



**THÈSE DE DOCTORAT**  
**DE L'UNIVERSITÉ PSL**

Préparée à l'Ecole Normale Supérieure  
Dans le cadre d'une cotutelle avec l'Université de Florence

**Peptides glucosylés dans les maladies auto-immunes:  
stratégies de synthèse et leur application à la détection  
et à la capture d'anticorps**

**Glucosylated peptides in autoimmune diseases:  
synthetic strategies and application to antibody  
detection and capture**

Soutenue par  
**Antonio MAZZOLENI**  
Le 05 mai 2020

Ecole doctorale n° 406  
**Chimie moléculaire de Paris-  
Centre**

Spécialité  
**Chimie moléculaire**

Composition du jury :

Boualem SENDID Professor, Lille University	<i>Président</i>
Laurence MULARD DR, Institut Pasteur	<i>Rapporteur</i>
Claudio DE FELICE Doctor, University of Siena	<i>Rapporteur</i>
Sara PELLEGRINO Professor, University of Milan	<i>Examineur</i>
Frédéric SCHMIDT DR, Institut Curie	<i>Examineur</i>
Paolo ROVERO Professor, University of Florence	<i>Examineur</i>
Jean-Maurice MALLET DR, Ecole Normale Supérieure	<i>Directeur de thèse</i>
Anna Maria PAPINI Professor, University of Florence	<i>Directeur de thèse</i>

*Alla mia Mamma.  
Ma anche al mio Babbo,  
di nuovo e per sempre.*

## **ACKNOWLEDGEMENTS**

First, I would like to express my profound gratitude to both my PhD advisors, Dr Jean-Maurice Mallet and Professor Anna Maria Papini for giving me the chance to participate in this project. I really want to thank them for their constant support, guidance, and inspiration. I will never forget this extraordinary time of my life, and I sincerely owe them for their kindness and trust.

I acknowledge the French-Italian University for financing such an international and interdisciplinary research project, and all the people who constantly work to provide new opportunities and perspectives for young adventurers.

I want to thank all the members of my thesis committee, for the time they spent reading my manuscript, and for their enthusiastic comments and suggestions. In particular, I would like to thank Professor Paolo Rovero for his mentoring throughout my PhD journey.

A giant thank goes to all my colleagues and friends from all the Laboratoire de Biomolécules's teams and the Franco-Italian PeptLab's group. It was my pleasure and honour to work in both these two amazing environments, full of such zealous, passionate, and excellent people. I am not able to provide a comprehensive list, everyone having been so important to me, but a special mention (randomly) goes to: Solange L., Olivier L., Laurent C., Agathe P., Roxanne O., Ghaniat G., Martha Z., Blaise D., Rodrigue M., Benjamin Z., Feliciano R.F., Fosca E., Claudia B., Mario C., Giuseppina S., Francesca N., Chiara T., Lorenzo A., Lorenzo P., Hendrik R..

I also truly need to thank my family, Mauro, Anna, Lollo and my long-time friends, because they constituted a constant source of strength and emotional support, making this unforgettable journey so rich and intense.

Last but not least, during the last three years, I have met so many people from all over the world (colleagues, friends, flatmates, collaborators, senior researchers, young scientists and not only), who I will never forget, because they contributed in different but always meaningful ways to this challenging and thrilling piece of work – and of life. It would be impossible for me to summarize with a list of names the enormous gratitude and affection I feel for each of them, so I will just conclude with my most sincere

*Grazie a tutti*

*Merci à tous*

*Thank you all*

Antonio Mazzoleni



## **TABLE OF CONTENTS**

<b>1. INTRODUCTION</b> .....	4
<b>1.1 AUTOIMMUNE DISEASES: a chemist's point of view</b> .....	4
<b>1.1.1 Immune system and autoimmune diseases</b> .....	4
<b>1.1.2 Infections and clinical relevance of molecular mimicry</b> .....	6
<b>1.1.3 Autoantibodies as biomarkers</b> .....	8
<b>1.1.4 Peptide-based approaches in autoimmune diseases</b> .....	13
<b>1.1.5 Isolation of antibodies and plasmapheresis</b> .....	15
<b>1.2 MULTIPLE SCLEROSIS: a complicate puzzle</b> .....	20
<b>1.2.1 Overview of Multiple Sclerosis</b> .....	20
<b>1.2.2 Cause and pathogenesis</b> .....	21
<b>1.2.3 Diagnosis</b> .....	22
<b>1.2.4 Current therapies</b> .....	23
<b>1.2.5 N-glycosylation and HMW1 adhesin</b> .....	27
<b>2. SCOPE OF THE PROJECT AND THESIS OUTLINE</b> .....	32
<b>3. HMW1(1347-1354)-DERIVED PEPTIDES</b> .....	35
<b>3.1 INTRODUCTION</b> .....	35
<b>3.2 RESULTS AND DISCUSSION</b> .....	37
<b>3.2.1 Glycosylated peptide synthesis</b> .....	37
<b>3.2.2 Immunological Assays</b> .....	41
<b>3.2.3 NMR and CD-based conformational investigations</b> .....	47
<b>3.3 CONCLUSIONS</b> .....	52
<b>4. PEPTIDE HOMODIMERS AND HMW1(Glc) EXPRESSION</b> .....	54
<b>5. DEXTRAN-BASED TENTACLES</b> .....	64
<b>5.1 INTRODUCTION</b> .....	64
<b>5.2 RESULTS AND DISCUSSION</b> .....	67
<b>5.2.1 Functionalization of dextran</b> .....	67
<b>5.2.2 Dextran-peptide conjugates</b> .....	72
<b>5.2.3 Applications to antibody detection and capture</b> .....	77
<b>5.3 CONCLUSIONS</b> .....	83
<b>6. S-ALKYLATION: TOWARD NEW GLYCOPEPTIDE CONJUGATES</b> ...87	
<b>6.1 INTRODUCTION</b> .....	87
<b>6.2 RESULTS AND DISCUSSION</b> .....	90
<b>6.2.1 Synthesis of relevant glycosyl building blocks</b> .....	90
<b>6.2.2 Alkylation of a model cysteinyl peptide</b> .....	96

6.3	CONCLUSIONS.....	100
7.	EXPERIMENTAL PART.....	103
7.1	ORGANIC SYNTHESIS .....	103
7.1.1	Materials .....	103
7.1.2	(2,3,4,6-Tetra-O-Acetyl)- $\beta$ -D-Glucopyranosyl-Azide (I).....	103
7.1.3	(2,3,4,6-Tetra-O-Acetyl)- $\beta$ -D-Glucopyranosyl-Amine (II) .....	104
7.1.4	N-Fmoc-L-Aspartic Anhydride (III) .....	104
7.1.5	(2,3,4,6-Tetra-O-Acetyl-N-[N-Fmoc-L-aspart-4-oyl])- $\beta$ -D-Glucopyranosylamine (Fmoc-Asn[Glc(OAc) <sub>4</sub> ]-OH) .....	105
7.1.6	4-azido-N-(3-hydroxypropyl)-benzamide (IV) .....	106
7.1.7	Synthesis of 4-(4-(hydroxymethyl)-1H-1,2,3-triazol-1-yl)-N-(3-hydroxypropyl)benzamide (V) .....	107
7.1.8	(2,3,4,6-Tetra-O-Acetyl)- $\beta$ -D-Glucopyranosyl- (2'-Br)-N-Acetyl-Amine (VI).....	107
7.1.9	(2,3,4,6-Tetra-O-Acetyl)- $\beta$ -D-Galactopyranosyl-Azide (VII) .....	108
7.1.10	(2,3,4,6-Tetra-O-Acetyl)- $\beta$ -D-Galactopyranosyl-Amine (VIII)....	109
7.1.11	(2,3,4,6-Tetra-O-Acetyl)- $\beta$ -D-Galactopyranosyl- (2'-Br)-N-Acetyl-Amine (IX).....	109
7.1.12	2,6-di-O-benzoyl- $\beta$ -D-Galactopyranosyl-Azide (X) .....	110
7.1.13	3-O-Benzyl- $\alpha/\beta$ -D-Glucose (XI) .....	111
7.1.14	1,2,4-Tri-O-Benzoyl-3-O-Benzyl-6-O- <i>tert</i> -butyldimethylsilyl- $\alpha/\beta$ -D-Glucopyranose (XII).....	112
7.1.15	Methyl 1,2,4-Tri-O-Benzoyl-3-O-Benzyl- $\alpha/\beta$ -D-Glucopyranosuronate (XIII).....	112
7.1.16	Methyl 1,2,4-Tri-O-Benzoyl-3-O-Levulinyl- $\alpha/\beta$ -D-Glucopyranosuronate (XIV) .....	113
7.1.17	Methyl 1-bromo-2,4-di-O-Benzoyl-3-O-Levulinyl- $\alpha$ -D-Glucopyranosuronate (XV) .....	114
7.2	PEPTIDE SYNTHESIS .....	114
7.2.1	General procedure for peptides synthesis.....	114
7.2.2	Synthesis of Peptides 1-4 .....	116
7.2.3	Synthesis of Peptides 5-6 .....	117
7.2.4	Synthesis of Peptide 7 .....	117
7.2.5	Synthesis of Peptide 8 .....	118
7.2.6	Synthesis of Peptide 9-10 .....	118
7.2.7	Synthesis of Peptide 11-12.....	119
7.3	PEPTIDE CONFORMATIONAL ANALYSIS .....	120

7.3.1	NMR of peptides 1-4 .....	120
7.3.2	CD of peptides 1-4.....	126
7.4	DEXTRAN CONJUGATES.....	127
7.4.1	Synthesis of Dex40-GP.....	127
7.4.2	Synthesis of Dex40-Pept6 .....	127
7.4.3	NMR analysis of dextran-based conjugates .....	128
7.5	PROTEIN EXPRESSION .....	136
7.5.1	General procedure for protein expression.....	136
7.5.2	General procedure for protein purification .....	137
7.5.3	General procedure for SDS-PAGE .....	138
7.6	INDIRECT COMPETITIVE ELISA .....	139
7.7	INDIRECT SP-ELISA .....	140
7.8	IMMUNOAFFINITY COLUMN .....	142
8.	ABBREVIATIONS.....	144
9.	REFERENCES .....	146

# **1. INTRODUCTION**

## **1.1 AUTOIMMUNE DISEASES: a chemist's point of view**

### **1.1.1 Immune system and autoimmune diseases**

The immune system is the host system composed of physical, biological, and chemical barriers and mechanisms whose purpose is to avoid the penetration and the subsequent damages caused by several types of exogenous molecules, e.g. pathogens such as viruses or bacteria. An immune response is based on a complex interplay of highly specialized components and can be humoral (antibody-mediated), cell-mediated, or both. In many species such as humans, there are two subtypes of immune system: innate and adaptive.

