylation pattern, substituted at positions 2, 3 and 2', 3' by two 3-hydroxydodecanoic acids, respectively. Instead, the secondary acylation pattern is characterized by the presence of two shorter decanoic acid (C10:0) chains linked to the distal non-reducing end glucosamine. In addition, the level of micro-heterogeneity in the *P. arcticus* structure is characterized by the presence of 3-hydroxy-tetradecanoic acid [C14:0(3-OH)] and minor amount of 3-hydroxy-undecadecanoic acid [C11:0(3-OH)] and 3-hydroxy-tridecanoic acid [C13:0(3-OH)] as primary acyl groups instead of the 3-hydroxy-dodecanoic acid [C12:0(3-OH)]. Biological activity assays performed in the human macrophage cell line THP-1 show an agonistic effect of *Psychrobacter arcticus* lipid A on tumor necrosis factor production in these cells and a slight agonistic effect at the highest concentration used.

5.3 Experimental section

5.3.1 Bacteria growth and LPS isolation

P. arcticus strain 273-4, was isolated from permafrost soil located in Siberia. Shake flask cultivation were performed in Luria-Bertani broth (Sambrook et al. 2001) at 4°C in aerobic condition. When the liquid cultures reached late exponential phase (about 90 h, OD_{600nm} 4) cells were collected by centrifugation for 15 minutes at 7000 rpm at 4 °C. Dried bacterial cells (3.1 g) were extracted first by PCP method to give very

poor yield of LOS (0.03% w/w of dried cells) and then by hot phenol/water method (Galanos et al. 1969; Westphal et al. 1965). A 240 mg amount of water extract was dialyzed (cut-off 3500 Da) and then digested with proteases, DNAses and RNAses to remove contaminating proteins and nucleic acids. The sample was dialyzed (cut-off 3500 Da) to afford 96 mg of LOS sample (yield 3.1 % w/w of dried cells).

5.3.2 Mild acid hydrolysis

The LOS (20 mg) was hydrolyzed with 1% aqueous CH_3COOH (2 mL, 100 °C for 4 h). The resulting suspension was then centrifuged (7500 rpm, 4 °C, 30 min) and the pellet was washed twice with water, and two fractions were recovered: lipid A (7.5 mg) as pellet and the saccharide portion (12.5 mg) as supernatant.

5.3.3 Chemical Analysis

The lipid A fraction (0.5 mg) was analyzed as derivatives methyl glycoside acetylated (AMG) as indicated in paragraph 4.3.3.

5.3.4 NH_4OH hydrolysis of lipid A

Lipid A (0.5 mg) was incubated with conc. aqueous NH_4OH (200 μL) as reported (Silipo et al. 2002). The sample was dried and analyzed by ESI FT-ICR mass spectrometry.

5.3.5 Mass spectrometry analysis

Performed by Prof. B. Lindner and Dr. D. Schwudke, Division of Bioanalytical Chemistry, Leibniz-Center for Medicine and Biosciences, Germany. Electrospray ionization Fourier transform ion cyclotron (ESI FT-ICR) mass spectrometry was performed in negative ion mode using an APEX QE (Bruker Daltonics) instrument equipped with a 7 Tesla actively shielded magnet. For negative ions mode, samples ($\approx 10 \text{ ng mL}^{-1}$)

were dissolved in 2-propan-ol/water/Et₃N (50:50:0.001, v/v/v) and sprayed at a flow rate of 2 $\mu\text{L min}^{-1}$. The capillary entrance voltage was set to 3.8 kV, and the dry gas temperature to 200°C.

5.3.6 Q-Exactive hybrid quadrupole orbitrap

For MS/MS analyses in the positive-ion mode, the lipid A samples were analyzed with a Q-Exactive Plus (Thermo, Bremen, Germany) by using the H-ESI source with 4.0 kV, a sheath gas flow rate of 5 Da, and a transfer capillary temperature of 250°C. The sample was dissolved in 2-propan-ol/water/30 mM ammonium acetate (50:50:4, v/v/v) adjusted with acetic acid to pH 4.5. For enhanced fragmentation analysis, triethylamine was added as described in literature (Nazarenko et al. 2003), and measurements were performed in direct infusion mode by using a syringe pump at a flowrate of 5 $\mu\text{L min}^{-1}$. The S-lens RF (Radio Frequency) level was set to 100, and a source fragmentation of 50 eV was applied. The complete isotopic cluster of a precursor was selected by using 10 Da isolation window. The resolution was set to 280000 full width at half maximum (FWHM) defined at m/z 200. For MS/MS, normalized collision energy (NCE) was set to either 13 or 14 NCE. For further in-depth MS/MS analysis of selected molecular species, nano-ESI experiments were performed by using the Nanomate Triversa (Advion, Ithaca, US). The spray voltage was set to 1.1 kV, and the additional backpressure to 7.6 kPa. The source fragmentation was set to 100 eV, and the precursor selection was performed with unit resolution specifically to isolate the monoisotopic signal.

5.3.7 Reagents and cell cultures

See paragraph 4.3.9.

5.3.8 Biological activity assay in human macrophage-like cell line THP1

See paragraph 4.3.10.

References

- Beynon, L.M.; Richards, J.C.; Perry, M.B. *Carbohydr. Res.* **1994**, *256*, 303-317.
- Casillo A. et al., *Mar. Drugs* **2015**, *13*, 4539-4555.
- Carillo, S. et al., *Chem. Eur. J.* **2011**, *17*, 7053-7060.
- Carillo, S. et al., *Eur. J. Org. Chem.* **2013**, 3771-3779.
- Chintalapati, S.; Kiran, M.D.; Shivaji, S. *Cell. Mol. Biol.* **2004**, *50*, 631-642.
- Corsaro, M.M. et al., *Eur. J. Biochem.* **2001**, *268*, 5092-5097.
- Corsaro, M.M. et al., *J. Bacteriol.* **2004**, *186*, 29-34.
- Corsaro, M.M. et al., *Chem. Eur. J.* **2008**, *14*, 9368-937.
- Driessen, J.M. et al., *Fems Microbiol. Rev.* **1996**, *18*, 139-148.
- Galanos, C., Lüderitz, O., Westphal, O. *Eur. J. Biochem.* **1969**, *9*, 245-249.
- Kenne, L., Lindberg, B. In *The Polysaccharides*. Ed. Aspinall G. O., Academic Press: New York, Volume 2, 287-363, **1983**.
- Korneev K.V., Kondakova A.N., Arbatsky N.P., et al. *Biochemistry* **2014**, *79*, 1333.
- Muldoon J. et al., *Carbohydr. Res.* **2003**, *338*, 459-462.
- Silipo A. et al., *J. Lipid Res.* **2002**, *43*, 2188-2195.
- Nazareno E.L. et al., *Carbohydr. Res.* **2003**, *338*, 2449-2457.
- Rivkina E.M. et al., *Appl. Environ. Microbiol.* **2000**, *66*, 3230-3233.
- Rodrigues D.F. et al., *ISMEJ.* **2009**, *3*, 658-665.
- Sambrook J.; Russell D.W. In *Molecular cloning: a laboratory*

manual, 3rd ed.; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, **2001**.

- Shi T. et al., *Microb. Ecol.* **1997**, *33*, 169–179.
- Vishnivetskaya T. et al., *Extremophiles* **2000**, *4*, 165–173.
- Westphal O. et al., *Methods Carbohydr. Chem.* **1965**, *5*, 83–91.
- Zych K. et al., *Fems Immunol. Med. Mic.* **1998**, *21*, 1–9.

SECTION III

Semisynthetic derivatives from *Escherichia coli*
lipid A

Chapter 6: Multigram-scale extraction and purification of lipid A from *E. coli*

Lipid A is involved in binding to Toll-like receptor 4 (TLR4) and responsible for the potent inflammatory effects of LPS, that may be associated with a variety of severe symptoms such as fever, diarrhea, blood pressure drop, and septic shock (see paragraph 1.4.2). Nonetheless, the studies about the structure–activity relationship of natural lipid A and various synthetic analogues have shown that the phosphorylation grade as well as the number, length, and distribution of acyl chains on lipid A are crucial for its endotoxicity (see paragraph 1.3.1) (Li et al. 2013; Rietschel et al. 1994). For example, it is well-known that diphosphorylated and highly acylated lipid As, such as the major hexa-acylated species from *Escherichia coli* (**1**, Figure 6.1), are generally optimally recognized by TLR4 (Park et al. 2009), but their clinical use is precluded due to high toxicity. However, their strong inflammatory activity can be significantly reduced by removing the anomeric phosphate group. The reduced toxicity of monophosphoryl lipid A (**2**, Figure 6.1) has been attributed to the selective induction of the TLR-4–TRIF signaling pathway over the TLR-4–MyD88 one (Mata-Haro et al. 2007), leaving the immunostimulatory properties unaffected. Furthermore, underacylated (three or four chains) or overacylated (seven chains) lipid A species can often inhibit TLR4, thus acting as TLR4-antagonists in humans. Therefore, therapeutic regulation of the innate immune response by suitable lipid A structures is an interesting approach towards TLR4-active immunomodulators, that can be useful for the treatment of several pathologies such as allergies, neuroinflammation and cancer, as well for a highly efficient formulation of vaccines (Marzabadi et al. 2017).

Indeed, the monophosphoryl lipid A from *Salmonella enterica* ser. Minnesota RC 595, selectively deacylated at position *O*-3 of GlcN I (MPL[®] **3**, Figure 6.1) (Mata-Haro et al. 2007), was recently introduced for human use (Casella et al. 2008) in HPV vaccine Cervarix and HBV vaccine Fendrix in combination with Alum, an already employed immunoadjuvant. Moreover, MPL[®] has entered clinical trials stage investigations for its employment as an adjuvant in the therapy of a number of different types of cancer too: stage IV melanoma, ovarian cancer, lung cancer, leukemia, sarcoma, Merkel cell carcinoma, and non-Hodgkins lymphoma (Marzabadi et al. 2017; for recent and ongoing studies see: <https://clinicaltrials.gov>).

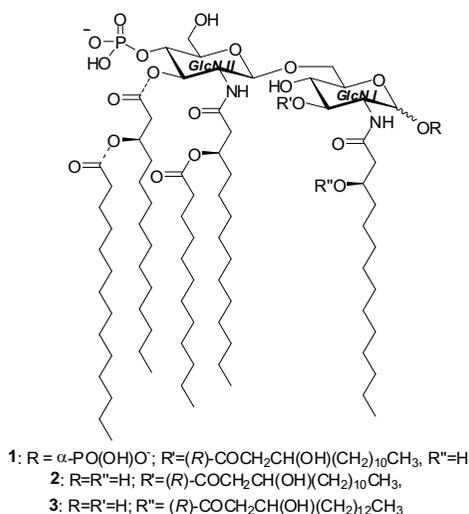


Figure 6.1: Chemical structure of the lipid A species extracted from *E. coli* (diphosphoryl lipid A **1**, monophosphoryl lipid A **2**) and MPL[®] from *S. Minnesota* **3**.

Given these preliminary remarks, the research of new, efficient immunomodulators has increased in the last years, not only in terms of purification, structural characterization and biological evaluation of lipid A species from wild-type or engineered bacterial extracts (Molinaro et al. 2015; Choma et al. 2017; Silipo et al. 2014; Casillo et al. 2017; Di

Lorenzo et al. 2017; Needhama et al. 2013; Carillo et al. 2011), but also of total chemical synthesis of suitably designed lipid A structures (Gao et al. 2017) including conjugates between lipid A and Tumor-associated carbohydrate antigens (TACAs) (Liao et al 2016; Zhou et al. 2015; Liao et al 2015; Wang et al 2012) or mimics thereof (Peri et al. 2014). Nonetheless, both approaches suffer some problems: in the former case they are mainly related to the difficulties of obtaining highly pure lipid A samples, due to frequent contamination with other biomolecules, such as glycerophospholipids. On the other hand, the total synthetic approach requires long and complex procedures with a high number of chemical steps, which hamper an application on a large scale. A third approach has recently emerged. It is based on semi-synthetic strategies, that combine the advantages of fermentative and scalable processes affording complex natural products in high quantity with the development of site-selective chemical reactions modifying these structures. (Pieretti et al. 2014) Indeed, the research of an optimized fed-batch fermentation/phenol-free purification procedure for the gram scale production of lipid A 1 from *Escherichia coli* K4 (Pieretti et al. 2014), and site-selective manipulations of its structure (D'Alonzo et al. 2016) represent an emerging, alternative and smart strategy for the obtainment of novel semisynthetic lipid A immunoadjuvant candidates.

6.1 Isolation, purification and compositional analysis of lipid A from *E. coli* K4 fermentation

Lipid As are typically isolated in milligram quantities from bacterial cells through LPS extraction followed by selective hydrolysis of the labile glycosidic bond generally linking the lipid A moiety from the rest of the

LPS macromolecule. In alternative, during this PhD work, a successful protocol for the gram-scale production of **1** was repeated several times. To this aim, the first step is focused on a fed- batch fermentation process that overproduces LPS. One of the most common fermentation strategies used to reach high cell density cultivations is the dissolved-oxygen (DO)-stat method. This strategy is based on the evidence that the concentration of dissolved oxygen increases rapidly when all the substrate in the medium is consumed; therefore, it allows to rapidly respond to the microorganisms metabolic needs avoiding the accumulation of toxic byproducts and addressing the metabolism towards the product of interest, also reducing production costs (Shay et al. 1987). This strategy was previously exploited on small-scale bioreactors for the fermentation of the recombinant *E. coli* K4/ EcK4r3 strain (Cimini et al. 2013) and then repeated successfully on a pilot scale (90 L). The LPS material was generally collected from the fermentation broth through centrifugation and protease digestion followed by concentration/diafiltration and ultrafiltration purification steps. Selective cleavage of the 3-deoxy-octulosonic acid glycosidic bond linking lipid A and inner core in the LPS was then exploited by mild acid hydrolysis to give lipid A as a pellet from the crude mixture. From the chemical derivatization of a small aliquot of this material to fatty acid methyl esters (FAMEs, for the analysis of lipids) and peracetylated methyl *O*-glycosides (for the analysis of carbohydrates) followed by GC-MS analysis, it was possible to reveal the presence of several contaminants together with lipid A. Indeed, together with the expected 12:0, 14:0, and 14:0 (3-OH) FAMEs, 16:0 and 18:1 FAME lipid chains, which are typical constituents of glycerophospholipids from bacterial membranes (Oursel et al. 2007), were also detected. Saccharide contaminants were found too, as

indicated by the presence of several monosaccharide peaks in the GC chromatogram (Figure 6.2). Glucose (Glc), galactose (Gal), heptose, ribose, and quinovosamine (QuiN) could be associated to the presence in the pellet of residual *E. coli* K4 LPS and/or O-chain and core fragments thereof (MacLean and Perry 1997). *E. coli* K4 capsular polysaccharide and/or fragments thereof are also detected, by the peaks of 2-amino-2-deoxy-galactose (GalN) and glucuronic acid (GlcA) the two monosaccharide constituents of the chondroitin backbone of the *E. coli* K4 capsular polysaccharide (Rodriguez et al. 1988; Bedini et al. 2011).

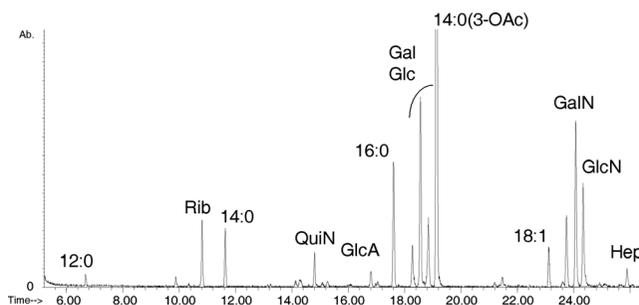


Figure 6.2: GC-MS analysis of crude lipid A from *E. coli* HCl hydrolysis step.

As GC-MS analysis revealed, further purification of the lipid A material was mandatory. Most of the saccharide contaminants were eliminated in two purification steps, by filtrating over Celite pad and then selectively precipitating the lipid A in a 3:1 v/v methanol-chloroform mixture at a 3 g/L concentration. In this way, the precipitate was enriched in the lipid A species, whereas the supernatant fraction (SF) contained glycerophospholipids and a minor quantity of lipid A. The latter could be recovered in turn by reiterating the precipitation process. The MALDI-TOF mass spectrometry analysis (Figure 6.3) of purified lipid A could confirm its structure, including a natural heterogeneity in the lipid

portion. Indeed, the most intense signal (1797.18 m/z) could be attributed to the major hexa-acylated species. The less intense signal at 1360.47 m/z could be associated to a minor species lacking the C14:0 (3-OH) chain at position O-3 of GlcN II, whereas the very small peak at 1587.19 m/z to another minor species lacking only the C14:0 one.

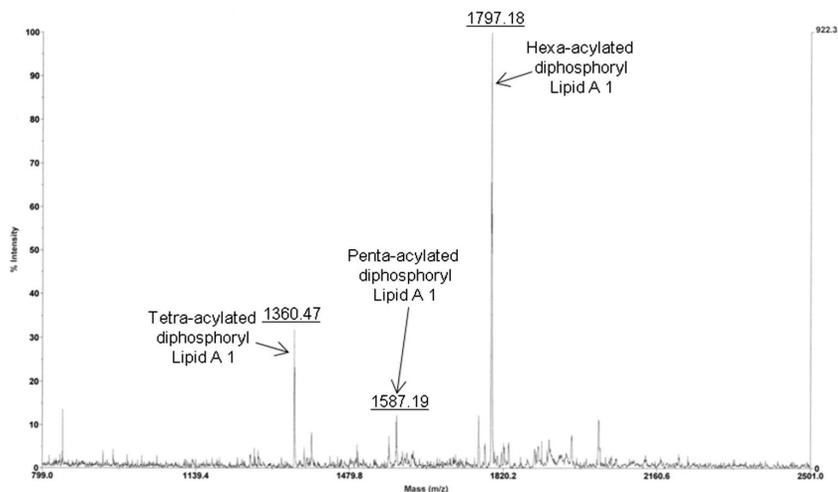


Figure 6.3: MALDI-TOF mass spectrum (negative ions mode) of lipid A 1.

6.2 Selective de-O-phosphorylation of lipid A

It is well-known that diphosphorylated and highly acylated lipid As, such as the major hexa-acylated species from *Escherichia coli* (1, Figure 6.1), are generally optimally recognized by TLR4,⁷ but their clinical use is precluded due to high toxicity. However, their strong inflammatory activity can be significantly reduced by removing the anomeric phosphate group. With *E. coli* diphosphoryl lipid A 1 in hands, the subsequent step was the regioselective cleavage of the phosphate group at the anomeric position of GlcN I, in order to weaken the endotoxic activity of the lipid A. The reaction conditions had to be mild enough not only to differentiate between the two phosphate groups but also to avoid the

cleavage of other acid-labile bonds such as the glycosidic connection between the two GlcNs or the ester linkages. Furthermore, the cleavage of the phosphate group at the anomeric position of GlcN I had to be quantitative. All these constraints were necessary in order to avoid the production of non-natural structural heterogeneity in monophosphoryl lipid A. Among the several reaction conditions tested, TLC and MALDI-TOF mass spectrometry analysis revealed that, on a gram scale, a regioselective and quantitative de-*O*-phosphorylation without modification of the acylation pattern was achieved exclusively by treating lipid A **1** with aqueous HF. The regioselective cleavage of the phosphate on the GlcN I over that at position *O*-4 on GlcN II was both explained with the well known higher reactivity of the anomeric position to the acidic hydrolysis conditions and the steric hindrance of the latter, due to the presence of a bulky acyloxy-acyl ester moiety at position *O*-3 of GlcN II (Tirsoaga et al. 2007).

6.3 Structural characterization and biological assays

The obtained MPLA **2** was purified by silica-gel column chromatography using a suitably optimized gradient elution. Pure MPLA was then subjected to MALDI-TOF mass spectrometry analysis and two-dimensional NMR (Ribeiro et al. 1999) characterization (Figures 6.4-6.5 and Table 6.1). In particular, the presence of signals for both α - and β -configured anomeric CH of GlcN I ($\delta_{\text{H}}=4.95$, $\delta_{\text{C}}=90.6$ ppm and $\delta_{\text{H}}=4.31$, $\delta_{\text{C}}=101.1$ ppm, respectively) confirmed the lack of the anomeric phosphate group and the presence of a free hemiacetal moiety.

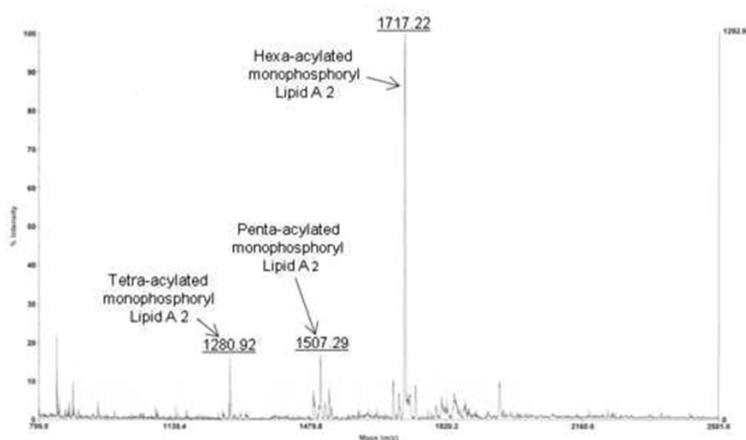


Figure 6.4: MALDI-TOF mass spectrum (negative ions mode) of MPLA 2.

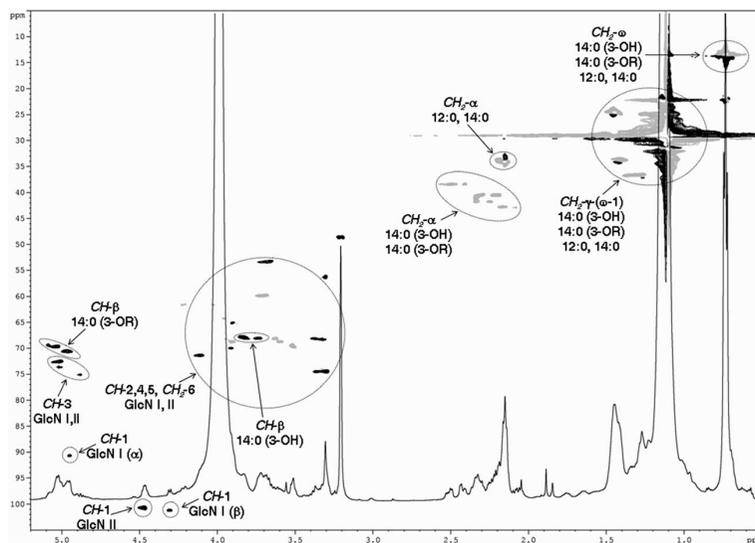


Figure 6.5: HSQC-DEPT (600MHz, 298 K, 4:1 v/v CDCl₃-CD₃OD) spectrum of MPLA 2.

Table 6.1: NMR characterization data of monophosphoryl lipid A **2**.

Sugar residue ^{a,b}	1	2	3	4	5	6	Fatty acid ^{a,b}	$\alpha\alpha$, $\alpha\beta$	β	γ	δ -(ω -1)	ω
GlcN I (α)	4.95, <i>90.6</i>	3.64, <i>53.2</i>	5.02, <i>73.6</i>	3.32, <i>68.2</i>	3.90, <i>65.1</i>	3.60, 3.91, <i>68.7</i>	14:0 (3-OH) 2- GlcN I	2.08, 2.17, <i>42.8</i>	3.73, <i>68.1</i>	1.23, 36.6	1.08–1.18, <i>29.1</i>	0.71–0.75, <i>13.3</i>
GlcN I (β)	4.31, <i>101.1</i>	3.64, <i>53.2</i>	4.88, <i>75.0</i>	3.37, <i>68.0</i>	3.91, <i>69.9</i>	3.63, 3.95, <i>68.0</i>	14:0 (3-OH) 3- GlcN I	2.25, 2.33, <i>41.7</i>	3.83, <i>67.9</i>	1.28, 36.5	1.08–1.18, <i>29.1</i>	0.71–0.75, <i>13.3</i>
GlcN II	4.47, <i>100.9</i>	3.68, <i>53.4</i>	5.03, <i>72.6</i>	4.12, <i>71.4</i>	3.32, <i>74.5</i>	3.70, <i>59.7</i>	14:0 (3-OR) 2- GlcN II	2.20, 2.30, <i>40.4</i>	4.96, <i>70.5</i>	1.41, 33.5	1.08–1.18, <i>29.1</i>	0.71–0.75, <i>13.3</i>
							14:0 (3-OR) 3- GlcN I	2.42, 2.51, <i>38.0</i>	5.05, <i>69.7</i>	1.41, 33.5	1.08–1.18, <i>29.1</i>	0.71–0.75, <i>13.3</i>
							12:0	2.15, <i>34.1</i>	1.44, <i>24.4</i>	1.08–1.18, <i>29.1</i>	1.08–1.18, <i>29.1</i>	0.71–0.75, <i>13.3</i>
							14:0	2.15, <i>34.1</i>	1.44, <i>24.4</i>	1.08–1.18, <i>29.1</i>	1.08–1.18, <i>29.1</i>	0.71–0.75, <i>13.3</i>

^{a1} H chemical shift in plain; ¹³C chemical shift in italic^b Data referred to the major hexa-acylated lipid A species

The endotoxic activity of monophosphoryl lipid A **2** was compared with the diphosphoryl lipid A **1** and the supernatant fraction (SF) contained glycerophospholipids, by profiling the cytokine production in vitro. An attenuated activity was detected for MPLA, which exhibited a behavior rather similar to the recently introduced immunomodulator monophosphoryl lipid A from *Salmonella* Minnesota R595 (MPL®) **3**, whereas **1** showed higher cytokine production levels (Pieretti et al., 2014) (Figure 6.6).

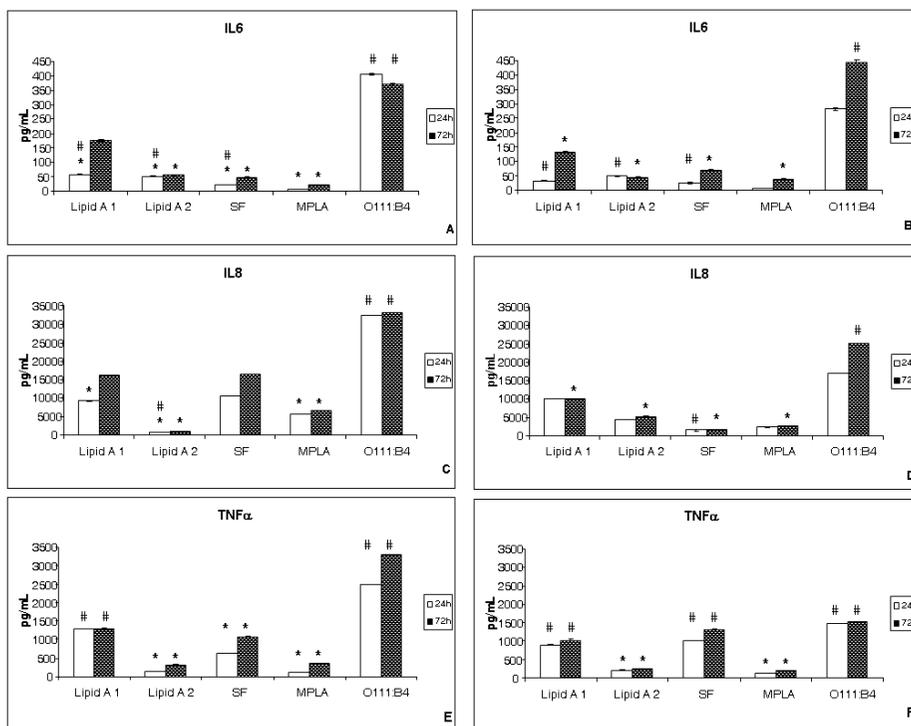


Figure 6.6: Cytokine concentrations in tissue-culture supernatants of THP1 cells stimulated by increasing concentrations of SF, **1** and **2** (1 μ g/mL panel A, C, E; 20 μ g/mL panel B, D, F). Significant differences between stimulations were shown (* $p < 0.05$ vs O111:B4; # $p < 0.05$ vs MPLA). The experiments were performed in triplicates and data were shown as means \pm standard error.

6.4 Conclusions

In conclusion, monophosphoryl lipid A **2** with weakened endotoxic activity was obtained from an optimized fed-batch fermentation of *E. coli* K4, followed by a gram-scale, phenol-free extraction protocol suitably developed for the purification of lipid A from the fermentation broth, and then by a regioselective de-*O*-phosphorylation. The sequence of downstream steps for the obtainment of pure monophosphoryl lipid A **2** from the fermentation broth is depicted in Figure 6.7.

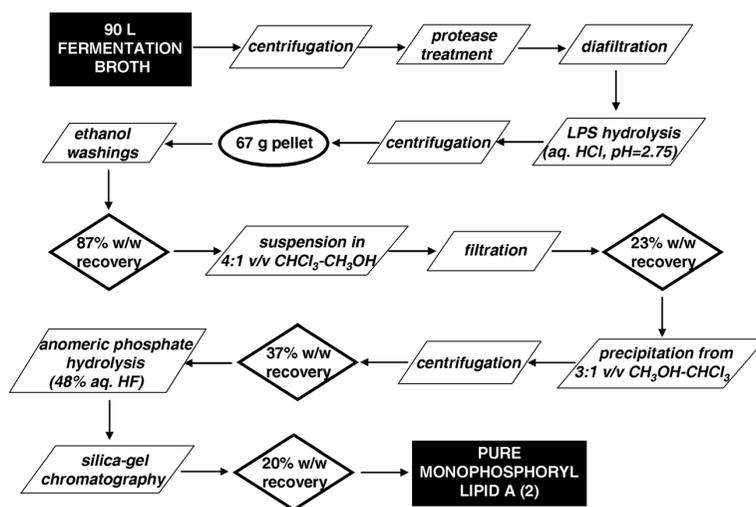


Figure 6.7: Sequence of purification and chemical transformation steps from the fermentation broth to pure monophosphoryl lipid A **2**.

The scalability of this combined fermentative-chemical approach allowed the obtainment of MPLA **2** in large amount and consequently its employment as a scaffold for the semi-synthesis of several potential immunostimulating lipid As and derivatives thereof through few tailored chemical modifications (see following chapters).

6.5 Experimental section

6.5.1 Fermentation of *E. coli* K4/EcK4r3

The wild-type strain was CCUG 11307 and the strain has been engineered as reported (Cimini et al. 2013). Fed-batch experiments were carried out in the laboratories of Prof. Schiraldi at the University of Campania “L. Vanvitelli” in a Biostat D100 reactor with a working volume of 90 L, connected to a DCU and a remote control unit. The standard medium used for fermenter production studies consisted of a

defined salt medium (KH_2PO_4 2 g/L; K_2HPO_4 9.7 g/L; $\text{Na}_3\text{C}_6\text{H}_5\text{O}_7$ 0.5 g/L; $(\text{NH}_4)_2\text{SO}_4$ 1 g/L; MgCl_2 0.1 g/L) supplemented with glucose (10 g/L) as the main carbon source and yeast extract (2 g/L). A constant pH of 7.5 was maintained via automated addition of 30% (v/v) NH_4OH and 30% (v/v) H_2SO_4 . Fed-batch experiments on a 90-L reactor started immediately after depletion of the carbon source in the medium, after approximately 6-7 h of batch growth. During the fed-batch phase, an exponential feeding strategy or a dissolved oxygen (DO)-stat method was applied. In both cases, the substrate concentration was maintained within a desired range by the addition of an established amount of glucose/yeast extract feeding solution when the DO value increased up to 30%. The concentrated feeding solution contained 350 g/L of glucose, 70 g/L of yeast extract, and inorganic salts according to the medium recipe. For the duration of all cultivations, 5-mL samples were withdrawn from the reactors at regular time intervals for the determination of biomass, substrates, extracellular metabolites, and LPS concentration. Quantification of the LPS was performed by capillary electrophoresis on a Beckman Coulter high-performance capillary electrophoresis instrument (P/ACE MDQ, Palo Alto, CA, USA) equipped with a diode array detector and a UV lamp, according to recently described method (Restaino et al. 2009).

6.5.2 Purification of lipid A

The fermentation broth was centrifuged at 9,000g in a continuous centrifuge (Clara 80, Alpha Laval). The supernatant was first subjected to a treatment with protease and subsequently to concentration/diafiltration steps on a tangential filtration system (Sartoflow alpha, Sartorius Stedim). Composite polyethersulfone (PES) ultrafiltration membranes (Fluxa Filtri) with nominal molecular mass cutoffs of 100 kDa were used in

order to remove low molecular mass contaminants and to reduce the volume of about 8–10-fold. The concentrated sample was then heated at 90°C for 2 h after acidification with HCl at pH 2.75 ± 0.05 . At the end of the hydrolysis process, centrifugation was carried out to recover the pellet containing the lipid A and separate it from the supernatant. The pellet (67 g) was then extracted/washed with 96% ethanol at 50°C for 30 min three times. After drying at 40°C for 36 h, 58 g of material was obtained. A 2.50-g aliquot was suspended in 4:1v/v CHCl_3 - CH_3OH (250 mL) and passed through a short Celite column (approx. 40 cm³). Roto-evaporation of the eluate gave a brownish powder (0.57 g) that was in turn dissolved in CHCl_3 (50 mL) and then treated with CH_3OH (150 mL). After standing at -20°C overnight, the formation of a white precipitate was observed. It was collected by centrifugation (Eppendorf 5804R) at 4°C (7600g, 20 min), then washed twice with CH_3OH (50 mL). After drying under a vacuum, a white powder (0.21 g, 1.5% w/w global yield from the pellet) was obtained.

6.5.3 Chemical derivatization for GC-MS analysis

The sample (2.0 mg) was suspended in 1.25 M HCl solution in CH_3OH (1.0 mL) and heated at 80°C overnight. Volatiles were then removed and the residue was dissolved in 1:1 v/v pyridine-acetic anhydride (100 μL). The solution was heated at 100°C for 30 min. Volatiles were then removed. The residue was dissolved in CHCl_3 (1 mL) and washed three times with water. The organic phase, containing both sugars as per-*O*-acetylated methyl *O*-glycoside derivatives and lipids as (3-*O*-acylated)-fatty acid methyl esters, was collected, filtered, and concentrated to give a residue that was injected for the GC-MS analysis as acetone solution.

6.5.4 Mass spectrometry

GC-MS analysis were performed with an Agilent Technologies 6850A gas chromatography apparatus equipped with a mass selective detector 5973N and a Zebron ZB-5 capillary column (Phenomenex, 30 m×0.25 mm i.d., flow rate 1 mL/min, He as carrier gas). The samples were analyzed using the following temperature program: 150°C for 3 min, 150→ 280°C at 3 °C/min. Matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS) spectra were recorded on a Voyager DE STR instrument (Applied Biosystems, Framingham, MA) in the negative ions mode: compounds were dissolved in 4:1 v/v CHCl₃-MeOH at a concentration of 0.1 mg/mL, and 1 μL of these solutions were mixed with 1 μL of a 20 mg/mL solution of 2,5- dihydroxybenzoic acid in 7:3 v/v CH₃CN/H₂O.

