



**University of Naples “Federico II” Polytechnic and Basic
Sciences School**

Ph.D. School in Chemical Sciences – Cycle XXXV

The elucidation of the chemical structure of
Lipopolysaccharides and other glycoconjugates isolated
from bacteria of the Human Microbiota

Molly Dorothy Pither

Tutors

Prof. Antonio Molinaro

Prof. Flaviana Di Lorenzo

Examiner

Prof. Angela Amoresano



This PhD *was* funded through the European Training Network: Sweet CrossTalk within the *European* Union's Horizon 2020 research and innovation program, grant agreement number: 814102.

Acknowledgments

Firstly, I would like to express my sincere gratitude to Professor Antonio Molinaro for the continuous support during my PhD and throughout my years in Naples. I would also like to extend my sincere thanks to Professor Alba Silipo and Professor Cristina De Castro who have both provided valuable insights and support during my PhD.

I would like to express my deepest gratitude to Professor Flaviana Di Lorenzo, for her expertise, encouragement, and patience throughout my PhD, both scientifically and as a true friend.

I would like to thank everyone past and present, within the SSCN research group in Naples. With a special thanks to Ferran Nieto Fabregat, Anna Notaro, Imma Speciale, Pilar García del Vello, Cristina Di Carluccio, Valeria Pistorio, Marta Tiemblo Martín, Tania Gerpe, Fabiana Esposito, and Roberta Cirella, all of whom I worked alongside in laboratory, kept me laughing, provided me a bed to sleep in, and even drove me to work.

I would like to express a special thanks to Serena Traboni and Marcello Ziaco who have continuously supported me throughout all my years in Naples. Their friendship and love will stay with me forever.

I am also grateful to Professor Sonsoles Martín Santamaría for giving me the opportunity to come and learn from her expertise. With a special thanks to Alejandra Matamoros Recio who patiently guided and taught me, but who has become a great friend.

I would like to thank my family back in the UK, especially my Mum. I could not have undertaken this journey without her support and her unconditional love. Thank you for always being available to talk, provide excellent advice and encouragement that kept me going when things were tough.

Lastly, I would like to thank Ciro Petricciuolo, for his infinite patience and love, that continues to push me to be the best version of myself. A thanks should also go to all the Petricciuolo family who gave me a home in Naples.

Summary

The human intestinal tract is colonised by a complex community of microorganisms, collectively known as the gut microbiota (GM). These microbes provide metabolic functions to the human host, which are also sustained and influenced by the host environment, such as diet, immune system, and exposure to antibiotics. Research into the microbiota field has increased in the last ten years. Nevertheless, there are still many questions pending, such as: *how do bacteria adapt into an environment as complex as the human gut? How do bacteria compete or work symbiotically together, among and/or within species? How do bacteria gain immune tolerance from the host immune system?* A better understanding of the interactions occurring between the GM and the human host could influence our current view on pathologies, normal physiology, but also on nutrition, catalysing nutritional recommendations and policies dependent on age, geography, diet, and health of individuals.

Gram-negatives make up a huge majority of the bacteria within the human GM. They intrinsically possess lipopolysaccharides (LPS) on their outer membrane, which are glycoconjugates universally considered “inflammatory”. Yet, humans and GM live in, not only a harmonious, but also in a mutualistic relationship. Deciphering how this occurs despite the constant presence of LPS is crucial and studies aimed at shedding light on this aspect promise results of unquestionable importance.

LPS molecules are composed of three distinct chemical components: the lipid A, the core oligosaccharide (core OS), and a polysaccharide repeating unit (the O-antigen). Depending on the nature of the saccharide moiety, an LPS can be classified as smooth or rough: smooth-type LPS (S-LPS) is built up of all the three above domains, whereas a rough-type LPS (R-LPS) does not express the O-antigen. Although all LPS possess this overall architecture, the chemical structure of LPS is extremely complex and variable. LPS are recognised in a structure dependent manner by the innate immunity receptor complex known as the Toll-Like receptor 4/Differentiation factor-2 (TLR4/MD-2). Thus, minute structural changes in the LPS structure can lead to

completely diverse host immunological outcomes. Given this, it is of utmost importance to characterise how LPS from harmless gut bacteria are recognised by the human immune system, which implicitly means determining their chemical structures. This can lead to understanding the key molecular features that render these LPS somehow “invisible” and/or “not dangerous” to the host and potentially even beneficial molecules. To this aim, this PhD entailed the isolation, purification, and full chemical characterisation of LPS from an assortment of different and widespread GM species. The complete LPS structures reported in this thesis were all attained through converging information from chemical analyses via gas liquid chromatography–mass spectrometry (GC–MS), matrix-assisted laser desorption/ionisation MS (MALDI-MS), and Nuclear Magnetic Resonance (NMR) spectroscopy.

The first project discussed in this thesis (*Chapter 3*) is dedicated to the chemical characterisation of the LPS from *Bacteroides vulgatus*, a prominent species within the human intestinal tract. The lipid A was confirmed to be a heterogeneous mixture of tetra- and penta-acylated species. As for the saccharide portion, the core region was built up of a novel hexa-saccharide (**Figure I**) and the O-antigen structure was determined as built up of repeating units of β -mannose and α -rhamnose. For the immunological properties, in human macrophages the LPS from *B. vulgatus* showed a weaker capability to elicit pro-inflammatory cytokine production compared to the potent LPS from *E. coli*. In addition, the LPS was able to highly stimulate the release of IL-10, a potent anti-inflammatory cytokine. Furthermore, experiments performed by using HEK-Blue™ TLR2 cell lines showed a specific and significant interaction of *B. vulgatus* LPS with this receptor. Interestingly, this interaction was upregulated by the co-expression of TLR4 and TLR2 when co-transfected HEK-Blue™ cell lines were used, thus showing a clear synergistic effect between the two TLRs in NF- κ B activation and CXCL-8 production.

Therefore, *B. vulgatus* was revealed to have an S-LPS with a unique structural chemistry that was reflected in an equally uncommon immunocompetency. Indeed,

from this study, a “paradigm shift” in our comprehension of the role of LPS could have been proposed for the first time, that is LPS not anymore as a toxic molecule but also as a potentially beneficial glycoconjugate able to modulate human immune response.

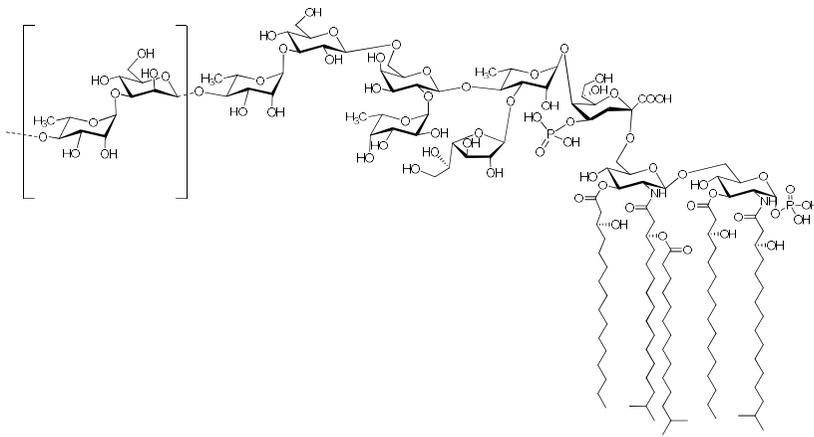


Figure I The full structure of the LPS from *B. vulgatus*

The second project argued in this thesis (*Chapter 4*) is centred on the LPS of *Alcaligenes faecalis*, which inhabits the gut-associated lymphoid tissues (GALT), Peyer's patches (PP). *Alcaligenes* spp. are able to create and maintain a homeostatic environment in PPs without triggering any harmful responses. Furthermore, it was previously demonstrated that *Alcaligenes* spp. LPS acts as a weak TLR4 agonist and it could promote interleukin-6 (IL-6) release from dendritic cells (DCs), which, in turn, enhanced Immunoglobulin A (IgA) production. On these bases, extraction and characterisation of this LPS from *A. faecalis* was completed. This bacterium was found to synthesise both S-LPS and R-LPS. The full structure is sketched in **Figure II**. Briefly the core OS was defined to be a complex nonasaccharide comprising Kdo, two *L-glycero-D-manno*-heptoses, two *N*-acetyl *D*-glucosamine, two *N*-acetyl *D*-galactosamine, one *D*-glucose, and one *D*-galactose. This novel core OS structure found in both the R-LPS and S-LPS of *A. faecalis* was also found to be more

heterogenous due to the presence of non-stoichiometrically linked sugar residues. In the S-LPS of *A. faecalis*, which inherently has an additional polysaccharide repeating unit (O-antigen), was disclosed to be a branched xylosylated rhamnan chain. Conclusively, the lipid A structure was also defined as a heterogenous mixture of *mono* and *bis*-phosphorylated tetra-, penta- and hexa-acylated species. This project then laid the bases for the chemical synthesis of the *A. faecalis* lipid A. This structure to function study was essential to appreciate that tetra- and penta-acylated lipid A species were unable to activate TLR4-mediated signalling, while the hexa-acylated ones were able to trigger it but at a less extent compared to *E. coli* LPS. Finally, these synthetic hexa-acylated lipid As were shown to activate IL-6 release in THP-1 cells which was comparable to what was previously observed after administration of the full *A. faecalis* LPS.⁴ Therefore, these results suggested that lipid A (in its hexa-acylated form, although present in minor amount) is the main responsible for most of the IL-6-mediated IgA production occurring in PPs. The intrinsic peculiarity of *A. faecalis* LPS to act as a weak TLR4-signalling activator but to be able, at the same, to potentially induce IgA production, are of paramount importance in the evaluation of this molecule as a potential adjuvant candidate in future vaccine production. As a matter of fact, thanks to this structure to function study, chemically synthesised *A. faecalis* lipid A have been recently proposed as potent and safe nasal vaccine adjuvants for the induction of *Streptococcus pneumoniae*-specific IgA and Th17 mediated protective immunity, augmenting both the innate and acquired arms of mucosal immunity against respiratory bacterial infection.

All in all, this study unveiled the novel and complex structure of the LPS from *A. faecalis*, with an in-depth structure to function investigation, which showed a peculiar immunological role that has enormous potential to be harnessed for future development of vaccine adjuvants inspired by molecules provided by our own GM.

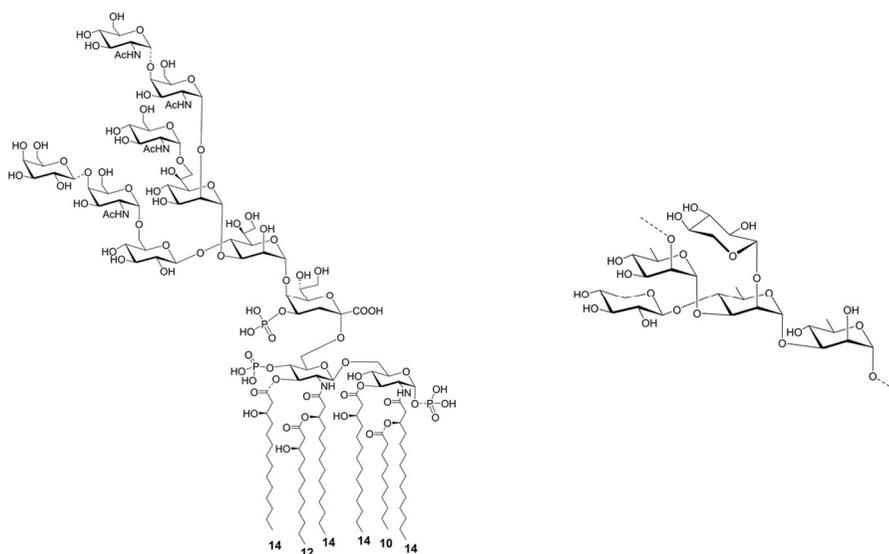


Figure II R-LPS structure from *A. faecalis* (Left) and the O-antigen structure of *A. faecalis* (Right)

In *Chapter 5* the characterisation of the chemical structure and immunological activity of the R-LPS from the prominent gut bacterium *Bacteroides thetaiotaomicron* is discussed. The extraction of the LPS was performed on a bacterial strain mutated in the gene cluster for the production of capsular polysaccharides (CPS) to enable a successful isolation of the sole LPS molecule. The extracted R-LPS underwent different chemical treatments (full de-acylation and mild acid hydrolysis) followed by multiple different analytical methods including 1D and 2D NMR experiments, MALDI-TOF MS and MS/MS, and GC-MS analysis, to define the chemical structure of the R-LPS from *B. thetaiotaomicron* (**Figure III**). The immunological activity of *B. thetaiotaomicron* LPS was investigated *in vitro* using HEK-Blue™ cells stably transfected with human TLR4 (hTLR4), MD-2, and CD14 genes or with human TLR2 (hTLR2) gene. These results demonstrated that the R-LPS from *B. thetaiotaomicron* was a weaker activator of the TLR4 signalling compared to the known pro-inflammatory LPS from the enteropathogen *Salmonella typhimurium* and showed to

activate in a dose dependent manner the TLR2-mediated signalling. Further investigation using multiple reaction monitoring mass spectrometry (MRM-MS) based method demonstrated that HEK-Blue™ cells transfected with hTLR2 and stimulated with the R-LPS from *B. thetaiotaomicron* were the main responsible for the release of pro-inflammatory IL-2, IL-6, and IL-1 β , whereas cells transfected with hTLR4 showed to be in charge of the release of the anti-inflammatory IL-10 upon their stimulation by this gut symbiotic R-LPS.

Thanks to this study, the unusual and highly heterogenous chemical structure of the R-LPS of *B. thetaiotaomicron* was deduced. This represents an extremely important step in the microbiota research field as *B. thetaiotaomicron* still remains the most studied mutualist of human intestine, although nothing was known, so far, about the chemistry of its LPS along with the immunological behaviour of the LPS devoid of any CPS contamination. Moreover, this study unveiled for the first time the diverse impact(s) and role(s) that *B. thetaiotaomicron* R-LPS has on activation of TLR2 and TLR4-mediated signalings, with the observation of induction of a clear differential cytokine release profile.

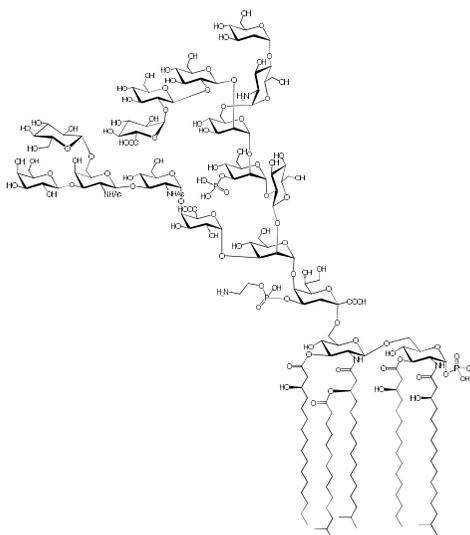


Figure III Structure of the R-LPS from *B. thetaiotaomicron*

Another project described in this PhD thesis is related to the investigation of the structure to function relationship of the LPS from two isolates of *Veillonella parvula*. This is a bacterium abundant in the normal healthy microbiota of various organs but that has also been found to be opportunistic in certain pathological situations, making it a “pathobiont” worth to be studied. The extraction and characterisation of two isolates of *V. parvula* was executed: a human intestinal isolate and an oral cavity isolate. Differences in the nature of the LPS from these two isolates was immediately observed, that is the production of an S-LPS in the intestinal isolate and of an R-LPS in the oral isolate. In addition, the fatty acid compositional analysis showed important differences mainly regarding the degree of branching, which was found to be higher in the oral isolate than in the intestinal one. A full structural study on the LPS from both isolates was performed by merging chemical analyses data, NMR and MS investigations. Two novel LPS structures were defined, both with fascinating chemical properties (**Figure IV**). An unprecedented O-antigen polysaccharide structure was deduced for the intestinal isolate. Given the peculiar chemistry, the immunological behaviour of both *V. parvula* LPS was also analysed. To this aim, among other systems considered within the frame of this project, HEK-Blue™ cells transfected with hTLR4 were used and clearly showed that the two LPS exert a differential activation of the signalling mediated by this innate immune receptor. Briefly, the R-LPS from the oral cavity strain acted as a weaker activator of the pro-inflammatory response compared to the S-LPS isolated from the intestinal isolate. To gain insights into this phenomenon, the molecular details and mechanisms underlying how the different lipid A structures from each isolate of *V. parvula* interact with TLR4/MD-2 was investigated by using molecular dynamic approaches. This work was completed during my three-month secondment in Madrid, Spain at the CSIC under the supervision of Prof. Martín-Santamaría.

This investigation into the LPS from *V. parvula* from different organs of the human host revealed important structural differences responsible for specific immunological properties, further highlighting how intricate and delicate is the balance between the

chemical structure of an LPS from human microbiota and its recognition and impact on the human immune system.

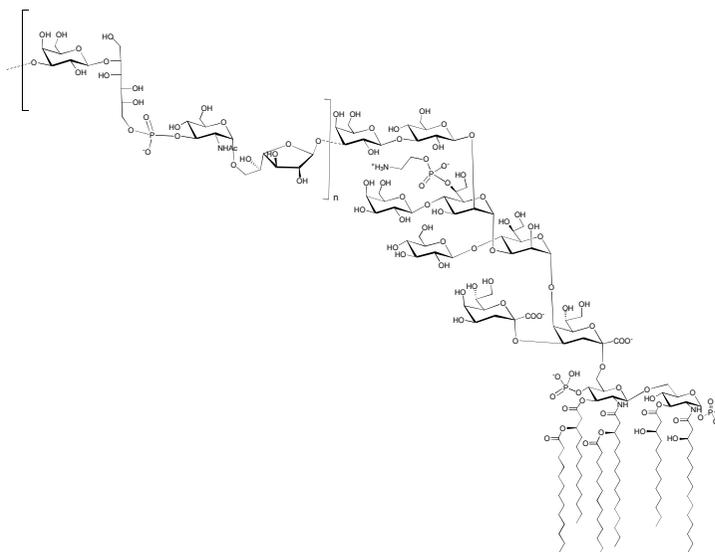


Figure IV Structure of the S-LPS from *V. parvula* (intestinal isolate)

The final project discussed is focused on *Bacteroides uniformis* (Chapter 7). This is another important bacterium of the healthy GM and is also extensively studied for its ability to induce the release of the anti-inflammatory cytokine IL-10. Thanks to this project, the R-LPS was successfully isolated from a few milligrams of starting bacterial pellet. Following the isolation and purification of the LPS, the lipid A species were characterised, and the saccharide moieties were analysed. Despite the low amount of material, given by both cell growth problems and occurrence of polysaccharidic natural contamination, this project was extremely important to set up a protocol to overcome obstacles in the extraction procedures, thus guiding future LPS isolation that might result in similar drawbacks.

An *Appendix* chapter has been added to the present thesis (Chapter 8: *Appendix*). This covers two additional projects on the full characterisation of the lipid A from a cystic fibrosis pathogen *Pandoraea pulmonicola*, and the full chemical characterisation of

the unusual lipid A from two species of *Echinicola* genus, which are marine Gram-negatives belonging to the Bacteroidetes phylum. The former project on *P. pulmonicola* was fundamental to prove the capability of some bacteria to extensively modify the chemistry of their LPS to allow the colonisation and persistence within the human body. Peculiar structural features were identified in the lipid A of the investigated *P. pulmonicola* strain that might help in the future preparation of new treatment or therapeutic strategies to deal with dangerous opportunistic pathogens in particular in the case of cystic fibrosis disease. On the other hand, the project on *Echinicola* strains, which was further expanded by the evaluation of the LPS immunological properties, was key to dissect the chemistry of lipid A from environmental Bacteroidetes. This study intriguingly showed that both *Echinicola* LPS exert a weak activation of the TLR4 mediated signalling, as observed for other Bacteroidetes studied in the frame of this PhD project.

Thanks to these two additional studies, it was further highlighted the importance of studying in parallel to commensal, also pathogenic and environmental LPS chemistry and immunity especially in the perspective of design of new molecules that could be used as antagonists (inhibitors) or immunotherapeutics, as well as in the conception of new antimicrobial compounds to treat opportunistic pathogen infections.

The results presented in this PhD thesis were adapted from the following articles co-written by the author of the present thesis:

- (1) Di Lorenzo, F., **Pither, M. D.**, Martufi, M., Scarinci, I., Guzmán-Caldentey, J., Łakomicz, E., Jachymek, W., Bruijns, S. C. M., Santamaría, S. M., Frick, J.-S., van Kooyk, Y., Chiodo, F., Silipo, A., Bernardini, M. L. & Molinaro, A. Pairing *Bacteroides vulgatus* LPS structure with its immunomodulatory effects on human cellular models. *ACS Cent. Sci.* 6, 1602–1616 (2020)
- (2) Shimoyama, A., Di Lorenzo, F., Yamaura, H., Mizote, K., Palmigiano, A., **Pither, M. D.**, Speciale, I., Uto, T., Masui, S., Sturiale, L., Garozzo, D., Hosomi, K., Shibata, N., Kabayama, K., Fujimoto, Y., Silipo, A., Kunisawa, J., Kiyono, H., Molinaro, A. & Fukase, K. Lipopolysaccharide from Gut-associated lymphoid-tissue-resident *Alcaligenes faecalis*: complete structure determination and chemical synthesis of its lipid A. *Angew. Chem. Int. Ed.* 60, 10023–10031 (2021)
- (3) **Pither, M.D.**, McClean, S., Silipo, A. et al. A chronic strain of the cystic fibrosis pathogen *Pandoraea pulmonicola* expresses a heterogenous hypoacylated lipid A. *Glycoconj J* 38, pp.135–144 (2021)
- (4) **Pither M.D**, Mantova G., Scaglione E., Pagliuca C., Colicchio R., Vitiello M., Chernikov O.V., Hua K-F., Kokoulin M.S., Silipo A., Salvatore P., Molinaro A., Di Lorenzo F. The unusual lipid A structure and immunoinhibitory activity of LPS from marine bacteria *Echinicola pacifica* KMM 6172^T and *Echinicola vietnamensis* KMM 6221^T. *Microorganisms.*; 9,12, pp. 2552, (2021)
- (5) **Pither, M. D.**, Illiano, A., Pagliuca, C., Jacobson, A., Mantova, G., Stornaiuolo, A., Colicchio, R., Vitiello, M., Pinto, G., Silipo, A., Fischbach, M. A., Salvatore, P., Amoresano, A., Molinaro, A. & Di Lorenzo, F. *Bacteroides thetaiotaomicron* rough-type lipopolysaccharide: the chemical structure and the immunological activity. *Carbohydr. Polym.* 297, 120040 (2022)
- (6) **Pither, M.**, Molinaro, A., Silipo, A. & Di Lorenzo, F. in *Methods Mol. Biol.* 2613, (Springer International Publishing) (2022). **In press.**

Other papers and chapters not described in this thesis are:

- (7) **Pither, M. D.**, Silipo, A., Di Lorenzo, F. & Molinaro, A. Glycans in Bacterial Infections: Gram-Negative Infections in the Respiratory Tract, in *Compr. Glycosci. Second Ed.* (ed. Barchi, J. J.) 233–249 (2021)
- (8) **Pither, M. D.**, Sun, M.-L., Speciale, I., Silipo, A., Zhang, Y.-Z., Molinaro, A. & Di Lorenzo, F. Structural determination of the lipid A from the deep-sea bacterium *Zunongwangia profunda* SM-A87: a small-scale approach. *Glycoconj. J.* 39, 565–578 (2022).

Manuscripts in preparation:

- (9) **Pither, M. D.**, Rocca, G., Balzarini F, Matamoros Recio A., Colicchio, R., Salvatore, P., Chiodo, F., Granucci, F., Molinaro, A., Di Lorenzo, F. Location within the human gastrointestinal tract and the impact on the LPS chemistry and immunology: the case of *Veillonella parvula*. **In preparation.**
- (10) Garcia-Vello, P., Tytgat, H.L.P., Elzinga, J., **Pither M.D.**, Plovier, H., Nicolardi, S., Cani, P., De Castro, C., Silipo, A., Di Lorenzo, F., Molinaro, A., de Vos W.M., The lipooligosaccharide of *Akkermansia muciniphila* has a unique structure and signals through Toll Like Receptor 2. **In preparation.**
- (11) **Pither, M. D.**, Pagliuca C, Colicchio R, Vitiello M, Kokoulin MS, Salvatore P, Molinaro A, Di Lorenzo F. The structure and immunological properties of the lipid A isolated from marine bacteria *Cellulophaga* sp.. **In preparation**

Abbreviations

AA	Acetylated Alditols
AMG	Acetylated Methyl Glycosides
COSY	Correlation Spectroscopy
CPS	Capsular Polysaccharide
DC	Dendritic Cells
DMSO	Dimethyl Sulphoxide
DNase	Deoxyribonuclease I From Bovine Pancreas
DOSY	Diffusion Ordered Spectroscopy
EPS	Exopolysaccharide
F/B Ratio	Firmicutes/Bacteroidetes Ratio
GC-MS	Gas Chromatography Mass Spectrometry
GF	Germ Free (Mice)
GM	Gut Microbiota
HEK	Human Embryonic Kidney (Cells)
HMBC	Heteronuclear Multiple Bond Correlation Spectroscopy
HSQC	Heteronuclear Single-Quantum Correlation Spectroscopy
IBD	Inflammatory Bowel Disease
IL	Interleukin
LBP	LPS Binding Protein
LipA	Lipid A
LPS	Lipopolysaccharides
LTA	Lipoteichoic Acid
MALDI -TOF	Matrix-Assisted Laser Desorption/Ionisation Mass Spectrometry
MAMP	Microbe Associated Molecular Pattern
MAPK	Mitogen-Activate Protein Kinase
MD-2	Myeloid Differentiation Factor 2
MS	Mass Spectrometry
MWCO	Molecular Weight Cut Off
NABD₄	Sodium borodeuteride
NMR	Nuclear Magnetic Resonance

NOE	Nuclear Overhauser Effect
NOESY	Nuclear Overhauser Spectroscopy
OM	Outer Membrane
OS	Oligosaccharide
PAGE	Polyacrylamide Gel Electrophoresis
PCP	Petroleum/Chloroform/Phenol
PDB	Protein Data Bank
PGN	Peptidoglycan
PMAA	Partially Methylated Acetylated Alditols
PS	Polysaccharides
RNase	Ribonuclease A
ROESY	Rotating Frame Nuclear Overhauser Effect
SDS	Sodium Dodecyl Sulphate
SEC	Size-Exclusion Chromatography
TFA	Trifluoroacetic Acid
TLR	Toll-Like Receptor
TNF-α	Tumour Necrosis Factor α
TOCSY	Total Correlation Spectroscopy
WT	Wild Type
WTA	Wall Teichoic Acid

Sugar abbreviations

Fuc	Fucose
Gal	Galactose
GalA	Galacturonic acid
GalN	Galactosamine
GalNAc	<i>N</i> -Acetylgalactosamine
Glc	Glucose
GlcA	Glucuronic acid
GlcN	Glucosamine
GlcNAc	<i>N</i> -Acetylglucosamine
Hep	Heptose
Kdo	3-deoxy-D- <i>manno</i> -octulosonic acid
Ko	D-glycero-D-talo-oct-2-ulosonic acid
Man	Mannose
MurNAc	<i>N</i> -acetylmuramic acid
Rha	Rhamnose

Table of Contents

Section 1: The gut microbiota	1
1.1 Overview of Gut Microbiota – The Forgotten Organ	2
1.2 History of microbiota research	3
1.3 The Gut Microbiota (GM)	7
1.3.1 GM and immune system	10
1.3.2 Gut Microbiota: the bacterial community	13
Section 2: Gram-negatives of the gut microbiota and their cell surface glycans	15
1.4 Gram-negative bacteria: an overview	16
1.5 Lipopolysaccharides	18
1.5.1 Lipid A	19
1.5.2 Core Oligosaccharide	21
1.5.3 O-antigen polysaccharides	23
1.6 LPS and elucidation of the host immune system	25
1.6.1 Innate and adaptive immunity: an overview	25
1.6.2 Innate immunity and LPS: Toll-like Receptor 4	25
1.6.3 Structure and activation of the TLR4/MD-2 receptor complex	28
1.6.4 Agonist ligand binding to TLR4/MD-2 complex	30
1.6.5 Antagonist ligand binding to TLR4/MD-2 complex	30
1.6.6 Innate immunity and LPS: TLR2 and lectins	33
1.7 Capsular polysaccharides and immunological functions	36
1.8 Objectives	37
1.9 References	38
Extraction, purification, and structural characterisation of bacterial glycoconjugates	45
2.1 Introduction	46

2.2	<i>Preparation of bacterial cells prior to LPS extraction</i>	46
2.3	<i>Extraction and purification of LPS</i>	48
2.3.1	Extraction of LPS	48
2.3.2	Purification of LPS	51
2.4	<i>Structural characterisation of the saccharide moieties</i>	52
2.4.1	Structural characterisation of the saccharide moieties using Gas Chromatography Mass Spectroscopy (GC-MS)	52
2.4.2	Chemical degradation techniques	57
2.4.3	Structural characterisation of the saccharide moieties using NMR Spectroscopy	59
2.4.4	Structural characterisation of the saccharide moieties using MALDI TOF mass spectrometry	63
2.5	<i>Structural determination of the lipid A moiety</i>	66
2.6	<i>Computational Techniques in the characterisation of lipid A ligands with TLR4/MD-2 complex</i>	67
2.6.1	Molecular docking	67
2.7	<i>References</i>	70

Pairing *Bacteroides vulgatus* LPS structure with its immunomodulatory effects on human cellular models **73**

3.1	<i>Introduction</i>	74
3.2	<i>Structural characterisation of <i>B. vulgatus</i> LPS</i>	76
3.2.1	Purification and chemical analyses of LPS from <i>B. vulgatus</i>	76
3.2.2	Structural characterisation of the saccharide portion <i>B. vulgatus</i> LPS by NMR	78
3.2.3	Structural characterisation of the lipid A from <i>B. vulgatus</i> LPS, by MALDI-MS	85
3.3	<i>Immunological evaluation of <i>B. vulgatus</i> LPS</i>	88

3.3.1	Activation of peripheral blood monocyte-derived macrophages and of peripheral blood-monocyte-derived dendritic cells by <i>B. vulgatus</i> LPS and lipid A	88
3.3.2	Human peripheral blood monocyte-derived macrophages (MoMs)	89
3.3.3	Peripheral blood-monocyte-derived dendritic cells (MoDC).	90
3.3.4	Activation of bone marrow monocyte-derived macrophages by <i>B. vulgatus</i> LPS and lipid A	91
3.3.5	Activation of TLR4 and TLR2 by LPS _{Bv} and lipid A _{Bv} in HEK293 cell models	92
3.3.6	Competition tests of LPS _{Bv} and lipid A _{Bv} against agonistic LPS in HEK-Blue™ cell model	95
3.4	<i>Discussion and conclusions</i>	96
3.5	<i>References</i>	99

Lipopolysaccharide from Gut-Associated Lymphoid-Tissue (GALT) resident *Alcaligenes faecalis*: complete structure determination and chemical synthesis of its lipid A **101**

4.1	<i>Introduction</i>	102
4.2	<i>Structure determination of the LPS of <i>A. faecalis</i></i>	104
4.2.1	Isolation and purification of <i>A. faecalis</i> R-LPS and S-LPS	104
4.2.2	Structural characterisation of the core OS from <i>A. faecalis</i> LPS	105
4.2.3	Structural characterisation of the O-antigen from <i>A. faecalis</i> S-LPS	109
4.2.4	Structural characterisation of the lipid A from <i>A. faecalis</i> LPS	110
4.3	<i>Chemical synthesis of the lipid A of <i>A. faecalis</i> and immunological properties</i>	114
4.3.1	Chemical synthesis of hexa-acylated <i>A. faecalis</i> lipid A	114
4.4	<i>Discussion and conclusions</i>	117
4.5	<i>References</i>	119

***Bacteroides thetaiotaomicron* rough-type lipopolysaccharide: the chemical structure and the immunological activity** **121**

5.1	<i>Introduction</i>	122
-----	---------------------	-----

5.2	<i>Structural characterisation of B. thetaiotaomicron R-LPS</i>	124
5.2.1	Extraction, purification, and compositional analysis of <i>B. thetaiotaomicron</i> R-LPS	124
5.2.2	Structural characterisation of the lipid A from <i>B. thetaiotaomicron</i> R-LPS	126
5.2.3	Structural characterisation of the core OS by NMR spectroscopy and MALDI-TOF	128
5.2.4	NMR Spectroscopy of the fully deacylated R-LPS of <i>B. thetaiotaomicron</i>	129
5.2.5	NMR on product obtained after mild acid hydrolysis (OS2) of the R-LPS of <i>B. thetaiotaomicron</i>	139
5.2.6	MALDI-TOF MS analysis	142
5.2.7	Structural characterisation by MALDI mass spectrometry of the intact R-LPS	144
5.3	<i>Immunological properties of B. thetaiotaomicron R-LPS</i>	146
5.3.1	Recognition by TLR4/MD-2 complex	146
5.4	<i>Conclusions</i>	150
5.5	<i>References</i>	153

A structural and immunological investigation into the lipopolysaccharides from two isolates of *Veillonella parvula* **156**

6.1	<i>Introduction</i>	157
6.2	<i>Structural investigation of LPS</i>	161
6.2.1	Extraction and purification of LPS from two <i>V. parvula</i> isolates	161
6.2.2	Structural characterisation of lipid A from <i>V. parvula</i>	161
6.2.3	Structural characterisation of LPS from <i>V. parvula</i> isolates: chemical analyses	166
6.2.4	Structural characterisation of LPS from <i>V. parvula</i> isolates: NMR of the core OS	167
6.2.5	Structural characterisation of LPS from <i>V. parvula</i> isolates: NMR of O-antigen and the establishment of the biological repeating unit	174
6.3	<i>Immunological investigation</i>	181

6.3.1	HEK-Blue™-hTLR4 expressing cells: stimulation assays	181
6.3.2	Computational docking studies on lipid A from <i>V. parvula</i>	182
6.3.3	Intestinal isolate lipid A: docking studies	182
6.3.4	Oral isolate lipid A: docking studies	184
6.4	<i>Discussion and future perspectives</i>	190
6.5	<i>References</i>	194
An investigation into the LPS from a clinical isolate of <i>Bacteroides uniformis</i>		198
7.1	<i>Introduction</i>	199
7.2	<i>Structural characterisation of <i>B. uniformis</i> LPS</i>	200
7.2.1	Purification of <i>B. uniformis</i> cells before LPS extraction	200
7.2.2	Extraction and compositional analysis of the R-LPS from <i>B. uniformis</i>	202
7.2.3	Structural investigation of the structure of the core oligosaccharide	205
7.3	<i>Structural characterisation of the lipid A</i>	206
7.4	<i>Conclusions and future perspectives</i>	208
7.5	<i>References</i>	209
Structural characterisation of lipid A from <i>Pandoraea</i> and <i>Echinicola</i> species		211
Section 1: Characterisation of lipid A from <i>Pandoraea pulmonicola</i>		214
8.1	<i>Introduction</i>	214
8.2	<i>Structural characterisation <i>P. pulmonicola</i> lipid A</i>	215
8.2.1	LPS extraction from <i>P. pulmonicola</i> strain RL8228 and compositional analysis of the lipid A	215
8.2.2	MALDI-TOF MS and MS ² analysis on the isolated lipid A from <i>P. pulmonicola</i> strain RL8228	216
8.3	<i>Conclusions</i>	224

Section 2: Unusual lipid A structure and immunoinhibitory activity of LPS from marine bacteria <i>Echinicola pacifica</i> KMM 6172^T and <i>Echinicola vietnamensis</i> KMM 6221^T	226
8.4 <i>Introduction</i>	226
8.5 <i>Structural characterisation of the lipid A from <i>E. pacifica</i> KMM 6172^T and <i>E. vietnamensis</i> KMM 6221^T</i>	227
8.5.1 Isolation and fatty acid compositional analysis of the LPS from <i>E. pacifica</i> KMM 6172 ^T and <i>E. vietnamensis</i> KMM 6221 ^T	227
8.5.2 MALDI-TOF MS and MS/MS analysis on the isolated lipid A from <i>E. pacifica</i> KMM 6172 ^T	229
8.5.3 MALDI MS and MS/MS analysis on the isolated lipid A from <i>E. vietnamensis</i> KMM 6221 ^T	236
8.6 <i>Immunological properties of the LPS from <i>E. pacifica</i> KMM 6172^T and <i>E. vietnamensis</i> KMM 6221^T</i>	238
8.6.1 HEK-Blue™ hTLR4 cell models: activation assay	238
8.6.2 HEK-Blue™ hTLR2 cell models: competition assay	241
8.6.3 HEK-Blue™ hTLR2 cell models	241
8.7 <i>Conclusions</i>	244
8.8 <i>References</i>	246
Materials and Methods	249
9.1 <i>Bacterial strains, growth conditions, and reagents</i>	250
9.2 <i>Extraction and purification of LPS</i>	251
9.2.1 Preliminary manipulations and checks	251
9.3 <i>Extraction of the R-LPS and S-LPS</i>	252
9.4 <i>Purification and isolation of saccharide moieties of the LPS</i>	253
9.4.1 Chapter 3: Purification of the S-LPS and isolation of the core OS and O-antigen from <i>B. vulgatus</i> mpk	253

9.4.2	<i>Chapter 4: Purification of the R-LPS and S-LPS of A. faecalis and the isolation of the core OS and the O-antigen fractions</i>	254
9.4.3	<i>Chapter 5: Purification of the R-LPS and isolation of core oligosaccharide from B. thetaiotaomicron</i>	255
9.4.4	<i>Chapter 6: Isolation of core OS and O-antigen from V. parvula intestinal and oral isolate</i>	256
9.4.5	<i>Chapter 7: Purification of LPS and isolation of core OS from B. uniformis</i>	257
9.4.6	<i>Chapter 8: Isolation of LPS from the two strains of Echinicola</i>	257
9.4.7	<i>Chapter 8: Isolation of LPS from the Pandoraea pulmonicola</i>	257
9.5	<i>Chemical Analyses</i>	258
9.5.1	Acetylated O-methyl glycosides (AMG)	258
9.5.2	Acetylated Alditols (AA)	258
9.5.3	Partially Methylated Alditol Acetates (PMAA)	258
9.5.4	Octyl-glycosides	259
9.5.5	Gas Chromatography Mass Spectrometry	259
9.6	<i>Isolation of lipid A</i>	260
9.6.1	Dephosphorylation of lipid A and derivatization to alditol acetates	260
9.7	<i>Nuclear Magnetic Resonance (NMR)</i>	261
9.8	<i>MALDI-TOF Mass Spectrometry and MS2</i>	263
9.9	<i>HEK-Blue™ hTLR4/CD14/MD-2 cell culture, transfection, and stimulation</i>	264
9.10	<i>Multiple Reaction monitoring (MRM) analysis of HEK-Blue™ hTLR4/CD14/MD-2 cell experiment ¹⁹</i>	265
9.10.1	In solution digestion of HEK-Blue™ cellular extracts following R-LPS stimulation	265
9.10.2	LC-MS/MS analysis by MRM ion mode	265
9.11	<i>Molecular Modelling</i>	267
9.11.1	Structure construction	267
9.11.2	Docking	267
9.12	<i>References</i>	268

CHAPTER 1

Section 1: The gut microbiota

1.1 Overview of Gut Microbiota – The Forgotten Organ

Bacteria have existed for over three billion years and their co-evolution the complex multi-cellular human host has created what is now known as the “**microbiota**”, a unique micro-ecosystem. These microorganisms can colonise open surfaces like that of the skin as well as the inside of specific organs such as the gastrointestinal, respiratory, and urogenital tract.¹ Based on both culture-dependent and -independent techniques, it has been estimated that around 150-400 species reside in the human gut, which are collectively named the “**gut microbiota**” (**GM**). GM is not solely of bacterial origin but also comprises viruses, yeast, fungi, and phages.^{2,3} This complex community of microbes live in a symbiotic relationship with the human host, meaning that they not only thrive within this highly competitive niche but also provide beneficial roles for the host.⁴ The significance of the GM to the human health along with its enormous complexity and “virtual” dimensions, has led to it being considered as an additional organ of the human body: the so-called “forgotten organ”.⁵ Some of the beneficial roles that the GM provide are outlined in **Figure 1.1**. The research activity devoted to the characterisation of the human bacterial community (microbiota or bacteriobiota) and its genome (microbiome) has increased dramatically within recent years thanks to technological advances and the production of more economic method of isolation and analysis. This has resulted in great insights into the composition of the healthy and unhealthy microbiota and their interactions and effects on the human host, with tremendous repercussions for future engineering of new diagnostic techniques, prevention strategies against diseases, and pharmaceutical drug design.

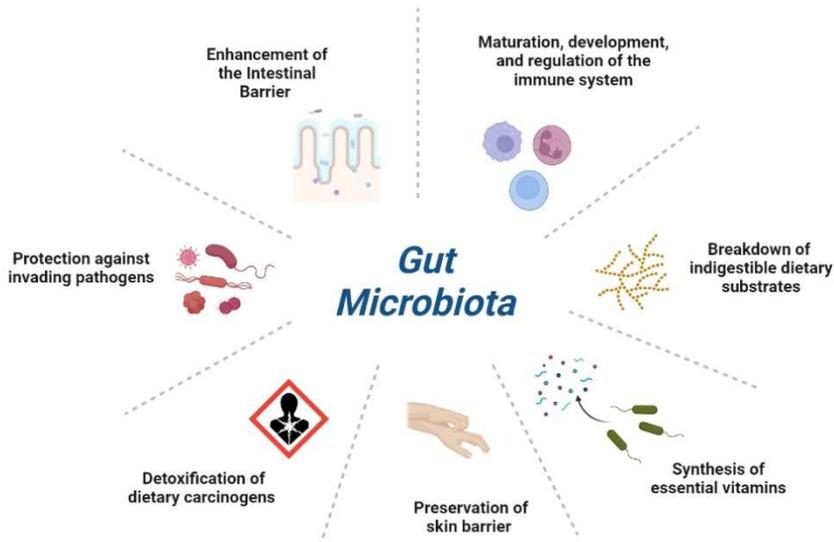


Figure 1.1 Graphic representation of the main beneficial roles of the gut microbiota towards the human host. *Created with BioRender.com*

1.2 History of microbiota research

The term microbiota has been claimed to be first defined by the Nobel laureate-microbiologist Joshua Lederberg in 2001. Yet this terminology was common in the field of microbiology for at least 50 years prior, being used to define a small ecological niche incorporating plant and animal life.⁶ Indeed, the co-existence of microbes colonising humans has been ongoing for thousands of years however, the depth of research in understanding this precious relationship is a contemporary area of research, and thus is quite limited compared to other fields.

Summarised in **Figure 1.2** is an overview of some of the most important scientific milestones achieved in the field of microbiota research. Dating back to 1683 our fascination with bacteria was first documented when Antonie van Leeuwenhoek developed a new handcrafted microscope that enabled him to observe different

bacteria within samples taken from his own and other people's mouths. Fast forward to the 1800s, Joseph Leidy published the book "*A Flora and Fauna within Living Animals*" which has been claimed to be the beginning of microbiota research.⁷ In 1860, the chemist Louis Pasteur laid the foundations of how we understand host-microorganism interactions and initiated the "germ theory" of disease. Intriguingly, in the twentieth century Pasteur, as well as others like Metchnikoff, hypothesised that non-pathogenic microorganisms might have an important role in normal human physiology.^{7,8} In 1890, Koch outlined the four criteria designed to establish a causative relationship between a microorganism and a disease, leading to microbiology research being focused on the identification of causative agents of pathologies. Later, in the twentieth century, huge developments in the field of microbiology and bacterial cultivation allowed for breakthroughs in anaerobic cultivation. This was crucial in providing a way to isolate and classify bacteria and thus tremendously boosted the understanding of bacteria that colonise the human body.⁹⁻¹¹ The exponential growth of sequencing technologies has led to massive advances in human microbiome research from species identification to metagenomic approaches of human-associated microbes, which now can even be used to reveal how microbial metabolic activities correlate with human health and disease.^{12,13}

At present, these techniques continue to develop and improve, with cutting-edge advances in high-throughput molecular strategies, such as 16S ribosomal RNA sequencing and metagenomics. Indeed, many large national and international projects, such as the *Human Microbiome Project* in the United States of America and the European Union Project on metagenomics of the GM called *The MetaHIT Consortium*, represent key milestones in microbiota research.¹⁴ These projects were in fact set up with the goal of generating resources that would enable the comprehensive characterisation of the human microbiome and analysis of its role in human health and disease.^{14,15}

Much of the current research in this field focuses on the role that microbiota and their by-products have on and how they interact with the host immune system. A key example in the field of glycobiology applied to microbiota analysis, is the work by Mazmanian and Kasper (2008) in which they isolated a specific polysaccharide (a capsular polysaccharide) from the symbiotic bacterium *Bacteroides fragilis* and demonstrated that it plays a fundamental role in protecting from inflammatory disease by suppression of pro-inflammatory interleukins release.²⁸ In this frame, modern research is also looking into the microbiome of tumours, defining their distinct composition, and the role that these bacteria and their metabolites play in the response to anti-cancer therapy as well as in worsening the cancer conditions and progression.¹⁶

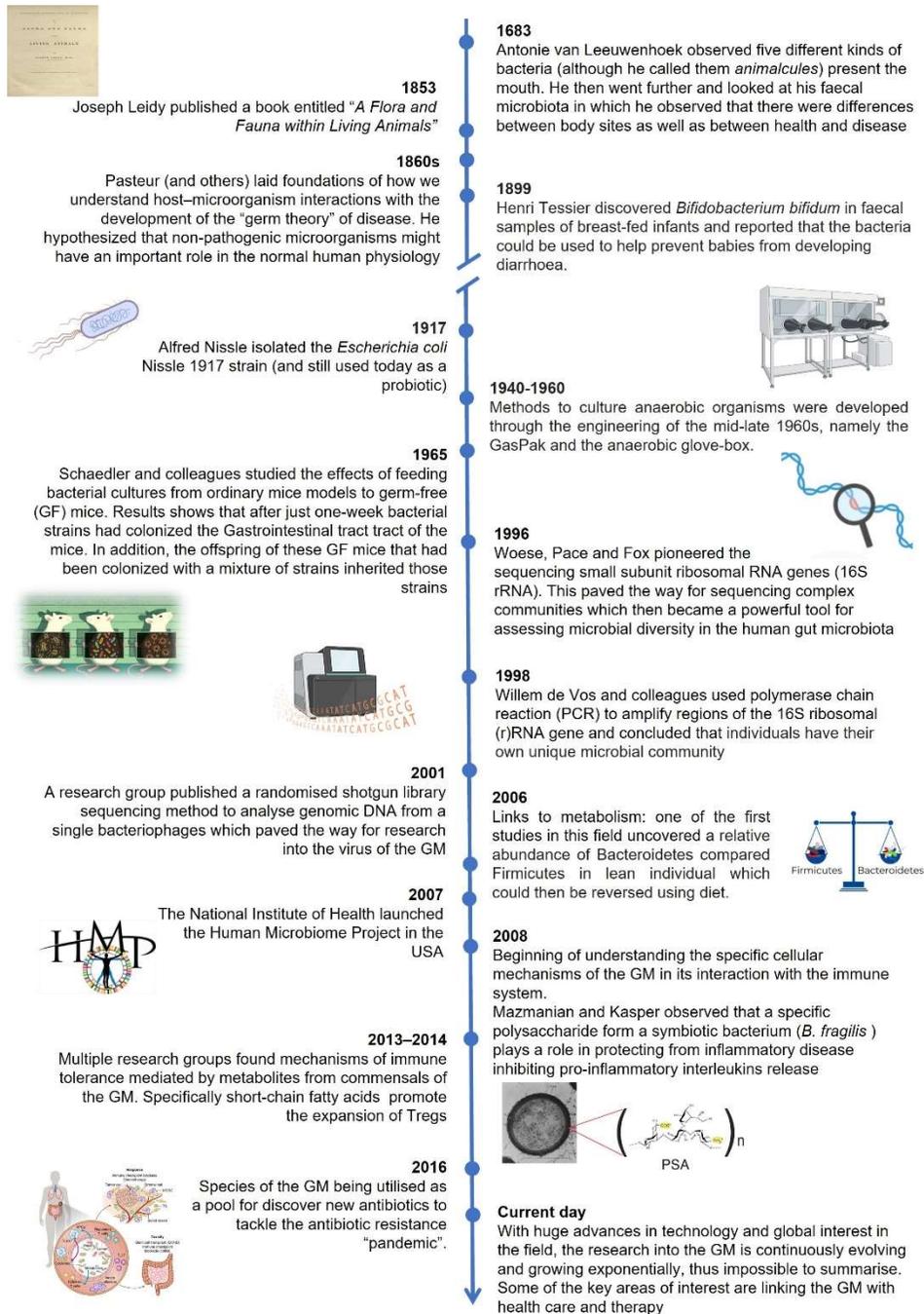


Figure 1.2 Timeline of some of the key milestones within microbiota research. GM: gut microbiota. Created with images from *BioRender.com*

1.3 The Gut Microbiota (GM)

The GM composition is delicately balanced within the intestines, and this homeostasis is vital to be maintained (eubiosis). When this balance is altered, it can lead to a pathological condition known as “dysbiosis”. This dysregulation of the GM composition means functional changes in the microbial transcriptome, proteome and/or metabolome, thus leading also to heightened immune responses by the host and dangerous inflammatory conditions (**Figure 1.3**). Among the main diseases that have been linked to GM dysbiosis are diabetes type 1, inflammatory bowel diseases (IBD), asthma, and even some types of cancer.¹⁷ The investigation of such an imbalance within the GM composition is, however, not an easy feat due to several reasons. First, it should be kept in mind that the relative proportions of each bacterial species vary between individuals and even within the same individual due to GM evolution during lifetime (see *Section 1.3.1*). In addition, the GM can vary greatly across populations due to diet and lifestyle.¹⁸ It is worth mentioning that there are continuously increasing research activities that are devoted to uncovering hallmarks in the GM across populations and the different disease states. These studies will have priceless repercussions for healthcare through manipulating the microbiota of sick patients to ameliorate a specific pathological disorder. In addition, a lot of research efforts are currently given to investigation of whether it is possible to improve human (and animal) health through lifestyle modifications, such as through diet or exercise or using clinical interventions like faecal microbiota transplantation, antibiotics, and the use of pre- and probiotics (**Table 1.1**).³

Table 1.1 Key terminologies used within the microbiota field and their definitions

Term	Definition
Microbiota	The sum of all microorganisms (including bacteria, archaea, eukaryotes and viruses) that reside in and/or on a host. This can also be defined into a specific organ such as the microbiota of the gastrointestinal tract: Gut microbiota
Microbiome	The collection of all genomes (DNA and RNA) of bacteria and other microorganisms in each environment
Prebiotics	Bacterial substrates that are selectively utilised by host conferring a health benefit.
Probiotics	Live microorganisms that when administered in adequate amounts act as a health benefit to the host
Dysbiosis	An imbalance in the structural and/or functional configuration of the microbiota, leading to a disruption of host–microorganism homeostasis

Numerous studies have focused their attention on defining the differences observable within the GM composition of healthy and diseased patients to harness this information for potential therapeutic purposes, both for prebiotics and probiotics (**Table 1.1**). It is important to mention that the cell-surface macromolecules such as proteins (surface layer associated proteins and mucin-binding proteins) and non-protein components (lipoteichoic acid, peptidoglycan, lipopolysaccharides and exopolysaccharides) are the main modulators between the probiotic strains and the host.^{20,21}

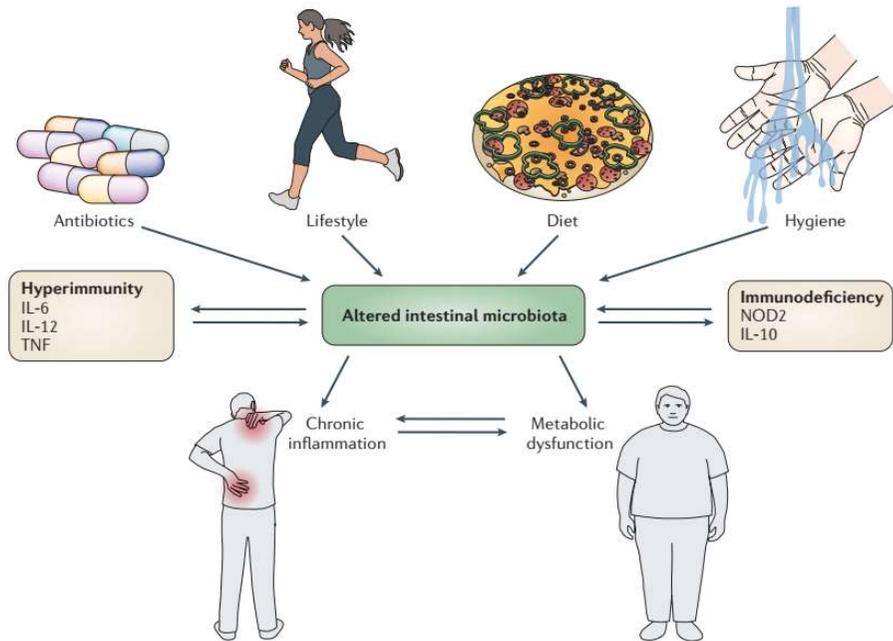


Figure 1.3 Factors that influence and shape the intestinal microbiota include various environmental factors, i.e., the use of antibiotics, lifestyle, diet, hygiene preferences, and the host's genetic disposition, which then in turn have consequences on the health of the host. Figure taken from Sommer and Bäckhed (2013).²²

1.3.1 GM and immune system

The gastrointestinal tract not only harbours microbes of the GM but also a huge pool of immune cells that are important for the regulation and monitoring of the GM.^{23,24} Before microbes encounter the immune system, the epithelial cells of the intestinal tract act as first guardians and communicate with the immune cells within the lamina propria.²³ This regulation and monitoring is mainly performed by the cells of the innate immune system through important innate immunity receptors known as pattern recognition receptors (PRRs), among which there are the Toll-like receptors. However, it is important to note that although the immune system impacts the GM, many publications have proven that the proper maturation and homeostasis of the immune system is significantly linked to microbiota itself. Indeed in 1930, studies performed on Germ-Free (GF) mice proved that mice completely devoid of a microbiota have a vastly underdeveloped immune system compared to that of wild type (WT) animals.⁴ Furthermore, other studies have demonstrated that antibiotic treated-mice have an immature immune system with a lower number of lymphocytes and decreased expression of cytokines.^{24,25}

Although the establishment of the microbiota of the neonate begins at birth, it is well known that the immune system of a foetus starts to develop throughout the pregnancy and is affected by the maternal microbiota.²⁶ Therefore, factors affecting the mother microbiota, such as exposure to antibiotics/drugs, dietary habits, and environment, can already start to create a great influence on a child's microbiota. During pregnancy, cell populations of the innate immune system of the foetus are shaped the most such as monocytes, innate lymphoid cells (ILCs) and neutrophils (**Figure 1.4**).²⁷ At birth the neonate has its first exposure to microbes and indeed, differences have been found between babies born vaginally or through C-section. Neonates born vaginally have a microbiota enriched with *Lactobacillus*, *Bacteroides*, and *Prevotella* spp. which resembles those of the maternal vaginal microbiota, whereas those born through C-section have similarities to the skin or environmental microbiota, such as

Staphylococcus, and *Streptococcus*.²⁶ As the immune system at birth is confronted with many bacteria, the neonate is dependent on maternal protection through passive immunisation *via* breast feeding.²⁷ As the infant grows, the beneficial interplay with the microbiota and the developing immune system aids in protecting against invading pathogens. The microbiota will then evolve over time shaped by various factors such as exposure to antibiotics, breast feeding, genetics, and the environment (**Figure 1.4**). After 2-3 years of age, the microbiota community of a child begins to stabilise and resemble that of an adult and remains almost unchanged throughout lifetime. Several studies have related dysbiosis in infants to the occurrence of chronic pro-inflammatory states like obesity, IBD, or psoriasis.²⁸ Of note, a ground breaking study showed a correlation between the depletion of *Fecalibacterium*, *Lachnospira*, *Rothia*, and *Veillonella* at 3 months of age and the development of asthma. These results have also been confirmed with GF mice that evidenced that the depletion of *Fecalibacterium*, *Lachnospira*, *Veillonella* and *Rothia* species were linked to atopy and wheeze.^{28,29}

Although the early life GM is essential for the maturation of the immune system, the education and priming of the host immune response is a lifelong process. As an individual ages, a decline of diversity and dysbiosis within the GM occurs which can lead to alterations in the functioning of the innate and adaptive system. The case of this dysbiosis can be the result of many factors such as increased use of antibiotics, lack of nutrients, reduced physical activity and alterations in hormones (i.e., menopause).²⁶ In fact, this aging of the microbiota has been termed as “microb-aging” and is characterised by a decrease in commensal bacterial species like *Clostridiales* and *Bifidobacterium*, and an overrepresentation of pathobionts, such as *Enterobacteriaceae*.^{30,31} However, it is still not clearly understood whether the dysbiosis of the aged GM is a cause or an effect of the aging host, but current research is indicating that restoration of the GM homeostasis may provide benefits and even increase longevity.²⁶

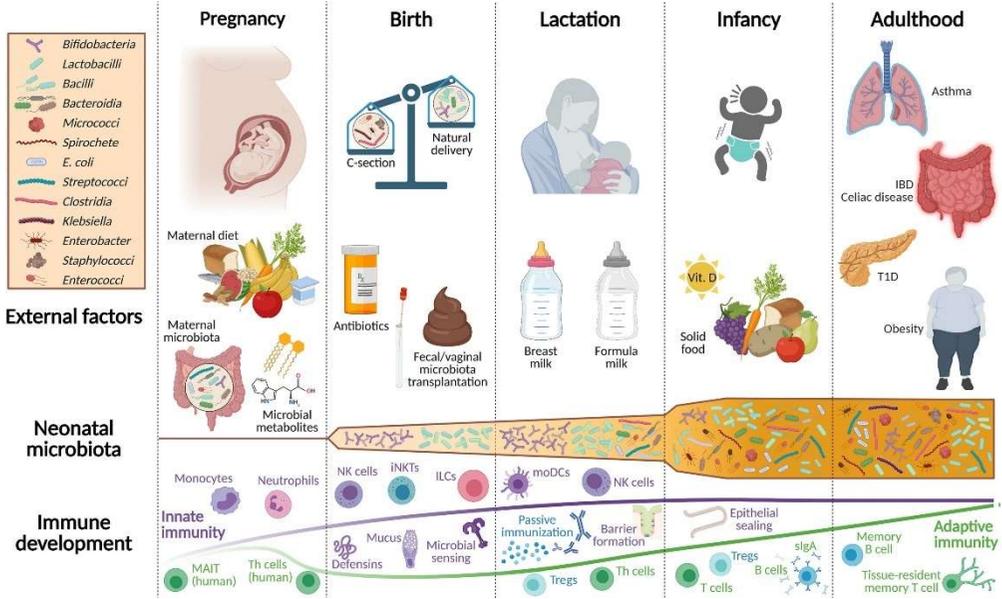


Figure 1.4 A cartoon representation of some of the changes that occur in the GM composition as the human host ages and the development of the immune system. Image taken from Kalbermatter *et al.* (2021)²⁷

1.3.2 Gut Microbiota: the bacterial community

The bacterial component within the GM is a complex consortium that has been classified into distinct phyla, of which the two most represented in the gut are Bacteroidetes and Firmicutes, with minor species belonging to Actinobacteria, Verrucomicrobia, and Proteobacteria phyla.²⁴ Much research has shown important correlations between the ratio of Bacteroidetes and Firmicutes, known as the “F/B ratio”. In this frame, various diseases were shown to be associated with changes in the F/B ratio, including obesity,³² cardiovascular disease³³, and IBD.³⁴

The majority of bacteria within the Firmicutes phylum are Gram-positive, with the exception of two small side branching classes called the Negativicutes and the Halanaerobiales, which consist of species possessing a didermic membrane and lipopolysaccharides, thus are Gram-negatives.^{35,36} The two largest orders of the Firmicutes, Lactobacillales, and Clostridiales, have been extensively studied providing evidence of their importance and beneficial functions for the host. Species from the Clostridiales have been cited as contributing to a breadth of pivotal activities, including resistance to the colonisation by enteric pathogens, protection against food allergy, promoting T cell differentiation, and amelioration of IBD. One key representative is the *F. prausnitzii*, a Gram-positive belonging to Ruminococcaceae family in Clostridium cluster IV. This microbe has been observed to protect from inflammation through blocking NF- κ B activation and IL-8 production and to improve intestinal barrier function.³⁷⁻³⁹

Moreover, species of the Lactorbacillales order have been found to interact with the mucosal immune system, enhancing the intestinal barrier defence by promoting mucus secretion and inhibiting the growth of some detrimental bacteria through production of lactic acid.⁴¹ As a consequence of these benefits, many species from these orders have been predicted to emerge in future microbial therapeutics^{39,41} A key examples is the *Lactobacillus rhamnosus* GG that has been extensively studied and shown to prevent gastrointestinal infections and stimulate immune responses aiding

in preventing allergic symptoms. The underlying molecular mechanisms involved in these functions are still being researched but have shown to entail the participation of exopolysaccharides (EPS) and lipoteichoic acids.⁴²

On the other hand, the Bacteroidetes phylum is composed of Gram-negative species which constitute between 5% and 60% of the intestinal microbiota and is mainly made up of 4 key families: *Bacteroidaceae*, *Prevotellaceae*, *Rikenellaceae*, and *Porphyromonadaceae*. There are conflicting and complex studies that show different functions by bacteria of the Bacteroidetes phyla. Several studies have demonstrated their beneficial roles in breakdown of complex polysaccharides, resistance to pathogens, and production of vitamin B.^{43,44} Yet, some species have been linked to disease states, such as *Bacteroides fragilis* that has been associated with colon cancer,⁴⁵ despite this same bacterium was known to produce, as stated above, an immunomodulatory polysaccharide able to prevent IBD and to positively impact on T-cell development.⁴⁵ These contradictory studies highlight the complexities of the GM research field and strengthen the view that we are still far from the full understanding of the mechanisms underlying the relationship between GM and host.

