Structural investigation of cell-surface glycoconjugates and their interactions

Cédric Laguri

Cédric Laguri

Habilitation à Diriger des Recherches

Structural investigation of cell-surface glycoconjugates and their interactions.

Université Grenoble-Alpes
Ecole doctorale chimie et sciences du vivant

L'HDR sera soutenue le 31 mai 2021 devant le jury compose de :
Madame Alba Silipo.
Professeur University of Napoli Federico II. Rapportrice
Madame Sophie Zinn-Justin
Directrice de recherche, I2BC Gif-sur-yvette. Rapportrice
Monsieur Christoph Rademacher
Professeur, University of Wien. Rapporteur
Madame Anne Imberty
Directrice de recherche, CERMAV Grenoble. Examinatrice
Monsieur Romain Vives
Directeur de recherche, IBS Grenoble. Examinateur

PREFACE

I studied biochemistry at University of Cergy-Pontoise (outside Paris), as I had a broad interest in science but no clear idea of what I wanted to do. The first three years were highly uninteresting, mostly because teaching remained very scholar. It became fascinating for me during the master's year, as we started to study research articles, biophysical methods and structural biology. I did a master 2 in Nancy University where I learned a lot about protein purification and characterization and had my first real course in NMR and the first look at an NMR machine. I then decided to shift gear and do another master 2 degree this time in Molecular Biophysics in Paris 6 University (Now Sorbonne University!). This move was not usual, and I was told by the head of the master's degree that he accepted me in the course but that I was already too old (25) and had no chance to join CNRS….

The master was great with excellent courses, good teachers, and motivated students. I must confess I understood little of the NMR courses but, being stubborn did anyway my M2 project in the DIEP CEA/Saclay with Bernard Gilquin and Sophie Zinn-Justin on the NMR of a small protein. I could solve my first assignment puzzle with a ruler and pencil but more interestingly I was lucky to be there at the time when development of automated structure determination using ambiguous restraints was starting. In the end we determined the structure in a few months which was quite quick at the time. Having missed the ministry of research scholarship I turned towards the UK to do a PhD in Biomolecular NMR.

Sheffield's NMR lab was great, and life in and around pubs was very stimulating and I made friends for life there. While UK system leaves the student mostly on their own with minimal interactions with their supervisors, it is the perfect opportunity to test its good (and bad) ideas and to acquire autonomy. In the end we did some structure and nice interactions with DNA. I did then my first post-doc back in Saclay, still on structure determination and interactions with Nucleic Acids.

With that background I was lucky to join the Institut de Biologie Structurale in Grenoble to start investigating protein-sugars interactions in 2005 that led me to be recruited by the CNRS in 2008 (Thanks Hugues). Since then, I could study these interactions by NMR and other techniques but could also develop methods to generate ^{13}C labelled sugars to be used in interactions. I recently switched to the study of bacterial glycoconjugates, their transport pathways as well as their interaction with proteins. NMR is of course particularly adapted to the study of heterogeneous, poorly ordered macromolecules even when they are insoluble by Magic Angle Spinning solid state NMR. Those challenging systems also require the use of integrated structural biology approaches that always lead to the use of new biophysical methods and satisfy my appetite for always learning new techniques.

Cell-surface glycoconjugates are ubiquitous in nature.

Virtually all cell surfaces are decorated with glycan chains. The nature and abundance of these chains is highly variable, but they serve many functions in cell-cell communication, host-pathogens interactions, and protection against the environment. They are found often in the form of glycoconjugates, covalently attached to proteins anchored to the cell-surface (N- or O- glycosylation) or to lipids involved in the membrane structure to form glycolipids. Glycans are of course not DNA

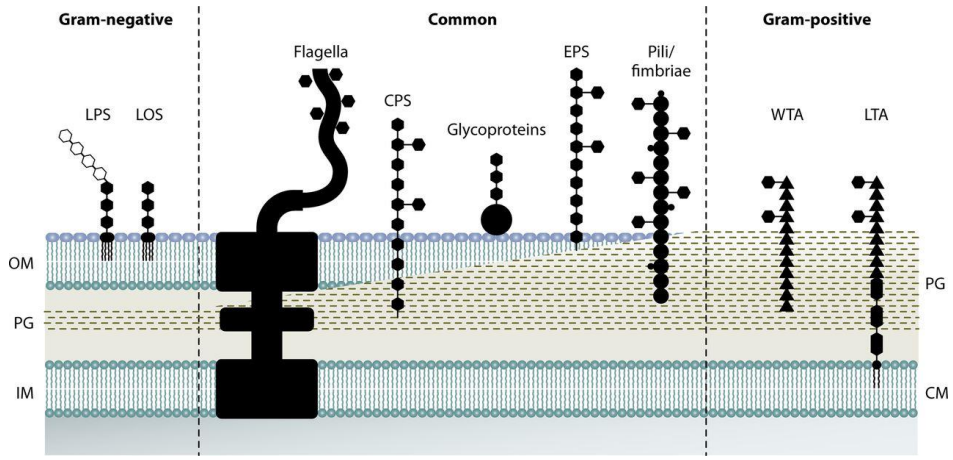
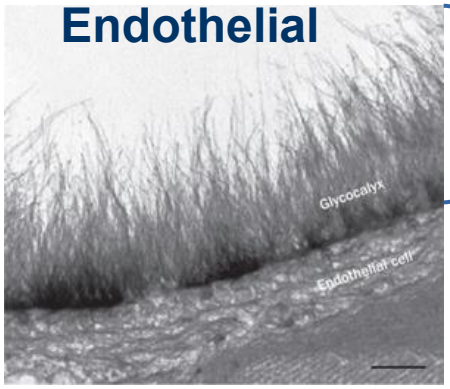


Figure 1 Bacterial glycoconjugates.

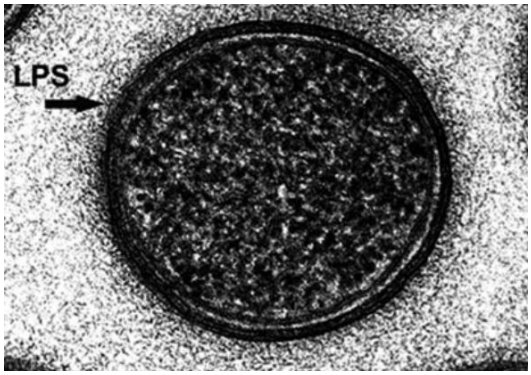
Bacteria synthesize a plethora of glycoconjugates. Characteristic for Gram-negative species are lipopolysaccharide (LPS) and lipooligosaccharide (LOS) structures, while several Gram-positive species are known to glycosylate their teichoic acids (TAs) (wall teichoic acid [WTA] and lipoteichoic acid [LTA]). Both Gram-negative and -positive species can produce glycosylated flagella, pili or fimbriae, capsular polysaccharide (CPS), exopolysaccharide (EPS), and glycoproteins and a peptidoglycan (PG) layer (from [44])

encoded and rely on complex biosynthetic machineries that are regulated depending on the cell context to vary sugars compositions at the cell surface. The first part of this document will focus on Heparan Sulfate Glycosaminoglycan, a complex sugar present at surfaces of higher eukaryote cells (Figure 2). The second part of this document will focus on LipoPolySaccharides, which compose the outside layer of gram-negative bacteria, are a key virulence factor and constitute now my main research activity.



Gly-
cocalix

Figure 2 : Electron micrographs of endothelial glycocalyx (top) that contains Heparan Sulfates and bacterial surface (bottom) showing LPS..



Gram -

bacterium

Hunter, Beveridge

2005

I. Structural basis of Heparan sulfates-protein interactions

Heparan sulfate are complex linear polysaccharides found covalently linked to proteins at the surface of all higher eukaryote's cells and in the extra-cellular matrix. They are composed of a repetition of a disaccharide motif comprising a glucuronic acid (GlcA) and an N-acetyl Glucosamine (GlcNAc) (Figure 3). The complexity of GAG relies on the presence in cells of a complex enzymatic machinery located in the Golgi system with NDST (N-deacetylase N-sulfotransferase) enzymes that convert GlcNAc into GlcNS and creates N-sulfated stretches. In these so-called S-domains GlcA is epimerized into IdoA (Iduronic acid) by the C5-epimerase and IdoA and more marginally GlcA can be sulfated in the 2- position by 2-O Sulfotransferase (2-OST). Then 6-OST and 3-OST add sulfate groups on GlcNS at position 6 and 3 respectively. Nevertheless, the final composition of the polysaccharides highly depends on the cell type and state. Indeed, the presence of several isoforms of the biosynthetic enzymes (NDST, 3-OST, 6-OST) that have different specificities and the regulation of their expression complicates the pattern of modifications found on HS on different cells. These different HS exposed by cells allow a fine tuning of the binding of hundreds of proteins that use Heparan sulfates as a primary receptor and thus mediate important cellular processes such as

differentiation, inflammation, cell migration. They are also exploited by pathogens to adhere to cells.

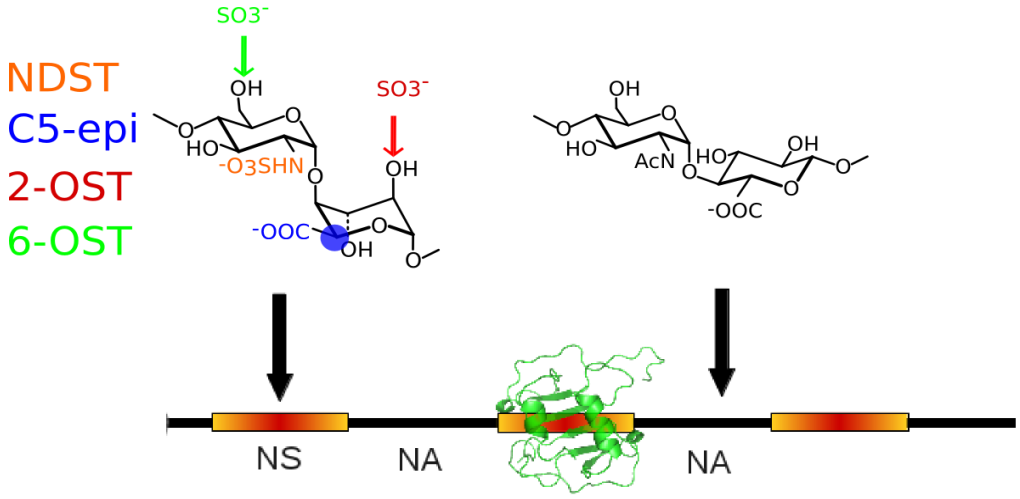


Figure 3 Scheme of HS biosynthesis

Top: Sequential Biosynthesis of Heparan Sulphates (From [1])

Bottom: Details of the positions modified by biosynthetic enzymes. Biosynthetic enzymes generate modified stretches (yellow-orange) called NS for N-sulfated domains while regions termed NA (N-acetylated) remain largely unmodified. NS regions mediate interactions with proteins.

While several hundred proteins have been shown to interact with Heparan Sulfates, the extraction of atomic-scale information about these interactions remain very elusive. First, it is difficult to obtain large quantities of homogeneous HS oligosaccharides necessary to perform structural studies. This is mostly achieved starting from a highly sulfated form of HS called Heparin that is found abundantly in pig's liver or by chemistry. Chemistry methods involve multiple steps of protection/ deprotection since sulfated positions are chemically equivalent. They are thus mostly applied for small oligosaccharides. Heparin fragments on the other hand can be generated from enzymatic digestion of polysaccharides extracted from pig's liver followed by size and charge separation. They can be obtained in large amounts, at least the fully sulfated species, but they constitute a poor representation of HS modification patterns found at cell-surfaces. Both sources of HS have been used successfully in numerous biochemical and X-ray/NMR studies of HS interacting proteins.

The second difficulty in obtaining atomic scale details on these interactions is to obtain stable HS/proteins complexes that can be crystallized or characterized by NMR with NOEs structural restraints. The inherent plasticity of oligosaccharides-protein interactions indeed hampers lots of structural studies. NMR titration studies, that rely on observing NMR frequencies changes on the protein resonances, are often the best way to map the interaction sites on a protein, even when interaction is weak. This approach does provide nevertheless only limited information on the glycans'

moieties important for the interactions.

The discovery of a capsular polysaccharide called Heparosan (or K5 polysaccharide,) extracted from a pathogenic *E. coli* strain, that possesses the same composition as unmodified Heparan Sulfate and can be chemically modified[2] combined with the use of recombinant HS modifying enzymes paved new ways for generation of HS oligosaccharides[3]. I will present here strategies to generate ^{13}C-labeled HS analogues of different sizes and sulfation patterns which can be used to study protein HS interactions by NMR.

A. Chemoenzymatic design of ^{13}C labelled heparan sulfate oligosaccharides

Tools were developed to produce oligosaccharides possessing different modification motifs and ^{13}C labelled for their study by NMR alone and in interactions with proteins. *E. coli* K5 bacteria are urinary tract pathogens and produce a capsular polysaccharide that possesses the same structure as unmodified HS (GlcNS-GlcNAc), called Heparosan (HSAN)(Figure 4). The glycan backbone is extracted from *E. coli* K5 bacteria grown in medium containing ^{13}C as the sole carbon source, with a yield approaching 50mg of polysaccharide per litre of culture in M9 medium[4]. Polysaccharides obtained are N-deacetylated and N-sulfated chemically (Figure 4) to generate 95% N-sulfated polymers, NS-HSAN (as checked by NMR), and then modified chemically and/or enzymatically.

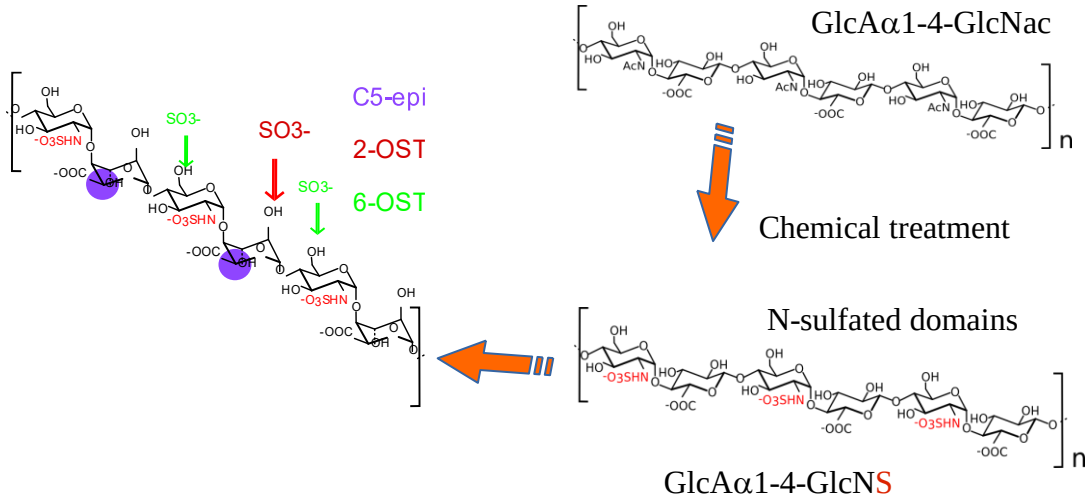


Figure 4 Chemo-enzymatic design of ^{13}C Heparan sulfates.

