University of Naples Federico II
Polytechnic and Basic Sciences School

Department of Chemical Sciences

Ph.D. School in Chemical Sciences – Cycle XXXV

Binding studies between *Neisseria gonorrhoeae* lipooligosaccharides (LOS) and monoclonal antibodies

Forgione Maria Concetta

Supervisors:

Prof. Alba Silipo

Prof. Roberta Marchetti

GSK Supervisor:

Dr. Francesco Berti

Dr. Silvia Martini

Examiner:

Prof. Luigi Paduano

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

AA	ACETILATED ALDITOLS
AA	AMINO ACID
Ab	ANTIBODY
AMG	ACETYLATED METHYL GLYCOSIDES
AMR	ANTIMICROBIAL RESISTANCE
APC	ANTIGEN PRESENTING CELLS
BCR	B-CELL RECEPTOR
BSA	BOVINE SERUM ALBUMIN
CAP	COMMUNITY-ACQUIRED PNEUMONIA
COSY	CORRELATION SPECTROSCOPY
DOSY	DIFFUSION ORDERED SPECTROSCOPY
GC-MS	GAS CROMATOGRAPHY MASS SPECTROMETRY
GPCA	*GENITALIUM* AND *PNEUMONIAE CYTOADHESINS*
GSL	GLYCOSPHINGOLIPIDS
HIV	HUMAN IMMUNODEFICIENCY VIRUS
HMBC	HETERONUCLEAR MULTIPLE-BOND CORRELATION
HPLC	HIGH PERFORMANCE LIQUID CROMATOGRAPHY
HSQC	HETERONUCLEAR SINGLE-QUANTUM COHERENCE
LGA	LAMARCKIAN GENETIC ALGORITHM
LNnT	LACTO-N-NEOTETRAOSE
LOS	LIPOOLIGOSACCHARIDE
LPS	LIPOPOLYSACCHARIDES
MAb	MONOCLONAL ANTIBODY
MC	MONTE CARLO
MD	MOLECULAR DYNAMICS
Mge	*MYCOPLASMA GENITALIUM*
MGL	MACROPHAGE GALACTOSE-TYPE LECTIN

MHC	MAJOR HISTOCOMPLEX
MM	MOLECULAR MECHANICS
Mpn	*MYCOPLASMA PNEUMONIAE*
MW	MOLECULAR WEIGHT
Ng	*NEISSERIA GONORRHOEAE*
NMR	NUCLEAR MAGNETIC RESSONANCE
NOE	NUCLEAR OVERHAUSER EFFECT
NOESY	NUCLEAR OVERHAUSER EFFECT SPECTROSCOPY
OM	OUTER MEMBRANE
OMV	OUTER MEMBRANE VESICLES
OS	CORE OLIGOSACCHARIDE
PAMP	PATHOGEN-ASSOCIATED MOLECULAR PATTERN
PBC	PERIODIC BOUNDARY CONDITIONS
PID	PELVIC INFLAMMATORY DISEASE
PBS	PHOSPHATE-BUFFERED SALINE
PCP	PETROLEUM ETHER-CHLOROFORM-PHENOL
PGC	POROUS GRAPHITIC CARBON
PRR	PATTERN RECOGNITION RECEPTOR
QM	QUANTUM MECHANICS
RMSD	ROOT MEAN SQUARE DEVIATION
ROESY	ROTATING FRAME OVERHAUSER EFFECT SPECTROSCOPY
SDS-PAGE	SODIUM DODECYL SULPHATE- POLYACRYLAMIDE GEL ELECTROPHORESIS
STD NMR	SATURATION TRANSFER DIFFERENCE NMR
TFA	TRIFLUOROACETIC ACID
TNF	TUMOUR NECROSIS FACTOR
TOCSY	TOTAL CORRELATION SPECTROSCOPY
Tr-NOESY	TRANSFERED NOESY

WB	WESTERN BLOT
WHO	WORLD HEALTH ORGANIZATION
WT	WILD TYPE

SUGAR ABBREVIATIONS

3'SLN	3'sialyl N-acetyl lactosamine
6'SL	6'sialyllactose
Gal	Galactose
GalNA	Amino galacturonic acid
GalNAc	N-acetyl galactosamine
Glc	Glucose
GlcNA	Aminoglucuronic acid
GlcNAc	N-acetyl glucosamine
Hep	Heptose
KDO	3-deoxy-D-manno-oct-2-ulopyranosonic acid
Neu5Ac	N-acetyl neuraminic acid
Sia	Sialic acid

SECTION I – INTRODUCTION

Chapter 1 – Bacterial-host interaction and principle of vaccination

1.1 Bacteria and the immune system

Bacteria are unicellular microorganisms of a few micrometres in length, composed by a cell structure simpler than many other organisms. Bacteria are able to reproduce themselves in the environment and also in various tissues of the human body. They are commonly present on the skin surface and on mucous membranes and some of them are not harmful to the host; on the contrary, their presence could be useful for certain metabolic functions and for immune defences. They are known as *commensals* and their whole is called *microbiota* [1]. Conversely, called *pathogens* can damage tissues and organs and are responsible for an enormous variety of infectious diseases. Structurally bacteria are composed of a plasma membrane and a semipermeable barrier consisting of phospholipids and proteins surrounded by the cytoplasm. Bacteria cells are generally surrounded by an *outer cell wall* and an *inner cell membrane*. Some microorganisms may even have a third, outermost protective layer, called capsule. The cell wall consists of a rigid structure around the plasma membrane, that gives structural packing and protection, providing the characteristic shape of the cell. Bacteria can be classified into five groups according to their basic shapes: spherical (cocci), rod (bacilli), spiral (spirilla), comma (vibrios) or corkscrew (spirochaetes). They can exist as single cells, in pairs, chains or clusters. A peculiar feature is the presence of an external dense carbohydrate layer, named *glycocalyx*, which makes bacteria resistant to phagocytosis and allows, in some cases, to evade the immune system. The term "*glycocalyx*" is an umbrella term to indicate the entirety of free glycans, glycoproteins, proteoglycans, and glycolipids present on the cell surface [2]. Free glycans are usually polymers called glycopolymers. The terms glycolipid and glycoprotein refer to a glycan covalently linked to a lipid and protein,

respectively. In the case of glycoprotein, if the attached glycans are long, polymeric sugar chains, the resulting conjugate is called proteoglycan. Cell envelopes of most bacteria fall into one of two major groups: Gram-negative and Gram-positive bacteria. This division refers to the different ability of bacteria to react to the Christian Gram's staining method, developed in 1884 [3]. One group of bacteria, the Gram-positive, retains Gram's stain while the other one, the Gram-negative, does not. Structurally, the two bacterial classes differ for the presence of the *outer membrane* (OM). Gram-negative bacteria are composed by an internal membrane, a thin peptidoglycan cell wall and an outer membrane containing lipopolysaccharides (LPSs), lipoproteins, and phospholipids. Otherwise, Gram-positive bacteria lack an outer membrane but they have a thick multilayer section of peptidoglycans, which surround the internal membrane (Figure 1.1) [4].

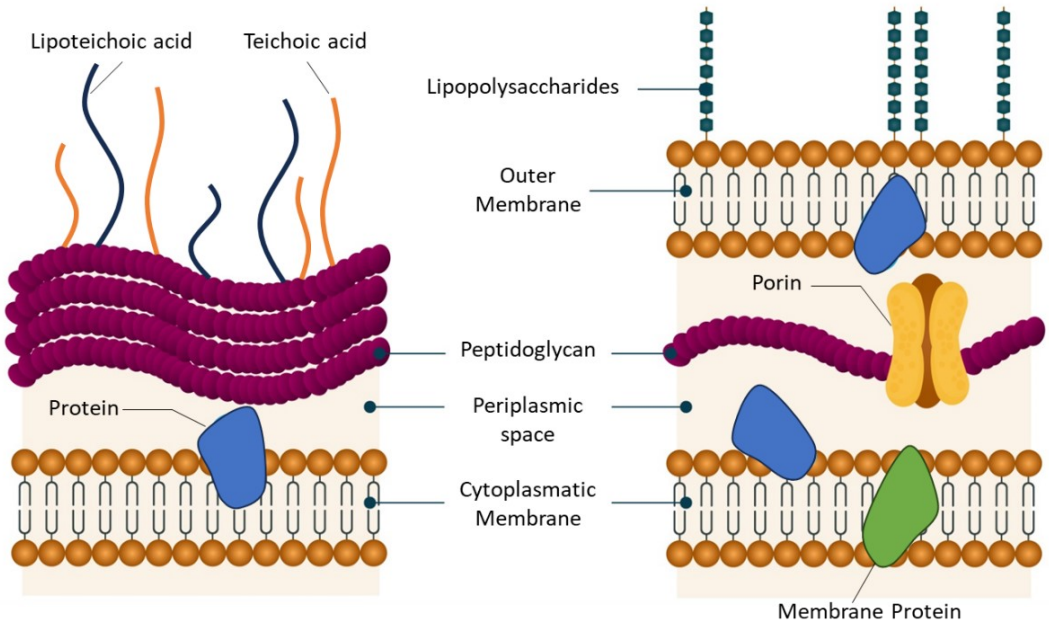


Figure 1.1. Schematic structure of Gram-positive (on the left) and Gram-negative (on the right) cell walls. Gram-positive cell walls contain only one lipid plasma membrane and a thick peptidoglycan layer interlinked with teichoic and lipoteichoic acids. Gram-negative

bacteria have an inner and an outer cell membrane and only a thin layer of peptidoglycan in the periplasmic space between the inner and outer membrane [3].

The molecular properties of the bacterial cell surface are crucial in the adhesion process. Through their dense and complex array of glycans which decorates the cells surface, bacteria communicate and interact with the host cells activating the *immune system*. The immune system is a host's protection mechanism against infections made up of a complex network of organs, cells and proteins and able to discriminate *self* from *non-self* molecules (present in microorganisms [5]). The immune system consists of two different responses: *innate* and *adaptative* response. The innate immune response, the early line of host immune defence, provides immediate defence against pathogens, but it does not confer long-lasting protective immunity to the host. Innate immunity occurs when the Pathogens (Microbes)-Associated Molecular Patterns (PAMPs; MAMPs) are recognized by specific receptors, known as Pattern Recognition Receptors (PRRs) [6]. PAMPs are a vast array of molecules, including glycans and glycoconjugates [7]. Among PRRs, four main categories are recognized: Toll-like receptors (TRLs), NOD-like receptors (NLRs), Triggering Receptors Expressed on Myeloid cells (TREMs) and RIG-like receptors (RLRs). On the other hand, the adaptative immune response generates a specific immune *stimulus* against microbes, producing specialized lymphocytes and specific antibodies. The adaptative immune system consists of T and B lymphocytes, which express on their plasmatic membrane a large repertory of antigen-specific receptors [8], known as T Cell Receptor (TCR) and B Cell Receptor (BCR), through which it is able to recognize *non-self* antigens [5]. The advantage of the adaptative immunity resides in the capability to remember antigens already encountered (*memory B cells*), becoming more effective each time the pathogen is met. B lymphocytes are responsible for the humoral response and T lymphocytes are involved in both

humoral and cell-mediated immunity. When the B cells interact with the antigen, they produce plasma cells and secrete specific antibodies against specific antigens. T cells produce T-cytotoxic (CD8+) and T-helper (CD4+) cells. The cooperation of these two lymphocytes (B and T cells) leads to memory B cells production. When memory B cells encounter the same antigen again, they produce specific antibodies deputies to recognize and bind these antigens.

1.2 Antimicrobial resistance (AMR): a new threat to life

Bacterial infections are commonly treated with antibiotics that in the last century increased considerably life expectancy [9]. However, over the years, a massive use of antibiotics, not only for therapeutic purposes, caused an exponential increase in antimicrobial resistance (AMR). Some bacteria (as well as viruses, fungi and parasites) developed the ability to bypass the mechanisms that drugs use against them, making treatments ineffective. Antimicrobial resistance or antibiotic resistance (exhibited by bacteria) can be categorized into two classes: *intrinsic* or *acquired* [10]. *Intrinsic* resistant consists in a natural phenomenon found within bacteria which does not depend on antibiotic exposure. *Klebsiella pneumoniae* spp., for instance, are intrinsically resistant to ampicillin or *Enterococcus faecium* and *faecalis* are resistant to cephalosporins. The mechanisms by which intrinsic resistance acts are various, including lack of outer membrane permeability and the absence of target sites required for antibiotic [11]. On the other hand, *acquired* resistance is related to environmental factors. It is defined as the ability of a bacterium, previously sensitive, to acquire a resistance mechanism through a mutation or by the acquisition of new genetic material from exogenous sources. The main mechanisms bacteria use to counteract the antibiotics action are based on *i*. inactivation of the drug through its hydrolysis or structural modification, *ii.* limiting the access by membrane

permeability reduction or overexpressing efflux pumps and *iii.* modification of antibiotic targets (Figure 1.2).

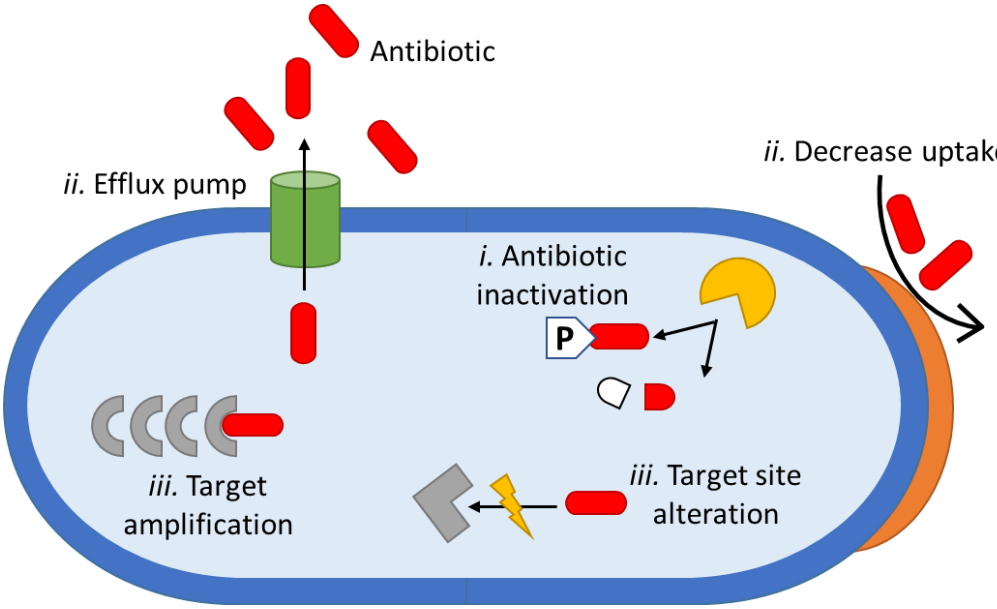


Figure 1.2. A schematic view of antimicrobial resistance mechanisms in bacteria.

Mainly in recent years, the AMR become a new emergency in global health care and previous uncomplicated infections have turned into more complex and severe diseases. Indeed, the AMR can result in treatment inefficiency or failures and persistent infections, leading to a high risk of severe diseases and transmission. Several medical procedures as immunosuppressive chemotherapy and organ transplantation are becoming more critical due to the lack of effective antibiotics against multidrug-resistant pathogens. Recently, it has been estimated that a reduction of about 30% in the effectiveness of antibiotic prophylaxis for chemotherapeutic and surgical treatments could lead to a drastic increase in the number of infections and more than 6,000 infection-related deaths per year in the U.S. [9]. New measures and medical treatments are therefore needed to take under

control and limit the increase in microbial resistance. In this scenario, vaccines could be a valuable weapon against AMR incurrence.

1.3 Vaccines and principle of vaccination

Vaccines represent one of the main medical applications in preventing infectious diseases [12]. At least 31 human diseases have been countered by effective vaccination strategies [13]. In 2020, WHO (World Health Organization) estimated that 2-3 million lives were saved each year thanks to current immunization programmes [14]. Immunology history began with the Edward Jenner's experiment in 1796. Jenner injected the purulent material from cowpox lesions into the patient affected by smallpox infection, inducing protection [15]. In 1885 Louis Pasteur, improving Jenner's work, developed the first human vaccine against rabies based on attenuated viruses [8]. Pasteur placed the basis of vaccinology meaning the principle of isolation, inactivation, and administration of disease-causing pathogens.

The principle of vaccination is to stimulate the immune response against pathogens by mimicking its natural interaction with the human immune system, without causing the relating disease. As a result of the immune system stimulation, lymphocytes B produce the memory B cells. In this way, when the natural infection occurs, the immune system responds quicker and more effectively to infection. Furthermore, vaccine reduces the risk of complications and mortality following subsequent exposure to an infectious agent [16].

Vaccination confers protection through the production of specific antibodies which are capable to recognize and specifically bind the antigens present on pathogen cell surface. Typically, antibodies bind a small region of antigens, called *epitope*. To generate memory B cells, the activation of T-helper cells (CD4+) which can be stimulated by T-dependent antigens, such as proteins is necessary. T-dependent antigens interact with antigen-presenting cells (APCs) as dendric cells, macrophages

and B cells. The internalized antigens are digested into smaller peptides containing epitopes, which are then presented to CD4+ T cells by the Major Histocompatibility Complex class II (MHC-II). Interaction with T cells drives B cells proliferation and their subsequential differentiation into plasma cells and memory B cells.

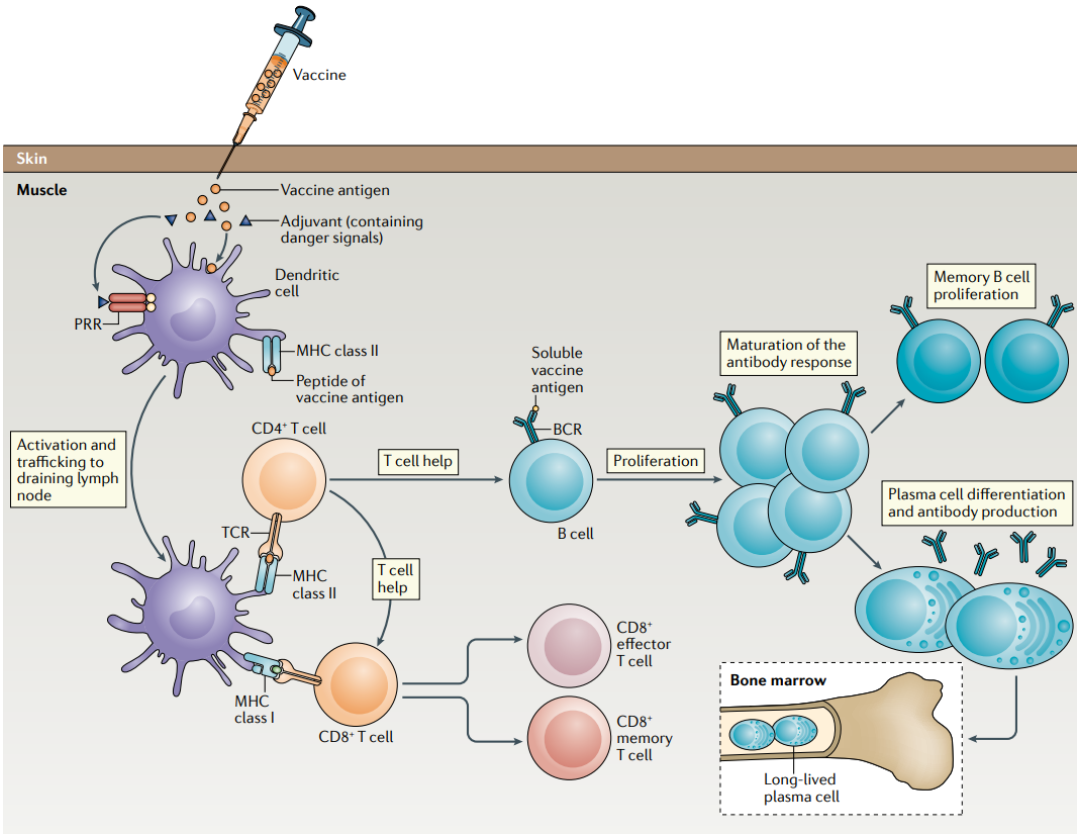


Figure 1.3: The immune response to a vaccination. Scheme of the immune response following vaccination. After injection, the pattern recognition receptors (PRRs) activate the dendritic cells (DCs), which take up the antigen. The antigen is then processed to the draining lymph node, where the exposition of the peptides (present in the antigen vaccine) by MHC molecules on the DCs activates T cells through T cells receptor (TCR). T cells drive B cells development in the lymph node. Here, B cells (which are T cell-dependent) determine the antibody response. In the two weeks following administration, take place an increase in serum antibody levels, thanks to the production of palm cells that secrete specific antibodies

for the vaccine protein. A proliferation of memory B cells occurs, on which immune memory depends. Long-lived plasma cells that reside in bone marrow, can continue to produce antibodies for a long time (Figure from Pollard *et al*, 2021) [14].

In contrast to T-dependent antigens, T-independent antigens, such as polysaccharides, do not require T cell help to elicit an immune response. Many vaccines currently in use consist of purified capsular polysaccharides from invasive bacteria such as *Streptococcus pneumoniae* and *Neisseria meningitidis* [17]. Polysaccharide-based vaccines provide protection against diseases, stimulating B cells by linking B-cell receptors (BCR) and producing immunoglobulins. However, the resulting process does not involve memory B cells production [18]. This category of vaccines is effective in adults but not in children, especially under 2 years of age. To circumvent this limitation, a new class of vaccines has been developed: protein–polysaccharide conjugate vaccines, which consist in poly- or oligosaccharides from the bacterial surface chemically conjugated to a protein carrier, like CRM_{197} [19]. The protein carrier (T-dependent antigens) is recognized by the T cells which provide help B cells, leading to the production of both plasma and memory B cells. As a result, glycoconjugates-based vaccines induce long-lasting protection and are effective even in infants and people who belong to high-risk categories [20]. Immunisation using glycoconjugates is one of the safest and most effective approaches for fighting bacterial diseases and containing AMR.

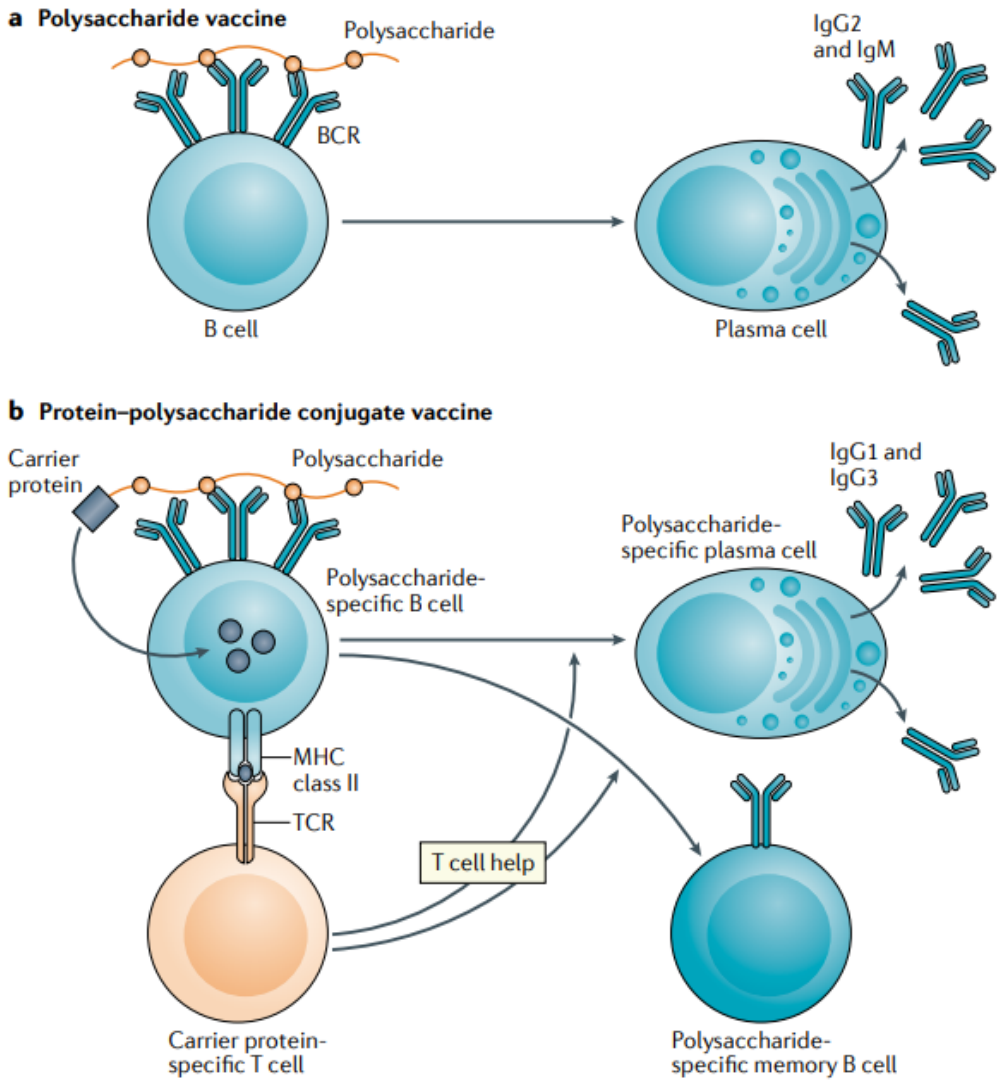


Figure 1.4. Differences in immune response between polysaccharide and protein–polysaccharide conjugate vaccines. a. Polysaccharide-vaccines induce B cells response by linking the B cells receptors BCR, inducing antibodies production. No generation of memory B cells occurs. **b.** Protein-polysaccharide conjugate vaccines thanks to the presence of carrier protein elicit the production of memory B cells and the plasma cells. (Picture from Pollard *et al*, 2021) [14]

Chapter 2 – *Neisseria gonorrhoeae*

The Neisseriaceae family consists of *Neisseria* genus as well as the heterogenous genera like *Kingella*, *Alysiella* and *Eikenella* [21]. The genus Neisseria is comprised of Gram-negative bacteria. To date, about 30 Neisseria species have been reported and they include species that are both commensal to human *e.g.*, *N. cinerea*, *N. mucosa*, *N. subflava*, *N. lactamica* and pathogenic *e.g.*, *Neisseria meningitidis* and *N. gonorrhoeae*. [22]

N. gonorrhoeae (gonococcus), discovered by Albert L. Neisser in 1879 in gonorrhea pus from male and female patients [23], is an obligate human pathogen [24] and the causative agent of gonorrhoeae. Gonorrhea is one of sexually transmitted infections (STI) and constitutes one of the major global public health concerns due to its increasing worldwide incidence coupled with the absence of an appropriate treatment. In 2016, World Health Organization (WHO) estimated 376.4 million incident global cases of STI-affected, of which 86.9 million were gonorrhoeae cases among 15-49 years of age [25]. In addition, several *N. gonorrhoeae* strains have developed antimicrobial resistance (AMR) becoming resistant to most available antibiotics recommended for *Ng* treatments [25]. Given the spread of the AMR and the gonorrhea growth worldwide, the need of a gonococcal vaccine is becoming increasingly urgent.

2.1 Gonococcal infection

N. gonorrhoeae primarily colonizes the mucosal epithelium of the male urethra and female endo/ectocervix, causing urethritis in men and cervicitis in women [26]. In women, gonococcal infection is mostly asymptomatic [27], while the majority of men are symptomatic [28]. The reason of this difference lies in the fact that symptomatology in men is easier to diagnose [29] and more specific (purulent exudate from the penis and painful urination). Due to the asymptomatic nature,

especially in women, the untreated gonococcus infection can result in complications and sequelae. Ascending infection can cause chronic complications, as well as pelvic inflammatory disease (PID), chronic pelvic pain, ectopic pregnancy and infertility [26] [27] [30]; gonococcal infections can also cause complications during pregnancy and during birth. Infected women, indeed, can transmit the pathology to the children. Gonorrhoea also facilitates the transmission and acquisition of HIV [25] [24] [31] and other STI. *Neisseria gonorrhoeae* infection consists in several steps: adherence (I), colonization (II), invasion (III), release of peptidoglycan, lipooligosaccharide (see paragraph 1.4.2) and OMVs (IV) and signalling (IV).

2.1.1 Bacterial adhesion, colonization, and invasion

The primary event to establish the infection resides in bacterium adherence to the mucosa epithelium (step I) and is mediated by type IV pili. Type IV pili, present in bacterial outer membrane, are an essential virulence factor of gonococcus which is responsible not only for host tissues adhesion, but also for autoagglutination, twitching motility [32], and immune evasion through *antigenic variation* [30, 32]. Antigenic variation or antigenic alteration is the mechanism by which a microorganism such as bacterium or virus alters the proteins or carbohydrates on its surface [33]. As a result of their surface structural heterogeneity, pathogens can evade the host immune response and are able to re-infect the same host. Together with type IV pili, other surface structures of gonococcus display the antigenic variation, *e.g.* opacity (Opa) proteins and lipooligosaccharides (LOSs). Opa proteins bind to a class of surface-exposed receptors called human carcinoembryonic antigen-related cell adhesion molecule (CEACAM). The interaction between Opa proteins and CEACAM promote the adhesion [34, 35] and the microcolony formation [30]. After adhesion to the mucosal epithelium, *Neisseria gonorrhoeae* forms microcolonies competing with resident microbiota (microorganisms that reside under the

superficial cells of the *stratum corneum*) [36] for nutrients acquisition [37-39]. In particular, iron, zinc and manganese are essential nutrients for several bacteria including gonococcus [40]. After proliferating, the bacterium penetrates the host cell tissue via a mechanism known as transcytosis [41]. During this phase, *N. gonorrhoeae* can also release fragments of peptidoglycans, LOS and outer membrane vesicles (OMVs) causing the activation of the host innate immune sensors: Toll-like receptors (TRL2 and TRL4) and nucleotide-binding oligomerization domain-containing protein 1. As response to bacterial stimulation, macrophages and dendritic cells release cytokines and chemokines (IL-1, IL-6, IL-8, IL-17 and TNF, figure 2.1). These immune effectors cause an influx of neutrophils at the site of infections causing an inflammatory damage, that manifests as a purulent exudate characteristic of symptomatic urethral gonorrhoea. It is unclear whether no neutrophil influx occurs during asymptomatic infection or whether this influx is not sufficient to produce observable symptoms. Nevertheless, *Neisseria gonorrhoeae* is able to evade mechanism on macrophages by modulating the apoptosis, the production of cytokine and polarizing macrophages to suppress T cell proliferation [42].

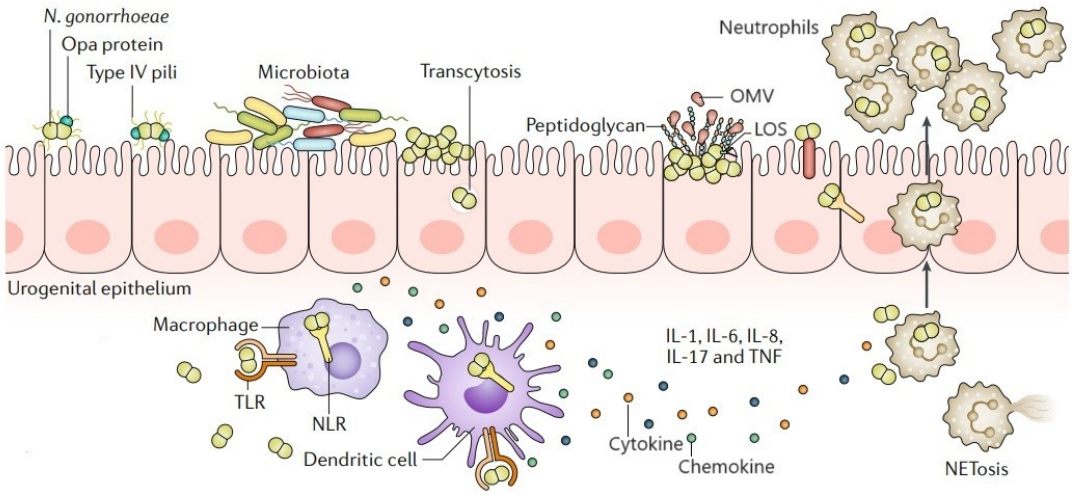


Figure 2.1. *Neisseria gonorrhoeae* infection and pathogenesis steps. Bacterial adhesion to the epithelium firstly requires type IV pili and then Opa proteins. Opa proteins is crucial for a more intimate adhesion and for bacterial proliferation. On the host epithelium *N. gonorrhoeae* forms microcolony competing with resident microbiota for nutrients acquisition. The release of peptidoglycan fragments, OMV and LOS active the Toll-like receptors (TRLs) and NOD. NOD and TRLs signalling active cytokines and chemokines and tumour necrosis factor (TNF). However, bacteria can survive phagocytosis or NETosis (the cell death mediated by neutrophil extracellular traps, NETs) (Figure adapted from Unemo M. *at al.*, 2019) [30].

2.1.2 Complement system bypassing by *Neisseria gonorrhoeae*

The *classical* and *alternative complement* pathways are two of the three distinct pathways of the innate immune system through which complement can be activated to kill microorganisms [43]. It is known that *Neisseria gonorrhoeae* has developed the ability to escape recognition and attack from human complement system. *N. gonorrhoeae* can prevent opsonization (immune process by which a pathogen is marked for phagocytosis [44]) and killing through two different mechanisms. The first mechanism involves the complement cascade components inactivation and the inhibition of membrane attack complex formation through the binding of complement proteins, such as C3b. Lipidic moiety of gonococcal LOS binds C3b and promotes its inactivation to iC3b through factor I, turning off the complement cascade [45]. In the second mechanism, *N. gonorrhoeae* evades complement activation by presenting itself as *self* by expressing molecules found in the host on its surface. For example, the bacterium hides itself by binding the alternative complement pathway regulator *factor H* through sialylated LOS. Furthermore, outer-membrane-localized porin (PorB) can bind factor H and C4b-binding protein (C4BP) blocking the bacterial recognition from complement (Figure 2.2).

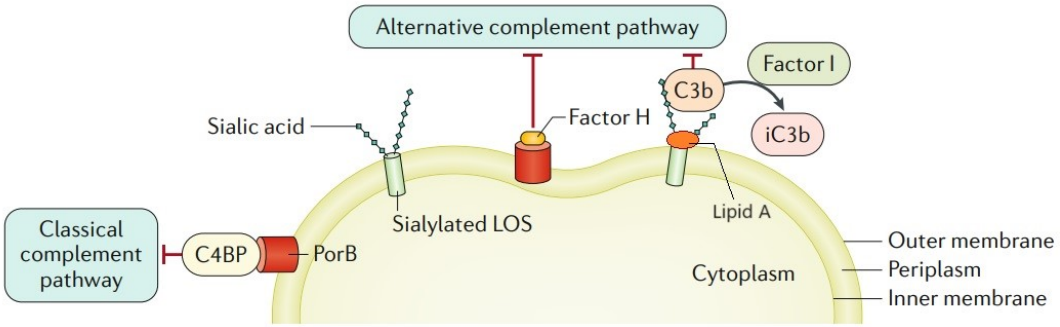


Figure 2.2. Modulation of the innate immune system during gonococcal infection. *N. gonorrhoeae* has developed the ability to prevent complement activation, opsonization and bacterial killing. *N. gonorrhoeae* binds host complement factor H and C4b-binding protein (C4BP) hiding itself from complement recognition and therefore blocks the alternative and classical complement pathways. *N. gonorrhoeae* can also inactivate the alternative complement system by binding the C3b receptor through lipid A on its LOS (Figure adapted from Unemo M. *at al.*, 2019) [30].

2.1.3 Adaptative immunity

Neisseria gonorrhoeae does not induce adaptative immunity and it is known that individuals who contract gonorrhoea do not develop *immunological memory* and consequently re-infection can occur. To evade the adaptative system, gonococcus has acquired multiple mechanisms, including the ability to manipulate multiple immune signalling pathways:

i. As previously discussed, *N. gonorrhoeae* displays antigenic and phase variation in surface-exposed factors, *e.g.* type IV pili, Opa proteins, and LOS [46].

ii. LOS structures mimic the host molecules and present themselves as self-molecules escaping the immune response.

iii. Murine studies showed that *Neisseria gonorrhoeae* suppresses the adaptative immune response through a selective inhibition of Type 1 and Type 2 T helper (Th1 and Th2) cells [47].

iv. *N. gonorrhoeae*-exposed dendritic cells fail to elicit antigen-induced CD4+ T cell proliferation [48].

Gonococcal factors involved in host immune suppression may be useful targets in developing vaccines that induce protective adaptive immune responses to this pathogen [48].

2.2 Gonococcal lipooligosaccharide

Lipooligosaccharide (LOS) is a major glycolipid component of the outer membrane of Gram-negative bacteria including members of the genera *Neisseria*, *Haemophilus*, *Bordetella*, and *Branhamella* [49]. LOS is analogous to the lipopolysaccharide (LPS) found in other Gram-negative families. LOSs displays similar structures and functional activities as LPS. It is composed by a lipid A moiety and oligosaccharide portion (*inner* and *outer core* oligosaccharide) and differs from the LPS structure due to the absence of the O-polysaccharide repeating units (*O-antigen*) (Figure 2.3).

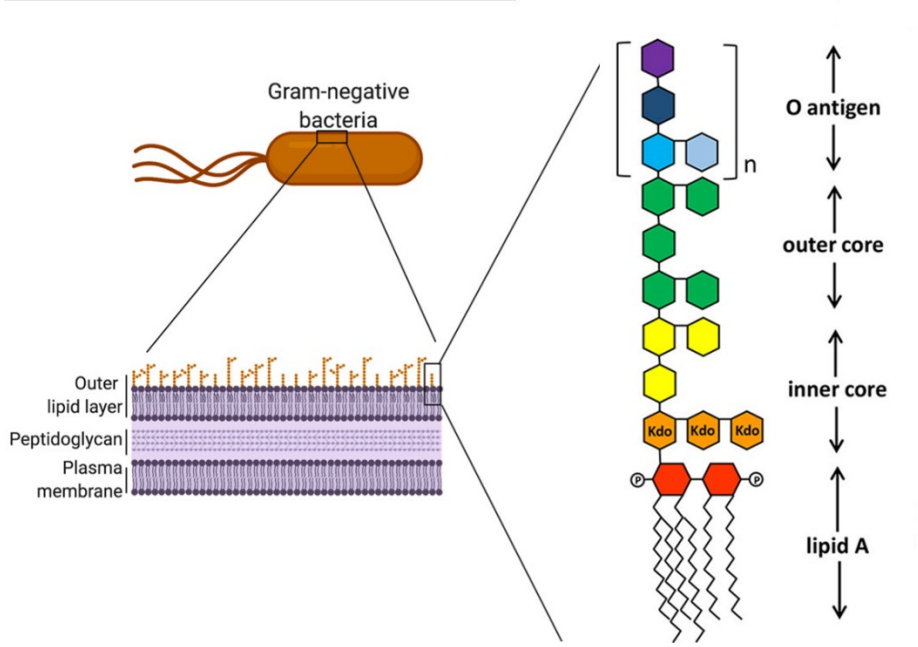


Figure 2.3. General structure of lipopolysaccharide and lipooligosaccharide from a Gram-negative bacterium. LPS is an amphipathic molecule with a general structure consisting of three different regions: hydrophobic lipid A, core polysaccharide, and O-antigen (repeats of polysaccharide chain, where n can be up to 40 repeats). Lipid A consist of bis-phosphorylated di-glucosamine backbone substituted with a variable number of acyl chains that are attached by ester or amide linkage (Figure adapted from Mazgaeen L, Gurung P., 2020) [50].

Neisseria gonorrhoeae contains lipooligosaccharide (LOS) as a major constituent within the outer membrane. The LOS plays a key role in the pathogenesis of Gram-negative bacteria infection by inducing host inflammatory responses and in the evasion mechanism of host innate immunity through sialylation [51]. Gonococcal LOSs are potential vaccine candidates because they are densely represented on the bacterial surface and readily accessible as targets of adaptative immunity [52]. In particular, LOS-derived oligosaccharides (OSs) are the major immune targets, although they undergo phase variation.

In general, gonococcal LOSs consist of three oligosaccharide chains attached to the lipid A portion. In details, two 2-keto-3-deoxy-mannooctolosonic acid (KDO) residues connect two heptose (Hep[I] and Hep[II]) residues to the Lipid A. The first OS chain elongates from the first heptose, Hep[I] and is also known as *α-chain*, while the second and third OS chains are linked to the second heptose, Hep[II] and are named *β*- and *γ-chain* (Figure 2.4).

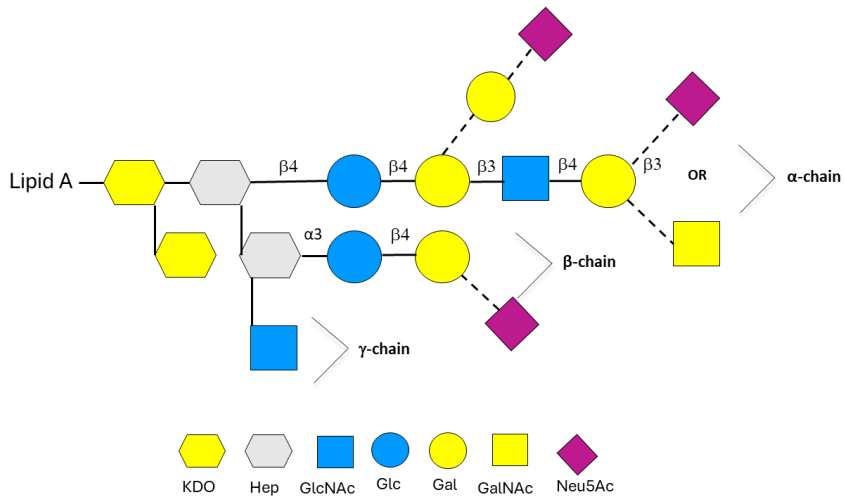


Figure 2.4. Schematic representation of *N. gonorrhoeae* LOS structure(s). Gonococcal LOS consists of lipid A portion and of three oligosaccharide (OS) chains. The OS chains elongate from two heptose residues attached to Lipid A via two 2-keto-3-deoxy-mannooctulosonic acid (KDO) molecules. One OS chain extends from the first heptose (Hep[I]) and two chains are linked to the second heptose (Hep[II]). All saccharide chains can express a sialic acid (Neu5Ac) as terminal residue.

The number of branches and the length of the oligosaccharides can vary between gonococcal strains, and in the same strain during growth *in vitro* and *in vivo* [52]. The genes responsible for LOS biosynthesis have been characterized. In details, the *rfaC* and *rfaF* genes encode for two different heptosyltransferases. Moreover that are

responsible for synthetising the core heptoses units. The *rfaC* gene links the Hep[I] to KDO and the *rfaF* adds the Hep[II] to Hep[I] [53]. The *rfaK* gene product is required for Hep[II] elongation with terminal GlcNAc (Figure 2.5). The expression of different LOS structures on the gonococcal surface is controlled by the phase variable expression of the LOS glycosyltransferases *lgt* genes [54]. In particular, glycan extensions from the core heptoses are modulated by genes *lgtF, lgtE, lgtA, lgtB and lgtD.* These genes are required for stepwise addition of each hexose, or hexosamine in the case of *lgtD*, to synthetize the α-chain [55, 56]. Instead, *lgt* gene encodes for α-galactosyl transferase responsible for adding a terminal galactose to the lactose linked to Hep[I], making an alternative α-chain called P^K-like LOS structure [57]. The *lgtG* gene product is required to add the first glucose unit to the Hep[II], to form the β-chain. The β-chain can be composed by a single monosaccharide or by a lactose unit. The latter is synthetized by *lgtE* gene. *N. gonorrhoeae* can synthetize a variable array of LOS structures by adding a glycose unit on the conserved core trisaccharide (Glc- NAc-Hep[II]-Hep[I]) [58] to express different elongation from Hep[I] only, from both Hep[I] and Hep[II], or occasionally from Hep[II] only [59].