The innate immune system consists of the first response mechanisms against most exogenous factors, thanks to its ability to recognize common non-self molecular profiles, e.g., carbohydrates moieties belonging to the cell wall of pathogens. Conversely, the adaptive immune system, which is found only in vertebrates, is characterized by a slower but much higher specific response against a given pathogen. It can produce an immunological memory after the first contact with the exogenous agent in order to increase its protective efficacy in case of subsequent encounters with the same pathogen. The clonal expansion of a specific subset of leukocytes, B and T lymphocytes (cell-mediated response), produced by stem cells in the bone marrow, composes the adaptive immune system. T lymphocytes (or T cells) expose T cell receptors (TCRs) on their surface and can recognize the pathogen only with the assistance of the antigen-presenting cells (APC), such as dendritic cells. T cells carry out their action by producing and releasing cytokines (CD4+ T cells) in order to enhance the immune response or by destroying infected cells (CD8+ T cells). Differently, B lymphocytes (or B cells) act through B cell receptors (BCRs) that are able to recognize the pathogen by detecting linear or conformational portions of it, called epitopes. They produce antibodies that specifically interact with pathogenic antigens, subsequently remaining in the immune system for several years leading to immunological memory.

In a healthy organism, the adaptive immune system must correctly distinguish self-from non-self. The adaptive immune system is partially supported in this recognition by the Major Histocompatibility Complex (MHC), or Human Leukocyte Antigen

(HLA) in humans, a set of polymorphic genes which encode for several cell-surface proteins or peptides. These molecules can be classified as MHC class I and II. MHC class I molecules are short peptides (8-9 amino acid residues) exposed on the cell surface and presented in order to indicate a possible intracellular infection to CD8<sup>+</sup> T cells and leading eventually to the apoptosis. MHC class II molecules, expressed by APCs, are longer than those of class I (15-30 amino acid residues). After a proteolytic cut of a non-self antigen, its epitope fragment is presented by these molecules to the CD4<sup>+</sup> T cells in order to activate B cells and produce specific antibodies.

On the other hand, B cells are capable of distinguishing self- and non-self without the assistance of MHC proteins, thanks to the specificity of their BCRs. The interaction between B cells and the corresponding non-self antigen plus the additional signal from CD4<sup>+</sup> leads to the production of specific antibodies<sup>[1,2]</sup>.

When the immune system loses the ability to correctly recognize self, unharmed molecules, leading to the autoreactivity against self-antigens, the outcome is the onset of an autoimmune condition. The production of autoantibodies and/or the appearance of activated lymphoid cells against the organism itself can cause critical damages to organs and tissues, ultimately promoting the dysregulation of numerous endogenous biochemical processes.

Autoimmune diseases are estimated to affect about 3-5% of the world population, with high variability in gender, age, and geographic distribution<sup>[3]</sup>. Autoimmune thyroid disease and type I diabetes are the most common of these conditions, but other critically disabling autoimmune diseases include multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis, demyelinating monoclonal gammopathies, and many more. They can be organ-specific or systemic, but the molecular mechanisms and the triggering causes underpinning the breaking of immune tolerance remain unclear for most autoimmune pathologies. Generally, autoimmune diseases are acknowledged as multifactorial diseases, including both genetic predisposition and environmental factors as possibly involved in their pathogenesis (Figure 1) <sup>[3]</sup>.

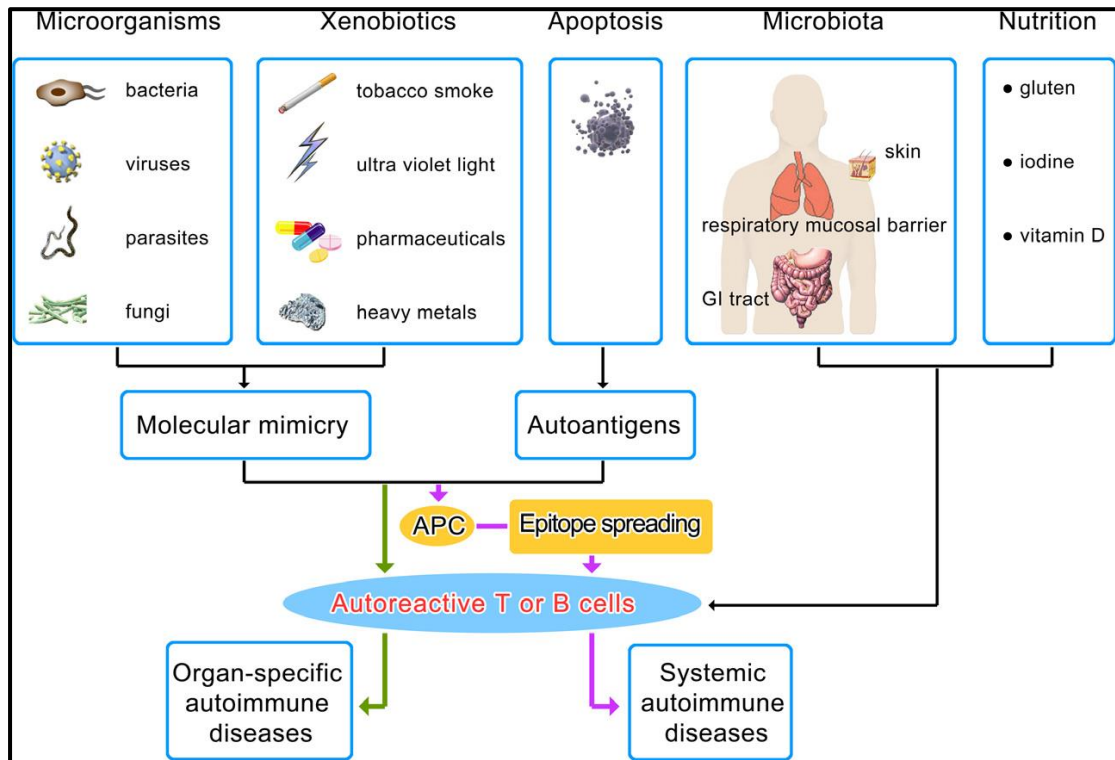


Figure 1 - Environmental factors in autoimmunity. Multiple environmental factors have been implicated in the development of autoimmune diseases. ‘Molecular mimicry’ is the most common mechanism that activates autoreactive T and B cells. ‘Epitope spreading’ is a mechanism that results in the generation of multiple neo-epitopes. In addition, by modulating innate and adaptive immunity, microbiota and nutrition (e.g. vitamin D, iodine and gluten) may also contribute to loss of tolerance.<sup>[3]</sup>

### 1.1.2 Infections and clinical relevance of molecular mimicry

Despite the tremendous ongoing research efforts in the field of immune-mediated diseases, little is known about the precise molecular mechanisms triggering the disruption of immune tolerance and misrecognition of “self” epitopes as foreign antigens<sup>[3,4]</sup>. Nonetheless, there is increasing evidence about the crucial role played by glycans and glycoproteins in the pathophysiology of several antibody-mediated diseases, and an association between alterations of the serum glycome and autoimmunity has been proposed<sup>[5]</sup>. More and more studies show the relevance of glycosylation to pathogen recognition, to the immune system control and immune cell homeostasis. Equally critical and maybe even connected is the glycans-antibodies interaction in the development of immune-diseases, including autoimmunity and cancer [6,7].

The occurrence of carbohydrate-protein interactions as the first cell surface flags of cell-cell communication in case of infections and many other processes involving the

immune response plays a crucial role in the discrimination between “self” and “non-self” recognition<sup>[8]</sup>.

Genetic predisposition is not sufficient alone to elicit the complex mechanisms at the base of many pathologies. It is now accepted that environmental factors such as infections have been implicated in the onset and/or promotion of aberrant immune response<sup>[9] [10] [11]</sup>.

Among the various biological processes that could break the physiological balance, i.e., tolerance, resulting in autoimmunity, the most relevant one is the so-called ‘molecular mimicry’ effect. Many eukaryotic and prokaryotic pathogens express glycoconjugates on their surface or secreted products. Glycans expressed by pathogens can include terminal glycostructures, that mimicking human biomolecules trigger an immune response to the pathogen leading to cross-reactive antibodies. The best-studied examples of pathogen-induced autoimmune conditions are Guillain-Barré syndrome (GBS) and its variant Miller Fisher syndrome <sup>[12]</sup>.