Section 2: Gram-negatives of the gut microbiota and their cell surface glycans

1.4 Gram-negative bacteria: an overview

Bacteria can be classified as Gram-negative or Gram-positive based on a staining technique developed by Hans Christian Gram in 1884.⁴⁶ The staining uses a crystal violet-iodine complex and a safranin counterstain to differentiate bacteria based on the cell envelope structure and chemical composition (**Figure 1.5**). Therefore Gram-positives retain the crystal-violet stain thus appearing purple, whereas Gram-negatives lose the purple-coloured dye after the counterstain and appear as a pale red colour. Both Gram-positives and Gram-negatives have a phospholipid bilayer known as the inner membrane (IM) which encloses the cytoplasm. External to the IM is the periplasmic space, which contains the peptidoglycan that provides rigidity, shape and osmotic strength to bacteria. The peptidoglycan is formed by β -(1 \rightarrow 4)-linked disaccharide chains, alternating *N*-acetylglucosamine (GlcNAc) and *N*-acetylmuramic acid (MurNAc) residues in turn cross linked by short peptide stems composed of alternating L- and D-amino acids.⁴⁷ It is possible to classify peptidoglycan on the basis of the third amino acid residue of the stem peptide; indeed, in Gram-negatives this residue is the unusual meso-diaminopimelic acid (m-DAP) while in Gram-positive bacteria is commonly a lysine.

Gram-negatives have a thin layer of peptidoglycan (reason why they lose the staining upon decolouration step foreseen in Gram staining protocol) and a distinguishing feature of an additional phospholipid bilayer known as the outer membrane (OM) which is responsible for the effective permeable barrier against hydrophobic compounds and high molecular weight hydrophilic compounds. The inner leaflet of the OM is essentially composed of phospholipids and proteins classified into either lipoproteins and β -barrel proteins⁴⁸, whereas the outer leaflet of the OM is rich in lipopolysaccharides (LPS), covering up to 75 % of the whole cell surface.^{47,49,50}

Gram-positives have a much thicker layer of peptidoglycan, as well as long anionic polymers known as teichoic acids (TA), which give stability and protection to the membrane. There are two major types of teichoic acids: wall teichoic acids (WTAs), which are covalently linked to peptidoglycan, and lipoteichoic acids (LTAs), which are linked to the cell membrane by a lipid anchor.⁴⁸ Both Gram-positives and Gram-negatives can also have additional features on the cell envelope, such as capsular polysaccharides (CPS), exopolysaccharides (EPS), flagella, and S-layers that provide multiple beneficial roles for bacteria including structural stability, permeability, motility, and defence against attacks from the host immune system.^{51,52}

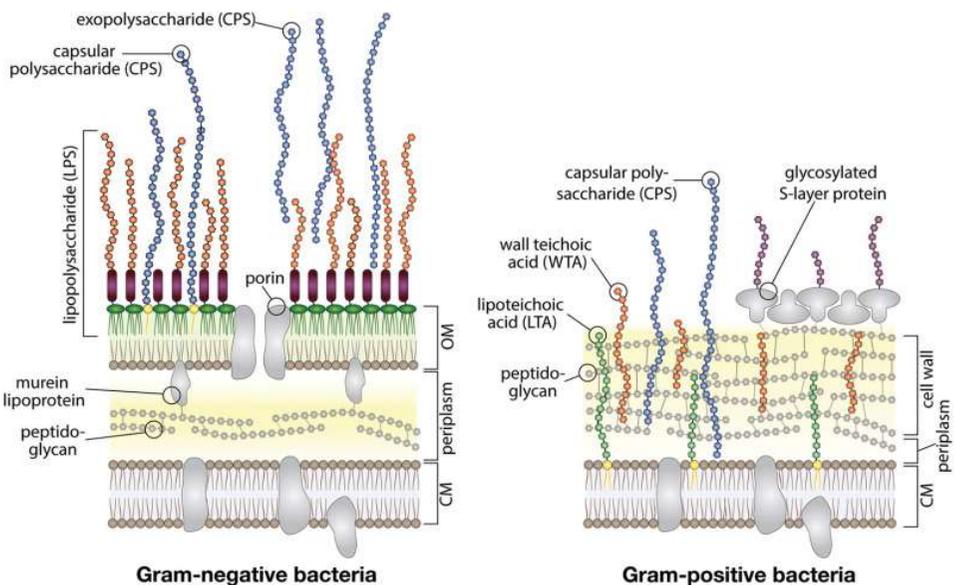


Figure 1.5 A cartoon representation of the Gram-negative and Gram-positive bacteria cell envelope (Image by Whitfield *et al.* 2022).⁵³

1.5 Lipopolysaccharides

Lipopolysaccharides (LPS) are amphiphilic macromolecules essential for the growth, viability and correct assembly of the external membrane of Gram-negatives.⁵² LPS molecules pack tightly together to form a highly structured monolayer held through strong electrostatic interactions between the charged groups of the LPS and the environmental divalent cations (*e.g.* Ca²⁺ and Mg²⁺). This results in an increased membrane permeability to hydrophobic compounds and resistance to external stressors, such as antibiotics and bile salts.^{54,55} Additionally, due to the external location of LPS, they are crucial in host-bacterium interaction events, such as recognition, adhesion, colonisation, virulence, and elicitation of both the animal and plant immune system.⁵⁶ However, among these roles, LPS are notoriously known for being the main cause of “endotoxic” shock associated with the pathogenic Gram-negatives-mediated septicaemia.⁵⁰

LPS are biosynthesised according to a common structural architecture, which entails three genetically, biologically, and chemically distinct moieties 1) the glycolipid anchor known as the **lipid A**, 2) the core oligosaccharide (**Core OS**), and 3) the O-antigen polysaccharide repeating unit (**O-antigen**) (**Figure 1.6**). The presence of the O-antigen defines whether the LPS is a “smooth-type LPS” (S-LPS), *i.e.*, it possesses an O-antigen, or a “rough-type LPS” (R-LPS), which lacks this polysaccharidic portion.⁵⁰ Despite this general architecture, LPS are in fact complex and highly heterogeneous molecules with chemical variations observable in all three moieties. These variations can occur not only between bacterial species but also among strains of the same species. Additionally, chemical modifications have been reported within the same species grown by using different conditions.⁵⁷ Of paramount importance is that these chemical heterogeneities have a profound impact on the cross-talk between LPS and the immune system.⁵⁰

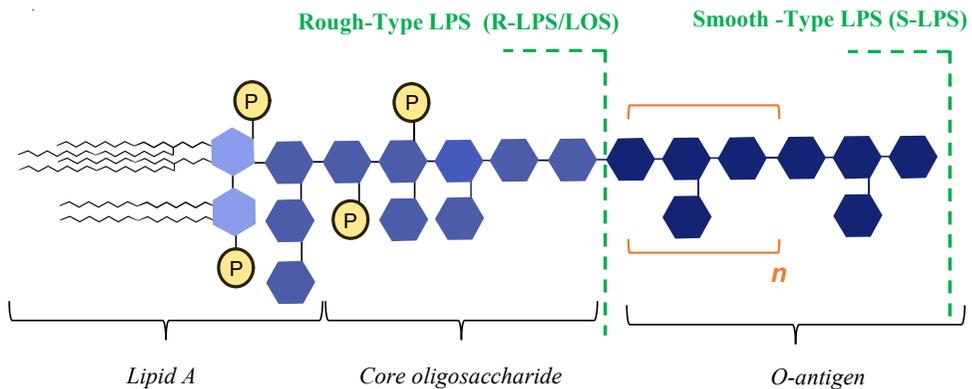


Figure 1.6 Cartoon of the general structure of an LPS. Indicating the two types: smooth-type LPS (S-LPS) and rough type LPS (R-LPS). “*n*” stands for an undefined number of repeating units

1.5.1 Lipid A

The lipid A is the glycolipid moiety of the LPS that anchors the LPS into the OM, and acts as the main ligand for the innate immunity receptor Toll Like Receptor 2/Myeloid Differentiation factor 2 (MD-2) (TLR4/MD-2, discussed in more detail in *Section 1.5*). The structure of the lipid A was first characterised by Westphal and Lüderitz (1954) and comprises of a β -(1 \rightarrow 6) linked glucosamine (GlcN) disaccharide, commonly *bis*-phosphorylated at positions 1 of the α -GlcN_p and 4' of the β -GlcN_p.⁵⁸ Both the glucosamines can be acylated at positions 2,2' and 3,3' with 3-hydroxy fatty acids by both amide and ester linkages.^{58,59} Despite the lipid A being the most structurally conserved moiety, its chemistry widely varies among bacterial species and strains.⁶⁰ Of note, a single bacterial species can even produce a heterogeneous mixture of lipid A structures.¹¹ Indeed, structural variations have been characterised within the acyl chain composition, their distribution and branching pattern, as well as in the phosphorylation degree and presence on the phosphate groups and/or directly on the glucosamine backbone of other carbohydrate and non-carbohydrate substituents.

The lipid A from *Escherichia coli* has the archetypal hexa-acylated bis-phosphorylated structure and is known as the most potent agonist for the activation of immune response mediated by TLR4/MD-2 complex. The lipid A from *E. coli* possesses 14:0 (3-OH) as primary acyl chains, with those on the non-reducing GlcN_p esterified by two secondary acyl chains, that is 12:0 and 14:0, thus resulting in an “asymmetric” 4+2 distribution of the fatty acids with the respect to the disaccharide backbone (**Figure 1.7a**). The composition and distribution of the acyl chains of *E. coli* lipid A allows it to strongly bind into the binding pocket of MD-2 with only five of its six acyl chains, while the sixth one lies on the surface and interacts with TLR4. This interaction is crucial for the proper activation of the TLR4/MD-2 mediated intracellular transduction signalling leading to the elicitation of the immune response (see *Section 1.6.3* for more details).⁶¹

Lipid A molecules having longer or shorter acyl chains have shown to exert a reduced TLR4 immunostimulatory activity.^{55,62,63} Also the hypo-acylation, like in the case of *Bacteroides* sp. which possess tetra- and penta-acylated lipid A species (**Figure 1.7b**), and the “symmetric” distribution of the acyl chains, such as the 3+3 symmetry of *N. meningitidis* (**Figure 1.7c**) lipid A, have been associated with a poor immunostimulatory activity.^{64,65} Differences in the substitution occurring on the phosphate group(s) have also been observed, including the presence of additional phosphate units or sugar residues: an example is the glucuronic acid (D-GlcpA) unit on the lipid A of *Echinicola pacifica* and *E. vietnamiensis*, (**Figure 1.7d**, and *Chapter 8*) or the 2-aminoethanol group (EtN) linked to the phosphate, like that observed on the lipid A of *N. meningitidis* and *P. pulmonicola* (**Figure 1.7b**, *Chapter 8*).^{65,66} Phosphate groups of the lipid A create an overall negative charge of the molecule and aid in the stability and permeability of the bacterial OM.⁶⁰ In rare cases, like in *Campylobacter jejuni*, the usual GlcN residue of the backbone is replaced by a 2,3-diamino-2,3-dideoxy-D-glucose (GlcN3N).⁶⁷

Because of their chemical structure, some lipid A species behave as TLR4/MD-2 antagonists, thus competing for the binding to the receptor with potent agonistic lipid A species of pathogenic bacteria, and therefore are able to impede/reduce the subsequent cascade of detrimental events associated to an uncontrolled inflammatory reaction. The typical antagonist for human TLR4/MD-2 is the biosynthetic precursor of *E. coli* lipid A, the lipid IV_A (**Figure 1.7e**).

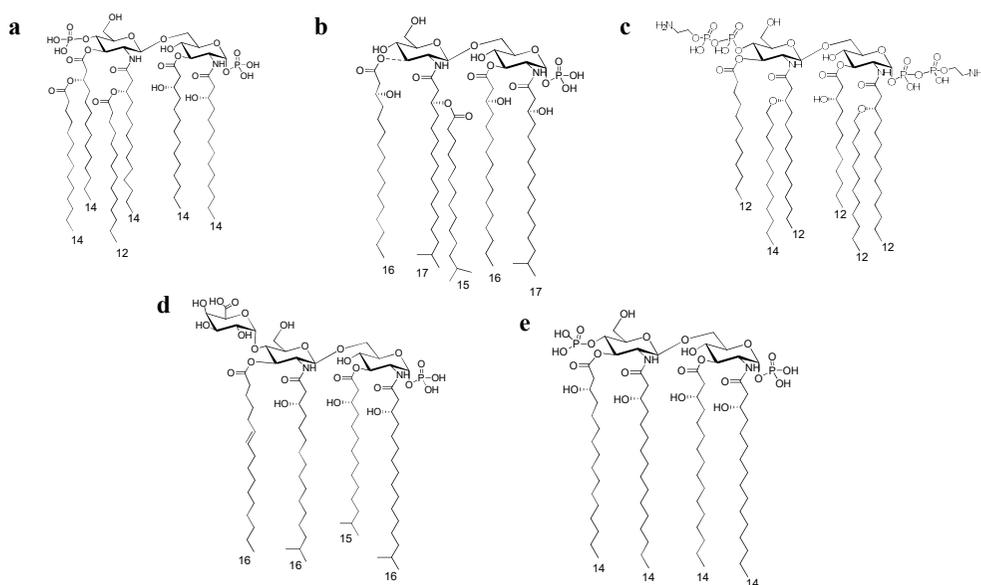


Figure 1.7 Chemical structures lipid A from different bacteria to illustrate heterogeneities that can be observed in the acylation (number, type and distribution of acyl chains), and phosphorylation patterns, and substitutions on the disaccharide backbone **a)** *Escherichia coli* **b)** *Bacteriodes vulgatus*, **c)** *Neisseria meningitidis*, **d)** *Echinicola pacifica* and **e)** lipid IV_A

1.5.2 Core Oligosaccharide

The core oligosaccharide (core OS) is a complex oligosaccharide composed of up to 15 monosaccharides of various nature that can be arranged as either branched or linear moieties. The core OS can be further distinguished into an inner core and an outer core.⁵² The inner core is linked to the lipid A through a well conserved sugar, a 3-

deoxy-D-manno-oct-ulosonic acid (Kdo) (**Figure 1.8a**), linked to position 6 of the non-reducing glucosamine of the lipid A. As the presence of Kdo is almost exclusively found within LPS, it is considered a diagnostic marker for Gram-negative bacteria.⁵⁰ In some unusual cases however, other sugars have been identified in place of Kdo, such as in the case of *Acinetobacter*⁶⁸ where the first sugar of the inner core is a D-glycero-D-talo-oct-2-ulosonic acid (Ko) (**Figure 1.8b**) or within the LPS of *Shewanella*⁶⁹ where there is an 8-amino-8-deoxy-manno-oct-2-ulosonic acid (Kdo8N). Kdo typically carries charged groups, such as phosphate or uronic acids, that often present in a non-stoichiometric manner.^{50,70} The residue commonly found linked to the Kdo is a heptose (L,D-Hepp and D,D-Hepp). However, it has been unveiled that some bacteria lack this sugar, an example being *B. vulgatus* where a rhamnose residue has been found directly linked at position 5' of the Kdo.⁷¹ As previously stated, the lipid A is the portion of the LPS that interacts with TLR4/MD-2. Nevertheless, it has been well documented that the inner core OS also plays a fundamental role in stabilising this interaction.⁷⁰ That said, structural variations within the core, especially substitutions like phosphate groups and amino-sugars, can alter the overall charge of the core OS, and thus profoundly affecting the interaction with the receptor. Furthermore, studies looking into mutants of the core region of *E. coli*, among others, demonstrated that phosphorylation of the inner core is critical for the stability and permeability of the OM.⁷² The extent of the structural variability of the inner core is limited within each genus or family, whereas the outer core region has higher structural heterogeneity.⁵⁰ Indeed, the outer core is the most exposed portion for a rough-type LPS and is the bridge between the core and the O-antigen in the smooth-types.

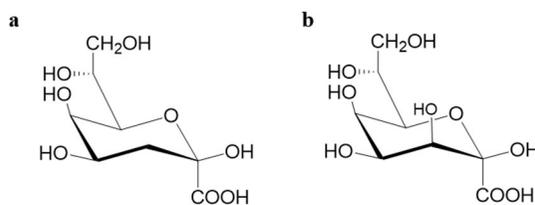


Figure 1.8 Structure of **a**) 3-deoxy-D-manno-octulosonic acid (Kdo) and **b**) D-glycero-D-talo-oct-2-ulosonic acid (Ko).⁵⁰

1.5.3 O-antigen polysaccharides

Gram-negatives expressing a smooth-type LPS have an additional polysaccharide portion covalently linked to the outer core known as the O-antigen polysaccharide. This terminology comes from the Kauffman–White serological classification system developed for *Salmonella* in the 1930.⁵³ The presence of the O-antigen can be visualised using SDS-PAGE and a silver nitrate stain that develops as a “ladder-like” pattern, where each band of the ladder represents a lipid A plus the core plus incrementing the number of O-antigen unit repeats.⁷³ The structural variability of the O-antigen is almost limitless, with structural heterogeneity able to arise from a varying composition and number of monosaccharide residues, branching or linear structures as well as non-stoichiometrically linked sugar and non-sugar residues. Such heterogeneity of O-antigens across bacterial species led them to be used for serological classifications, as in the case of *E. coli* that has more than 170 O-serotypes and *Salmonella* having 46 O-serogroups.^{74,75}

In S-LPS, the O-antigen is the most exposed moiety of the molecule and therefore is at the interface between the bacterium and its environment, thus it plays a vital role in the evasion from the host immune surveillance. Several studies have demonstrated that some Gram-negatives (*i.e.* *Helicobacter pylori*, *Neisseria gonorrhoeae*, *Neisseria meningitidis* and *Haemophilus influenza*) possess O-antigen structures that resemble

those of found in human glycoconjugates. In the case of *H. pylori*, the O-antigen structure mimics that of Lewis blood groups, hence greatly contributing to the invasiveness, colonisation, and antigenicity of the bacteria.⁷⁵ It is also well studied that the chain length of the O-antigen is crucial for the protection of complement system-mediated lysis. An example is represented by *V. cholerae* where a R-LPS significantly reduces its ability to colonise the intestinal epithelia compared to the smooth-type strain, due to an increased sensitivity to complement system.⁷⁶ Furthermore, some O-antigen structures have been characterised as providing protection from environmental stresses such as oxidative stress, bile salts and CAMPs (Cationic antimicrobial peptides), all of which are advantageous for microbes to colonise the intestine.⁷⁷ However, in some cases the O-antigen also contributes to LPS recognition by means of carbohydrate binding proteins, known as lectins, which are other actors of the immune system influencing the elicitation of inflammatory responses. An example is given by the mannose rich O-antigen of *Hafnia alvei* PCM 1223 which is recognised by the dendritic cell-associated C-type lectin-2 (Dectin-2) leading to the augmentation of TLR4 response in myeloid cells.⁷⁸

1.6 LPS and elucidation of the host immune system

1.6.1 Innate and adaptive immunity: an overview

The immune system can be considered a complex collection of organs, cells and proteins that protect the human body from intruders. The immune system is divided into two branches: the innate immune system and the acquired immune system, of which the innate immune system can be viewed as the first line of defence against invading pathogens. The innate immune system includes physical barriers such as the skin and the mucus layer that covers the epithelium of the intestinal, respiratory, and genitourinary tract. It also consists of soluble proteins and small bioactive molecules that are released (complement proteins, defensins, ficolins, cytokines and chemokines) as well as membrane bound receptors and cytoplasmic proteins which use an efficient recognition strategy through the detection of conserved microbial molecular target that are widespread across microbes.^{59,79} These targets are known as MAMPs (Microbe Associated Molecular Patterns) or previously termed as PAMPs (Pathogen Associated Molecular Patterns), which are recognised by the innate immune receptors called PRRs (Patten Recognition Receptors). Upon recognition of MAMPs by PRRs an intracellular signalling cascade is triggered, leading to changes in the expression of receptor proteins and transcription factors, enabling activation of the innate immune response and then supporting the acquired immunity.

The acquired immune responses are slower to develop on first exposure to a new pathogen but are a more sophisticated system of defence, as it creates a memory that enables it to be more effective when re-exposed to the same stimuli. The main components of the acquired system are composed of four major populations of mature lymphocytes: B cells, T cells, natural killer (NK) cells, and NK-T cells.⁷⁹

1.6.2 Innate immunity and LPS: Toll-like Receptor 4

One of the most studied PRRs is represented by the large family of Toll-Like Receptors (TLR). These receptors are transmembrane receptors that are characterised

by an extracellular leucine-rich repeat unit (LRR) domain and an intracellular Toll/IL-1 (TIR) domain. The LPS from Gram-negatives are recognised as a DAMP by a specific TLR: the TLR4. The TLR4 is the only TLR that requires the accessory protein Myeloid Differentiation factor 2 (MD-2), thus forming the TLR4/MD-2 complex. TLR4 is expressed on different types of immune cells and non-immune cells, with differences in function varying on the cell type. For example, TLR4 on hepatocytes is necessary for the uptake of LPS whereas on macrophages TLR4 is essential for enhancing phagocytosis.⁸⁰

LPS is anchored to the OM through the lipid A thus for recognition by the host immune system it must be removed from the bacterial membrane. This is achieved either through cell death leading to the release of the LPS or through a direct extraction mechanism performed by the soluble LPS-binding protein (LBP) (**Figure 1.9**). Important to note here is that LBP is secreted into the intestinal tract where it can sense both pathogenic and harmless (commensal) Gram-negatives, yet the host immune system is able to distinguish between a pathogen and a commensal and is vital in maintaining the GM-host homeostasis. The N-terminal domain of LBP binds to the lipid A moiety while the C-terminal domain recognises another protein called CD14, which consequently is where LBP transfers the LPS molecule to. The CD14 can be either membrane-anchored or soluble and plays an essential role in the sensing of LPS, to the extent that there is an increase by a factor of 10^2 to 10^3 in the susceptibility of the TLR4/MD-2 expressing cell when CD14 is present. The CD14 acts as a chaperone for LPS to the MD-2, which in turn creates the formation of a heterooligomeric protein complex consisting of two TLR4 and two MD2 molecules (TLR4/MD2-TLR4'MD2'). Of note, to form the active TLR4/MD-2 conformation, the ligand must be a pro-inflammatory lipid A species, such as that of *E. coli*.^{61,81} This extracellular dimerisation causes an intracellular dimerisation of the TIR-domains which can cause the activation of two separate immune signalling pathways: 1) the myeloid differentiation primary response gene 88 (MyD88)-dependent pathway and 2) the TIR domain containing adapter-inducing interferon- β (TRIF)/TRIF related

adaptor molecule (TRAM)-dependent pathway. The activation of these pathways leads to activation of the mitogen-activated protein kinases (MAPKs), interferon regulatory factors (IRFs), and the nuclear transcription factor κ B (NF- κ B) that results in the transcription of pro-inflammatory genes. This consequently results in the production of pro-inflammatory cytokines and chemokines such as TNF- α , IL-1 β , IL-6, IL-8, and IL-12. It is noteworthy to mention that the binding of an antagonistic ligand (such as that described for the lipid IV_A or Eritoran) hinders the dimerisation thus preventing the intercellular signalling events to occur and further immune system activation.¹⁵

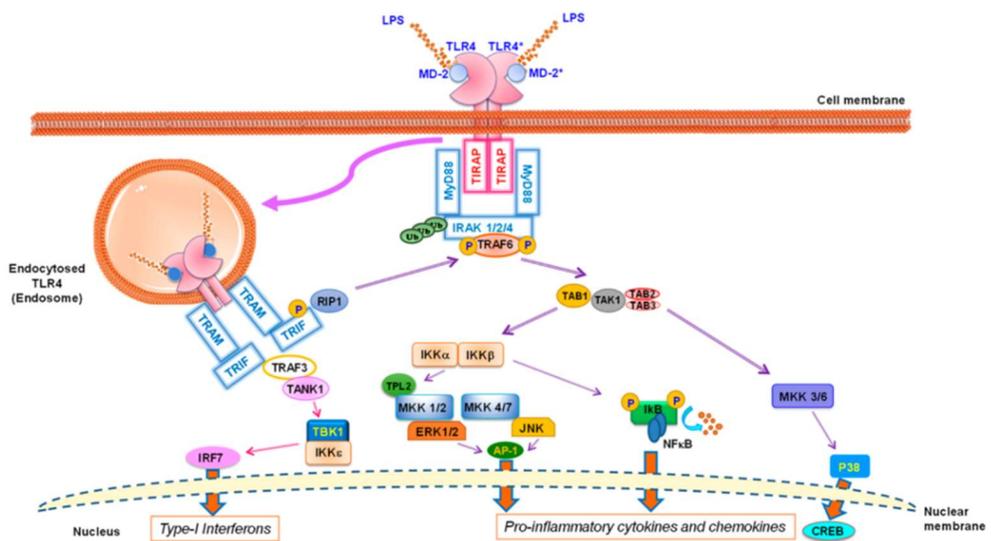


Figure 1.9 Graphical representation of the activation of TLR4-dependent signaling of extracellular LPS. When an agonist LPS binds to the TLR4/MD-2, it causes the heterodimerisation of the complex resulting in the activation of two intracellular signalling pathways: Myd88-dependent pathway and the TRIF-dependent pathway. The former from the cell membrane whereas the later from the endosome. Image taken from Di Lorenzo *et al.* (2021).⁵⁰

1.6.3 Structure and activation of the TLR4/MD-2 receptor complex

As mentioned, the host immune response must be able to distinguish between a pathogenic and a commensal bacterium and therefore it is crucial to highlight that the binding mode of an LPS to the TLR4/MD-2 receptor from pathogenic bacteria compared to commensals LPS must be in some ways different. These differences seen in the binding mode and ability to activate the downstream events originate in the chemical structure of the LPS, or more specifically the lipid A, as the moiety that directly binds the receptor. The molecular structure of the TLR4 can be divided into three domains: the N-terminal domain, the central domain, and the C-terminal domain. The N-terminal and central domains bind the TLR4 with the co-receptor the MD-2, forming the TLR4/MD-2 heterodimer. The molecular structure of the MD-2 is composed of two anti-parallel beta-sheets which form a β -cup-like fold that creates a hydrophobic pocket to encompass the fatty acid (FA) chains of the lipid A.⁶¹ The structure of the overall TLR4/MD-2 complex depends on whether it is bound to an agonist (activated) (**Figure 1.10**) or to an antagonist (inactivated). Indeed, activation leads to the dimerisation of the TLR4/MD-2 heterodimer whereas an antagonistic is unable to confer this dimerisation.^{63,81}

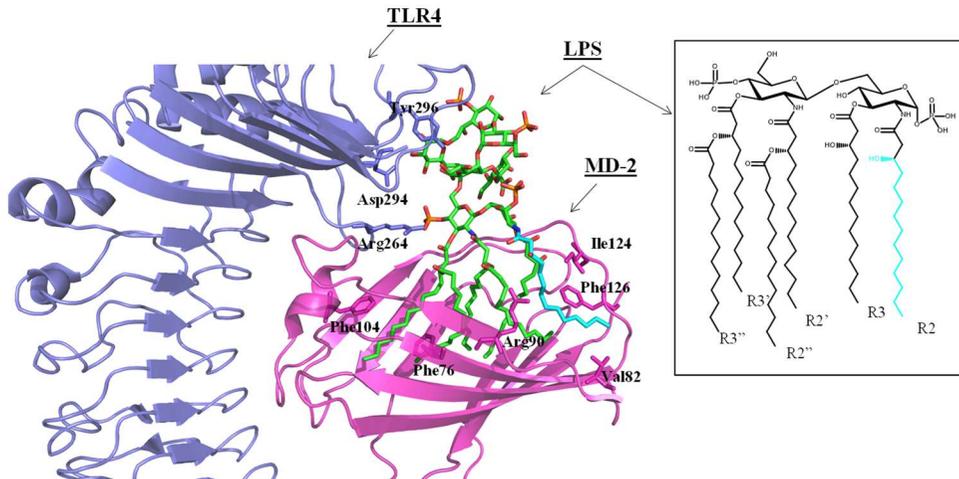


Figure 1.10 Graphical representation of the LPS from *E. coli* (in green) with the TLR4/MD-2 (TLR4 in blue, MD-2 in magenta) from the x-ray crystallographic structure (PDB ID: 3FXI). Sketch of the lipid A structure is shown in the inset. The R2 FA of the lipid A coloured in cyan to indicate how this FA protrudes out of the MD-2 pocket to create the dimerisation of the TLR4/MD-2 complex.

1.6.4 Agonist ligand binding to TLR4/MD-2 complex

With structural information of the TLR4/MD-2 complex with the agonist ligand of the LPS from *E. coli* (EH100) now available on the protein data bank (PDB) (ID: 3FXI), information at an atomic level of the agonistic TLR4/MD-2 complex can be appreciated. This data confirms the binding behaviour of the lipid A of *E. coli* into the MD-2 through accommodating up to 5 FA chains in the hydrophobic pocket. The sixth FA chain (R2 in **Figure 1.10**) is found to be protruding out of the MD-2 hydrophobic pocket, creating a local conformational change involving the Phe126 side chain and the surrounding residues (Ile124 and Leu87). The stabilisation of this binding is aided by interactions between the phosphate groups of the lipid A glucosamine backbone and the polar MD-2 rim (e.g., Arg90, Lys91, Ser118 and Lys122). Furthermore, the inner core of *E. coli* LPS establishes a network of polar interactions with TLR4 amino acids, emphasising the importance of all the structural components of the LPS molecule in the binding to this receptor. All these interactions thus enable the dimerisation and activation of the TLR4/MD-2/TLR4'/MD-2' complex (**Figure 1.11a**).⁸¹

1.6.5 Antagonist ligand binding to TLR4/MD-2 complex

Several computational studies have been performed to clarify the binding mode of different lipid A structures and synthetic compounds with TLR4/MD-2. The antagonist ligand lipid IV_A has its' X-ray crystallographic structure interacting with the MD-2 molecule are available on the Protein Data Bank (ID: 2E56) thus providing great insights in the difference at a molecular level when this receptor interacts with an antagonistic ligand. The most obvious difference between the structures from agonists and antagonists is in the 180-degree rotation of the ligand in the binding site. The binding of lipid IV_A to the receptor creates a re-orientation of the “on/off” switch loop containing Phe126, causing destruction in the side chains in proximity thus hindering the dimerisation of the receptor (**Figure 1.11b**).⁸¹ In this context, molecular modelling, docking and MD simulations studies have enabled these ligand/receptor

interactions to be more easily studied. Of note, a great deal of work in this sense has provided various relevant contributes leading to promising novel drug design procedures; a key example can be seen in the development of lipid A based vaccine adjuvants. An example can be seen in the case of Eritoran, a synthetic lipid A mimetic and a potent TLR4 antagonist which reached phase III clinical trials as an antiseptic drug. Unfortunately, Eritoran failed to succeed further as it did not meet its primary endpoint of reduction of mortality in patients with severe sepsis in a 28-day period.⁸² Resatorvid (TAK-242) was another antagonist but able to bind to the intracellular domain of TLR4 and therefore blocked the cytoplasmic signalling pathways. Although promising results were seen in relieving symptoms of sepsis it was unable to pass the clinical phase III trials. Moreover, a synthetic glucolipid (FP7) was shown to down regulate several TLR4-dependent pro-inflammatory proteins making it a promising new candidate as a TLR4 antagonist.⁵⁹

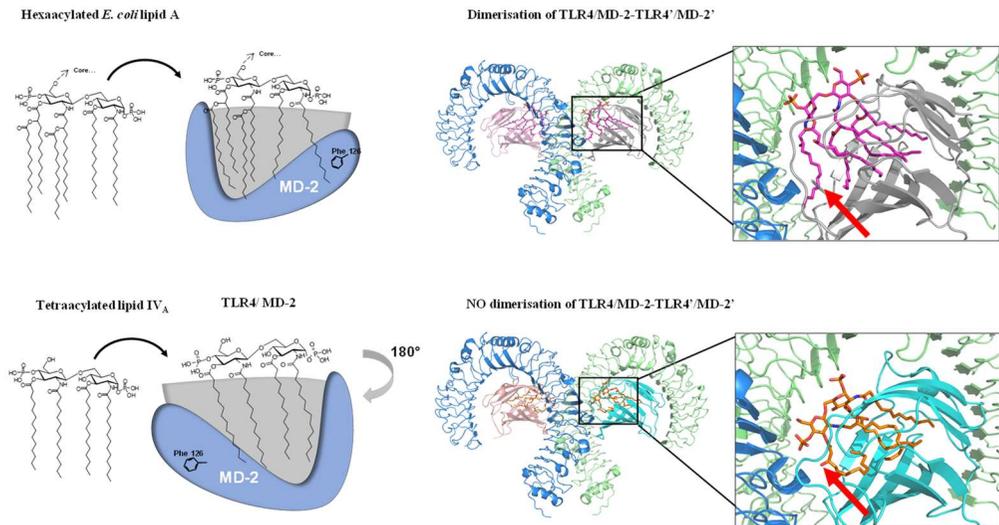


Figure 1.11 Scheme showing the modes of interaction of agonist and antagonist lipid A variants with the TLR4/MD-2 complex. **a)** Lipid A from *E. coli* as an example of an agonistic ligand; when bound, the R2 of *E. coli* Lipid A is exposed to the surface and creates hydrophobic interactions with the Phe126 loop. Additional stabilising interactions are also made, leading to the dimerisation of the TLR4·MD-2-ligand complex. **b)** Lipid IV_A as an example of an antagonistic ligand. The orientation of the ligand is inverted by 180° and the loop containing the Phe126 points outward, which prevents dimerisation of the TLR4·MD-2-ligand complex. PDB ID: 3FXI and 2E56

1.6.6 Innate immunity and LPS: TLR2 and lectins

It has recently come to light that LPS from some bacterial species are also able to interact with TLR2, another receptor from the Toll-like receptor family. TLR2 mediated signalling occurs via heterodimerisation with either TLR1 or TLR6, of which the heterodimers are thought to be pre-formed on the cell surface (TLR2/TLR1 is shown in **Figure 1.12**).^{83,84} These different heterodimers are known to trigger specific immune responses, both pro and anti-inflammatory. It has been hypothesised that this diversity in immune response is due to time-dependent structure-activity relationship between the TLR and the ligand. The LPS from *L. interrogans*, *L. pneumophila*, and *Rhizobium spp.* have been reported to induce TLR2-mediated inflammatory responses in immune cells.⁸⁵⁻⁸⁷ Other species such as the LPS from *B. vulgatus* (as presented in *Chapter 3* of this thesis) showed a synergistic effect between TLR4 and TLR2 when studied in HEK cell lines.⁷¹ Furthermore, LPS from *O. intermedium* has shown a weak activation of TLR4 yet when investigated in a preassembled receptor complex TLR4/TLR2/MD-2, a high release of pro-inflammatory cytokines was observed. This suggested that also this LPS requires TLR4 together with TLR2 to activate the innate immune response.⁸⁵ Advancing our understanding of how LPS and TLR2 interact, plus its biological role needs to be addressed. This key information could lead to the development of new pharmacological strategies against diseases that are caused or exacerbated by TLR-mediated inflammation.

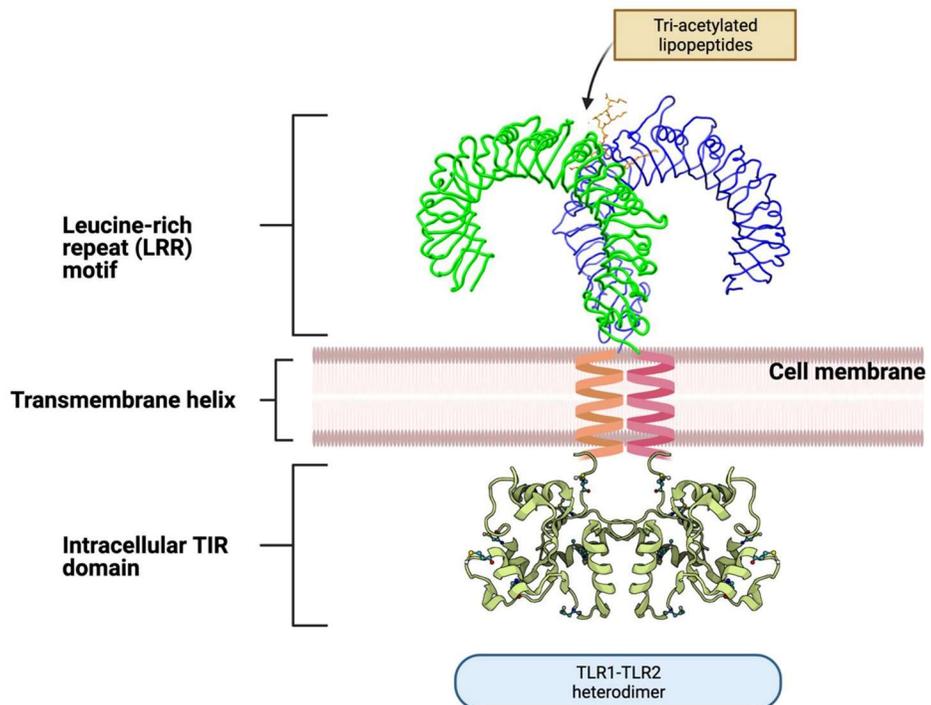


Figure 1.12 Schematic representation of the TLR2/TLR1 heterodimer. Labelled at the main structural elements of the leucine-rich repeat (LRR) motif, transmembrane helix and intracellular TIR domain. Images taken from Jastrzab *et al.* (2021).⁸⁴ The LRR structure is based on the interaction of the receptor with Tri-acetylated lipopeptides, one of the key ligands studied and recognised by this receptor.

In addition to TLR, lectins are a major class of PRR involved in glycan recognition and activation of the immune system. Lectins are carbohydrate binding proteins largely subdivided into three classes: galectins which bind to galactose containing substrates, Siglecs which are specific for recognition of sialic acid containing glycans, and C-type lectins (CLRs) which are known for their need for Ca^{2+} for glycan binding (Figure 1.13).⁸⁸ Many C-type lectins are expressed on antigen presenting cells where they act as PRRs and recognise specific MAMPs on the surface of bacteria. Following this recognition event, internalisation and presentation via the major histocompatibility complex (MHC) of the bacterium occurs, hence linking the innate immune system to the acquired system.⁸⁹ Indeed, the triggered immune outcome differs depending on the respective lectin and/or bacteria species, as highlighted in Figure 1.13.

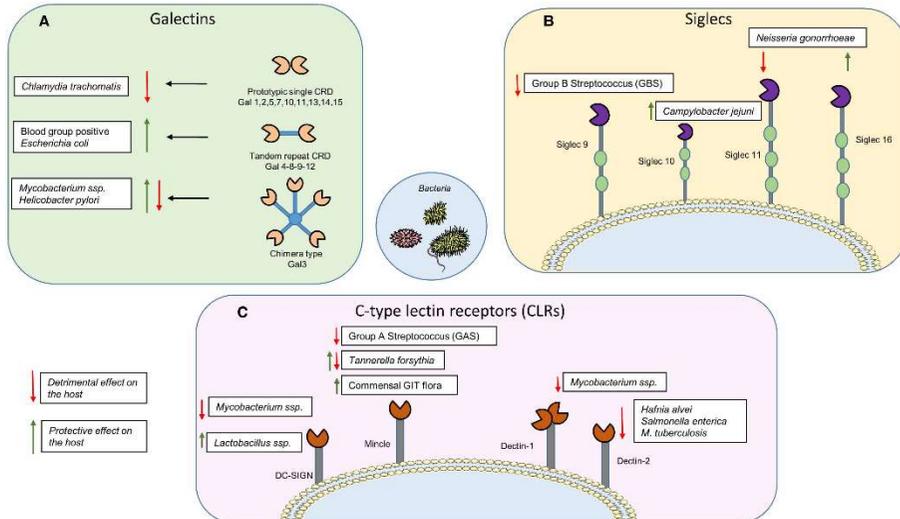


Figure 1.13 Graphical representation of different families of lectins found within the immune system. Siglecs a subgroup of I-type lectins that carry a sialic acid recognition domain, C-type lectins that recognise glycans in a calcium dependent manner and galectins. Image from Acosta *et al.* (2019).⁸⁸

1.7 Capsular polysaccharides and immunological functions

Both Gram-negative and Gram-positive bacteria can also synthesise capsular polysaccharides (CPS). These are long chain polysaccharides tightly attached to the surface of bacteria either via covalent bond or through a linkage with a phospholipid. The composite polysaccharide repeating units of CPS are typically negatively charged and create an extensive hydrated layer surrounding the cell which can be viewed under microscopy.⁹⁰ CPS specific composition and structure are impressively diverse and unique across bacterial species but also between strains within a single bacterial species. Bacteria which produce CPS have been associated with increased resistance to complement and anti-microbial attacks.⁴⁴ *B. fragilis* is one of the most studied gut commensal species for its CPSs, with the most studied one called the polysaccharide A (PS A), isolated from strain 9343. Indeed PS A can be seen as the archetypical example of a molecule from a commensal microbe that has immunoregulatory properties (**Figure 1.14**).^{91–93} The chemical structure of PS A displays a high but equal density of both positive and negative charged groups thus rendering this CPS a zwitterionic polysaccharide (ZWP).^{94,95} Although some species have been greatly investigated for their CPS and their potential therapeutic, the discovery and characterisation of other bacterial species CPS is still greatly understudied considering that most gut microbes produce a massive pool of CPS structures with immune-potential roles still to be discovered.

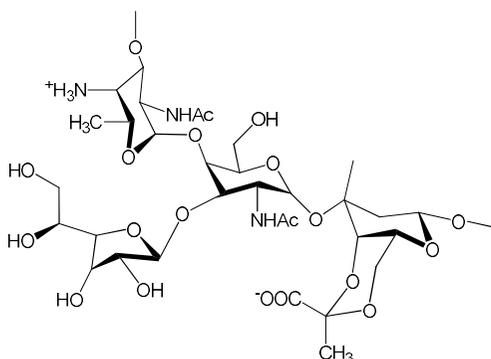


Figure 1.14 Sketch of the capsular polysaccharide (PS A) from *B. fragilis* NCTC9343

1.8 Objectives

The human intestinal tract is colonised by thousands of microorganisms which are collectively known as the gut microbiota (GM) key for the proper functioning of the human body. Gram-negatives make up a huge majority of the bacterial population within the GM, and intrinsically possess lipopolysaccharides (LPS) on their outer membrane. LPS are glycoconjugates universally considered “inflammatory”, as they are sensed by the host’s innate immune system as a warning of a potential treat. Yet, humans live in a harmonious and also a mutualistic relationship with these Gram negatives. With the knowledge that the immunological response is dependent on the chemical structure of the LPS, there must be a difference between dangerous and harmless bacteria. This PhD project thesis aimed at shedding light on this by:

- Defining the chemical structure of LPS from key members of the human GM, thus increasing the knowledge of diversity and complexity of LPS chemistry.
- Investigating the immunological behaviour of these LPS, using both in vitro and computational methods to gain insights into the key molecular features that render these LPS somehow “invisible” and/or “not dangerous” to the host. This unprecedented information on how these bacteria not only persist in our intestine but also whether their LPS could be beneficial for the development of the immunological system.
- Improving and updating existing protocols for the extraction and purification of LPS from GM species to overcome frequent complications, such as the presence of various other carbohydrate molecules such capsular polysaccharides or exopolysaccharides, which can be easily co-extracted with LPS.

1.9 References

1. Jiao, Y., Wu, L., Huntington, N. D. & Zhang, X. Crosstalk Between Gut Microbiota and Innate Immunity and Its Implication in Autoimmune Diseases. *Front. Immunol.* **11**, 282 (2020).
2. Di Lorenzo, F., De Castro, C., Silipo, A. & Molinaro, A. Lipopolysaccharide structures of Gram-negative populations in the gut microbiota and effects on host interactions. *FEMS Microbiol. Rev.* **43**, 257–272 (2019).
3. Quigley, E. M. M., Gajula, P., Van De Wiele, T., Guarner, F. & Val D’hebron, H. Recent advances in modulating the microbiome. *F1000Research 2020 946* **9**, 46 (2020).
4. Surana, N. K. & Kasper, D. L. Deciphering the tête-à-tête between the microbiota and the immune system. *J. Clin. Invest.* **124**, 4197–4203 (2014).
5. O’Hara, A. M. & Shanahan, F. The gut flora as a forgotten organ. *EMBO Rep.* **7**, 688–693 (2006).
6. Prescott, S. L. History of medicine: Origin of the term microbiome and why it matters. *Hum. Microbiome J.* **4**, 24–25 (2017).
7. A field is born. *Nat. Portfolio* (2019). at <<https://www.nature.com/articles/d42859-019-00006-2>>
8. Mackowiak, P. A. Recycling metchnikoff: probiotics, the intestinal microbiome and the quest for long life. *Front. Public Health* **1**, 52 (2013).
9. Aranki, A., Syed, S. A., Kenney, E. B. & Freter, R. Isolation of anaerobic bacteria from human gingiva and mouse cecum by means of a simplified glove box procedure. *Appl. Microbiol.* **17**, 568–576 (1969).
10. Sonnenwirth, A. C. Evolution of anaerobic methodology. *Am. J. Clin. Nutr.* **25**, 1295–1298 (1972).
11. Clark, H. Culturing anaerobes. *Nat. Res. 2021* (2019). at <<https://www.nature.com/articles/d42859-019-00007-1>>
12. Wang, W. L., Xu, S. Y., Ren, Z. G., Tao, L., Jiang, J. W. & Zheng, S. S. Application of metagenomics in the human gut microbiome. *World J. Gastroenterol.* **21**, 803–814 (2015).
13. Tang, L. Sequence-based identification of human-associated microbiota. *Nat. Res. 2021* (2019). at <<https://www.nature.com/articles/d42859-019-00011-5>>
14. Turnbaugh, P. J., Ley, R. E., Hamady, M., Fraser-Liggett, C. M., Knight, R. & Gordon, J. I. The Human Microbiome Project. *Nature* **449**, 804–810 (2007).
15. MetaHIT. *Gut Microbiota Health* (2012). at <<https://www.gutmicrobiotaforhealth.com/metahit/>>
16. Nejman, D., Livyatan, I., Fuks, G., Gavert, N., Zwang, Y., Geller, L. T., Rotter-Maskowitz, A., Weiser, R., Mallel, G., Gigi, E., Meltser, A., Douglas, G. M., Kamer, I., Gopalakrishnan, V., Dadosh, T., Levin-Zaidman, S., Avnet, S., Atlan, T., Cooper, Z. A., Arora, R., Cogdill, A. P., Khan, M. A. W., Ologun, G., Bussi, Y., Weinberger, A., Lotan-Pompan, M., Golani, O., Perry, G., Rokah, M., Bahar-Shany, K., Rozeman, E. A., Blank, C. U., Ronai, A., Shaoul, R., Amit, A.,

- Dorfman, T., Kremer, R., Cohen, Z. R., Harnof, S., Siegal, T., Yehuda-Shnaidman, E., Gal-Yam, E. N., Shapira, H., Baldini, N., Langille, M. G. I., Ben-Nun, A., Kaufman, B., Nissan, A., Golan, T., Dadiani, M., Levanon, K., Bar, J., Yust-Katz, S., Barshack, I., Peeper, D. S., Raz, D. J., Segal, E., Wargo, J. A., Sandbank, J., Shental, N. & Straussman, R. The human tumor microbiome is composed of tumor type-specific intracellular bacteria. *Science* **368**, 973–980 (2020).
17. Zeng, M. Y., Inohara, N. & Nuñez, G. Mechanisms of inflammation-driven bacterial dysbiosis in the gut. *Mucosal Immunol.* *2017 101* **10**, 18–26 (2016).
 18. Reyman, M., van Houten, M. A., van Baarle, D., Bosch, A. A. T. M., Man, W. H., Chu, M. L. J. N., Arp, K., Watson, R. L., Sanders, E. A. M., Fuentes, S. & Bogaert, D. Impact of delivery mode-associated gut microbiota dynamics on health in the first year of life. *Nat. Commun.* **10**, 1–12 (2019).
 19. Hill, C., Guarner, F., Reid, G., Gibson, G., Merenstein, D., Pot, B., Morelli, L., Berni Canani, R., Flint, H., Salminen, S., Calder, P. & Sanders, M. The International Scientific Association for Probiotics and Prebiotics consensus statement on the scope and appropriate use of the term probiotic. *Nat. Rev. Gastroenterol. Hepatol.* **11**, (2014).
 20. Sanders, M. E., Merenstein, D. J., Reid, G., Gibson, G. R. & Rastall, R. A. Probiotics and prebiotics in intestinal health and disease: from biology to the clinic. *Nat. Rev. Gastroenterol. Hepatol.* **16**, 605–616 (2019).
 21. Cani, P. D. & Van Hul, M. Novel opportunities for next-generation probiotics targeting metabolic syndrome. *Curr. Opin. Biotechnol.* **32**, 21–27 (2015).
 22. Sommer, F. & Bäckhed, F. The gut microbiota-masters of host development and physiology. *Nat. Rev. Microbiol.* **11**, 227–238 (2013).
 23. Cani, P. D. Human gut microbiome: hopes, threats and promises. *Gut* **67**, 1716–1725 (2018).
 24. Hill, D. A., Hoffmann, C., Abt, M. C., Du, Y., Kobuley, D., Kirn, T. J., Bushman, F. D. & Artis, D. Metagenomic analyses reveal antibiotic-induced temporal and spatial changes in intestinal microbiota with associated alterations in immune cell homeostasis. *Mucosal Immunol.* **3**, 148–158 (2010).
 25. Gordon, H. A. Morphological and physiological characterization of germfree life. *Ann. N. Y. Acad. Sci.* **78**, 208–220 (1959).
 26. Nagpal, R., Mainali, R., Ahmadi, S., Wang, S., Singh, R., Kavanagh, K., Kitzman, D. W., Kushugulova, A., Marotta, F. & Yadav, H. Gut microbiome and aging: Physiological and mechanistic insights. *Nutr. Healthy Aging* **4**, 267–285
 27. Kalbermatter, C., Fernandez Trigo, N., Christensen, S. & Ganai-Vonarburg, S. C. Maternal Microbiota, Early Life Colonization and Breast Milk Drive Immune Development in the Newborn. *Front. Immunol.* **12**, (2021).
 28. Tamburini, S., Shen, N., Wu, H. C. & Clemente, J. C. The microbiome in early life: implications for health outcomes. *Nat. Med.* **22**, 713–722 (2016).
 29. Arrieta, M.-C., Stiemsma, L. T., Dimitriou, P. A., Thorson, L., Russell, S., Yurist-Doutsch, S., Kuzeljevic, B., Gold, M. J., Britton, H. M., Lefebvre, D. L.,

- Subbarao, P., Mandhane, P., Becker, A., McNagny, K. M., Sears, M. R., Kollmann, T., Mohn, W. W., Turvey, S. E., Brett Finlay, B. & Finlay, B. B. Early infancy microbial and metabolic alterations affect risk of childhood asthma. *Sci. Transl. Med.* **7**, 307ra152-307ra152 (2015).
30. Vemuri, R., Gundamaraju, R., Shastri, M. D., Shukla, S. D., Kalpurath, K., Ball, M., Tristram, S., Shankar, E. M., Ahuja, K. & Eri, R. Gut Microbial Changes, Interactions, and Their Implications on Human Lifecycle: An Ageing Perspective. *BioMed Res. Int.* **2018**, e4178607 (2018).
 31. Bosco, N. & Noti, M. The aging gut microbiome and its impact on host immunity. *Genes Immun.* **22**, 289–303 (2021).
 32. Grigor'eva, I. N. Gallstone Disease, Obesity and the Firmicutes/Bacteroidetes Ratio as a Possible Biomarker of Gut Dysbiosis. *J. Pers. Med.* **2021 Vol 11 Page 13** **11**, 13 (2020).
 33. Spychala, M. S., Venna, V. R., Jandzinski, M., Doran, S. J., Durgan, D. J., Ganesh, B. P., Ajami, N. J., Putluri, N., Graf, J., Bryan, R. M. & McCullough, L. D. Age-related changes in the gut microbiota influence systemic inflammation and stroke outcome. *Ann. Neurol.* **84**, 23–36 (2018).
 34. Duan, R., Zhu, S., Wang, B. & Duan, L. Alterations of Gut Microbiota in Patients With Irritable Bowel Syndrome Based on 16S rRNA-Targeted Sequencing: A Systematic Review. *Clin. Transl. Gastroenterol.* **10**, (2019).
 35. Taib, N., Megrian, D., Witwinowski, J., Adam, P., Poppleton, D., Borrel, G., Beloin, C. & Gribaldo, S. Genome-wide analysis of the Firmicutes illuminates the diderm/monoderm transition. *Nat. Ecol. Evol.* **4**, 1661-1672 (2020)
 36. Sorbara, M. T. & Pamer, E. G. Microbiome-based therapeutics. *Nat. Rev. Microbiol.* **20**, 365–380 (2022).
 37. Carlsson, A. H., Yakymenko, O., Olivier, I., Håkansson, F., Postma, E., Keita, A. V. & Söderholm, J. D. *Faecalibacterium prausnitzii* supernatant improves intestinal barrier function in mice DSS colitis. *Scand. J. Gastroenterol.* **48**, 1136–1144 (2013).
 38. Dikeocha, I. J., Al-Kabsi, A. M., Chiu, H.-T. & Alshawsh, M. A. *Faecalibacterium prausnitzii* Ameliorates Colorectal Tumorigenesis and Suppresses Proliferation of HCT116 Colorectal Cancer Cells. *Biomedicines* **10**, 1128 (2022).
 39. Guo, P., Zhang, K., Ma, X. & He, P. Clostridium species as probiotics: Potentials and challenges. *J. Anim. Sci. Biotechnol.* **11**, 1–10 (2020).
 40. Maioli, T. U., Borrás-Nogues, E., Torres, L., Barbosa, S. C., Martins, V. D., Langella, P., Azevedo, V. A. & Chatel, J.-M. Possible Benefits of *Faecalibacterium prausnitzii* for Obesity-Associated Gut Disorders. *Front. Pharmacol.* **12**, 740636 (2021).
 41. Dempsey, E. & Corr, S. C. *Lactobacillus* spp. for Gastrointestinal Health: Current and Future Perspectives. *Front. Immunol.* **13**, 1534 (2022).
 42. Segers, M. E. & Lebeer, S. Towards a better understanding of *Lactobacillus rhamnosus* GG - host interactions. *Microb. Cell Factories* **13**, S7 (2014).