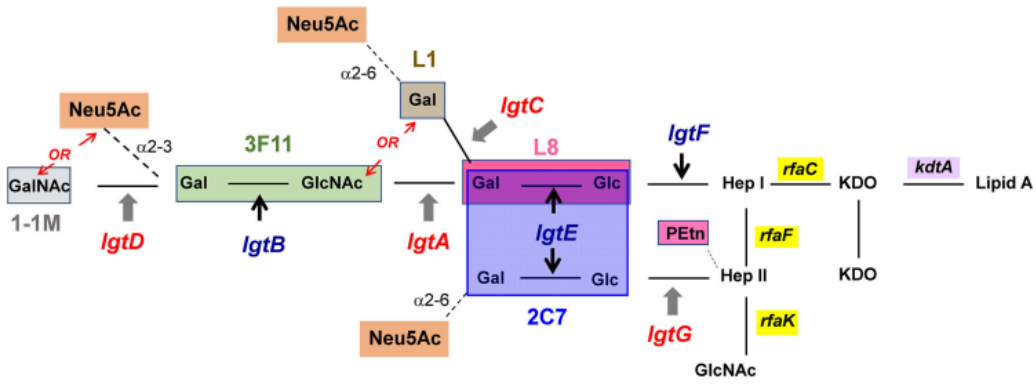


Figure 2.5. General structure of gonococcal LOS and LOS epitopes. In the boxes are highlighted the LOS epitopes which are defined by MAbs. Blue box reports the minimal structure for MAb 2C7 binding. MAb L8 epitope, in pink, in addition to the two hexoses

(Gal, Glc), it also requires the 3-phosphoethanolamine on Hep[II]. MAb 3F11 epitope, in green, is composed by GlcNAc and Gal of Hep[I] chain elongation. The extension (GalNAc or Neu5Ac) abrogates binding. MAbs L1 and 1-1M recognize respectively the Gal sugar in the brown box and GalNAc in the grey box [52].

Among the *lgt* genes mentioned above, four involve gonococcal phase variation mechanism: *lgtG*, *lgtA, lgtC*, and *lgtD* [54]. The genes *lgtA, lgtC, and lgtD* contain homopolymeric tracts of guanine poly (G), and the *lgtG* gene, holds a poly cytosine (C) tract [54, 56, 60]. During DNA replication, a slipped strand mispairing can lead to alteration in coding sequences resulting in truncated LOS structures. This mechanism provides a widespread LOS heterogeneity. Among the different possible predominant glycoforms, some share structure with human glycosphingolipids (GSLs) [57, 59, 61]. The lacto-N-neotetraose that consists of four sugars extending from Hep[I], Galβ1-4GlclNAcb1-3Galβ1-4Glcβ1-4-, displays the same structure as human erythrocyte GSLs. The alternative α-chain, P^K -like LOS structure (Galα1-4Galβ1-4Glcβ1-4-), is similar to the human paraglobosides. Moreover, the fully extended α-chain from Hep[I], composed by five sugar units GalNAcβ1-3Galβ1-4GlclNAcβ1-3Galβ1-4Glcβ1-4- shares the identical structure of human asialo-G3 ganglioside. This structural similarity between gonococcal LOS and human molecules avoids the bacterial immune recognition and the evasion of the immune response. Therefore, these antigenic determinants may not elicit an organism-specific response and consequently their role as possible vaccine antigens has been questioned. However, two epitopes of *N. gonorrhoeae* LOSs do not cross react with human GSL antigens and they could be investigated as potential vaccine antigens. The first epitope is recognized by the anti-meningococcal MAb L8 [62] and consists of lactose (Galβ1-4Glc), β-(1→4) linked to Hep[I] [63, 64] and of phosphoethanolamine (PEtn) substitution at the 3- cyclic position on Hep[II] [65]

(Figure 1.9). The second epitope comprises the two lactose (Galβ1-4Glc) elongation, one β-linked to Hep[I] and the other α-linked to Hep[II] (Figure 1.9). This sugar composition represents the minimal epitope required for binding of MAb 2C7 [59] and it is known as 2C7 epitope.

2.3 The promising epitope of MAb 2C7

During the last decades, MAb 2C7 epitope has been widely explored due to its interesting and promising immunotherapeutic applications [66] and for its key role in virulence. Previous studies, performed by using the truncated LOS, have reported that the lack of lactose linked to Hep[II] (and therefore the loss of the 2C7 epitope) results in an attenuation of gonococcal infection in the experimental mouse cervico/vaginal colonization model [60, 65, 66]. This result suggests that 2C7 expression may be an important virulence factor necessary for both survival and productive infection in humans [52]. The synthesis of gonococcal 2C7 epitope involves *lgtF* and *lgtE* genes. The addition of α-linked glucose unit to Hep[II] is mediated by the phase variable gene *lgtG.* Despite phase variable gene *lgtG,* 2C7 epitope is highly conserved and wildly expressed by 94% of gonococci that reside in the human genital tract (*in vivo*) and by 95% of first passaged isolates [56, 65-67].

Among the different LOS structures able to bind MAb 2C7, some have been extensively studied for their immunogenic activities. All these structures have the same oligosaccharide core in common (two lactose attached to the Hep[I] and Hep[II]) (Figure 2.6). Especially interesting is the structure 5-Hex/G+ (Figure 9) which displays the pentasaccharide **GalNAc-LNnT** (GalNAc-Gal-GlcNAc-Gal-Glc) linked to the Hep[I]. Terminal GalNAc, indeed, plays a key role in the shifting of the immune response in favour to bacterial survival. It interacts to specific C-type lectin, named macrophage galactose-binding lectin (MGL) modifying the immunity toward T helper cell type 2 (Th2) lineage [68]. In addition, the loss of terminal

GalNAc from pentasaccharide structure (5-Hex/G+) decreases the binding of natural IgM in human serum and increases resistance to complement [68, 69].

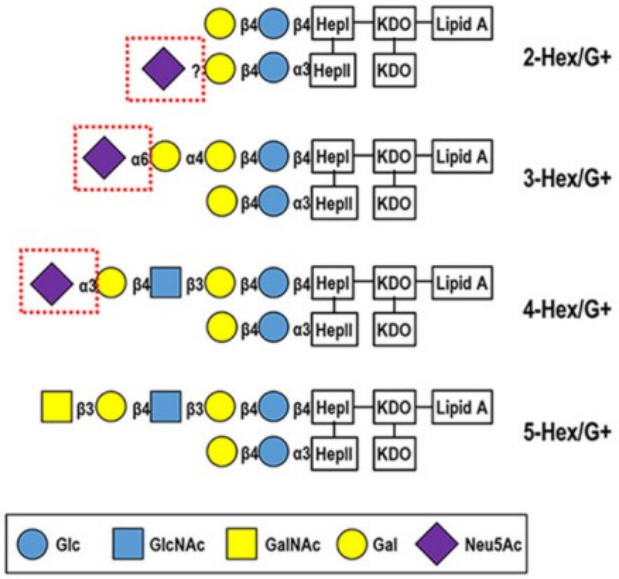


Figure 2.6. LOS glycan extensions from heptose I (Hep[I]) and heptose II (Hep[II]) produced by different strains of *N. gonorrhoeae*. Glycan extensions are shown utilizing the symbol nomenclature for the graphic representation [70]. All the reported structures are recognised and bind by the monoclonal antibodies 2C7. Figure adapted from reference [65].

2.4 The aim of the study

Neisseria gonorrhoeae has been considered a challenge for researchers for decades. This pathogen causes illness in millions of humans and uses an arsenal of antimicrobial resistance strategies and highly variable surface structures that make treatment and prevention a huge challenge. Among the several gonococcus surface molecules that could be an appropriate vaccine targets, the LOS represents an excellent target for a potential protective gonococcal vaccine [71, 72]. Although LOS undergoes phase variations [73], it is an easily accessible target on the gonococcal

surface, where it is present in abundance. Furthermore, while many LOS antigens share structures with human glycosphingolipids [74, 75], making them ineligible for targeting, two distinct LOS epitopes, recognised by MAb L8 and MAb 2C7, [59] do not suffer from this drawback. Recently, Ram *et al* [65] proved that 2C7 epitope, which is widely conserved among gonococci (~95%), is immunogenic during natural infection. The accurate structural characterization of 2C7 epitope and the study of the mechanisms underlying antibody binding are a key point for designing a gonococcal vaccine. The objective of this thesis is to *i.* elucidate the LOSs structures from *Ng* strains and *ii.* investigate the *Ng* LOS-MAb 2C7 interactions at atomic level by means of ligand-based NMR techniques and computational approaches. In particular, the study is focused on two gonococcal strains: FA1090 wild type and 15253.

The work has been structured into two main sections:

1. LOS extraction from *Ng* FA1090 and its chemical characterisation by several techniques (mostly Mass Spectrometry and NMR Spectroscopy) and molecular interaction analysis of the complex FA1090 OS-MAb 2C7 via STD NMR spectroscopy and *in silico* approaches (SECTION II, Chapter 4).
2. Purification and NMR characterization of the LOS from *Ng* 15253 and the investigation of the binding that occurred between 15253 OS and MAb 2C7 (SECTION II, Chapter 5).

Chapter 3 – Tools to disclose ligand-receptor interactions

The glycoconjugates coat, or *glycocalix*, that covers cell surface of microorganisms including bacteria, accomplishes several biological functions such as adherence to host cells and their invasion [76]. Especially lipo- and oligosaccharides, central components of Gram-negative bacteria outer membrane [77], are involved in host-pathogen interactions with the immune system playing a key role in health and disease. These molecules are considered a major antigenic target; therefore the understanding of the molecular basis underlying ligand-macromolecules recognition is pivotal to fully understand their biological functions and to develop new drug design approaches. Interactions between glycans and their receptors are often characterized by rather low binding affinity. This weak binding affinity in combination with the high complexity of carbohydrate structures makes the detection and analysis of these systems complicated especially *in vitro*. Therefore, the combination of orthogonal, complementary techniques, including biophysical, spectroscopic, spectrometric and computational methods, constitutes the best strategy to depict such recognition processes. X-ray crystallography and Nuclear Magnetic Resonance (NMR) together with the emerging cryogenic electron microscopy (cryo-EM) allow not only the detection of the binding but also the characterization of the three-dimensional structure of the interacting systems. In addition, computational techniques as well as homology modelling, Docking and Molecular Dynamics simulations are crucial to confirm, compare and refine the experimental data. The combined use of the above-mentioned techniques allows the description of the binding mode at atomic level and to set up a 3D model of protein-ligand complex.

In the present chapter, the principal techniques required to achieve a detailed description of these complexes are described, starting from the isolation of bacterial glycans to the ligand-based NMR and molecular modelling methods.

3.1 Isolation and purification of glycoconjugates

The first step for studying glycans consists of their isolation from the bacterial biomass through extractive protocols that preserve their chemical structure. Two main extraction protocols are generally employed depending on the nature of the macromolecule to be treated. As for LPSs, which consist of a long polysaccharide chain and are therefore more hydrophilic, the hot phenol-water method is used. While, the LOSs, composed by a shorter saccharide chain than LPS and consequently present a higher hydrophobic behavior, are extracted by using a phenol/chloroform/petroleum ether-based method, named PCP.

The LOS in the present thesis work is firstly treated with a mixture composed of petroleum ether, chloroform and 96% aqueous phenol in a proportion 8:5:2 (v/v/v). Solid phenol is added to the solution, which initially appears opalescent, until a limpid solution is obtained. LOS is then extracted by precipitation from the pure phenolic phase (obtained after removing the other solvents present) through the addition of ultra-pure water droplets. In the second step, the pure phenol phase undergoes to a second extraction with hot phenol-water method. Cells are treated with a mixture of water and phenol at 70°C. After centrifugation three distinct phases are obtained: a lower phenol phase, an upper water phase and an interphase layer. The phenol and water phases are recovered, dialyzed and freeze dried. To remove nucleic acids and proteins the extracted materials are subjected to enzymatic hydrolysis with DNAse, RNAse and protease. The purified sample is analysed by Sodium Dodecyl Sulphate - PolyAcrylamide Gel Electrophoresis (SDS-PAGE) followed by silver nitrate staining. The presence of LOS was confirmed by the presence of bands at the bottom of the gel and the binding with the monoclonal antibody 2C7 was confirmed by Western Blot (WB) experiment.

Prior MS and NMR analysis further purification steps, like size exclusion chromatography, HPLC and Solid Phase Extraction (SPE) are also performed.

3.2 Quali-quantitative composition of monosaccharide residues

Once isolated and purified, the LOS can undergo further treatments and derivatizations in order to characterize its chemical composition. Gas Chromatography (GC) coupled with Mass Spectrometry (MS) is a useful and versatile tool in carbohydrate chemistry [78]. Via GC-MS, the nature of the sugar can be established thanks to specific fragmentation patterns and retention times. GC-MS analysis requires the volatile sample, thus, prior analysis the derivatizations to make it volatile is needed. Several GC-MS experiments exist and depending on the nature of the sample, the most appropriate derivatization should be select. The monosaccharide composition can be detected by means of two distinct protocols named Acetylated Alditols (AA) and Acetylated Methyl Glycosides (AMG).

3.2.1 Acetylated Alditols (AA)

The AA method is suitable for the identification of aldoses, ketoses and aminosugars monosaccharides, but is restricted to neutral or basic species, so it is not possible to detect acid monosaccharides unless the reduction of the carboxyl group is carried out. LPS/OS are firstly hydrolyzed in trifluoroacetic acid (TFA) to cleave the glycosidic linkage. The resulting monosaccharides are reduced with sodium borohydride and the resulting alditols are acetylated using acetic anhydride in pyridine.

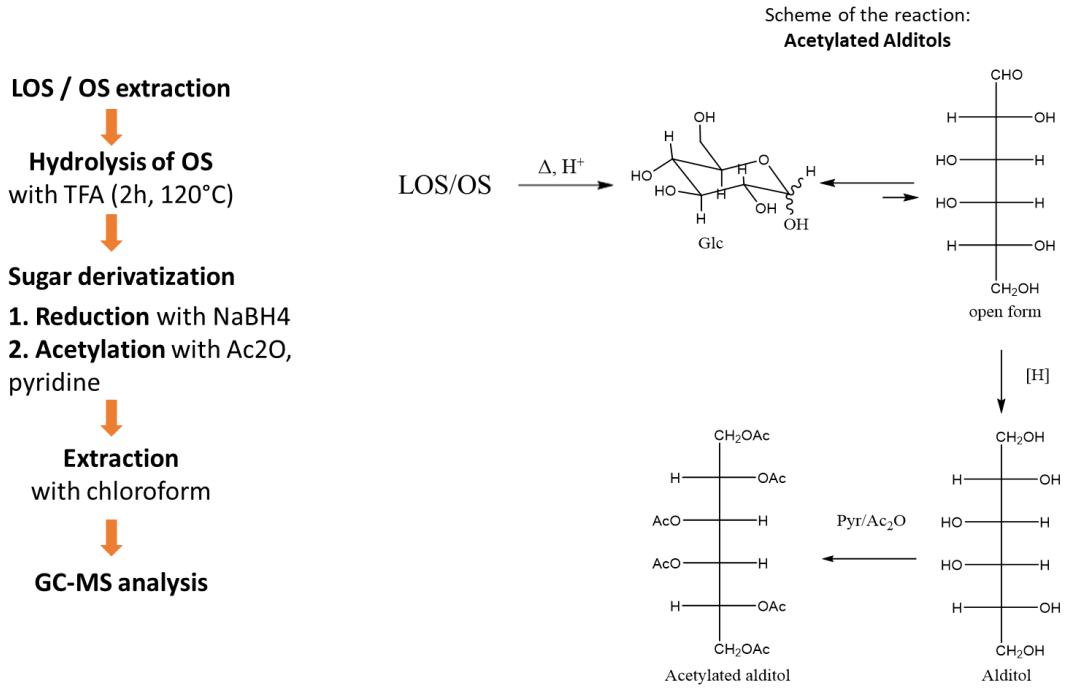


Figure 3.1. Schematic of the reaction that occurs upon derivatisation of glycans into acetylated alditols. Figure adapted from De Castro, Gas Chromatography-Mass Spectrometry for Glycosciences, Glycopedia [79].

One of the major advantages of AA protocol lies in the fact that the aldoses give a single peak, increasing the sensitivity of detection.

3.2.2 Acetylated Methyl Glycosides (AMG)

The AMG method is suitable for detecting almost all types of monosaccharides, including neutral or basic aldoses, uronic and ulosonic acids, but it is not appropriate for recognition of ketoses, like fructose. In the first step the sample is treated with HCl/CH_3OH for hydrolysis and methanolysis in anhydrous condition. In the second step the free hydroxyl groups are acetylated with Ac_2O in pyridine (Figure 3.2). The acetylated methyl glycosides are solubilized in acetone prior injection. The sugar

identification occurs by comparing the fragmentation pattern and retention time of the standards.

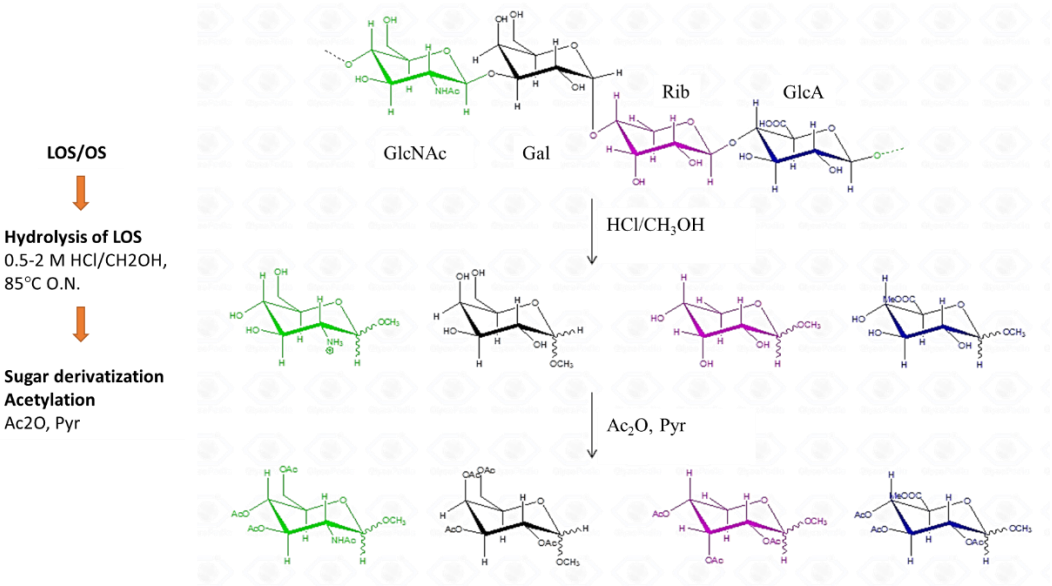


Figure 3.2. Scheme of the derivatization reaction to obtain the acetylated methyl glycosides. Picture adapted from reference [79].

3.3 NMR analysis to solve carbohydrates structure

Nuclear Magnetic Resonance spectroscopy is the elective technique to characterize the carbohydrates in their native state. NMR offers the possibility to study a wide range of molecules in solution and determine not only their structure, but also their conformation, geometry and dynamic. Structural analysis is obtained via combination of homo- and hetero- nuclear mono/ multi-dimensional NMR methods. The assignment of the entire spin systems and the connectivity between monosaccharide units, is obtained from a set of complementary experiments. As a first step, the monosaccharide composition is determined by the evaluation of NMR chemical shifts (δ) and coupling constants (J) in ^{1}H- NMR and ^{13}C-NMR spectra. Some diagnostic ^{1}H and ^{13}C resonances of carbohydrates are listed in Table 3.1.

Table 3.1. Diagnostic ^{1}H and ^{13}C chemical shift values of carbohydrates.

Chemical shift (ppm)	**^{1}H**
8.5 - 7.5	Amide resonances
5.5 - 4.2	Anomeric protons
4.5 - 2.8	Sugar ring protons
2.6 - 1.8	α-methylene protons of deoxy sugars
2.0 - 1.0	Methyl protons of 6-deoxy sugars and of the acetyl groups
2.2 – 0.8	Methyl groups
Chemical shift (ppm)	**^{13}C**
180 - 160	Carbonyl carbons
107 - 95	Anomeric carbons
80 - 60	Sugar ring carbons
60 - 45	Nitrogen bearing carbon signals
~ 30	Aliphatic methylene carbons of deoxy sugars
20 - 17	Methyl carbons of deoxy sugars, acetyl groups

The presence of the signals in the spectral region where anomeric protons resonate (shift range 5.5-4.2 ppm) provides a first indication about the number of different monosaccharide residues. Heteronuclear 2D spectra analysis such as $^{1}H^{13}C$-HSQC can confirm the monosaccharide number anomeric carbon resonances. Moreover, the distinctive values of coupling constants $^{1}J_{C1,H1}$ and $^{3}J_{H1,H2}$ discriminates between the anomeric proton orientation in the sugar units. For example, pyranose rings in *gluco* or *galacto* configuration show $^{3}J_{H1,H2}$ above 8 Hz, which is indicative of **β-configuration** at the anomeric carbon, while $^{3}J_{H1,H2}$ below 3 Hz is indicative of **α-configuration**. *Manno* configured sugars are characterized by $^{3}J_{H1,H2}$ below 2 Hz.

Anomeric configuration can be also evaluated from the magnitude of $^{1}J_{C1,H1}$, as a value under 165 Hz denotes the β-anomer whereas beyond 170 Hz indicates the α-anomeric configuration. A combination of homo- and heteronuclear 2D NMR experiments is required for the unambiguous identification of the sugar sequence, the nature of non-carbohydrate substituents and the conformational features.

In details, COSY (Correlation Spectroscopy) detects the signals of neighbouring protons that are connected by two or three bonds, which share a geminal ($^{2}J_{H-H}$) or vicinal ($^{3}J_{H-H}$) *coupling constant*. TOCSY (Total Correlation Spectroscopy) experiment also measures the protons correlations within the same monosaccharide unit. TOCSY spectrum not only identifies nuclei directly coupled, but also nuclei connected by a chain of couplings, which means that correlations from H1 to H2, H3, H4 or even beyond can be seen. TOCSY experiments provide more information than COSY but they are less specific since the order of correlation is not immediately evident in TOCSY. Therefore, COSY and TOCSY are often used in tandem, as an overlay. This way TOCSY provides signals of the spin system, while COSY brings information on how some of those signals are connected. The degree of information provided by these two experiments is related to the identity of each sugar residue and the mixing time of the NMR experiment. The identity is determined by the axial/equatorial configurations of the ring protons, which in turn determine the coupling constants between ring protons. This means, for example, that the magnetisation transfer in the TOCSY of a β-galactose residue terminates in H4, as equatorial H4 results in an extremely small $J_{H4,H5}$ coupling constant. In contrast, a β-glucose that has exclusively axial ring hydrogens will have a larger magnetisation transfer that will allow correlations up to H6 to be observed. The mixing time is the time during which magnetisation transfer takes place and it is directly proportional to the distance covered by the magnetisation around the ring. Longer mixing times result in correlations through a greater number of bonds. However long mixing times

result in spin relaxation, which causes a decrease in signal intensity. NOE-based experiments, like Nuclear Overhauser Effect spectroscopy (NOESY), display correlations that rely on dipolar couplings through space, meaning that intra-residue correlations between ring protons that are spatially close to each other, such as diaxial H1-H3 or -H5, can be detected. Since NOESY allows to detect inter-residue correlations, it is very useful to determine the conformation of the molecule. An overlay of COSY, TOCSY and NOESY often contains enough information to assign most, if not all, of the spin system cross-peaks. For intermediate sized molecules, such as disaccharides or trisaccharides, the NOE intensity may be close to zero, and ROESY (Rotating frame Overhauser Effect Spectroscopy) experiments, in which homonuclear Nuclear Overhauser effects (NOEs) are measured under spin-locked conditions, can be used in place of NOESY.

Further structural determinants could be deduced from heteronuclear 2D experiments as Heteronuclear Single-Quantum Correlation Spectroscopy (HSQC), which displays the correlations between carbons and the directly linked protons, and Heteronuclear Multiple Bond Correlation (HMBC) that gives signals between 13C and 1H nuclei that are correlated through several bonds. HSQC spectrum gives also information about the presence of a non-carbohydrate substituent, such as methyl, acetyl or phosphate residue. Indeed, the presence of substituents can affect both proton and carbon resonances. For example, in a HSQC spectrum the typical down-field shift of 4–10 ppm of the carbon resonances at the anomeric position (glycosylation shift) can be observed [80]. Also, the acylation sites can be easily located by analyzing the downfield shift of ring proton signals (acylation shift). Figure 3.3 displays some correlations which can be detected by 2D NMR experiments described.

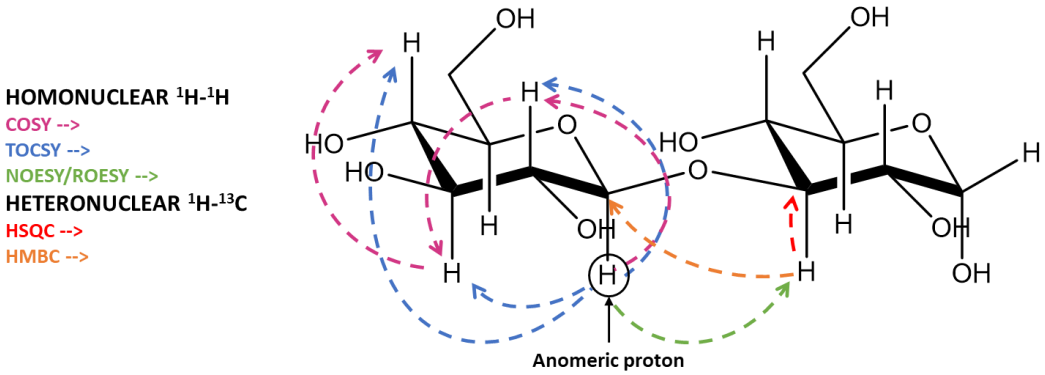


Figure 3.3 Example disaccharide displaying the correlations that would be seen from the anomeric proton using the array of NMR experiments described.

3.4 NMR methods for studying protein-ligand interactions

NMR spectroscopy has a leading role in the drug discovery process [81], mainly because it represents a powerful and versatile tool to investigate proteins-ligands interactions in solution. NMR methods offer important experimental information to detect and characterize binding events and to provide key dynamic information over a wide range of system interactions. Two different approaches can be used to detect the binding events between carbohydrate (or small molecules) and their receptors: the so called *ligand-based* and *receptor-based* approaches. The ligand-based methods show more advantages than the receptor-based ones. First, the ligand-based methods are not limited by the size of the macromolecule, do not require huge amount of receptor (protein concentrations are typically in the micromolar range) and no labelling of the sample is needed. They are particularly useful in the medium-low affinity range, characterized by dissociation binding constant K_D > 100 μM. Among the various ligand-based techniques the tools extensively used in the present Thesis work are the saturation Transfer Difference Spectroscopy (STD) NMR and the transferred NOE (tr-NOE). Deep information can be derived, such as the epitope

map of the ligands involved in the interaction, a quali- and quantitative analysis of the binding and the conformational analysis of the ligand in its free and bound states, named *bioactive conformation*.

3.4.1 NMR ligand-based approach

As discussed above, NMR methodologies are important drug discovery tools widely used to obtain information on the dynamic interactions between small molecules, such as LOS and their macromolecular target, like proteins and antibodies. In particular, ligand-based NMR techniques are very attractive because they are broadly applicable and yield several advantageous: 1) there are no constraints in receptor upper size; 2) labelling of the macromolecule is not required, as only the ligand signals are detected; 3) slight amount of receptor is required (typically in the micromolar range) in the presence of 5–100 fold excess of the ligand [82]. Ligand-based approaches rely on the transfer of NMR attributes from the protein in the bound state to the ligand in the free state. They are very useful in the medium–low affinity range, characterized by dissociation constants (K_D) ≥ 100 μM [83, 84]. Upon binding to a target protein, the molecular tumbling rates of interacting ligands in solution change [85]. Generally, small molecules exhibit slow relaxation rates (T1 and T2), fast diffusion coefficients and usually positive NOE. When bound to a macromolecule, ligands share some of the properties of the much larger receptors, such as the correlation time which characterizes the dynamic properties of the ligand itself. Nuclear spin relaxation time and molecular diffusion coefficient drastically change, as well as NOE. In details, T1 and T2 values increase, molecular diffusion coefficient decreases and 2D-NOESY cross peaks share the same sign of the diagonal peaks, *i.e.* NOE becomes negative. These changes gain access to the ligand-protein interaction by directly observing the NMR signals of the ligand. Most ligand-based NMR experiments detect the binding by exploiting the differential mobility of the

ligand in the free versus bound state, causing change in the diffusion coefficient or by exploiting a ^{1}H magnetization transfer process from the receptor, thus changing in relaxation properties of the ligand, and in the sign of NOE. The process underlying most NMR-based ligand binding screening experiments is described by protein–ligand binding kinetics. In solution, when a protein (P) and a small ligand (L) characterized by mutual affinity are mixed, the time-dependent association between them can be described as [86]:

$$P + L \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} PL \qquad (3.1)$$

where PL represents the protein–ligand complex, k_{on} and k_{off} are the kinetic rate constants that indicate the forward binding and reverse unbinding (or dissociation) reaction, respectively. At equilibrium, the forward binding reaction (P + L → PL) is balanced by the reverse reaction (PL → P + L), as follows:

$$k_{on}\,[P]\,[L] = k_{off}\,[PL] \qquad (3.2)$$

where all molecular species are at equilibrium concentration. The binding affinity (K_b in unit of M−1), is given by the following equation:

$$\mathbf{K_b} = k_{on} / k_{off} = [PL] / [P][L] = 1/ K_{\mathbf{D}} \qquad (3.3)$$

where K_D is the dissociation constant (in unit of M). If the system is in a *slow-exchange* regime, exhibiting a high binding constant, the NMR spectrum will contain two different signals of ligand proton and K_D will be low. On the contrary, if the system is in a *fast-exchange* regime with a low K_b, the spectrum of the complex will

show a single signal of the ligand proton and the K_D value will be high. Systems of interaction in *fast-exchange* regime are characterized by dissociation rate constant k_{off} in the range $1000<k_{off}<100\ 000\ s^{-1}$[84].

3.4.2 NOE and transferred NOE

The Nuclear Overhauser Effect (NOE) is widely used to determinate molecular 3D structure of ligands which bind to proteins in solution. NOE effect results from relaxation processes caused by dipole–dipole interactions between nuclei [87]. Low-medium molecular weight molecules (MW ≤ 2 KDa) exhibit a short correlation time, τ_c, which corresponds to a fast tumbling (random rotation) in solution and positive NOE contacts. High molecular weight molecules, like proteins, exhibit longer correlation times, *i.e.* slow rotation which cause negative NOEs (Figure 3.4). The correlation time is influenced by several factors including temperature (the higher the temperature, the shorter τ_c), solvent viscosity and in solution aggregation [83].

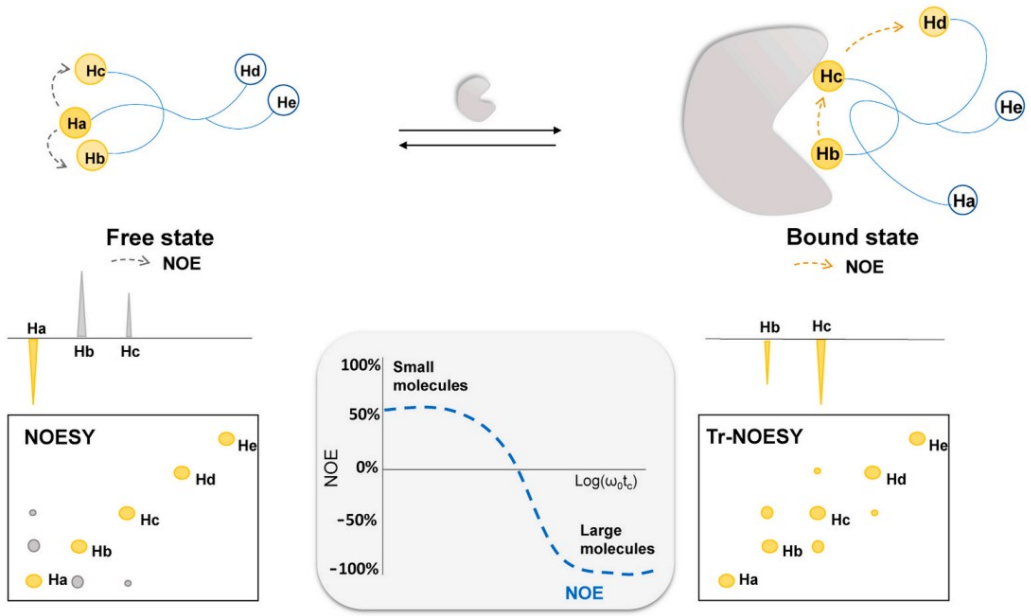


Figure 3.4. Schematic representation of NOE (left) and tr-NOE (right) effects. In the free state, small molecules exhibit positive NOEs. NOESY spectrum displays cross-peaks with opposite sign to the diagonal peaks). In the bound state, small molecules adopt the behaviour of the receptor and show negative NOEs. In the tr-NOESY spectrum of the interacting protein–ligand system in dynamic exchange, the NOE turns negative. The maximum NOE is a function of molecular tumbling rates in logarithmic scale, defined by $\omega 0\tau c$, with $\omega 0$ being the spectrometer observation frequency and τc the rotational correlation time (correlated to the molecular size). Figure from reference [82].

When NOESY sequence is applied to a protein-ligand system in dynamic exchange, the NOE is called transferred-NOE (tr-NOE). Such an experiment provides key information on the ligands' binding mode and allows to determine the bioactive conformation in the bound state. The observation of tr-NOE relies on different tumbling times of free and bound ligand. Indeed, upon binding to a receptor, the ligand acts as part of the macromolecule and adopts the corresponding NOE behaviour. Consequently, the NOE sign undergoes a drastic change leading to the observation of transferred NOEs (Figure 3.4).

3.4.3 Saturation Transfer Difference NMR (STD-NMR) spectroscopy

Saturation Transfer Difference NMR (STD-NMR) spectroscopy represents one of the most powerful and versatile method for detection and characterization of transient ligand-protein interactions in solution [88-90]. STD experiment reveals the presence of the interaction and identifies ligand binding epitope(s) at atomic resolution (*e.g.* small molecules, peptides and oligosaccharides), while interacting with their receptors, such as proteins or nucleic acids [91]. This technique is applicable to solution systems in the fast-exchange regime (small ligands with a medium-weak affinity to the receptors exhibiting a K_D generally in the millimolar to

micromolar range (10^{-6}–10^{-3} M). STD-NMR experiments are based on saturation transfer from the protein to the bound ligand which determines intermolecular NOEs [92, 93]. The nuclear spin energetic levels of the receptor are saturated by a selective radio frequency-pulse train in the range of seconds named *saturation time*. The magnetization, induced by saturation pulse, rapidly spreads over the entire protein through spin diffusion and intramolecular NOEs [94]. Saturation is then transferred from the receptor to the ligand during its residence time in the protein binding pocket via intermolecular saturation transfer and chemical exchange. Ligand protons which are in close contact to the receptor binding site will receive a higher degree of magnetization by intermolecular ^{1}H–^{1}H cross relaxation pathways and, therefore, will generate stronger STD signals. In contrast, ligand protons located far from the binding site will receive little or no saturation (Figure 3.5).

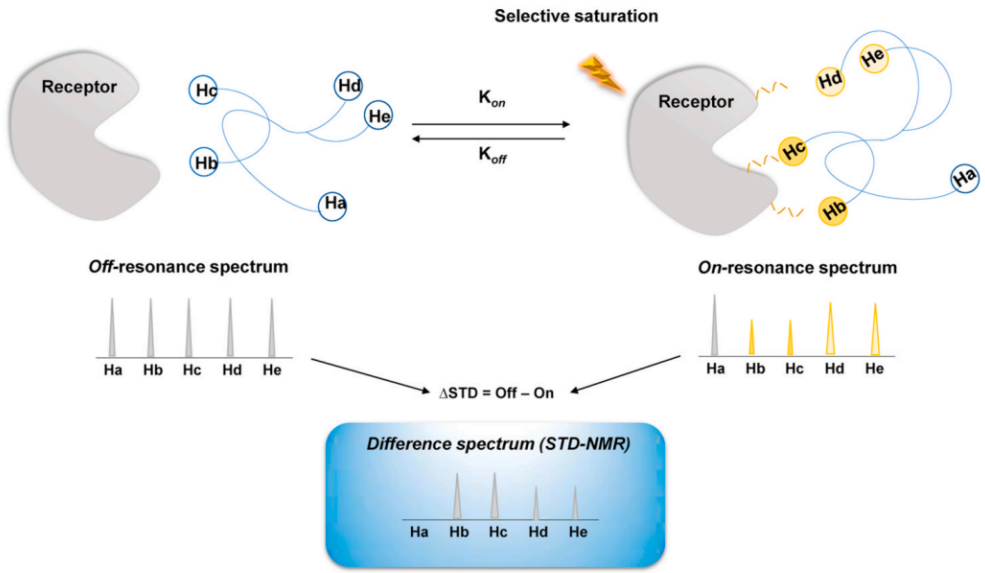


Figure 3.5. Schematic representation of Saturation Transfer Difference (STD)-NMR method. When the receptor and unbound small molecule do not receive receptor saturation (on the left), the *off-resonance* spectrum (recorded by irradiating far from both protein and ligand signals, *i.e.*, at 40 ppm), do not show a decrease in signals intensities (Ha-Hb-Hc-Hd-

He). Conversely, when the receptor is selectively saturated with RF pulses (on the right), the magnetization is transferred from the receptor to the ligand protons (highlighted in yellow in the figure) by intermolecular NOE. Protons in close contact to the binding pocket (Hb and Hc) receive a higher degree of saturation with a following decrease of ligand signal intensities, indicative of the binding. While protons far from the binding pocket (Ha) do not receive magnetization and are not reduced in signal intensities. The STD-NMR spectrum (in the blue panel) is obtained by subtracting the on-from the off-resonance spectrum and displays only the ligand signals directly involved in the binding event. Figure from reference [82].

Two 1D NMR spectra are acquired on a mixture containing a low concentration of protein (nM to μM) and a large molar excess of ligand(s), (typically from 1:10 up to 1:1000). The first is the reference spectrum, named *on-resonance* experiment, where the irradiation occurs in a chemical shift region free from signals of both protein and ligand, typically around 40 ppm. The second is the so called *on-resonance* experiment where only protein signals are selectively irradiated, typically in the aromatic or aliphatic spectral regions (6.5–8.5 ppm and 0 to − 2 ppm). STD-NMR spectrum is given by the difference between the off- and on-resonance experiments and exhibits only the ligand(s) signals that have been saturated by the protein. In the on-resonance experiments, ligand protons that receive magnetization from the receptor, decrease their signal intensities, due to an accumulation of saturated free ligand in the bulk solution. The subtraction of the off- and on-resonance spectra leads to the difference spectrum (STD) which contains only signals of interacting ligand protons. The STD intensities (I_{STD}) are calculated by the following equation:

$$I_{STD} \cdot = (I_0 - I_{sat})/I_0 \quad (3.4)$$

where I_0 is the intensity of a signal in the off-resonance experiment and I_{sat} is the intensity of the signal in the on-resonance experiment. Since the saturation degree of

each ligand proton depends on their proximity to the binding site, it is possible to identify the interacting epitope of the ligand [93]. Furthermore, STD% values are calculated by normalizing the STD intensities obtained from equation (4), against the most intense signal. Ligand protons closer to the receptor pocket will exhibit a higher STD% and *vice versa*.

3.5 Computational approaches to investigate the carbohydrate-protein complexes

Drug design is one of the essential steps in the pharmaceutical industry. Thanks to the rapid development of molecular modelling softwares and algorithms, the drug screening process and design has become much faster [95]. The term *Molecular modelling* encompasses all theoretical methods and computational approaches used to mimic and solved the structure and behaviour of molecules, ranging from small systems (such as carbohydrates) to large biological molecules. In recent years, the application of biomolecular simulations has undergone a rapid development [96-98]. Nowadays molecular modelling methods are used routinely to identify drug binding sites and to understand the molecular mechanisms of biological processes. These methods are also extensively used to investigate the structure, dynamics properties and the carbohydrates ligands in both free and bound states. Depending on the biological processes under investigation, multiscale models are used in biomolecular simulations. In general, to predict the possible binding sites, namely the bound conformations and to deliver a first estimation of the binding energy of each model, docking techniques were used. Molecular mechanics (MM) and molecular dynamics (MD) allows for the assessment of the type and strength of the interactions between ligands and their biological targets. The most common molecular modelling protocol for studying carbohydrates are based on the application of molecular mechanics and molecular dynamics simulations using *ad hoc* force fields. Indeed, research in this

field has been very active in recent years, leading to the development of several force fields capable of accurately reproducing the unique dynamic and electronic properties of oligo/polysaccharides [99-101]. In the following paragraphs, the main approaches used in this study to characterise ligands (LOSs) in both free and bound states will be described.

