Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN

BIOLOGIA CELLULARE E MOLECOLARE

Ciclo XXX

Settore Concorsuale: 05/E2

Settore Scientifico Disciplinare: SSD BIO/11

Understanding The MHC-II Presentation Mechanism For The Rational Design Of Glycoconjugate Vaccines

Presentata da: Luigi Capriotti

Coordinatore Dottorato	**Supervisore**
Prof. Giovanni Capranico	*Prof. Vincenzo Scarlato*
	Co-Supervisore
	Dott.ssa Nathalie Norais

Esame finale anno 2018

Abstract

Despite the great advantages in using glycoconjugates for generating a protective immunity, the molecular mechanism of T cell epitope recruitment and MHC-II presentation has to be clarified. Dissecting the mechanism that controls glycoconjugates/peptide-MHC-II interactions will allow defining a rational for a better design of glycoconjugate vaccines. It is well accepted that polysaccharides function as T cell–independent antigens, since they fail to induce T cell–mediated immune responses (IgM to IgG class switching, booster antibody response and T cell memory). On the contrary, when a polysaccharide is linked to a carrier protein, the protein provides the T cell epitopes that engage the T cell receptor (TCR) and trigger the release of cytokines that help the B cell to differentiate and proliferate. With the aim to analyze glycoconjugate/MHC-II interactions and evaluating the efficacy of glycoconjugate MHC-II processing and presentation, different glycoconjugates were synthesized. ß-1,3-glucans oligosaccharides were covalently linked to the lysine side chains of recombinant proteins from *Neisseria meningitidis* and *Streptococcus pneumoniae*. Testing the glycoconjugates in mice, we highlighted differences in the immune response probably due to different pattern of glycosylation that can lead to a different MHC-II peptide interaction. Using proteomic and glycoproteomic approaches, we evidenced differences in the pattern and extent of conjugation. Mutated recombinant carriers lacking the identified conjugation-sites were produced, conjugated, and tested as carrier in mice. This represents a first step in the design of experiments that will provide insight in the understanding of the peptide/glycopeptide-MHC-II interaction.

Contents

1. Introduction

1.1 History of vaccination

Vaccination represents one of the key developments in fighting infectious diseases. It is estimated to prevent 2.5 million of deaths each year[1]. The history of vaccine began in 1796 when Edward Jenner discovered that people could be protected against smallpox by being previously injected with the cowpox virus. Jenner's idea of using a poxvirus from animal to prevent smallpox was essentially based on believing that an agent virulent for animals might be attenuated in human[2]. The concept of attenuation (i.e. bacteria killed by heat or chemical treatment) was well described and applied by Pasteur almost 100 years later. At the end of the 19th century, he successively developed attenuated rabies virus in animals and humans, anthrax in sheep, and *Pasteurella multocida* in chicken[3]. The concept of isolation, inactivation and injection of a causative microorganism, developed by Pasteur, provided the scientific basis for vaccinology and its evolution during the 20th century. During the 1900s, many discoveries and successes have been registered on vaccine development and the effectiveness of vaccination has been proved on the human health. The impact on public health has been clearly proven by the eradication of smallpox and the reduction of mortality and morbidity caused by various diseases. Over the last three centuries, vaccinology has enabled the realization of numerous vaccines; furthermore modern approaches appear promising for the selection and the design of novel protective antigens that will increase the number of preventable diseases in the future. Understanding the concept of 'antigen' and that the whole pathogen is not always needed to induce immunity, lead to increase vaccines' safety and efficacy[4].

As reported in **Table 1**, the development of vaccines has involved different approaches (e.g the traditional whole pathogens live attenuated; killed or inactivated pathogens; and toxin isolated from the pathogen and inactivated). Subunit vaccines such as proteins produced by recombinant DNA or from genetically detoxified microorganism as well as capsular polysaccharides (CPSs), alone or covalently linked to a carrier protein (glycoconjugates), have been used.

Live attenuated	Killed whole organisms	Purified proteins or polysaccharides	Genetically engineered
18th Century			
Smallpox (1798)			
19th Century			
Rabies (1885)	Typhoid (1896)		
	Cholera (1896)		
	Plague (1897)		
Early 20th Century, first half			
Tuberculosis (bacilli Calmette-Guérin) (1927)	Pertussis (1926)	Diphtheria toxoid (1923	
Yellow fever (1935)	Influenza (1936)	Tetanus toxoid (1926)	
	Rickettsia (1938)		
20th Century, second half			
Polio (oral) (1963)	Polio (injected) (1955)	Anthrax secreted proteins (1970)	Hepatitis B surface antigen recombinant (1986)
Measles (1963)	Rabies (cell culture) (1980)	Meningococcus polysaccharide (1974)	Lyme OspA (1998)
Mumps (1967)	Japanese encephalitis (mouse brain) (1992)	Pneumococcus polysaccharide (1977)	Cholera (recombinant toxin B) (1993)
Rubella (1969)	Tick-borne encephalitis (1981)	*Haemophilus influenzae* type B polysaccharide (1985)	
Adenovirus (1980)	Hepatitis A (1996)	*H.influenzae* type b conjugate (1987)*	
Typhoid (*Salmonella* TY21a) (1989)	Meningococcal conjugate (group C) (1999)	Acellular pertussis (1996)	
Varicella (1995)		Hepatitis B (plasma derived) (1981)	
Rotavirus reassortants (1999)			
Cholera (attenuated) (1994)			
Cold-adapted influenza (1999)			
21st Century			
Rotavirus (attenuated and new reassortants) (2006)	Japanese encephalitis (2009) (Vero cell)	Pneumococcal conjugates*(heptavalent) (2000)	Human papillomavirus recombinant (quadrivalent) (2006)
Zoster (2006)	Cholera (WC only) (2009)	Meningococcal conjugates* (quadrivalent) (2005)	Human papillomavirus recombinant (bivalent) (2009)
		Pneumococcal conjugates*(13-valent) (2010)	Human papillomavirus recombinant (9-valent) (2014)
			Meningococcal group B proteins (Fh-factor) (2014)
			Meningococcal group B (reverse vaccinology) (2015)
*Capsular polysaccharide conjugated to a carrier protein			

Table 1: Outline of the development of Human vaccine [5]

1.2 Carbohydrate based vaccines

Carbohydrates are the most abundant and structurally diverse molecules in nature. They are largely present on pathogens surface such as viruses, bacteria, parasites and fungi. They can form an amorphous layer of extracellular polysaccharide surrounding the cell that may be further organized into a distinct structure termed capsule[6]. Capsular polysaccharides (CPs) have several functions: they protect bacteria from desiccation during transmission from host to host and mediate adhesion of the bacteria to surfaces and to each other. Polysaccharides can also confer resistance to the specific/nonspecific host immunity allowing the pathogen to escape the host immune system by either mimicking the host structures or conferring resistance to the complement-mediated killing[7]. Even though a small set of polysaccharides are poorly immunogenic, most of them can elicit an immune response and be used as pathogen signatures.

CPs are homo or heteropolymers, composed of monosaccharide repeated units joined with glycosidic linkages[7]. These can be recognized by the immune system and induce an immune response[8]. Avan and Heidelberg already described this effect in the 1920, showing that *Streptococcus pneumoniae* CPs are immune reactive components[9-10]. However, the parallel introduction of antibiotics put aside the development of carbohydrate vaccines. Indeed, the first polysaccharide vaccine, PneumoVax (Merck & Co.), was introduced only in the 1983. It consisted of CPs extracted from 14 pneumoniae serotypes. Today, the Pneumo Vax 23 includes 23 out of 90 known serotypes.

(*Merck & Co., Inc. Pneumovax 23 (pneumococcalvaccinepolyvalent).*

Merckwebsite [online] http://www.merck.com/product/usa/pi_circulars/p/

pneumovax_23/pneumovax_pi.pdf)

Unfortunately, CPs vaccines have failed to induce an immune memory response in young children below two years of age and only partially have induced protection in the elderly and patients with chronic diseases[11]. Moreover, every few years, a boost is needed regardless of age because of the vaccine's inability to induce immune memory[12]. The absence of immune memory is mainly due to the mechanism involved in the carbohydrate immune response. Antigens can follow

different pathways of internalization and processing (**Figure 1**). Typically, proteins are thymus –dependent antigens (Td), in which the antigen is internalized and processed by an antigen presenting-cell (APC). Then, the antigen-peptides are presented to the MHC-II complex molecules for the binding with an antigen-specific CD4+T cells via their specific T cell receptor (TCR)[13,14]. The activated CD4+T cells promote the release of cytokines leading to B cell proliferation, maturation, isotype switching and, immunological memory formation[15] (**Figure 1.b**). CPs act as thymus independent antigen (Ti) and do not active T cells[16-17]. They are not presented on the MHC-II but directly activate B cell independently of the interaction of the CD4+ T-cells, avoiding the affinity maturation and isotype switching, as well as the generation of T and B cell memory[18] (**Figure 1.a**). CD4+ T-cell epitopes must be usually provided in the form of a carrier protein in order to recruit CD4+T-cell for antibody responses to the glycan. Avery and Goedel, as early as 1931, reported that the immunogenicity of carbohydrates could be enhanced when the glycans were conjugated to protein scaffolds[19]. Due to the lake of immune-memory with the vaccination of pure CPs the studies moved towards the development of glycoconjugates.
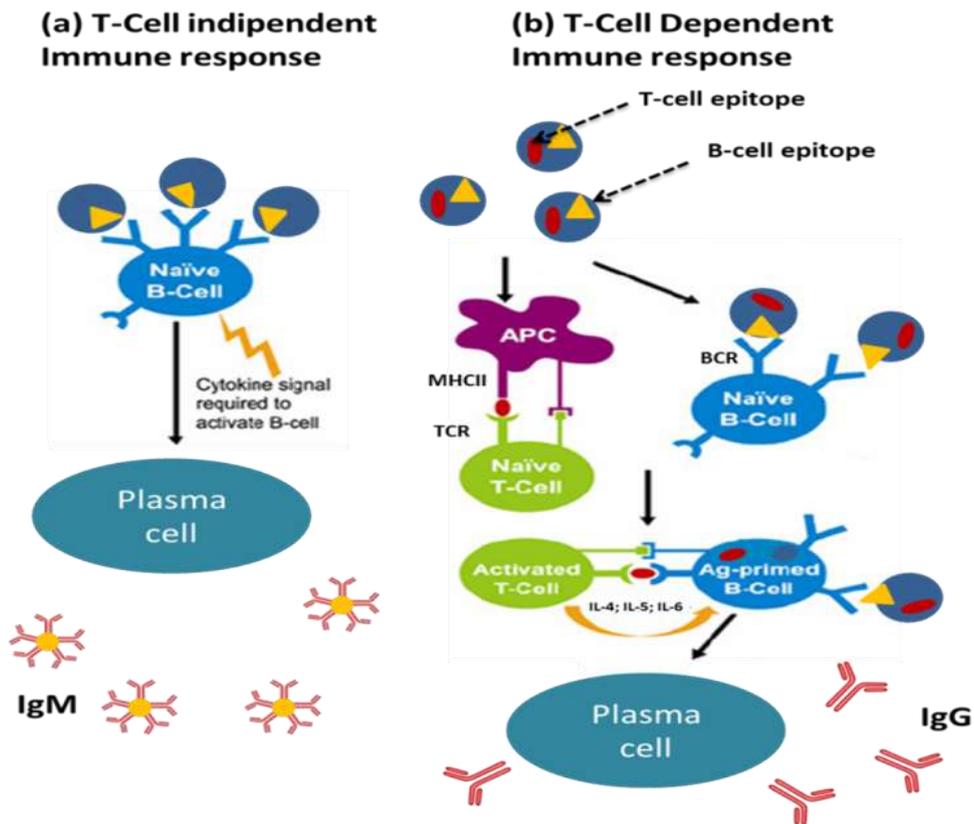
Figure 1 - Pathways of immune cell activation by protein: (1.a) T-*cell-independent pathway.* *(1.b)T-cell-dependent pathway. The generation of immune response by the T-cell-dependent pathway requires the presence of both B-cell and T-cell epitopes. After the recognition of the B-cell epitope by BCR, the B-cell will process and present the T-cell epitope to the MHC class II molecule. Meanwhile the uptake and processing by professional APCs lead to the presentation of the T-cell epitope on MHC class II to naïve T-cells that will be activated for the recognition of the antigen-primed B-cell, releasing cytokine signals required for the switching of B-cells into plasma cells. The T-cell-independent response occurs instead with the activation of the B cell via the BCRs, maturating into plasma cells after the release of cytokine signal[20] (Adapted from Kumar S. et al.).*

1.3 Glycoconjugate vaccines

It is now well known that immunization with capsular polysaccharides covalently linked to an immunogenic protein carrier (Glycoconjugates) induces memory response against the encapsulated bacteria, even among people in high risk groups (neonates, children under two years of age, and elderly).[21] In 1930, Avery & Goebel were pioneers in demonstrating the ability to enhance the immunogenicity of polysaccharide[19]. The authors showed that the poor immunogenicity of the *S. pneumoniae* type III polysaccharide could be enhanced by conjugation of the polysaccharide to a carrier protein.

Earlier studies in 1980s led to the development of the first licensed conjugate based vaccine, in which a capsular polysaccharide of *H.influenzae* type b (Hib) was covalently conjugated to diphtheria toxoid (DT) protein[22]. The success of the Hib conjugate has accelerated the design of other capsule-based vaccines. For two decades, conjugate vaccines against Meningococcus serogroup C, and against seven serotypes of pneumococcus were licensed; in particular in the past decade, the same technology was used to develop vaccines against meningococcus serogroup A, C, Y, W[23,24] and against six extra pneumococcus serotypes leading to a 13-valent pneumococcal conjugate vaccine[25].

Currently, the carrier proteins used for glycoconjugation are typically proteins obtained in sufficient amount and purity. These are generally not toxic and stable under the chemical conditions applied during the conjugation procedures (pH, concentration, ionic strength, temperature)[26]. Specifically, the most used carrier proteins consist in denatured bacteria toxoids, i.e. tetanus toxin (TT), diphtheria toxin (DT) and a nontoxic DT mutant, CRM_{197}[27]. Different chemistries can be applied for the conjugation of polysaccharides to these carrier proteins. All of them take advantage of a suitable spacer and a functional group for conjugation that can be aldehyde, thiol, activated ester, hydrazide, carboxyl acid or amine that can be either presented or generated on the polysaccharides for the subsequent linkage to the protein[28,29]. Until very recently, non-specific methods for site-conjugation of the glycan antigen have been available and most of the glycoconjugates still result in different glycoforms. This heterogeneity may show differences in the immunological properties and pharmacokinetic due to a

different epitope density and glycosylation sites. Hence, increasing attentions have been putting on the chain length, presence of ramification points and charge on the polysaccharide motif. A great effort has been done to improve the conjugation chemistry and the type of spacer employed, which may direct the immune response away from the antigen carbohydrate. At the present, Glycoconjugates represent a promising strategy to combine the antigenic of polysaccharide with the use of a carrier protein for the induction of a T cell dependent response based on the carrier protein-derived peptide presentation by the **MHC-II** molecules on the **APCs.**

1.3.1 Characteristics of Carriers proteins in conjugate vaccines

Production of glycoconjugate vaccines involves the chemical conjugation of glycans to an immunogenic carrier protein to elicit a strong and long-lasting immune response (T-cell dependent). To date, 5 carrier proteins have been used in licensed conjugate vaccines: meningococcal outer membrane protein complex (OMPC), diphtheria toxoid (D), tetanus toxoid (T), *H. influenzae* protein D (HiD) and CRM_{197} a genetically modified cross reacting material of diphtheria toxin. All of these carriers are effective in increasing vaccine immunogenicity but they differ in the quantity and avidity of antibodies produced, the ability to carry multiple antigens in the same product and the efficiency to be given with other vaccines[30]. The first-generation of carrier proteins used are represented by diphtheria toxin and tetanus toxin that required detoxification with formaldehyde. They were initially selected as carriers because of safety track record established over decades of vaccination against tetanus and diphtheria[31]. The OMPC of *N. meningitis* serogroup B is used as carrier for Hib conjugate developed by Merck. Protein D, which derived from non-typeable H. influenza (NTHi), is used as carrier for most of polysaccharides included into the multivalent pneumococcal conjugate vaccine by GSK. At the moment, the most wildly used and highly effective carrier protein is the $CRM_{197,}$ a non-toxic mutant of diphtheria toxin. It differs from the wild-type diphtheria toxin in that a point mutation at the amino acid position 52 substitute glycine with glutamic acid, which eliminates enzymatic activity and toxicity[32]. CRM_{197} shows several advantages as protein used for the conjugation: it is nontoxic and has more lysine side-chains available for conjugation comparing

to the diphtheria toxin. It is still extensively used as carrier for licensed conjugate vaccines against *Streptococcus pneumoniae, Neisseria meningitis* and *H. influenza* type b. Among the 5 carriers, CRM_{197} shows a greater versatility when it is used in combination with other vaccines and also on multiple polysaccharides conjugates[30].

1.3.2 Glycosylation and antigen processing

1.3.2.1 Major histocompatibility complex structure

As already described in this introduction, the MHC-II molecule is involved in the regulation of the adaptive immune system against invading pathogens detecting the presence of foreign pathogens.

The MHC-II histocompatibility complex is a heterodimer composed of two membrane chains, α and β. The proximal membrane domains (α_2 and β_2) are folded into Ig-like domain and a trans-membrane sequence for anchoring. The distal membrane region α_1 and β_1 are folded together in a single groove-shape, which is defined by a β sheets floor and two parallel helical sides[33].

The MHC-II molecule binds peptides of around 12-24 aa in length[34]. It is possible to distinguish two main sub structures of interaction: a central core of binding represented by the groove, which establishes a nonameric binding region and, residues outside the nonameric binding region called peptide flanking residues (PFRs), which is also able to interact with the MHC-II complex (**Figure 2**).

The peptide flanking residues, not bound in the groove, can enhance the binding and stability conformation of the complex[35]. It is also widely accepted that they can contribute to the T cell mediated peptide-MHC-II recognition, and influence the TCR activation as well as the residues in the core binding region[36].
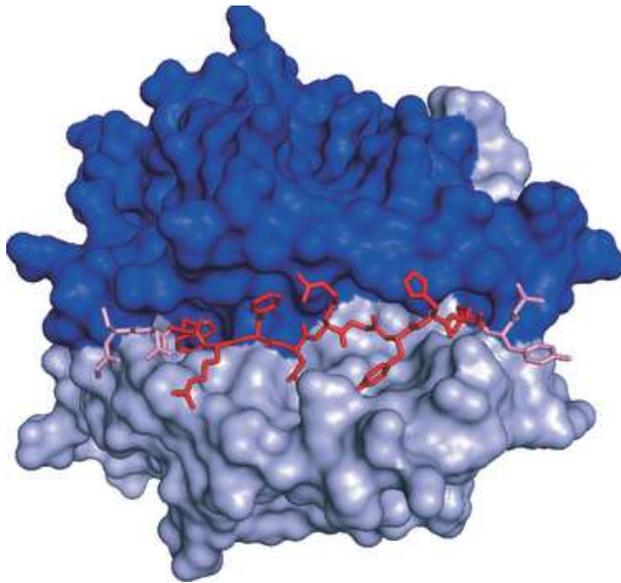
Figure 2 - structure and peptide interaction to the MHC-II: Protein crystal structure of MHC-II molecule (HLA-DRB1*0101) interacting with peptide (PDB id: 1AQD). The α-chain is shown in dark blue, while the β-chain in grey. The peptide (GSDWRFLRGYHQYA) is shown as sticks: in red the peptide binding core and in pink the flanking amino acids.

1.3.2.2 Peptide glycopeptide antigen presentation

The capacity of the T-cell-dependent antigen to recruit CD4[+]T cells during the primary immune response represents a critical step to the generation of a memory immune response[38]. The T cell dependent antigen can be internalized by the B cell using its B cell receptor (BCR) and then processed and presented within the MHC-II on the surface of the B cell[13]. This allows the B cell to interact with the related CD4[+]T cells following the recognition of the peptide/MHC-II complex by the T-cell receptor (TCR). In contrast, T cell independent type 2 (TI-2) antigen including isolated polysaccharides of encapsulated bacteria such as *Neisseria meningitidis* and *Streptococcus pneumoniae* cannot, in general, be presented to the MHC-II complex and are thus unable to recruit CD4[+]T cells for the classical isotype switching of memory B cells[39]. Conjugates vaccines, targeting encapsulated bacteria, are able to overcome the limitation of polysaccharide immune response through their covalent linkage to an immunogenic carrier protein for the presentation to the MHC-II complex[40]. As previously discussed, the current dogma states that polysaccharides act as T cell-independent antigens. Being highly hydrophilic, they cannot dock to the groove of the MHC molecules; therefore, they cannot be presented to the T cell receptor[41] (**Figure 3.a**). Nevertheless, some studies suggest that the polysaccharide can be presented to the MHC-II complex after depolymerization within the APC endosomes/lysosomes by nitric oxide-derived reactive nitrogen species (RNs) and/or superoxide-derived reactive oxygen species (ROSs)[42,43].Few years ago,

Avici et. al.[44] described also a new molecular mechanism, in which the glycoconjugates internalized in the B cells are processed and then presented to the MHC-II as glycopeptides.

The presentation of the carbohydrate motif is possible thanks to the peptide that act as an anchor binding to the MHC-II. This new proposed mechanism also contributes to indicate that the carbohydrate can interact with a specific carbohydrate specific T cell clone (**Figure 3.b**). It is necessary to specify that only small saccharides form 1 to 4 carbohydrate units could be inserted within the binding site of T cell receptor[45], whereas the B cell receptor could theoretically bind up to six units[46]. Even though 4-6 repeating units of carbohydrate can interact with the T cell receptor, the glycoconjugate vaccines present on the market show a different scenario. Except for the *Cuban anti–H influenzae*[47] vaccine, the glycoconjugate vaccines, commercially available, present safety issue due to the inherent lot-to-lot variability. From an immunological point of view, the carbohydrate structures on these glycoconjugates are too large and randomly distributed at the carrier surface, limiting the possibility that carbohydrate processed in the lysosome could result in a sufficient quantity of homogenous glycopeptide suited for the T cell binding[48].