43. Rowland, I., Gibson, G., Heinken, A., Scott, K., Swann, J., Thiele, I. & Tuohy, K. Gut microbiota functions: metabolism of nutrients and other food components. *Eur. J. Nutr.* **57**, 1–24 (2018).
44. Cheng, J., Hu, J., Geng, F. & Nie, S. Bacteroides utilization for dietary polysaccharides and their beneficial effects on gut health. *Food Sci. Hum. Wellness* **11**, 1101–1110 (2022).
45. Mazmanian, S. K., Round, J. L. & Kasper, D. L. A microbial symbiosis factor prevents intestinal inflammatory disease. *Nature* **453**, 620–625 (2008).
46. Coico, R. Gram Staining. *Curr. Protoc. Microbiol.* **00**, A.3C.1-A.3C.2 (2006).
47. Vollmer, W., Blanot, D. & De Pedro, M. A. Peptidoglycan structure and architecture. *FEMS Microbiol. Rev.* **32**, 149–167 (2008).
48. Silhavy, T. J., Kahne, D. & Walker, S. The Bacterial Cell Envelope. *Cold Spring Harb. Perspect. Biol.* **2**, a000414 (2010).
49. Raetz, C. R. H. Biochemistry of endotoxins. *Annu. Rev. Biochem.* **59**, 129–170 (1990).
50. Di Lorenzo, F., Duda, K. A., Lanzetta, R., Silipo, A., De Castro, C. & Molinaro, A. A Journey from Structure to Function of Bacterial Lipopolysaccharides. *Chem. Rev.* **122**, 15767–15821 (2022).
51. Fagan, R. P. & Fairweather, N. F. Biogenesis and functions of bacterial S-layers. *Nat. Rev. Microbiol.* **2014 123** **12**, 211–222 (2014).
52. Silipo, A., De Castro, C., Lanzetta, R., Parrilli, M. & Molinaro, A. in *Prokaryotic Cell Wall Compd.* 133–153 (Springer Berlin Heidelberg, 2010).
53. Whitfield, C., Szymanski, C. M., Lewis, A. L. & Aebi, M. in *Essent. Glycobiol.* (eds. Varki, A., Cummings, R. D., Esko, J. D., Stanley, P., Hart, G. W., Aebi, M., Mohnen, D., Kinoshita, T., Packer, N. H., Prestegard, J. H., Schnaar, R. L. & Seeberger, P. H.) (Cold Spring Harbor Laboratory Press, 2022).
54. Nikaido, H. Molecular Basis of Bacterial Outer Membrane Permeability Revisited. *Microbiol. Mol. Biol. Rev.* **67**, 593–656 (2003).
55. Herrera, C. M., Voss, B. J. & Trent, M. S. Homeoviscous adaptation of the acinetobacter baumannii outer membrane: Alteration of lipooligosaccharide structure during cold stress. *mBio* **12**, (2021).
56. Westphal, O. & Jann, K. Bacterial Lipopolysaccharides Extraction with Phenol-Water and Further Applications of the Procedure. *Methods Carbohydr. Chem.* **5**, 83–91 (1965).
57. Bertani, B. & Ruiz, N. Function and biogenesis of lipopolysaccharides. *EcoSal Plus* **8**, (2018).
58. Molinaro, A., Holst, O., Di Lorenzo, F., Callaghan, M., Nurisso, A., D’Errico, G., Zamyatina, A., Peri, F., Berisio, R., Jerala, R., Jiménez-Barbero, J., Silipo, A. & Martín-Santamaría, S. Chemistry of Lipid A: At the Heart of Innate Immunity. *Chem. - Eur. J.* **21**, 500–519 (2015).
59. Garcia-Vello, P., Di Lorenzo, F., Zucchetta, D., Zamyatina, A., De Castro, C. & Molinaro, A. Lipopolysaccharide lipid A: A promising molecule for new immunity-based therapies and antibiotics. *Pharmacol. Ther.* 107970 (2021).

60. Steimle, A., Autenrieth, I. B. & Frick, J. S. Structure and function: Lipid A modifications in commensals and pathogens. *Int. J. Med. Microbiol.* **306**, 290–301 (2016).
61. Park, B. S., Song, D. H., Kim, H. M., Choi, B. S., Lee, H. & Lee, J. O. The structural basis of lipopolysaccharide recognition by the TLR4–MD-2 complex. *Nat.* **458**, 1191–1195 (2009).
62. Arenas, J., Pupo, E., De Jonge, E., Pérez-Ortega, J., Schaarschmidt, J., Van Der Ley, P. & Tommassen, J. Substrate specificity of the pyrophosphohydrolase LpxH determines the asymmetry of *Bordetella pertussis* lipid A. *J. Biol. Chem.* **294**, 7982–7989 (2019).
63. Park, B. S. & Lee, J. O. Recognition of lipopolysaccharide pattern by TLR4 complexes. *Exp. Mol. Med.* **2013** 4512 **45**, e66–e66 (2013).
64. Kulshin, V. A., Zahringer, U., Lindner, B., Frasch, C. E., Tsai, C. M., Dmitriev, B. A. & Rietschel, E. T. Structural characterization of the lipid A component of pathogenic *Neisseria meningitidis*. *J. Bacteriol.* **174**, 1793–1800 (1992).
65. Di Lorenzo, F., Billod, J. M., Martín-Santamaría, S., Silipo, A. & Molinaro, A. Gram-Negative Extremophile Lipopolysaccharides: Promising Source of Inspiration for a New Generation of Endotoxin Antagonists. *Eur. J. Org. Chem.* **2017**, 4055–4073 (2017).
66. Pither, M. D., Mantova, G., Scaglione, E., Pagliuca, C., Colicchio, R., Vitiello, M., Chernikov, O. V., Hua, K.-F., Kokoulin, M. S., Silipo, A., Salvatore, P., Molinaro, A. & Di Lorenzo, F. The Unusual Lipid A Structure and Immunoinhibitory Activity of LPS from Marine Bacteria *Echinicola pacifica* KMM 6172^T and *Echinicola vietnamensis* KMM 6221^T. *Microorganisms* **9**, 2552 (2021).
67. Moran, A. P. Biological and serological characterization of *Campylobacter jejuni* lipopolysaccharides with deviating core and lipid A structures. *FEMS Immunol. Med. Microbiol.* **11**, 121–130 (1995).
68. Vinogradov, E. V., Müller-Loennies, S., Petersen, B. O., Meshkov, S., Thomas-Oates, J. E., Holst, O. & Brade, H. Structural Investigation of the Lipopolysaccharide from *Acinetobacter Haemolyticus* Strain NCTC 10305 (ATCC 17906, DNA Group 4). *Eur. J. Biochem.* **247**, 82–90 (1997).
69. Vinogradov, E., Korenevsky, A. & Beveridge, T. J. The structure of the core region of the lipopolysaccharide from *Shewanella* algae BrY, containing 8-amino-3,8-dideoxy-d-manno-oct-2-ulosonic acid. *Carbohydr. Res.* **339**, 737–740 (2004).
70. Cochet, F. & Peri, F. The Role of Carbohydrates in the Lipopolysaccharide (LPS)/Toll-Like Receptor 4 (TLR4) Signalling. *Int. J. Mol. Sci.* **18**, 2318 (2017).
71. Di Lorenzo, F., Pither, M. D., Martufi, M., Scarinci, I., Guzmán-Caldentey, J., Łakomic, E., Jachymek, W., Bruijns, S. C. M., Santamaría, S. M., Frick, J.-S., van Kooyk, Y., Chiodo, F., Silipo, A., Bernardini, M. L. & Molinaro, A. Pairing *Bacteroides vulgatus* LPS Structure with Its Immunomodulatory Effects on Human Cellular Models. *ACS Cent. Sci.* **6**, 1602–1616 (2020).

72. Klein, G., Müller-Loennies, S., Lindner, B., Kobylak, N., Brade, H. & Raina, S. Molecular and Structural Basis of Inner Core Lipopolysaccharide Alterations in *Escherichia coli*: Incorporation Of Glucuronic Acid And Phosphoethanolamine In The Heptose Region. *J. Biol. Chem.* **288**, 8111 (2013).
73. Kittelberger, R. & Hilbink, F. Sensitive silver-staining detection of bacterial lipopolysaccharides in polyacrylamide gels. *J. Biochem. Biophys. Methods* **26**, 81–86 (1993).
74. Liu, B., Knirel, Y. A., Feng, L., Perepelov, A. V., Senchenkova, S. N., Reeves, P. R. & Wang, L. Structural diversity in *Salmonella* O antigens and its genetic basis. *FEMS Microbiol. Rev.* **38**, 56–89 (2014).
75. De Jong, H., Wösten, M. M. S. M. & Wennekes, T. Sweet impersonators: Molecular mimicry of host glycans by bacteria. *Glycobiology* **32**, 11–22 (2022).
76. Nesper, J., Lauriano, C. M., Klose, K. E., Kapfhammer, D., Kraiß, A. & Reidl, J. Characterization of *Vibrio cholerae* O1 El Tor galU and galE Mutants: Influence on Lipopolysaccharide Structure, Colonization, and Biofilm Formation. *Infect. Immun.* **69**, 435 (2001).
77. Whitfield, C., Williams, D. M. & Kelly, S. D. Lipopolysaccharide O-antigens—bacterial glycans made to measure. *J. Biol. Chem.* **295**, 10593–10609 (2020).
78. Wittmann, A., Lamprinaki, D., Bowles, K. M., Katzenellenbogen, E., Knirel, Y. A., Whitfield, C., Nishimura, T., Matsumoto, N., Yamamoto, K., Iwakura, Y., Saijo, S. & Kawasaki, N. Dectin-2 Recognizes Mannosylated O-antigens of Human Opportunistic Pathogens and Augments Lipopolysaccharide Activation of Myeloid Cells. *J. Biol. Chem.* **291**, 17629–17638 (2016).
79. Chaplin, D. D. Overview of the Immune Response. *J. Allergy Clin. Immunol.* **125**, S3-23 (2010).
80. Deng, M., Scott, M. J., Loughran, P., Gibson, G., Sodhi, C., Watkins, S., Hackam, D. & Billiar, T. R. Lipopolysaccharide Clearance, Bacterial Clearance, and Systemic Inflammatory Responses Are Regulated by Cell Type-Specific Functions of TLR4 during Sepsis. *J. Immunol. Author Choice* **190**, 5152–5160 (2013).
81. Billod, J.-M., Lacetera, A., Guzmán-Caldentey, J. & Martín-Santamaría, S. Computational Approaches to Toll-Like Receptor 4 Modulation. *Mol. Basel Switz.* **21**, E994 (2016).
82. Barochia, A., Solomon, S., Cui, X., Natanson, C. & Eichacker, P. Q. Eritoran tetrasodium (E5564) Treatment for Sepsis: Review of Preclinical and Clinical Studies. *Expert Opin. Drug Metab. Toxicol.* **7**, 479 (2011).
83. Oliviera-Nascimento, L., Massari, P. & Wetzler, L. The Role of TLR2 in Infection and Immunity. *Front. Immunol.* **3**, (2012).
84. Jastrzab, R., Graczyk, D. & Siedlecki, P. Molecular and Cellular Mechanisms Influenced by Postbiotics. *Int. J. Mol. Sci.* **22**, 13475 (2021).
85. Francisco, S., Billod, J. M., Merino, J., Punzón, C., Gallego, A., Arranz, A., Martín-Santamaria, S. & Fresno, M. Induction of TLR4/TLR2 Interaction and

- Heterodimer Formation by Low Endotoxic Atypical LPS. *Front. Immunol.* **12**, (2022).
86. Werts, C., Tapping, R. I., Mathison, J. C., Chuang, T. H., Kravchenko, V., Saint Girons, I., Haake, D. A., Godowski, P. J., Hayashi, F., Ozinsky, A., Underhill, D. M., Kirschning, C. J., Wagner, H., Aderem, A., Tobias, P. S. & Ulevitch, R. J. Leptospiral lipopolysaccharide activates cells through a TLR2-dependent mechanism. *Nat. Immunol.* **2001** *24* **2**, 346–352 (2001).
 87. Sun, K. Y., Xu, D. H., Xie, C., Plummer, S., Tang, J., Yang, X. F. & Ji, X. H. *Lactobacillus paracasei* modulates LPS-induced inflammatory cytokine release by monocyte-macrophages via the up-regulation of negative regulators of NF-kappaB signaling in a TLR2-dependent manner. *Cytokine* **92**, 1–11 (2017).
 88. Acosta, M. P. & Lepenies, B. Bacterial glycans and their interactions with lectins in the innate immune system. *Biochem. Soc. Trans.* **47**, 1569–1579 (2019).
 89. Mnich, M. E., van Dalen, R. & van Sorge, N. M. C-Type Lectin Receptors in Host Defense Against Bacterial Pathogens. *Front. Cell. Infect. Microbiol.* **10**, 309 (2020).
 90. Cress, B. F., Englaender, J. A., He, W., Kasper, D., Linhardt, R. J. & Koffas, M. A. G. Masquerading microbial pathogens: Capsular polysaccharides mimic host-tissue molecules. *FEMS Microbiol. Rev.* **38**, 660–697 (2014).
 91. Coyne, M. J., Kalka-Moll, W., Tzianabos, A. O., Kasper, D. L. & Comstock, L. E. *Bacteroides fragilis* NCTC9343 produces at least three distinct capsular polysaccharides: Cloning, characterization, and reassignment of polysaccharide B and C biosynthesis loci. *Infect. Immun.* **68**, 6176–6181 (2000).
 92. Tzianabos, A. O., Pantosti, A., Baumann, H., Brisson, J. R., Jennings, H. J. & Kasper, D. L. The capsular polysaccharide of *Bacteroides fragilis* comprises two ionically linked polysaccharides. *J. Biol. Chem.* **267**, 18230–18235 (1992).
 93. Erturk-Hasdemir, D. & Kasper, D. L. Finding a needle in a haystack: *Bacteroides fragilis* polysaccharide a as the archetypical symbiosis factor. *Ann. N. Y. Acad. Sci.* **1417**, 116–129 (2018).
 94. Knirel, Y. A. & Van Calsteren, M.-R. *Bacterial Exopolysaccharides. Ref. Module Chem. Mol. Sci. Chem. Eng.* **1**, 21-95 (Elsevier Inc., 2020).
 95. Blandford, L. E., Johnston, E. L., Sanderson, J. D., Wade, W. G. & Lax, A. J. Promoter orientation of the immunomodulatory *Bacteroides fragilis* capsular polysaccharide A (PSA) is off in individuals with inflammatory bowel disease (IBD). *Gut Microbes* **10**, 569–577 (2019).

CHAPTER 2

Extraction, purification, and structural characterisation of bacterial glycoconjugates

2.1 Introduction

Most of the studies performed during this PhD were on the characterisation of LPS from gut microbiota (GM) and therefore in this chapter it will be outlined the techniques used to successfully achieve structural determination of

LPS. However, the methods explained for the analysis of the carbohydrate moieties of the LPS after their extraction, have been exploited for the full characterisation of other GM polysaccharides, such as capsular polysaccharides (CPS), although their results have not been reported in this thesis.

2.2 Preparation of bacterial cells prior to LPS extraction

Before the initiation of the LPS characterisation, the isolation and purification of the LPS must first be carried out. The extraction of LPS from pathogenic and environmental microbes is well published^{1,2} but still is understudied how these techniques should be modified to isolate LPS from GM. Hence, additional work throughout this PhD was also to adapt and modify, through a trial-and-error approach, the typical protocols for LPS extraction/purification to the GM LPS world.³ These modifications were mandatory for several reasons. First, gut microbes are often anaerobic (both strict or partial), thus require specific and sometimes very complex growth conditions; this typically implies that only reduced amounts of starting cellular material is obtained, thus resulting in a low yield of LPS. Second, GM species are notorious for expressing various other carbohydrate molecules on their cell surfaces or secreting them in the surrounding environment, such as CPS or exopolysaccharides (EPS) which can be easily co-extracted with LPS; this obviously would create forthcoming problems for the detailed structural characterisation. Certainly, to have some prior knowledge of the bacterial species under investigation and of its capability to produce (and how much/how many) CPS and/or EPS before starting any LPS extraction procedures would reduce these drawbacks. However, by making some few practical tricks, it is possible to protect the quality and the robustness of an LPS

structural analysis. These “tricks” include cell washes, enzymatic digestions, and ultracentrifugation, as described below. These initial steps were all employed during this PhD work but varied depending on the bacterial species analysed, the total amount of starting material, and the chemical properties of the LPS itself. Furthermore, these refined protocols led to a publication of a book chapter on “Extraction, Purification, and Chemical Degradation of LPS from Gut Microbiota Strains” in *Methods of Molecular Biology*.

2.3 Extraction and purification of LPS

2.3.1 Extraction of LPS

There are several conventional procedures to achieve LPS extraction which depend on the nature of the LPS (R-LPS or S-LPS), refer to *Section 1.5*. Two methods are largely used and exploit the chemical properties of the two LPS types, that is the presence of the O-antigen renders the S-LPS more hydrophilic than R-LPS that lacks this carbohydrate portion. Indeed, a method known as the Phenol/Chloroform/Petroleum ether (PCP) procedure, is selective for R-LPS; indeed, after the extraction has been performed, the evaporation of the chloroform and petroleum ether causes the R-LPS to be retained in the phenol phase, which can be then easily precipitated by adding a few drops of water (Example of the gel-like precipitate formed **Figure 2.1a**). The PCP procedure results in a sample that is generally free of cell contaminants allowing for a subsequent easier purification of the extracted material.⁴ Another extraction method is the hot phenol/water procedure that consists of suspending dried cells in 90% phenol/water 1:1 (v/v) at 68 °C.⁵ The LPS is then typically isolated in the water phase. However, based on the intrinsic chemistry of a specific LPS, it is possible in some cases, to isolate LPS also in the phenol phase (Example of the different phases from this method in **Figure 2.1b**). To be sure of isolating the LPS, both water and phenol phases are typically collected and subject to further purification steps including dialysis to remove phenol traces and enzymatic digestion with DNAase, RNAase and proteases to remove nucleic acids and proteins. It should be noted that, although the PCP method is mostly effective on R-LPS, the S-LPS can also be extracted by this approach.⁶

The detection of LPS nature and purity can be achieved through polyacrylamide gel electrophoresis with sodium dodecyl sulphate (SDS-PAGE) as the denaturing agent for the disaggregation of the LPS micelles, which is a necessary step to allow the migration of such big molecules through the gel pores.⁷ A silver nitrate gel staining procedure is then used to reveal the nature of the LPS. The presence of a S-LPS is

determined by the observation of a “ladder-like” pattern on the gel, which is due to the migration of the O-antigen polysaccharide repeating units. The LPS molecules devoid of this polysaccharide part migrate to the bottom of the gel due to their lower molecular weight.⁸ (example given in **Figure 2.1**)

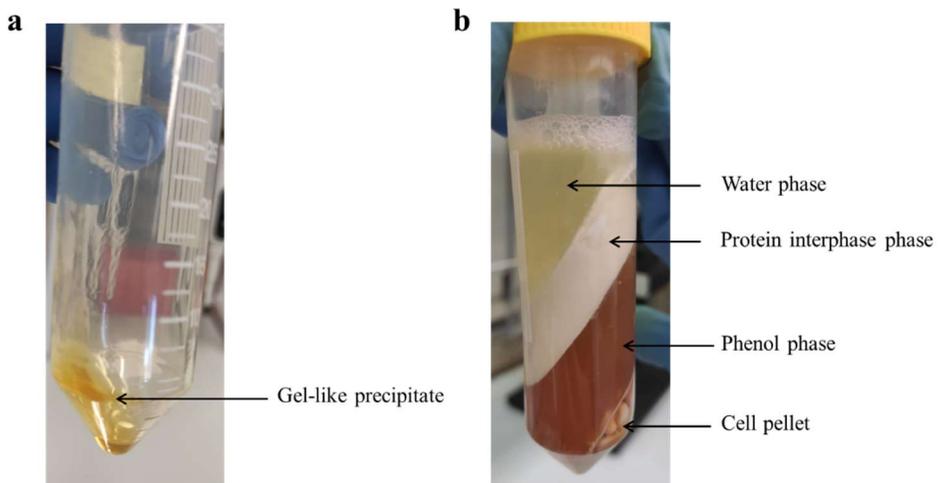


Figure 2.1 Visual examples of both LPS extraction techniques **a)** A gel-like precipitate that contains the LPS, that can form during the PCP extraction method **b)** The result after the centrifugation step during the hot-phenol/water extraction method. This centrifugation step enables the separation of the different phases: a pellet containing the cellular material, a lower phase containing phenol, a white milky interphase containing proteins, and clear supernatant containing water and also typically the LPS.

Other gel staining techniques can be employed to distinguish other molecules possibly co-extracted. One such used frequently is a Coomassie Brilliant Blue staining that enables detection of protein and lipoprotein contaminations. In this frame, a quantitative analysis of protein contamination can also be performed by loading a series of protein concentrations and comparing the staining to that of the sample.^{8,9} Another additional staining method is given using Alcian blue that is able to stain acid groups and therefore often helpful in identifying CPS.¹⁰



Figure 2.2 Example of the two types of LPS on a 13% SDS-PAGE with silver nitrate staining **a)** A smooth-type LPS (S-LPS), a ladder-like pattern is formed due to the presence of the O-antigen. **b)** An example of a rough-type LPS (R-LPS), lacking the O-antigen polysaccharide and thus only a single low band is formed.

2.3.2 Purification of LPS

Multiple purification methods can be applied once the LPS is successfully extracted, including enzymatic digestions (as previously explained), ultracentrifugation, and size-exclusion chromatography (SEC). SEC is an efficient and fast method to purify LPS and is easily performed with various commercially available resins (**Table 2.1**). Indeed, the choice of the resin must be carefully considered in a case-to-case manner depending on the size of the molecules (LPS and/or contaminants) and to maximize the yield. This method separates molecules depending on their hydrodynamic dimensions by passing the sample in a mobile phase (eluent) through a solid phase made of a densely packed porous material (resin). The smaller molecules penetrate the pores and take longer to be eluted compared to larger molecules which can freely diffuse around the pores.¹¹ A detector, such as a refractive index can be used to sense the difference in the refractive index between the chosen mobile phase and the presence of the eluted analyte.³ From this, a chromatogram can be obtained and each fraction can be collected, pooled and then analysed through several techniques, as explained below in *Section 2.4*.

Table 2. Overview of some commercially available resins used for the purification of LPS.

Resin	Matrix	Example of use in LPS purification
Sephacryl (i.e., Sephacryl S-300, GE Life Sciences)	Cross-linked allyl dextran with <i>N,N'</i> -methylenebisacrylamide.	Macromolecule separation: Purification of extracted LPS
Bio-gel (i.e., Bio-Gel ® P-4 (BioRad))	Polyacrylamide	For separation and purification: After chemical treatments have been used for the isolation of saccharide part of LPS.
Sephadex (i.e., Sephadex G-10)	Cross-linked dextran with epichlorohydrin	Often used for desalting after full de-acylation treatment of R-LPS

2.4 Structural characterisation of the saccharide moieties

Achieving the chemical characterisation of carbohydrates comes with its complexities because their primary structure is depending on many parameters. Furthermore, bacterial glycoconjugates like LPS are particularly complex as their structures are highly heterogenous. Variations can be found in composition of the sugar monomers, most of which are widely distributed in nature such as: pentoses, hexoses, hexosamines, 6-deoxyhexoses, and hexuronic acids, yet, bacteria can also synthesise other monosaccharides that are unusual and strain specific.¹² In addition, a polysaccharide structure can also vary in the stereochemistry (D or L) of its monosaccharide constituents, as well as in their ring size (pyranose or furanose) and/or anomeric configuration (α or β). It should also be considered the presence and nature of any noncarbohydrate substituents commonly found in LPS.¹³ All the above structural possibilities clearly highlight how complex can be the structural characterisation of LPS, which requires delicate and accurate chemical analyses integrated with powerful instrumental techniques, such as Nuclear Magnetic Resonance (NMR) and several Mass Spectrometry (MS) techniques.

2.4.1 Structural characterisation of the saccharide moieties using Gas Chromatography Mass Spectroscopy (GC-MS)

For the chemical analyses, which entail the investigation of sugar and lipid content, the sugar linkage pattern, and the absolute configuration, Gas Chromatography Mass Spectrometry (GC-MS) is the leading analytical technique.¹⁴ However, in order to use this instrument and exploit its potential, it is necessary to render both sugars and lipids volatile and therefore they must be “modified” to obtain proper derivatives that can be inspected via GC-MS .

One method typically employed in the qualitative analysis of the carbohydrate residues of an LPS entails the treatment of the poly/oligosaccharide with MeOH/HCl at 85 °C to cleave the glycosidic bonds and form *O*-methyl glycosides. This is followed by acetylation with acetic anhydride in pyridine which produces Acetylated

Methyl Glycosides (AMG) (Figure 2.3).^{1,15} By comparison of the retention times provided by the GC analysis and the fragmentation pattern of the obtained mass spectra, it is possible to identify the type of monosaccharide residues. One drawback to the AMG method is, however, the production of isomers, which cannot be distinguished apart, and thus could lead to miss-quantification.

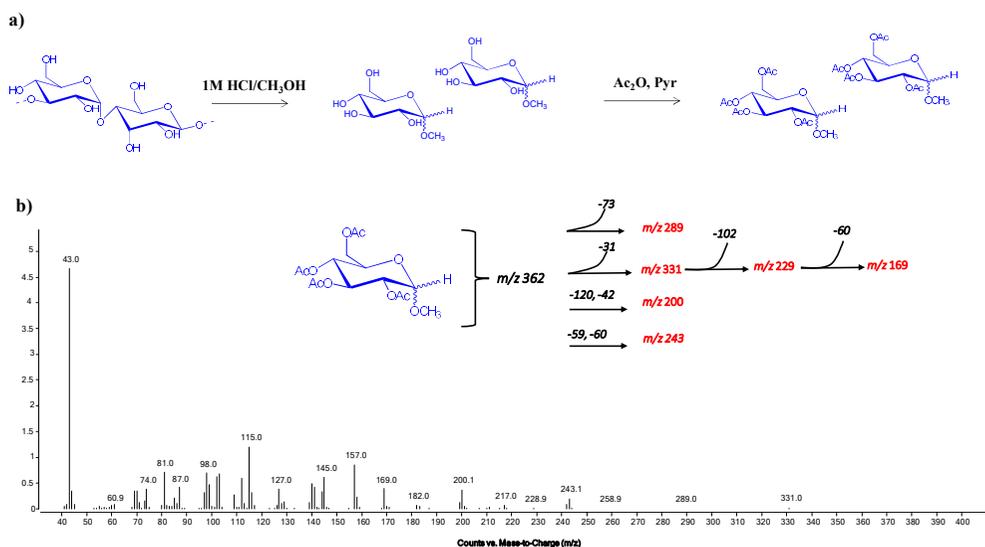


Figure 2.3 a) Scheme of the reactions taking place during derivatisation of a glycan [\rightarrow 4)-D-Glcp(1 \rightarrow] into the corresponding acetylated methylhexopyranose **b)** MS spectrum of a fully acetylated methylhexopyranose and the primary fragments identified.

To overcome this problem, a parallel method that is used entails the derivatisation of sugars into Acetylated Alditols (AA), which is performed using trifluoroacetic acid (TFA) for a strong acid hydrolysis reaction followed by reduction of the carbonyl moiety of the free sugar residues with sodium tetrahydroborate (NaBH_4) (**Figure 2.4**). An advantage of the AA derivatisation method is the production of “only one peak per residue” in the resultant GC spectra (excluding ketoses and aldoses as the reduction step can yield the same alditol)¹⁴, which is an important advantage for quantification analysis. An unfavourable occurrence of this method is the production of a reactive aldehyde group that can create unwanted reactions thereby reducing the

recovery yield of each component. In fact, the AA technique has an additional drawback of being unable to detect acidic residues, unless the reduction of the carboxylic group is performed.¹⁶

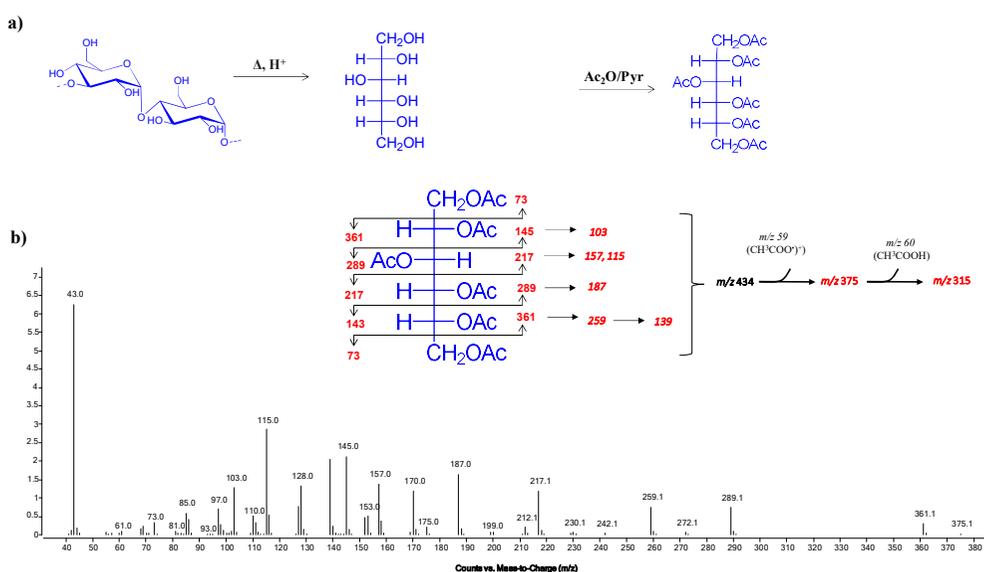


Figure 2.4 a) Scheme of the reactions taking place during derivatisation of a glycan [$\rightarrow 4$] α -D-Glcp (1 \rightarrow) into the corresponding Acetylated Alditol (AA). **b)** MS spectrum of a fully acetylated hexitol (glucitol) together with of the primary fragments

The GC-MS can also be employed to distinguish the absolute configuration of individual monosaccharides within the LPS structure. This is performed through the distinction between enantiomers by solvolysis with an enantiomerically pure alcohol such as 2-(+)-octanol or 2-(+)-butanol. After the acetylation step, a comparison between the retention time of the acetyl 2-(+)-octyl glycosides and one of a standard mixture of *O*-2-(\pm)-octyl-glycosides in D or L configuration allows the assignment of the monosaccharide configuration (**Figure 2.5**).¹⁷

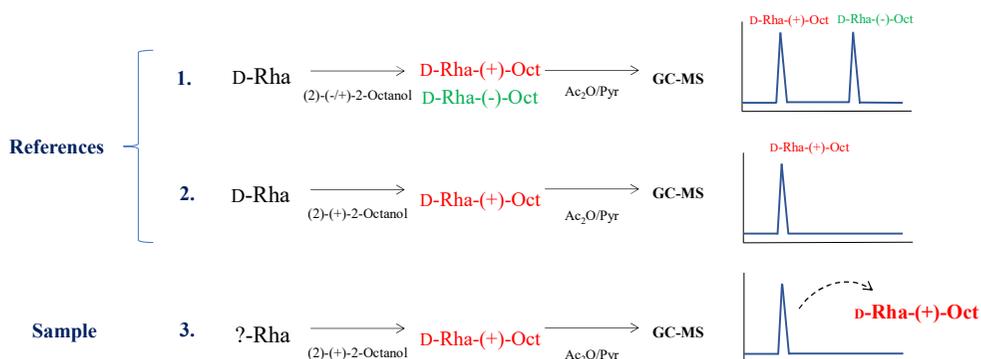


Figure 2.5 Strategy used for the construction of the appropriate octyl glycoside standard used to identify the absolute configuration of a monosaccharide. D-Rha (rhamnose) is used as an example: **1)** The reaction of 2-(+/-)-octanol with D-Rha produces a mixture of diastereoisomers: D-Rha-(+)-oct. and D-Rha(-)-oct. Mixture is acetylated in pyridine and injected into GC-MS. The retention time of these compounds is the same as the corresponding enantiomers L-Rha(-)-oct. and L-Rha-(+)-oct. respectively. **2)** The same reaction is repeated on D-Rha using the enantiopure 2-octanol to determine the retention time of D-Rha-(+)-oct. and L-Rha-(+)-oct. **3)** Absolute configuration of the Rha (in the sample) is determined through comparison of the retention times after the reaction with the enantiopure 2-octanol

Derivatisation into Partially Methylated Acetylated Alditols (PMAA) is a crucial method to determine the ring size (pyranose or furanose) as well as the linkage pattern of the monosaccharides. This is an invaluable tool although the method is a longer and more extensive in its reaction steps than those previously described. Indeed, this procedure consists of four steps: complete methylation of polysaccharide, successive hydrolysis, reduction with NaBD₄, and acetylation. This procedure consists in a and successive hydrolysis, reduction with NaBD₄, and acetylation (**Figure 2.6a**). When injected into the GC-MS, these PMAA can be easily analysed following their substitution groups (acetyl and methoxyl groups) because the molecules break

through ionisation preferably leaving the methoxyl group with a positive charge (Figure 2.6b).

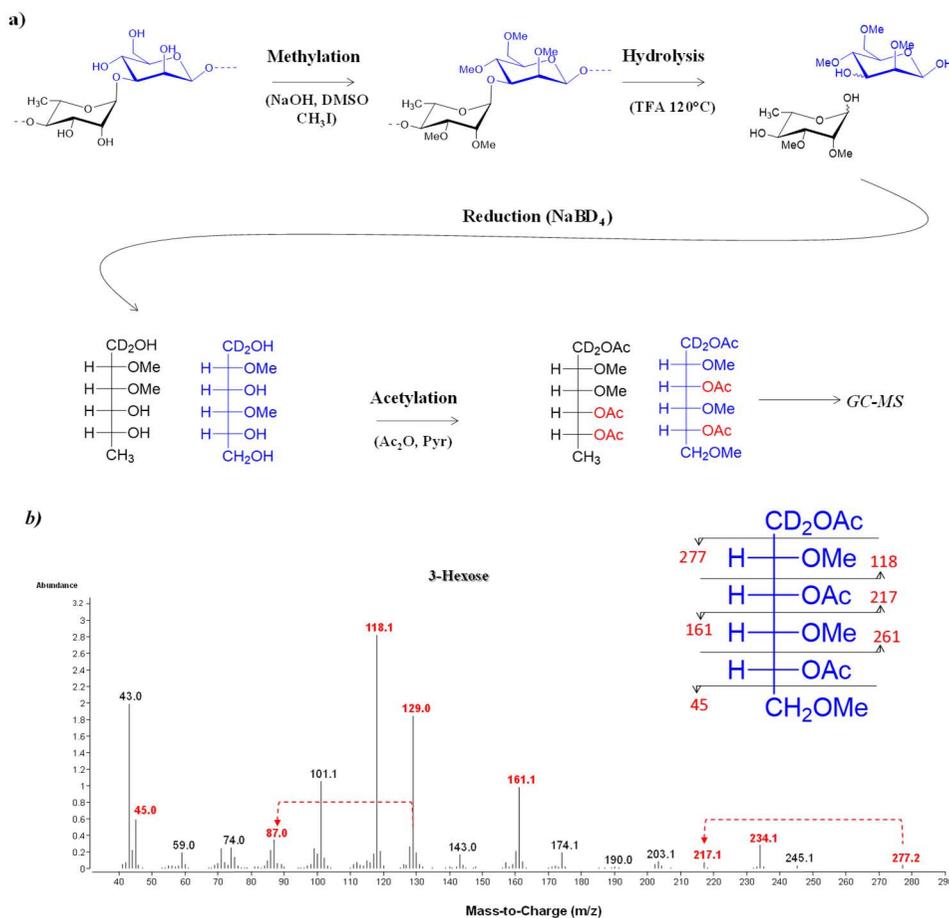


Figure 2.6 a) Scheme of the reactions occurring during derivatisation of a sugar into the corresponding Partially Methylated Acetylated Alditol. Example given here of a disaccharide of [\rightarrow 4] α -L-Rhap (1 \rightarrow 3) α -D-Manp(1 \rightarrow) **b)** MS spectrum of the PMAA from a 3-linked hexose annotated with the alditol backbone fragmentations which enable the identification of the substitution pattern

For a full analysis of a complex mixture of sugars like those present in LPS, the composition can be derived using a combination of these methods. A comparison

through GC-MS injection against known standards is highly recommended as a final confirmation of the nature of the sugar residues.

2.4.2 Chemical degradation techniques

Due to the amphiphilic nature of the LPS, micelle formation occurs thus resulting in low solubility in aqueous and apolar solvents. This is one of the biggest drawbacks of the structural investigation of LPS and therefore several strategies have been developed to overcome these hindrances. Briefly, the carbohydrate part (O-antigen and core OS) in a S-LPS or the core OS in a R-LPS can be isolated through from the lipid A by means of two chemical approaches: one entails a two-step deacylation procedure and another entails a mild acid hydrolysis. The first step of the deacylation protocol is an *O*-deacylation using anhydrous hydrazinolysis to remove ester-bound acyl chains of the lipid A. The second reaction step is a *N*-deacylation through a strong alkaline hydrolysis to cleave the amide-linked fatty acids.¹⁸

The mild acid hydrolysis, instead, leads to separation of highly hydrophobic lipid A fraction from the water-soluble saccharide moiety due to the acid-labile bond bridging the first monosaccharide of the inner core, the Kdo, and the non-reducing glucosamine residue of the lipid A disaccharide backbone. The separation of the insoluble lipid A from the saccharide mixture can be accomplished by centrifugation. The insoluble lipid A portion can be further purified using the Bligh and Dyer method.¹⁹ As for the saccharide mixture obtained, it is essential that a SEC is performed to further purify the product and to possibly separate the core OS from the O-antigen allowing the structural investigation of each single LPS moiety.

The main disadvantage of the acid hydrolysis treatment is however formation of microheterogeneity due to the reducing Kdo residue (formation of α and β , furanose and pyranose forms) hindering the analysis of the core OS by NMR and therefore is often desirable to do in parallel the *O*- and *N*-deacylation treatment previously described. The latter procedure has, by contrast, other intrinsic drawbacks as the harsh

alkaline treatment can cause loss of important structural data. For example, a β -elimination process typically occurs in 4-substituted uronic acids leading to the loss of information on the saccharide component linked at this position. In addition, in presence of di- and trisaccharide of Kdo residues, the *N*-deacylation treatment can cause linkage cleavage, leading to the formation of di- and monosaccharide residues, respectively. Finally, alkaline labile compounds, such as phosphodiester, pyrophosphate, diphosphodiester, acetyl, and carbamoyl groups, might be cleaved by this approach.

Then, a combination of state-of-art NMR experiments and soft-ionisation mass spectrometry techniques together with compositional and linkage analyses (*Section 2.4.3*) is the key approach to fully determine the primary structure of the saccharide portions obtained after these chemical treatments.

2.4.3 Structural characterisation of the saccharide moieties using NMR Spectroscopy

NMR is a powerful technique for the characterisation of carbohydrates for its non-destructive nature, its broad applicability, and the large number of different experimental sets available. By contrast, this method comes with its limitations, which include a high economic cost of the instrument, the need for adequate amount and purity of sample, relatively long experimental time needed for the acquisition of all necessary experiments, and a high level of expertise and experience for the spectra interpretation.

The first set of experiments for the characterisation of an oligosaccharide or polysaccharide is given by mono-dimensional ^1H -NMR (proton) spectra which furnish key information regarding the nature of the sample under investigation.²⁰ In general the spectra can be divided into four regions (**Figure 2.7** ^1H NMR example showing the different regions of the spectra).

2.7–1.0 ppm	The aliphatic region: important for the detection of deoxy sugars. The methyl group of 6-deoxyresidues, such as rhamnose or fucose, can be found at about 1.1–1.3 ppm. As well as the diastereotopic protons of the methylene group at C-3 of Kdo (3-deoxy-2-keto-D-manno-octulosonic acid).
4.4–3.0 ppm	The carbinolic region: reports the ring protons of the monosaccharide residues and is generally considered as the “fingerprint” of a specific glycan.
5.6–4.4 ppm	The anomeric region: reports the anomeric protons of any sugar units and is considered the most informative part of a proton spectra.
8–7 ppm	The aromatic region: this range of the spectrum comprises aromatic signals for the detection of glycosides with an

aromatic aglycon or for sugar-nucleotides in biosynthetic studies

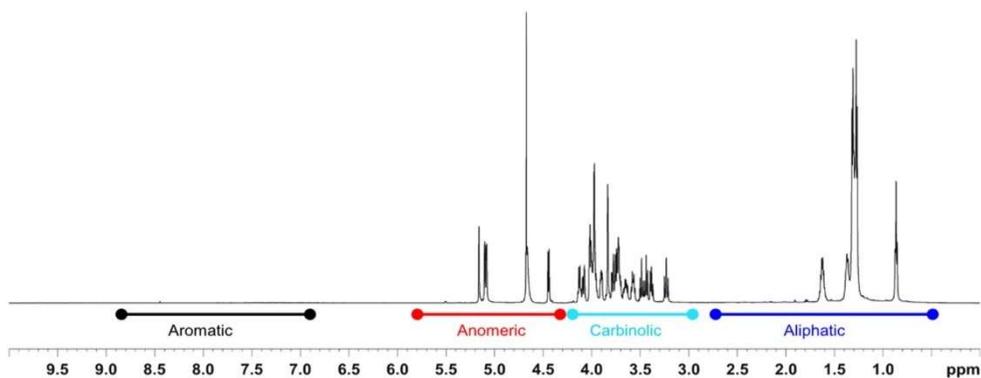


Figure 2.7 ^1H NMR example showing the different regions of the spectra. Figure taken from Speciale *et al.* (2022)²⁰

The acquisition of a combination of both homo and heteronuclear 2D NMR experiments is essential to complete the characterisation of an unprecedented structure. Usually the use of homonuclear through-bond experiments like Correlation Spectroscopy (COSY) and Total Correlation Spectroscopy (TOCSY) is applied to allow for identification and assignment of all ring protons. The Heteronuclear Single Quantum Coherence (^1H , ^{13}C -HSQC) spectrum allows for the assignment of all ^{13}C resonances and can also be used to locate the positions of glycosylation due to the typical down-field of the chemical shifts observed when a monosaccharide is glycosylated by another one. It is also key to appreciate the presence of *O*- and *N*-acyl groups beyond enabling for the identification of carbons resonating at 20-17 ppm, which are diagnostic for the presence of deoxy sugars.

NOESY (Nuclear Overhauser Effect Spectroscopy) and ROESY (Rotating Frame Overhauser Enhancement Spectroscopy) are experiments that correlate spatially close protons, thus can relate protons within the same residue (*intra*-residue NOE effect) to provide information on the specific anomeric configurations (α and β): NOE correlations between H-1/H-3/H-5 are typical of β -*gluco*-configured residues. Whereas NOE-contacts seen between H-1 and only H-2 defines an α -*gluco* configuration. NOE-residue contacts can also be useful for understanding the linkage between residues by tracing the *inter*-residue dipolar correlations. However, to define the primary sequence of an oligo-/polysaccharide, heteronuclear multiple-bond correlation (^1H , ^{13}C -HMBC) experiments should be recorded. This experiment allows the analysis of scalar long-range correlations among the ^1H and ^{13}C nuclei involved in glycosidic linkages. In addition, the HMBC spectrum provides information regarding the cyclisation of the ring with scalar long-range correlations detectable between H-1/C-5 and C-1/H-5 for pyranosidic rings or between H-1/C-4 and C-1/H-4 for furanosidic rings. Therefore, the information gained from the 2D NOESY and ROESY spectra combined with the scalar long-range correlations visible in the ^1H , ^{13}C -HMBC spectrum can be utilised to establish the full sequence of the poly/oligosaccharide. Finally, by recording a ^1H , ^{31}P -HSQC experiment it is possible

to identify the location of any phosphorylation sites likely occurring on the saccharide moiety under investigation.²⁰

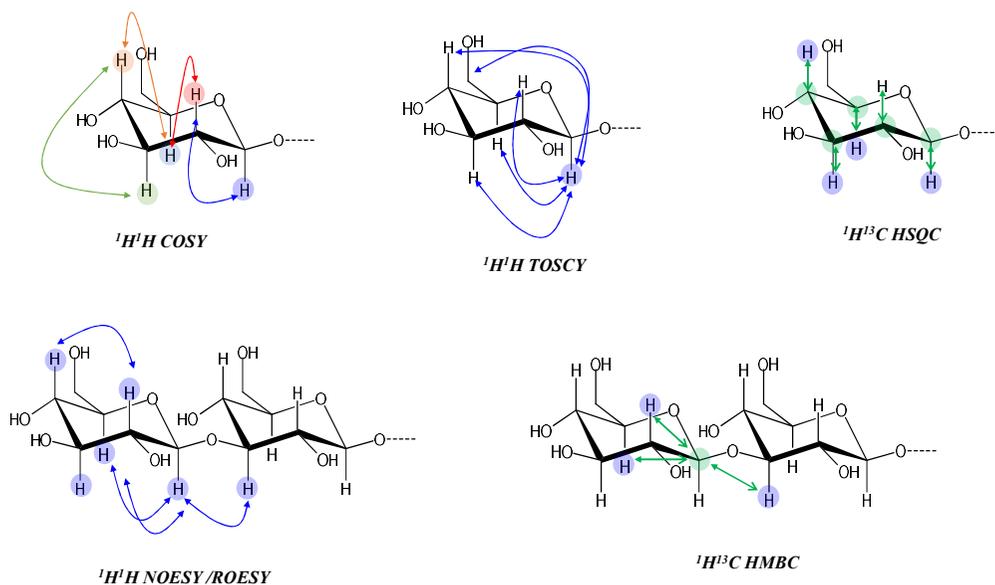


Figure 2.8 Schematic representation of nuclei involved in the main NMR techniques for the analysis of carbohydrates.

2.4.4 Structural characterisation of the saccharide moieties using MALDI TOF mass spectrometry

Advances in the development of mass spectrometers leading to high sensitivity, high precision, and fast acquisition time, have allowed to reach tremendous progresses in the structural characterisation of carbohydrates and glycoconjugates. In this frame, mass spectrometry (MS) using a soft ionisation method, such as either electrospray ionisation (ESI) or matrix-assisted laser desorption/ionisation (MALDI), are invaluable tools for the characterisation of poly/oligosaccharides and full R-LPS. In addition, MS can be ‘personalised’ through choice of analysers such as Quadrupole (Q), Quadrupole Ion Trap (QIT), Fourier transform Ion Cyclotron Resonance (FT-ICR) and Time-of-Flight (TOF).²¹ MS techniques are usually utilised in support to data obtained through NMR investigation, gaining new structural insights and thus additional information that could be lost during the NMR analysis.

Above all, MALDI-TOF analysis provides an ideal profile of the sample by defining the composition of all the molecules, thus allowing to capture a full picture of the presence of various glycoforms and their composition along with the presence of non-saccharide substituents. This type of analysis allows for the study of mixtures of oligosaccharides (such as the core OS of LPS) differing by minimal structural subunits. An additional benefit of this method is the lack of need for prior purification of the samples before the measurement.²²

The matrix is necessary to aid in the ionisation process of the analytes, but additionally it absorbs the laser energy and therefore plays an essential role in protecting the sample and assuring for a soft ionisation with formation of single charged ions.^{15,23} In addition, by using a MALDI-TOF/TOF system, it is possible to perform tandem mass spectrometry experiments, by selecting a specific ion and inducing its fragmentation. To define the fine structure of a carbohydrate or a glycoconjugate under investigation in the sample, the analysis of the fragmentation pattern in this MS/MS spectra is fundamental. However, the resulting MS/MS spectra is not always trivial as the

fragmentation does not necessarily occur at each glycosidic linkage. In addition, other different fragment types can be formed that are not easily distinguishable. The systematic nomenclature to describe these fragments was developed by Domon and Costello (**Figure 2.9**) and is based on distinguishing them into two types: derived from glycosidic cleavages (B-, C-, Y-, Z-type) and from cross-ring fragmentations (A- and X-type). More specifically, the A, B and C type are used to designate fragments containing a terminal (non-reducing end) sugar moiety, whereas X, Y and Z represent ions still containing the reducing end of the carbohydrate portion or the aglycon.^{15,24}

Great advances have been made in MALDI-TOF MS approaches enabling the study of an intact R-LPS, without any loss of structural information that can arise when the MS analysis is executed on derivatives of selective degradations. These approaches have also been shown to produce high sensitivity and resolution even at relatively high molecular masses. The laser is in fact able to cleave the acid labile linkage between the Kdo and the GlcN of the lipid A, as it occurs when a mild acid hydrolysis (explained in *Section 2.3.4*) is performed on a LPS to physically separate lipid A and saccharide moieties.^{25,26} In the resulting MALDI-TOF MS spectra there are typically three regions of ions originated by the lipid A, the core OS and the intact R-LPS. In S-LPS, due to the dispersion of molecular weight of the O-antigen, information on the size of the repeating unit can be deduced.²¹ Furthermore, recent developments have arisen through applying the use of ion mobility MS (IMS-MS) and tandem MS (IMS-MS/MS) which showed to provide a much higher resolution of the intact R-LPS.^{15,24}

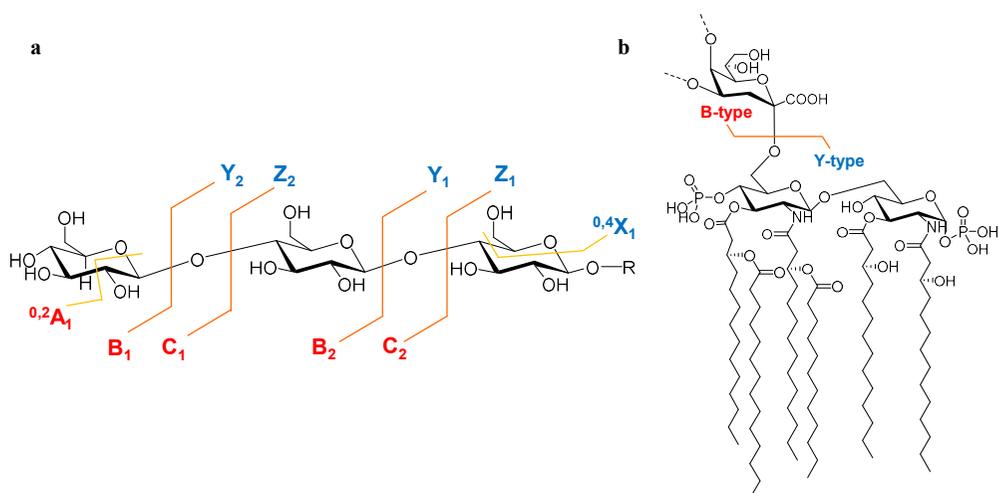


Figure 2.9 a) A scheme showing the systematic nomenclature for carbohydrate product ions generated by tandem MS approach.²⁴ The glycosidic cleavages labelled B-, C-, Y-, Z-type and the A- and X-type of the cross-ring cleavages. Subscript numbers indicate the position relative to the termini, whereas superscript numbers indicate cleavages within sugar rings **b)** A schematic representation indicating the cleavage occurring in MALDI-TOF MS analysis of a R-LPS. Indicated are the B-type (core OS) and Y-type (lipid A) ions which occur in the spectrum. Image taken from Garcia-Vello *et al.* (2021).¹⁵

2.5 Structural determination of the lipid A moiety

For the characterisation of the lipid A, first it is necessary to isolate it from the saccharide moieties of the LPS which can be achieved through a mild acid hydrolysis. A centrifugation step then enables the precipitation of the lipid A that can then be purified using the Bligh and Dryer method.^{3,19} The determination of fatty acids content is a key analysis in the process of LPS, and of lipid A structural characterisation, which can be executed by following several procedures. The lipid A can be characterised compositionally using the previously mentioned AMG or AA methods and GC-MS (*Section 2.3.3*), to analyse the sugar backbone but also to gain information about other atypical sugars that might be found as part of the saccharide backbone of some lipid As.

Another method is through studying fatty acid as methyl esters upon their proper derivatisation and (*Section 2.3.3*) hexane extraction. The total fatty acid content can be obtained upon treatment of the lipid A with HCl followed by neutralisation with NaOH. The free acyl chains are then extracted in CHCl₃, methylated with diazomethane, and investigated by GC-MS. By analysis of the mass spectra and through comparisons with properly prepared standards, the fatty acids composition and nature (straight or branched) can be defined.

MALDI-TOF MS/MS is applied to the isolated lipid A to detailly define its structure. In addition, the amide-linked fatty acids can be deduced through MALDI-TOF MS and MS/MS analysis of the *O*-deacylated lipid A by using anhydrous hydrazine¹⁸ or by treating the lipid A with NH₄OH, which cleaves the ester linked acyl chains leaving the amide-linked ones unaffected. This analysis is therefore crucial in the determination of the secondary fatty acid distribution whose definition is not always a trivial task to carry out.

2.6 Computational Techniques in the characterisation of lipid A ligands with TLR4/MD-2 complex

Computational techniques can be employed to characterise and investigate the structure and dynamic behaviour of molecules by combining the principles from chemistry and quantum physics. Indeed, these techniques can be used in combination with *in vitro* experimentation to either help demonstrate a hypothesis or create a new one. Molecular docking is one of the most rapidly developing computational tools, with its use in computational chemistry, biology, and drug design. This method can provide a robust answer to some of the more complex biological problems, to improve existing drugs to better bind the target protein or find novel molecules with enhanced affinity. In this thesis, molecular docking was employed to predict the binding mode of a ligand (lipid A) to the target site of a protein (TLR4/MD-2) to shed light on the mode of interaction between these two molecules and provide answers to the observed immunological properties.

2.6.1 Molecular docking

The aim of molecular docking is to predict if a given molecule (a ligand) will be able to bind with a biological target (a protein, also known as a macromolecule) and how strong their interactions are. Indeed, in biological systems interactions between two molecules play an important role in signal transduction. Furthermore, the orientation and shape of the interacting molecules can greatly influence this signalling event, and therefore molecular docking is an invaluable tool for predicting biological responses.

Docking is achieved in two stages: the first is by sampling the conformation of the ligands in the active site of the receptor, and second to rank these conformations based on a scoring function. The defining of the score function is proportional to the binding affinity of the ligand towards a target. There are three different classifications of scoring functions: physics-based, knowledge-based, and empirical.²⁷ The algorithms of these two steps should be experimentally reproducible, meaning that predicted poses should be biologically relevant. For a successful docking, the ligand is often

allowed flexibility while the protein is held rigid.²⁸ The flexibility of the protein at the binding site can be allowed at a later stage when performing the molecular dynamic (MD) simulations. There are a variety of available programs to use for docking calculations, with the Autodock being one of the most common in academic and commercial use. Within this work, two docking programs have been employed: Autodock4 and Vina.

AutoDock4

Autodock4.0 is the most widely employed molecular docking software for fast and accurate results on predicted bound poses of a ligand with a chosen receptor. This program uses a grid-based approach where the target site of the protein is embedded in a grid. A Lamarckian Genetic Algorithm is used to search for the conformation of the ligand to the protein. Autodock then produces a list of results which are ranked by their predicted binding energies through a semiempirical function. This function works in a two-step process: first, the program estimates the *intra*-molecular energetics for the transition from *unbound* to *bound* state for both the ligand and the protein; second, it involves the *inter*-molecular energetics resulting from a combination of these two molecules.^{27,29}

Vina

AutoDock Vina (Vina) is another open-source software developed by the Scripps Institution of Oceanography which has no graphical interface but uses the MGL-Tools program. Vina does not require parameter or grid maps files like in AutoDock4 but functions with just 3D structures of the molecules (ligand and proteins) and the specification of the search space in the protein (active site coordinates). A limitation of Vina is that only 20 poses can be predicted per ligand but benefits from being remarkably faster than Autodock4. Vina uses a hybrid scoring function but can be majorly classified as empirical in nature.³⁰

Once docking methods have been completed, estimation of the strength of the intermolecular interaction between the ligands and its biological target can be completed through MD studies. These can also provide information regarding small conformational changes that occur in the macromolecule upon its interaction with the ligand.