Polysaccharides with a GlcA-GlcNAc repetition are extracted from bacteria and subsequently modified and digested

Following N-sulfation, the next modification steps cannot all be achieved selectively by chemistry methods and require the natural biosynthetic enzymes.

1. *Setup of C5-epimerization and 2-O sulfation of N-sulfated ^{13}C-labeled HS precursors*

C5-epi, 2-OST and 6-OST1 were expressed and purified from *E. coli* as recombinant proteins (as MBP fusions) and tested for their activity on ^{13}C-labelled NS-HSAN. For the sulfation reaction to occur, a constant supply of the sulfate donor PAPS (3'-Phosphoadenosine-5'-phosphosulfate) is provided by a regeneration system using Aryl-Sulfo-Transferase (AST-IV) and PNPS(Potassium 4-nitrophenyl sulfate)[5]. Enzymatic reactions are easily followed in 2D ^{1}H-^{13}C correlation experiments, benefiting from the ^{13}C labelling of the polysaccharides (Figure 5). Incubation of NS-HSAN with C_5-epi converts GlcA to IdoA and 2-OST sulfates mostly IdoA sugars to IdoA2S and more marginally on GlcA to GlcA2S. Enzymatic reactions can also be followed in real time, confirming that C_5-epi converts GlcA to IdoA in a reversible way up to a 30%IdoA/ 70% GlcA equilibrium[6], and that when both enzymes are introduced together 2-OST has no influence on C5-epi catalytic rate(Figure 5). Nevertheless while the rate of C5-epi is not influenced by 2-OST presence, we could demonstrate the two proteins assemble (Kd=80nM) and that, when associated, C5-epi and 2-OST generate extended GlcNS-IdoA2S stretches[7]. 6-OST activity was also confirmed by NMR and the recombinant enzymes proved to be sufficiently active for the purpose of modifying milligrams amount of ^{13}C labelled polysaccharides.

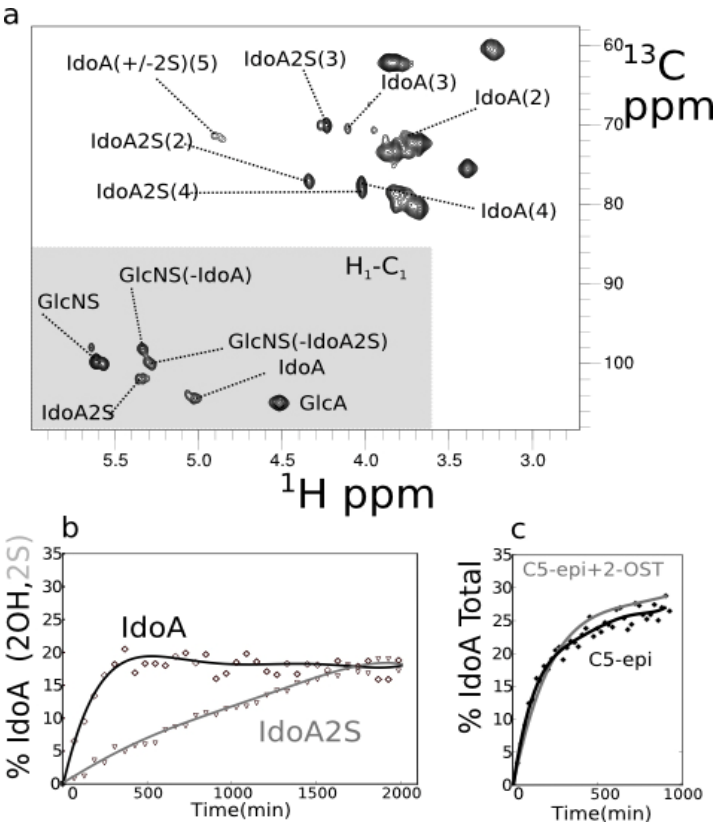


Figure 5 Monitoring of the activities of the C_5-Epi and 2OST activities by NMR.

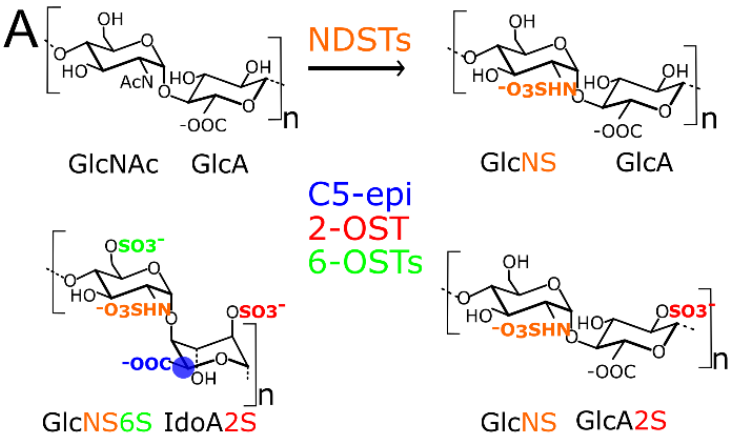
(A) Correlation Spectrum $^{13}C^{1}H$ of NS-HSAN after reaction by C_5-Epi and 2OST. (b) Evolution of IdoA and IdoA2S species during enzymatic reaction. (c) Total amount of IdoA (± 2s) generated by C5-Epi alone (black) or in the presence of 2-OST (grey).

2. *Generation of libraries of size defined ^{13}C-labeled HS oligosaccharides*

We developed two related strategies to engineer HS. In the first approach, we combined the capacity of the C5-epi and 2-OST enzymes to generate extended GlcNS-IdoA2S repeats when they function as a complex [7] and the specificity of Heparinase III to isolate ΔHexA-(GlcNS-IdoA2S/GlcA2S)$_n$-GlcNS resistant domains which are size fractionated, modified with 6-OST-1, then purified by HPLC. The second approach consists first in generating size defined (GlcNS-GlcA)$_n$ repeats through controlled Heparinase II digestion of N-sulfated Heparosan followed by sequential modification by the C5-epi/2-OST and 6-OST-1 enzymes and purification by HPLC.

Oligosaccharides compositions were assessed by ^{1}H-^{13}C NMR experiments, using the ^{1}H and ^{13}C chemical shifts of the pyranose ring that are highly sensitive on its modifications (presence of O-sulfate substituents or epimerization) and on the modifications on the neighbouring residue[4,7–9].

A



B

	1	2	3	4	5	6	sulfates
dp4[1]	ΔHexA	GlcNS	**IdoA2S**	GlcNS			3
dp4[2]	ΔHexA	**GlcNS6S**	**IdoA2S**	GlcNS			4
dp4[3]	ΔHexA	**GlcNS6S**	GlcA	GlcNS			3
dp4[4]	ΔHexA	**GlcNS6S**	GlcA	**GlcNS6S**			4
dp6[1]	ΔHexA	**GlcNS6S**	GlcA	GlcNS	GlcA	GlcNS	4
dp6[2]	ΔHexA	GlcNS	**GlcA2S**	GlcNS	**IdoA2S**	GlcNS	5
dp6[3]	ΔHexA	**GlcNS6S/GlcNS**	**GlcA/IdoA/IdoA2S**	**GlcNS6S**	**GlcA/IdoA/IdoA2S**	**GlcNS/GlcNS6S**	>5
dp6[4]	ΔHexA	**GlcNS6S**	**IdoA2S**	**GlcNS6S**	**IdoA2S**	GlcNS	7
dp6[5]	ΔHexA	GlcNS	**IdoA2S**	GlcNS	**IdoA2S**	GlcNS	5

Figure 6(A) Heparan sulfate modifications. N-deacetylation, N-sulfation step and examples of modifications produced by C5-epi and 2-or 6- OST enzymes. (B) Table summarizing the tetra- and hexa-saccharides chemo-enzymatically produced purified, and their sequence determined by NMR, including dp63 that remained heterogeneous.

B. Interaction of CXCL12α, CXCL12γ chemokines with diverse 13C labelled HS oligosaccharides

1. *Interaction of chemokines with four diverse tetrasaccharides*

As a model system to investigate their structure-binding activity relationships, we used two related chemokines, CXCL12α and CXCL12γ whose interactions with HS have been functionally characterized by NMR. They display distinct affinity for heparin, used as an HS proxy in binding

studies, of 93 and 0.9nM. [10–13] CXCL12α and CXCL12γ occur from alternative splicing of the same gene, possess the same first 68 amino acids, with a 30 amino-acids long C-terminal unstructured and highly basic extension for the γ isoform. CXCL12γ possesses two HS binding sites, one strictly required, found in the core structured domain of the protein shared with CXCL12α, the other one in the C-terminus which, by enhancing the half-life of the complexes, contributes to a strong retention on cell surfaces [11].

The interaction of four oligosaccharides ($dp4^{1\text{-}4}$ Figure 6B) was tested with adding increasing concentrations of CXCL12γ and CXCL12α to the oligosaccharide up to 4:1 protein:oligosaccharide molar ratios while monitoring the sugars resonances by ^{13}C-^{1}H correlation NMR experiments. Changes in the environment of the oligosaccharides nuclei due to protein interaction is reflected by a change in their chemical shift termed hereafter CSP (chemical shift perturbation).

Interaction with $dp4^1$ that contains an IdoA2S in position 3 with CXCL12α and CXCL12γ show CSPs around the IdoA2S residue, reflecting binding induced at this position in the presence of the proteins (Figure 7). Binding to $dp4^1$ can be compared with an oligosaccharide of different structure $dp4^3$ (here a 6-sulfate at position 2 and no IdoA2S), but with the same number of sulfate groups. No significant CSP can be observed suggesting low binding to this oligosaccharide. Addition of one 6-O sulfate group on the reducing end glucosamine of $dp4^3$, leading to $dp4^4$, significantly increases CSP and thus binding around the additional 6-O sulfate, especially for CXCL12γ (Figure 7). While this 6-O sulfate group at position 4 has a clear effect on protein binding, addition of a 6-O sulfate to $dp4^1$, leading to $dp4^2$, did not change CSP compared to $dp4^1$ for any protein, suggesting that this sulfate group is not important for binding, compared to IdoA2S group.

Using only 4 different tetrasaccharides it was possible to observe significant differences in the binding of the proteins with respect to overall sulfate numbers and to additions of 6-O sulfates at several positions, and to deduce the relative importance of several sulfate groups for binding. Tetrasaccharides have a low affinity for proteins and are shorter that the optimal size for binding for

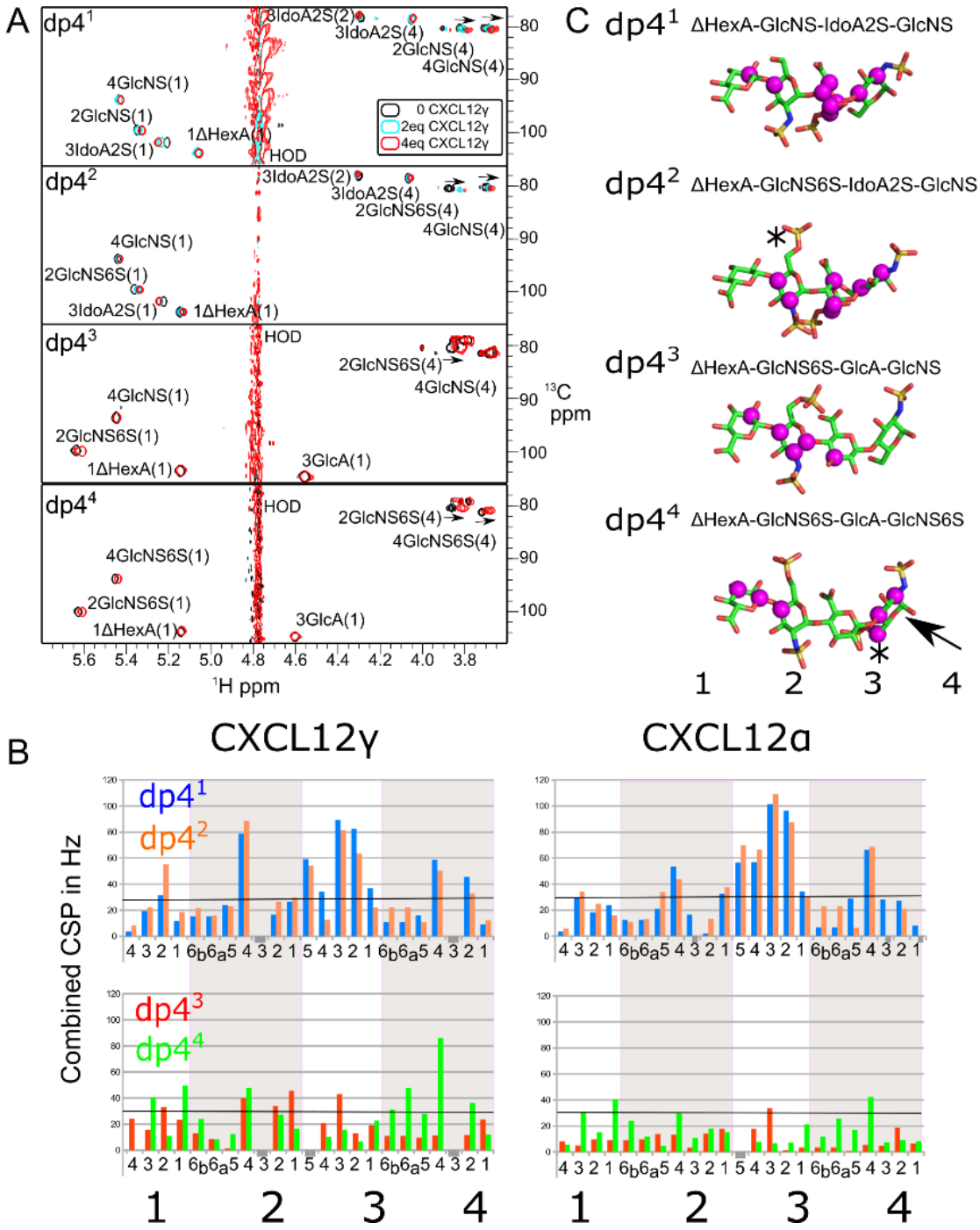


Figure 7 : Interaction of 4 different tetrasaccharides with CXCL12γ.

a) NMR spectra of dp4$^{1-4}$ upon CXCL12γ addition. b) Significant chemical shift perturbation visualized on the oligosaccharides structures in magenta and in c) as histogram

proteins (dp6-dp8 size). We thus investigated binding to different hexasaccharides.