3.5.1 Molecular Docking

Molecular docking is one of the most common approaches in Structure-Based Drug Design (SBDD) due to its ability to predict, with a considerable degree of accuracy, the possible orientations of small molecules within their appropriate target binding sites [102, 103]. Docking calculations also provide an estimation of the ligand-macromolecule binding affinity and is useful to determine the bioactive conformations of the ligand. Computational docking protocols require structural parameters of both ligand and receptor of interest. Ligands are typically built by using molecular mechanisms methods or directly taken from structural databases. Energy parameters suitable for energy minimization and/or molecular dynamics of protein-carbohydrate complex are available for different force fields. The 3D structures of the receptors are currently obtained by X-ray crystallography, NMR spectroscopy or cryogenic electron microscopy. If the 3D structure of the target has not been elucidated, it can be predicted by homology modelling using templates that share significant sequence similarity. Knowledge about the receptor binding sites is preferable, otherwise it may be assumed by comparison of other related target proteins or through complementary bioinformatic analyses. In the context of protein-ligand recognition and drug design, several docking programs are available including AutoDock [104], AutoDock Vina [105] and GLIDE among others. Although the docking programmes (or servers) operate differently, they share the overall workflow that consists of two main steps: 1. the conformational search predicting possible

ligand conformations; 2. the energy scoring of the resulting binding poses, in which the predicted ligand-receptor complexes are ranked by applying a scoring function that evaluates the docking energy. The main differences between existing docking programmes concern the type of computational search algorithm and the nature of the scoring function applied to classify docked poses [106, 107]. The step that most impacts accuracy and computational cost is the scrutiny for possible conformations and thus the number of rotable ligand bonds. In fact, as the number of rotatable bonds increases, the computational cost increases exponentially. The most common employed search algorithms for conformational screening are shape matching, systematic search and stochastic algorithms. Among the stochastic algorithms, Monte Carlo Simulated Annealing (MC-SA) and Evolutionary Algorithms (EA) are considered the best in term of time-efficiency. They rely on the induction of random ligand changes, *i.e.*, translation, rotations, orientations. In MC methods, the Boltzmann probability function is applied for optimisation, while EA protocol is based on genetics and evolutionary processes of biological evolution systems [108]. EA class includes the popular Genetic Algorithm (GA), the Local Search method (LS) and the hybrid Lamarckian genetic algorithm (LGA) [109]. Regarding the second step, the energy step, the scoring and ranking of the predicted binding modes may be mainly performed by the scoring functions. Scoring function are typically divided into three main classes: *force field-based* (ff), *knowledge-based* (kb) and *empirical* [110]. Force filed-based function is a sum of energy terms of a classical force field (van der Waals and electro-static energies as well as stretching, bending, and torsional energies) which usually considers the interaction energies of the protein-ligand complex and the internal energy of the ligand [111]. Knowledge-based scoring functions extract the binding energy and they are based on the statistical analysis of pairs of interacting atoms from protein-ligand complexes with available three-dimensional structures [112]. Empirical scoring functions simulate

experimental affinity data, combining several weighted empirical energy terms including hydrogen bond and hydrophobic interactions [113]. In the present thesis the AutoDock 4.2 molecular modelling simulation software was used. This software is especially effective for protein-ligand molecular docking and works through an automated procedure by which the grid maps of interaction energy are precalculated. The use of these maps, necessary to evaluate the total binding energy of the complex, significantly reduces the time required for docking calculations. In AutoDock, the conformational search is performed by a Lamarckian genetic algorithm and the docked binding positions are evaluated and classified by a semi-empirical function (equation 2.5), where the free energy of binding is calculated as the difference between the potential energy of the ligand and protein in the complex (bound state) and the potential energy of the ligand and protein in a "free" or unbound state:

$$\Delta G = (V_{bound}^{L-L} - V_{unbound}^{L-L}) + (V_{bound}^{P-P} - V_{unbound}^{P-P}) + (V_{bound}^{P-L} - V_{unbound}^{P-L} + \Delta Sconf) \quad (3.5)$$

Each potential energy contribution is obtained by the sum of van der Waals, hydrogen bonds, electrostatic and solvation terms. The loss of entropy (ΔS_{conf}) associated to the binding is also taken into account. This contribution is strictly related to the number of rotatable bonds (Ntors) in the ligand (Equation 2.6):

$$\Delta S_{conf} = W_{conf} N_{tors} \quad (3.6)$$

3.5.2 Molecular Mechanics and Molecular Dynamics Simulations

Classical-mechanics method known as Molecular mechanics or force-field method is a branch of computational chemistry that uses classical physics to simulate the behaviour of molecules, especially the conformational flexibility. MM is involved in most studies of Computational Structure-Based Drug Design (CSBDD) [114] and allows to predict the structure, properties, and dynamics of molecules, as well as their interactions with other molecules and their environment. This technique is capable of speeding up calculation times and, consequently, of performing analyses of systems containing a significant number of atoms. In MM, molecules are described as a set of bound atoms, whose interactions can be modelled using standard classical physics, which combines electrostatic and mechanical tools (springs, tensors, rotators). Atoms are thus treated as charged spheres, whose size is usually proportional to the van der Waals radius of the atom in question. These spheres are bound by covalent bonds and unbound interactions, which include van der Waals interactions and electrostatics. Atoms and nuclei are assumed to have a static electron distribution and to be connected by elastic (harmonic) forces, according to Hook's law. Atomic interactions, can be modelled with *force fields*, *i.e.* simple parametrized functions which measure the energy of the system. These force fields are typically obtained from *ab initio* or semi-empirical quantum mechanical (QM) calculations or from fitting experimental data. Molecular mechanics expresses the potential energy function as sum of two main terms: bonded and non-bonded potential energy [114]. The bonded potential term describes the covalent contribution to the total energy, coming from bending, stretching energy and dihedral and improper torsions. While, the non-bonded term explains the repulsive and van der Waals interactions and the Coulombic interactions, as described by the following equation (2.7):

$$V = \Sigma V\ stretch + \Sigma V\ bend + \Sigma V\ torsion + \Sigma V\ non\text{-}bond \quad (3.7)$$

First term is related to the bond stretching between directly bonded atoms:

- $V_{stretch} = \frac{1}{2}\, k_s (l - l_0)^2$

(3.8)

where ks is the stretching force constant that is determined empirically. l is the actual bond length in the molecule and l_0 is the "natural" bond length. The second term describe the angle bending between atoms that are geminal to each other:

- $V_{bending} = \frac{1}{2}\, k_b (\theta - \theta_0)^2$

(3.9)

kb is the bending force constant and θ is the actual bond angle in the molecule and θ0 is the "natural" bond angle. The angle-bending term is summed over all bond angles in the molecule. Third sum of the equation 2.7 consists in the torsion energy given by equation 2.10:

- $V_{torsion} = \frac{1}{2}\, V_0 (1 + \cos(n\omega - \gamma))$

(3.10)

Here V_0 is the dihedral constant, n the periodicity parameter, ω the instantaneous dihedral angle and γ the phase term. The last term represents the non-bond interactions, *i.e.*, atoms interacting through Van der Waals attraction, steric repulsion, and electrostatic attraction/repulsion depending on their distance from each other. For two approaching non-bonded atoms, the interaction is attractive (London dispersion force) until the atoms get too close and start to repel each other (van der Waals repulsion/steric strain).

Another very important tool that became a standard approach for the conformational dynamics investigation of the system is Molecular Dynamics (MD) Simulation [115]. MD is widely used to analyse both molecules and complex at different timescales in solution. In the CSBDD MD area is very useful to study the molecular

properties as well as dynamic events, mainly receptor-ligand interactions. An MD simulation is set up by assigning initial rates and positions to all the atoms in a biomolecular system. The forces acting on each atom are calculated and by applying Newton's second law of motion, it is then possible to predict the directions and positions of each atom as a function of time [116]. The process is repeated for defined time intervals and the forces evaluated at each step are used to determine the position and velocity of each atom. This results in a trajectory that describes the configuration of the system at the atomic level at each point during the desired period of time [114]. To run molecular dynamic simulations several programs are nowadays available, such as GROMACS, AMBER, NAMD and CHARMM. Before starting the simulation, the right force field should be selected. The choice depends mainly on the type of system to be studied. AMBER package [117], used in this Thesis , contains force fields optimized for lipids (to date Lipid14) [118], for sugars (GLYCAM06) [100], for proteins, nucleic acids and water molecules (ff14SB) [119] and for organic molecules (GAFF) [120].

MD simulation protocol can be divided into mainly five steps: 1. *system preparation*, 2. *energy minimization*, 3. *heating*, 4. *equilibration* and 5. *production*. First step, system preparation, consists of preparing the system to analyze by setting the initial geometry. The structure can be obtained from both experimental and computational techniques or by combining the two approaches. Ideally the initial configuration should be in the equilibrium configuration, thus reducing the equilibration step. During this phase, the solvent environment is crucial. Different methods can be applied, including *in vacuo* simulation, in implicit and explicit solvent. With explicit water, the simulation takes place into a box of a defined shape and dimension in which solvent molecules surround the simulated molecules and the periodic boundary condition (pbc) may be applied to avoid surface artefacts. To properly refine the system, a neutralization is carried out by adding the appropriate

counterions (Na^+, Cl^-) which will take place at solvent molecules. Second step consists in system energy minimization in which, starting from initial energy state, the search of a minimum in the energy landscape of a system is performed. The minimization is achieved through the steepest descendent method, that executes the geometry optimization until reaching the local minimum [121]. The aim of third step, the heating phase, is to remove the unfavourable contacts, including steric clashes and donor/donor or acceptor/acceptor bonds between solvent and solute by increasing the temperature with a consequent atom rate enhancement. In the equilibration stage, the system is brought to equilibrium keeping under control energy, temperature, pression and volume conditions. These parameters should be allowed to flocculate over the simulation time, but they should not deviate excessively from the starting minimized structure. The parameter that indicates if the system is in equilibrium is the root mean square deviation (RMSD). By plotting the RMDS measures as a function of time, it is possible to estimate if the system is deviating from the original configuration. The final step of MD simulation is the production phase, in which the final trajectory is produced.

In the context of receptor-carbohydrate interactions, MM and MD simulations represent two focal tools for obtaining insights into the conformational behaviour in solution of both glycans alone and in complex with the biological target. Conformational analysis, which is fundamental for determining three-dimensional features, done by means of computational studies has now become routine. The resulting structure can then be compared to information obtained from experimental techniques such as NMR, *e.g.* proton distances derived from NOE, to obtain a reliable 3D topology of the glycans of interest. The joint use of modelling and NMR protocols is becoming very common due to their high complementarity [122]. Indeed, the synergic use of both methods is extremely effective for inferring the

conformational and dynamic properties of free and bound carbohydrate molecules, thus obtaining the 3D complex of the system under study [123].

SECTION II – RESULTS & DISCUSSION

Chapter 4 – Characterization of binding between Ng strain FA1090 wild type and MAb 2C7

4.1 Introduction

Neisseria gonorrhoeae strain FA1090 is, together with FA19, MS11, and F62 [124], one of the main strain widely used for the majority of experimental infection studies conducted over the past 20 years [125]. Ng FA1090 wild type can expresses several LOS structures as a result of the combination of phase variation (the switch between on and off states of gene expression) and antigenic variation (the expression of different antigen versions of a structure)[65]. Despite the phase variation, this strain raised interest due to its ability to bind the MAb 2C7.

The structural elucidation of the OSs expressed by *N. gonorrhoeae* FA1090 wild type is essential to further characterize the binding mode between the bacterial LOS and the MAb 2C7. The detailed analysis of the mechanisms underlying the interaction is pivotal for the design of new targeted therapeutic strategy.

4.2 LOS extraction and purification

Neisseria gonorrhoeae FA1090 LOS was extracted from dried cells (provided by GlaxoSmithKline) following two different protocols: PCP extraction and hot phenol-water (see chapter 7 – experimental section) according to the modified procedure of Westphal[126]. In a first step the PCP procedure was plied and then the hot phenol-water method. The extracted material underwent enzymatic treatment with DNAse, RNAse and proteinase K and extensive dialysis to remove phenol. Both phenol and water phases were freeze-dried and analysed by 12 and 13.5 % SDS-PAGE followed by silver nitrate staining. FA1090 LOS migrated as three distinct bands (Figure 4.1 a). Separation over SDS-PAGE gel followed by silver staining is a highly sensitive method capable to detect as low as 1 ng LPS and is routinely used for visualization

of the band pattern of purified LOS. When immunoblotted using MAb 2C7, all three bands of FA1090 LOS reacted (Figure 4.1 b), proving the presence of the 2C7 epitope in all detected bands.

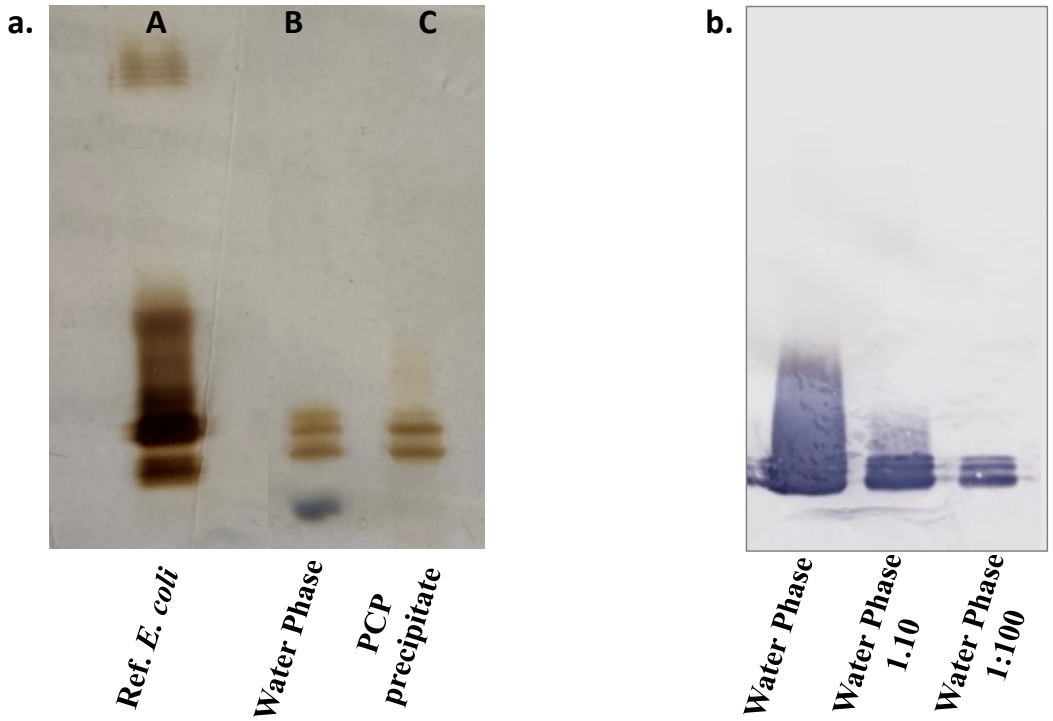


Figure 4.1. **Silver staining and Western Blot of purified *Ng* FA1090 LOS. a.** SDS-PAGE electrophoresis followed by silver staining. Positive control from *E. coli* (lane A) and the LOS isolated by modified hot water-phenol extraction protocol (1ine B) and by PCP extraction method (line C). **b.** Western Blot analysis of FA1090 wt LOS, carried out by using as primary antibody the MAb 2C7 and as a secondary antibody an Anti IgG mouse- whole molecule.

4.3 Chemical characterization

The monosaccharide composition was determined by GC-MS. In details, the sugar analysis was achieved through Acetylated Methyl Glycosides (AMG) derivates and revealed the following sugar units: Galactose (Gal), Glucose (Glc), N-Acetylglucosamine (GlcNAc), Heptose (Hep), 3-deoxy-D-manno-oct-2-ulosonic acid acid (KDO) (Figure 4.2).

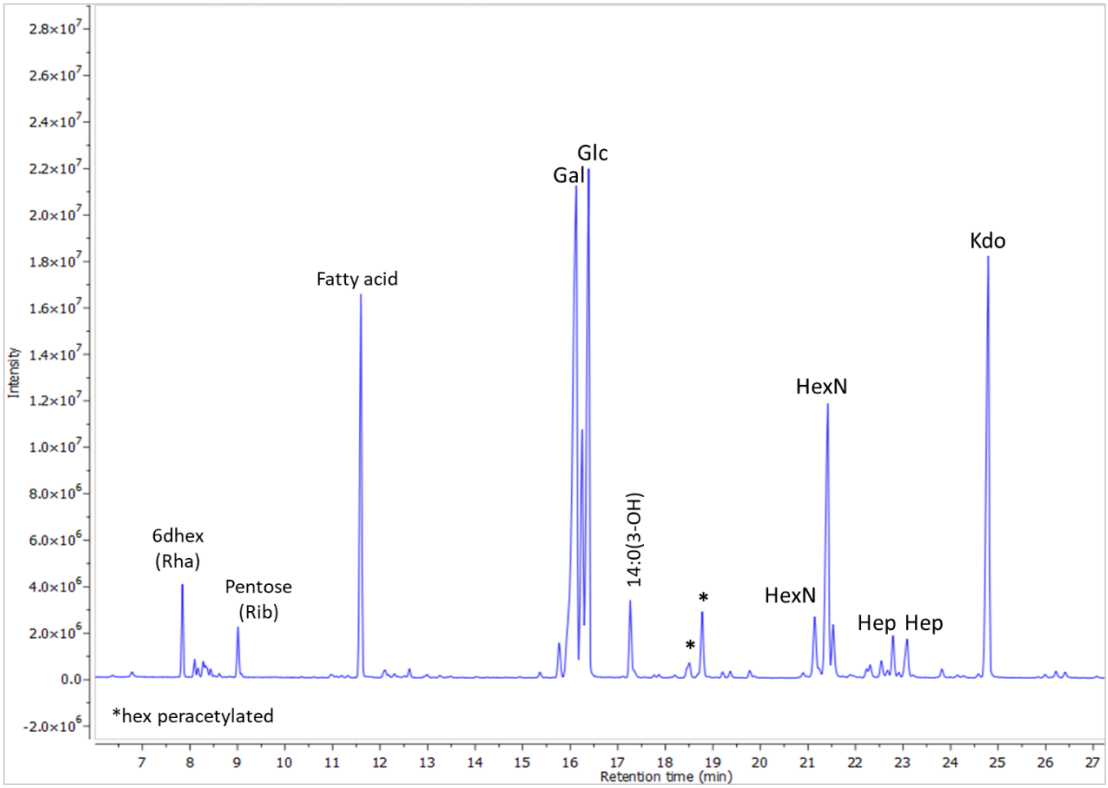


Figure 4.2. AMG compositional analysis of the *N. gonorrhoeae* FA1090 wt LOS. The peaks were assigned by injecting AMG standards. The symbol (*) indicates the hexoses peracetylated.

In order to investigate the relative abundance of the different OS species expressed, the monosaccharide quantification was performed by acetylated alditols (AA)

derivates. Gas-liquid chromatography of acetylated alditols is a technique widely used for determining the monosaccharide composition, especially those resulting from the hydrolysis of poli- and oligosaccharides [127].

Relative quantification of monosaccharides was achieved by built two external calibration curves. Since the four sugars may not respond similarly, due to their different structure, the Gal and Glc monosaccharides concentrations (in nmol) were extrapolated from the external calibration curve obtained by using glucose as standard. While GalNAc and GlcNAc moles were calculated on the curve made by using GlcNAc as external standard. The standards and the oligosaccharide were hydrolysed and derivatised according to the same AA protocol [128].

Calibration curves were performed by plotting the areas of the standards solutions against their relative concentrations (see details in Table 4.1 and Table 4.2).

Table 4.1. Galactose calibration curve parameters: sample concentration for each level, peaks areas and corresponding moles.

Standard Solutions	**Standard conc. (mg/mL)**	**Peak area**	**Standard conc. (nmol)**
1	0.500	118504987.6	**1387.66**
2	0.250	58365013.83	**693.83**
3	0.100	16425351.49	**277.53**
4	0.050	5423827.10	**138.77**
5	0.025	1068065.19	**93.38**

Table 4.2. Parameters for building N-Acetylglucosamine calibration curve: sample concentration for each level, peaks area and corresponding moles.

Standard Solutions	**Standard conc. (mg/mL)**	**Peak area**	**Standard conc. (nmol)**

1	0.500	216844240.6	**1130.00**
2	0.250	94707997.31	**565.10**
3	0.100	41283617.93	**339.00**
4	0.050	20156493.19	**226.00**
5	0.025	5315564.33	**113.00**

After OS derivatization the monosaccharide separation was obtained via GC-MS (Figure 4.3).

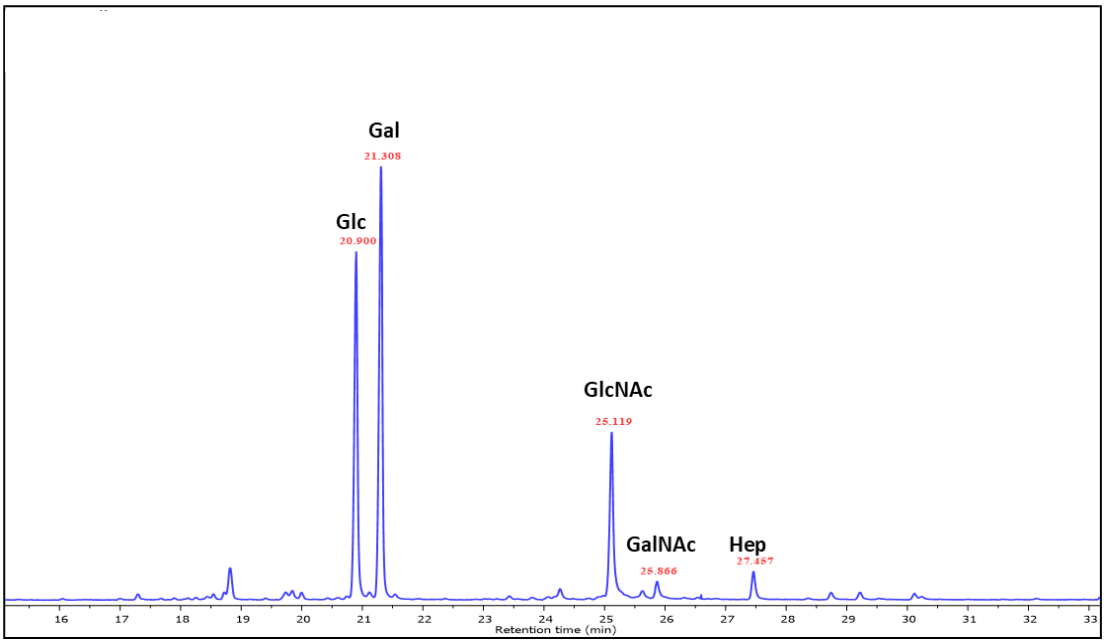


Figure 4.3. Separation of alditol acetates produced by reducing and then acetylating the monosaccharides. The peaks were assigned by comparing the retention time and fragmentation patterns of AA standards.

Monosaccharides peaks areas and the relative concentrations (in nmol) were reported in Table 4.3.

Table 4.3. Retention times, peaks area and moles of acetylated alditols. Molar ratio have been obtained by setting the Gal moles equal to 2 and normalizing all values on the Gal residue.

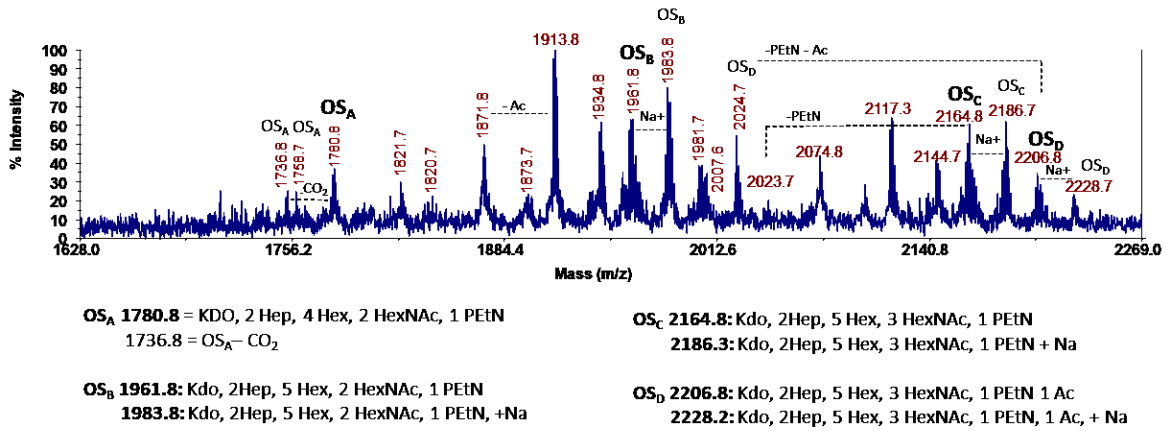
Residue	Retention time (min)	Peak area	Conc. (nmol)	Molar ratio
Glc	20.9	25320886.5	261.3	2
Gal	21.3	32813365.8	330.5	2.5
GlcN	25.1	14305120.0	159.5	1.2
GalN	25.9	1478039.5	40.9	0.3

From the results, it can be assumed that the predominantly abundant species expressed by the *N. gonorrhoeae* FA1090 consists in the smallest structure which is composed by the two **Galβ-(1,4)-Glc** chains from both Hep[I] and Hep[II]. The extended structure which carries the GalNAc as terminal sugar in the α-chain is the less abundant.

4.4 MALDI and ESI-MS characterization

To determine the chemical structure, the LOS was subjected to a mild acid hydrolysis in 1% acetic acid to selectively cleave the acid-labile glycosidic linkage between the KDO and the non-reducing GlcN of the lipid A. The lipid A was obtained in the precipitate and while the oligosaccharide fraction was obtained in the supernatant. As for the supernatant further purification steps were required. Specifically, as first, a size exclusion chromatography with Biogel P-6 resin and then the HPLC purification step with TSKG3000 were performed from which pure oligosaccharides fractions was obtained.

A negative mode MALDI-TOF MS analysis of OS fraction was also performed (Figure 4.4). Briefly, the MALDI-TOF MS spectrum showed four several species named **OS_A**, **OS_B**, **OS_C** and **OS_4**. The lowest mass ion peaks at *m/z* 1780.8 and 1736.8 were interpreted as OS_A made up of KDO, 2 Hep, 4 Hex, 2 HexNAc, 1 PEtN with loss of H_2O and CO_2 respectively. The species at *m/z* 1758.7 was interpretated as the sodium adduct of **OS_A**. Ions observed at *m/z* 1961.8 corresponded to an oligosaccharide species having KDO, 2 Hep, 5 Hex, 2 HexNAc, 1 PEtN assigned as **OS_B**; the ion with *m/z* 2164.8, following this hypothesis, could correspond to a species consisting of KDO, 2 Hep, 5 Hex, 3 HexNAc, 1 PEtN, named **OS_C** and the ion mass at *m/z* 2186.7 was the sodium adduct of OS_C unit. The higher mass fragment ion with *m/z* 2206.8 was in agreement with a oligosaccharide chain composed by KDO, 2 Hep, 5 Hex, 3 HexNAc, 1 PEtN 1 Ac.



OS_A 1780.8 = KDO, 2 Hep, 4 Hex, 2 HexNAc, 1 PEtN
1736.8 = OS_A – CO_2

OS_B 1961.8: Kdo, 2Hep, 5 Hex, 2 HexNAc, 1 PEtN
1983.8: Kdo, 2Hep, 5 Hex, 2 HexNAc, 1 PEtN, +Na

OS_C 2164.8: Kdo, 2Hep, 5 Hex, 3 HexNAc, 1 PEtN
2186.3: Kdo, 2Hep, 5 Hex, 3 HexNAc, 1 PEtN + Na

OS_D 2206.8: Kdo, 2Hep, 5 Hex, 3 HexNAc, 1 PEtN 1 Ac
2228.2: Kdo, 2Hep, 5 Hex, 3 HexNAc, 1 PEtN, 1 Ac, + Na

Figure 4.4. Negative-ion MALDI TOF mass spectrum of oligosaccharide from *N. gonorrhoeae* strain FA1090 wt.

Electrospray ionization Mass Spectrometry (ESI-MS) analysis was carried out to unveil the complete OS structures expressed by *N. gonorrhoeae* FA1090 wt. Mass spectral ionization of oligosaccharide chain, after acetic acid release, has proven

successful for a detailed understanding of the structures. Mass Spectrometry results confirmed the oligosaccharide structure heterogeneity observed from MALDI-TOF analysis (Figure 4.5).

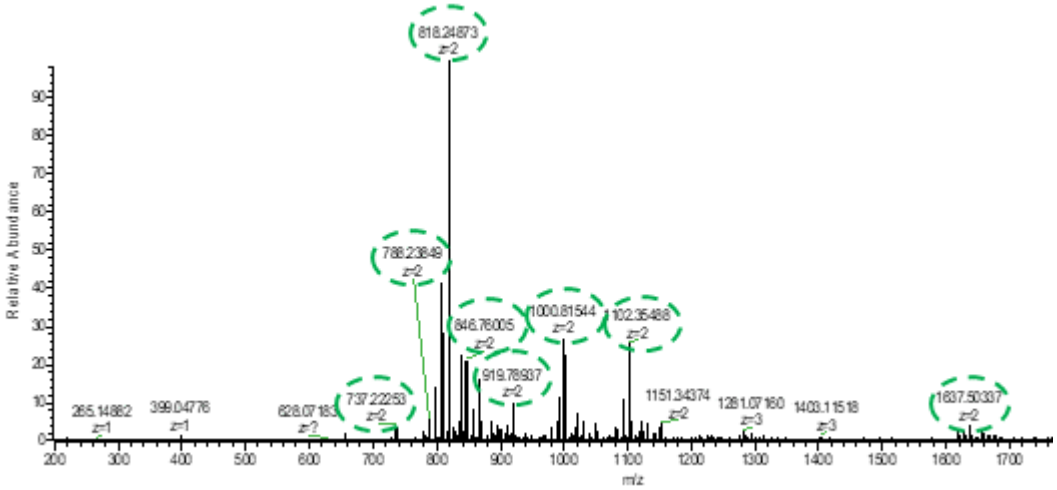


Figure 4.5. ESI-MS mass spectrum of *N. gonorrhoeae* FA1090 OS acquired in negative ion mode. The presence of several ions species, related to the different OS structures, support published data that the *Ng* strain FA1090 expresses multiple LOS structures.

By combining MALDI and ESI-MS data, four different OS structures, named **OS_1**, **OS_2**, **OS_3** and **OS_4**, were characterized. The highest molecular weight ion, *m/z* 1102.35, fits a carbohydrate composition that includes 1KDO, 2Hep, 1PEtN, 3HexNAc, 5Hex and 1Ac (OS_4). The ions with *m/z* 1000.82 and 919.79 differ from the previous species for one HexNAc (OS_3) and for 1Hex and 1HexNAc (OS_2) respectively. An ion composition of 1KDO, 2Hep, 4Hex, 1HexNAc, 1PEtN and 1Ac (OS_1) could account for the mass ion, *m/z* 818.25, which was also the most abundant species (**Figure 4.5**). A loss of acetyl group was observed for ion with *m/z* 797.24 and z=2. Instead, the peak having *m/z* equal to 737.22 could derive from the OS_1 structure which has lost one Hex unit.

Data for the *N. gonorrhoeae* strains FA1090 wt are summarized in Table 4.4.

Table 4.4. Structures and molecular masses of oligosaccharides from *N. gonorrhoeae* FA1090 wt. Calculated masses were based on monoisotopic ions: CH3COOH, 42.0105647; PEA, 141.019097; Hex, 162.052823; Hep, 192.063388; HexNAc, 203.079372, KDO, 220.058303.

OS spp.	Proposed structure	z	Experimental (*m/z*)	Theoretical mass (*m/z*)	Exact mass (Da)	Error %	Error ppm
OS_1	1KDO; 2Hep; 5Hex; 3HexNAc; 1PEtN; 1Ac	2	1102.3488	1102.35121	2206.71697	-0.000333	3.33
OS_2	1KDO; 2Hep; 5Hex; 2HexNAc; 1PEtN; 1Ac	2	1000.81544	1000.81152	2003.63760	-0.000392	3.92
OS_3	1KDO; 2Hep; 4Hex; 2HexNAc; 1PEtN; 1Ac	2	919.78933	919.78511	1638.50540	-0.000459	4.58
OS_4	1KDO; 2Hep; 4Hex; 1HexNAc; 1PEtN; 1Ac	2	818.24868	818.24542	1838.50540	-0.000398	3.98
$OS_{1\text{-}COCH3}$	1KDO; 2Hep; 4Hex; 1HexNAc; 1PEtN	2	797.24372	797.24014	1596.49484	-0.000449	4.49

OS_{1-Hex}	1KDO; 2Hep; 3Hex; 1HexNAc; 1PEtN	2	737.2221	737.21901	1476.45258	-0.000419	4.19

Therefore [59, 60, 65], four oligosaccharide structures were characterized. Figures 5.6, 5.7, 5.8 and 5.9 report the OSs proposed structures of *N. gonorrhoeae* FA1090 wt.

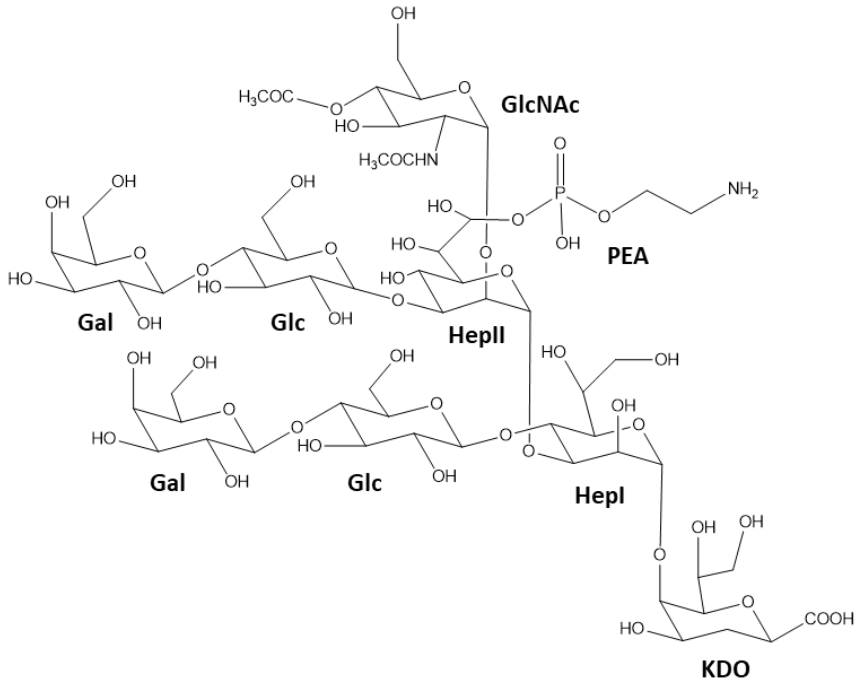


Figure 4.6. OS structure corresponding to *m/z* = 1638.5; OS structure is composed by 1KDO unit, Hep[I], Hep[II], PEtN (Phosphorylethanolamine) on Hep[II]; the α-chain and β-chain which elongate from Hep[I] and Hep[II] respectively, are composed by two lactoses (βGal1-4βGlc).

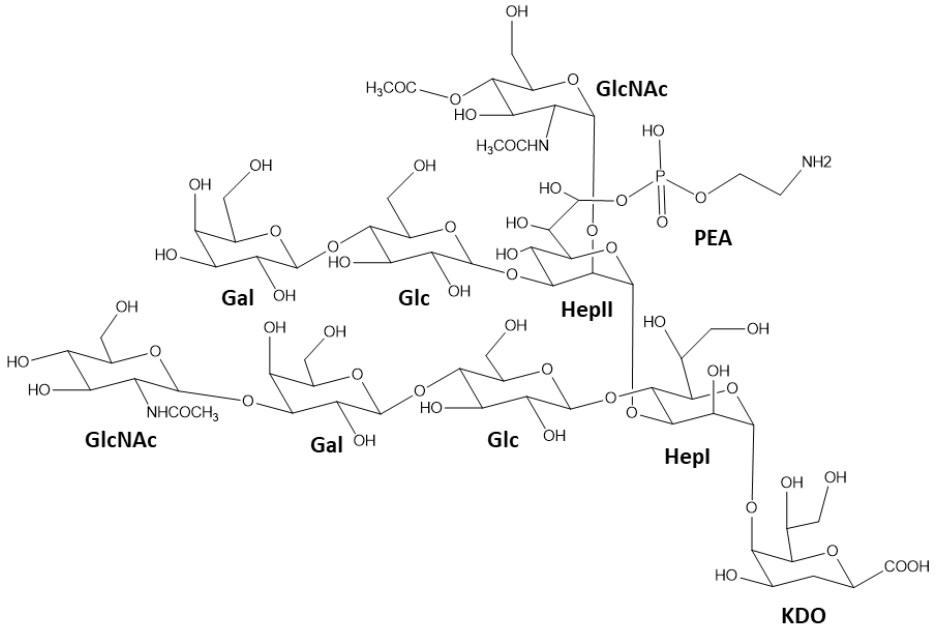


Figure 4.7: **OS structure with *m/z* 1841.6**; OS structure is comprised 1KDO moiety, Hep[I] and Hep[II], PEtN (Phosphorylethanolamine) on the Hep[II]; in addition to the lactose moiety, the Hep[I] branch contains the GlcNAc unit. Otherwise, the chain branched to the Hep[II] displayed a lactose (βGal1-4βGlc).

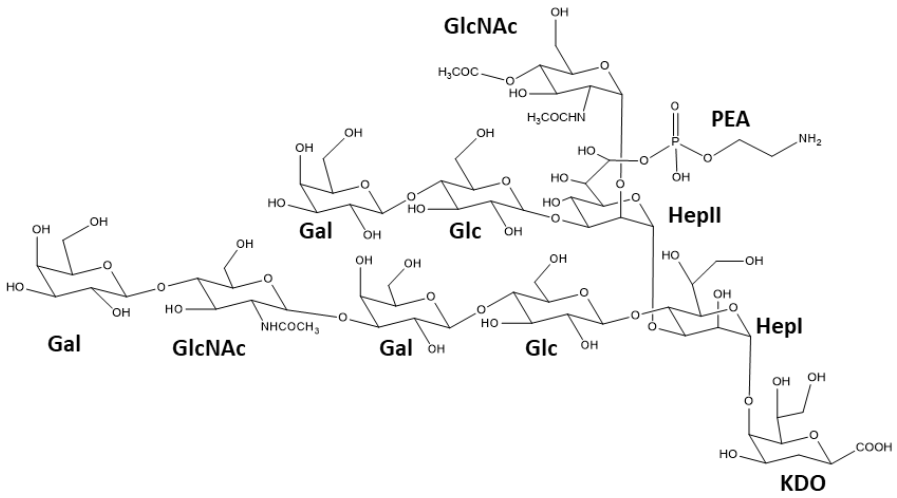


Figure 4.8: OS structure with *m/z* 2003.6; OS structure consists of LNnT structure (Gal-GlcNAc-Gal-Glc) linked to the Hep[I] and of lactose unit is linked to Hep[II] together with GlcNAc. Hep[II] is decorated with PEtN (Phosphorylethanolamine).

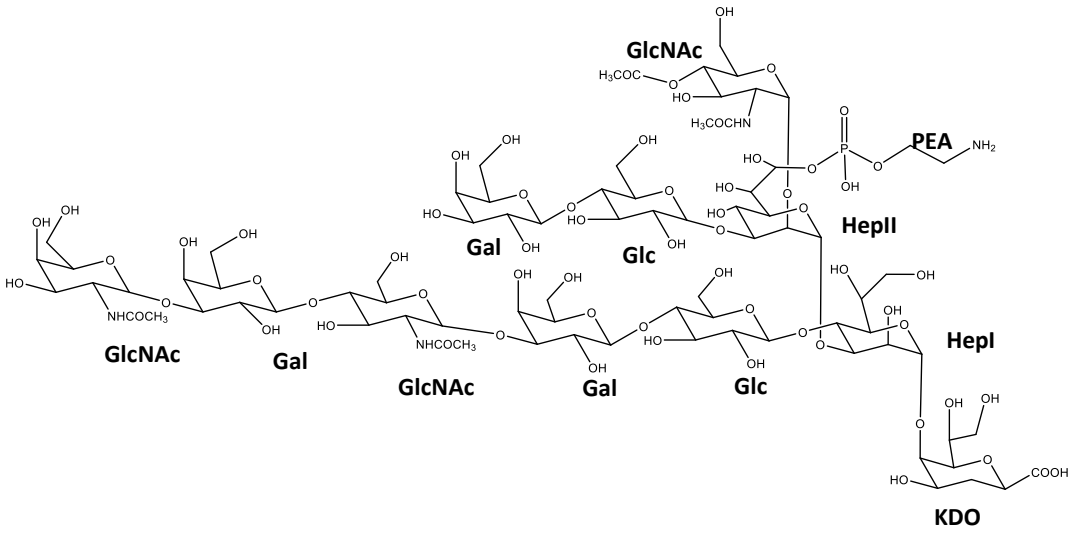


Figure 4.9. OS structure with *m/z* 2206.8; in Hep[I] branch resides the pentasaccharide GalNAc-LNnT (GalNAc-Gal-GlcNAc-Gal-Glc) and the Hep[II] elongation is composed by two sugar units: 1 Gal and 1 Glc. The entire structure is also composed of 1 KDO, 1 GlcNAc residue which constitutes the γ chain and 1PEtN as decoration on Hep[II].

4.5 NMR structural characterization

The *Ng* OS FA1090 wt, obtained after mild acid hydrolysis and purified *via* size exclusion chromatography (Bio-gel P6), was also analysed by means NMR spectroscopy (Figure 4.10, Figure 4.11, and Table 4.5). The full NMR assignment was performed combining homonuclear and heteronuclear NMR experiments. Figure 4.11 reports the ^{1}H-NMR and HSQC spectra. Nine anomeric signals (5.83, 5.32, 5.10, 5.05, 4.56, 4.67, 4.47, 4.41 and 4.38 ppm) were identified and assigned

to Hep[II] (A), αGlc (C), αGlcNAc (B), Hep[I] (H), βGlcNAc (L), βGalNAc (E), βGlc (D), βGal (F) and βGal (G) and βGal (Y) respectively. Sugar residues A, B, H and C resonated above 4.9 ppm and where consistent with **α** configuration, while residues D, E, F, Y, G and L resonated below 4.9 ppm suggesting the **β**-linked hexoses. Indeed, chemical shift values of the anomeric carbons related to A, B, H and C resonated below 101 ppm, in contrast to those of D, E, F, Y, G and L units, which resonated at lower field above 102 ppm [129]. The HSQC spectrum also validated the identification of residues B, L and E as amino sugars since carbons in position 2 resonated at 54.21 ppm, 52.51 ppm and 55.17 ppm [130]. Using 2D correlation experiments, spin systems where fully assigned. 2D NOESY experiments aided in identifying as yet unassigned peaks through *intra*-residue correlations. The C-(1→3)-A connection was confirmed by a correlation between H1 C and H3 A. *Inter*-residue correlations H1 D - H4 H and H1 B to H2 A confirmed the linkages βD-(1→4)-H and αB-(1→2)-A.

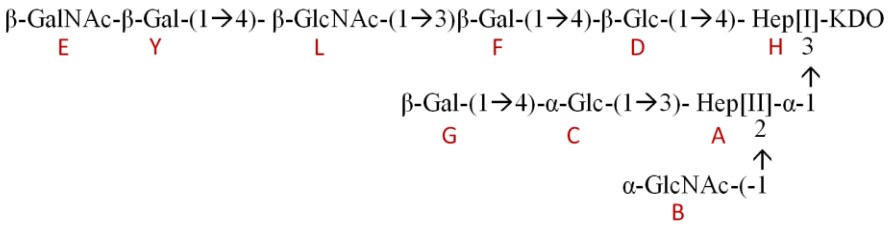


Figure 4.10. The most elongated oligosaccharide structure from *N. gonorrhoeae* FA1090 wt.

Table 4.5. ^{1}H, ^{13}C chemical shift values of the oligosaccharide of *Neisseria gonorrhoeae* FA1090 wt.