2.7 References

1. De Castro, C., Parrilli, M., Holst, O. & Molinaro, A. Microbe-Associated Molecular Patterns in Innate Immunity: Extraction and Chemical Analysis of Gram-Negative Bacterial Lipopolysaccharides. *Methods Enzymol.* **480**, 89–115 (2010).
2. Di Lorenzo, F., De Castro, C., Lanzetta, R., Parrilli, M., Silipo, A. & Molinaro, A. in *Drug Discov.* (eds. Jimenez-Barbero, J., Canada, F. J. & Martin-Santamaria, S.) 38–63 (Royal Society of Chemistry, 2015).
3. Pither, M.D, Molinaro, A., Silipo, A. & Di Lorenzo, F. in *Methods Mol. Biol.* **2613**, (Springer International Publishing). **In press**.
4. Galanos, C., Luderitz, O. & Westphal, O. A New Method for the Extraction of R Lipopolysaccharides. *Eur. J. Biochem.* **9**, 245–249 (1969).
5. Westphal, O. & Jann, K. Bacterial Lipopolysaccharides Extraction with Phenol-Water and Further Applications of the Procedure. *Methods Carbohydr. Chem.* **5**, 83–91 (1965).
6. Hancock, R. *Pseudomonas aeruginosa* isolates from patients with cystic fibrosis: a class of serum-sensitive, nontypable strains deficient in lipopolysaccharide O side chains. *Infect Immun* **42**, 170–177 (1983).
7. Peterson' And, A. A. & Mcgroarty², E. J. High-Molecular-Weight Components in Lipopolysaccharides of *Salmonella typhimurium*, *Salmonella minnesota*, and *Escherichia coli*. *J. Bacteriol.* **162**, 738–745 (1985).
8. Tsai, C.-M. & Frasch, C. E. A sensitive silver stain for detecting lipopolysaccharides in polyacrylamide gels. *Anal. Biochem.* **119**, 115–119 (1982).
9. Kittelberger, R. & Hilbink, F. Sensitive silver-staining detection of bacterial lipopolysaccharides in polyacrylamide gels. *J. Biochem. Biophys. Methods* **26**, 81–86 (1993).
10. Corzo, J., Pérez-Galdona, R., León-Barrios, M. & Gutiérrez-Navarro, A. M. Alcian blue fixation allows silver staining of the isolated polysaccharide component of bacterial lipopolysaccharides in polyacrylamide gels. *ELECTROPHORESIS* **12**, 439–441 (1991).
11. Deb, P. K., Kokaz, S. F., Abed, S. N., Paradkar, A. & Tekade, R. K. in *Basic Fundam. Drug Deliv.* (ed. Tekade, R. K.) 203–267 (Academic Press, 2019).
12. Herget, S., Toukach, P. V., Ranzinger, R., Hull, W. E., Knirel, Y. A. & Von Der Lieth, C. W. Statistical analysis of the bacterial carbohydrate structure data base (BCSDB): Characteristics and diversity of bacterial carbohydrates in comparison with mammalian glycans. *BMC Struct. Biol.* **8**, (2008).
13. Kamerling, J. P. & Gerwig, G. J. in *Compr. Glycosci.* (ed. Kamerling, H. B. T.-C. G.) 1–68 (Elsevier, 2007).
14. Ruiz-Matute, A. I., Hernández-Hernández, O., Rodríguez-Sánchez, S., Sanz, M. L. & Martínez-Castro, I. Derivatization of carbohydrates for GC and GC-MS

- analyses. *J. Chromatogr. B Analyt. Technol. Biomed. Life. Sci.* **879**, 1226–1240 (2011).
15. Garcia-Vello, P., Speciale, I., Di Lorenzo, F., Molinaro, A. & De Castro, C. in *Lipopolysaccharide Transp. Methods Protoc.* (ed. Sperandeo, P.) 181–209 (Springer US, 2022).
 16. Pettolino, F. A., Walsh, C., Fincher, G. B. & Bacic, A. Determining the polysaccharide composition of plant cell walls. *Nat. Protoc.* **7**, 1590–1607 (2012).
 17. Leontein, K., Lindberg, B. & Lonngren, J. Assignment of absolute configuration of sugars by g.l.c. of their acetylated glycosides formed from chiral alcohols. *Carbohydr. Res.* **62**, 359–362 (1978).
 18. Holst, O. Deacylation of lipopolysaccharides and isolation of oligosaccharide phosphates. *Methods Mol. Biol. Clifton NJ* **145**, 345–353 (2000).
 19. Bligh, E.G. and Dyer, W. J. Liquid Extraction: Bligh and Dyer. *Can. J. Biochem. Physiol.* **37**, (1959).
 20. Speciale, I., Notaro, A., Garcia-Vello, P., Di Lorenzo, F., Armiento, S., Molinaro, A., Marchetti, R., Silipo, A. & De Castro, C. Liquid-state NMR spectroscopy for complex carbohydrate structural analysis: A hitchhiker’s guide. *Carbohydr. Polym.* **277**, 118885 (2022).
 21. Harvey, D. J. Matrix-assisted laser desorption/ionization mass spectrometry of carbohydrates and glycoconjugates. *Int. J. Mass Spectrom.* **226**, 1–35 (2003).
 22. Smolira, A. & Wessely-Szponder, J. Importance of the Matrix and the Matrix/Sample Ratio in MALDI-TOF-MS Analysis of Cathelicidins Obtained from Porcine Neutrophils. *Appl. Biochem. Biotechnol.* **175**, 2050–2065 (2015).
 23. El-Aneed, A., Cohen, A. & Banoub, J. Mass Spectrometry, Review of the Basics: Electrospray, MALDI, and Commonly Used Mass Analyzers. *Appl. Spectrosc. Rev.* **44**, 210–230 (2009).
 24. Domon, B. & Costello, C. E. A systematic nomenclature for carbohydrate fragmentations in FAB-MS/MS spectra of glycoconjugates. *Glycoconj. J.* **5**, 397–409 (1988).
 25. Sturiale, L., Garozzo, D., Silipo, A., Lanzetta, R., Parrilli, M. & Molinaro, A. New conditions for matrix-assisted laser desorption/ionization mass spectrometry of native bacterial R-type lipopolysaccharides. *Rapid Commun. Mass Spectrom.* **19**, 1829–1834 (2005).
 26. Phillips, N. J., John, C. M. & Jarvis, G. A. Analysis of Bacterial Lipooligosaccharides by MALDI-TOF MS with Traveling Wave Ion Mobility. *J. Am. Soc. Mass Spectrom.* **27**, 1263–1276 (2016).
 27. Murugan, N. A., Podobas, A., Gadioli, D., Vitali, E., Palermo, G. & Markidis, S. A Review on Parallel Virtual Screening Softwares for High-Performance Computers. *Pharmaceuticals* **15**, 63 (2022).
 28. Lin, X., Li, X. & Lin, X. A Review on Applications of Computational Methods in Drug Screening and Design. *Molecules* **25**, 1375 (2020).

29. Murugan, N. A., Podobas, A., Gadioli, D., Vitali, E., Palermo, G. & Markidis, S. A Review on Parallel Virtual Screening Softwares for High Performance Computers. Preprint at <http://arxiv.org/abs/2112.00116> (2021)
30. Trott, O. & Olson, A. J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. *J. Comput. Chem.* **31**, 455–461 (2010).

CHAPTER 3

Pairing *Bacteroides vulgatus* LPS structure with its immunomodulatory effects on human cellular models

3.1 Introduction

This first chapter will be focused on the characterisation of the LPS from *Bacteroides vulgatus*. *B. vulgatus* is found with the genus *Bacteroides*, comprising the predominant anaerobes in the gut of the healthy American and Western European population, whose members can be considered as commensal, mutualist, or even pathobiont microorganisms. The cell surface structures of various *Bacteroides* species have been shown to exert immunomodulatory effects.^{1,2} Examples can be given by the CPS of *B. fragilis* that has been identified for its beneficial effects, reported to induce the proliferation of macrophages, promote the generation of T regulatory (Treg) cells, induce immune tolerance in the intestine, and protect against various sterile inflammatory disorders, including viral encephalitis and colitis.^{3,4} *B. vulgatus* strain mpk was shown to exert a strong immune-modulating properties leading to prevention of colitis-induction in several mouse models for experimental colitis. In 2019, through a collaboration with Professor Frick at the University of Tübingen in Germany, our group showed that the isolated LPS from *B. vulgatus* (LPS_{Bv}) does not induce TLR4-mediated expression of pro-inflammatory cytokines yet does create hypo-responsiveness toward subsequent LPS-stimuli in intestinal CD11c+ cells in murine *in vitro* systems.⁵ Therefore, this LPS fairly merged properties of TLR4 antagonists and agonists. Furthermore, this work proved that the administration of LPS_{Bv} re-established intestinal immune homeostasis in a mouse model for experimental colitis. This evidenced that LPS_{Bv} acts differently to what is historically thought to be exclusively a pro-inflammatory molecule. Given that LPS immunological properties are reflected in its chemical structure, such unusual effects on the host immune system must of course be attributed to the LPS chemistry. Therefore, the focus of this project was to characterise the full structure of LPS_{Bv} through the combined use of chemical, spectroscopic, and spectrometric techniques. Finally, the immunological properties of this LPS were evaluated by using human *in vitro* systems, such as HEK cell lines, human macrophages, and dendritic cells.

**The full report of this research was published in *ACS Central Science Journal*,
6, 1602-1616, in 2020, Doi: [10.1021/acscentsci.0c00791](https://doi.org/10.1021/acscentsci.0c00791)**

3.2 Structural characterisation of *B. vulgatus* LPS

3.2.1 Purification and chemical analyses of LPS from *B. vulgatus*

LPS_{Bv} was extracted from lyophilised bacterial cells by the hot phenol/water procedure.⁶ The extracted LPS underwent an enzymatic digestion by DNase, RNase and proteases in order to remove cell contaminants, followed by purification by size-exclusion chromatography and ultracentrifugation. The nature and purity of the LPS_{Bv} was determined by SDS-PAGE analysis after silver nitrate gel staining,⁷ disclosing the smooth nature of the extracted LPS, as proven by the ladder-like pattern in the upper part of the gel (**Figure 3.1**). In parallel, no bands were visible in a parallel SDS-PAGE analysis followed by Coomassie Brilliant Blue gel staining (**Figure 3.1b**), which was essential to prove the absence of any contaminating protein/lipoprotein in the isolated LPS material. These results were also further confirmed by a Micro BCA protein assay that proved a protein content of less than 1%.

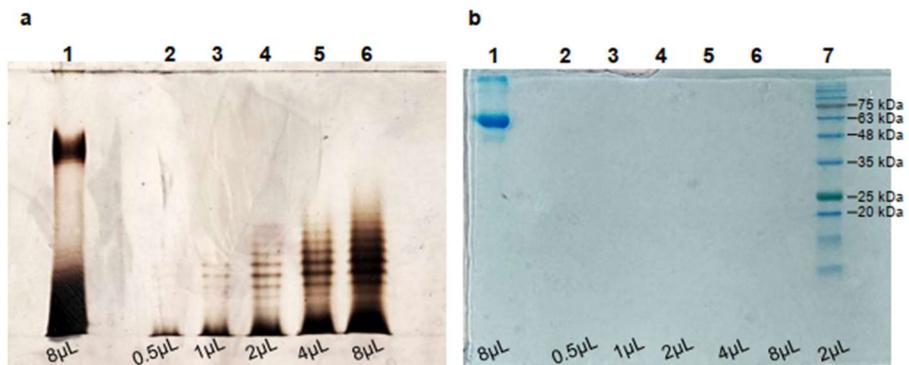


Figure 3.1 13.5% SDS-PAGE of LPS_{Bv} stained with silver nitrate (**a**) and Coomassie Brilliant Blue (**b**). 1 mg/mL solution of LPS_{Bv} were loaded on the gel in increasing volumes: 0.5, 1, 2, 4, 8 µL (Lanes 2-6). In the silver stained SDS-PAGE gel (**a**) LPS from *E. coli* O127:B8 (8 µL) (Lane 1) was used as a benchmark. In the Coomassie Brilliant Blue stained SDS-PAGE gel (**b**) bovine serum albumin (BSA) (8 µL) (Lane 1) and BLUeye Pre-stained Protein Ladder (2 µL) (Lane 7) were used as references.

Chemical analyses executed on isolated and purified LPS revealed the presence of the following monosaccharide residues: terminal, 4-substituted and 3,4-disubstituted L-rhamnopyranose (L-Rhap), terminal L-fucopyranose (L-Fucp), terminal galactofuranose (D-Galf), 3-substituted D-mannopyranose (D-Manp), 3-substituted D-glucopyranose (D-Glcp), 2,6- disubstituted D-galactopyranose (D-Galp), 6-substituted amino D-glucopyranose (D-GlcpN), and 5-substituted Kdo.⁸ Fatty acids were: tetradecanoic acid (14:0), pentadecanoic acid (15:0), hydroxypentadecanoic acid (15:0(3-OH)), hydroxyhexadecanoic acid (16:0(3-OH)), and hydroxyheptadecanoic acid (17:0(3-OH)). These results were in full agreement with those previously reported by Hashimoto *et al.* (2002) on the composition and structure of lipid A from *B. vulgatus* strain IMCJ 1204.^{9,10}

3.2.2 Structural characterisation of the saccharide portion *B. vulgatus* LPS by NMR

An aliquot of the pure LPS isolated from *B. vulgatus* mpk underwent a full deacylation¹¹ furnishing the complete saccharide portion of the LPS. The deacylated LPS fraction was then purified by gel permeation chromatography and then analysed by 1D and 2D NMR spectroscopy. A combination of homo- and heteronuclear 2D NMR experiments (DQF-COSY; TOCSY; NOESY; ROESY; ¹H, ¹³C HSQC; ¹H, ¹³C HMBC; and ³¹P and ¹H, ³¹P HSQC) were performed to elucidate the complete saccharide sequence of LPS_{Bv}. Use of the double quantum-filtered correlation spectroscopy (DQF-COSY) and the total correlation spectroscopy (TOCSY) spectra enabled the spin system of each residue to be assigned. The heteronuclear single-quantum coherence (HSQC) spectrum allowed to identify each carbon atom (**Table 3.1**). The anomeric configuration of each sugar residue was defined by *intra*-residual NOE contacts, detectable in the nuclear Overhauser effect spectroscopy (NOESY) spectrum and the ³J_{H-1,H-2} coupling constants from the DQF-COSY spectrum. The assignment of the relative configuration of each sugar unit was obtained through the observation of the vicinal ³J_{H,H} coupling constant values. Combining the information from the rotating-frame NOE spectroscopy (ROESY), NOESY, and HMBC (heteronuclear multiple-bond correlation spectroscopy) spectra, it was possible to characterise the entire primary structure of the LPS saccharide portion. Finally, ³¹P and ³¹P, ¹H HSQC experiments were pivotal to establish the location of the phosphate group decorating the LPS_{Bv}.

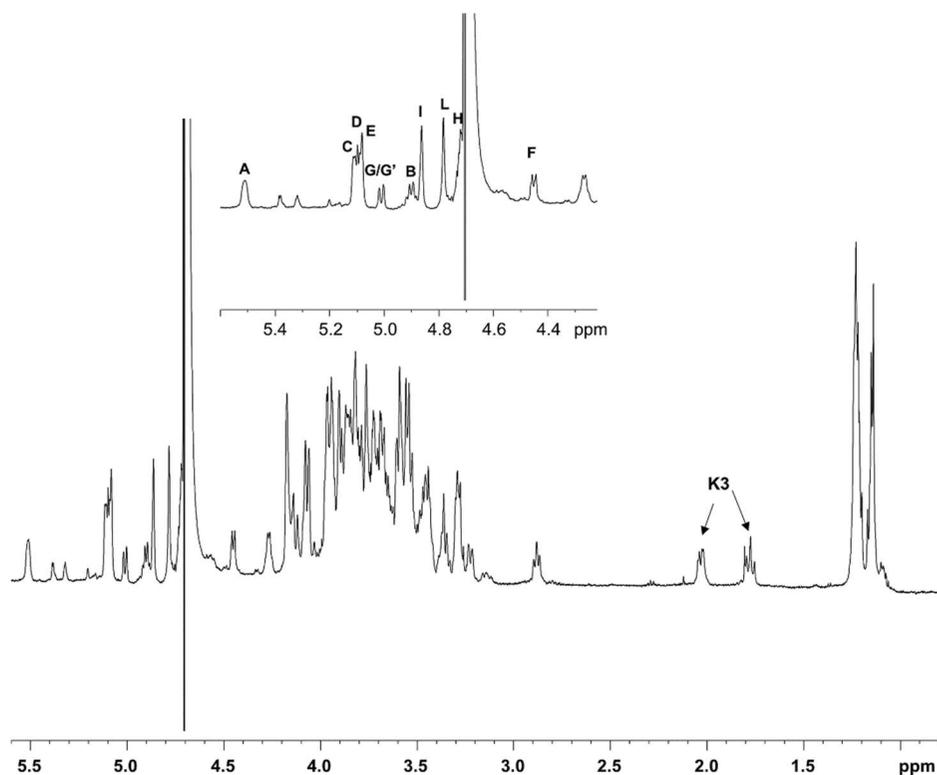


Figure 3.2 ¹H NMR spectra of the deacylated LPS from *B. vulgatus*. Inset shows a zoom of the anomeric region. Letters indicate the 11 anomeric signals (A–L).

Figure 3.2 shows the ¹H NMR spectrum with the 11 anomeric signals labelled as reported in **Table 3.1**. The spectrum also annotates the signals at δ_{H} 1.77 and 2.02 ppm that were attributed to the H-3 methylene protons of the Kdo unit (**K3** in **Figure 3.2**). All units except for residue **D**, were present as pyranose rings, according to both ¹³C chemical shift values and the presence of long-range correlations between C-1/H-1 and H-5/C-5 observable in the ¹H, ¹³C HMBC spectrum. Residue **D** was found to be a furanose ring, through the occurrence of low-field shifted ring carbon signals resonating around δ_{C} 82.0 ppm and anomeric carbon over δ_{C} 108 ppm.

Spin systems **A** (H-1 δ 5.50 ppm) and **B** (H-1 δ 4.90 ppm) were identified as the α -Glc_pN and β -Glc_pN of the lipid A moiety based on their H-2 proton signals, which correlated with two nitrogen-bearing carbon atoms at δ_C 54.3 and δ_C 55.6 ppm, respectively (**Table 3.1**). Furthermore, the occurrence of a correlation seen in the ^{31}P , ^1H HSQC spectrum between the signal at δ 2.99 ppm and the anomeric proton signal of residue **A** (δ_H 5.50 ppm, **Table 3.1**), allowed the allocation of a phosphate group at this position. The analysis of spin systems of **C**, **G**, **G'** and **I** enabled their identification as α -rhamnopyranose residues as proven by the correlations, in the TOCSY spectrum, with the methyl group signals resonating at δ_H 1.23 and 1.24 ppm (δ_C 16.9 and 17.0 ppm; **Table 3.1**).

Residues **D**, **E** and **H** were assigned as *galacto*-configured sugar units. Starting from residue **D** (H-1 δ 5.07 ppm, **Table 3.1**) that was revealed to be a β -galactofuranose (anomeric carbon signal at δ 108.6 ppm, **Table 3.1**) based on its chemical shifts and *intra* residual scalar and dipolar correlations. By contrast, spin system **H** (H-1 δ 4.72 ppm) was assigned to a β -galactopyranose, as proven by the chemical shift values of ring protons and the $^3J_{\text{H,H}}$ ring coupling constants; furthermore, the large $^3J_{\text{H-1,H-2}}$ values and the NOE correlations of H-1 with H-3 and H-5, were indicative of the β -anomeric configuration. The residue **E** (H-1 δ 5.09 ppm) was identified as an α -fucopyranose as shown by the correlations, in the TOCSY spectrum, with the methyl proton signal at δ_H 1.14 ppm (δ_C 16.2 ppm) and the $^3J_{\text{H-1,H-2}}$ coupling constant values that confirmed the α -anomeric configuration.

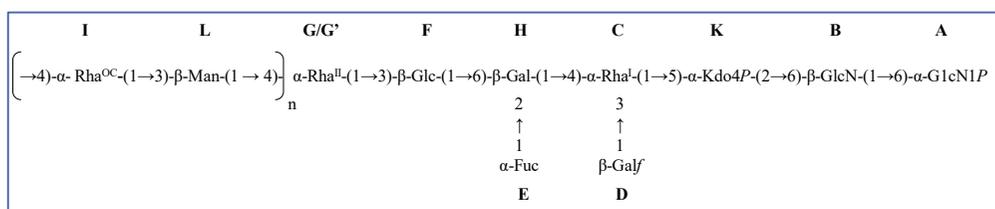
Spin system **F** (H-1 δ 4.47 ppm) was assigned to a β -glucopyranose as indicated by the large $^3J_{\text{H,H}}$ ring coupling constants and the chemical shift values of ring protons, in agreement with *gluco*-configuration of pyranose rings (**Table 3.1**). In addition, the large $^3J_{\text{H-1,H-2}}$ values and the NOE contacts of H-1 with H-3 and H-5, were diagnostic for the β -anomeric configuration. Finally, the spin system **L** (H-1 δ 4.78 ppm) was attributed to a β -mannopyranose, as proved by the small $^3J_{\text{H-1,H-2}}$ and $^3J_{\text{H-2,H-3}}$ values and the *intra*-residual correlations between H-1 and H-5 in the NOESY and ROESY

spectra which confirmed the β -anomeric configuration. The assignment of the Kdo unit was achieved starting from the diastereotopic methylene signals. The anomeric α -configuration was assigned on the basis of the chemical shift values of H-3 (δ 1.77/2.02 ppm, **Table 3.1**), and of the $^3J_{\text{H-7,H8a}}$ and $^3J_{\text{H-7,H-8b}}$ coupling constants.¹² Moreover, the Kdo residue was shown to be phosphorylated at position O-4, due to the occurrence of a correlation with a signal at 1.64 ppm in the ^{31}P , ^1H HSQC spectrum.

Table 3.1 ^1H , ^{13}C (*italic*) and ^{31}P (**bold**) chemical shift values of the saccharide region of the LPS_{Bv}.

Chemical shifts (δ)								
Unit	1	2	3	4	5	6	7	8
A	5.50	3.22	3.81	3.27	4.07	4.12/3		
6-α-Glc_pN	<i>90.4</i>	<i>54.3</i>	<i>69.9</i>	<i>70.1</i>	<i>73.0</i>	.80		
	2.99					<i>68.0</i>		
B	4.90	2.86	3.52	3.43	3.55	3.69/3		
6-β-Glc_pN	<i>99.6</i>	<i>55.6</i>	<i>72.3</i>	<i>69.8</i>	<i>74.5</i>	.47		
					<i>61.3</i>			
C	5.11	4.16	3.96	3.85	3.85	1.23		
3,4-α-Rhap (α -Rha ^I)	<i>101.0</i>	<i>70.2</i>	<i>78.9</i>	<i>73.6</i>	<i>68.6</i>	<i>16.9</i>		
D	5.07	4.06	3.96	3.92	3.72	3.56		
<i>t</i>-β-Gal_f	<i>108.6</i>	<i>81.8</i>	<i>77.0</i>	<i>82.8</i>	<i>70.3</i>	<i>62.7</i>		
E	5.09	3.67	3.76	3.71	4.27	1.14		
<i>t</i>-α-Fuc_p	<i>99.4</i>	<i>68.3</i>	<i>71.4</i>	<i>73.2</i>	<i>66.9</i>	<i>16.2</i>		
F	4.47	3.29	3.48	3.36	3.36	3.82/3		
3-β-Glc_p	<i>102.5</i>	<i>73.5</i>	<i>82.4</i>	<i>69.9</i>	<i>75.6</i>	.62		
					<i>60.8</i>			
G	5.01	3.93	3.81	3.34	3.96	1.22		
<i>t</i>-α-Rhap (α -Rha ^{II})	<i>100.8</i>	<i>70.4</i>	<i>69.7</i>	<i>71.9</i>	<i>67.4</i>	<i>17.0</i>		
G'	5.00	3.93	ND	3.58	3.88	1.22		
4-α-Rhap (α -Rha ^{III})	<i>100.8</i>	<i>70.4</i>	<i>ND</i>	<i>79.4</i>	<i>67.5</i>	<i>17.0</i>		
H	4.72	3.45	3.75	3.80	3.71	3.93		
2,6-β-Gal_p	<i>100.2</i>	<i>76.8</i>	<i>71.4</i>	<i>69.7</i>	<i>70.6</i>	<i>69.0</i>		
I	4.86	3.90	3.94	3.58	3.89	1.22		
4-α-Rhap (α -Rha ^{OC})	<i>99.7</i>	<i>70.3</i>	<i>70.3</i>	<i>79.4</i>	<i>67.5</i>	<i>17.0</i>		

L	4.78	4.17	3.58	3.54	3.29	3.82/3		
3-β-Manp	100.3	66.6	76.6	65.0	76.1	.64		
						60.8		
K	-	-	1.77/2	4.15	4.07	3.68	3.96	3.95/3
5-α-Kdop			.02	67.2	73.3	70.5	70.2	.86
			34.5	1.64				65.6



Scheme 3.1 The structure of the O-antigen, core OS and the glucosamine disaccharide backbone of the lipid A of LPS_{Bv}, Monosaccharide lettering is indicated as in **Table 3.1**

The linkage between each residue and thus the conclusion of the LPS structure was achieved through combining information of the *inter*-residual NOE contacts (observable in both the NOESY and ROESY spectra) and the long-range correlations in the HMBC spectrum (**Figure 3.3**). This was also confirmed and concluded by the down-field shifted carbon signals, which are indicative of glycosylation sites. The glucosamine backbone of the lipid A was assigned to residues **A** and **B**, with the latter found to be substituted at O-6 by Kdo, as proven by the weak downfield shift of signal of C-6 of residue **B** (δ 61.3 ppm, **Table 3.1**), which is consistent with the α -(2→6) ketosidic linkage of Kdo with the β -GlcN of the lipid A. The Kdo was then found to be substituted at O-5 by the residue **C** due to the long-range correlation between the anomeric proton signal of **C** and the C-5 of Kdo visible in the HMBC spectrum (**Figure 3.3**). Residue **C**, was identified as a 3,4- α -Rhap that at O-3 is substituted by residue **D** and at O-4 by the residue **H**. These linkages were proved by the occurrence

of the strong NOE contacts between H-1 of **D** and H-3 of **C** and between H-1 of **H** and H-4 of **C**. These glycosidic linkages were also confirmed by the observation of the respective long-range correlations in the HMBC spectrum (**Figure 3.3**). Residue **H**, in turn, was substituted by the α -Fuc (residue **E**) at position O-2 and by the β -Glc (residue **F**) at position O-6, as proven by the long-range correlations between the H-1 of **E** and the C-2 signal of **H**, and H-1 of **F** and C-6 of **H** (**Figure 3.3**). This latter residue **F**, turned out to be substituted at position O-3 by rhamnose **G**, as suggested by the long-range contact between the anomeric proton signal of **G** and the carbon atom signal at δ 82.4 ppm (**Table 3.1**) of β -glucose **F**. As can be noted from the **Table 3.1**, it was identified a second slightly different glycoform of rhamnose **G**, referred to here as **G'**, differing for the presence of a glycosylated position, that is position O-4.

The final sugars units **I** and **L** were attributed to the components of the disaccharide O-antigen moiety, which were in full agreement with the LPS O-antigen repeating unit previously elucidated by Hashimoto *et al.* (2002).⁹ Finally, I was able also to establish the biological repeating unit, in fact, residue **G'**, assigned to a 4-substituted α -Rha, was identified as the sugar likely connecting the O-antigen moiety to the core region, as proven by the long-range correlation between H-1 of **L** and C-4 of **G'** (**Figure 3.3**).

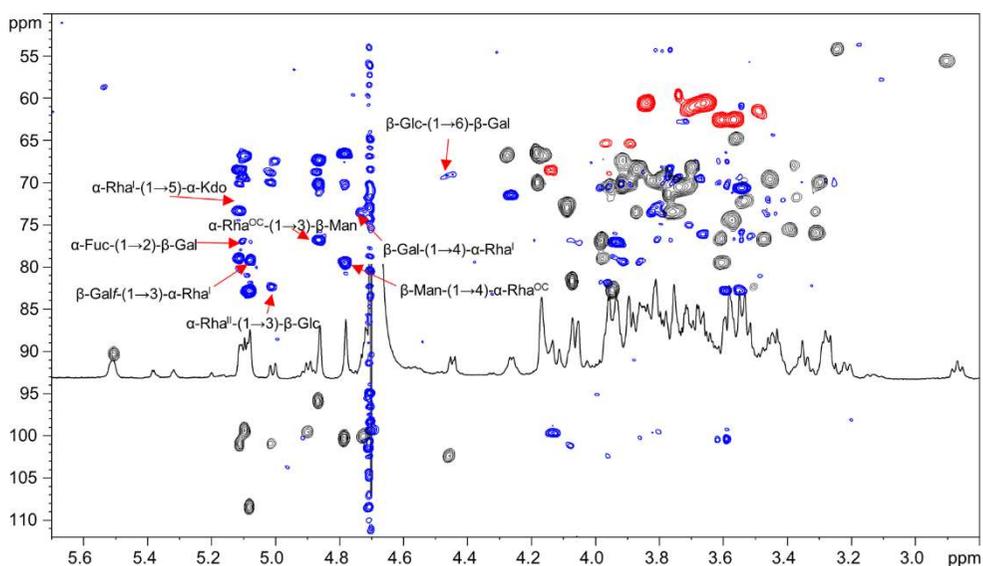


Figure 3.3 Overlap of ^1H , ^1H , ^{13}C HMBC (blue) and ^1H , ^{13}C HSQC (black and red) NMR spectra. Most relevant *inter-residue* scalar long-range correlations have been reported in the spectrum.

3.2.3 Structural characterisation of the lipid A from *B. vulgatus* LPS, by MALDI-MS

To define the structure of the lipid A of the LPS_{Bv}, an aliquot of pure LPS underwent a mild acid hydrolysis to isolate the lipid A fraction. **Figure 3.4** shows the MALDI MS spectrum acquired in negative polarity and portrays a high heterogeneity of lipid A species that was immediately apparent through the occurrence of two clusters of peaks related to lipid A species differing in the number of the fatty acid chains. Furthermore, each of these clusters of ions was characterised by the presence of mass differences of 14 and 28 amu diagnostic for lipid A species differing in the acyl chains length. The intense cluster of ions in the mass range m/z 1617–1717 Da was consistent with *mono*-phosphorylated penta-acylated lipid A species. Within this cluster, the ion peak at m/z 1688.28, was determined as a lipid A made up of two 17:0(3-OH) as *N*-linked acyl chains, two C6:0(3-OH) as primary *O*-linked fatty acids, and 15:0 as a secondary *O*-acyl substitution. The other cluster of ions visible in the mass range m/z 1385–1479 Da was attributed to *mono*-phosphorylated tetra-acylated lipid A species

with the main ion peak at m/z 1436.02 assigned to a lipid A possessing the above-mentioned fatty acid composition but lacking one primary *O*-linked 16:0(3-OH). Overall, this analysis matches which was previous published for another *B. vulgatus* strain.⁹

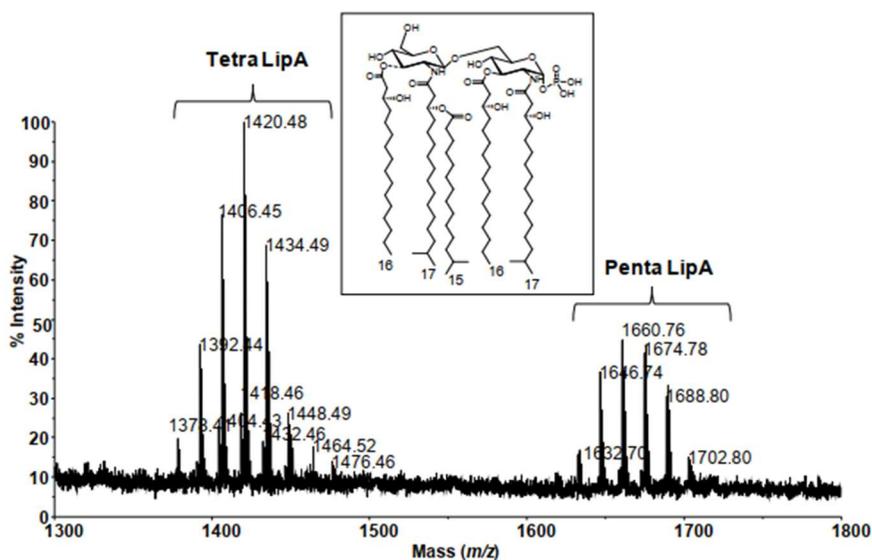


Figure 3.4 MALDI MS spectrum of *B. vulgatus* lipid A, acquired in negative polarity. Assignment of the lipid A species is depicted as TetraLipA (tetra-acylated) and PentaLipA (penta-acylated). The inset shows the structure of a representative *B. vulgatus* mpk lipid A species, relative to ion at m/z 1688.3 is depicted.

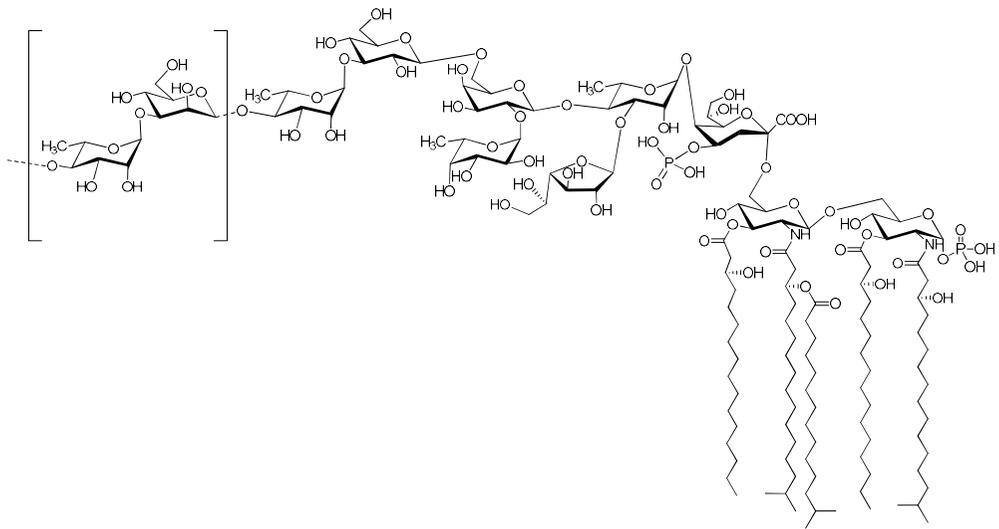


Figure 3.5 Sketch of the completed structure of the LPS from *B. vulgatus* mpk.

3.3 Immunological evaluation of *B. vulgatus* LPS

3.3.1 Activation of peripheral blood monocyte-derived macrophages and of peripheral blood-monocyte-derived dendritic cells by *B. vulgatus* LPS and lipid A

Our investigation into the immunological role of LPS_{Bv} was performed in collaboration with Professor Bernardini and Professor Chiodo at Sapienza-University of Rome and Vrije University Medical Center, Amsterdam, respectively. The first evaluations of the LPS_{Bv} were performed by using human Peripheral Blood Monocyte-Derived Macrophages (MoMs) and on Peripheral Blood-Monocyte-Derived Dendritic Cells (MoDC). These studies used both the purified LPS_{Bv} and the lipid A (lipid A_{Bv}). Cells were stimulated with LPS_{Bv} and lipid A_{Bv} at concentrations of 1, 10 and 100 ng/mL. As a positive control, the hexa-acylated *Escherichia coli* O111:B4 LPS was used while for a negative control, the synthetic tetra-acylated lipid IV_A was used. The release of cytokines from the stimulated cells were quantified by ELISA after 12 hours and assessed for the following cytokines: Tumor Necrosis Factor- α (TNF- α), Interleukin-6 (IL-6), the chemokine IFN- γ -Inducible Protein-10 (IP-10), and the Interleukin-10 (IL-10).

The choice of these cytokines/chemokines was because they provide a good representation of inflammatory (TNF- α and IL-6) and anti-inflammatory (IL-10) responses following an LPS stimuli. As explained in *Chapter 1*, following the recognition of LPS by the receptor TLR4, two signalling pathways can be activated which are mediated by the adaptor molecules MyD88 and TRIF. Both pathways lead to the activation of the NF- κ B and the production of the pro-inflammatory molecules, while the only TRIF pathway promotes the production of type I interferon and the interferon-related cytokines/chemokines, such as the chemokine IP-10.¹³ Therefore, the measurement of this chemokine was chosen to assess whether LPS_{Bv} can activate the TRIF pathway in addition to the MyD88 pathway.

3.3.2 Human peripheral blood monocyte-derived macrophages (MoMs)

Firstly, observing the MoMs after stimulation with the LPS_{BV}, (**Figure 3.6a**), the amount of TNF- α triggered at 100 ng/mL was significantly reduced with respect to that of *E. coli* LPS at the same concentration ($p < 0.5$) but comparable to that of *E. coli* LPS at 10 ng/mL. As for the chemokine IP-10, **Figure 3.6c** shows that its release, triggered by 100 ng/mL of LPS_{BV}, was appreciable but significantly lower ($p < 0.001$) than that of *E. coli* LPS at the same and minor concentration (10 ng/mL). Of note, the anti-inflammatory cytokine IL-10 production upon stimulation with 100 ng/mL LPS_{BV} or *E. coli* LPS, IL-10 was comparable while at 10 ng/mL the difference between these values was significant ($p < 0.01$), see **Figure 3.6d**. As for the stimulation of cells with the lipid A_{BV}, there was no release of the cytokines, with exception of IL-6 that was scantily but significantly produced ($p < 0.001$, for lipid A_{BV} vs untreated cells) at the concentration of 100 ng/mL (**Figure 3.6b**).

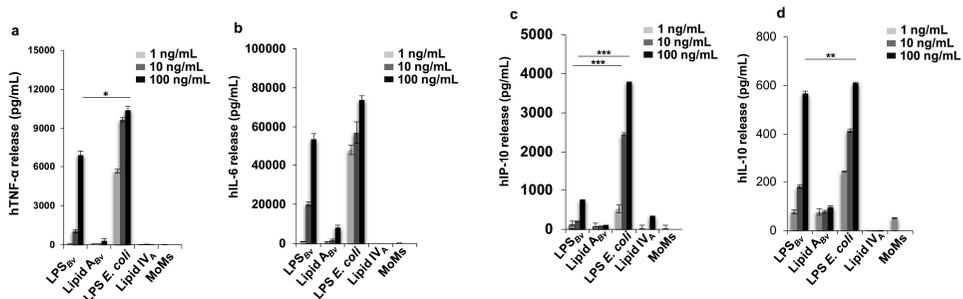


Figure 3.6 Cytokine and chemokine release induced by LPS_{BV} and lipid A_{BV} in MoMs. Release of TNF- α (**a**), IL-6 (**b**), IP-10 (**c**) and IL-10 (**d**) in cell free supernatants of MoMs. The cells were stimulated with LPS_{BV} or lipid A_{BV}, with *E. coli* LPS and with lipid IV_A at the concentration of 1, 10 and 100 ng/mL for 12 hours. Untreated cells are the negative control in these experiments. Untreated cells are the negative control in these experiments. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

3.3.3 Peripheral blood-monocyte-derived dendritic cells (MoDC).

Figure 3.7 shows the results of the ELISA when observing the MoDCs upon LPS_{Bv} and lipid A_{Bv} stimulation. The first graph (**Figure 3.7a**) shows that the production of TNF- α after stimulation with LPS_{Bv} is visibly triggered but was significantly lower than that of *E. coli* ($p < 0.001$ for all concentrations LPS_{Bv} vs *E. coli* LPS). When looking at the results for the MoDCs and lipid A_{Bv}, no release of TNF- α and IL-6 was seen following stimulation at all the concentrations evaluated. Looking at **Figure 3.7d**, showing the IL-10 response in MoDCs stimulated with LPS_{Bv}, the release of this cytokine was lower than that elicited by *E. coli* LPS at 1 and 10 ng/mL ($p < 0.01$ and $p < 0.05$, respectively, for LPS_{Bv} vs *E. coli* LPS at the same concentrations) while at 100 ng/mL no statistical difference between the two values was found. Likewise, lipid A_{Bv} elicited an exiguous but still detectable production of IL-10 ($p < 0.01$ for lipid A_{Bv} vs untreated cells). Therefore, the findings in MoDCs consolidated those achieved in MoMs and highlighted the peculiar ability of the LPS_{Bv} to stimulate the release the anti-inflammatory cytokine IL-10.

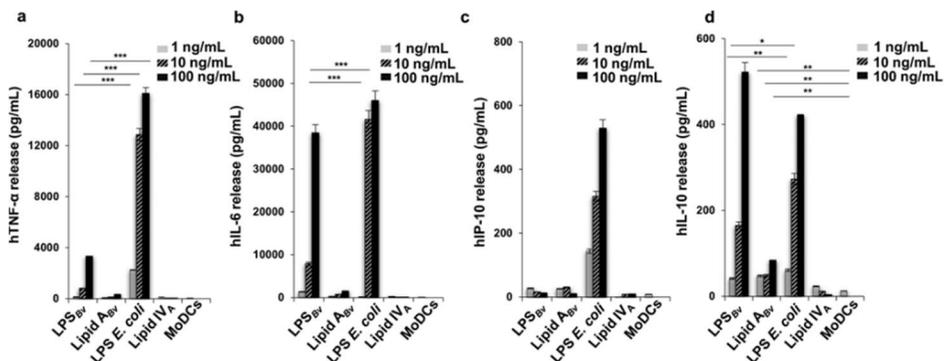


Figure 3.7 Cytokine and chemokine release induced by LPS_{Bv} and lipid A_{Bv} in MoDCs. Release of TNF- α (a), IL-6 (b), IP-10 (c) and IL-10 (d) in cell free supernatants of MoDCs. The cells were stimulated with LPS_{Bv} or lipid A_{Bv}, with *E. coli* LPS and with lipid IV_A at the concentration of 1, 10 and 100 ng/mL for 12 hours. Untreated cells are the negative control in these experiments. Untreated cells are the negative control in these experiments. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

3.3.4 Activation of bone marrow monocyte-derived macrophages by *B. vulgatus* LPS and lipid A

A common model to analyse the immune potential of LPS is the use of murine Bone Marrow Monocyte-derived Macrophages (BMDMs). Therefore, we also investigated the LPS_{Bv} and lipid A_{Bv} on this model by using ELISA and 12 hours of stimulation. In contrast to human cells, in BMDMs stimulated with LPS_{Bv} at all concentrations the TNF- α production (**Figure 3.8a**) was scant and significantly lower ($p < 0.001$, for LPS_{Bv} vs *E. coli* LPS at all concentrations) than that induced by *E. coli* LPS at all concentrations tested. Likewise, IL-6, IP-10, and IL-10 values were undetectable or scantily detectable ($p < 0.05$ for IL-10: LPS_{Bv} at 100 ng/mL vs untreated cells). Conclusively, human but not murine cells were particularly sensitive to the immunopotential of LPS_{Bv} although a particular immunological behaviour was observed also in human cell models, that is a selective induction of the anti-inflammatory IL-10 without the stimulation of the pro-inflammatory TNF- α production.

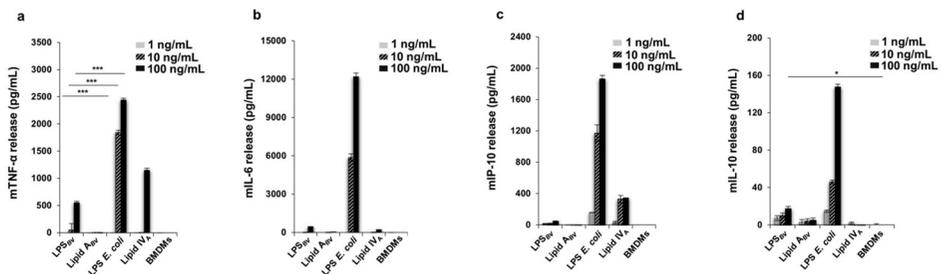


Figure 3.8 Cytokine and chemokine release induced by LPS_{Bv} and lipid A_{Bv} in BMDMs. Cytokine and chemokine releases were analysed through ELISA. Data are expressed as mean \pm SD of four independent experiments in triplicate. Untreated cells are the negative control in this experiment. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

3.3.5 Activation of TLR4 and TLR2 by LPS_{Bv} and lipid A_{Bv} in HEK293 cell models

TLR4

Further analyses were performed to understand in more detail the biological properties of the LPS_{Bv} and to gain information on the receptors involved in its recognition. The LPS_{Bv} and lipid A_{Bv} were in fact tested in the model of HEK293 cell line stably transfected with the human CD14/MD-2/TLR4 genes. LPS_{Bv} and lipid A_{Bv} were analysed at the concentration of 1, 10 and 100 ng/mL and after 6 hours the activation of the nuclear factor- κ B (NF- κ B) (**Figure 3.9a**) and the production of Chemokine (C-X-C motif) ligand 8 (CXCL-8 or IL-8) were measured (**Figure 3.9b**). *E. coli* LPS and the lipid IV_A were used at the same concentration and were the controls. Briefly, in HEK293 hTLR4 the level of NF- κ B activation induced by LPS_{Bv} (100ng/mL) reached 7 % compared to that of *E. coli* LPS at the same concentration while the cells were not responsive to lipid A_{Bv}. No CXCL-8 release was detected upon LPS_{Bv} and lipid A_{Bv} exposure.

TLR2

As explained in *Chapter 1*, research by Chavarría-Velázquez *et al.* (2018) reported that *Helicobacter pylori* LPS can activate TLR2,¹⁴ underlining that some atypical LPS show a TLR2-activating capacity. Therefore, it was decided to also evaluate whether LPS_{Bv} or lipid A_{Bv} could induce a TLR2 response. expressing hTLR2 were used to assess the immunologic impact of LPS_{Bv} and lipid A_{Bv} at the concentration of 1, 10 and 100 ng/mL as described above. *E. coli* LPS and Lipid IV_A were used in parallel. Pam3CSK4 (Pam3 CSK4) (500 ng/mL), a synthetic tri-acylated lipopeptide (LP), which mimics the acylated amino terminus of bacterial lipoprotein, represented the positive control for a TLR2 ligand. No activation of NF- κ B (**Figure 3.9c**) or CXCL-8 (**Figure 3.9d**) release was detected upon stimulation with *E. coli* LPS and lipid IV_A, as expected. LPS_{Bv} (100 ng/mL), but not lipid A_{Bv}, elicited a visible degree of NF- κ B activation ($p < 0.001$ vs untreated cells). With LPS_{Bv} CXCL-8 production was consistent with NF- κ B results ($p < 0.05$ for LPS_{Bv} vs *E. coli* LPS, both at 100 ng/mL)

and lipid A_{Bv} triggered only a scant yield. As expected, Pam3 determined high NF-κB activation and CXCL-8 production.

TLR4 plus TLR2

Finally, to evaluate the mutual influence of both TLRs (TLR4 and TLR2), the HEK293 hTLR4 cells were transfected with hTLR2 and analysed under the same conditions as stated above. These transfected cells were evaluated upon stimulation with LPS_{Bv} and lipid A_{Bv} (**Figure 3.9e and 3.9f**). In contrast to the results in HEK293 hTLR4, activation of NF-κB and CXCL-8 release induced a certain degree of activation following stimulation with LPS_{Bv} (100 ng/mL) although they were significant lower ($p < 0.001$ and $p < 0.01$ for NF-κB and CXCL-8, respectively: LPS_{Bv} vs *E. coli* LPS, both at 100 ng/mL) than those of *E. coli* LPS at the same concentration.

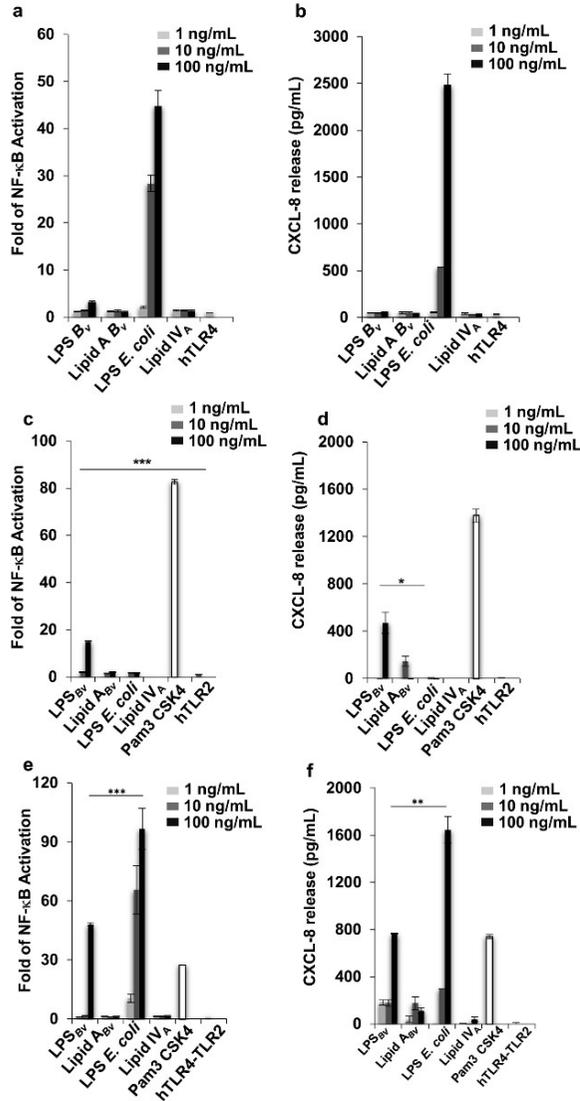


Figure 3.9 Analysis of immunopotential of LPS_{Bv} and lipid A_{Bv} in HEK293 cells expressing hTLR4/MD2-CD14 (a, b), hTLR2 (c, d) or hTLR4/MD2-CD14 and hTLR2 (e, f). Cells were treated for 6 hours with LPS_{Bv} or lipid A_{Bv}, with *E. coli* LPS, with the synthetic antagonist tetraacylated lipid IV_A at the concentration of 1, 10 and 100 ng/mL, or with the synthetic TLR2 agonist Pam3 CSK4 (500 ng/mL) (c, d, e, f). Fold of NF-κB activation in (a, c and e) and release of CXCL-8 release supernatants of HEK293 (b, d and f)

3.3.6 Competition tests of LPS_{Bv} and lipid A_{Bv} against agonistic LPS in HEK-Blue™ cell model

It is well known that under-acylated lipid A and/or unusual LPS show an inhibitory activity against endotoxically active LPS in human cells.^{15,16} Therefore, it was investigated the ability of LPS_{Bv}/lipidA_{Bv} to compete with the hexa-acylated fully immunocompetent *E.coli* LPS for binding to the TLR4 receptor. This was performed through pre-incubating HEK-Blue™ hTLR4 cells with LPS_{Bv} or lipid A_{Bv} which were then stimulated with 10 and 100 ng/mL of *E. coli* LPS for 4 h. NF-κB activation and CXCL-8 production were the read-out of these experiments. **Figure 3.10a** and **3.10b** shows that LPS_{Bv} and lipid A_{Bv} did not significantly inhibit the activity of NF-κB upon *E. coli* LPS stimulation at all the concentrations tested, albeit the trend of NF-κB was reduced with respect to the only stimulation with *E. coli* LPS. As for the CXCL-8 readout, the lipid A_{Bv} (1 and 10 ng/mL) plus *E. coli* LPS (10 ng/mL) significantly reduced the production of CXCL-8 of *E. coli* LPS (10 ng/mL) alone ($p < 0.01$) (**Figure 3.10c**). Likewise, LPS_{Bv} and lipid A_{Bv} (1 and 10 ng/mL) plus *E. coli* LPS (100 ng/mL) induced a lower release of CXCL-8 than *E. coli* LPS (100 ng/mL) alone ($p < 0.01$) (**Figure 3.10d**). As expected, lipid IV_A showed an inhibitory activity on *E. coli* LPS for NF-κB and CXCL-8 production at all the concentrations tested ($p < 0.001$). It could be concluded that LPS_{Bv} and lipid A_{Bv} show a weak inhibitory activity towards *E. coli* LPS, which agrees with the observation that LPS_{Bv} is a slight competitive inhibitor of *E. coli* LPS for binding to murine TLR4 as showed previously.⁵

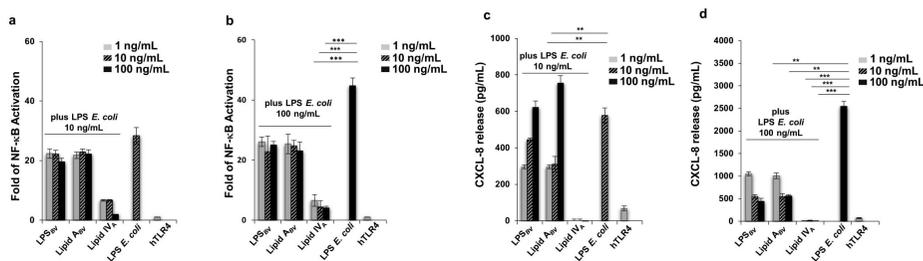


Figure 3.10 Competition assay between LPS_{Bv} or lipid A_{Bv} with *E. coli* LPS for hTLR4 binding. Fold of NF-κB activation (**a** and **b**) and CXCL-8 release (**c** and **d**) in HEK293 hTLR4/MD2-CD14 cells primed with LPS_{Bv} or lipid A_{Bv} or with lipid IV_A at the concentration of 1, 10 and 100 ng/mL for 1 hour and then stimulated with 10 (**a** and **c**) or 100 ng/mL (**b** and **d**) of *E. coli* LPS for 4 hours

3.4 Discussion and conclusions

The research activity presented here provides the complete chemical structure of the S-LPS from the intestinal bacterial species *B. vulgatus* strain mpk, which uncovered its particular and novel structural features. The lipid A species was determined to be comprised of a heterogeneous mixture of tetra- and penta-acylated *mono*-phosphorylated species. It is widely known that the immunogenicity, or the strength of the intracellular signalling, triggered by LPS via TLR4 is mostly mediated by the lipid A structure and a reduction in the acyl chain number typically correlates with a decreased immunopotency.¹⁷ *B. vulgatus* lipid A was indeed a mixture of hypo-acylated lipid A species, i.e. tetra- and penta-acylated forms, thus likely being responsible of a low agonism towards TLR4 activation. Furthermore, both 1- and 4'-phosphates on the lipid A disaccharide backbone are known to be important moieties for MD-2/TLR4 receptor complex activation,⁴ and since *B. vulgatus* lipid A possesses only one phosphate at position 1 of the reducing glucosamine, this may also contribute to its weaker (compared to *E. coli* LPS) agonistic effects as a missing 4'-phosphate was demonstrated to result in a 100-fold reduction in endotoxic activity.¹⁰ Of note, also the carbohydrate portion of LPS_{Bv} showed to possess novel and uncommon characteristics. Briefly, the core OS revealed a rhamnose substituted at the O-5

position of the Kdo, an unusual feature because almost always a heptose sugar is found at this position. Furthermore, the core OS contains a five-member ring sugar, that is Gal f , which is rare for a core OS and is known to be immunogenic to mammals, thus representing an innovative signature for bacterial LPS. It is also worth noting that Gal f can be recognised by specific lectins of the eukaryotic innate immunity system, such as intelectin,¹⁸ whose interaction would be of great interest to be further investigated, which will be a logical follow up of this project. The O-antigen structure was built up of repeating units made up of β -Man p and α -Rhap, with α -Rhap as the sugar binding the polysaccharide chain to the core OS

The investigation of the immunological properties of LPS_{Bv} in human macrophages showed a weaker capability to elicit pro-inflammatory cytokines compared to that of the potent agonist LPS from *E. coli*. It was also shown that the lipid A from *B. vulgatus* had an even weaker effect compared to the whole LPS molecule thus testifying that the full structure is pivotal for the innate immunity modulation. It can be hypothesised that the LPS_{Bv} acts by providing a basic anti-inflammatory intracellular transcription program, without exceeding a pro-inflammatory threshold, which could exert a “protective” role for the colonised host. In addition, our results showed that LPS_{Bv} was able to highly stimulate the release of IL-10, a potent anti-inflammatory cytokine that restrains pathogenic inflammation in the gut. By using the HEK cell lines, it was understood that the LPS_{Bv} acts as a weaker TLR4 agonist compared to the *E. coli* LPS but also that LPS_{Bv} interacts in a clear synergistic manner when both TLR4 and TLR2 are present on immune cells. This is extremely important as it has been demonstrated that TLR2 pathway is key in the establishment of commensals colonisation in the human gut.

This chapter is a summary of the work we published in 2020 in *ACS Centr. Sci.* and that launched our investigation into the determination of LPS structures from various species of the GM and their interaction with the human host.² From this research, a “paradigm shift” in our comprehension of the role of LPS began: from a toxic

molecule to a potentially beneficial glycoconjugate to modulate immune response. Indeed, the long-term vision is to gain insights into the molecular motifs from harmless and beneficial gut bacterial species that may play an important role in the training, education, and maturation of our immune system.

3.5 References

1. Wexler, H. M. Bacteroides: the Good, the Bad, and the Nitty-Gritty. *Clin Microbiol Rev* **20**, 593–621 (2007).
2. Di Lorenzo, F., Pither, M. D., Martufi, M., Scarinci, I., Guzmán-Caldentey, J., Łakomic, E., Jachymek, W., Bruijns, S. C. M., Santamaría, S. M., Frick, J.-S., van Kooyk, Y., Chiodo, F., Silipo, A., Bernardini, M. L. & Molinaro, A. Pairing *Bacteroides vulgatus* LPS Structure with Its Immunomodulatory Effects on Human Cellular Models. *ACS Cent. Sci.* **6**, 1602–1616 (2020).
3. Neff, C. P., Rhodes, M. E., Arnolds, K. L., Collins, C. B., Donnelly, J., Nusbacher, N., Jedlicka, P., Schneider, J. M., McCarter, M. D., Shaffer, M., Mazmanian, S. K., Palmer, B. E. & Lozupone, C. A. Diverse Intestinal Bacteria Contain Putative Zwitterionic Capsular Polysaccharides with Anti-inflammatory Properties. *Cell Host Microbe* **20**, 535–547 (2016).
4. Erturk-Hasdemir, D., Oh, S. F., Okan, N. A., Stefanetti, G., Gazzaniga, F. S., Seeberger, P. H., Plevy, S. E. & Kasper, D. L. Symbionts exploit complex signaling to educate the immune system. *Proc Natl Acad Sci U S A* 201915978 (2019).
5. Steimle, A., Michaelis, L., Di Lorenzo, F., Kliem, T., Münzner, T., Maerz, J. K., Schäfer, A., Lange, A., Parusel, R., Gronbach, K., Fuchs, K., Silipo, A., Öz, H. H., Pichler, B. J., Autenrieth, I. B., Molinaro, A. & Frick, J. S. Weak Agonistic LPS Restores Intestinal Immune Homeostasis. *Molecular Therapy* **27**, 1974–1991 (2019).
6. Westphal, O. & Jann, K. Bacterial Lipopolysaccharides Extraction with Phenol-Water and Further Applications of the Procedure. *Methods in Carbohydrate Chemistry* **5**, 83–91 (1965).
7. Kittelberger, R. & Hilbink, F. Sensitive silver-staining detection of bacterial lipopolysaccharides in polyacrylamide gels. *J. Biochem. Biophys. Methods.* **26**, 81–86 (1993).
8. De Castro, C., Parrilli, M., Holst, O. & Molinaro, A. Microbe-Associated Molecular Patterns in Innate Immunity: Extraction and Chemical Analysis of Gram-Negative Bacterial Lipopolysaccharides. *Meth. Enzymol.* **480**, 89–115 (2010).
9. Hashimoto, M., Kirikae, F., Dohi, T., Adachi, S., Kusumoto, S., Suda, Y., Fujita, T., Naoki, H. & Kirikae, T. Structural study on lipid A and the O-specific polysaccharide of the lipopolysaccharide from a clinical isolate of *Bacteroides vulgatus* from a patient with Crohn's disease. *Eur J Biochem* **269**, 3715–3721 (2002).
10. Rietschel, E. T. Absolute configuration of 3-hydroxy fatty acids present in lipopolysaccharides from various bacterial groups. *Eur J Biochem* **64**, 423–428 (1976).
11. Holst, O. Deacylation of lipopolysaccharides and isolation of oligosaccharide phosphates. *Methods in mol. biol. (Clifton, N.J.)* **145**, 345–353 (2000).