6.5.5 De-O-phosphorylation of lipid A 1 to 2

Lipid A 1 (1.00 g) was suspended in 48% w/w aqueous HF (50 mL). The mixture was stirred at 5°C for 48 h, then diluted with water (100 mL) and extracted three times with CHCl₃ (150 mL). The organic phases were collected by centrifugation (Eppendorf 5804R) at 5°C (7600g, 20 min), then pooled, dried with anhydrous Na₂SO₄, filtered, and concentrated to give a residue that was dissolved in the minimum amount of 9:1 v/v CHCl₃-CH₃OH and charged on a Kieselgel 60 (63– 200 mesh; Merck) silica-gel column (approx. 90 cm³). Elution was performed with 9:1 v/v CHCl₃-CH₃OH (240 mL), 8:2 v/v CHCl₃-CH₃OH (50 mL), and 18:6:1 v/v/v CHCl₃-CH₃OH-H₂O (400 mL). The elution of the products was detected by analytical thin-layer chromatographies (TLCs) performed on aluminum plates precoated with Merck Silica Gel 60 F254 as the adsorbent and developed with 10% H₂SO₄ ethanolic solution and then

heated to 150°C. As the first eluted fraction, a mixture of glycerolipids was obtained (0.57 g) as a yellowish oil. As the second eluted fraction, pure lipid A **2** (0.20 g) was obtained as a white powder.

6.5.6 NMR spectroscopy

NMR spectra were recorded on a Bruker DRX-600 (¹H 600 MHz, ¹³C 150 MHz) instrument equipped with a cryoprobe, in 4:1 v/v CDCl₃-CD₃OD (residual CHCl₃ as internal standard, ¹H δ 7.26 ppm; ¹³C δ 77.0 ppm) at 298K. All two-dimensional homo and heteronuclear experiments (COSY, TOCSY, ROESY, and HSQC-DEPT) were performed by using standard pulse sequences available in the Bruker software. The mixing time for TOCSY and ROESY was 100 ms.

6.5.7 Biological assays

Biological activity of the obtained products was evaluated by stimulating the human monocytic leukemia cell line THP-1 (ATCC TIB-202). The enzyme-linked immunosorbent assay (ELISA) was selected as a specific and highly sensitive method for quantification of cytokine production, *in vitro*. In particular, monocytic THP-1 cells (LGC, Promochem, Milan, Italy) were grown in suspension in RPMI 1640 medium supplemented with 10 % fetal bovine serum at 37 °C in a humidified incubator with a 5% CO₂ atmosphere and stimulated with the 4:1 v/v CHCl₃-MeOH supernatant fraction (SF), lipid A **1** and **2**. LPS from *E. coli* O111:B4 (Sigma, Milan, Italy) and monophosphoryl lipid A (MPLA) from *Salmonella Minnesota* R595 (Avanti Polar Lipids, Spectra 2000, Milan, Italy) were used as positive controls. Mabtech's ELISAPRO kits and Enzo Life Science EIA Kit (BioRad, Milan, Italy) were used to quantify IL-6, IL-8, and TNF-α production in tissue-culture supernatants, following manufacturer's instructions.

References

- Bedini E. et al., *Angew Chem Int Ed.* **2011**, *50*, 6160– 6163.
- Carillo S. et al., *Chem. Eur. J.* **2011**, *17*, 7053–7060.
- Casella C. R., Mitchell T. C. *Cell. Mol. Life Sci.* **2008**, *65*, 3231–3240.
- Casillo A. et al., *Chem. Bio. Chem.* **2017**, (*in press*)
doi: 10.1002/cbic.201700287.
- Choma A., Komaniecka I., Zebracki K. *Biochim. Biophys. Acta* **2017**, *1862*, 196–209.
- Cimini D. et al., *Microb Cell Factories.* **2013**, *12*, 46.
- D’Alonzo D. et al., *Chem. Eur. J.* **2016**, *22*, 1053–11063.
- Di Lorenzo F. et al., *Chem. Eur. J.* **2017**, *23*, 3637–3647.
- Gao J., Guo Z. *Med. Res. Rev.* **2017**, (*in press*)
doi: 10.1002/med.21447.
- Li Y. et al., *Mar. Drugs* **2013**, *11*, 3197–3208.
- Liao G. et al., *ACS Cent. Sci.* **2016**, *2*, 210–218.
- Liao G., Zhou Z., Guo Z., *Chem. Commun.* **2015**, *51*, 9647–9650.
- MacLean L.M., Perry M.B. *Biochem Cell Biol* **1997**, *75*, 199–205.
- Marzabadi C. H., Franck R. W., *Chem. Eur. J.* **2017**, *23*, 1728–1742.
- Mata-Haro V. et al., *Science* **2007**, *316*, 1628–1632.
- Molinaro A. et al., *Chem. Eur. J.* **2015**, *21*, 500–519.
- Needhama B. D. et al., *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 1464–1469.
- Oursel D. et al., *Rapid Commun Mass Spectrom* **2007**, *21*, 1721– 1728.

- Park B. S. et al., *Nature* **2009**, *458*, 1191-1196.
- Peri F., Calabrese V., *J. Med. Chem.* **2014**, *57*, 3612-3622.
- Pieretti G. et al., *Appl. Microbiol. Biotechnol.* **2014**, *98*, 7781-7791.
- Restaino O.F. et al., *Electrophoresis*, **2009**, *30*, 3877-3883.
- Ribeiro A.A., Zhou Z., Raetz C.R.H., *Magn Reson Chem* **1999**, *37*, 620-630.
- Rietschel E. T. et al., *FASEBJ.* **1994**, *8*, 217-225.
- Rodriguez M.L., Jann B., Jann K., *Eur. J. Biochem.* **1988**, *177*, 117-124.
- Shay L.K., Hunt H.R., Wegner G.H., *J. Ind. Microbiol.* **1987**, *2*, 79-85.
- Silipo A. et al., *Nat. Commun.* **2014**, *5*, 5106.
- Tirsoaga A. et al., *J Lipid Res* **2007**, *48*, 2419-2427.
- Wang Q. et al., *ACS Chem. Biol.* **2012**, *7*, 235-240.
- Zhou Z. et al., *Chem. Sci.* **2015**, *6*, 7112-7121.

Chapter 7: Functionalization on the C-6 of GlcN II of *E. coli* monophosphoryl lipid A

The development of a protocol to access **2** in gram-scale quantities gave us the chance to use it as a synthetic building block, exploring its chemistry through the transformations conceived in Figure 7.1.

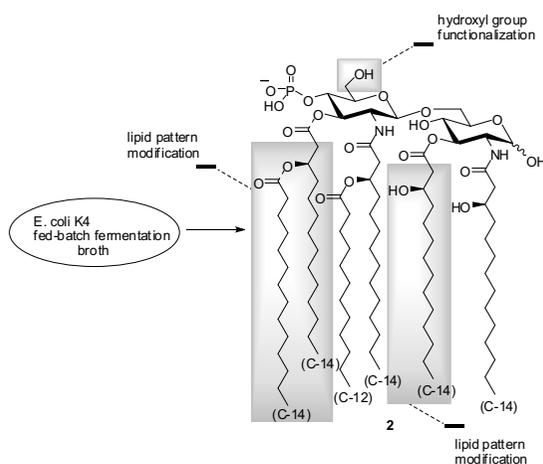


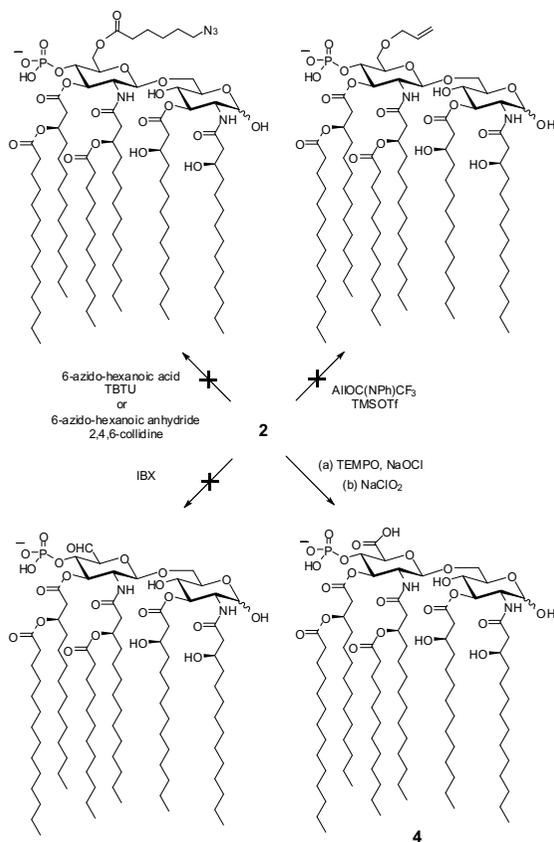
Figure 7.1: Structural modifications planned on MPLA from *E. coli* (**2**).

Notably, despite the number of approaches leading to **2** and its congeners, even in large scale (Holst et al. 1993; Pieretti et al. 2014; Coler et al. 2011), to our knowledge, no systematic investigation of their reactivity has been performed so far. This is partly due to the difficulties in obtaining considerable amounts of lipid A compounds in high purity. In addition, the structural peculiarity of all lipid A derivatives (highly polar molecules with amphiphilic character having both base- and acid-labile functionalities) renders their manipulation a challenging topic. A first look at lipid A reactivity was provided on the hydroxyl at C-6 of GlcN II (Figure 7.1). The modification of the lipid A scaffold at this position was attractive not only for the expected higher reactivity of such primary alcohol with respect to the secondary hydroxyls, but also

because it could be biologically well-tolerated. Indeed, lipid A derivatives and the core part of LOSs or LPSs are linked together through a glycosidic bond between GlcN II O-6 atom and the anomeric carbon of a 2-keto-octonic acid (3- deoxy-D-manno-2-octulosonic acid, Kdo or, much more rarely, D-glycero-D-talo-2-octulosonic acid, Ko).

7.1 Site-selective chemical reactions

To insert a chemical handle that would be easily and orthogonally derivatizable with respect to the other structural moieties present on lipid A, we attempted the installation of an azide or olefinic functionality on **2** under several different conditions. By esterification with 6-azido-hexanoic acid (Grandjean et al. 2005) in the presence of TBTU and Hünig's base (Twibanire et al. 2011), a sluggish conversion was detected. Instead, the reaction with 6-azido-hexanoic anhydride (Montanez et al. 2011) in the presence of 2,4,6-collidine and catalytic 4-(*N,N*-dimethylamino)pyridine (DMAP) in CHCl₃ (Yoshida et al. 2010) gave no reaction at all. No result was obtained also in the case of TMSOTf-catalyzed etherification with allyl *N*-phenyl trifluoroacetimidate (Tsabedze et al. 2013) (Scheme 7.1).



Scheme 7.1: First attempts to derivatize the primary hydroxyl group of MPLA **2**.

The lower nucleophilicity of the *C*-6 hydroxyl prompted us to transform it into a different functionality. Its oxidation to the corresponding aldehyde by treatment with 2-iodoxybenzoic acid (IBX) in dimethyl sulfoxide (DMSO) (Zhdankin et al. 2011) was tested, but no reaction was observed. Eventually, **2** was treated with 2,2,6,6-tetramethylpiperidine-1-oxyl (TEMPO) and NaOCl under phase-transfer conditions (NaBr, NaHCO₃, TBAB, 2:1 v/v CH₂Cl₂/H₂O), followed by one-pot oxidation of the aldehyde with NaClO₂ (Huang et al. 2006). Under these conditions, a quantitative transformation of the starting material into a more polar product could be observed by TLC. MALDI-MS analysis suggested the oxidation of a single hydroxyl to carboxylic acid, accompanied by a

partial loss of the C14:0(3-OH) lipid chain at position 3 of GlcNAc I, that is known to be the most base-labile (Myers et al. 1990) (m/z 1730.28 and 1752.15 for the major hexa-acylated species and its sodium salt, respectively; m/z 1504.39 and 1526.27 for the minor de-*O*-acylated counterpart and its sodium salt, respectively). (Figure 7.2A).

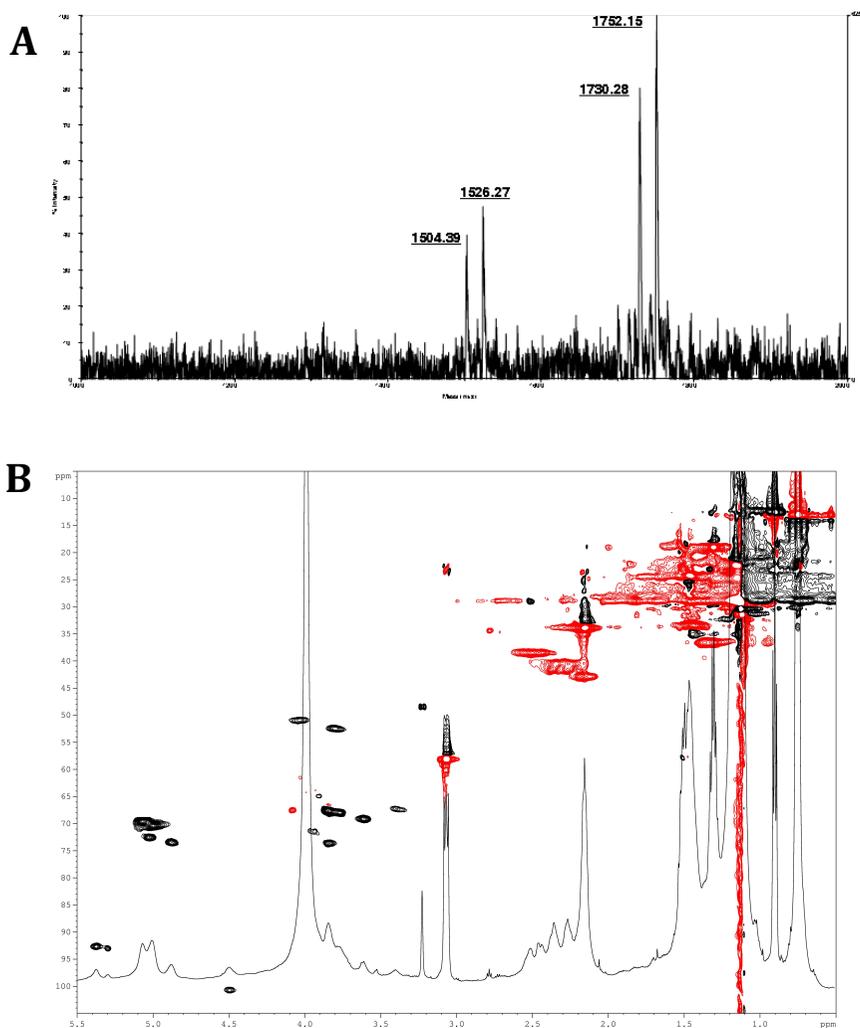


Figure 7.2 : A) MALDI-TOF mass spectrum (negative ions mode) of pure **4**; B) ^1H and HSQC-DEPT NMR spectra (600 MHz, 298K, 4:1 v/v $\text{CDCl}_3\text{-CD}_3\text{OD}$) of pure **4**.

Table 7.1: NMR data referred to lipid A derivative **4** (major hexa-acylated specie)

	Sugar portion ^a							Lipid portion ^a						Phosphate ^b
	1	2	3	4	5	6	α	β	γ	$\delta-(\omega-1)$	ω			
4	GlcN I (α)	5.38	4.04	5.01	3.38	3.92	4.08	C14:0(3-OH) [N-2 GlcN I]	2.09, 2.17 <i>42.8</i>	3.76 <i>68.0</i>	1.30 <i>36.6</i>	1.09-1.18 <i>29.0</i>	0.71-0.75 <i>13.5</i>	-0.6
		<i>92.5</i>	<i>51.0</i>	<i>72.4</i>	<i>67.5</i>	<i>71.5</i>	<i>67.5</i>	C14:0(3-OH) [O-3 GlcN I]	2.26, 2.35 <i>41.5</i>	3.84 <i>67.6</i>	1.36 <i>36.6</i>			
								C14:0(3-OR) [N-2 GlcN II]	2.26, 2.35 <i>40.6</i>	5.00 <i>70.4</i>	1.43 <i>33.4</i>			
								C14:0(3-OR) [O-3 GlcNII]	2.43, 2.50 <i>38.4</i>	5.06 <i>69.6</i>	1.43 <i>33.4</i>			
	GlcN II	4.49	3.80	4.87	3.60	3.84	---	C12:0	2.15	1.46	1.09-1.18			
		<i>100.7</i>	<i>52.5</i>	<i>73.5</i>	<i>69.1</i>	<i>74.1</i>	<i>169.1</i>	C14:0	33.9	24.3	<i>29.0</i>			

^a ¹H (600 MHz, 298K) and ¹³C (125 MHz, 298K) chemical shifts (in plain and in italic, respectively) expressed in δ relative to residual CHCl₃ (¹H: δ 7.26 ppm; ¹³C: δ 77.0 ppm) in 4:1 v/v CDCl₃-CD₃OD.

^b ³¹P (160 MHz, 298K) chemical shifts expressed in δ relative to 85% phosphoric acid (external standard).

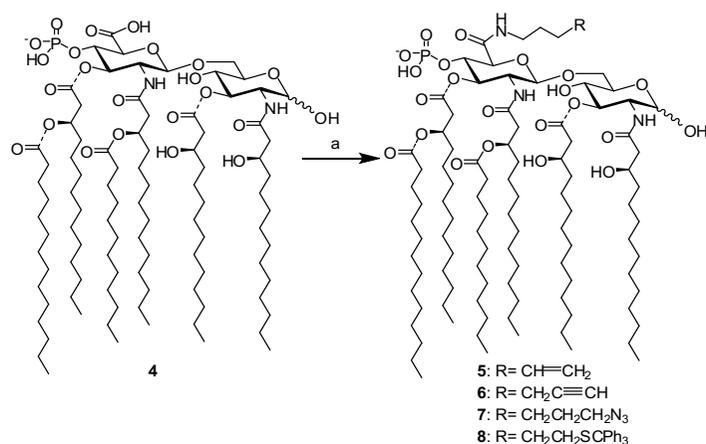
From the full NMR characterization of the major hexa-acylated oxidized derivative by 2D-techniques (COSY, TOCSY, HSQC- DEPT and HMBC) (Figure 7.2B; Table 7.1), it emerged that GlcN II CH₂OH signals were missing with respect to NMR data of starting lipid A **2** (Figure 6.4 and Table 6.1), whereas the GlcN I anomeric CH signal was still present. This allowed the assignment of structure **4** to the oxidized lipid A derivative, with the GlcN II primary alcohol transformed into a carboxylic acid and the GlcN I hemiacetal moiety unaffected. Interestingly, the carboxylic acid function at position C-5 of GlcN II could not only serve as a suitable moiety for further derivatizations, but also adds a further negative charge to the molecule, similarly to lipid A-Kdo structures, that have been demonstrated to induce stronger immune responses than their parent lipid A (Zhang et al. 2008; Zughailer et al. 2004; Yoshizaki et al. 2001).

The lack of reactivity of the primary hydroxyl group of **2** to most of the reaction conditions mentioned above represented one of the major obstacles to the synthetic manipulation of **2** en route to novel lipid A derivatives. Although an unambiguous explanation for this unusual reactivity profile was not found, we hypothesized the existence of inverse micellar aggregates in organic solvents as a consequence of the amphiphilic nature of **2**. Accordingly, the polar character of the hydroxyl group at the O-6 position of GlcN II as well as its proximity to the charged phosphate group could direct it inside the inverse micelle, thus impeding any interaction with reagents in solution.

7.2 Insertion of a clickable moiety

In order to insert a clickable moiety on lipid A structure, derivative **4** was treated with diverse amines carrying a suitable functionalization at their

ω -position (double or triple bond, azide, *S*-trityl as protected thiol; see 5-8, Scheme 7.2 and Table 7.2) (Ziaco et al., 2017). The first aim was to find the best conditions for the amide formation reaction using a model amine such as 4-penten-1-amine (Yazici et al., 2013). The coupling was tested under typical peptide coupling conditions, in the presence of a base [4-(dimethylamino)pyridine (DMAP), *N,N*-diisopropylethylamine (DIPEA) or 4-methylmorpholine (NMM)] and different kinds of carboxylic acid activators (El-Faham et al., 2011): i) a carbodiimide reagent such as *N,N'*-diisopropylcarbodiimide (DIPC), ii) a triazine activator such as 2-chloro-4,6-dimethoxy-1,3,5-triazine (CDMT), iii) an uronium salt such as *O*-(benzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium tetrafluoroborate (TBTU), and iv) a phosphonium salt such as (benzotriazol-1-yloxy)tripyrrolidinophosphonium hexafluorophosphate (PyBOP[®]), in some cases with a protic additive such as 1-hydroxybenzotriazole (HOBt).



Scheme 7.2: Insertion of a clickable moiety on *E. coli* monophosphoryl lipid A 4,

(a) see Table 7.2

A crucial issue for the optimization of the reaction conditions to transform derivative 4 into carboxyl amide 5 was to avoid, or at least to

minimize, the concomitant or even alternative amidation of the phosphate group. Indeed, studies pointed towards the synthesis of oligopeptides containing phosphotyrosine or its phosphonic analogue without preliminary protection of the phosphoric or phosphonic acid functionalities, revealed that such moieties can be activated under some standard carboxyl amide formation conditions. In particular, it was shown that carbodiimide-mediated reactions give amidation on both phosphoric and carboxylic acid of phosphotyrosine, whereas conditions employing uronium- or phosphonium-type reagents are selective for carboxylic acid derivatization, with the phosphoric/phosphonic moiety reacting more sluggishly (Perich 2004; Campagne et al. 1995).

In order to investigate the behaviour of the carboxylic and phosphoric acid moiety of lipid A derivative **4**, we firstly performed a MALDI mass spectrum (negative ions mode) of the crude reaction mixtures of **4** with 4-penten-1-amine under different reaction conditions (Table 7.2, entries 1-6). to find out signals of the desired 4-pentenyl amide derivative **5**, the residual starting lipid A **4** and the potential byproducts due to concomitant activation of phosphate group too (see Figures 7.3-7.8).

Table 7.2 Amide formation reactions on lipid A derivative **4**.

Entry	Amine	Coupling conditions	Solvent	Product	Yield ^a	Phosphate activation ^b
1	4-penten-1-amine	DIPC, DMAP	1:1 v/v DMF-CH ₂ Cl ₂	5	0%	No
2	4-penten-1-amine	CDMT, NMM	5:1 v/v THF-DMF	5	39%	Yes (byproduct 9) ^c
3	4-penten-1-amine	PyBOP [®] , HOBT, DIPEA	1:1 v/v DMF-CH ₂ Cl ₂	5	69%	Yes (byproduct 10a or 10b) ^c
4	4-penten-1-amine	PyBOP [®] , HOBT, H ₂ O, DIPEA	1:1 v/v DMF-CH ₂ Cl ₂	5	82%	Yes (byproduct 10a or 10b) ^c
5	4-penten-1-amine	TBTU, DIPEA	1:1 v/v DMF-CH ₂ Cl ₂	5	68%	Traces (byproduct 10a or 10b) ^c
6	4-penten-1-amine	TBTU, HOBT, DIPEA	1:1 v/v DMF-CH ₂ Cl ₂	5	82%	No
7	5-hexyn-1-amine	TBTU, HOBT, DIPEA	1:1 v/v DMF-CH ₂ Cl ₂	6	62%	No
8	6-azido-hexan-1-amine	PyBOP [®] , HOBT, H ₂ O, DIPEA	1:1 v/v DMF-CH ₂ Cl ₂	7	35%	Yes (byproduct 11a or 11b) ^c
9	5-(tritylthio)- pentan-1-amine	PyBOP [®] , HOBT, H ₂ O, DIPEA	1:1 v/v DMF-CH ₂ Cl ₂	8	83%	Traces (byproduct 12a or 12b) ^c

^a Isolated yield

^b As indicated by MALDI-MS (negative ions mode) analysis of the crude reaction mixture (see Figures 7.3-7.8)

^c See Figure 7.12 for the putative structure of byproducts.

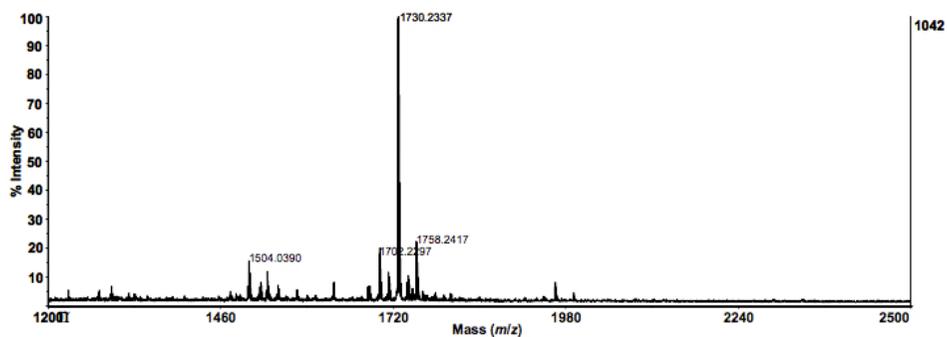


Figure 7.3: MALDI-TOF mass spectrum (negative ions mode) of the crude reaction mixture of Table 7.1, entry 1 (**4** with 4-pentenyl-1-amine in the presence of DIPC and DMAP).

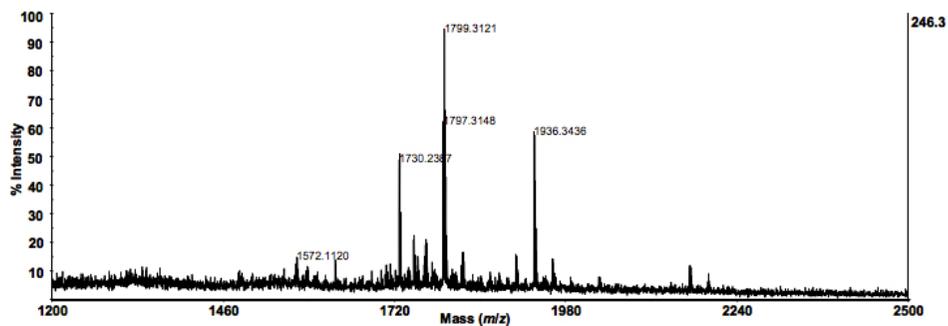


Figure 7.4: MALDI-TOF mass spectrum (negative ions mode) of the crude reaction mixture of Table 7.1, entry 2 (**4** with 4-pentenyl-1-amine in the presence of CDMT and NMM).

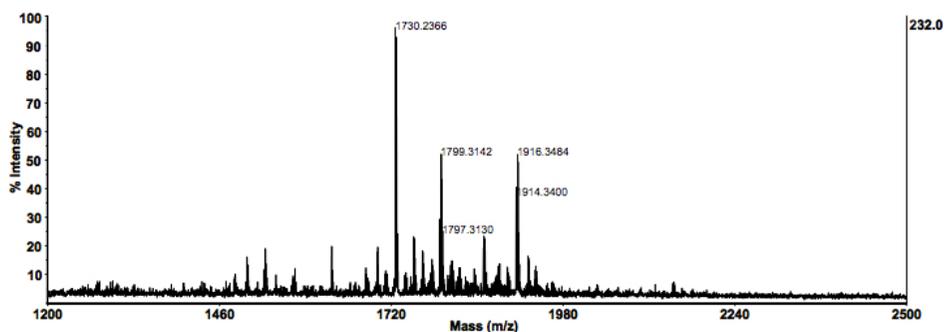


Figure 7.5: MALDI-TOF mass spectrum (negative ions mode) of the crude reaction mixture of Table 7.1, entry 3 (**4** with 4-pentenyl-1-amine in the presence of PyBOP®, HOBt, and DIPEA).

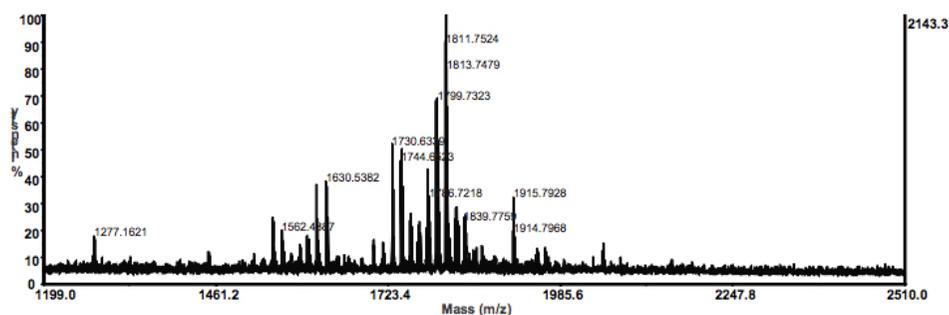


Figure 7.6: MALDI-TOF mass spectrum (negative ions mode) of the crude reaction mixture of Table 7.1, entry 4 (**4** with 4-pentenyl-1-amine in the presence of PyBOP®, HOBt, H₂O and DIPEA).

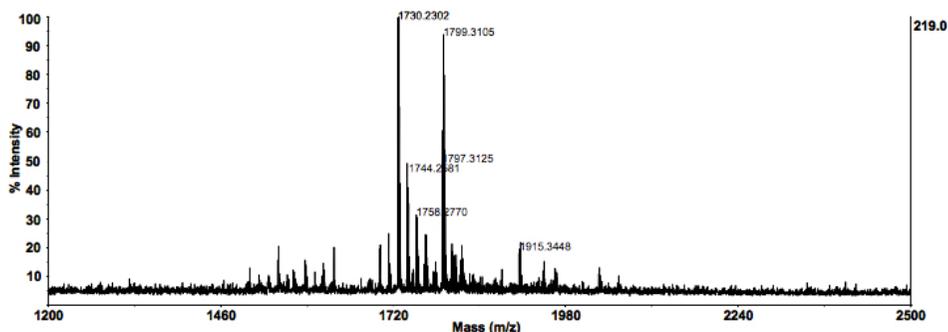


Figure 7.7: MALDI-TOF mass spectrum (negative ions mode) of the crude reaction mixture of Table 7.1, entry 5 (**4** with 4-pentenyl-1-amine in the presence of TBTU and DIPEA).

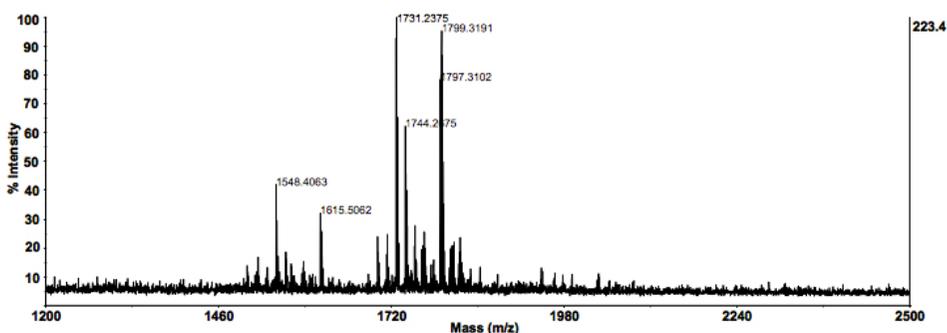


Figure 7.8: MALDI-TOF mass spectrum (negative ions mode) of the crude reaction mixture of Table 7.1, entry 6 (**4** with 4-pentenyl-1-amine in the presence of TBTU, HOBT and DIPEA).

A carbodiimide reagent such as DIPC in combination with DMAP did not give any type of activation on the lipid **A** (entry 1, Figure 7.3), whereas uronium-, phosphonium- and triazine-type activators did it. Indeed, in the presence of a base such as NMM, CDMT activated both carboxylic and phosphoric acid moieties of **4** (entry 2, Figure 7.4), affording, as major hexa-acylated species, the desired product **5** (Figure 7.9a, m/z 1797.31 for $[M-H]$) and related minor species with a higher or lower number of methylenes in the lipid portion, together with the putative phosphotriazine ester byproduct **9** (Figure 7.12) (m/z 1936.34

for [M-H]) and the residual starting compound **4** (m/z 1730.24 for [M-H]) and related minor species as indicated for **5**).

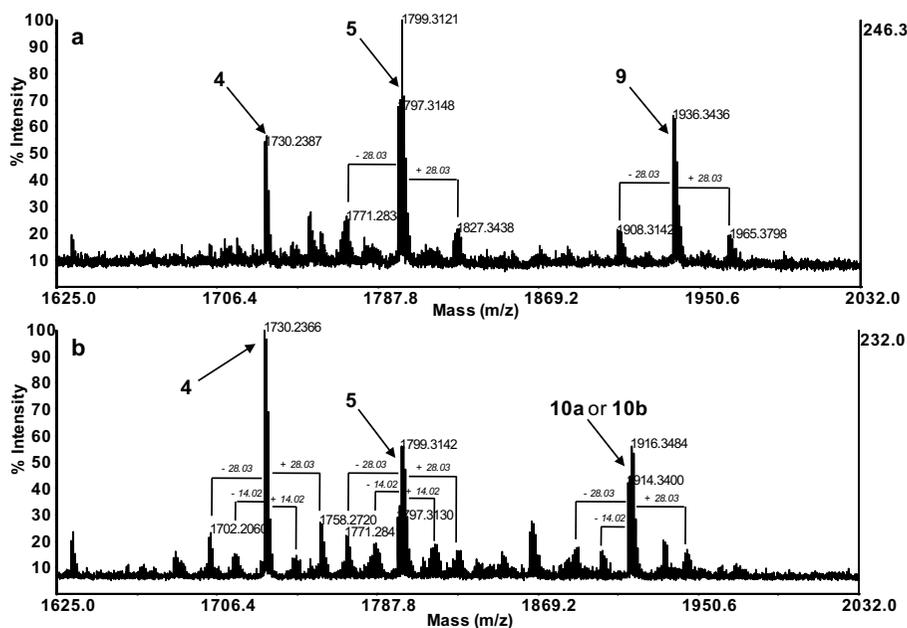


Figure 7.9: MALDI-TOF mass spectrum (negative ions mode) of the crude reaction mixture of **4** with 4-pentenyl-1-amine in the presence of (a) CDMT and NMM or (b) PyBOP[®], HOBt and DIPEA (Table 7.1, entries 2 and 3, respectively; zoom on hexa-acylated species).

Purification of the crude reaction mixture was possible through gel filtration chromatography on a Sephadex[®] LH-20 column, affording the target double bond-equipped lipid **A** **5** in 39% yield (Figure 7.10).

The covalent linkage of the 4-pentenyl moiety on the lipid **A** through a carboxamide instead of a phosphoramidate bond was confirmed by correlations of both GlcN II *CH*-5 ($\delta_{\text{H/C}}$ 3.65/72.1 ppm) and 4-pentenoyl amide α -CH₂N ($\delta_{\text{H/C}}$ 3.12/38.2 ppm) signals with the same carboxamide density (δ_{C} 170.2 ppm) in the heteronuclear multiple bond correlation (HMBC) NMR spectrum (Figure 7.11).

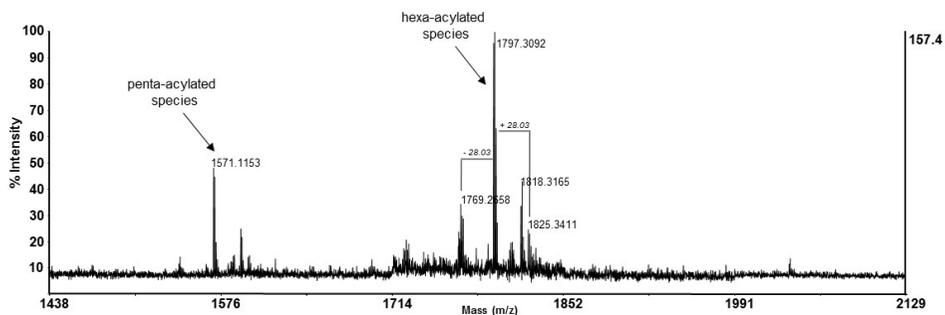


Figure 7.10: MALDI-TOF mass spectrum (negative ions mode) of pure **5**.