	Chemical shift (δ)					
Residue	**1**	**2**	**3**	**4**	**5**	**6**

Hep[II]-**A**	^{1}H	5.83	4.04	4.16	3.64		
	^{13}C	97.79	79.46	47.63	71.51		
Hep[I]-**H**	^{1}H	5.05	4.09	3.96	4.25		
	^{13}C	98.85	70.29	71.74	72.69		
Glc-**C**	^{1}H	5.32	3.58	3.72	3.64	3.79	
	^{13}C	99.97	71.28	71.24	77.86	71.37	
GlcNAc-**B**	^{1}H	5.10	3.83	3.21	3.83	4.26	4.48/4.11
	^{13}C	98.50	54.21	71.02	71.43	70.39	64.02
Glc-**D**	^{1}H	4.47	3.21	3.55	3.44	3.49	3.74/3.67
	^{13}C	102.26	71.01	74.57	79.26	75.26	61.04
Glc- **G**	^{1}H	4.38	3.51	3.65	3.86	3.66	
	^{13}C	102.90	70.99	72.50	68.56	75.36	
Glc - **Y**	^{1}H	4.38	3.52	3.52	3.71/3.88	3.75	
	^{13}C	102.90	70.03	80.15	82.03/67.69	75.04	
Glc – **F**	^{1}H	4.41	3.48	3.60	3.86	3.74	
	^{13}C	102.93	71.02	72.56	68.56	75.35	
GlcNAc- **L**	^{1}H	4.56	3.88	3.68			
	^{13}C	103.38	52.51	70.94			
GalNAc- **E**	^{1}H	4.67	3.74	3.67	3.54		
	^{13}C	102.77	55.17	72.26	74.52		

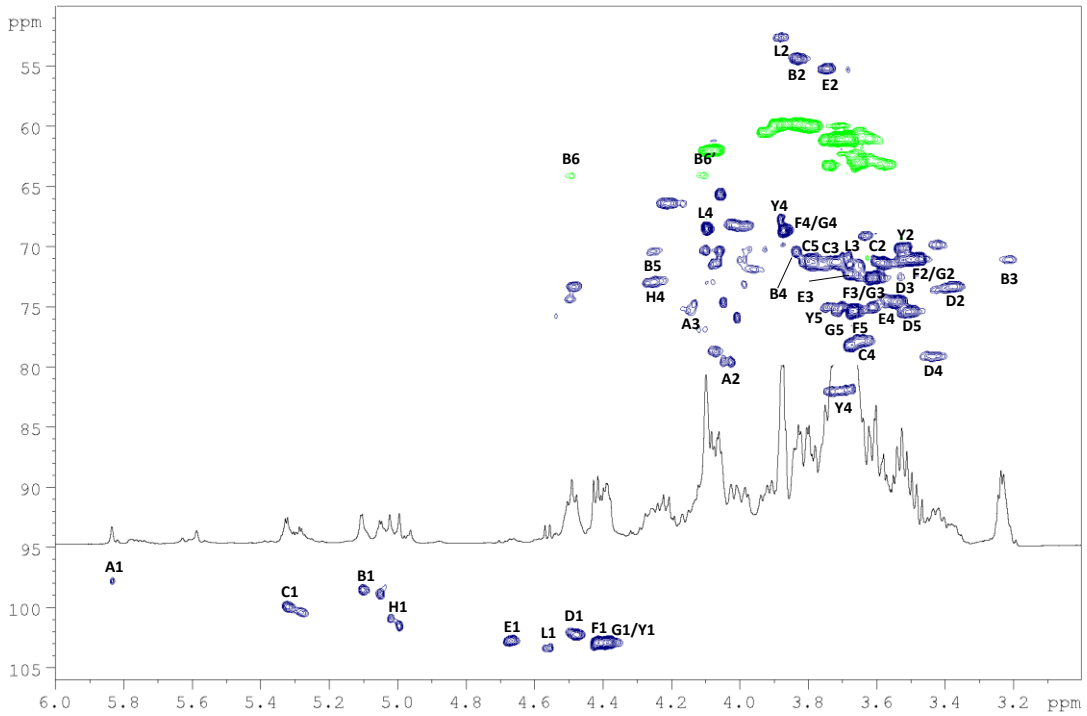


Figure 4.11. $^{1}H,^{13}C$ HSQC spectrum of the oligosaccharide of *N. gonorrhoeae* FA1090 wt.

4.6 Binding analysis between *N. gonorrhoeae* FA1090 OS and MAb 2C7

Ng OS FA190 wt, as previously demonstrated, is able to bind the monoclonal antibody 2C7. Indeed, all the structures expressed by this gonococcal strain consist of the 2C7 core, *i.e.* the minimum epitope required for 2C7 recognition. In order to better determine the epitope map of the oligosaccharide, the binding studies were performed on the complex 2C7/OS FA1090 wt.

4.6.1 STD NMR analysis

The isolated oligosaccharide(s) from *Ng* FA1090 wt was used to perform the binding studies with MAb 2C7. OS FA1090 gave rise to STD effects, being indicative of the interaction. Preliminary studies indicated that Hep[I] elongation chain (*i.e.*, residues

E, Y and G) and residues G and B linked to Hep[II] participate to the binding (Figure 4.12). In details, galactose residues, named F and G, participated to the binding displaying a high STD enhancement for protons in position 2 and 3. A good magnetization transfer was also observed for H4 belong to the same sugar units together with H2 of GlcNAc (Y). GlcNAc (B) residue also interact to the MAb via hydrogens in position 2 and 4. While, H4 of terminal GalNAc (E) underwent a slight magnetization transfer. In addition, a minor contribution was given by protons in position 4 belong to GlcNAc (Y) and by H5 of terminal galactose from Hep[II]-lactose.

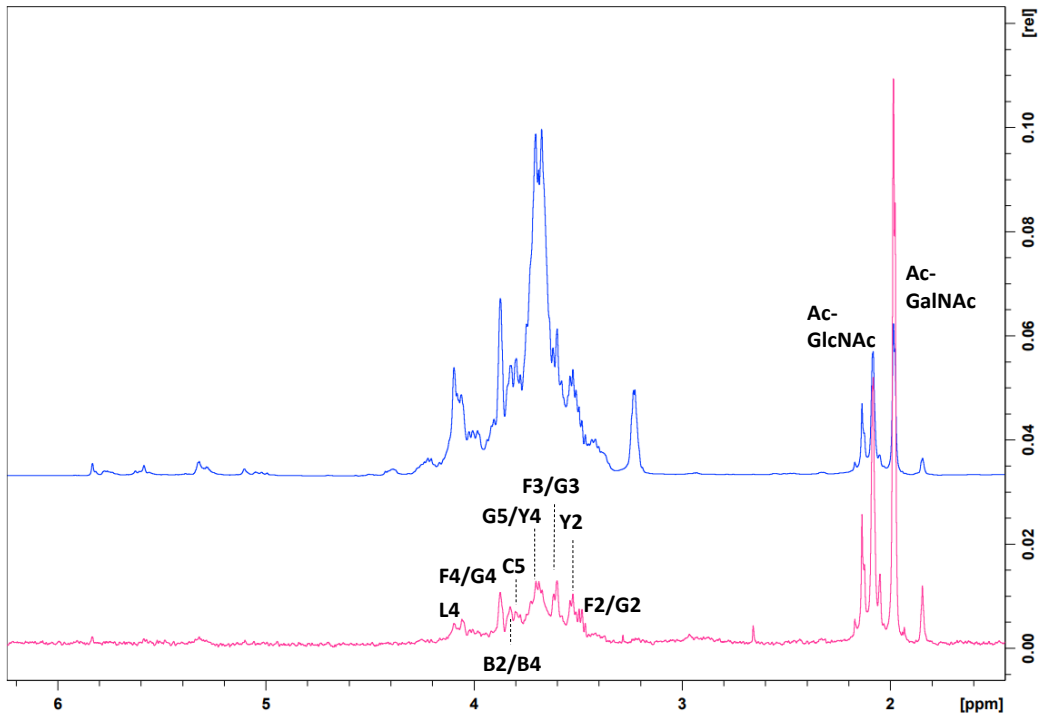


Figure 4.12. **STD NMR analysis of *N. gonorrhoeae* FA1090 bound to MAb 2C7**. 1D STD NMR spectrum at the bottom (pink) with the on top ^{1}H NMR reference spectrum (blue) of the 1:100 mixture for MAb 2C7-OS FA1090 wt complex with some of the key proton resonance signals labelled.

4.6.2 Computational approaches

Signals overlapping in the STD-NMR spectrum represented an issue toward the proper attribution of the protons involved during the interaction. Therefore, computational studies were carried out to acquire data in order to better describe the interaction mechanism underlying the recognition process between the antibody and the oligosaccharide. To lead the *in silico* studies, the most elongated OS structure was selected, which consist of GalNAc-Gal-GlcNAc-Gal-Glc chain from Hep[I]. First step was the MM studies carried out by using Maestro software for evaluating the conformation features. In oligo- or polysaccharide the chain conformation is mainly defined by the relative orientation of the sugar moieties, *i.e.* by torsion angles **φ** and **ψ** (Figure 4.13). Torsion angles around the glycosidic bonds are defined as follows: **φ** (H1-C1-O-CX') and **ψ,** (C1-O-CX'-HX'). Consequently, the potentially energy surface for each disaccharide connected by a glycosidic linkage were built (Figure 4.14) and then the OS FA1090 wt was built and modelled according to the energetic minima permitted. The OS structure was further minimized within AMBER 18 (Figure 4.15).

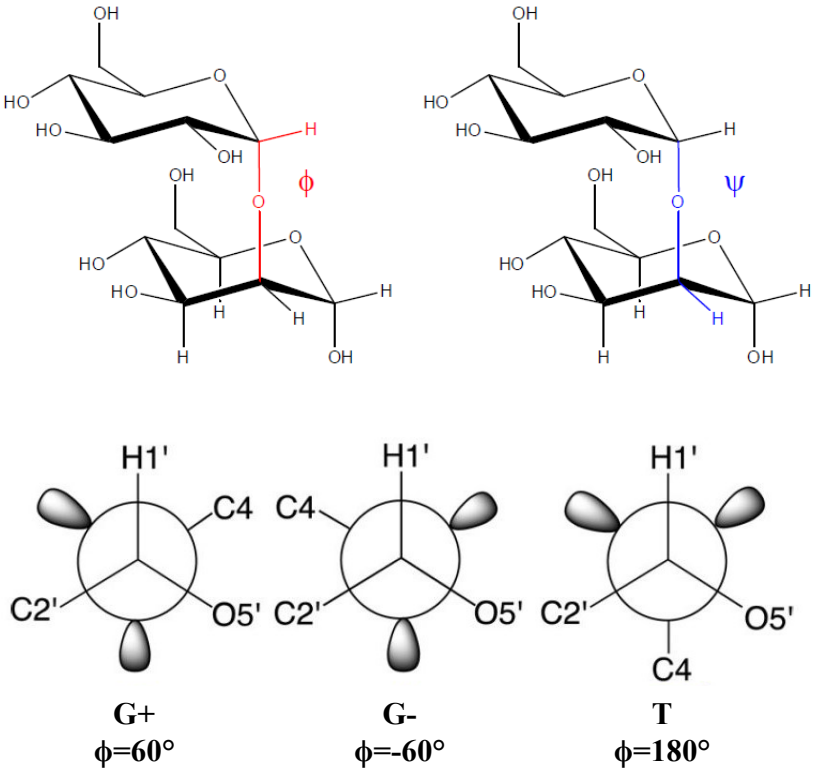
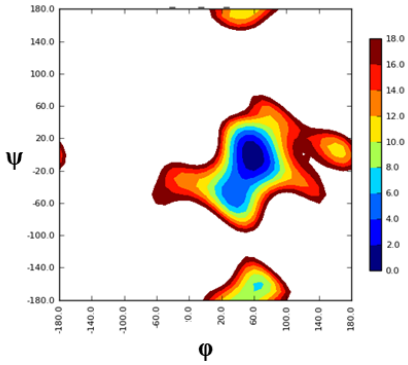
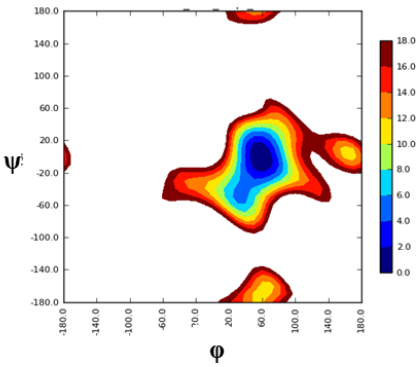


Figure 4.13. Representation of the ϕ, ψ dihedral angles in a disaccharide and Newman projections representation of the three staggered rotameric conformers of the ϕ dihedral angle.

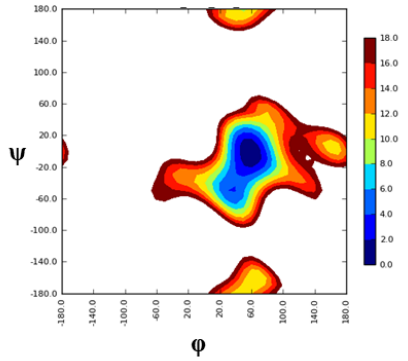
β-Gal-(1,4)-β-Glc



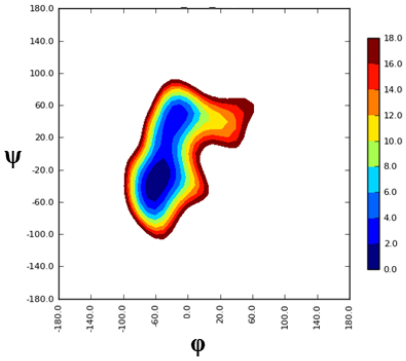
β-Glc-(1,4)-α-Hep[I]



β-Gal-(1,4)-α-Glc



α-Glc-(1,3)-Hep[II]



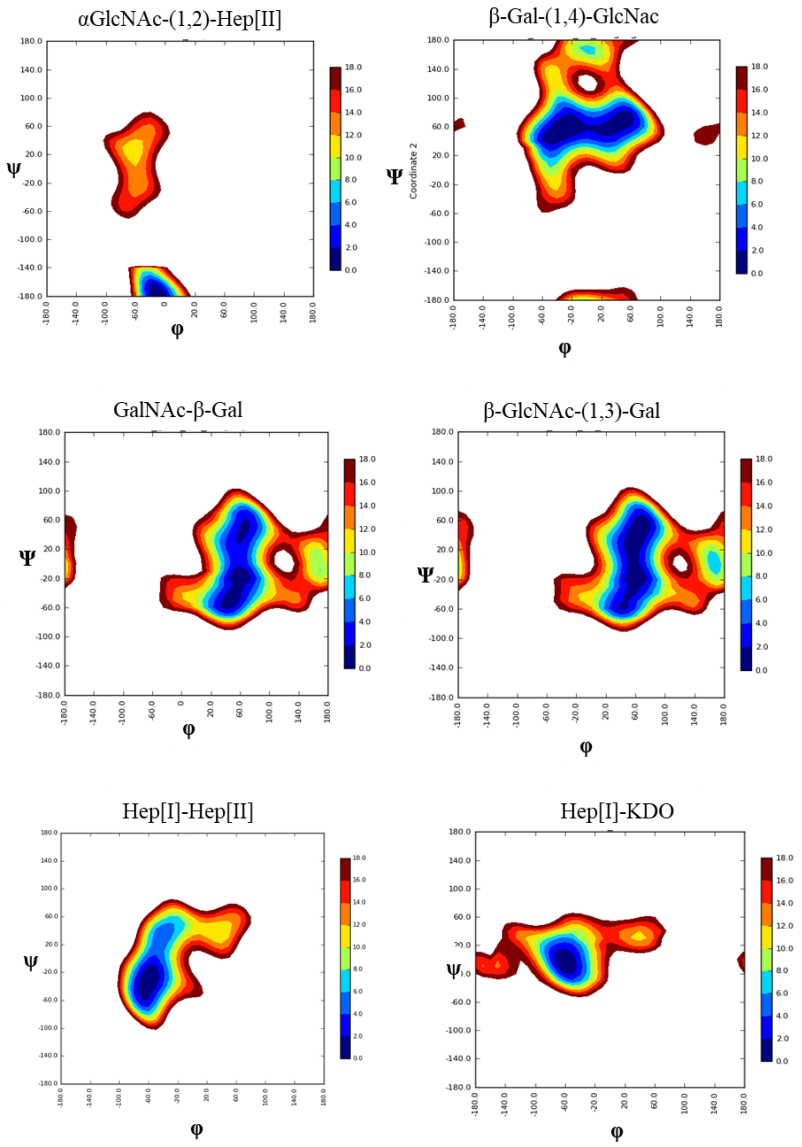


Figure 4.14. Adiabatic energy maps built within Maestro software for each disaccharide connected by a glycosidic linkage, defined by the torsion angles ϕ (H1-C1-O-CX') and ψ (C1-O-CX'-HX').

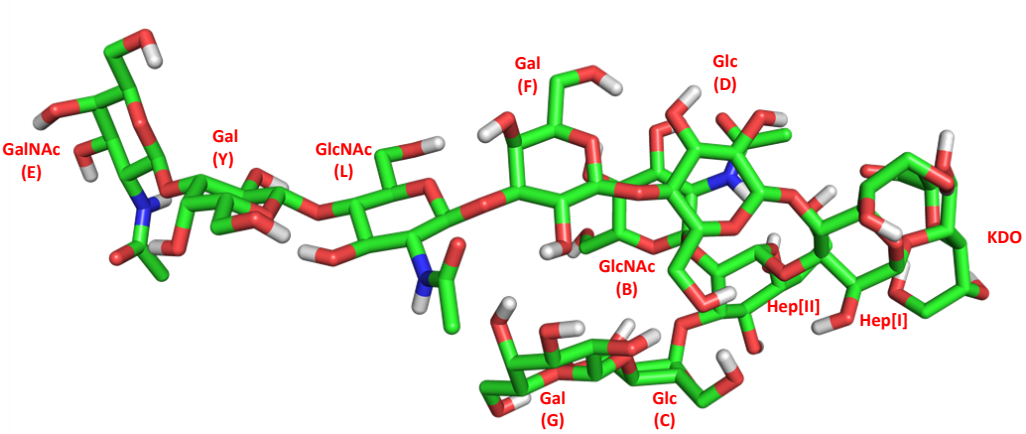


Figure 4.15. Oligosaccharide FA1090 wt minimized structure.

The minimized OS structure was modelled in the 2C7 binding site by taking into account the interactions found in the previous system studied (2C7-*Ng* OS 15253). The resulting complex between OS FA1090 wt and 2C7 was subjected to MD simulation od 100 ns in explicit water. The complex during MD simulation proved to be stable.

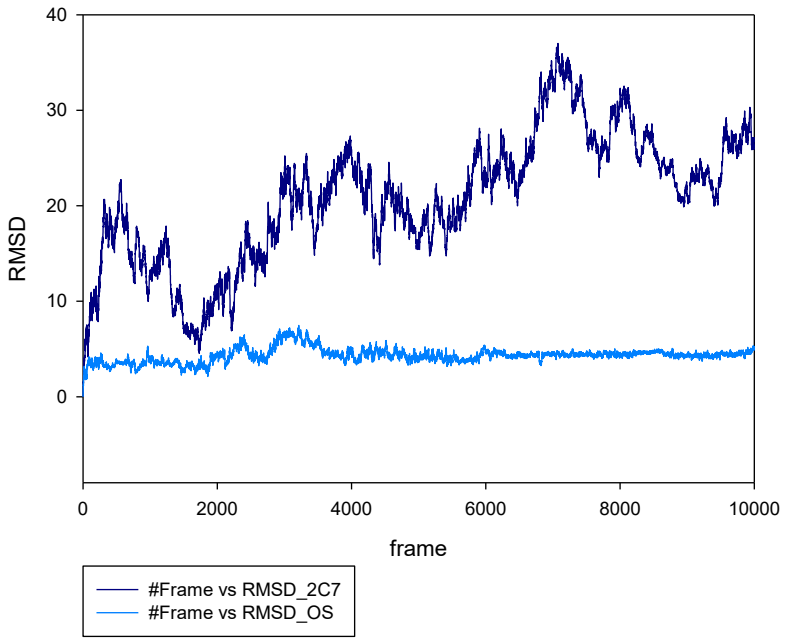


Figure 4.16. RMSD of the MAb 2C7 (dark blue) and OS FA1090 wt (light blue) after MD simulation having the protein as reference.

In the representative pose of the complex OS FA1090 - 2C7 obtained from a cluster analysis of the MD simulation, the main contacts took place through the recognition of residues B, G, Y and E (Figure 4.17). Hydroxyl groups in position 6 and 4 of Gal (Y) unit established a very stable H bond with Tyr72 and Tyr315 respectively. GalNAc (E) residue participated to the binding making three hydrogen bonds with His318, Gln74 and Ser316. Asn317 established contacts with the carbonyl group and OH in position 3 of GlcNAc (B) and with hydroxyl moiety at position 4 of Gal (Y). Additionally, Gal (Y) unit was also recognized by Arg77 residue, which set up a hydrogen bond with hydroxyl group at position 6. Moreover, Asn64 interacted with Gal (G) sugar unit by establishing a hydrogen bond via the OH at position 6.

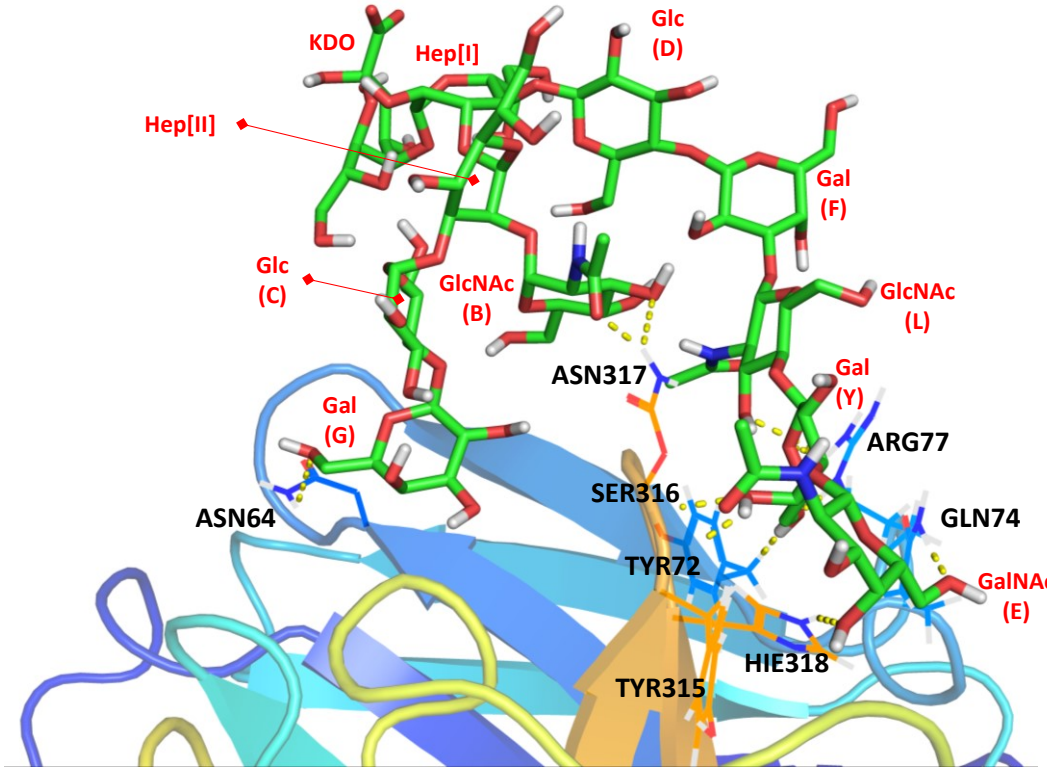


Figure 4.17. Representative pose of the 3D complex involving OS FA1090 wt and MAb 2C7. The protein residues involved in in H-bonds formation at protein-ligand interface are displayed as lines and highlighted in black, while the sugar unit are shown as sticks and coloured in green. The interactions are highlighted in yellow.

4.7 Discussion

In this chapter, the investigation of the interaction between *N. gonorrhoeae* FA1090 oligosaccharide and MAb 2C7 was undertaken with the aim to describe their binding mode. With this aim, as first, the LOS was extracted from *Ng* bacterial cells and then purified by applying several techniques. The second step involved the structural characterization of the oligosaccharide moiety and in particular the GC-MS experiments allowed the quali- and quantitative analysis of the monosaccharides,

while by MALDI-TOF and ESI-MS were very useful in determining the different OSs structure expressed by this *Ng* strain. After NMR characterization the basis of the interaction have been investigated by combining NMR and *in silico* techniques. Preliminary STD NMR studies have shown the occurrence of the binding and the involvement of OS residues E, Y, G, F and B, from both heptoses elongation and terminal Gal residue from the Hep[II] chain. It is important to highlight that the NMR studies were performed on a mixture of oligosaccharides in solution, which therefore generated a high degree of signal overlap. Therefore, to enrich the experimental data and better define the ligand binding epitope, computational studies were carried out. *In silico* model proven the engagement of all three branches into the binding event. In details, the involvement of terminal GalNAc (E) and galactose (Y) from Hep[I] elongation, and galactose (G) from Hep[II]-lactose together with GlcNAc (B) residues interacted with MAb 2C7 were observed in both STD and MD simulations approaches. By combining all data obtained, we can hypothesize a binding epitope composed by Y, G, E and B residues (Figure 5.18).

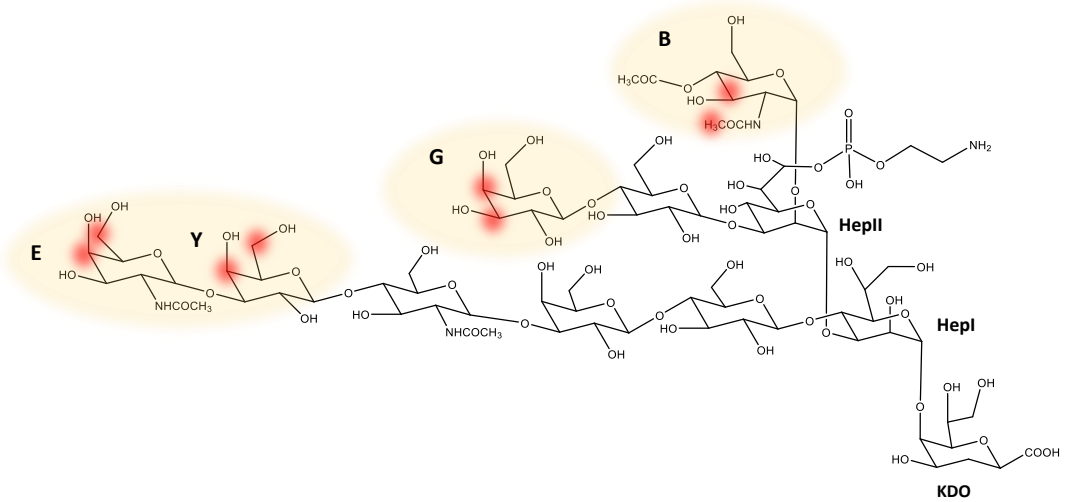


Figure 4.18. Proposed epitope map of the interacting oligosaccharide FA1090 wt from *N. gonorrohoeae* obtained by combining STD and computational data.

Chapter 5 – Molecular recognition of *Neisseria gonorrhoeae* strain 15253 and MAb 2C7

5.1 Introduction

Neisseria gonorrhoeae strain 15253 produces predominately a single LOS component [63] (Figure 5.1) which is composed by two lactose extensions attached to the conserved core trisaccharide GlcNAc–Hep II-Hep I. One lactose is β-(1-4)-linked to Hep[I] and the other is α-(1-3)-linked to Hep[II].

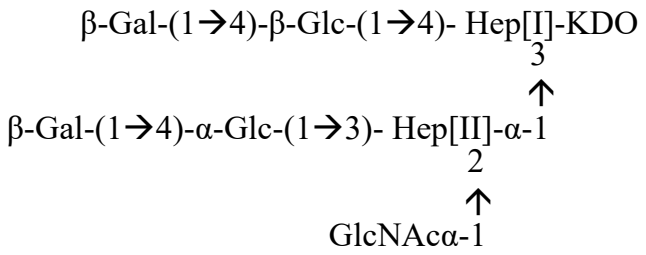


Figure 5.1. Predominantly OS structure expressed by *Neisseria gonorrhoeae* 15253. This structure corresponds to the minimal carbohydrate core required for MAb 2C7 binding [131].

5.2 OS isolation and purification

N. gonorrhoeae LOS 15253, provided by Professor S. Ram, was firstly stripped of its lipid portion by mild acid hydrolysis. The saccharide moiety, soluble in water, was then purified by size exclusion chromatography using a Biogel P4 column. Fractions were checked by NMR and then pooled to isolate the pure oligosaccharide portion.

5.3 NMR structural characterization

The oligosaccharide obtained after purification of the supernatant of the hydrolysis in acetic acid 1% of the LOS was analysed by NMR Spectroscopy. By combining

homonuclear and heteronuclear NMR techniques it was possible to determine the structure of the obtained oligosaccharide fraction (Figure 5.2, Figure 5.3 and Table 5.1). Figure 5.2 reports the ^{1}H-NMR spectrum of the oligosaccharide and the proposed structure. Seven anomeric signals (5.81, 5.33, 5.10, 5.05, 4.47, 4.42 and 4.38 ppm) were identified and assigned to Hep II (A), αGlc (C), αGlcNAc (B), Hep I (H), βGlc (D), βGal (G) and βGal (F) respectively. Sugar residues A, B, H and C resonated above 4.9 ppm and where consistent with **α** configuration, while residues D, F and G resonated below 4.9 ppm suggesting the **β**-linked hexoses. Indeed, chemical shift values of the anomeric carbons related to A, B, H and C resonated below 101 ppm, in contrast to those of D, F and G units, which resonated at lower field above 102 ppm [129]. The HSQC spectrum (Figure 5.3) also validated the identification of residue B as amino sugar since proton H-2 correlated with a nitrogen bearing carbon signal resonating at 54.21 ppm [130], N-acetylated at testified by the down-field shift of H-2 (3.83 ppm).

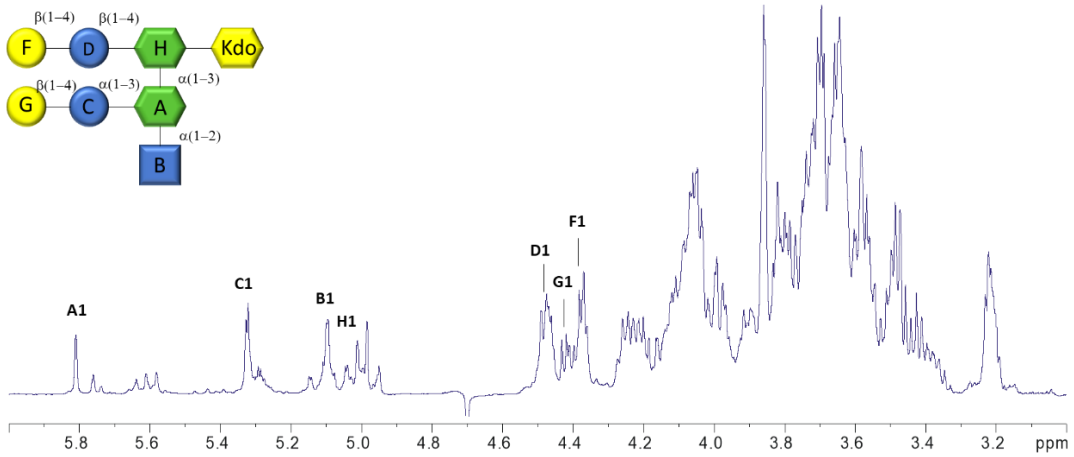


Figure 5.2. 1H spectrum of OS of *Neisseria gonorrhoeae* strain 15253 recorded at 298 K with anomeric peaks labelled.

Using 2D correlation experiments, spin systems where fully assigned. 2D NOESY experiments aided in identifying as yet unassigned peaks through *intra*-residue correlations. The C-(1→3)-A connection was confirmed by a correlation between H1 C and H3 A. *Inter*-residue correlations H1 D - H4 H and H1 B to H2 A confirmed the linkages βD-(1→4)-H and αB-(1→2)-A (Figure 5.4).

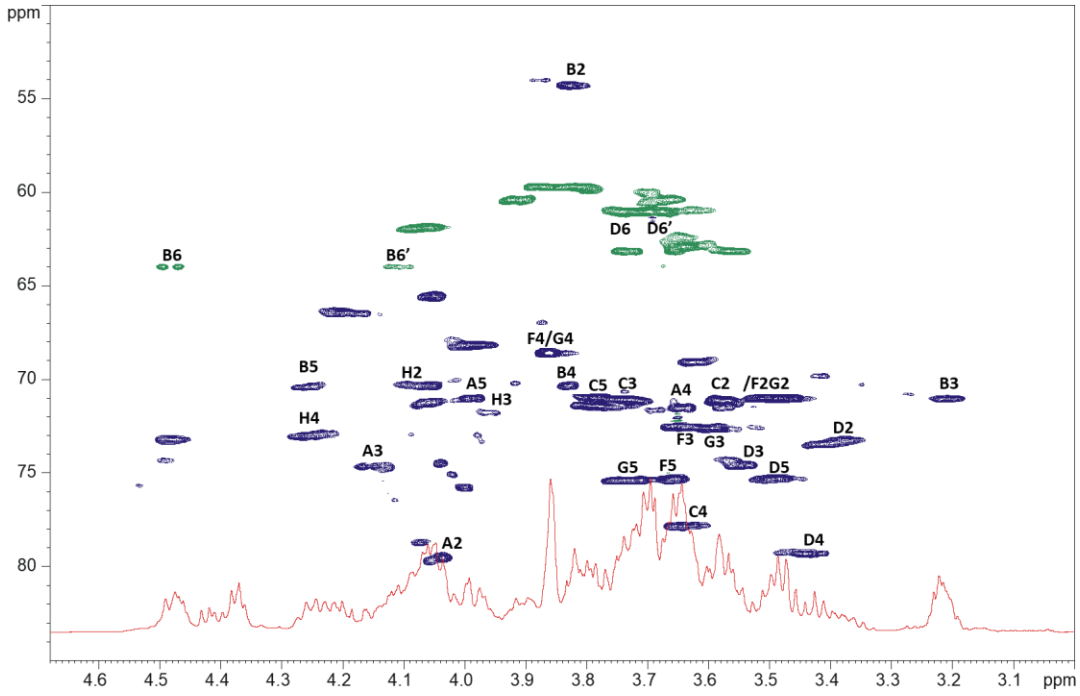


Figure 5.3. ^{1}H- ^{13}C HSQC spectrum of the *N. gonorrhoeae* 15253 oligosaccharide in PBS/D_2O, at 298K.

The chemical shift data is collected in Table 5.1.

Table 5.1. ^{1}H, ^{13}C chemical shifts (δ) values of *Neisseria gonorrhoeae* 15253 oligosaccharide. β-Gal-(1→4)-β-Glc-(1→4) linked to Hep[I] and β-Gal-(1→4)-α-Glc-(1→3) Hep[II]; α-GlcNAc-(1→2)-α-Hep[II]-(1→3)-Hep[I]-KDO.

Residue		1	2	3	4	5	6
αHep II	^{1}H	5.81	4.04	4.16	3.64		
(A)	^{13}C	97.89	79.46	74.63	71.51		
αGlcNAc	^{1}H	5.10	3.83	3.21	3.83	4.26	4.48/4.11
(B)	^{13}C	99.50	54.21	71.02	70.35	70.37	64.02
αHep I	^{1}H	5.05	4.09	3.96	4.25		
(H)	^{13}C	98.85	70.29	71.74	72.69		
αGlc	^{1}H	5.33	3.58	3.73	3.64	3.79	
(C)	^{13}C	99.82	71.12	71.08	77.78	71.37	
βGlc	^{1}H	4.47	3.21	3.55	3.44	3.49	3.74/3.67
(D)	^{13}C	102.35	71.01	74.57	79.26	75.26	61.04
βGal	^{1}H	4.38	3.51	3.65	3.86	3.66	
(F)	^{13}C	102.85	70.99	72.50	68.56	75.36	
βGal	^{1}H	4.42	3.48	3.60	3.86	3.74	
(G)	^{13}C	103.03	71.02	72.56	68.56	75.35	

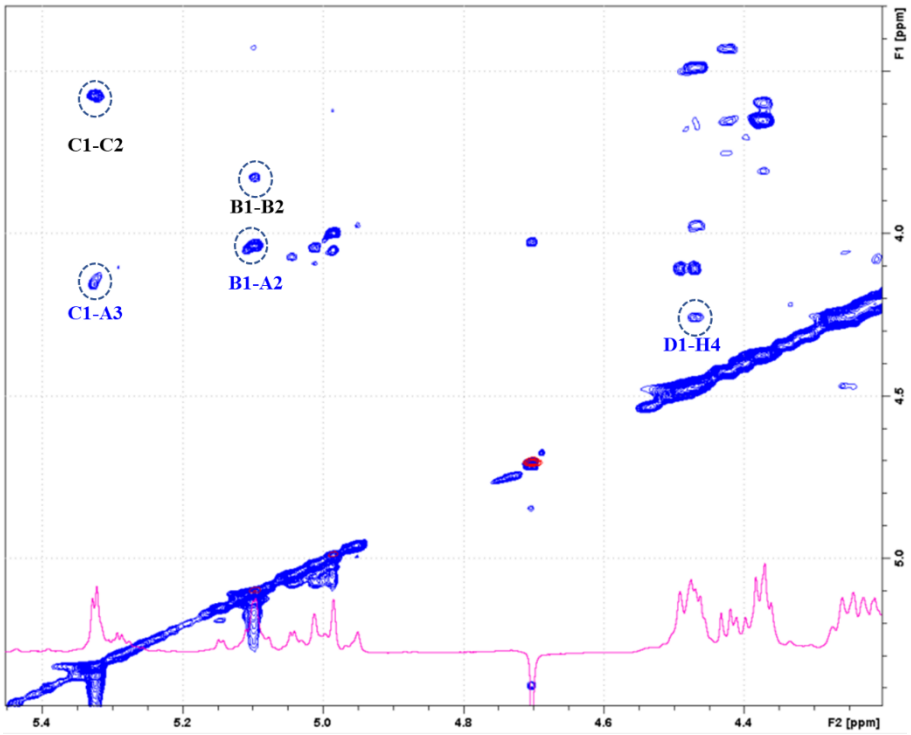


Figure 5.2. Anomeric region of the 2D NOESY spectrum of OS from *N. gonorrhoeae* 15253 with identified intra- and inter- residue NOESY correlations labelled.

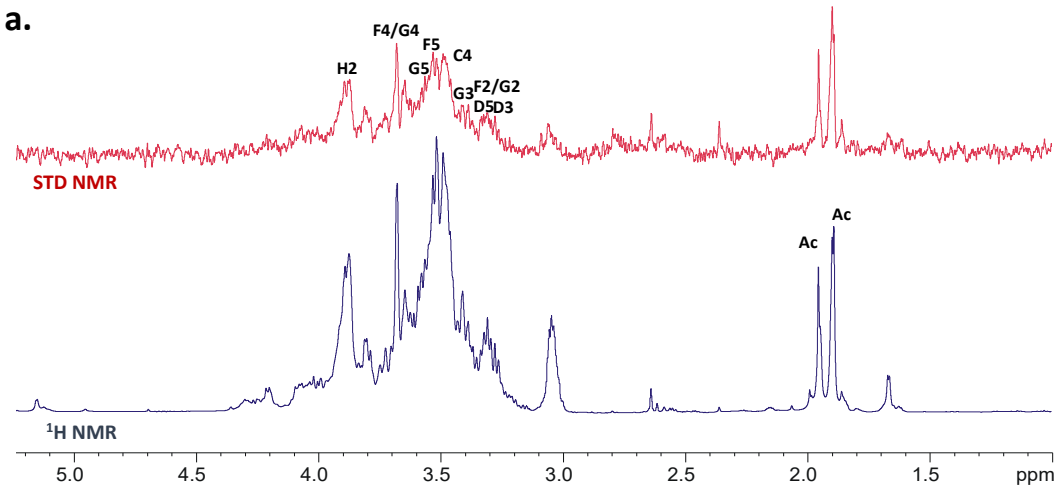
5.4 Binding studies between *N. gonorrhoeae* 15253 oligosaccharide and MAb 2C7

5.4.1 STD-NMR analysis

To dissect the molecular mechanisms at the basis of gonococcal oligosaccharide 15253 recognition by monoclonal antibody 2C7, its interaction was proven first through Saturation Transfer Difference Spectroscopy. STD NMR was employed to define the OS binding epitope upon interaction with MAb. Before acquiring STD experiments on the mixture, STD spectra were acquired on the OS and MAb in the free state in order to establish the optimal acquisition parameters, such as temperature, *power pulse*, *duration pulse* and *on-resonance frequency*. As for OS,

spectra at -0.5, 0, 6.5 and 7.5 ppm as irradiation frequency were run; the best on-resonance frequency was chosen at -0.5 ppm, because no STD signals were observed. Conversely the experiment carried out on the MAb in its free state at the same on-resonance frequency showed a good and uniform saturation of the protein (data not shown). The off-resonance frequency was set at 40 ppm and the temperature was set at 283K.

As shown in the spectrum reported in Figure 5.4, changes in the multiplicity and relative intensity of STD signals with respect to those of the reference (Figure 5.4) were observed, indication of specific binding.



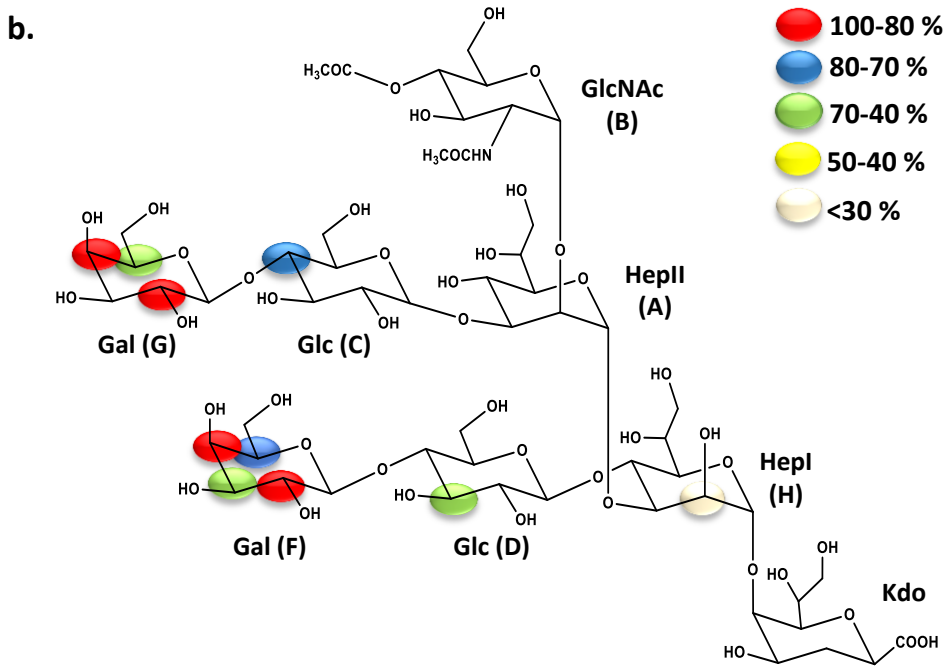


Figure 5.4. a. STD NMR analysis of *N. gonorrhoeae* 15253 bound to MAb 2C7. Ligand/protein ratio 1:100.

b. Epitope map of oligosaccharide calculated by (I0−Isat)/I0, where (I0−Isat) is the intensity of the signal in the STD NMR spectrum and I0 is the peak intensity of the unsaturated reference spectrum (off-resonance).

Preliminary STD results shown that the most involved sugars during the binding are the two terminal galactoses. Indeed, the highest STD signals belonged to the galactose residues (G) and (F) and especially to their protons in position 4 and 2 indicating a major involvement in the interaction. Other STD enhancement was observed for other protons of residue G, such as H3 and H5. Less intense STD effects were also ascribable to the H5 proton of residue F (close to 60%) and to the protons H3 and H4 of sugars D and C respectively. A low degree of magnetization transfer was also detected for proton in position 2 of the Hep[I]. However, given the high

signals overlap these data should be consolidated and confirmed by further binding experiments and by computational approaches.

5.4.2 Docking & Molecular Dynamic simulations

5.4.2.1 Conformational analysis

The investigation of the conformational features and the available conformational space for the glycosidic angles of the OS was performed through molecular mechanics and molecular dynamic (MM and MD) simulations. Molecular mechanics provided a first estimation of the conformational regions energetically accessible. First, the seven basic constituent disaccharides of the OS structure β-Gal-(1→4)-Glc, β-Glc-(1→4)-Hep[I], β-Gal-(1→4)-Glc, α-Glc-(1→3) Hep[II], α-GlcNAc-(1→2)-Hep[II], α-Hep[II]-(1→3)-Hep[I] and α-Hep[I]- (1→3)-KDO) were built and MM calculation were performed to analyse the energetically accessible conformational regions by means of adiabatic energy maps. The resulting energy maps, obtained by using Maestro suit programme, for the glycosidic torsions **φ** (H1-C1-O-CX') and **ψ** (C1-O-CX'-HX') [132] are reported in Figure 5.5. The values of dihedral angles φ and ψ are enlisted in Table 5.2.