12. Birnbaum, G. I., Roy, R., Brisson, J.-R. & Jennings, H. J. Conformations of Ammonium 3-Deoxy-D- *manno* -2-octulosonate (KDO) and Methyl α - and β -Ketopyranosides of KDO: X-Ray Structure and ^1H NMR Analyses. *J. Carbohydr. Chem.* **6**, 17–39 (1987).
13. Yamamoto, M., Sato, S., Hemmi, H., Hoshino, K., Kaisho, T., Sanjo, H., Takeuchi, O., Sugiyama, M., Okabe, M., Takeda, K. & Akira, S. Role of adaptor TRIF in the MyD88-independent toll-like receptor signaling pathway. *Science* **301**, 640–643 (2003).
14. Chavarría-Velázquez, C. O., Torres-Martínez, A. C., Montaña, L. F. & Rendón-Huerta, E. P. TLR2 activation induced by *H. pylori* LPS promotes the differential expression of claudin-4, -6, -7 and -9 via either STAT3 and ERK1/2 in AGS cells. *Immunobiology* **223**, 38–48 (2018).
15. Pither, M. D., Mantova, G., Scaglione, E., Pagliuca, C., Colicchio, R., Vitiello, M., Chernikov, O. V., Hua, K.-F., Kokoulin, M. S., Silipo, A., Salvatore, P., Molinaro, A. & Di Lorenzo, F. The Unusual Lipid A Structure and Immunoinhibitory Activity of LPS from Marine Bacteria *Echinicola pacifica* KMM 6172T and *Echinicola vietnamensis* KMM 6221T. *Microorganisms* **9**, 2552 (2021).
16. Lembo-Fazio, L., Billod, J.-M., Di Lorenzo, F., Paciello, I., Pallach, M., Vaz-Francisco, S., Holgado, A., Beyaert, R., Fresno, M., Shimoyama, A., Lanzetta, R., Fukase, K., Gully, D., Giraud, E., Martín-Santamaría, S., Bernardini, M.-L. & Silipo, A. Bradyrhizobium Lipid A: Immunological Properties and Molecular Basis of Its Binding to the Myeloid Differentiation Protein-2/Toll-Like Receptor 4 Complex. *Front Immunol* **9**, 1888 (2018).
17. Brandenburg, K., Mayer, H., Koch, M. H., Weckesser, J., Rietschel, E. T. & Seydel, U. Influence of the supramolecular structure of free lipid A on its biological activity. *Eur J Biochem* **218**, 555–563 (1993).
18. Nonnecke, E. B., Castillo, P. A., Dugan, A. E., Almalki, F., Underwood, M. A., De La Motte, C. A., Yuan, W., Lu, W., Shen, B., Johansson, M. E. V., Kiessling, L. L., Hollox, E. J., Lönnerdal, B. & Bevins, C. L. Human intelectin-1 (ITLN1) genetic variation and intestinal expression. *Sci Rep* **11**, 12889 (2021).

CHAPTER 4

Lipopolysaccharide from Gut-Associated Lymphoid-Tissue (GALT) resident *Alcaligenes faecalis*: complete structure determination and chemical synthesis of its lipid A

4.1 Introduction

Peyers patches (PPs) are the major gut-associated lymphoid tissues (GALT) lying immediately under the intestinal epithelium over the entire length of the small intestine (see **Figure 4.1**).^{1,2} PPs are also known as the largest sites for the initiation and regulation of intestinal Immunoglobulin A (IgA) responses by crosstalk via cytokines and cell–cell interactions comprising dendritic, T, and B cells. It is widely accepted that microbial stimuli are required for the development and maintenance of intestinal IgA production.

Alcaligenes faecalis is the predominant Gram-negative bacterium inhabiting PPs and thus is a good candidate for such stimuli. *Alcaligenes* spp. were classified as opportunistic bacteria but it has been demonstrated that they create and maintain a homeostatic environment in PPs without triggering any harmful responses.³ Furthermore, the LPS from *Alcaligenes* spp. revealed to act as a weak TLR4 agonist and could promote IL-6 release from dendritic cells (DCs), which, in turn, enhanced IgA production.³ From this knowledge, it can be hypothesized that *Alcaligenes* spp. LPS tends to favour bacterial persistence in PPs by promoting homeostasis rather than inflammation. Acting as such, *Alcaligenes* spp. LPS would take part in the host immune system vigilance through the production of IgA, which, in turn, might favour commensal persistence in PPs.^{3,4} These remarkable immunomodulating properties of the LPS from *Alcaligenes* spp. must be necessarily attributed to its chemical structure. All these premises boosted my interest in determining the complete structure of the LPS isolated from *A. faecalis* thanks to a collaboration with Prof. Kiyono and Prof. Fukase of Tokyo and Osaka University respectively.

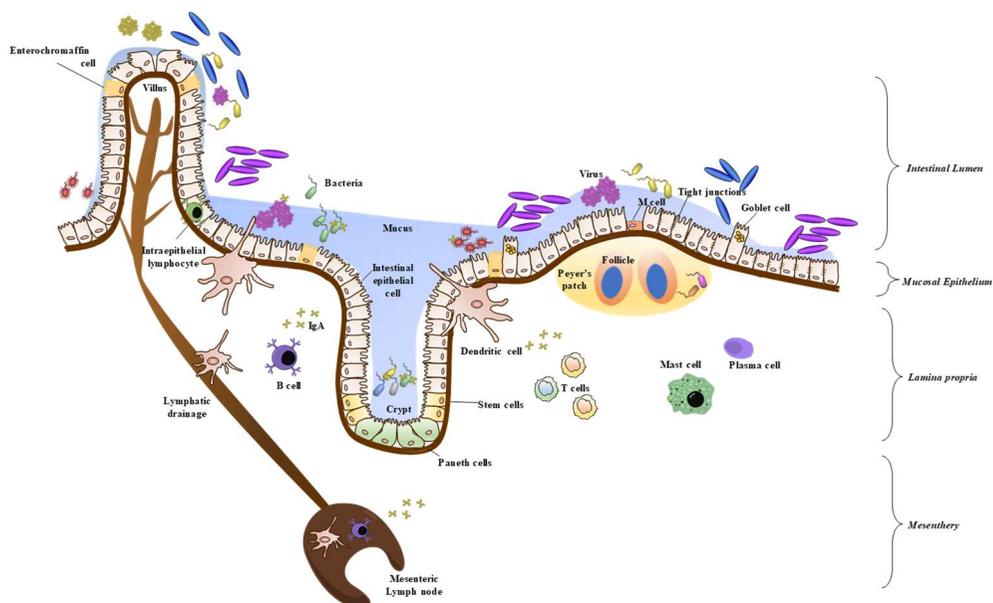


Figure 4.1 Schematic representation of the intestinal barrier with a view of the Peyer’s patch and mesenteric lymph node. Image from Di Lorenzo *et al.* (2019)²

This chapter will present the results of the full characterisation of the lipid A, core OS and the O-antigen from this prominent and important Gram-negative of the human microbiota. With the perspective of a structure-to-function relationship study, in the frame of this project, the chemical synthesis of *A. faecalis* lipid A was also performed, and the so-obtained product(s) were immunologically tested. The chemical synthesis was performed thanks to my collaboration with research group of Prof. Fukase.

The full research described in this chapter has been published in *Angewandte Chemie (International ed. in English)* in 2021, Doi: 10.1002/anie.202012374

4.2 Structure determination of the LPS of *A. faecalis*

4.2.1 Isolation and purification of *A. faecalis* R-LPS and S-LPS

The rough form LPS (R-LPS) of *A. faecalis* was isolated from dried bacterial cells by using the phenol/chloroform/light petroleum extraction protocol⁵ whereas the smooth form (S-LPS) was isolated through the hot phenol/water extraction method.⁵ The extracts were extensively dialysed against distilled water and cell material was removed by means of enzymatic digestion with nucleases followed by dialysis. SDS-PAGE analysis was used to understand the nature of the purified extracts.⁶ The LPS material (R-LPS and S-LPS) were then further purified by means of size-exclusion chromatography. Compositional analysis of the monosaccharides present in both LPS were detected as acetylated *O*-methyl glycoside derivatives.⁷ The absolute configuration was defined through analysis of the *O*-octylglycoside derivatives.⁷ The linkage pattern of the saccharide moieties was defined by analysis of the partially methylated and acetylated alditol derivatives of each monosaccharide unit.⁷ By converging the information from these analyses it was possible to identify for the R-LPS: terminal D-galactose (Gal), 6-substituted D-glucose (Glc), terminal D-glucosamine (GlcN), 6-substituted D-GlcN, terminal D-galactosamine (GalN), 4-substituted D-GalN, 3,4-disubstituted L,D-heptose (Hep), 2,7-disubstituted L,D-Hep, and 5-substituted 3-deoxy-D-manno-oct-2-ulosonic (Kdo). In contrast, the monosaccharide analysis of S-LPS highlighted the presence mainly of 2-substituted, 3-substituted, and 2,3,4- trisubstituted D-rhamnose (Rha) and terminal D-xylose (Xyl). Fatty acid content, identical for both S-LPS and R-LPS, revealed the presence of (*R*)-3-hydroxytetradecanoic acid (14:0(3-OH)), either in both ester and amide linkages, and of ester-linked (*R*)-3-hydroxydodecanoic acid (12:0(3-OH)), decanoic (10:0), dodecanoic (12:0), and tetradecanoic acids (14:0).

4.2.2 Structural characterisation of the core OS from *A. faecalis* LPS

The core OS structure was determined on the fully deacylated R-LPS after alkaline treatment and assessed by combining data from the compositional analyses and NMR spectroscopy. It was also key to perform an additional NMR investigation of the core OS structure on the material obtained after mild acid hydrolysis of the R-LPS. In particular, by analysing the acid treated product it was possible to identify the acetyl groups occurring on the huge number of aminohexoses composing the core OS of *A. faecalis*.

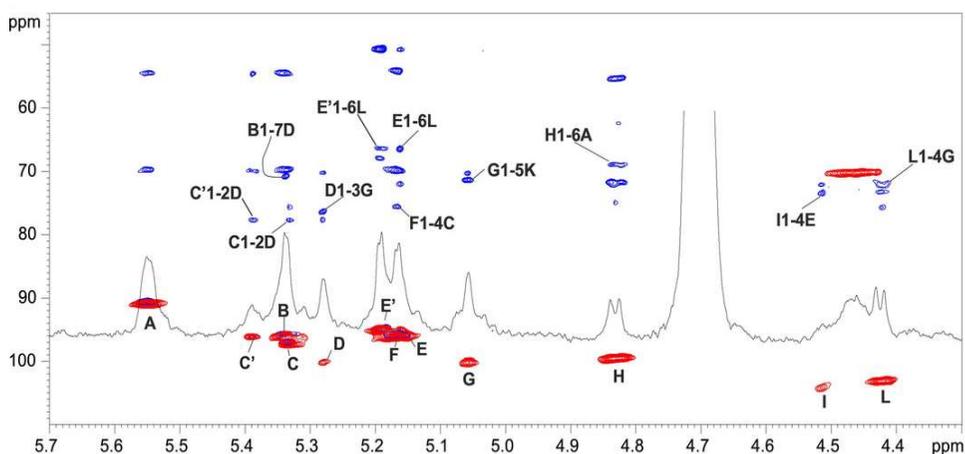


Figure 4.2 Zoom of the anomeric region of the overlapped ^1H , ^1H ^{13}C HSQC and ^1H ^{13}C HMBC recorded for the deacylated R-LPS from *A. faecalis*. Key inter-residue correlations are labelled as reported in **Table 4.1**.

NMR investigation of the fully deacylated R-LPS revealed the occurrence of twelve anomeric signals, indicative of twelve spin systems (**A–L**), plus the up field-shifted signals attributable to the H-3 methylene resonances of the Kdo unit (**K**). These residues were then fully assigned by tracing the spin connectivity derived from the DQF-COSY and the TOCSY spectra; the precise identification of each carbon atom was attained through the ^1H , ^{13}C HSQC spectrum. All the chemical shifts so deduced are reported in **Table 4.1**. The core OS was concluded to be composed of one α -Kdo

(**K**), two α -L,D-Hep (**D** and **G**), two α -D-GalNAc (**E/E'** and **C/C'**), two α -D-GlcNAc (**B** and **F**), one β -D-Gal (**I**), and one β -D-Glc (**L**). The primary sequence was inferred by using NOE contacts visible in the NOESY spectrum (**Figure 4.3**) and the ^1H , ^{13}C HMBC spectra, which is shown, together with the HSQC, in **Figure 4.2**.

Table 4.1 ^1H , ^{13}C (*italic*) and ^{31}P (**bold**) chemical shifts (ppm) of fully deacylated R-LPS from *A. faecalis*.

		1	2	3	4	5	6	7	8
A	^1H	5.54	3.26	3.80	3.37	4.05	4.17/3.91		
6- α -D-GlcN1P	^{13}C	<i>90.7</i>	<i>54.1</i>	<i>69.7</i>	<i>69.6</i>	<i>72.1</i>	<i>69.0</i>		
	^{31}P	3.60							
B	^1H	5.34	3.27	3.82	3.42	3.69	3.88/3.79		
<i>t</i> - α -D-GlcN	^{13}C	<i>95.9</i>	<i>54.1</i>	<i>69.7</i>	<i>69.5</i>	<i>72.2</i>	<i>60.2</i>		
C	^1H	5.33	3.30	3.83	4.00	3.89	3.67		
4- α -D-GalN	^{13}C	<i>97.2</i>	<i>54.2</i>	<i>69.8</i>	<i>75.6</i>	<i>71.4</i>	<i>60.9</i>		
C'	^1H	5.38	3.27	3.83	3.70	3.89	3.67		
<i>t</i> - α -D-GalN	^{13}C	<i>95.9</i>	<i>54.2</i>	<i>69.8</i>	<i>69.7</i>	<i>71.4</i>	<i>60.7</i>		
D	^1H	5.27	4.02	3.92	3.75	3.65	4.06	3.75	
2,7- α -L,D-Hep	^{13}C	<i>100.0</i>	<i>77.7</i>	<i>70.1</i>	<i>69.2</i>	<i>72.2</i>	<i>69.3</i>	<i>70.7</i>	
E	^1H	5.15	3.42	3.93	3.98	4.02	3.66		
4- α -D-GalN	^{13}C	<i>94.8</i>	<i>50.8</i>	<i>71.9</i>	<i>73.6</i>	<i>71.7</i>	<i>60.8</i>		
E'	^1H	5.18	3.42	3.91	3.97	4.07	3.66		
<i>t</i> - α -D-GalN	^{13}C	<i>94.8</i>	<i>50.8</i>	<i>68.0</i>	<i>66.6</i>	<i>71.7</i>	<i>60.8</i>		
F	^1H	5.16	3.22	3.82	3.37	3.69	3.77/3.68		
<i>t</i> - α -D-GlcN	^{13}C	<i>95.9</i>	<i>54.1</i>	<i>69.7</i>	<i>69.5</i>	<i>72.0</i>	<i>60.2</i>		
G	^1H	5.05	3.98	4.10	4.03	3.67	3.80	3.65/3.48	
3,4- α -L,D-Hep	^{13}C	<i>100.0</i>	<i>70.2</i>	<i>76.3</i>	<i>72.1</i>	<i>69.9</i>	<i>69.8</i>	<i>62.5</i>	
H	^1H	4.82	2.98	3.76	3.64	3.43	3.62/3.44		
6- β -D-GlcN4P	^{13}C	<i>99.2</i>	<i>55.4</i>	<i>71.8</i>	<i>73.9</i>	<i>72.0</i>	<i>62.3</i>		
	^{31}P				2.01				
I	^1H	4.51	3.54	3.75	3.99	3.66	3.69		
<i>t</i> - β -D-Gal	^{13}C	<i>103.7</i>	<i>72.2</i>	<i>73.5</i>	<i>70.4</i>	<i>75.7</i>	<i>60.3</i>		
L	^1H	4.41	3.21	3.36	3.35	3.50	3.88		
6- β -D-Glc	^{13}C	<i>102.8</i>	<i>73.3</i>	<i>75.8</i>	<i>69.5</i>	<i>74.4</i>	<i>66.2</i>		
K	^1H	-	-	1.90/2.16	4.46	4.18	3.69	3.95	3.86/3.56
5- α -D-Kdo4P	^{13}C	<i>171.7</i>	<i>95.6</i>	<i>34.2</i>	<i>70.2</i>	<i>71.2</i>	<i>72.1</i>	<i>70.1</i>	<i>63.6</i>
	^{31}P				-0.76				

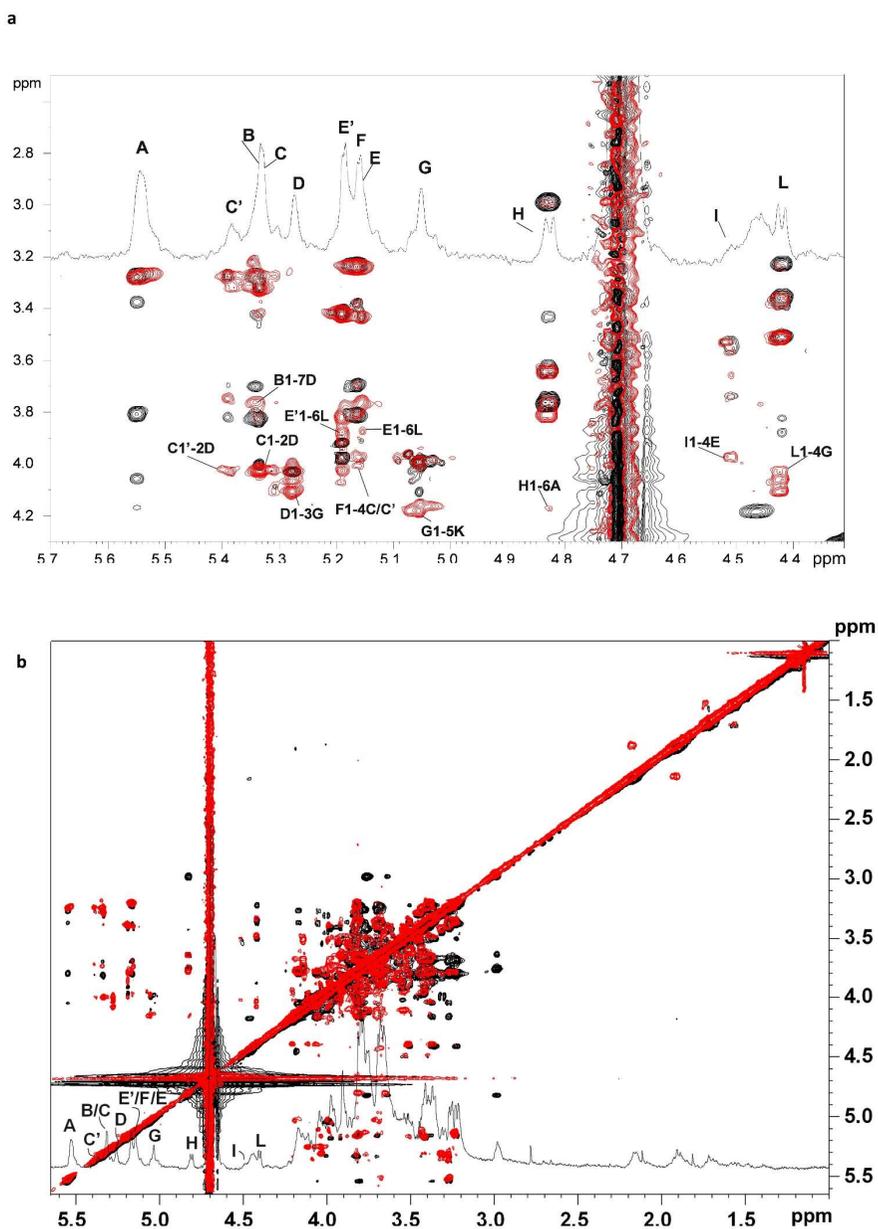


Figure 4.3 **a)** Zoom of the overlapped ^1H , NOESY (red) and TOCSY (black) spectra of the fully deacylated R-LPS of *A. faecalis*. Indicated are the key *inter-residue* NOE contacts. **b)** The superimposition of the full proton, NOESY (red) and TOCSY (black) spectra. (600 MHz, 298 K, D_2O , Table 4.1).

Briefly, starting from the lipid A sugar backbone, built up of GlcN residues **A** and **H**, the latter was found to be substituted at position O-6 by Kdo unit (**K**). This was, in turn, substituted at O-5 by α -Hep **G** as proven by the *inter*-residue NOE correlation of H-1 **G** with H-5 **K** (**Figure 4.3**), and by the corresponding long-range correlation observed in the HMBC spectrum (**Figure 4.1**). α -Hep **G** was, in turn, substituted at positions O-3 and O-4 by α -Hep **D** and β -Glc **L**, respectively. The latter β -Glc **L** was, in turn, substituted at its O-6 by terminal α -GalN unit **E/E'**. Residue **E** was found to be non-stoichiometrically substituted at its position O-4 by β -Gal **I**. Moreover, the α -Hep **D** was substituted at O-2 by β -GalN **C/C'** and at O-7 by α -GlcN **B**. Finally, residue **C** was found to be non-stoichiometrically substituted at its position O-4 by a GlcN residue **F**. The final structure of the R-LPS with the lipid A (explained further on) is sketched in **Figure 4.4**

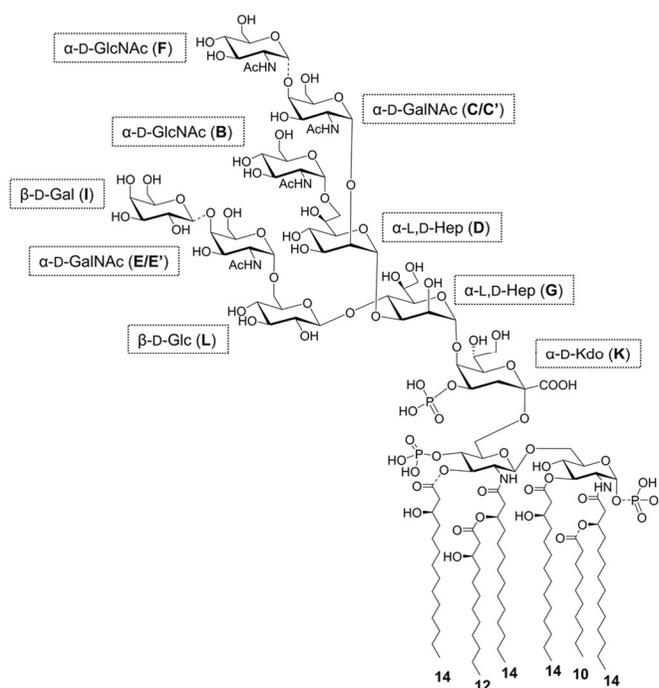


Figure 4.4 Structure of the R-LPS form *A. faecalis* with hexa-acylated lipid A species

4.2.3 Structural characterisation of the O-antigen from *A. faecalis* S-LPS

For the S-LPS which is also produced by *A. faecalis*, the O-antigen was fully characterised after a mild acid hydrolysis performed on the S-LPS, which cleaves the acid labile linkage between the Kdo residue and the β -GlcN (residue **H**) of the lipid A. The O-antigen fraction was then isolated, purified, and investigated through NMR. The final structure of the deduced xylosylated rhamnan chain can be seen in **Figure 4.5** and matches with previously published data.⁸ Chemical shifts allowing this structural determination are reported in **Table 4.2**. Briefly, a linear rhamnan chain was defined as composed of the sequence **M1** \rightarrow **3N1** \rightarrow **3O1** \rightarrow **2M** (**Figure 4.5**) as proven by the relative long-range and *inter*-residue NOE correlations. The β -Xyl units **Q** and **R** were found sitting at positions O-2 and O-4 of unit **N** respectively, as proven by the long-range correlations between H-1 of **Q** and C-2 of **N**, and H-1 of **R** with C-4 of **N**, in addition to the corresponding *inter*-residue NOE connectivities. As stated above, this branched repeating unit was identical to the one reported by Knirel *et al.* (1986),⁸ nevertheless, a minor substructure was also identified. Information about the latter could be drawn by the occurrence of the long-range correlations between H-1 of **P** and C-2 of **M**, in addition to the one between H-1 of **O'** and C-2 of **P**, suggesting the existence of the substructure **O'1** \rightarrow **2P1** \rightarrow **2M**, i.e. α -Rha-(1 \rightarrow 2)- β -Rha-(1 \rightarrow 2)- α -Rha.

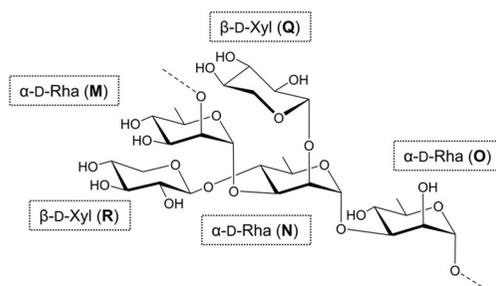


Figure 4.5 Sketch of the O-antigen from the S-LPS of *A. faecalis*

Table 4. 2 (600 MHz, 2983 K, D₂O) ¹H and ¹³C (*italic*) chemical shifts (ppm) of the O-antigen isolated through mild acid hydrolysis of the *A. faecalis* LPS fraction.

		1	2	3	4	5	6
M	¹ H	5.25	4.01	4.06	3.52	3.88	1.33
2-α-D-Rha	¹³ C	<i>100.3</i>	<i>78.5</i>	<i>69.3</i>	<i>72.1</i>	<i>69.1</i>	<i>16.6</i>
N	¹ H	5.11	4.16	4.03	3.65	3.93	1.39
2,3,4-α-D-Rha	¹³ C	<i>102.3</i>	<i>79.3</i>	<i>75.5</i>	<i>79.7</i>	<i>69.5</i>	<i>16.6</i>
O	¹ H	4.94	4.12	3.99	3.51	3.81	1.28
3-α-D-Rha	¹³ C	<i>102.0</i>	<i>69.4</i>	<i>80.2</i>	<i>72.9</i>	<i>69.0</i>	<i>16.5</i>
O'	¹ H	4.96	4.19	3.99	3.51	3.79	1.28
3-α-D-Rha	¹³ C	<i>102.0</i>	<i>70.0</i>	<i>80.2</i>	<i>72.9</i>	<i>69.0</i>	<i>16.5</i>
P	¹ H	4.83	4.17	3.72	3.71	3.47	1.33
2-β-D-Rha	¹³ C	<i>100.2</i>	<i>79.4</i>	<i>76.6</i>	<i>72.3</i>	<i>71.9</i>	<i>16.7</i>
Q	¹ H	4.47	3.37	3.43	3.49	3.98/3.22	
<i>t</i> -β-D-Xyl	¹³ C	<i>105.2</i>	<i>73.4</i>	<i>75.5</i>	<i>72.0</i>	<i>64.9</i>	
R	¹ H	4.48	3.32	3.38	3.48	3.98/3.27	
<i>t</i> -β-D-Xyl	¹³ C	<i>103.0</i>	<i>73.1</i>	<i>75.7</i>	<i>72.0</i>	<i>64.9</i>	

4.2.4 Structural characterisation of the lipid A from *A. faecalis* LPS

Detailed MALDI-TOF MS and MS² analyses were performed on the isolated lipid A obtained as a precipitate after mild acid hydrolysis of the S-LPS and R-LPS. The reflectron MALDI-TOF mass spectrum, recorded in negative polarity, of the lipid A is reported in **Figure 4.6**. The spectrum showed a pattern of peaks in the range m/z 1347.8–1784.9 relative to deprotonated [M-H]⁻ lipid A species differing both in the nature and number of acyl chains and phosphate content. Three distinct families of peaks around m/z 1375.8, 1601.9 and 1756.1 were clearly identified and matched with tetra-, penta-, and hexa- acylated lipid A species respectively, whose heterogeneous nature was also clearly apparent due to the occurrence of mass differences of 28 amu (-(CH₂)₂-unit). To establish the detailed structure of *A. faecalis* lipid A, i.e., revealing

the exact location of the acyl chains as well as of the phosphate units with respect to the glucosamine disaccharide backbone, a negative-ion MS² analysis on several ion peaks was conducted. This MS² analysis showed that *A. faecalis* hexa-acylated lipid A was decorated by four 14:0 (3-OH) moieties as the primary fatty acids, whereas 12:0 (3-OH) and 10:0 were the secondary acyl chains. As sketched in **Figure 4.4**, the hexa-acylated species were revealed to have a 3 + 3 distribution of the acyl chains with respect to the GlcN disaccharide backbone. This was then further unequivocally confirmed by a positive-ion MALDI-TOF MS and MS² analysis (**Figure 4.6**). Indeed, positive ion MS² analysis of peak at m/z 1320.2, relative to a sodiated *mono*-phosphorylated tetra-acylated lipid A species carrying three 14:0 (3-OH) and one 12:0 (3-OH), underwent a positive-ion MS² investigation. The MS² spectrum of precursor ion at m/z 1320.2 (**Figure 4.7**), among other product ions, showed a very diagnostic peak for structural elucidation at m/z 688.4 attributed to the oxonium ion. This fundamental fragment ion arose from the cleavage of the glycosidic linkage and gave a first indication of the 12:0 (3-OH) moiety linked as a secondary acyl chain to the primary amide bound 14:0 (3-OH) of the non-reducing glucosamine. In support of this hypothesis, the ion at m/z 852.7 (**Figure 4.7**), derived from a loss of 467 mass units, matched with the sequential loss of the phosphate group plus 368 mass units; the observation of such a peak led to the conclusion that a rearrangement, promoted by the free 3-OH groups on both primary 14:0 (3-OH), drove the loss of 184 mass units (C₁₂H₂₄O) from each primary acyl chain. Because such a fragmentation can only occur if the fatty acids have a free 3-OH group, this further suggested that the 12:0 (3-OH) was a secondary acyl moiety of the primary *N*-linked 14:0 (3-OH) of the non-reducing glucosamine unit. Therefore, combining fatty acid analysis with both positive- and negative-ion MALDI-TOF MS and MS² data on isolated lipid A, it was possible to define the *A. faecalis* R-LPS/S-LPS as expressing a blend of tetra- to hexa-acylated *mono* and *bis*-phosphorylated lipid A species with the uncommon occurrence also of 3-hydroxylated fatty acids as secondary substituents. This is a very unusual

characteristic as, when present, secondary hydroxylated fatty acids carry the hydroxyl group at their position 2. Here only 3-hydroxylated acyl chains were identified.

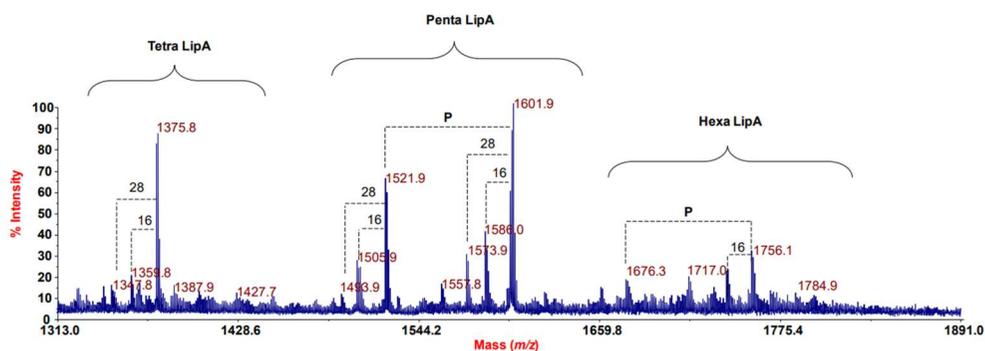


Figure 4.6 Reflectron MALDI-TOF mass spectrum, measured in negative polarity, of lipid A from *A. faecalis* LPS. The lipid A species clusters are indicated as Tetra LipA, Penta LipA and Hexa LipA. **P** indicates phosphate group

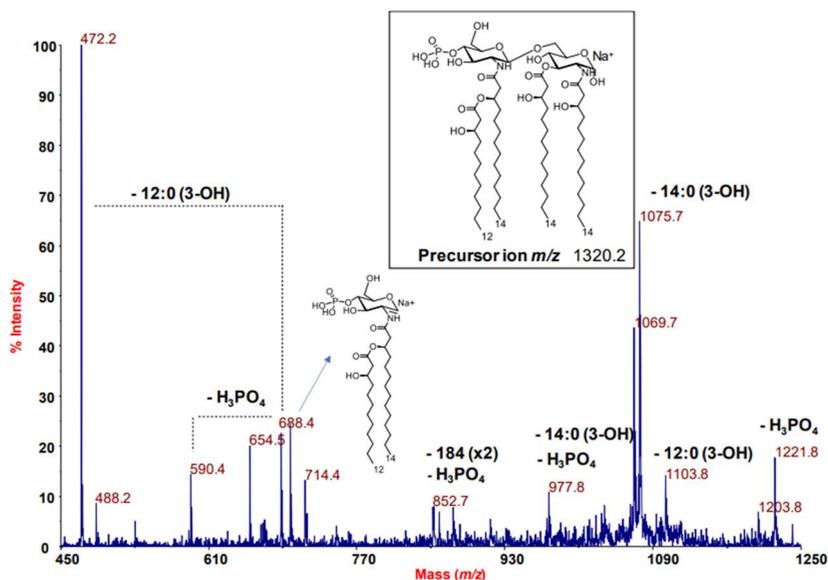


Figure 4.7 Positive-ion MALDI-TOF MS² analysis of precursor ion at m/z 1320.2 relative to a monophosphorylated lipid A species from *A. faecalis* R-LPS/S-LPS. The proposed structure

for the tetra-acylated lipid A species is sketched in the inset. The structure for the oxonium ion identified at m/z 688.4 is also reported

4.3 Chemical synthesis of the lipid A of *A. faecalis* and immunological properties

4.3.1 Chemical synthesis of hexa-acylated *A. faecalis* lipid A

The remarkable TLR4-mediated immunomodulating properties of *Alcaligenes* spp. LPS previously shown by Shibata *et al.* (2018),³ must be attributable to the lipid A chemical structure as this is the moiety which directly interacts with this innate immune receptor. Prof. Fukase and his team at Osaka University chemically synthesised the *bis*-phosphorylated tetra-, penta-, and hexa-acylated *A. faecalis* lipid A. The Scheme of the synthesis is shown in **Figure 4.8** and the full explanation can be found in the publication.⁴ Briefly, the chemical synthesis approach entailed a diversity-oriented synthetic strategy based on a key disaccharide, intermediate **1**. On this intermediate each protecting group, that is, 1-*O*-allyl, 2-*N*-allyloxycarbony (Alloc), 2'-*N*-2,2,2-trichloroethyloxycarbonyl (Troc), 3'-*O*-*p*-methoxybenzyl (MPM), and 4',6'-*O*-benzylidene groups could be selectively removed to enable the sequential introduction of acyl or phosphate groups to the proper position. By a similar approach, and starting with the same intermediate **1**, penta- and tetra-acylated lipid A were also synthesised.

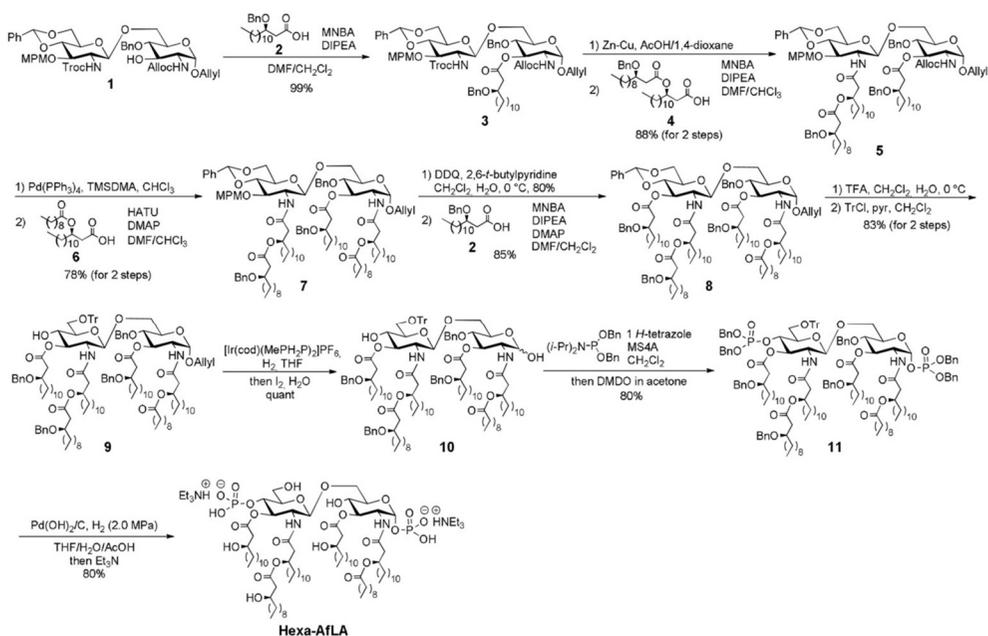


Figure 4.8 Synthesis of the hexa-acylated *Alcaligenes faecalis* lipid A

Once synthesised, the immunomodulating properties of lipid A derivatives were investigated by using HEK-Blue TLR4 to evaluate the NF-κB activation. The results in **Figure 4.9a** showed that the activation of the HEK-Blue hTLR4 cells by the synthetic hexa-acylated *A. faecalis* lipid A was weaker than *E. coli* LPS, and comparable to the previously reported data of the full LPS of *A. faecalis*. On the other hand, synthetic penta- and tetra- acylated lipid As did not show any immunostimulatory effect. Following this, the ability of synthetic lipid As were evaluated for their ability to release IL-6, TNF-α, IL1β, and IL-18 in phorbol 12-myristate 13- acetate (PMA)-differentiated THP-1 human monocytic cells and evaluated using ELISA. The synthetic hexa-acylated lipid As induced IL-6 release, although significantly less than that observed for upon *E. coli* LPS stimulation (**Figure 4.9c**). Likewise, all other cytokines exhibited low immunostimulant activity upon stimulation with synthetic lipid As compared to *E. coli* LPS.

Finally, a competition assay against *E. coli* LPS was also performed (**Figure 4.9d**), but the hexa- lipid A from *A. faecalis* exhibited almost no inhibitory and/or synergistic effect in the presence of *E. coli* LPS.

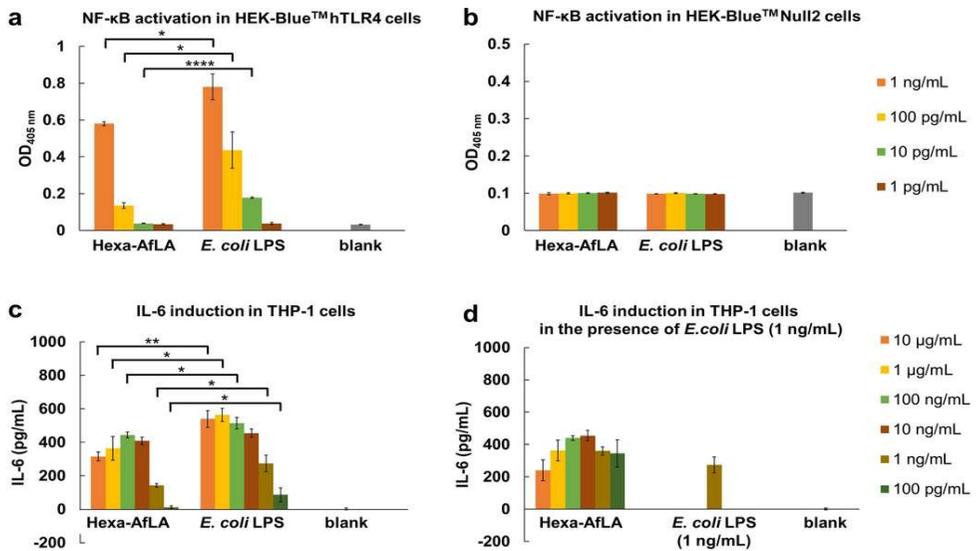


Figure 4.9 a) NF-κB activation in HEK Blue™ hTLR4 cells b) NF-κB activation in HEK-Blue™ Null2 cells, that is cells devoid of the TLR4 genes, c) cytokine IL-6 induction activity in PMA-differentiated THP-1 cells d) cytokine IL-6 release in the presence of *E. coli* LPS upon the stimulation of synthetic hexa-acylated lipid A from *A. faecalis*

4.4 Discussion and conclusions

In this chapter is presented the full structural determination of the LPS from *A. faecalis* the predominant Gram-negative inhabiting the gut-associated lymphoid tissue called PPs. From previous work it was demonstrated that *Alcaligenes* spp. can establish symbiotic relationships with the host in PPs despite the continuous contact with immune cells, and fascinatingly the LPS of *Alcaligenes* spp. was proved to be directly involved in the maintenance of homeostatic immunological conditions by induction of IgA production through an IL-6-dependent mechanism.⁴ Given these premises, the determination of the chemical structure of this unusual LPS was initiated. This bacterium revealed to present two typologies of LPS, i.e., R-LPS and S-LPS, which possess uncommon chemical features with interesting peculiarities both in the lipid A and the saccharide domains. The R-LPS is made up of the lipid A attached to a nonasaccharide core OS composed of Kdo, two *L-glycero-D-manno*-heptoses, two *N*-acetyl D-glucosamine, two *N*-acetyl D-galactosamine, one D-glucose, and one D-galactose residue. The S-LPS, which inherently has an additional polysaccharide repeat unit, was also defined as a branched xylosylated rhamnan chain. As for the lipid A component, a mixture of tetra-, penta- and hexa-acylated species were identified. The hexa-acylated forms were composed of a *bis*-phosphorylated glucosamine disaccharide backbone carrying 14:0 (3-OH) as primary and 12:0 (3-OH) and 10:0 as secondary fatty acids, which were distributed in a 3+3 fashion.

In addition, this project led to further research being undertaken, as the lipid A moieties were chemically synthesised and evaluated for their immunological properties. These results confirmed the weak agonistic behaviour towards the TLR4 activation as was previously shown by the full LPS of *A. faecalis*.³ The release of IL-6 cytokine significantly increased in THP-1 cells and were comparable to those previously observed after administration of the full *A. faecalis* LPS.⁴ Therefore, these results suggest that lipid A is the main responsible for most of the IL-6-mediated IgA production occurring in PPs. Indeed, due to its physiological location and persistence

(a gut lymphoid tissue), a *mono*-phosphorylated hypo acylated lipid A was expected for *A. faecalis*, like that seen in other gut bacterial species like. However, this structural investigation revealed the presence, as minor species, also of hexa-acylated *bis*-phosphorylated forms and not only tetra and penta-acylated. Although in the hexa-acylated form, the presence of two phosphates, the number and nature of most of the acyl chains is identical to that of *E. coli*, several different structural features might be responsible for the diverse immunological functions observed between *A. faecalis* LPS and *E. coli* LPS. In this frame, **1)** the symmetrical distribution of the acyl chains in the *A. faecalis* lipid A (3 + 3 distribution, in its hexa-acylated form, vs 4 + 2 symmetry for the *E. coli* lipid A), **2)** the presence of a shorter secondary fatty acid (10:0 in place of 14:0) and in a different position compared to *E. coli*, **3)** the presence of a hydroxylated secondary acyl moiety, and **4)** the occurrence of hypo-acylated and *mono*-phosphorylated lipid A species are notable. These chemical features, in fact, could globally result in a weak elicitation of the TLR4-mediated immune response.⁹⁻

¹¹ The structure-related properties of the lipid A from *A. faecalis* of inducing only weak activation of TLR4 signalling in combination with the potent IgA production-inducing activity are significant in the evaluation of this molecule as a potential adjuvant candidate in vaccine production. In other words, LPS from *A. faecalis* and possibly other GALT-resident mutualistic bacteria might be of inspiration for the chemical synthesis of a pool of safer and effective adjuvants for vaccine development. In addition, by analysing synthetic derivatives of LPS and/or its substructures, as done within the frame of this project, it is possible to deduce the structure to function relationship of an LPS. This can aid in understanding the common chemical features owned by symbiotic bacteria of the GM that might be essential to maintain physiological homeostasis and to provide beneficial roles to the host.

As a follow up of the present project, the saccharide moieties of *A. faecalis* LPS, especially the core OS moiety with its huge number of *N*-acetyl aminohexoses, are currently under investigation for their role and involvement in the interaction(s) with the host immune system receptors.

4.5 References

1. Kunisawa, J. & Kiyono, H. *Alcaligenes* is Commensal Bacteria Habituating in the Gut-Associated Lymphoid Tissue for the Regulation of Intestinal IgA Responses. *Front Immunol* **3**, 65 (2012).
2. Di Lorenzo, F., De Castro, C., Silipo, A. & Molinaro, A. Lipopolysaccharide structures of Gram-negative populations in the gut microbiota and effects on host interactions. *FEMS Microbiology Reviews* **43**, 257–272 (2019).
3. Shibata, N., Kunisawa, J., Hosomi, K., Fujimoto, Y., Mizote, K., Kitayama, N., Shimoyama, A., Mimuro, H., Sato, S., Kishishita, N., Ishii, K. J., Fukase, K. & Kiyono, H. Lymphoid tissue-resident *Alcaligenes* LPS induces IgA production without excessive inflammatory responses via weak TLR4 agonist activity. *Mucosal immunology* **11**, 693–702 (2018).
4. Shimoyama, A., Di Lorenzo, F., Yamaura, H., Mizote, K., Palmigiano, A., Pither, M. D., Speciale, I., Uto, T., Masui, S., Sturiale, L., Garozzo, D., Hosomi, K., Shibata, N., Kabayama, K., Fujimoto, Y., Silipo, A., Kunisawa, J., Kiyono, H., Molinaro, A. & Fukase, K. Lipopolysaccharide from Gut-Associated Lymphoid-Tissue-Resident *Alcaligenes faecalis*: Complete Structure Determination and Chemical Synthesis of Its Lipid A. *Angew. Chem. Int. Ed.* **60**, 10023–10031 (2021).
5. Galanos, C., Luderitz, O. & Westphal, O. A New Method for the Extraction of R Lipopolysaccharides. *Eur. J. Biochem.* **9**, 245–249 (1969).
6. Kittelberger, R. & Hilbink, F. Sensitive silver-staining detection of bacterial lipopolysaccharides in polyacrylamide gels. *J. Biochem. Biophys.* **26**, 81–86 (1993).
7. De Castro, C., Parrilli, M., Holst, O. & Molinaro, A. Microbe-Associated Molecular Patterns in Innate Immunity: Extraction and Chemical Analysis of Gram-Negative Bacterial Lipopolysaccharides. *Meth. Enzymol.* **480**, 89–115 (2010).
8. Knirel', I. A., Zdorovenko, G. M., Shashkov, A. S., Zakharova, I. I. & Kochetkov, N. K. [Antigenic polysaccharides of bacteria. 19. The structure of the O-specific polysaccharide chain of *Alcaligenes faecalis* lipopolysaccharide]. *Bioorg Khim* **12**, 1530–1539 (1986).
9. Molinaro, A., Holst, O., Di Lorenzo, F., Callaghan, M., Nurisso, A., D'Errico, G., Zamyatina, A., Peri, F., Berisio, R., Jerala, R., Jiménez-Barbero, J., Silipo, A. & Martín-Santamaría, S. Chemistry of Lipid A: At the Heart of Innate Immunity. *Eur. J. Chem.* **21**, 500–519 (2015).

10. Di Lorenzo, F., Billod, J. M., Martín-Santamaría, S., Silipo, A. & Molinaro, A. Gram-Negative Extremophile Lipopolysaccharides: Promising Source of Inspiration for a New Generation of Endotoxin Antagonists. *EurJOC* , **2017**, 4055–4073 (2017).
11. Tanimura, N., Saitoh, S.-I., Ohto, U., Akashi-Takamura, S., Fujimoto, Y., Fukase, K., Shimizu, T. & Miyake, K. The attenuated inflammation of MPL is due to the lack of CD14-dependent tight dimerization of the TLR4/MD2 complex at the plasma membrane. *Int Immunol* **26**, 307–314 (2014).

CHAPTER 5

Bacteroides thetaiotaomicron rough-type
**lipopolysaccharide: the chemical structure and the
immunological activity**

5.1 Introduction

Bacteroides play a vital role in the microbiota of the human host, providing protection from pathogens and supplying nutrients to other microbial residents of the gut.^{1,2} *B. thetaiotaomicron* is an abundant symbiont of the normal healthy gut, and one most extensively studied microorganisms of the GM. *B. thetaiotaomicron* has shown to degrade complex polysaccharides and to stimulate the initiation of epithelial and immunological development.^{3,4} A decrease of these bacteria in the intestinal tract has shown to correlate with gut inflammation and disease emergence.^{5,6}

B. thetaiotaomicron colonisation within the gut has been possible thanks to intricate strategies developed by the bacterium to interact with the host, including modification of its surface. In this frame, *B. thetaiotaomicron* dedicates a significant portion of its genome (eight distinct genomic loci) to production of capsular polysaccharides (CPS) resulting in the expression of eight structurally different CPS, which can reach several hundred nanometres thick^{7,8}. Therefore, most of the crosstalk between the immune system and *B. thetaiotaomicron* is supposed to involve these complex glycans, although the underlying mechanisms remain poorly understood. As a Gram-negative, *B. thetaiotaomicron* is also equipped with LPS, and although a number of publications have shown that *Bacteroides* species express R-LPS with immune inhibitory activities, an investigation into the LPS chemistry to function had yet to be completed. The lack of structural information about the LPS of such a largely studied gut mutualist likely derived by the fact that it produces a broad array of heterogenous CPS that renders the isolation of the LPS a complex procedure.⁹

This chapter is thus based on the structural and immunological investigation of the LPS from *B. thetaiotaomicron*. This work took advantage of a mutant previously constructed by Martens and co-workers and provided by Prof. Michael Fischbach of Stanford University, in which all eight CPS gene clusters were deleted to have a successful isolate of the sole LPS and thus enabling its chemical characterisation

through NMR and MALDI-TOF MS.¹⁰ Subsequently, an investigation of the immunological properties of the isolated *B. thetaiotaomicron* LPS was accomplished.

The full research described in this chapter has been published *Carbohydrate Polymers*, in 2022. Doi: 10.1016/j.carbpol.2022.120040

5.2 Structural characterisation of *B. thetaiotaomicron* R-LPS

5.2.1 Extraction, purification, and compositional analysis of *B. thetaiotaomicron* R-LPS

The LPS material was extracted using the hot phenol/water extraction method ¹¹, followed by purification steps including enzymatic digestion, dialysis, and size exclusion chromatography. The confirmation of R-LPS extraction in the water phase was achieved by using SDS-PAGE followed by gel staining with silver nitrate.¹² Indeed, the gel did not exhibit the ladder-like pattern typical of an S-LPS, but it showed two main strongly stained, rapidly migrating bands in the lower part of the gel (**Figure 5.1**). This suggested the occurrence of R-LPS with different molecular weight likely due to the length of the saccharide portion. This was also in accordance with previous Tricine SDS-PAGE gel analysis performed on a micro-scale extracted R-LPS from this mutant strain.¹⁰

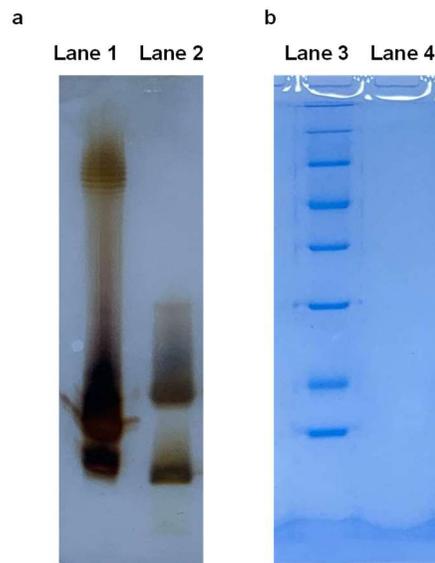


Figure 5.1 SDS-PAGE after silver **a**) and Coomassie brilliant blue **b**) staining of R-LPS from *B. thetaiotaomicron* (lanes 2 and 4). The S-LPS from *E. coli* (lane 1) was used as a benchmark in (a). In the Coomassie brilliant blue stained SDS-page gel (b), BLUeye prestained protein ladder (2 μ l) (lane 3) was used as a reference.

For the composition of monosaccharides composing the isolated R-LPS, a set of chemical analyses on the purified water phase extract was performed.¹³ By merging of data from compositional analyses, through acetylated methyl glycosides derivatisation, and absolute configuration, through acetylated octyl glycosides derivatisation, revealed the presence of D-galacturonic (D-GalpA), D-glucuronic (D-GlcpA), D-mannose (D-Manp), D-glucose (D-Glcp), D-galactose (D-Galp), 2-amino-2-deoxy-D-glucose (D-GlcpN), 2-amino-2-deoxy-D-galactose (D-GalpN), and Kdo. Linkage analysis, through partially methylated acetylated alditols derivatization proved all sugar residues are in the pyranose form. This technique also allowed the identification of their branching points as follows: terminal and 2-substituted D-Glcp, terminal D-Galp, 2-substituted, 2,3-disubstituted and 2,6-disubstituted D-Manp, terminal D-GlcpA, 4-substituted D-GalpA, terminal, 3-substituted and 4-substituted D-GlcpN, 6-substituted and 3,6 disubstituted D-GalpN, and 5-substituted Kdo were detected.

5.2.2 Structural characterisation of the lipid A from *B. thetaiotaomicron* R-LPS

Fatty acid compositional analysis of *B. thetaiotaomicron* R-LPS revealed the occurrence of tetradecanoic acid (14:0), *anteiso*-pentadecanoic acid (*a*15:0), *iso*-15:0 (*i*15:0), 15:0, hydroxypentadecanoic acid (3-OH 15:0), hydroxyhexadecanoic acid (3-OH 16:0), *iso*-hydroxyheptadecanoic acid (3-OH *i*17:0), and 3-OH *a*17:0, in full agreement with previously reported data on the structure of lipid A from *Bacteroides* sp, as also described in *Chapter 3*.

For the characterisation of the lipid A, an aliquot of the R-LPS underwent a mild acid hydrolysis followed by purification and then detailed investigation using MALDI-TOF MS and MS². The MALDI mass spectrum was recorded in negative polarity and is presented in **Figure 5.2**. Two main clusters of peaks were detected which were identified as penta-acylated and tetra-acylated *mono*-phosphorylated species, and clearly showed to have a highly heterogeneous nature due to mass differences of 14 amu, indicative of lipid A species differing in their acyl chain length. A sketch of the proposed structure from the ion peak m/z 1688.6 can also be seen in **Figure 5.2**, with the presence of branched fatty acid chains also drawn. This structural deduction matched with that of previously reported lipid A for other *Bacteroides* sp. as also described in *Chapter 3*.^{10,14-16}

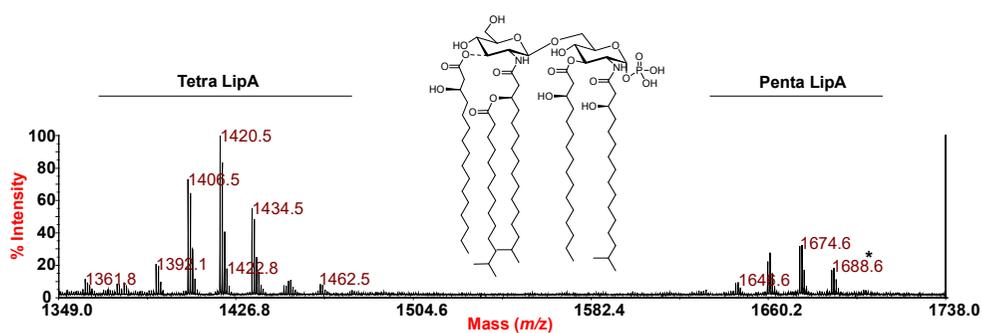
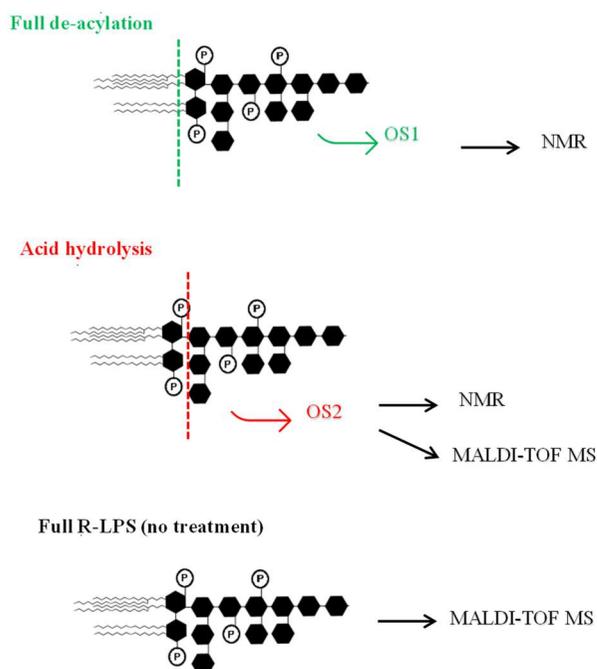


Figure 5.2 Negative-ion MALDI-TOF MS spectrum of lipid A from *B. thetaiotaomicron* R-LPS. The lipid A fraction was obtained as a precipitate after centrifugation of the mild acid hydrolysis executed on an aliquot of R-LPS. Tetra LipA and Penta LipA indicate the acylation degree. The asterisk indicates the ion peak (m/z 1688.6) whose proposed structure is sketched.

5.2.3 Structural characterisation of the core OS by NMR spectroscopy and MALDI-TOF

The complete structure of the core OS of *B. thetaiotaomicron* R-LPS was achieved by merging data from NMR spectroscopy and MALDI-TOF MS investigations executed both on the fully deacylated R-LPS (**OS1**) and on the product obtained after mild acid hydrolysis (**OS2**) of the R-LPS. Finally, the MALDI-TOF MS analysis of the intact R-LPS was then pivotal to obtain complete information regarding the heterogenous structure of *B. thetaiotaomicron* R-LPS. For the sake of simplicity, in **Scheme 1** is reported the procedure adopted for this structural characterisation.



Scheme 1 An overview of the different procedures performed and the techniques used for the full characterisation of the R-LPS from *B. thetaiotaomicron*.

5.2.4 NMR Spectroscopy of the fully deacylated R-LPS of *B. thetaiotaomicron*

To elucidate the primary structure of the saccharide portion of *B. thetaiotaomicron* R-LPS, an aliquot of pure R-LPS underwent a full deacylation.¹⁷ Following this treatment, the complete oligosaccharide, named as **OS1**, was obtained and then subject to various purification methods. **OS1** was then analysed via 1D and 2D NMR spectroscopy. All sugar units were present as pyranose rings, as proven by the ¹³C chemical shift values observable in the HSQC spectrum. The inspection of both the ¹H NMR and the HSQC spectra (**Figure 5.3** and **5.4** respectively) of the pure **OS1** showed the presence of thirteen anomeric signals assigned to thirteen different spin systems (**A-O**, **Table 5.1**). The presence of the diastereotopic H-3 methylene proton signals at δ_{H} 1.95/2.13 ppm (**K**, δ_{C} 34.2 ppm) were indicative of the occurrence of a Kdo unit.