This complex mechanism is still under discussion and the current literature is not able to clarify the mechanism of glycan processing and carrying yet.
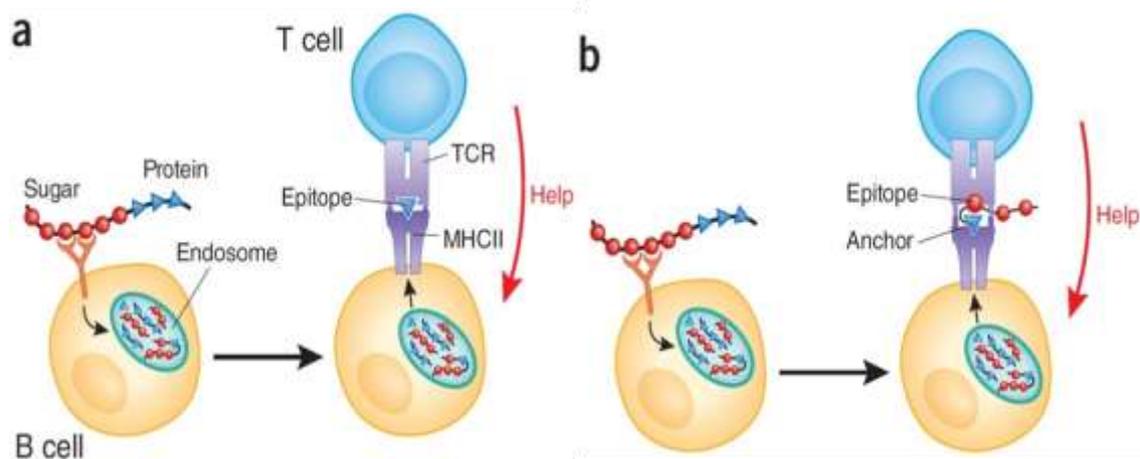
*Figure 3 - **Model proposed for antigen processing and presentation** [41]: (3.a) Current dogma: The peptide epitope is processed, presented by the MHC-II and recognized by a peptide-specific T cell; (3.b) the new mechanism presented by Avci et al.: the peptide presented to the MHC-II acts as an anchor for the sugar epitope, allowing the presentation of the sugar epitope to a polysaccharide-specific T cell. (Adapted from Rappuoli et al. 2011).*

1.4 Background for the study

In a previous study performed in the host laboratory, a group of twenty-seven proteins, derived from different pathogens and expressed as recombinant proteins in *E. coli*, were conjugated to the polysaccharide Laminarin (Lam)[49]. Lam is a β-(1, 3) glucan with β-(1, 6) branches, that has been studied as vaccine candidate against *C. albicans* infections[50-51-52]. The carrier proteins of this study were selected using the following criteria:

- not expected toxicity based on available information

- solubility in physiological buffers at concentrations of 0.4–6 mg/mL

- presence of a sufficient number of lysines for conjugation

- proteins with molecular weight between 40 and 100 kDa

These recombinant carriers were conjugated with Lam modified at its end reducing group. The succinimido diester of adipic acid (SIDEA) was used as linker obtaining in such a way a terminal succinimido ester group available to react with the primary amines on the side-chain of the lysine on the proteins.

The conjugates were tested in mice without the use of any adjuvant highlighting the intrinsic value of the carrier only. Seven proteins (pathogenic *E. coli* Orf3526 and Upec-5211, S. pneumoniae spr907, spr1418, and spr1875 *N. meningitidis* GNA2091-fHbp[53],and *S. agalactiae* RrgB I-II-III, were demonstrated to function as proteins carrier for Lam with a level of anti-Lam antibodies inferior or equal to **1145 (GMT ± 95% CI)** which is comparable to the level of anti-Lam antibodies produced by CRM_{197}–Lam. The only conjugate able to induce an IgG anti-Lam antibody titer significantly higher than the CRM_{197}–Lam was the spr96/2021. Spr96/2021 is a recombinant fusion composed of two pneumococcal proteins: spr2021 (secreted PcsB protein) and spr0096 (LysM protein) [54-55]. Curiously, no response was induced by the spr0096 conjugate and very low by the spr2021 one, indicating that some characteristics derived from the fusion of the two proteins were critical for the carrier activity. By conjugated to meningococcal serogroup C (Men C) oligosaccharides, the spr96/2021 conjugate induced anti Men C polysaccharide IgG titers comparable to those developed by CRM197 conjugate; while Upec-5211, Orf3526, and spr1875 conjugate responses were comparable to the positive control and RrgB I-II-III, spr907, spr1418 and

GNA2091-fHbp conjugates were inferior to the control. Moreover, spr96/2021 was also tested as carrier for serogroup A, W and Y meningococcal saccharides, which are currently part of licensed vaccines[56-57] and serogroup meningococcal X polysaccharide (Men X), recently identified as a potential candidate for vaccine development[58], at a comparable level to CRM_{197} conjugates[49].

As already mentioned in this section, several factors are important in determining the ability of a protein to work as carrier, but the main role of the carrier in glycoconjugate vaccines is to provide T-cell epitopes which bind to the class II MHC. At the time, the data were published and no correlation between the numbers of MHC-II predicted peptides and IgG production was identified.

2. Aim of the study

Glycoconjugates have been demonstrated to very safe and efficient vaccines. In spite of their success, their design is still very empirical, and their mode of action is not well understood. In particular, the influence of the glycosylation on the molecular mechanism of T cell epitope recruitment and MHC-II presentation has to be clarified. Dissecting the mechanism that controls glycopeptide/peptide MHC-II interactions will allow defining a rational for a better design of glycoconjugate vaccines. For this purpose, the aim of the thesis is to:

- Draw a hypothesis on the parameters able to influence the efficacy of a protein to function as a carrier for glycoconjugate vaccine by comparing production of anti-saccharide IgG of a large number of potential carriers reported in literature to the presence of prediction of MHC-II-presenting peptides on these carriers.

- Select a carrier protein model to validate the hypothesis that efficacy of a protein carrier is correlated to predicted MHC-II-presenting peptides with high affinities and that when these predicted MHC-II-presenting peptides contained lysine residues, these ones should not be conjugated.

- Use mass spectrometry (proteomics and structural mass spectrometry) to assess the pattern and extent of conjugation of the carriers (*S. pneumoniae* carrier protein spr2021 alone or in fusion constructs).

- Correlate the efficacy in mice of the *S. pneumoniae* carrier protein spr2021 with the absence of conjugation on lysine residues present in predicted MHC-II-presenting peptides, by assessing the pattern and extent of conjugation of the carrier alone or in fusion constructs by proteomics.

- Based on the hypothesis and analytical data, construct and validate mutated protein carriers with a higher efficacy in terms of IgG production.

3. Materials and methods

3.1 Mutants' generation of GNA2091-fHbp, spr2021/96 and spr2021

Site-directed mutagenesis on GNA2091-fHbp, spr2021/96 and spr2021 DNA sequences was performed to produce the set of mutants listed below. Mutants were constructed using the polymerase incomplete primer extension cloning method (PIPE). The mutagenesis reactions were transformed into chemically competent *E. coli* Mach1™-T1^R (Thermo Scientific). After sequencing, each plasmid was used to chemically transform *E. coli* BL21(DE3) T1R cells (NEB) for protein production.

spr2021 mutant	Plasmid characteristics
Spr2021_Lys204Arg	pET151/D-TOPO derivative for expression of recombinant spr2021, containing a Lys204Arg mutation, AmpR
Spr2021_Lys248Arg	pET151/D-TOPO derivative for expression of recombinant spr2021, containing a Lys248Arg mutation, AmpR
Spr2021_Lys204Arg_Lys246Arg_Lys248Arg	pET151/D-TOPO derivative for expression of recombinant spr2021, containing a Lys201Arg, Lys256Arg,Lys248Arg mutations, AmpR
Spr2021/96 mutant	**Plasmid characteristics**
Spr2021/96_Lys172Arg	pET21b(+), derivative for expression of recombinant spr2021, containing a Lys172Arg mutation AmpR
Spr2021/96_Lys216Arg	pET21b(+), derivative for expression of recombinant spr2021/96, containing a Lys216Arg mutation AmpR
Spr2021/96_Lys172Arg_Lys216Arg	pET21b(+), derivative for expression of recombinant spr2021/96, containing a Lys172Arg and Lys216Arg mutations AmpR
Spr2021/96_Lys172Arg_Lys214Arg_Lys216Arg	pET21b(+), derivative for expression of recombinant spr2021/96, containing a Lys172Arg, Lys214Arg and Lys216Arg mutations AmpR
Spr2021-CRM$_{197}$-96	pET21b(+), derivate for expression of recombinant spr2021/96 in which the linker GSGSGGGG have been substituted with GIALSSLMVAQAIPLVG , AmpR
GNA2091-fHbp mutant	**Plasmid characteristics**
GNA2091-CRM$_{197}$-fHbp	pET24 derivative for expression of recombinantGNA2091-fHbp, which the linker GSGSGGGG have been substituted with GIALSSLMVAQAIPLVG, AmpR
GNA2091-2021-fHbp	pET24 derivative for expression of recombinantGNA2091-fHbp, which the linker GSGSGGGG have been substituted with GRRASQQQSVLASANTG, AmpR

3.1.1 Protein expression and purifications

Escherichia coli strain BL21(DE3) cells (Novagen) were used for protein expression. Cells were grown using BioSilta medium (Enpresso B Animal-free growth systems), at 30 °C for 12 h, and recombinant protein expression was induced by the addition of 1mM isopropyl β-D-1-thiogalactopyranoside (**IPTG**) at 25 °C. After an additional 24 h, cells were harvested by centrifugation and re-suspended in 50 mM NaH_2PO_4, 250 mM NaCl, pH 7.4, followed by lysis via sonication (Qsonica Q700). Cell lysates were clarified by centrifugation at 10000 g for 30 min, and the supernatant, containing the expressed protein, was filtered using a 0.22 µm membrane filters (EMD Millipore filters) before starting the first chromatography step. All proteins were purified at room temperature (RT, 25 °C) using an AKTA purifier 100 system (GE Healthcare). A first purification step by using Co^{2+} affinity chromatography (5 mL HiTrap TALON crude, GE Healthcare) was performed for those proteins that presented His tag motif. Conversely, a first purification step by ionic exchange chromatography, either anionic or cationic depending on the isoelectric point of the proteins by 5 mL SP or Q HP (GE healthcare) columns, was carried out if the His tag motif was not evidenced.. A second step of purification was performed using Hydrophobic chromatography (4.7 mL Hiscreen Phenyl HP), followed by size-exclusion chromatography on a Superdex 200 16/60 column equilibrated in 20 mM NaH_2PO_4, 50 mM NaCl, pH 7,4). The purity of the proteins was checked using 4–12% **SDS–PAGE** gradient gels in **MES** buffer. The content of lipopolysaccharide (LPS) on the purified protein was checked using the Endosafe nexgen-PTS system (Charles River). When the content of LPS was out of the range, they were removed using either Q HP (GE healthcare) columns or EndoTrap Red columns (Hyglos).

3.1.2 Differential Scanning Calorimetry

The thermal stability of the proteins was assessed by differential scanning calorimetry (DSC) using a MicroCal VP-Capillary DSC instrument (GE Healthcare). A protein concentration of 0.4 mg/mL in PBS buffer was used to prepare samples. The DSC temperature scan ranged from 20 °C to 140 °C, with a thermal ramping rate of 150 °C/h and a 4 s filter period. Data were analysed by

subtracting the reference buffer data through the use of the Origin 7 software. The mean values of the melting temperature (T_m) were determined.

3.2 Preparation of oligosaccharide

3.2.1 Reductive amination

Laminarin (Sigma-Aldrich) was treated by reductive amination to introduce an NH_2 group at the reducing end of the sugar through the use of the following conditions: Polysaccharide with a final concentration of 4 mg/mL was added with sodium acetate to a final concentration of 300 g/L. Sodium cyanoborohydrate was added in a quantity of 1/5 mole of ammonium acetate. The pH of the reaction was checked and constrained between 7.4 and 7.6. The polysaccharide was incubated for 5 days at 50 °C. The aminated laminarin was then purified from the reaction mixture by Tangential fast flow on regenerated cellulose membrane (cut-off 1 kDa; Sartorius Stedim) and the amount of amino group introduced was determined. Amino group was determined by colorimetric assay as described in paragraph 3.7

3.2.2 Derivatization of Laminarin amino polysaccharide to active ester

The purified aminated polysaccharide was dried under-vacuum and re-suspended to a concentration of amine of 40 µM/mL in *H_2O:DMSO* 1:9 (*v/v*). Thriethylamine was added in a molar excess of 5 fold compared to the primary amine, and Succinoamido diester of adipic acid (*SIDEA*) was added in a 12 fold molar excess compared to the primary amine. The reaction was kept under gentle stirring for 2 h at room temperature. The activated polysaccharide was purified form the reagents by precipitation with ethylacetate. The precipitate was washed 10 times and then dried under vacuum. Content of N-hydroxysuccinimide ester introduced was determined as described in paragraph 3.7

3.3 Conjugation of polysaccharide with carrier proteins and purification of glycoconjugate

The SIDEA activated Laminarin was conjugated to the carrier proteins. The conjugation was carried out in 20 mM NaH_2PO_4, 50 mM NaCl pH 7.2; the reaction was carried overnight with gentle stirring. The molar ratio carrier protein: active ester polysaccharide was kept at 20:1. The conjugates obtained were purified from the unreactive sugar by ultrafiltration on polyethersulphone membrane with a cut-off of 10 kDa (Vivaspin, Millipore) using PBS pH 7.4.

3.3.1 Determination of the total and free saccharide in the conjugates preparation

Total and free laminarin saccharide was determined by anionic exchange pulsed amperometric detection (HPAE-PAD, (Dionex ICS 3000 system).

3.3.1.1 Laminarin saccharide determination

Total laminarin was determined by HPAE-PAD analysis using a CarboPAC1 column coupled to a CarboPAC1 guard column connected to a Dionex ICS3000 system. Sample was treated with 2 M trifluoroacetic acid (TFA), heated at 100 ºC for 4 h in a closed screw-cup tube. The sample was then dried under vacuum, dissolved in water and then filtered. The separation was performed with a flow rate of 1 mL/min using an isocratic elution of 40 mM NaOH for 12 min and followed by a washing step of 5min with 500 mM NaOH. The chromatographic data were processed using Dionex ChromeleonTM software. The glucose was used for generating the calibration curve, at a concentration range of 0.5–10.0 µg/mL.

3.3.1.2 Determination of laminarin free saccharide content in the conjugates preparations

Determination of free saccharide in the conjugate formulation was determined through the separation of the free saccharide and the conjugated by using the

cartridges Bioselect C4 300 Å. The saccharide quantification has been then performed as described in the previous paragraphs.

3.4 Protein and glycoconjugate characterization

3.4.1 Entire protein mass analysis

To verify the molecular of the proteins and mutants produced entire mass was performed. The samples were diluted with formic acid (Sigma-Aldrich) 0.1%, and injected on a Waters SynaptG2 ESI mass spectrometer equipped with a standard ESI source (Waters®). The protein samples were trapped and desalted for 2 min at a flow rate of 800 µL/min using a Protein Micro Trap column (Michrom BioResources) equilibrated with 100% buffer A (0.1% formic acid in water (v/v)). Proteins were directly eluted into the mass spectrometer at a flow rate of 600 µL/min with 55% solvent B (acetonitrile/water, 0.1% formic acid (v/v)). The ESI source was set as following: capillary voltage, 3.0 kV; sampling cone, 35V; extraction cone, 4 V; source temperature, 80 °C; desolvation gas flow and temperature, 600 L/h and 180 °C, respectively; cone gas flow, 20 L/h; trap collision energy, 4 V. Mass spectra were acquired in resolution mode (m/z 100-2000) and the calibration was performed in positive mode using a 2 mg/mL cesium iodide (Sigma) solution prepared in 50% isopropanol. The spectra were processed with MassLynx 4.1 software (Waters).

3.4.2 Chemical deglycosylation

For the identification of the sites of conjugation a chemical method using Trifluoromethanesulphonic acid (TFMSA) for a specific solvolysis of glycosidic bonds was chosen for the removal of carbohydrates from the glycoconjugates. In detail the glycoconjugates were desalted by ultrafiltration using a polyethersulphone membrane with a cut-off of 10 KDa (Vivaspin, Millipore) and then freeze-dried for at least 24 h. Dried glycoproteins were treated with pre-cooled TFMSA containing 10% (*v/v*) anisole in anhydrous condition and incubated at -20 ºC for 2 h. The TFMSA/conjugates solution is then neutralized in a dry ice/

ethanol bath, by a 40% pyridine/water (v/v) a ratio v/v 3:1 to the volume of TFMSA initially added. The de-glycosylated carriers were isolated from the reagents by dialysis against a solution of 20 mM Sodium phosphate pH 7.4.

3.4.3 Glycoconjugates digestion and LC-MS/MS analysis

Glycoconjugates and de-glycosylated proteins were characterized by LC-MS/MS analysis. 25 μg of sample was denatured 10 min at 100 ºC with Rapigest® (Waters). If a protein contains disulphide bridges 5 mM DTT were added to the denaturation solutions for the sequent alkylation of side chain of cysteine, iodoacetamide at 50 mM final concentration was added and the reaction was allowed to proceed in the dark for 30 min.

Trypsin (Gold Mass Spectrometry Grade Promega®) was added in a ratio protein/enzyme of 1:20 (w/w). Digestion was carried out at 37 °C overnight. The digestion was stopped by adding formic acid FA at 0.1% (v/v) solution to acidic pH (pH~2).

The peptide mixtures were then desalted using OASIS cartridges (Waters) following the manufacturer's protocol. Desalted peptides were concentrated with a Speedvac (Eppendorf) and re-suspended in 50 μl of 0.1% (v/v) FA.

Peptides solution were analyzed by LC-MS/MS performed on a nanoAcquity UPLC system (Waters®) coupled a Waters SynaptG2 ESI mass spectrometer equipped with a nanospray source (Waters®). Samples were loaded onto a trap Symmetry C18 180 μm x 20 mm, 5 μm (Waters®) using a full loop injection at a flow rate of 600 nL/min in a mobile phase A (0.1% FA). Peptide were than separated on a nano Acquity UPLC Peptide BEH C18 Column 75 μm x 100 mm (Waters®) using a 70 min gradient 3-98% mobile phase B (98% (v/v) ACN, 0.1% (v/v) FA) at a flow rate of 300 nL/min.

An automated data-dependent acquisition (DDA) using the Mass Lynx software (waters) was used for the acquisition of the eluted peptides. An MS survey scan was used to automatically select multi charged peptides over the *m/z* ratio range of 300–2,000 for further MS/MS fragmentation. Up to eight different peptides were individually subjected to MS/MS fragmentations following each MS survey scan. After data acquisition, individual MS/MS spectra were combined, smoothed, and centroided using ProteinLynx, (Waters®) to obtain the peak list file (pkl).

Peptide and glycopeptides identification was carried from the generated peak list using the Mascot engine software (Matrix Science). Search parameters were:

i. Fixed modification was the carboamidomehtylation on the Cysteine residues (when cysteine alkylation was performed).

ii. Variable modifications for carriers and deglycosylated carriers were: methionine oxidation, glutamine and asparagine deamidation; SIDEA-Glc (+291,13180 Da) was added to the variable modification for the analysis of the deglycosylated carriers.

iii. Peptide mass tolerance as 0.5 Da, peptide MS/MS tolerance as 0.5 Da, missed cleavage = 2, ion charge states: +2, +3, +4). Only significant hits were considered as defined by the Mascot scoring and probability system >15.

3.4.4 Epitope Mapping and HDX

3.4.4.1 Peptidic map

Neisseria meningitidis proteins GNA20191-fHbp and fHbp (which do not contain cysteine residue) or Streptococcus pneumoniae proteins spr2021, spr96-2021 and spr2021-96 (which contain disulfide bridge) were treated for 1 h at 60 ºC with 2 M guanidinium chloride or 3.5 M Urea, 0.4 M TCEP, 0.2 M Guanidinium chloride. Following this denaturing treatment, proteins were subsequently injected into nanoACQUITY ultra-performance liquid chromatography system for the digestion, that was performed online for 2.5 min at 20 °C with a pepsin column (Poroszyme® Immobilized Pepsin Cartridge). The generated peptides were trapped, concentrated and desalted using a pre-column (VanGuard BEH 1.7 µm, 2.1x5 mm) and separated using a ACQUITY UPLC BEH C18 reverse phase column, 1.7 µm, 1.0x100 mm (Waters) with a linear gradient from 3 to 45% of acetonitrile/water, 0.1% formic acid over 6.8 min at 40 µl/min. The mass spectra were acquired in a resolution mode (m/z 100-2000) on a Waters SynaptG2 mass spectrometer equipped with a standard ESI source. Mass accuracy was ensured by continuously infusing a GluFib solution (600 fmol/µL in 50% acetonitrile, 0.1% formic acid) through the reference probe of the ESI source. The identity of each peptide was confirmed by MS^e analyses. MS^e was directly performed by a succession of low (6 V) and high collision (25 V) energies in the transfer region of

the mass spectrometer. Data were processed using Protein Lynx Global Server 3.0 (Waters). The DynamX 3.0 software (Waters) was used to select the considered peptides. Only the peptides present in at least four samples out of five were considered for the analysis.

3.4.4.2 Protein deuteration

54 pmol of protein was diluted in PBS deuterated buffer (excess of deuterium 90%). The labeling exchange was monitored at 6 different time points: 30 s, 5 min, 10 min, 30 min, 24 h. The labeling reaction was quenched either with 7 M Urea, 0.8 M TCEP, 0.4 M guanidinium chloride pH 2.4 or 2 M guanidinium chloride at pH 2.4 depending on the proteins immediately frozen in liquid nitrogen and stored in dry ice. Labelled proteins were thawed rapidly to 0 ºC and injected into the Synapt_G2 mass spectrometry system, where the AQUITY UPLC BEH C18 reverse phase column, 1.7 μm, 1.0x100 mm (Waters) and associated tubing were kept to 0 ºC to limit the back exchange. The Dynamix 3.0 software was used for the subsequent analysis of the labelled peptides.