Table 5.2 Approximated average values of each glycosidic linkage dihedrals of the oligosaccharide obtained through adiabatic maps.

Residues	ϕ	ψ
Hep I – KDO	-60°	-10°
Hep II – Hep I	-60°	-20°/ -60°
GlcNAc – Hep II	-30°	-180°
Gal β – Glc β	60°	0°

Glc β – HepI	60°	0°
Gal β – Glc α	60°	0°
Glc α - HepII	-60°	-30°/30°

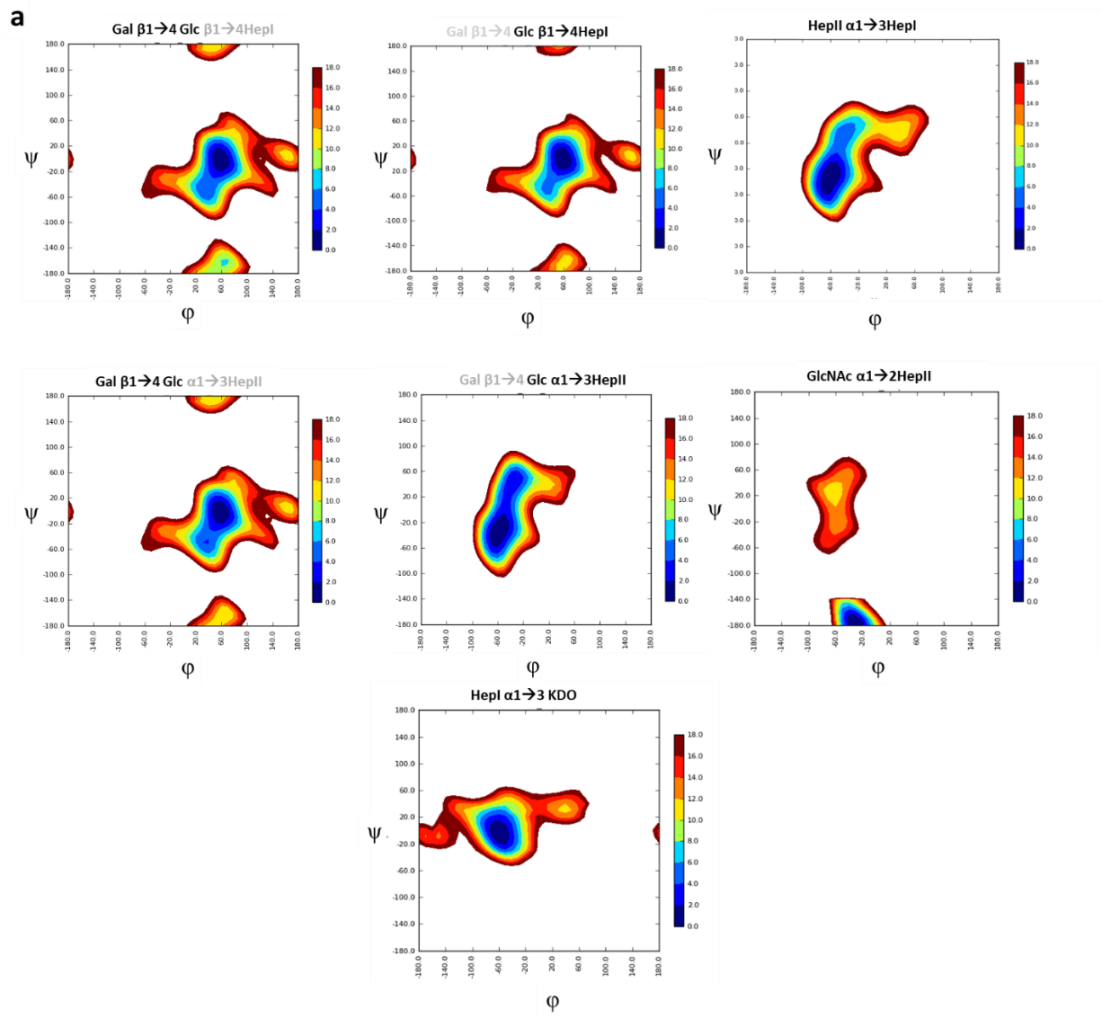


Figure 5.5. MM study of *N. gonorrhoeae* OS 15253. a. Adiabatic energy maps of the seven disaccharides connected by a glycosidic linkage, performed with Maestro suit programme.

The map for each disaccharide is a display of the energy plotted against the two torsion angles. **b**. Values of each glycosidic linkage dihedrals of the disaccharides.

Once the optimal values for φ and ψ dihedral angles had been estimated for each disaccharide, the oligosaccharide structure was built and subjected to 100 ns MD simulations in explicit water with AMBER 18 (Figure 5.6 a). It is worth to note that as the KDO is not a standard sugar implemented in GLYCAM06-j forcefield, its parametrization was performed in our research laboratory.

The MD simulation analysis in explicit solvent highlighted the presence of different conformational families. In details, two different families, defined by a different set of dihedral angles were obtained from cluster analysis, named **Rep1** and **Rep2** (Figure 5.6 c). **Rep1** family was characterized by a "folded" form, while **Rep2** family exhibited a more "extended" conformation. Indeed, in Rep1 conformation the GlcNAc residue (γ chain) was enclosed between the chains α and β. Conversely, in Rep2 the GlcNAc unit was more exposed. The stability of each family and the variation of the dihedral angles along each glycosidic linkage was monitored. In detail, Rep1 conformation (in green in Figure 5.6 c) is representative of the clusters 0 and 2 (Figure 5.6 b) and represent the most populated family. Rep2 conformation (in cyan in Figure 5.6 c) is representative of the clusters 1 and 3.

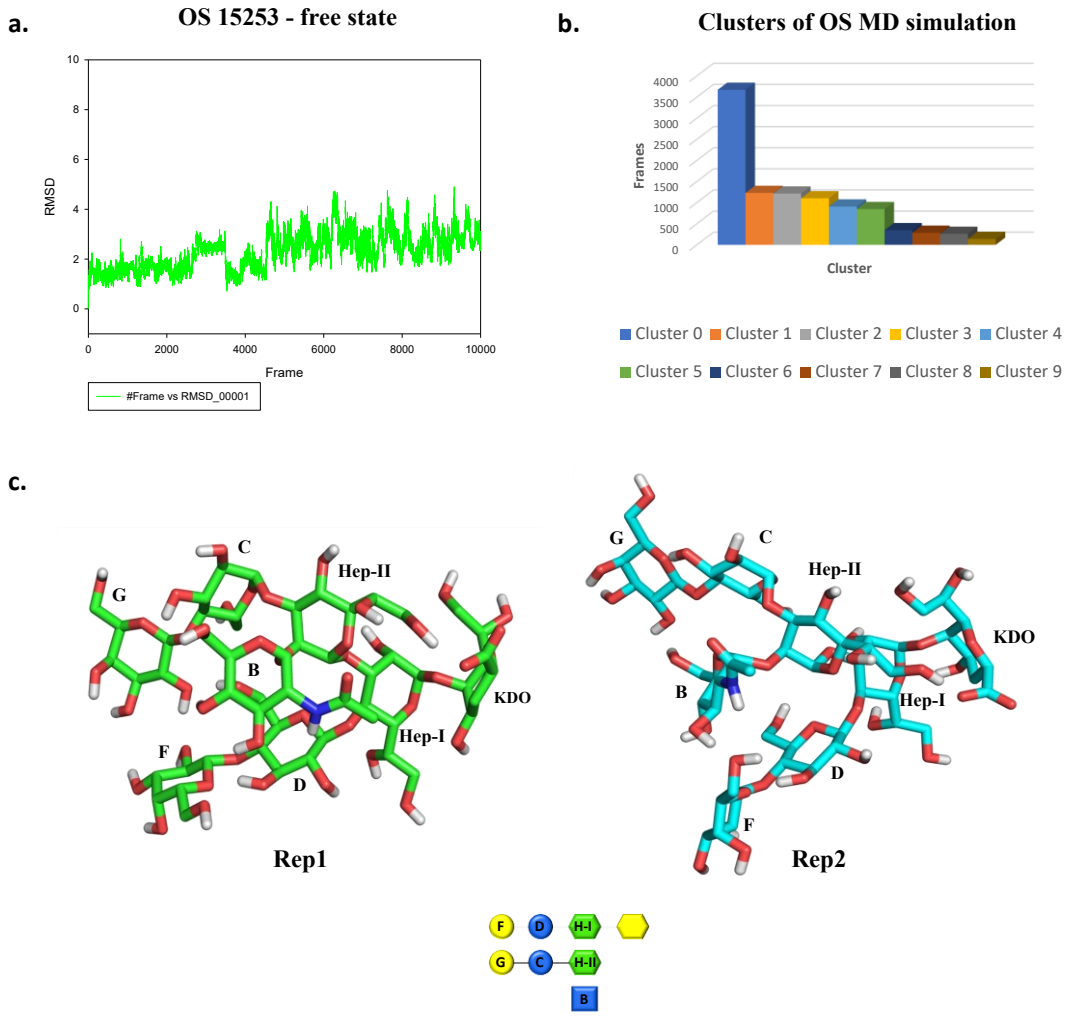


Figure 5.6. Molecular dynamic simulation analysis of the oligosaccharide 15253 in its free state. **a.** RMSD analysis of the *Ng* 15253 OS in the free state. **b.** Graph showing all frames of the 100 ns dynamics grouped into 10 clusters. Among the most populated clusters (0, 1, 2 and 3), the first one is the most representative cluster. **c.** The two main conformational families of the oligosaccharide, named Rep1 and Rep2. Rep1 conformation was characterized by a rather bent topology and is representative of clusters 0 and 2, while Rep2 shown a more extended topology and is indicative of clusters 1 and 3.

The torsion angles of both Rep1 and Rep2 were measured and reported in Table 5.2.

Table 5.2. Torsion angles ϕ and ψ values measured for Rep1 and Rep2 of the oligosaccharide from N. gonorrhoeae 1525 in the free state.

Residues	Rep1		Rep2	
	ϕ	ψ	ϕ	ψ
HepI – Kdo	-53°	-14°	-22°	34°
HepII – HepI	-45°	-34°	-65°	-56°
GlcNAc – HepII	0°	177°	-71°	-49°
Gal β – Glc β	44°	-28°	89°	36°
Glc β – HepI	63°	0°	60°	23°
Gal β – Glc α	34°	13°	58°	-8°
Glc α – HepII	-60°	-12°	-56°	-20°

5.4.2.2 Docking analysis

Once the oligosaccharide conformation was evaluated, the docking calculation with the MAb 2C7 was performed. Rep1 family conformation was used as starting ligand geometry, which matched with the set of minima around each glycosidic linkage and the most stable conformation in water. The docking calculations were performed with AutoDock 4.2 [133]. The most populated and less energetic cluster (Figure 5.7) obtained from the docking analysis gave an indication of the interactions which occurred in the modelled complex. In the representative docking pose the two galactoses were accommodated into the MAb binding pocket, while the conserved core trisaccharide GlcNAc–Hep II-Hep I was located far from the interacting site.

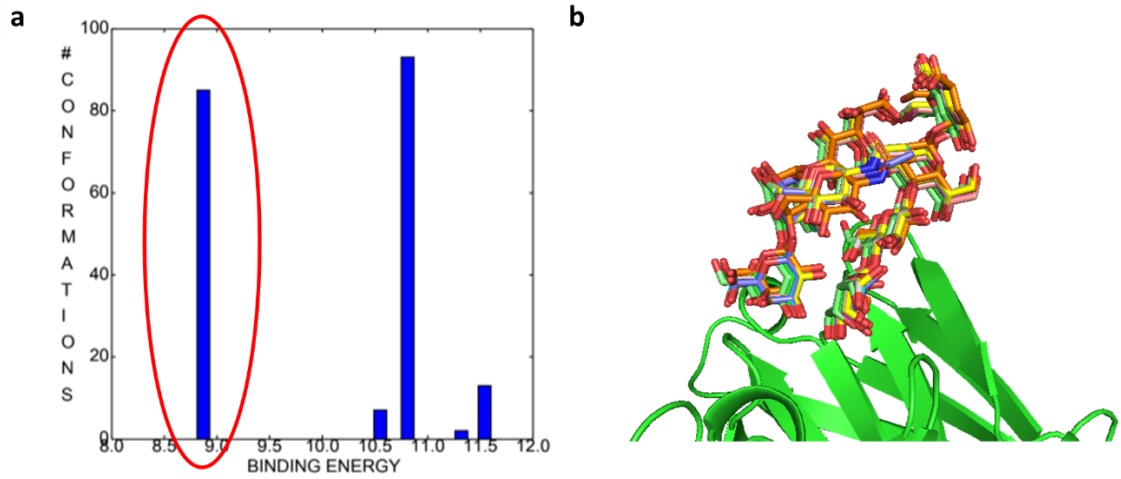


Figure 5.7. Docking analysis of the complex between Ng OS 15253 and MAb 2C7. **a.** Graph reporting the clusters of all conformations. **b.** Superposition of the main poses from the cluster at lower binding energy.

5.4.2.3 MD simulation of the complex involving OS 15253 and MAb 2C7

The representative pose underwent the minimization and MD simulation of 100 ns in explicit water, to evaluate the stability and the most relevant protein/ ligand interactions. MD simulations were run by using AMBER 18 [134]. The complex proved to be stable for the entire simulation time. In Figure 5.8a the RMSD of the MAb 2C7 (in blue) and the OS 15253 (in green) were reported. The torsion angles in the bound state showed that the oligosaccharide preferentially adopted a bent topology (Figure 5.9) similar to **Rep1** (Figure 5.6 c). The dihedral angels ϕ and ψ of the most representative pose were measured (Table 5.3). In Figure 5.8b, as an example, the plots showing the dihedral angels ϕ and ψ around β-Gal-(1,4)-α-Glc glycosidic linkage have been reported.

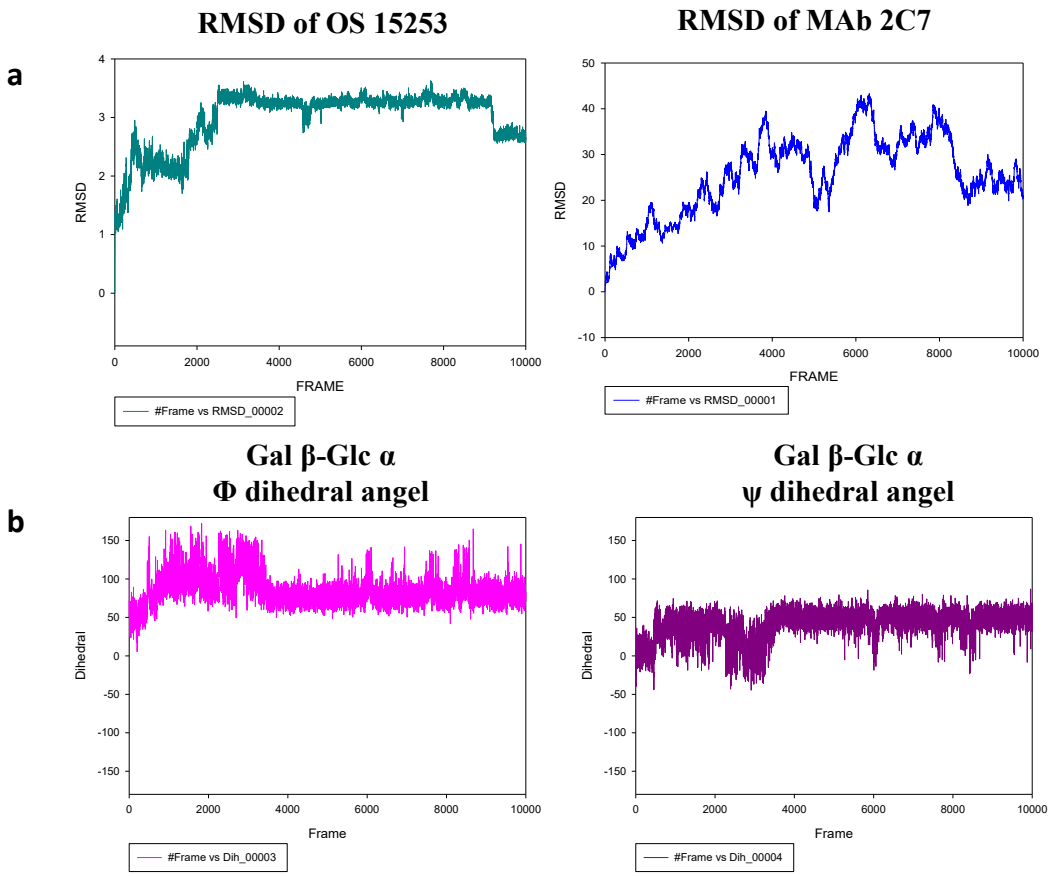


Figure 5.8. MD simulation analysis of OS 15253 /MAb 2C7 complex a. Root-mean square deviation (RMSD) plots of the oligosaccharide 15253 and MAb 2C7. **b.** Torsion angles around the β-Ga/α-Glc glycosidic linkage monitored during the MD simulation.

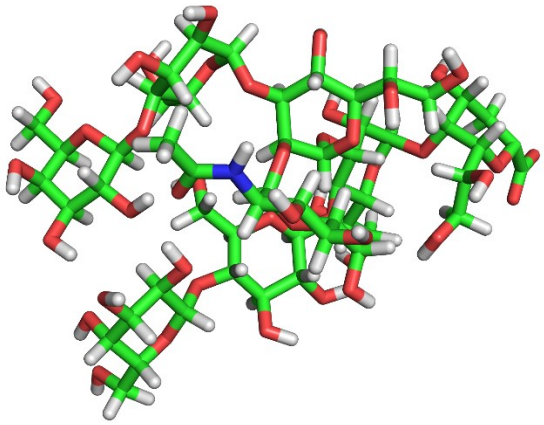


Figure 5.9. *N. gonorrhoeae* 15253 oligosaccharide conformation adopted during the binding with the MAb 2C7.

Table 5.3. Torsion angles ϕ and ψ values measured for the most representative pose of the OS 15253 in the bound state with MAb 2C7.

Residues	ϕ	ψ
Hep [I] – KDO	60°	56°
Hep [II] – Hep[I]	-49°	-40°
GlcNAc – Hep [II]	-55°	33°
β Gal – β Glc	-22°	-31°
β Glc – Hep[I]	38°	0°
β Gal – α Glc	78°	52°
α Glc – Hep [II]	-46°	-29°

MD simulation showed the involvement of the two lactose branches in the recognition process (Figure 5.9). A stable H-bond took place between the H in position 6 of **Gal G** (from Hep[II] elongation) and Thr39 present for about 70% of

MD simulation. The same galactose residue also engaged stable hydrogen bonds with Asn61 and Trp110 (48% and 25% respectively). Asn63 established contact with O6 of **Glc C** residue and **Glc D** was involved in hydrogen bonds with Asn64 as well as in Pi-Sigma interaction with Phe66. Hep I also participated to binding making the interactions with Asn64. It is worth to note that no relevant interactions or contacts specifically involving the KDO, Hep II and GlcNAc were found, since they were placed far from the binding site.

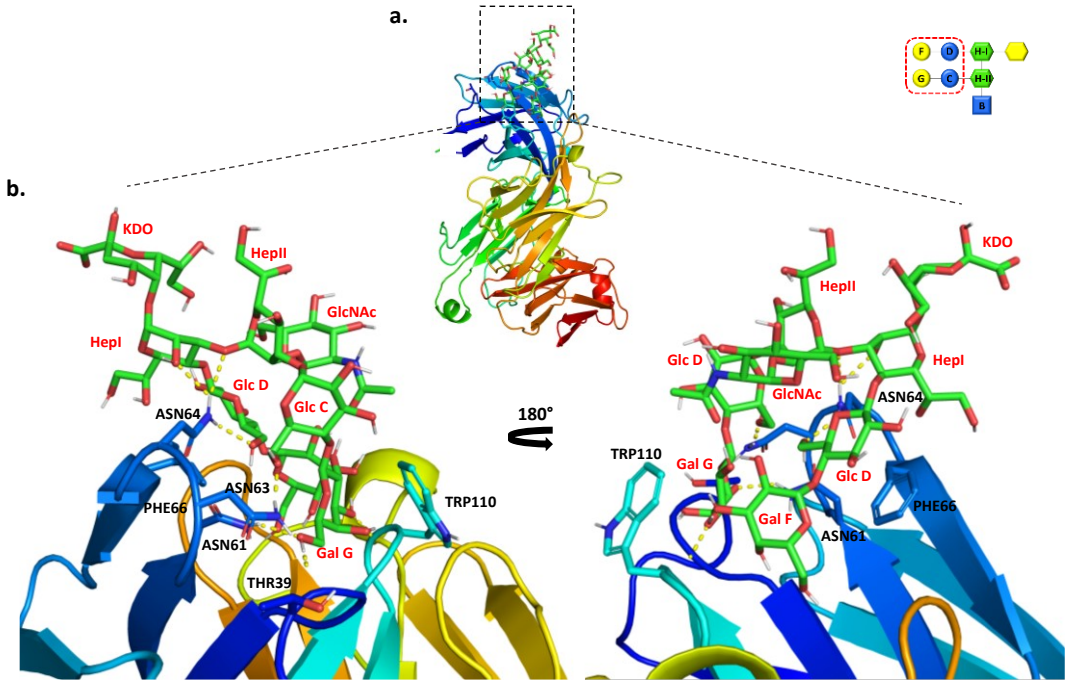


Figure 5.9. a. 3D model coming from 100 ns MD for the complex MAb 2C7 - *N. gonorrhoeae* 15253 oligosaccharide; **b.** Close up view of oligosaccharide binding mode at MAb 2C7 binding site. The main amino acid residues involved in the binding are shown in sticks. OS chain is depicted in green, and the sugar residues were highlighted in red.

5.5 Discussion

Anti-LOS monoclonal antibody 2C7 was elected as a promising anti-gonococcal immunotherapy, capable of eliciting complement-dependent bactericidal activity. Previous studies have shown binding affinity between LOSs expressed by *N. gonorrhoeae* and monoclonal antibody 2C7. The LOSs showing affinity towards the MAb share the same conserved epitope composed by the trisaccharide core (GlcNAc–Hep [II]-Hep[I]) and two lactose elongations from each heptose. Interestingly, it has been noted that the loss of lactose unit from Hep [II] [62] inhibited interaction with 2C7. In this chapter we explored the interaction between the oligosaccharide fraction from *Neisseria gonorrhoeae* 15253 LOS and MAb 2C7 with the aim of describing at molecular level the mechanisms underlying the binding. We first characterize the ligand by NMR spectroscopy and then we performed the STD NMR analysis to detect and study the interaction. STD results revealed that the binding occurred between the OS 15253 and 2C7. The most involved residues within the interaction were Gal F and Gal G from Hep[I] and Hep[II] respectively. The Glc moieties underwent a minor magnetization transfer by the MAb, whereas the heptoses and KDO appear not to be recognized. However, due to the high signals overlap in NMR spectrum it was not possible to correctly establish whether only one lactose unit or both were recognised. In order to enrich the NMR data and provide a more detailed analysis of the binding the computational studies were carried out on the complex. First, the OS conformation was investigated in its free state and two different conformational families were found, Rep1 and Rep2. Rep 1, corresponding to the most populated family, displayed a rather bent topology, while Rep2 shown a more extended conformation. The Rep1 was used as starting point for running the MD simulation on the complex (OS 15253-MAb 2C7). MD simulation of 100ns revealed involvement of both chains. However, the strongest interaction was observed for Gal G coming from the Hep[II] elongation. Notably, the Hep[II] branch

recognition is consistent with the previously published findings. Indeed, as discussed before the MAb does not detect the LOSs structure which lack the *lgtG* glycosyltransferase gene, suggesting the key role in the recognition. Furthermore, the OS core composed by the two Hep and GlcNAc together with the KDO were far from the binding pocket. These results are in agreement with the preliminary STD experimental results. However, to better define the interaction further experiments are needed, such as the STD-TOCSY experiment.

5.6 Concluding remarks

Neisseria gonorrhoeae represents one of the leading global health concerns due to its worldwide spread. Furthermore, the uncontrolled increase in antibiotic resistance necessitated the development of alternative and targeted treatments to antibiotic therapy to fight gonococcal infection and related disease. The development of an effective vaccine, based on the use of highly conserved surface components of *N. gonorrhoeae* strains, is a promising option to fight the emerging epidemic disease. Among the different surface components, LOS has proven to be a valid candidate. The LOS is the major component of the outer membrane, and it is readily accessible as a target of adaptive immunity.

In the present thesis work two different gonococcal strains were studied. In details, the LOSs isolation from *Ng* strains FA1090 and 15253 and their structural analysis allow to characterize the binding mode between the LOS-derived oligosaccharides and the monoclonal antibody 2C7 by both experimental (STD-NMR spectroscopy) and in silico approaches. 3D models of the complexes (OSs-MAb) were derived for each system studied depicting the OS binding epitope.

In conclusion, the results obtained provide an important starting point for developing innovative targeted therapies.

SECTION III – OTHER PROJECTS

Chapter 6 – Molecular basis of recognition process involving *Mycoplasma pneumoniae* and *Mycoplasma genitalium* cytoadhesins and eukaryotic glycans

6.1 Introduction

Mycoplasmas are the smallest self-replicating prokaryotic cells [135] which lack the rigid cell wall having an unusually small genome size of 580-2200 kilobase pair [136]. *Mycoplasmas* belong to *Mycoplasmataceae* family and order *Mycoplasmatales* [137-140]. These pathogens are phylogenetically related to Gram-positive bacteria and are able to invade and replicate in eukaryotic cells [141]. Several species of human *mycoplasma* exist, but only six are able to cause disease; among these species, the most important and predominant pathogens are *Mycoplasma pneumoniae (Mpn)* and *Mycoplasma genitalium (Mge)* [137]. *Mpn* and *Mge* are two emerging human pathogens which are responsible for respiratory and urogenital tract infection respectively. These pathogens express cytoadhesins (adhesion molecules) on their bacterial surface for mediating the attachment to sialylated host glycan receptors and initiating the bacterial infections.

The high incidence of *M. genitalium* and *M. pneumoniae*-related diseases together with the increase in antibiotic resistance [142-144], and the ability of pathogens to elude the host immune system [143], have made the development of more targeted therapeutic treatments essential.

6.2 *Mycoplasma pneumoniae* (*Mpn*) and *Mycoplasma genitalium* (*Mge*)

M. pneumoniae triggers acute and lower tract infections [145]. It is known to be one of the leading cause of about 40% of community-acquired pneumonia (CAP) both in children and young adults [146, 147] and it has also been estimated that 18% of paediatric cases require hospitalisation [148]. In addition, *Mpn* can cause atypical

CAPs involving extra-pulmonary organs, including head, eyes, skin or intestinal tract [149].

On the other hand, *Mycoplasma genitalium* (*Mge*) is a sexually transmitted bacterial pathogen [150] that often co-infect the urogenital tract simultaneously with other pathogens such as human immunodeficiency virus (HIV), *Chlamydia trachomatis*, *N. gonorrhoeae* [151] and *Trichomonas vaginalis* [152]. *Mge* infections are one of the main causes (about 30-40% of cases) of urethritis in men; while, in women has been associated with several diseases related to the reproductive system such as cervicitis, endometritis, pelvic inflammatory disease (PID) [153, 154]. It can also result in infertility and susceptibility to the HIV [155].

Mpn and *Mge* encode a variety of virulence factors, which include adhesins, glycolipids, toxic metabolites, exotoxins, pathogenic enzymes, and some membrane components [156]. They initiate host invasion by adhering to the surface of epithelial cells of the respiratory/urogenital tract by means of a polar tip structure known as attachment organelle (Figure 6.1). The attachment organelle consists in a network of adhesins and cytoadherence accessory proteins [157, 158]. The main cytoadhesins which decorate the membrane of the attachment organelle constitute a transmembrane adhesion complex named Nap. Nap plays a central role in the infection mechanism and share a tetrameric organization which consist of dimer of heterodimers. In *Mpn*, Nap is composed by heterodimers P1 and P40/P90, while in *Mge* is made up heterodimers P140 and P110 [159, 160] Figure 6.1).

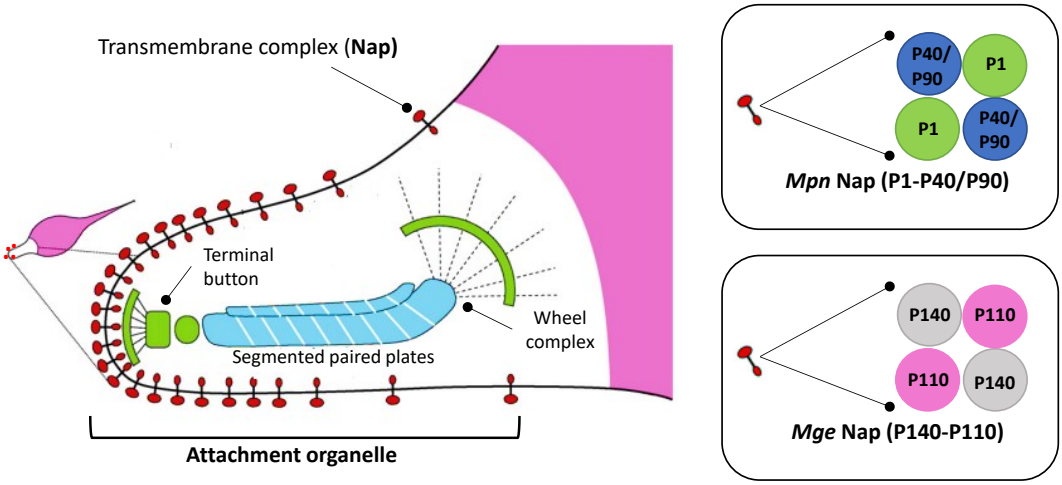


Figure 6.1. General schematic representation of the attachment organelle of *M. pneumoniae* and *M. genitalium*. An outline of the entire cell is presented on the lefthand side, alongside an enlargement of the attachment organelle. The entire cell is shown on the left, while on the right is an enlargement of the attachment organelle. The cytoplasm is magenta coloured. The *mycoplasmas* structures surrounding the attachment organelle can be classified as follows: 1. Transmembrane complex, **Nap**, comprising adhesin P1, P90 and P40 in *Mpn* and P140 and P110 in *Mge*; 2. segmented paired plate, comprising HMW1-HMW2 in *Mpn* and MG312-MG218 in *Mge*; 3. terminal button, comprising P65, P30 and HMW3 for Mpn and MG217-MG317 for *Mge*; 4. wheel complex, including P41, P24 and P200 in *Mpn* and MG491, MG386, MG200 and MG269 in *Mge* [151, 161]. Figure adapted from reference [161].

For years, P1 in *M. pneumoniae* has been identified as one of a major immunodominant proteins in *Mpn* cells, playing a key role in cytoadhesion. However, recently, it has been discovered that the binding site for sialylated oligosaccharides lies on the interface of the two heterodimers P40/P90 [160, 162]. As for *M. genitalium*, the binding pocket is found on the P110 protein [159, 163]. *M. genitalium* and *pneumoniae* cytoadhesins (GPCAs) share a similar structural organization which consists of a large extracellular N-domain, a small C-terminal

domain (C-domain) and a transmembrane helix. The N-terminal domain is organised into two regions: seven-blade β-propeller and the crown (Figure 5.2). The sialic acid binding site is located in the terminal crown.

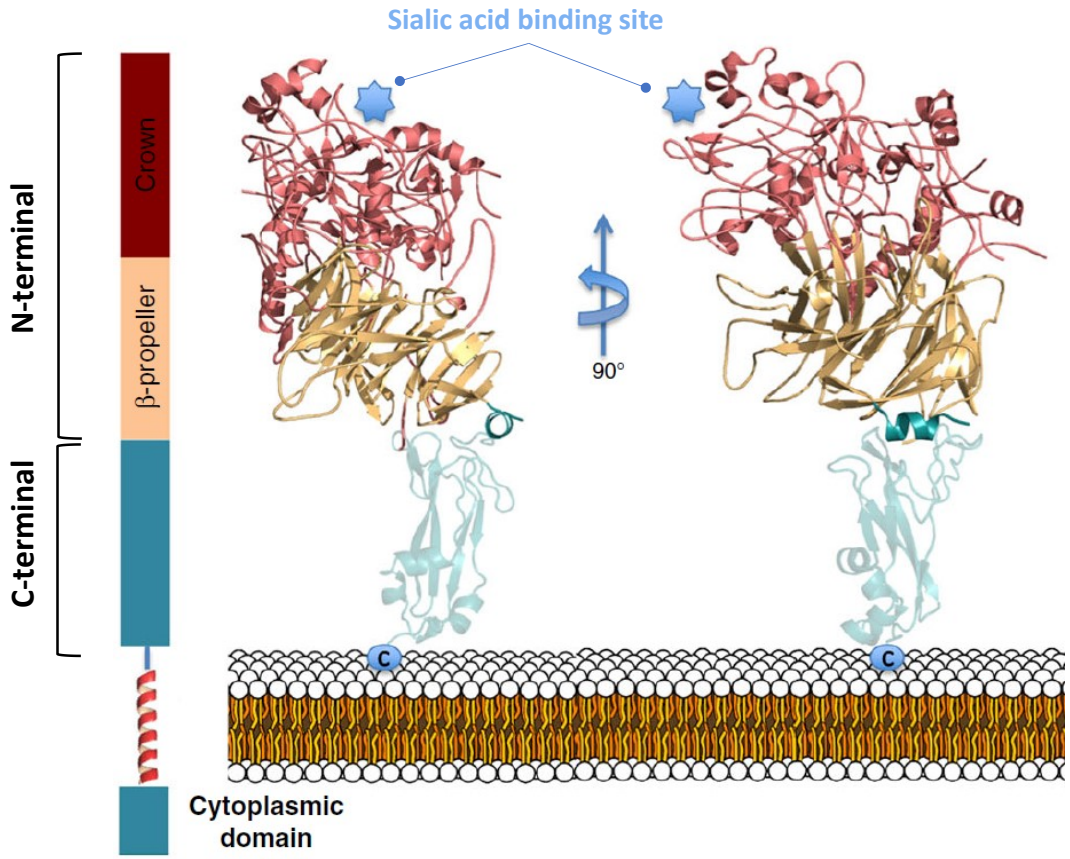
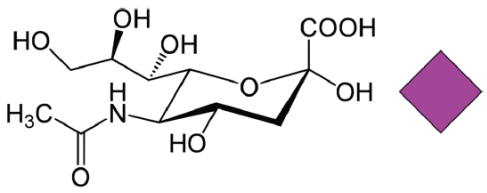


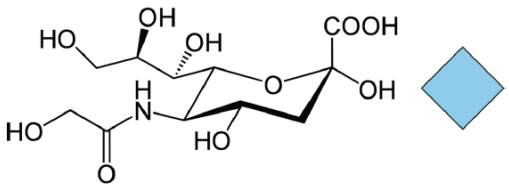
Figure 6.2. Ribbon representation of the crystal structure from P40/P90, with three 90° apart views. Extracellular N-terminal domain is composed by seven-blade β-propeller (light orange) from where emerges the crown (pink). The crown hosts the binding site for sialylated glycans (indicated in the Figure in light blue). The predicted C-terminal domain is also show in cyan. The sialic binding site found in P40/P90 is indicated by the star symbol in light blue. Figure adapted from reference [162].

6.3 Biological role of sialylated glycans in bacterial adhesion

Bacterial infection consists in a multi-factorial process, and it can be divided into several stages [164]; the first key step is the microorganism attachment to a target host cell. This contact is mediated by bacterial factors such as adhesins and a plethora of dedicated surface host structures. It is well known that the carbohydrates moiety and in particular their terminal sialic acid, are involved in several biological functions and may serve as points of adherence for bacterial cells [165]. Sialic acids (Sia) or neuraminic acids, which are found in abundance on cell surface of humans and other mammals, are a family of nonulosonic sugars which are usually linked to the terminal moieties of N-glycans, O-glycans and glycosphingolipids [166]. They share a nine-carbon sugar backbone and encompass over 50 natural derivates with different types of modifications of the sugar at various positions [167, 168]. The most abundant sialic acids found in mammals are N-acetylneuraminic acid (Neu5Ac) and N-glycolylneuraminic acid (Neu5Gc) (Figure 6.3) [169-171].



***N*-acetylneuraminic acid, Neu5Ac**



***N*-glycolylneuraminic acid, Neu5Gc**

Figure 6.3. Structure of the most abundant sialic acid species: N-acetylneuraminic acid (Neu5Ac) and N-glycolylneuraminic acid (Neu5Gc). On the right of the panel are reported the graphical representation of these glycans as per SNFG guidelines [172]. Figure from reference [165].

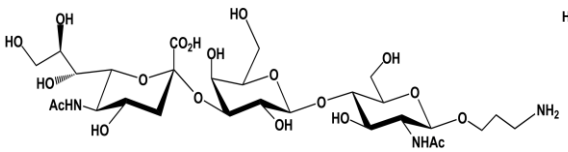
Sialic acid can modulate and mediates a variety of health and pathological process [171] including cell-cell adhesion, signalling, immunity regulation and discrimination of *self* from *non-self.* Is involved in tolerance mechanisms by presenting itself as determinant of *self* and, on the other hand, may also serve as primary receptor for infection. Some bacteria and viruses, like Mumps virus hemagglutinin-neuraminidase [173], bind to Sia-terminated glycans on the host cell surface, either as primary receptors for infection or by using them as accessory receptors in host cell binding uptake [174].

Mycoplasma pneumoniae and *genitalium* represent two emerging human pathogens leading to various inflammations. Given the high incidence of diseases related to these microorganisms, both in adults and children, and the increase in antibiotic resistance that makes current therapies ineffective, the development of alternative therapies is necessary. An alternative therapeutic approach could be based, for example, on inhibiting the bacterium adhesion to the host cell, counteracting mycoplasma-associated infections. Therefore, the detailed knowledge of the interactions processes behind adhesion and gliding is crucial for this purpose.
In this chapter, the recognition of different sialylated O- and N-glycans by the cytoadhesins P40/P90 in *Mpn* and P110 in *Mge* is investigated by using complementary techniques. In particular, to derive the ligand epitope maps of the ligand and to define their conformational behaviour NMR techniques such as STD NMR and tr-NOESY are used respectively. After all, the 3D models of the complexes were provided by applying computational approaches.

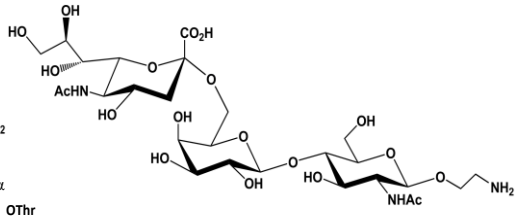
6.4 Sialoglycans used for interaction studies

Aiming to investigate glycans recognition mode of *M. pneumoniae* and *genitalium*, the interaction of two cytoadhesins, from *Mpn* and *Mge,* with α–(2,6)/α-(2,3)-linked sialic acid containing ligands, (Figure 6.4) was explored. In details, we investigated both *N*- and O-glycans: **ligand 1 and 2** (namely 3'SLn and 6'SLn, respectively), which are two trisaccharides representing the terminal end of common host N-glycans, and **ligands 4** and **5,** which are longer biantennary complex type N-glycans. **Ligand 3** (namely **sTa-Thr)** corresponds, instead, to *ad hoc* synthetized sialyl-T-antigen linked to a threonine residue.

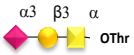
Ligand 1

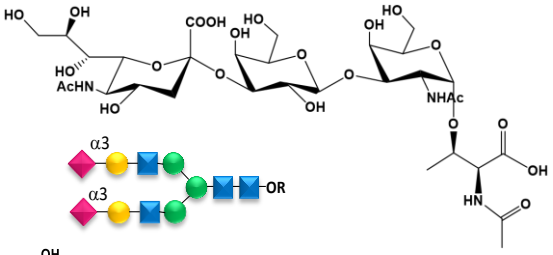


Ligand 2

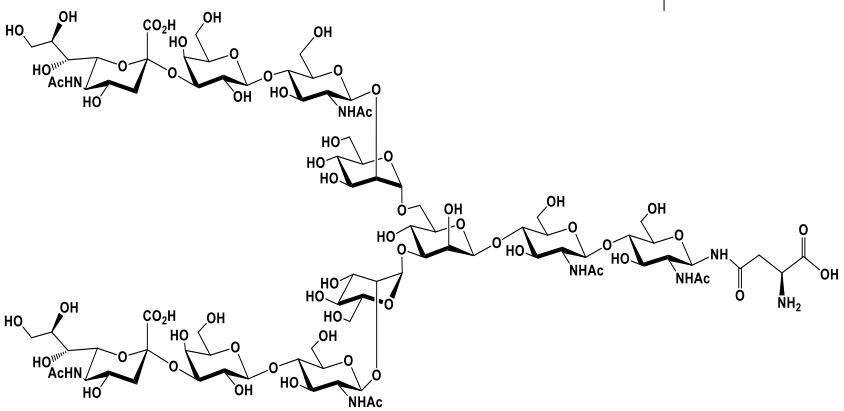


Ligand 3

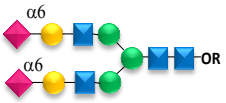




Ligand 4



Ligand 5



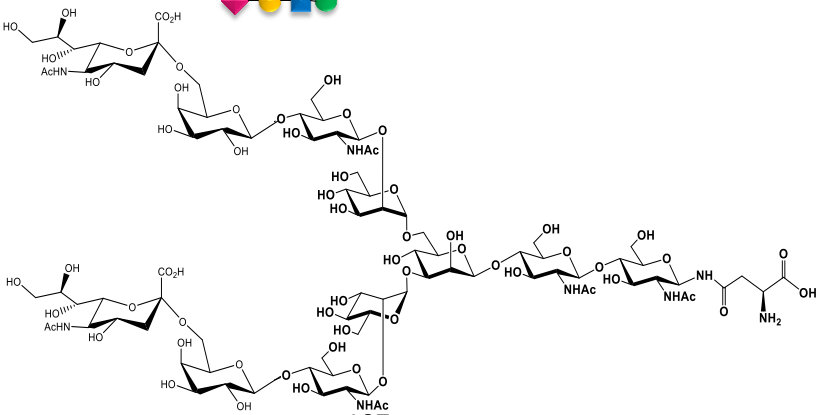


Figure 6.4. Structure of sialylated ligands employed in this chapter together with their SNFG representation. L**igand 1**, trisaccharide Sia-α-(2→3)-Gal-β-(1→4)-GlcNAc and **ligand 2**, trisaccharide Sia-α−(2→6)-Gal-β-(1→4)-Glc-NAc; the two trisaccharides 1 and 2 differ in the type of glycosidic linkage at the non-reducing end; **ligand 3**, trisaccharide sialyl T-antigen and **ligands 4** and **5** represent the longer glycans commonly exposed on mammalian cells.

6.5 Association measurement by fluorescence spectroscopy

To first determine the association between *mycoplasma* cytoadhesins, P110 (from *Mge*) and P40/P90 (from *Mpn*), and both *N*- and *O*-glycans typically decorating the host cell surface fluorescence spectroscopy experiments have been performed in collaboration with the colleague A. Marseglia. A titration of both *cytoadhesins* was conducted by adding increasing concentrations of ligands (3'SLn, 6'SLn and sTa-Thr) to a fixed concentration of both proteins. The binding isothermal curves for all protein-ligand mixtures were derived, obtaining information about the binding affinities. Florescence spectroscopy results and in particular the values of the binding constants (Kb) in the micromolar range, proved the ability of both P110 and P40/P90 *cytoadhesins* to similarly recognize the different ligands (Figure 6.5).