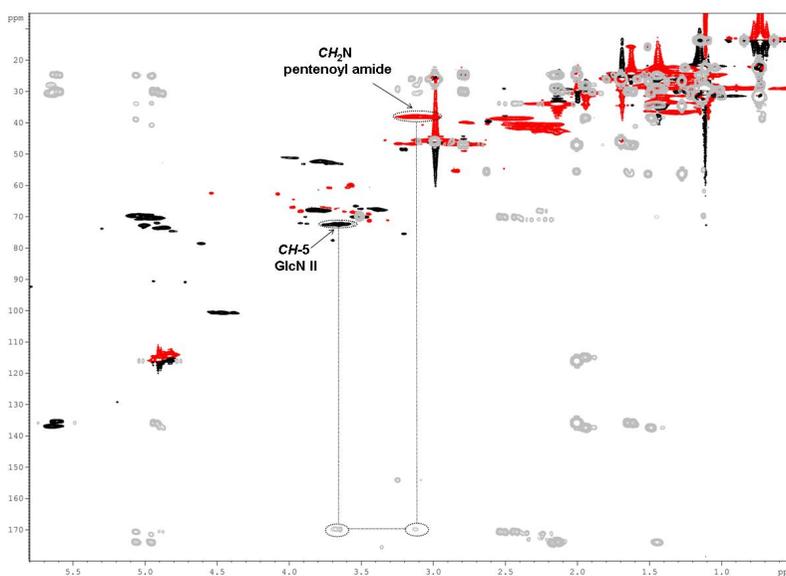


Figure 7.11: HSQC-DEPT (red and black) and HMBC (grey) NMR spectra (600 MHz, 298K, 4:1 v/v CDCl₃:CD₃OD) of pure **5**.

The coupling of **4** and 4-penten-1-amine under PyBOP[®] activation in the presence of HOBT and DIPEA (entry 3, Figure 7.5) gave again a reaction on both carboxylic and phosphoric moieties. The desired lipid **5** was isolated in 69% yield from residual starting material and a byproduct showing a peak at m/z 1914.34 (and at ± 14.02 or 28.03 m/z , indicating minor species with a higher or lower number of methylenes in the lipid portion; Figure 7.9b) by means of a chromatography on Sephadex[®] LH-

20. The structure of the byproduct was tentatively associated to a lipid A with a phosphate diester moiety (Figure 7.12, **10a**). Indeed, benzotriazolyl phosphoesters have been shown to originate in some cases from activation of a phosphoric acid functionality in the presence of uronium and phosphonium reagents. They do not react further in the presence of amine nucleophiles, while activated carboxylic acids do it. Since phosphonic acid benzotriazolyl esters are analogously known to be poorly reactive with amines, but very prone to react with alcohol and water (Campagne et al. 1995), we tested the coupling between **4** and 4-penten-1-amine under the PyBOP[®]/HOBT/DIPEA conditions with the addition of water, in order to hydrolyze *in situ* the putative phosphoric acid benzotriazolyl ester upon formation (Figure 7.6). An enhancement of yield from 69% to 82% was observed (Table 7.1, entries 3 and 4). Interestingly, by treating the crude reaction mixture of entry 3 with DIPEA and water in 1:1 v/v DMF-CH₂Cl₂, hydrolysis of byproduct **10** was not observed by MALDI-MS analysis (data not shown). This suggested that the benzotriazolyl ester **10a** had rearranged during the course of the coupling reaction into the more stable phosphoramidate **10b** (Figure 7.12) (Musiol et al. 1994).

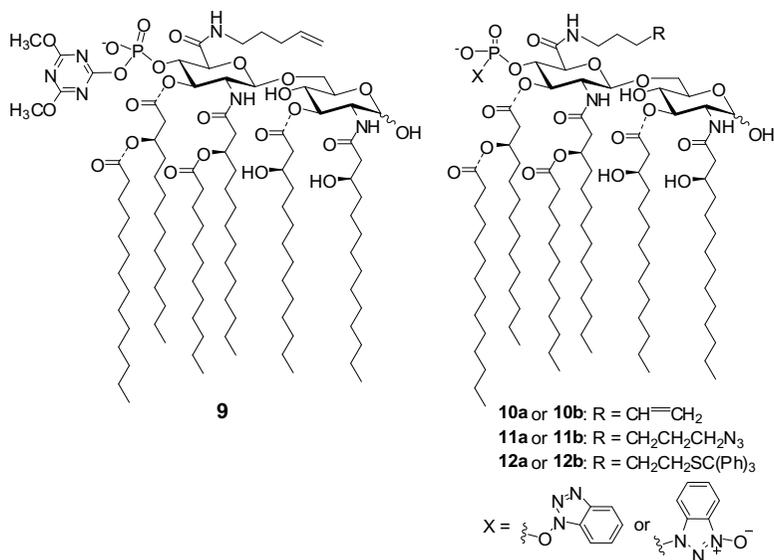


Figure 7.12: Putative structure for the byproducts obtained from phosphate activation of **4** under carboxyl amide formation reactions (see Table 7.1)

A switch from a phosphonium- to an uronium-type activator such as TBTU afforded almost exclusively activation at carboxylic site with only trace amounts of **10a/10b** (Table 7.1, entry 5; see also Figure 7.7). Nonetheless, the isolated yield of **5** was not excellent (68%), due to a significant amount of residual starting compound **4**. By adding a protic additive to the reaction mixture, the yield could be increased to 82% (entry 6, Figure 7.8). After having detected PyBOP[®] and TBTU as the best coupling reagent under appropriate conditions (entries 4 and 6, respectively), these were employed to install clickable functionalities other than a double bond on lipid A **4**. By reaction with 5-hexyn-1-amine (Saito et al., 2008) and 6-azido-hexan-1-amine (Lee et al. 2001), triple bond- and azide-equipped lipid A derivatives **6** and **7** could be obtained (entries 7 and 8, Figures 7.13-7.16; Table 7.1). A tritylthio moiety, as thiol group precursor, was also installed on **4** by reaction with 5-

(tritylthio)-pentan-1-amine (Jagadish et al. 2012) to afford derivative **8** in 83% yield (entry 9, Figures 7.17-7.18).

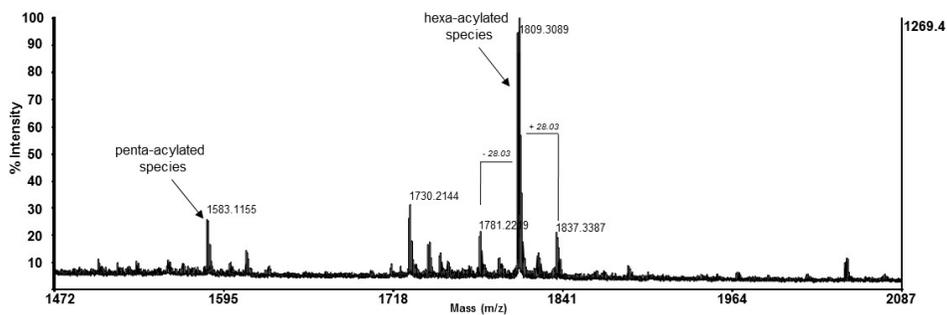


Figure 7.13: MALDI-TOF mass spectrum (negative ions mode) of pure **6**.

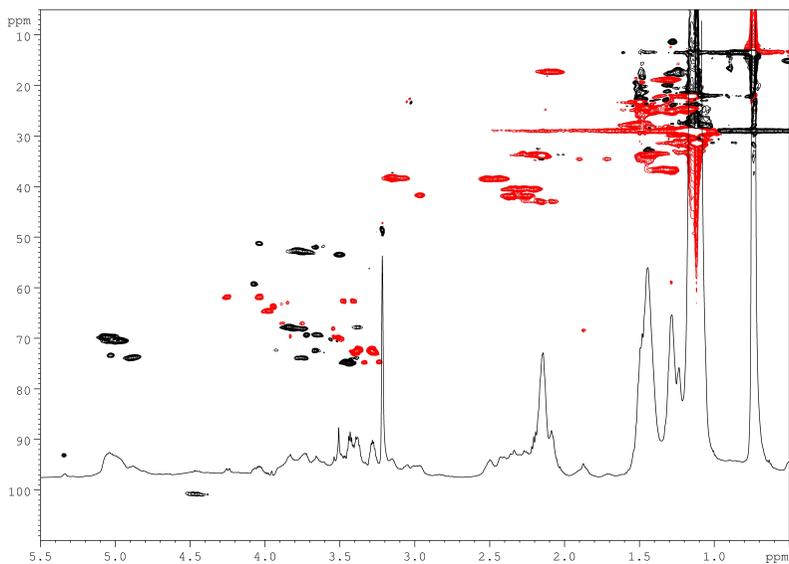


Figure 7.14: ¹H and HSQC-DEPT NMR spectra (600 MHz, 298K, 4:1 v/v CDCl₃-CD₃OD) of pure **6**.

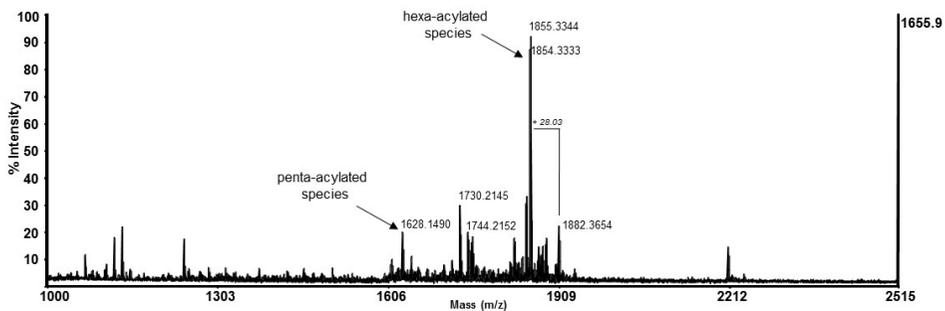


Figure 7.15: MALDI-TOF mass spectrum (negative ions mode) of pure **7**.

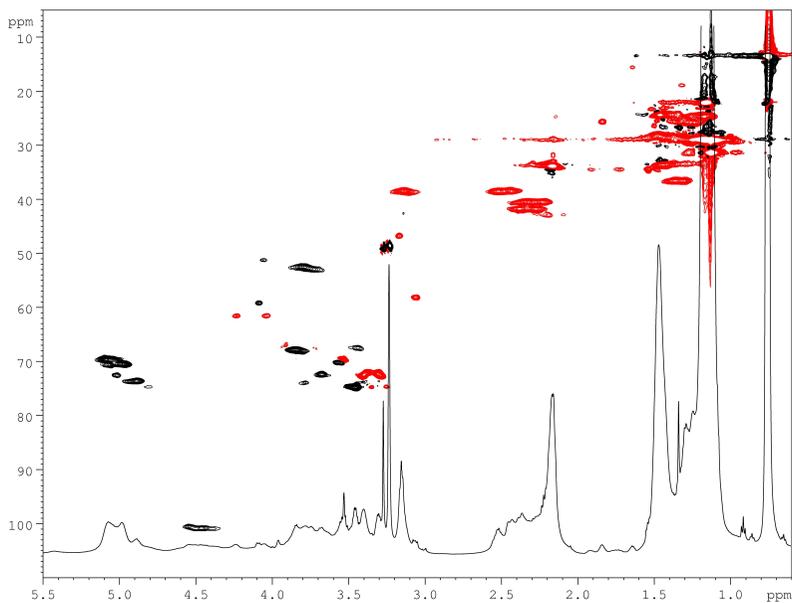


Figure 7.16: ^1H and HSQC-DEPT NMR spectra (600 MHz, 298K, 4:1 v/v CDCl_3 - CD_3OD) of pure **7**.

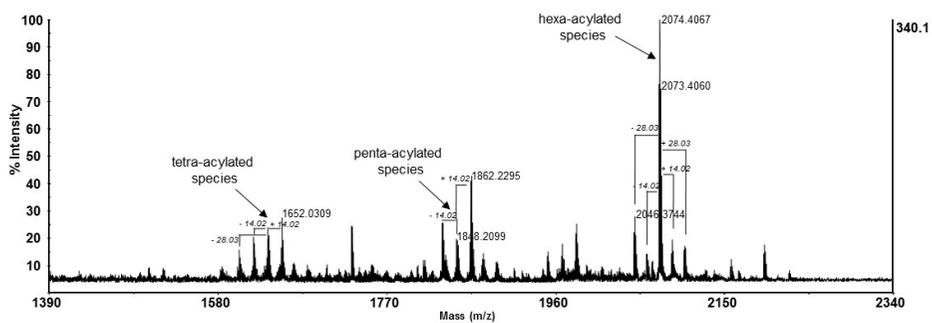


Figure 7.17: MALDI-TOF mass spectrum (negative ions mode) of pure **8**.

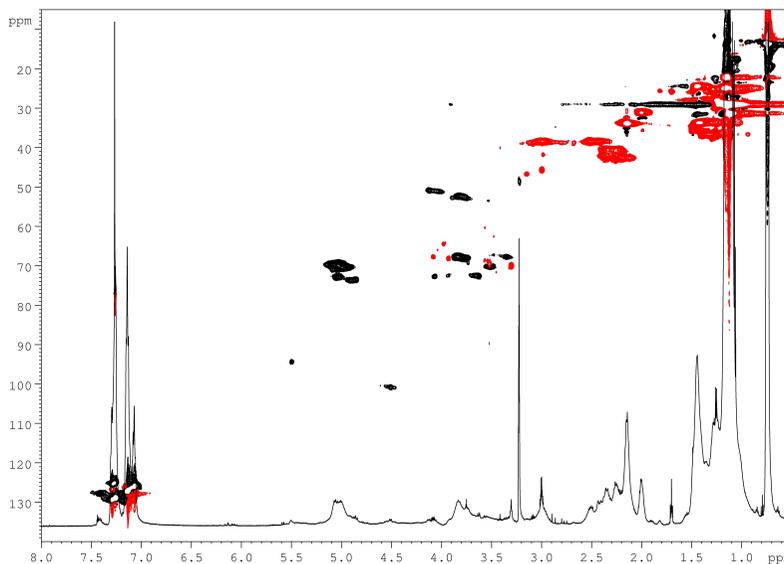


Figure 7.18: ^1H and HSQC-DEPT NMR spectra (600 MHz, 298K, 4:1 v/v CDCl_3 - CD_3OD) of pure **8**.

The obtained four novel clickable lipid A derivatives were completely characterized by NMR spectroscopic analysis (See Table 7.3).

The production of these clickable lipid A derivatives opens a straightforward access to their conjugation with other interesting biomolecules, such as tumor-associated carbohydrate antigens (TACAs). This methodology stands as a valuable alternative to recently reported total synthetic approaches, based on highly complex syntheses of lipid A carrying at anomeric position of their pseudo-reducing end a suitable moiety for conjugation with TACA antigens (Liao et al. 2015/2016; Zhou et al. 2014/2015; Wang et al. 2009/2012; Tang et al. 2010).

Table 7.3: NMR data referred to lipid A derivatives **5-8** (major hexa-acylated species)

	Sugar portion ^a							Lipid portion ^a							Other signals ^a	Phosphate ^b	
	1	2	3	4	5	6	C=O	α	β	γ	δ -(ω -1)	ω					
5	GlcN I (α)	4.95 90.5	4.99 72.8	3.37 67.7	3.73 72.8	3.72, 3.98 67.1	C14:0(3-OH) [N-2 GlcN I]	--- 171.3	2.08, 2.18 42.8	3.75 68.0	1.30 36.7	1.09-1.18	0.71-0.75 13.5	δ -pentenoyl amide: α : 3.12/38.2 β : 1.48/24.4 γ : 1.94/30.7 δ : 5.65/136.8 ϵ : 4.80, 4.91/114.5	-13.4		
					C14:0(3-OH) [O-3 GlcN I]		--- 171.1	2.24, 2.35 41.7	3.83 67.8	1.37 36.7							
		GlcN I (β)	4.38 100.7	4.82 74.6	3.61 72.4	3.74, 3.92 68.3	C14:0(3-OR) [N-2 GlcN II]	--- 171.2	2.22, 2.33 40.7	4.96 70.6	1.39 33.5						
	C14:0(3-OR) [O-3 GlcN III]						--- 170.4	2.42, 2.50 38.7	5.05 69.8	1.46 33.5							
	GlcN II		4.46 100.8	3.79 53.3	4.87 73.6	3.52 69.3	3.65 72.1	---	170.2	C12:0	---	173.7	2.14			1.45	1.09-1.18
		C14:0								---	173.6	34.1	24.5			22.0, 29.0, 31.4	
6	GlcN I (α)	5.35 93.1	4.06 51.3	5.05 73.4	3.38 67.9	3.74 73.9	3.76, 3.90 67.1	C14:0(3-OH) [N-2 GlcN I]	---	170.8	2.09, 2.17 43.1	3.75 68.2	1.30 36.7	1.09-1.18	0.71-0.75 13.5	ϵ -hexenoyl amide: α : 3.08/38.5; β : 1.53/27.8 γ : 1.42/24.9; δ : 2.07/17.2; ϵ : ---/83.4; ζ : 1.88/68.4	-1.8
					C14:0(3-OH) [O-3 GlcN I]			---	170.0	2.27, 2.36 41.9	3.84 67.8	1.34 36.7					
		GlcN II	4.47 100.8	3.77 52.8	4.90 73.9	3.55 70.2	3.67 72.4	---	170.2	C14:0(3-OR) [N-2 GlcN II]	---	170.8	2.21, 2.33 40.7				
	C14:0(3-OR) [O-3 GlcN III]									---	170.2	2.44, 2.51 38.6	5.07 69.9	1.44 33.7			
	C12:0		---	173.7	2.16	1.46	1.09-1.18										
	C14:0	---	173.6	33.9	24.5	22.0, 29.0, 31.4											
7	GlcN I (α)	5.43 93.8	4.05 51.6	5.01 72.7	3.44 67.7	3.90 72.7	3.73, 3.94 67.0	C14:0(3-OH) [N-2 GlcN I]	---	170.3	2.10, 2.20 42.9	3.86 68.0	1.29 36.6	1.09-1.18	0.71-0.75 13.5	ζ -azido-hexanoyl amide: α : 3.15/38.6; β : 1.48/28.4 γ : 1.27/26.0; δ : 1.27/26.0 ϵ : 1.82/25.6; ζ : 3.18/46.9	-3.3
					C14:0(3-OH) [O-3 GlcN I]			---	170.3	2.27, 2.38 41.8	3.86 67.8	1.36 36.6					
		GlcN II	4.46 100.6	3.80 52.5	4.88 73.5	3.56 70.2	3.67 72.4	---	169.5	C14:0(3-OR) [N-2 GlcN II]	---	171.5	2.23, 2.34 40.6				
	C14:0(3-OR) [O-3 GlcN III]									---	170.4	2.45, 2.52 38.5	5.08 69.8	1.46 33.6			
	C12:0		---	173.7	2.16	1.47	1.09-1.18										
	C14:0	---	173.7	34.1	24.6	22.0, 29.0, 31.4											
8	GlcN I (α)	5.51 94.6	4.14 51.0	5.05 73.1	3.36 68.0	4.08 72.8	3.73, 3.93 68.3	C14:0(3-OH) [N-2 GlcN I]	---	171.6	2.24, 2.34 42.6	3.75 68.0	1.28 36.8	1.09-1.18	0.71-0.75 13.5	ϵ -(tritylthio)-pentanoyl amide: α : 3.01/38.6; β : 1.71/25.9 γ : 1.29/25.2; δ : 1.45/24.6 ϵ : 2.01/31.2 Aromatics: 7.06/125.8; 7.13/127.1 7.26/128.9; ---/144.9 CPHs: ---/66.3	-3.2
					C14:0(3-OH) [O-3 GlcN I]			---	172.4	2.27, 2.36 42.0	3.85 67.9	1.36 36.8					
		GlcN II	4.52 101.0	3.83 52.4	4.87 73.7	3.50 70.4	3.64 72.8	---	169.6	C14:0(3-OR) [N-2 GlcN II]	---	172.2	2.27, 2.34 40.7				
	C14:0(3-OR) [O-3 GlcN III]									---	170.8	2.44, 2.51 38.6	5.07 69.8	1.44 33.8			
	C12:0		---	173.8	2.15	1.46	1.09-1.18										
	C14:0	---	174.0	34.0	24.6	22.0, 29.0, 31.4											

^a ¹H (600 MHz, 298K) and ¹³C (150 MHz, 298K) chemical shifts (in plain and in italic, respectively) expressed in δ relative to residual CHCl₃ (¹H: δ 7.26 ppm; ¹³C: δ 77.0 ppm) in 4:1 v/v CDCl₃-CD₃OD.

^b ³¹P (160 MHz, 298K) chemical shifts expressed in δ relative to 85% phosphoric acid (external standard).

7.3 Semi-synthetic “self-adjuvant” anticancer vaccine candidates

TACAs are useful targets for the development of therapeutic cancer vaccines, but they are typically poorly immunogenic and T-cell independent antigens. To make them more immunogenic and T-cell dependent antigens, they have to be covalently linked to immunologically active carrier molecules to form conjugate vaccines. In recent years, fully synthetic TACA-based conjugate cancer vaccines have become a hot topic, because they possess a number of advantages, such as defined structures, reproducible physical, chemical and biological properties, and promising immunologic activity (Danishefsky et al. 2000; Guo et al. 2009; Chen et al. 2013; Kaiser et al. 2010). For the development of “self-adjuvant” synthetic cancer vaccine, MPLA can be a particularly useful carrier molecule. First, MPLA is a strong immunostimulator. Second, owing to its bacterial origin, MPLA can function as a “danger signal” to the immune system to enhance the immune responses to TACAs covalently linked to it (Buskas et al. 2005). Third, as mentioned before, the endotoxicity of lipid A could be dramatically reduced after removal of its anomeric phosphate group, whereas its beneficial immunostimulatory activity was unaffected. In reality, MPLA has been approved for clinical use as a vaccine adjuvant and immunostimulator (Casella et al. 2008).

To probe the potential of MPLA as a carrier molecule for the development of self-adjuvant conjugate cancer vaccines, Guo and co-workers designed and synthesized a monophosphoryl derivative of *N. meningitidis* lipid A, which had lipid chains symmetrically distributed on

the two sugar units, coupled with a modified form of ganglioside monosialic acid (GM3), NPhAcGM3 **13** (Figure 7.19).

GM3 is a trisaccharide TACA, Neu5Ac α (2 \rightarrow 3)Galp β (1 \rightarrow 4)Glc_p, with a Neu5Ac residue attached to the nonreducing end (Hakomori et al., 1991; Hakomori et al., 1997). It exists in gangliosides or glycoproteins with the lipid tails or peptide chains inserting into cell membranes, and the trisaccharide is displayed on the cell surface, where it can be recognized by antibodies. GM3 is abundantly expressed on a number of tumors (Bitton et al., 2002), especially malignant melanoma.

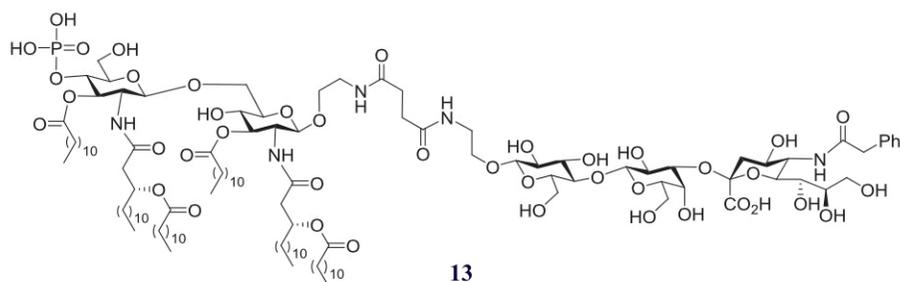


Figure 7.19: Chemical structure of the synthetic *N. meningitidis* MPLA-NPhAcGM3 conjugate.

Immunological studies of **13** in mice revealed that it alone, namely without the use of an external adjuvant, elicited robust IgG antibody responses. The results have demonstrated that MPLA could act not only as a carrier molecule to significantly improve the immunogenicity of TACAs but also as a built-in adjuvant to form self-adjuvanting vaccines (Wang et al. 2012; Pan et al. 2005). Therefore, MPLA was identified as a highly promising carrier for the design and development of synthetic self-adjuvant glycoconjugate cancer vaccines.

Guo and Co-workers have also developed a convergent and synthetic strategy, for a monophosphorylated *N. meningitidis* lipid A derivative

coupled with globo H, a TACA expressed by several tumors (**14**, Figure 7.20) (Zhou et al. 2015).

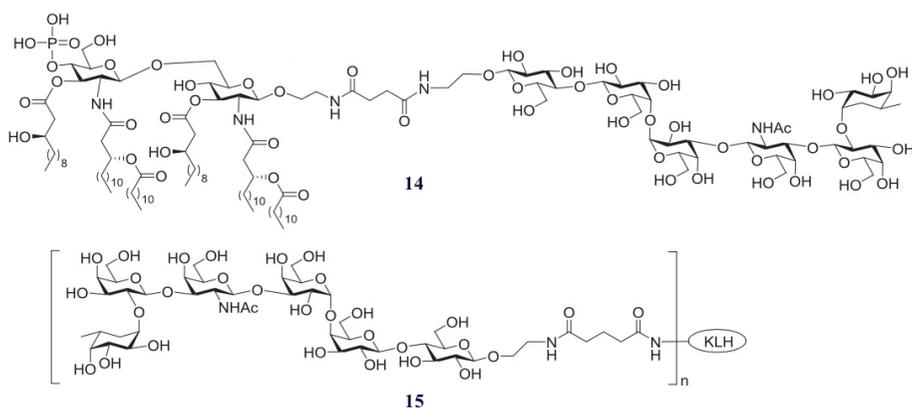
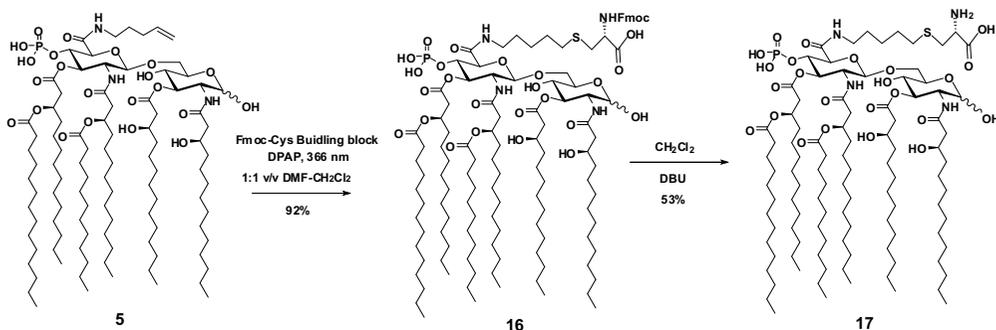


Figure 7.20: Chemical structure of the synthetic *N. meningitidis* MPLA-globo H conjugate **14** and the KLH-globo H conjugate **15**.

It was evaluated in mouse and compared with the Keyhole Limpet Hemocyanin (KLH) conjugate of globo H **15**, as the positive control, which is at Phase III clinical trial for the treatment of colon cancer. It was discovered that MPLA-globo H conjugate **14** elicited much faster and stronger immune responses than that of the KLH conjugate **15**, in terms of both total and IgG antibodies. It was demonstrated that both **14** and **15** induced the similar patterns of immune responses. These results suggested the great promise of MPLA-TACA conjugates as therapeutic cancer vaccines. These two examples, fully synthesized by Guo and co-workers, suggested that the introduction of unique functional groups, such as an azido or alkyne group, to MPLA can provide the flexibility for its regioselective coupling with TACAs or other biomolecules, such as carbohydrates, peptides, and proteins, to quickly generate complex multicomponent conjugates in a convergent manner for the development of synthetic self-adjuvant anticancer glyco-conjugate vaccines.

To this aim, we performed the synthesis of two different TACA antigens, in order to achieve two novel semi-synthetic self-adjutant anticancer glyco-conjugates following a different strategy for the coupling. Indeed, in all the Guo's total synthetic conjugates the TACA antigens were conjugated to the lipid A scaffold through a linker placed at the pseudoreducing anomeric site of GlcN I, whereas in our semi-synthetic approach the conjugation site was the C-6 atom of GlcN II. The strategy described in this thesis should be more biologically well-tolerated, because intact LPS and LOS structures always show this GlcN II C-6 position involved in a linkage with the anomeric carbon of a carboxylic acid containing sugar residue such as Kdo or, more rarely, Ko.

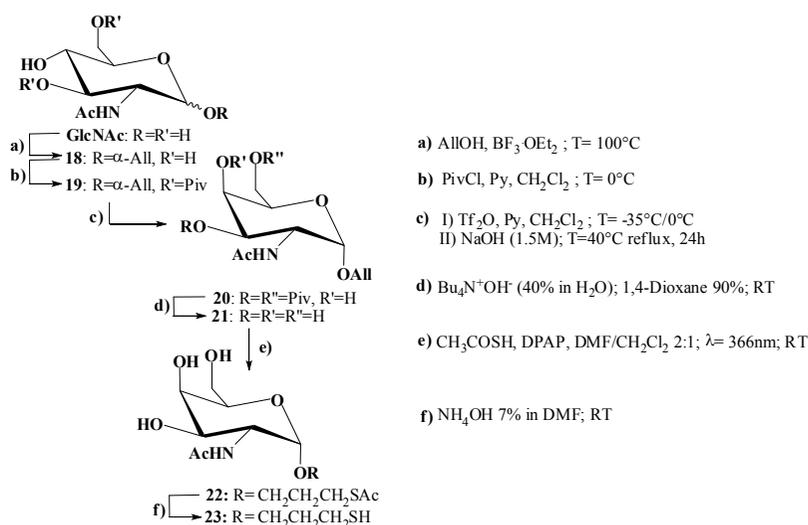
To demonstrate the feasibility of this approach, a first attempt was carried out through a thiol-ene coupling (Dondoni et al. 2008) of **5** with an amino acid. Indeed, by irradiation of the mixture of **5** and Fmoc-cysteine at $\lambda=366$ nm in 1:1 v/v DMF/CH₂Cl₂ in the presence of 2,2-dimethoxy-2-phenylacetophenone (DPAP) as photo-initiator (Fiore et al. 2009) and after mild basic treatment, the conjugate **17** was obtained as sole product in 53% (over two steps) yield after extensive chromatographic purification (Scheme 7.3) and then fully characterized by MALDI-TOF mass spectrometry and 2D-NMR spectroscopy.



Scheme 7.3: Coupling of clickable lipid A **5** with Fmoc-cysteine to give derivative **17** after Fmoc deprotection.

7.3.1 *E. coli* MPLA - Tn antigen conjugate

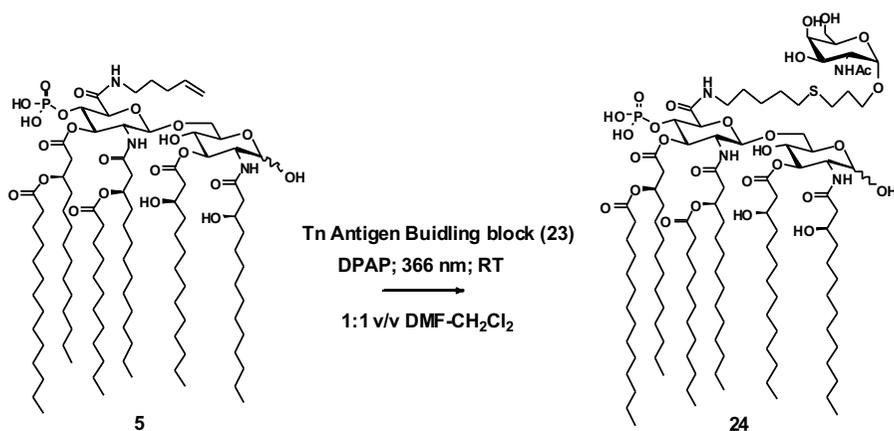
Another common feature of tumours is the overexpression of truncated *O*-glycans. The GalNAc-type *O*-glycans, also called mucintype *O*-glycans, are frequently found in most transmembrane and secreted glycoproteins. During malignancy, aberrant glycosylation also occurs in glycoproteins that display abnormal expression of shortened or truncated glycans, such as the disaccharide Thomsen-Friedenreich antigen (T antigen) and the monosaccharide GalNAc (also known as Tn) and their sialylated forms (ST and STn (Neu5Ac α 2-6GalNAc α -*O*-R), respectively), which result from the incomplete synthesis of *O*-glycans (Kudelka et al., 2015). For the first coupling of **5** with a TACA, Tn antigen -the structurally most simple TACA- was selected. To this end, the α -allyl glycoside of 2-acetamido-2-deoxy-D-galactopyranose (GalNAc) was suitably synthesized, starting from commercially available GlcNAc, to give compound **21** (Feng et al. 2004) the allyl aglycone was then converted into a 3-thioacetyl-propyl group to give **22**, which was then treated with ammonium hydroxide to afford **23** (Scheme 7.4).



Scheme 7.4: Synthesis of the Tn antigen derivative **23**.

Coupling of **5** and crude **23** under the photochemical reaction conditions described above furnished conjugate **24** after extensive chromatographic purification (Scheme 7.5).

Notably, compound **24** represents the first example of a lipid A-TACA derivative, as self-adjutant anticancer vaccine candidate, obtained by conjugation at the *C*-6 of GlcN II instead of *C*-1 of GlcN I (Zhou et al. 2014; Zhou et al. 2015; Tang et al. 2010). MALDI-TOF mass spectrum and full NMR characterization of the major hexa-acylated species confirmed its structure (Figures 7.21-7.22; Table 7.4).



Scheme 7.5: Coupling of clickable lipid A **5** with Tn Antigen derivative **23** to give compound **24**.

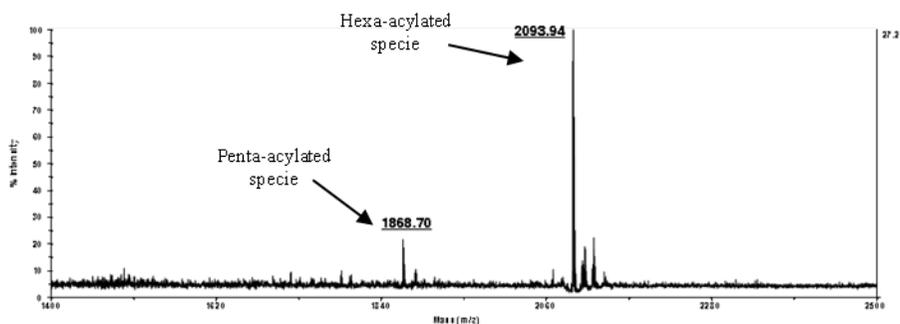


Figure 7.21: MALDI-TOF mass spectrum (negative ions mode) of pure **24**.