GBS patients present antibodies against gangliosides, i.e., sialic acid-containing glycolipids and major constituents of the nerve cell membrane. The molecular mimics are glycans expressed on lipooligosaccharides (LOS) of preceding infectious organisms, in particular *Campylobacter jejuni*, that can induce an antibody response to these carbohydrate antigens<sup>[13] [14]</sup>. The specificity of the anti-ganglioside autoantibodies is closely related to the nature of the preceding infections in GBS. Specific anti-ganglioside GM1, GM1b GD1a, and GalNAc-GD1a antibodies are related to a GBS form affecting only motor nerves, whereas antibodies against ganglioside GQ1b are associated with the Miller Fisher syndrome. The Miller Fisher syndrome is a subform of GBS affecting predominantly the nerves that innervate muscles governing eye movements. Other identified ganglioside mimetics are sialyllactose derivatives in non-typeable *Haemophilus influenzae* and *Haemophilus influenzae* b-type<sup>[15] [16]</sup>.

Despite many epidemiological and animal studies supporting the mimicry hypothesis<sup>[17]</sup>, the role of autoantibodies and the connection with exogenous infective agents remains only speculative for most autoimmune diseases. For some of them, such as systemic lupus erythematosus (SLE) and rheumatoid arthritis (RA), self-antigens recognized by the host immune system have been identified but the etiology has not been clarified<sup>[4] [18]</sup>. However, the crucial role of infective agents in triggering autoimmunity is highly suspected, yet the lack of a cross-reactive exogenous antigen hampers the assessment of a precise response pathway<sup>[19] [20]</sup>.

The initial host response to a viral or microbial infection can alternatively produce a cross-reaction with an appropriate host-antigen, leading to a “molecular mimicry” mechanism that may degenerate in an autoimmune disease. Molecular mimicry was firstly postulated because of the evidence of existing cross-reactivities between exogenous elements and host ‘self’ determinants while generating monoclonal antibodies towards viral proteins<sup>[21]</sup>. Other epidemiologic and clinical evidence, such as infections often preceding autoimmune diseases or homozygote twins seldom suffering from the same autoimmune disorders, also contributed to developing the hypothesis of a mimicry mechanism. In fact, many pieces of biological evidence support the existence of a mismatch between exogenous pathogen antigens and “self” cellular components. Therefore, molecular mimicry is highly contemplated as the primary cause for many autoimmune pathologies and is a proof-of-concept to uncover its etiologic agents<sup>[22]</sup>.

### **1.1.3 Autoantibodies as biomarkers**

Autoantibodies with pathogenic potential targeting extracellular protein domains have been associated with a vast plethora of diseases of both the central nervous system (CNS)<sup>[23]</sup> and the peripheral one (PNS)<sup>[15]</sup>. Their isolation and characterization are crucial to understanding their role in a specific disease and would provide a fundamental achievement both for diagnostics and therapeutics.

In general, antibodies (Abs) or immunoglobulins (Igs), that are secreted by B lymphocytes, have two distinct functions: one involves the specific non-covalent binding to their target antigens (Ags); the other is to elicit an immune response against the bound Ag by recruiting other cells and molecules<sup>[24]</sup>.

From a structural point of view, an Ig molecule is a quaternary, Y-shaped protein composed of two identical light chains (L,  $\approx$  25 kDa) coupled to two identical heavy chains (H,  $\approx$  55 kDa) through disulfide bridges. Each chain has a variable (V) region that interacts with the antigen and a constant (C) region. The two N-terminal portions of the Ab are called antigen-binding fragments (Fab), and they are composed of one constant and one variable domain of each of the heavy and the light chain (VL, CL, VH1, CH1). The C-terminal fraction of the Ig molecule is the crystallizable fragment (Fc), which is composed of two or three (depending on Ab isotype) identical polypeptide chains (Figure 2).

Functionally, the CH domains of Fc region confer effector properties such as complement binding, half-life length, interactions with Fc receptors, and the class, or isotype, of the Ig. In contrast, the V domains provide specificity to the Ig molecule by functioning as the direct contact between the Ig and its Ags. In fact, the terminal part of Fab region contains the six hypervariable loops commonly termed complementarity determining regions (CDRs), that are often thought to contain the paratope, i.e., the portion of the Ab which specifically interacts with the epitope of the corresponding antigen, even though recent studies suggested that this clear functional separation between the V and C regions may be an oversimplification<sup>[24,25]</sup>.

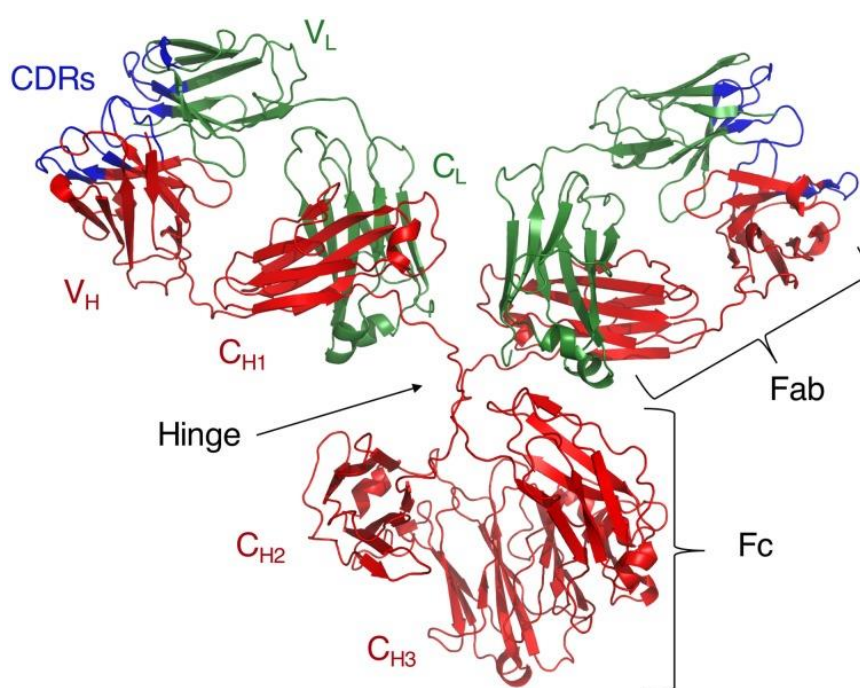


Figure 2 - Crystal structure of intact IgG<sub>1</sub> (PDB code: 1IGY)<sup>[25]</sup>

The reversible Ag-Ab interaction is non-covalent, and when the equilibrium is reached, the rate of [Ab-Ag] complex formation is equal to the rate of dissociation into its components [Ab] + [Ag]. The ratio of the reaction rate constants  $k_{on}/k_{off}$  can be used to define an equilibrium or affinity constant ( $K_A = 1/K_D$ ), between the antibody and its antigen. The  $K_D$  value relates to the concentration of antibody and so the lower the  $K_D$  value (lower concentration) and thus the higher the affinity of the antibody. When describing the strength of the antigen-antibody complex, affinity and avidity are always mentioned.

Affinity is the strength of binding between an epitope and an antibody's antigen binding site. It is typically measured and reported by the equilibrium dissociation constant ( $K_D$ ),

which is used to evaluate and rank order strengths of bimolecular interactions. The binding of an antibody to its antigen is a reversible process, and the rate of the binding reaction is proportional to the concentrations of the reactants.  $K_D$  and affinity are inversely related.

Avidity gives a measure of the overall strength of an antibody-antigen complex and depends on the affinity of the antibody for the epitope, the valency of both the antibody and antigen and the structural arrangement of the parts that interact.

All antibodies are multivalent, and the greater an immunoglobulin's valency (number of antigen binding sites), the greater the amount of antigen it can bind. Similarly, antigens can demonstrate multivalency because they can bind to more than one antibody. Multimeric interactions between an antibody and an antigen help their stabilization.

A favorable structural arrangement of antibody and antigen can also lead to a more stable antibody-antigen complex. Many immunoassays like Western blotting and ELISA exploit this principle, and therefore they are routinely used in autoimmune research.



*Figure 3 - An immobilized antigen (a high local concentration of available epitopes) provides more opportunity for the antibody-antigen complex to form than free antigen in solution over the same time period. Once the first antigen binding arm of an antibody attaches to an antigen on a solid support, the chances of a bivalent interaction are greatly improved<sup>[26]</sup>.*

Since most Ags are highly complex and may present numerous epitopes, Abs are also classified as polyclonal and monoclonal. Whereas polyclonal Abs are produced by different B cell lines and target the same antigen by interacting with different epitopes, monoclonal Abs are secreted by a single B lymphocyte clone that reacts against a unique epitope of a specific antigen<sup>[27]</sup>.

Additionally, immunoglobulins are produced against intact antigens in soluble form, and thus preferentially identify surface epitopes that can represent conformational structures that are non-contiguous in the antigen's primary sequence. This ability to identify components of the antigen independently of the rest makes it possible for the B cell to discriminate between two closely related antigens, each of which can be viewed as a collection of epitopes. It also permits the same antibody to bind different antigens that share equivalent or similar epitopes, a phenomenon that is named "cross-reactivity"<sup>[28]</sup>.

The overall abundance of Igs in the serum of most individuals is approximately 20 mg/mL, but the presence of each isotype varies greatly depending on several conditions. In humans, the antibody isotypes consist of four immunoglobulin G (IgG) subclasses (IgG1, IgG2, IgG3, and IgG4), two IgA subclasses (IgA1, IgA2), IgM, IgD, and IgE. All the different isotypes are assembled with the similar Ig basic unit consisting of two heavy chains and two light chains, but both IgM and IgA can form multimers. IgM molecules are usually secreted as pentamers (more rarely as hexamers), in which single Ig units are linked to each other by disulfide bonds in the CH4 region, while an additional polypeptide chain (J-chain), which is bound to two of the monomers through disulfide bonds, facilitates secretion at mucosal surfaces. IgA is normally present as a dimer, whereas IgG subtypes and IgE differ only slightly in their conformation (Figure 4). Noteworthy, the constant domains of the H chain can be switched to allow altered effector function while maintaining antigen specificity.