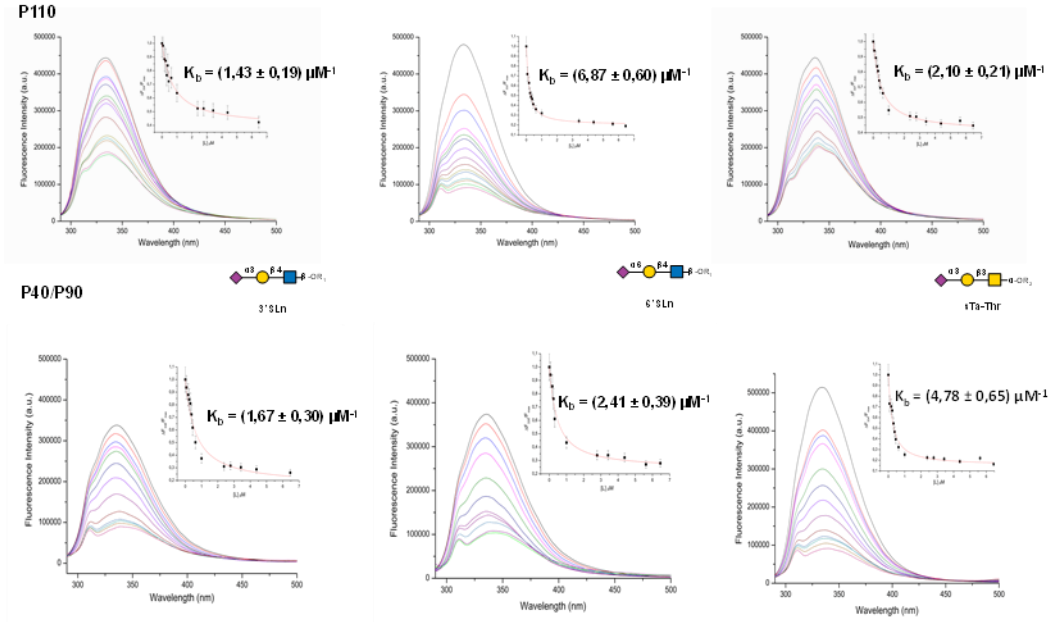


Figure 6.5: Fluorescence titration experiments. Fluorescence spectra and the binding isotherm with the Kb (binding constant) of Mge P110 (upper panel) and Mpn P40/P90 (lower panel) in the presence of increasing amounts of the ligands 1-3.

Prior performing analysis on the protein/ligand mixtures, experiments were acquired on both protein and ligand in the free state, in order to determine the optimal irradiation frequency.

6.6 Sialylated glycans recognition by P110 adhesin (*Mge*)

6.6.1 3'SLn recognition by P110

STD NMR analysis of the P110/3'SLn mixture clearly showed that the protein binding pocket was selective towards Neu5Ac moiety (Figure 6.6). As suggested by the STD enhancements, the entire Neu5Ac was recognised and received a good magnetisation transfer from the P110 receptor. The strongest relative STD effect belongs to the acetyl group of the Neu5Ac unit, followed by the H6 and H7 protons

with an STD percentage above 40%. Other protons of Neu5Ac (H4, H5, H8 and H9) showed STD signals but to a lesser extent. The other sugar residues, however, did not participate in the recognition process. These results are in agreement with the interactions observed in the crystal structure of P110 in complex with 3'SLn by Aparicio and colleagues [159].

Given the similar topology of the *cytoadhesins Mge* and *Mpn* with the neuraminidase proteins of influenza A viruses, the possible presence of a sialidase catalytic site close to protein binding pocket has been previously hypothesized [163]. However, experimental data did not suggest a catalytic activity in the crown region and a Neu5Ac hydrolysis was not observed from the ligand in the presence of the protein (data not shown).

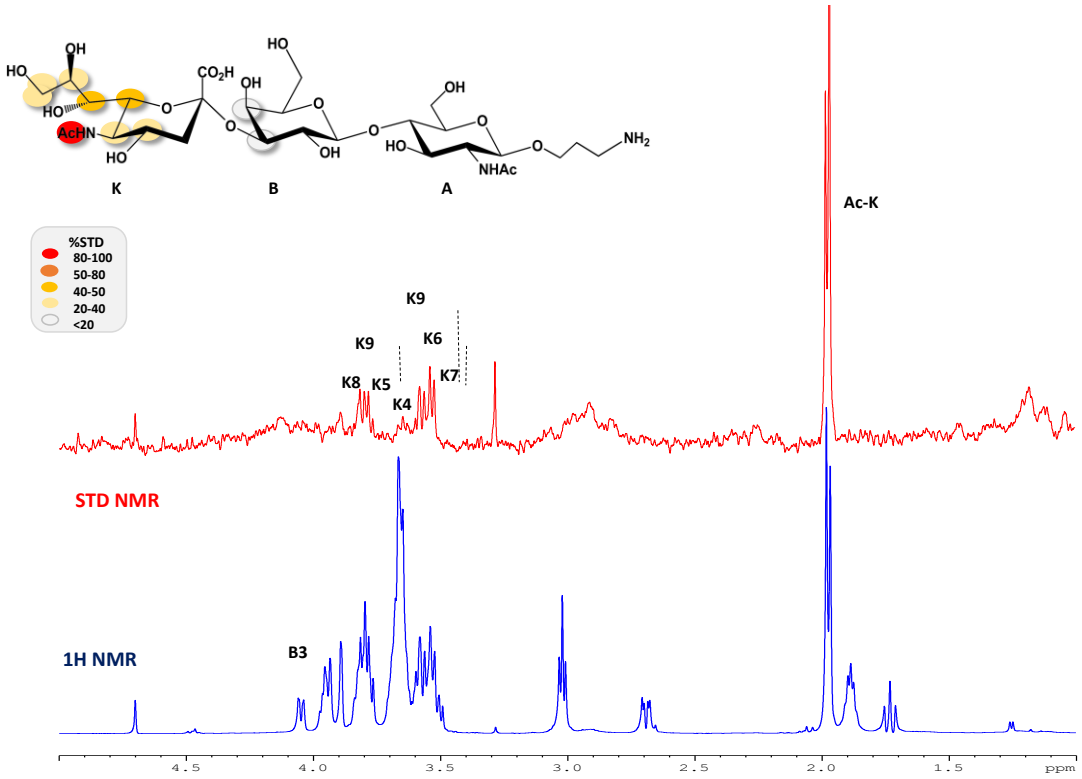


Figure 6.6. **NMR analysis of 3'SLn bound to P110.** STD NMR spectrum (red) and the unsaturated spectrum (blue) together with the epitope map of the ligand interacting with the P110, obtained by (I0-Isat)/I0, where (I0-Isat) represents the intensity of the signal in the STD NMR spectrum and the I0 is the signal intensity of the unsaturated spectrum. The STD percentages are obtained by normalizing with respect to the highest signal.

Tr-NOESY experiments, and MD simulations allowed to describe the ligand conformational behaviour. Firstly the glycan conformation in its free state was investigated and the results indicated that, in absence of the protein, the 3'SLn could adopt three main different conformations, namely *t*, *g* and *-g,* according to the different values of Φ(C1-C2-O-C3') dihedral angle around Neu5Ac-α-(2,3)-Gal glycosidic linkage [175]. While 3'SLn when accommodated into the P110 binding pocket, adopted preferentially the so called *t* conformation. These findings were revealed either by tr-NOESY experiments and by MD simulation. In particular, the key NOEs between the diastereotopic protons at position 3 of the Neu5Ac and the proton H3 of the Gal residue (Figure 6.7a) suggested a conformer selection upon binding. In addition, once modelled the 3'SLn into the P110 binding site, the MD simulation of 100 ns in explicit water was performed and the glycosidic torsion trajectories were monitored (Figure 6.7b).

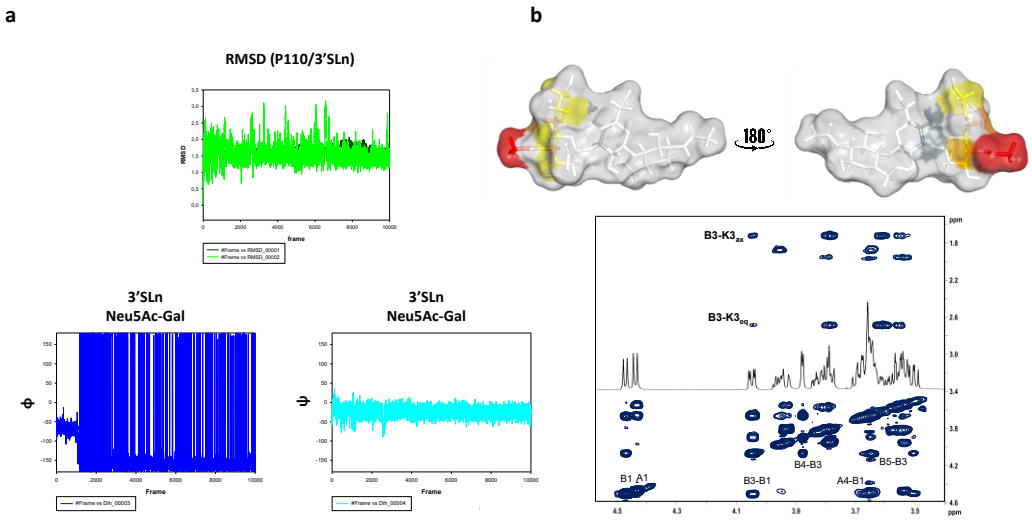


Figure 6.6. MD simulation and tr-NOESY analysis. Left panel: RMSD of the protein (black) and the ligand (red) are reported. The ligand RMSD was calculated in reference to the protein. 3'SLn dihedral angle ϕ and ψ analysis in the bound state with P110, during the MD simulation. RMSF measures the average amplitude of each atom's motions relative to a mean reference position. The higher value of RMSF correlates to a higher flexibility. Right panel: STD-derived epitope map of 3'SLn in its bioactive conformation and tr-NOESY spectrum of P110/3'SLn mixture.

MD simulation analysis also permitted to accurately describe the binding mode and the main molecular interactions established between the protein and the ligand (Figure 6.7). MD results were consistent with the STD NMR data and numerous hydrogen bonds were observed between the protein and Neu5Ac. In detail, H-bonds were established between Neu5Ac carboxyl group and Arg600 and Ser458. Asn200 exhibited two interactions with Neu5Ac: one with the OH at position 4 and the second with its acetamide moiety, while Pro197 formed a H-bond with the hydroxyl proton 7 of Neu5Ac. In addition, Asn475 established crucial H-bonds with hydroxyl protons 8 and 9 of the glycerol chain of Neu5Ac, which strongly contributed to

stabilize the complex and resulted in the flexible loop region (471-482 aa) moving closer to the Neu5Ac binding site. As shown in Figure 6.7. Notably, the acetamide group of the Neu5Ac was positioned in a hydrophobic pocket close to Phe457, contributing to stabilize the binding. Another significant contact was observed between the amide nitrogen of Neu5Ac and Ser456, forming a hydrogen bond, stable during the MD simulation, in accordance with the high STD contribution.

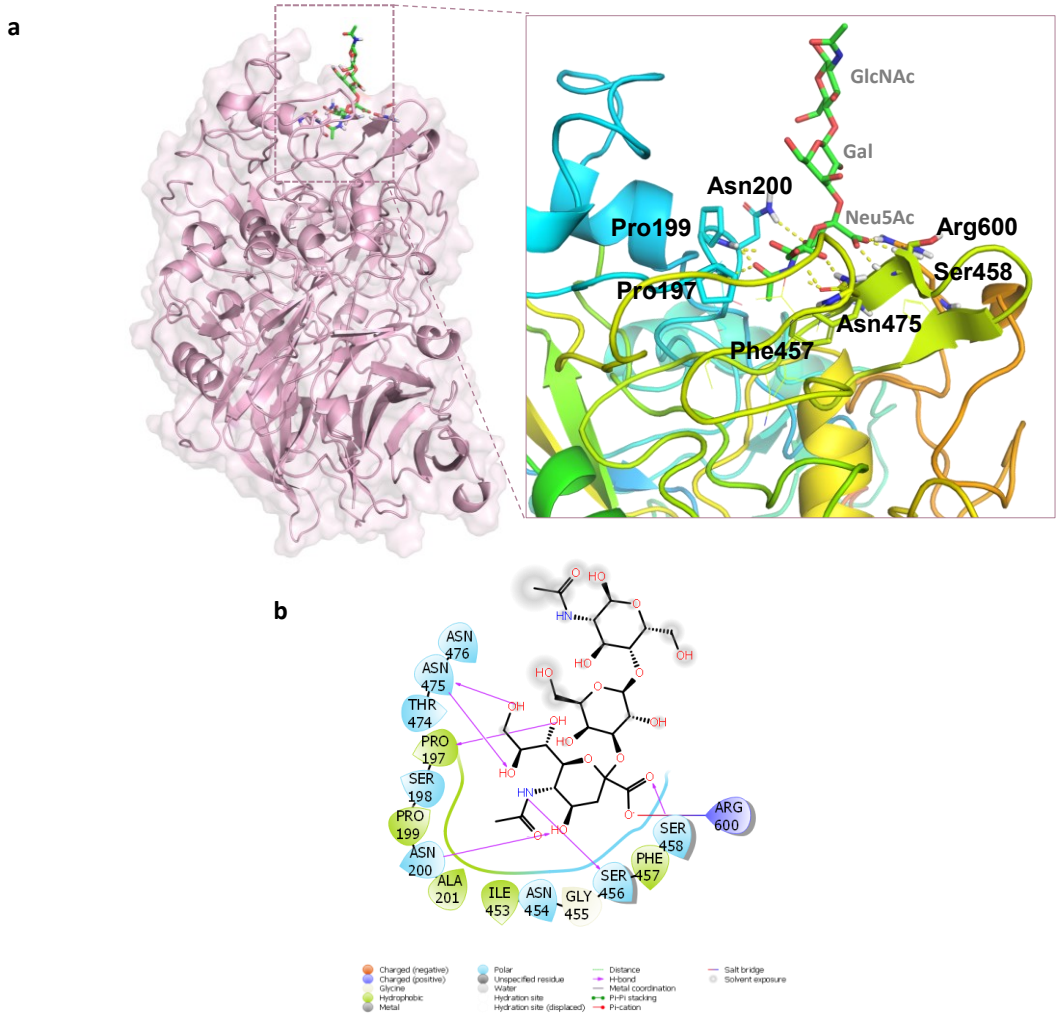


Figure 6.7. 3D view of the P110-3'SLn complex (a). The aa of the binding pocket involved in the binding with Neu5Ac are represented as sticks. The flexible loop moving closer to the binding site upon binding is represented in yellow. Two-dimensional plots highlighting the main protein-ligand interactions are also reported. Solid arrows represent hydrogen bonds with the functional groups of the backbone; the other residues in the binding pocket participate in polar and hydrophobic interactions (b).

6.6.2 6'SLn recognition by P110 (*Mge*)

The interaction mechanism between 6'SLn and P110 and the ligand conformational behaviour were also investigated by combining NMR experiments and computational approaches. As for the previous system, the protein selectively recognised the sialic acid residue. The STD NMR results, and the resulting epitope map (Figure 5.8) shown the highest magnetization transfer to the Acetyl group of Neu5Ac from the protein. Hydrogen at position 7 of Sia also exhibited a high STD enhancement close to 50%. A minor contribution was given by protons at position 5, 6, and protons 8 and 9 belong to glycerol chain.

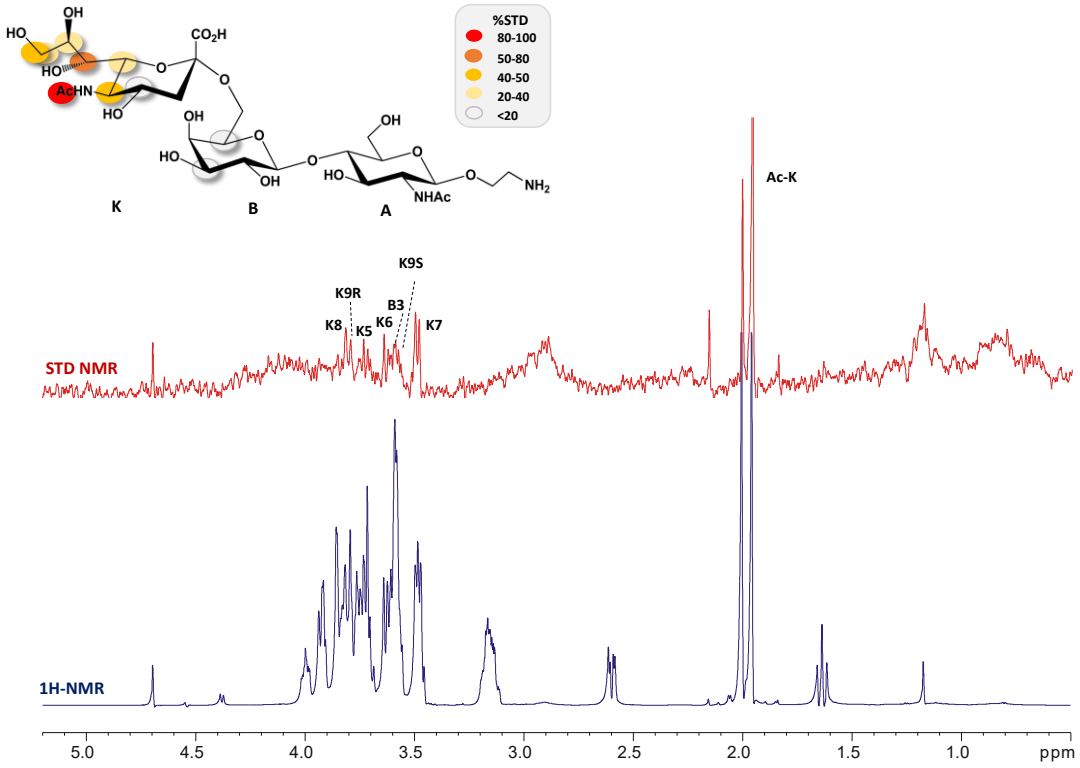


Figure 6.8. NMR analysis of 6'SLn bound to P110. STD NMR spectra (red) and the unsaturated reference spectrum (blue) together with the epitope map of the ligand interacting with P110 protein were reported together with the ligand epitope map showing the protons directly involved in the binding process. The STD percentage were obtained by (I0-Isat)/I0, where (I0-Isat) represents the intensity of the signal in the STD NMR spectrum and the I0 is the signal intensity of the unsaturated spectrum. The STD percentages are obtained by normalizing with respect to the highest signal.

STD experimental results were than supported with the molecular dynamics simulation data. MD simulation analysis showed numerous H-bonds established at the protein-ligand interface. Stable interactions above 90% of the simulation time, were settled between Pro197 and Neu5Ac hydroxyl group in position 7. Polar

interactions were also established between Ser456 and the amide group, and Asn200 interact with carbonyl oxygen of acetyl moiety and OH at position 4. Hydroxyl protons at position 8 and 9 of the glycerol chain of Neu5Ac established contacts with Asn475 (Figure 6.9). A crucial interaction also occurred between the hydroxyl group of Ser458 and carboxyl moiety of Neu5Ac. Other interactions, involving the GalNAc residue and some amino acids of P110 were observed in different clusters over the MD simulation.

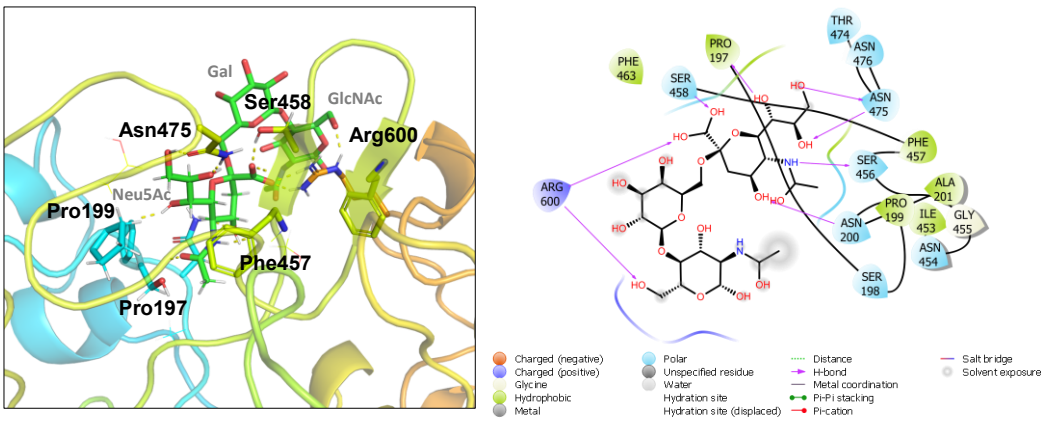


Figure 6.9. 3D view of the P110-6'SLn complex. The aa of the binding pocket involved in the binding with Neu5Ac are represented as sticks. Two-dimensional plots highlighting the main protein-ligand interactions are also reported. Solid arrows represent hydrogen bonds with the functional groups of the backbone; the other residues in the binding pocket participate in polar and hydrophobic interactions.

6'SLn, was characterized by the presence of an additional torsion angle named **ω** (O'6−C'6−C5'−O'5). Consequently, the ligand population distribution at the equilibrium in its free state was described by different rotamers namely ***gt***, ***tg*** and ***gg***, according to the ω angle values (60°, 180° and -60° respectively). While, in the

bound state, MD analysis shown that 6'SLn accommodated the protein binding site by adopting only two conformations, *tg* and *gt* (Figure 6.10).

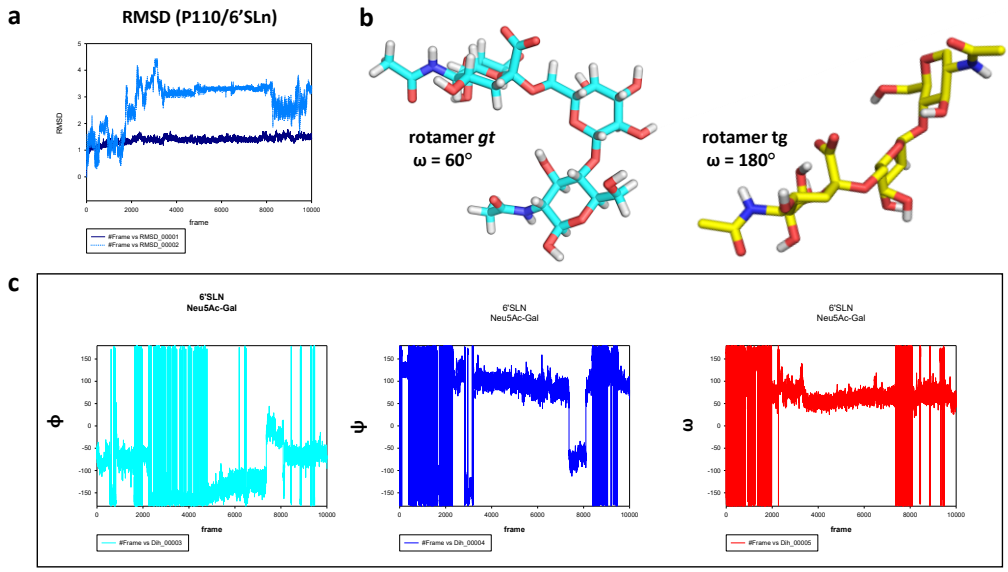


Figure 6.10. analysis of MD simulation of P110/6'SLn complex. a) RMSD of the protein (dark blue) and the ligand (blue) was reported on the left. The ligand RMSD was calculated in reference to the protein. b) The two selected ligand rotamers, *gt* and *tg*, during the binding process were shown. c) Panel showing the monitored dihedral angles φ, ψ and ω between the glycosidic linkage Neu5Ac-Gal.

6.6.3 sTa-Thr recognition by P110

The recognition mode of O-glycan was also investigated and in particular was studied the interaction mechanism of the complex sTa-Thr/P110. STD NMR experiment (Figure 6.11) reveal that the P110 protein recognized the O- linked sialyl-T-antigen. As with the previous ligands analysed, the binding was driven by the sialic acid recognition. The highest contribution to the binding comes from Neu5Ac, while the galactose unit contributes slightly and the reducing GalNAc, together with the Thr residue, did not receive any magnetization transfer from P110.

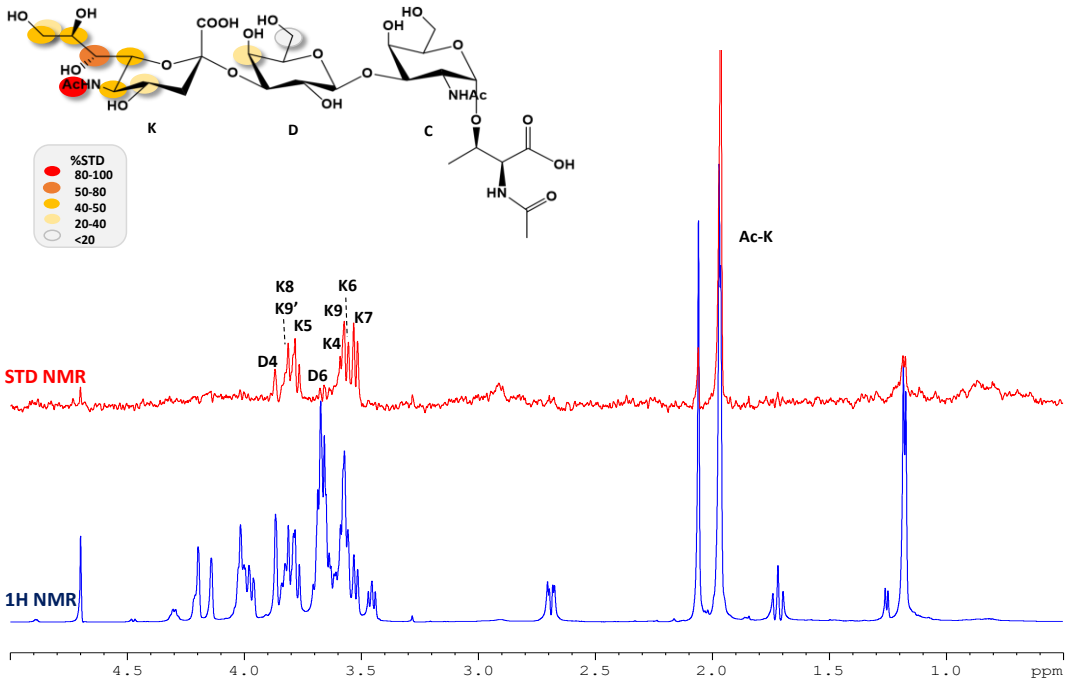


Figure 6.11. STD NMR analysis of sTa-Thr bound to P110. STD NMR spectrum (in red) and the unsaturated spectrum (in blue) are shown. The epitope map of the ligand was also reported to the top, showing the protons most involved during the binding.

STD analysis clearly showed that the acetyl group of Neu5Ac undergoes the highest magnetisation transfer. Proton at position 7 also shown a good degree of saturation close to 70% following by the H5, H6, H8 and H9 which displayed an STD relative percentage upon 50%. Conversely, the proton of Neu5Ac at position 4 together with the protons at positions 4 and 6 of galactose were less involved in the interaction sharing a STD enhancement less than 40%.

Once the ligand interacting epitope was assessed, the binding mode was also analysed through both computational studies and tr-NOESY experiments. The 100 ns MD simulation, performed in explicit water on the P110/sTa-Thr complex,

revealed that the protein binding pocket was the same as that observed for N-glycans and the main interactions occurred with Neu5Ac (Figure 6.12). In details, Pro197 interacted with hydroxyl group at position 7 and Ser456 with amide nitrogen of the sialic acid. The Asn200 residue provided an H-bond with OH at position 4 of Neu5Ac and Ser458 established two interactions with the carboxyl group of this residue as also observed in the crystal structures [159, 162].

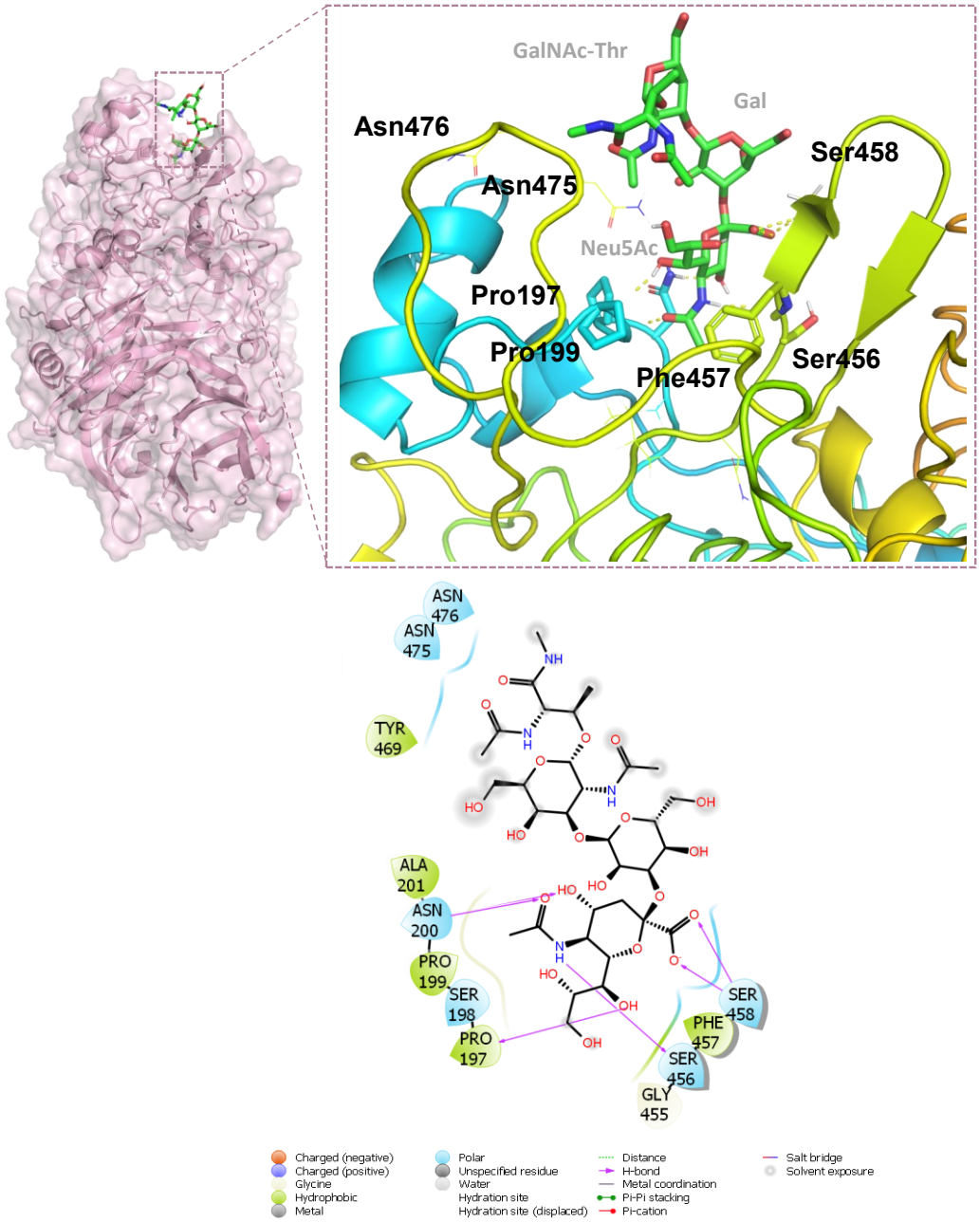


Figure 6.12. 3D model of the P110/sTa-Thr complex. The amino acids involving in the binding with Neu5Ac were shown as stick. To the bottom, the 2D plot showing the main

protein-ligand interactions was reported. Solid arrows represent hydrogen bonds with the functional groups of the backbone; the other residues in the binding pocket participate in polar and hydrophobic interactions.

To finalise the depiction of the complex, the conformational behaviour of the ligand in both the free and bound state was also studied. First estimation was performed *via* tr-NOESY experiment and then the trajectories of the sTa-Thr glycosidic torsion angles were monitored during the entire MD simulations in free and bound state. The ligand conformation was influenced by the glycosidic torsion angles namely ϕ (C1-C2-O-C3')/ψ (C2-O-C3'-H3') around Neu5Ac-α-(2,3)-Gal and ϕ (H1–C1–O–C3')/ ψ (C1–O–C3'–H3') around the Gal-β-(1,3)–GalNAc bonds. In the bound state with P110, a ligand conformer selection was observed and only the minimum characterized by the value of ϕ torsional angle at -60° resulted to be populated (Figure 6.13).

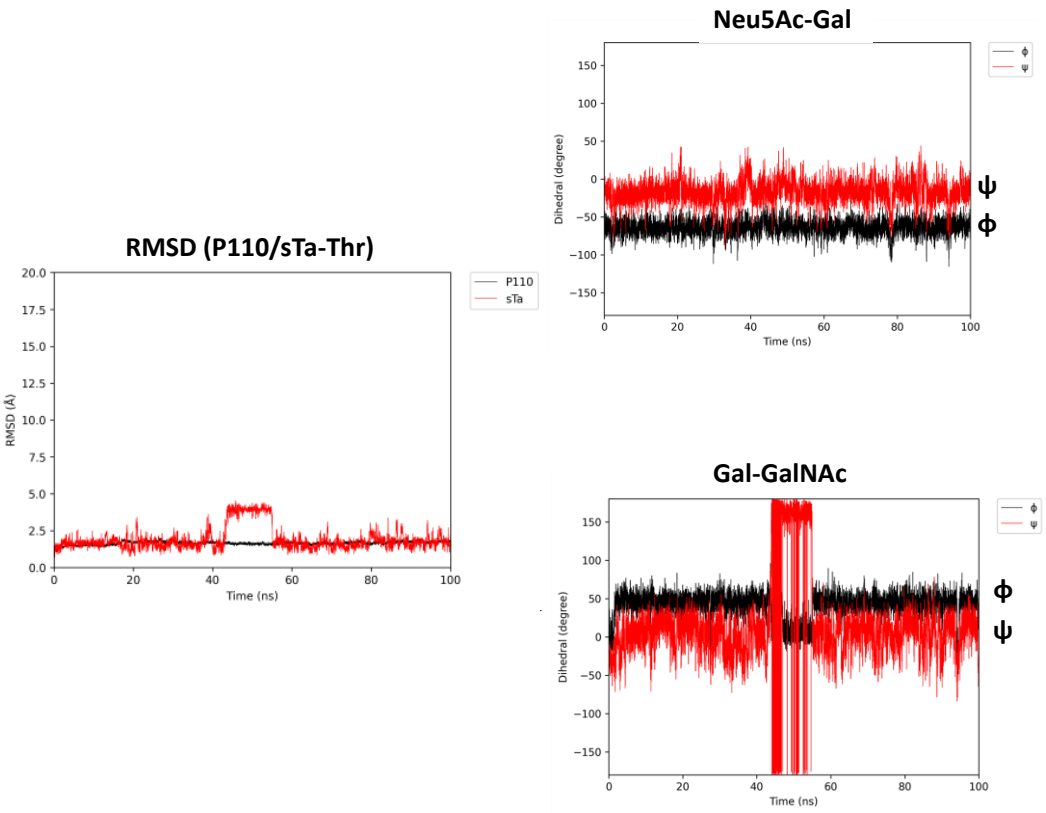


Figure 6.13. MD analysis of P110 in the interaction with sTa-Thr. RMSD of the protein (black) and the ligand (red) are reported on the left. The ligand RMSD was calculated in reference to the protein. sTa-Thr dihedral angles analysis in the bound state with P110 presented as scatter plots of the ϕ torsion against ψ, during the MD simulation.

6.7 Adhesion mechanisms of P40/P90 (*Mpn*) to sialylated glycans

Similarly to P110, *cytoadhesin* P40/P90 from *Mpn* recognises sialylated glycans and to depict the binding events underlying these interaction, several NMR experiments and *in silico* studies were performed.

6.7.1 Adhesion mechanism of P40/P90 to 3'SLn

The association between P40/P90 and the trisaccharide 3'SLn was first analysed by STD NMR experiments. *Mpn cytoadhesin* showed a selective binding mode which involve the sialic acid moiety. The highest STD enhancement was observed for the acetyl group of Neu5Ac, followed by the proton at position 7 with an STD effect above than 50%. Protons H5, H6 and H9 from Sia unit contributed less to the binding displaying an STD percentage within the range of 50%-20% (Figure 6.14). Molecular dynamic simulations, corroborated the NMR experimental data, showed that the Neu5Ac residue was well accommodated in the P40/P90 binding pocket, otherwise the galactose and N-acetylglucosamine residues were located farther from the protein binding site (Figure 6.15). In details, in the main representative pose, a contact occurred between the amide group at position 5 of Neu5Ac and the Leu630, stable for about 70% of MD simulation time. The Neu5Ac carboxyl group also established H-bonds with the side chain of Thr632. Hydroxyl group in position 7 interacted with Pro235 and the oxygen of the acetyl moiety established a contact with Lys238. These findings were consistent with the STD data. In addition, the presence of hydrophobic residues close to the binding site, such as Phe631, promoted the stability of the complex by anchoring the residue to the protein via hydrophobic contacts.

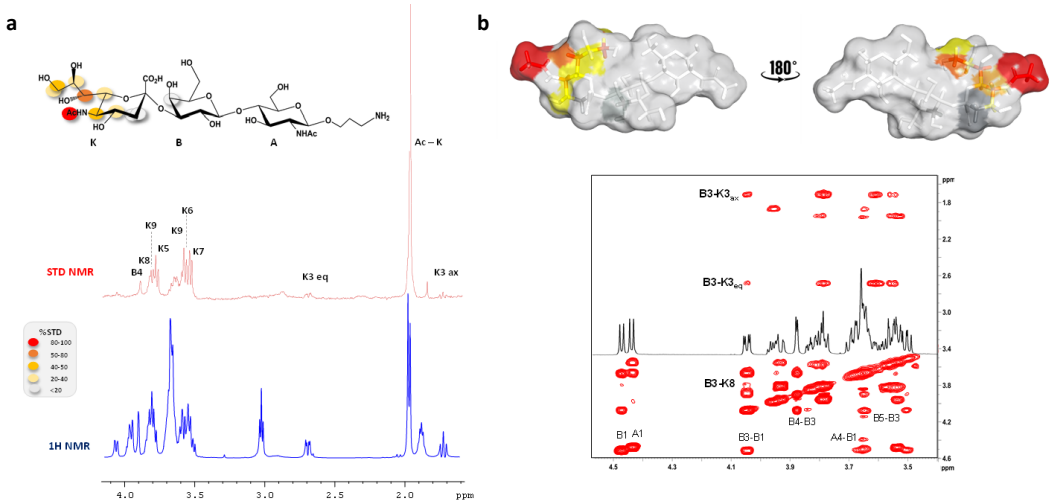


Figure 6.14. NMR analysis of 3'SLn bound to P40/P90. STD NMR spectrum (in red) and unsaturated spectrum (in blue) together with the epitope map of 3'SLn interacting with P40/P90 are reported on the left. The STD-derived epitope map of 3'SLn in its bioactive conformation reporting the surface coloured according to the STD effect and Tr-NOESY spectrum of P40/P90 mixture were also reported on the right.

The conformational behaviour was also evaluated. In accordance with the tr-NOESY experiment, MD simulation showed that upon binding the 3'SLn joined the binding pocket adopting two different conformations *t* and *-g* with Φ dihedral angle between the glycosidic linkage Neu5Ac-Gal, respectively 180° and 0°.

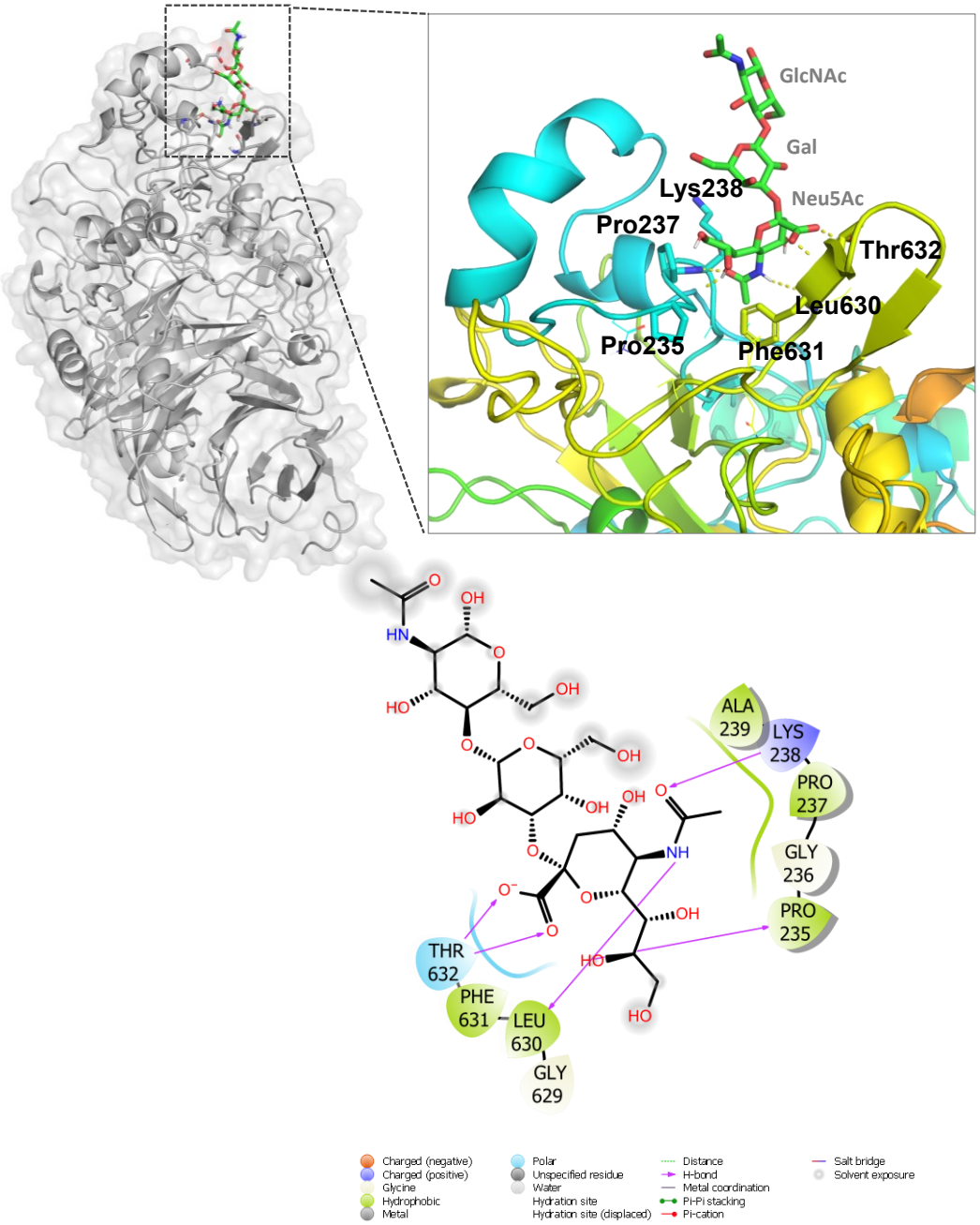
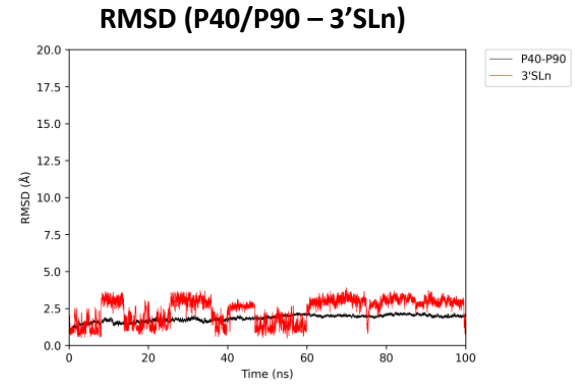
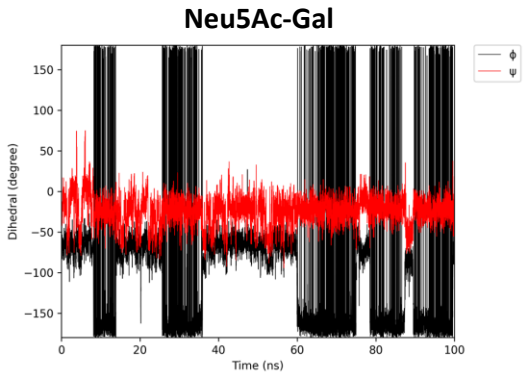
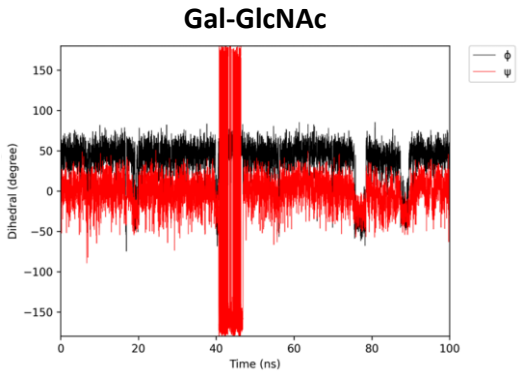


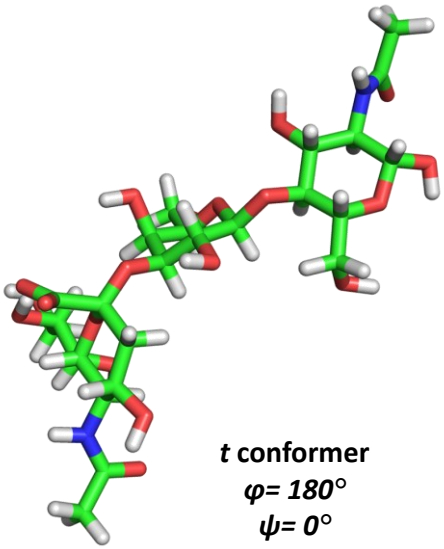
Figure 6.15. P40/P90–3'SLn complexes. To the top was reported the 3D view of the P40/P90-3'SLn complex. Two-dimensional plots highlighting the main protein-ligand

interactions were also reported to the bottom. Solid arrows represent hydrogen bonds with the functional groups of the backbone; the other residues in the binding pocket participate in polar and hydrophobic interactions.