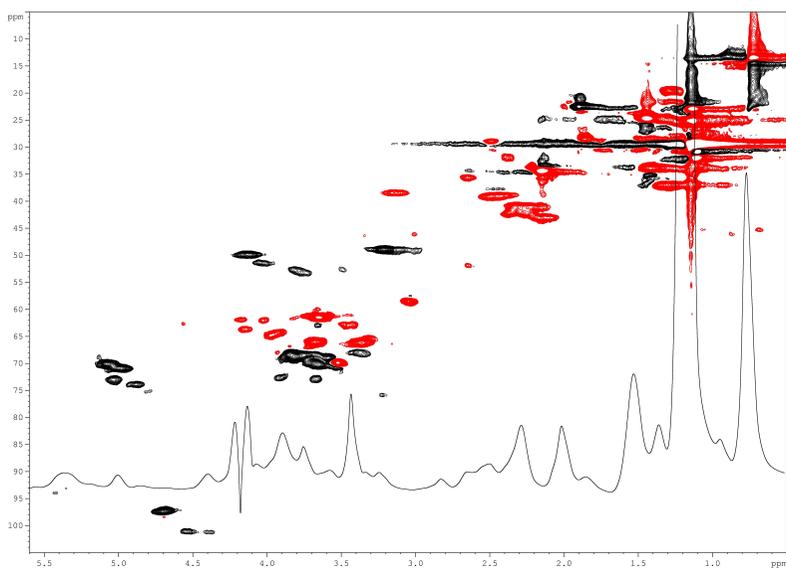


Figure 7.22: ¹H and HSQC-DEPT NMR spectra (600 MHz, 298K, 4:1 v/v CDCl₃-CD₃OD) of pure **24**.

Table 7.4: NMR data referred to lipid A derivative **24** (major hexa-acylated specie)

	Sugar portion ^a							Lipid portion ^a						Other signals ^a	Phosphate ^b
	1	2	3	4	5	6	α	β	γ	δ -[ω -1]	ω				
24	GlcN I (α)	5.43	4.03	5.03	3.38	3.91	3.51	C14:0(3-OH)	2.12, 2.21	3.76	1.28	1.08- 1.16 29.3	0.71- 0.75 13.5	5-S-Pentanoyl amide α : 3.14/38.4; β : 1.45/24.5; γ : 1.31/25.0; δ : 1.48/28.6; ϵ : 2.40/31.7 SCH ₂ CH ₂ CH ₂ O: 2.65/35.6 SCH ₂ CH ₂ CH ₂ O: 1.86/28.2 SCH ₂ CH ₂ CH ₂ O: 3.37, 3.67/66.3	-1.7
		<i>93.8</i>	<i>51.3</i>	<i>73.0</i>	<i>67.9</i>	<i>72.4</i>	<i>69.7</i>	[N-2 GlcN I]	<i>42.9</i>	<i>68.5</i>	<i>37.0</i>				
	GlcN II	4.54	3.80	4.88	3.56	3.68	---	C14:0(3-OH)	2.26, 2.36	3.85	1.30				
		<i>100.7</i>	<i>52.8</i>	<i>73.6</i>	<i>70.4</i>	<i>72.5</i>	<i>169.8</i>	[O-3 GlcN I]	<i>42.0</i>	<i>68.6</i>	<i>37.0</i>				
	GalNAc (α)	4.69	4.14	3.59	3.82	3.66	3.67	C14:0(3-OR)	2.26, 2.36	4.98	1.43				
		<i>97.0</i>	<i>49.8</i>	<i>68.9</i>	<i>68.6</i>	<i>70.0</i>	<i>61.3</i>	[N-2 GlcN II]	<i>41.2</i>	<i>70.7</i>	<i>33.8</i>				
							C14:0(3-OR)	2.45, 2.52	5.08	1.45					
							[O-3 GlcNII]	<i>39.1</i>	<i>69.8</i>	<i>33.8</i>					
							C12:0	2.15	1.44	1.08-1.16					
							C14:0	<i>34.4</i>	<i>24.5</i>	<i>29.3</i>					

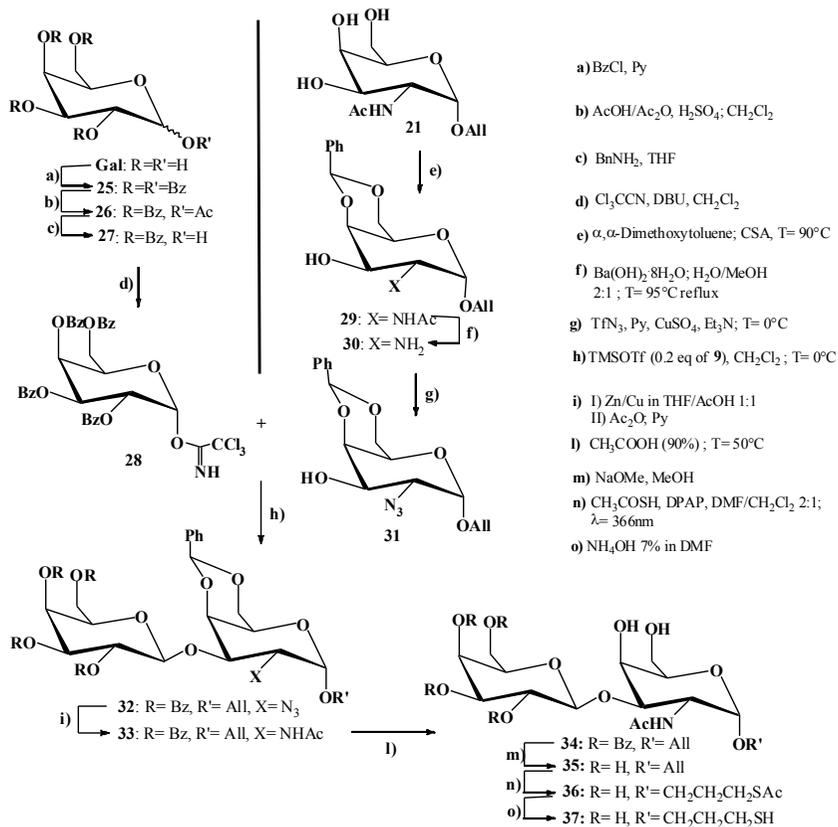
^a ¹H (600 MHz, 298K) and ¹³C (125 MHz, 298K) chemical shifts (in plain and in italic, respectively) expressed in δ relative to residual CHCl₃ (¹H: δ 7.26 ppm; ¹³C: δ 77.0 ppm) in 4:1 v/v CDCl₃-CD₃OD.

^b ³¹P (160 MHz, 298K) chemical shifts expressed in δ relative to 85% phosphoric acid (external standard).

7.3.2 *E. coli* MPLA - TF antigen conjugate

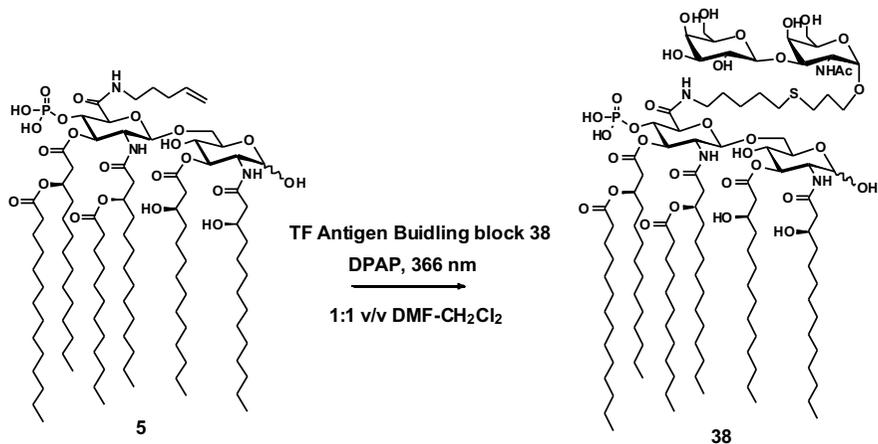
In order to obtain our second self-adjuvant anticancer vaccine candidate, a TF antigen derivative equipped with a thiolated aglycone (**37**, Scheme 7.6) was firstly synthesized from known galactosyl donor **28** (Ivanova et al., 1999) and 2-azido-galactosyl acceptor **31** (Bedini et al., 2012). Their glycosylation by catalytic trimethylsilyl trifluoromethanesulfonate (TMSOTf) activation of the trichloroacetimidate leaving group of **28**, furnished β -linked disaccharide **32** in good yield (84%). The β -stereoselectivity was proven by NMR ($^3J_{\text{H1H2}}$) and allowed through an anchimeric assistance of the benzoyl group on the C-2. Azide reduction with Zn/Cu couple in 1:1 v/v tetrahydrofuran (THF)-acetic acid (AcOH), followed by *N*-acetylation afforded **33** (52% over two steps), that was in turn subjected to benzylidene acetal hydrolysis in 9:1 v/v AcOH-H₂O at 50°C to afford **34** in 95% yield.

The sequence of protecting groups removal was completed by Zémpfen transesterification of Bz esters to afford deprotected disaccharide **35** (98%), carrying an allyl aglycone that was converted in a 3-thiopropyl group in two steps. Firstly, by irradiation of a mixture of **35** and thioacetic acid at $\lambda=366$ nm in 2:1 v/v DMF-CH₂Cl₂ in the presence of 2,2-dimethoxy-2-phenylacetophenone (DPAP) as photoinitiator (Fiore et al., 2009; Dondoni et al., 2009), a thiol-ene reaction converted the allyl aglycone into a 3-thioacetyl-propyl one (**36**, 35%; 69% based on reacted **35**). The latter was then subjected to thioester hydrolysis under mild alkaline conditions (7% aq. NH₄OH in DMF), to avoid dimerization of the resulting thiol to disulfide.



Scheme 7.6: Synthesis of TF antigen derivative **37**.

The thiolated TF antigen (**37**) was not purified, but immediately used for a further thiol-ene coupling with lipid A derivative **5** to furnish conjugate **38** (27%) after extensive purification (Scheme 7.7). MALDI-MS spectrum and full NMR characterization of the major hexa-acylated species confirmed its structure (Figures 7.23-7.24; Table 7.5).



Scheme 7.7: Coupling of clickable lipid A **5** with TF Antigen derivative **37** to give compound **38**.

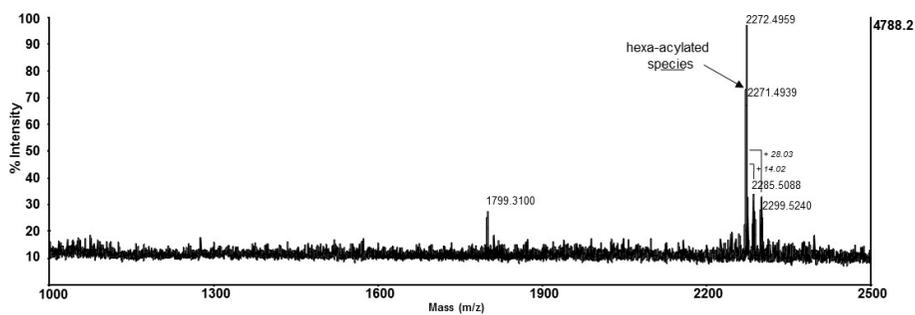


Figure 7.23: MALDI-TOF mass spectrum (negative ions mode) of pure **38**.

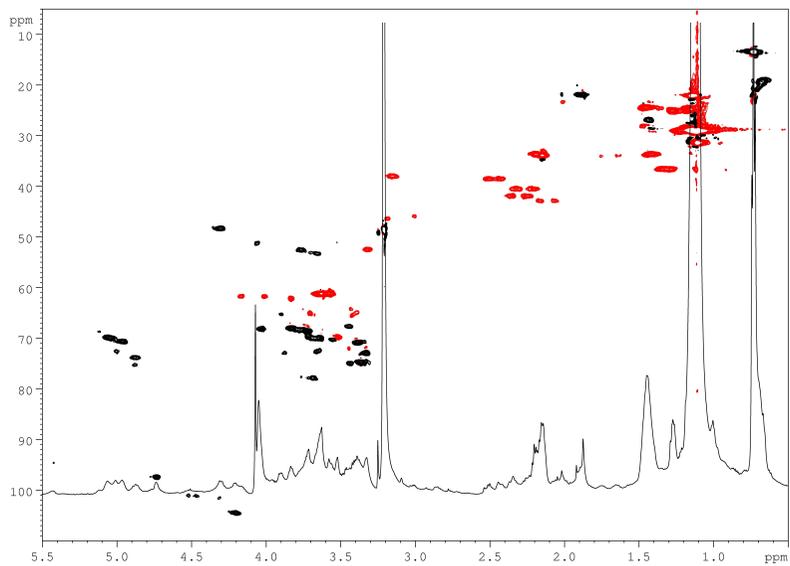


Figure 7.24: ¹H and HSQC-DEPT NMR spectra (600 MHz, 298K, 4:1 v/v CDCl₃-CD₃OD) of pure **38**.

Table 7.5: NMR data referred to lipid A derivative **38** (major hexa-acylated species)

	Sugar portion ^a						Lipid portion ^a						Other signals ^a	Phosphate ^b		
	1	2	3	4	5	6	C=O	α	β	γ	δ -(ω -1)	ω				
38	GlcN I (α)	5.43	4.06	5.00	3.45	3.88	3.53	C14:0(3-OH) [N-2 GlcN I]	--- <i>171.0</i>	2.08, 2.19 <i>42.8</i>	3.73 <i>68.4</i>	1.27 <i>36.6</i>	1.09-1.18 <i>22.0</i> <i>29.0</i> <i>31.4</i>	0.71-0.75 <i>13.5</i>	<p>ϵ-S-pentanoyl amide:</p> <p>α: 3.15/38.0; β: 2.02/23.4; γ: 1.15/24.7; δ: 1.24/29.0; ϵ: 3.01/45.8</p> <p>SCH₂CH₂CH₂O: 3.19/46.1 SCH₂CH₂CH₂O: 2.01/23.3 SCH₂CH₂CH₂O: 3.39, 3.70/64.8</p> <p>NHAc: CH₃: 1.88/21.5 CO: ---/172.6</p>	-0.9
		C14:0(3-OH) [O-3 GlcN I]	--- <i>170.1</i>	2.25, 2.36 <i>41.9</i>	3.84 <i>67.8</i>	1.29, 1.34 <i>36.6</i>										
	GlcN II	4.47 <i>100.9</i>	3.66 <i>53.1</i>	4.87 <i>73.7</i>	3.77 <i>68.4</i>	3.65 <i>72.4</i>	---	C14:0(3-OR) [N-2 GlcN II]	--- <i>170.4</i>	2.21, 2.32 <i>40.5</i>	4.96 <i>70.6</i>	1.41 <i>33.7</i>				
								C14:0(3-OR) [O-3 GlcN II]	--- <i>170.1</i>	2.44, 2.52 <i>38.5</i>	5.06 <i>69.8</i>	1.45 <i>33.7</i>				
	GalNAc	4.73 <i>97.2</i>	4.31 <i>48.3</i>	3.70 <i>77.6</i>	4.03 <i>68.0</i>	3.34 <i>74.9</i>	3.58, 3.64 <i>61.0</i>	C12:0	--- <i>173.3</i>	2.15	1.45	1.09-1.18 <i>22.0, 29.0, 31.4</i>				
								C14:0	--- <i>174.2</i>	33.9	24.4					
Gal	4.21 <i>104.2</i>	3.39 <i>70.8</i>	3.33 <i>73.0</i>	3.72 <i>69.4</i>	3.44 <i>75.0</i>	3.83 <i>62.2</i>										

^a ¹H (600 MHz, 298K) and ¹³C (150 MHz, 298K) chemical shifts (in plain and in italic, respectively) expressed in δ relative to residual CHCl₃ (¹H: δ 7.26 ppm; ¹³C: δ 77.0 ppm) in 4:1 v/v CDCl₃-CD₃OD.

^b ³¹P (160 MHz, 298K) chemical shifts expressed in δ relative to 85% phosphoric acid (external standard).

7.4 Biological assays

A panel of immunological assays, on the bone marrow-derived macrophages (BMDMs) and THP-1 cells, incubated with our lipid A derivatives **4** - **7**, **17**, **24**, **38**, *E.coli* Monophosphoryl lipid A (**2**) and *Salmonella* Minnesota MPL[®] (**3**) obtained by extraction in our laboratory, together with commercially available *E. coli* LPS and *S. Minnesota* MPL[®] **3**, were performed, in the laboratories of Prof. Gamian at the “L. Hirszfeld” Institute of Immunology and Experimental Therapy in Poland, in order to assess their cytotoxicity and immunostimulatory activity. Since one of the mechanisms of action of immunoadjuvants is to create a local pro-inflammatory environment, the ability of our compounds to modulate monocytes and macrophages for cytokine production was tested. Activation of macrophages is essential for induction of adaptive immune responses. Increased expression of MHC class II, activation marker CD86, and maturation marker CD83 leads to enhanced ability of antigen presenting cells to induce T lymphocyte activation and differentiation. When BMDMs were incubated with the lipid A derivatives, we observed very low level of induction of MHC class II and costimulatory molecules CD40, CD80, and CD86. Expression levels of these molecules were similar to levels in untreated controls and substantially lower than those induced by LPS or **3** (data not shown). However, BMDMs co-cultured for 6, 24, 48 and 72 h with the tested lipid A compounds were found to induce pro-inflammatory cytokines in comparison to untreated controls. Among the semi-synthetic lipid As, derivatives **4-6** and **38** influenced high level of interleukin-6 (IL-6) (Figure 7.25A). Most importantly, **6** induced significantly higher levels of IL-6 than **3**. We observed that almost all tested compounds induced

TNF- α production, however **38** showed minimal statistically non-significant cytokine induction. It is worth noting that **4**, **6**, and **24** induced higher levels of TNF- α than **3** (Figure 7.25B). Both cytokines were detected from 6 hours after stimulation onward and increased thereafter. All tested derivatives induced neither the production of IL-2, -4, -5, -10, -12, interferon gamma (IFN- γ), nor granulocyte macrophage colony-stimulating factor (GM-CSF) in BMDMs (data not shown).

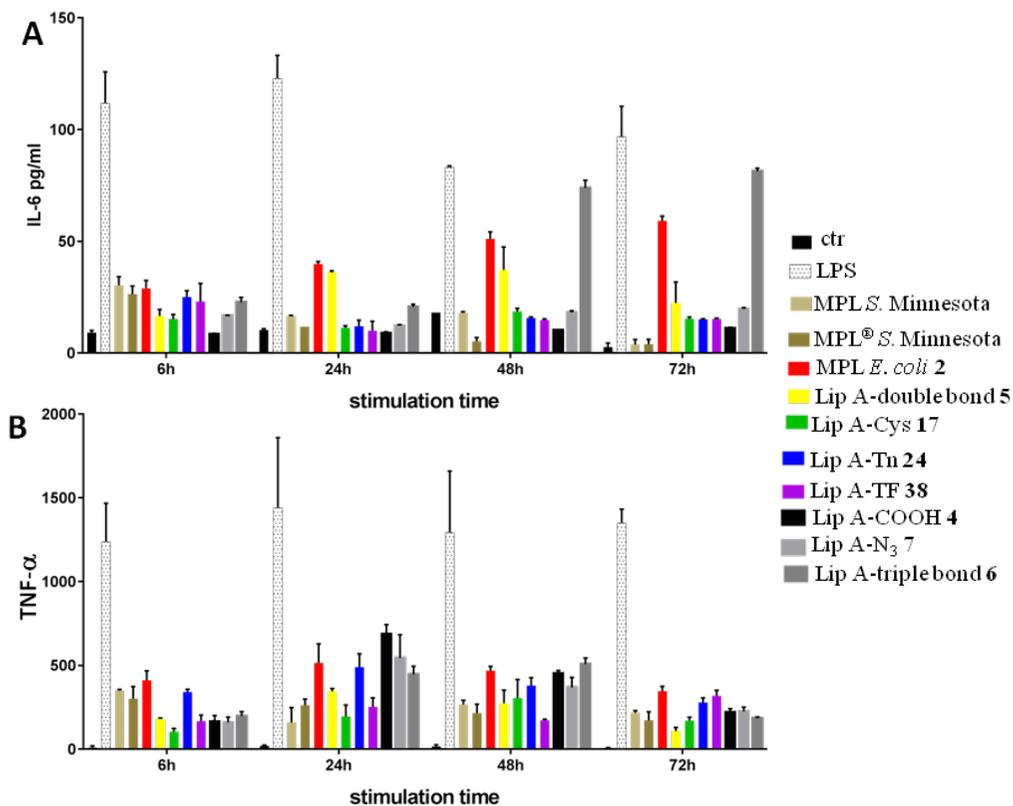
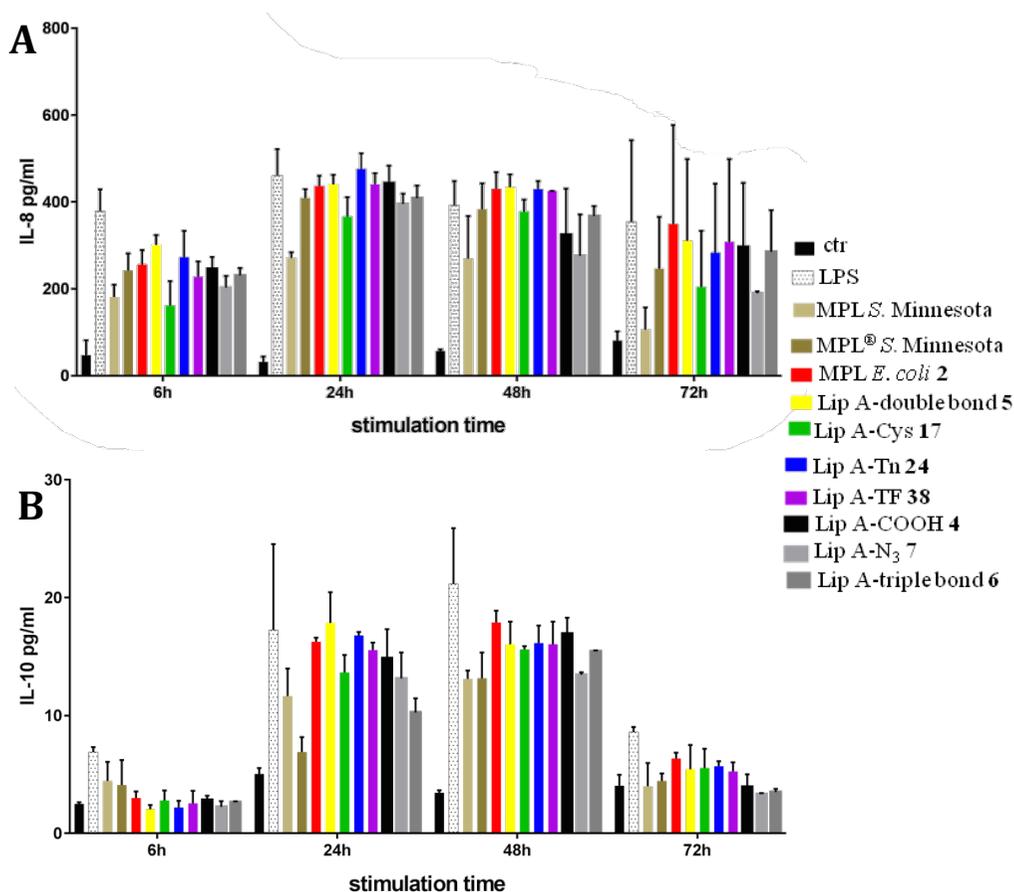


Figure 7.25: Cytokine IL-6 (A) and TNF- α (B) concentrations in tissue-culture supernatants of BMDM cells stimulated with lipid A derivatives (20 $\mu\text{g/mL}$). Pooled results from three independent experiments are shown. Data are means \pm standard deviation (SD) (* $P < 0.05$, ** $p < 0.005$, *** $P < 0.001$, **** $P < 0.0001$).

In supernatants collected from THP-1 cells (human monocytic cell line derived from an acute monocytic leukemia patients) cultured with semi-

synthetic lipid A derivatives, the cytokine production patterns appeared to be stimulus-dependent, and we observed increased production of IL-8, IL-6, TNF- α , IL-10, and IFN- γ . In general, all the tested compounds stimulated THP-1 cells to produce large amounts of IL-8 (Figure 7.26A) and low level of IL-10, but with highest up-regulation than **3** (Figure 7.26B). Compounds **4-7**, and **38** notably increased the level of TNF- α (Figure 7.26C) compared to control. Moreover, after stimulation with **4**, **5** and **24** a significant level of IL-6 was detected (Figure 7.26D). Remarkably, most of the tested compounds poorly induced IFN- γ secretion (Figure 7.26E), with exceptions of those treated with **5**, **24** and **38**, which induced significantly higher level of this cytokine compared to the control.



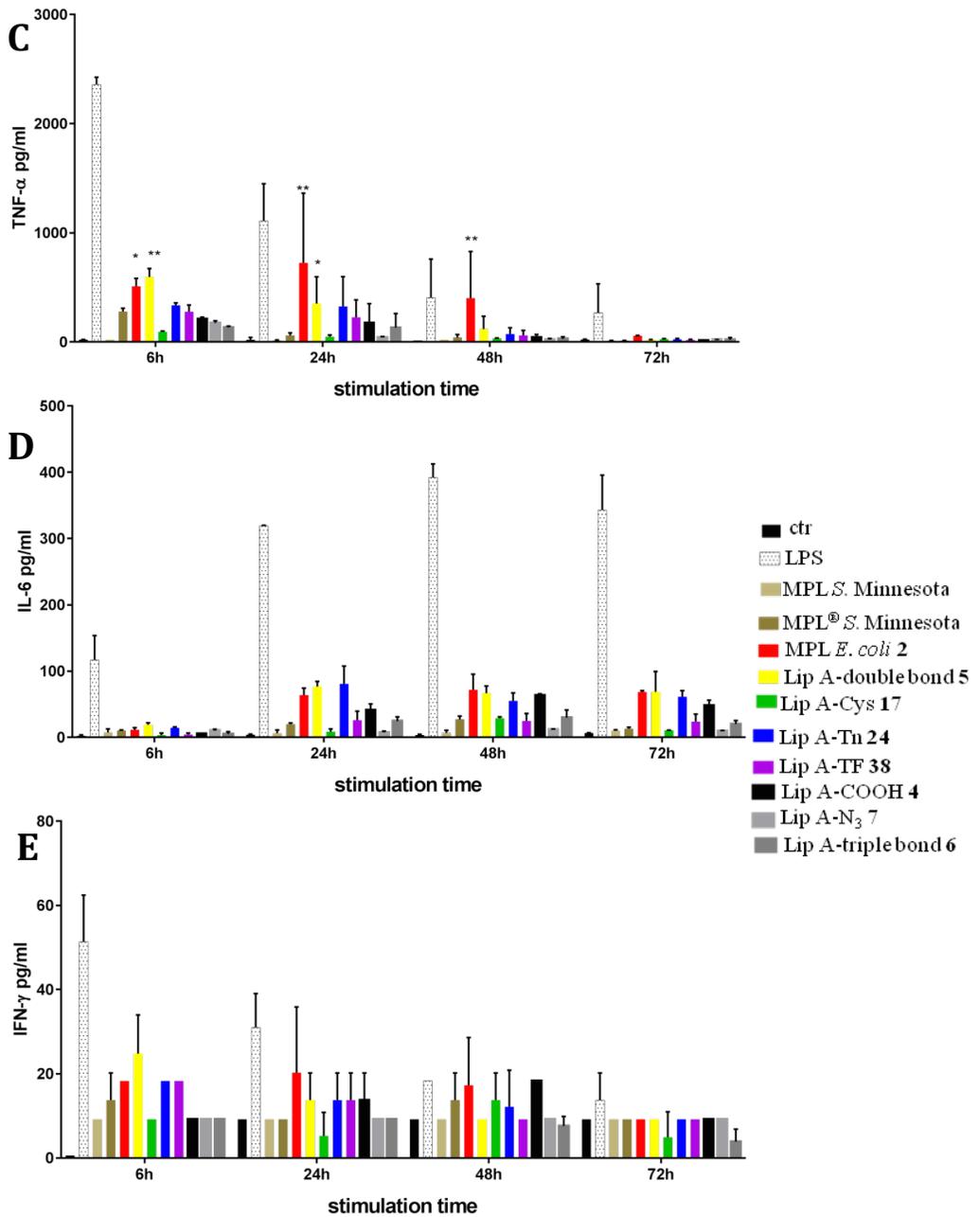


Figure 7.26: Cytokine IL-8 (A), IL-10 (B), IL-6 (C), TNF- α (D) and IFN- γ (E) concentrations in tissue-culture supernatants of THP-1 cells stimulated with lipid A derivatives (20 μ g/mL). Pooled results from three independent experiments are shown. Data are means \pm SD (* $P < 0.05$, ** $p < 0.005$, *** $P < 0.001$, **** $P < 0.0001$).

Our study shows that some of the tested compounds are able to induce the production of pro-inflammatory cytokines at least on the same level as commercially available **3** (Powell et al. 1995), thus suggesting that they could act as immunostimulant by promoting the generation of a specific immune response.

Vaccine delivery composition should include adjuvants for bolstering the immune response, but without the associated cytotoxicity. First and foremost, unacceptable side-effects and toxicity preclude the use of many adjuvant candidates, and this is particularly true for prophylactic vaccines where safety issues are paramount. The cytotoxic effect of the semi-synthetic lipid A derivatives was examined by performing sulforhodamine B (SRB) assay. It is a colorimetric, nondestructive and indefinitely stable assay, widely used for assessing cell viability in chemosensitivity testing (Skehan et al. 1990). The cell lines, THP-1 and BMDM, were incubated with the tested compounds at different concentrations: 20, 10, 1 and 0.1 $\mu\text{g/mL}$. It can be noticed that all semi-synthesized lipid A derivatives are non-toxic toward THP-1 cells (Figure 7.27). However, results show that **5**, **6**, and **24** (20 $\mu\text{g/mL}$) are slightly toxic against BMDM with **5** showing the highest cytotoxicity effect (Figure 7.28).

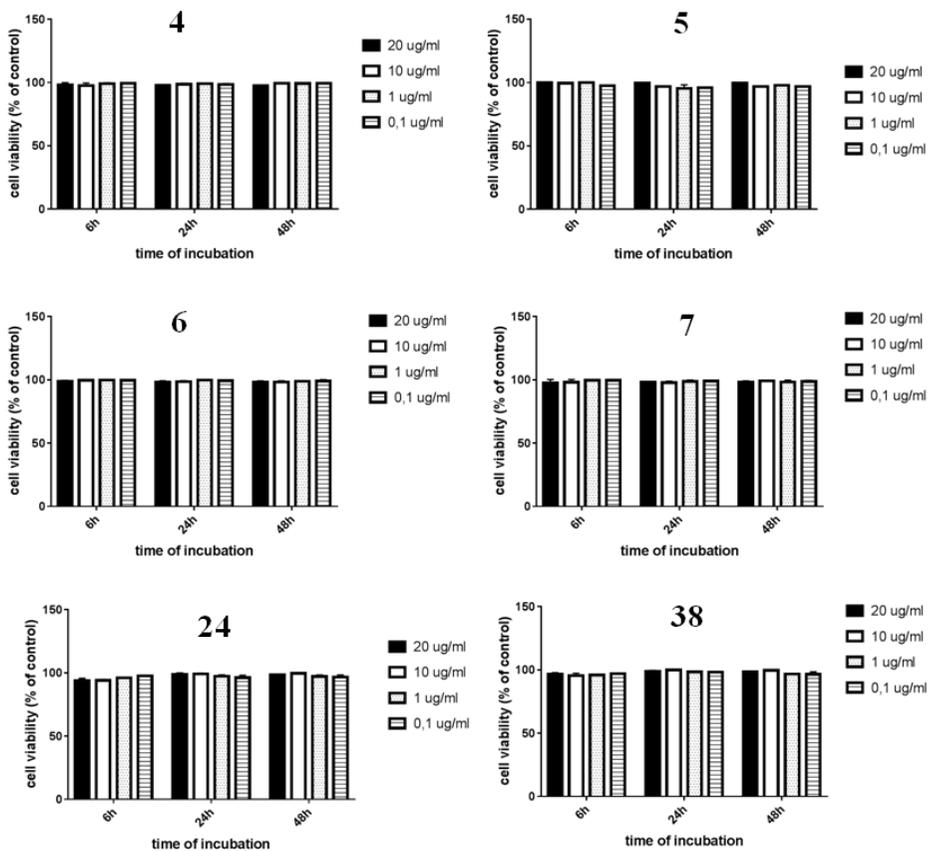


Figure 7.27: Cytotoxic effect of semi-synthetic lipid A derivatives toward THP-1 cells. Pooled results from three independent experiments are shown. Data are means \pm SD

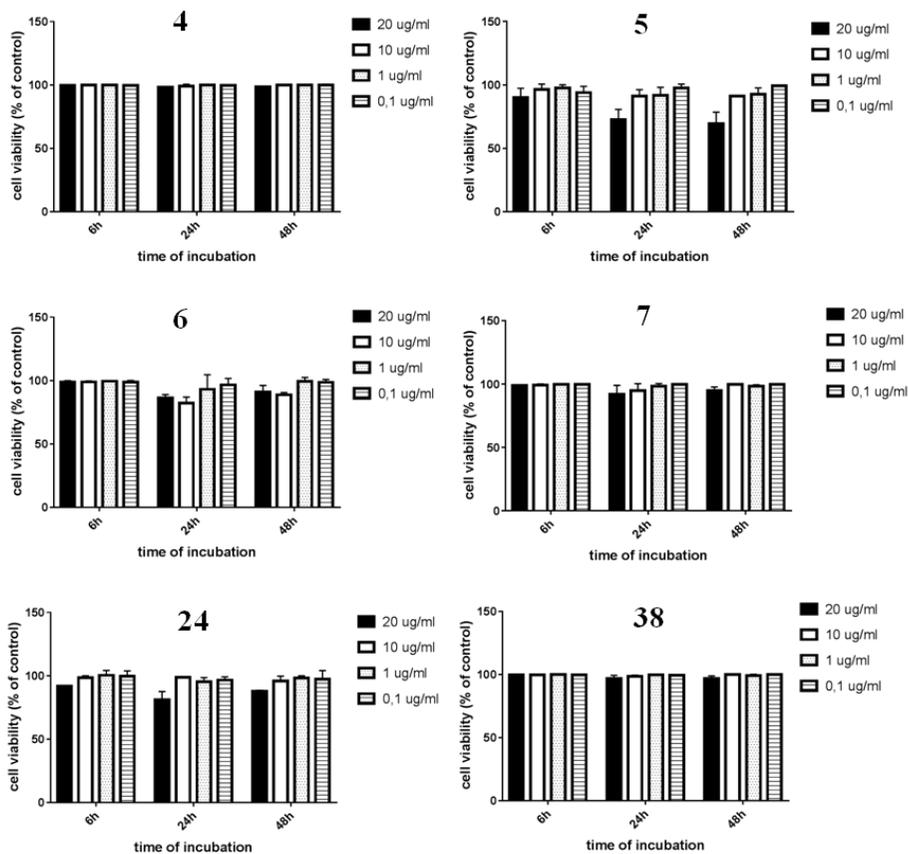


Figure 7.28: Cytotoxic effect of semi-synthetic lipid A derivatives toward BMDM cells. Pooled results from three independent experiments are shown. Data are means \pm SD.

7.5 Multivalent Tn antigens conjugates - MPLA

Following the idea to probe the potential of MPLA as a carrier molecule for the development of conjugate cancer vaccines, in the last part of the Ph.D. work, we started a collaboration with Prof. Renaudet, (Univ. Grenoble Alpes, CNRS, France). The group of Dumy and Renaudet have selected cyclopeptide carriers to construct multivalent synthetic vaccines (Grigalevicius et al. 2005). In particular, they focused their attention to the RAFT cyclodecapeptide platform (Regioselectively Addressable Functionalised Template) (Dumy et al. 1995; Dumy et al.