	IgG	IgM	IgA	IgE	IgD
<b>Molecular weight</b>	150,000	900,000	320,000 (secretory)		
<b>Heavy chain: Type MW</b>	$\gamma$ 53,000	$\mu$ 65,000	$\alpha$ 55,000	$\epsilon$ 73,000	$\delta$ 70,000
<b>Concentration in serum (approximate)</b>	10-16mg/mL	0.5-2mg/mL	1-4mg/mL	0.00001-0.0004mg/mL	0-0.4mg/mL
<b>Percent of total IgG</b>	80	6	13	0.002	0.2
<b>Carbohydrate (approximate)</b>	3%	12%	10%	12%	13%
<b>Distribution</b>	Intravascular and extravascular	Mostly intravascular	Intravascular and secretions	Basophils and mast cells in saliva and nasal secretions	Lymphocyte surface
<b>Function</b>	Secondary response	Primary response	Protect mucous membranes	Protect against parasites	Unknown
<b>Structure</b>					

Figure 4 – Antibody isotypes and their important properties.

Autoantibodies, especially IgG and IgM types, are found in many autoimmune disorders, including diseases affecting the peripheral nervous system (e.g. Guillain-Barré syndrome and other peripheral neuropathies), the central nervous system (e.g.

neuromyelitis optica, multiple sclerosis and limbic encephalitis) and the neuromuscular junction (e.g. myasthenia gravis), as well as in neuropsychiatric disorders and in paraneoplastic syndromes. Some of these antibodies are likely to play an important pathogenic role in disease development, whereas in some cases others might be merely useful biomarkers for the disease, without direct pathological relevance<sup>[29]</sup>. The antibody isotype and the target antigen against which the antibody is directed are essential aspects of an autoantibody and they must be thoroughly unveiled while considering its pathogenicity. Pathogenic autoantibodies may act through different mechanisms, including antibody-mediated cell lysis, opsonization of target proteins for attack by macrophages, crosslinking of Fc receptors, blocking or destroying receptors involved in neurotransmission or cellular homeostasis, and blocking of repair mechanisms such as remyelination.

Functionally, IgM is the very first Ab class produced by the immune system following the infection.

Generally, while monomeric IgM molecules have low affinity due to their immaturity, high avidity can be attained by means of multimeric interactions between the pentameric secreted antibody and the antigen, especially if that antigen contains multiple repeating epitopes itself. IgM act by opsonizing (coating) antigen for destruction and anchoring complement. The pentameric nature of the antibody renders it very efficient in this process. IgM antibodies are associated with a primary immune response and are frequently used to diagnose acute exposure to an immunogen or pathogen. IgM antibodies tend to be more poly-reactive than other isotypes, which allows IgM-bearing B cells to respond quickly to a variety of antigens. Some of these relatively low-affinity IgM antibodies not only participate as a first line of defense but are also thought to play a protective regulatory role in several diseases<sup>[30,31]</sup>. In fact, some studies suggested that despite the greatly elevated levels of autoreactive IgM in various autoimmune diseases, they are questionably responsible for pathogenesis. Rather, this effect is hinted to reflect a compensatory mechanism by which the increased IgM might diminish systemic chronic inflammation, accumulation of autoantigens, and/or the potential for increased exposure of antigen-reactive B and T cells to self-antigens. According to this hypothesis, the presence of autoantibodies of the IgM isotype seem to be protective, while antibodies with similar specificity, but of the IgG isotype, are often pathogenic<sup>[32,33]</sup>.

In other words, even though IgG and IgM autoantibodies in patients with autoimmune disease have similar specificities and can bind to the same antigens, mostly the self-reactive IgG antibody is considered to be pathogenic. This conclusion is in accordance with the thesis that the Fc region, rather than the antigen-specificity of Fab region, is the real responsible of the biological functions of Abs, probably through binding to specific Fc receptors. Even though it is still speculative, pathogenic autoantibodies tend to be drawn from the somatically mutated, high affinity IgG population.

During the secondary immune response, the most stable IgG become the major Igs class produced by the immune system, about 75% of the circulating Igs, becoming the most abundant class of antibodies in serum. It has the longest serum half-life of all immunoglobulin isotypes. It is also the most extensively studied class of immunoglobulins. Based on structural, antigenic and functional differences in the constant region of the heavy chain, particularly CH1 and CH3, we can distinguish four IgG subclasses (IgG1, IgG2, IgG3 and IgG4). IgG1 and IgG3 antibodies are generally induced in response to protein antigens whereas IgG2 and IgG4 are associated with polysaccharide antigens.

Pathogenic or not, the characterization of autoantibodies as biomarkers of an autoimmune condition is in any case a major goal toward which many efforts still need to be done in order to unveil the biomolecular mechanisms and to obtain diagnostic and therapeutic improvements.

When investigating these autoantibodies, it is important to choose appropriate methods to detect them, to check the isotype of the antibody, as well as the specificity, as that will determine the functional role that the antibody could play, and to think about the normal function of the target antigen, and consider how the antibody could cause the clinical symptoms and signs shown by the patients<sup>[29]</sup>.

#### **1.1.4 Peptide-based approaches in autoimmune diseases**

An epitope is defined as the portion of the antigen interacting with the paratope of the antibody. This definition is strictly operational<sup>[34]</sup> and includes the concept of cross-reacting epitopes, i.e., two different molecules recognized by the same antibody. At the immunological level, the concept of “molecular mimicry”, underpinned by this definition, has been used to explain the break of immunological tolerance leading to

autoimmune disorders. A molecule, such as a peptide, which mimics the structure of an epitope is called “mimotope”. This term was coined by H.M. Geysen in 1986 referring to peptide sequences causing an antibody response similar to the one elicited by the discontinuous protein antigenic determinant<sup>[35]</sup>.

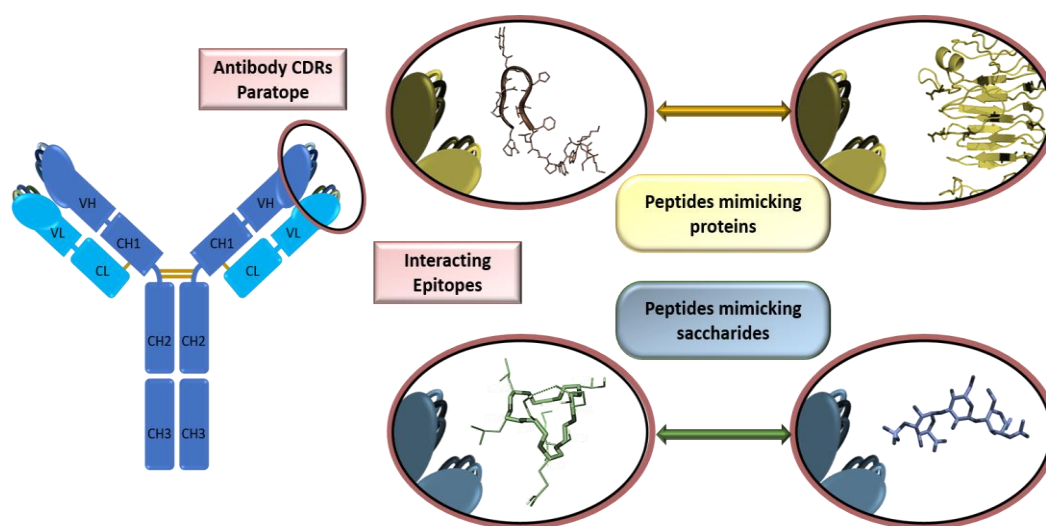


Figure 5 – Cartoon representing antibody paratopes and the concept of cross-reactivity. Two structurally unrelated molecules can be recognized by the same antibody, as in the case of the synthetic probe CSF114(Glc) cross-reacting with bacterial hyperglucosylated HMW1 adhesin (up) and a cyclic octapeptide that binds the murine monoclonal antibody specific for HNK-1 trisaccharide (down).<sup>[36]</sup>

Peptides represent unique chemical tools because of the amino acid diversity and of the different type of synthetic modifications, which can affect the chemical structure, the spatial conformation and the biological activity. Thanks to their ability to mimic proteins, glycans and/or glyconjugates, several “peptide mimetics” have been used as synthetic immunological probes to mimic the native antigens, especially when post-translational modifications are supposed to trigger the autoimmune response<sup>[37]</sup>. In this perspective, the development of synthetic peptides interacting with high affinity with autoantibodies, biomarkers of autoimmune diseases, can help the set up of diagnostic/prognostic peptide-based immunoassays particularly useful to guide therapeutic treatments<sup>[38]</sup>.

Additionally, peptide-based immunotherapy has been evaluated with success in several appropriate experimental animal models, although very few peptides are currently evaluated in clinical trials for the treatment of human autoimmune diseases<sup>[39]</sup>.

Glycans are now emerging as preferred epitopes being involved in many processes such as cell-cell communication and pathogenic recognition. Because of the intrinsic difficulty in obtaining pure carbohydrate derivatives both by isolation from a natural source and via chemical, and/or enzymatic synthesis, alternative strategies to experimentally reproduce and investigate glycan-protein interactions are challenging<sup>[36]</sup>.