2. *Interaction with IdoA2S containing hexasaccharide and docking with the core region of CXCL12α and CXCL12γ.*

The simplest IdoA2S containing dp6 (dp6^5), with a ΔHexA-(GlcNS-IdoA2S)$_2$-GlcNS structure was tested in interaction with the chemokines (Figure 8). The hexasaccharides showed CSP mostly

on the 2 IdoA2S sugars. It must be noted that increase in the size of the oligosaccharide generates signal overlap, mostly in Glucosamines for which signals can be only partly followed by NMR. CXCL12 α and γ, cause particularly perturbations in the reducing end IdoA2S. On the other hand ^{15}N-^{1}H NMR of the 2 proteins showed significant CSPs upon dp6[5] interaction located around residues K24 and R41 with additional CSPs in the basic C-terminus of CXCL12γ.

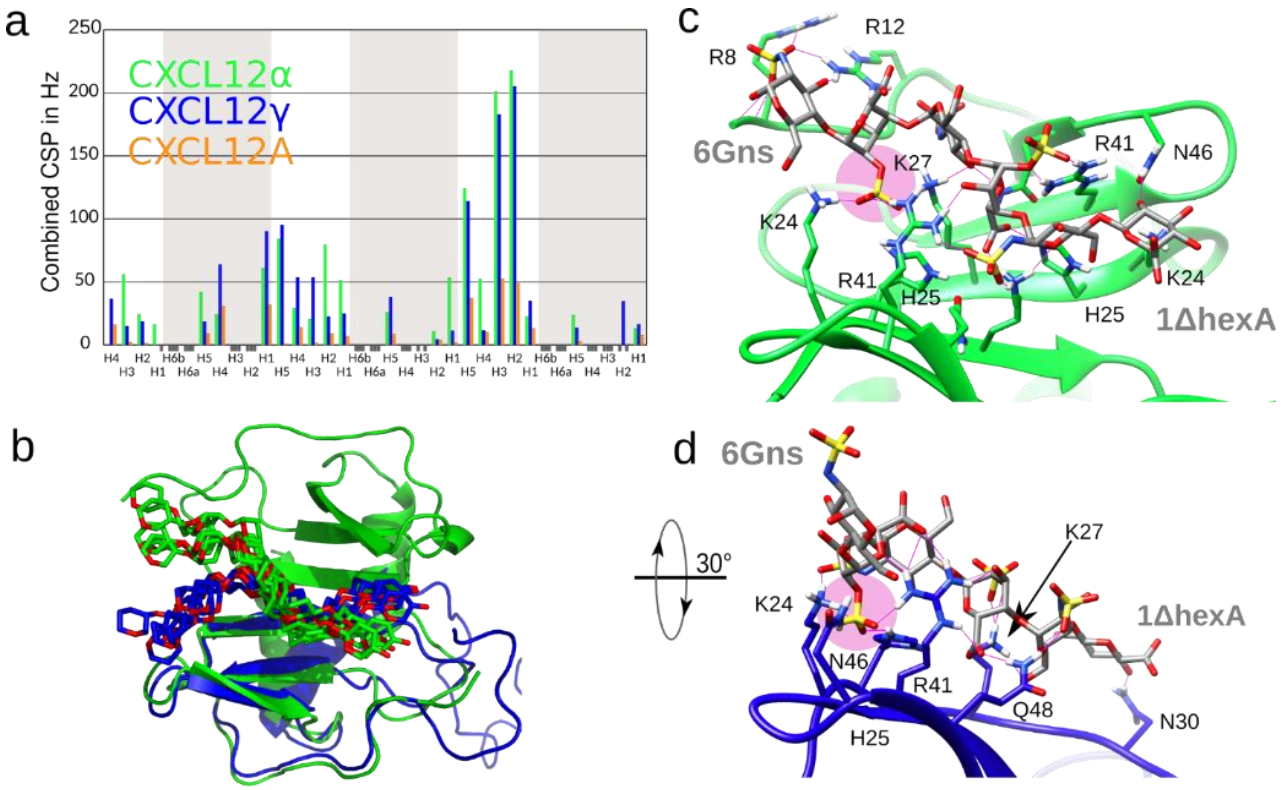


Figure 8 Interaction of the three proteins with ΔHexA-(GlcNS-IdoA2S)$_2$-GlcNS hexasaccharide a) CSP induced by addition of CXCL12α (green), CXCL12γ(blue) and drCXCL12A(homolog from zebrafish with weak binding orange). Grey bars represent groups not followed because of overlaps. (ΔHex= unsaturated hexuronic acid).b) 4 best models of dp6[5] docked onto overlayed CXCL12α (green) and CXCL12γ(blue). Only the carbons and oxygens of the pyranose ring are shown for clarity. c) and d) CXCL12α and CXCL12γ in the same orientation showing intermolecular interactions with reducing end IdoA2S(magenta).

Docking of CXCL12α and CXCL12γ with dp6[5] was performed with HADDOCK[14] in order to examine the role of the reducing end IdoA2S. To this purpose the 3D structure of CXCL12γ was solved by NMR (PDB: 6EHZ). Haddock procedures combined use of residues on both protein (^{15}N-^{1}H) and ligand (^{13}C-^{1}H) that are perturbed by the interaction as input and docks the molecules by creating ambiguous interaction restraints between these residues to form the complex. Complexes obtained are then refined in explicit solvent and sorted in clusters.

The best models are shown on Figure 8. The orientation of the dp6 with respect to the protein is slightly different due to the monomeric and dimeric natures of CXCL12α and CXCL12γ respectively that create different binding surfaces (Figure 8b). As expected from the CSP observed

on the proteins as well as previous NMR and mutational studies[4,11,13], the same residues of both proteins are involved in HS binding, primarily on β1 strand. The reducing end IdoA2S, that shows high CSP in dp6[5], is located at the same position in CXCL12α and CXCL12γ, in contact with K24, R41 and K27 (from the other monomer for CXCL12α). The NMR data, supported by the models of the two complexes, clearly underline the importance of the reducing end IdoA2S that contacts residues K24 and K27 which are essential for HS binding[13].

3. *Effect of epimerization in the interaction with CXCL12α*

One of the hexasaccharide produced (dp6[2] figure6B) possesses a GlcA2S instead of a IdoA2S in position 3. It offers a unique opportunity to observe the influence of the epimerization state at that position on CXCL12α binding. CSPs pattern upon CXCL12α addition is similar to dp6[5] with additional perturbation on the GlcNS preceding the GlcA2S. Docking of dp6[2] to CXCL12α following the same haddock protocol allows comparison of the two complexes. Models containing dp6[2] present more extensive intermolecular contacts than with dp6[5], better physical energy and convergence of solutions, indicating an overall better model than with dp6[5]. The 180° flip of the 2-O sulfate position in GlcA2S versus IdoA2S (Figure 9), due to the absence of epimerization, positions the GlcA2S and the consecutive GlcNS so that they both insert their sulfate groups into the groove formed at the

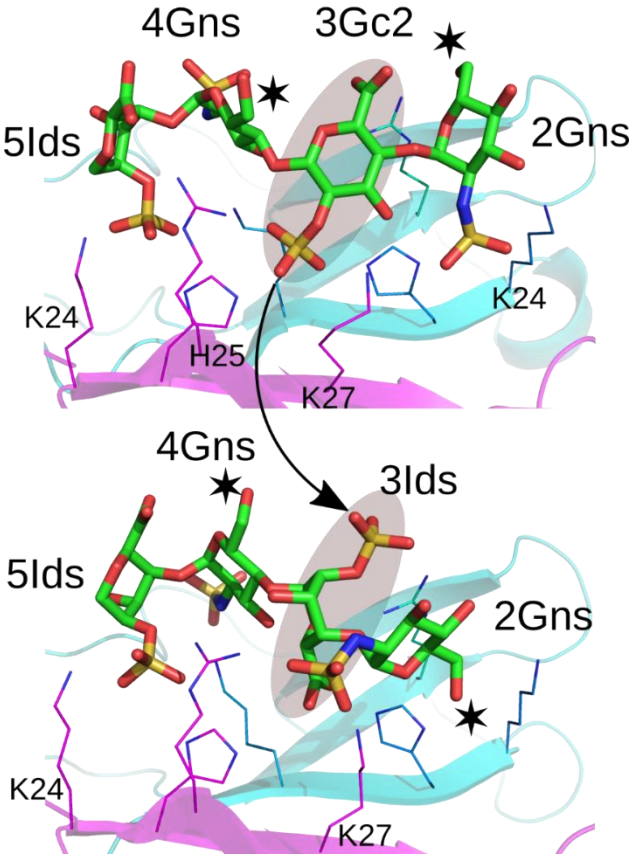


Figure 9 Effect of epimerization on CXCL12α binding.

Close up on the best models showing only sugars 2-5 for clarity. GlcA2S in dp6[2] (top) allows insertion of 3 sulfate groups into the basic groove, compared to dp6[5] (bottom). Stars represent the position of putative 6-O sulfation of glucosamines.

CXCL12α dimer interface (Figure 9). The reducing end IdoA2S (position 5) still interacts with K24, K27 and R41 and GlcNS and GlcA2S at position 2 and 3 interact with the same residues on the symmetrical K24, K27 and R41 site. The ability of this oligosaccharide to extensively binds the two K24, K27, R41 sites formed by CXCL12α dimerization probably explains it forms a better complex than dp6[5].

The increased availability of libraries of Heparan sulfate oligosaccharides is key to the understanding of protein-HS recognition. Here we could investigate at the molecular level the influence of the position of a sulfate group or epimerization in binding, thanks to ^{13}C labelled oligosaccharides of different modification patterns. NMR and docking show that CXCL12 core domain, either in CXCL12α or CXCL12γ, recognizes primarily an IdoA2S residue on the reducing end that establishes contacts with residues important for HS binding (K24, R41). Addition of a 6-sulfate group in the preceding glucosamine residue does not increase binding as it would point to the solvent in the models. While some combinations of sulfation could not be produced, the model with dp6[5] allows to predict that addition of a 6-O sulfation at GlcNS in position 2 would point to the second K24, K27, R41 site and could strengthen the binding (Figure 9). Change of the IdoA2S to GlcA2S at position 3 completely inverts the conformation of the oligosaccharide, maintains interaction of IdoA2S at position 5 with the first K24, K27, R41 and position of the 2-sulfate group and enables extensive interactions with the second K24, K27, R41 site forming a better interaction model for the same number of sulfate groups than dp6[5].

More generally while this methodology could be applied to many functionally important HS binding proteins (cytokines, growth factors), our results concur with the idea that the three-dimensional arrangement of the oligosaccharide[15] and hence the way it presents it sulfate groups to a given protein motif (flat motif, grooved motif) is key to the recognition and that the most important sulfate groups are masked in highly sulfated molecules like heparin [16].

II. Structural Study of LipoPolySaccharides (LPS)

The cytoplasm of Gram-negative bacteria is typically surrounded by two membranes with different lipid composition that are separated by an aqueous compartment called the periplasm. While the inner membrane (IM) is almost exclusively made of phospholipids, the outer membrane (OM) is highly asymmetric and contains phospholipids in its inner leaflet and an unusual glycolipid, lipopolysaccharide (LPS), in its outer leaflet. The LPS molecule is made of three covalently linked moieties: the lipid A, which is the hydrophobic anchor in the membrane, a core oligosaccharide, and a variable O-antigen sugar chain (Figure 10). The presence of the tightly packed LPS layer at the outer leaflet makes the OM fairly impermeable and protects Gram-negative bacteria from harmful compounds such as detergents and lipophilic antibiotics.

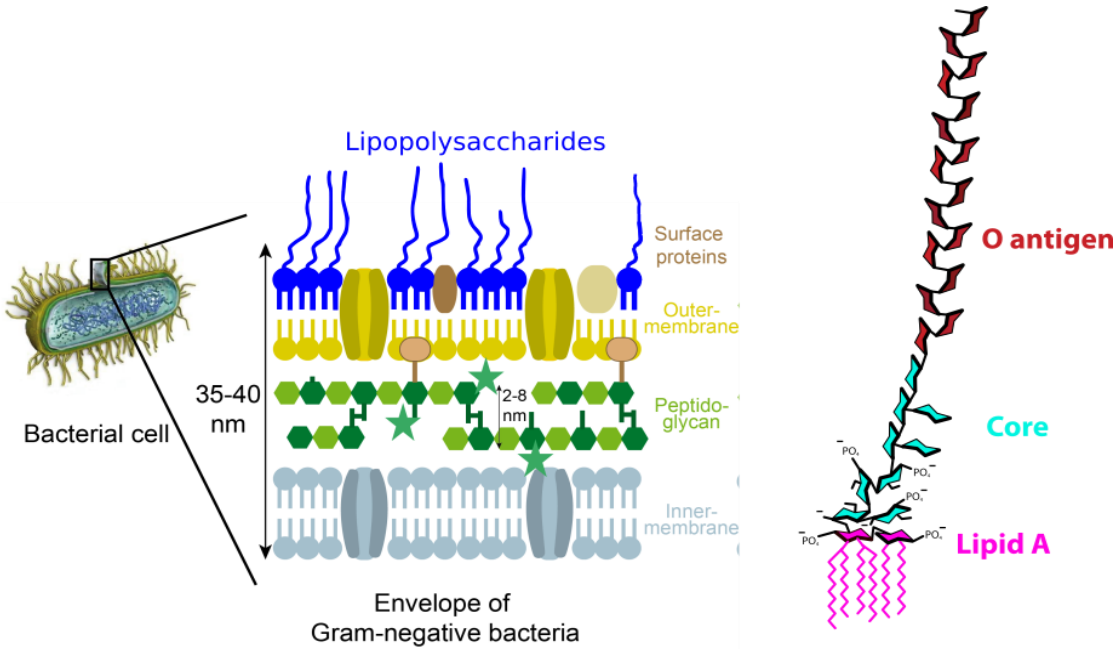


Figure 10 : Organisation of Gram negative bacteria envelope. Lipopolysaccharides are the major component of outer-membrane external leaflet of gram-negative bacteria (left). LPS are composed of three parts, the lipid A, core oligosaccharides and O-antigens(right).