1996; Peluso et al. 2001) which is composed of two adjacent proline-glycine as β -turn inducers that constrain the backbone conformation into an antiparallel β sheet.

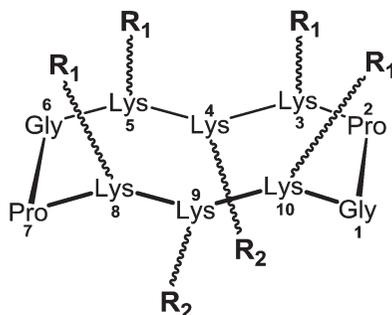


Figure 7.29: RAFT platform.

Up to six lysine residues can be introduced in the decapeptide sequence, in positions 3-4-5-8-9 and 10. The resulting scaffold thus presents two separate spatial domains where the Lys 3-5-8-10 are oriented in the upper plane, while the Lys 4-9 residues, where sometimes one of them is substituted by an Ala residue, delineate a second domain in the opposite plane (Figure 7.29). The cyclic nature of such scaffolds offers an improved stability towards enzymatic degradation compared to linear analogues, which makes cyclopeptides attractive scaffolds for biological applications. Renaudet and co-workers have developed an R4(GalNAc)-alkyne prototype by grafting four Tn antigen analogues as B-cell epitope, and a suitable alkyne-peptide for the coupling with our clickable Lipid A equipped with an azide moiety (7) as immuno-enhancing element, in order to increase and stimulate the T cell dependent response (the procedures for the coupling are still under investigation and they will not be shown in this thesis) (Figure 7.30). The native Tn antigen is depicted as a α -GalNAc unit linked to a Ser/Thr residue through a glycosidic bond. For synthetic considerations, Renaudet and Dumy have decided to introduce an oxyamine function on the anomeric carbon of the

α GalNAc unit to conjugate four copies of this Tn analogue on the upper side of the RAFT scaffold by oxime ligation (Ulrich et al. 2014).

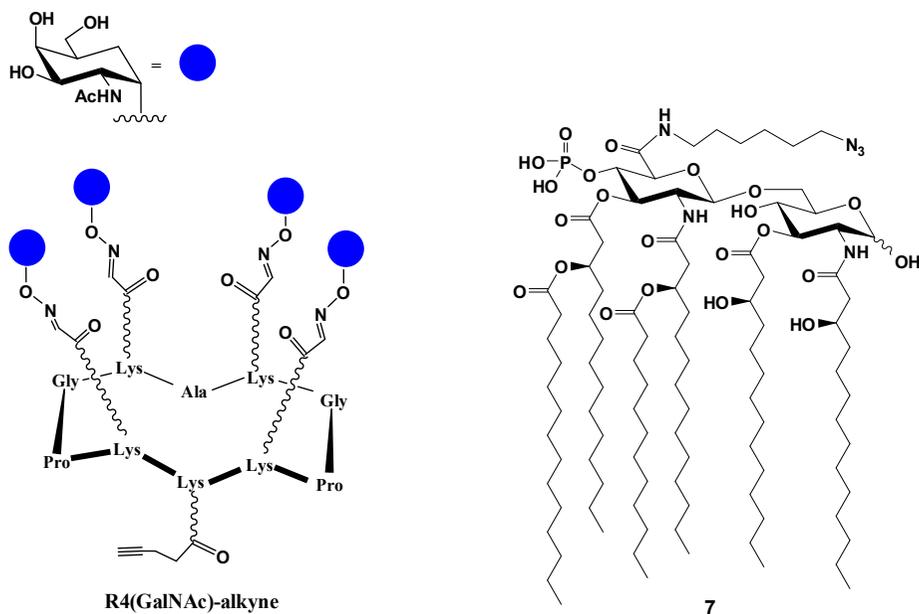


Figure 7.30: RAFT Tn-scaffold and clickable Lipid A 7.

Other structural modifications such as the presentation of a different display and/or higher density of carbohydrate epitopes, or combinations of different antigenic peptides or TACA glycopeptide mucin fragments represent interesting directions for future development towards immunological optimization (Pifferi et al. 2017).

Altogether, this strategy opens interesting perspectives to employ the RAFT scaffold-MPLA platform in engineering potent immunogenic conjugates as vaccine candidates. Indeed, the modularity and versatility of both the molecular assembly of RAFT and the derivatization of MPLA to make it clickable as well as the efficiency of MPLA as T-helper cells adjuvant might be exploited to improve the immune response by

combining TACAs with immunoadjuvants in order to develop multicomponent therapeutic anti-cancer vaccines.

7.6 Conclusions

A development of a semi-synthetic strategy for the insertion of clickable moieties (azide, alkyne, double bond, and a thiol precursor) into monophosphoryl lipid A obtained from *E. coli* has been accomplished. The key step was the coupling between a lipid A derivative carrying a carboxylic acid moiety at C-6 position of GlcN II and suitably ω -functionalized amines. This reaction was carefully investigated in order to find the best conditions for the selective obtainment of carboxyl amides and avoid, or at least minimize, the concomitant phosphate group amidation. The production of clickable lipid A derivatives opens a straightforward access to their conjugation with other interesting biomolecules, as, for example, the TACAs. This methodology stands as a valuable alternative to recently reported total synthetic approaches, based on highly complex syntheses of lipid As carrying at anomeric position of their pseudoreducing end a suitable moiety for conjugation with TACA antigens (Liao et al. 2016). In order to demonstrate the feasibility of our approach, the thiol-equipped derivatives of Tn and TF antigens were synthesized and then conjugated with a clickable lipid A derivative carrying a double bond moiety through a UV-mediated thiol-ene reaction, to give two potential self-adjuvant anticancer vaccine candidates. A preliminary evaluation of the immunological activity of these lipid A-TACA antigen conjugates, as well as of other semi-synthesized lipid A derivatives, including a Cys-appended lipid A derivative, was performed (Ziaco et al., 2017). It was shown that they

ameliorated production of proinflammatory cytokines IL-6, TNF- α , and IFN- γ , which are crucial in the process of activating immune cells and fighting infection. The tested derivatives also induce chemokine CXCL8 (IL-8) which provokes neutrophil migration to the site of infection, and immunoregulatory cytokine IL-10, an important B-cell differentiation factor, thus highlighting the promising immunobiological applications. As it is stated above, the mechanism of action of the examined lipid A derivatives is complex, but only such a versatile immunostimulation can give a therapeutic effect. Additionally, the described compounds exhibit very low toxicity which could facilitate their use in vaccines and give them an advantage over the currently used adjuvants. However, further experiments to evaluate these concepts are needed. Following this strategy, a collaboration with Prof. Renaudet has been started, in order to develop the first RAFT scaffold-MPLA conjugates as potent multivalent immunogenic vaccine candidates.

7.7 Experimental section

7.7.1 Lipid A derivative 4

A solution of **2** (183 mg, 106 μ mol) in CH_2Cl_2 (2.42 mL) was cooled to 0°C and then treated with H_2O (415 mL), 2,2,6,6-tetramethyl-1-piperidinyloxy free radical (TEMPO, 10.7 mg, 68.5 μ mol), 1M aqueous NaBr (60 μ L), 1M aqueous Bu_4NBr (120 μ L), 1M aqueous NaHCO_3 (300 μ L) and finally with an aqueous solution of NaOCl (365 μ L, minimum 4% chlorine content). The biphasic mixture was stirred at RT for 1 h, then it was neutralized by addition of 1M aqueous HCl. *t*BuOH (1.69 mL), a 2M solution of 2-methylbut-2-ene in THF (3.42 mL), and a solution of NaClO_2 (122 mg, 1.35 mmol) and NaH_2PO_4 (93.8 mg, 782

mmol) in H₂O (320 μL) were successively added. After 3 h stirring at RT, the mixture was diluted with sat. aq. NaH₂PO₄ (15 mL) and extracted with CHCl₃ (20 mL). The organic phase was collected, dried over anhydrous Na₂SO₄, filtered and concentrated to give **4**. For NMR data see Figure 7.2B and Table 7.1. MALDI TOF-HRMS for C₉₄H₁₇₅N₃O₂₃P (major hexa-acylated specie) (*m/z*): *M* (calcd) 1731,23, *M* (found) 1730.28 (M-H)⁺ (Figure 7.2A).

7.7.2 Lipid A derivative 5

Procedure A (see Table 7.2, entry 2): Compound **4** (25.0 mg, 14.5 μmol) was treated under argon atmosphere at 0°C with a solution of CDMT (5.6 mg, 32 μmol) in THF (0.5 mL), and then with a solution of NMM (3.50 μL, 32 μmol) in THF (35 μL). After stirring at 0°C for 7 hours, a solution of 4-pentene-1-amine (2.7 mg, 31.8 μmol) in THF (0.2 mL) was added and the suspension was stirred at rt overnight. The reaction mixture was then diluted with CHCl₃ (2 mL) and washed with H₂O (2 mL). The organic phase was collected, dried over anhydrous Na₂SO₄, filtered and concentrated. The residue was subjected to gel filtration chromatography on Sephadex[®] LH-20 (CHCl₃/MeOH, 3:2 v/v) to afford **5** (10.2 mg, 39%) as a yellowish oil.

Procedure B (see Table 7.2, entries 3 and 4): Compound **4** (20.0 mg, 11.6 μmol) was treated under argon atmosphere with a solution of PyBOP[®], (7.2 mg, 14 μmol) and HOBT (1.9 mg, 14 μmol) in 1:1 v/v DMF-CH₂Cl₂ (0.6 mL) and then a solution of 4-pentene-1-amine (3.0 mg, 35 μmol) in 1:1 v/v DMF-CH₂Cl₂ (0.6 mL) was added. In the case of reaction of entry 4, H₂O (2.0 μL, 111 μmol) was also added. Finally, the mixture was treated with DIPEA (9.0 μL, 51 μmol) and stirred at rt for 3 hours. The reaction mixture was then diluted with CHCl₃ (2 mL) and

washed with 0.1 M HCl (2 mL). The organic phase was collected, dried over anhydrous Na₂SO₄, filtered and concentrated. The residue was subjected to gel filtration chromatography on Sephadex[®] LH-20 (CHCl₃/MeOH, 3:2 v/v) to afford **5** (14.4 mg, 69% for entry 3, 17.0 mg, 82% for entry 4) as a white amorphous solid. Procedure C (see Table 7.2, entry 5 and 6): Compound **4** (20.0 mg, 11.6 μmol) was treated under argon atmosphere with a solution of TBTU (4.6 mg, 14 μmol) [also containing HOBt (1.9 mg, 14 μmol) in the case of reaction of entry 6] in 1:1 v/v DMF-CH₂Cl₂ (0.6 mL) and then a solution of 4-pentene-1-amine (3.0 mg, 35 μmol) in 1:1 v/v DMF-CH₂Cl₂ (0.6 mL) was subsequently added. Finally, the solution was treated with DIPEA (9.0 μL, 51 μmol) and stirred at rt for 3 hours. The reaction mixture was then diluted with CHCl₃ (2 mL) and washed with 0.1 M HCl (2 mL). The organic phase was collected, dried over anhydrous Na₂SO₄, filtered and concentrated. The residue was subjected to gel filtration chromatography on Sephadex[®] LH-20 (CHCl₃/MeOH, 3:2 v/v) to afford **5** (14.0 mg, 68% for entry 5, 17.1 mg, 82% for entry 6) as a yellowish oil. For NMR data see Figure 7.11 and Table 7.3. MALDI TOF-HRMS for C₉₉H₁₈₄N₃O₂₂P (major hexa-acylated species) (*m/z*): *M*_r (calcd) 1798.3109, *M*_r (found) 1797.3089 (M-H)⁺ (Figure 7.10).

7.7.3 Lipid A derivative 6

Compound **4** (20.0 mg, 11.6 μmol) was treated under argon atmosphere with a solution of TBTU (4.6 mg, 14 μmol) and HOBt (1.9 mg, 14 μmol) in 1:1 v/v DMF-CH₂Cl₂ (0.6 mL) and then a solution of 5-hexyn-1-amine (3.4 mg, 35 μmol) in 1:1 v/v DMF-CH₂Cl₂ (0.6 mL) was subsequently added. Finally, the mixture was treated with DIPEA (9.0 μL, 51 μmol) and stirred at rt for 3 hours. The reaction mixture was then

diluted with CHCl_3 (2 mL) and washed with 0.1 M HCl (2 mL). The organic phase was collected, dried over anhydrous Na_2SO_4 , filtered and concentrated. The residue was subjected to gel filtration chromatography on Sephadex[®] LH-20 ($\text{CHCl}_3/\text{MeOH}$, 3:2 v/v) to afford **6** (13.1 mg, 62%) as a yellowish oil. For NMR data see Figure 7.14 and Table 7.3. MALDI TOF-HRMS for $\text{C}_{100}\text{H}_{184}\text{N}_3\text{O}_{22}\text{P}$ (major hexa-acylated species) (m/z): M (calcd) 1810.3109, M (found) 1809.3092 ($M-H$) (Figure 7.13).

7.7.4 Lipid A derivative 7

Compound **4** (15.0 mg, 8.67 μmol) was treated under argon atmosphere with a solution of PyBOP[®] (5.4 mg, 11 μmol) and HOBt (1.4 mg, 11 μmol) in 1:1 v/v DMF- CH_2Cl_2 (0.6 mL) and then a solution of 6-azido-hexan-1-amine (3.7 mg, 26 μmol) and H_2O (1.5 μL , 84 μmol) in 1:1 v/v DMF- CH_2Cl_2 (0.6 mL) was added. Finally, the mixture was treated with DIPEA (6.8 μL , 39 μmol) and stirred at rt for 3 hours. The reaction mixture was then diluted with CHCl_3 (2 mL) and washed with 0.1 M HCl (2 mL). The organic phase was collected, dried over anhydrous Na_2SO_4 , filtered and concentrated. The residue was subjected to gel filtration chromatography on Sephadex[®] LH-20 ($\text{CHCl}_3/\text{MeOH}$, 3:2 v/v) to afford **7** (5.6 mg, 35%) as a yellowish oil. For NMR data see Figure 7.16 and Table 7.3. MALDI TOF-HRMS for $\text{C}_{100}\text{H}_{187}\text{N}_6\text{O}_{22}\text{P}$ (major hexa-acylated species) (m/z): M (calcd) 1855.3436, M (found) 1854.3333 ($M-H$) (Figure 7.15).

7.7.5 Lipid A derivative 8

Compound **4** (20.0 mg, 11.6 μmol) was treated under argon atmosphere with a solution of PyBOP[®] (7.2 mg, 14 μmol) and HOBt (1.9 mg, 14 μmol) in 1:1 v/v DMF- CH_2Cl_2 (0.6 mL) and then a solution of 5-

(tritylthio)pentan-1-amine (12.6 mg, 34 μmol) and H_2O (2.0 μL , 111 μmol) in 1:1 v/v DMF- CH_2Cl_2 (0.6 mL) was added. Finally, the mixture was treated with DIPEA (9.0 μL , 51 μmol) and stirred at rt for 3 hours. The reaction mixture was then diluted with CHCl_3 (2 mL) and washed with 0.1 M HCl (2 mL). The organic phase was collected, dried over anhydrous Na_2SO_4 , filtered and concentrated. The residue was subjected to gel filtration chromatography on Sephadex[®] LH-20 ($\text{CHCl}_3/\text{MeOH}$, 3:2 v/v) to afford **8** (19.8 mg, 83%) as a white amorphous solid. For NMR data see Figure 7.18 and Table 7.3. MALDI TOF-HRMS for $\text{C}_{118}\text{H}_{200}\text{N}_3\text{O}_{22}\text{PS}$ (major hexa-acylated species) (m/z): M (calcd) 2074.4082, M_r (found) 2073.4060 (M-H) (Figure 7.17).

7.7.6 Lipid A derivative 17

A mixture of **5** (20.2 mg, 11.2 μmol), Fmoc-cysteine (35.6 mg, 103 μmol), and 2,2-dimethoxy-2-phenylacetophenone (DPAP; 2.7 mg, 10.3 μmol) was dissolved in DMF/ CH_2Cl_2 (500 μL , 1:1 v/v) under argon atmosphere. The solution was briefly degassed and then irradiated at $\lambda=366$ nm for 1 h. The residue obtained after rotary evaporation was subjected to silica-gel column chromatography (CHCl_3 , then $\text{CHCl}_3/\text{acetone}$, 9:1 to 7:3 v/v, then $\text{CHCl}_3/\text{MeOH}$, 9:1 to 8:2 v/v, then $\text{CHCl}_3/\text{MeOH}/\text{H}_2\text{O}$, 18:6:1 v/v/v) to afford **16** (22.2 mg, 92%) as a slightly yellow amorphous solid. Compound **16** was dissolved in dry CH_2Cl_2 (400 μL), treated with 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU; 4 μL , 16.5 μmol) and stirred at rt for 30 minutes. The reaction mixture was neutralized with Dowex 50WX8 (H⁺ form) and then filtered and concentrated by rotary evaporation. The obtained residue was subjected to silica-gel column chromatography (CHCl_3 , then $\text{CHCl}_3/\text{MeOH}$, 9:1 to

7:3 v/v, then CHCl₃/MeOH/H₂O, 18:6:1 v/v/v) to afford **17** (10.2 mg, 53%) as a slightly yellow amorphous solid.

7.7.7 3'-Thioacetyl-propyl 2-acetamido-2-deoxy- α -D-galactopyranoside (**22**)

Allyl galactoside **21** (225 mg, 0.861 mmol) and DPAP (66.3 mg, 0.258 mmol) were dissolved in DMF/CH₂Cl₂ (3.6 mL, 2:1 v/v) and then thioacetic acid (183 μ L, 2.58 mmol) was added under argon atmosphere. The solution was briefly degassed and then irradiated at $\lambda=366$ nm for 4 h. The residue obtained after rotary evaporation was subjected to silica-gel column chromatography (CHCl₃/MeOH, 94:6 to 85:15 v/v) to afford, as first eluted fraction, **22** (123 mg, 42%) as a white powder. As second eluted fraction, unreacted **21** (106 mg, 0.406 mmol) was recovered as a white powder. Compound **22**: $[\alpha]_D^{25} = +146$ (c=0.7, MeOH); ¹H NMR (400 MHz, D₂O): δ =4.91 (d, *J* 3.5 Hz, 1H; H-1), 4.18 (dd, *J* 10.9, 3.5 Hz, 1H; H-2), 4.01–3.77 (m, 6H; H-3, H-4, H-5, H-6a, H-6b, OCHHCH₂CH₂SAc), 3.51 (dt, *J* 10.2, 5.9 Hz, 1H; OCHHCH₂CH₂SAc), 3.02 (t, *J* 6.6 Hz, 2H; OCH₂CH₂CH₂SAc), 2.40 (s, 3H; SCOCH₃), 2.07 (s, 3H; NCOCH₃), 1.91 ppm (quintet, *J* 6.4 Hz, 2H; OCH₂CH₂CH₂SAc); ¹³C NMR (100 MHz, D₂O): δ =174.4 (NCOCH₃, SCOCH₃), 96.8 (C-1), 70.7, 68.3, 67.6, 66.0, 61.0 (C-3, C-4, C-5, C-6, OCH₂CH₂CH₂SAc), 49.8 (C-2), 29.8, 28.0, 25.5, 21.7 ppm (NCOCH₃, OCH₂CH₂CH₂SCOCH₃).

7.7.8 Lipid A derivative **24**

Compound **22** (76.0 mg, 0.226 mmol) was suspended under argon atmosphere in a previously degassed 4:1 v/v mixture of DMF and

aqueous 33% NH_4OH (1.7 mL). After 3 h stirring at RT, the reaction mixture was neutralized with Dowex[®] 50WX8 (H^+ form) and then filtered and concentrated by rotary evaporation. The residue was mixed with **5** (20.0 mg, 11.1 μmol) and DPAP (21.2 mg, 82.7 μmol) and then dissolved in $\text{DMF}/\text{CH}_2\text{Cl}_2$ (2.0 mL, 1:1 v/v) under argon atmosphere. The solution was briefly degassed and then irradiated at $\lambda=366$ nm for 3 h. The residue obtained after rotary evaporation was subjected to silica-gel column chromatography (CHCl_3 , then $\text{CHCl}_3/\text{acetone}$, 9:1 to 7:3 v/v, then $\text{CHCl}_3/\text{MeOH}$, 98:2 to 7:3 v/v, then $\text{CHCl}_3/\text{MeOH}/\text{H}_2\text{O}$, 18:6:1 v/v/v). The fractions containing the product were pooled, concentrated and the obtained residue was further purified on Sephadex[®] LH-20 ($\text{CHCl}_3/\text{MeOH}$, 3:2 v/v) to afford **24** (7.4 mg, 32%) as a white powder. For NMR data see Figure 7.22 and Table 7.4.

7.7.9 Allyl 2,3,4,6-tetra-*O*-benzoyl- α -D-galactopyranosyl-(1 \rightarrow 3)-2-azido-4,6-*O*-benzylidene-2-deoxy- α -D-galactopyranoside (**32**)

A mixture of **28** (1.14 g, 1.54 mmol) and **31** (0.284 mg, 0.850 mmol) was coevaporated three times with dry toluene (5 mL). The residue was dried, mixed with freshly activated AW-300 4 \AA molecular sieves, and then suspended under argon atmosphere in CH_2Cl_2 (10 mL) at 0 $^\circ\text{C}$. A 0.84 M TMSOTf solution in CH_2Cl_2 (366 μL , 0.308 mmol) was then added. The mixture was stirred for 2 h at 0 $^\circ\text{C}$ and then quenched by adding one drop of Et_3N , filtered over a Celite pad, and concentrated. Silica-gel chromatography (4:1 to 2:1 v/v hexane-ethyl acetate) afforded **32** (0.657 g, 85%) as a yellowish solid. $[\alpha]_{\text{D}}^{25} +97$ (c 0.8, CHCl_3); ^1H NMR (600 MHz, CDCl_3): δ 8.08 (d, 2H, J 7.1 Hz, H_{ortho} Bz), 8.03 (d, 2H, J 7.6 Hz, H_{ortho} Bz), 8.01 (d, 2H, J 7.2 Hz, H_{ortho} Bz), 7.80 (d, 2H, J 7.1 Hz, H_{ortho} Bz), 7.58-7.25 (m, 17H, H-Ar), 5.99 (d, 1H, J 3.4 Hz, H-4 $_{\text{B}}$), 5.91 (dd,

1H, *J* 10.3, 7.9 Hz, H-2_B), 5.87 (m, 1H, OCH₂CH=CH₂), 5.58 (dd, 1H, *J* 10.3, 3.4 Hz, H-3_B), 5.49 (s, 1H, OCHPh), 5.31 (dd, 1H, *J* 17.2, 1.6 Hz, *trans* OCH₂CH=CHH), 5.22 (dd, 1H, *J* 10.4, 1.6 Hz, *cis* OCH₂CH=CHH), 5.15 (d, 1H, *J* 7.9 Hz, H-1_B), 5.01 (d, 1H, *J* 3.5 Hz, H-1_A), 4.77 (m, 1H, H-5_B), 4.46 (d, 1H, *J* 3.1 Hz, H-4_A) 4.40 (m, 2H, H-3_A, H-6_{AB}), 4.17-4.12 (m, 3H, H-6_A, H-6_B, OCHHCH=CH₂), 4.02 (dd, 1H, *J* 12.6, 5.6 Hz, OCHHCH=CH₂), 3.83 (dd, 1H, *J* 10.7, 3.4 Hz, H-2_A), 3.73 (d, 1H, *J* 12.3 Hz, H-6_B), 3.50 (bs, 1H, H-5_A); ¹³C NMR (100 MHz, CDCl₃): δ 165.9-165.2 (4 CO Bz), 137.7-133.1 (C_{ipso}, OCH₂CH=CH₂), 130.0-126.1 (C-Ar), 118.2 (OCH₂CH=CH₂), 103.0, 100.6, 97.6 (C-1_A, C-1_B, OCHPh), 76.3, 76.0, 72.0, 71.6, 69.6, 68.9, 68.8, 68.3, 63.1, 62.5, 58.7 (C-2_A, C-2_B, C-3_A, C-3_B, C-4_A, C-4_B, C-5_A, C-5_B, C-6_A, C-6_B, OCH₂CH=CH₂). MALDI TOF-HRMS for C₅₀H₄₅N₃O₁₄ (*m/z*): *M*⁺ (calcd) 911.2902, *M* (found) 934.2811 (M+Na)⁺. Anal. Calcd: C, 65.86, H, 4.97, N, 4.61; Found: C, 65.78, H, 5.01, N, 4.59.

7.7.10 Allyl 2,3,4,6-tetra-*O*-benzoyl- α -D-galactopyranosyl-(1 \rightarrow 3)-2-acetamido-2-deoxy- α -D-galactopyranoside (34)

To a solution of **32** (0.657 g, 0.721 mmol) in THF (25 mL) and AcOH (25 mL), Zn/Cu alloy (1.97 g) was then added and the mixture was vigorously stirred at rt overnight. The mixture was then diluted with ethyl acetate (10 mL), filtered over a Celite pad, and concentrated. The residue was coevaporated two times with toluene (50 mL) then diluted with ethyl acetate (150 mL) and washed with a saturated NaHCO₃ aqueous solution (150 mL). The organic phase was collected, dried over anhydrous Na₂SO₄, filtered and concentrated. The residue was then dissolved in pyridine (3 mL) and treated with acetic anhydride (3 mL). The solution was stirred at rt overnight and then quenched by dilution

with CH_2Cl_2 (100 mL) and washing with 0.1M HCl (100 mL). The organic phase was collected, dried over anhydrous Na_2SO_4 , filtered and concentrated. Silica-gel chromatography (3:1 to 1:3 v/v hexane-ethyl acetate) afforded **33** (0.349 g, 52% over two steps from **14**) as a yellowish solid. ^1H NMR (600 MHz, CDCl_3): δ 8.07 (d, 2H, J 7.3 Hz, H_{ortho} Bz), 8.00 (d, 2H, J 7.6 Hz, H_{ortho} Bz), 7.95 (d, 2H, J 7.6 Hz, H_{ortho} Bz), 7.78 (d, 2H, J 7.7 Hz, H_{ortho} Bz), 7.61-7.24 (m, 17H, H-Ar), 5.99 (d, 1H, J 3.4 Hz, H-4_B), 5.83 (m, 2H, H-2_B, $\text{OCH}_2\text{CH}=\text{CH}_2$), 5.60 (dd, 1H, J 10.1, 3.4 Hz, H-3_B), 5.42 (d, 1H, J 8.0 Hz, NH), 5.39 (s, 1H, OCHPh), 5.25 (d, 1H, J 17.4 Hz, *trans* $\text{OCH}_2\text{CH}=\text{CHH}$), 5.18 (d, 1H, J 10.1 Hz, *cis* $\text{OCH}_2\text{CH}=\text{CHH}$), 5.16 (d, 1H, J 7.8 Hz, H-1_B), 5.11 (d, 1H, J 3.4 Hz, H-1_A), 4.70 (dd, 1H, J 11.5, 6.8 Hz, H-6_{ab}), 4.62 (m, 1H, H-5_B), 4.46 (dd, 1H, J 11.4, 5.5 Hz, H-2_A), 4.40 (m, 2H, H-4_A, H-6_{bB}), 4.14 (m, 2H, H-6_{aA}, $\text{OCHHCH}=\text{CH}_2$), 4.08 (dd, 1H, J 11.3, 3.4 Hz, H-3_A), 3.98 (dd, 1H, J 13.1, 5.9 Hz, $\text{OCHHCH}=\text{CH}_2$), 3.76 (d, 1H, J 11.7 Hz, H-6_{ab}), 3.54 (bs, 1H, H-5_A), 1.41 (s, 3H, COCH_3). MALDI TOF-HRMS for $\text{C}_{52}\text{H}_{49}\text{NO}_{15}$ (m/z): M_r (calcd) 927.3102, M_r (found) 950.3008 ($\text{M}+\text{Na}$)⁺.

A solution of **33** (0.349 g, 0.376 mmol) in 9:1 v/v AcOH- H_2O (20 mL) was stirred at 50°C for 7h, then cooled to rt, concentrated and coevaporated two times with toluene (5 mL). Silica-gel chromatography (99:1 to 8:2 v/v CH_2Cl_2 -MeOH) afforded **34** (0.300 g, 95% from **33**) as a yellowish solid. $[\alpha]_{\text{D}} +139$ (c 0.4, CHCl_3); ^1H NMR (600 MHz, CDCl_3): δ 8.10 (d, 2H, J 8.1 Hz, H_{ortho} Bz), 8.02 (d, 2H, J 8.2 Hz, H_{ortho} Bz), 7.96 (d, 2H, J 8.3 Hz, H_{ortho} Bz), 7.75 (d, 2H, J 8.2 Hz, H_{ortho} Bz), 7.66-7.23 (m, 12H, H-Ar), 5.97 (d, 1H, J 3.2 Hz, H-4_B), 5.83 (m, 2H, H-2_B, $\text{OCH}_2\text{CH}=\text{CH}_2$), 5.58 (dd, 1H, J 10.4, 3.4 Hz, H-3_B), 5.18 (m, 3H, H-1_B, $\text{OCH}_2\text{CH}=\text{CH}_2$), 4.95 (d, 1H, J 7.9 Hz, NH), 4.89 (d, 1H, J 3.5 Hz, H-1_A), 4.56 (m, 3H, H-2_A, H-5_B, H-6_{ab}), 4.38 (m, 1H, H-6_{bB}), 4.18 (s, 1H,

H-4_A), 4.10 (dd, 1H, *J* 12.7, 5.3 Hz, OCH₂CH=CH₂), 3.92 (dd, 1H, *J* 12.7, 5.3 Hz, OCH₂CH=CH₂), 3.81 (dd, 1H, *J* 8.2, 2.0 Hz, H-6_{aA}), 3.74 (m, 2H, H-3_A, H-5_A), 3.52 (dd, 1H, *J* 8.2, 2.0 Hz, H-6_{bA}), 1.30 (s, 3H, COCH₃); ¹³C NMR (100 MHz, CDCl₃): δ 169.7 (COCH₃), 166.1, 165.6, 165.5, 164.7 (4 CO Bz), 133.8-133.4 (C_{ipso}, OCH₂CH=CH₂), 130.1-128.3 (C-Ar), 117.8 (OCH₂CH=CH₂), 102.0, 97.0 (C-1_A, C-1_B), 79.9, 72.1, 71.4, 69.7, 69.3, 68.9, 68.4, 68.1, 62.7, 62.5 (C-2_B, C-3_A, C-3_B, C-4_A, C-4_B, C-5_A, C-5_B, C-6_A, C-6_B, OCH₂CH=CH₂), 47.7 (C-2_A), 22.4 (COCH₃). MALDI TOF-HRMS for C₄₅H₄₅NO₁₅ (*m/z*): *M_r* (calcd) 839.2789, *M_r* (found) 862.2700 (M+Na)⁺. Anal. Calcd: C, 64.36, H, 5.40, N, 1.67; Found: C, 64.29, H, 5.44, N, 1.66.

7.7.11 Allyl α-D-galactopyranosyl-(1→3)-2-acetamido-2-deoxy-α-D-galactopyranoside (35)

A solution of **34** (0.330 g, 0.393 mmol) in dry CH₃OH (11 mL) was treated with a 0.94 M methanolic solution of CH₃ONa (2.1 mL, 2.0 mmol). After 3 h stirring at rt, the solution was diluted with H₂O (10 mL), neutralized with Amberlist-15 (H⁺ form), concentrated and coevaporated two times with toluene (5 mL). Silica-gel chromatography (99:1 to 7:3 v/v CH₂Cl₂-MeOH) afforded **35** (0.163 g, 98%) as a white fluffy solid. [α]_D +9.4 (c 1.3, H₂O); ¹H NMR (600 MHz, D₂O): δ 5.98 (m, 1H, OCH₂CH=CH₂), 5.35 (dd, 1H, *J* 15.8, 1.5 Hz, *trans* OCH₂CH=CH₂), 5.26 (dd, 1H, *J* 10.4, 1.5 Hz, *cis* OCH₂CH=CH₂), 4.94 (d, 1H, *J* 3.7 Hz, H-1_A), 4.46 (d, 1H, *J* 7.7 Hz, H-1_B), 4.24-3.35 (m, 14H, H-2_A, H-2_B, H-3_A, H-3_B, H-4_A, H-4_B, H-5_A, H-5_B, H-6_{aA}, H-6_{bA}, H-6_{aB}, H-6_{bB}, OCH₂CH=CH₂), 2.02 (s, 3H, NHCOCH₃); ¹³C NMR (100 MHz, D₂O): δ 174.6 (COCH₃), 133.7 (OCH₂CH=CH₂), 117.9 (OCH₂CH=CH₂), 104.7, 96.4 (C-1_A, C-1_B), 77.2, 75.0, 72.5, 70.6, 70.5,

68.7, 68.6, 68.4, 61.2, 61.0 (C-2_B, C-3_A, C-3_B, C-4_A, C-4_B, C-5_A, C-5_B, C-6_A, C-6_B, OCH₂CH=CH₂), 48.6 (C-2_A), 22.0 (COCH₃). MALDI TOF-HRMS for C₁₇H₂₆NO₁₁ (*m/z*): *M_r* (calcd) 423.1741, *M_r* (found) 446.1629 (M+Na)⁺. Anal. Calcd: C, 48.22, H, 6.90, N, 3.31; Found: C, 48.08, H, 6.96, N, 3.28.

7.7.12 3'-Thioacetyl-propyl β-D-galactopyranosyl-(1→3)-2-acetamido-2-deoxy-α-D-galactopyranoside (36)

Derivative **35** (0.190 g, 0.449 mmol) and DPAP (34.5 mg, 0.135 mmol) were dissolved in 2:1 v/v DMF-CH₂Cl₂ (1.8 mL) and then thioacetic acid (95.0 μL, 1.35 mmol) was added under argon atmosphere. The solution was briefly degassed and then irradiated at λ=366 nm for 6 h. The residue obtained after rotoevaporation was subjected to a silica-gel column chromatography (7:3:0 to 7:3:0.3 v/v/v CH₂Cl₂-MeOH-H₂O) to afford, as first eluted fraction, **36** (66.2 mg, 35%; 69% based on reacted **35**) as a yellowish powder and starting compound **35** (82.0 mg, 0.194 mmol) as second eluted fraction. [α]_D +19.6 (c 1.0, H₂O); ¹H NMR (600 MHz, D₂O): δ 4.85 (d, 1H, *J* 3.0 Hz, H-1_A), 4.44 (d, 1H, *J* 7.6 Hz, H-1_B), 4.30-3.47 (m, 14H, H-2_A, H-2_B, H-3_A, H-3_B, H-4_A, H-4_B, H-5_A, H-5_B, H-6_A, H-6_B, H-6_{AB}, H-6_{BB}, OCH₂CH₂CH₂S), 2.98 (m, 2H, OCH₂CH₂CH₂S), 2.34 (s, 3H, SCOCH₃), 2.00 (s, 3H, NHCOCH₃), 1.87 (m, 2H, OCH₂CH₂CH₂S); ¹³C NMR (100 MHz, D₂O): δ 202.5 (SCOCH₃), 175.2 (NHCOCH₃), 105.4, 97.8 (C-1_A, C-1_B), 77.9, 75.6, 73.2, 71.3, 71.2, 69.4, 69.2, 66.8, 61.8, 61.6 (C-2_B, C-3_A, C-3_B, C-4_A, C-4_B, C-5_A, C-5_B, C-6_A, C-6_B, OCH₂CH₂CH₂S), 49.3 (C-2_A), 30.7, 28.9, 26.4, 22.7 (OCH₂CH₂CH₂S, SCOCH₃, NHCOCH₃). MALDI TOF-HRMS for C₁₉H₃₃NO₁₂S (*m/z*): *M_r* (calcd) 499.1723, *M_r* (found) 522.1609

(M+Na)⁺. Anal. Calcd: C, 45.68, H, 6.66, N, 2.80, S, 6.42; Found: C, 45.49, H, 6.73, N, 2.76, S, 6.38.