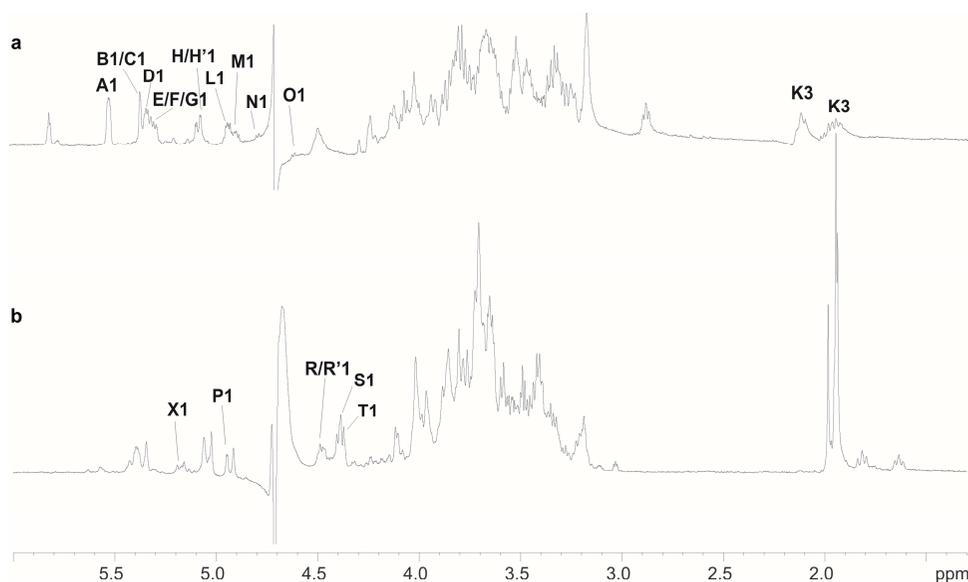


Figure 5.3 ¹H NMR spectra of **OS1** (a) and **OS2** (b) products derived from the full deacylation and mild acid hydrolysis of *B. thetaiotaomicron* R-LPS, respectively. Anomeric signals of spin systems are labelled as in **Table 5.1**. Diastereotopic methylene proton signals of Kdo residue (**K**) are also indicated. Spin system **X** is the 4-substituted α -GalpA residue that underwent β -elimination and that originated the terminal hex-4-eneuronic acid designated as **D** in **OS1** product.

At this point, the complete assignment of all spin systems was achieved as described for other projects discussed in this PhD thesis (see *Chapters 2, 3 and 4*).

Spin system **A** (H-1/C-1 at 5.51/90.3 ppm) and **L** (H-1/C-1 at 4.93/99.5 ppm, **Table 5.1**) were assigned to α -Glc

N and β -Glc

N of the lipid A on the basis of their H-2 protons that showed a correlation with two nitrogen-bearing carbon atoms at 54.3 and 55.5 ppm, respectively. The observation of an *inter*-residue connectivity between H-1 of **L** and H-6_{a,b} of **A**, identified in the NOESY spectrum (**Figure 5.5**), corroborated the assignment of residues **A** and **L** to the lipid A backbone. Likewise, the large $^3J_{\text{H,H}}$ coupling constant values, the NOE correlation between H-1 and H-2, as well as the C-2 signals resonating at 54.0 ppm indicated that residues **H** and **H'** (H-2 at 3.19 and 3.18 ppm, respectively) were both α -Glc

N. Spin systems **B**, **E**, and **F** were all identified as α -Man

residues, as proven by the small $^3J_{\text{H-1,H-2}}$ (~1.2 Hz), indicative of a H-2 equatorial orientation, and by the *intra*-residue NOE contact of H-1 only with H-2. Spin systems **C**, **M**, **N** and **O** were identified as glucose residues due to the chemical shifts of ring protons that agreed with the *gluco*-configuration of pyranose rings (**Table 5.1**, **Figure 5.4**). However, the large $^3J_{\text{H-1,H-2}}$ values, together with the NOE contacts of H-1 with H-3 and H-5, were diagnostic of the anomeric β -configuration for **M**, **N** and **O**, whereas the *intra*-residue NOE contact of H-1 with H-2, and the small $^3J_{\text{H-1,H-2}}$ coupling constant (3.3 Hz) were indicative of an α -anomeric configuration for residue **C**. Due to the presence of a correlation of H-5 signal of **G** with a carboxyl group whose carbon atom resonated at 177.3 ppm, it was possible to assign this spin system as an uronic acid. This was identified as a Glc

A due to the TOCSY correlations that ended at H-5; in addition, the *intra*-residue NOE contact of H-1 with H-2 and the $^1J_{\text{C-1,H-1}}$ and $3J_{\text{H-1,H-2}}$ values (about 171 Hz and 3.2 Hz, respectively) led to assign the anomeric α -configuration for this Glc

A residue **G**. Of note, the signal at 5.33 ppm, labeled as **D**, was identified as the anomeric proton of a hex-4-enuronic residue, which resulted from β -elimination in a 4- substituted uronic acid as a consequence of the alkaline treatment. Therefore, the olefine signal clearly visible in the HSQC spectrum at H-1/C-1 5.80/106.7 ppm (**Figure 5.4**, **D4**) was

assigned as H-4 of residue **D**. This was further confirmed by the observation that H-1 and H-4 of **D** were correlated in the TOCSY and COSY spectra via two other proton resonances (**Table 5.1**). Finally, the Kdo unit (**K**) was assigned starting from the diastereotopic methylene signals, while its anomeric α -configuration was established on the basis of the chemical shift values of H-3 and of the $^3J_{\text{H-7,H8a}}$ and $^3J_{\text{H-7,H-8b}}$ coupling constants.¹⁹

Table 5.1 ¹H, ¹³C (*Italic*), and ³¹P (**bold**) chemical shifts of the **OS1 (A-O)** and **OS2 (P-T/X)** products derived from mild acid hydrolysis and complete deacylation respectively of the R-LPS from *B. thetaiotaomicron*.

Chemical Shifts (δ)							
Unit	1	2	3	4	5	6	7 8
A	¹ H 5.51	3.25	3.78	3.27	4.07	4.13/3.77	
6- α -D-GlcN1P	¹³ C <i>90.3</i>	<i>54.3</i>	<i>69.8</i>	<i>70.0</i>	<i>72.8</i>	<i>68.7</i>	
	³¹ P 1.60						
B	¹ H 5.36	4.29	3.85	3.61	3.71	3.92/3.80	
2,6- α -D-Man	¹³ C <i>98.2</i>	<i>79.0</i>	<i>71.0</i>	<i>67.1</i>	<i>71.5</i>	<i>66.2</i>	
C	¹ H 5.36	3.36	3.69	3.40	3.95	3.82/3.61	
<i>t</i> - α -D-Glc	¹³ C <i>97.3</i>	<i>71.7</i>	<i>72.6</i>	<i>69.4</i>	<i>71.74</i>	<i>60.6</i>	
D	¹ H 5.33	3.88	4.02	5.80	-	-	
α - Δ 4,5-GalA	¹³ C <i>97.3</i>	<i>69.3</i>	<i>65.4</i>	<i>106.7</i>	<i>n.d.</i>	<i>n.d.</i>	
E	¹ H 5.31	4.49	4.10	3.62	3.65	3.65	
2,3- α -D-Man	¹³ C <i>98.4</i>	<i>75.2</i>	<i>77.0</i>	<i>67.2</i>	<i>72.1</i>	<i>60.5</i>	
F	¹ H 5.29	4.24	4.16	3.70	3.63	3.93/3.82	
2- α -D-Man3P	¹³ C <i>97.5</i>	<i>79.0</i>	<i>73.6</i>	<i>65.2</i>	<i>72.1</i>	<i>61.5</i>	
	³¹ P	0.52					
G	¹ H 5.28	3.41	3.90	3.64	4.51	-	
<i>t</i> - α -D-GlcA	¹³ C <i>97.6</i>	<i>71.5</i>	<i>72.0</i>	<i>72.8</i>	<i>75.0</i>	<i>177.3</i>	
H	¹ H 5.08	3.19	3.82	3.39	3.66	3.82	
4- α -D-GlcN	¹³ C <i>95.1</i>	<i>54.0</i>	<i>69.8</i>	<i>76.8</i>	<i>72.2</i>	<i>61.6</i>	
H'	¹ H 5.06	3.18	3.81	3.36	3.66	3.82	
<i>t</i> - α -D-GlcN	¹³ C <i>94.8</i>	<i>54.0</i>	<i>69.7</i>	<i>69.4</i>	<i>72.2</i>	<i>61.6</i>	
L	¹ H 4.93	2.89	3.52	3.47	3.54	3.53/3.44	
6- β -D-GlcN	¹³ C <i>99.5</i>	<i>55.5</i>	<i>72.1</i>	<i>69.6</i>	<i>74.3</i>	<i>62.3</i>	

M	¹ H	4.88	3.33	3.51	3.32	3.59	3.86		
2-β-D-Glc	¹³ C	101.6	77.1	75.3	69.8	74.8	60.7		
N	¹ H	4.78	3.37	3.49	3.33	3.55	3.63/3.47		
2-β-D-Glc	¹³ C	101.8	76.7	74.4	69.8	74.8	61.2		
O	¹ H	4.61	3.35	3.46	3.37	3.54	3.63		
2-β-D-Glc	¹³ C	101.9	77.9	75.6	69.5	74.4	60.6		
K	¹ H	-	-	1.95/2.13	4.49	4.24	n.d.	3.86	3.83/3.54
5-α-D-Kdo4PEtN	¹³ C	-	-	34.2	70.9	72.5	n.d.	70.9	63.3
	³¹ P				-0.80				
PEtN	¹ H	4.00	3.18						
	¹³ C	61.9	40.0						
P	¹ H	4.94	3.76	3.57	3.41	3.69	3.70		
3-α-D-GlcNAc	¹³ C	96.6	53.0	79.7	71.6	71.5	60.2		
R	¹ H	4.49	3.73	3.78	4.10	3.85	3.99/3.79		
3,6-β-D-GalNAc	¹³ C	101.5	54.6	79.9	67.5	73.9	68.9		
R'	¹ H	4.47	3.71	3.65	4.08	3.85	3.99/3.78		
6-β-D-GalNAc	¹³ C	101.3	54.6	72.7	67.5	73.9	67.7		
S	¹ H	4.39	3.32	3.69	3.41	3.37	3.56		
<i>t</i> -β-D-Glc	¹³ C	104.0	73.2	76.3	69.3	75.5	60.0		
T	¹ H	4.37	3.40	3.51	3.80	3.58	3.65		
<i>t</i> -β-D-Gal	¹³ C	103.1	70.8	74.3	70.1	75.1	60.8		
X	¹ H	5.18	4.07	4.01	4.24	4.60	-		
4-α-D-GalA	¹³ C	97.4	67.5	71.7	80.9	73.2	175.3		

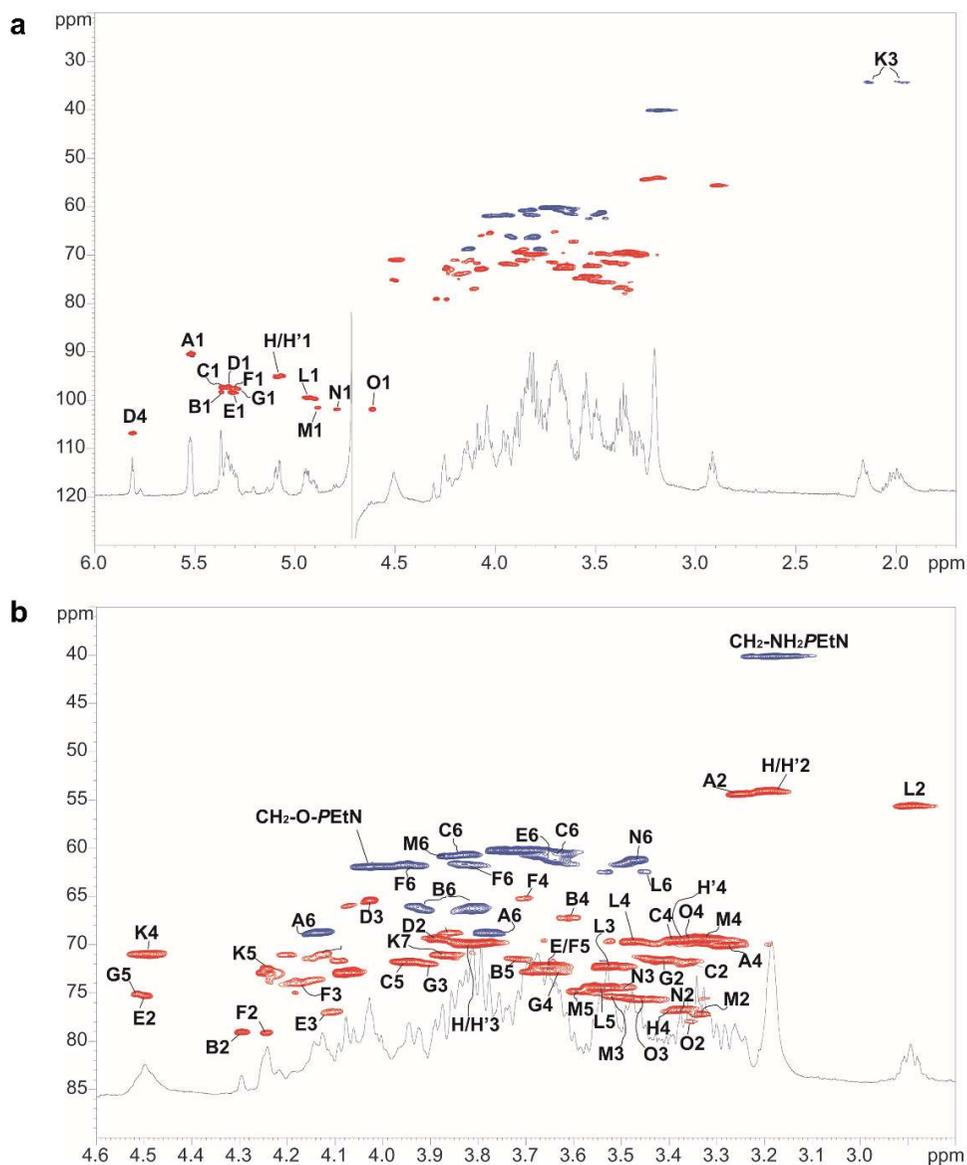


Figure 5.4 a) Superimposition of ^1H and HSQC spectra recorded for **OS1** product of *B. thetaiotaomicron* R-LPS fraction after a full deacylation treatment. b) Zoom of the carbinolic region of the overlapped ^1H and HSQC spectra.

Downfield shifted carbon signals (**Table 5.1**) were useful to identify substitution at O-2 and O-3 of α -Manp **E**, at O-2 and O-6 of α -Manp **B**, at O-2 of β -Glc p **N**, **M** and **O** and α -Manp **F**, at O-4 of α -Glc p **H**, at O-6 of α -Glc p **A** and β -Glc p **L**, and at O-5 of Kdo; on the other hand, residues **C**, **H'** and **G** were ascribed to terminal sugar units. The complete sequence of the monosaccharides composing **OS1** (**Scheme 5.2**) was deduced by using both NOE contacts attained from the NOESY spectrum and the long-range correlations visible in the HMBC spectrum. In detail, starting from the lipid A disaccharide backbone made up of residues **A** and **L**, the latter was substituted at position O-6 by the Kdo, as confirmed by the weak downfield shift of C-6 signal of β -Glc p **L** (δ 62.3 ppm, **Table 5.1**), which matches with the α -(2 \rightarrow 6) ketosidic linkage of Kdo with the non-reducing Glc p **N** units of the lipid A. The Kdo residue was, in turn, substituted at its position O-5 by the α -Manp **E**, as suggested by both the long-range correlation between the anomeric proton signal of residue **E** and the carbon atom at δ 72.5 ppm of residue **K** visible in the HMBC spectrum (**Figure 5.5b**) and by the related NOE contact between H-1 of **E** and H-5 of **K** (**Figure 5.5a**). Following *inter*-residue contacts, unit **E** was in turn substituted at O-2 by β -Glc p **O** and at O-3 by the hex-4-enuronic acid **D**. Residue **O** carried at O-2 α -Manp **F**, in turn substituted at its O-2 by α -Manp **B**. The latter was substituted at O-2 by β -Glc p **N** and at O-6 by α -Glc p **H**, in turn substituted at its O-4 by α -Glc p unit **C**, as shown by the long-range correlation of C-1 of **C** with H-4 of **H** (**Figure 5.5b**). Here, the observation that also terminal α -Glc p **H'** showed a long-range correlation with O-6 of **B** suggested that **H'** was an alternative spin system for residue **H** generated by the non-stoichiometric presence of the α -Glc p unit **C**. Finally, β -Glc p **N** was substituted at O-2 by β -Glc p **M** in turn carrying at O-2 α -Glc p **A** residue **G**.

In conclusion, the ^{31}P NMR analysis of **OS1** revealed the presence of three phosphate signals at 1.60 ppm, 0.52 ppm, and -0.80 ppm. The ^{31}P - ^1H HSQC spectrum (**Figure 5.6**) indicated correlations for the ^{31}P with ^1H signals as follows: H-1 of α -Glc p **A** of the lipid A with ^{31}P signal at δ 1.60 ppm; H-3 of α -Manp **F** with ^{31}P signal

at δ 0.52 ppm, and H-4 of Kdo with ^{31}P signal at δ -0.80 ppm. The latter ^{31}P signal also correlated with methylene proton resonances at δ_{H} 4.00 and 3.18 ppm, which were diagnostic for the presence of a 2-aminoethyl phosphate (*PEtN*) unit; thus, it was concluded that a *PEtN* decorated the Kdo residue at its position O-4 (**Scheme 5.2**). In support to this structural interpretation, the chemical shifts of the carbon resonances visible in the HSQC spectrum (C-1 of **A** 90.3 ppm, C-3 of **F** 73.6 ppm, and C-4 of **K** 70.9 ppm, **Table 5.1**) were clearly shifted to lower field due to substitution with a phosphate group.

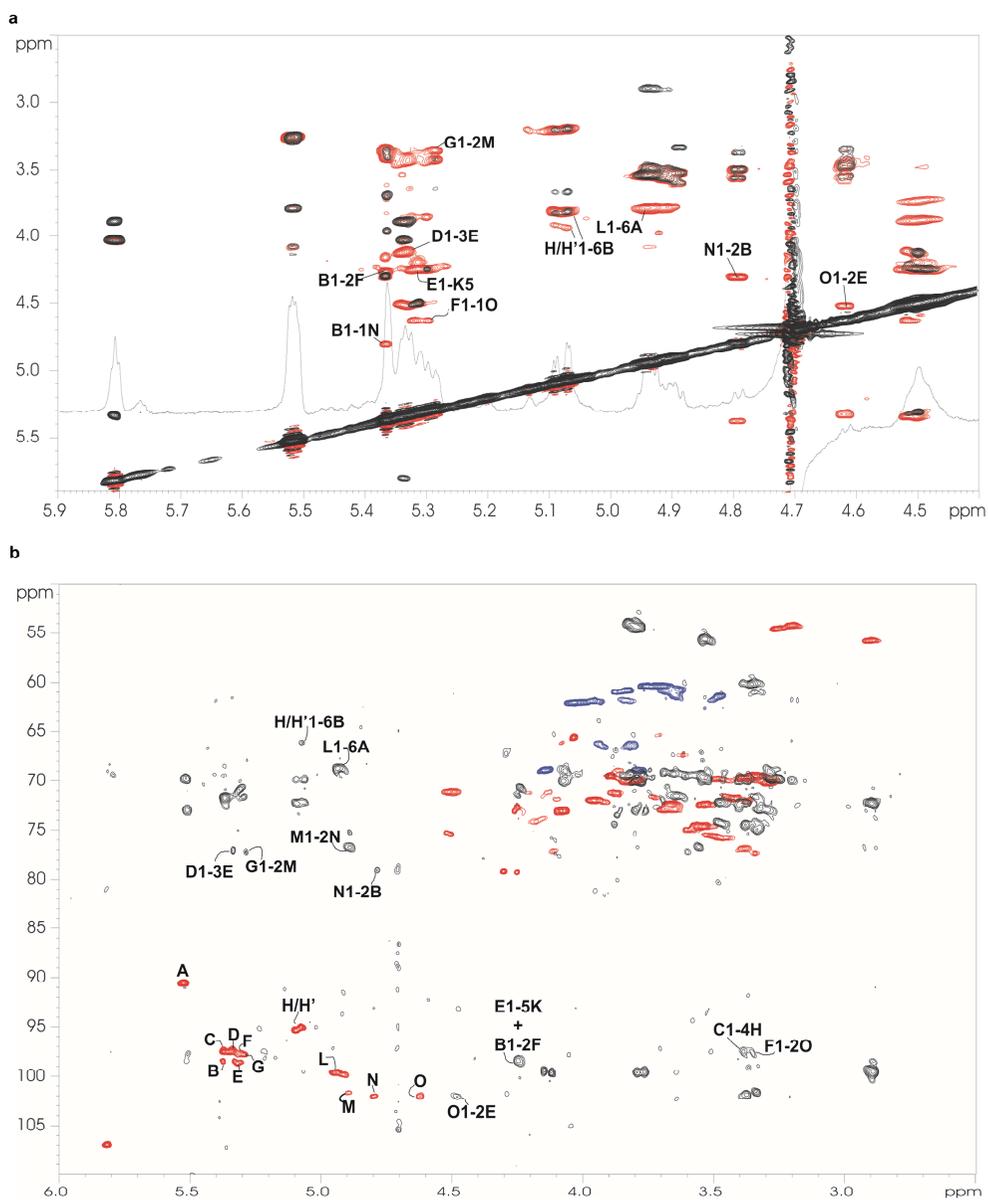


Figure 5.5 a) A zoom of the overlapped ^1H , NOESY (red) and TOCSY (black) spectra of the OS1 product. The *inter*-residue NOE correlations that provided structural information on the primary sequence of the *B. thetaiotaomicron* core OS after complete deacylation are indicated. b) Superimposition of ^1H and HMBC spectra recorded for OS1.

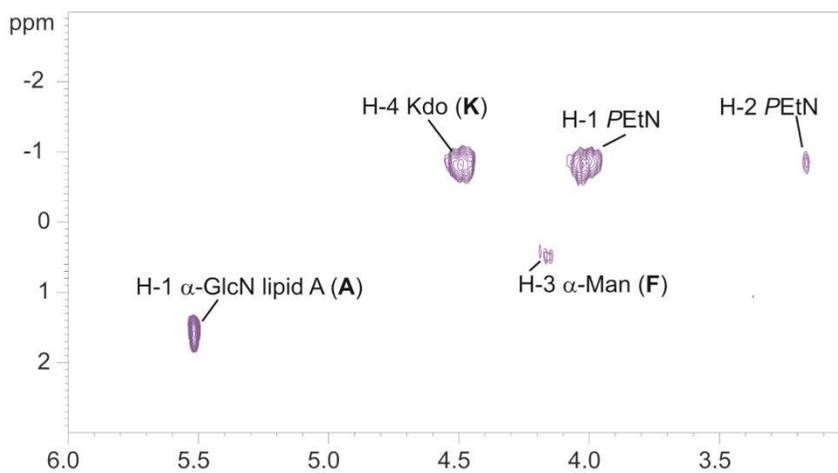


Figure 5.6 Section of $^1\text{H}, ^{31}\text{P}$ HSQC spectrum of **OS1**. The localisation of the phosphate and the 2-aminoethyl phosphate (*PEtN*) group is indicated. Anomeric signals of spin system are designated as in **Table 5.1**.

5.2.5 NMR on product obtained after mild acid hydrolysis (OS2) of the R-LPS of *B. thetaiotaomicron*

To establish the structure of the saccharide component that was lost by β -elimination that occurred upon strong alkaline treatment (of the full deacylation reaction), an aliquot of R-LPS was also subjected to a mild acid hydrolysis. The so-obtained product (**OS2**) underwent a complete 1D and 2D NMR. Despite the great heterogeneity observed in the product, the NMR analysis of **OS2** allowed to identify a tetrasaccharide as the portion that was lost after the alkaline treatment (**Scheme 5.2**, coloured in red).

Briefly, five additional anomeric proton signals were identified in the ^1H NMR spectrum compared to those found for **OS1** and they were assigned to five different spin systems (**P-T**, **Figure 5.3b** and **Table 5.1**). **R**, **R'**, and **T** were identified as *galacto*-configured residues based on their $^3J_{\text{H-3,H-4}}$ and $^3J_{\text{H-4,H-5}}$ values (3 and 1 Hz, respectively) and all with an anomeric β -configuration due to the NOE correlations occurring between H-1 and H-3 and H-5. Finally, the occurrence of signals for C-2 of **R** and **R'** at 51.0 ppm and 54.6 ppm (**Table 5.1**) indicated that C-2 of these sugar residues were nitrogen-bearing carbon atoms, thus allowing their final assignment as β -GalpN units, while **T** was identified as a β -Galp. The observation of the downfield displacement of carbon signals led to the assignment of **R** as a 3,6-disubstituted β -GalpN, **R'** as a 6-substituted β -GalpN, whereas **T** was identified as a terminal β -Galp. Likewise, spin systems **P** and **S** were attributed to *gluco*-configured units due to their high $^3J_{\text{H,H}}$ ring proton values; in addition, due to the occurrence of a correlation of H-2 of **P** with a nitrogen-bearing carbon signal at 53.0 ppm, it was possible to identify this sugar unit as an α -Glc pN, whereas **S** was assigned to a β -Glc p due to the chemical shifts of H-1 and C-1 (4.37 and 103.1 ppm respectively) and the *intra*-residue NOE correlation of H-1 with H-3 and H-5. Low field shift of C-3 of **P** allowed its identification as a 3-substituted α -Glc pN, whereas **S** was identified as a terminal sugar

unit. By analysing the *inter*-residue NOE contacts visible in the transverse ROESY (T-ROESY) spectrum (**Figure 5.7c**), it was possible to deduce the sequence **T1**→**3R1**→**3P1**, with **R** also carrying at O-6 β -GlcP **S**. Spin system **R'** was identified as an alternative spin system for **R** originated by the non-stoichiometric presence of β -Galp **T** at its O-3 position. Further on, it was possible to observe a clear *inter*-residue NOE contact between H-1 of **P** and H-4 of spin system designated as **X** ($\delta_{H/C}$ 5.18/97.4 ppm, **Figure 5.7c**, **Table 5.1**). The latter was identified as a 4-substituted α -GalpA as proven by the chemical shifts, the $^3J_{H-3,H-4}$ and $^3J_{H-4,H-5}$ values (3.6 and 2 Hz, respectively), and the low-field displacement of C-4 (80.9 ppm, **Table 5.1** and **Figure 5.7b**). It was reasonable to assign residue **X** as the uronic acid that underwent β -elimination after the alkaline treatment leaving the terminal hex-4-eneuronic acid designated as **D** in **OS1** product (**Table 5.1**, **Figure 5.3**). Finally, ^{31}P NMR analysis of **OS2** product showed no signals attributable to the presence of phosphate groups that were lost due to the acid treatment. In agreement, it was possible to analyse **OS2** product by MALDI-TOF MS only in the positive-ion mode, as explained below. Therefore, by combining data attained from both the fully deacylated (**OS1**) and the mild acid hydrolysis (**OS2**) products, it was possible to deduce the structure of the core OS from *B. thetaiotaomicron* R-LPS, as reported in **Scheme 5.2**.

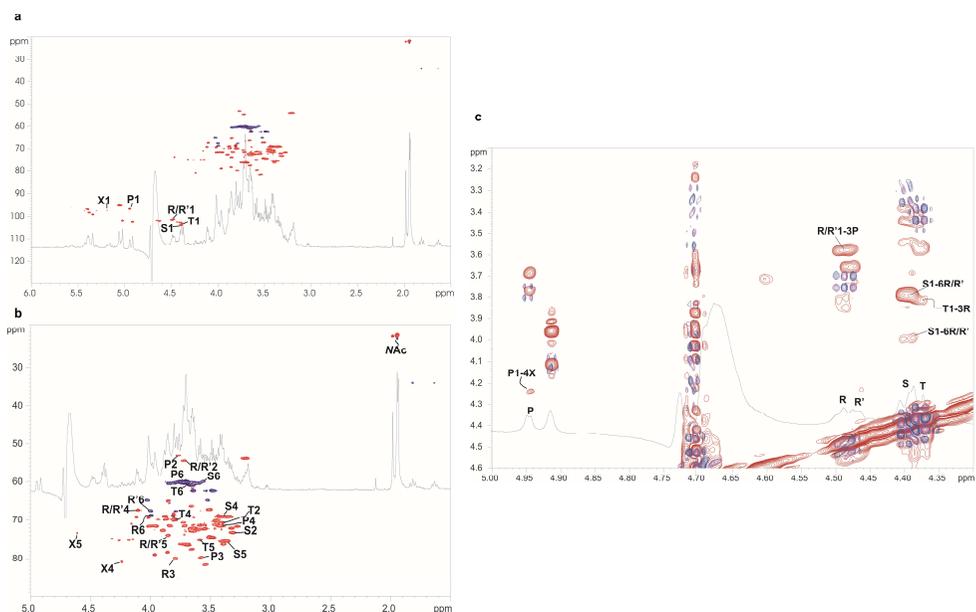
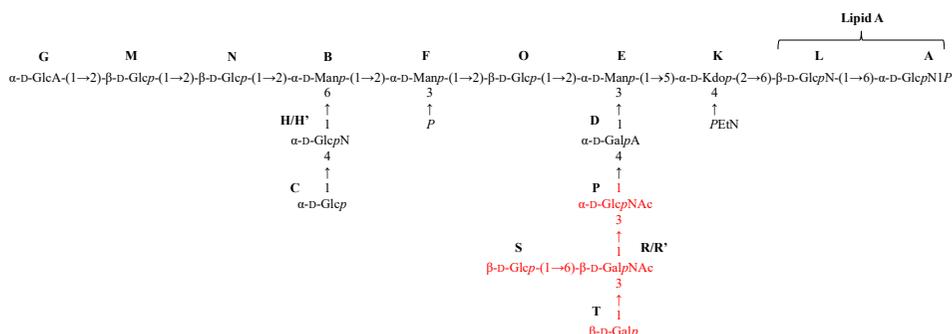


Figure 5.7 a) Superimposition of ^1H and HSQC spectra recorded for **OS2** product of *B. thetaiotaomicron* R-LPS fraction after mild acid treatment. b) Zoom of the carbinolic region of the ^1H and HSQC spectra. Key one-bond heteronuclear correlations are indicated and numbered as reported in **Table 5.1**. Densities related to the acetyl groups (NAc) have been also indicated. c) Section of the overlapped ^1H , tROESY (red) and COSY (purple/blue) spectra of the **OS2** product. The *inter*-residue NOE contacts that were essential to define the primary sequence of the tetrasaccharide lost after strong alkaline treatment are indicated. Spin system **X** is the 4-substituted α -GalpA residue that was subjected to β -elimination and resulting in the terminal hex-4-eneuronic acid designated as **D** in **OS1** product.



Scheme 5.2 Structural assessments of the **OS1** and **OS2** products. The complete structure of the core OS from *B. thetaiotaomicron* was achieved by combination of data attained from the NMR investigation of the LPS after full deacylation (yielding **OS1**) and mild acid hydrolysis (yielding **OS2**) procedures. Coloured in red is reported the tetrasaccharide portion of the core OS that was lost due to the β -elimination occurred as a consequence of the alkaline conditions used to deacylate the R-LPS. The core OS structure is labelled using letters in **Table 5.1**. “p” indicates the pyranose form of sugar residues, “P” indicates the phosphate group, and PEtN stands for 2-aminoethyl phosphate.

5.2.6 MALDI-TOF MS analysis

A positive-ion MALDI-TOF MS analysis of **OS2** (**Figure 5.8a**) was also performed. Briefly, the MALDI-TOF MS spectrum showed a dominant species at m/z 2296.7 $[M+Na]^+$ (**OS_A**), which was interpreted as the sodium adduct of an oligosaccharide made up of 7 hexoses (Hex), 2 hexuronic acids (HexA), 2 *N*-acetyl hexosamines (HexNAc), one HexN, and one Kdo (in its anhydro form), which was in agreement with **OS2** structure (**Scheme 5.2**), deduced by NMR analysis, minus two Hex units. This was likely due to the non-stoichiometric presence of α -Glcp (labelled as **C**, Scheme 2) and β -Galp (designated as **T**, **Scheme 5.2**) on the core OS. In line with this hypothesis, a minor peak at m/z 2458.8 $[M+Na]^+$ (**OS_B**) was corresponded to an oligosaccharide having the same sugar composition of **OS_A** plus an additional Hex unit. Of note, other core OS species (**OS_C**-**OS_H**) were identified and attributed to structural heterogeneities within the core OS structure.

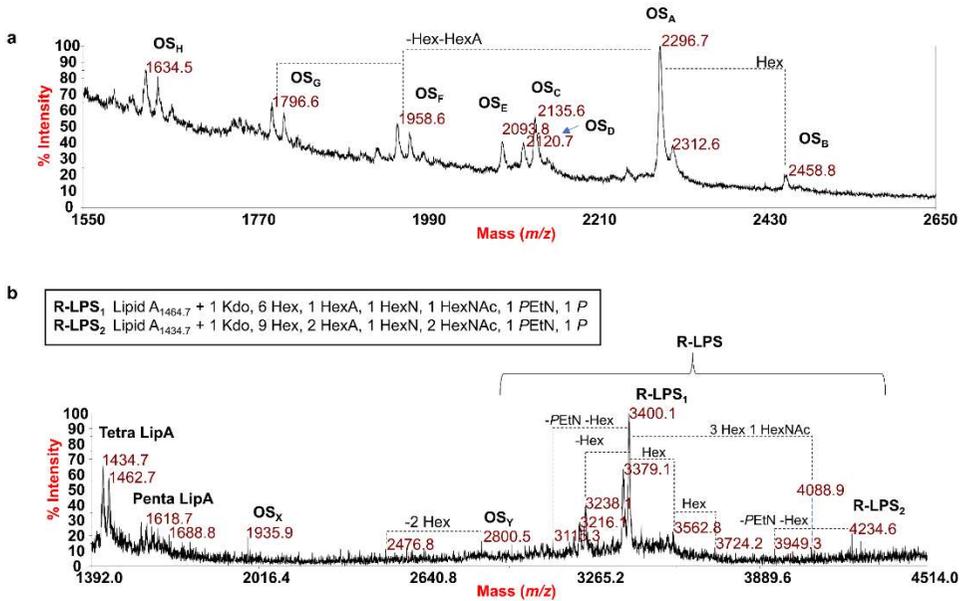


Figure 5.8 a) Positive-ion MALDI-TOF MS spectrum of OS₂ product obtained after mild acid hydrolysis of *B. thetaiotaomicron* R-LPS. The structural heterogeneities within the core OS structure have been depicted as OS_A-OS_H **b)** Negative-ion MALDI-TOF MS spectrum of intact R-LPS from *B. thetaiotaomicron*, recorded in linear mode. R-LPS molecular ions (R-LPS₁ and R-LPS₂) and their ion fragments, attributable to the core OS (OS_X, OS_Y) and lipid A (Tetra and Penta LipA) species are visible and indicated in the spectrum.

5.2.7 Structural characterisation by MALDI mass spectrometry of the intact R-LPS

To a complete characterisation of the structure of *B. thetaiotaomicron* R-LPS, an aliquot of intact R-LPS was analysed by negative-ion MALDI-TOF MS (**Figure 5.8b**). This analysis was key to fully dissect the structure of such a complex R-LPS. In particular, the spectrum showed at higher molecular masses (between m/z 3100 and 4500) a series of $[M-H]^-$ ions corresponding to the whole R-LPS mixture, composed of species differing by the composition of both the acyl and the sugar moieties. Two main peak clusters were identified at about m/z 3400.1 and m/z 4234.6, labelled as (**R-LPS₁** and **R-LPS₂**) respectively. These species were assigned as R-LPS differing in the length of the carbohydrate portion, that corresponds to what was observed in the SDS-PAGE analysis (**Figure 5.1**). In addition, at the lower mass ranges (m/z 1900-2800), these peaks come from ion fragments originating from the cleavage of the labile glycosidic bond between Kdo and the lipid A. This type of insource fragmentation (β -elimination) yielding both oligosaccharide(s) and lipid A ions was key to further appreciate the heterogeneity of both the lipid A and the core OS of *B. thetaiotaomicron* R-LPS beside confirming the NMR-derived structure.²⁰ Indeed, the main ion fragments at these lower mass ranges, were in agreement with previous studies and with our MS inspection of the isolated lipid A fraction (see *Section 5.2.2*).^{10,14} Moreover, it was possible to identify two core OS fragments at m/z 1935.9 (**OS_X**) and m/z 2800.5 (**OS_Y**). In detail, **OS_X** species detected at m/z 1935.9 was interpreted as an oligosaccharide composed of 6 Hex, 1 HexA, 1 HexN, 1 HexNAc, 1 Kdo, 1 phosphate, and 1 PEtN, whereas **OS_Y** detected at m/z 2800.5 was inferred as a core OS made up of 9 Hex, 2 HexA, 1 HexN, 2 HexNAc, 1 Kdo, 1 phosphate, and 1 PEtN, thus perfectly matching with the one identified by merging NMR data of **OS1** and **OS2** (**Scheme 2**). A minor species at m/z 2476.8 was assigned to a core OS composed of 7 Hex, 2 HexA, 1 HexN, 2 HexNAc, 1 Kdo, 1 phosphate, and 1 PEtN, that is an oligosaccharide likely devoid of the two non-stoichiometric Hex residues. Finally, the main R-LPS species were consistent with a combination of **OS_X** and **OS_Y** ion peaks and differently acylated lipid A species. Through the merging of NMR and MS

information it was possible to define the complete structure of *B. thetaiotaomicron* R-LPS which has been sketched in **Figure 5.9**, and here with a penta-acylated lipid A species.

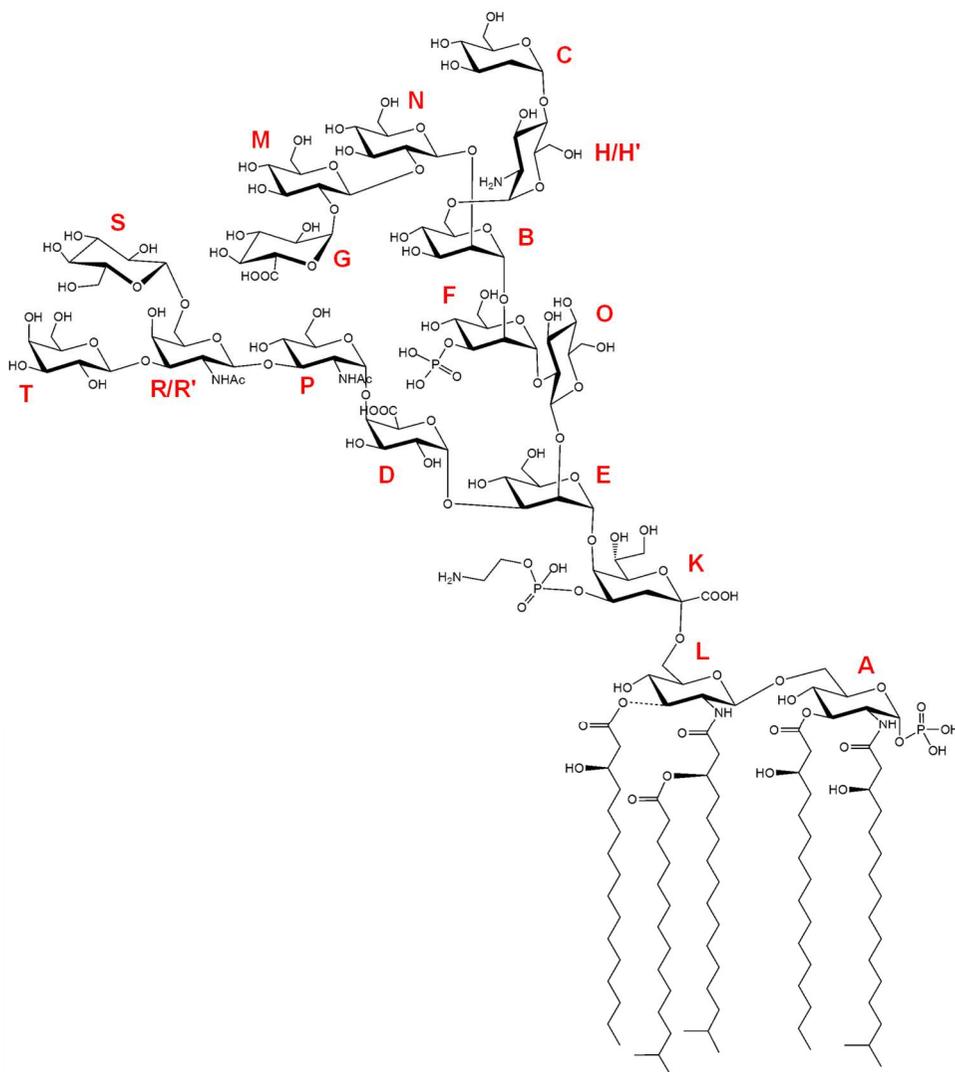


Figure 5.9 Chemical structure of *B. thetaiotaomicron* R-LPS. The letter labels used for NMR investigations are as reported in **Table 5.1**, **Scheme 5.2**. *N*-Acetyl groups were deduced by NMR analysis of the R-LPS product after mild acid hydrolysis (**OS2**). The lipid A is heterogeneous, composed of a mixture of *mono*-phosphorylated penta- and tetra-acylated species. The dotted lines indicate the non-stoichiometric substitution.

5.3 Immunological properties of *B. thetaiotaomicron* R-LPS

5.3.1 Recognition by TLR4/MD-2 complex

Given the structural similarities of lipid A from *B. thetaiotaomicron* and *B. vulgatus* (Chapter 3), it is tempting to hypothesise that also *B. thetaiotaomicron* R-LPS might act providing only a weak pro-inflammatory intracellular transcription program, without exceeding an inflammatory threshold, that could exert a “protective” role for the colonised host. On the basis of this hypothesis, the immunological activity of *B. thetaiotaomicron* R-LPS was established *in vitro* in HEK-Blue™ cells stably transfected with human TLR4, MD-2, and CD14 genes. The NF-κB activation was used as the read-out of this experiment. HEK-Blue™ hTLR4 cells were stimulated with different concentrations (1, 10, and 100 ng/mL) of *B. thetaiotaomicron* R-LPS and compared with the pro-inflammatory *S. typhimurium* SH 2201 S-LPS. The results, presented in **Figure 5.10a** showed that *B. thetaiotaomicron* R-LPS induced a significantly lower NF-κB activation compared to *S. typhimurium* S-LPS (*B. thetaiotaomicron* R-LPS vs. *S. typhimurium* S-LPS, $p < 0.01$ at 1 ng/mL and 10 ng/mL, $p < 0.05$ at 100 ng/mL).

In addition, we analysed whether *B. thetaiotaomicron* R-LPS could stimulate the TLR2-mediated signalling, as previously observed for LPS of other *Bacteroides* species, such as *B. vulgatus* (Chapter 3).^{14,21} Therefore, HEK-Blue™ hTLR2 cells were stimulated with *B. thetaiotaomicron* R-LPS and with S-LPS from *S. typhimurium* at different concentrations and analysed as above. It was visible that *B. thetaiotaomicron* R-LPS can activate the TLR2 in a potent dose-dependent manner, which by contrast was not observed upon stimulation with *S. typhimurium* S-LPS (**Figure 5.10b**). To evaluate if the activation of the cells was dependent on TLR receptors, cells without TLR4/MD-2/CD14 or TLR2 expression, known as HEK-Blue™ Null2™, were used. These results revealed a lack of any NF-κB activation in HEK-Blue™ Null2™ cells, indicated that NF-κB activation by *B. thetaiotaomicron* R-LPS were TLR-dependent.

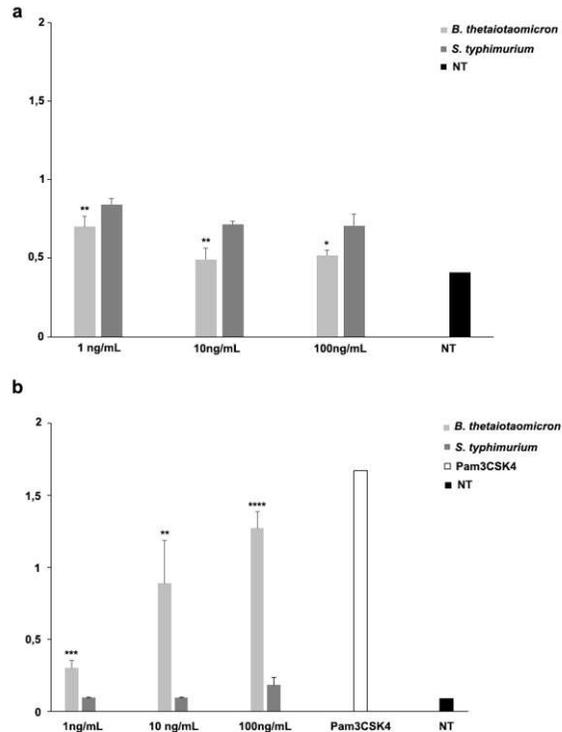


Figure 5.10 a) Stimulation of HEK BlueTM hTLR4 and **b)** HEK BlueTM hTLR2 cells. SEAP levels (OD) upon stimulation with R-LPS of *B. thetaiotaomicron* at 1, 10, and 100 ng/mL; *S. typhimurium* SH 2201 S-LPS was used as a positive control. Significant differences between *B. thetaiotaomicron* R-LPS and *S. typhimurium* S-LPS values are indicated. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$ by the Student *t*-test. NT, not treated cells. Data are expressed as mean \pm SD of three independent experiments in triplicate.

Through collaboration with Prof. Amoresano at the University of Naples, Federico II and a novel Multiple Reaction Monitoring Mass Spectrometry (MRM/MS)based method recently developed by her group, more information was gained about the cytokine release profile upon stimulation of HEK293 cells.²² Five proteins were selected for this study, interferon γ (IFN- γ), IL-2, IL-6, IL-10 and IL-1 β , which were then monitored in LC-MS/MS in MRM ion mode through the selection of 28 proteotypic peptides and the monitoring of 141 precursor ion to product ion transitions. To compare the protein expression along all samples, the peak area of each

peptide was averaged and the area values of stimulated HEK-Blue™ hTLR2 and HEK-Blue™ hTLR4 cells were compared to untreated HEK-Blue™ hTLR2 and HEK-Blue™ hTLR4 cells, respectively. The results were reported as a fold change respect to untreated HEK-Blue™ hTLR2 and HEK-Blue™ hTLR4 cells. This approach was adopted to analyse both the supernatants and the cells upon stimulation with *B. thetaiotaomicron* R-LPS to evaluate both the extracellular and the intracellular cytokine profile.²³

The result of this preliminary immunological study shows that *B. thetaiotaomicron* R-LPS exerted a dose-dependent decrease in the production of IL-1 β , IL-6, and IL-2 upon stimulation of HEK-Blue™ hTLR2 cells (**Figure 5.11a**), and only a small amount of IL-10 was released, with 100 ng *B. thetaiotaomicron* R-LPS exerting the strongest IL-10 production. By contrast, a dose-dependent increase followed by a drastic decrease of INF- γ expression (at 100 ng) was observed in the supernatants upon stimulation of HEK-Blue™ hTLR2 cells with *B. thetaiotaomicron* R-LPS. When MRM/MS was applied to directly analyse HEK-Blue™ hTLR2 cells, an opposite trend was observed.

Focusing on the HEK-Blue™ hTLR4 cells, this analysis revealed that the expression trend of IL-1 β , IL-6, and IL-2 (**Figure 5.11b**) in supernatants was maintained relatively constant regardless the dose used. On the other hand, a dose-dependent increase of IL-10 release and INF- γ expression IL-1 β , IL-6, and IL-2 was observed. Also in this case, the MS analysis of the expression trend of IL-10, IL-6, IL-2, and INF- γ directly on cells reflected the opposite of what was observed in the supernatants.

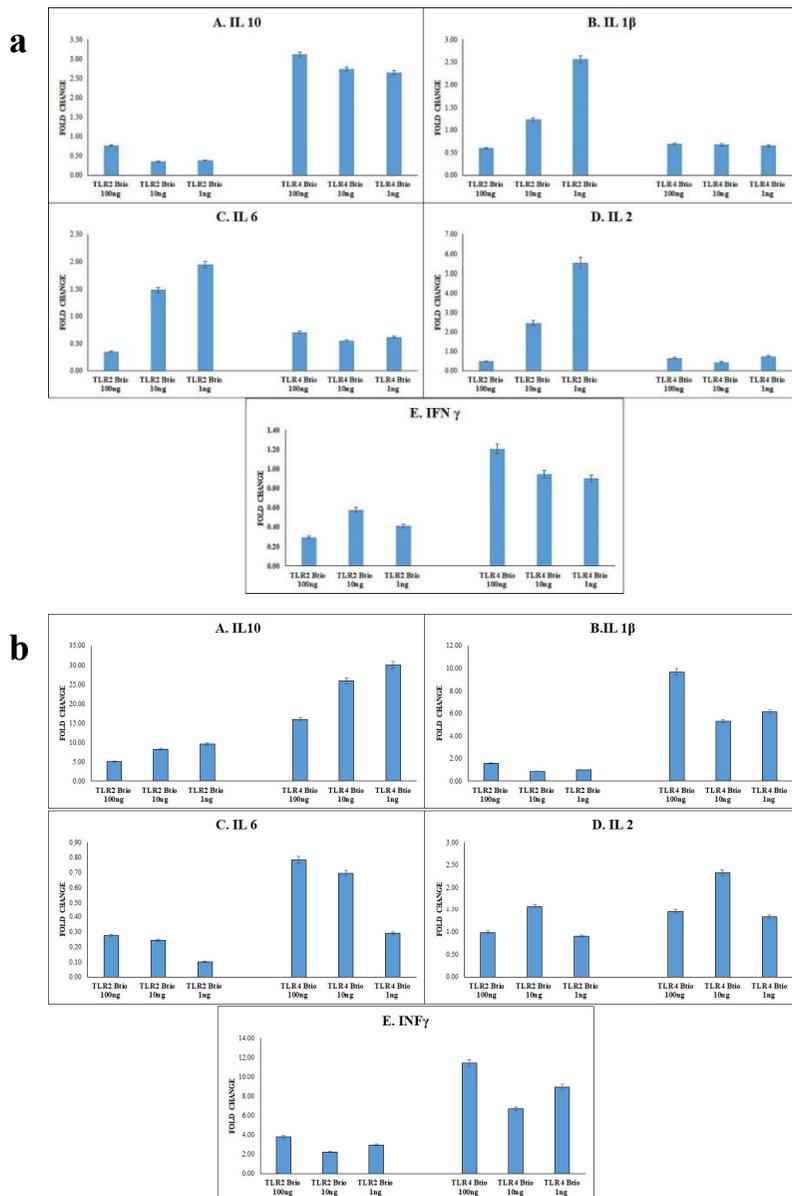


Figure 5.11 LC-MRM-MS fold change histograms of 5 different cytokines: IL-10, IL-1 β , IL-6, IL-2, and IFN- γ . **a)** Observed in the supernatants of HEK Blue™ hTLR4 and HEK Blue™ hTLR2 cells after stimulation with different amount (1, 10 and 100 ng) of *B. thetaiotaomicron* R-LPS (Btio). **b)** Observed directly on HEK Blue™ hTLR4 and HEK Blue™ hTLR2 cells after stimulation with different amount (1, 10 and 100 ng) of *B. thetaiotaomicron* R-LPS (Btio). Results are expressed as Fold change with a variation coefficient of (CV%) around 3 %.

5.4 Conclusions

This multifaceted project concluded the complex and heterogeneous structure of the R- LPS from one of the most famous gut-derived Gram-negative, *B. thetaiotaomicron*, which was obtained by taking advantage of an acapsular mutant previously described by Rogers *et al.*²⁴ and by Jacobson *et al.*¹⁰ The lipid A was confirmed to be a heterogeneous blend of *mono*-phosphorylated tetra- and penta-acylated species, which matched with previously reported lipid A structures from both *B. thetaiotaomicron* and other *Bacteroides* sp. (see also *Chapter 3*).^{10,14,15} The OS moiety was analysed through NMR and MALDI-TOF MS of either the full R-LPS or following specific chemical degradations.

It was highlighted the extreme complexity of the carbohydrate portion of *B. thetaiotaomicron* R-LPS, which can be described as made up of two oligosaccharides that branch off from the first sugar linked to the Kdo, which was an α -Manp residue. The inner core motif in most LPS comprises linear or branched sequences of L(D)-glycero-D(L)-manno-heptosyl residues linked to Kdo. These heptoses are often decorated with non-carbohydrate substituents, such as phosphate, PEtN, pyrophosphate, and carbamoyl groups. In the case of this R-LPS, no heptoses were found and the α -Manp residue bears none of the above substituents. The absence of heptoses or phosphorylated heptosyl residues in the core OS of *B. thetaiotaomicron* R-LPS is, however, not unprecedented and agrees with previous studies on other species of the Bacteroidales order.^{14,25} Still, this R-LPS is highly negatively charged due to the presence of two uronic acids and a phosphate group. By contrast, two neutral amino hexoses and a positively charged GlcpN were also found at the two opposite sides of the core OS, respectively. The presence of a GlcpN is noteworthy as the amino group is rarely found in the free form, while it is commonly found capped with an acetyl moiety. Therefore, this uncapping confers a positive charge at the outermost part of the core OS, likely balancing the negative charge of the carboxyl group of the close glucuronic acid. Likewise, the occurrence of a PEtN residue linked

to the Kdo might serve to decrease the negative charge in proximity of the bacterial membrane. The tendency of masking negative charges of LPS is a well-known strategy adopted by some pathogenic bacteria to hide themselves from the immune surveillance,^{26,27} which also entails the release of cationic antimicrobial peptides (CAMPs). Due to their cationic nature, CAMPs are attracted by negatively charged LPS, which, once bound to CAMPs, are no longer able to trigger TLR-mediated response.^{28,29} Therefore, as previously suggested for hypo-phosphorylation of *B. thetaiotaomicron* lipid A,³⁰ the partial masking of negative charges on the core OS might contribute to enhance bacterial resistance to these antimicrobial peptides, allowing *B. thetaiotaomicron* resilience in the intestine. In addition, the complexity of the carbohydrate part of this R-LPS raises questions about the significance it might have on the whole molecule recognition by the human immune system. As example, it was previously demonstrated that the LPS from *B. vulgatus* (Chapter 3) has a selective and potent affinity with DC-SIGN, a carbohydrate-binding protein associated with the gut lymphoid tissue, which might favour bacterial colonisation and persistence of *B. vulgatus* in our body.³¹ Future studies will be performed to shed light on this crucial interplay.

This chapter also presents the results from an initial structure to function study to verify whether the chemistry of the R-LPS of *B. thetaiotaomicron* is reflected in any peculiar immunological behaviour that might be relatable to the ability of this bacterium to finely tune the host immune responses. Indeed, *B. thetaiotaomicron* R-LPS showed it is a weaker activator of the TLR4-mediated immune response compared to the S-LPS from the enteropathogenic *S. typhimurium*, while it was able to activate the TLR2-dependent signalling unlike *S. typhimurium* S-LPS. Moreover, by employing a recently developed method based on MRM/MS²², the differential cytokine release profile was evaluated upon stimulation with *B. thetaiotaomicron* R-LPS of HEK293 cells transfected with TLR2 or TLR4 genes. By analysing the pattern of cytokines in two different TLR cellular models, it was possible to establish that

each TLR, on equal terms of stimulation time, elicits a distinctive type of response to *B. thetaiotaomicron* R-LPS stimulation. Interestingly, the release of the cytokines analysed in this study, except for IL-10, was stronger in the supernatants of transfected hTLR2 cells than in hTLR4 cells, whereas a significant dose-dependent increase of IL-10 production and INF- γ expression was observed in the supernatants of transfected hTLR4 cells. By contrast, IL-1 β , IL-2 and IL-6 released by hTLR4 cells remained almost constant regardless of the dose. TLR2 and TLR4 are known to be involved in immune response activating pro-inflammatory cytokine production by NF- κ B. However, an increasing number of studies have described a critical role for TLR2 and TLR4 in the prevention of intestinal inflammation.^{32,33} with a particular attention to the regulation of the anti-inflammatory cytokine IL-10 production. In this frame, it has been demonstrated that resident gut bacteria stimulate IL-10 release through TLR2 activation resulting in suppression of aggressive immune responses and preservation of mucosal homeostasis.³⁴ Therefore, it was interesting to observe that, in presence of *B. thetaiotaomicron* R-LPS, the production of this anti-inflammatory mediator almost entirely relied on TLR4 activation. Unfortunately, contrary to TLR4 pro-inflammatory signaling, the anti-inflammatory response mechanisms induced by this TLR activation are much less characterised. Further investigation is therefore required on the possible differential role of *B. thetaiotaomicron* R-LPS in TLR4-mediated anti-inflammatory responses as well as on the impact of TLR2 activation in the intracellular signalling activated by this R-LPS.