A. Structural study of LPS by Magic Angle Spinning Solid state NMR (ssNMR)

LPS molecules are amphiphilic and inherently heterogeneous; they differ between species, especially in the O-antigen region, but also depending on the bacteria growth conditions. Their heterogeneity and low solubility in water complicates their structural investigation. Their structural analysis is performed after chemical treatment to study separately lipidic and hydrophilic parts of the

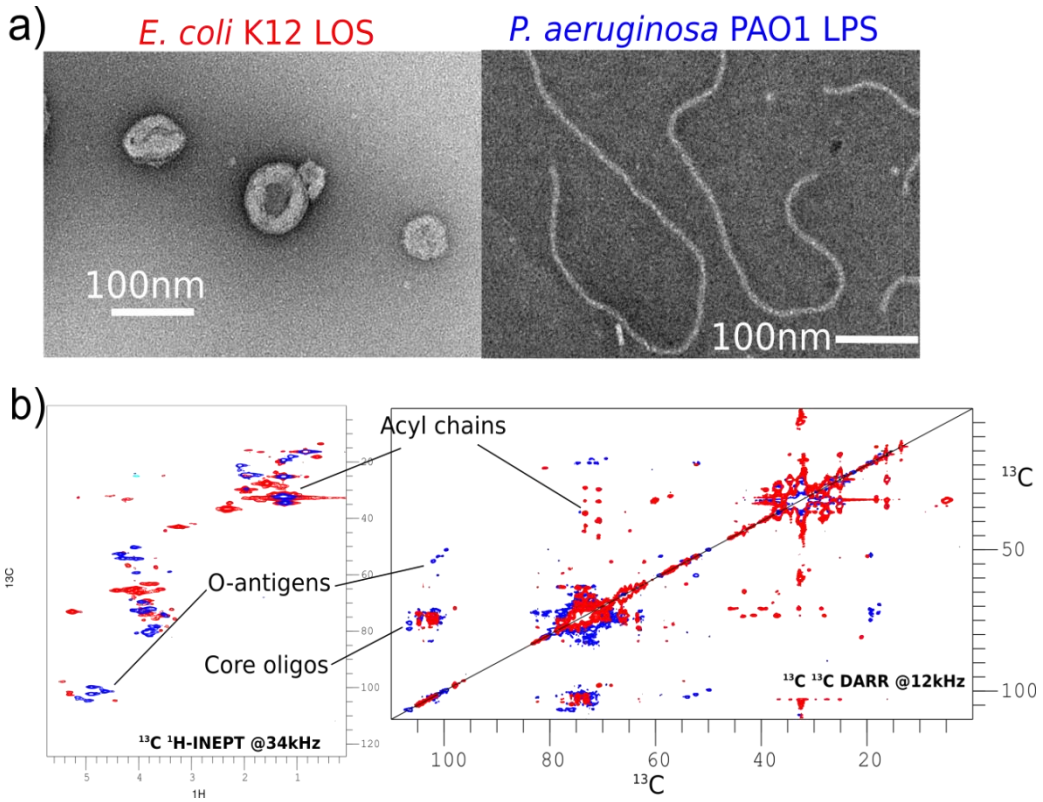


Figure 11: LPS from E. coli and P. aeruginosa EM and ssNMR comparison a) Negative stain Electron micrographs (23000x magnification) of E. coli LOS (left) that forms mainly large vesicles in solution and P. aeruginosa LPS (right) which forms structures corresponding to elongated micelles. b) Overlay of E. coli (red) and P. aeruginosa (blue) ssNMR spectra

molecule by combinations of mass spectrometry and NMR [17]. In order to study the intact LPS we used solid state NMR with Magic Angle Spinning[18]. Purified ^{13}C labelled LPS from *E. coli* lacking O-antigen (LipoOligoSoccharide LOS) and from pathogenic *P. aeruginosa* (with O-antigen) were prepared by the laboratory of A. Molinaro in Napoli University. In solution these LPS form respectively large vesicles and elongated micelles (Figure 11), and ^{13}C-^{13}C and ^{13}C-^{1}H correlation experiments could be recorded. They allowed assignment of most resonances and thus observation of the different parts of the molecules, O-antigens, core oligosaccharide and Lipid A.

Assignment of LPS resonances with ssNMR provides interesting probes to observe interactions with molecules, in particular antibiotics. Gentamicin is an antibiotic from the aminoglycoside family and was shown to strongly associate with *P. aeruginosa* LPS[19]. Interaction of gentamicin was thus tested on our preparation of LPS. While gentamicin does not affect the macromolecular organization of LPS in solution (Figure 12 a), NMR signals specific to the O-antigen sugars are affected by the presence of the antibiotic. Strong signal decrease and significant frequency changes are observed in the trisaccharide repeat of the O-antigen (Figure 12 b). We could nevertheless not pinpoint the sugars

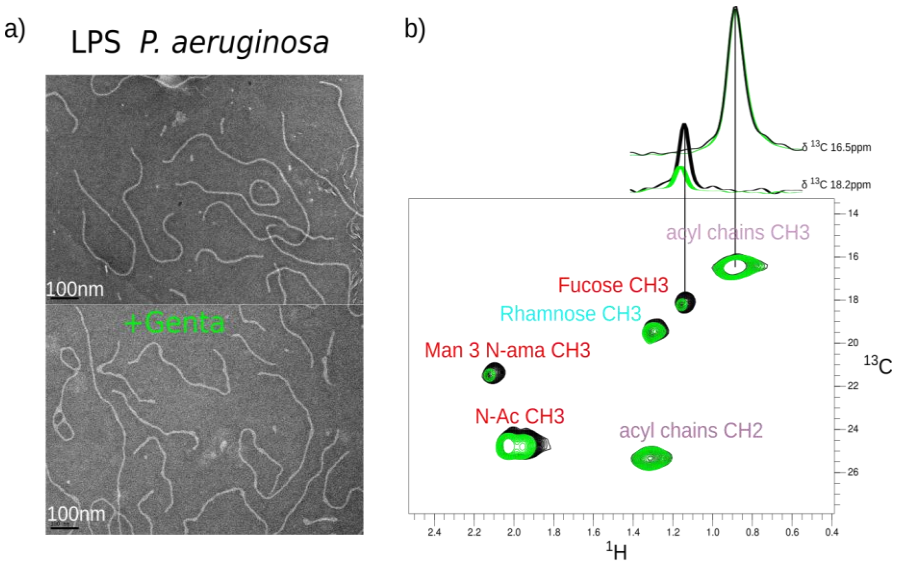


Figure 12 : P. aeruginosa LPS O-antigen is affected by interaction with gentamicin. a) EM pictures with negative staining of P. aeruginosa LPS in absence (up) or presence of 25mM Gentamicin(down) b) ^{1}H-^{13}C INEPT ssNMR spectrum of LPS in absence or presence (green) of 25mM gentamicin. Spectra are normalised to the intensity of lipid A CH3 peak. 1D slices of Lipid A CH_3 and fucose CH_3 peaks are shown above the spectrum.

important for binding as all O-antigen is perturbed by the interaction. The final target of gentamicin is the 30S ribosome subunit but binding to the O-antigen might increase the local concentration of the antibiotic near the membrane and facilitate intake by bacteria.

Solid-state NMR is an especially useful method to examine large systems such as proteins in either crystalline state, amyloid fibres, or more generally multimeric assemblies, membranes and

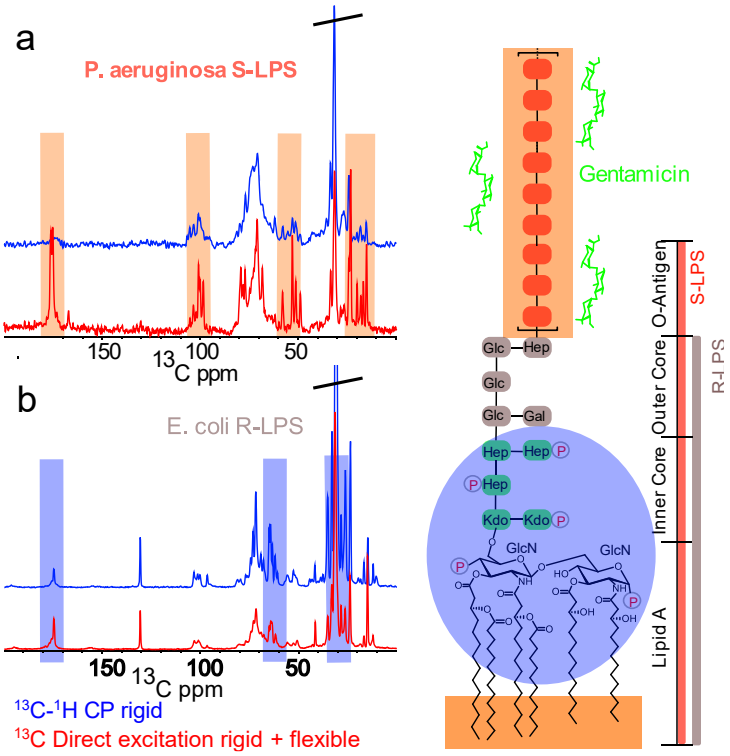


Figure 13 ssNMR-highlighted regions of different flexibilities and regions of interactions in LPS. (a) Comparison of 1D 13C spectra of LPS with direct 13C excitation (red) or through 13C−1H cross-polarization experiment (blue) shows that P. aeruginosa S-LPS, which assembles into micelles, displays mainly flexible signals (orange) belonging to O-antigens with less efficient CP transfer. These regions were identified as interacting with gentamicin antibiotic. (b) E. coli R. LPS in vesicles CP spectrum shows rigid portions of the LPS (blue), mainly core oligosaccharides and lipids.

bacterial cell-walls. For complex glycoconjugates like peptidoglycan[20,21] and LPS we can exploit the capacity of ssNMR to record experiments through J-couplings, sensitive to flexible components and through dipolar couplings, via Cross Polarisation (CP) schemes, for rigid parts of the molecule. Examples of the observables (rigid, flexible) on LOS/LPS are shown in figure 13. On LOS, which assemble into vesicles we can observe lipid A and core oligosaccharides, mostly from CP experiments. Those parts of the molecules are not visible in LPS where mostly dynamic components (extremity of lipid chains and O-antigens) are observed. The same behaviour was recently seen in our lab on another *E. coli* LPS (O157:H7), also organised in solution as elongated micelles. We still do not know if the differences we observe are due to a lack of sensitivity (O-antigen in LPS constitutes the most abundant component in LPS) or to the differences in macromolecular organisation (micelles vs vesicles). Furthermore, investigation of *E. coli* LOS vesicles from other strains by ssNMR, including at high MAS speed (>50kHz) with ^{1}H detection, produced low quality spectra. Purified LPS/LOS reconstituted in solution are thus to this point difficult objects to investigate by ssNMR and generalisation of this method to LPS study will require developments at the sample preparation level.

B. The LPS transport system in E. coli

In *Escherichia coli* the LPS molecules are synthesized in the cytoplasm and at the Inner Membrane, flipped across the IM by the dedicated ABC transporter MsbA and then exported to the OM by the Lpt pathway (Figure 14). This machinery, discovered by our collaborator A. Polissi in Milan University, is made up of seven proteins (LptA-G) that physically interact and assemble in a complex spanning the entire envelope. At the IM the $LptB_2FG$ complex, associated to LptC, utilizes

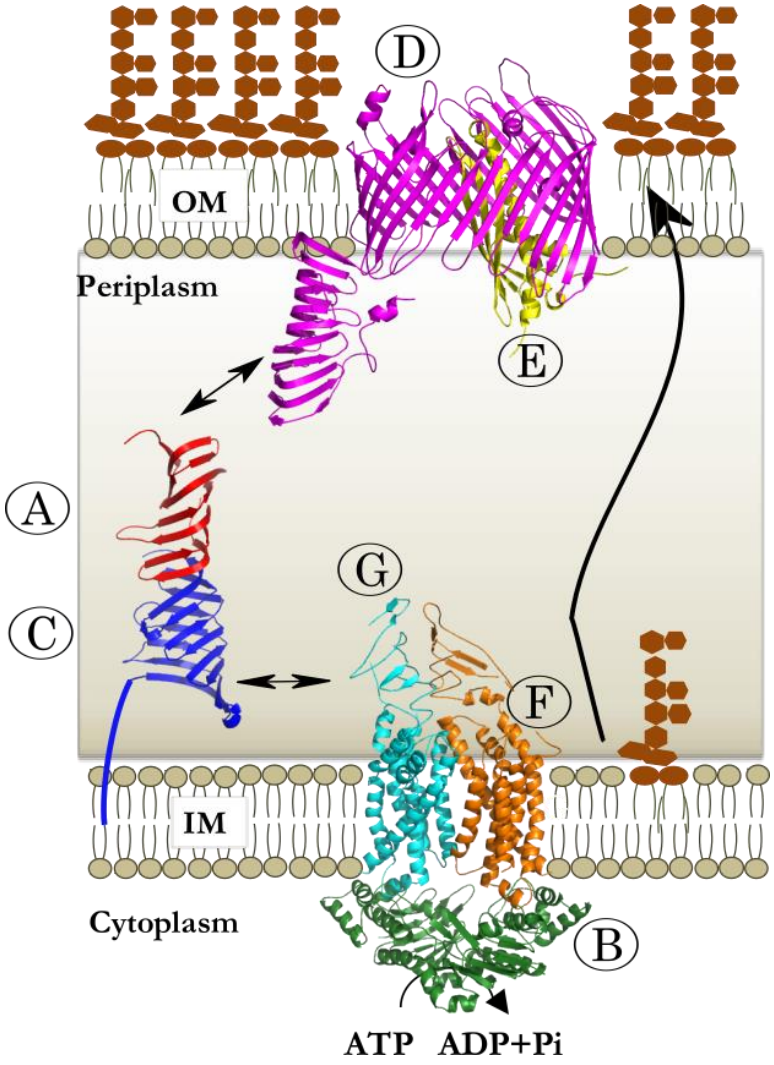


Figure 14 Export of LPS to the cell surface. The Lpt molecular machinery (LptA-G) transports LPS across the periplasm and inserts it into the outer membrane (OM) at the expense of ATP hydrolysis.

ATP hydrolysis to promote the transport of LPS across the periplasm. At the OM the LptD/E complex deliver the LPS to the cell surface. LptA is a key protein that connects the $LptB_2CFG$ complex to LptD/E by interacting with LptC and the periplasmic region of LptD via its N- and C-terminal ends, respectively. The crystal structures of all (LptA, $LptB_2FG$, LptC, and LptD/E) Lpt proteins are known. All periplasmic domains contain a similar β-jellyroll fold that is predicted to accommodate the lipid chains of the LPS. Despite the availability of a wealth of structural and functional data many

important details are still missing for both the architecture of the Lpt machinery and the mechanism of LPS transport.