7.7.13 Lipid A derivative 38

Compound **36** (30.0 mg, 60.1 μmol) was suspended under argon atmosphere in a previously degassed 4:1 v/v mixture of DMF and aqueous 33% NH₄OH (1.0 mL). After 3 h stirring at rt, the reaction mixture was neutralized with Dowex[®] 50WX8 (H⁺ form), then filtered and concentrated by rotoevaporation. The residue was immediately mixed with **5** (13.0 mg, 7.23 μmol) and DPAP (1.9 mg, 7.3 μmol) and then dissolved in 1:1 v/v DMF-CH₂Cl₂ (2.0 mL) under argon atmosphere. The solution was briefly degassed and then irradiated at $\lambda=366$ nm overnight. The residue obtained after rotoevaporation was subjected to column chromatography (CHCl₃, then 9:1 to 7:3 v/v CHCl₃/acetone, then 97:3 to 8:2 v/v CHCl₃/MeOH, then 18:6:1 to 12:6:1 v/v/v CHCl₃/MeOH/H₂O). The fractions containing the product were pooled, concentrated and the obtained residue was further purified by gel filtration chromatography to afford **38** (4.4 mg, 27%) as a white powder. For NMR data see Figure 7.24 and Table 7.5. MALDI TOF-HRMS for C₁₁₆H₂₁₅N₄O₃₃PS (major hexa-acylated species) (*m/z*): *M* (calcd) 2255.4727, *M* (found) 2271.4939 (M+NH₄⁺-2H)⁺ (Figure 7.23).

7.7.14 Cell lines and culture

To evaluate the potential cytotoxicity and immunomodulatory effect of the semi-synthetic lipid A derivatives, the human monocytic leukemia cell line THP-1 and BMDMs from wild type of C57BL/6 mice were stimulated. BMDMs were a gift of dr. Jakub Siednienko (Siednienko et al. 2010), whereas THP-1 cell line (ATCC) was obtained from a cell

bank at the Ludwik Hirszfeld Institute of Immunology and Experimental Therapy, Polish Academy of Sciences. Cells were grown on 75 cm² cell culture flasks (BD Bioscience) in Roswell park memorial institute (RPMI) medium (Sigma-Aldrich), supplemented with 10 % fetal bovine serum (FBS) (v/v, Sigma-Aldrich), penicillin (100 µg/mL, Sigma-Aldrich), streptomycin (100 µg/mL, Sigma-Aldrich), and, for BMDMs cultivation, additionally with gentamycin (5 mg/mL, Sigma-Aldrich). Cells were maintained at 37 °C in a humidified atmosphere containing 5% CO₂.

7.7.15 Activation of THP-1 and BMDMs cells

THP-1 or BMDMs cells were seeded (0.5×10^6 cells/well) on 48-well plates and grown for 24 h. The complete medium was replaced and cells were then treated with a medium solution of the tested derivatives (20 µg/mL) for 6, 24, 48 and 72 h. **2**, **3** (Sigma-Aldrich), **3** obtained in our laboratory and LPS from *Escherichia coli* 0111:B4 (Sigma-Aldrich) were used as positive controls. After incubation period, the cell-free supernatants were removed and analysed for cytokine content according to the manufacturer's instructions. For THP-1 the following cytokines were measured: IL-8, IL-6 (ELISA Ready-Set-Go! Kit, BD Bioscience) and IL-2, IL-4, IL-5, IL-10, IL-12, IL-13, IFN-γ, GM-CSF, TNF-α (Bio-Plex Pro Human Cytokine Th1/Th2 Immunoassay, Bio-Rad). For BMDMs the level of following cytokines was determined: IL-6, TNF-α (ELISA Ready-Set-Go! Kit, Invitrogen) and IL-2, IL-4, IL-5, IL-10, IL-12, IFN-γ, GM-CSF (Bio-Plex Pro Mouse Cytokine Th1/Th2 Immunoassay, Bio-Rad). Induced BMDMs cells were additionally labeled with monoclonal antibodies CD40, CD80, or CD86 (each conjugated to phycoerythrin; eBioscience) and MHC class II

(conjugated to allophycocyanin; eBioscience). Appropriate isotype antibodies were used as controls to determine nonspecific binding. Cells were analyzed using a FACSCalibur[®] flow cytometer (Becton-Dickinson), and obtained data were analyzed with Flowing Software 2.5.1. Three independent experiments were carried out. Statistical analysis was performed by two-way analysis of variance (ANOVA) followed by Dunnet's multiple comparison test using Prism 5.04 software (GraphPad, San Diego, CA). *P* values of < 0.05 were considered significant.

7.7.16 Cytotoxicity of semi-synthetic lipid A derivatives

To study the cytotoxic effect of semi-synthetic lipid A derivatives, the SRB assay was performed. In our research the following parameters of the assay were applied. THP-1 or BMDMs cells were seeded (2×10^4 cells/well) on 96-well plates and grown for 24 h. The complete medium was replaced and cells were then treated with a medium solution of the tested derivatives (20, 10, 1 or 0.1 $\mu\text{g/mL}$ in DMSO) for 6, 24 and 48 h. Then, cultures fixed with trichloroacetic acid were stained for 30 minutes with 0.4% (wt/vol) sulforhodamine B (SRB, Sigma-Aldrich) dissolved in 1% acetic acid. Unbound dye was removed by four washes with 1% acetic acid, and protein-bound dye was extracted with 10 mM unbuffered Tris base. The absorbance was measured at 510 nm using microtiter plate reader (Biotek, Tekane). Three independent experiments were carried out.

References

- Bedini E., Cirillo L., Parrilli M., *Carbohydr. Res.* **2012**, *349*, 24-32.
- Bitton R.J. et al., *Oncol Rep.* **2002**, *9*, 267-276.
- Buskas T. et al., *Angew Chem Int Ed* **2005**, *44*, 5985-5988.
- Campagne J. M., Coste J., Jouin P., *Tetrahedron Lett.* **1995**, *36*, 2079-2082.
- Casella C. R., Mitchell T. C., *Cell. Mol. Life Sci.* **2008**, *65*, 3231-3240.
- Chen Y.A. et al., *Chem Soc Rev* **2013**, *42*, 4532-4542.
- Coler R.N. et al., *PLoS One* **2011**, *6*, 16333.
- D'Alonzo D. et al., *Chem. Eur. J.* **2016**, *22*, 1053-11063.
- Danishefsky S.J. et al., *Angew Chem Int Ed* **2000**, *39*, 836-863.
- Dondoni A. et al., *Angew. Chem. Int. Ed.* **2008**, *47*, 8995-8997; *Angew. Chem.* **2008**, *120*, 9133-9135.
- Dondoni, A., Massi, A., Nanni, P., Roda, A., *Chem. Eur. J.* **2009**, *15*, 11444-11449.
- Dumy P. et al., *Tetrahedron Lett.*, **1995**, *36*, 1255-1258.
- Dumy P. et al., *Biopolymers*, **1996**, *39*, 297-308.
- El-Faham A., Albericio F., *Chem. Rev.* **2011**, *111*, 6557-6602.
- Feng F. et al., *Org. Biomol. Chem.* **2004**, *2*, 1617 -1623.
- Fiore M., Marra A., Dondoni A., *J. Org. Chem.* **2009**, *74*, 4422-4425.
- Gao J., Guo Z., *Med. Res. Rev.* **2017**, (*in press*)
doi: 10.1002/med.21447.
- Grandjean C. et al., *J. Org. Chem.* **2005**, *70*, 7123-7132.
- Grigalevicius S. et al., *Bioconjugate Chem.*, **2005**, *16*, 1149-1159.

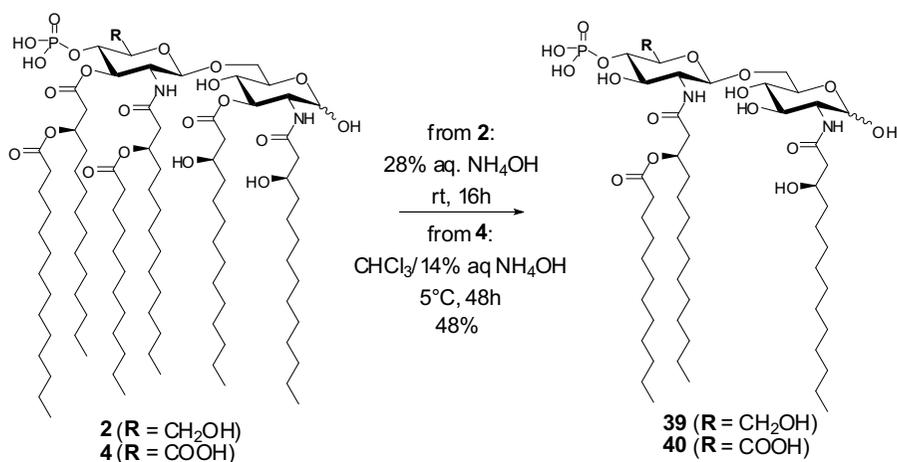
- Guo Z.W. et al., *Curr Opin Chem Biol* **2009**; *13*, 608–617.
- Hakomori S., *Curr Opin Immunol.* **1991**, *3*, 646–653.
- Hakomori S., Zhang Y., *Chem Biol.* **1997**, *3*, 97–104
- Holst O. et al., *Eur. J. Biochem.* **1993**, *214*, 695–701.
- Huang L. et al., *Chem. Eur. J.* **2006**, *12*, 5246–5252.
- Ivanova I. A. et al., *J. Chem. Soc., Perkin Trans.* **1999**, *1*, 1743–1753.
- Jagadish B. et al., *J. Med. Chem.* **2012**, *55*, 10378–10386.
- Kaiser A. et al., *Angew Chem Int Ed* **2010**, *49*, 3688–3692.
- Komaniecka I. et al., *Chem. Eur. J.* **2010**, *16*, 2922–2929.
- Komaniecka I. et al., *J. Biol. Chem.* **2014**, *289*, 35644–35655.
- Kudelka M.R. et al., *Adv. Cancer Res.* **2015**, *126*, 53–135.
- Lee J.W., Jun S. I., Kim K., *Tetrahedron Lett.* **2001**, *42*, 2709–2711.
- Liao G. et al., *ACS Cent. Sci.* **2016**, *2*, 210–218.
- Liao G., Zhou, Z., Guo Z., *Chem. Commun.* **2015**, *51*, 9647–9650.
- Marzabadi C. H., Franck, R. W., *Eur. J.* **2017**, *23*, 1728–1742.
- Montanez M.I. et al., *Biomacromolecules* **2011**, *12*, 2114–2125.
- Myers K.R., Truchot A.T., *Ribi ImmunoChem Research Inc. Hamilton, MT, USA, U.S. Patent 4912094*, **1990**.
- Musiol H.J. et al., *J. Org. Chem.* **1994**, *59*, 6144–6146.
- Pan Y.B. et al., *J Med Chem* **2005**, *48*, 875–883.
- Peluso S. et al., *ChemBioChem*, **2001**, *2*, 432–437.
- Perich J. W., Houben-Weyl Methods in Organic Chemistry, Volume E22, Synthesis of Peptides and Peptidomimetics; Goodman, M., Toniolo, C., Moroder, L., Arthur, F., Eds.; Thieme: Stuttgart, 375–424, **2004**.

- Phillips N. J. et al., *J. Biol. Chem.* **2011**, *282*, 21203–21219.
- Pieretti G. et al., *Appl. Microbiol. Biotechnol.* **2014**, *98*, 7781–7791.
- Powell M. F., Newman M. J., Vaccine design: the subunit and adjuvant approach; Eds.; Plenum Press: New York, **1995**.
- Rietschel E. T. et al., *FASEB J.* **1994**, *8*, 217–225.
- Saito Y. et al., *Tetrahedron* **2008**, *64*, 3578–3588.
- Schwudke D. et al., *J. Biol. Chem.* **2003**, *278*, 27502–27512.
- Siednienko J. et al., *Eur. J. Immunol.* **2010**, *40*, 3150–3160.
- Skehan P. et al., *J. Natl. Cancer Inst.* **1990**, *82*, 1107–1112.
- Tang S., Wang Q., Guo Z., *Chem. Eur. J.* **2010**, *16*, 1319–1325.
- Tsabedze S.B. et al., *Tetrahedron Lett.* **2013**, *54*, 6983 – 6985.
- Twibanire J.K. et al., *Org. Lett.* **2012**, *14*, 3909–3911;
- Wang Q. L., Xue J., Guo Z.W., *Chem. Commun.* **2009**, 5536–5537.
- Wang Q. et al., *ACS Chem. Biol.* **2012**, *7*, 235–240.
- Yazici A., Pyne S. G., *Org. Lett.* **2013**, *15*, 5878–5881.
- Yoshida K. et al., *Tetrahedron Lett.* **2010**, *51*, 4830– 4832.
- Yoshizaki H. et al., *Angew. Chem.* **2001**, *113*, 1523–1528.
- Ulrich S. et al., *Chem. Eur. J.*, **2014**, *20*, 34–41.
- Zhang Y. et al., *Chem. Eur. J.* **2008**, *14*, 558–569;
- Zhdankin V.V. *J. Org. Chem.* **2011**, *76*, 1185 – 1197.
- Zhou Z. et al., *Org. Biomol. Chem.* **2014**, *12*, 3238–3245.
- Zhou Z. et al., *Chem. Sci.* **2015**, *6*, 7112–7121.
- Ziacco M. et al., *J. Med. Chem.* **2017**, *60*, 9757–9768.
- Zughaiier S.M. et al., *Infect. Immun.* **2004**, *72*, 371–380.

Chapter 8: Lipid pattern modification of *E. coli* monophosphoryl lipid A

8.1 Selective deacylation reactions

To further explore the possibility of site-selective transformations on lipid A, the modification of the lipid pattern was then investigated. According to the general guideline of a lower TLR4 agonist activity associated with underacylated lipid A structures (Li et al. 2013; Rietschel et al. 1994), compound **2** was subjected to mild ester hydrolysis with 28% aqueous NH_4OH .



Scheme 8.1: Deacylation reaction on monophosphoryl lipid A **2** and **4**.

In line with previous studies conducted on analytical scale (Silipo et al. 2002), the reaction proceeded with high regioselectivity, with tri-acylated lipid A **39** being obtained as the sole product (Scheme 8.1; Figures 8.1-8.2) as a result of the lower stability of acyl and acyloxyacyl esters to hydrolysis with respect to acyl esters of acyloxyacyl amides.

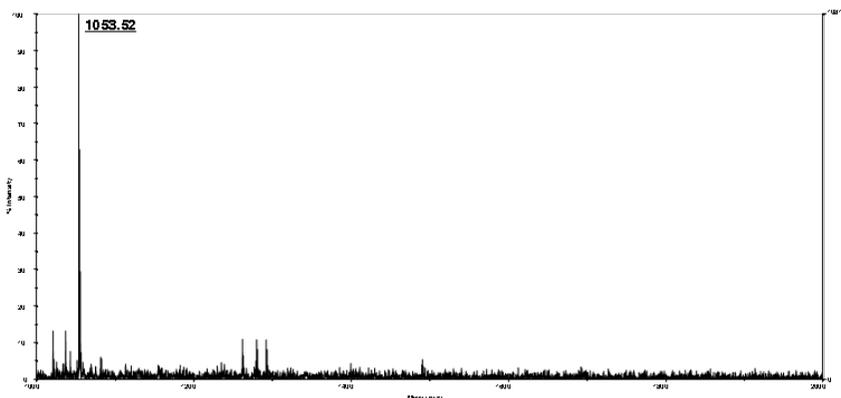


Figure 8.1: MALDI-TOF mass spectrum (negative ions mode) of **39**.

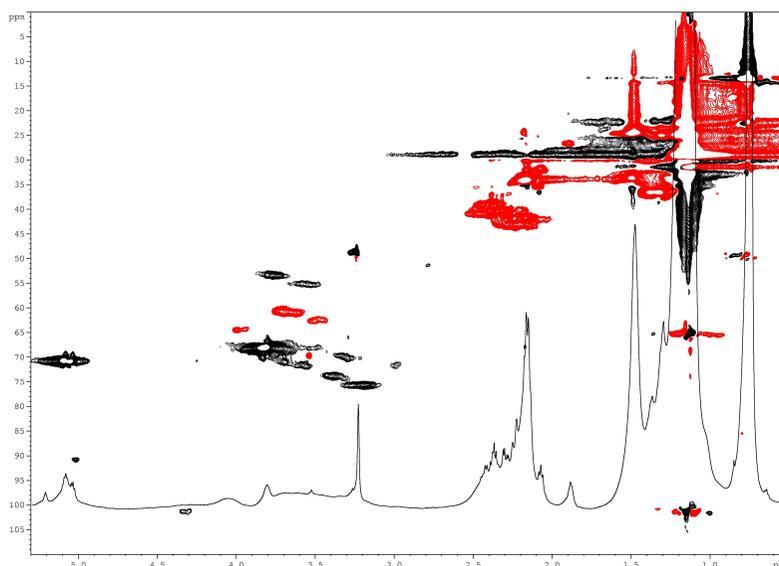


Figure 8.2: ^1H and HSQC-DEPT NMR spectra (600 MHz, 298K, 4:1 v/v CDCl_3 - CD_3OD) of **39**.

Furthermore, this reaction was also performed on oxidized lipid **A 4** under slightly modified conditions ($\text{CHCl}_3/14\%$ aq. NH_4OH , 1:1 v/v, at 5°C), again affording the tri-acylated lipid A derivative **40** as the sole product, as detected by MALDI-MS and NMR spectroscopic analysis (Figure 8.3-8.4 and Table 8.1).

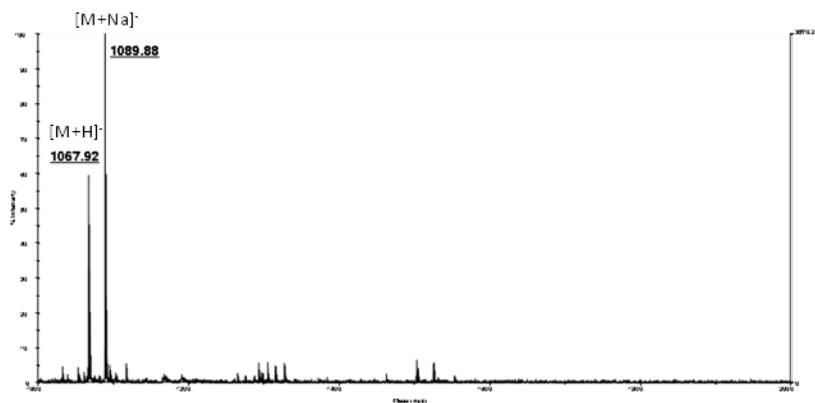


Figure 8.3: MALDI-TOF mass spectrum (negative ions mode) of **40**.

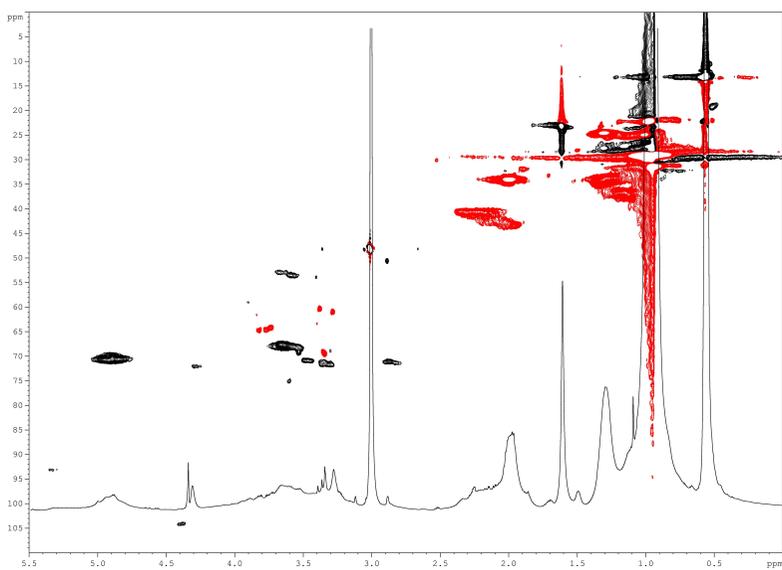


Figure 8.4: ^1H and HSQC-DEPT NMR spectra (600 MHz, 298K, 4:1 v/v CDCl_3 - CD_3OD) of **40**.

Table 8.1: NMR data referred to lipid A derivatives **39** and **40**.

	Sugar portion ^a							Lipid portion ^a					Phosphate ^b	
	1	2	3	4	5	6	α	β	γ	δ -(ω -1)	ω			
39	GlcN I (α)	5.01 <i>90.8</i>	3.75 <i>53.5</i>	3.29 ^c <i>70.1</i>	3.57 <i>71.7</i>	3.32 ^c <i>69.7</i>	3.94, 3.99 <i>64.4</i>	C14:0(3-OH) [N-2 GlcN I]	2.14, 2.22 <i>41.8</i>	3.82 <i>67.9</i>	1.34 <i>36.5</i>	1.07-1.18 <i>29.0</i>	0.71-0.74 <i>13.5</i>	-0.4
	GlcN II	4.32 <i>101.4</i>	3.55 <i>55.3</i>	3.39 <i>73.9</i>	3.69 <i>71.0</i>	3.18 <i>75.7</i>	---	C14:0(3-OR) [N-2 GlcN II]	2.36, 2.42 <i>39.1</i>	5.09 <i>70.4</i>	1.47 <i>33.6</i>			
								C12:0	2.18 <i>33.9</i>	1.49 <i>24.4</i>	1.07-1.18 <i>29.0</i>			
40^d	GlcN I (α)	5.32 <i>93.1</i>	3.67 <i>52.6</i>	3.31 ^c <i>71.7</i>	3.46 <i>70.8</i>	3.36 ^c <i>71.8</i>	3.77, 3.83 <i>64.5</i>	C14:0(3-OH) [N-2 GlcN I]	1.97 <i>43.3</i>	3.66 <i>67.6</i>	1.15 <i>36.8</i>	0.93-0.98 <i>29.1</i>	0.53-0.58 <i>13.2</i>	3.3
	GlcN II	4.38 <i>104.2</i>	3.59 <i>53.5</i>	3.46 <i>70.8</i>	3.54 <i>68.3</i>	2.88 <i>71.0</i>	---	C14:0(3-OR) [N-2 GlcN II]	2.22 <i>40.7</i>	4.90 <i>70.2</i>	1.31 <i>33.9</i>			
								C12:0	2.00 <i>34.0</i>	1.31 <i>24.5</i>	0.93-0.98 <i>29.1</i>			

a ¹H (600 MHz, 298K) and ¹³C (125 MHz, 298K) chemical shifts (in plain and in italic, respectively) expressed in δ relative to residual CHCl₃ (¹H: δ 7.26 ppm; ¹³C: δ 77.0 ppm) in 4:1 v/v CDCl₃-CD₃OD, except where differently indicated.

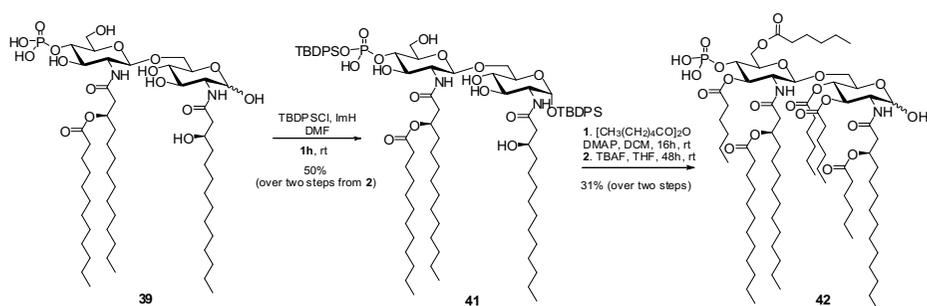
b ³¹P (160 MHz, 298K) chemical shifts expressed in δ relative to 85% phosphoric acid (external standard).

c Interchangeable assignments.

d Chemical shift values measured in 1:1 v/v CDCl₃-CD₃OD.

8.2 Selective re-acylation with shorter lipid chains

Lipid A derivatives with some shortened acyl chains have previously been obtained by total synthesis and shown to possess interesting immunostimulating activities (Jiang et al. 2007; Johnson et al. 1999; Fukase et al. 1998). Along these lines, in collaboration with Dr. D'Alonzo (Department of Chemical Sciences, University of Naples Federico II), during this PhD work, we finely tuned the conditions suitable for the semisynthesis of some new lipid A derivatives from the tri-acylated lipid A **39**. Given that the direct acetylation or succinylation of a tetra-acylated lipid A very similar to **39** afforded exclusively per-*O*-acylated products (Tanamoto 1994), a three-step approach based on a regioselective protection-acylation-deprotection sequence was considered. The most reactive functionalities of **39** could be identified as the hydroxyl groups at the anomeric position of GlcN I and at position 6 of GlcN II as well as the phosphate group. By reacting **39** under common silylation conditions (TBDPSCl/ImH), a double protection product **41** was obtained in 50% yield after 1h (Scheme 8.2; Figures 8.5-8.6).



Scheme 8.2: Reactions sequence on the Tri-acylated monophosphoryl lipid A **39**.

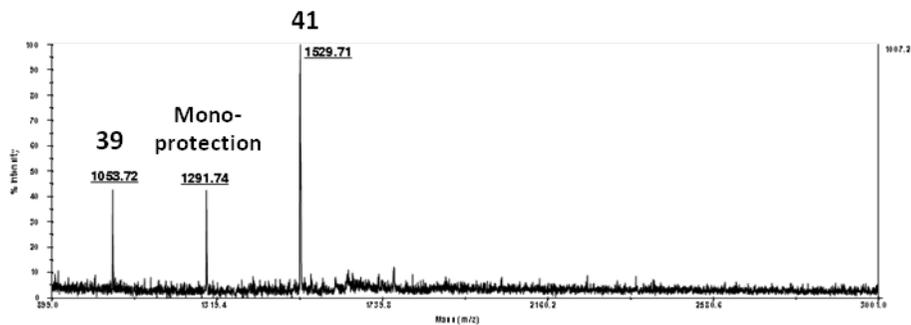


Figure 8.5: MALDI-TOF mass spectrum (negative ions mode) of crude **41**.

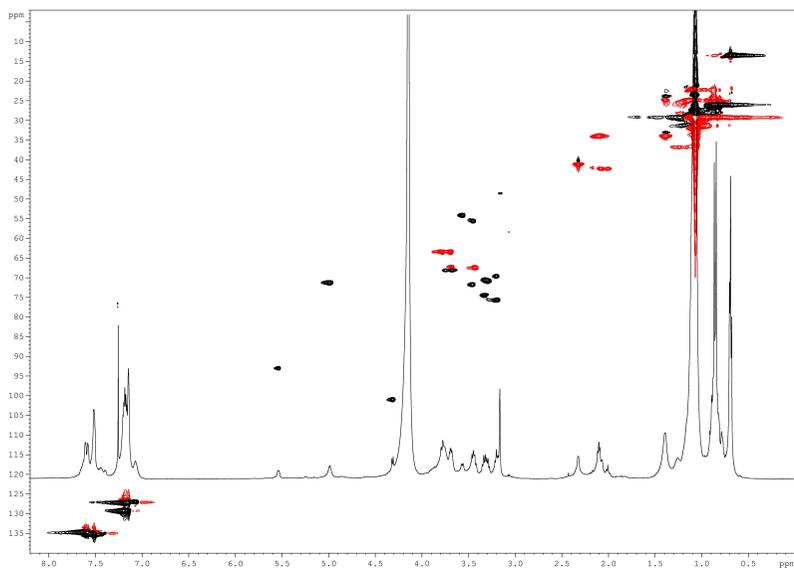


Figure 8.6: ^1H and HSQC-DEPT NMR spectra (600 MHz, 298K, 4:1 v/v CDCl_3 - CD_3OD) of pure **41**.

Furthermore, the ROESY spectrum clearly indicated through-space correlations between one TBDPS aromatic signal ($\delta=7.60$ ppm) and GlcN I anomeric ($\delta=5.57$ ppm) and one of the two H-6 signals ($\delta = 3.45$ ppm), as well as correlations between the second TBDPS aromatic signal ($\delta=7.52$ ppm) and GlcN II H-4, H-5 and both H-6 signals ($\delta=3.32, 3.23, 3.72, \text{ and } 3.82$ ppm, respectively) (Figure 8.7-a).

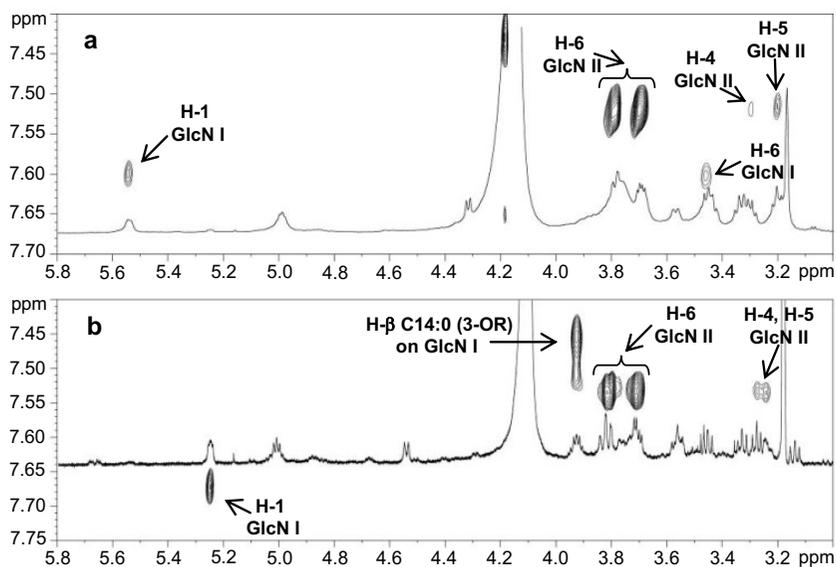
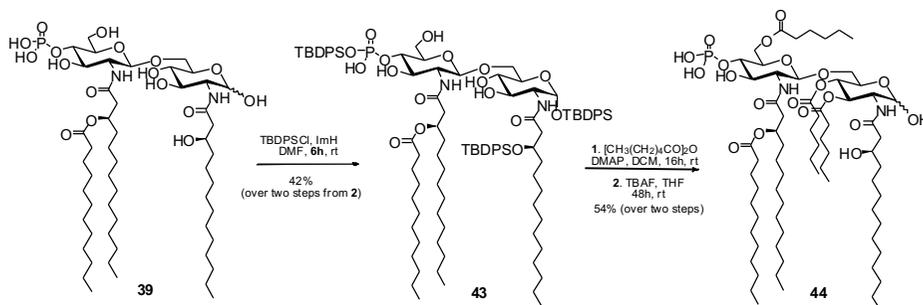


Figure 8.7: Zoom of ^1H and ROESY NMR spectra (600 MHz, 298K, 4:1 v/v $\text{CDCl}_3\text{-CD}_3\text{OD}$) for **41** (a) and **42** (b).

This strongly suggested that one TBDPS was placed at the anomeric position (singularly, with an axial orientation) of GlcN I and the second group was either at position *O*-6 of GlcN II or on the phosphate group. Conversely, at this stage, we could not clearly assign the position of the latter. Interestingly, we also observed that when the silylation reaction was carried out for longer times (6 h), a tri-silylated compound **43** was isolated as the major product (42% yield) (Scheme 8.3; Figures 8.8-8.9).



Scheme 8.3: Second reactions sequence on the tri-acylated monophosphoryl lipid A **39**

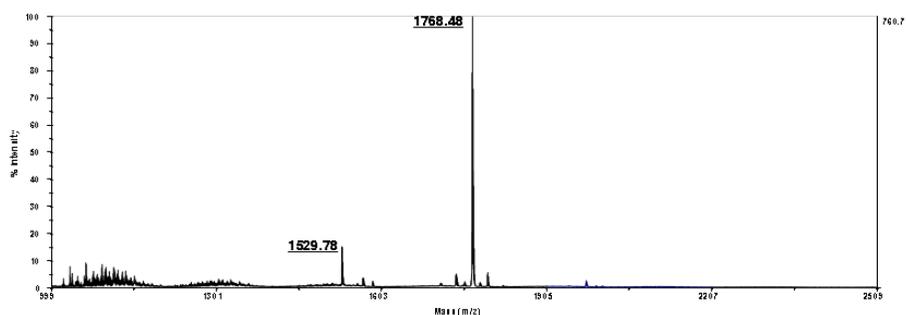


Figure 8.8: MALDI-TOF mass spectrum (negative ions mode) of **43**.

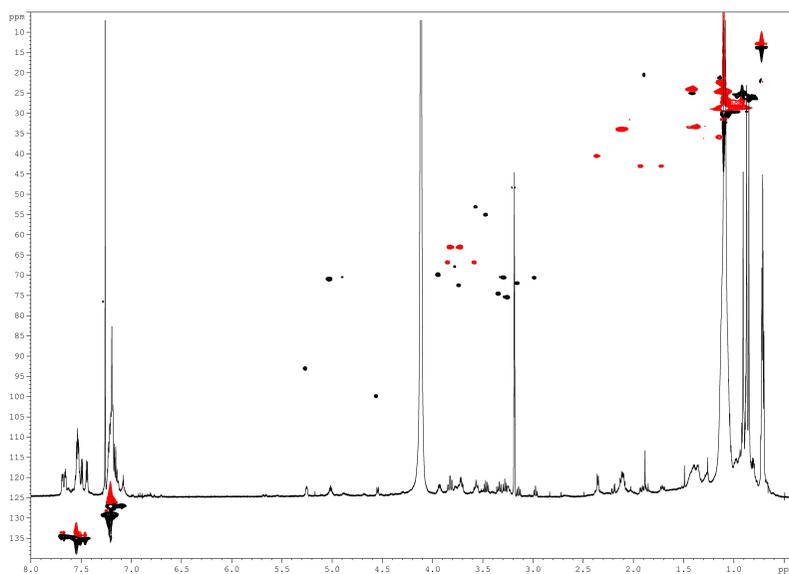


Figure 8.9: ^1H and HSQC-DEPT NMR spectra (600 MHz, 298K, 4:1 v/v CDCl_3 - CD_3OD) of **43**.

In this case, ROESY experiments showed correlations similar to those found for the di-silylated derivative, and an additional interaction between a TBDPS aromatic signal ($\delta=7.46\text{ppm}$) and H-b of C14:0 (β -OR) chain on GlcN I ($\delta=3.96\text{ ppm}$) (Figure 8.7-b), thus suggesting that installation of the third TBDPS group occurred at the secondary hydroxyl at the β -position of the acyl chain on GlcN I. Both silylated derivatives were employed in preliminary acylation studies (Scheme 8.2-8.3). Di-silylated lipid A **41** was treated with freshly prepared hexanoic

anhydride and *N,N*-dimethylaminopyridine (DMAP) in dichloromethane. After 16h at room temperature, a complex mixture (NMR) was obtained, although MALDI-MS analysis indicated the presence of one major reaction product having five hexanoyl chains and only one TBDPS group (Scheme 8.2). De-*O*-silylation of the crude mixture under standard conditions [tetrabutylammonium fluoride (TBAF)] yielded, after chromatographic purification, lipid A derivative **42** as the major species in 31% yield over two steps (Figures 8.10-8.11).