The exploitation of peptides and glycopeptides reproducing antigen-antibody interactions holds the great potential to answer many unresolved questions in chemical immunology. Compared to proteins and glycans, peptides are more accessible to produce and to modify, i.e., efficiently managing their molecular structure to fine-tuning biological activity and stability. Rational design of mimetic molecules as antibody binders can include the incorporation of non-proteinogenic amino acids and altered backbones, providing an expanding list of different peptidomimetics, such as peptoids,  $\beta$ -peptides, retro-inverso peptides, etc.<sup>[40-42]</sup>.

These peptide mimetics may be very precious at different levels:

- 1) to investigate the molecular nature of the antibody-antigen interaction, to reveal the actual epitope and to understand the underlying mechanisms of the immune tolerance disruption;
- 2) as probes for diagnostic applications and to monitor the disease activity;
- 3) to produce peptide drugs either blocking glycan-protein interactions or stimulating the immune system for vaccine purposes (internal use);
- 4) as molecular “baits” to fish out specific autoantibodies from patients’ sera (external use for immunoaffinity columns).

### **1.1.5 Isolation of antibodies and plasmapheresis**

Unravelling the multi-layered network architecture of the immune system in diseases, by elucidating the intricate communication between immune cells and the cryptic interactions between antigen epitopes and antibodies, is the ultimate ambition for immunologists. Despite the tremendous advances that have been achieved in the last decades, the goal of decoding the mechanisms and the causes underlying the disruption of immune tolerance remains unfulfilled for most of autoimmune pathologies. Therefore, it is not surprising that basically none of them has a cure. However, for many diseases, treatments based on the control of the overactive immune response and the

reduction of pain and inflammation are currently available. Along with physical therapy, immunosuppressants, corticosteroids, anti-inflammatory drugs, and cell or tissue transplantation can all contribute to alleviate the symptoms, but they do not alter the overall chronic course of diseases<sup>[43]</sup>. For example, immunosuppressants (i.e., cyclosporine A) inhibit the activity of the immune system by reducing the proliferation and function of cells associated with immune reactions and show partial efficacy in many autoimmune diseases. However, their therapeutic effects are dependent on chronic drug administration that can lead to systemic immune suppression, with the potential risk of development of cancer and infections<sup>[44]</sup>.

An alternative that could be able to modify the outcome of the diseases by restoring self-tolerance toward autoantigens may be found in antigen-specific immunotherapy (ASI). With a similar approach to vaccines, ASI is based on the introduction of low levels of stimulatory autoantigens to antigen presenting cells (APCs) in order to correct the immune responses<sup>[45,46]</sup>. The great advantage of this intervention is specificity, since ASI selectively targets disease-relevant T cells, acting directly through T cell receptor (TCR) on effector T cells and/or via regulatory T cells that secrete anti-inflammatory cytokines, while leaving the normal immune system intact. Many initial experimental approaches, such as oral antigen administration, particulate autoantigen delivery, altered peptide ligands, and dose escalating immunotherapy, have shown efficacy in preclinical models, but the results were not as encouraging in human clinical trials. However, the modification of autoantigen through functional conjugation to antibodies, polymers, nanoparticles (NPs), or small molecules have already been successfully applied in vaccines, hence it may hold promising potential for clinical translation also for autoimmune diseases. Therefore, there is an increasing interest in developing bioconjugate strategies to incorporate additional immunomodulatory functions into autoantigens, with the final aim to induce autoantigen-specific tolerance even in humans<sup>[44]</sup>.

Nevertheless, skepticisms about the disease specificity, mechanistic underpinnings, developability and translational potential in autoimmune conditions for most of these approaches persist<sup>[47]</sup>. So, to date, no FDA-approved ASI is available for treating patients with autoimmune diseases and induction of antigen-specific tolerance to dominant immune responses driving autoimmunity remains an unmet challenge.

Considering antibody-mediated autoimmune diseases, another striking triumph would be the selective removal of pathogenic autoantibodies. The tremendous flaw of current

therapies that aim to deplete pathogenic autoantibodies, decrease expanded autoantibody-producing B-cell clones, and/or interfere with antibody-effector mechanisms, is the lack of efficiency and selectivity. Therefore, so far very few satisfactory treatments are available. A more efficient and safer therapy might be achieved with antigen-specific agents that selectively target the autoantibodies. In general, there is a strong need for personalized, new disease-specific therapy to avoid nonspecific immunosuppression. An effective demonstration of autoantibody-specific treatments could be achieved by the preparation of specific autoantigens-grafted columns for plasmapheresis devices.

Plasmapheresis or therapeutic apheresis (TA) is a technique consisting in separating the plasma from the cellular part of the blood and then either replacing or cleaning it. For therapeutic purposes, plasma centrifugation or membrane-based plasma filtration are used in a variety of diseases to remove plasma components such as antibodies, proteins, cryoprecipitates and circulating immune complexes<sup>[48]</sup>.

Therapeutic plasma exchange (TPE) surely represents a valid treatment option in patients with systemic autoimmune diseases because most of their clinical manifestations are related to the presence of antibodies or immune complex deposition<sup>[49]</sup>. More and more retrospective studies assess that TPE is safe and effective in patients with severe manifestations of several autoimmune diseases, such as myasthenia gravis, neuromyelitis optica, Guillain-Barré variants<sup>[50,51]</sup>. However, this procedure is usually nonselective, leading to the removal of all the key plasma components that are discarded and replaced with some combination of plasma substitutes such as albumin, donor plasma or colloids. Moreover, the administration of these compounds could expose patients to potential side effects.

Aiming to selectively remove any pathogenic substances, immunoadsorption (IAS) therapy was suggested as an alternative to plasma exchange for removing pathogenic substances from the plasma of patients with autoimmune diseases refractory to conventional treatments<sup>[52]</sup>. Unlike TPE, IAS is a blood-purification technique that enables the selective removal of immunoglobulins from separated plasma through high-affinity adsorbents, without removing plasma proteins such as albumin and clotting factors. IAS is currently used for treatment of a large variety of antibody-mediated or immunological diseases (e.g., humoral transplant rejection, lupus nephritis, multiple sclerosis, Guillain-Barré syndromes)<sup>[53-55]</sup>. Adsorption generally is based on columns

containing nanoparticles linked to the high affinity absorber, but also alternative adsorbents such as membranes, monoliths and cryogels have been developed<sup>[56,57]</sup>. During the last years, tremendous efforts have been made to develop more and more selective apheresis procedures, targeting a specific molecule, antibody, or cellular element, thereby preserving plasma proteins and avoiding replacement solutions. In particular, columns containing immobilized antigens can be designed to remove only the pathogenic autoantibodies that are reactive toward that specific antigen, leaving all other immunoglobulins and plasma constituents untouched (Figure 6). The greatest limitation of these therapies is that many of the developed plasma processing technologies are still investigational, costly and not currently approved for clinical use, mainly because of the uncertainty of shelf life and stability of immunoabsorption columns<sup>[48]</sup>.

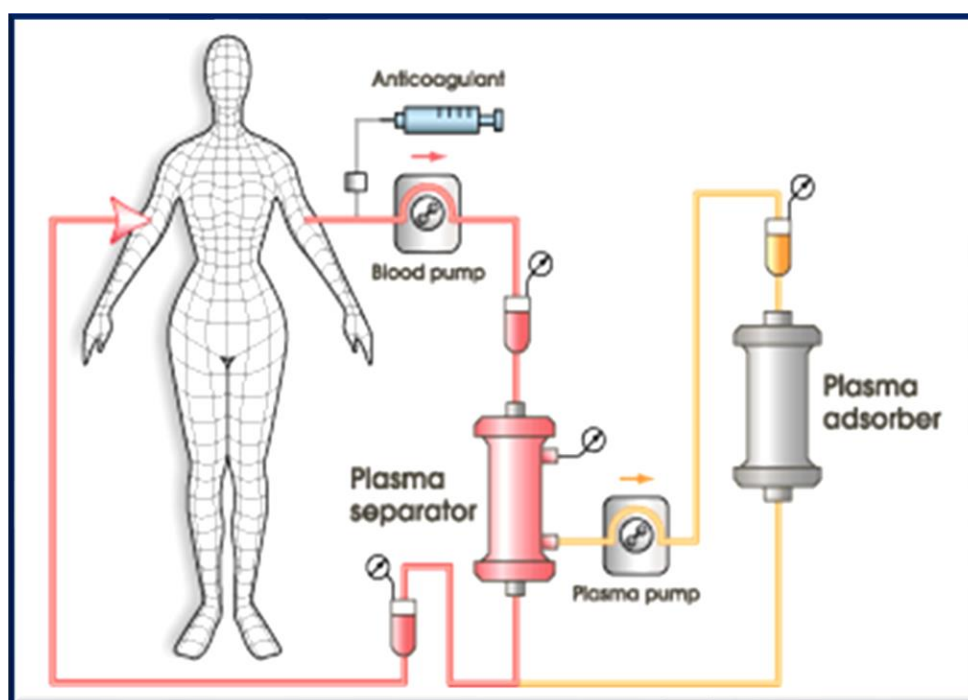


Figure 6 – Schematic representation of antigen-specific immunoabsorption therapy, where the patient's plasma separated with a plasma separator flows into a plasma adsorption column which selectively removes harmful substances.<sup>[58]</sup>

However, the antigen-specific immunoabsorption would provide the selective depletion of only the pathogenic autoantibodies, reducing the possibility of side effects while maximizing the benefit. Therefore, developing such a system, i.e. antigen-grafted nanoparticles, membranes or gels for the stationary phase of high affinity columns, could be of great impact for the treatment of antibody-mediated autoimmune diseases.