1. *Structural characterization of LptC-LptA-LPS complex*

Periplasmic LptC and LptA proteins of the LPS transport system (Lpt) are responsible for LPS transfer between the Lpt inner and outer membrane complexes. We studied the interaction of these two proteins as well as with LPS to get insights into the periplasmic LPS transfer mechanism. LptC is composed of an N-terminal transmembrane helix and a β-jellyroll domain. The construct used lacked the transmembrane domain (ΔTM-LptC). Because of the tendency of LptA to oligomerise we engineered a monomeric *E. coli* LptA mutant ($LptA_m$) and showed *in vivo* that a stable LptA

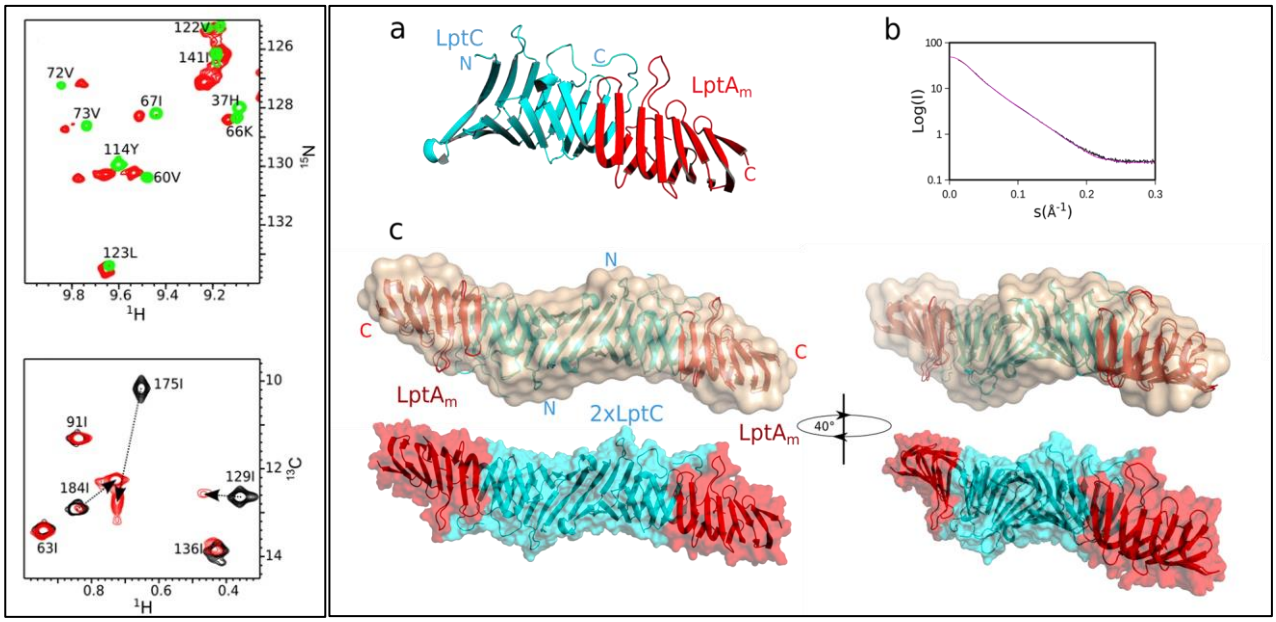


Figure 15 Combined NMR-SAXS model of the LptC-LptAm complex. Left NMR CSP on $LptA_m$(top) and LptC (bottom) upon complex formation (red). Right combined model of the complex (A) Representation of LptC-LptAm half complex. (B) Overlay of the backcalculated SAXS curve from the best energy NMR-SAXS model (magenta) with respect to the raw data (black). The chi-squared value is χ2 = 9.25. (C) Ribbon representation of the best NMR-SAXS LptC-$LptA_m$ complex fitted into the ab-inito envelope of the complex (wheat). In the lower panel the best NMR-SAXS structure is represented in cyan-red as a surface for comparison with the upper-panel SAXS envelope. LptC and LptAm molecules are shown in cyan and in red, respectively.

oligomeric form is not strictly essential for bacteria[22]. To investigate the molecular structure of the LptC-$LptA_m$ complex, NMR experiments were performed using [^{2}H,^{13}C,^{15}N]- labelled samples or [^{2}H,^{12}C,^{15}N]-samples specifically ^{13}C-labeled and protonated on $A^{\beta}I^{\delta1}L^{\delta1}V^{\gamma1}$ methyl groups. An LptC-$LptA_m$ low-resolution complex was characterized by a combination of SAXS (SAXS envelope) and NMR (chemical shift perturbations on LptC and $LptA_m$) using Haddock (Figure 15).

Focus was then shifted to the interaction with LPS. An LPS from *E. coli* pathogenic bacterium

O157:H7[23] was tested for interaction as its long O-antigen increases its solubility. Chemical shift perturbations could be observed on LptC amide and methyl groups (Figure 16) as well as appearance of new methyl resonances assigned to the end of the LPS lipid chain, in the ^{13}C-labeled LPS. It must be noted that titration experiments suggest only a very partial saturation of LptC by LPS, probably because of the low availability of monomeric LPS in solution. A LptC-LPS complex was built based

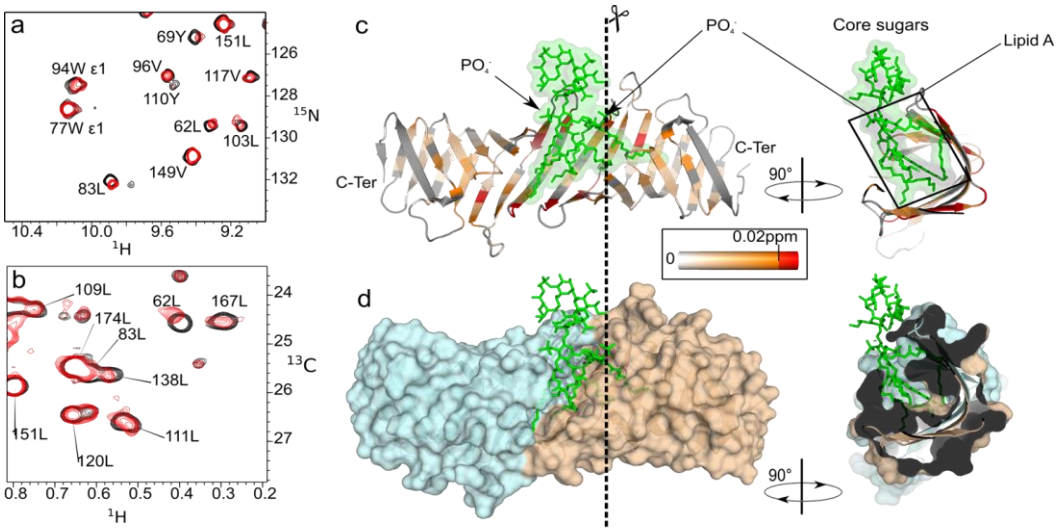


Figure 16 (A) [^{1}H,^{15}N]-correlation spectrum of [^{2}H ,^{15}N]-labelled LptC in presence (red) and absence (black) of 0.8 mg/ml of LPS. (B) [^{1}H,^{13}C]-correlation spectrum of [^{2}H,^{15}N, ^{1}H/^{13}C-($A^{\beta}I^{\delta 1}L^{\delta 1}V^{\gamma 1}$)]-labelled LptC in presence (red) and absence (black) of 0.8 mg/ml of LPS. (C) Combined ^{1}H and ^{15}N CSP observed on LptC upon LPS binding are displayed on a ribbon representation of the best HADDOCK LptC-LPS model calculated from these CSP with (D) Surface representation.

on CSP NMR data in which the lipid moiety of the LPS is buried at the interface of the two β-jellyrolls of the LptC dimer (Figure 16). We were also able to observe interaction of LPS with $LptA_m$ as well as with the LptC-$LptA_m$ complex.

Altogether the determination of a LptC-LptA model, LPS binding experiments and analysis of structures suggest that the three intermolecular cavities LptC-LptC, LptC-LptA and LptA-LptA are important sites for LPS binding. These sites, present only when specific complexes between β-jellyroll proteins are assembled could be essential to understand the mechanism of the LPS flow through the Lpt system[22].

2. *LptB$_2$FG has ATPase and Adenylate kinase activity, both regulated by LptB$_2$FGCA complex assembly.*

The transfer of LPS from the outer leaflet of the inner membrane of *E. coli* to the periplasmic components LptC and LptA is still poorly understood. While two crystal structures of the $LptB_2FG$

complex have been determined, structural information on LPS binding/transfer is still lacking. In collaboration with A. Polissi in Milan University and in the framework of an ITN European project, we have explored the activity and interactions of the $LptB_2FG$ ABC transporter. One of our goal in this study was to use Methyl-labelled LptC (described above) to monitor its interaction with $LptB_2FG$ complex and LPS transfer. $LptB_2FG$ complex could be expressed and purified in DDM micelles[24]. DDM-$LptB_2FG$ complex was pure and functional with respect to interaction with LptC and ATPase activity but highly unstable above 20°C, which made it incompatible with the study of large complexes by NMR.

Styrene-maleic acids polymers (SMA) are increasingly used to form nanodiscs of membrane proteins, either from detergent purified proteins or by extracting directly proteins from the natural membrane[25,26]. The latter method allows to extract the protein with its native lipid surrounding and generates nanodiscs of about 10nm. In the meantime, several structures of $LptB_2FG$ in complex with LptC were determined by X-ray crystallography and Cryo-Electron microscopy[27–29] so we turned our attention on the characterisation of $LptB_2FG$ in SMA-nanodiscs.

SMA-$LptB_2FG$ was prepared from *E. coli* inner membranes and showed good stability for several hours at 37°C and ATPase activity (Figure 17). SMA-$LptB_2FGCA$ complex (with ΔTM-LptC and $LptA_m$ cf 1.) ATPase activity was assessed in real-time by ^1H NMR (Figure 17). It showed, as expected fast ATP consumption and ADP release. Nevertheless, initial ADP build-up was followed

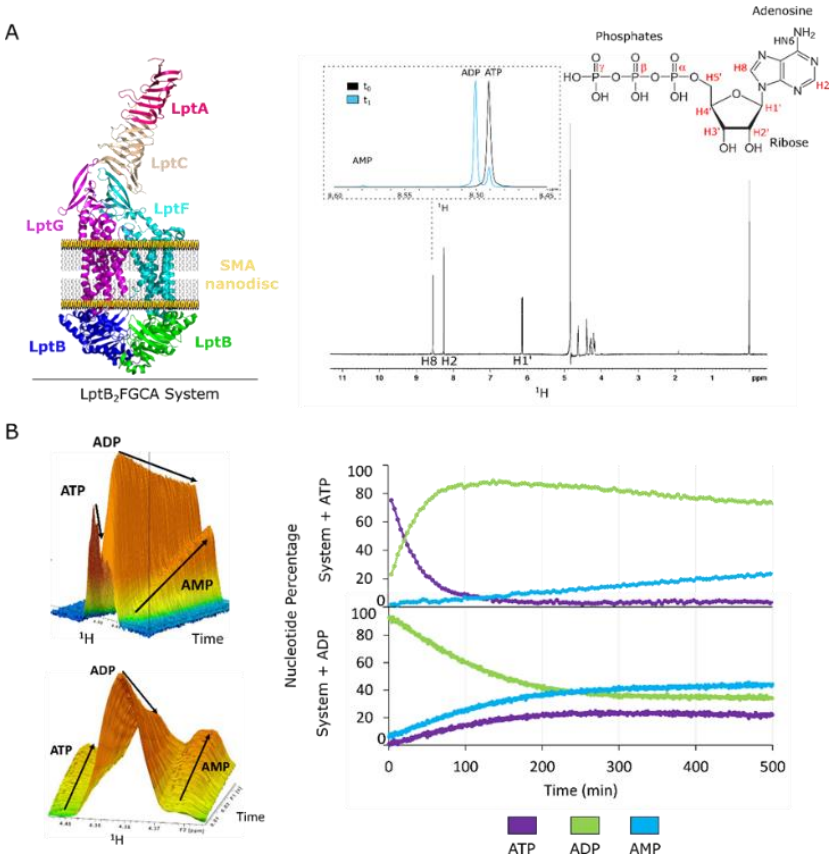


Figure 17 : (A) Representation of $LptB_2FGCA$ (using ΔTM-LptC and $LptA_m$) system extracted in SMA nanodiscs. LptB activity was followed by observing variations in nucleotide frequencies, here ATP is represented with H8 proton chemical shifts of the different nucleotides observed; (B) Real-time kinetics of $LptB_2FGCA$ observed by ^1H-NMR (left) with either ATP (top) or ADP (bottom) as substrate over time with evolution of the characteristic peaks of ATP, ADP and AMP (H4'). Nucleotide quantification along the experiment is represented (right).

with a decrease and concomitant appearance of a new adenosine species, identified by its ^{1}H and ^{31}P chemical shifts as AMP (Adenosine Mono-Phosphate). Appearance of AMP is consistent with an Adenylate Kinase (AK) activity of $LptB_2FG$, 2ADP ⇔ AMP + ATP. This activity has been already described for several ABC-transporters ATPase domains including MsbA, the inner membrane complex that flips LPS to the outer layer of the inner-membrane leaflet[30]. When incubated with ADP as substrate SMA-$LptB_2FGCA$ generates ATP and AMP, confirming the AK activity (Figure 17 B).

SMA-$LptB_2FG$ complexes alone have ATPase activity and little AK activity (Figure 18). Those two activities are upregulated by the establishment of complexes on the periplasmic side with ΔTM-LptC and $LptA_m$. Full length LptC with both membrane and periplasmic domains downregulates ATPase activity as already reported [29] but the inhibition is partly relieved by $LptA_m$ binding on the periplasmic side (Figure 18B).

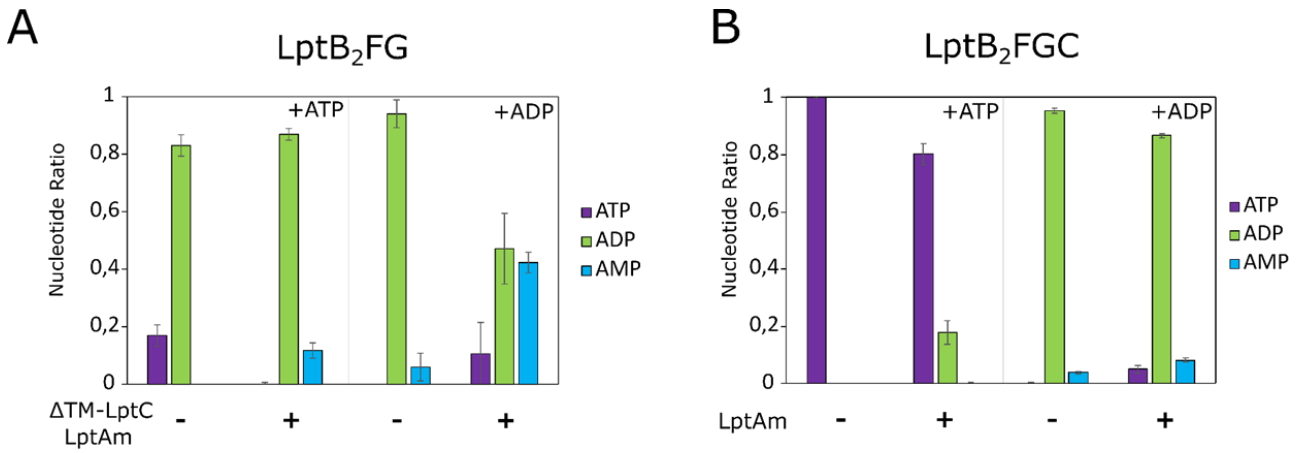


Figure 18 Effect of ΔTM-LptC, LptC and LptAm ATPase and AK activity (A) SMA-$LptB_2FG$ wt without/with ΔTM-LptC/LptAm, (B) with full length LptC, both in presence of either ATP or ADP. Nucleotide levels (ATP, ADP, and AMP, in colour-code) were detected using a 1D ^{1}H-NMR experiment in 2 independent experiments.