5.5 References

1. Zafar, H. & Saier, M. H. Gut *Bacteroides* species in health and disease. *Gut Microbes* **13**, 1–20 (2021).
2. Wexler, A. G. & Goodman, A. L. An insider's perspective: *Bacteroides* as a window into the microbiome. *Nat. Microbiol.* **2017** 2:5 **2**, 1–11 (2017).
3. Hooper, L. V., Stappenbeck, T. S., Hong, C. V. & Gordon, J. I. Angiogenins: a new class of microbicidal proteins involved in innate immunity. *Nat Immunol.* **2003** 4:3 **4**, 269–273 (2003).
4. Zocco, M. A., Ainora, M. E., Gasbarrini, G. & Gasbarrini, A. *Bacteroides thetaiotaomicron* in the gut: Molecular aspects of their interaction. *Dig Liver Dis.* **39**, 707–712 (2007).
5. Porter, N. T., Luis, A. S. & Martens, E. C. *Bacteroides thetaiotaomicron*. *Trends Microbiol.* **26**, 966–967 (2018).
6. Frank, D. N., St. Amand, A. L., Feldman, R. A., Boedeker, E. C., Harpaz, N. & Pace, N. R. Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases. *PNAS*, **104**, 13780 (2007).
7. Martens, E. C., Roth, R., Heuser, J. E. & Gordon, J. I. Coordinate Regulation of Glycan Degradation and Polysaccharide Capsule Biosynthesis by a Prominent Human Gut Symbiont. *JBC*, **284**, 18445–18457 (2009).
8. Hsieh, S. A. & Allen, P. M. Immunomodulatory Roles of Polysaccharide Capsules in the Intestine. *Front. Immunol.* **11**, 690, (2020)
9. Pither, M.D, Molinaro, A., Silipo, A. & Di Lorenzo, F. in *Methods Mol. Biol.* **2613**, (Springer International Publishing). **In press**
10. Jacobson, A. N., Choudhury, B. P. & Fischbach, M. A. The biosynthesis of lipooligosaccharide from *Bacteroides thetaiotaomicron*. *mBio* **9**, (2018).
11. Westphal, O. & Jann, K. Bacterial Lipopolysaccharides Extraction with Phenol-Water and Further Applications of the Procedure. *Methods in Carbohydrate Chemistry* **5**, 83–91 (1965).
12. Kittelberger, R. & Hilbink, F. Sensitive silver-staining detection of bacterial lipopolysaccharides in polyacrylamide gels. *J. Biochem. Biophys. Methods* **26**, 81–86 (1993).
13. De Castro, C., Parrilli, M., Holst, O. & Molinaro, A. Microbe-associated molecular patterns in innate immunity: Extraction and chemical analysis of Gram-negative bacterial lipopolysaccharides. *Meth enzymol.* **480**, 89–115 (2010).
14. Di Lorenzo, F., Pither, M. D., Martufi, M., Scarinci, I., Guzmán-Caldentey, J., Łakomic, E., Jachymek, W., Bruijns, S. C. M., Santamaría, S. M., Frick, J.-S., van Kooyk, Y., Chiodo, F., Silipo, A., Bernardini, M. L. & Molinaro, A. Pairing

- Bacteroides vulgatus* LPS Structure with Its Immunomodulatory Effects on Human Cellular Models. *ACS Cent. Sci* **6**, 1602–1616 (2020).
15. Berezow, A. B., Ernst, R. K., Coats, S. R., Braham, P. H., Karimi-Naser, L. M. & Darveau, R. P. The structurally similar, penta-acylated lipopolysaccharides of *Porphyromonas gingivalis* and *Bacteroides* elicit strikingly different innate immune responses. *Microbial Pathogenesis* **47**, 68–77 (2009).
 16. Weintraub, A., Zähringer, U., Wollenweber, H. W., Seydel, U. & Rietschel, E. T. Structural characterization of the lipid A component of *Bacteroides fragilis* strain NCTC 9343 lipopolysaccharide. *Eur. J. Biochem* **183**, 425–431 (1989).
 17. Holst, O. Deacylation of lipopolysaccharides and isolation of oligosaccharide phosphates. *Methods mol. biol. (Clifton, N.J.)* **145**, 345–353 (2000).
 18. Speciale, I., Notaro, A., Garcia-Vello, P., Di Lorenzo, F., Armiento, S., Molinaro, A., Marchetti, R., Silipo, A. & De Castro, C. Liquid-state NMR spectroscopy for complex carbohydrate structural analysis: A hitchhiker’s guide. *Carbohydrate Polymers* **277**, 118885 (2022).
 19. Birnbaum, G. I., Roy, R., Brisson, J.-R. & Jennings, H. J. Conformations of Ammonium 3-Deoxy-D- *manno* -2-octulosonate (KDO) and Methyl α - and β -Ketopyranosides of KDO: X-Ray Structure and ^1H NMR Analyses, *J. Carbohydr. Chem* **6**, 17–39 (1987).
 20. Sturiale, L., Garozzo, D., Silipo, A., Lanzetta, R., Parrilli, M. & Molinaro, A. New conditions for matrix-assisted laser desorption/ionization mass spectrometry of native bacterial R-type lipopolysaccharides. *RCM* **19**, 1829–1834 (2005).
 21. Alhawi, M., Stewart, J., Erridge, C., Patrick, S. & Poxton, I. R. *Bacteroides fragilis* signals through Toll-like receptor (TLR) 2 and not through TLR4. *J. Med. Microbiol.* **58**, 1015–1022 (2009).
 22. Illiano, A., Pinto, G., Gaglione, R., Arciello, A. & Amoresano, A. Inflammation protein quantification by multiple reaction monitoring mass spectrometry in lipopolysaccharide-stimulated THP-1 cells. *RCM* **35**, (2021).
 23. Serban, A. I., Stanca, L., Geicu, O. I. & Dinischiotu, A. AGEs-Induced IL-6 Synthesis Precedes RAGE Up-Regulation in HEK 293 Cells: An Alternative Inflammatory Mechanism? *Int. J. Mol. Sci.* **16**, 20100–20117 (2015).
 24. Rogers, T. E., Pudlo, N. A., Koropatkin, N. M., Bell, J. S. K., Moya Balasch, M., Jasker, K. & Martens, E. C. Dynamic responses of *Bacteroides thetaiotaomicron* during growth on glycan mixtures. *Mol. Microbiol.* **88**, 876–890 (2013).
 25. Paramonov, N. A., Aduse-Opoku, J., Hashim, A., Rangarajan, M. & Curtis, M. A. Structural Analysis of the Core Region of O-Lipopolysaccharide of *Porphyromonas gingivalis* from Mutants Defective in O-Antigen Ligase and O-Antigen Polymerase. *Journal of Bacteriology* **191**, 5272–5282 (2009).

26. Di Lorenzo, F., Kubik, Ł., Oblak, A., Lorè, N. I., Cigana, C., Lanzetta, R., Parrilli, M., Hamad, M. A., De Soyza, A., Silipo, A., Jerala, R., Bragonzi, A., Valvano, M. A., Martín-Santamaría, S. & Molinaro, A. Activation of Human Toll-like Receptor 4 (TLR4)·Myeloid Differentiation Factor 2 (MD-2) by Hypoacylated Lipopolysaccharide from a Clinical Isolate of *Burkholderia cenocepacia*. *JBC*, **290**, 21305–21319 (2015).
27. Pither, M. D., Silipo, A., Di Lorenzo, F. & Molinaro, A. Glycans in Bacterial Infections: Gram-Negative Infections in the Respiratory Tract, in *Compr. Glycosci. Second Ed.* (ed. Barchi, J. J.) 233–249 (2021)
28. Needham, B. D. & Trent, M. S. Fortifying the barrier: the impact of lipid A remodelling on bacterial pathogenesis. *Nat Rev Microbiol* **11**, 467–481 (2013).
29. Loutet, S. A., Di Lorenzo, F., Clarke, C., Molinaro, A. & Valvano, M. A. Transcriptional responses of *Burkholderia cenocepacia* to polymyxin B in isogenic strains with diverse polymyxin B resistance phenotypes. *BMC Genomics* **12**, 472 (2011).
30. Cullen, L. & McClean, S. Bacterial adaptation during chronic respiratory infections. *Pathogens* **4**, 66–89 (2015).
31. Zhu, Q., Shen, Z., Chiodo, F., Nicolardi, S., Molinaro, A., Silipo, A. & Yu, B. Chemical synthesis of glycans up to a 128-mer relevant to the O-antigen of *Bacteroides vulgatus*. *Nat Commun* **11**, 4142 (2020).
32. Jia, Y., Wang, K., Zhang, Z., Tong, Y., Han, D., Hu, C., Li, Q., Xiang, Y., Mao, X. & Tang, B. TLR2/TLR4 activation induces Tregs and suppresses intestinal inflammation caused by *Fusobacterium nucleatum* in vivo. *PLOS ONE* **12**, e0186179 (2017).
33. Latorre, E., Layunta, E., Grasa, L., Pardo, J., García, S., Alcalde, A. I. & Mesonero, J. E. Toll-like receptors 2 and 4 modulate intestinal IL-10 differently in ileum and colon. *UEG Journal* **6**, 446–453 (2018).
34. Mishima, Y., Oka, A., Liu, B., Herzog, J. W., Eun, C. S., Fan, T.-J., Bulik-Sullivan, E., Carroll, I. M., Hansen, J. J., Chen, L., Wilson, J. E., Fisher, N. C., Ting, J. P. Y., Nochi, T., Wahl, A., Garcia, J. V., Karp, C. L. & Sartor, R. B. Microbiota maintain colonic homeostasis by activating TLR2/MyD88/PI3K signaling in IL-10-producing regulatory B cells. *J Clin Invest* **129**, 3702–3716 (2019).

CHAPTER 8

Appendix

Structural characterisation of lipid A from *Pandoraea* and *Echinicola* species

Preface

Although the lipid A moiety is the most conserved part of the LPS structure, many subtle chemical differences can arise within the architecture due to bacterial adaptation, selective pressure, incomplete biosynthesis, changes of environmental conditions, and the presence of external stimuli. Alterations in the lipid A chemistry can dramatically influence the bioactivity of the whole molecule, which can be harnessed to understand pathogenesis or utilised for design of therapeutic interventions.

Pathogenic bacteria alter their lipid A structure i.e., the degree and pattern of acylation; this can render the LPS undetectable by the host immune sensors or can alter the permeability of the OM, thus providing resistance to harsh pH and antibiotics influencing the outcomes of the treatment of a bacterial infection. The characterisation of lipid A structures can also provide insights into the chemical structural features observable in some of the most serious and persistent infections caused by Gram-negatives. On the other hand, environmental bacteria especially those living in uninhabitable or even lethal habitats for most eukaryotic cells (termed extremophiles), usually modify the LPS structure to reinforce their membrane system. These modifications have shown to ensure protection and adaptation but have also shown to exert only a weak agonistic behaviour towards TLR4. Therefore, these chemical patterns and their bioactivity can then be exploited to develop therapeutics such as TLR4 antagonists, vaccine adjuvants or antibiotics.

This chapter focuses on the chemical characterisation of the lipid A isolated from bacteria not strictly related to the human gut but that are a clear example of the chemical alterations detectable as a consequence of the adaptation phenomena enabling bacterial colonisation of the host or of a specific environment. Section 1 covers the full structural determination of the lipid A from *Pandora* *pulmonicola*, an emerging cystic fibrosis pathogen able to chronically infect patients, while Section

2 outlines the elucidation of the unusual lipid A structure and immunoinhibitory activity of LPS from two marine *Echinicola* species.

Section 1: Characterisation of lipid A from *Pandoraea pulmonicola*

8.1 Introduction

Cystic fibrosis (CF) is an incurable, chronic disease, caused by defects in genes encoding for the CF transmembrane conductance regulator (CFTR), a chloride channel that regulates the activity of other chloride and sodium channels at the cell surface epithelium.¹ A defective CFTR protein results in irregular movement of salts and water across the cells and leads to the creation of a thick mucus that blocks the airways. This airway blockage leads to problems breathing and susceptibility to serious (multiple) lung infections. Indeed, the main cause of morbidity and mortality for CF patients is associated to infection by bacterial pathogens and opportunistic pathogens encountered throughout their lives. Beside *Pseudomonas aeruginosa* and bacteria of the *Burkholderia cepacia* complex, which are well-known Gram-negative opportunistic pathogens able to persistently colonise CF respiratory tract, *Pandoraea* species are considered as emerging CF pathogens leading to the worsening of the lung disease.²

In this section, I report the structural elucidation of the lipid A from the LPS of *P. pulmonicola* strain RL8228, isolated from a chronically colonised CF Irish patient who died 52 months after their first *Pandoraea* isolation.³ The structure has been defined by merging information that were attained from compositional analyses executed on pure LPS and isolated lipid A with data from a MALDI-TOF MS and MS² investigation executed on the lipid A fraction and directly on the bacterial pellet. **These results were published in *Glycoconjugate Journal* in 2021.**⁴

8.2 Structural characterisation *P. pulmonicola* lipid A

8.2.1 LPS extraction from *P. pulmonicola* strain RL8228 and compositional analysis of the lipid A

The LPS material was isolated from lyophilised bacterial cells and visualised via SDS-PAGE and silver nitrate staining.^{5,6} This analysis showed that *P. pulmonicola* strain RL8228 expresses an S-LPS as proven by the ladder-like pattern in the upper part of the gel (**Figure 8.1**), indicative of the presence of the O-antigen moiety. After purification of the LPS, a detailed compositional analysis was performed to establish the fatty acid content, revealing the occurrence of (*R*)-3-hydroxytetradecanoic acid (14:0 (3-OH)) in both ester and amide linkages, whereas (*S*)-2-hydroxytetradecanoic (14:0 (2-OH)), (*S*)-2-hydroxydodecanoic (12:0 (2-OH)), dodecanoic (12:0) and tetradecanoic acid (14:0) were found only as ester-bound acyl chains.



Figure 8.1 SDS-PAGE after silver nitrate staining of S-LPS from *P. pulmonicola* RL8228 (lanes 1 - 3) in various concentrations. The S-LPS from *E. coli* (lane 4) was used as a benchmark reference.

To elucidate the structure of the lipid A moiety, an aliquot of the LPS underwent a mild acid hydrolysis, which was then centrifuged to obtain the lipid A fraction. An aliquot of the lipid A underwent a compositional analysis to define the nature of the lipid A sugar backbone. This analysis revealed the occurrence of GlcN as the only

sugar composing the lipid A moiety. To establish the fine structure of the lipid A, a detailed MALDI-TOF MS and MS² analysis was performed on another aliquot of the isolated mild acid hydrolysis product. In addition, MALDI-TOF MS and MS² investigation was also applied to an aliquot of bacterial cell pellet.⁷

8.2.2 MALDI-TOF MS and MS² analysis on the isolated lipid A from *P. pulmonicola* strain RL8228

The reflectron MALDI-TOF MS spectrum, recorded in negative ion polarity, of the isolated lipid A from *P. pulmonicola* RL8228 is reported in **Figure 8.2**. This spectrum shows an extremely complex pattern of signals relative to deprotonated [M-H]⁻ lipid A species differing in the nature and number of the fatty acid chains and the phosphate content. Moreover, the spectrum displayed additional peaks differing in 161 amu, which likely suggested the occurrence of lipid A species characterised by hexosamine (HexN) modification on one or both phosphate groups of the disaccharide backbone. In the spectra shown in **Figure 8.2**, in the mass range m/z 1082.1–1946.9, *mono*- and *bis*-phosphorylated tri- to penta-acylated lipid A species have been identified. The high heterogeneity due to differences of 16 amu occurring between most of the peaks is indicative of lipid A species differing in the absence or presence of hydroxylated acyl chains. The main peak at m/z 1376.3 was matched with a *bis*-phosphorylated tetra-acylated lipid A species carrying two *N*-linked primary 14:0 (3-OH), one *O*-linked primary 14:0 (3-OH) and one secondary 12:0 (2-OH), whose *mono*-phosphorylated form was detected at m/z 1296.4. Moreover, the related *bis*-phosphorylated penta-acylated lipid A form bearing an additional secondary 14:0 (2-OH) moiety with respect to species at m/z 1376.3, matched with peak at m/z 1602.5. Interestingly, starting from the *mono*-phosphorylated tetra-acylated lipid A species at m/z 1296.4, it was possible to identify a related species decorated by one additional HexN matching with peak at m/z 1457.5; the corresponding penta-acylated form, carrying another secondary 14:0 (2-OH), was identified at m/z 1683.7. Starting from the lipid A species at m/z 1398.4, additional lipid A forms were identified as decorated

by one HexN (m/z 1559.5), one HexN and one 14:0 (OH) (m/z 1785.7), and two HexNs and one 14:0 (OH) (m/z 1946.9). Notably, the negative-ion MALDI mass spectrum, which was recorded directly on the intact bacterial cells, confirmed structural determination attained from the isolated lipid A fraction. This observation consolidated the results obtained by analysing the mild acid hydrolysis product, thus excluding any lack of structural information due to the chemical treatment performed to isolate the lipid A.

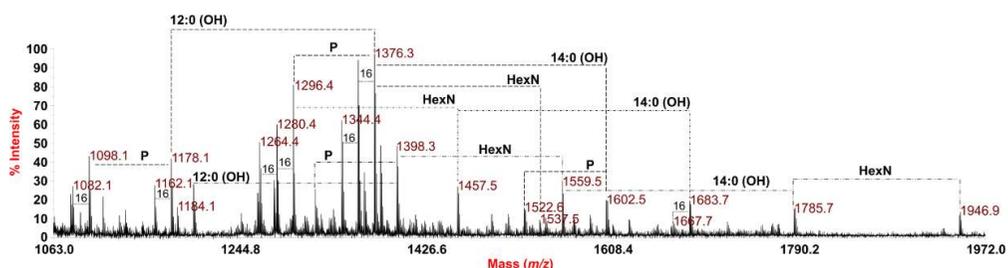


Figure 8.2 Negative-ion MALDI-TOF (reflectron mode) mass spectrum of the lipid A from *P. pulmonicola* RL8228 obtained by mild acid treatment of the purified LPS. Differences of 16 amu are reported in the spectrum. Differences in the nature of the acyl chains have been also indicated. “P” indicates the phosphate group; “HexN” indicate differences of 161 amu (i.e. a hexosamine unit)

A negative-ion MS² analysis was performed on various peaks, to delineate the structure of *P. pulmonicola* RL8228 lipid A, and therefore defining the exact location of the acyl chains, phosphates and additional HexN with respect to the diglucosamine backbone. This data defined that hydroxylated 14:0 fatty acids were found as the primary acyl chains, while various secondary hydroxylated and non-hydroxylated fatty acids were found.

To meticulously define the structure of *P. pulmonicola* RL8228 lipid A, that is to define the exact location of acyl chains, phosphates and of the additional HexN with respect to the diglucosamine backbone, a negative-ion MS² analysis was performed on various peaks. The first precursor ion at m/z 1296.4 was chosen for MS² (**Figure 8.3**), as representative of *mono*-phosphorylated tetra-acylated lipid A species. The spectra showed an intense peak at m/z 1052.4 attributable to an ion derived from the loss of one primary 14:0 (3-OH). However, an important peak, in terms of structural characterisation, was identified at m/z 812.4 which was attributed to an ion originating from the sugar ring fragmentation (^{0,2}A₂);⁸ the observation of the ion demonstrated that **1**) two hydroxylated 14:0 chains and the phosphate are on the reducing glucosamine unit, and that **2**) the hydroxyl group at position 3 of the reducing glucosamine was free. This, in turn, proved that the 12:0 (2-OH) acyl chain was present as a secondary substituent of the primary amide-bound 14:0 (3-OH) of the reducing glucosamine. A further sugar ring fragmentation (^{0,4}A₂)^{8,9} at m/z 752.4 was also identified and concurred to confirm the above structural hypothesis.

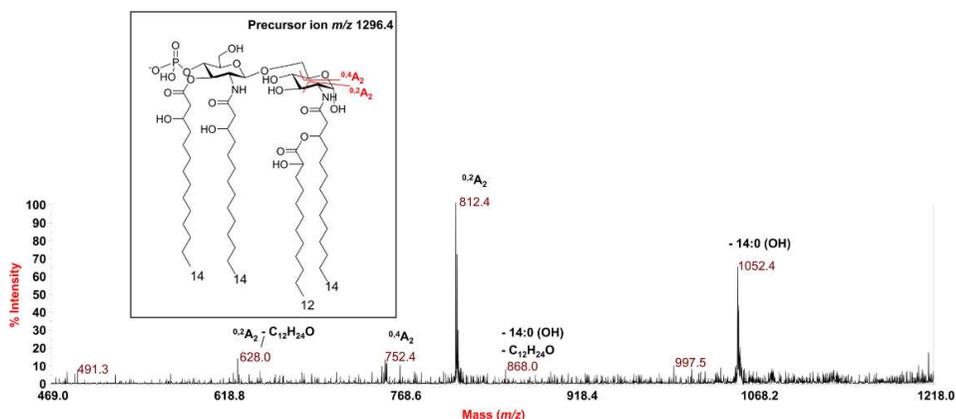


Figure 8.3 Negative-ion MALDI MS² spectra of precursor ion at m/z 1296.4 of the lipid A isolated from *P. pulmonicola* RL8228. This is a representative ion peak of the cluster ascribed to tetra-acylated lipid A species decorated by one phosphate. The main fragments' assignment is indicated in the spectrum. The proposed structure is reported in the inset with the observed sugar ring fragmentations (^{0,4}A₂ and ^{0,2}A₂). The loss of C₁₂H₂₄O (184 mass units) is also indicated and was due to a rearrangement typically occurring on primary 14:0 (3-OH) acyl chains only when their 3-OH group is free, thus contributing to the establishment of the location of the secondary acyl substitution

The negative-ion MS² spectrum of precursor ion at m/z 1522.6 (**Figure 8.4**), identified as a *mono*-phosphorylated penta-acylated lipid A species, was chosen to settle the location of the secondary acyl moieties. In detail, the MS² spectrum showed an intense peak at m/z 1278.6 matching with an ion derived from the loss of one hydroxylated 14:0 moiety (**Figure 8.4**). In addition, the peak at m/z 1038.5, assignable to an ion originating from the sugar ring fragmentation ^{0,2}A₂, proved that the non-reducing glucosamine unit was decorated by three acyl chains, that is plausibly two primary 14:0 (3-OH) and one secondary 14:0 (2-OH). The absence of fragments matching with the loss, from the precursor ion, of a whole unit of a hydroxylated 14:0 fatty acid carrying a secondary hydroxylated 14:0 unit, suggested that the secondary acyl substitution occurred on the *N*-linked primary acyl chain. Likewise, the absence of a peak matching with an ion derived from the sequential loss of one 14:0 (OH) and one 12:0 (OH), suggested that the latter was in an acyloxyacyl moiety. Finally, and further supporting the structural assessment, the peaks at m/z 812.3 and m/z 794.2 were assigned to the sugar ring fragmentation ^{0,2}A₂ plus the loss of one hydroxylated 14:0, likely the primary ester-bound acyl moiety, eliminated as either a free fatty acid (m/z 794.2) and as a ketene derivative (m/z 812.3).

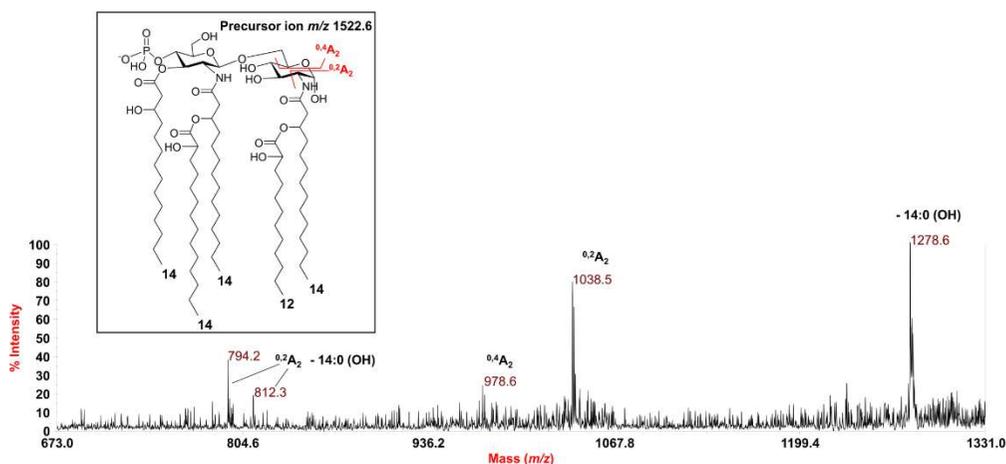


Figure 8.4 Negative-ion MALDI MS² spectrum of precursor ion at m/z 1522.6 of the *P. pulmonicola* RL8228 lipid A. This is a representative ion peak of the cluster attributed to penta-acylated lipid A species decorated by one phosphate. The main fragments' assignment is indicated in the spectrum and in the inset is sketched the proposed structure is given in the inset with the observed sugar ring fragmentations ($^{0,4}A_2$ and $^{0,2}A_2$)

Since data from MS² analyses led to locate the secondary acyl substitution only on the primary amide-bound fatty acids, to further confirm this structural deduction, an aliquot of lipid A was subjected to a treatment with NH₄OH,¹⁰ which is a procedure to selectively cleave the acyl and acyloxyacyl esters, while leaving the acyl and acyloxyacyl amides unaltered. The NH₄OH-treated lipid A was subject to MALDI-TOF MS in negative polarity (**Figure 8.5**). It clearly showed three main peaks at m/z 1279.9, 1359.9 and 1375.9 matching with *mono-* (m/z 1279.9) and *bis-*phosphorylated (m/z 1359.9 and 1375.9) tetra-acylated lipid A species carrying the sole primary *N*-linked hydroxylated 14:0 acyl chains in turn substituted by the secondary fatty acids, which were unaffected by the NH₄OH treatment. This definitively proved that the secondary acyl substitution in the lipid A from *P. pulmonicola* RL8228 exclusively occurs on the primary amide-bound fatty acids.

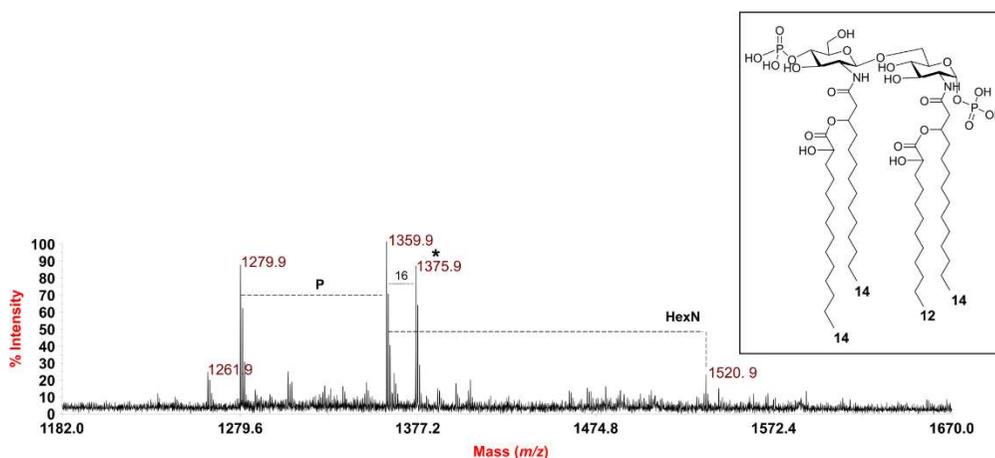


Figure 8.5 Negative ion MALDI-TOF (reflectron mode) mass spectrum of the lipid A after NH_4OH hydrolysis. * Indicates the structure in the inset. “P” indicates the phosphate group; “HexN” indicates a hexosamine unit (differences of 161 amu).

In conclusion, despite chemical analyses only showed the presence of GlcN in the isolated lipid A fraction, in order to unequivocally define the nature of the additional HexN decorating some lipid A species of *P. pulmonicola* RL8228 LPS, an aliquot of lipid A underwent a dephosphorylation treatment; this was followed by derivatisation to alditol acetates of the unknown HexN properly separate from the lipid A moiety, which was then investigated by means of GC-MS (**Figure 8.6**). This analysis revealed the occurrence of glucosaminitol acetate, thus finally demonstrating that the additional HexN unit decorating the disaccharide backbone of some lipid A species from *P. pulmonicola* RL8228 was a GlcN. An overview of the lipid A species which were identified and confirmed through scrutiny of the MALDI TOF MS² can be seen in **Figure 8.7**.

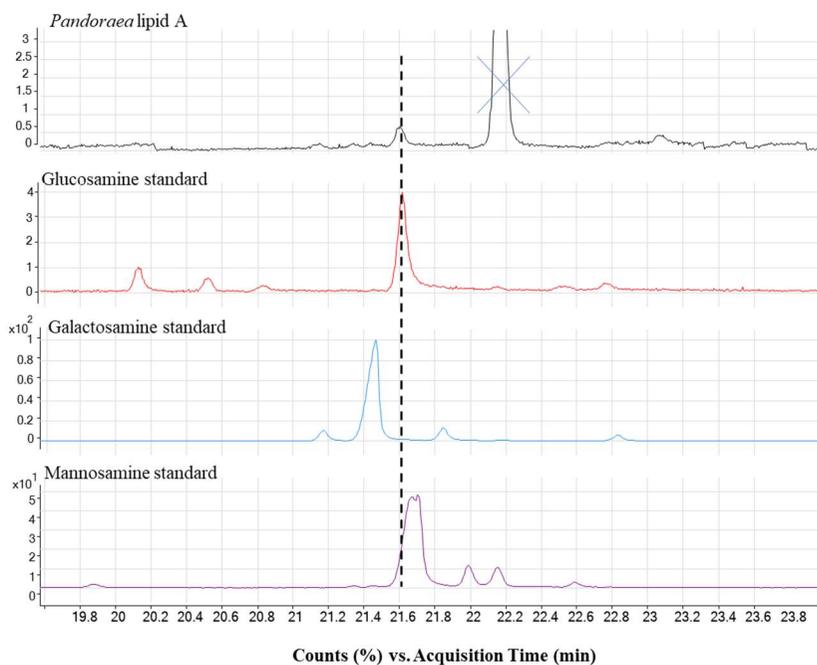


Figure 8.6 a) The GC-MS chromatogram profile of the alditol acetate derivative of the Hex isolated from the *P. pulmonicola* RL8228 lipid A by de-phosphorylation and proper extraction. The GC-MS chromatogram profiles of b) glucosaminitol, c) galactosaminitol and d) mannosaminitol acetates prepared from authentic standards and used as a reference. The comparison of the retention times clearly proved that the unknown HexN decorating the lipid A of *P. pulmonicola* RL8228 was a GlcN

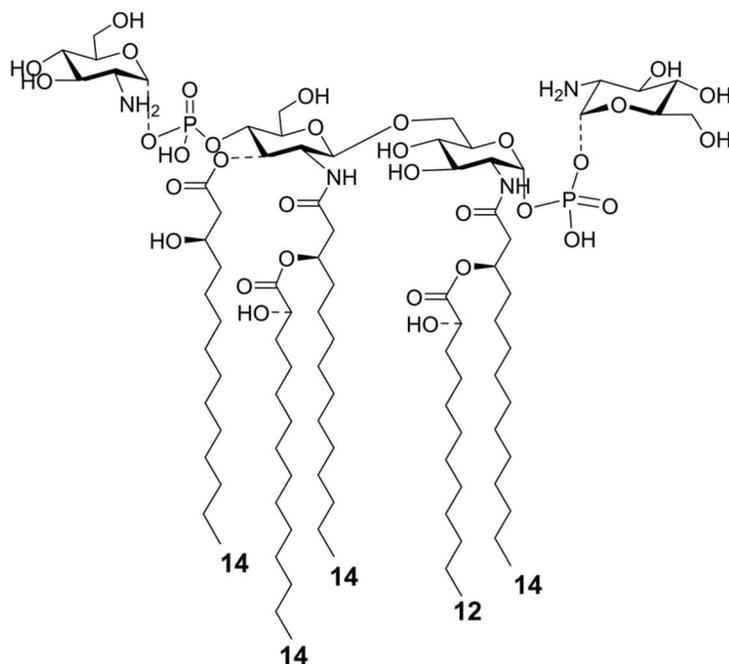


Figure 8.7 A summary of the variable structural features observed *P. pulmonicola* RL8228 lipid A. Some of the observed incompleteness of the substitutions are reported as dotted-line bonds. The secondary acyl moieties can also be non-hydroxylated and the possible occurring fatty acids have been indicated as suggested also by chemical analyses data. The α configuration at the anomeric centre of the GlcN units was tentative.

8.3 Conclusions

Pandoraea species are not only able to chronically colonise the respiratory tract of CF individuals, but they are also easily transmitted between patients causing an aggravation of the lung disease and bacteremia.¹¹ In this frame, *P. pulmonicola* strains have been identified as the most virulent among the other *Pandoraea* species reported so far.³ To gain an insight into the potential virulence mechanisms of *P. pulmonicola*, we focused the attention on the main constituent of its OM, the LPS molecule. As a potent elicitor of the host immune system, the LPS surely has a role in the virulence of this bacterium. Here we characterised the lipid A structure isolated from the S-LPS of the chronic strain *P. pulmonicola* RL8228. The lipid A turned out to be a complex mixture of species differing in nature and number of the acyl chains, but also for the phosphate content and the additional decoration on one or both the phosphate groups with GlcN.

The structural elucidation of the lipid A from *P. pulmonicola* RL8228 S-LPS already revealed several structural features that could explain the exceptional virulence of this strain as well as its capability to chronize in CF lungs. These being **(i)** the capability to express different lipid A species simultaneously with varying abundance, **(ii)** the low degree of acylation (less than six acyl chains), **(iii)** the unusual fatty acid asymmetry and variability at position 3 and 3', **(iv)** the 2-hydroxylation of both 14:0 and 12:0 in the acyloxyacyl positions at 2 and 2', as well as **(v)** the occurrence of the additional decoration by GlcN of the phosphate groups. Most of these characteristics are considered as a strategy adopted by bacteria to escape the host immune surveillance and thus enabling their persistence in the infected tissue.

Notably, some structural features found for the *P. pulmonicola* RL8228 lipid A resemble those observed for another respiratory tract pathogens, *i.e.*, *Bordetella* sp. Indeed, a similar fatty acid composition and distribution with the respect to the glucosamine disaccharide backbone has been observed for some *Bordetella* species.¹² However, in contrast to *P. pulmonicola* RL8228, in *Bordetella* sp. short-chain fatty

acids (10:0 (OH)) are commonly found, in addition to the occurrence in some strains, of 16:0 as a modification required for persistent colonisation of the mouse respiratory tract, and for protection from complement-mediated killing during *Bordetella* respiratory infection.¹³ How and at what extent the above structural features determined for *P. pulmonicola* RL8228 lipid A are involved in the pathogenicity of such an emerging CF pathogen is currently under investigation in the context of a detailed structure to function relationship study.

These results were as part of a full study that was published in *Glycoconjugate Journal* and can be found at: Doi: 10.1007/s10719-020-09954-8

Section 2: Unusual lipid A structure and immunoinhibitory activity of LPS from marine bacteria *Echinicola pacifica* KMM 6172^T and *Echinicola vietnamensis* KMM 6221^T

8.4 Introduction

Marine environments host remarkably high and diverse microbial populations and are considered as the oldest life on the planet. These microorganisms have adapted, over millions of years, to survive and proliferate in seas and oceans, which are extraordinarily delicate and highly mutable, besides being often characterised by one or more extreme chemical or physical parameters (i.e., high salinity, low pressure, and low temperature).¹⁴ Indeed, marine bacteria represent a fascinating taxonomic lineage and a source of natural biologically active compounds comprising a broad range of antitoxins, antibiotics, antitumor, antimicrobial agents, and enzymes with a wide activity spectrum.¹⁵ The cell envelope of marine bacteria is in constant contact with the stress factors of the surrounding environment, which necessarily implies that its main constituents (membrane lipids, proteins, and glycoconjugates) must be precisely organised to maintain the proper physiology and functionality, even in the “extreme” aquatic habitat. At the same time, these membrane components must be versatile enough to face the constantly changing environment.

Marine bacteria are predominantly Gram-negatives and thus possess LPS on their outer membrane. It is expected that LPS from marine bacteria expresses unusual chemical characteristics enabling bacterial life in the extremely mutable aquatic environment. Of note, LPS expressing unusual structural features are considered as potential inhibitors of the TLR4-mediated signaling. Manifold studies have indeed revealed peculiar and uncommon LPS chemical structures from marine bacteria, which were reflected in an interesting immunological behavior that comprises a very

weak immunostimulant activity or even antagonistic properties.^{16–19} Altogether, the structural elucidation of the lipid A from marine bacteria is certainly a fundamental starting point for the comprehension of the molecular details at the basis of adaptation processes, but it is also extremely important in the perspective of the realisation of new generation immune-therapeutics inspired by a natural source. In light of this, we have investigated the *Echinicola* genus that actually belongs to the phylum Bacteroidetes (central in this PhD thesis project) and that also comprises Gram-negative pigmented bacteria of marine origin.

In particular, this section describes the structural characterisation of the lipid A moiety of *E. pacifica* KMM 6172^T isolated from the sea urchin *Strongylocentrotus intermedius* (from Sea of Japan)²⁰, and *E. vietnamensis* KMM 6221^T isolated from seawater collected in a mussel farm located in a lagoon of Nha Trang Bay (South China Sea).¹⁵ Within this study the immunoactivity of the isolated LPS was evaluated and showed to act both as weak TLR4 agonists and as inhibitors of the inflammatory response triggered by inflammatory LPS on HEK-Blue™ human (h) TLR4 cells.

8.5 Structural characterisation of the lipid A from *E. pacifica* KMM 6172^T and *E. vietnamensis* KMM 6221^T

8.5.1 Isolation and fatty acid compositional analysis of the LPS from *E. pacifica* KMM 6172^T and *E. vietnamensis* KMM 6221^T

Following the extraction of the LPS and purification, a detailed compositional analysis was performed to establish the fatty acid content (results are summarised in **Table 8.1**). This analysis revealed an extremely heterogenous composition, which was however qualitatively equal for both *Echinicola* strains. The same compositional analysis was also performed on the isolated lipid A confirming the fatty acid content as seen in **Table 8.1** and enabling the observation of the acetylated methyl glycoside derivative 2-amino-2-deoxy-D-glucose (D-glucosamine) and also of D-galacturonic

acid (D-GalA), thus, suggesting the presence of an additional decoration of the lipid A of both *Echinicola* strains by this acid sugar residue.

Table 8.1 Fatty acid content of the LPS isolated from the two *Echinicola* strains examined in the current study. Both strains displayed the same fatty acid composition and were both characterised by a disaccharide of D-glucosamine as the lipid A sugar backbone carrying an additional saccharide moiety, i.e., a D-galacturonic acid. For the unsaturated acyl chains, the position of the double bond or the stereochemistry remain to be defined.

Fatty Acid Component
9:0 <i>iso</i> (3-OH) [<i>i</i> 9:0(3-OH)]
10:0(3-OH)
14:00
15:1 <i>iso</i> (<i>i</i> 15:1)
15:0 <i>iso</i> (<i>i</i> 15:0)
15:0 <i>anteiso</i> (<i>a</i> 15:0)
16:01
16:0
15:0 <i>iso</i> (3-OH) [<i>i</i> 15:0(3-OH)]
15:0(3-OH)
17:0 <i>iso</i> (<i>i</i> 17:0)
17:1 <i>iso</i> (<i>i</i> 17:1)
16:0(3-OH)
17:0 <i>iso</i> 3-OH [<i>i</i> 17:0(3-OH)]

8.5.2 MALDI-TOF MS and MS/MS analysis on the isolated lipid A from *E. pacifica* KMM 6172^T

Figure 8.5a reports the MALDI-TOF MS spectrum recorded in negative ion polarity of the lipid A from *E. pacifica* KMM 6172^T. Here two clusters of peaks can be seen that were identified as being deprotonated [M-H]⁻ tri-acylated (m/z 1343.6–1371.6) and tetra-acylated lipid A species (m/z 1499.8–1635.8). From the occurrence of peaks differing in 14 amu (-CH₂- unit) and 28 amu (-CH₂CH₂-) it was immediately apparent the heterogeneity of fatty acid chains composing the lipid A species. Moreover, the spectrum showed additional peaks, in the m/z range 1417.7–1445.7, that differ in 176 amu from the main tetra-acylated lipid A species, which suggested the occurrence of lipid A forms characterised by the presence of a hexuronic acid (HexA) modification on the lipid A saccharide backbone, also in accordance with chemical analyses. The main peak at m/z 1607.8 from **Figure 8.5a** was assigned to a tetra-acylated lipid A species composed of the typical glucosamine disaccharide backbone substituted by one phosphate and one HexA unit, bearing two *i*17:0(3-OH), one 15:0(3-OH) [or *i*15:0(3-OH)], and one 16:1; whereas the peak at m/z 1527.7 matched with a *mono*-phosphorylated lipid A species, decorated by HexA, carrying two *i*17:0(3-OH), one 15:0(3-OH) [or *i*15:0(3-OH)], and one *i*9:0(3-OH). Minor peaks in the lower mass region (m/z 1343.6 -1371.6) represents the tri-acylated lipid A species carrying one phosphate and one HexA unit. These analyses were also observed by analysing directly the bacterial pellet following the Larrouy-Maumus *et al.* (2016) protocol.

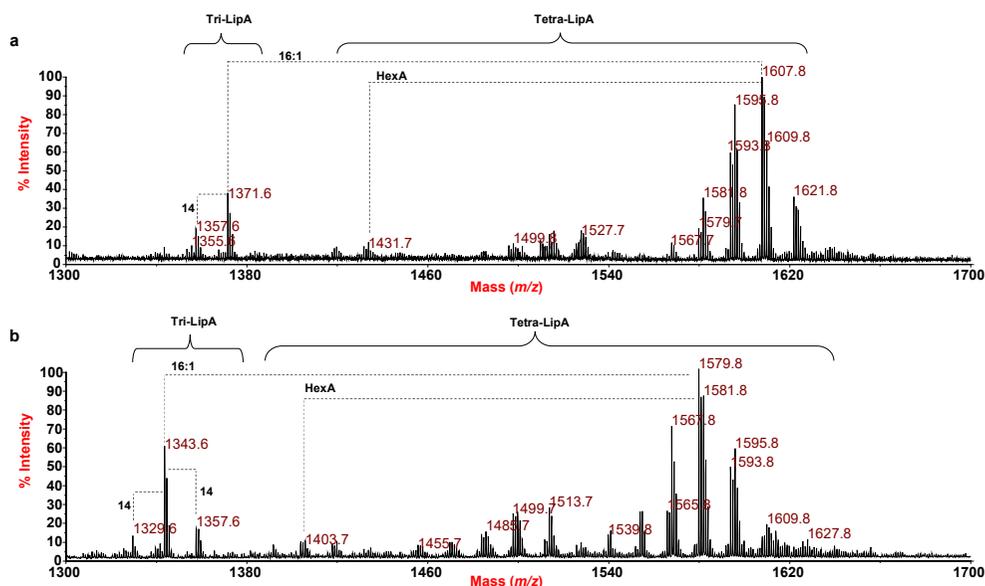


Figure 8.5 Reflectron MALDI-TOF MS, recorded in negative polarity. **a)** of lipid A from *E. pacifica* KMM 6172^T and **b)** *E. vietnamensis* KMM 6221^T. The data was obtained after acetate buffer treatment of each LPS. The lipid A species are labelled as tri- and tetra-lip A indicating the degree of acylation. “hexa” indicates the hexuronic acid (i.e., the D-GalA)

A negative-ion MS/MS investigation was conducted to unveil the exact location of the lipid A acyl moieties with respect to the glucosamine disaccharide backbone as well as the position of the phosphate and the HexA residue. Briefly, the MS/MS spectrum of the precursor ion at m/z 1607.8 (**Figure 8.6**) showed two intense peaks at m/z 1349.7 and 1353.7 attributed to ions derived from the loss of a 15:0(3-OH) and a 16:1 fatty acid, respectively. The peaks observed at m/z 1509.8 and m/z 1431.8 were assigned to fragments that originated from the loss of the phosphate group and the HexA unit, respectively (**Figure 8.6**). The peak at m/z 1251.7 was attributed to a fragment devoid of the phosphate group and the 15:0(3-OH) unit; whereas, the peak at m/z 1177.7 was identified as an ion originated from the loss of the HexA unit and the 16:1. The observation of the ion peak at m/z 1095.6, matching with a fragment caused by the sequential loss of one 15:0(3-OH) and one 16:1, was important for the structural characterisation. Indeed, the presence of this peak gave a first indication

that the unsaturated acyl moiety (16:1) was not bound as a secondary substituent of the primary ester linked 15:0(3-OH). In parallel, the occurrence of the ion at m/z 882.2, originating from the sugar ring fragmentation ($^{0,4}A_2$)⁸, was fundamental to define the nature of the fatty acids that decorated the nonreducing glucosamine (namely, one *i*17:0(3-OH) and one 16:1) as well as the location of the HexA on such a glucosamine unit. In support of this hypothesis, an ion originated from the sugar ring fragmentation $^{0,4}A_2$ plus the loss of the 16:1 acyl moiety was also assigned to the peak at m/z 628.1. Therefore, the three main observations key to locate the 16:1 as a primary ester-bound fatty acid of the nonreducing glucosamine were (1) the presence of several peaks relative to fragments originated from the loss of the 16:1 moiety; (2) the fact that, compared to their acyl and acyloxyacyl counterparts, the loss of secondary acyl substituents of primary amide-linked fatty acids are less commonly observed in MS/MS investigation; (3) the absence of any peak indicative of the loss of a whole unit of a hydroxylated fatty acid bearing 16:1 as a secondary acyl substituent. This structural hypothesis was then definitively corroborated by the observation of the peak at m/z 1155.6 assigned to a fragment originating from a rearrangement occurring only if the *N*-linked fatty acids at positions C-2 and C-2' have a free 3-OH group, i.e., they do not carry secondary acyl substituents. Indeed, an enamine to imine tautomerisation followed by a six-membered ring-based rearrangement justifies the generation of the ion peak at m/z 1155.6, which derives from the loss of a $C_{15}H_{30}O$ neutral fragment (226 amu) from each primary *N*-linked acyl chain possessing a free 3-OH group. Since such a rearrangement can occur only if no secondary acyl substituents are linked to position 3 of the amide-linked acyl chains,²¹ the occurrence of this peak unequivocally located the 16:1 as a primary ester-linked moiety. Finally, the location of the phosphate group was defined based on the detection of the Y-type ion (m/z 766.3)⁸ originated from the cleavage of the glycosidic linkage of the glucosamine backbone, which in turn also confirmed the location on the reducing glucosamine unit of a *i*17:0(3-OH) and a 15:0(3-OH).

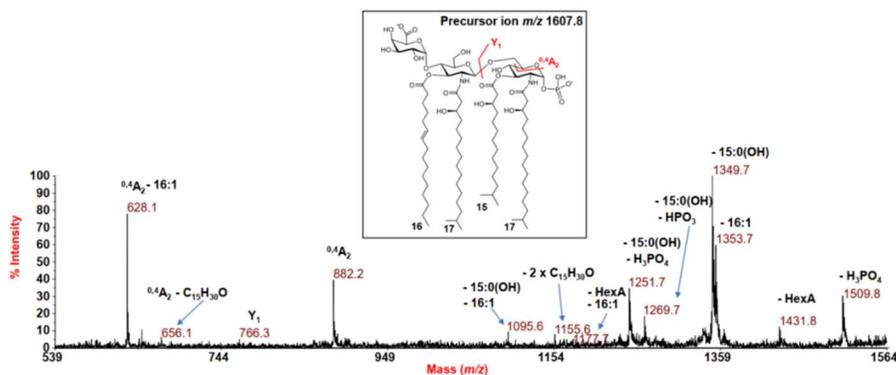


Figure 8.6 Negative-ion MALDI MS² spectrum of precursor ion at m/z 1607.8, a representative ion peak of the cluster ascribed to tetra-acylated lipid A species from *E. pacifica* KMM 6172^T. The assignment of main fragments is reported in the spectrum. The proposed structure for the lipid A species is reported in the inset, where the representation of 15:0(3-OH) in its branching form is tentative.

Likewise, the structure of the lipid A species detected at m/z 1595.8 and m/z 1621.8 from the MS spectra **Figure 8.5a** were also elucidated (**Figure 8.7**). These results revealing a tetra-acylated lipid A carrying two *i*17:0(3-OH), one 15:0(3-OH), and one 15:0 for the species detected at m/z 1595.8, and a tetra-acylated species bearing two *i*17:0(3-OH), one 15:0(3-OH), and one 17:1 for the species at m/z 1621.8; in both cases, a phosphate decorates the reducing glucosamine and a HexA decorates the nonreducing glucosamine.

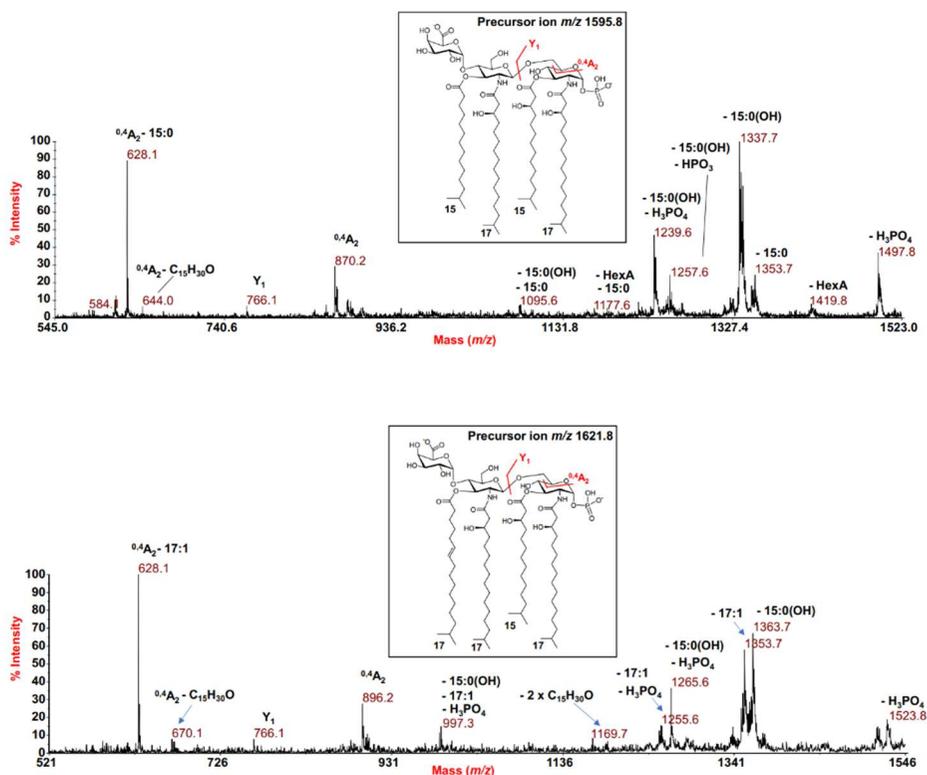


Figure 8.7 Negative-ion MALDI MS/MS spectrum of precursor ion at a) m/z 1595.8 and b) m/z 1621.8. As two other representative ion peak of the cluster assigned to tetra-acylated lipid A species from *E. pacifica* KMM 6172^T

In order to further dissect the structure of such a heterogenous lipid A, the precursor ion at m/z 1527.7 was also chosen for negative-ion MS/MS analysis. The MS/MS spectrum (**Figure 8.8**) showed an intense peak at m/z 1269.6 indicative of a fragment derived from the loss of a 15:0(3-OH) moiety, whose fragment lacking the phosphate group was detected at m/z 1171.6. The peak detected at m/z 1353.6 was attributed to an ion originated from the loss of a *i*9:0(3-OH) moiety. Moreover, in this case, the cross-ring fragmentations $^{0,4}A_2$ (m/z 802.2) and $^{0,4}A_2$ minus an *i*9:0(3-OH) (m/z 628.1) were crucial to define the location and the nature of the acyl chains, namely one *i*17:0(3-OH) and one *i*9:0(3-OH), on the nonreducing glucosamine as well as the

position of the HexA unit. Finally, the investigation of the negative-ion MS/MS spectrum of the precursor ion at m/z 1371.6 (**Figure 8.8b**), chosen as a representative of the tri-acylated lipid A species, disclosed a lipid A composed of two *N*-linked *i*17:0(3-OH) and the O-linked 15:0(3-OH) acyl chains on a backbone made up of the glucosamine disaccharide substituted by HexA and one phosphate unit. This was suggested by the detection of the peak at m/z 646.2, derived from the cross-ring fragmentation $^{0,4}A_2$, as well as of the Y1 ion peak at m/z 766.3 (**Figure 8.8b**), derived from the cleavage of the glycosidic bond.

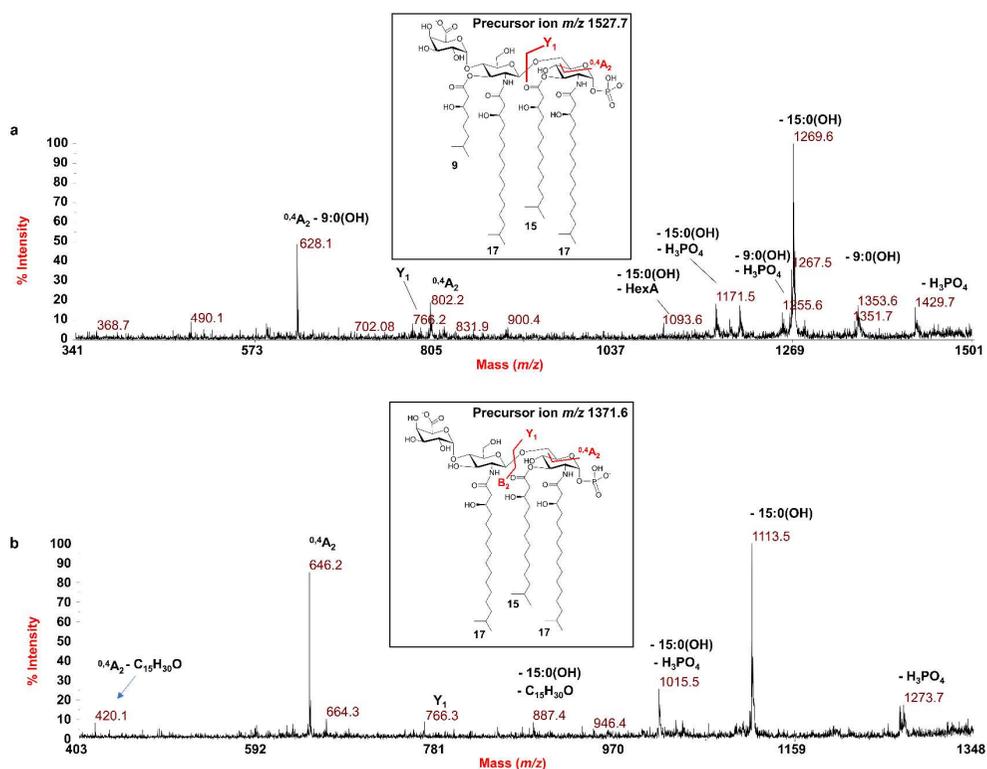


Figure 8.8 MALDI MS/MS analysis of the lipid A species from *E. pacifica* KMM 172^T. **a)** Negative ion MALDI MS/MS spectrum of precursor ion at m/z 1527.7, a tetra-acylated lipid A species **b)** Negative-ion MALDI MS/MS spectrum of precursor ion at m/z 1371.6, a tri-acylated lipid A species. The proposed structure for the lipid A species is reported in each inset.

Finally, to unequivocally establish the location of the acyl chains, an aliquot of lipid A was treated with NH_4OH which selectively removed the acyl and acyloxyacyl esters, while leaving the acyl and acyloxyacyl amides unaltered. A negative-ion MALDI-TOF MS was recorded on this NH_4OH -treated lipid A (**Figure 8.9a**). This spectrum showed a main peak at m/z 1131.7 matching with a lipid A carrying two *N*-linked *i*17:0(3-OH) besides being decorated by one phosphate and the HexA residue. Indeed, from this spectrum it was also possible to notice the lipid A forms differing in 14 amu ($-\text{CH}_2-$). This analysis definitively confirmed that no secondary acyl substitutions occur in the *E. pacifica* KMM 6172^T lipid A and therefore by combining data from the MALDI-TOF MS and MS/MS analysis of the mild acid hydrolysis product with information attained from compositional analyses of both the whole LPS and the isolated lipid A fraction, it was possible to establish that the LPS from *E. pacifica* KMM 6172^T is made up of a complex blend of mono-phosphorylated tri- and tetra-acylated lipid A species mostly decorated by one phosphate on the reducing glucosamine unit and one D-GalA on the nonreducing glucosamine unit.

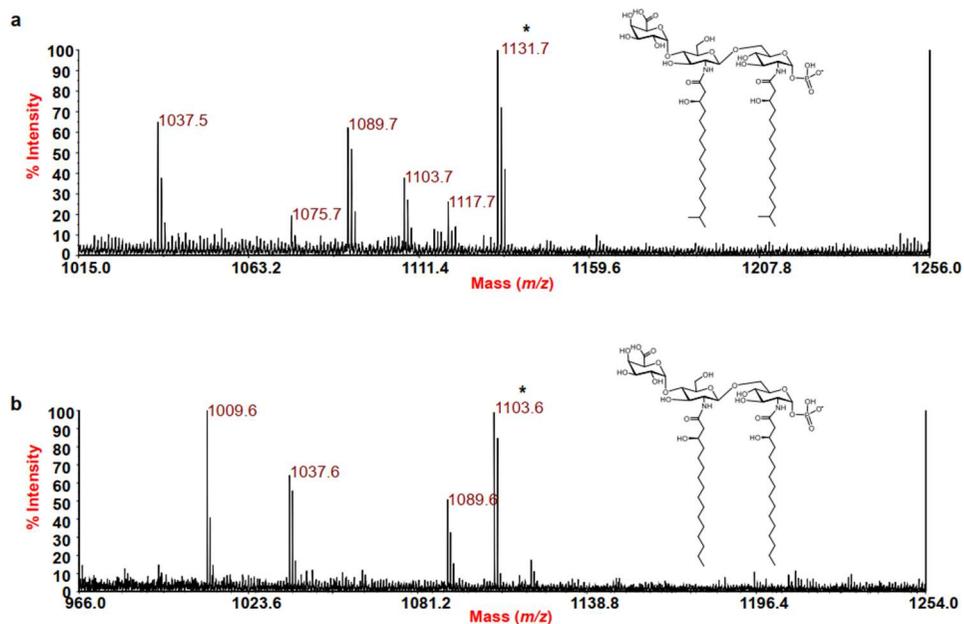


Figure 8.9 Reflectron MALDI-TOF mass spectrum, recorded in negative polarity, of lipid A from *E. pacifica* KMM 6172^T **a)** and *E. vietnamensis* KMM 6221^T **b)** after treatment with NH₄OH. * Indicates the ion peak whose proposed structure is sketched in the inset, where the negative charge held by the phosphate is tentative.

8.5.3 MALDI MS and MS/MS analysis on the isolated lipid A from *E. vietnamensis* KMM 6221^T

The same MS and MS² approach was used to define the structure of the lipid A from *E. vietnamensis* KMM 6221^T. The MALDI TOF MS of the lipid A, shown in **Figure 8.5b**, after a mild acid hydrolysis of the extracted LPS, displayed a similar pattern of peaks to that observed for *E. pacifica* KMM 6172^T. The lipid A species were understood to be a heterogenous mixture of *mono*-phosphorylated tri- and tetra-acylated lipid A species carrying the HexA decoration. However, differences in the predominant species between the two strains were observed. For example, the main peak at *m/z* 1579.8 was attributed to a tetra-acylated lipid A species carrying 16:0(3-OH), as primary *N*-linked fatty acids, and 15:0(3-OH) and 16:1 as the *O*-linked acyl chains. In relation to this peak, a tetra-acylated lipid A species devoid of the HexA

was attributed to peak at m/z 1403.7; whereas a tri-acylated lipid A form lacking the 16:1 unit was assigned to peak at m/z 1343.6.