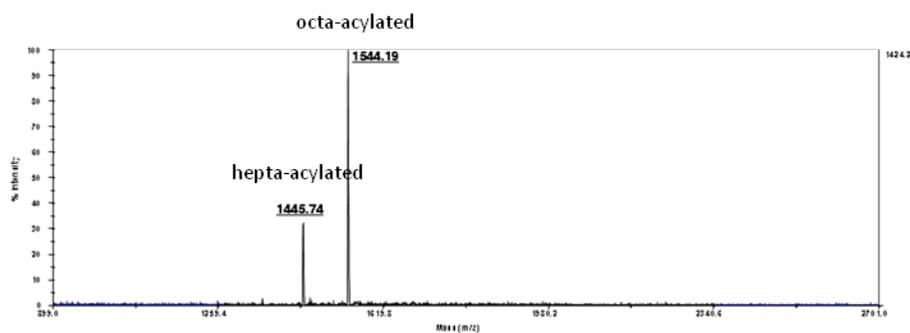


Figure 8.10: MALDI-TOF mass spectrum (negative ions mode) of pure **42**.

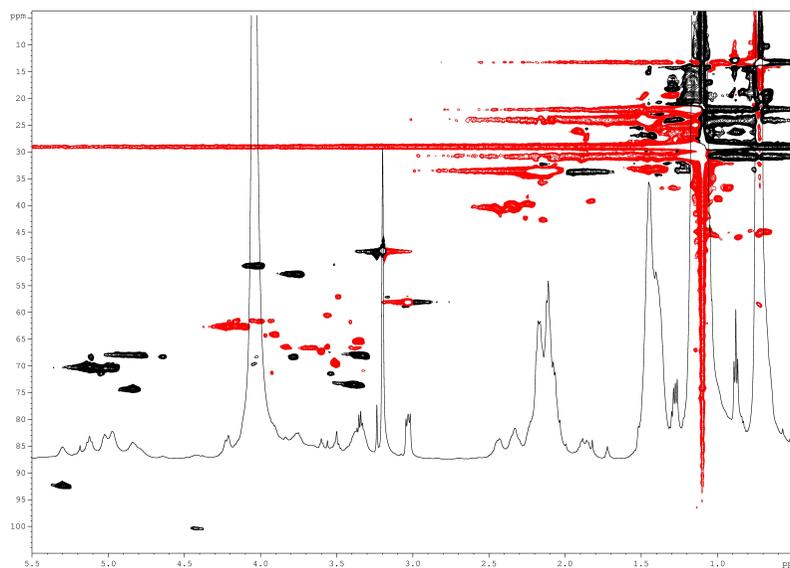


Figure 8.11: ¹H and HSQC-DEPT NMR spectra (600 MHz, 298K, 4:1 v/v CDCl₃-CD₃OD) of **42**.

The latter was characterized by the presence of eight acyl chains; 2D-NMR studies clearly established that the five shorter ones were placed at the *C*-3 and *C*-4 positions of GlcN I, the *C*-3 position of the *N*-acyl chain of GlcN I and the *C*-3 and *C*-6 positions of GlcN II (Table 8.2). Likewise, this suggested a structure with TBDPS group at position *O*-1 of GlcN I and on phosphate group for di-silylated synthetic intermediate **41**. Interestingly, the lack of TBDPS protection on the hydroxyl at position 6 of GlcN II was in agreement with the sluggish reactivity of this alcohol as already highlighted above. Acylation of tri-silylated lipid A **43** derivative was eventually studied (Scheme 8.3). Its treatment with hexanoic anhydride and DMAP again provided a complex mixture (MALDI-MS, NMR). Therefore, the crude reaction mixture was directly subjected to de-*O*-silylation, as in the previous case, affording hexa-acylated lipid A derivative **44** as the major species (54% yield over two steps) (Figures 8.12-8.13; Table 8.2), with three hexanoyl chains at positions *C*-3 and *C*-4 of GlcN I and *C*-6 of GlcN II, as detected by 2D-NMR spectroscopic analysis.

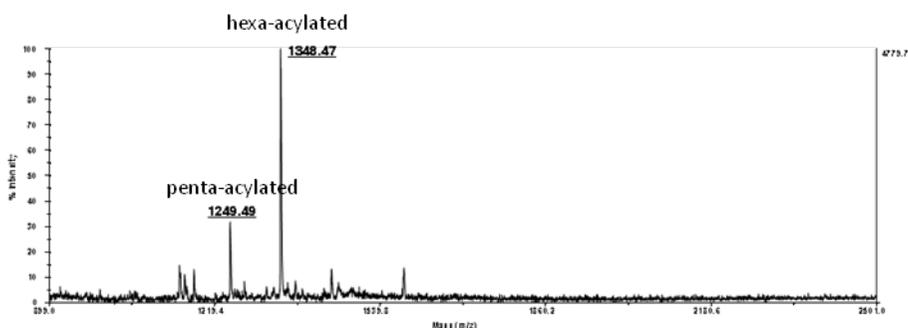


Figure 8.12: MALDI-TOF mass spectrum (negative ions mode) of **44**.

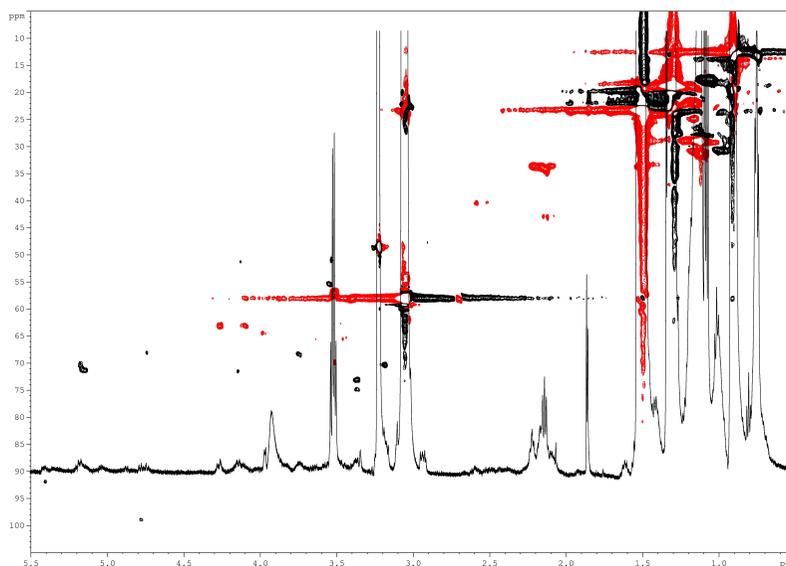


Figure 8.13: ^1H and HSQC-DEPT NMR spectra (600 MHz, 298K, 4:1 v/v CDCl_3 - CD_3OD) of **44**.

The presence of the acyl chain on *O*-6 of GlcN II was due again to the absence of TBDPS protection at this position on tri-silylated synthetic intermediate **43**. The lack of an ester group at position *C*-3 of GlcN II of **44** could not be easily explained, although a steric hindrance of the TBDPS group at the β -position of the acyl chain on GlcN I could be invoked.

Table 8.2: NMR data referred to lipid A derivatives 41-44.

	Sugar portion ^a							Lipid portion ^a						Other signals ^a	Phosphate ^b
	1	2	3	4	5	6	α	β	γ	δ -(ω -1)	ω				
41	GlcN I (α)	5.57 <i>92.9</i>	3.61 <i>53.8</i>	3.35 <i>70.6</i>	3.49 <i>71.7</i>	3.26 <i>69.5</i>	3.45, 3.71 <i>67.3</i>	C14:0(3-OH) [N-2 GlcN I]	2.05, 2.10 <i>42.2</i>	3.70 <i>68.0</i>	1.30 <i>36.8</i>	1.07-1.18 <i>29.0</i>	0.71-0.74 <i>13.5</i>	TBDPS CH ₃ : 0.85/ <i>26.1</i> 0.87/ <i>26.1</i> Ph: 7.15/ <i>129.3</i> 7.18/ <i>127.1</i> 7.20/ <i>129.3</i> 7.52/ <i>135.1</i> 7.60/ <i>134.8</i>	-12.7
	GlcN II	4.34 <i>100.6</i>	3.48 <i>55.2</i>	3.37 <i>74.5</i>	.32 <i>70.5</i>	.23 <i>75.6</i>	3.72, 3.82 <i>63.4</i>	C14:0(3-OR) [N-2 GlcN II]	2.34, 2.39 <i>41.2</i>	5.02 <i>71.2</i>	1.41 <i>33.9</i>				
								C12:0	2.13 <i>34.0</i>	1.39 <i>24.7</i>	1.07-1.18 <i>29.0</i>				
42	GlcN I (α)	5.30 <i>92.2</i>	4.00 <i>51.1</i>	5.12 <i>70.1</i>	4.80 <i>67.4</i>	3.35 <i>67.9</i>	3.58, 3.71 <i>66.6</i>	C14:0(3-OR) [N-2 GlcN I]	2.25, 2.35 <i>39.6</i>	5.02 <i>70.3</i>	1.41 <i>33.1</i>	1.08-1.18 <i>29.1</i>	0.71-0.75 <i>13.3</i>	---	0.7
	GlcN II	4.41 <i>100.5</i>	3.75 <i>52.9</i>	4.84 <i>74.3</i>	4.03 <i>68.7</i>	3.39 <i>73.5</i>	4.11, 4.22 <i>62.6</i>	C14:0(3-OR) [N-2 GlcN II]	2.33, 2.45 <i>40.5</i>						
								C6:0, C12:0	2.09, 2.19 <i>33.6</i>	1.38-1.46 <i>24.1</i>	1.08-1.18 <i>29.1</i>				
43	GlcN I (α)	5.27 <i>93.0</i>	3.56 <i>53.1</i>	3.14 <i>71.8</i>	2.99 <i>70.7</i>	3.73 <i>72.6</i>	3.59, 3.85 <i>66.8</i>	C14:0(3-OTBDPS) [N-2 GlcN I]	1.70, 1.92 <i>43.1</i>	3.96 <i>69.8</i>	1.12 <i>36.0</i>	1.07-1.18 <i>28.9</i>	0.71-0.74 <i>13.4</i>	TBDPS CH ₃ : 0.86/ <i>26.0</i> 0.88/ <i>25.8</i> 0.91/ <i>25.1</i> Ph: 7.15-7.23/ <i>127.0</i> 7.21-7.24/ <i>129.0</i> 7.46/ <i>135.2</i> 7.51/ <i>135.3</i> 7.55/ <i>134.9</i> 7.69/ <i>134.4</i>	-11.8
	GlcN II	4.56 <i>99.9</i>	3.47 <i>55.1</i>	3.36 <i>74.6</i>	3.30 <i>70.5</i>	3.26 <i>75.5</i>	3.73, 3.82 <i>63.1</i>	C14:0(3-OR) [N-2 GlcN II]	2.33, 2.38 <i>40.6</i>	5.04 <i>70.8</i>	1.37 <i>33.4</i>	1.07-1.18 <i>28.9</i>			
								C12:0	2.12 <i>34.0</i>	1.40 <i>24.2</i>	1.07-1.18 <i>28.9</i>				
44	GlcN I (α)	5.41 <i>91.8</i>	4.13 <i>51.1</i>	5.18 <i>70.3</i>	4.74 <i>68.1</i>	4.14 <i>71.6</i>	3.46, 3.64 <i>65.5</i>	C14:0(3-OH) [N-2 GlcN I]	2.08, 2.14 <i>42.9</i>	3.75 <i>68.2</i>	1.27 <i>n.d.</i>	1.08-1.18 <i>29.1</i>	0.72-0.77 <i>13.3</i>	---	-1.5
	GlcN II	4.78 <i>98.8</i>	3.55 <i>55.5</i>	3.37 <i>73.0</i>	3.18 <i>70.2</i>	3.37 <i>73.8</i>	4.11, 4.27 <i>62.9</i>	C14:0(3-OR) [N-2 GlcN II]	2.52, 2.58 <i>40.3</i>	5.15 <i>71.3</i>	1.48 <i>33.3</i>				
								C6:0, C12:0	2.09, 2.21 <i>33.6</i>	1.42-1.47 <i>24.1</i>	1.08-1.18 <i>29.1</i>				

^a ¹H (600 MHz, 298K) and ¹³C (125 MHz, 298K) chemical shifts (in plain and in italic, respectively) expressed in δ relative to residual CHCl₃ (H: δ 7.26 ppm; ¹³C: δ 77.0 ppm) in 4:1 v/v CDCl₃-CD₃OD.

^b ³¹P (160 MHz, 298K) chemical shifts expressed in δ relative to 85% phosphoric acid (external standard).

8.3 Biological assays

Preliminary immunological assays were performed on lipid A derivatives **40**, **42** and **44**, in collaboration with Prof. Giuliano at the University of Campania “L. Vanvitelli”, with the aim of studying whether these molecules be easily detected by human monocytic cell line (THP-1) and could trigger cellular responses in terms of cytokine amounts and length of stimulation (Figure 8.14). The amounts of different cytokines were examined in tissue culture supernatants of THP-1, and detected by using Enzyme-Linked Immunosorbent Assay (ELISA). *E. coli* O111-B4 LPS and MPL[®] from *S. Minnesota* R595 (**3**) were used as positive controls for an agonist and immunoadjuvant response profile, respectively. However, as the former showed an inflammatory potential exceedingly greater than that of the latter and of the synthesized compounds (e.g., IL-6, up to >4000 pg mL⁻¹; IL-8, up to >10000 pg mL⁻¹; TNF- α , up to >10000 pg mL⁻¹) data in Figure 8.14 are shown related to **3** only. The semisynthetic lipid A derivatives showed cytokine profiles comparable to those of the immunoadjuvating drug **3**, but with interesting peculiarities. Antimicrobial substances and proinflammatory cytokines such as TNF- α , IL-6 and IL-8 create a local proinflammatory environment, which helps to recruit and activate phagocytic cells, activate the complement cascade, contain the invading pathogen and chemo-attract the effector cells of the adaptive immune response. However, overstimulation can also result in septic proinflammatory responses such as secretion of TNF- α and IL-8, which, in severe cases, can be detrimental to the host. In this regard, all the tested derivatives further reduced IL-8 production, which is a key mediator associated with inflammation, driving towards a less proinflammatory profile than the one induced by **3**.

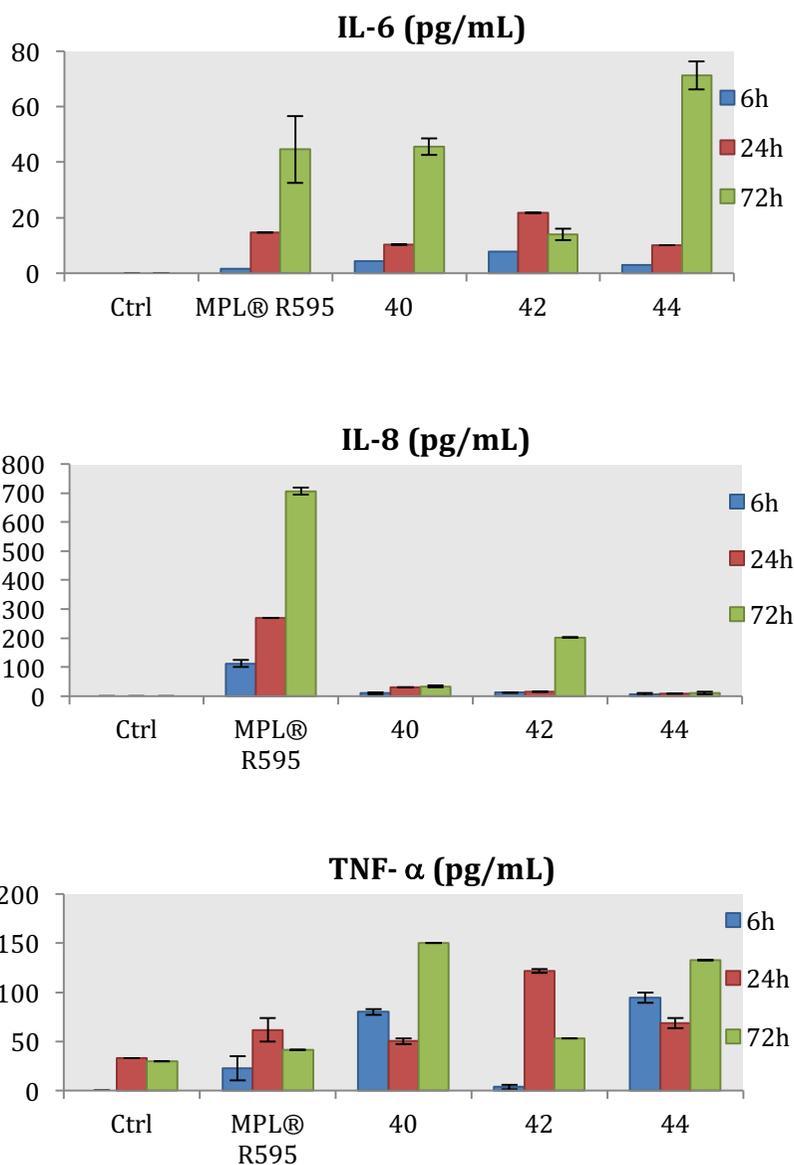


Figure 8.14: Cytokine concentrations (Mabtech's ELISA^{PRO}) in tissue-culture supernatants of THP-1 cells stimulated with lipid A derivatives (20 $\mu\text{g/mL}$). Ctrl represents the cells not treated.

8.4 Conclusions

Investigations into the peculiar reactivity of lipid A, leading to novel derivatives with immunostimulating activity, has been carried out through site-selective chemical reactions on monophosphoryl lipid A **2** and its oxidized **4**. In particular, we finely tuned the conditions suitable for the semisynthesis of some new lipid A derivatives, by selective modification of the lipid pattern (selective deacylation and re-acylation with shorter chains). In this way, a small collection of new lipid A derivatives was rapidly prepared, avoiding lengthy total synthetic approaches. Notably, potential access to several other modified lipid A structures could be opened by this strategy, because the reactions developed on *E. coli* lipid A scaffold should also be applicable on lipid A substrates from different bacterial sources. Some of the new compounds were assayed for preliminary in vitro immunological tests, showing somewhat reduced inflammatory activity and a cytokine production profile similar to the already used immunomodulant lipid A derivative **3**, thus highlighting the promising immunobiological applications.

8.5 Experimental section

8.5.1 Lipid A derivative (**40**)

A suspension of **4** (33.5 mg, 20.5 μmol) in 1:1 v/v CHCl_3 -14% aqueous NH_4OH (4.7 mL) was stirred at 5°C. After 48 h stirring at 5°C, the reaction mixture was concentrated. The residue was subjected to silica-gel column chromatography ($\text{CHCl}_3/\text{MeOH}$, 97:3 to 3:2 v/v) to give **40** (10.5 mg, 48%) as a white powder.

8.5.2 Lipid A derivative (41)

A suspension of compound **2** (131 mg, 76.3 μmol) in 28% aqueous NH_4OH (3.0 mL) was stirred at RT overnight. The reaction was worked up by concentration to give crude **39** (130 mg) as a white powder. The latter was co-evaporated with anhydrous DMF (3 x 5 mL). The resulting oil was dissolved in further anhydrous DMF (5 mL), the mixture was cooled to 0°C and treated with TBDPSCl (339 μL , 1.2 mmol) and imidazole (84 mg, 1.2 mmol). The solution was then warmed to RT and stirred for 1 h. Afterwards, the reaction mixture was diluted with chloroform (50 mL) and washed with brine. The organic phase was dried over anhydrous Na_2SO_4 , filtered and concentrated. Silica-gel chromatography of the crude residue ($\text{CHCl}_3/\text{MeOH}$, 100:0 to 95:5 v/v) gave pure **41** (58 mg, 50% overall yield from **2**) as a colorless oil.

8.5.3 Hexanoic anhydride

DCC (89 mg, 430 μmol) was added to a solution of hexanoic acid (100 mg, 860 μmol) in anhydrous CH_2Cl_2 (5 mL). The resulting suspension was stirred at RT for 16 h; then it was filtered, washing the solid with cold hexane (50 mL). The filtrate was concentrated under reduced pressure to give hexanoic anhydride (92 mg, 99% yield) as a colorless oil. ^1H and ^{13}C NMR data of the latter were fully in line with those reported elsewhere (Park et al. 2005).

8.5.4 Lipid A derivative (42)

Hexanoic anhydride (92 mg, 429 μmol) and DMAP (3.5 mg, 2.9 μmol) were added to a stirring solution of silylated lipid A derivative **41** (45 mg, 29 μmol) in anhydrous CH_2Cl_2 (3 mL). The mixture was stirred at RT for 16 h; then it was transferred to a separatory funnel, extracted with

CHCl₃ (3 x 25 mL) and washed with brine. The combined organic layers were dried over anhydrous Na₂SO₄, filtered and concentrated. To a solution of the crude acylated product (25 mg) was added a solution of TBAF (11 μL, 1M in THF) and the mixture was stirred at RT for 48 h. The solvent was then removed under reduced pressure. Silica-gel chromatography of the crude residue over silica gel (CHCl₃/MeOH, 100:0 to 85:15 v/v) provided lipid A derivative **42** (12 mg, 31% overall yield over two steps) as a colorless oil.

8.5.5 Lipid A derivative (**43**)

TBDPSCl (260 μL, 1 mmol) and imidazole (68 mg, 1 mmol) were added to a cooled (0°C) solution of crude **39** (100 mg, obtained from 130 mg of **2**) in anhydrous DMF (5 mL). The reaction mixture was warmed to RT and stirred at the same temperature for 6 h. The solution was then diluted with chloroform (50 mL) and washed with brine. The organic phase was dried over anhydrous Na₂SO₄, filtered, and concentrated. Silica-gel chromatography of the crude residue (CHCl₃/MeOH, 100:0 to 95:5 v/v) gave **43** (56.2 mg, 42% overall yield from **2**) as a colorless oil.

8.5.6 Lipid A derivative (**44**)

Under the same conditions reported for the synthesis of **42**, lipid A derivative **44** (12 mg, 9 μmol) was obtained from TBDPS ether **43** (29 mg, 17 μmol) after acylation of the latter (hexanoic anhydride, 14 mg, 66 μmol; DMAP, 0.2 mg, 2 μmol) and subsequent de-*O*-silylation (TBAF, 17 μL, 1M in THF) of the resulting acylated product (52% overall yield over two steps).

8.5.7 Immunological assays

Immunological activities were evaluated by stimulating the human monocytic leukemia cell line THP-1 (ATCC TIB-202) (Pieretti et al. 2014). Briefly, monocytic THP-1 cells (LGC, Promochem, Milan, Italy) were grown in suspension in RPMI 1640 medium supplemented with 10% fetal bovine serum at 37°C in a humidified incubator with a 5% CO₂ atmosphere and stimulated with the lipid A derivatives at 20 µg mL⁻¹. LPS from *E. coli* O111:B4 (Sigma, Milan, Italy) and MPLA from *S. Minnesota* R595 (Avanti Polar Lipids, Spectra 2000, Milan, Italy) were used as positive controls. Mabtech's ELISA^{PRO} kits and Enzo Life Science EIA Kit (BioRad, Milan, Italy) were used to quantify IL-6, IL-8 and TNF-α production in tissue-culture supernatants, following manufacturer's instructions.

References

- Fukase K. et al., *Tetrahedron*, **1998**, *54*, 4033 -4050.
- Li Y. et al., *Mar. Drugs* **2013**, *11*, 3197 -3208.
- Jiang Z.H. et al., *Carbohydr. Res.* **2007**, *342*, 784 -796.
- Johnson D.A. et al., *J. Med. Chem.* **1999**, *42*, 4640 - 4649.
- Park Y.D. et al., *Synth. Commun.* **2005**, *35*, 371 -378.
- Pieretti G. et al., *Appl. Microbiol. Biotechnol.* **2014**, *98*, 7781 - 7791.
- Rietschel E.T. et al., *FASEB J.* **1994**, *8*, 217- 225.
- Silipo A. et al., *J. Lipid Res.* **2002**, *43*, 2188 -2195.
- Tanamoto K. *FEBS Lett.* **1994**, *351*, 325- 329.

Chapter 9: Fluorescent derivatives of *E. coli* lipid A

The last chapter of this thesis is dedicated to an ongoing study that we have in collaboration with Prof. Robert Tampé from the Institute of Biochemistry, at the Goethe-University in Frankfurt. Tampé and co-workers have performed several studies during these last years in the field of membrane and trans-membrane proteins. Among the plethora of factors that govern the stability and function of membrane proteins, the local lipid environment is emerging as one of the most important. The topological organization of protein subunits, as well as their stability and activity have all been linked directly to the effects of neighbouring lipids (Bao et al., 2013; Dowhan et al., 2011; Betaneli et al., 2012; Obara et al., 2005; Leganowsky et al., 2014). The first shell of lipid molecules, in direct contact with the membrane protein, is often referred to as the lipid annulus (Lee et al., 2011). Spin-labelling and fluorescence-quenching studies have demonstrated that these annular lipids form only weak, non-specific interactions with membrane proteins (Lee et al., 2003). Annular lipids are therefore often lost during solubilization and/or purification of membrane proteins and are usually not observed in X-ray structures. Lipids identified through crystallography in detergent micelles are bound specifically to the protein and therefore represent a special case of annular lipids. To distinguish lipids that induce a direct effect on membrane proteins from loosely associated ones remains a challenging task and raises questions about whether or not the lipid annulus is constant in composition or variable.

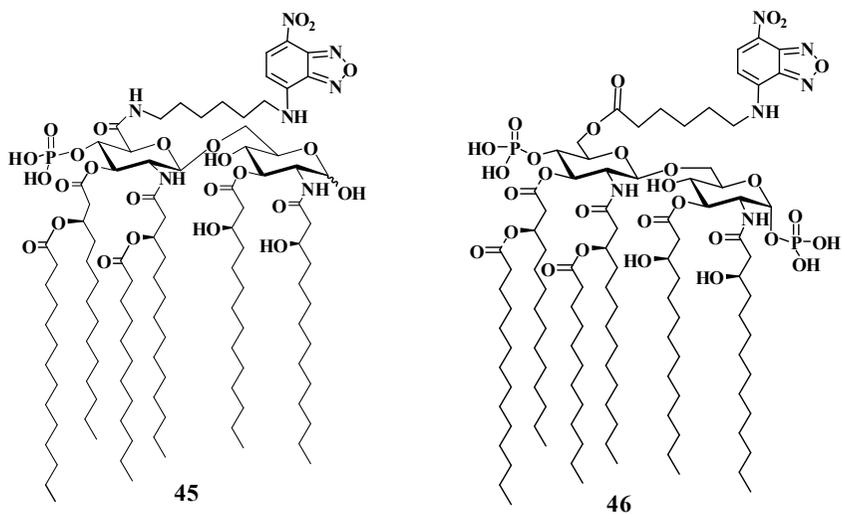
9.1 Flippases

The ATP-binding cassette (ABC) exporters reside in the membrane lipid bilayer of all cells and are responsible for multidrug resistance phenotypes, extruding antibiotics or chemotherapeutics from target cells (Schweizer et al., 2012; Huang et al., 2006; Gottesman et al. 2002). This renders them a prime target for clinical research. It has been established that some ABC transporters (as P-glycoprotein (MDR1, ABCB1), TmrAB, LmrA and MsbA) act as lipid pumps or ‘flippases’, either effluxing or transporting lipids from the inner to the outer leaflet, respectively (Borst et al., 2000; Reuter et al., 2003; Hendrich et al., 2003; King et al., 2012). In particular, MsbA was found in Gram-negative bacteria and is an ABC transporter involved in the lipid A and LPS transport from the cytoplasmic leaflet (inward-facing) to the periplasmic leaflet (outward-facing) of the plasma membrane (Karow et al. 1993; Polissi et al. 1996; Wang et al. 2004). Almost half of the 48 human ABC transporters have been linked directly or indirectly with lipid transport (Tarling et al., 2013). Moreover, lipids have been shown to bind directly to some ABC transporters or to have an effect on the activity of others (Bao et al., 2013; Shintre et al., 2013; Doerrler et al., 2002; Marek et al., 2011; Eggensperger et al., 2014). Tampé and co-workers have recently identified and monitored the composition of part of an endogenous annular lipid belt bound to a heterodimeric membrane protein complex (Bechara et al. 2015). Twenty-four different phospholipid species were detected together with various tetra- and hepta-acylated lipid A species in four independent protein preparations. The majority of phospholipids in close contact with the protein had similar acyl chain lengths (between 16 and 18 carbon atoms). Classifying lipids and their potential role in

contributing to protein stability, or in mediating drug export, is therefore of paramount interest for drug discovery and for understanding the mechanisms of ABC transporters.

9.2 Fluorescent mono/diphosphoryl lipid As and ABC transporters

A further step for the ABC transporters mechanisms comprehension was the synthesis of fluorescent-lipid As in order to develop some *in vitro* fluorescence-based lipid A flippase assays. Towards this direction, in this thesis two different kind of nitro-benzofurazan (NBD)-labeled lipid A derivatives were synthesized. The soundness of the preliminary data and a pending patent, forces us to describe this topic only as an overview without furnishing some experimental data and results in full details. However, the first derivative that was synthesized, according to the strategy showed in the previous chapters, was obtained from the mono-phosphoryl lipid A derivative oxidized on the position C-6 of the GlcNII (**4**, see paragraph 7.1). Indeed, after having synthesized an amine-appended NBD derivative from commercially available 4-Chloro-7-nitrobenzofurazan (Yamaguchi et al. 2014), the optimized coupling conditions (described in paragraph 7.2) gave the first semi-synthetic fluorescent mono-phosphoryl lipid A derivative (NBD-MPLA) **45**. MALDI-TOF mass spectrum and full NMR characterization of the major hexa-acylated species confirmed its structure (Scheme 9.1; Figures 9.1-9.2; Table 9.1).



Scheme 9.1: Semi-synthesis of fluorescent mono- and di-phosphoryl lipid As **45** and **46**.

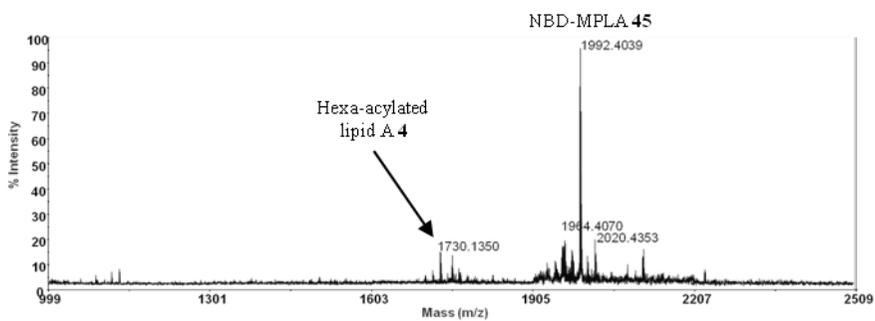


Figure 9.1: MALDI-TOF mass spectrum (negative ions mode) of **45**.

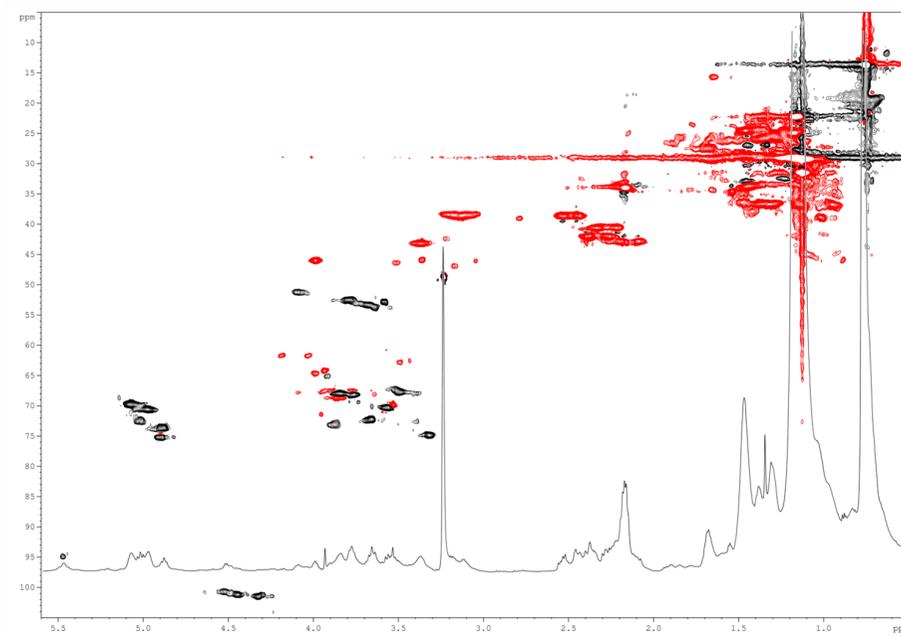


Figure 9.2: ^1H and HSQC-DEPT NMR spectra (600 MHz, 298K, 4:1 v/v CDCl_3 - CD_3OD) of **45**.

It is worth noting that, similarly to lipid A derivatives described in chapter 7, the strategy for the coupling was focused on the possibility to tag the lipid A, with a fluorescent moiety without any derivatization on some moieties (phosphate group, anomeric position or acyl pattern) which could be important for the interactions with the proteins. Indeed, also in this case, the coupling through the C-6 of GlcNII should be highly bio-mimetic.

In collaboration with Prof. Tampé, the fluorescent monophosphoryl lipid A **45** was exposed to several biological assays. They first tested whether our NDB-MPLA **45** (0.1, 1.0 and 1.5 mol%) could affect the transport activity of TmrAB. The hypothesis was that lipid A and other substrates compete for the same translocation pathway. Derivative **45** was offered in 1, 10, 100 nM and 1 μM to TmrAB 1 μM in order to understand if it was associated or loaded onto isolated detergent-

solubilized TmrAB. In these studies we were excited to see that NBD-MPLA **45** co-migrates with TmrAB (200 kDa) in a fluorescence-based SEC (Figure 9.3).

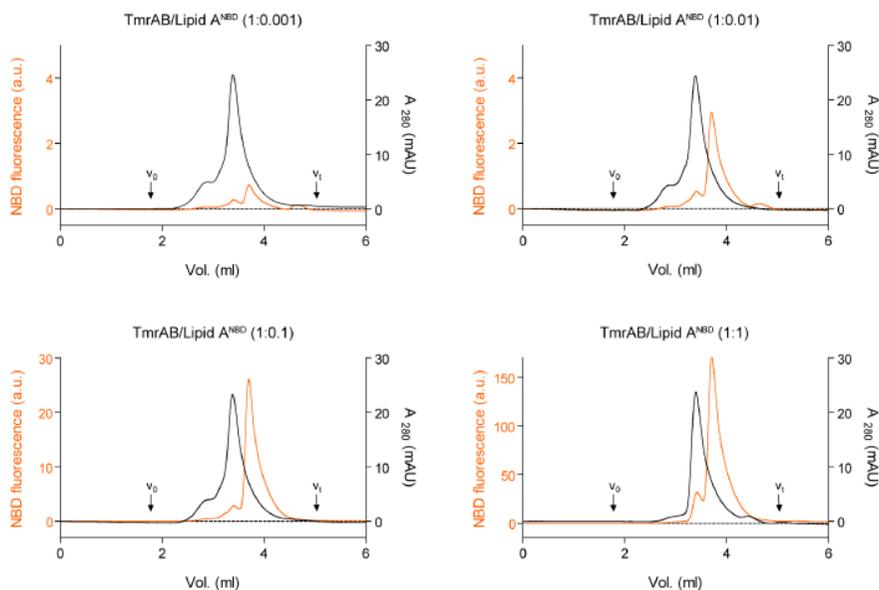


Figure 9.3: Co-migration of TmrAB/NBD-MPLA (**45**).