$LptB_2FG$ ATPase and AK activity are both regulated by the establishment of complexes, and more specifically they are activated by interacting with LptC and LptA domains in the periplasm. The mechanisms that propagate the binding information on the periplasmic side to the LptB subunit on the cytoplasmic side, that carries both ATPase and AK activities, through the membrane are unknown. Those mechanisms are probably important to ensure no LPS is not transported unless the Lpt transport machinery is assembled, to avoid leaking of LPS in the periplasm.

To probe the location of the AK site, several mutations were introduced in LptB (expressed alone) around the conserved ATPase site (Figure 19 A). Isolated LptB mutant proteins showed decreased ATPase activity and little AK activity (Figure 19). Two mutations E163Q and Y13W show a moderate and high AK activity increase, respectively. STD experiments on wt and mutant LptB were recorded with a non-substrate nucleotide ADPβS. STD is an NMR technique that allows observation of the free ligand in transient binding to a protein[31]. The protein is selectively saturated, and part of the saturation is transferred to the ligand during binding and a decrease of intensity is observed on the free ligand. STD on wild-type LptB shows significant saturation on protons H2, H8 and H1' of the nucleotide (Figure 19). Y13W shows significantly reduced STD that can be ascribed to a decrease in affinity for the ligand. This is consistent with the influence of this mutant in other ABC-transporters [32,33]. This Y13W mutation confers LptB similar AK activity as full $LptB_2FGCA$(with ΔTM-LptC/LptAm) and we can hypothesize that the full complex might have reduced affinity for ADP. Unfortunately, the Y13W mutant could not be purified as a complex with LptF and LptG in SMA.

The role of the AK activity in $LptB_2FG$ ABC transporter (and the related MsbA) remains

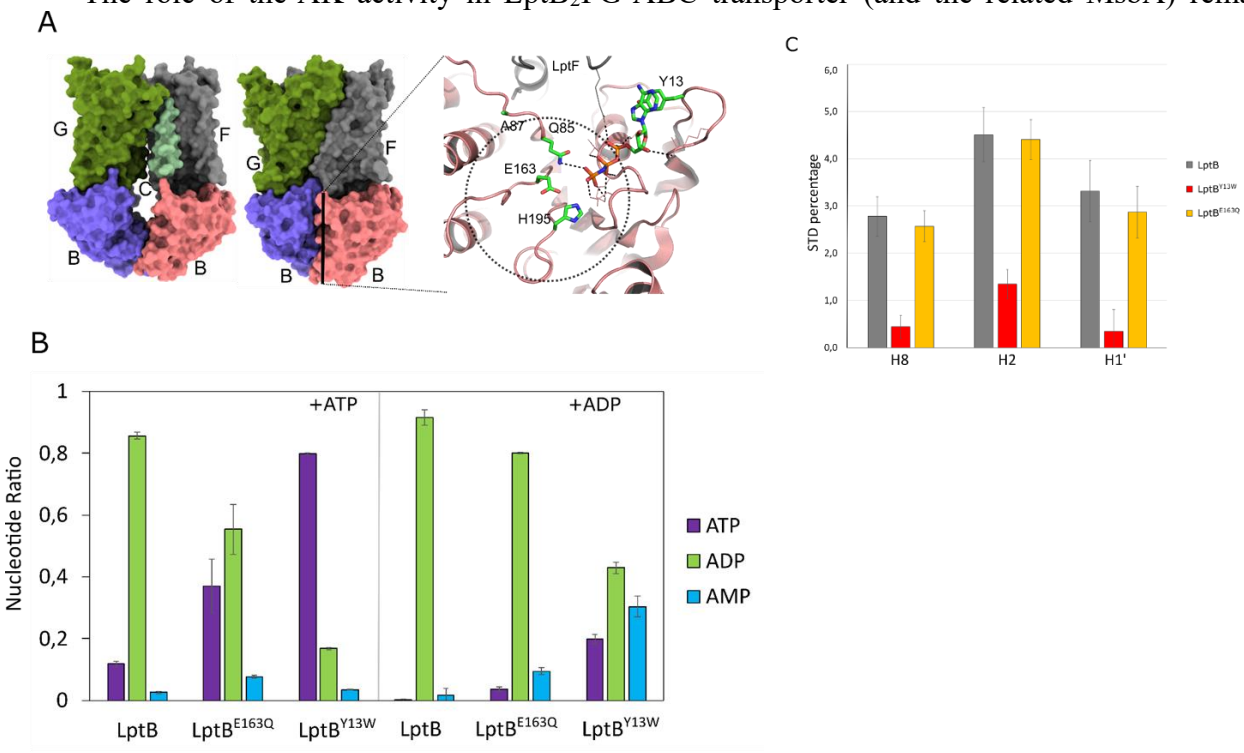


Figure 19 (A) Structures of $LptB_2FGC$ in an open and closed conformation (with bound ATP analog), respectively Protein Data Bank (PDB) codes 6S8N (left structure) and 6S8G (central structure), the jellyroll domains were not resolved in those cryoEM structures. Right panel shows ATPase site with putative location of the second ADP in the AK reaction (dashed line area). Vicinity residues mutated in this study are shown (B) ATPase/AK of $LptB_2$ and $LptB_2^{E163Q}$ and $LptB_2^{Y13W}$ mutants, around canonical ABC motifs, in presence of either ATP or ADP, at 20°C. Nucleotide levels (ATP, ADP and AMP, in colour-code) were detected using a 1D ^{1}H-NMR experiment in 3 mm tubes at 20°C, extracting peak intensities for each specie in 2 independent experiments; (C) Saturation Transfer Different (STD) experiments using ADP-βS. The three resonances shown are from H8, H2 and H1' on the adenosine of ADP-βS. Data are represented as (Iref-Isat/Iref). Error bars represent two times the standard deviation of the NMR experimental noise.

unknown. Mutants with altered ATPase activity but AK activity like E163Q cannot sustain *E. coli*

growth[34], AK activity is thus not sufficient for LptB activity. Nevertheless, AK activity can generate ATP which can in turn be used for LPS transport. The relationships between AK, ATPase and LPS transport will have to be examined to understand the exact role of AK in $LptB_2FG$. The AK activity, which is regulated by assembly of the $LptB_2FGCA$ bridge, must be considered to review this activity in relation to the abundant structural, biochemical, and cellular studies already published on LPS transport system. This work was performed by a PhD student I cosupervised, T. Baeta and a manuscript describing these results will be submitted shortly.

C. Molecular characterization of bacterial LPS recognition by C-type lectin receptors

LPS structural variability and tight assembly protect bacteria against host defences and uptake of antimicrobials making gram-negative bacteria a major challenge for human health (B. cenocepacia, P. aeruginosa., K. pneumoniae). Both lipidic and glycans parts of LPS constitute Pathogen Associated Molecular Patterns (PAMPs) detected by the immune system. Among Pathogen Recognition Receptor (PRR), MD2/TLR4 recognize mostly the lipid moieties of LPS, while C-type Lectins Receptors (CLRs), at the surface of dendritic cells, interact with the glycan part of LPS.

The interaction of these CLRs with their ligands, discriminating non-self and self-molecular

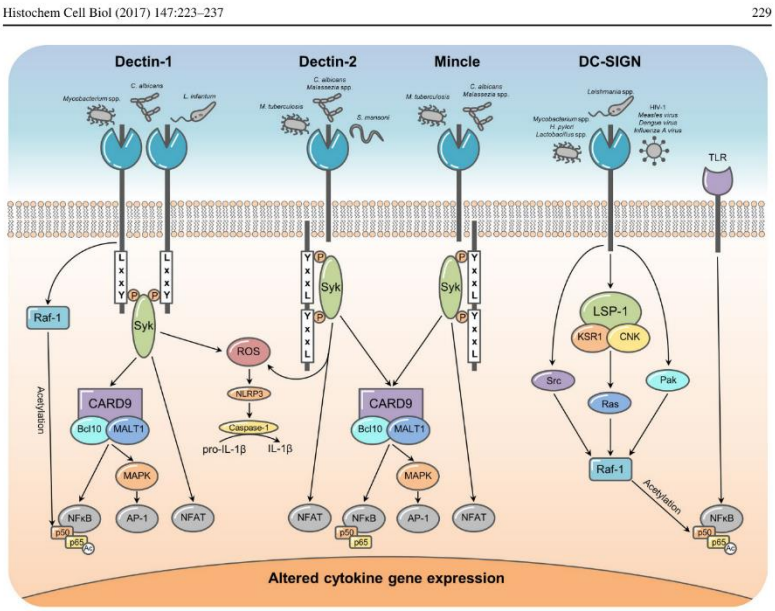


Figure 20 Pathogen recognition by myeloid C-type lectins. Myeloid CLR recognize pathogens, such as bacteria, parasites, viruses, and fungi. Recognition triggers several signalling pathways and modulate immune response. from (Mayer et al . 2017)

motifs, allows Dendritic cells to modulate the immune response towards either activation or tolerance (Figure 20)[35]. Oligomerisation and clustering of CLRs enable high affinity interaction with glycans

at the surface of pathogens. The multivalency generated between cell surface exposed CLRs and LPS is key to ensure strong specific interactions and thus pathogen recognition. The CLRs are known to recognise mostly, but not exclusively the core oligosaccharide of the LPS that present only subtle structural variations between bacteria[36,37]. Understanding how those glycan structures impact bacteria recognition by the CLRs is a key step towards deciphering the underlying mechanism of defence against pathogenic bacteria while commensal bacteria are being tolerated and should not elicit inflammatory response. This project was conducted in the framework of a shared PhD studentship (M. Maalej) funded by glyco@alps and Napoli University in collaboration with A. Molinaro in Napoli and F. Fieschi group at IBS, Grenoble.

We focused our LPS/CLR interactions on MGL (Macrophage Galactose-type Lectin) is a C-

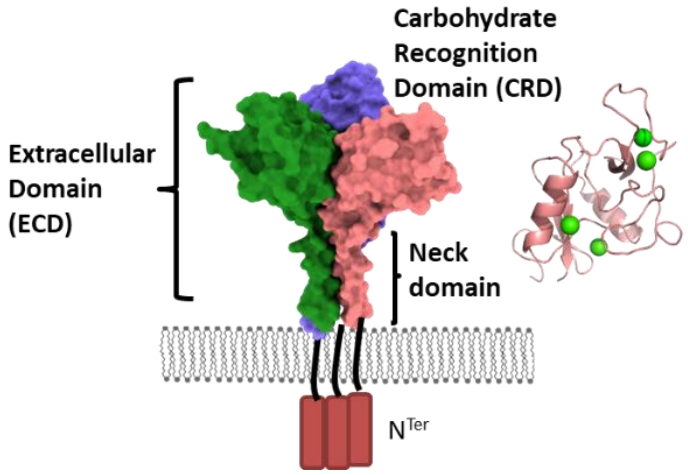


Figure 21 Schematic representation of MGL based on langerin structure. The CRD of MGL is shown in salmon with calcium ions in green. The two top-right calciums are part of the conserved calcium sites that participate in glycan binding through a Gln-Pro-Asp motif.

type lectin expressed on the cell surface of macrophages and Dendritic cells from skin and lymphoid organs. It is a type II C-type lectin composed of a short N-terminal cytoplasmic fragment, a transmembrane helix, a helical Neck domain followed by the Carbohydrate Recognition Domain (CRD) which binds several calcium ions (Figure 21). MGL is reported to be a trimer that has high affinity for terminal GalNAc residues, in particular in Tumor Necrosis antigens [38–40].

3. *Macrophage Galactose-type Lectin(MGL) specifically binds to E. coli R1 core oligosaccharide.*

The extracellular domain (ECD) of MGL was expressed, purified and labelled with AlexaFluor647 fluorophore and incubated with three *E. coli* strains expressing (non-pathogenic) LipoOligoSaccharides. *E. coli* BL21(DE3) IS the common laboratory strain and *E. coli* R1 and R3 express LOS with core oligosaccharides structures found in *E. coli* pathogen clinical isolates. Those three bacteria LOS differ only by the terminal core oligosaccharides structures (Figure 22).