A final analysis was performed to unequivocally establish the location of the acyl chains. An aliquot of lipid A was treated with NH_4OH to selectively remove the acyl and acyloxyacyl esters, while leaving the acyl and acyloxyacyl amides unaltered and then a negative-ion MALDI-TOF MS analysis of these NH_4OH -treated lipid A was done (**Figure 8.9**). Both species of *Echinicola* lipid A were definitively confirmed to have no secondary acyl substitutions occurring on both lipid As.

8.6 Immunological properties of the LPS from *E. pacifica* KMM 6172^T and *E. vietnamensis* KMM 6221^T

8.6.1 HEK-Blue™ hTLR4 cell models: activation assay

To assess the immunological activity of LPS from the two *Echinicola* strains, *in vitro* experiments using the model of HEK-Blue™ hTLR4 cells stably transfected with human TLR4, MD-2, and CD14 genes were performed. The read-outs of this experiment were the NF-κB activation and IL-8 release (**Figure 8.10**). The HEK-Blue™ hTLR4 cells were stimulated with different concentrations (1, 10, and 100 ng/mL) of *E. pacifica* KMM 6172^T LPS or *E. vietnamensis* KMM 6221^T LPS. *S. typhimurium* SH 2201 LPS, containing fully hexa- and hepta-acylated lipid A, was used at the same concentrations, and evaluated as a positive control. As a negative control, untreated cells were used.

Both LPS from *Echinicola* strains induced a significantly lower NF-κB activation compared to cells treated with *S. typhimurium* LPS (*E. pacifica* KMM 6172^T LPS vs. *S. typhimurium* LPS, $p < 0.001$ at 1 ng/mL, 10 ng/mL, and 100 ng/mL, **Figure 8.10a**; *E. vietnamensis* KMM 6221^T LPS vs. *S. typhimurium* LPS, $p < 0.001$ at 1 ng/mL, 10 ng/mL, and 100 ng/mL, **Figure 8.10a**). In accordance with this outcome, IL-8 secretion was also lower after stimulation with both *Echinicola* LPS than after treatment with the LPS from *S. typhimurium* (*E. pacifica* KMM 6172^T LPS vs. *S. typhimurium* LPS, $p < 0.001$ at 1 ng/mL and 10 ng/mL; $p < 0.01$ at 100 ng/mL, **Figure 8.10b**; *E. vietnamensis* KMM 6221^T LPS vs. *S. typhimurium* LPS, $p < 0.01$ at 1 ng/mL; $p < 0.001$ at 10 ng/mL and 100 ng/mL **Figure 8.10b**). Finally, the lack of any NF-κB activation and IL-8 production in HEK-Blue™ Null2 cells, indicated that NF-κB activation and IL-8 release by *E. pacifica* KMM 6172^T LPS and *E. vietnamensis* KMM 6221^T LPS were TLR4-dependent (**Figure 8.11**).

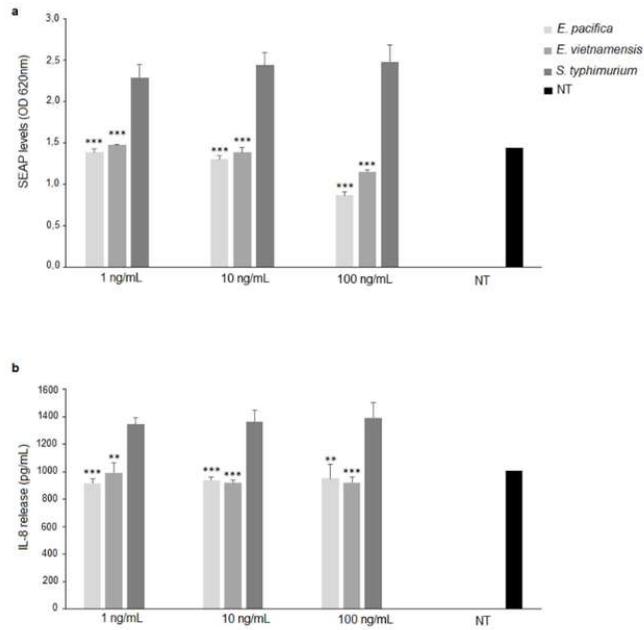


Figure 8.10 Stimulation of HEK Blue™ hTLR4 cells. **a)** SEAP levels OD and **b)** IL-8 release (pg/mL) upon stimulation with LPS of *E. pacifica* KMM 6172^T (1, 10, 100 ng/mL) or *E. vietnamensis* KMM 6221^T LPS (1, 10, and 100 ng/mL); *S. typhimurium* SH 2201 LPS was used as positive control

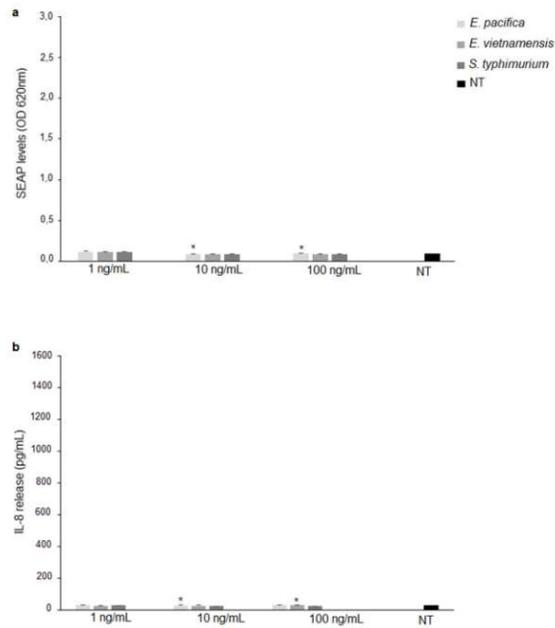


Figure 8.11 Stimulation of HEK Blue™ Null2™ cells. SEAP levels (OD). **a)** and IL-8 release **b)** upon stimulation with LPS of *E. pacifica* KMM 6172^T (1, 10, 100 ng/mL) or *E. vietnamensis* KMM 6221^T LPS (1, 10, 100 ng/mL); *S. typhimurium* SH 2201 LPS was used as positive control

8.6.2 HEK-Blue™ hTLR2 cell models: competition assay

It is understood that LPS isolated from marine bacteria have been proven to be potential inhibitors of the pro-inflammatory action of LPS from pathogenic bacteria. Therefore, the capability of LPS from the two *Echinicola* strains to interfere with the TLR4-mediated signaling triggered by *S. typhimurium* LPS was investigated. To appraise this ability, HEK-Blue™ hTLR4 cells were pretreated with different concentrations of *E. pacifica* KMM 6172^T LPS (1, 10, and 100 ng/mL) or *E. vietnamensis* KMM 6221^T LPS, and then stimulated with *S. typhimurium* LPS (1 ng/mL) (**Figure 8.12**). This study clearly showed that both *Echinicola* LPS significantly inhibited *S. typhimurium* LPS-dependent TLR4-mediated NF-κB activation and IL-8 production at all the concentrations tested but with *E. pacifica* KMM 6172^T LPS displaying stronger antagonistic properties than *E. vietnamensis* KMM 6221^T LPS.

8.6.3 HEK-Blue™ hTLR2 cell models

Finally, we investigated whether *Echinicola* LPS could stimulate the TLR2-mediated signaling, as already observed for LPS of other bacteria like *B. vulgatus* and as *B. thetaiotaomicron* (Chapter 3 and 5).^{22,23} For this investigation, HEK-Blue™ hTLR2 cells, obtained by co-transfection of the human TLR2 and SEAP genes into HEK-Blue™ cells, were stimulated with LPS from the two *Echinicola* strains. These experiments were also carried on using the LPS from *S. typhimurium*, as a negative, and with the synthetic tri-acylated lipoprotein Pam3CSK4 as a positive control. However, there was no significant TLR2 activation observed with either *E. pacifica* KMM 6172^T LPS or *E. vietnamensis* KMM 6221^T LPS or *S. typhimurium* LPS (**Figure 8.13**).

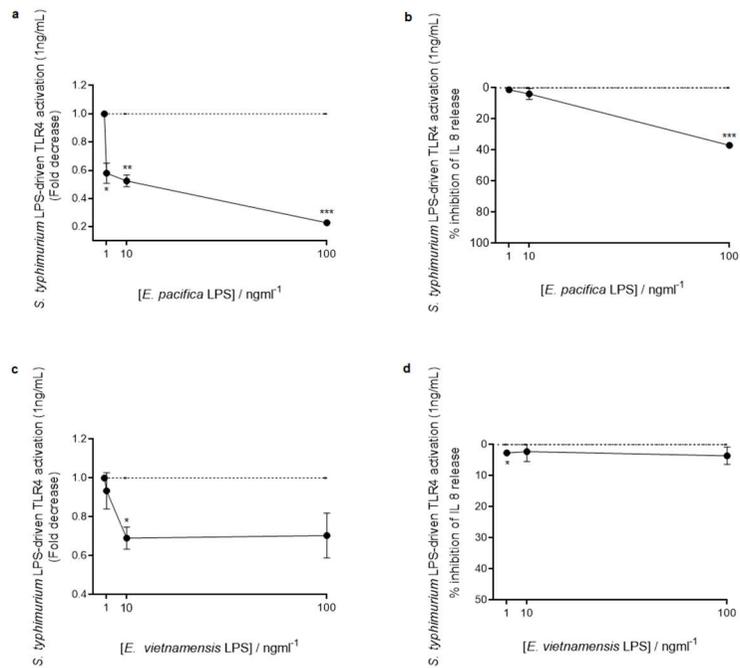


Figure 8.12 Competition assay in HEK Blue™ hTLR4 cells pretreated with different concentrations (1, 10, and 100 ng/mL) of *E. pacifica* KMM 6172^T LPS (a,b) or *E. vietnamensis* KMM 6221^T LPS (c,d) and then stimulated with *S. typhimurium* LPS (1 ng/mL). Fold decrease in NF-κB and inhibition percentage of IL-8 release. Significant differences between the pretreatment with the two *Echinicola* LPS and *S. typhimurium* LPS alone values are indicated. * $p < 0.05$, ** $p < 0.01$ and *** $p < 0.001$ by the Student's t-test

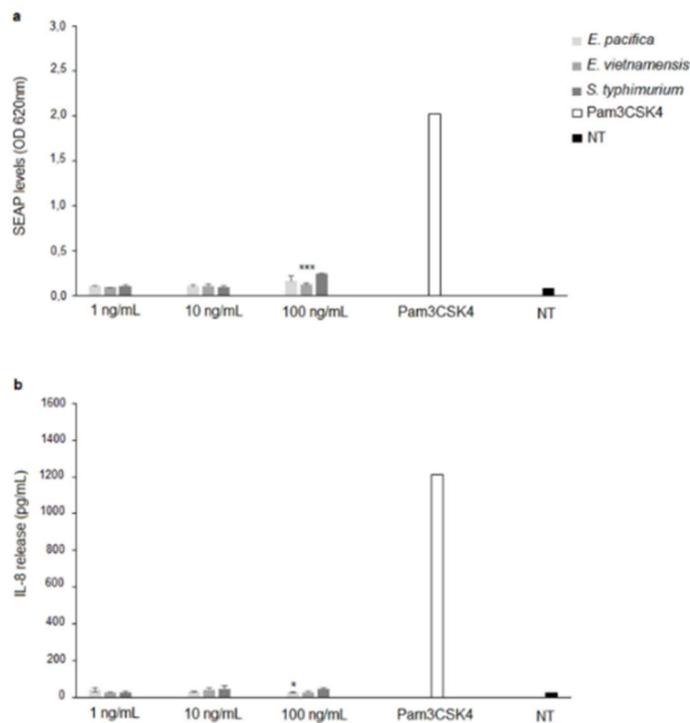


Figure 8.13 Stimulation of HEK Blue™ hTLR2 cells. SEAP levels (OD) (a) and IL-8 release (pg/mL) (b) upon stimulation with LPS of *E. pacifica* KMM 6172^T (1, 10, 100 ng/mL) or *E. vietnamensis* KMM6221^T LPS (1, 10, 100 ng/mL); *S. typhimurium* SH2201 LPS or Pam3CSK4 (500 ng/mL) were used as positive and negative controls, respectively. Significant differences between *E. pacifica* KMM 6172^T LPS or *E. vietnamensis* KMM 6221^T LPS and the corresponding *S. typhimurium* LPS values are indicated. * $p < 0.05$, *** $p < 0.001$ by Student's *t*-test.

8.7 Conclusions

Various environmental pressures within the marine environment can potentially impact the structure of the membrane constituents of Gram-negative bacteria. As the portion of the LPS embedded in the Gram-negative OM, the lipid A is one of the main bacterial components subjected to these selective pressures. Therefore, lipid A from marine bacteria present several structural peculiarities. For example, hypo-acylation (tetra- and penta-acylated) and a tendency toward desaturation of the acyl chains, reduction in their length, and increase in their branching are commonly observed in several marine LPS lipid A.^{16,17} Moreover, some marine species express *mono*- or even non-phosphorylated lipid A.

E. pacifica KMM 6172^T and *E. vietnamensis* KMM 6221^T were shown to express hypo-acylated and hypo-phosphorylated lipid A species. In addition, a D-GalA was found linked to the nonreducing glucosamine of the lipid A disaccharide backbone. Likewise, both strains also decorate their lipid A with unsaturated acyl chains, which are linked to the saccharide backbone as primary ester-linked moieties. This tendency of both *Echinicola* strains to increase the desaturation level of their membrane lipids can be seen as an ingenious strategy to increment the outer membrane fluidity by decreasing the lipid packing. Likely for the same purpose, *E. pacifica* and *E. vietnamensis* express lipid A that can be acylated also by short acyl moieties, such as *i*9:0(3-OH) or 10:0(3-OH), which establish lower interchain Van der Waals interactions than long-chain fatty acids, resulting in an additional contribute to preserving the membrane fluidity.

Because of the unconventional structural features of these lipid A species and the unexplored possibilities of utilising chemical structures from marine environments as leads in a pharmaceutical setting, this study investigated whether such unusual chemical structures are reflected in equally particular immunological properties. By leveraging HEK-Blue™ cell lines stably expressing human TLR4, MD-2, and CD-

14, it was highlighted the very weak capability of both *Echinicola* LPS to elicit a TLR4-mediated immune response while keeping the ability to bind the receptor. This interesting immunological behaviour finds its reason to be in the hypo-acylated and hypo phosphorylated chemical structure of the lipid A of both marine bacteria. Even more fascinating was the observation of the potent inhibitory activity shown by both *Echinicola* LPS towards the pro-inflammatory response induced by the *S. thypimurium* LPS on HEK-Blue™ cells. Many studies have showed that these features are associated to a scant immune-elicitation potency,²⁴ such as the lipid A from the marine species *Halobacteroides lacunaris*¹⁷ and *Cobetia pacifica*. In this context, the present work encourages the investigation of marine LPS as a continual source of natural molecules that can be potentially used as inspiration for the development immunomodulatory compounds exploitable as therapeutics for inflammation-related pathologies.

These results were as part of a full study that was published in *Microorganisms* Journal and can be found at Doi: [10.3390/microorganisms9122552](https://doi.org/10.3390/microorganisms9122552)

8.8 References

1. Rafeeq, M. M. & Murad, H. A. S. Cystic fibrosis: current therapeutic targets and future approaches. *J. Transl. Med.* **15**, 84 (2017).
2. Mahenthiralingam, E. Emerging cystic fibrosis pathogens and the microbiome. *Paediatr. Respir. Rev.* **15**, 13–15 (2014).
3. Costello, A., Herbert, G., Fabunmi, L., Schaffer, K., Kavanagh, K. A., Caraher, E. M., Callaghan, M. & McClean, S. Virulence of an emerging respiratory pathogen, genus *Pandoraea*, in vivo and its interactions with lung epithelial cells. *J. Med. Microbiol.* **60**, 289–299 (2011).
4. Pither, M. D., McClean, S., Silipo, A., Molinaro, A. & Di Lorenzo, F. A chronic strain of the cystic fibrosis pathogen *Pandoraea pulmonicola* expresses a heterogenous hypo-acylated lipid A. *Glycoconj. J.* **38**, 135–144 (2021).
5. Kittelberger, R. & Hilbink, F. Sensitive silver-staining detection of bacterial lipopolysaccharides in polyacrylamide gels. *J. Biochem. Biophys. Methods* **26**, 81–86 (1993).
6. Pither, M., Molinaro, A., Silipo, A. & Di Lorenzo, F. in *Methods Mol. Biol.* 2613, (Springer International Publishing) (2022). **In press.**
7. Larrouy-Maumus, G., Clements, A., Filloux, A., McCarthy, R. R. & Mostowy, S. Direct detection of lipid A on intact Gram-negative bacteria by MALDI-TOF mass spectrometry. *J. Microbiol. Methods* **120**, 68–71 (2016).
8. Domon, B. & Costello, C. E. A systematic nomenclature for carbohydrate fragmentations in FAB-MS/MS spectra of glycoconjugates. *Glycoconj. J.* **5**, 397–409 (1988).
9. Garcia-Vello, P., Speciale, I., Di Lorenzo, F., Molinaro, A. & De Castro, C. in *Lipopolysaccharide Transp. Methods Protoc.* (ed. Sperandeo, P.) 181–209 (Springer US, 2022)
10. Silipo, A., Lanzetta, R., Amoresano, A., Parrilli, M. & Molinaro, A. Ammonium hydroxide hydrolysis: aluable support in the MALDI-TOF mass spectrometry analysis of Lipid A fatty acid distribution. *J Lipid Res* **43**, 2188–2195 (2002).
11. Kokcha, S., Bittar, F., Reynaud-Gaubert, M., Mely, L., Gomez, C., Gaubert, J.-Y., Thomas, P. & Rolain, J.-M. *Pandoraea pulmonicola* chronic colonization in a cystic fibrosis patient, France. *New Microbes New Infect.* **1**, 27–29 (2013).
12. Marr, N., Hajjar, A. M., Shah, N. R., Novikov, A., Yam, C. S., Caroff, M. & Fernandez, R. C. Substitution of the *Bordetella pertussis* lipid A phosphate groups with glucosamine is required for robust NF- κ B activation and release of proinflammatory cytokines in cells expressing human but not murine toll-like receptor 4-MD-2-CD14. *Infect. Immun.* **78**, 2060–2069 (2010).

13. Pilione, M. R., Pishko, E. J., Preston, A., Maskell, D. J. & Harvill, E. T. pagP Is Required for Resistance to Antibody-Mediated Complement Lysis during *Bordetella bronchiseptica* Respiratory Infection. *Infect. Immun.* **72**, 2837–2842 (2004).
14. Rothschild, L. J. & Mancinelli, R. L. Life in extreme environments. *Nature* **409**, 1092–1101 (2001).
15. Debbab, A., Aly, A. H., Lin, W. H. & Proksch, P. Bioactive Compounds from Marine Bacteria and Fungi. *Microb. Biotechnol.* **3**, 544–563 (2010).
16. Ialenti, A., Di Meglio, P., Grassia, G., Maffia, P., Di Rosa, M., Lanzetta, R., Molinaro, A., Silipo, A., Grant, W. & Ianaro, A. A novel lipid A from *Halomonas magadiensis* inhibits enteric LPS-induced human monocyte activation. *Eur. J. Immunol.* **36**, 354–360 (2006).
17. Di Lorenzo, F., Palmigiano, A., Paciello, I., Pallach, M., Garozzo, D., Bernardini, M. L., La Cono, V., Yakimov, M. M., Molinaro, A. & Silipo, A. The deep-sea polyextremophile *Halobacteroides lacunaris* TB21 rough-type LPS: Structure and inhibitory activity towards toxic LPS. *Mar. Drugs* **15**, (2017).
18. Di Lorenzo, F., Palmigiano, A., Albitar-Nehme, S., Pallach, M., Kokoulin, M., Komandrova, N., Romanenko, L., Bernardini, M. L., Garozzo, D., Molinaro, A. & Silipo, A. Lipid A Structure and Immunoinhibitory Effect of the Marine Bacterium *Cobetia pacifica* KMM 3879T. *Eur. J. Org. Chem.* **2018**, 2707–2716 (2018).
19. Di Lorenzo, F., Billod, J. M., Martín-Santamaría, S., Silipo, A. & Molinaro, A. Gram-Negative Extremophile Lipopolysaccharides: Promising Source of Inspiration for a New Generation of Endotoxin Antagonists. *Eur. J. Org. Chem.* **2017**, 4055–4073 (2017).
20. Tomshich, S. V., Kokoulin, M. S., Kalinovsky, A. I., Nedashkovskaya, O. I. & Komandrova, N. A. Structure of the O-specific polysaccharide from a marine bacterium *Echinicola pacifica* KMM 6172(T) containing 2,3-diacetamido-2,3-dideoxy-D-glucuronic acid. *Carbohydr. Res.* **425**, 22–27 (2016).
21. Di Lorenzo, F., Palmigiano, A., Albitar-Nehme, S., Sturiale, L., Duda, K. A., Gully, D., Lanzetta, R., Giraud, E., Garozzo, D., Bernardini, M. L., Molinaro, A. & Silipo, A. The Lipid A from *Rhodopseudomonas palustris* Strain BisA53 LPS Possesses a Unique Structure and Low Immunostimulant Properties. *Chem. – Eur. J.* **23**, 3637–3647 (2017).
22. Di Lorenzo, F., Pither, M. D., Martufi, M., Scarinci, I., Guzmán-Caldentey, J., Łakomiec, E., Jachymek, W., Bruijns, S. C. M., Santamaría, S. M., Frick, J.-S., van Kooyk, Y., Chiodo, F., Silipo, A., Bernardini, M. L. & Molinaro, A. Pairing

- Bacteroides vulgatus LPS Structure with Its Immunomodulatory Effects on Human Cellular Models. *ACS Cent. Sci.* **6**, 1602–1616 (2020).
23. Pither, M. D., Illiano, A., Pagliuca, C., Jacobson, A., Mantova, G., Stornaiuolo, A., Colicchio, R., Vitiello, M., Pinto, G., Silipo, A., Fischbach, M. A., Salvatore, P., Amoresano, A., Molinaro, A. & Di Lorenzo, F. *Bacteroides thetaiotaomicron* rough-type lipopolysaccharide: The chemical structure and the immunological activity. *Carbohydr. Polym.* **297**, 120040 (2022).
24. Di Lorenzo, F., Duda, K. A., Lanzetta, R., Silipo, A., De Castro, C. & Molinaro, A. A Journey from Structure to Function of Bacterial Lipopolysaccharides. *Chem. Rev.* **122**, 15767–15821 (2022).

CHAPTER 9

Materials and Methods

9.1 Bacterial strains, growth conditions, and reagents

All bacteria analysed in this PhD thesis were kindly provided by external collaborators.

B. vulgatus mpk (Chapter 3) was provided by Prof. J. S. Frick and grown in brain-heart-infusion (BH) medium in anaerobic conditions at 37 °C as previously described.¹ For *Alcaligenes faecalis* (Chapter 4), the bacterial strain was purchased from the National Institute of Technology and Evaluation Biological Resource Center and grown in Tryptic Soy Broth (BD Diagnostics, Sparks, MD, USA) at 37 °C by our collaborators in Prof. Fukase group at Osaka University. In Chapter 5, the acapsular *B. thetaiotaomicron* strain VPI 5482 was cultured and grown by Prof. M. Fischbach at the Department of Bioengineering and ChEM-H at Stanford University, as described in detail in Pither *et al.*, (2022)² with a final mass of 4.3 g lyophilised bacterial cell material. The two *Veillonella parvula* isolates (Chapter 6) were provided by Prof. P. Salvatore of the Department of Molecular Medicine and Medical Biotechnologies, Hospital University of Naples Federico II. The DSM 2008 strain (intestinal isolate) was purchased from the DSMZ/German Collection of Microorganisms and Cell Cultures GmnH (<https://www.dsmz.de>). The *V. parvula* clinical oral isolate was taken from a throat swab of a healthy subject, at the Policlinico Hospital of Naples, Federico II, Italy. Both these strains were grown and treated under the same conditions, using Brain Heart Infusion Broth (BHI) + 0.6% Sodium Lactate for 48 hours in anaerobic conditions. For Chapter 7, the *B. uniformis* dried cellular pellet was kindly provided by Prof. Salonen and Dr. Matharu of the University of Helsinki, in the Faculty of Biology and Environmental Sciences. The bacterium was isolated and grown in anaerobic conditions using a protocol set up within their group using anaerobic growth bags, in Bacteroides Bile Esculin (BBE) media (Anaerobe systems Inc). In Chapter 8, *P. pulmonicola* was purchased from the LMG bacterial collection of Ghent University and grown by Dr. S. McClean at the Centre of Microbial Host Interactions, Institute of Technology Tallaght, as described in Pither

et al. (2021).³ *E. pacifica* type strain KMM 6172^T and *E. vietnamensis* type strain KMM 6221^T were both from the Collection of Marine Microorganisms (KMM) of the G.B. Elykov Pacific Institute of Bioorganic Chemistry, Far Eastern Branch of Russian Academy of Sciences. The strains were cultivated and kindly provided by Dr. M. Kokoulin.⁴

9.2 Extraction and purification of LPS

9.2.1 Preliminary manipulations and checks

Prior to starting any LPS extraction protocol, preliminary manipulations are undertaken. These were in the form of cell washing or ultracentrifugation of the cells. Bacterial cell washes involved a first wash with distilled water and the centrifuged ($10,000 \times g$, $10\text{ }^{\circ}\text{C}$, 20 mins), supernatant collected and checked through ^1H NMR experimentation. This step was repeated with ethanol, acetone, and light petroleum. Ultracentrifugation was achieved through suspension of dried bacterial cells in water, divided evenly in ultracentrifuge tubes using mechanical pipettes, and centrifuged in a Beckman Coulter's ultracentrifuge (Brea, CA, USA) at $100,000 - 500,000 \times g$ at $4\text{ }^{\circ}\text{C}$ for 6 hours.

A preliminary check was also performed through the preparation and run of an electrophoresis gel (13.5 % SDS-PAGE) of the LPS upon its selective extraction from a little aliquot of bacterial cells. Briefly, the cells were lysed by enzymatic digestion with proteinase K [5mg/mL] at $56\text{ }^{\circ}\text{C}$ for 10 minutes in sample buffer [0.5 M Tris-HC (pH 6.8), glycerol (20% v/v), mercaptoethanol (4% v/v), SDS (4% w/v), and bromophenol blue (0.2% w/v)]. The LPS was then visualised using 13.5 % SDS-PAGE followed by silver nitrate staining.^{Error! Bookmark not defined.}

9.3 Extraction of the R-LPS and S-LPS

Two methods for LPS extraction were employed during this PhD thesis work. The PCP extraction method, that involved the preparation of a solution of Phenol/Chloroform/Light Petroleum (2:5:8 v/v/v). The cells were dissolved at room temperature in this solution and stirred for 90 minutes. The mixture was then centrifuged for 30 minutes at $8,000 \times g$. The supernatant was collected in a round bottom flask and the extraction was repeated two more times. The solution was concentrated under vacuum, removing the chloroform and the petroleum, until only the phenol remains. The solution was then transferred in a falcon and few drops of cold water were added. A centrifugation was performed ($8,000 \times g$, 1 h, RT) to pellet any precipitate containing the LPS.⁵ The other extraction method used was the hot phenol/water procedure; warm water was added to the cells following the PCP extraction or directly to a freeze-dried bacterial pellet, followed by equal volume of warm phenol. The suspended cells were stirred for 120 minutes at $70 \text{ }^\circ\text{C}$ and then centrifuged ($8,000 \times g$, 30 mins, $4 \text{ }^\circ\text{C}$). The water phase (supernatant) was collected, and the extraction was repeated two more times. Both the collected water phase and the phenol phases are dialysed against water (12-14 kDa). Phases were collected and freeze-dried.⁶ To remove cell contaminations, a treatment with RNase (Roth), DNase (Roth), and Proteinase K (Roth) at 37 and $56 \text{ }^\circ\text{C}$ followed by dialysis against distilled water was executed.⁷

Following the extraction, the LPS was checked by preparing a 13,5% SDS-PAGE gel followed by Alcian blue⁸ and silver nitrate stainings.^{7,9} Protein contamination was evaluated using staining with Coomassie Blue Solution (0.5% w/v Coomassie Brilliant Blue G-250 w/v in 40% methanol, 5% acetic acid).¹⁰

Further purification methods of the LPS material were also carried out. The methods for each project are summarised below.

9.4 Purification and isolation of saccharide moieties of the LPS

9.4.1 Chapter 3: Purification of the S-LPS and isolation of the core OS and O-antigen from *B. vulgatus* mpk

S-LPS extracted from *B. vulgatus* was further purified by several ultracentrifugation steps (4 °C, 100,000 × g, 24 h) and a gel-filtration chromatography Sephacryl S-300 (GE Healthcare, 1.5 × 90 cm, eluent 50 mM NH₄HCO₃). An aliquot of the pure S-LPS was treated with anhydrous hydrazine (2 mL, stirred at 37 °C, 90 min), cooled, poured into ice-cold acetone, and allowed to precipitate. The precipitate was centrifuged (4000 × g, 30 min), washed with ice-cold acetone, dried, dissolved in water, and lyophilised. The *O*-deacylated product was then *N*-deacylated with 4 M KOH (120 °C, 16 h). The removal of salts was executed by gel-filtration chromatography on a Sephadex G-10 column (Pharmacia, 50 × 1.5 cm, eluent distilled water). The fully deacylated product was further purified on a Toyopearl TSK HW-50 instrument (Tosoh Bioscience, 1.5 × 90 cm, eluent 50 mM NH₄HCO₃).

9.4.2 Chapter 4: Purification of the R-LPS and S-LPS of *A. faecalis* and the isolation of the core OS and the O-antigen fractions

The R-LPS was isolated from *A. faecalis* through the PCP method and purified by means of size-exclusion chromatography on a Sephacryl HR-S300 and the S-LPS from the hot phenol/water extraction on a Sephacryl HR-S400 columns (GE Healthcare Life Sciences, 1.5 × 90 cm, 50 mM NH₄HCO₃).

The R-LPS from *A. faecalis* was treated with anhydrous hydrazine (1.5 mL), stirred at 37 °C for 90 min, cooled, poured into ice-cold acetone and allowed to precipitate. The precipitate was then centrifuged (4,500 ×g, 30 min), washed with ice-cold acetone, dried, dissolved in water and lyophilized. The O-deacylated product was then N-deacylated with 4 M KOH (120 °C, 16 h). Salts were removed by gel-filtration chromatography on a Sephadex G-10 column (Pharmacia, 50 x 1.5 cm, eluent distilled water). The fully deacylated product was then further purified on a Toyopearl TSK HW-40S column (Tosoh Bioscience, 1.5 × 90 cm, eluent 50 mM NH₄HCO₃). An aliquot of R-LPS fraction also underwent a mild acid hydrolysis with acetate buffer (pH 4.4, 2 h, 100 °C) followed by a centrifugation step (8,000 × g, 30 min, 4 °C) that allowed the separation of the saccharide component in the supernatant which was then further purified by means of a gel filtration chromatography Toyopearl TSK HW-40S column (Tosoh Bioscience, 1.5 × 90 cm, eluent 50 mM NH₄HCO₃). The above acetate buffer treatment was also performed on the S-LPS fraction (10 mg) derived from the hot phenol/water extraction, to isolate the O-antigen moiety. The O-antigen fraction was then purified through gel-filtration chromatography on a Sephacryl HR-S100 column (GE Healthcare Life Sciences, 1.5 × 90 cm, eluent 50 mM NH₄HCO₃).

9.4.3 *Chapter 5: Purification of the R-LPS and isolation of core oligosaccharide from B. thetaiotaomicron*

The R-LPS from *B. thetaiotaomicron* was purified after extraction with the hot phenol/water method, using several ultracentrifugation steps (4 °C, 208,000 × g, 24 h) and followed by a gel-filtration chromatography on a Sephacryl S-300 (GE Healthcare, 1.5 × 90 cm, eluent 50 mM NH₄HCO₃).

An aliquot of pure R-LPS (~15 mg) treated as described above to obtain a full deacylated product. The fully deacylated product (OS1) was then further purified on a Toyopearl TSK HW-50S (Tosoh Bioscience, 1.5 × 90 cm, eluent 50 mM NH₄HCO₃). In parallel, an aliquot of R-LPS (~10 mg) was dissolved in 1 % acetic acid and hydrolysed at 100 °C for 2 h, the reaction was centrifuged (4 °C, 8,800 × g, 60 min) to allow for the supernatant containing the saccharide portion (OS2) to be collected. The OS2 fraction was then purified by gel-filtration chromatography on a Bio-Gel P4 column (Bio-Rad, 1.5 × 90 cm, eluent distilled water).¹¹

9.4.4 *Chapter 6: Isolation of core OS and O-antigen from V. parvula intestinal and oral isolate*

Both S-LPS and R-LPS were subject to extensive purification techniques: enzymatic digestions, as previously explained, and ultracentrifugation (4 °C, 208,000 ×g, 24 h). An aliquot of the S-LPS (~10 mg) was dissolved in 1 % acetic acid and hydrolysed at 100 °C for 4 h. The lipid A precipitate was pelleted using centrifugation (4 °C, 8,800 × g, 60 min) and then the supernatant containing the saccharide mixture of the LPS was loaded on Toyopearl TSK HW40S (Tosoh Bioscience, 1.5 × 90 cm, eluent 50 mM NH₄HCO₃). Fractions containing the O-antigen and the core OS were pooled and freeze-dried. This acid hydrolysis treatment was repeated on an aliquot of R-LPS from the oral isolate (~10 mg). Followed by SEC chromatography on a Toyopearl TSK HW4-0S (Tosoh Bioscience, 1.5 × 90 cm, eluent 50 mM NH₄HCO₃). Fractions containing the core OS were collected and freeze dried. An aliquot of both pure R-LPS (~8 mg) and S-LPS (5 mg) was fully deacylated and purified as described above.^{7,11}

9.4.5 Chapter 7: Purification of LPS and isolation of core OS from *B. uniformis*

After the S-LPS was extracted, additional purification of the extracted LPS was performed using enzymatic digestions using pullulanase and amylase (56 °C, 16 hours) followed by dialysis against milli-Q water (1 kDa, Sigma-Aldrich Co.). An aliquot of S-LPS (~3 mg) was dissolved in 1 % acetic acid and hydrolysed at 100 °C for 4 h. The lipid A precipitate was pelleted using centrifugation (4 °C, 8,800 × g, 60 min) and then the supernatant containing the saccharide mixture of the LPS was dialysed using a Pur-A-Lyzer™ Midi Dialysis Kit (1 kDa, Merck) against water. The extracted OS was then loaded onto a HyperSep™ Hypercarb™ (Thermo Scientific™).

9.4.6 Chapter 8: Isolation of LPS from the two strains of *Echinicola*

Following the extraction of the LPS from bacterial cells of *E. pacifica* KMM 6172^T and *E. vietnamensis* KMM 6221^T, both the LPS were subject to ultracentrifugation step (200,000 × g, 4 °C, 16 h) and a size-exclusion chromatography on a Sephacryl high-resolution S-500 column (GE Healthcare, 1.5 × 90 cm, eluent 50 mM NH₄HCO₃).

9.4.7 Chapter 8: Isolation of LPS from the *Pandoraea pulmonicola*

For the characterisation of the lipid A from *P. pulmonicola*, the S-LPS was extracted using the hot phenol/water method and purified using ultracentrifugation (200,000 × g, 16 h, 4 °C) and then loaded onto gel-filtration chromatography (Sephacryl High Resolution S-400, GE-Healthcare, 1.5 × 90 cm, eluent 50 mM NH₄HCO₃).

9.5 Chemical Analyses

9.5.1 Acetylated *O*-methyl glycosides (AMG)

An aliquot (~ 0.5 mg) of the sample to be investigated (purified LPS, OS, O-antigen) was treated with 300 μL of 1.25 M HCl/CH₃OH (80 °C, 16 h). The mixture was then neutralised using methanol washes and then acetylated with acetic anhydride (25 μL) in pyridine (25 μL) (80 °C, 30 min). The acetylated sugars were extracted using chloroform and water washes. The final extraction was then dried and diluted in 100 μL of acetone and analyzed by GC-MS.^{12,13}

9.5.2 Acetylated Alditols (AA)

An aliquot of freeze-dried sample to be analysed (~ 0.5 mg) was treated with 2 M TFA for 2 hours 120 °C and then dried under airflow. The monosaccharides were reduced with NaBH₄ in ethanol to produce alditols and acetylated with acetic anhydride (25 μL) in pyridine (25 μL) (80 °C, 30 minutes). The acetylated alditols were dried under airflow and diluted in 100 μL of acetone and analysed by GC-MS.

9.5.3 Partially Methylated Alditol Acetates (PMAA)

An aliquot of freeze-dried sample (~1 mg) suspended in DMSO, and powdered NaOH added, this mixture was then treated with 300 μL CH₃I (stirred, 16h, RT). The sample was placed under airflow for one hour. Washes were performed with chloroform and water, then the chloroform was extracted and dried under airflow. The sample was hydrolysed with trifluoroacetic acid (TFA) (4 M, 100 °C, 4 h) and dried under airflow, and neutralised with methanol washes. The carbonyl functions were then reduced using NaBD₄ in ethanol (stirring, 16 h, RT). The reaction was stopped by adding a few drops of HCl (2 M) and then neutralised by means of methanol washes. The free hydroxyl groups were acetylated using acetic anhydride in pyridine (80 °C, 30 minutes) followed by GC-MS study.

9.5.4 Octyl-glycosides

Enantiopure 2-(+ or -)-octanol (100 μ L) was added to aliquot of the freeze-dried sample (~1 mg). Racemic 2-(\pm)-octanol racemic alcohol (100 μ l) was added to another aliquot in a separate tube. Acetyl chloride (15 μ L) was added and heated at 60 °C overnight. The samples were dried under airflow and then acetylated with acetic anhydride in pyridine (80 °C, 30 minutes). Injection into GC-MS for comparison with known standards.^{12,13}

9.5.5 Gas Chromatography Mass Spectrometry

The analyses were all performed on an Agilent Technologies Gas Chromatograph 7820A equipped with a mass selective detector 5977B and an HP-5 capillary column (Agilent, Milan, Italy 30 m \times 0.25 mm, flow rate 1 mL/min, He as carrier gas). The temperature program used to analyse AMG and OGA was: 140 °C for 3 min, then 140 \rightarrow 240 °C at 3 °C/min. The temperature program for PMAA was: 90 °C for 1 min, then 90 \rightarrow 140 °C at 25 °C/min, then 140 \rightarrow 200 °C at 5 °C/min, then 200 \rightarrow 280 °C at 10 °C/min, and finally 280 °C for 10 min. Finally, to analyse fatty acids content the following temperature program was used: 150 °C for 5 min, 150 to 280 °C at 3 °C/min, and 280 °C for 5 min.^{2,13}

9.6 Isolation of lipid A

An aliquot of extracted and purified LPS was dissolved in 1 % acetic acid and hydrolysed at 100 °C for 4 h. Centrifugation ($8,800 \times g$, 60 min, 4 °C,) allowed to separate in the precipitate the lipid A and in the supernatant the oligo-/polysaccharide (as previously explained). The lipid A precipitate was collected and washed with freshly prepared Bligh/Dyer mixture ($\text{CHCl}_3/\text{CH}_3\text{OH}/\text{H}_2\text{O}$, 2:2:1.8 (v/v/v) ratio).^{7,14} This mixture was then shaken and centrifuged ($8,800 \times g$, 20 min, 4 °C). The chloroform phase, containing the lipid A fraction was pooled, dried, and analysed by MALD-TOF MS and MS² to define the chemical structure of the lipid A.

9.6.1 Dephosphorylation of lipid A and derivatization to alditol acetates

This method was utilised in *Chapter 8*, to establish the nature of the additional hexosamine(s) decorating the *P. pulmonicola* RL8228 lipid A. A small aliquot (0.6 mg) of the mild acid hydrolysis product, i.e. the isolated lipid A fraction, was treated with 100 μL of 48% aqueous hydrofluoric acid (HF) (Sigma Aldrich®, St. Louis, MO, USA) at 4° C for 16 h to remove the phosphate groups. The sample was placed in an ice bath and the HF was then evaporated until dry. The HF-treated sample was then dissolved in distilled water and lyophilized

9.7 Nuclear Magnetic Resonance (NMR)

All structural assignments of the core OS and the O-antigen from *B. vulgatus* LPS (Chapter 3), the core and O-antigen of *A. faecalis* LPS (Chapter 4), OS1 and OS2 fractions of *B. thetaiotaomicron* LPS (Chapter 5), core OS and O-antigen of *V. parvula* LPS (Chapter 6), core OS from *B. uniformis* LPS (Chapter 7) were carried out as followed.

1D and 2D NMR spectra were recorded in D₂O at 298 K at pD = 7 with a Bruker 600 DRX spectrometer equipped with a cryoprobe. The spectra were calibrated with internal acetone ($\delta_{\text{H}} = 2.225$ ppm; $\delta_{\text{C}} = 31.45$ ppm). ³¹P NMR experiments were carried out with a Bruker DRX-400 spectrometer; aqueous 85 % phosphoric acid was used as external reference ($\delta = 0.00$ ppm). Total correlation spectroscopy (TOCSY) experiments were recorded with spinlock times of 100 ms by using data sets ($t_1 \times t_2$) of 4096×256 points. Rotating frame Overhauser enhancement spectroscopy (ROESY) and Nuclear Overhauser enhancement spectroscopy (NOESY) experiments were performed by using data sets ($t_1 \times t_2$) of 4096×256 points and by using mixing times between 100 and 300 ms, acquiring 16 scans. Double-quantum-filtered phase sensitive correlation spectroscopy (DQF-COSY) experiments were carried out by using data sets of 4096×512 points. The data matrix in all the homonuclear experiments was zero-filled in both dimensions to give a matrix of $4\text{K} \times 2\text{K}$ points and was resolution-enhanced in both dimensions by a cosine-bell function before S5 Fourier transformation. Coupling constants were determined by 2D phase-sensitive DQF-COSY.^{15,16} Heteronuclear single quantum coherence (¹H,¹³C HSQC) and heteronuclear multiple bond correlation (¹H,¹³C HMBC) experiments were executed in ¹H-detection mode by single-quantum coherence with proton decoupling in the ¹³C domain using data sets of 2048×256 points. ¹H,¹³C HSQC was performed using sensitivity improvement and in the phase-sensitive mode using Echo/Antiecho gradient selection, with multiplicity editing during selection step. ¹H,¹³C HMBC was optimised on long range coupling constants, with lowpass J-filter to suppress one-

bond correlations, using gradient pulses for selection. Moreover, a 60 ms relay was used for the evolution of long-range correlations. ^1H , ^{13}C HMBC was optimised for a 6 Hz coupling constant, and ^1H , ^{31}P HSQC was optimized for an 8 Hz coupling constant. The data matrix in all the heteronuclear experiments was extended to 2048×1024 points by using forward linear prediction extrapolation.

9.8 MALDI-TOF Mass Spectrometry and MS2

The MS investigation was performed on ABSCIEX TOF/ TOF™ 5800 Applied Biosystems mass spectrometer equipped with an Nd:YAG laser ($\lambda = 349$ nm), with a 3 ns pulse width and a repetition rate of up to 1000 Hz.. Each spectrum was the result of the accumulation of 1500 laser shots, whereas 6000–7000 shots were summed for the MS² data acquisitions. All reported experiments were acquired in both negative and positive polarity, reflectron mode. For MS spectra accumulation of 2000 laser shots was used, whereas 5000–7000 shots were summed for the MS/MS spectra. Each experiment was performed in replication

Lipid A preparation: for all cases reported in this PhD thesis, lipid A samples were solved in CHCl₃/CH₃OH (50:50, v/v). 2,4,6-trihydroxyacetophenone (THAP) in CH₃OH/0.1% trifluoroacetic acid/CH₃CN (7:2:1, v/v) at a concentration of 75 mg/mL was typically the matrix solution used. A 1:1, v/v sample/matrix solution mixture (<1 μ L) was deposited onto a stainless-steel MALDI sample plate and left to dry at room temperature.¹⁷

R-LPS preparation: As in the case of R-LPS of *B. thetaiotaomicron* (Chapter 5). An aliquot of intact R-LPS was suspended in 5 mM ethylenediaminetetraacetic acid (EDTA), shaken and left a few seconds in an ultrasonic bath. Then, R-LPS was desalted with cation exchange beads (Dowex 50 \times , NH₄⁺) and then added to an equal volume of 2,5-dihydroxybenzoic acid (DHB) in 0.1 % citric acid. 1 μ L of R-LPS/DHB mixture per well was transferred to the stainless-steel MALDI plate. MS spectrum of intact R-LPS was recorded in linear mode and negative ion polarity. As for the analysis of OS2 fraction, sample was dissolved in water (1 mg/mL) and DHB (10 mg/mL in CH₃CN/0.2 M citric acid, 1:1) was used as the matrix. 0.5 μ L of sample and 0.5 μ L of matrix solution were spotted on the plate and left to dry at room temperature. The spectra were recorded in both negative and positive polarity, but only in the positive-ion polarity it was possible to analyze the OS2 product.

9.9 HEK-Blue™ hTLR4/CD14/MD-2 cell culture, transfection, and stimulation

HEK-Blue™ Null2, hTLR4 and hTLR2 Stimulation Assays were performed as previously described.^{2,4,18} Briefly, HEK-Blue™ hTLR4, HEK-Blue™ hTLR2, and HEK-Blue™ Null2™ cells (InvivoGen, San Diego, CA, USA) were cultured at 37 °C in 5% CO₂ using Dulbecco's minimal essential media (DMEM) supplemented with 10% heat-inactivated fetal bovine serum (FBS) (Microgem, Naples, Italy), 1% glutamine (Himedia, Einhausen, Germany), 1% penicillin/streptomycin (Himedia, Einhausen, Germany), and 100 µg/mL Normocin (InvivoGen). To select the plasmids in HEK-Blue™ hTLR4 and hTLR2 cells, a mixture of selective antibiotics (HEK-Blue™ Selection) (InvivoGen) was used, whereas for HEK-Blue Null2™ cells 100 µg/mL Zeocin (InvivoGen) was employed.

HEK-Blue™ hTLR4 cells were seeded into 96-well plates (3×10^5 cells per well) and incubated with different concentrations of the extracted and purified LPS (S-LPS) and isolated lipid A from *B. vulgatus* (Chapter 3), R-LPS from *B. thetaiotaomicron* and the LPS from *E. pacifica*, *E. vietnamensis* (Chapter 8) and positive control (*S. typhimurium* SH 2201 LPS or *E. coli* (0127:B8)) (1, 10, 100 ng/mL). After 18 h incubation, the Nuclear Factor kappa B (NF-κB) activation was analysed by evaluating NF-κB-dependent secreted alkaline phosphatase (SEAP) using QUANTI-Blue™ (InvivoGen). Likewise in Chapters 3, 5, 6 and 8, the LPS were tested also on HEK-Blue™ hTLR2 and HEK-Blue™ Null2™ cell lines were exposed to the pure LPS or *S. typhimurium* SH 2201 LPS (1, 10, 100 ng/mL) for 18 h. For this experiment, also Pam3CSK4 (500 ng/mL) (InvivoGen) was analysed as above.

9.10 Multiple Reaction monitoring (MRM) analysis of HEK-Blue™ hTLR4/CD1/MD-2 cell experiment ¹⁹

9.10.1 In solution digestion of HEK-Blue™ cellular extracts following R-LPS stimulation

Cell growth media for every sample (four replicate each) were subjected to the in-solution digestion protocol. 200 µL of each cell growth medium (supernatant) were added to 800 µL of cold acetone (1:4, v/v) to precipitate proteins. The reduction reaction was performed by adding 20 µL of 100 mM 1,4-dithiothreitol (DTT) dissolved in 50 mM NH₄HCO₃ to each sample and conducted at 60°C for one hour. 20 µL of 120 mM iodoacetamide, dissolved in 50 mM NH₄HCO₃, were added to the sample and the reaction was left for 45 min at room temperature, in the dark. 500 µL of CH₃OH was used to precipitate proteins. The solution was then centrifuged for 15 min at 10,000 × g. The supernatants were removed, and the pellets dried under vacuum. The dried pellets were then suspended in 50 µL of 50 mM NH₄HCO₃ and mixed with 2 µL of trypsin (1 µg/µL). The hydrolysis reaction was performed at 37 °C for 15 h. Finally, the peptide mixture obtained for these supernatants was suspended in 12 µL of 2 % CH₃CN, 0.2 % formic acid solution to perform LC-MS/MS analysis. 100 µL of protein lysis buffer (7 M urea, 3.5 mM SDS, 0.01 M TRIS, 2 mM EDTA, 10 mM DTT) and 2 µL of protease inhibitors cocktail were added to each cellular pellet to conduct the lysis. Three freeze-thaw cycles of 30 minutes followed by a sonication step (30 min) were performed to enhance the cell lysis. Subsequently, protein hydrolysis was carried out as reported for the cellular growth media. The peptide mixture obtained after hydrolysis was suspended in 20 µL of 2% CH₃CN 0.2% formic acid solution to perform LC-MS/MS analysis.

9.10.2 LC-MS/MS analysis by MRM ion mode

The peptide mixtures were analysed by LC-MS/MS analysis by using a Xevo TQ-S (Waters Corporation) equipped with an IonKey UPLC Microflow Source coupled to a Waters UPLC Acquity System. For each run, peptide mixture (1 µL) was separated

on a iKey BEH C18 130, 1.7 μm , 150 μm x 50 mm (Waters, Milford, MA, USA) at 45 °C at flow rate of 3 $\mu\text{L}/\text{min}$. Water acidified with 0.1 % formic acid containing 2 % CH_3CN was used as eluent A, while CH_3CN with 5 % water and 0.1 % formic acid was eluent B. Peptides were eluted (starting 1 min after injection) with a linear gradient of eluent B in A from 7 % to 95 % in 20 min. The column was re-equilibrated at initial conditions for 4 min. The mass spectrometric analyses were performed in positive-ion mode by using a multiple reaction monitoring (MRM) detection; the duty cycle and dwell times were automatically set to the MassLynx. Cone voltage was set to 35 V. Skyline software (3.7, 64 bit version MacCoss Lab Software, University of Washington, USA) was used for setting m/z precursor ion, m/z product ions, and relative collision energy for each peptide in order to develop a targeted MRM method. The best two to six transitions per peptide were selected from the top ranked Y- and B-fragments. The selection of peptides to be monitored and the best transitions were carried out by comparing the Skyline output and SRM Atlas platform.

9.11 Molecular Modelling

9.11.1 Structure construction

The full structure of the human TLR4/MD-2 receptor (hTLR4/MD-2/E.coli)₂ was taken from the Protein Data Bank, PDB ID: 3FXI and antagonist pose of MD-2 ID: 2E59. The three-dimensional structures of the lipid A species were constructed by the molecular graphic and modelling package PyMOL (www.pymol.org) based on the lipid A from *E. coli* taken from Protein Data Bank (PDB ID 3FXI). The lipid A structures then went through restrained minimisation with Maestro.

9.11.2 Docking

The docking calculations were performed using Autodock VINA1.1.2 and then the best poses were selected for redocking with Autodock 4.2 (<http://autodock.scripps.edu>).^{20,21} The Lamarckian evolutionary algorithm was chosen for the Autodock 4.2. The structures of the TLR4/MD-2 were kept rigid but the lipid A structures remained partially flexible by providing freedom to some of the appropriately selected dihedral angles. The docking boxes were set to 1Å for VINA and 0.375Å for Autodock. The size of the box covered the centre mass of residues Arg90 (MD-2), Lys122(MD-2) and Arg264(TLR4) in the hTLR4/MD-2 receptor by being set to 33.00 Å x-axis, 40.50 Å y-axis, and 35.25 Å z-axis. The determination of the best result from each docking was based on the predicted binding energy and mode of interaction. All docked poses were then analysed, and images were then created using PyMOL.²²

9.12 References

1. Steimle, A., Michaelis, L., Di Lorenzo, F., Kliem, T., Münzner, T., Maerz, J. K., Schäfer, A., Lange, A., Parusel, R., Gronbach, K., Fuchs, K., Silipo, A., Öz, H. H., Pichler, B. J., Autenrieth, I. B., Molinaro, A. & Frick, J. S. Weak Agonistic LPS Restores Intestinal Immune Homeostasis. *Mol. Ther.* **27**, 1974–1991 (2019).
2. Pither, M. D., Illiano, A., Pagliuca, C., Jacobson, A., Mantova, G., Stornaiuolo, A., Colicchio, R., Vitiello, M., Pinto, G., Silipo, A., Fischbach, M. A., Salvatore, P., Amoresano, A., Molinaro, A. & Di Lorenzo, F. *Bacteroides thetaiotaomicron* rough-type lipopolysaccharide: The chemical structure and the immunological activity. *Carbohydr. Polym.* **297**, 120040 (2022).
3. Pither, M. D., McClean, S., Silipo, A., Molinaro, A. & Di Lorenzo, F. A chronic strain of the cystic fibrosis pathogen *Pandoraea pulmonicola* expresses a heterogenous hypo-acylated lipid A. *Glycoconj. J.* **38**, 135–144 (2021).
4. Pither, M. D., Mantova, G., Scaglione, E., Pagliuca, C., Colicchio, R., Vitiello, M., Chernikov, O. V., Hua, K.-F., Kokoulin, M. S., Silipo, A., Salvatore, P., Molinaro, A. & Di Lorenzo, F. The Unusual Lipid A Structure and Immunoinhibitory Activity of LPS from Marine Bacteria *Echinicola pacifica* KMM 6172^T and *Echinicola vietnamensis* KMM 6221^T. *Microorganisms* **9**, 2552 (2021).
5. Galanos, C., Luderitz, O. & Westphal, O. A New Method for the Extraction of R Lipopolysaccharides. *Eur. J. Biochem.* **9**, 245–249 (1969).
6. Westphal, O. & Jann, K. Bacterial Lipopolysaccharides Extraction with Phenol-Water and Further Applications of the Procedure. *Methods Carbohydr. Chem.* **5**, 83–91 (1965).
7. Pither, M., Molinaro, A., Silipo, A. & Di Lorenzo, F. in *Methods Mol. Biol.* **2613**, (Springer International Publishing). **In press**.
8. Corzo, J., Pérez-Galdona, R., León-Barrios, M. & Gutiérrez-Navarro, A. M. Alcian blue fixation allows silver staining of the isolated polysaccharide component of bacterial lipopolysaccharides in polyacrylamide gels. *ELECTROPHORESIS* **12**, 439–441 (1991).
9. Tsai, C.-M. & Frasch, C. E. A sensitive silver stain for detecting lipopolysaccharides in polyacrylamide gels. *Anal. Biochem.* **119**, 115–119 (1982).
10. Dong, W. H., Wang, T. Y., Wang, F. & Zhang, J. H. Simple, time-saving dye staining of proteins for sodium dodecyl sulfate-polyacrylamide gel electrophoresis using coomassie blue. *PLoS ONE* **6**, (2011).

11. Holst, O. Deacylation of lipopolysaccharides and isolation of oligosaccharide phosphates. *Methods Mol. Biol. Clifton NJ* **145**, 345–353 (2000).
12. De Castro, C., Parrilli, M., Holst, O. & Molinaro, A. Microbe-Associated Molecular Patterns in Innate Immunity: Extraction and Chemical Analysis of Gram-Negative Bacterial Lipopolysaccharides. *Methods Enzymol.* **480**, 89–115 (2010).
13. Garcia-Vello, P., Speciale, I., Di Lorenzo, F., Molinaro, A. & De Castro, C. in *Lipopolysaccharide Transp. Methods Protoc.* (ed. Sperandeo, P.) 181–209 (Springer US, 2022).
14. Bligh, E.G. and Dyer, W. J. Canadian Journal of Biochemistry and Physiology. *Can. J. Biochem. Physiol.* **37**, (1959).
15. Rance, M., Sørensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R. & Wüthrich, K. Improved spectral resolution in COSY (1)H NMR spectra of proteins via double quantum filtering. 1983. *Biochem. Biophys. Res. Commun.* **425**, 527–533 (2012).
16. Piantini, U., Sorensen, O. W. & Ernst, R. R. Multiple quantum filters for elucidating NMR coupling networks. *J. Am. Chem. Soc.* **104**, 6800–6801 (1982).
17. Di Lorenzo, F. The lipopolysaccharide lipid A structure from the marine sponge-associated bacterium *Pseudoalteromonas* sp. 2A. *Antonie Van Leeuwenhoek Int. J. Gen. Mol. Microbiol.* **110**, 1401–1412 (2017).
18. Di Lorenzo, F., Pither, M. D., Martufi, M., Scarinci, I., Guzmán-Caldentey, J., Łakomicz, E., Jachymek, W., Bruijns, S. C. M., Santamaría, S. M., Frick, J.-S., van Kooyk, Y., Chiodo, F., Silipo, A., Bernardini, M. L. & Molinaro, A. Pairing *Bacteroides vulgatus* LPS Structure with Its Immunomodulatory Effects on Human Cellular Models. *ACS Cent. Sci.* **6**, 1602–1616 (2020).
19. Illiano, A., Pinto, G., Gaglione, R., Arciello, A. & Amoresano, A. Inflammation protein quantification by multiple reaction monitoring mass spectrometry in lipopolysaccharide-stimulated THP-1 cells. *Rapid Commun. Mass Spectrom.* **35**, (2021).
20. Trott, O. & Olson, A. J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. *J. Comput. Chem.* **31**, 455–461 (2010).
21. Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S. & Olson, A. J. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.* **30**, 2785–2791 (2009).
22. Lorenzo, F. D., Kubik, Ł., Oblak, A., Lorè, N. I., Cigana, C., Lanzetta, R., Parrilli, M., Hamad, M. A., Soysa, A. D., Silipo, A., Jerala, R., Bragonzi, A., Valvano, M. A., Martín-Santamaría, S. & Molinaro, A. Activation of Human Toll-like

Receptor 4 (TLR4)·Myeloid Differentiation Factor 2 (MD-2) by Hypoacylated Lipopolysaccharide from a Clinical Isolate of *Burkholderia cenocepacia*. *J. Biol. Chem.* **290**, 21305–21319 (2015).