Of course, an essential feature is the knowledge of the antigen recognized by autoantibodies in a certain autoimmune condition, which is unfortunately not granted yet for the vast majority of them. Nevertheless, few studies going in this direction reported very promising initial results of this proof-of-concept applied to autoimmune diseases such as Miller Fisher syndrome and other GBS variants<sup>[56,59]</sup>, membranous nephropathy<sup>[60]</sup>, myasthenia gravis<sup>[61]</sup>.

## 1.2 MULTIPLE SCLEROSIS: a complicate puzzle

### 1.2.1 Overview of Multiple Sclerosis

Multiple sclerosis (MS) is a chronic inflammatory demyelinating and neurodegenerative disorder of the central nervous system (CNS). It is the most frequent cause of nontraumatic neurological disability among young adults in the developed countries. MS is an autoimmune disease, but indeed has a very complex pathophysiology, with a highly heterogeneous clinical presentation and course. According merely to phenomenological evidences, it is currently classified into three main types of clinical subsets: relapsing–remitting (RRMS), primary progressive (PPMS), and secondary progressive (SPMS). In the majority of cases (85–90%), MS presents with a relapsing-remitting course, characterized by discrete episodes of neurological dysfunction (relapses or exacerbations) separated by clinically quiescent periods (remissions). The frequency of relapses can vary widely among patients as well as during different periods in an individual patient’s disease course. The rate, severity, and symptoms of relapses are highly variable and unpredictable. At present no clinical features or biomarkers that are predictive of relapse rate have been identified. The signs and symptoms that occur during relapses are also diverse and unpredictable, since lesions can form at any site in the CNS, spanning the cerebrum, brainstem, cerebellum, optic nerves, and spinal cord. MS lesions are readily visualized in CNS white matter via magnetic resonance imaging (MRI). Symptomatic lesions generally occur in locations where nerve fibers converge to subserve a common function. During RRMS, relapses decrease in frequency over time and sometimes disappear completely. However, most of the times a gradual accumulation of disability leads to the *secondary progressive (SP) stage*, with different symptoms and signs that characterize neurological decline. The current management of SPMS involves alleviation of symptoms, optimization of residual functions, and prevention of complications.

PPMS is distinguished from SPMS by the absence of an antecedent RR phase, but the clinical features are the same. However, PPMS tends to appear later in age and without a distinction between male and females.

The hallmark of MS pathology is the focal demyelinated lesion, or “plaque,” with perivascular inflammatory infiltration and BBB breakdown. Inflammation, demyelination, remyelination, neurodegeneration and glial scar formation occur either focally or diffusely affecting both the white and grey matter in the brain and spinal cord.

These pathological features are present in both RRMS and SPMS, as well as in PPMS, although they vary over time both quantitatively and qualitatively between these three forms of MS and among individuals with the same form of the disease, thus contributing to the heterogeneity in phenotypic expression of the disease and response to therapies.<sup>[62]</sup>

### **1.2.2 Cause and pathogenesis**

Although the aetiology of MS unknown, evidence suggests that the disease may result from a complex interaction between the environmental factors, the genetic background that defines individual susceptibility, and the immunological and physiological settings of the individual. The majority of MS susceptibility loci map to regions containing genes implicated in immunological pathways, including human leukocyte antigen (HLA) class II molecules, the interleukin-2 (IL-2) receptor, and the IL-17 receptor. Relapse rates decline during the third trimester of pregnancy, in association with high serum levels of estriol. Environmental risk factors include low vitamin D levels, exposure to foreign pathogens in adulthood, cigarette smoking, and childhood obesity. Although MS is indeed a heterogeneous disease with disparate inflammation processes throughout its progression, autoimmune component, i.e. T- and B-cell responses, against CNS antigens remain a central feature in the current view of MS pathogenesis. Mechanisms of the pathophysiology of MS involve mainly three physiological compartments: 1) the peripheral blood, in which immune processes mainly take place; 2) the blood brain barrier (BBB), which breaks down to a point so that immune cells can pass into the CNS; and 3) the CNS, in which lesions mark acute sites of inflammation and neural damage, leading to the phenotypic displayed symptoms of disability.

More and more evidences indicate that B cells, plasma cells, and, in particular, antibodies contribute to development and progression of MS. Critical goals of the MS research community at present are to better define the cellular and molecular mechanisms that link neuroinflammation to demyelination and axonopathy and to elucidate the pathogenic pathways that underlie clinical progression. Myelin antigens are the prevalent putative targets studied in the field due to the often primarily demyelinating nature of inflammatory CNS lesions. Within candidate myelin antigens, myelin oligodendrocyte glycoprotein (MOG) is probably the one that is most

investigated due to its extracellular location on the outermost myelin lamellae, which makes it an exposed target accessible to an initial autoimmune attack against properly myelinated axons. To date, no conclusive understanding exists on possible CNS target antigens against which the antibody response may be raised in MS. Empirical support for a pathogenic role of antibodies derives from the therapeutic approach of plasma exchange, which was found to be beneficial in a subgroup of MS patients with severe therapy-refractory relapses and antibody deposition within inflammatory CNS lesions. These data suggest that, in a subgroup of patients, peripherally produced CNS-directed antibodies may indeed contribute to MS pathogenesis. Besides these findings in MS patients, several observations derived from experimental CNS autoimmunity suggest that B cells and CNS reactive antibodies contribute in a pathogenic manner.

In parallel to autoantibody studies, many researchers have reported on elevated antibody titers to a broad range of pathogens in MS patients. Along with the autoimmune hypothesis, it is assumed that infectious agents may play a significant role in the pathogenesis of MS. The antibody reactivities against various numbers of pathogens have been investigated in CSF and serum of MS patients.<sup>[63]</sup>

Beside the prevalence of specific antibodies in MS, it is crucial to demonstrate a biological consequence of antibodies and their participation in one effector function leading to demyelination.

### **1.2.3 Diagnosis**

The most widely used guideline for diagnosing MS is the *McDonald criteria*<sup>[64]</sup>, although there is a desperate lack of univocal clinical features or biomarkers. Biomarkers are measurable indicators of normal biological and pathogenic processes, or pharmacological responses to a therapeutic intervention. A good biomarker should be precise and reliable, able to distinguish between MS disease and control, can detect inflammatory activity, as well as the degree of neurodegeneration and demyelination/remyelination, in order to get a more accurate picture of the disease status. Unfortunately, the absence of validation is a common problem with biomarkers of complex diseases such as MS. This could be due to a bias in statistical analysis or a scarcity of available data, but it also indicates difficulties in performing clinical validation studies.

In MS, most studies search for biomarkers within the CSF with the view that this is more likely to reflect CNS disease. The presence of unique oligoclonal bands and/or elevated IgG index in the cerebrospinal fluid CSF, indicative of primary antibody production in the CNS, supports a diagnosis of MS, but those findings are observed in a wide range of different neuroinflammatory conditions. However, blood-based biomarkers are of great clinical value, because of the ease with which blood can be obtained in a minimally invasive manner.

Blood biomarkers may exist in MS if there is a systemic component of the disease, or if peripheral changes mimic central disease. Peripheral blood biomarkers can give important information regarding immune triggers of MS, as well as therapeutic efficacy of drugs administered. Additionally, blood has two properties that make it attractive for the search for biomarkers: 1) it is more easily accessible than other body tissues; and 2) the perfusion of blood through different organs and tissues can result in the addition of new proteins, or modification of existing proteins, which may vary according to specific physiological or pathological conditions. Thus, the blood can carry molecules derived from other tissues, reflecting the biological status of the body.

For decades, self-reactive antibodies have been implicated to participate in the pathogenesis of multiple sclerosis (MS). A plethora of investigations attempted to identify the target antigens which these antibodies may be raised against. Some findings implicated that antibodies within the inflamed CNS may eventually recognize components of the myelin sheath; however, no unequivocal evidence of such CNS reactivity has been established to date. Pathophysiological and clinical complexity of MS inevitably leads to a great variety of potential biomarkers specific for diagnostics, prediction of disease course and optimization of therapeutic responses.<sup>[65]</sup>

#### **1.2.4 Current therapies**

Despite the tremendous research in the field and the massive improvements achieved, there is no definitive cure for MS. Nevertheless, with the approval of interferon- $\beta$ 1b by the US Food and Drug Administration (FDA) in 1993 for the management of MS, a new era of treatment of this incurable disease began. Before that date, corticosteroids were the only class of drugs available to treat MS. Corticosteroids can often accelerate the rate of recovery from multiple sclerosis (MS) relapses, but there is little evidence that they impact the ultimate degree of recovery or the future clinical course.

Nowadays, several different disease-modifying agents (DMAs) for patients with MS with different mechanisms of action and side effect profiles exist. Most of these drugs act on the immune system and suppress immune cells so that auto-reactive immune cells will be unable to attack myelin sheaths of neurons (Table 1 and Table 2). They contributed significantly to reduce the annualized relapse rate and the frequency of gadolinium enhancing (acutely inflamed) MRI lesions. However, in spite of the clinical efficacy of these drugs, numerous and severe adverse effects have been reported in experimental studies and clinical trials. Therefore, the choice of DMAs must be customized on an individual basis, taking into account disease activity and risk tolerance. These medications include interferon- $\beta$ 1a intramuscular (Avonex), interferon- $\beta$ 1a subcutaneous (Rebif), interferon- $\beta$ 1b subcutaneous (Betaseron/Extavia), glatiramer acetate (Copaxone), natalizumab (Tysabri), fingolimod (Gilenya), teriflunomide (Aubagio), and mitoxantrone (Novantrone). Moreover, several clinical trials are being conducted to test various experimental agents in patients with MS, including alemtuzumab, dimethyl fumarate, laquinimod, rituximab, daclizumab, and cladribine<sup>[66-68]</sup>.