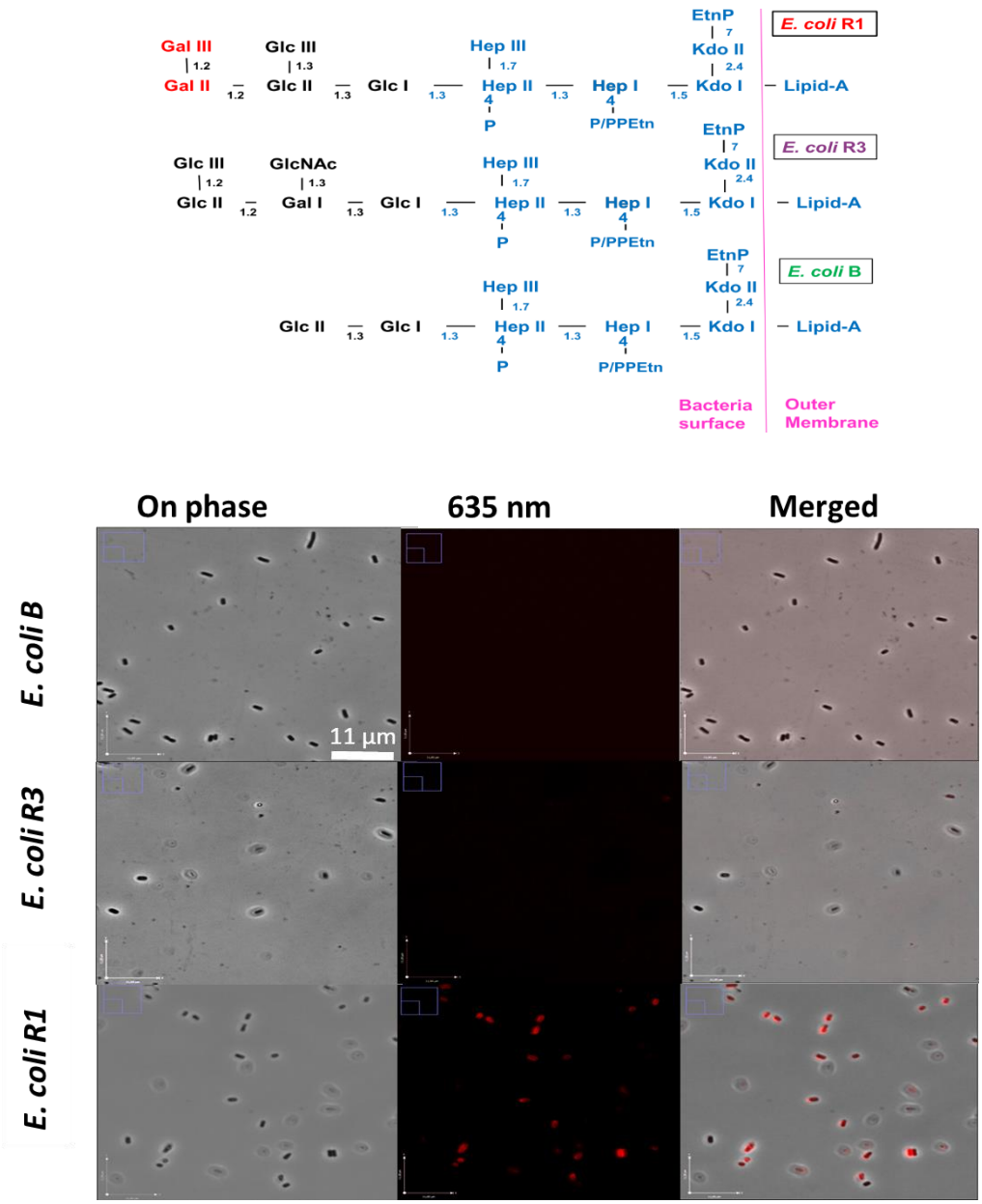


Figure 22 (Top) schematic structure of the B, R1 and R3 core oligosaccharides. (Bottom) labelling of the three bacterial strains with MGL-Alexa647 by epifluorescence microscopy. On phase images on the left, middle alexa647 fluorescence and right merged images.

MGL protein selectively binds to the surface of *E. coli* R1 and not B or R3 strains (Figure 22). Those results were quantitatively confirmed with flow cytometry analysis. An affinity in the low nanomolar range could also be measured by Biolayer Interferometry between immobilised MGL and purified vesicles of R1 LOS.

To determine the sugars or R1 core oligosaccharide involved in MGL interaction, interaction of human MGL was performed by saturation-transferred-difference NMR. In Antonio Molinaro's

laboratory they could obtain delipidated R1 oligosaccharide (OS R1) and determine by STD that MGL binds to the two terminal galactoses that are specific to R1 strain [41].

^{15}N-labelled CRD domain of MGL was titrated with R1 and R3 oligosaccharide, as well as with GalNAc as control. GalNAc binds to the canonical C-type lectin calcium site (Figure 23) with µM affinity as reported [42,43]. OS R1 and OS R3 interact both with MGL-CRD with low chemical shifts perturbations and a low affinity in the mM range. Moreover, while CSP are observed in the canonical Calcium binding site, another surface is affected by OSR1 and OSR3 oligosaccharides suggesting a putative secondary binding site.

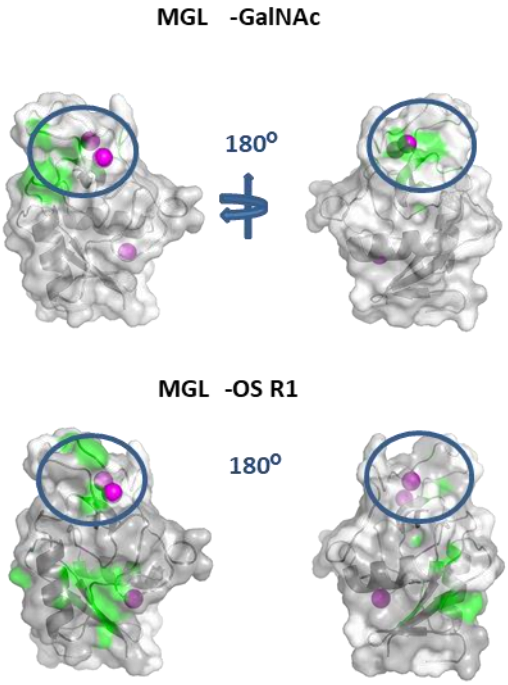


Figure 23 Combined ^{1}H-^{15}N Chemical shift perturbations of GalNAc and OSR1 (10 molar equivalents in green) on MFL-CRD structure. In blue is underlined the conserved Calcium binding site. A new surface with CSP is systematically found at the bottom.

MGL-LPS interactions are still under investigation. Mutants of the conserved calcium site have been produced to evaluate the contribution of a putative secondary glycan binding site on MGL. The presentation of the three CRD domains on the ECD of MGL is probably the key for tight binding observed on whole cells and the ^{13}C-^{15}N labelled ECD of MGL will also be investigated. Further fluorescence microscopy experiments (Confocal and superresolution) will be recorded to localise MGL binding at the surface of *E. coli* R1 bacteria, as preliminary experiment show differential binding of MGL depending on the growth phase.

III. FUTURE PROJECTS

My main studies revolve now around the study of LPS and their interaction with proteins using biophysical methods. NMR is nowadays able to tackle large protein systems with a combination of advanced labelling methods like deuteration and methyl labelling but also with solid-state NMR at high speed (100kHz) and ^{1}H detection. Nevertheless, the study of LPS by NMR, but this is also valid for other biophysical techniques remain particularly challenging. Of course, this is due to the hydrophobic nature of LPS and its assembly as membrane structures in solution. Also the nature of the bacterial outer membrane is asymmetric with LPS and phospholipids located in the outer and inner leaflets respectively. The focus of my future studies will be to setup the preparation of asymmetrical LPS particles resembling as much as possible to the outer membrane. This will enable high-resolution studies of LPS in its native environment and its interaction with proteins and antibiotics.

Bibliography

1. Li, J. P. & Kusche-Gullberg, M. Heparan Sulfate: Biosynthesis, Structure, and Function. *Int. Rev. Cell Mol. Biol.* **325**, 215–273 (2016).
2. Casu, B. *et al.* Heparin-like compounds prepared by chemical modification of capsular polysaccharide from E. coli K5. *Carbohydr. Res.* **263**, 271–84 (1994).
3. Zhang, X. *et al.* Chemoenzymatic synthesis of heparan sulfate and heparin oligosaccharides and NMR analysis: Paving the way to a diverse library for glycobiologists. *Chem. Sci.* **8**, 7932–7940 (2017).
4. Laguri, C. *et al.* 13C-labeled heparan sulfate analogue as a tool to study protein/heparan sulfate interactions by NMR spectroscopy: Application to the CXCL12α chemokine. *J. Am. Chem. Soc.* **133**, 9642–9645 (2011).
5. Zhang, Z. *et al.* Solution structures of chemoenzymatically synthesized heparin and its precursors. *J. Am. Chem. Soc.* **130**, 1–28 (2008).
6. Hagner-Mcwhirter, A., Lindahl, U. & Li, J. P. Biosynthesis of heparin/heparan sulphate: mechanism of epimerization of glucuronyl C-5. *Biochem. J.* **347 Pt 1**, 69–75 (2000).
7. Préchoux, A., Halimi, C., Simorre, J. P., Lortat-Jacob, H. & Laguri, C. C5-epimerase and 2- O -sulfotransferase associate in vitro to generate contiguous epimerized and 2- O -sulfated heparan sulfate domains. *ACS Chem. Biol.* **10**, 1064–1071 (2015).
8. Yates, E. A. *et al.* 1H and 13C NMR spectral assignments of the major sequences of twelve systematically modified heparin derivatives. *Carbohydr. Res.* **294**, 15–27 (1996).
9. Wang, Z., Zhang, Z., McCallum, S. A. & Linhardt, R. J. Nuclear magnetic resonance quantification for monitoring heparosan K5 capsular polysaccharide production. *Anal. Biochem.* **398**, 275–277 (2010).
10. Connell, B. J. *et al.* Heparan sulfate differentially controls CXCL12a-And CXCL12g-mediated cell migration through differential presentation to their receptor CXCR4. *Sci. Signal.* **9**, 1–12 (2016).
11. Laguri, C. *et al.* The novel CXCL12γ isoform encodes an unstructured cationic domain which regulates bioactivity and interaction with both glycosaminoglycans and CXCR4. *PLoS One* **2**, e1110 (2007).
12. Rueda, P. *et al.* The CXCL12γ chemokine displays unprecedented structural and functional properties that make it a paradigm of chemoattractant proteins. *PLoS One* **3**, e2543 (2008).
13. Sadir, R., Baleux, F., Grosdidier, A., Imberty, A. & Lortat-Jacob, H. Characterization of the Stromal Cell-derived Factor-1α-Heparin Complex. *J. Biol. Chem.* **276**, 8288–8296 (2001).
14. Dominguez, C., Boelens, R. & Bonvin, A. M. J. J. HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. *J. Am. Chem. Soc.* **125**, 1731–1737 (2003).
15. Kjellén, L. & Lindahl, U. Specificity of glycosaminoglycan–protein interactions. *Curr. Opin. Struct. Biol.* **50**, 101–108 (2018).
16. Préchoux, A., Simorre, J.-P., Lortat-Jacob, H. & Laguri, C. Deciphering the structural attributes of protein–heparan sulfate interactions using chemo-enzymatic approaches and NMR spectroscopy. *Glycobiology* (2021) doi:10.1093/glycob/cwab012.
17. Di Lorenzo, F. *et al.* Persistent cystic fibrosis isolate Pseudomonas aeruginosa strain RP73 exhibits an under-acylated LPS structure responsible of its low inflammatory activity. *Mol. Immunol.* **63**, 166–175 (2015).
18. Laguri, C. *et al.* Solid State NMR Studies of Intact Lipopolysaccharide Endotoxin. *ACS Chem. Biol.* acschembio.8b00271 (2018) doi:10.1021/acschembio.8b00271.
19. Kadurugamuwa, J. L., Clarke, A. J. & Beveridge, T. J. Surface action of gentamicin on Pseudomonas aeruginosa. *J. Bacteriol.* **175**, 5798–5805 (1993).
20. Kern, T. *et al.* Toward the characterization of peptidoglycan structure and protein-

peptidoglycan interactions by solid-state NMR spectroscopy. *J. Am. Chem. Soc.* **130**, 5618–5619 (2008).
21. Schanda, P. *et al.* Atomic model of a cell-wall cross-linking enzyme in complex with an intact bacterial peptidoglycan. *J. Am. Chem. Soc.* **136**, 17852–17860 (2014).
22. Laguri, C. *et al.* Interaction of lipopolysaccharides at intermolecular sites of the periplasmic Lpt transport assembly. *Sci. Rep.* **7**, 9715 (2017).
23. Miyashita, A. *et al.* Lipopolysaccharide O-antigen of enterohemorrhagic Escherichia coli O157:H7 is required for killing both insects and mammals. *FEMS Microbiol. Lett.* **333**, 59–68 (2012).
24. Baeta, T. Regulated activity of a bacterial transenvelope machinery : the LPS Transport System. (Université Grenoble Alpes [2020-....], 2020).
25. Knowles, T. J. *et al.* Membrane proteins solubilized intact in lipid containing nanoparticles bounded by styrene maleic acid copolymer. *J. Am. Chem. Soc.* **131**, 7484–7485 (2009).
26. Dörr, J. M. *et al.* The styrene–maleic acid copolymer: a versatile tool in membrane research. *Eur. Biophys. J.* **45**, 3–21 (2016).
27. Tang, X. *et al.* Cryo-EM structures of lipopolysaccharide transporter LptB2FGC in lipopolysaccharide or AMP-PNP-bound states reveal its transport mechanism. *Nat. Commun.* **10**, 4175 (2019).
28. Li, Y., Orlando, B. J. & Liao, M. Structural basis of lipopolysaccharide extraction by the LptB2FGC complex. *Nature* **567**, 486–490 (2019).
29. Owens, T. W. *et al.* Structural basis of unidirectional export of lipopolysaccharide to the cell surface. *Nature* vol. 567 550–553 (2019).
30. Kaur, H. *et al.* Coupled ATPase-adenylate kinase activity in ABC transporters. *Nat. Commun.* **7**, 13864 (2016).
31. Mayer, M. & Meyer, B. Characterization of ligand binding by saturation transfer difference NMR spectroscopy. *Angew. Chemie - Int. Ed.* **38**, (1999).
32. Carrier, I., Urbatsch, I. L., Senior, A. E. & Gros, P. Mutational analysis of conserved aromatic residues in the A-loop of the ABC transporter ABCB1A (mouse Mdr3). *FEBS Lett.* **581**, 301–308 (2007).
33. Ambudkar, S. V., Kim, I. W., Xia, D. & Sauna, Z. E. The A-loop, a novel conserved aromatic acid subdomain upstream of the Walker A motif in ABC transporters, is critical for ATP binding. *FEBS Letters* vol. 580 1049–1055 (2006).
34. Sherman, D. J. *et al.* Decoupling catalytic activity from biological function of the ATPase that powers lipopolysaccharide transport. *Proc. Natl. Acad. Sci.* **111**, 4982–4987 (2014).
35. Mayer, S., Raulf, M. K. & Lepenies, B. C-type lectins: their network and roles in pathogen recognition and immunity. *Histochemistry and Cell Biology* vol. 147 223–237 (2017).
36. Hanske, J. *et al.* Bacterial polysaccharide specificity of the pattern recognition receptor langerin is highly species-dependent. *J. Biol. Chem.* **292**, 862–871 (2017).
37. Geissner, A. *et al.* Microbe-focused glycan array screening platform. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 1958–1967 (2019).
38. Iida, S. I., Yamamoto, K. & Irimura, T. Interaction of human macrophage C-type lectin with O-linked N- acetylgalactosamine residues on mucin glycopeptides. *J. Biol. Chem.* **274**, 10697–10705 (1999).
39. Marcelo, F. *et al.* Delineating binding modes of Gal/GalNAc and structural elements of the molecular recognition of tumor-associated mucin glycopeptides by the human macrophage galactose-type lectin. *Chem. - A Eur. J.* **20**, 16147–16155 (2014).
40. van Vliet, S. J. *et al.* Carbohydrate profiling reveals a distinctive role for the C-type lectin MGL in the recognition of helminth parasites and tumor antigens by dendritic cells. *Int. Immunol.* **17**, 661–669 (2005).
41. Maalej, M. *et al.* Human Macrophage Galactose-Type Lectin (MGL) Recognizes the Outer

Core of Escherichia coli Lipooligosaccharide. *ChemBioChem* **20**, 1778–1782 (2019).
42. Marcelo, F. *et al.* Identification of a secondary binding site in human macrophage galactose-type lectin by microarray studies: Implications for the molecular recognition of its ligands. *J. Biol. Chem.* **294**, 1300–1311 (2019).
43. Diniz, A. *et al.* The Plasticity of the Carbohydrate Recognition Domain Dictates the Exquisite Mechanism of Binding of Human Macrophage Galactose-Type Lectin. *Chem. - A Eur. J.* **25**, 13945–13955 (2019).
44. Tytgat, H. L. P. & Lebeer, S. The Sweet Tooth of Bacteria: Common Themes in Bacterial Glycoconjugates. *Microbiol. Mol. Biol. Rev.* **78**, 372–417 (2014).