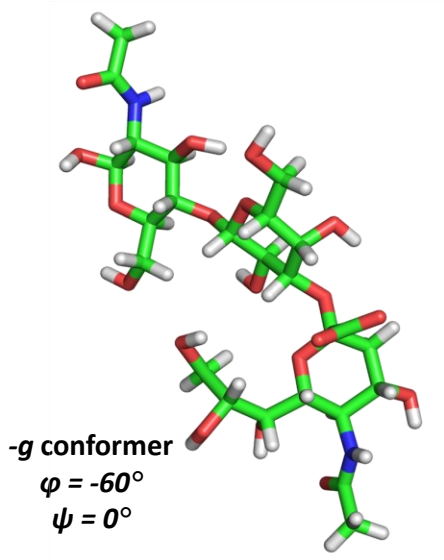


Figure 6.16. MD analysis of the P40/P90 - 3'SLnv interaction. RMSD of the protein (black) and the ligand (red) are reported. The ligand RMSD was calculated in reference to the protein. 3'SLn dihedral angles analysis in the bound state with P40/P90 represented as scatter plots of the ϕ torsion against ψ, during the MD simulation. The two main conformational families the 3'SLn found in the bound state.

6.7.2 Adhesion mechanism of P40/P90 to 6'SLn

The binding mode involving the trisaccharide 6'SLn and the *cytoadhesin* P40/P90 from *Mpn* was also investigated and STD NMR results, as for the above systems, indicated a specific recognition for the Neu5Ac residue. In fact, in STD spectrum (Figure 6.17) a highest magnetisation transfer was observed for the acetyl group of Sia residue. Afterwards, protons H6, H5 and H7 displayed a high STD enhancement, above 50%, while the glycerol chain and proton in position 4 less contributed to the binding. The two other units of galactose and N-acetylglucosamine did not undergo magnetisation transfer from the protein thus indicating they did not contribute to the binding.

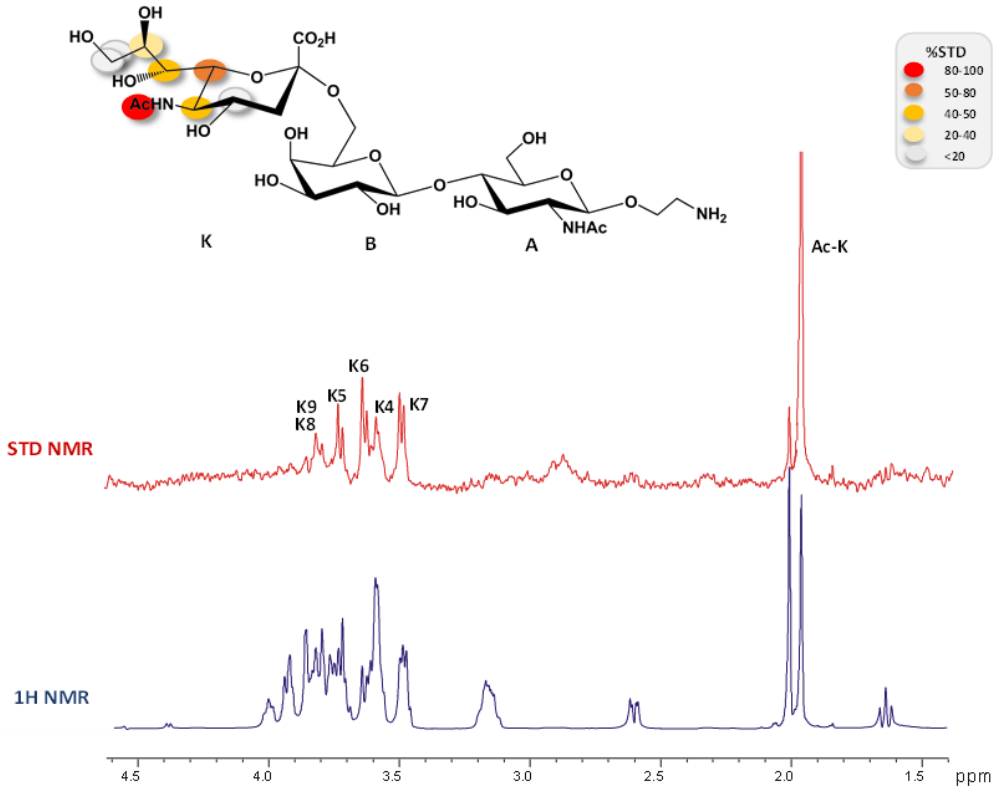


Figure 6.17. NMR analysis of 6'SLn bound to P40/P90. STD NMR spectra (red) and the unsaturated reference spectrum (blue) together with the epitope map of the ligand interacting with P40/P90 are reported together with the ligand epitope map showing the protons directly involved in the binding process. The STD percentages were obtained by (I0-Isat)/I0, where (I0-Isat) represents the intensity of the signal in the STD NMR spectrum and the I0 is the signal intensity of the unsaturated spectrum. The STD percentages are obtained by normalizing with respect to the highest signal.

Experimental data were then confirmed by computational results. MD simulation performed on the complex involving the protein P40/P90 and 6'SLn clearly demonstrated the selective engagement of the Neu5Ac into the binding and the absence of interactions between the protein's amino acids and the galactose and N-

acetylglucosamine residues (Figure 6.18). In particular, a stable H-bonds were observed between the acetamide moiety and Lys238 and Leu630, confirming the experimental finding. Pro235 established an interaction with hydroxyl group at position 8 of the glycerol side chain. In addition, the Sia carboxyl group engaged a polar interactions with Thr632, consistent with the crystal structure of P40/P90 in complex with 3'SLn [160].

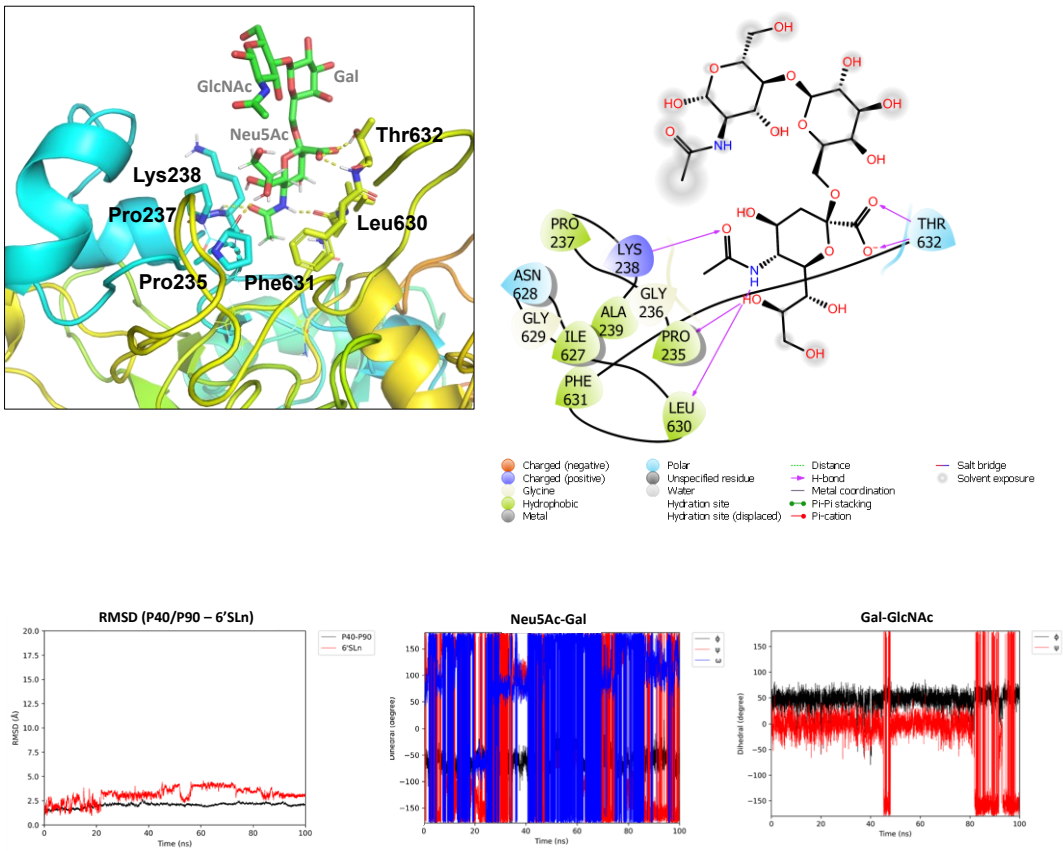


Figure 6.18. 3D view of the P40/P90-6'SLn complex (upper panel). The aa of the binding pocket involved in the binding with Neu5Ac are represented as sticks. Two-dimensional

plots highlighting the main protein-ligand interactions are also reported. Solid arrows represent hydrogen bonds with the functional groups of the backbone; the other residues in the binding pocket participate in polar and hydrophobic interactions. In the lower panel, the MD analysis of P40/P90 in the interaction with 6'SLn. RMSD of the protein (black) and the ligand (red) are reported. The ligand RMSD was calculated in reference to the protein. 6'SLn dihedral angles analysis in the bound state with P40/P90 represented as scatter plots of the ϕ torsion against ψ or ω, during the MD simulation.

The MD simulation analysis also revealed greater ligand flexibility upon binding than observed for the complex with P110. Indeed, when bound to P110, 6'SLn adopted preferentially the *gt* conformation (with ω torsion angle of 60°), while in complex with P40/P90 the ligand was in equilibrium between different rotamers.

6.7.3 Adhesion mechanism of P40/P90 to sTa-Thr

Similarly to what observed for P110-sTa-Thr complex, a high magnetisation transfer was achieved by the ligand complex with P40/P90. The STD NMR experiment showed a high STD effect for the acetyl group and the proton at position 7 of Neu5Ac. An STD percentage greater than 50% was observed for protons in position 4, 5 and 9, while for H8, H6 and H3 belonging to the sialic acid an STD effect lower than 50% was found. Differently from the complex with P40/P90, sTa-Thr participated to the binding by engaging also the galactose unit. A slight STD effect (< 20%) was actually observed for H4, H3, H6 and H2 of galactose (Figure 6.19).

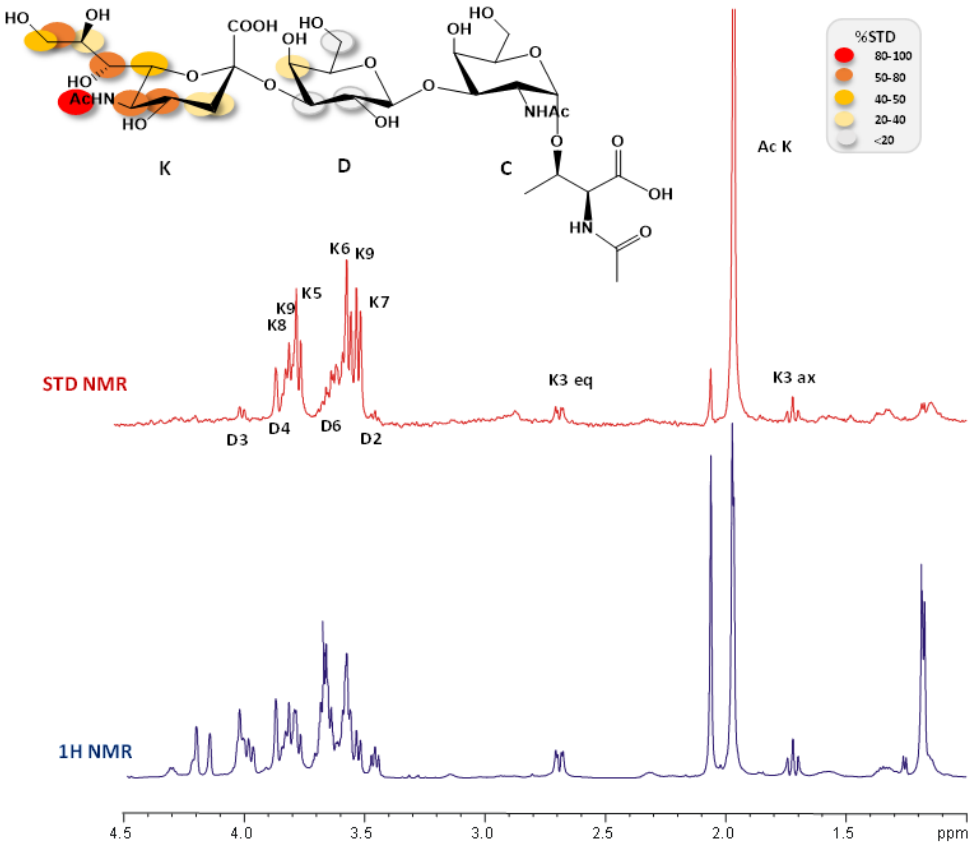


Figure 6.19. NMR analysis of sTa-Thr bound to P40/P90: STD NMR spectra (red) and the unsaturated reference spectrum (blue). The epitope map of the ligand interacting with P40/P90 are reported together with the ligand epitope map showing the protons directly involved in the binding process. The STD percentages were obtained by (I0-Isat)/I0, where (I0-Isat) represents the intensity of the signal in the STD NMR spectrum and the I0 is the signal intensity of the unsaturated spectrum. The STD percentages are obtained by normalizing with respect to the highest signal.

The conformational behaviour of the ligand was studied by tr-NOESY and *in silico* studies. In contrast to the system sTa-Thr/P110, here no significant conformational differences were found between the free and bound state. sTa-Thr accommodated in

the protein binding pocket adopting preferentially two conformations. By analyzing the MD simulation data, indeed, two main conformational families were observed: the first one was characterized by φ torsion angles of 180° along the Neu5Ac-α-(2,3)-Gal unit, while the second conformation was characterized by φ torsion angles of -60° (Figure 6.20).

The MD simulation of P40/P90 in complex with sTa-Thr also confirmed the selective recognition towards Neu5Ac (Figure 6.20). Several contacts were found between the aa of the protein and the sialic acid unit. In detail, the most stable contact occurred between Leu630 and the amide group of Neu5Ac, in complete agreement with the high STD value observed. The carboxyl moiety of Neu5Ac also provided stable H-bonds with Thr632. Other significant contacts were observed for the hydroxyl group at C7 of Neu5Ac, forming a hydrogen bond with Pro235, which was stable during the MD simulation, supporting the high STD contribution of H7 (about 80%). Lastly, Lys238 interacted with the acetyl group, forming a hydrogen bond.

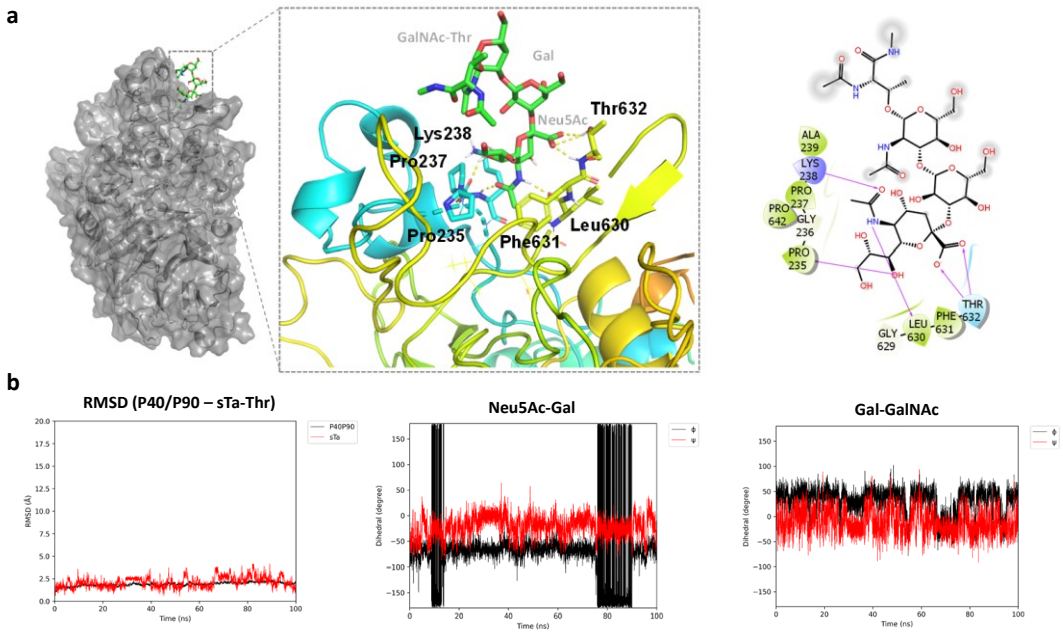


Figure 6.20. MD simulation analysis. a) 3D view of the P40/P90-sTa-Thr complex. The aa of the binding pocket involved in the binding with Neu5Ac are represented as sticks. Two-dimensional plots highlighting the main protein-ligand interactions are also reported. Solid arrows represent hydrogen bonds with the functional groups of the backbone; the other residues in the binding pocket participate in polar and hydrophobic interactions. **b)** RMSD of the protein (black) and the ligand (red) are reported. The ligand RMSD was calculated in reference to the protein. sTa-Thr dihedral angles analysis in the bound state with P40/P90 represented as scatter plots of the ϕ torsion against ψ or ω, during the MD simulation.

6.8 Recognition of complex type N-glycans by P110 and P40/P90

In order to further investigate the recognition of N-glycans by P110 and P40/P90, longer ligands mimicking natural branched-type complex N-glycans containing the Neu5Ac-α-(2,3)-Gal or Neu5Ac-α-(2,6)-Gal epitopes at the terminal end were also studied. Remarkably, STD NMR experiments conducted on the branched undecasaccharides (ligand 4 and 5 in Figure 6.4) demonstrated selective protein recognition for Neu5Ac residues, whereas the other ligand units were not recognized by the proteins during the binding process. Like for the smallest sialylated glycans, the acetyl moiety shown a highest STD effect, proving to be the driving force of the interaction. Other protons of Neu5Ac sugar backbone, in particular H7 and H6 displayed a high STD enhancements. On the contrary, STD spectra did not shown signals belong to other sugar units, except for some protons of the galactose residue close to the Neu5Ac, suggesting a minor contribution to the binding.

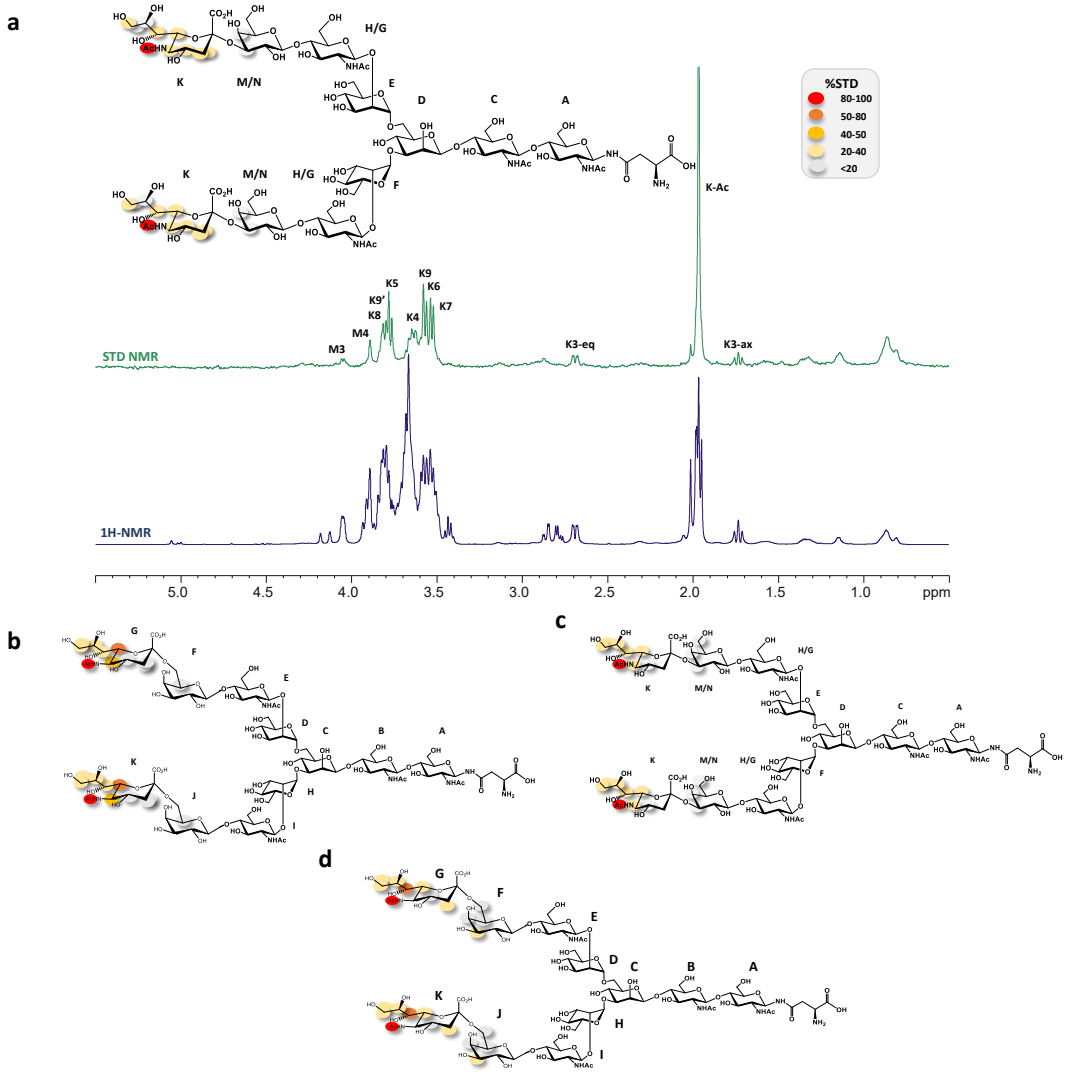


Figure 6.21. NMR analysis of complex N-glycans bound to P40/P90 and P110. a) STD experiments of complex involving P40/P90 and undecasaccharide linked α-(2,3) together with the ligand epitope map; b) epitope map of the undecasaccharide linked α-(2,6) bound to P40/P90; c) and d) epitope maps of the N-complex glycans lined α-(2,3) and α-(2,6) respectively bound to P110. The STD percentages were obtained by (I0-Isat)/I0, where (I0-Isat) represents the intensity of the signal in the STD NMR spectrum and the I0 is the signal

intensity of the unsaturated spectrum. The STD percentages are obtained by normalizing with respect to the highest signal.

However, given the symmetry of the glycan structure, it was not possible from the STD analysis to determine whether only one or both arms were involved in the interaction. Deeper insights were thus obtained by the MD analysis. Once docked the ligand into the protein binding pocket, a 100 ns MD simulation was run. Interestingly, MD results showed that both branches were recognized by the protein. Figure 6.22 clearly shows how the ligand could be anchored to the protein surface by using the sialic acids from both chains. In particular, one antenna mainly established polar contacts between Neu5Ac and Leu634 and Thr632 and hydrophobic interactions involving Phe631, as previously observed for the trisaccharide ligand. The second glycan branch was anchored to the protein by establishing additional interactions between the Neu5Ac unit positioned at the terminal end and the amino acid residues Arg226 and Ser227.

NMR and *in silico* results allowed us, for the first time, to model a 3D view of the interaction between the cytoadhesins *Mge* and *Mpn* and the biantenary N-glycan complexes.

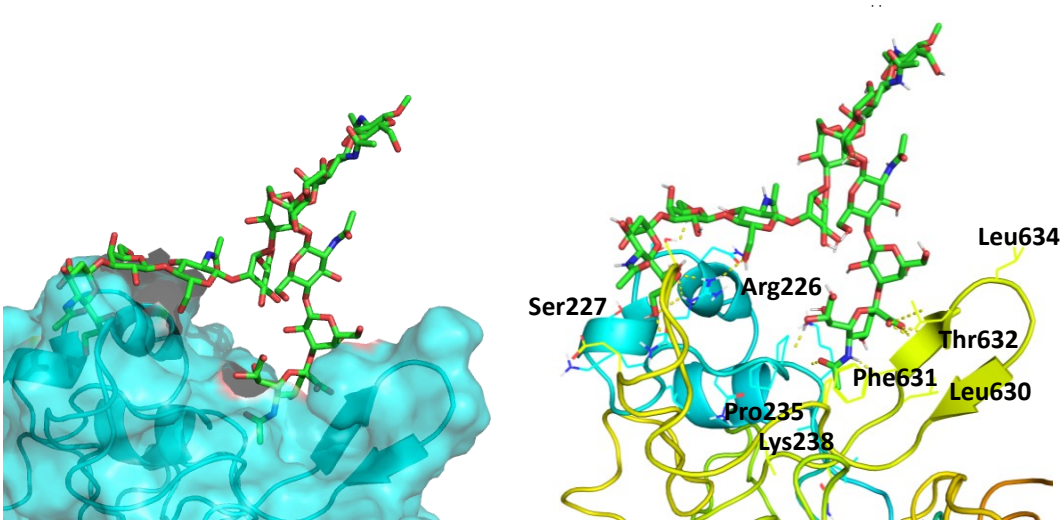


Figure 6.22. MD simulation analysis of the longer N-glycan bound to P40/P90.

6.9 Discussion

Sialylated glycans, which are highly abundant on vertebrate cell surfaces, are involved in a plethora of biological functions relevant to both health and disease. They act as essential modulators in cell signalling, cell-cell recognition and therefore regulate a wide range of immune responses. It is known that many pathogens, including several species of *Mycoplasma*, have evolved the ability to exploit host sialylated glycans to facilitate adhesion and infect into target cells. Among *Mycoplasma* species, *M. pneumoniae* and *M. genitalium* interact with host sialoglycans found in the eukaryotic mucosal surfaces to adhere and glide into the host cells initiating bacterial infection [176-178].

In the present chapter, a detailed description of the interaction mechanism underlying the adhesion between the two cytoadhesins P110 and P40/P90 expressed by *M. genitalium* and *M. pneumonia*, respectively, and several sialylated N- and O-glycans has been reported. By means of STD NMR analysis, epitope maps of the ligands

were obtained, and a selective recognition of sialic acid by both proteins was observed for all systems analysed. The tr-NOESY experiments together with the results obtained from the computational studies allowed to study the bioactive conformation of the ligands and to define 3D models for all complexes, providing a good starting point for new drug development.

SECTION IV – EXPERIMENTAL SECTION

Chapter 7 – Experimental section

7.1 Materials

7.1.1 *Neisseria gonorrhoeae* FA1090 wt bacterial cells grow

N. gonorrhoeae wt was grown in Gonococcal (GC) medium base (Proteose Peptone No. 3 15g/L, Corn Starch 1g/L, Dipotassium Phosphate 4g/L, Monopotassium Phosphate 1g/L, Sodium Chloride 5g/L, Agar 10g/L). A pre-inoculum was done in a 250 mL flask using 60 mL of sterile medium base and 1.8 mL of inoculum. A second growth was conducted using a 2L flask with 500 mL of GC complete medium and 55 mL of preculture. The incubation step was performed at 37°C, 160 rpm overnight. The final liquid culture arrived to an OD600 nm of 2.28 and centrifuged at 4°C, 8000 rpm for 30 minutes. The pellet was collected and resuspended in sterile water and inactivated. For inactivation, the pellet was kept at 99°C for 15 minutes, at -20°C for 5 minutes and at 4°C for 30 minutes. The pellet inactivation was then verified by plating the material on GC + isoviatlex plates.

7.1.2 LOS extraction and purification (SECTION II – chapter 4)

The freeze-dried cells were extracted according to the petroleum ether-chloroform-phenol (PCP) and hot phenol-water protocols as previously described [179-181]. Each phase was dialyzed against distilled water, freeze-dried and analysed by 12% and 13.5 % SDS-PAGE followed by silver nitrate staining [182] revealing the presence of lipooligosaccharide material. The phases containing the LOS were purified by enzymatic digestion with DNAse, RNAse and proteinase K as previously described [183, 184], and dialyzed against distilled water to remove cell contaminations.

In order to perform the NMR experiment a mild acid hydrolysis was performed with acetic acid 1% (100°C, 3 h) in order to separate the oligosaccharide (OS) from the lipid A portion. The resulting oligosaccharide fraction, obtained after centrifugation (14000 rpm, 30 min, 4°C), was further purified using size exclusion chromatography on Biogel P6, from which oligosaccharide OS was obtained. The eluate was monitored by a refractive index detector (Knauer GmbH – WellChrom Differential Refractometer K-2301) and all fractions were checked by NMR spectroscopy and pooled to separate the OS. Prior NMR analysis OS fraction was subjected on Porous Graphitic Carbon (PGC) Solid Phase Extraction (SPE) (Hypercarb™ SPE Cartridges Thermo Scientific™).

7.1.3 Proteins production (SECTION III – Chapter 5)

For the Ectodomain Region of P110: The region corresponding to the MG_192 gene from Mge (strain G37, residues 23–938) was amplified from a synthetic clone (Genscript) using forward primer P110F and reverse primer P110R. These PCR fragments were subsequently cloned into the pOPINE54 expression vector (gift from Ray Owens, Addgene plasmid #26043) to create a C-terminal Histidine-tagged protein.

The recombinant protein was produced through over-night expression in B834(DE3) cells (Merck) at 20°C after induction with 1 mM IPTG at an OD600 nm of 0.6. Cell lysis was carried out in 1× PBS buffer *via* sonication. After-ward, the cell extract was centrifuged at 20 rpm at 4°C, and the supernatant was applied to a 5 mL Histrap column (GE Healthcare) equilibrated with 1× PBS as a binding buffer and 1× PBS with 500 mM imidazole as elution buffer. Soluble aliquots of His6-tagged P110 were pooled and loaded onto a HiLoad Superdex 200 16/60 column (GE Healthcare) in a buffer containing 50 mM Tris pH 7.4 and 150 mM NaCl.

For the Ectodomain Region of P40/P90: The ectodo-main region of P40/P90 (23-1114) was amplified from the synthetic clone of the MPN_142 gene (GenScript) from Mpn using forward primer P40P90_F and reverse primer P40P90_R. The PCR fragment was subsequently cloned into the pOPINE expression vector. The recombinant protein was obtained after expression in B834(DE3) cells (Merck) induced with 0.8 mM IPTG at 22°C over-night. The cell pellets were lysed in a buffer containing 40 mM Imidazole and 30 mM Tris-HCl at pH 7.4 (binding buffer) and then centrifuged at 20,000 rpm at 4°C. The supernatant was loaded onto a HisTrap 5 mL column (GE Healthcare) and eluted with a buffer containing 400 mM imidazole. Soluble aliquots were concentrated and loaded onto a Superdex 200 GL 10/300 column (GE Healthcare) pre-equilibrated with 20 mM Tris·HCl buffer at pH 7.4 and 150 mM NaCl.

7.1.4 Synthetic glycoconjugates (SECTION III – Chapter 5)

The 3'-sialylactosamine (3'-SLn) and 6'-sialylactosamine (6'-SLn) were purchased from Tokyo Chemical Industry Co., Ltd. The sialyl-T-antigen linked to the threonine was chemically synthesized.26 The biantennary N-glycans were provided from GlyTech, Inc.27.

7.2 Methods

7.2.1 Western blotting (SECTION II – Chapter 4)

Mouse MAb 2C7 was provided by Professor Sanjay Ram (Division of Infectious Diseases and Immunology, University of Massachusetts Medical School).

0.5 mg of extracted and purified LOS was suspended in 0.5 mL of distilled water. Resulting solution was diluted 1:100. 3 µL of DTT and 3 µL of loading buffer (LDS Sample buffer 4X) were added to 5 µL of LOS solution. The solution was placed at 90°C for 1 min and then loaded on SDS-PAGE. Sample was run on 16% Tris-glycine

SDS-PAGE gel by using a Tris-glycine buffer. The marker was Ultra-Low range and consists of the following protein markers: Bradykinin (1060 Da), Insulin Chain B (3496 Da), Aprotinin (6500 Da), α-Lactalbumin (14200 Da), Myoglobin (17000 Da) and Triosephosphate Isomerase (26600 Da). LOS was transferred to nitrocellulose membranes (iBlotTM Kit Thermofisher) and membranes were blocked with PBS+ BSA 2% (v/v) + Tween20 0.1% for 1h at room temperature. Anti-LOS MAbs (diluted 1:1000 in PBS 1x + Tween20 0.1%) were incubated with membranes for 1h at RT on stirring plate. MAb-reactive LOS bands were visualized with anti-mouse IgG alkaline phosphatase (AP) (diluted 1:2000 in PBS 1x + Tween20 0.1%) incubated for 30 minutes at room temperature, followed by AP Conjugate Substrate kit (Biorad) for 5 minutes at room temperature.

7.2.2 Fluorescence microscopy (SECTION III-chapter 5)

Steady-state fluorescence analysis was performed on a Fluoromax-4 spectrofluorometer (Horiba Scientific, Edison, USA). All the measurements were performed at fixed temperature of 25°C. The emission spectra were recorded in the emission range of 290 and 500 nm upon excitation at 280 nm. The slits widths were set to 5 nm for both the excitation and emission monochromators. All spectra were recorded by using a quartz cuvette with a path length of 1 cm after an equilibration time of 3 minutes. A fixed protein concentration of 0.07 μM in 1.2 mL PBS buffer (pH 7.4) was titrated by adding small aliquots of 1-100 μL of a ligand stock solution of 84 μM, covering a range from 0 to 6.5 μM. The protein fluorescence was found to quench in the presence of the ligands. Data analysis was performed using the software Origin 8.1 and the binding curve was obtained by plotting ΔF/F0 values *versus* ligand concentration, where F and F0 are fluorescence intensities in presence and in absence of the ligands respectively. The binding constant for each system, Kb,

was determined by non-linear regression with One Site-Specific Binding model [185-187].

7.3 NMR methods

NMR spectra were recorded on a Bruker AVANCE NEO 600 MHz equipped with a cryo probe. Data acquisition and processing were performed with TOPSPIN 4.1.1 software. All samples were prepared in phosphate saline deuterated buffer (10 mM Na_2HPO_4, 2.7 mM KCl, 137 mM NaCl, 10 mM NaN_3 in D_2O), pH 7.4 at 298 K. [D4](trimethylsilyl)propionic acid, sodium salt (TSP, 10 μM) was used as internal reference.

7.3.1 Free ligand spectroscopic characterization

The NMR assignment of the ^{1}H and ^{13}C resonances, for all ligand analyzed in the present work Thesis, was achieved combining the information from 1D-NMR and 2D-NMR DQF-COSY, TOCSY, NOESY, t-ROESY and 1H-13C HSQC/HMBC experiments. ROESY and NOESY spectra were measured using data sets (t1xt2) of 4096x256 points with mixing times between 100 ms and 400 ms. Double quantum filtered phase sensitive COSY experiments were performed using data sets of 4096x512 points. Total correlation spectroscopy experiments (TOCSY) were performed with a spinlock time of 100 ms, using data sets (t1 xt2) of 4096x256 points.

In all homonuclear experiments the data matrix was zero-filled in the F1 dimension to give a matrix of 4096x2048 points and was resolution enhanced in both dimensions by a cosine-bell function before Fourier transformation. Coupling constants were determined on a first- order basis from 2D phase sensitive double quantum filtered correlation spectroscopy (DQF-COSY) Heteronuclear single

quantum coherence (HSQC) and heteronuclear multiple bond correlation (HMBC) experiments were measured in the ^{1}H-detected mode *via* single quantum coherence with proton decoupling in the ^{13}C domain, using data sets of 2048x256 points. Experiments were carried out in the phase-sensitive mode. A 60 ms delay was used for the evolution of long-range connectivity in the HMBC experiment. In all heteronuclear experiments the data matrix was extended to 2048x1024 points using forward linear prediction extrapolation.

7.3.2 tr-NOESY

tr-NOESY spectra were recorded by using protein/ligand molar ratios of 1:10→1:20. Homonuclear 2D 1H-1H NOESY and ROESY experiments were acquired by using data sets of 2048x512 points with mixing times of 600–700 ms for the free states and of 300-400 ms for the bound states. The concentration of the protein was 10-20 μM.

7.3.2 STD NMR experiments

Protein/ligand molar ratio varied from 1:50 to 1:150 and saturation times of 2s was used. The on-resonance frequency was set at different ppm values according to the analyzed system, since it has to fall in a region in which ligand signals must be absent, necessary condition to avoid false-positive results in the STD spectrum. The off-resonance frequency was always maintained at δ=40 ppm, where no protein or ligand signals resonated. For protein saturation, a train of 40 Gaussian shaped pulses each of 50 ms was employed. The saturated and reference spectra were acquired simultaneously by creating a pseudo-2D experiment. Control experiments with the ligands and proteins in the absence of protein and ligand respectively were carried out to optimized STD NMR parameters. STD spectra were performed with 16 k and 64 k data points. The original free induction decay (FID) was zero-filled to 64 k, and

Fourier transformation with use of an exponential window function was applied (lb = 1–2 Hz).

The 2D STD NMR spectra were recorded by using data sets of 2048/256 points. To determine the magnitude of the STD effects, the intensity of the signals in the STD NMR spectra were compared with the signal intensities of a reference spectrum (off-resonance). The %STD displayed in the epitope maps of the ligands were obtained by the ratio of the STD signals (I_0 – Isat) and each relative peak intensity of the unsaturated reference spectrum (off-resonance, I_0). The highest STD signal was set to 100% and all the other STD were normalized to this value.

7.4 Mass Spectrometry experiments

7.4.1 GC-MS analysis (SECTION II – Chapter 4)

Monosaccharide composition of pure extracted LOS from *N. gonorrhoeae* strain FA1090 wild type was established by evaluation of the acetylated O-methyl glycoside (AMG) derivatives. AMG derivatization was obtained by treatment with HCl/CH_3OH (1.25 M, 85 °C, 18 h) followed by acetylation with acetic anhydride in pyridine (85 °C, 30 min) [183]. The quantitative analysis of sugar residues was achieved by acetylated alditols (AA) derivatization. As for AA, Sample was treated with trifluoroacetic acid (TFA 2M) (120°C, 2h) to hydrolyse the LOS and then subjected to reduction reaction with ethanol and $NaBH_4$ (room temperature overnight.). Acetylation reaction occurred by treatment with pyridine and Ac_2O (85° C for 30 min). The derivatized sugars were obtained after chloroform-water extraction after centrifugation (1500 rpm, 5 min). All of these sugar derivatives were solved in acetone and analysed by GC-MS. GC-MS experiments were performed with Agilent Technologies gas chromatograph 6850A equipped with a mass selective detector 5973N and a Zebron ZB-5 capillary column (Phenomenex, 30 m x 0.25

mm.d., flow rate 1 cm^3/min, He as carrier gas). The temperature program was set as follow: 150 °C for 3 min, 150 °C → 240 °C at 3 °C/min.

7.4.2 Maldi-TOF analysis (SECTION II – Chapter 4)

MALDI-TOF MS spectra were all recorded on an ABSCIEX TOF/ TOF™ 5800 Applied Biosystems mass spectrometer equipped with an Nd:YAG laser (λ = 349 nm), with a 3 ns pulse width and a repetition rate of up to 1000 Hz. 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 μL of sample and 0.5 μL 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. For MS experiments each spectrum was a result of the accumulation of 2000 laser shots, whereas 5000–7000 shots were summed for the MS/MS spectra. Each experiment was performed in duplicate.

7.4.3 ESI-MS characterization (SECTION II – Chapter 4)

Extracted and purified LOS was hydrolysed in acetic acid solution at 1% (v/v) (100°C for 2h). Sample was then centrifuged (14000 rpm, 4°C for 15) and the supernatant was recovered and dried under nitrogen flow. Dried OS was resuspended in ultrapure water and purified through a solid phase extraction (SPE) by using Oasis HLB column (Waters). Before loading the sample, the HLB column was activated by washes with a 0.1% (v/v) acetonitrile and 1% (v/v) formic acid solution. Prior injection the sample was diluted 1:2 with 1% (v/v) formic solution and then analyzed by UPLC-Mass Spectrometry. The UPLC-MS experiments were performed on ACQUITY UPLC I-Class (Waters) coupled with Q-Exactive Plus Biopharma (Thermo Scientific). Ions were generated by electrospray ionization (ESI) source in

negative ion mode. The chromatographic separation was performed using a ACQUITY UPLC Glycan BEH Amide column, 130 Å, 1.7 µm, 2.1mm x 100 mm (Waters) and the mobile phases were 3% (v/v) acetonitrile with 0.1 % (v/v) formic acid and 100% (v/v) acetonitrile with 1% (v/v) formic acid.

7.5 Molecular mechanics and molecular dynamics simulations

The determination of the conformation and dynamics of glycans, in free and bound state, can be performed by using molecular modelling techniques supported by NMR spectroscopy data. The individual monosaccharides show a limited flexibility and within an oligosaccharide, they are assumed as rigid ring structures, so the conformation of a glycan can be established just defining the behaviour of each glycosidic linkage through molecular mechanics and dynamic simulation.