3.5 Vaccine Formulation and Immunoassay

3.5.1 Immunization schedule

Formulations were made using PBS and isotonic saline solution, under sterile wood. All the formulation had a pH range of 7.4 ± 0.5 and an osmolarity of 300 ± 60 mOsm/Kg. When allum phosphate was added as adjuvant to the formulations the pH range was kept to 6.5 ± 0.5 and the osmolarity to 300 ± 60 mOsm/Kg. Proteins have been administered to mice in 5 μg protein content for dose; Glycoconjugates with laminarin were administered to mice in 5 μg for dose in polysaccharide content.

Balb/c mice female were immunized subcutaneously at day 1, 14 and 28. Bleeding was performed at day 0 (pre-immune) day 28 (post2) and day 42 (post3).

All animal studies were carried out in compliance with the arrive guidelines, the current Italian legislation about the care and use of animals in experimentation (Legislative Decree 116/92), and with the GSK Animal Welfare Policy and Standards. Protocols were approved by the Italian Ministry of Health (authorization 249/2011-B).

3.5.2 Antibody titer determination by ELISA assay

The Antibody response induced by protein carriers and glycoconjugates were measured by ELISA. The coating was done overnight at 4 °C on Maxisorp plates by adding 100 µL/well of either protein or polysaccharide. Coating for the proteins was done using 2 µg/mL protein solution in PBS buffer at pH 7.2. Instead for the polysaccharide coating, was performed at 50 µg/mL in 0.05 M sodium carbonate buffer at pH 9.6. The coating solution was removed by three washes with PBS buffer containing 0.05% of tween 20 (**TPBS**) (Sigma-Aldrich). A blocking step was then performed using adding a 100 µL/well solution of 3% (w/v) BSA in TPBS and incubating the plates at 37 °C for 1 h. The blocking solution was then removed by washing three times the plates with TPBS. 200 µL of pre-diluted serum (diluted in TPBS containing 0.3% (w/v) of BSA) is added on the first well of each column while on the others wells 100 µL of TPBS 0.3% (w/v) BSA was dispensed. A 2-fold dilution on each well was performed transferring 100 µL of sera solution from well to well. The plates were incubated for 2 h at 37 °C. After incubation the solutions were removed and the plates were washed three times with TPBS, and a secondary alkaline phosphate conjugate antibodies (antimouse IgG 1:1000) in 100 µL TPBS containing 0.3% (w/v) BSA solutions is added for 1 h at 37°C. After three washes with TPBS 100 µL /well 0.5 M diethanolammine buffer pH 9.6 supplemented with 1 mg/mL *para*-Nitrophenylphosphate (p-NPP) (Sigma-Aldrich) was added and the plates were incubated at room temperature for 30 min and then read at 405 nm by using plate reader Biorad. Raw data were acquired with Microplate Software (BioRad). Sera titers were expressed as the reciprocal of sera dilution corresponding to a cut-off Ottical density (OD) of 1 in the case of carrier protein IgG titer determination or OD = 0.2 in the case of polysaccharide IgG titer determination. The immunization group was represented with the

geometrical mean of the single mouse titer with a confidential interval of 95%. The graphical analysis was done using GraphPad Prism 6.0 software.

3.6 Bioinformatic analysis for MHC-II prediction

The prediction analysis was performed using the IEDB analysis source. The tested proteins were analyzed applying both the Consensus method(63) (SMM/NN) and the SMM-align methods[64.].

3.7 Colorimetric assays for the determination of proteins concentration and oligosaccharides activation

During the purification process, protein content was determined by Bradford colorimetric assay. In the conjugates, it was determined by BCA protein assay[59]. Amino groups were determined by colorimetric assay[60].

Active ester groups, introduced in the oligosaccharides, were determined by the analysis of released N-hydroxy-succinimido groups[61].

4. Results

As extensively reviewed in the introductive section, several factors are important to determine the ability of a protein to function as a carrier. The main role of the carrier in glycoconjugate vaccines is thought to provide T-cell epitopes, which bind to the class II MHC. In the specific case of glycoconjugate vaccines, understanding the role of presenting peptides as well as the impact on the carrier efficacies, when potential MHC-II peptide binders are glycoconjugated, is fundamental for the selection of efficient new protein carriers and for the rational design of conjugation strategies.

4.1 Selection of carrier proteins for the study

4.1.1 Efficacy of the protein carriers previously tested to produce IgG anti-saccharide

In order to draw a working hypothesis, concerning on the importance of the MHC presenting peptides and their glycosylation, we analyzed the efficacy, in term of production of IgG anti-saccharide, of a set of 27 different carrier proteins reported in *Tontini et al.*, as a function of the affinity of the predicted MHC-II binder peptides. The investigated carrier proteins derived from different bacteria: extra-intestinal pathogenic *E.coli* (ExPEC), *N. meningitidis* serogroup B, group A *Streptococcus* (GAS), group B *Streptococcus* (GBS), and *S. pneumoniae* (spr); while the sugar antigens used for the conjugation to these carriers was laminarin. These proteins were selected on the basis of an absence of expected toxicity considering the available information: solubility in physiological buffers at concentrations 0.3 - 5.0 mg/mL; presence of a sufficient number of lysines for conjugation and molecular weight between 40 and 100 kDa. These were all conjugated with modified laminarin at the end reducing group. The succinimido diester of adipic acid (SIDEA) was used as linker to have a terminal succinimido ester group available to react with the primary amine group of lysine residues on the proteins. All the conjugates were tested in mice without the use of adjuvant in

order to highlight the intrinsic value of the carriers only. CRM$_{197}$–Lam was used in each immunogenicity study as benchmark. As determined by ELISA, 15 of the 27 conjugates did not induce anti-Lam antibody titers or induce them to a level of 5 time inferiors[49].

4.1.2 MHC-II binding prediction of the 27 carrier proteins studied

Access to bioinformatics analytic tools and curated databases on immune reactions and specific pathogens have been becoming more frequently crucial for the discovery of epitopes and the vaccine design. Currently, the Immune Epitope Database and Analysis Resource is the most common and used bioinformatics tools due to its freely availability on the IEDB website (http://www.iedb.org//). It includes several practical tools and methods to analyze and forecast epitopes, but also a broad amount of immune epitopes experimentally measured. For instance, the IEDB analysis resource software was applied for a direct prediction of the peptides presented to the MHC-II of all the carrier proteins analyzed in the study of *Tontini et al.* Since the 27 carrier proteins were tested on the Balb/c mice; the specific haplotype of interest, H2-IAd was used for the prediction. According to the peer-reviewed scientific literature, new available data and information are periodically used to update the database. Particularly, the last update reports has showed a number over 18940 of curated references, a database containing over 424000 epitopes and over 1500000 B cell, T cell, MHC binding, and MHC ligand elution assays (positive and negative)[62].

The software, directly available on the IEDB website at the following url:http://tools.iedb.org/mhcii, allows comparing the predicted affinity of every peptide with that of a greater sample randomly selected. It calculates the percentile rank by using various available prediction methods including, Consensus[63-64], netMHCIIpan[65], SMM-align, Sturniolo, and NN-align[66].

In this study we tested proteins using the SMM-align method[67], which parsed the input protein sequences into a series of 15-mers that overlap by 10 residues, and submited these series instead of the entire protein sequence. With an overlap of 10 amino acids it was possible to capture the minimal number of 15-mers with all possible 9-mers binding cores with at least one flanking residues (PFRs) on both sides and predicts the binding affinity for each peptide. The results of these

analyses were provided as a nominal SMM-align rank (nominal value) and a prediction SMM-align (IC50 nM). Briefly the lower is the SMM-align rank, the higher is the binding interaction with the MHC-II. The SMM-align method predicts the IC50 of the peptide MHC binding affinity values, making it suitable for rational epitope discovery[67]. As guide, we considered a good and medium affinity for the predicted peptide ≤1000 nM[68].

As examples of the outcome of the software, **Figure 4** shows the predictions for the proteins CRM$_{197}$ (see **panel a**), spr96/2021(see **panel b**) and GNA2091-fHbp (see **panel c**). For a better visualization, the data in the figure are shown as reciprocal of the nominal percentile rank, where the higher is the value, the stronger is the affinity for the specific MHC-II haplotype. The CRM$_{197}$ protein showed a high affinity for the MHC-II, displaying IC50 between 539 nM and 615 nM for the three peptides with high SMM-align rank. The spr96/2021 (see **panel b**) exhibited stronger binder affinity of predicted peptides compared to the CRM$_{197}$ protein and it had a high affinity with a low IC50 (see **panel b**).The GNA2091-fHbp (see **panel c**) had a lower number of predicted MHC-II binders compared to the CRM$_{197,}$ and spr96/2021, having only few peptides with an IC50 lower than 1000 nM.

Because the side chain of lysine residues consisted in the site of conjugation, a specification were needed: in the case of CRM$_{197}$ the predicted epitopes did not present any lysine residues in their sequences, while spr96/2021 and GNA2091-fHbp, contained lysine residues. It is remarkable to stress it out considering that the MHC-II presentation could be influenced by the poly/oligosaccharide conjugation.
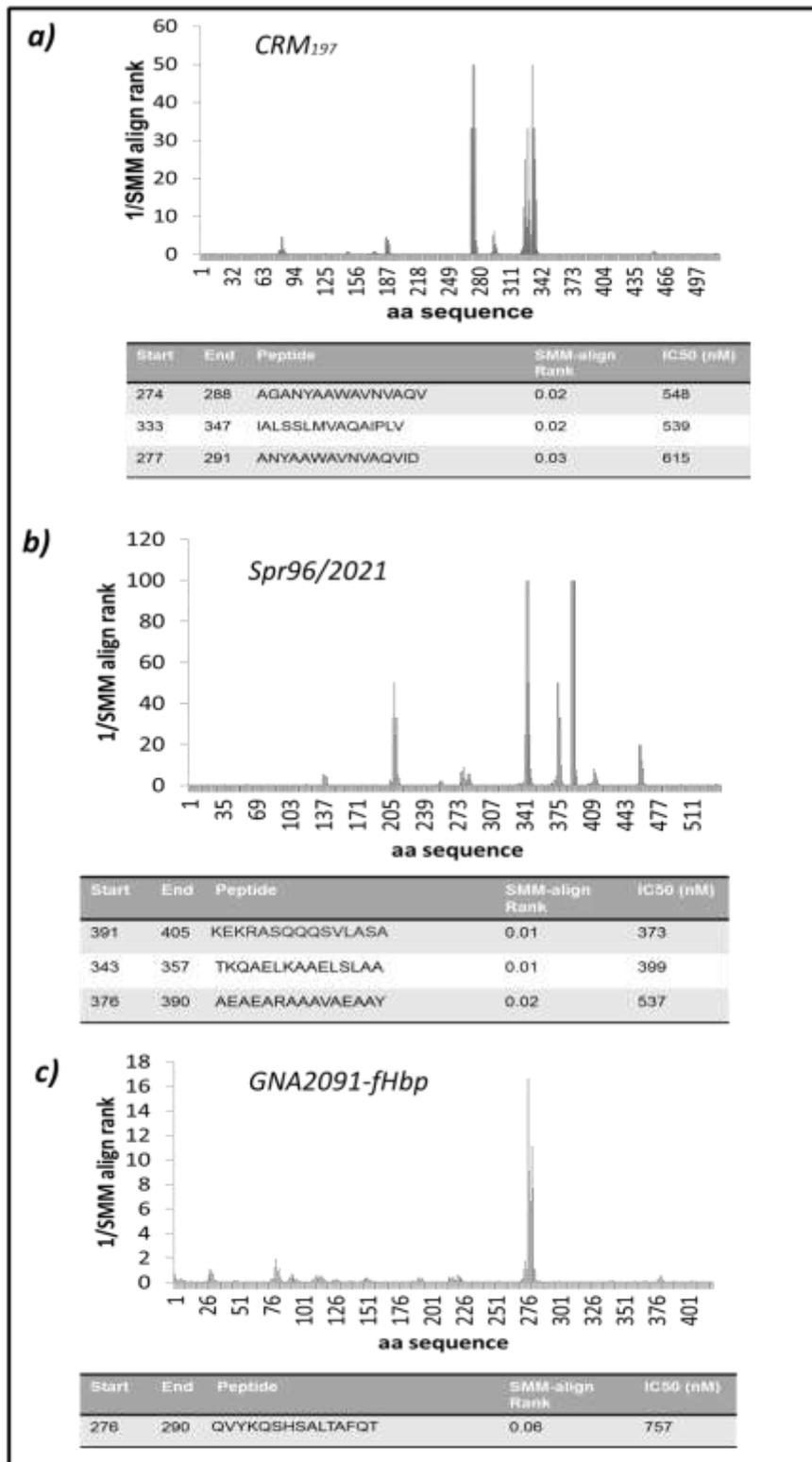
a)

CRM$_{197}$

Start	End	Peptide	SMM-align Rank	IC50 (nM)
274	288	AGANYAAWAVNVAQV	0.02	548
333	347	IALSSLMVAQAIPLV	0.02	539
277	291	ANYAAWAVNVAQVID	0.03	615

b)

Spr96/2021

Start	End	Peptide	SMM-align Rank	IC50 (nM)
391	405	KEKRASQQQSVLASA	0.01	373
343	357	TKQAELKAAELSLAA	0.01	399
376	390	AEAEARAAAVAEAAY	0.02	537

c)

GNA2091-fHbp

Start	End	Peptide	SMM-align Rank	IC50 (nM)
276	290	QVYKQSHSALTAFQT	0.06	757

Figure 4 - *Examples of SMM-align predictions: CRM$_{197}$ (**panel a**), spr96/202 (**panel b**), GNA2091-fHbp (**panel c**) data prediction are shown as reciprocal of the nominal percentile rank (y axis), while the x axis represents the amino acid composition starting*

4.1.3 Predicted MHC-II peptide affinities as a function of carrier efficiencies

The emitted working hypothesis was very simple. We wondered if a read-out as macroscopic and easy that the measurement of IgG amount could reflect a process as complex as the binding, internalization, processing, presentation on MHC-II and creation of the synapse with the T-cell receptor. For this task, the anti-laminarin IgG titer, for each glycoconjugate, was plotted as the affinity-value of the carrier peptide with the highest affinity (**Figure 5**). We arbitrary defined a protein as an efficient carrier whether the anti-laminarin IgG was equal or superior to 1/5 of the anti-laminarin IgG produced by the benchmark CRM_{197}, and according to the recommendation of the IEDB, we selected as high-affinity peptide those with a 1/SMM-align rank of 8, i.e. a value equivalent to IC50 not higher than 1000 nM. Based on the threshold used, four groups of carriers have been identified:

I. Eleven carrier proteins able to induce high anti-Lam IgG titers that we arbitrary defined as superior or equal at 1/5 of the anti-Lam IgG titers induced by CRM_{197} contain at least one predicted high affinity MHC-II peptide (**Figure 5**, group **a**).

    This group indicated a potential correlation between the presence of at least one high MHC-II affinity peptide and a high efficacy of the protein as glyco-carrier. It is also interesting to notice that, with the exception of the two fusion proteins RrgB III-II-I and RrgB I-II-III, the spr2021 (alone or fused with the protein spr0096), and the GNA2091-fHbp, all the other carriers of this group did not contain lysine residues so as to indicate that the conjugation should not influence the binding of the predicted peptides. The protein spr2021was particularly interesting because it could function as an efficient or even highly efficient carrier.

II. Six carriers, unable to induce high IgG titers, contained only low affinity predicted MHC-II peptides (**Figure 5**, group **c**).

    As for the previous group, this observation evidenceed a potential correlation between the presence/absence of high affinity MHC-II peptides

and level of IgG titers. In this particular group, the absence of predicted MHC-II peptides resulted in the incapability of the carrier to induce anti-saccharide IgG.

III.   Seven carrier proteins, unable to induce high IgG titer, contained **high** affinity predicted MHC-II peptides (**Figure 5**, group **b**).

At priory, this set of protein carriers did not fit with a model of "prediction of high affinity MHC-II peptide correlated with high efficacy of the carrier protein". Specifically all those carrier predicted MHC-II peptides, except one, contained high affinity predicted peptides carrying lysine residues, to indicate the importance of lysines and their potential state of conjugation in the immune response.

IV.   Two carrier proteins able to induce high IgG titer contained low affinity predicted MHC-II peptides (**Figure 5,** group **d**).

These two carriers did not fit with the model and might represent an error in the prediction of the MHC-II binder peptides
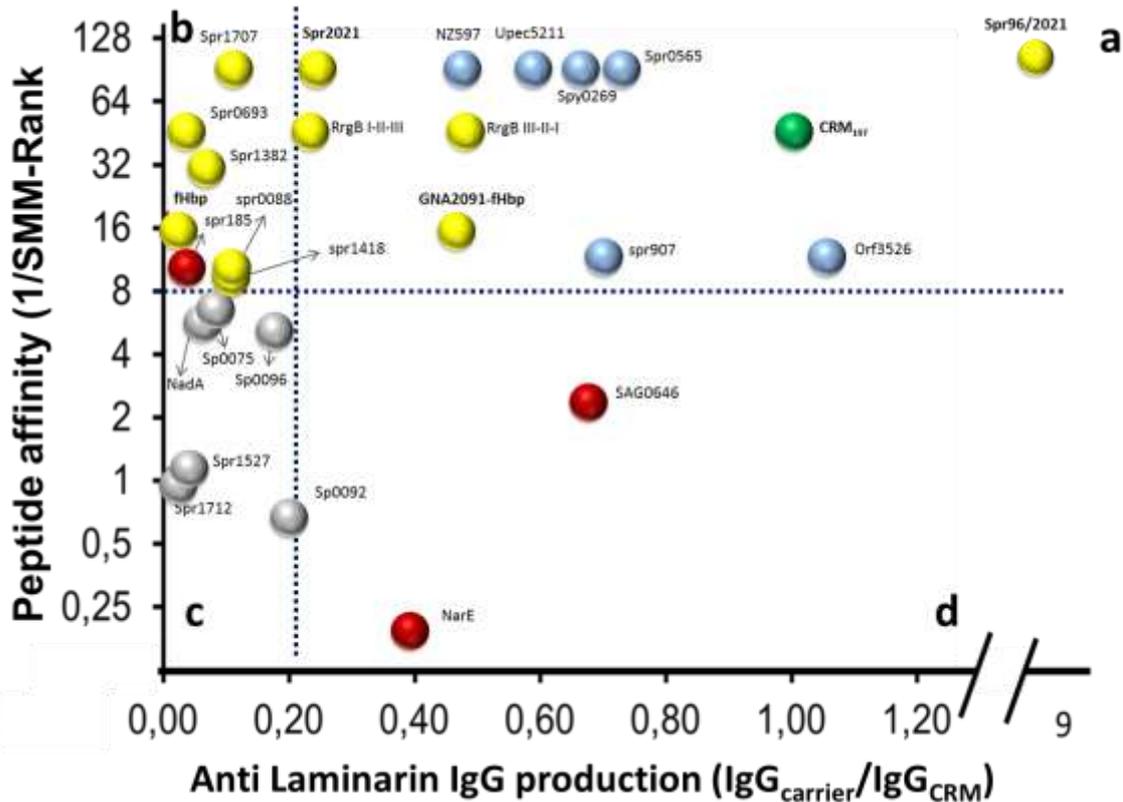
Figure 5 - *The Y axis is the reciprocal of the SMM rank of the carrier peptide with the highest affinity (the higher the value is, the stronger the MHC-II prediction binding is); the X axis is the anti laminarin IgG titer normalized by CRM$_{197}$ which is used as benchmark The vertical dashed line represents the separation between efficient (superior to) and non-efficient (inferior to) carriers. The horizontal dashed line represents the separation between high and low affinity peptides. Each dot represents a different protein. The CRM$_{197}$ is represented in **green**, while in **blue** and **yellow** respectively the proteins containing high affinity predicted peptides with and without lysine residues. In **grey** the proteins with low immune affinity and Immune response, in **red** the proteins that do not fit the model*

4.1.4 Elaboration of a working hypothesis and design of the study

The following observations allowed drawing a working hypothesis:

1. The software of prediction applied in this analysis allowed the hypothesis of correlation between presence on the carrier proteins of a high affinity MHC-II binder peptides and IgG production efficacy could be evidenced.

2. The MHC-II peptide prediction allowed speculating that conjugated/unconjugated lysine residues in the predicted MHC-II peptides may drive the production of IgG.

3. The production of IgG seemed to be driven by the peptides of the highest affinity.

4. The fact that a correlation between predicted MHC-II peptides and carrier efficacy could be established implied that glycoconjugates were efficiently internalized and processed. These steps were not limited and only the MHC-II association was monitored.

To test these hypotheses, the protein spr96/2021 was selected as a carrier model. The spr96/2021 exhibited stronger binder affinity predicted peptides compared to the CRM_{197} protein (**Figure 4**). In fact CRM_{197} had SMM-align affinity IC50 of 539 nM with peptides not containing lysine residues. Whereas spr96/2021 showed a SMM-align affinity IC50 of 373 nM and 399 nM for the two peptides with the strongest affinity, which contained three lysine residues. Regarding the efficacy of the proteins as carrier, the conjugate spr96/2021 induced an anti-Lam antibody titer significantly higher than the one induced by CRM_{197}–Lam (**Figure 5**). Spr96/2021 is a recombinant fusion of the two pneumococcal proteins: spr0096 (LysM domain secreted protein) and spr2021 (secreted PcsB protein)[69,70,71]. To address the question whether one of the two moieties could be responsible for this strong carrier effect, Lam conjugates with spr0096 and spr2021 respectively, were tested in mice. No response was induced by the spr0096 conjugate and very low by the spr2021 one, indicating that some feature derived from the fusion of the two proteins was critical for the carrier activity. The planed acitivities were:

▪ Include the fusion protein spr2021/96

▪ Empirically determine if the predicted MHC-II peptides bind the MHC-II complex

▪ Evaluate the immunogenicity of the carrier proteins

▪ Evaluate a relationship between carrier protein and pattern of glycosylation.

▪ Generation of mutant constructs based on the evaluation of different pattern of glycosylation in the protein carriers.