*Table 1 – Features of clinical usage of MS drugs (mechanism of action, form, dosage, route and frequency)<sup>[69,70]</sup>*

<b>Drug name</b>	<b>Usage</b>	<b>Mechanism</b>	<b>Dosage</b>
<b>Interferon beta 1a</b>	Indicated for first line treatment of RRMS	Effects on the endothelial cells of BBB	Injection, 30 $\mu$ g I.M. once a week or 44 $\mu$ g S.C. three times a week
<b>Interferon beta 1b</b>	Indicated for first line treatment of RRMS, SPMS	Effects on the endothelial cells of BBB	Injection, 250 $\mu$ g S.C. every other day
<b>Mitoxantrone</b>	Immunosuppressive agent for second line treatment of RRMS, SPMS and PPMS	Intercalating with DNA repair and inhibiting the topoisomerase II	Injection, 12 mg/m <sup>2</sup> I.V. every three months or 8 mg/m <sup>2</sup> I.V. every month
<b>Natalizumab</b>	Immunosuppressive agent for second line treatment of RRMS	Targeting the $\alpha$ 4-chain of $\alpha$ 4 $\beta$ 1 integrin	Injection, 300 mg I.V. every month
<b>Fingolimod</b>	Immunosuppressive agent for second line treatment of RRMS	Sphingosine 1-phosphate (S1P) receptor modulator	Cap.0.5 mg PO Qid, every day

<b>Glatiramer Acetate</b>	Immunomodulatory agent for first line treatment of RRMS	Binding to the major histocompatibility complex class II molecules	Injection, 20 mg S.C. every day
<b>Alemtuzumab</b>	Immunosuppressive agent for second or third line treatment of RRMS, SPMS and PPMS	Humanized monoclonal antibody of the IgG1 subclass against CD52	Injection, 12 mg/d I.V. for five days followed by 12 mg/d I.V. for three days one year after the first course
<b>Dimethyl Fumarate</b>	Immunomodulatory agent for first line treatment of RRMS	It's unclear by inducing lymphocytopenia	Cap. Starting dose: 120 mg Bid for 7 days. P.O maintenance dose 240 mg Bid. P.O
<b>Terifluonamide</b>	Immunomodulatory agent for first line treatment of RRMS	A noncompetitively and reversibly inhibition of mitochondrial enzyme dihydro-orotate dehydrogenase (DHODH)	Coated tab.14 mg P.O every day
<b>Laquinimod</b>	Immunomodulatory agent for first line treatment of RRMS and SPMS	Protection of neurons by decreases IL-17 levels and migration of leucocytes to CNS	Cap, 0.6-1.2 mg P.O once daily for every day
<b>Rituximab</b>	Immunosuppressive agent for second line treatment of RRMS	Depleting CD20+ B lymphocytes via cell mediated and complement-dependent cytotoxic effects	Infusion, 4×375 mg/m <sup>2</sup> or 1000 mg at week
<b>Daclizumab</b>	Immunosuppressive agent for second or third line treatment of RRMS and SPMS	Inhibiting of activation of lymphocytes (anti-CD25)	S.C injection, 5 ml once monthly
<b>Cladribine</b>	Immunomodulatory agent for first line treatment of RRMS	Depleting both CD4+ and CD8+ lymphocytes	0.07 mg/kg/day sc. 5 days/month or 0.1 mg/kg/day iv

			days for 4 months
--	--	--	-------------------

BBB: blood-brain barrier, RRMS: relapsing-remitting multiple sclerosis, PPMS: Primary-progressive multiple sclerosis, SPMS: secondary-progressive multiple sclerosis, SC: sub cutaneous, IV: intra venous, IM: intra muscular, Amp: ampule, Cap: capsule, CNS: central nervous system, P.O: per os, Qid: quater in die.

Table 2 – Most common adverse effects of MS drugs. For less common adverse effects see references <sup>[69,70]</sup>

Drug name	Adverse effects
<b>Interferon beta 1a</b>	Headache, flu-like symptoms, muscle aches, nausea, fever, asthenia, chills and diarrhea, injection site reactions, autoimmune phenomena
<b>Interferon beta 1b</b>	Leukopenia, flu-like symptoms, elevated hepatic transaminases, injection site reactions, headache, fever, malaise and myalgia, autoimmune phenomena
<b>Mitoxantrone</b>	Cardiotoxicity, malignancy and hepatotoxicity
<b>Natalizumab</b>	Headache, fatigue, urinary tract infection, lower respiratory infection, arthralgia, gastroenteritis, vaginitis, diarrhea, and hypersensitivity reactions
<b>Fingolimod</b>	Cardiac abnormalities such as dose dependent bradycardia, blood pressure effects, macular edema, some laboratory abnormalities in the liver, enzymes level in blood and possible infection risks
<b>Glatiramer Acetate</b>	Injection site reactions or symptoms of a systemic immediate post-injection reaction including flushing, chest pain, palpitations, anxiety, dyspnea, tachycardia, throat constriction and urticarial
<b>Alemtuzumab</b>	Infusion related symptoms, cytokine storm, increased risk of autoimmune diseases and increased risk of infections
<b>Dimethyl Fumarate</b>	Gastrointestinal symptoms and flushing
<b>Terifluonamide</b>	Gastrointestinal symptoms, hair thinning, neutropenia and lymphopenia
<b>Laquinimod</b>	Elevation of ALT, abdominal pain, back pain, cough, respiratory tract infections, headache, asthenic conditions, insomnia, nausea and vomiting, dizziness, arthralgia, and diarrhea
<b>Rituximab</b>	fever, chills, headache, urticarial and infections
<b>Daclizumab</b>	Transient elevation of liver enzymes, infections and cutaneous AEs
<b>Cladribine</b>	lymphopenia, leucopenia, upper respiratory infection, muscle weakness, hypertonia, purpura, rhinitis, ataxia, Injection site pain, injury, dizziness, and tremor

MS has been considered for a long time as a T cell-dominated disease, mainly because of the T cell-driven animal model of experimental autoimmune encephalomyelitis (EAE). Instead, thanks to the success of B cell-directed therapies and emerging

experimental evidence of antibody production involvement, the crucial role B cells is now recognized, allowing for more diverse and personalized treatment options for MS patients. In fact, newer generation DMAs either deplete lymphocytes, inhibit their expansion, or block their migration to the CNS<sup>[71]</sup>.

Based on the evidence that B-cell-derived antibodies against CNS autoantigen are important in the pathogenesis of certain MS subtypes, plasma exchange has also been proposed as a therapeutic strategy to remove or reduce self-reactive antibodies in treatment of acute MS relapses<sup>[63,72,73]</sup>. Interestingly, the respective MS subtype appears to be a predictor for the achievable benefit from plasma exchange. This technique was also successfully used to treat fulminant demyelinating attacks with MS patients unresponsive to steroids, and therapy-responsive patients were distinguished by histopathological analysis of antibody deposition and enrichment of B cells within the CNS<sup>[74]</sup>. Taken all together, the data evidence the urgent need of personalized treatments and the strong demand for intravital biomarkers distinguishing individual MS subtypes. Moreover, this finding strongly supports the hypothesis that in a subgroup of patients CNS-directed autoantibodies indeed contribute to MS pathogenesis. Finally, whereas the introduction of DMAs has represented a major advance in the treatment of MS, currently available drugs have global effects on lymphocytes, thereby compromising the whole immune system and increasing the risk of infection.

Developing tools or drugs that specifically target pathogenic and/or autoreactive molecules will be a critical goal for the future<sup>[63,75]</sup>.

### **1.2.5 N-glycosylation and HMW1 adhesin**

Nowadays, a large number of protein post-translational modifications have been identified and characterized in bacteria. Compared to eukaryotes, most post-translational protein modifications occur in a relatively low number of bacterial proteins, with substoichiometric levels of modification, hampering their structural and functional analysis<sup>[76]</sup>. In particular, bacterial protein glycosylation is emerging as a key player for many vital functions including virulence<sup>[77]</sup>.

When bacterial structures mimicking host molecules elicit an immune response leading to autoantibody production, the result is the onset of autoimmunity. As already discussed in previous chapters, the relevance of saccharide misrecognition by the antibodies in the development of many autoimmune diseases can be ascribed to several

mechanisms such as molecular mimicry and epitope spreading<sup>[9,78]</sup>. Moreover, the chemical modifications of proteins, i.e. aberrant glycosylations resulting from bacterial or viral infections, could trigger pathological reactions of the immune system, thus leading to the production of circulating antibodies directed towards “self” antigens.<sup>[12,79–81]</sup>

Among the many pathogen-related autoimmune diseases, only a few of them has been directly associated with bacterial glycosylation patterns, and yet for most of these the corresponding human antigen has not been depicted. Nevertheless, among the identified bacteria producing surface glycoconjugates that mimic host structures, there are many pathogenic organisms possibly involved in autoimmune diseases, such as *Campylobacter jejuni*, *Helicobacter pylori*, and *Haemophilus influenzae*<sup>[12]</sup>.