7.5.1 Molecular mechanics simulations

Molecular mechanics studies were performed using Maestro software [188]. The saccharidic chain conformation is mainly defined by the relative orientation of the sugar units, *i.e.* by Φ and Ψ torsion angles. Therefore, the first step in the conformational analysis was the building of the potential energy surfaces for each disaccharide connected by a glycosidic linkage. MM simulations provided a first estimation of the conformational regions energetically accessible and the resulting adiabatic energy maps shown the global and local minima. Φ represents the torsion angle about H1- C1-O-CX' while Ψ about C1-O-CX'-HX'. MM calculations were carried out utilizing MM3* force field included in MacroModel [189] and the solvent bulk was simulated by a dielectric constant of 80. For each disaccharide both Φ and Ψ dihedral angles were varied incrementally using a grid step of 18 degrees. The

corresponding flexible maps were drawn as 2D contours plots using the graphical tools of MacroModel tool.

7.5.2 Docking

Docking calculations were performed with AutoDock 4.2.2 and analyzed with AutoDock tools [133]. For the docking calculations, Gasteiger charges of protein and ligands were assigned by AutoDockTools. With AutoDock 4.2.2, the Lamarckian genetic algorithm was used to sample different conformations of the ligands, by randomly changing all the torsion angles and overall orientation of the molecule. A 3D grids was defined selecting them considering the volume necessary to cover the ligand and the key amino acids of protein binding site. The grid point spacing thus centred to the ligand. A distance-dependent dielectric constant and the original Lennard-Jonnes and hydrogen-bonding potentials provided by AutoDock were used. A total of 200 runs using Lamarckian Genetic algorithm was performed, with a population size of 100, and 250000 energy evaluations. After docking, the 200 solutions were clustered in groups with root-mean-square deviation less than 1.0 Å. The clusters were ranked by the lowest energy representative of each cluster.

7.5.3 Non canonical sugars parametrization

The non-standard residue KDO were parametrized by following a home-made protocol developed in the research group of prof. Marchetti.

7.5.4 Molecular dynamics simulations

The oligosaccharides were built with the glycam website (https://www.glycam.org) [190] carbohydrate builder utility. The torsion angles were chosen based on the values calculated from the molecular mechanics calculations. For protein

preparation, missing hydrogen atoms were added, and the protonation state of ionizable groups and cap termini were added using PyMol software.

Molecular dynamic calculations were performed with AMBER 18 software packages [134], San Francisco, CA. MD simulations were performed in explicit waters using AMBER ff14SB,237 Glycam06j-1,236 and TIP3P force fields for the protein residues, the saccharide ligand, and the water solvent molecules, respectively. Systems were hydrates by using an octahedral box of water (TIP3P) of the proper size and the remote interactions were calculated using a cut off of 10 Angstroms and counterions were added to neutralize the system.

Molecular dynamics simulations are typically carried out in four steps under isotermic isobaric conditions. The input files were generated using the *tleap* modules of the software package. The minimization steps were performed using Sander module and molecular dynamic calculations were performed using the PMEMD module. Once prepared the input files, an energy minimization process was carried out to refine the initial structure. Calculation employed SHAKE for C-H bonds and 1 fs of integration step. Periodic boundary conditions were applied, as well as the smooth particle mesh Ewald method to represent the electrostatic interactions, with a grid space of 1 Å. First minimization of the system was performed holding the complex while a second minimization was performed on the entire system. Furthermore, the whole system was slowly heated from 0 to 300 K using a weak restrain on the solute.

In details, the temperature was increased from 0 to 100 K at constant volume and then, from 100 to 300 K in an isobaric ensemble. Then the system was equilibrated at 300 K.

The system coordinates were saved and used for the 100ns simulations using the PMEMD module implemented in AMBER. The coordinate trajectories were recorded each 2 ps throughout production runs, yielding an ensemble of 10,000

structures for each complex, which were finally analyzed. Trajectories were analyzed using the *ptraj* module within AMBER 18. The VMD program was used to analyze the MD results [191]. Each trajectory was submitted to cluster analysis with respect to the ligand RMSD using the K-mean algorithm implemented in the *ptraj* module. The representative structure of the most populated cluster was considered to describe the complexes interactions. The hydrogen bonds analysis was performed using the CPPTAJ module in AMBER 18 [192]. The H-bond is defined as occurring between an acceptor heavy atom A, a donor hydrogen atom H, and a donor heavy atom D. The distance cut-off was set to 3 Å and the A-H-D angle cut-off was 135°. The 3D pictures were created with PyMol [193], and VMD [191] and the 2D interaction diagrams with Maestro [188].

This work was sponsored by GlaxoSmithKline Biologicals SA.

BIBLIOGRAPHY

1. Ruff, W.E., T.M. Greiling, and M.A. Kriegel, *Host-microbiota interactions in immune-mediated diseases.* Nat Rev Microbiol, 2020. **18**(9): p. 521-538.
2. Mockl, L., *The Emerging Role of the Mammalian Glycocalyx in Functional Membrane Organization and Immune System Regulation.* Front Cell Dev Biol, 2020. **8**: p. 253.
3. Silhavy, T.J., D. Kahne, and S. Walker, *The bacterial cell envelope.* Cold Spring Harb Perspect Biol, 2010. **2**(5): p. a000414.
4. Rajagopal, M. and S. Walker, *Envelope Structures of Gram-Positive Bacteria.* Curr Top Microbiol Immunol, 2017. **404**: p. 1-44.
5. Gonzalez, S., et al., *Conceptual aspects of self and nonself discrimination.* Self Nonself, 2011. **2**(1): p. 19-25.
6. Giamarellos-Bourboulis, E.J. and M. Raftogiannis, *The immune response to severe bacterial infections: consequences for therapy.* Expert Rev Anti Infect Ther, 2012. **10**(3): p. 369-80.
7. Maverakis, E., et al., *Glycans in the immune system and The Altered Glycan Theory of Autoimmunity: a critical review.* J Autoimmun, 2015. **57**: p. 1-13.
8. Netea, M.G., et al., *Innate and Adaptive Immune Memory: an Evolutionary Continuum in the Host's Response to Pathogens.* Cell Host Microbe, 2019. **25**(1): p. 13-26.
9. Rosini, R., et al., *Vaccines Against Antimicrobial Resistance.* Front Immunol, 2020. **11**: p. 1048.
10. Morrison, L. and T.R. Zembower, *Antimicrobial Resistance.* Gastrointest Endosc Clin N Am, 2020. **30**(4): p. 619-635.
11. Cox, G. and G.D. Wright, *Intrinsic antibiotic resistance: mechanisms, origins, challenges and solutions.* Int J Med Microbiol, 2013. **303**(6-7): p. 287-92.
12. Finco, O. and R. Rappuoli, *Designing vaccines for the twenty-first century society.* Front Immunol, 2014. **5**: p. 12.
13. Stern, P.L., *Key steps in vaccine development.* Ann Allergy Asthma Immunol, 2020. **125**(1): p. 17-27.
14. Pollard, A.J. and E.M. Bijker, *A guide to vaccinology: from basic principles to new developments.* Nat Rev Immunol, 2021. **21**(2): p. 83-100.
15. Willis, N.J., *Edward Jenner and the eradication of smallpox.* Scott Med J, 1997. **42**(4): p. 118-21.
16. Canoui, E. and O. Launay, *[History and principles of vaccination].* Rev Mal Respir, 2019. **36**(1): p. 74-81.
17. Lesinski, G.B. and M.A. Westerink, *Novel vaccine strategies to T-independent antigens.* J Microbiol Methods, 2001. **47**(2): p. 135-49.
18. Pollard, A.J., K.P. Perrett, and P.C. Beverley, *Maintaining protection against invasive bacteria with protein-polysaccharide conjugate vaccines.* Nat Rev Immunol, 2009. **9**(3): p. 213-20.

19. Madore, D.V., et al., *Safety and immunogenicity of Haemophilus influenzae type b oligosaccharide-CRM197 conjugate vaccine in infants aged 15 to 23 months.* Pediatrics, 1990. **86**(4): p. 527-34.
20. Costantino, P., R. Rappuoli, and F. Berti, *The design of semi-synthetic and synthetic glycoconjugate vaccines.* Expert Opin Drug Discov, 2011. **6**(10): p. 1045-66.
21. Tone Tønjum, J.v.P., *Neisseria*, ed. F. Edition. Vol. Infectious Diseases. 2017: Elsevier.
22. Vanbaelen, T., et al., *Global epidemiology of antimicrobial resistance in commensal Neisseria species: A systematic review.* Int J Med Microbiol, 2022. **312**(3): p. 151551.
23. Ligon, B.L., *Albert Ludwig Sigesmund Neisser: discoverer of the cause of gonorrhea.* Semin Pediatr Infect Dis, 2005. **16**(4): p. 336-41.
24. Rice, P.A., et al., *Neisseria gonorrhoeae: Drug Resistance, Mouse Models, and Vaccine Development.* Annu Rev Microbiol, 2017. **71**: p. 665-686.
25. Rowley, J., et al., *Chlamydia, gonorrhoea, trichomoniasis and syphilis: global prevalence and incidence estimates, 2016.* Bull World Health Organ, 2019. **97**(8): p. 548-562P.
26. Humbert, M.V. and M. Christodoulides, *Atypical, Yet Not Infrequent, Infections with Neisseria Species.* Pathogens, 2019. **9**(1).
27. Walker, C.K. and R.L. Sweet, *Gonorrhea infection in women: prevalence, effects, screening, and management.* Int J Womens Health, 2011. **3**: p. 197-206.
28. Edwards, J.L. and M.A. Apicella, *The molecular mechanisms used by Neisseria gonorrhoeae to initiate infection differ between men and women.* Clin Microbiol Rev, 2004. **17**(4): p. 965-81, table of contents.
29. Quillin, S.J. and H.S. Seifert, *Neisseria gonorrhoeae host adaptation and pathogenesis.* Nat Rev Microbiol, 2018. **16**(4): p. 226-240.
30. Unemo, M., et al., *Gonorrhoea.* Nat Rev Dis Primers, 2019. **5**(1): p. 79.
31. Cohen, M.S., et al., *Reduction of concentration of HIV-1 in semen after treatment of urethritis: implications for prevention of sexual transmission of HIV-1. AIDSCAP Malawi Research Group.* Lancet, 1997. **349**(9069): p. 1868-73.
32. Stohl, E.A., et al., *Neisseria gonorrhoeae metalloprotease NGO1686 is required for full piliation, and piliation is required for resistance to H2O2- and neutrophil-mediated killing.* mBio, 2013. **4**(4).
33. van der Woude, M.W. and A.J. Baumler, *Phase and antigenic variation in bacteria.* Clin Microbiol Rev, 2004. **17**(3): p. 581-611, table of contents.
34. James-Holmquest, A.N., et al., *Differential attachment by piliated and nonpiliated Neisseria gonorrhoeae to human sperm.* Infect Immun, 1974. **9**(5): p. 897-902.
35. Jerse, A.E., et al., *Multiple gonococcal opacity proteins are expressed during experimental urethral infection in the male.* J Exp Med, 1994. **179**(3): p. 911-20.
36. in *WHO Guidelines on Hand Hygiene in Health Care: First Global Patient Safety Challenge Clean Care Is Safer Care*. 2009: Geneva.

37. Spurbeck, R.R. and C.G. Arvidson, *Lactobacillus jensenii surface-associated proteins inhibit Neisseria gonorrhoeae adherence to epithelial cells.* Infect Immun, 2010. **78**(7): p. 3103-11.
38. St Amant, D.C., I.E. Valentin-Bon, and A.E. Jerse, *Inhibition of Neisseria gonorrhoeae by Lactobacillus species that are commonly isolated from the female genital tract.* Infect Immun, 2002. **70**(12): p. 7169-71.
39. Spurbeck, R.R. and C.G. Arvidson, *Inhibition of Neisseria gonorrhoeae epithelial cell interactions by vaginal Lactobacillus species.* Infect Immun, 2008. **76**(7): p. 3124-30.
40. Cassat, J.E. and E.P. Skaar, *Iron in infection and immunity.* Cell Host Microbe, 2013. **13**(5): p. 509-519.
41. Mendes, A.C., et al., *Epithelial Haven and Autophagy Breakout in Gonococci Infection.* Front Cell Dev Biol, 2020. **8**: p. 439.
42. Ortiz, M.C., et al., *Neisseria gonorrhoeae Modulates Immunity by Polarizing Human Macrophages to a M2 Profile.* PLoS One, 2015. **10**(6): p. e0130713.
43. Janeway, C., K.M. Murphy, and C. Weaver, *Janeway's Immunobiology*. 2017: Garland Science.
44. Thau, L., E. Asuka, and K. Mahajan, *Physiology, Opsonization*, in *StatPearls*. 2023: Treasure Island (FL).
45. Edwards, J.L. and M.A. Apicella, *The role of lipooligosaccharide in Neisseria gonorrhoeae pathogenesis of cervical epithelia: lipid A serves as a C3 acceptor molecule.* Cell Microbiol, 2002. **4**(9): p. 585-98.
46. Cahoon, L.A. and H.S. Seifert, *Focusing homologous recombination: pilin antigenic variation in the pathogenic Neisseria.* Mol Microbiol, 2011. **81**(5): p. 1136-43.
47. Liu, Y., et al., *Neisseria gonorrhoeae selectively suppresses the development of Th1 and Th2 cells, and enhances Th17 cell responses, through TGF-beta-dependent mechanisms.* Mucosal Immunol, 2012. **5**(3): p. 320-31.
48. Zhu, W., et al., *Neisseria gonorrhoeae suppresses dendritic cell-induced, antigen-dependent CD4 T cell proliferation.* PLoS One, 2012. **7**(7): p. e41260.
49. Preston, A., et al., *The lipooligosaccharides of pathogenic gram-negative bacteria.* Crit Rev Microbiol, 1996. **22**(3): p. 139-80.
50. Mazgaeen, L. and P. Gurung, *Recent Advances in Lipopolysaccharide Recognition Systems.* Int J Mol Sci, 2020. **21**(2).
51. Christodoulides, M., *Preparation of Lipooligosaccharide (LOS) from Neisseria gonorrhoeae.* Methods Mol Biol, 2019. **1997**: p. 87-96.
52. Gulati, S., et al., *Targeting Lipooligosaccharide (LOS) for a Gonococcal Vaccine.* Front Immunol, 2019. **10**: p. 321.
53. Gibson, B.W., et al., *Investigation of the structural heterogeneity of lipooligosaccharides from pathogenic Haemophilus and Neisseria species and of R-type lipopolysaccharides from Salmonella typhimurium by electrospray mass spectrometry.* J Bacteriol, 1993. **175**(9): p. 2702-12.

54. Jennings, M.P., et al., *Molecular analysis of a locus for the biosynthesis and phase-variable expression of the lacto-N-neotetraose terminal lipopolysaccharide structure in Neisseria meningitidis.* Mol Microbiol, 1995. **18**(4): p. 729-40.
55. Kahler, C.M., et al., *Two glycosyltransferase genes, lgtF and rfaK, constitute the lipooligosaccharide ice (inner core extension) biosynthesis operon of Neisseria meningitidis.* J Bacteriol, 1996. **178**(23): p. 6677-84.
56. Gotschlich, E.C., *Genetic locus for the biosynthesis of the variable portion of Neisseria gonorrhoeae lipooligosaccharide.* J Exp Med, 1994. **180**(6): p. 2181-90.
57. John, C.M., et al., *The structural basis for pyocin resistance in Neisseria gonorrhoeae lipooligosaccharides.* J Biol Chem, 1991. **266**(29): p. 19303-11.
58. Brooks, G.F. and C.J. Lammel, *Humoral immune response to gonococcal infections.* Clin Microbiol Rev, 1989. **2 Suppl**(Suppl): p. S5-10.
59. Yamasaki, R., et al., *Structural and immunochemical characterization of a Neisseria gonorrhoeae epitope defined by a monoclonal antibody 2C7; the antibody recognizes a conserved epitope on specific lipo-oligosaccharides in spite of the presence of human carbohydrate epitopes.* J Biol Chem, 1999. **274**(51): p. 36550-8.
60. Banerjee, A., et al., *Identification of the gene (lgtG) encoding the lipooligosaccharide beta chain synthesizing glucosyl transferase from Neisseria gonorrhoeae.* Proc Natl Acad Sci U S A, 1998. **95**(18): p. 10872-7.
61. Schneider, H., et al., *Expression of paragloboside-like lipooligosaccharides may be a necessary component of gonococcal pathogenesis in men.* J Exp Med, 1991. **174**(6): p. 1601-5.
62. Chakraborti, S., et al., *Phase-Variable Heptose I Glycan Extensions Modulate Efficacy of 2C7 Vaccine Antibody Directed against Neisseria gonorrhoeae Lipooligosaccharide.* J Immunol, 2016. **196**(11): p. 4576-86.
63. Yamasaki, R., et al., *The structure of lipooligosaccharide produced by Neisseria gonorrhoeae, strain 15253, isolated from a patient with disseminated infection. Evidence for a new glycosylation pathway of the gonococcal lipooligosaccharide.* Journal of Biological Chemistry, 1994. **269**(48): p. 30345-30351.
64. Ngampasutadol, J., et al., *Characterization of a peptide vaccine candidate mimicking an oligosaccharide epitope of Neisseria gonorrhoeae and resultant immune responses and function.* Vaccine, 2006. **24**(2): p. 157-70.
65. Ram, S., et al., *A Novel Sialylation Site on Neisseria gonorrhoeae Lipooligosaccharide Links Heptose II Lactose Expression with Pathogenicity.* Infect Immun, 2018. **86**(8).
66. Gulati, S., et al., *Immunization against a saccharide epitope accelerates clearance of experimental gonococcal infection.* PLoS Pathog, 2013. **9**(8): p. e1003559.
67. Gulati, S., et al., *Immunogenicity of Neisseria gonorrhoeae lipooligosaccharide epitope 2C7, widely expressed in vivo with no immunochemical similarity to human glycosphingolipids.* J Infect Dis, 1996. **174**(6): p. 1223-37.

68. van Vliet, S.J., et al., *Variation of Neisseria gonorrhoeae lipooligosaccharide directs dendritic cell-induced T helper responses.* PLoS Pathog, 2009. **5**(10): p. e1000625.
69. Balthazar, J.T., et al., *Lipooligosaccharide Structure is an Important Determinant in the Resistance of Neisseria Gonorrhoeae to Antimicrobial Agents of Innate Host Defense.* Front Microbiol, 2011. **2**: p. 30.
70. Varki, A., et al., *Symbol Nomenclature for Graphical Representations of Glycans.* Glycobiology, 2015. **25**(12): p. 1323-4.
71. Maurakis, S.A. and C.N. Cornelissen, *Recent Progress Towards a Gonococcal Vaccine.* Front Cell Infect Microbiol, 2022. **12**: p. 881392.
72. Gulati, S., et al., *Strategies for mimicking Neisserial saccharide epitopes as vaccines.* Int Rev Immunol, 2001. **20**(2): p. 229-50.
73. Kurzyp, K. and O.B. Harrison, *Bacterium of one thousand and one variants: genetic diversity of Neisseria gonorrhoeae pathogenicity.* Microb Genom, 2023. **9**(6).
74. Harvey, H.A., W.E. Swords, and M.A. Apicella, *The mimicry of human glycolipids and glycosphingolipids by the lipooligosaccharides of pathogenic neisseria and haemophilus.* J Autoimmun, 2001. **16**(3): p. 257-62.
75. Mandrell, R.E., et al., *Lipooligosaccharides (LOS) of some Haemophilus species mimic human glycosphingolipids, and some LOS are sialylated.* Infect Immun, 1992. **60**(4): p. 1322-8.
76. Jacques, M., *Role of lipo-oligosaccharides and lipopolysaccharides in bacterial adherence.* Trends Microbiol, 1996. **4**(10): p. 408-9.
77. Maldonado, R.F., I. Sa-Correia, and M.A. Valvano, *Lipopolysaccharide modification in Gram-negative bacteria during chronic infection.* FEMS Microbiol Rev, 2016. **40**(4): p. 480-93.
78. Jörgen Lönngren, S.S., *Mass Spectrometry in Structural Analysis of Natural Carbohydrates*. Advances in Carbohydrate Chemistry and Biochemistry, ed. D.H. R. Stuart Tipson. Vol. 29. 1974, Academic Press
79. De Castro, C., *Gas Chromatography - Mass Spectrometry for Glycosciences.* Ed. Glycopedia
80. Bubb, W.A., *NMR spectroscopy in the study of carbohydrates: Characterizing the structural complexity.* Inc. Concepts Magn Reson, 2003 **19A**(1): p. 1-19.
81. Nitsche, C. and G. Otting, *NMR studies of ligand binding.* Curr Opin Struct Biol, 2018. **48**: p. 16-22.
82. Di Carluccio, C., et al., *Investigation of protein-ligand complexes by ligand-based NMR methods.* Carbohydr Res, 2021. **503**: p. 108313.
83. Marchetti, R., et al., *"Rules of Engagement" of Protein-Glycoconjugate Interactions: A Molecular View Achievable by using NMR Spectroscopy and Molecular Modeling.* ChemistryOpen, 2016. **5**(4): p. 274-96.
84. Meyer, B. and T. Peters, *NMR spectroscopy techniques for screening and identifying ligand binding to protein receptors.* Angew Chem Int Ed Engl, 2003. **42**(8): p. 864-90.

85. Zhang, X., et al., *Structure-based drug design: NMR-based approach for ligand-protein interactions.* Drug Discov Today Technol, 2006. **3**(3): p. 241-5.
86. Du, X., et al., *Insights into Protein-Ligand Interactions: Mechanisms, Models, and Methods.* Int J Mol Sci, 2016. **17**(2).
87. Koos, M.R.M., K.H.G. Schulz, and R.R. Gil, *Reference-free NOE NMR analysis.* Chem Sci, 2020. **11**(36): p. 9930-9936.
88. Haselhorst, T., A.C. Lamerz, and M. Itzstein, *Saturation transfer difference NMR spectroscopy as a technique to investigate protein-carbohydrate interactions in solution.* Methods Mol Biol, 2009. **534**: p. 375-86.
89. Haselhorst, T., et al., *Saturation transfer difference (STD) 1H-NMR experiments and in silico docking experiments to probe the binding of N-acetylneuraminic acid and derivatives to Vibrio cholerae sialidase.* Proteins, 2004. **56**(2): p. 346-53.
90. Wagstaff, J.L., S.L. Taylor, and M.J. Howard, *Recent developments and applications of saturation transfer difference nuclear magnetic resonance (STD NMR) spectroscopy.* Mol Biosyst, 2013. **9**(4): p. 571-7.
91. Bhunia, A., S. Bhattacharjya, and S. Chatterjee, *Applications of saturation transfer difference NMR in biological systems.* Drug Discov Today, 2012. **17**(9-10): p. 505-13.
92. Mayer, M. and B. Meyer, *Characterization of Ligand Binding by Saturation Transfer Difference NMR Spectroscopy.* Angew Chem Int Ed Engl, 1999. **38**(12): p. 1784-1788.
93. Mayer, M. and B. Meyer, *Group epitope mapping by saturation transfer difference NMR to identify segments of a ligand in direct contact with a protein receptor.* J Am Chem Soc, 2001. **123**(25): p. 6108-17.
94. Brender, J.R., et al., *Binding Moiety Mapping by Saturation Transfer Difference NMR.* Methods Mol Biol, 2018. **1824**: p. 49-65.
95. Lin, X., X. Li, and X. Lin, *A Review on Applications of Computational Methods in Drug Screening and Design.* Molecules, 2020. **25**(6).
96. Abel, R., et al., *Advancing Drug Discovery through Enhanced Free Energy Calculations.* Acc Chem Res, 2017. **50**(7): p. 1625-1632.
97. De Vivo, M., et al., *Role of Molecular Dynamics and Related Methods in Drug Discovery.* J Med Chem, 2016. **59**(9): p. 4035-61.
98. Jorgensen, W.L., *The many roles of computation in drug discovery.* Science, 2004. **303**(5665): p. 1813-8.
99. Guvench, O., et al., *CHARMM Additive All-Atom Force Field for Glycosidic Linkages between Hexopyranoses.* J Chem Theory Comput, 2009. **5**(9): p. 2353-2370.
100. Kirschner, K.N., et al., *GLYCAM06: a generalizable biomolecular force field. Carbohydrates.* J Comput Chem, 2008. **29**(4): p. 622-55.
101. Lins, R.D. and P.H. Hunenberger, *A new GROMOS force field for hexopyranose-based carbohydrates.* J Comput Chem, 2005. **26**(13): p. 1400-12.
102. Ferreira, L.G., et al., *Molecular docking and structure-based drug design strategies.* Molecules, 2015. **20**(7): p. 13384-421.

103. Meng, X.Y., et al., *Molecular docking: a powerful approach for structure-based drug discovery.* Curr Comput Aided Drug Des, 2011. **7**(2): p. 146-57.
104. Morris, G.M., et al., *Distributed automated docking of flexible ligands to proteins: parallel applications of AutoDock 2.4.* J Comput Aided Mol Des, 1996. **10**(4): p. 293-304.
105. Trott, O. and A.J. Olson, *AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading.* J Comput Chem, 2010. **31**(2): p. 455-61.
106. Li, J., A. Fu, and L. Zhang, *An Overview of Scoring Functions Used for Protein-Ligand Interactions in Molecular Docking.* Interdiscip Sci, 2019. **11**(2): p. 320-328.
107. Pagadala, N.S., K. Syed, and J. Tuszynski, *Software for molecular docking: a review.* Biophys Rev, 2017. **9**(2): p. 91-102.
108. Revard, B.C., W.W. Tipton, and R.G. Hennig, *Structure and stability prediction of compounds with evolutionary algorithms.* Top Curr Chem, 2014. **345**: p. 181-222.
109. Fuhrmann, J., et al., *A new Lamarckian genetic algorithm for flexible ligand-receptor docking.* J Comput Chem, 2010. **31**(9): p. 1911-8.
110. Wang, R., Y. Lu, and S. Wang, *Comparative evaluation of 11 scoring functions for molecular docking.* J Med Chem, 2003. **46**(12): p. 2287-303.
111. Guedes, I.A., F.S.S. Pereira, and L.E. Dardenne, *Empirical Scoring Functions for Structure-Based Virtual Screening: Applications, Critical Aspects, and Challenges.* Front Pharmacol, 2018. **9**: p. 1089.
112. Gohlke, H., M. Hendlich, and G. Klebe, *Knowledge-based scoring function to predict protein-ligand interactions.* J Mol Biol, 2000. **295**(2): p. 337-56.
113. Bohm, H.J., *The development of a simple empirical scoring function to estimate the binding constant for a protein-ligand complex of known three-dimensional structure.* J Comput Aided Mol Des, 1994. **8**(3): p. 243-56.
114. Vanommeslaeghe, K., O. Guvench, and A.D. MacKerell, Jr., *Molecular mechanics.* Curr Pharm Des, 2014. **20**(20): p. 3281-92.
115. Hansson, T., C. Oostenbrink, and W. van Gunsteren, *Molecular dynamics simulations.* Curr Opin Struct Biol, 2002. **12**(2): p. 190-6.
116. Hollingsworth, S.A. and R.O. Dror, *Molecular Dynamics Simulation for All.* Neuron, 2018. **99**(6): p. 1129-1143.
117. D. A. Case, H.M.A., K. Belfon, I.Y. Ben-Shalom, J.T. Berryman, S.R. Brozell, D.S. Cerutti, T.E. Cheatham, III, G.A. Cisneros, V.W.D. Cruzeiro, T.A. Darden, N. Forouzesh, G. Giambaşu, T. Giese, M.K. Gilson, H. Gohlke, A.W. Goetz, J. Harris, S. Izadi, S.A. Izmailov, K. Kasavajhala, M.C. Kaymak, E. King, A. Kovalenko, T. Kurtzman, T.S. Lee, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, M. Machado, V. Man, M. Manathunga, K.M. Merz, Y. Miao, O. Mikhailovskii, G. Monard, H. Nguyen, K.A. O'Hearn, A. Onufriev, F. Pan, S. Pantano, R. Qi, A. Rahnamoun, D.R. Roe, A. Roitberg, C. Sagui, S. Schott-Verdugo, A. Shajan, J. Shen, C.L. Simmerling, N.R. Skrynnikov, J. Smith, J. Swails, R.C.

Walker, J. Wang, J. Wang, H. Wei, X. Wu, Y. Wu, Y. Xiong, Y. Xue, D.M. York, S. Zhao, Q. Zhu, and P.A. Kollman, *Amber 2023*. 2021, University of California, San Francisco.
118. Dickson, C.J., et al., *Lipid14: The Amber Lipid Force Field.* J Chem Theory Comput, 2014. **10**(2): p. 865-879.
119. Maier, J.A., et al., *ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB.* J Chem Theory Comput, 2015. **11**(8): p. 3696-713.
120. Wang, J., et al., *Development and testing of a general amber force field.* J Comput Chem, 2004. **25**(9): p. 1157-74.
121. Arias, T.A., M.C. Payne, and J.D. Joannopoulos, *Ab initio molecular dynamics: Analytically continued energy functionals and insights into iterative solutions.* Phys Rev Lett, 1992. **69**(7): p. 1077-1080.
122. Fisette, O., et al., *Synergistic applications of MD and NMR for the study of biological systems.* J Biomed Biotechnol, 2012. **2012**: p. 254208.
123. Sapay, N., A. Nurisso, and A. Imberty, *Simulation of carbohydrates, from molecular docking to dynamics in water.* Methods Mol Biol, 2013. **924**: p. 469-83.
124. Jordan, P.W., L.A. Snyder, and N.J. Saunders, *Strain-specific differences in Neisseria gonorrhoeae associated with the phase variable gene repertoire.* BMC Microbiol, 2005. **5**: p. 21.
125. Hobbs, M.M., et al., *Experimental Gonococcal Infection in Male Volunteers: Cumulative Experience with Neisseria gonorrhoeae Strains FA1090 and MS11mkC.* Front Microbiol, 2011. **2**: p. 123.
126. Westphal, O.a.J., K., *Bacterial Lipopolysaccharides. Extraction with Phenol-Water and Further Applications of the Procedure*. Vol. 5. 1965: Methods in Carbohydrate Chemistry.
127. Oxley, D., G. Currie, and A. Bacic, *Monosaccharide composition analysis: alditol acetates.* CSH Protoc, 2006. **2006**(1).
128. De Castro, C., *Acetylated Alditols : Advantages & Limitations*. Glycopedia
129. Agrawal, P.K., *NMR spectroscopy in the structural elucidation of oligosaccharides and glycosides.* Phytochemistry, 1992. **31**(10): p. 3307-30.
130. Speciale, I., et al., *Liquid-state NMR spectroscopy for complex carbohydrate structural analysis: A hitchhiker's guide.* Carbohydr Polym, 2022. **277**: p. 118885.
131. Chakraborti, S., et al., *Bypassing Phase Variation of Lipooligosaccharide (LOS): Using Heptose 1 Glycan Mutants To Establish Widespread Efficacy of Gonococcal Anti-LOS Monoclonal Antibody 2C7.* Infect Immun, 2020. **88**(2).
132. Tyrikos-Ergas, T., et al., *Structural Studies Using Unnatural Oligosaccharides: Toward Sugar Foldamers.* Biomacromolecules, 2020. **21**(1): p. 18-29.
133. Morris, G.M., et al., *AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility.* J Comput Chem, 2009. **30**(16): p. 2785-91.
134. Case, D.A.B.-S., I. Y.; Brozell, S. R.; Cerutti, D. S.; Cheatham, T. E.; Cruzeiro, V. W. D.; Darden , T. A.; Duke, R. E.; Ghoreishi, D.; Gilson, M. K.; Gohlke, H.; Goetz, A. W.;

Greene, D.; Harris, R.; Homeyer, N.; Huang, Y.; Izadi, S.; Kovalenko, A.; Kurtzman, T.; Lee, T. S.; LeGrand, S.; Li, P.; Lin, C.; Liu, J.; Luchko, T.; Luo, R.; Mermelstein, D. J.; Merz, K. M.; Miao, Y.; Monard, G.; Nguyen, C.; Nguyen, H.; Omelyan, I.; Onufriev, A.; Pan, F.; Qi, R.; Roe, D. R.; Roitberg, A.; Sagui, C.; Schott-Verdugo, S.; Shen, J.; Simmerling, C. L.; Smith, J.; SalomonFerrer, R.; Swails, J.; Walker, R. C.; Wang, J.; Wei, H.; Wolf, R. M.; Wu, X.; Xiao, L.; York, M.; Kollman, P. A., *AMBER 2018*. 2018, University of California, San Francisco.

135. Razin, S., D. Yogev, and Y. Naot, *Molecular biology and pathogenicity of mycoplasmas.* Microbiol Mol Biol Rev, 1998. **62**(4): p. 1094-156.
136. Kumar, S., *Mycoplasma pneumoniae: A significant but underrated pathogen in paediatric community-acquired lower respiratory tract infections.* Indian J Med Res, 2018. **147**(1): p. 23-31.
137. Waites, K.B., et al., *In vitro susceptibilities to and bactericidal activities of garenoxacin (BMS-284756) and other antimicrobial agents against human mycoplasmas and ureaplasmas.* Antimicrob Agents Chemother, 2003. **47**(1): p. 161-5.
138. Domingues, D., et al., *[Mycoplasmas: what is the role in human infections?].* Acta Med Port, 2005. **18**(5): p. 377-83.
139. Matas Andreu, L., et al., *[Serologic diagnosis of Mycoplasma pneumoniae infections].* Enferm Infecc Microbiol Clin, 2006. **24 Suppl 1**: p. 19-23.
140. Waites, K.B. and D.F. Talkington, *Mycoplasma pneumoniae and its role as a human pathogen.* Clin Microbiol Rev, 2004. **17**(4): p. 697-728, table of contents.
141. Qin, L., Y. Chen, and X. You, *Subversion of the Immune Response by Human Pathogenic Mycoplasmas.* Front Microbiol, 2019. **10**: p. 1934.
142. Lee, H., et al., *Antimicrobial therapy of macrolide-resistant Mycoplasma pneumoniae pneumonia in children.* Expert Rev Anti Infect Ther, 2018. **16**(1): p. 23-34.
143. Gautier-Bouchardon, A.V., *Antimicrobial Resistance in Mycoplasma spp.* Microbiol Spectr, 2018. **6**(4).
144. Chernova, O.A., et al., *Mycoplasmas and Their Antibiotic Resistance: The Problems and Prospects in Controlling Infections.* Acta Naturae, 2016. **8**(2): p. 24-34.
145. Jiang, Z., et al., *Mycoplasma pneumoniae Infections: Pathogenesis and Vaccine Development.* Pathogens, 2021. **10**(2).
146. Korppi, M., T. Heiskanen-Kosma, and M. Kleemola, *Incidence of community-acquired pneumonia in children caused by Mycoplasma pneumoniae: serological results of a prospective, population-based study in primary health care.* Respirology, 2004. **9**(1): p. 109-14.
147. Bradley, J.S., et al., *The management of community-acquired pneumonia in infants and children older than 3 months of age: clinical practice guidelines by the Pediatric Infectious Diseases Society and the Infectious Diseases Society of America.* Clin Infect Dis, 2011. **53**(7): p. e25-76.

148. Waites, K.B., *New concepts of Mycoplasma pneumoniae infections in children.* Pediatr Pulmonol, 2003. **36**(4): p. 267-78.
149. Kashyap, S. and M. Sarkar, *Mycoplasma pneumonia: Clinical features and management.* Lung India, 2010. **27**(2): p. 75-85.
150. Aguila, L.K.T., et al., *Ascending Reproductive Tract Infection in Pig-Tailed Macaques Inoculated with Mycoplasma genitalium.* Infect Immun, 2022. **90**(6): p. e0013122.
151. Yueyue, W., et al., *Pathogenicity and virulence of Mycoplasma genitalium: Unraveling Ariadne's Thread.* Virulence, 2022. **13**(1): p. 1161-1183.
152. Gnanadurai, R. and H. Fifer, *Mycoplasma genitalium: A Review.* Microbiology (Reading), 2020. **166**(1): p. 21-29.
153. Ona, S., R.L. Molina, and K. Diouf, *Mycoplasma genitalium: An Overlooked Sexually Transmitted Pathogen in Women?* Infect Dis Obstet Gynecol, 2016. **2016**: p. 4513089.
154. Haggerty, C.L. and B.D. Taylor, *Mycoplasma genitalium: an emerging cause of pelvic inflammatory disease.* Infect Dis Obstet Gynecol, 2011. **2011**: p. 959816.
155. McGowin, C.L. and C. Anderson-Smits, *Mycoplasma genitalium: an emerging cause of sexually transmitted disease in women.* PLoS Pathog, 2011. **7**(5): p. e1001324.
156. Yiwen, C., et al., *Infection strategies of mycoplasmas: Unraveling the panoply of virulence factors.* Virulence, 2021. **12**(1): p. 788-817.
157. Krause, D.C., *Mycoplasma pneumoniae cytadherence: unravelling the tie that binds.* Mol Microbiol, 1996. **20**(2): p. 247-53.
158. Krause, D.C. and J.B. Baseman, *Mycoplasma pneumoniae proteins that selectively bind to host cells.* Infect Immun, 1982. **37**(1): p. 382-6.
159. Aparicio, D., et al., *Mycoplasma genitalium adhesin P110 binds sialic-acid human receptors.* Nat Commun, 2018. **9**(1): p. 4471.
160. Vizarraga, D., et al., *Immunodominant proteins P1 and P40/P90 from human pathogen Mycoplasma pneumoniae.* Nat Commun, 2020. **11**(1): p. 5188.
161. Miyata, M., *Centipede and inchworm models to explain Mycoplasma gliding.* Trends Microbiol, 2008. **16**(1): p. 6-12.
162. Aparicio, D., et al., *Structure and mechanism of the Nap adhesion complex from the human pathogen Mycoplasma genitalium.* Nat Commun, 2020. **11**(1): p. 2877.
163. Vizarraga, D., et al., *The Sialoglycan Binding Adhesins of Mycoplasma genitalium and Mycoplasma pneumoniae.* Trends Microbiol, 2021. **29**(6): p. 477-481.
164. Sakarya, S. and S. Oncu, *Bacterial adhesins and the role of sialic acid in bacterial adhesion.* Med Sci Monit, 2003. **9**(3): p. RA76-82.
165. Jennings, M.P., C.J. Day, and J.M. Atack, *How bacteria utilize sialic acid during interactions with the host: snip, snatch, dispatch, match and attach.* Microbiology (Reading), 2022. **168**(3).
166. Dwek, R.A., *Glycobiology: Toward Understanding the Function of Sugars.* Chem Rev, 1996. **96**(2): p. 683-720.

167. Chang, Y.C. and V. Nizet, *The interplay between Siglecs and sialylated pathogens.* Glycobiology, 2014. **24**(9): p. 818-25.
168. Varki, A., R.L. Schnaar, and R. Schauer, *Sialic Acids and Other Nonulosonic Acids*, in *Essentials of Glycobiology*, A. Varki, et al., Editors. 2015: Cold Spring Harbor (NY). p. 179-95.
169. Kawanishi, K., et al., *Human species-specific loss of CMP-N-acetylneuraminic acid hydroxylase enhances atherosclerosis via intrinsic and extrinsic mechanisms.* Proc Natl Acad Sci U S A, 2019. **116**(32): p. 16036-16045.
170. Bergfeld, A.K., et al., *Metabolism of vertebrate amino sugars with N-glycolyl groups: elucidating the intracellular fate of the non-human sialic acid N-glycolylneuraminic acid.* J Biol Chem, 2012. **287**(34): p. 28865-81.
171. Mahajan, V.S. and S. Pillai, *Sialic acids and autoimmune disease.* Immunol Rev, 2016. **269**(1): p. 145-61.
172. Lewis, A.L., et al., *Cataloging natural sialic acids and other nonulosonic acids (NulOs), and their representation using the Symbol Nomenclature for Glycans.* Glycobiology, 2023. **33**(2): p. 99-103.
173. Kubota, M. and T. Hashiguchi, *Unique Tropism and Entry Mechanism of Mumps Virus.* Viruses, 2021. **13**(9).
174. Wasik, B.R., K.N. Barnard, and C.R. Parrish, *Effects of Sialic Acid Modifications on Virus Binding and Infection.* Trends Microbiol, 2016. **24**(12): p. 991-1001.
175. Forgione, R.E., et al., *Unveiling Molecular Recognition of Sialoglycans by Human Siglec-10.* iScience, 2020. **23**(6): p. 101231.
176. Manchee, R.J. and D. Taylor-Robinson, *Utilization of neuraminic acid receptors by mycoplasmas.* J Bacteriol, 1969. **98**(3): p. 914-9.
177. Sobeslavsky, O., B. Prescott, and R.M. Chanock, *Adsorption of Mycoplasma pneumoniae to neuraminic acid receptors of various cells and possible role in virulence.* J Bacteriol, 1968. **96**(3): p. 695-705.
178. Kasai, T., et al., *Role of binding in Mycoplasma mobile and Mycoplasma pneumoniae gliding analyzed through inhibition by synthesized sialylated compounds.* J Bacteriol, 2013. **195**(3): p. 429-35.
179. Rezania, S., et al., *Extraction, Purification and Characterization of Lipopolysaccharide from Escherichia coli and Salmonella typhi.* Avicenna J Med Biotechnol, 2011. **3**(1): p. 3-9.
180. Hofstad, T. and T. Kristoffersen, *Lipopolysaccharide from Bacteroides melaninogenicus isolated from the supernatant fluid after ultracentrifugation of the water phase following phenol-water extraction.* Acta Pathol Microbiol Scand B Microbiol Immunol, 1971. **79**(1): p. 12-8.
181. Leone, S., et al., *Molecular structure of endotoxins from Gram-negative marine bacteria: an update.* Mar Drugs, 2007. **5**(3): p. 85-112.

182. Kittelberger, R. and F. Hilbink, *Sensitive silver-staining detection of bacterial lipopolysaccharides in polyacrylamide gels.* J Biochem Biophys Methods, 1993. **26**(1): p. 81-6.
183. De Castro, C., et al., *Microbe-associated molecular patterns in innate immunity: Extraction and chemical analysis of gram-negative bacterial lipopolysaccharides.* Methods Enzymol, 2010. **480**: p. 89-115.
184. Vanacore, A., et al., *Role of EPS in mitigation of plant abiotic stress: The case of Methylobacterium extorquens PA1.* Carbohydr Polym, 2022. **295**: p. 119863.
185. Ribeiro, M.M., et al., *Molecular interaction studies of peptides using steady-state fluorescence intensity. Static (de)quenching revisited.* J Pept Sci, 2008. **14**(4): p. 401-6.
186. Forgione, R.E., et al., *Unveiling Molecular Recognition of Sialoglycans by Human Siglec-10.* iScience, 2020. **23**(8): p. 101401.
187. Rosario Oliva, F.B., Serena Cozzolino, Eugenio Notomista, Roland Winter, Pompea Del Vecchio, Luigi Petraccone, *Encapsulating properties of sulfobutylether-β-cyclodextrin toward a thrombin-derived antimicrobial peptide.* J Therm Anal Calorim, 2019. **138**: p. 3249–3256.
188. *Schrödinger Release 2022-3, Maestro, Schrödinger, LLC: New York*. 2021.
189. *Schrödinger Release 2022-3, MacroModel, Schrödinger, LLC: New York*. 2021.
190. WoodsGroup. *GLYCAM Web, Complex Carbohydrate Research Center, University of Georgia, Athens, GA, 2005-2022*

191. Humphrey, W., A. Dalke, and K. Schulten, *VMD: visual molecular dynamics.* J Mol Graph, 1996. **14**(1): p. 33-8, 27-8.
192. Roe, D.R. and T.E. Cheatham, 3rd, *PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data.* J Chem Theory Comput, 2013. **9**(7): p. 3084-95.
193. *The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.*