4.2 Evaluation of the *S. pneumoniae* spr96/2021, 2021, 2021/96 proteins immunogenicity in mice

To evaluate the ability of the *S. pneumoniae* carriers spr96/2021, 2021, 2021/96 at generating an immune response, the proteins (5 µg) were used to immunize mice. The study was performed with and without the presence of alum adjuvant and CRM_{197} was used as positive control. The anti-protein antibodies IgG titers were evaluated after the third immunization. All the *S. pneumoniae* proteins tested showed a significant higher IgG titers than the one induced by the CRM_{197} (as shown in **Figure 6**). The protein spr2021 alone or in fusion with the spr0096 as spr96/2021 or 2021/96 were highly immunogenic in mice showing a high IgG titer (GMT value) without the use of adjuvant, compared to the IgG titer for the CRM_{197}. The GMT value obtained for the spr96/2021, 2021/96 and 2021 were respectively 5290, 8947, 2706, while the one obtained for the CRM_{197} was 31. The three forms of the proteins spr2021 were able to produce a similar immunological response, indicating that without any modification of the lysine residues, the three proteins were highly immunogenic. We also attempted to correlate the efficacy in the production of IgG to the affinities to the predicted MHC-II peptides (370 nM for spr2021 and 548 nM for CRM_{197}, **Figure 4**), although it could not be excluded that other factors such as the binding, the internalization and the processing could affect the immunogenicity of the proteins. Indeed, alum may exert its role as a delivery system to immune cells directly being engulfed by immune cells within the draining lymph nodes[72], supporting the internalization of the antigen protein. In fact the CRM_{197} protein in presence of Aluminium phosphate had a significant increase in its immunogenicity. It had a GMT IgG titer value of 10773 in presence of Alum while a GMT of only 31 when administered without it. It could be speculated that even if the CRM_{197} had high immunogenic predicted peptide, its ability to generate an immune response could be affected by the internalization and processing. Instead the proteins spr2021, spr96/2021 and 2021/96 in presence of alum adjuvant did not show a substantial increase in immunogenicity (**Figure 6**). The following step of the study was to identify the lysine residues involved in the conjugation, for doing this the spr2021,

the fusions spr96/2021 and 2021/96 were conjugated with the Laminarin polysaccharide.
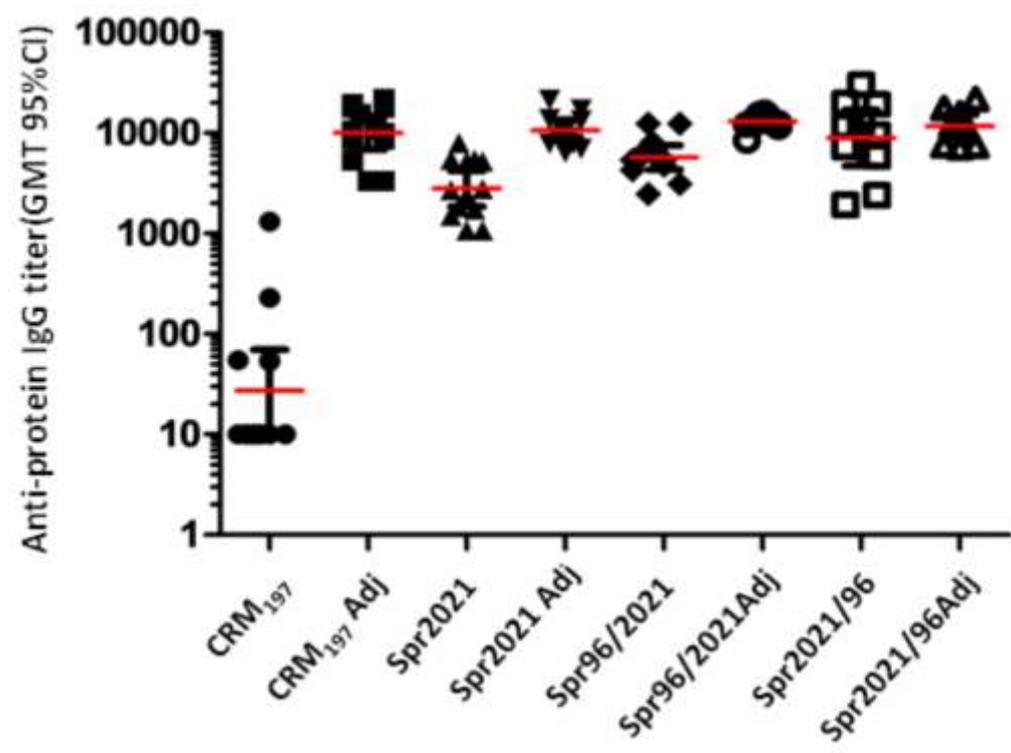

Figure 6 *- Anti-protein IgG response induced by spr2021, spr96/2021 or spr2021/96 without and in presence of adjuvant, compared with the CRM$_{197}$ used as positive control.*
Anti-carrier IgG represented by each dot indicates a single mouse ELISA titer; the horizontal bar(red) represents the GMT of the group, instead the vertical bar shows the statistical 95% CI.

4.3 Conjugates preparation of laminarin glycoconjugates spr96/2021, spr2021 and spr2021/96

The Laminarin glycoconjugates were generated by chemical derivatization at the reducing end of the polysaccharide with an adipic acid linker (succinimido diester of adipic acid) as described in the materials and methods section. The activated polysaccharide was reacted with the different protein carriers using the same activated polysaccharide/protein molar ratio.

Figure 7 reports a schematic description of the chemical reactions to produce the laminarin conjugates.

Figure 7: *Schematic reaction of Laminarin glycoconjugates formation*
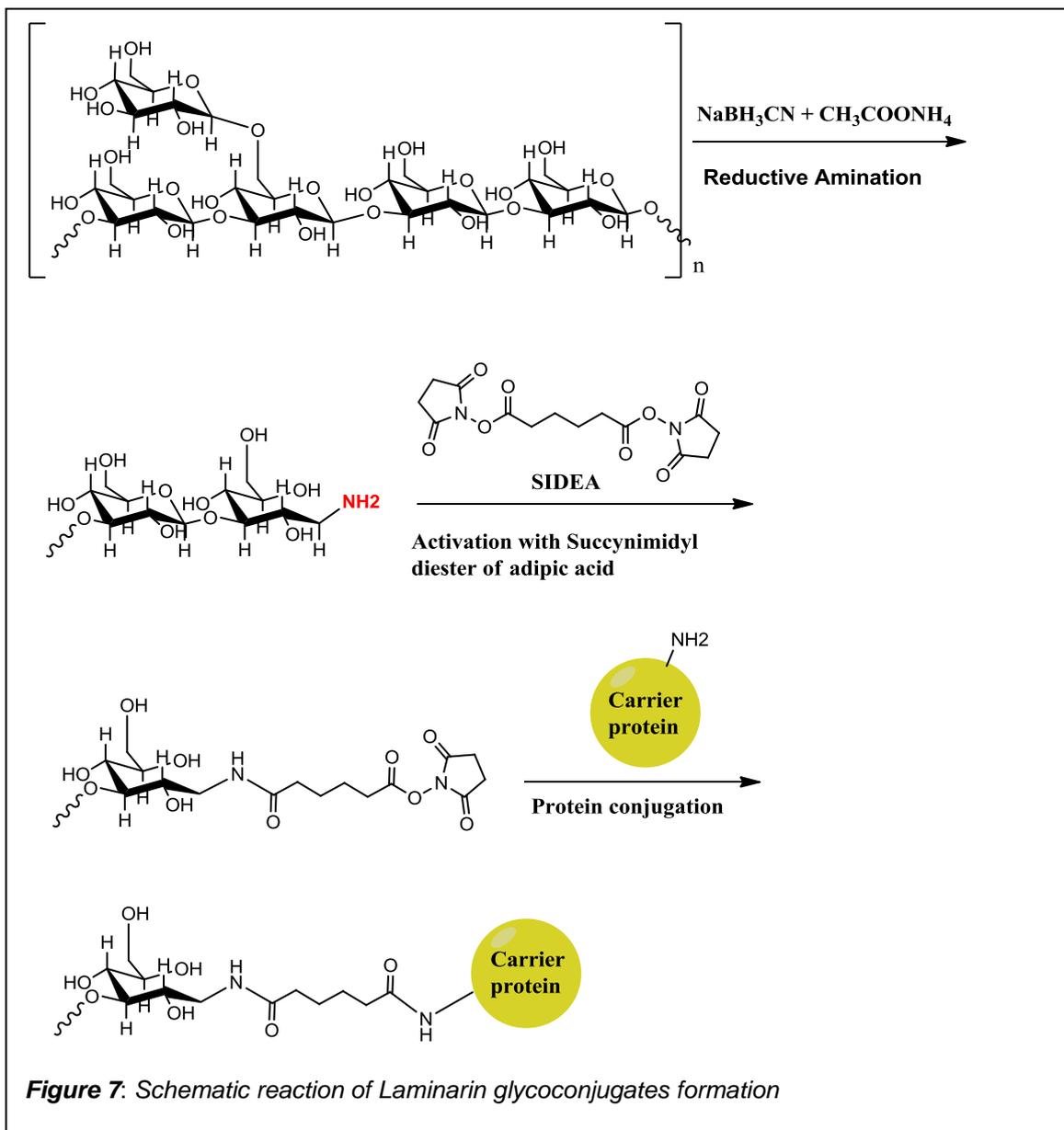
The conjugates spr96/2021, 2021, and 2021/96 were analyzed by SDS-PAGE electrophoresis to control the efficacy of the reaction (**Figure 8**). **Figure 8** overleaf shows the classical smear of glycoconjugates due to variable sugar chains bound to the protein, and the band corresponding to the starting recombinant protein is not anymore visible to indicate a complete conjugation of the carrier.

As described in chapter 2 – Material and methods, the glycoconjugates were purified from the unreacted polysaccharide and the glycosylation degree was calculated in term of total saccharide and free saccharide (not conjugated). The glycosylation degree for protein was between 0.22 and 0.36 in term of saccharide/protein ratio (w/w); the content in free saccharide was quite

homogenous between 9.2 and 12.1 % (**Table 2**), in agreement with the previously published data[49].
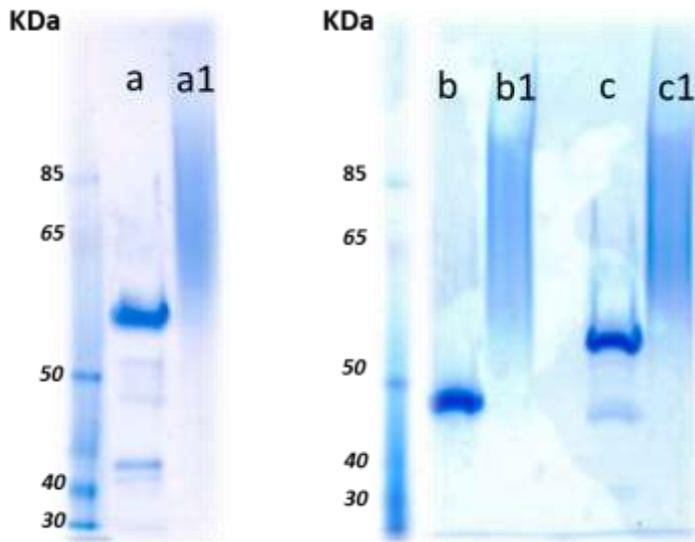

***Figure 8**: SDS-page of spr96/2021, 2021, and 2021/96 proteins (respectively **a, b, c**) and their relative laminarin glycoconjugates(**a1, b1, c1**).*

Laminarin conjugates	Total Saccharide µg/mL	Saccharide/protein (w/w)	Free saccharide %
Spr2021	**550.9**	*0.36*	*11.1*
Spr96/2021	**412.9**	*0.22*	*12.1*
Spr2021/96	**541.5**	*0.29*	*9.2*

Table 2: *Free saccharide and saccharide/protein (w/w) ratio determination in Laminarin glycoconjugates.*

4.4 Immune response evaluation of the *S.pneumoniae* spr96/2021, 2021/96, 2021 Laminarin conjugates

The *S.pneumoniae* laminarin conjugates generated were tested in mice and the anti laminarin response was evaluated monitoring the antilam IgG titer by ELISA assay. The CRM$_{197}$-Lam conjugate was used as control. The study was performed without adjuvant with an immunization dose of 5 µg in term of saccharide. The anti-laminarin antibodies IgG titers were evaluated after the third immunization. Regarding the efficacy of the conjugate spr96/2021 the anti-laminarin IgG titer (**Figure 9**) was significantly higher than the one induced by

CRM$_{197}$-Lam as previously described (p value =0.0023). It was also confirmed that the spr2021-Lam had a lower immunogenicity when compared to the fused conjugate spr96/2021-Lam (p value=0.007). Interestingly, the fused conjugate spr2021/96-Lam showed a lower immune response compared to the spr96-2021-Lam (p value=0.0011). These results accentuated the possibility that some features derived from the fusions were critical for the glycoconjugates activity. A different rearrangement in the two fused protein structures and the possible different glycosylation pattern in the conjugates may drive to a different immune response.
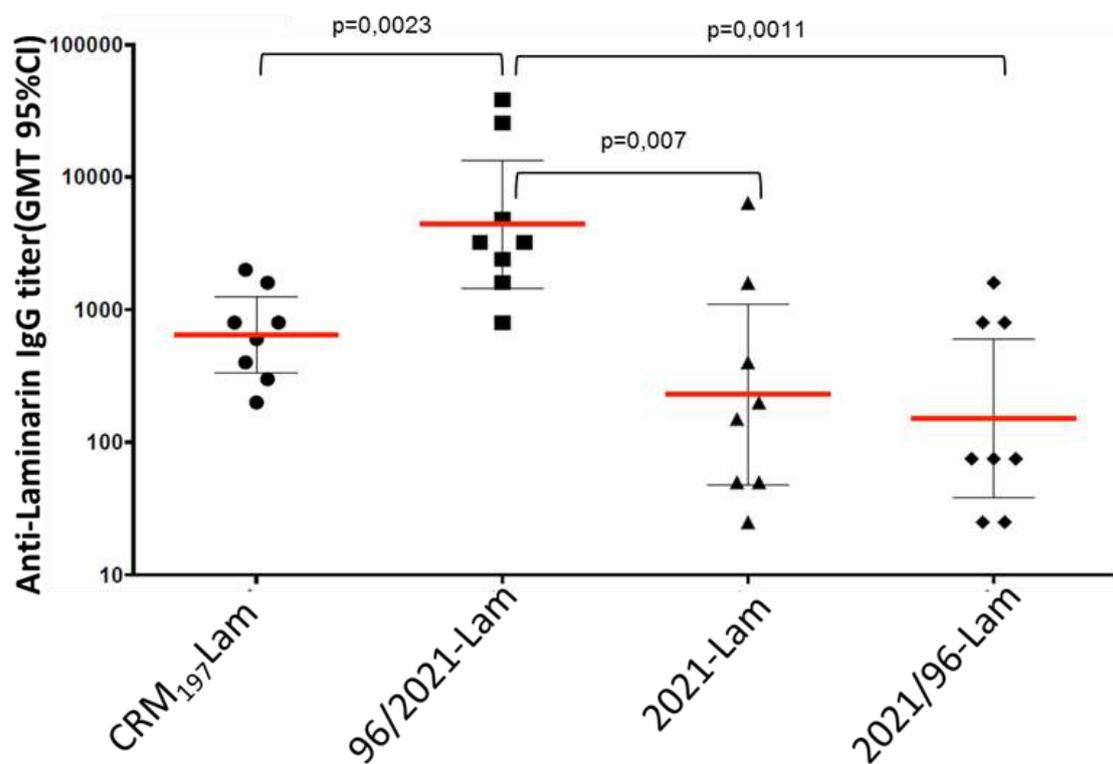

Figure 9: Anti-polysaccharide IgG antibody titer induced by Laminarn conjugates in *S. pneumoniae* conjugates, CRM$_{197}$–Lam is used as benchmark. Each spot indicates a single mouse ELISA titer, the horizontal red bar refers to the GMT of the group and the vertical bar shows the 95% CI. Non parametric Mann–Whitney test has been used for p value calculation.

4.5 Determination of glycosylation factors in *S. pneumoniae* constructs

Because differences in term of immunogenicity against the polysaccharide antigen were identified for the three different carrier constructs spr96/2021, spr2021 and spr2021/96, we tried to identify features in the protein structures and glycosylation in the conjugates that could lead to a different immune response. Two approaches were used:

a. Determination of the glycosylation pattern.
b. Determination of differences in the dynamics of the proteins leading to a different glycosylation.

4.5.1 Pattern of glycosylation for the spr2021, spr96/2021 and spr2021/96 conjugates

The different glycoconjugates were analyzed to identify possible differences in the glycans distribution especially regarding the lysine residues present in the predicted MHC-II peptide binder. Chemical deglycosylation method and Mass Spectrometry techniques were combined to identify the glycosylation sites. We took advantages of the differences between the glycosidic bonds and the covalent linking of the glycan to the protein. To be more specific, the amide bond of the protein-SIDEA-Sugar and the glycosidic bond of the rest of the oligosaccharide are very different from the chemistry point of view (**Figure10**). After the chemical deglycosylation process, only the SIDEA-monosaccharide was retained on the glycosylated lysines allowing to distinguish between glycosylated/unglycosylated lysine residues by mass spectrometry.
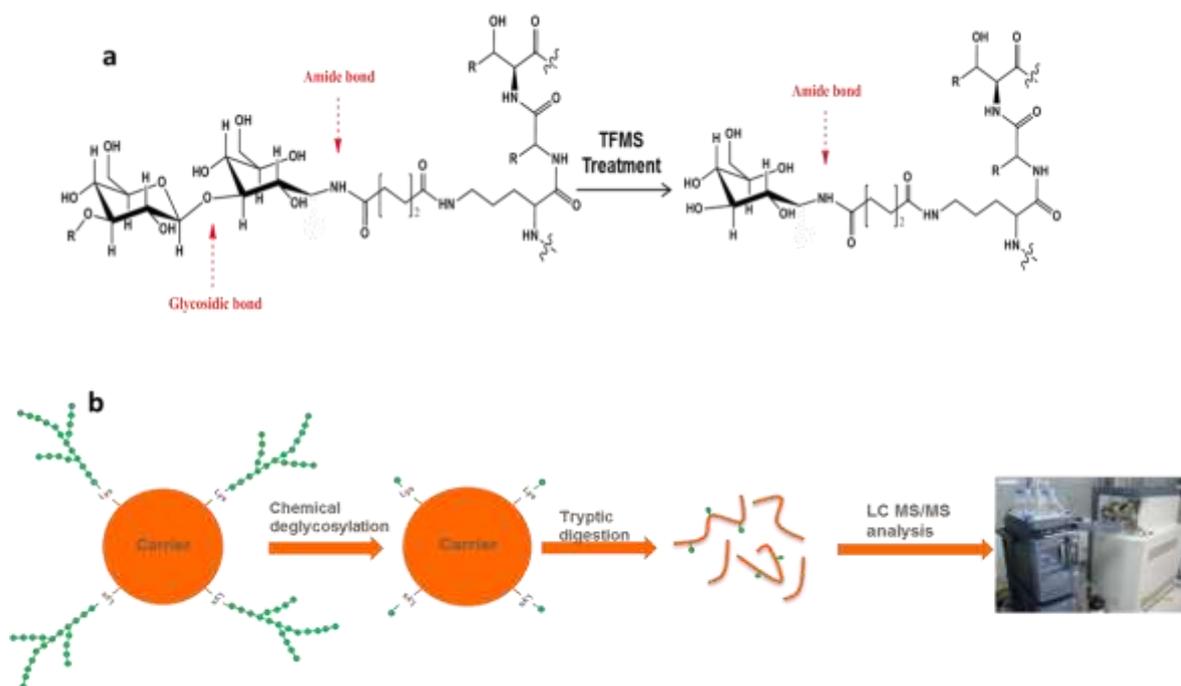
Figure 10 - *(a) Different chemical effect of TFMS treatment on glycosidic bond and N-linkage of reduced glycan bound to SIDEA. (b) Principal of the glycosylation method coupled with MS analysis.*

4.5.2 Chemical deglycosylation

The chemical deglycosylation was performed by chemical solvolysis as described in section 3. The deglycosylation procedure removed all sugars that contained covalently bound sugar chains. Amide bonds were stable to TFMS and so the integrity of the protein was maintained. After the deglycosylation of the conjugates, sugars were removed and only the SIDEA-monosaccharide linked to the side chains of lysine residues remained. This adduct served as label for the identification of conjugated lysines .

The integrity of the proteins after the deglycosyation were checked by SDS-PAGE (**Figure 11**). The characteristic smear of the glycoconjugates was not anymore present after the chemical reaction and to assess that the carriers did not show evident degradations due to the harsh deglycosylation reaction, the starting recombinant proteins were also loaded on the SDS-PAGE gel. No difference in the apparent molecular weight of the carriers and the deglycosylated carriers were observed.
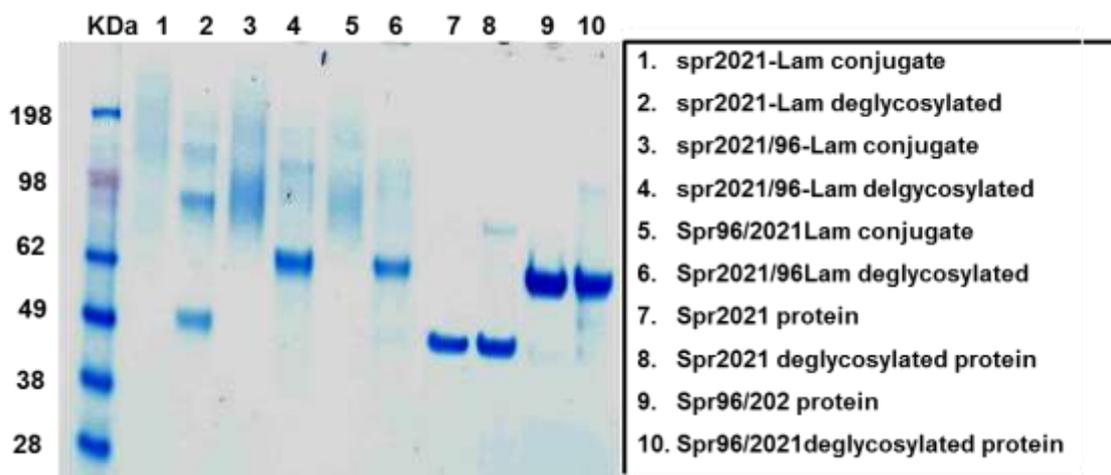
Figure 11 - *Deglycosylated fractions analyzed by SDS-PAGE the single band characteristic of the spr96/2021, spr2021 and 2021/96 proteins is shown after the treatment with TFMS on the glycoconjugates confirming the removal of the sugar.*

4.5.3 Glycosylation pattern identified by mass spectrometry

Almost all currently marketed glycoconjugate vaccines contains large and heterogeneous glycans, conjugated by different methods, to the carrier proteins (**Section 1.3**). Many parameters such as length of the glycan, incorporation into the protein, conjugation linker and the distance from the protein can influence the immunogenicity of the vaccine conjugate[73,74]. Although the correct arrangements of the polysaccharide antigens make it available to interact with the APCs, the attachment site of the carbohydrates to the carrier protein can also be crucial for the induction of a T-cell response[75,44]. Most of the glycoconjugate in clinical use are prepared by covalent linkage of the carbohydrate to the side chain of lysine. This conjugation generally is considered highly random, and hard to control, also inducing batch-to-batch variability in the immunological properties of the vaccine. Analysis and localization of the glycosylation sites on conjugate vaccines are very important for the understanding of their mechanism of action[76]. The identification of the lysine residues involved in conjugation has only recently started to be explored[76,77]. Different groups have studied the possible conjugation on the common carrier protein CRM_{197}, demonstrating that the conjugation reaction take place with preferences for particular sites lysine residues. *Crotti et al.* and more recently *Möginger et al.*, identified that not all the lysine residues are equally reactive and prone to the conjugation. It has been demonstrated that the steric

accessibility, the local amino acid environment and the protein secondary structure are the most relevant parameters; influencing the conjugation reaction under not-denaturing conditions.