IV. CURRICULUM VITÆ

Cédric LAGURI

45 y/o

cedric.laguri@ibs.fr

0457428577

Biomolecular NMR group
Institut de Biologie Structurale UMR 5075
71 avenue des martyrs
38044 Grenoble Cedex 9

CNRS researcher in cell wall and antibiotic resistance team

Professional experience

2013-Now	**CNRS researcher**, cell wall and antibiotic resistance team (J.P. Simorre)
2008-2013	**CNRS researcher**, SAGAG group IBS (Grenoble).
2006-2008	**Post-Doctoral position**, SAGAG group IBS (Grenoble). Dir H. Lortat-jacob.
2004-2006	**Post-Doctoral position,** Laboratoire de Structure des Protéines, DIEP CEA Saclay.
2000-2004	**PhD** NMR laboratory of the Molecular Biology and Biotechnology department, Sheffield University (UK). Dir. Pr M.P. Williamson.
1999-2000	**M2 degree**, Molecular Biophysics, Université Pierre et Marie Curie, Paris 6.
1998-1999	**M2 degree**, Protein purification and characterization. Université Henri Poincaré/INPL, Nancy.
1994-1998	**M1 Biochemistry degree,** Université de Cergy-Pontoise (France).

Supervision

2018-2021	**PhD** student. Analysis of the determinants of lectins interactions with LPS sugars by NMR, cotutelle with Napoli (M. Maalej) Cosupervisor.
2017-2020	**PhD** student. Characterization of the Lpt machinery and its binding to LPS at atomic scale using NMR (T. Baeta) Cosupervisor.
2016	**M2** Erasmus student. Expression of inner membrane proteins from the lpt transport system (A. Burt 3 months)
2016	**M2** BBS (biochemistry and Structural Biology) Student. LptC-LptA transport system in *E. coli* (K. Pounot 6 months)
2012	Research engineer. ANR project Heplibscreen (A. Prechoux, 1 year)

2012 **M2** BBS student. Study by NMR of the activity of two enzymes essential in Heparan-sulfate metabolism (C. Halimi, 6 months)

2011 M1 student. Expression, purification, and characterization of Heparan-sulfate C5-epimérase enzyme (C. Halimi, 6 months).

2009 BTS Bioanalyses and controls. Purification of CXCL12A and B chemokines and interaction with Heparan sulfates (J. Ianni 3 months)

2009 M1 project student. Ligand synthesis for the study of protein-Heparan sulfates interaction (N. Dumas, 3 months).

2004 **M2** student: Backbone assignment of human PWWP domain, study of its phosphorylation and screenings for new protein partners by GST-Pulldown (N. Friedrich 6 months).

2002 2 M1 project students (6 months). Characterization of the link between phosphorylation and dimerisation in PrrA proteins from R. sphaeroides.

Teaching

2015 Practicals for Advanced Isotopic Labeling Methods workshop (January 2015 Grenoble)

2010-2018 Teaching and practicals in the "Protein structure determination by NMR"and "protein-ligand interaction by NMR" workshops in Grenoble and Gif-sur-Yvette.

2001-2002: Practicals (60 hours *biochemistry*) for *L1* students in biochemistry. Sheffield University

Scientific Events organization

2015, 2018 Workshop "Protein structure determination by NMR" IBS/Grenoble

2018 Structural Glycosciences Workshop IBS/Grenoble

Grants

2021 GRAL PhD studentship NanoLPS.

2018-2021 Glyco@alps shared PhD studentship.

2017-2020 Train2Target (ITN) project from the European Union's Horizon 2020 framework program for research and innovation (Project #721484)

2015-2017 AGIR pole CBS MEMTARG (coordinator) - Understanding and targeting of a bacterial complex of lipopolysaccharides transport

2013-2015 ANR Young researcher grant (coordinator): Heparan sulfate libraries screening by NMR spectroscopy (HepLibScreen)

Publications

Laguri, C., Gilquin, B., Wolff, N., Romi-Lebrun, R., Courchay, K., Callebaut, I., S. Zinn-Justin . (2001). Structural Characterization of the LEM Motif Common to Three Human Inner Nuclear Membrane Proteins. Structure London England 1993, 9(01), 503–511.

Potter, C. A., Ward, A., Laguri, C., Williamson, M. P., Henderson, P. J. F., & Phillips-Jones, M. K. (2002). Expression, purification and characterisation of full-length histidine protein kinase RegB from Rhodobacter sphaeroides. Journal of Molecular Biology, 320, 201–213. doi:10.1016/S0022-2836(02)00424-2

Laguri, C., Phillips-Jones, M. K., & Williamson, M. P. (2003). Solution structure and DNA binding of the effector domain from the global regulator PrrA (RegA) from Rhodobacter sphaeroides: insights into DNA binding specificity. Nucleic Acids Research, 31(23), 6778–6787. doi:10.1093/nar/gkg891

Laguri, C. , Stenzel, R. A., Donohue, T. J. T., Phillips-Jones, M. K., & Williamson, M. P. (2006). Activation of the global gene regulator PrrA (RegA) from Rhodobacter sphaeroides. Biochemistry, 45(25), 7872–7881. doi:10.1021/bi060683g

Laguri, C. , Sadir, R., Rueda, P., Baleux, F. F., Gans, P., Arenzana-Seisdedos, F., & Lortat-Jacob, H. (2007). The novel CXCL12gamma isoform encodes an unstructured cationic domain which regulates bioactivity and interaction with both glycosaminoglycans and CXCR4. PloS One, 2(10), e1110. doi:10.1371/journal.pone.0001110

Laguri, C., Arenzana-Seisdedos, F., & Lortat-Jacob, H. (2008). Relationships between glycosaminoglycan and receptor binding sites in chemokines-the CXCL12 example. Carbohydrate Research. doi:10.1016/j.carres.2008.01.047

Laguri, C. , Duband-Goulet, I., Friedrich, N., Axt, M., Belin, P., Callebaut, I., Couprie, J. J. (2008). Human mismatch repair protein MSH6 contains a PWWP domain that targets double stranded DNA. Biochemistry, 47(23), 6199–207. doi:10.1021/bi7024639

Rueda, P., Balabanian, K., Lagane, B., Staropoli, I., Chow, K., Levoye, A., Laguri, C., Sadir, R., Delaunay, T., Izquierdo, E., Pablos, J. L., Lendinez, E., Caruz, A., Franco, D., Baleux, F., Lortat-Jacob, H., and Arenzana-Seisdedos, F. (2008) The CXCL12γ chemokine displays unprecedented structural and functional properties that make it a paradigm of chemoattractant proteins. PLoS One 3, e2543.

Jasnin, M., van Eijck, L., Koza, M. M., Peters, J., Laguri, C., Lortat-Jacob, H., & Zaccai, G. (2010). Dynamics of heparan sulfate explored by neutron scattering. Physical Chemistry Chemical Physics: PCCP, 12, 3360–3362. doi:10.1039/b923878f

Boldajipour, B., Doitsidou, M., Tarbashevich, K., Laguri, C., Yu, S. R., Ries, J., Raz, E. (2011). Cxcl12 evolution--subfunctionalization of a ligand through altered interaction with the chemokine receptor. Development (Cambridge, England), 138(14), 2909–2914. doi:10.1242/dev.068379

Laguri, C., Sapay, N., Simorre, J.-P., Brutscher, B., Imberty, A., Gans, P., & Lortat-Jacob, H. (2011). 13C-labeled heparan sulfate analogue as a tool to study protein/heparan sulfate interactions by NMR spectroscopy: application to the CXCL12alpha chemokine. Journal of the American Chemical Society, 133(25), 9642–9645. doi:10.1021/ja201753e

Hou, Y., Genua, M., Tada Batista, D., Calemczuk, R., Buhot, A., Fornarelli, P., Koubachi, J., Bonnaffé, D., Saesen, E., Laguri, C., Lortat-Jacob, H., and Livache, T. (2012) Continuous evolution profiles for electronic-tongue-based analysis. *Angew. Chemie - Int. Ed. 51*, 10394–10398.

Seffouh, A., Milz, F., Przybylski, C., Laguri, C., Oosterhof, A., Bourcier, S., Vives, R. R. (2013). HSulf sulfatases catalyze processive and oriented 6-O-desulfation of heparan sulfate that differentially regulates fibroblast growth factor activity. FASEB Journal : Official Publication of the Federation of American Societies for Experimental Biology, 27(6), 2431–2439. doi:10.1096/fj.12-226373

Saesen, E., Sarrazin, S., Laguri, C., Sadir, R., Maurin, D., Thomas, A., Lortat-Jacob, H. (2013). Insights into the mechanism by which interferon-gamma basic amino acid clusters mediate protein binding to heparan sulfate. Journal of

the American Chemical Society, 135(25), 9384–9390. doi:10.1021/ja4000867

Schanda, P., Triboulet, S., Laguri, C., Bougault, C. M., Ayala, I., Callon, M., Simorre, J.-P. (2014). Atomic model of a cell-wall cross-linking enzyme in complex with an intact bacterial peptidoglycan. Journal of the American Chemical Society, 136(51), 17852–60. doi:10.1021/ja5105987

Prechoux, A., Halimi, C., Simorre, J.-P., Lortat-Jacob, H., & Laguri, C. (2015). C5-epimerase and 2-O-sulfotransferase associate in vitro to generate contiguous epimerized and 2-O-sulfated heparan sulfate domains. ACS Chemical Biology. doi:10.1021/cb501037a

Fichou, Y., Schirò, G., Gallat, F.-X., Laguri, C., Moulin, M., Combet, J., Lortat-Jacob, H.,... Weik, M. (2015). Hydration water mobility is enhanced around tau amyloid fibers. Proceedings of the National Academy of Sciences of the United States of America, 112(20), 6365–70. http://doi.org/10.1073/pnas.1422824112

Triboulet, S., Bougault, C.M., Laguri, C., Hugonnet, J.E., Arthur, M., and Simorre, J.P. (2015). Acyl acceptor recognition by Enterococcus faecium l,d-transpeptidase Ldtfm. Mol. Microbiol. *98*, 90–100.

Connell, B.J., Sadir, R., Baleux, F., Laguri, C., Kleman, J.-P., Luo, L., Arenzana-Seisdedos, F., and Lortat-Jacob, H. (2016). Heparan sulfate differentially controls CXCL12a-And CXCL12g-mediated cell migration through differential presentation to their receptor CXCR4. Sci. Signal. *9*, 1–12.

Manuse, S., Jean, N.L., Guinot, M., Lavergne, J.-P., Laguri, C., Bougault, C.M., VanNieuwenhze, M.S., Grangeasse, C., and Simorre, J.-P. (2016). Structure–function analysis of the extracellular domain of the pneumococcal cell division site positioning protein MapZ. Nat. Commun. *7*, 12071.

Laguri C., Sperandeo P., Pounot K., Ayala I. , Silipo A., Bougault C.M., Molinaro A., Polissi A. and Simorre J.P. (2017). Interaction of lipopolysaccharides at intermolecular sites of the periplasmic Lpt transport assembly. *Sci. Rep.* 2017, *7* (1), 9715.

Laguri, C., Silipo, A., Martorana, A. M., Schanda, P., Marchetti, R., Polissi, A., Molinaro, A., and Simorre, J.-P. (2018) Solid State NMR Studies of Intact Lipopolysaccharide Endotoxin. *ACS Chem. Biol.* acschembio.8b00271.

Maya-Martinez R, Alexander JAN, Otten CF, Ayala I, Vollmer D, Gray J, Bougault CM, Burt A, Laguri C, Fonvielle M, Arthur M, Strynadka NCJ, Vollmer W, Simorre JP (2019). Recognition of Peptidoglycan Fragments by the Transpeptidase PBP4 From Staphylococcus aureus. Front Microbiol. doi: 10.3389/fmicb.2018.03223.

Maalej M, Forgione RE, Marchetti R, Bulteau F, Thépaut M, Lanzetta R, Laguri C, Simorre JP, Fieschi F, Molinaro A, Silipo A. (2019). Human Macrophage Galactose-Type Lectin (MGL) Recognizes the Outer Core of Escherichia coli Lipooligosaccharide. Chembiochem. doi: 10.1002/cbic.201900087.

Moura, E. C. C. M., Baeta, T., Romanelli, A., Laguri, C., Martorana, A. M., Erba, E., Simorre, J. P., Sperandeo, P., and Polissi, A. (2020) Thanatin Impairs Lipopolysaccharide Transport Complex Assembly by Targeting LptC–LptA Interaction and Decreasing LptA Stability. *Front. Microbiol.*

Caveney, N. A., Egan, A. J. F., Ayala, I., Laguri, C., Robb, C. S., Breukink, E., Vollmer, W., Strynadka, N. C. J., and Simorre, J. P. (2020) Structure of the Peptidoglycan Synthase Activator LpoP in Pseudomonas aeruginosa. *Structure*.

Dekoninck, K., Létoquart, J., Laguri, C., Demange, P., Bevernaegie, R., Simorre, J. P., Dehu, O., Iorga, B. I., Elias, B., Cho, S. H., and Collet, J. F. (2020) Defining the function of OmpA in the Rcs stress response. Elife.

Pazos, M., Peters, K., Boes, A., Safaei, Y., Kenward, C., Caveney, N. A., Laguri, C., Breukink, E., Strynadka, N. C. J., Simorre, J. P., Terrak, M., & Vollmer, W. (2020). Spor proteins are required for functionality of class a penicillin-binding proteins in escherichia coli. *MBio*, *11*(6), 1–16. https://doi.org/10.1128/mBio.02796-20

Préchoux, A., Simorre, J.-P., Lortat-Jacob, H., & Laguri, C. (2021). Deciphering the structural attributes of protein–heparan sulfate interactions using chemo-enzymatic approaches and NMR spectroscopy. *Glycobiology*. https://doi.org/10.1093/glycob/cwab012