Subsequently, it was examined if NBD-MPLA was released from TmrAB in an ATP-dependent manner, similarly to what observed for endogenous phosphoglycolipids by ESI-MS (Bechara et al. 2015). However, as demonstrated by several independent experiments, NBD-MPLA appears just to participate in the detergent-protein micelles and no ATP- and TmrAB-dependent release of NBD-MPLA **45** was observed under different conditions tested. Probably, a reason of this difference with respect to similar endogenous phosphoglycolipids observed in the previous work (Bechara et al. 2015) could be explained with the lack on the NBD-MPLA **45** of the anomeric phosphate moiety. Indeed, typical MsbA substrates are thought to contain two phosphate groups. The lack of the anomeric phosphate group on the lipid A

reduces the toxicity, due to the selective induction of the TLR-4-TRIF signaling pathway over the TLR-4-MyD88 one, but leaving a solid interaction with the receptors and the immunostimulatory properties unaffected, whereas the same structural difference could presumably not leave the interaction with the ABC transporters unaltered. To clarify this point, the semi-synthesis of a fluorescent diphosphoryl lipid A (NBD-DPLA) was performed (**46**, Scheme 9.1). At first, several attempts, focused on the use of diphosphoryl lipid A **1** (see chapter 6) and its oxidation on the C-6 of GlcNII following the same conditions used on the monophosphoryl lipid A **2** (see paragraph 7.1), were tested without good results. Probably, in the conditions tested, the presence in the diphosphoryl lipid A structure of two polar heads, increases the formation of aggregates like inverse micells which hamper the exposition of the C-6 moiety to the oxidation reagents. Alternatively, we tried to install a further phosphate group on the anomeric position of **45**, but even in this case we did not achieve the desired compound. Finally, working on NBD derivative alternative to the amine obtained for the previous case, together with a different approach for the reaction on **1**, we achieved the synthesis of **46**, MALDI-TOF mass spectrum and full NMR characterization of the major hexa-acylated species confirmed its structure (Scheme 9.1; Figure 9.4-9.5; Table 9.1). Contrary to the results obtained for NBD-MPLA (**45**), the first biological tests on NBD-DPLA (**46**) clearly showed a better interaction/release with the ABC transporters protein, but the biological assays are still in progress and will not be showed in this thesis.

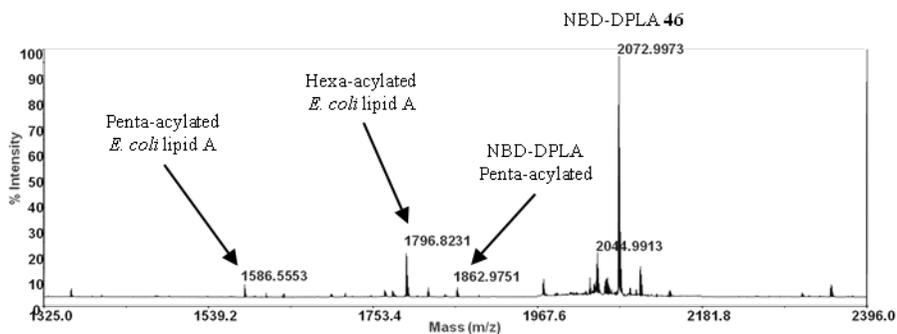


Figure 9.4: MALDI-TOF mass spectrum (negative ions mode) of **46**.

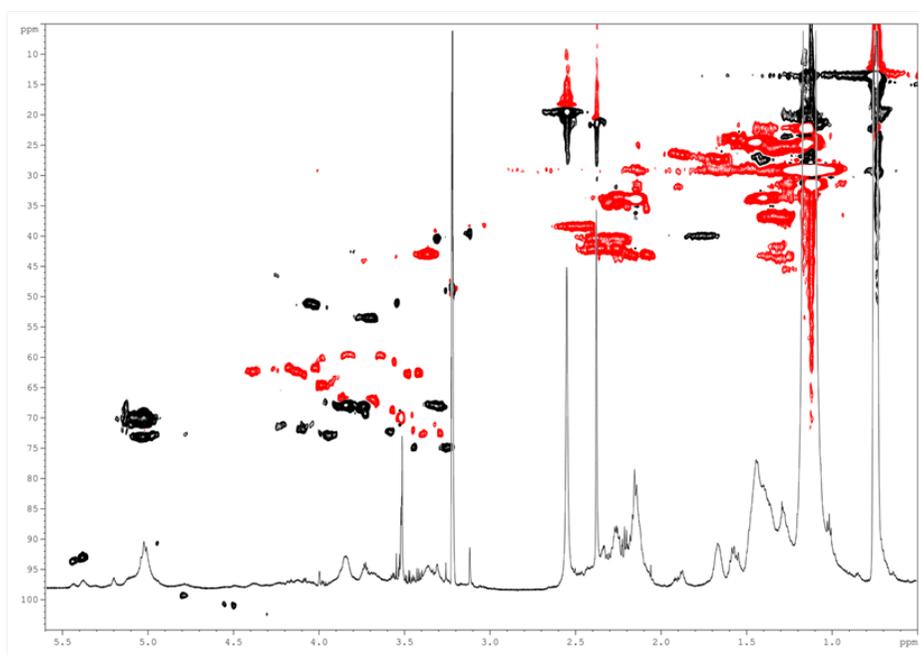


Figure 9.5: ^1H and HSQC-DEPT NMR spectra (600 MHz, 298K, 4:1 v/v CDCl_3 - CD_3OD) of **46**.

Table 9.1: NMR data referred to lipid A derivative **45** and **46** (major hexa-acylated specie)

	Sugar portion ^a						Lipid portion ^a					Other signals ^a	Phosphate ^b		
	1	2	3	4	5	6	α	β	γ	δ -(ω -1)	ω				
45	GlcN I (α)	5.01	3.98	4.94	3.57	3.62	3.74/3.99	C14:0(3-OH) [N-2 GlcN I]	2.10, 2.20 <i>41.8</i>	3.72 <i>67.5</i>	1.31 <i>36.8</i>	1.07-1.18 29.1	0.71-0.74 13.5	Linker: α : 2.30/34.2; β : 1.61/24.7; γ : 1.27/26.5; δ : 1.45/27.6; ω : 1.70/29.1; ζ : 3.45/44.6 NBD: I:--/145.2; II:--/144.8; III:--/143.7; IV:--/122.5; V: 6.19/97.7; VI: 8.50/136.0	-1.1
		<i>90.8</i>	<i>51.5</i>	<i>72.8</i>	<i>70.1</i>	<i>72.2</i>	<i>67.5</i>	C14:0(3-OH) [O-3 GlcN I]	2.24, 2.35 <i>41.0</i>	3.82 <i>69.1</i>	1.31 <i>36.9</i>				
								C14:0(3-OR) [N-2 GlcN II]	2.23, 2.34 <i>40.8</i>	4.97 <i>70.5</i>	1.39 <i>33.6</i>				
	GlcN II	4.74	3.99	5.00	3.62	3.72	---	C14:0(3-OR) [O-3 GlcNII]	2.42, 2.51 <i>38.8</i>	5.06 <i>69.7</i>	1.47 <i>33.5</i>	1.07-1.18 29.1	0.71-0.74 13.5		
		<i>101.3</i>	<i>52.5</i>	<i>73.6</i>	<i>71.1</i>	<i>71.8</i>	<i>173.2</i>	C12:0	2.15 <i>33.9</i>	1.46 <i>24.3</i>	1.07-1.18 29.1				
								C14:0							
46	GlcN I (α)	5.38	4.05	5.03	3.31	3.94	3.69/3.87	C14:0(3-OH) [N-2 GlcN I]	2.09, 2.17 <i>43.1</i>	3.75 <i>68.2</i>	1.30 <i>36.6</i>	1.09-1.18 29.0	0.71-0.75 13.5	Linker: α : 2.27/33.5; β : 1.57/24.0; γ : 1.37/25.7; δ : 1.66/27.4; ω : 3.38/43.1 NBD: I:--/145.1; II:--/144.6; III:--/143.4; IV:--/122.1; V: 6.08/97.8; VI: 8.37/136.9	1.2/-0.4
		<i>93.2</i>	<i>51.3</i>	<i>70.2</i>	<i>68.3</i>	<i>73.2</i>	<i>66.9</i>	C14:0(3-OH) [O-3 GlcN I]	2.27, 2.37 <i>41.7</i>	3.85 <i>67.8</i>	1.35 <i>36.8</i>				
								C14:0(3-OR) [N-2 GlcN II]	2.21, 2.32 <i>40.6</i>	4.99 <i>70.7</i>	1.41 <i>33.8</i>				
	GlcN II	4.80	3.70	5.05	4.14	3.60	4.40/---	C14:0(3-OR) [O-3 GlcN II]	2.45, 2.52 <i>38.5</i>	5.07 <i>69.9</i>	1.44 <i>33.7</i>	1.09-1.18 29.0	0.71-0.75 13.5		
		<i>99.5</i>	<i>53.6</i>	<i>73.3</i>	<i>72.0</i>	<i>72.4</i>	<i>63.0</i>	C12:0	2.17 <i>33.9</i>	1.47 <i>24.5</i>	1.09-1.18 29.0				
								C14:0							

^a ¹H (600 MHz, 298K) and ¹³C (125 MHz, 298K) chemical shifts (in plain and in italic, respectively) expressed in δ relative to residual CHCl₃ (¹H: δ 7.26 ppm; ¹³C: δ 77.0 ppm) in 4:1 v/v CDCl₃-CD₃OD.

^b ³¹P (160 MHz, 298K) chemical shifts expressed in δ relative to 85% phosphoric acid (external standard).

9.3 Imaging on HEK and macrophages cells

Compounds **45** and **46** were subjected to a preliminary cell imaging test, in collaboration with Prof. Huser (Department of Physics, University of Bielefeld, Germany), aimed to observe their staining with cells on a deconvolution microscope.

Two different cell lines were used, Human Embryonic Kidney (HEK) blue cells and J774 macrophages (received from the Center for Cooperative Research in Biomaterials-CIC biomaGUNE, Spain). The former are commonly used in immunological studies, equipped with the entire receptor complex (TLR4/MD-2/CD14), in response to LPS stimuli and also reactive to the (commercially available) LPS conjugated with the fluorescent dye Alexa Fluor 568 (LPS568). Such conjugate was used as positive control (Figure 9.6 A), while the (commercially available) 4-chloro-7-nitrobenzofurazan (NBD-Cl) was used as negative one (Figure 9.6 B), with a dosage of $1\mu\text{g mL}^{-1}$ and $2\mu\text{g mL}^{-1}$ respectively.

Cells were grown on glass-bottom chamber slides and incubated with **45** and **46**, at a concentration of $2\mu\text{g mL}^{-1}$, and then prepared for imaging on DV Elite microscope system. In the case of the HEK blue cells, both types of lipid As showed positive staining, which indicates their uptake in endocytotic vesicles on the cells.

Surprisingly, **45** emitted strong signals as small green spots and had general localization on nearly all the observed cells (Figure 9.6 C), while **46** had weaker intensity in overall cell population, where only a few cells were found to be preferentially reactive (Figure 9.6 D). On J774 macrophages cells, although LPS568 was not reactive (Figure 9.7 A), staining results of **45** and **46** were comparable with those obtained on HEK cells (Figure 9.7 C and D), which confirm their recognition. These studies are still in progress and further experiments will be performed on

derivatives, in order to optimize the conditions for the obtainment of super-resolution microscopy and 3D imaging.

HEK blue cells

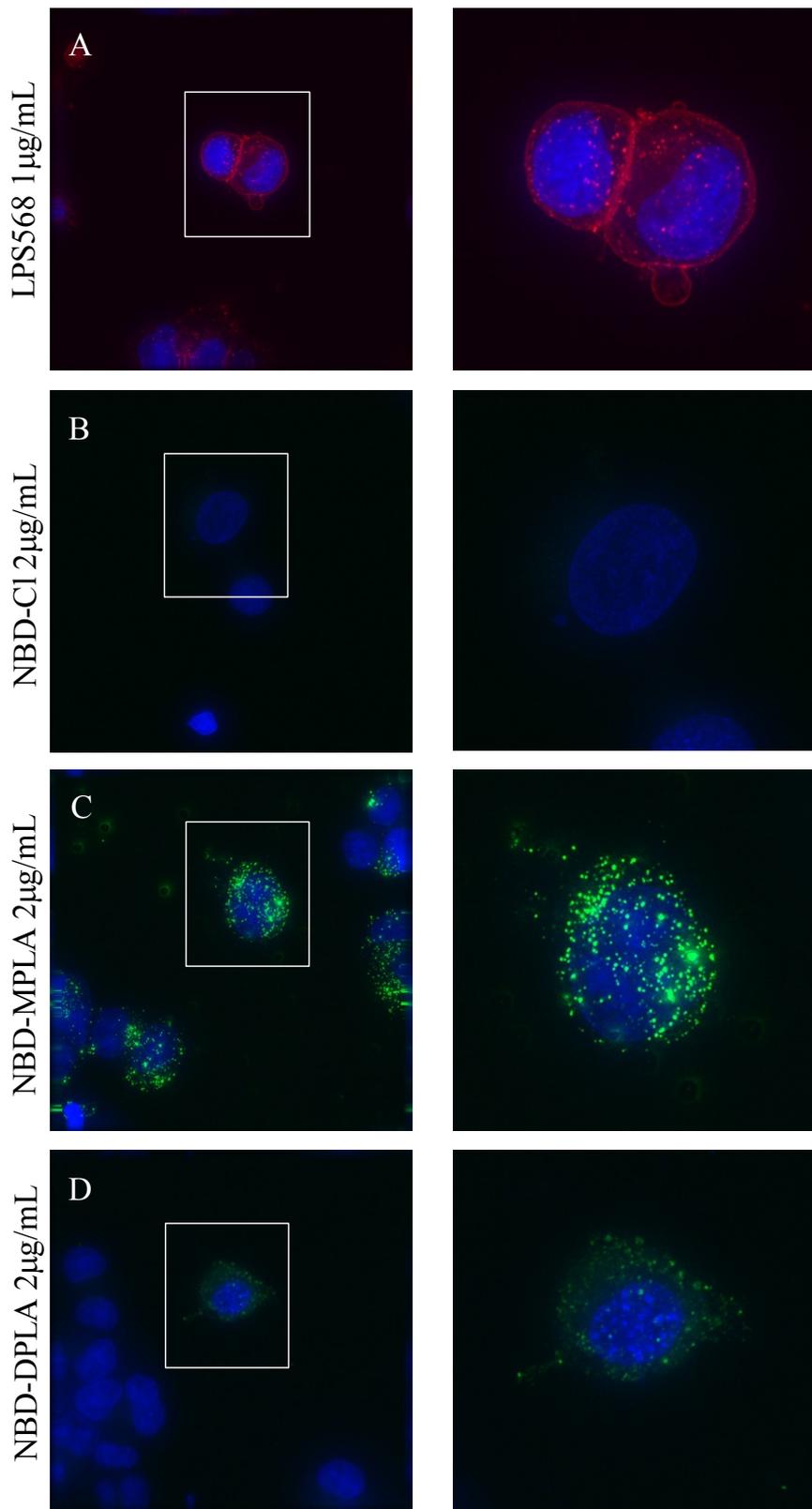


Figure 9.6: Fluorescent staining on HEK blue cells of LPS568, NBD-Cl, 45 and 46.

J774 Macrophages

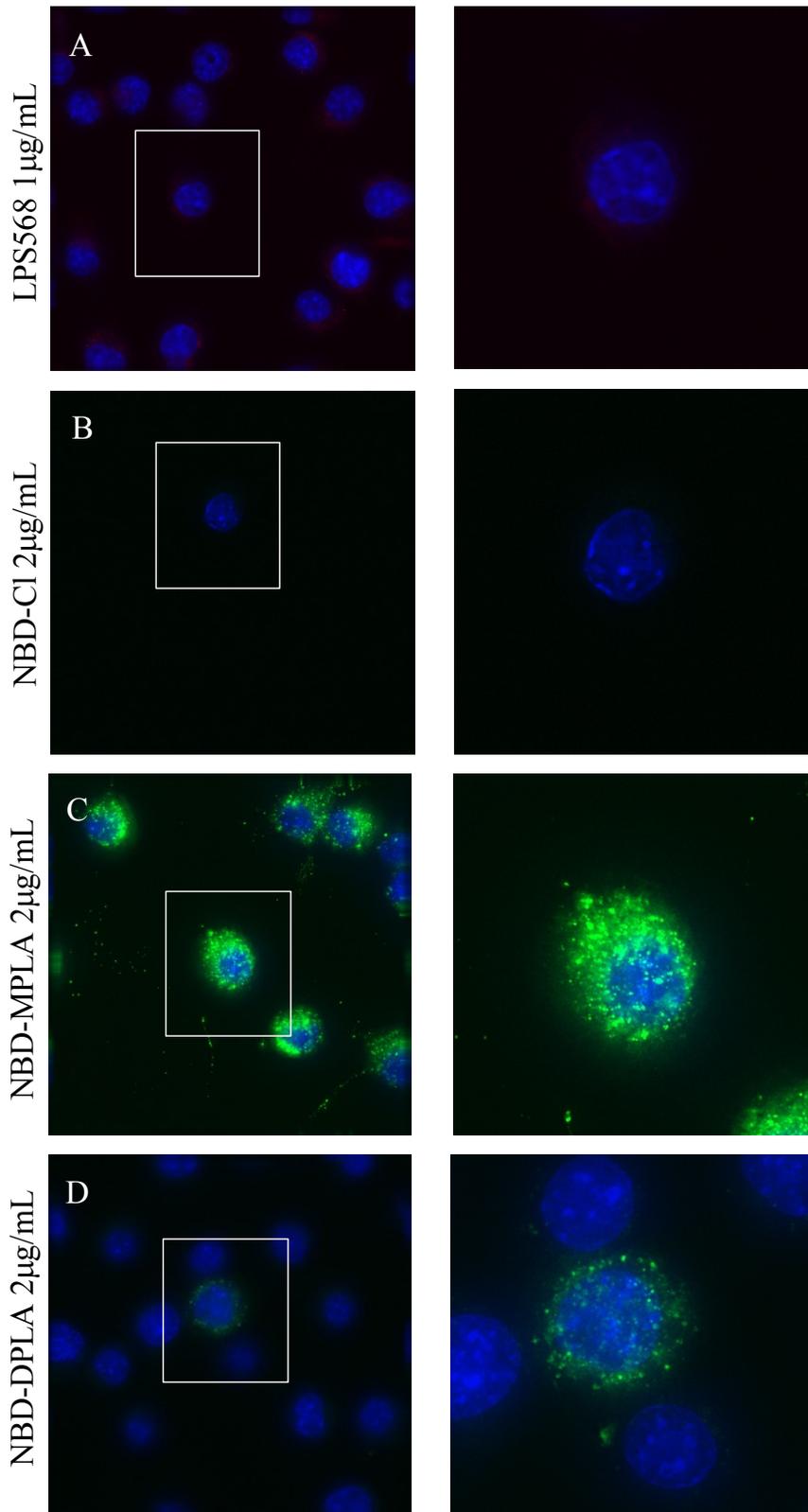


Figure 9.7: Fluorescent staining on J774 macrophages cells of LPS568, NBD-Cl, 45 and 46.

9.4 Conclusions

In the last part of this PhD work, two different fluorescent lipid A derivatives were for the first time semi-synthesized, in order to improve the understanding of lipids role in the mechanisms of ABC transporters. The first one was the mono-phosphoryl lipid A tagged with a NBD dye (NBD-MPLA; **45**). In collaboration with Prof. Tampé, **45** was employed for the development of some *in vitro* fluorescence-based lipid A-flippase assays, focused on the ABC transporters mechanisms comprehension. After the first data collected, we were excited to see that NBD-MPLA **45** co-migrates with TmrAB (200 kDa) in a fluorescence-based SEC, but unfortunately, by several independent experiments. It was observed, that **45** does not give a perfect interaction with the ABC transporters system, probably due to the lack of the anomeric phosphate. Therefore, after several attempts, a second fluorescent lipid A derivative was obtained. The *E. coli* diphosphoryl lipid A **1** was coupled with a suitable synthesized fluorescent tag to give the second semi-synthetic fluorescent diphosphoryl lipid A derivative (NBD-DPLA; **46**), which clearly showed a better interaction/release with the ABC transporters protein (biological assays are still in progress).

Finally, **45** and **46**, were subjected to a preliminary cell imaging test in collaboration with Prof. Huser, in order to observe their staining with two different cell lines on a deconvolution microscope (experiments are still in progress).

9.5 Experimental section

9.5.1 Lipid A derivative **45**

Compound **4** (20.0 mg, 11.6 μmol) was treated under argon atmosphere

with a solution of TBTU (4.6 mg, 14 μmol) and HOBt (1.9 mg, 14 μmol) in 1:1 v/v DMF- CH_2Cl_2 (0.6 mL). The synthesized amine-appended NBD derivative, obtained from commercially available 4-Chloro-7-nitrobenzofurazan (Yamaguchi et al. 2014), was subsequently added (9.6 mg, 34.8 μmol). Finally, the solution was treated with DIPEA (9.0 μL , 51 μmol) and stirred at rt for 3 hours. The reaction mixture was then diluted with CHCl_3 (2 mL) and washed with 0.1 M HCl (2 mL). The organic phase was collected, dried over anhydrous Na_2SO_4 , filtered and concentrated. The residue was subjected to gel filtration chromatography on Sephadex[®] LH-20 ($\text{CHCl}_3/\text{MeOH}$, 3:2 v/v) to afford **45** (17.9 mg, 78%) as a orange wax. For NMR data see Figure 9.2 and Table 9.1. MALDI TOF-HRMS for $\text{C}_{106}\text{H}_{190}\text{N}_7\text{O}_{25}\text{P}$ (major hexa-acylated species) (m/z): M (calcd) 1993.3583, M (found) 1992.4039 (M-H) (Figure 9.1).

9.5.2 Lipid A derivative 46

Compound **1** (11.5 mg, 6.4 μmol) was treated under argon atmosphere with a solution of DMAP (0.8 mg, 3.8 μmol) and 2,4,6-Collidine (3.6 mg, 28 μmol) in 1:1 v/v DMF- CHCl_3 (0.8 mL). The synthesized anhydride NBD derivative starting from the suitable synthesized 6-((7-nitrobenzo[*c*][1,2,5]oxadiazol-4-yl)amino)hexanoic acid, obtained from commercially available 4-Chloro-7-nitrobenzofurazan (Woodland et al. 2016), was subsequently added (6 mg, 10.6 μmol). The solution was stirred at rt for 16 hours. The reaction mixture was then diluted with CHCl_3 (2 mL) and washed with 0.1 M HCl (2 mL). The organic phase was collected, dried over anhydrous Na_2SO_4 , filtered and concentrated. The residue was subjected to gel filtration chromatography on Sephadex[®] LH-20 ($\text{CHCl}_3/\text{MeOH}$, 3:2 v/v) to afford **46** (7.2 mg, 30%) as a orange

wax. For NMR data see Figure 9.5 and Table 9.1. MALDI TOF-HRMS for $C_{106}H_{190}N_6O_{29}P_2$ (major hexa-acylated species) (m/z): M_r (calcd) 2073.9086, M_r (found) 2072.9973 (M-H)⁻ (Figure 9.4).

References

- Bao H. et al., *Biochim. Biophys. Acta* **2013**, *1828*, 1723–1730.
- Betaneli V., Petrov E. P., Schwille P., *Biophys. J.* **2012**, *102*, 523–531.
- Bechara C. et al., *Nat. Chemistry* **2015**, *7*, 255 – 262.
- Borst P., Zelcer N., van Helvoort A. *Biochim. Biophys. Acta* **2000**, *1486*, 128–144.
- Doerrler W.T., Raetz C.R., *J. Biol. Chem.* **2002**, *277*, 36697–36705.
- Dowhan W., Bogdanov M., *Biochem. Soc. Trans.* **2011**, *39*, 767–774.
- Eggenesperger S. et al., *J. Biol. Chem.* **2014**, *289*, 33098–33108.
- Gottesman M.M., Fojo T., Bates S.E., *Nature Rev. Cancer* **2002**, *2*, 48–58.
- Hendrich A.B., Michalak K., *Curr. Drug Targets* **2003**, *4*, 23–30.
- Huang Y., Sadee W., *Cancer Lett.* **2006**, *239*, 168–182.
- Karow et al., *Molecular Microbiology* **1993**, *7*, 69–79.
- King G., Sharom F.J. *Crit. Rev. Biochem. Mol. Biol.* **2012**, *47*, 75–95.
- Laganowsky A. et al., *Nature* **2014**, *510*, 172–175.
- Lee A.G., *Trends Biochem. Sci.* **2011**, *36*, 493–500.
- Lee A.G., *Biochim. Biophys. Acta* **2003**, *1612*, 1–40.
- Marek M. et al., *J. Biol. Chem.* **2011**, *286*, 21835–21843.

- Obara K. et al., *Proc. Natl Acad. Sci. USA* **2005**, *102*, 14489–14496.
- Polissi et al., *Molecular Microbiology* **1996**, *20*, 1221-1233.
- Reuter G. et al., *J. Biol. Chem.* **2003**, *278*, 35193–35198.
- Schweizer H.P., *Exp. Opin. Drug Discov.* **2012**, *7*, 633–642.
- Shintre C.A. et al., *Proc. Natl Acad. Sci. USA* **2013**, *110*, 9710–9715.
- Tarling E.J., De Aguiar Vallim T.Q., Edwards P.A., *Trends Endocrinol. Metab.* **2013**, *24*, 342–350.
- Wang et al., *PNAS* **2004**, *101*, 9774 –9779.
- Woodland et al., *Org. and Biom. Chemistry* **2016**.
- Yamaguchi T. et al., *Chem. Sci.* **2014**, *5*, 1021–1029.

Appendix

PhD Course Activity Summary

1) Attended Courses:

- Tecniche estrattive Solido-Liquido impiegate nella preparazione del campione per l'analisi chimica e nella produzione di estratti per usi industriali; Naviglio D.; 09-17/12/2014; 3 CFU.
- Sintesi, struttura ed applicazioni di oligonucleotidi naturali e modificati; Montesarchio D.; 04-13/02/2015; 3 CFU.
- Glicoscienza; Parrilli M., Bedini E.; 6-10/07/2015; 3 CFU.
- Spettrometria di Massa; Pucci P.; from 6/07/2015 to 10/07/2015; 4 CFU.
- Neutron scattering techniques; C. Zorn; 27-30/10/2015; 3 CFU.
- Applications of neutron scattering techniques in soft matter and bio-relevant systems. How to study their structure and dynamics; Y. Gerelli; 9-13/11/2015; 3 CFU.

2) Attended Seminars:

Title	Speaker	Place	Date
Multifunctional bioinspired catechol - based coatings and nanoparticles	Daniel Ruiz-Molina	Dept Chemical Sciences, University of Naples Federico II	05/11/2014
Le fitossine: un'avventura lunga 40 anni	Antonio Evidente	Dept Chemical Sciences, University of Naples Federico II	10/12/2014
New chiral catalysts derived from Iron(II) and Bismut (III) for asymmetric synthesis	Thierry Ollevier	Dept Chemical Sciences, University of Naples Federico II	15/12/2014
DNA-based strategies for the inhibitor of HMGB1 cytokine activity	Domenica Musumeci	Dept Chemical Sciences, University of Naples Federico II	22/01/2015
Gli accertamenti sulle impronte del Ra.C.I.S	Luca Niola	Dept Chemical Sciences, University of Naples Federico II	17/03/2015

Dall'analisi dei reflui alla depurazione degli impianti	Antonio Riccio	Dept Chemical Sciences, University of Naples Federico II	30/04/2015
Matrix metalloproteinase: inter-domain flexibility and recognition of substrates	Marco Fragai	Dept Chemical Sciences, University of Naples Federico II	08/05/2015
European large scale facilities: neutron and synchrotron source	Serge Perez	Dept Chemical Sciences, University of Naples Federico II	06/05/2015
Lectins from bacteria and fungi: therapeutical targets and research tools	Anne Imberty	Dept Chemical Sciences, University of Naples Federico II	15/05/2015
Characterizing biomolecular interactions: Biacore experience	Silvia Pileri	Dept Chemical Sciences, University of Naples Federico II	19/05/2015
Pharmaceutical companies: external manufacturing and quality assurance	D. Demasi	Dept Chemical Sciences, University of Naples Federico II	25/05/2015
Biosensing and bioelectronics based on organic electrochemical devices: from monitoring drug dynamics to hybrid bio-organic devices	Salvatore Iannotta	Dept Chemical Sciences, University of Naples Federico II	11/06/2015
The bacterial cell wall by liquid state, standard and DNP solid state NMR	Jean-Pierre Simorre	Dept Chemical Sciences, University of Naples Federico II	29/06/2015
Synthetic and biofunctional studies of microbial and animal glycan toward immunoregulation	Koichi Fukase	Dept Chemical Sciences, University of Naples Federico II	29/09/2015
Innovative approaches for polysaccharide-based vaccines	Francesco Berti	Dept Chemical Sciences, University of Naples Federico II	20/11/2015
Industria e ricerca nel settore biofarmaceutico: bisogni attuali e sviluppi futuri	Sara Carillo	Dept Chemical Sciences, University of Naples Federico II	25/02/2016
Basics of detergents formulations and challenges	Giulia Bianchetti	Dept Chemical Sciences, University of Naples Federico II	16/03/2016
La Chimica: un diamante	Mario Marzullo	Dept Chemical Sciences, University of Naples Federico II	21/04/2016
Chemical biology and medicinal chemistry of glycosphingolipid metabolism	Herman Overkleeft	Dept Chemical Sciences, University of Naples Federico II	19/12/2016

Modified nucleosides and oligonucleotides for biomedical applications	Daniela Montesarchio	Dept Chemical Sciences, University of Naples Federico II	18/01/2017
Antigens on the move - Structure and mechanism of translocation machineries in adaptive immunity	Robert Tampé	Dept Chemical Sciences, University of Naples Federico II	20/09/2017

3) Publications:

- Ziaco M, Górska S, Traboni S, Razim A, Casillo A, Iadonisi A, Gamian A, Corsaro MM and Bedini E “Development of clickable monophosphoryl lipid As towards semi-synthetic conjugates with tumor-associated carbohydrate antigens”. *J. Med. Chem.* **2017**, *60*, 9757-9768.
- Ziaco M, Hank S, Molinaro A, Corsaro MM, Tampé R and Bedini E “Fluorescence-based lipid A-flippases assays and ABC transporters mechanisms comprehension” (in preparation).
- Casillo A, Parrilli E, Ziaco M, Tutino ML, Corsaro MM “Psychrophiles lipopolysaccharides: isolation, characterization and biological activity” (submitted).
- Casillo A, Ziaco M, Lindner B, Parrilli E, Schwudke D, Holgado A, Verstrepen L, Sannino F, Beyaert R, Lanzetta R, Tutino ML and Corsaro MM “Unusual Lipid A from a Cold-Adapted Bacterium: Detailed Structural Characterization” *ChemBioChem* **2017**, *18*, 1845 - 1854.
- Casillo A, Papa R, Ricciardelli A, Sannino F, Ziaco M, Tilotta M, Selan L, Marino G, Corsaro MM, Tutino ML, Artini M, and Parrilli E “Anti-Biofilm Activity of a Long-Chain Fatty Aldehyde from Antarctic *Pseudoalteromonas haloplanktis* TAC125 against *Staphylococcus epidermidis* Biofilm” *Front Cell Infect Microbiol.* **2017**, *7*, 46.
- Casillo A, Ziaco M, Lindner B, Merino S, Mendoza-Barberá E, Tomás JM and Corsaro MM “Structural Characterization of Core Region in *Erwinia amylovora* Lipopolysaccharide” *Int. J. Mol. Sci.* **2017**, *18*, 559.

- D'Alonzo D, Cipolletti M, Tarantino G, Ziaco M, Pieretti G, Iadonisi A, Palumbo G, Alfano A, Giuliano M, De Rosa M, Schiraldi C, Cammarota M, Parrilli M, Bedini E and Corsaro MM “A Semisynthetic Approach to New Immunoadjuvant Candidates: Site-Selective Chemical Manipulation of *Escherichia coli* Monophosphoryl Lipid A” *Chem. Eur. J.* **2016**, *22*, 1 - 12.
- Ziaco M, De Castro C, Silipo A, Corsaro MM, Molinaro A, Iadonisi A, Lanzetta R, Parrilli M and Bedini E “Synthesis of the tetrasaccharide outer core fragment of *Burkholderia multivorans* lipooligosaccharide” *Carbohydrate Research* **2015**, *403*, 182-191.

4) **Attended congresses/workshops/summer schools/contribution:**

- 1st Napoli - Osaka Meeting; 10/07/2017; Napoli (Italy) University of Naples Federico II, Centro Congressi Partenope - **Oral Communication:** Site selective derivatization of *Escherichia Coli* Monophosphoryl Lipid A (M. Ziaco, S. Górska, E. Bedini and M.M. Corsaro)
- 19th European Carbohydrate Symposium - EUROCARB 2017; 2 - 6 /7/2017; Barcelona (Spain); **Poster session:** Site selective derivatization of *Escherichia Coli* Monophosphoryl Lipid A (M. Ziaco, S. Górska, E. Bedini and M.M. Corsaro).
- 4th Meeting TOLLerant: Molecular Aspects of Host/Microbe dialogue 5 - 7 /06/2017; Napoli (Italy) University of Naples Federico II, Complesso dei SS. Marcellino e Festo.
- BMMC VII - 7th Baltic Meeting on Microbial Carbohydrates; 25 - 29/09/2016; Güstrow (Rostock); **Oral Communication:** Site selective derivatization of *Escherichia Coli* Monophosphoryl Lipid A (M. Ziaco, D. D'Alonzo, G. Palumbo, M. Giuliano, M. De Rosa, C. Schiraldi, M. Parrilli, E. Bedini and M.M. Corsaro).
- IEIIS 2016: 14th Biennial Meeting International Endotoxin and Innate Immunity Society; 22 - 24/09/2016; Hamburg (Germany); **Poster session:** Site selective derivatization of *Escherichia Coli* Monophosphoryl Lipid A (M. Ziaco, D. D'Alonzo, G. Palumbo, M. Giuliano, M. De Rosa, C. Schiraldi, M. Parrilli, E. Bedini and M.M. Corsaro).

- XIV Convegno - Scuola Sulla chimica dei carboidrati; 19 - 22/06/2016; Certosa di Pontignano (Siena); **Oral Communication:** Site selective derivatization of *Escherichia Coli* Monophosphoryl Lipid A (M. Ziaco, D. D'Alonzo, G. Palumbo, M. Giuliano, M. De Rosa, C. Schiraldi, M. Parrilli, E. Bedini, M.M. Corsaro).
- XV Edizione delle Giornate Scientifiche C.I.N.M.P.I.S.; Challenges in Organic Synthesis: Efficient Processes fo Novel Applications; 11 - 12 /12/2015; Napoli (Italy)
- COST Training School BM 1003: Microbial cell surface determinants of virulence as target for new therapeutics in Cystic Fibrosis; 03-05/06/2014; University of Naples Federico II, Complesso di Monte Sant'Angelo.

5) Other Activities:

- In May 2016, elected as PhD student representative in the Department of Chemical Sciences at the University of Naples Federico II, and therefore engaged as member of Departmental Council and Joint Committee.