Considering all these aspects, we identified which were the most abundant glycosylation on the three different conjugates. The approach used in this study had the advantage to easily determine the glycosylation sites obtaining intact proteins tagged on the conjugated lysine residues, minimizing the problem related to hindrance of the polysaccharide that can reduce the efficacy of the trypsin proteolitic enzyme used for the digestion.

The deglycosylated carriers were analyzed by LC-MS/MS analysis after tryptic digestion performed as described in section 3. The peptides and glycopeptides identification was carried out from the generated peak list using Mascot engine software or the peaksoftware adding the modifications of SIDEA-Glc(+291.13180 Da) corresponding to the modification due to the presence on the side chain of lysine of SIDEA- glucose (SIDEA-Glc) leftover from the deglycosylation. To assign glycosylation side localization, we only considered the peptides in which, the modified peptide, was shown in MS/MS spectra as fragment ion displaying the b and/or y fragments of the modified lysines (**Figure 12**).
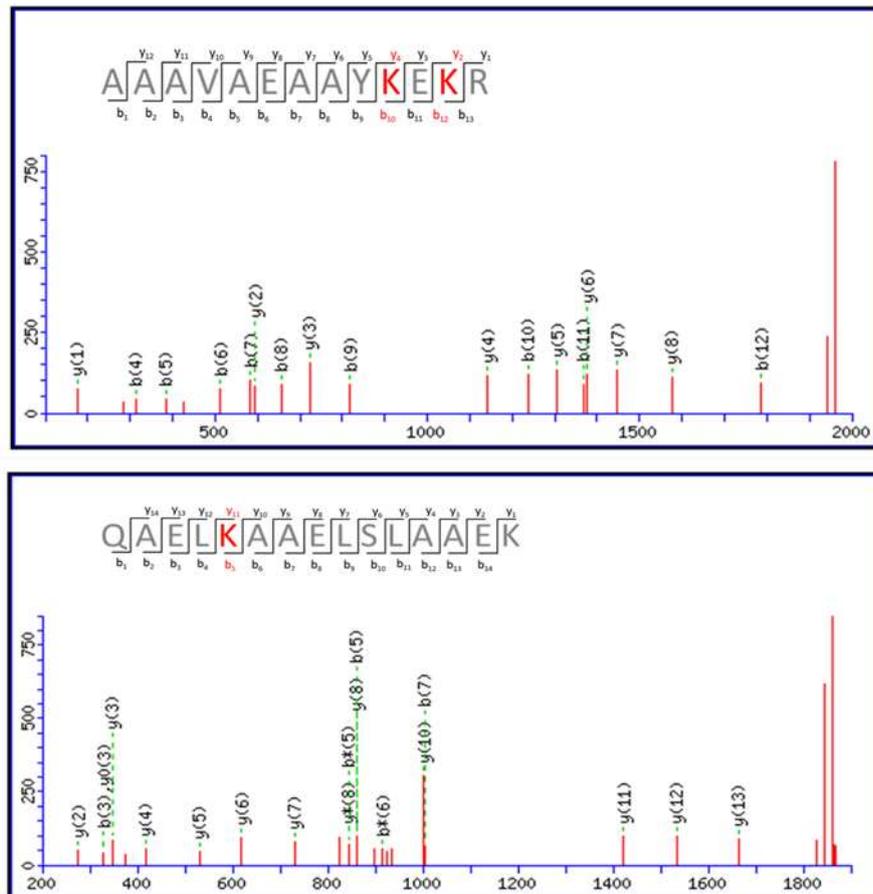
Figure 12 - *example of tandem mass spectrum of the identified glycopeptides (R.QAELK#AAELSLAAEK.A; R.AAAVAEAAYK#EK#R.A) from protein spr2021 and spr2021/96. In red on the sequences are highlighted the b and y transitions carrying the modified lysines.*

Besides the results obtained using the MS/MS ion search, for increasing the identification of the unmatched spectra and monitored potential uncharacterized modifications on the modified lysine residues, the *Mascot Error Tolerant Search* was used, or the *de novo* sequencing function in peaks software was used. These tools allowed increasing the coverage of the proteins and identifying other modifications on the lysine residues. In particular, three other modifications were highly recurrent: +128.1075 Da associated to the modification of the lysine residues with the SIDEA linker, +218.1671 Da and +100.0160 Da. It was not possible to define these unknown modifications. These modifications probably derived from a rearrangement of the glycan structure after the treatment of solvolysis.

The **Table 3** below lists some of the glycosylation that were identified in this work for the *S. pneumoniae* protein carriers. It is interesting to notice that the peptides predicted to have the highest affinity for the MHC-II complex also presented a different pattern of glycosylation. In particular, we identified that the peptides AAAVAEAAYKEKR and QAELKAAELSLAAEK were found mainly glycosylated in spr2021 and spr2021/96, while these modified peptides were not identified in spr96/2021 according to the stringent conditions we applied to the identification of modified peptides (**Figure 13**).

Conjugate	Sites	Ion score	Peptide sequence[#]
Spr2021-Lam	245	93.4	R.AAAVAEAAYK**#**EKR.A
	245, 247	79	R.AAAVAEAAYK**#**EK**#**R.A
	203	102.1	K.QAELK**#**AAELSLAAEK.A
Spr2021/96-Lam	213, 215	80.8	R.AAAVAEAAYK**#**EK**#**R.A
	171	66.7	K.QAELK**#**AAELSLAAEK.A

Table 3 - Example of peptides presenting lysine residues glycosylated (#) in spr2021-Lam, and spr2021/96-Lam.
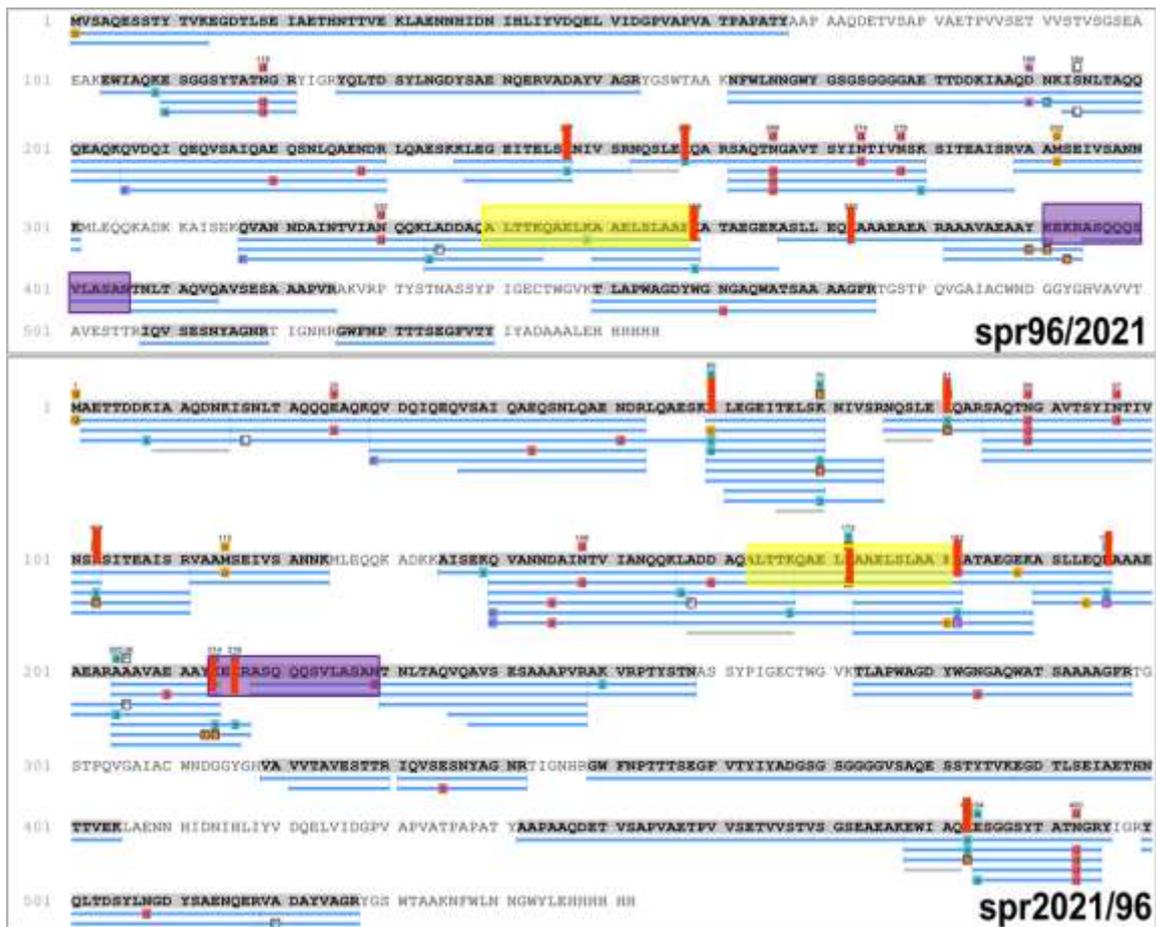
Figure 13 - Glycosylation pattern for spr96/2021 and 2021/96: In red are marked the glycosylated lysine in the two fusion proteins. In violet and yellow are highlighted the peptides with highest affinity for the predicted MHC-II peptides.

4.5.4 HDX-MS analysis on spr2021, spr96/2021 and spr2021/96 carrier proteins for structural differences

After the identification of a different pattern of glycosylation for the three different constructs we tried to better characterize the factors that could have lead to a different glycosylation. Due to differences in the steric interactions among the three constructs, the lysine residues could not be equally accessible for conjugation resulting in a different pattern of glycosylation. The spr2021 alone or in the two fusions were analyzed by Hydrogen deuterium exchange (HDX) technique with on-line pepsin proteolysis digestion, to evaluate the possible differences in conformation and dynamics structure among the proteins.

The rate of backbone amide hydrogen exchange in solution is directly dependent on the structure and dynamics of the protein. Therefore regions not exposed will exchange more slowly than regions fully exposed to the solvent.

The spr96/2021 was used for the pepsin peptide map, to generate a full map covering the full length of the protein in order to monitor the deuterium uptake in the entire sequences of spr2021, 96/2021, 2021/96. The peptides were assigned to the sequence according to their fragmentation in MS/MS. n° of 65 peptides were identified in at least 4 of 5 repeated analysis and represented a coverage of 85%. The HDX-MS analysis was performed at 6 deuteration points (30 s, 5min, 10min, 30min, 1 h and 24 h). When the experiment was performed at room temperature, no significant differences in incorporation were verified for the three *S. pneumoniae* proteins (data not shown). A difference in the incorporation was instead identified at 0 °C when the dynamics of the protein is slowed down. **Figure 14** illustrates the difference index of incorporation between the spr96/2021 and the spr2021/96. The data are plotted as the difference in relative uptake of each peptide (dots) in the selected state (lines). The closer the dots and lines are to the zero, the lower is the difference (in term of incorporation) between the proteins compared. 1 Da of difference in the incorporation was used as threshold to define a difference considered to be significant. Differences in the dynamics of the proteins were pointed out on the spr0096 region comparing the fusion proteins spr96/2021 and 2021/96 with a maximum difference of incorporation of 5 Da when the analysis was performed a 0 °C. The peptides in the C terminus path of the spr0096 region in the fusion of spr96/2021 present a faster kinetics of labeling due to a more accessible structure compared to the spr2021/96; however this difference was displayed only when the analysis was performed at 0°C, when the dynamic of the proteins was reduced.
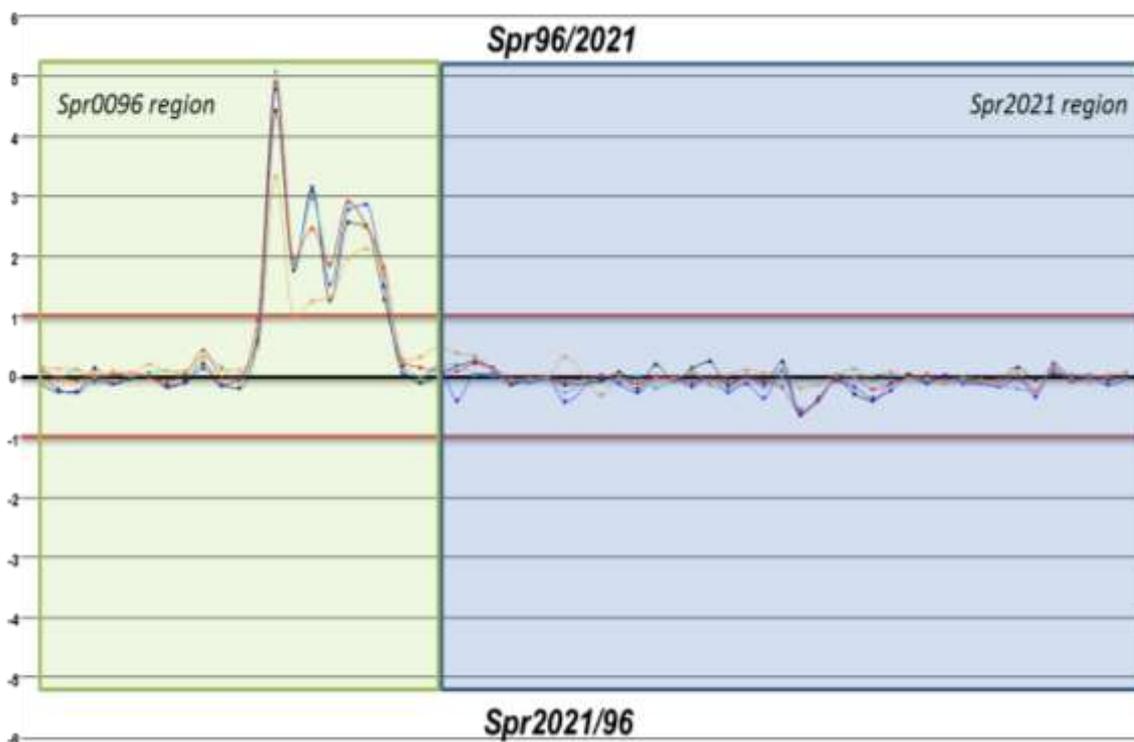
Figure 14 - *Difference index of incorporation for the protein spr96/2021 and spr2021/96 each dot represents a peptide and each line represents a different state of deuteration. In the green box is represented the difference in the incorporation for the spr0096 in the two fusion proteins. In the blue box the difference in the incorporation for the spr2021.*

Focusing our attention on the spr2021region of the proteins, that was the one carrying the predicted MHC-II peptides found differently glycosylated in the three constructs, we observed that the proteins appear at highly dynamics, displaying a high level of deuterium incorporation. **Figure 15 panel a** shows the deuterium incorporation for the protein spr2021. Briefly, each line represents a different time point and each dot represents a peptide. The lower is the uptake of deuterium, the closer are the line and dot to the zero. A lower incorporation is related to a more structured region, in which the peptide is not accessible to the solvent and viceversa.

We distinguished among high, medium and low dynamics of the peptides considering a high incorporation≥40%, 15%≤medium≤40%, low≤15% respectively in term of relative uptake. When the experiment was performed at room temperature, almost all the spr2021 protein showed a relative incorporation higher than 40% (data not reported). The analysis performed a 0 °C, allowed to better

characterize the protein structure, and in the identification of regions with a lower incorporation. At this temperature, the uptake of deuterium was slowed down in the regions more prone to interact. This aspect was in agreement with the crystallography structure available for the spr2021 protein[78].

In **Figure 15 panel b** the crystal structure of the spr2021 is shown and compared to the relative uptake of deuterium for each region identified. The region with high, medium, and low incorporation assigned in the relative uptake **Figure 15 in panel a** (indicated as A, B, C) are highlighted to the crystallography structure in **panel** b. In blue and green on the crystallography structure are showed the regions with an incorporation of deuterium ≤15% (blue) and between 15 and 40% (in green) are reported. The interhelix interactions in the coil-coil domain, mainly mediated by hydrophobic interactions and hydrogen bonds, leads to a lower incorporation.

The C-terminus domain presented β-sheet with a robust network of hydrogen-bond interactions also leading to a more structured region in agreement with the lower deuterium incorporation. The high dynamic of the proteins did not allow identifying differences, in term of protein structure that could help the understanding of the different pattern of glycosylation among the three constructs. As described, the spr2021 showed a similar incorporation among the three constructs (**Figure 14**).The proteins were highly dynamics and it was not possible to define a correlation between differences in the pattern of glycosylation and the deuterium exchange. However, it could be observed, comparing the pattern of glycosylation and the structure analysis performed by HDX that, in all the three constructs the lysines involved in the conjugation were presented on the more dynamic regions of the proteins. Lysines presented in the interhelix interaction in the coil-coil domains, and the C terminus region highlighted in blue and green on the crystallography structure (**Figure 15 panel b**), were found not glycosylated. In agreement with the possibility that lysines more exposed to the solvent were also more likely glycosylated.
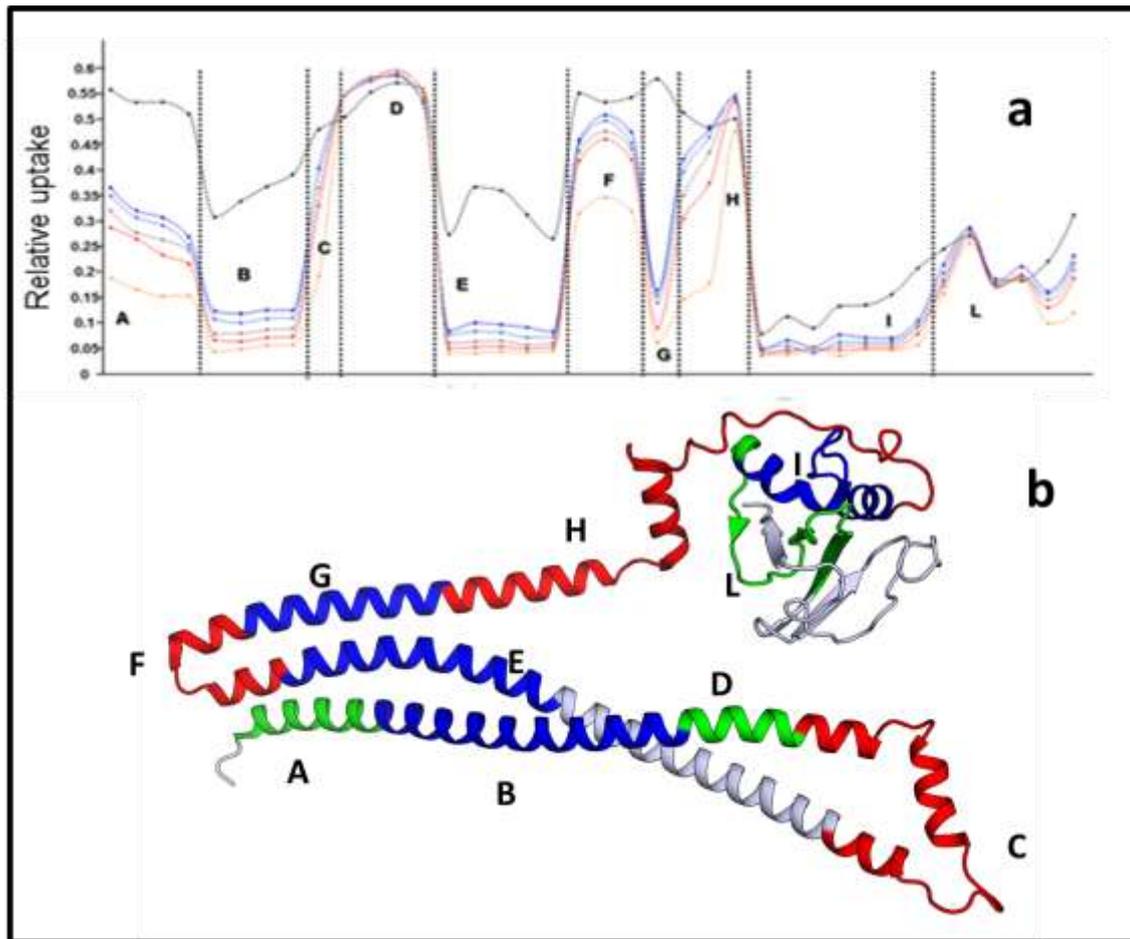
*Figure 15 - Dynamics of the spr2021 at 0°C. In **panel a** is shown the relative uptake for the protein spr2021, each dot represents a peptide and each line represents a different state of deuteration. Incorporation: high ≥40%, 15%≤medium≤40%, low≤15%. Each region with high , medium, and low incorporation assigned in the relative uptake figure in **panel a** (indicated as A, B, C) are assigned to the crystallography structure in **panel b** highlighted in red (high incorporation, green(medium), and blue(low).*

4.6 Generation of proteins mutant and laminarin glycoconjugates

Differences in the pattern of glycosylation were identified for the three carrier proteins spr96/2021, 2021 and 2021/96. The two peptides predicted to have the highest affinity for the MHC-II complex presented different pattern of glycosylation (**Figure 13**). The peptides AAAVAEAAYKEKR and QAELKAAELSLAAEK were found mainly glycosylated in spr2021 and spr2021/96, the two constructs with the lower immunogenicity against the polysaccharide antigen. These modified lysines

were not identified in spr96/2021, the carrier protein that leads to a higher immunogenicity against laminarin. Based on correlation that predicted MHC-II peptides were mainly glycosylated on the proteins that raise a lower anti-laminarin IgG titer, we speculated that the glycosylation on the presented MHC-II peptide could affect the MHC-II CD4$^+$T cells interaction and so the immune memory against the polysaccharide antigen. For evaluating the rule played by the glycosylation on the MHC-II predicted peptides two strategies were used. The first strategy considered:

I.   Generating a series of mutants for the spr2021 and 2021/96 protein in which the lysines residues on the predicted MHC-II peptides AAAVAEAAYKEKR and QAELKAAELSLAAEK were substituted with an Arginine residue. In this way it can be possible to prevent the glycosylation on the specific site and evaluate the rule played by these peptides.

II.   Insertion of the MHC-II predicted peptide from the CRM$_{197}$ protein IALSSLMVAQAIPLV as linker between the spr2021 and spr0096 proteins.

As a second strategy we selected another carrier protein, the GNA2091-fHbp previously analyzed in term of immune response against laminarin antigen. This glycoconjugate displayed a low immune response compared to the CRM$_{197}$ conjugates used as control. The GNA2091-fHbp showed in is structure low affinity MHC-II predicted peptides (IC50 772 nM) also containing lysine residues (see **Figure 4**). We produced mutants of this protein in which the linker between the GNA2091 and fHbp proteins were substituted with the predicted CRM MHC-II peptide IALSSLMVAQAIPLV and with the predicted spr2021 MHC-II peptides KRASQQQSVLASANT. In this last case, the lysine residues presented on the peptides were also substituted with Arginine residues (i.e. RRASQQQSVLASANT) avoiding the possible conjugation on the specific lysine residue.

### 4.6.1 Proteins production and characterization

To perform the conjugation, high pure amount of proteins was needed. Recombinant constructs of these proteins were produced and purified as described in materials and methods. The purified proteins showed a high degree of purity, which was verified by analytical size-exclusion high-performance liquid-

chromatography (SEC-UPLC); this was around 90-97% among all the proteins produced. The SEC-UPLC chromatography of the purified spr96/2021 and 2021 are shown in **Figure16** as an example of the analysis. To verify the molecular weight of the proteins entire mass were performed verifying the correct molecular weight. The entire mass, assessed by MS, are shown in **Table 4**; some of the proteins showed N-terminal methionine excision (NME)[79] with a loss of 131.1986 Da from the theoretical molecular weight. The entire masses obtained were in agreement with primary sequences of the proteins produced.

The content in lipopolysaccharide was evaluated using the endotoxin testing system (Charles River). When the content of endotoxins was higher than 0.05 EU/mL, the lipopolysaccharides were removed as described in Chapter 3.

Protein	Expected Mass (Da)	Obtained Mass (Da)
spr2021	42743.16	42745.77±2.59
Spr2021_Lys204Arg	42771.18	42772.08±0.65
Spr2021_Lys248Arg	42771.18	42771.70±0.35
Spr2021_Lys204Arg_Lys246Arg_Lys 248Arg	42827.20	42827.53±0.32
Spr2021/96	58854.37	*58727.85±1.76
Spr2021/96_Lys172Arg	58882.38	*58750.22±3.38
Spr2021/96_Lys216Arg	58882.38	*58753.96±1.69
Spr2021/96_Lys172Arg_Lys216Arg	58910.39	*58778.29±3.83
Spr2021/96_Lys172Arg_Lys214Arg_ Lys216Arg	58938.41	*58805.56±5.79
Spr2021-(CRM$_{197}$)-96	59959.90	*59827.77±5.54
GNA2091-fHbp	46282.00	*46149.89±2.75
GNA2091-(2021)-fHbp	47531.66	*47401.07±2.42
GNA2091-(CRM$_{197}$)-fHbp	47622.50	*47491.13±2.97

Table 4 - *Entire mass of the S.pneumoniae and N. meningitidis proteins produced, the star (*) is indicating the loss of methionin.*
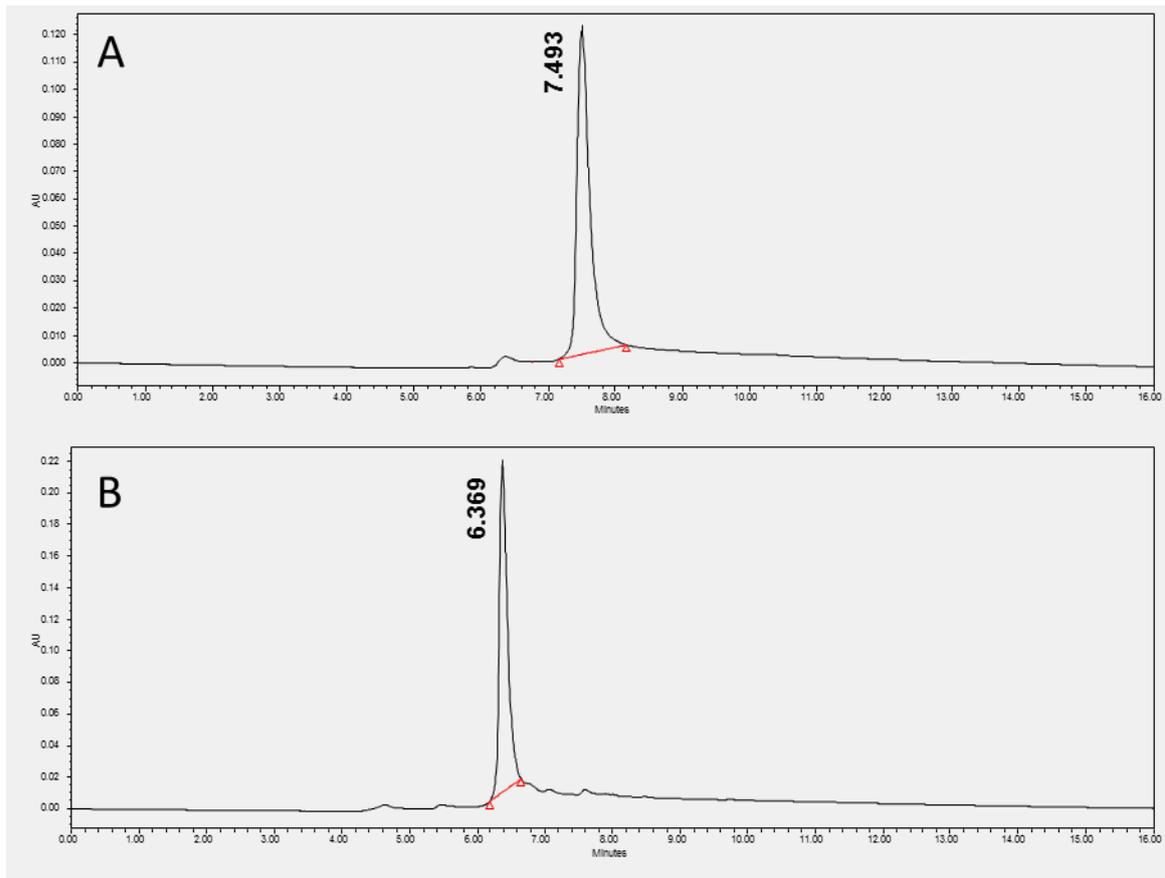
Figure 16 - *SEC-UPLC profile of purified spr2021(**A**), 2021/96_Lys172Arg(**B**), the absorbance (AU) was acquired at 280nm wavelength, the elution time for each peak is indicated in each panel. The peak % area indicates the % of the peak component among the total 'detectable' peaks.*

To further characterize possible disorders in the protein structures correlated to the mutations on the native sequences of the *Streptococcus pneumoniae* spr2021-96 and spr2021, and *Neisseria meningitis* GNA20191-fHbp, the thermal stability of the generated proteins was analyzed by differential scanning calorimetry (DSC). The same thermal stabilities for the spr2021, spr2021/96 and their relative mutants were verified. Two peaks corresponding to melting temperature transitions (T_m) at 29.96 and 52.68 ºC were showed for the spr2021/96 and its relative mutants (**Figure 17**). The same analysis was performed for the spr2021 and its mutants; a T_m of 29.96 and 55.02 ºC was showed (**Figure 18**). All the point mutation did not significantly affect the melting temperature of the proteins. Different thermal stability were instead showed for

the spr2021-(CRM)-96 mutant in which the linker GSGSGGG was substituted with the sequence GIALSSLMVAQAIPLVG corresponding to the CRM_{197} MHC-II predicted peptide. In this case the T_{m1} was maintained at 30 °C but the T_{m2} was instead lower from 52.68°C to the initial construct to 47.68 °C (**Figure 19**). The modification of the linker between the two proteins srp2021 and spr96 led to a change in the structure conformation. Same thermal stability was verified for the GNA2091-fHbp and its relative mutant. Two different peaks were observed at T_{m1} 66 °C and T_{m2} 83 °C (**Figure 20**). In both the two constructs the change of the linker did not affect the melting temperature of the GNA2091-fHbp protein.
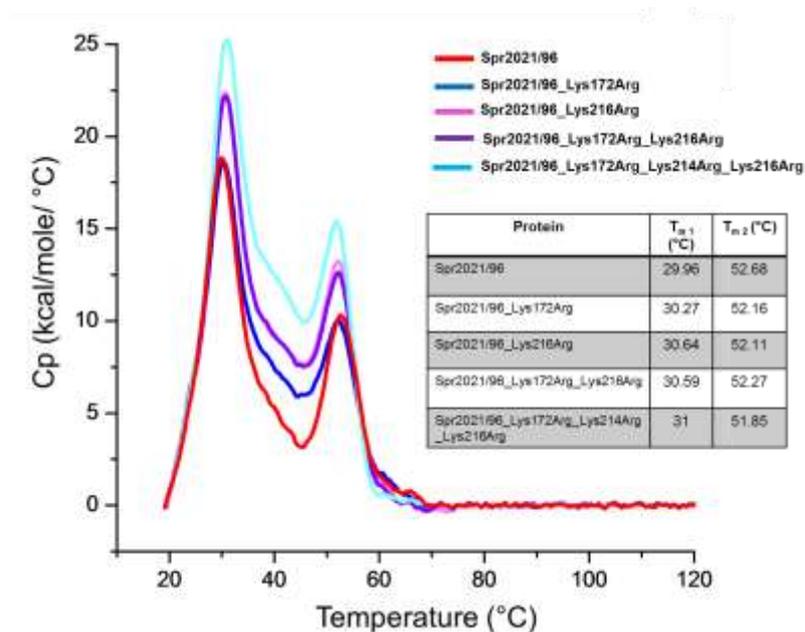

Figure 17 - *DSC profiles of spr2021/96 protein and mutants with the relative T_m associated to each peak. Samples are indicated as shown in the legend.*
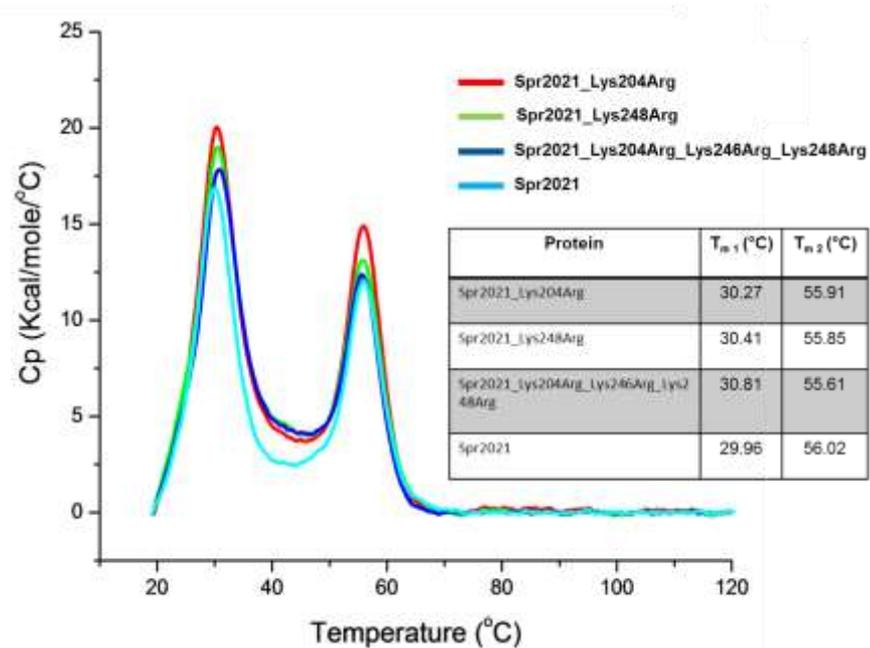
Figure 18 - DSC profiles of spr2021 protein and the respective mutants with the relative T_m associated to each peak. Samples are colored as follow: red spr2021_Lys204Arg, green spr2021_Lys248Arg, blue spr2021_lys204Arg_Lys246Arg_Lys248Arg and light blue for the spr2021

Figure 19 - DSC profiles of spr2021/96 protein and the spr2021-(CRM)-96 mutants with the relative T_m associated to each peak. Samples are indicated as shown in the legend.
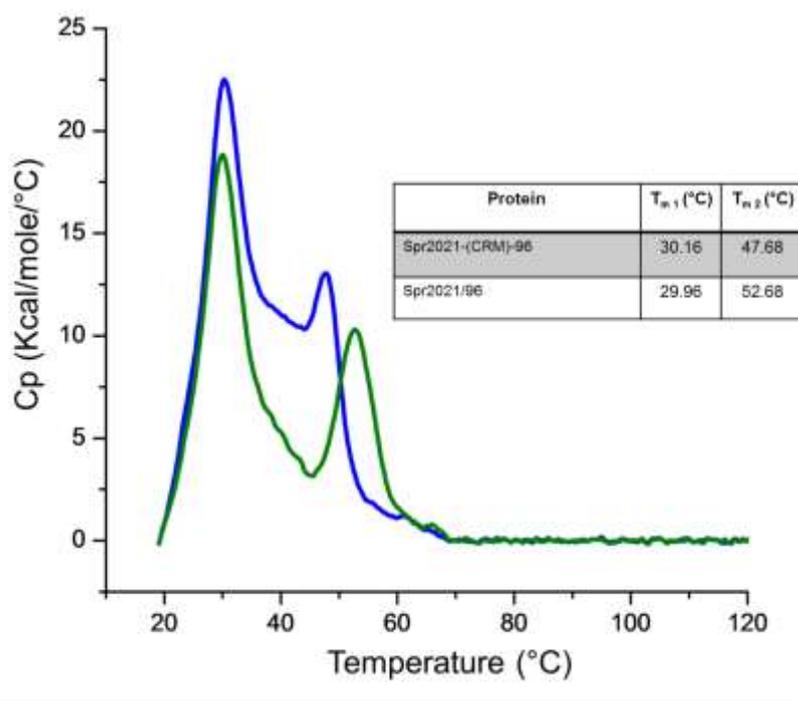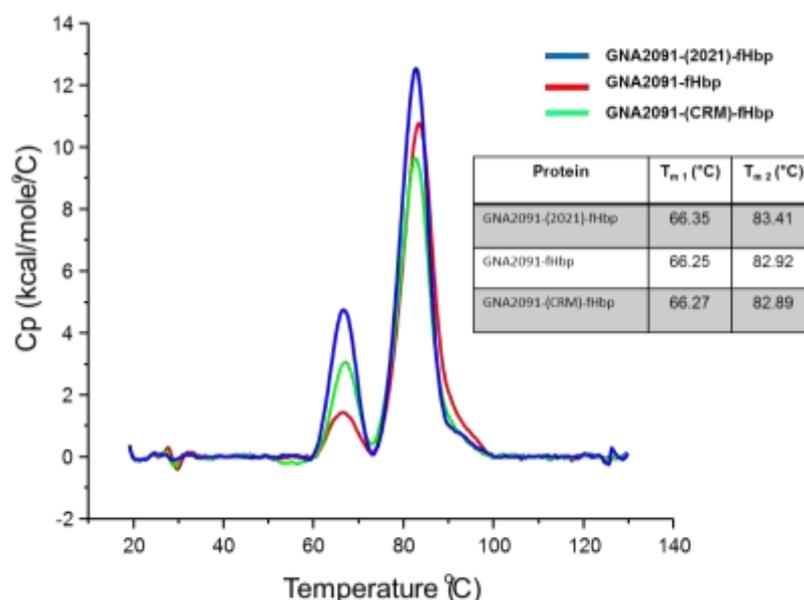
Fig20: *DSC profiles of GNA2091-fHbp protein and its mutants GNA2091-(2021)-fHbp, GNA2091-(CRM)-fHbp with the relative T_m associated to each peak. Samples are indicated as showed in the legend.*

4.6.2 Immune response evaluation of *S. pneumoniae* and *N. meningitidis* mutated proteins in mice

For evaluating the ability of the mutated *S. pneumoniae* and *N. meningitidis* produced to maintain an immune response, the mutants were tested in mice; the spr2021 and spr2021/96 were used as controls from which the immune response should be increased on the mutated proteins. Studies were performed without the use of adjuvants for all the proteins and mutants with an immunization dose of 5 µg in term of total protein. The anti-protein antibodies IgG titers were evaluated after the third immunization (**Figure 21**). All the *S. pneumoniae* mutants tested showed an IgG titer that is comparable to the spr2021 and spr2021/96 proteins (**Figure 21 panel a, b**). The same behavior was showed for the *N. meningitidis* mutants when compared to the control GNA2091-fHbp (**Figure 21 panel c**). Non significant differences were obtained in term of immunogenicity for mutated carriers. The mutations did not significantly affect the immunogenicity of the carrier proteins. It was possible to speculate that the immunogenicity of the modifications made, did not affect the process of presentation and internalization of the new carriers produced.
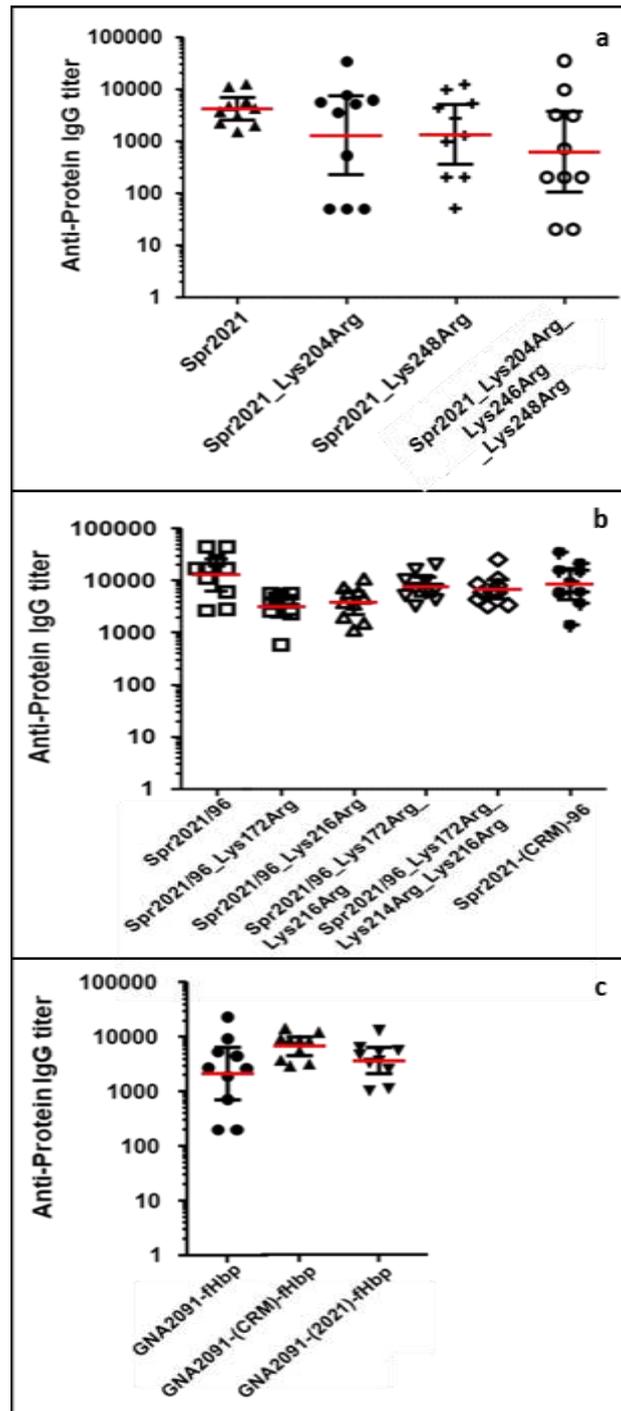
Figure 21 - *Anti-carrier IgG response induced by spr2021, spr2021/96 or GNA2091-fHbp proteins in comparison with the corresponding mutants.*

In panel (a) is shown the anti-carrier IgG response induce by mutated spr2021 without presence of adjuvant in comparison with the positive control spr2021.

In panel(b) is shown the anti-carrier IgG response induce by mutated spr2021/96 without presence of adjuvant in comparison with the spr2021/96 original construct.

In panel(c) is shown the anti-carrier IgG response induce by mutated GNA2091-fHbp without presence of adjuvant in comparison with the GNA2091-fHbp wild type.

4.6.3 Conjugates preparation of laminarin glycoconjugates for spr2021, 2021/96 and GNA-fHbp wild type and mutants

The Laminarin glycoconjugates were obtained as already described in Section 4.3 where **Figure 7** illustrates schematic description of the chemical reactions to obtain the laminarin conjugates. The conjugates were analyzed by SDS-PAGE electrophoresis to control the efficacy of the reaction. The classical smear of glycoconjugates due to variable sugar chains bound to the protein was observed, and the band corresponding to the starting recombinant protein was not any more visible indicating the complete conjugation of the carriers (**Figure 22).** The glycoconjugates were purified from the unreacted polysaccharide as described in Section 3, and the glycosylation degree was calculated in term of total saccharide and free saccharide (not conjugated). The saccharide content among the *S. pneumoniae and N. meningitidis* conjugates to Laminarin polysaccharide is reported in **Table 5.** The glycosylation degree for protein was between 0.19 and 0.7 in term of saccharide/Protein ratio(w/w), the content in free saccharide was quite homogenous waving between 6.1 and 13.7 % for the spr2021/96 conjugates and between 5.3 and 6.5% for the GNA2091-fHbp proteins and mutants.
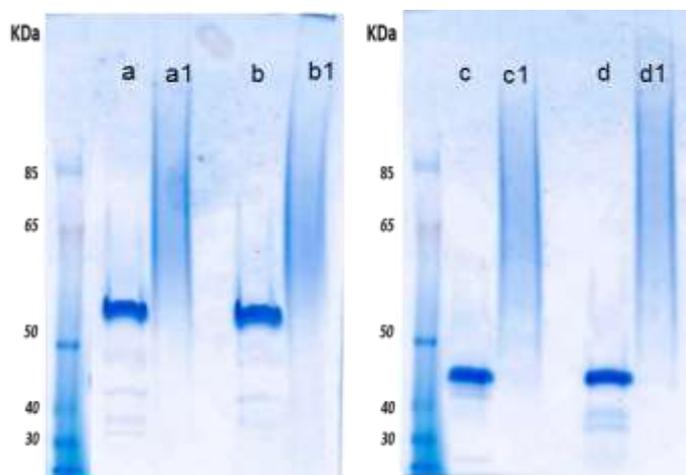

Figure 22 - Examples of SDS-page for the generated conjugates: Proteins spr2021/96_Lys172Arg,spr2021/96__Lys172Arg_Lys214Arg_Lys216Arg,spr2021_Lys204Arg, and spr2021_Lys204Arg_Lys248Arg (respectively a, b, c, d) and their relative laminarin glycoconjugates(a1, b1, c1,d2).

Laminarin conjugates	Total Saccharide µg/mL	Saccharide/protein (w/w)	Free saccharide %
Spr2021_Lys204Arg	*354.4*	*0.21*	*12.1*
Spr2021_Lys248Arg	*269.7*	*0.61*	*10.0*
Spr2021_Lys204Arg_Lys246 Arg_Lys248Arg	*421.4*	*0.41*	*16.6*
Spr2021/96_Lys172Arg	*465.1*	*0.34*	*10.7*
Spr2021/96_Lys216Arg	*569.3*	*0.20*	*13.2*
Spr2021/96_Lys172Arg_Lys 216Arg	*491.3*	*0.39*	*6.1*
Spr2021/96_Lys172Arg_Lys 214Arg_Lys216Arg	*437.7*	*0.19*	*13.7*
Spr2021-(CRM)-96	*582.1*	*0.35*	*10.9*
Spr2021/96	*541.5*	*0.29*	*9.2*
GNA2091-fHbp	*663.6*	*0.41*	*6.5*
GNA2091-(CRM)-fHbp	*201.3*	*0.47*	*5.3*
GNA2091-(2021)-fHbp	*581.5*	*0.7*	*-*

Table 5 - *S. pneumoniae and N. meningitidis Laminarin glycoconjugates: total content of saccharide, free saccharide and saccharide/protein (w/w) ratio.*

4.7 Immune response evaluation of mutants Laminarin conjugates in mice

For testing the ability of the new glycoconjugates produced to increase the immune response against the polysaccharide antigen, Laminarin conjugates spr2021, spr2021/96 and mutants were tested in mice. The study was performed without the use of adjuvants with an immunization dose of 5µg in term of total saccharide. The anti-Laminarin antibodies (IgG) were evaluated after the third immunization. Comparing the spr2021 conjugates and its relative mutants (**Figure 23 panel b)**, the point mutations made on the predicted MHC-II peptides did not allow a significant increase in the immunogenicity against the polysaccharide antigen. The same effect was showed in the case of the spr2021/96 and its relative mutants (**Figure 23 panel a**). The spr2021-CRM-96 conjugate, in which a

predicted MHC-II peptide from the CRM$_{197}$ sequence was added to the fusion protein spr2021/96 also showed a similar immune response compared to the original construct spr2021/96-Lam. In all constructs we obtained a similar immune response and a non significant increase among them was identified. Regarding the *N. meningitidis* conjugates, predicted MHC-II peptide from the CRM$_{197}$ protein and the spr2021 were added to the sequence with the aim of increasing the immune response of GNA2091-fHbp. Comparing the mutant tested to the original construct GNA2091-fHbp a non significant increase in the immunogenicity was observed. (**Figure 23 panel c**).
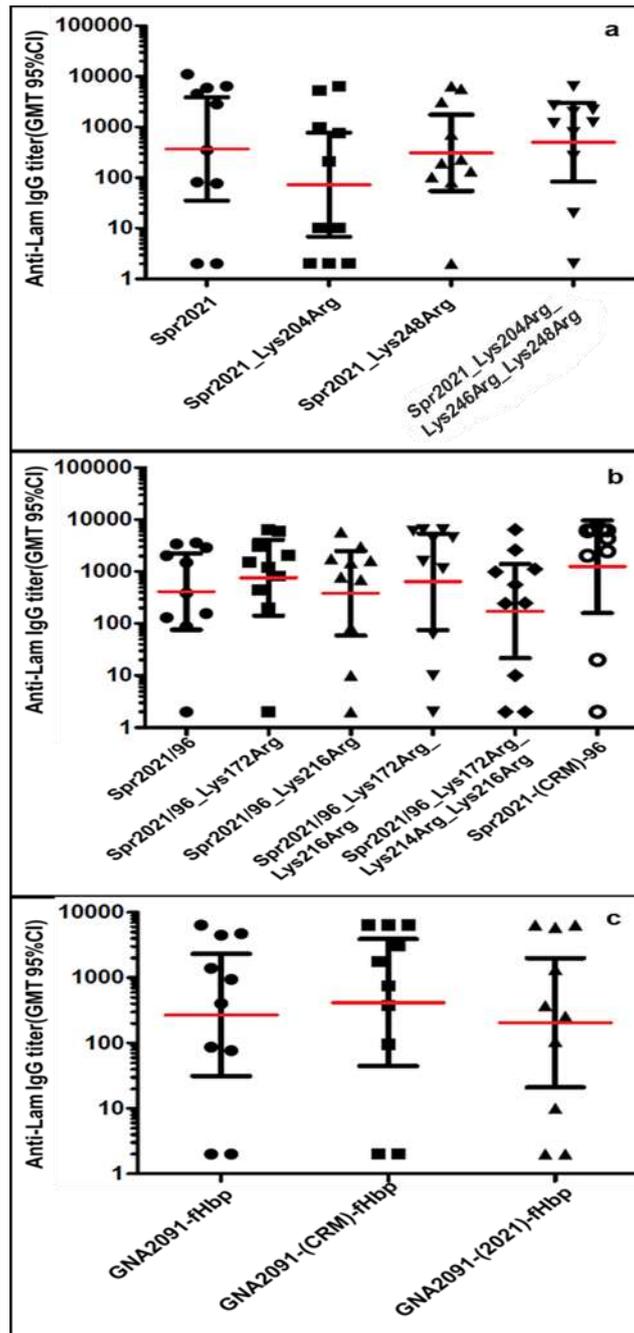
Figure 23 - *Anti-laminarin IgG response induced by spr2021 spr2021/96, or GNA2091-fHbp conjugates in comparison with the corresponding mutants.*

In panel (a) is shown the anti-laminarin IgG response induce by mutated spr2021 without presence of adjuvant in comparison with the spr2021.

In panel(b) is shown the anti-laminarin IgG response induce by mutated spr2021/96 without presence of adjuvant in comparison with the spr2021/06.

In panel(c) is shown the anti-laminarin IgG response induce by mutated GNA2091-fHbp without presence of adjuvant in comparison with the GNA2091-fHbp.

Anti-carrier IgG represented by each dot indicate a single mouse ELISA titer, the horizontal bar(red) is the geometric mean of the group, instead the vertical bar shows the statistical 95% CI.

5. Conclusions

Currently, glycoconjugate vaccines are among the safest and most efficacious vaccines developed so far, however they are still not completely characterized. Data show that the immunogenicity of conjugate vaccines can be affected by many variables such as the carrier protein, the conjugation chemistry and the carbohydrate antigen, resulting in different immunological properties. For instance, several studies have been published focusing on the influence of the carbohydrate on the immunogenicity of glycoconjugate vaccines and suggesting that chain length, conjugation process, saccharide loading onto the protein could somehow influence the glycoconjugate immunogenicity[80,81,82]. While these factors are often stated in publications, the attention on the protein carrier as the key factor on vaccine profile[83] has not been widely reported. The mechanism of action of glycoconjugate vaccines and in particular the role played by both the carbohydrate and the protein carrier during antigen presentation, the interaction with the MHC-II complex and the following interactions with the T cell receptor are still not completely clarified. Polysaccharides are generally considered to be T cell independent antigens and so not able to bind the MHC-II complex[84]. For this reason, the elicitation of T cell helpers by glycoconjugates is generally attributed to the peptide moieties, which is generated from the protein carrier processing. *Avici et al.* showed that carbohydrate motifs can be presented to the MHC-II thanks to the peptide that act as anchor for the binding to the MHC-II. In fact, the authors identified specific T cell clones for the specific carbohydrate antigen. A further consideraton is that most of the conjugate vaccines are produced from large poly/oligosaccharides, randomly distributed at the carrier surface, limiting the possibility that carbohydrate processing in the lysosome results in a sufficient quantity of homogenous glycopeptide suited for the T cell binding[48]. Aiming at bringing light to the influence of conjugation on processing and presentation of glycoconjugates, we first clustered the efficacy of 27 protein carriers reported in litterature[49], as a function of the presence in their sequences of peptides predicted to bind MHC-II with high affinities, as well as the presence/absence in these peptides of lysine residues, the amino acid involved in the conjugation[49]. We were able to evidence that carriers that did not induce anti-Lam antibody titers or

induce them to a level 5 time inferiors to the CRM_{197}-Lam used as benchmark, either do not bear predicted MHC-II-binder peptides of high affinity, or contain predicted MHC-II-binder peptides of high affinity that carry lysine residues. From this observation we made the hypothesis that the conjugation of the lysine residues present in predicted MHC-II-binder peptides of high affinity might considerably reduce the capability of a protein to function as an efficient carrier. The hypothesis was also supported by analyzing the set of proteins that function as efficient carriers and bear high affinity MHC-II binder peptides. Most of these peptides do not contain lysines (and the conjugation should not influence their high efficacy as carrier), but it was interesting to evidence that spr96/2021 that contains in the sequence of the spr2021 predicted high affinity MHC-II binder peptides with lysines induced an anti-Lam antibody titer significantly high, that is not the case of the spr2021. The same senario was observed from the GNA2091-fHbp carrier protein.

With the attempt to confirm the influence of glycosylation extent of MHC-II presenting peptides on the efficacy of the carrier, spr96/2021, spr2021 as well as the new fusion protein spr2021/96 were selected as a model to validate our hypothesis. We first showed that the three proteins induced a similar strong immune response in agreement with the presence of high affinity predicted MHC-II peptides identified. The observed immunogenicity did not increase when adjuvant was co-administrated, a situation that is quite different from what observed with CRM_{197}. These aspects led us to speculate that even if factors, such as binding, internalization or processing may be involved in the immune response, the selected model is not highly affected by them.

The three protein carriers were randomly conjugated with laminarin and their ability to induce an immune response against the polysaccharide antigen was tested. As previously showed, the fusion spr96/2021 showed a high IgG anti-laminarin antibody titer when compared to the spr2021. Curiously, the new conjugate tested spr2021/96 did not display a similar immunogenicity of its counterparty spr96/2021. To evaluate the extend of conjugation of the carrier, we used a new in house method, combining a chemical deglycosylation and mass spectrometry technique. We were able to identify a different pattern of glycosylation for the three constructs. In agreement to the hypothesis that

glycosylation on the MHC-II predicted peptides influences the efficacy of the carrier.

With the aim to improve to efficacy of the spr2021/96 and spr2021 carriers, we generated mutants in which the lysine residues involved in the conjugation were substituted with arginine residues. Moreover, insertion of predicted MHC-II peptide from the CRM_{197} protein was also inserted as spacer between the spr2021 and spr0096 proteins. As a second strategy, we selected another carrier protein, the GNA2091-fHbp in which the linker between the GNA2091 and fHbp protein was substituted with the predicted CRM_{197} MHC-II peptide and with one of the predicted spr2021 MHC-II peptide. All the mutated proteins showed a high level of immunogenicity comparable to the native forms. All the mutant conjugates were tested in mice for evaluating their ability to generate an immune response against the polysaccharide antigen. No increase in the immunogenicity was observed when the mutant conjugates were tested in mice. In all the tested group mice, a broad range of immune response was observed with the administration of the same conjugate; this wide range of immunogenicity together with the low IgG titer obtained for different mice in the same group increased the confidence interval for the GMT value obtained, implying the impossibility of applying any statistical analyses in order to identify differences among the constructs. The conjugates generated were deeply characterized and no differences were identified in term of sugar and protein/sugar ratio compared to the previous reported data[49]. Also the same strand, age and sex of mice was maintained to avoid any change in the immunogenicity response. The still low immunogenicity after the modification made on the carrier proteins could be due to the high number of predicted MHC-II peptides with a medium affinity, identified for the *S. pneumoniae* constructs that have not been taken into account in this study. Much effort was put in the identification of the peptides presented to the MHC-II for confirming the data obtained from the prediction software and the binding of the predicted peptides to the MHC-II. Mass spectrometry was extensively used for the detection of peptides presented by major histocompatibility complex class I and II, allowing rapid identification of a high number of peptides in a single experiment[85,86].

However, the analysis of MHC I and II peptides ligands from cells isolated in mouse was very challenging especially regarding the cost and amount of work

involved in each experiment. A high number of APCs were generally needed (in the order of 5 to 10 billion of APCs). Collecting a so high number of APCs required sacrificing from 50 to 100 mice[87]. Nevertheless, some laboratories were able to reduce the number of mice used allowing a reduction in term of amount of work. For instance, *in vivo* enrichment of mouse spleen DCs by Flt3L-B6 melanoma treatment was described, expanding the subset of dendritic cells of around 10 fold[88]. Even if it is possible to reduce the amount of mice necessary, other limitations came with the subsequent step of purification and identification. The techniques used for the purification of the MHC bounded peptides required an immunoaffinity purification, which is highly specific but come with the cost of a very low yield[89]. A further restriction of this approach, is the unpredictable cleavage sites of the peptides generated for the presentation to the MHC complex. The cleavage generated by specific proteases, led to have known C-terminus amino acid facilitating the MS/MS data analysis. Because of the impossibility of predicting the C-terminus region on these analyses, all the 20 amino acid must have been considered as potential terminal residues, increasing the possibility in misidentifying the peptides. The main limitation in our case was due to the unfeasibility to have a so high number of mice and/or to use hematopoietic factors such as the Flt3L for enhancing the number of APCs. We were forced to limit our experiments to the use of 10 mice, in which all the splenocytes (dendritic cells, B lymphocytes, macrophages) were collected. Briefly, the splenocytes were infected with the protein spr96/2021, after 6h of infection, the cells were lysated and the MHC-II bounded peptides were purified using an immunoaffinity purification. The sample was then analyzed by mass spectrometry but no peptides from spr96/2021 were identified. The impossibility to perform experiments with a higher number of mice and the intrinsic limitation of the experiment above described prevented to verify that the two peptides predicted with a high predicted affinity for the MHC-II complex were empirically presented on the MHC-II. Due to the failure in identifying the MHC-II bounded peptides we were forced to rely only on the prediction software. Future study will be necessary to well define the role played by glycosylation in generating a memory response. Other approaches that may be used to study the influence of glycosylation in developing an immune response, and the role played by the glycosylation on MHC-II presented peptides, may be accomplished by using specific glycosylation

on the T cell epitopes of interest. Recently, an *in vivo* conjugation process for the generation of the so called bioconjugate vaccines have been developed[90]. This technology allows the biosynthesis of polysaccharide and carrier proteins in *E. coli* cells with the subsequent *in vivo* coupling of the carrier to the polysaccharide by using a specific oligosaccharyl transferase enzyme from the N-linked glycosylation system of *Campylobacter*[91,92]. Without going to much in detail on this engineered glycosylation process, the antigenic oligosaccharide can be enzymatically introduced on the Asparagine residues of the consensus sequence D/E-Xaa-N-Xaa-S/T, where Xaa can be any amino acid except Proline. The consensus sequence was inserted into different proteins such as EPA, Hla that are not glycosylated in their original organism, leading to develop of bioconjugate candidate vaccines against *Staphylococcous aureus* and *Shigella flexneri 2a*[93,94]. The consensus sequence used to generate the bioconjugates could be inserted in carrier proteins containing predicted MHC-II peptides of high affinity, modulating in a specific way the site of conjugation. Using this technology it will be possible to understand the rule played by the glycosylation on specific T cell epitope when glycosylated and correlate the IgG titer to the affinity of the predicted MHC-II peptides.

References

[1] WHO *Global Vaccine Action Plan 2011-2020*

[2] Baxby, Derrick. 1999. 'Edward Jenner's Inquiry after 200 years', *BMJ : British Medical Journal*, 318: 390-90; ibid.; ibid.; ibid.; Bonanni, Paolo, and José Ignacio Santos. 2011. 'Vaccine evolution', *Perspectives in Vaccinology*, 1: 1-24.

[3] Plotkin, Stanley. 2014. 'History of vaccination', *Proceedings of the National Academy of Sciences*, 111: 12283-87.

[4] Strugnell, Richard, Fred Zepp, Anthony Cunningham, and Terapong Tantawichien. 2011. 'Vaccine antigens', *Perspectives in Vaccinology*, 1: 61-88.

[5] Plotkin, S.A., W.A. Orenstein, and P.A. Offit. 2012. *Vaccines* (Elsevier/Saunders).

[6] Roberts, Ian S. 1996. 'THE BIOCHEMISTRY AND GENETICS OF CAPSULAR POLYSACCHARIDE PRODUCTION IN BACTERIA', *Annual Review of Microbiology*, 50: 285-315.

[7] Taylor, C. M., and I. S. Roberts. 2005. 'Capsular polysaccharides and their role in virulence', *Contrib Microbiol*, 12: 55-66.

[8] Lepenies, B. 2015. *Carbohydrate-based Vaccines: Methods and Protocols* (Humana Press).

[9] Heidelberger, M., and O. T. Avery. 1923. 'THE SOLUBLE SPECIFIC SUBSTANCE OF PNEUMOCOCCUS', *The Journal of Experimental Medicine*, 38: 73-79.

[10] Avery, Oswald T., and René Dubos. 1931. 'THE PROTECTIVE ACTION OF A SPECIFIC ENZYME AGAINST TYPE III PNEUMOCOCCUS INFECTION IN MICE'ibid., 54: 73-89.

[11] Mond, J. J., and J. F. Kokai-Kun. 2008. 'The multifunctional role of antibodies in the protective response to bacterial T cell-independent antigens', *Curr Top Microbiol Immunol*, 319: 17-40.

[12] Bonanni, Paolo, and José Ignacio Santos. 2011. 'Vaccine evolution', *Perspectives in Vaccinology*, 1: 1-24.

[13] Constant, Stephanie L. 1999. 'B Lymphocytes as Antigen-Presenting Cells for CD4⁺T Cell Priming In Vivo', *The Journal of Immunology*, 162: 5695-703.

[14] Lanzavecchia, A. 1985. 'Antigen-specific interaction between T and B cells', *Nature*, 314: 537-9.

[15] Pulendran, Bali, and Rafi Ahmed. 2011. 'Immunological mechanisms of vaccination', *Nature Immunology*, 12: 509-17.

[16] Mond, James J., Andrew Lees, and Clifford M. Snapper. 1995a. 'T Cell-Independent Antigens Type 2', *Annual Review of Immunology*, 13: 655-92.

[17] Snapper, C M, and J J Mond. 1996. 'A model for induction of T cell-independent humoral immunity in response to polysaccharide antigens', *The Journal of Immunology*, 157: 2229-33.

[18] Hoffman, William, Fadi G. Lakkis, and Geetha Chalasani. 2016. 'B Cells, Antibodies, and More', *Clinical Journal of the American Society of Nephrology : CJASN*, 11: 137-54.

[19] Avery, Oswald T., and Walther F. Goebel. 1931. 'CHEMO-IMMUNOLOGICAL STUDIES ON CONJUGATED CARBOHYDRATE-PROTEINS : V. THE IMMUNOLOGICAL SPECIFITY OF AN ANTIGEN PREPARED BY COMBINING THE CAPSULAR POLYSACCHARIDE OF TYPE III PNEUMOCOCCUS WITH FOREIGN PROTEIN', *The Journal of Experimental Medicine*, 54: 437-47.

[20] Kumar, Sandeep, Satish K. Singh, Xiaoling Wang, Bonita Rup, and Davinder Gill. 2011. 'Coupling of Aggregation and Immunogenicity in Biotherapeutics: T- and B-Cell Immune Epitopes May Contain Aggregation-Prone Regions', *Pharmaceutical Research*, 28: 949.

[21] Gessner, B. D., and R. A. Adegbola. 2008. 'The impact of vaccines on pneumonia: key lessons from Haemophilus influenzae type b conjugate vaccines', *Vaccine*, 26 Suppl 2: B3-8.

[22] Ward, Joel, Carol Berkowitz, John Pescetti, Kelly Burkart, Oel Samuelson, and Lance Gordon. 1984. 'ENHANCED IMMUNOGENICITY IN YOUNG INFANTS OF A NEW HAEMOPHILUS INFLUENZAE TYPE B[lpar]HIB[rpar] CAPSULAR POLYSACCHARIDE[lpar]PRP[rpar]-DIPHTHERIA TOXOID[lpar]D[rpar] CONJUGATE VACCINE', *Pediatr Res*, 18: 287A-87A.

[23] Cooper, B., L. DeTora, and J. Stoddard. 2011. 'Menveo(R)): a novel quadrivalent meningococcal CRM197 conjugate vaccine against serogroups A, C, W-135 and Y', *Expert Rev Vaccines*, 10: 21-33.

[24] Croxtall, J. D., and S. Dhillon. 2012. 'Meningococcal quadrivalent (serogroups A, C, W135 and Y) tetanus toxoid conjugate vaccine (Nimenrix)', *Drugs*, 72: 2407-30.

[25] Plosker, Greg L. 2015. '13-Valent Pneumococcal Conjugate Vaccine: A Review of Its Use in Adults'ibid., 75: 1535-46.

[26] Adamo, Roberto, Alberto Nilo, Bastien Castagner, Omar Boutureira, Francesco Berti, and Gonçalo J. L. Bernardes. 2013a. 'Synthetically defined glycoprotein vaccines: current status and

future directions†This perspective is dedicated to the memory of Professor David Y. Gin', *Chemical Science*, 4: 2995-3008.

[27] Bröker, Michael. 2016. 'Potential protective immunogenicity of tetanus toxoid, diphtheria toxoid and Cross Reacting Material 197 (CRM197) when used as carrier proteins in glycoconjugates', *Human Vaccines & Immunotherapeutics*, 12: 664-67.

[28] Berti, Francesco, and Roberto Adamo. 2013. 'Recent Mechanistic Insights on Glycoconjugate Vaccines and Future Perspectives', *ACS Chemical Biology*, 8: 1653-63.

[29] Adamo, Roberto, Alberto Nilo, Bastien Castagner, Omar Boutureira, Francesco Berti, and Gonçalo J. L. Bernardes. 2013a. 'Synthetically defined glycoprotein vaccines: current status and future directions†This perspective is dedicated to the memory of Professor David Y. Gin', *Chemical Science*, 4: 2995-3008.

[30] Pichichero, Michael E. 2013. 'Protein carriers of conjugate vaccines: Characteristics, development, and clinical trials', *Human Vaccines & Immunotherapeutics*, 9: 2505-23.

[31] Neumuller, Carola. 1954. 'Detoxification of Diphtheria Toxin with Formaldehyde mixed with an Amino-Acid', *Nature*, 174: 405-06.

[32] Giannini, G., R. Rappuoli, and G. Ratti. 1984. 'The amino-acid sequence of two non-toxic mutants of diphtheria toxin: CRM45 and CRM197', *Nucleic Acids Research*, 12: 4063-69.

[33] Ferrante, Andrea. 2013. 'Thermodynamics of Peptide-MHC Class II Interactions: Not all Complexes are Created Equal', *Frontiers in Immunology*, 4: 308.

[34] Chicz, R. M., R. G. Urban, W. S. Lane, J. C. Gorga, L. J. Stern, D. A. Vignali, and J. L. Strominger. 1992. 'Predominant naturally processed peptides bound to HLA-DR1 are derived from MHC-related molecules and are heterogeneous in size', *Nature*, 358: 764-8.

[35] Rotzschke, O., K. Falk, J. Mack, J. M. Lau, G. Jung, and J. L. Strominger. 1999. 'Conformational variants of class II MHC/peptide complexes induced by N- and C-terminal extensions of minimal peptide epitopes', *Proc Natl Acad Sci U S A*, 96: 7445-50.

[36] Arnold, P. Y., N. L. La Gruta, T. Miller, K. M. Vignali, P. S. Adams, D. L. Woodland, and D. A. Vignali. 2002. 'The majority of immunogenic epitopes generate CD4+ T cells that are dependent on MHC class II-bound peptide-flanking residues', *J Immunol*, 169: 739-49.

[37] Nielsen, Morten, Ole Lund, Søren Buus, and Claus Lundegaard. 2010. 'MHC Class II epitope predictive algorithms', *Immunology*, 130: 319-28.

[38] McHeyzer-Williams, M., S. Okitsu, N. Wang, and L. McHeyzer-Williams. 2011. 'Molecular programming of B cell memory', *Nat Rev Immunol*, 12: 24-34.

[39] Mond, J. J., A. Lees, and C. M. Snapper. 1995b. 'T cell-independent antigens type 2', *Annu Rev Immunol*, 13: 655-92.

[40] Pollard, A. J., K. P. Perrett, and P. C. Beverley. 2009. 'Maintaining protection against invasive bacteria with protein-polysaccharide conjugate vaccines', *Nat Rev Immunol*, 9: 213-20.

[41] Rappuoli, Rino, and Ennio De Gregorio. 2011. 'A sweet T cell response', *Nat Med*, 17: 1551-52.

[42] Duan, Jinyou, Fikri Y. Avci, and Dennis L. Kasper. 2008. 'Microbial carbohydrate depolymerization by antigen-presenting cells: Deamination prior to presentation by the MHCII pathway', *Proc Natl Acad Sci U S A*, 105: 5183-88.

[43] Cobb, Brian A., Qun Wang, Arthur O. Tzianabos, and Dennis L. Kasper. 2004. 'Polysaccharide Processing and Presentation by the MHCII Pathway', *Cell*, 117: 677-87.

[44] Avci, F. Y., X. Li, M. Tsuji, and D. L. Kasper. 2011. 'A mechanism for glycoconjugate vaccine activation of the adaptive immune system and its implications for vaccine design', *Nat Med*, 17: 1602-9.

[45] Pellicci, D. G., A. J. Clarke, O. Patel, T. Mallevaey, T. Beddoe, J. Le Nours, A. P. Uldrich, J. McCluskey, G. S. Besra, S. A. Porcelli, L. Gapin, D. I. Godfrey, and J. Rossjohn. 2011. 'Recognition of beta-linked self glycolipids mediated by natural killer T cell antigen receptors', *Nat Immunol*, 12: 827-33.

[46] Kabat, E. A. 1960. 'The upper limit for the size of the human antidextran combining site', *J Immunol*, 84: 82-5.

[47] Verez-Bencomo, V., V. Fernández-Santana, Eugenio Hardy, Maria E. Toledo, Maria C. Rodríguez, Lazaro Heynngnezz, Arlene Rodriguez, Alberto Baly, Luis Herrera, Mabel Izquierdo, Annette Villar, Yury Valdés, Karelia Cosme, Mercedes L. Deler, Manuel Montane, Ernesto Garcia, Alexis Ramos, Aristides Aguilar, Ernesto Medina, Gilda Toraño, Iván Sosa, Ibis Hernandez, Raydel Martínez, Alexis Muzachio, Ania Carmenates, Lourdes Costa, Félix Cardoso, Concepción Campa, Manuel Diaz, and René Roy. 2004. 'A Synthetic Conjugate Polysaccharide Vaccine Against Haemophilus influenzae Type b', *Science*, 305: 522-25.

[48] Polonskaya, Zinaida, Shenglou Deng, Anita Sarkar, Lisa Kain, Marta Comellas-Aragones, Craig S. McKay, Katarzyna Kaczanowska, Marie Holt, Ryan McBride, Valle Palomo, Kevin M. Self, Seth Taylor, Adriana Irimia, Sanjay R. Mehta, Jennifer M. Dan, Matthew Brigger, Shane Crotty, Stephen P. Schoenberger, James C. Paulson, Ian A. Wilson, Paul B. Savage, M. G. Finn, and Luc Teyton. 2017. 'T cells control the generation of nanomolar-affinity anti-glycan antibodies', *The Journal of Clinical Investigation*, 127: 1491-504.

[49] Tontini, M., M. R. Romano, D. Proietti, E. Balducci, F. Micoli, C. Balocchi, L. Santini, V. Masignani, F. Berti, and P. Costantino. 2016. 'Preclinical studies on new proteins as carrier for glycoconjugate vaccines', *Vaccine*, 34: 4235-42.

[50] Torosantucci, Antonella, Carla Bromuro, Paola Chiani, Flavia De Bernardis, Francesco Berti, Chiara Galli, Francesco Norelli, Cinzia Bellucci, Luciano Polonelli, Paolo Costantino, Rino Rappuoli, and Antonio Cassone. 2005. 'A novel glyco-conjugate vaccine against fungal pathogens', *The Journal of Experimental Medicine*, 202: 597-606.

[51] Adamo, Roberto, Marta Tontini, Giulia Brogioni, Maria Rosaria Romano, Gabriele Costantini, Elisa Danieli, Daniela Proietti, Francesco Berti, and Paolo Costantino. 2011. 'Synthesis of Laminarin Fragments and Evaluation of a β-(1,3) Glucan Hexasaccharide-CRM197 Conjugate as Vaccine Candidate against Candida albicans', *Journal of Carbohydrate Chemistry*, 30: 249-80.

[52] Bromuro, C., M. Romano, P. Chiani, F. Berti, M. Tontini, D. Proietti, E. Mori, A. Torosantucci, P. Costantino, R. Rappuoli, and A. Cassone. 2010. 'Beta-glucan-CRM197 conjugates as candidates antifungal vaccines', *Vaccine*, 28: 2615-23.

[53] Giuliani, M. M., J. Adu-Bobie, M. Comanducci, B. Arico, S. Savino, L. Santini, B. Brunelli, S. Bambini, A. Biolchi, B. Capecchi, E. Cartocci, L. Ciucchi, F. Di Marcello, F. Ferlicca, B. Galli, E. Luzzi, V. Masignani, D. Serruto, D. Veggi, M. Contorni, M. Morandi, A. Bartalesi, V. Cinotti, D. Mannucci, F. Titta, E. Ovidi, J. A. Welsch, D. Granoff, R. Rappuoli, and M. Pizza. 2006. 'A universal vaccine for serogroup B meningococcus', *Proc Natl Acad Sci U S A*, 103: 10834-9.

[54] Costantino, P., and M.R. Romano. 2013. "Carrier molecule comprising a spr0096 and a spr2021 antigen." In.: Google Patents.

[55] Olaya-Abril, A., I. Jimenez-Munguia, L. Gomez-Gascon, I. Obando, and M. J. Rodriguez-Ortega. 2013. 'Identification of potential new protein vaccine candidates through pan-surfomic analysis of pneumococcal clinical isolates from adults', *PLoS One*, 8: e70365.

[56] Croxtall, J. D., and S. Dhillon. 2012. 'Meningococcal quadrivalent (serogroups A, C, W135 and Y) tetanus toxoid conjugate vaccine (Nimenrix)', *Drugs*, 72: 2407-30.

[57] Cooper, B., L. DeTora, and J. Stoddard. 2011. 'Menveo(R)): a novel quadrivalent meningococcal CRM197 conjugate vaccine against serogroups A, C, W-135 and Y', *Expert Rev Vaccines*, 10: 21-33.

[58] Micoli, F., M. R. Romano, M. Tontini, E. Cappelletti, M. Gavini, D. Proietti, S. Rondini, E. Swennen, L. Santini, S. Filippini, C. Balocchi, R. Adamo, G. Pluschke, G. Norheim, A. Pollard, A. Saul, R. Rappuoli, C. A. MacLennan, F. Berti, and P. Costantino. 2013. 'Development of a glycoconjugate vaccine to prevent meningitis in Africa caused by meningococcal serogroup X', *Proc Natl Acad Sci U S A*, 110: 19077-82.

[59] Smith, P. K., R. I. Krohn, G. T. Hermanson, A. K. Mallia, F. H. Gartner, M. D. Provenzano, E. K. Fujimoto, N. M. Goeke, B. J. Olson, and D. C. Klenk. 1985. 'Measurement of protein using bicinchoninic acid', *Anal Biochem*, 150: 76-85.

[60] Habeeb, A. F. 1966. 'Determination of free amino groups in proteins by trinitrobenzenesulfonic acid'ibid., 14: 328-36.

[61] Miron, Talia, and Meir Wilchek. 1982. 'A spectrophotometric assay for soluble and immobilized N-hydroxysuccinimide esters', *Anal Biochem*, 126: 433-35.

[62] Fleri, Ward, Sinu Paul, Sandeep Kumar Dhanda, Swapnil Mahajan, Xiaojun Xu, Bjoern Peters, and Alessandro Sette. 2017. 'The Immune Epitope Database and Analysis Resource in Epitope Discovery and Synthetic Vaccine Design', *Frontiers in Immunology*, 8: 278.

[63] Wang, P., J. Sidney, C. Dow, B. Mothe, A. Sette, and B. Peters. 2008. 'A systematic assessment of MHC class II peptide binding predictions and evaluation of a consensus approach', *PLoS Comput Biol*, 4: e1000048.

[64] Wang, P., J. Sidney, Y. Kim, A. Sette, O. Lund, M. Nielsen, and B. Peters. 2010. 'Peptide binding predictions for HLA DR, DP and DQ molecules', *BMC Bioinformatics*, 11: 568.

[65] Andreatta, M., E. Karosiene, M. Rasmussen, A. Stryhn, S. Buus, and M. Nielsen. 2015. 'Accurate pan-specific prediction of peptide-MHC class II binding affinity with improved binding core identification', *Immunogenetics*, 67: 641-50.

[66] Nielsen, M., and O. Lund. 2009. 'NN-align. An artificial neural network-based alignment algorithm for MHC class II peptide binding prediction', *BMC Bioinformatics*, 10: 296.

[67] Nielsen, M., C. Lundegaard, and O. Lund. 2007. 'Prediction of MHC class II binding affinity using SMM-align, a novel stabilization matrix alignment method'ibid., 8: 238.

[68] Southwood, S., J. Sidney, A. Kondo, M. F. del Guercio, E. Appella, S. Hoffman, R. T. Kubo, R. W. Chesnut, H. M. Grey, and A. Sette. 1998. 'Several common HLA-DR types share largely overlapping peptide binding repertoires', *J Immunol*, 160: 3363-73.

[69] Cardaci, Angela, Salvatore Papasergi, Angelina Midiri, Giuseppe Mancuso, Maria Domina, Veronica Lanza Cariccio, Francesca Mandanici, Roberta Galbo, Carla Lo Passo, Ida Pernice, Paolo Donato, Susanna Ricci, Carmelo Biondo, Giuseppe Teti, Franco Felici, and Concetta Beninati. 2012. 'Protective Activity of Streptococcus pneumoniae Spr1875 Protein Fragments Identified Using a Phage Displayed Genomic Library', *PLoS One*, 7: e36588.

[70] Olaya-Abril, A., I. Jimenez-Munguia, L. Gomez-Gascon, I. Obando, and M. J. Rodriguez-Ortega. 2013. 'Identification of potential new protein vaccine candidates through pan-surfomic analysis of pneumococcal clinical isolates from adults', *PLoS One*, 8: e70365.

[71] Costantino, P., and M.R. Romano. 2013. "Carrier molecule comprising a spr0096 and a spr2021 antigen." In.: Google Patents.

[72] Giusti, Fabiola, Anja Seubert, Rocco Cantisani, Marco Tortoli, Ugo D'Oro, Ilaria Ferlenghi, Romano Dallai, and Diego Piccioli. 2015. 'Ultrastructural Visualization of Vaccine Adjuvant Uptake In Vitro and In Vivo', *Microscopy and Microanalysis*, 21: 791-95.

[73] Adamo, Roberto, Alberto Nilo, Bastien Castagner, Omar Boutureira, Francesco Berti, and Goncalo J. L. Bernardes. 2013b. 'Synthetically defined glycoprotein vaccines: current status and future directions', *Chemical Science*, 4: 2995-3008.

[74] Pozsgay, Vince, Chiayung Chu, Lewis Pannell, Jennifer Wolfe, John B. Robbins, and Rachel Schneerson. 1999. 'Protein conjugates of synthetic saccharides elicit higher levels of serum IgG lipopolysaccharide antibodies in mice than do those of the O-specific polysaccharide from Shigella dysenteriae type 1', *Proceedings of the National Academy of Sciences*, 96: 5194-97.

[75] Berti, Francesco, and Roberto Adamo. 2013. 'Recent Mechanistic Insights on Glycoconjugate Vaccines and Future Perspectives', *ACS Chemical Biology*, 8: 1653-63.

[76] Crotti, Stefano, Huili Zhai, Jing Zhou, Martin Allan, Daniela Proietti, Werner Pansegrau, Qi-Ying Hu, Francesco Berti, and Roberto Adamo. 2014. 'Defined Conjugation of Glycans to the Lysines of CRM197 Guided by their Reactivity Mapping', *ChemBioChem*, 15: 836-43.

[77] Möginger, U., A. Resemann, C. E. Martin, S. Parameswarappa, S. Govindan, E. C. Wamhoff, F. Broecker, D. Suckau, C. L. Pereira, C. Anish, P. H. Seeberger, and D. Kolarich. 2016. 'Cross Reactive Material 197 glycoconjugate vaccines contain privileged conjugation sites', *Sci Rep*, 6.

[78] Bartual, Sergio G., Daniel Straume, Gro Anita Stamsås, Inés G. Muñoz, Carlos Alfonso, Martín Martínez-Ripoll, Leiv Sigve Håvarstein, and Juan A. Hermoso. 2014. 'Structural basis of PcsB-mediated cell separation in Streptococcus pneumoniae', 5: 3842.

[79] Bradshaw, Ralph A., William W. Brickey, and Kenneth W. Walker. 1998. 'N-Terminal processing: the methionine aminopeptidase and Nα-acetyl transferase families', *Trends in Biochemical Sciences*, 23: 263-67.

[80] Peeters, J. M., T. G. Hazendonk, E. C. Beuvery, and G. I. Tesser. 1989. 'Comparison of four bifunctional reagents for coupling peptides to proteins and the effect of the three moieties on the immunogenicity of the conjugates', *J Immunol Methods*, 120: 133-43.

[81] Rana, R., J. Dalal, D. Singh, N. Kumar, S. Hanif, N. Joshi, and M. K. Chhikara. 2015. 'Development and characterization of Haemophilus influenzae type b conjugate vaccine prepared using different polysaccharide chain lengths', *Vaccine*, 33: 2646-54.

[82] Wang, Chia-Hung, Shiou-Ting Li, Tzu-Lung Lin, Yang-Yu Cheng, Tsung-Hsien Sun, Jin-Town Wang, Ting-Jen R. Cheng, Kwok Kong Tony Mong, Chi-Huey Wong, and Chung-Yi Wu. 2013. 'Synthesis of Neisseria meningitidis Serogroup W135 Capsular Oligosaccharides for Immunogenicity Comparison and Vaccine Development', *Angewandte Chemie International Edition*, 52: 9157-61.

[83] Bröker, Michael, Francesco Berti, and Paolo Costantino. 2016. 'Factors contributing to the immunogenicity of meningococcal conjugate vaccines', *Human Vaccines & Immunotherapeutics*, 12: 1808-24.

[84] Duan, Jinyou, Fikri Y. Avci, and Dennis L. Kasper. 2008. 'Microbial carbohydrate depolymerization by antigen-presenting cells: Deamination prior to presentation by the MHCII pathway', *Proc Natl Acad Sci U S A*, 105: 5183-88.

[85] Tan, C. T., N. P. Croft, N. L. Dudek, N. A. Williamson, and A. W. Purcell. 2011. 'Direct quantitation of MHC-bound peptide epitopes by selected reaction monitoring', *Proteomics*, 11: 2336-40.

[86] Lemmel, C., S. Weik, U. Eberle, J. Dengjel, T. Kratt, H. D. Becker, H. G. Rammensee, and S. Stevanovic. 2004. 'Differential quantitative analysis of MHC ligands by mass spectrometry using stable isotope labeling', *Nat Biotechnol*, 22: 450-4.

[87] Karunakaran, Karuna P., Hong Yu, Xiaozhou Jiang, Queenie Chan, Michael F. Goldberg, Marc K. Jenkins, Leonard J. Foster, and Robert C. Brunham. 2017. 'Identification of MHC-Bound Peptides from Dendritic Cells Infected with Salmonella enterica Strain SL1344: Implications for a Nontyphoidal Salmonella Vaccine', *Journal of Proteome Research*, 16: 298-306.

[88] Bozzacco, Leonia, Haiqiang Yu, Henry A. Zebroski, Jörn Dengjel, Haiteng Deng, Svetlana Mojsov, and Ralph M. Steinman. 2011. 'Mass Spectrometry Analysis and Quantitation of Peptides Presented on the MHC II Molecules of Mouse Spleen Dendritic Cells'ibid., 10: 5016-30.

[89] Hassan, Chopie, Michel G. D. Kester, Gideon Oudgenoeg, Arnoud H. de Ru, George M. C. Janssen, Jan W. Drijfhout, Robbert M. Spaapen, Connie R. Jiménez, Mirjam H. M. Heemskerk, J. H. Frederik Falkenburg, and Peter A. van Veelen. 2014. 'Accurate quantitation of MHC-bound peptides by application of isotopically labeled peptide MHC complexes', *Journal of Proteomics*, 109: 240-44.

[90] Hatz, Christoph F. R., Bettina Bally, Susanne Rohrer, Robert Steffen, Stefanie Kramme, Claire-Anne Siegrist, Michael Wacker, Cristina Alaimo, and Veronica Gambillara Fonck. 2015. 'Safety and immunogenicity of a candidate bioconjugate vaccine against Shigella dysenteriae type 1 administered to healthy adults: A single blind, partially randomized Phase I study', *Vaccine*, 33: 4594-601.

[91] Wacker, Michael, Dennis Linton, Paul G. Hitchen, Mihai Nita-Lazar, Stuart M. Haslam, Simon J. North, Maria Panico, Howard R. Morris, Anne Dell, Brendan W. Wren, and Markus Aebi. 2002. 'N-Linked Glycosylation in Campylobacter jejuni and Its Functional Transfer into E. coli', *Science*, 298: 1790-93.

[92] Kowarik, Michael, Shin Numao, Mario F. Feldman, Benjamin L. Schulz, Nico Callewaert, Eva Kiermaier, Ina Catrein, and Markus Aebi. 2006. 'N-Linked Glycosylation of Folded Proteins by the Bacterial Oligosaccharyltransferase'ibid., 314: 1148-50.

[93] Kampf, M. M., M. Braun, D. Sirena, J. Ihssen, L. Thony-Meyer, and Q. Ren. 2015. 'In vivo production of a novel glycoconjugate vaccine against Shigella flexneri 2a in recombinant Escherichia coli: identification of stimulating factors for in vivo glycosylation', *Microb Cell Fact*, 14: 12.

[94] Wacker, Michael, Linhui Wang, Michael Kowarik, Meghan Dowd, Gerd Lipowsky, Amir Faridmoayer, Kelly Shields, Saeyoung Park, Cristina Alaimo, Kathryn A. Kelley, Martin Braun, Julien Quebatte, Veronica Gambillara, Paula Carranza, Michael Steffen, and Jean C. Lee. 2014. 'Prevention of Staphylococcus aureus Infections by Glycoprotein Vaccines Synthesized in Escherichia coli', *The Journal of Infectious Diseases*, 209: 1551-61.