The identification and optimization of glycosylated synthetic peptides as antigenic probes able to detect circulating autoantibodies in MS was achieved for the first time in the Laboratory of Peptide and Protein Chemistry and Biology (PeptLab) of the University of Florence, following a “Chemical Reverse Approach”. This experimental mindset is based on the screening of focused libraries of aberrantly modified peptide sequences, with the aim to develop optimized peptide antigens to detect autoantibodies specific of the autoimmune disease under investigation. Peptide epitopes designed by this approach, if selectively and specifically recognizing autoantibodies in a statistical significant number of patients, can be used as antigenic probes in immunoenzymatic assays to detect disease biomarkers in an efficient way. The strategy was defined “Reverse” because the screening of the synthetic antigenic probe is guided by autoantibodies circulating in autoimmune disease patients’ blood. “Chemical” because autoantibody recognition drives selection and optimisation of the “chemical” structure from defined peptide libraries<sup>[37]</sup>. Specifically, while investigating the possible role of MOG glycosylation in pathogenic autoantibodies generation, it was found that a very simple glucosyl moiety of a MOG peptide encompassing the immunodominant region and containing the native glycosylation site resulted in improved detection of specific autoantibodies in humans in a number of MS patients<sup>[82]</sup>. The MOG-derived peptide [Asn<sup>31</sup>(N-β-Glc)]hMOG(30–50), with sequence K<sup>30</sup>N(Glc)ATGMEVGWYRPPFSRVVHL<sup>50</sup>, was the first sequence described to detect putative demyelinating autoantibodies in humans and it was used as a template for the design of a new generation of peptide drugs with the aim to block circulating autoantibodies in patients affected by MS. From these following investigations, the

response to antibodies correlating with MS progression was proven to be even more specific for a synthetic N-glycosyl asparagine (N-Glc) glucopeptide, named CSF114(Glc), which was successfully used to detect, isolate, and characterize antibodies in a subpopulation of patients<sup>[83–85]</sup>. The IgM antibody level to CSF114(Glc) was significantly increased in MS patients versus healthy donors or versus other autoimmune diseases ( $p < 0.001$ ). The IgG response was restricted to the subclass IgG2. Antibodies to CSF114(Glc) were found in 30% of diagnosed RRMS patients and, at lower levels, in subjects affected by meningitis/encephalitis. The high specificity and affinity with which the CSF114(N-Glc) is recognized by antibodies in MS patient sera were ascribed to two crucial features of the biomarker: the N-linked  $\beta$ -D-glucopyranosyl moiety and a  $\beta$ -hairpin structure, which optimally exposes the minimal Asn(Glc) epitope, both essential for autoantibody recognition<sup>[41,86–88]</sup>.

The discovery of this 21-aminoacid antigenic probe TPRVERN(Glc)GHSVFLAPYGWMVK (Figure 7) strengthened the hypothesis that an aberrant asparagine-glycosylation might be implicated in triggering formation of autoantibodies in MS. However, with the only exception of an identified N-glucose moiety in mammalian laminin, an extracellular basement membrane glycoprotein<sup>[89]</sup>, the presence of Asn(Glc) in eukaryotic proteins was virtually unknown at that time in eukaryotes, and found only rarely in archaea<sup>[90]</sup>.

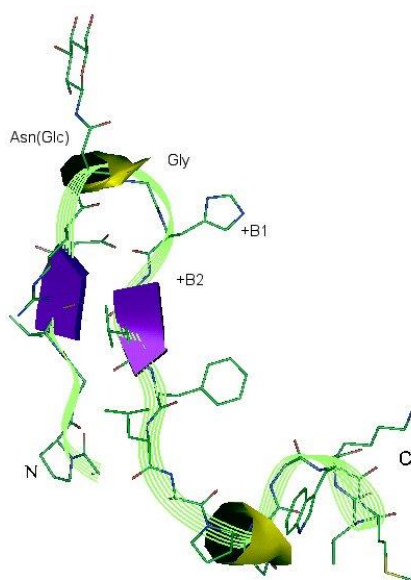


Figure 7 - Ribbon diagram of the lowest energy conformer of 200 calculated structures of TPRVERN(Glc)GHSVFLAPYGWMVK sequence, named CSF114(Glc), derived from NMR data<sup>[85]</sup>.

Therefore, when this unusual post-translational modification was found on cell-surface adhesin HMW1 protein of the pathogen non-typeable *Haemophilus influenzae* (NTHi)<sup>[91]</sup>, the rationale that bacterial infections are pivotal in many autoimmune conditions encouraged to evaluate a connection between NTHi infection and MS.

The NTHi HMW1 adhesin is a high-molecular weight protein that undergoes N-linked glycosylation at multiple asparagine residues with simple glucose units rather than N-acetyl glucose units, revealing an unusual N-glycosidic linkage and a new glycosyltransferase activity. The enzyme responsible for glycosylation of HMW1 is a protein called HMW1C, which shares homology with a group of bacterial proteins that are generally associated with two-partner secretion systems<sup>[92]</sup>. Glycosylation of HMW1 adhesin is necessary for protein stability, secretion, and efficient adherence of NTHi to host cells, hence vital to assure infectivity of the bacterium<sup>[91]</sup>.

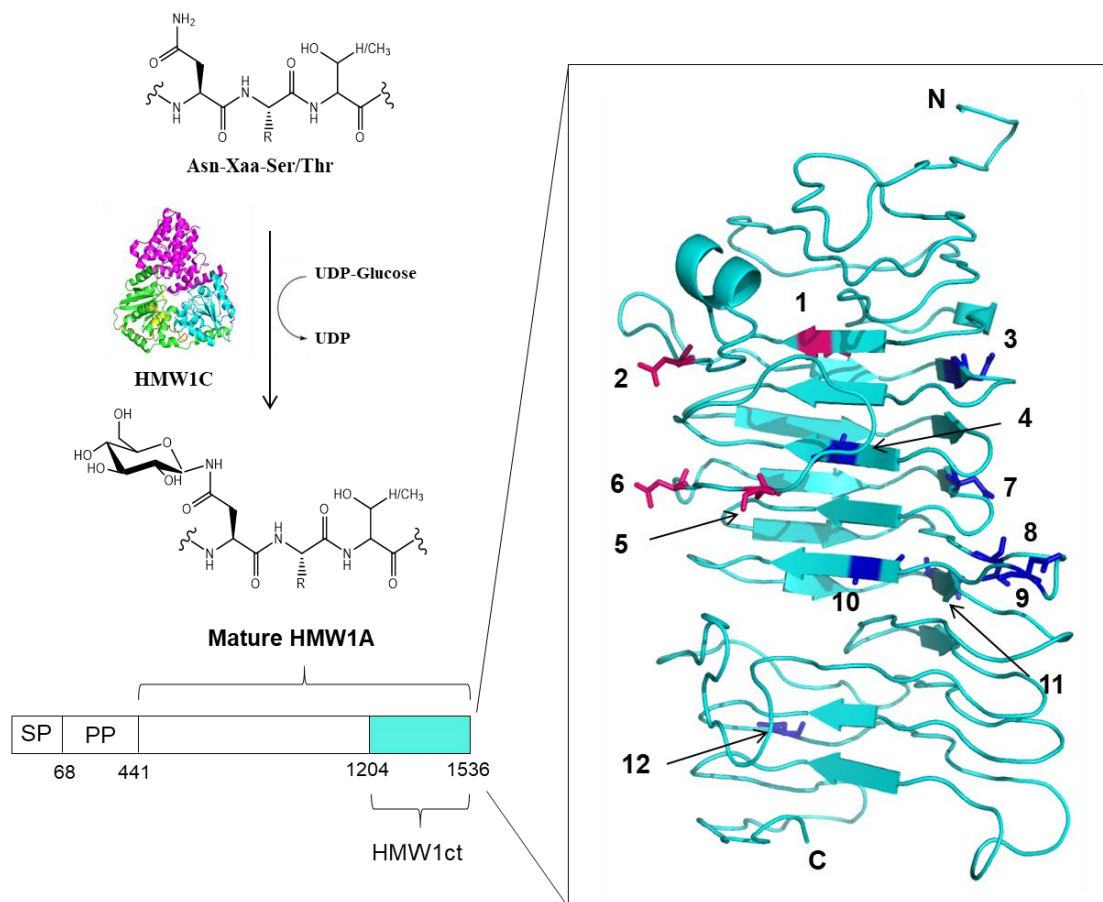


Figure 8 – Schematic representation of the enzymatic reaction by the N-glycosyl transferase HMW1C and the 12 N-glycosylation sites on HMW1ct with the relevant glycosylation sites 1, 2, 5, 6 in magenta, and remaining sites in blue<sup>[93]</sup>.

By using the simultaneous co-expression system of adhesin C-terminal fragment HMW1(1205-1536), named HMW1ct, and N-glycosyltransferase HMW1C in *E. coli*, hyperglucosylated HMW1ct(Glc) was obtained (Figure 8). It was revealed that the expression using this biosynthetic machinery provided HMW1ct(Glc) as a mixture of three N-glycosylated variants containing seven, eight, and nine glucose moieties on Asn out of twelve NX(S/T) glycosylation sites, in a 1:1:1 ratio. This N-Glc protein antigen was essential for the detection of high affinity antibody binding in a subpopulation of MS patients and allowed to purify IgG antibodies from MS patient sera cross-reacting with anti-CSF114(Glc) antibodies<sup>[93]</sup>.

Immunohistochemistry experiments on mouse spinal cord preparations showed that when anti-hyperglucosylated HMW1ct antibodies from MS patients were incubated with spinal cord sections of naïve healthy mice and grade 4 of the MS mouse model EAE, selective staining was observed compared to controls. NTHi glucosylated adhesin is the first example of an N-Glc antigen that can be considered a relevant candidate for triggering demyelinating antibodies in MS, supporting the hypothesis of a strong connection between MS and exogenous pathogens. Therefore, the groundwork was established for determining the nature of the molecular mimicry mechanism, and for elucidating the human protein target(s), which could be cryptic mimics recognized by anti-hyperglucosylated adhesin antibodies in Multiple Sclerosis<sup>[93]</sup>.

## **2. SCOPE OF THE PROJECT AND THESIS OUTLINE**

In the context of glycopeptide-based research in autoimmunity, the main purpose of this thesis project, started in January 2017, is to develop high-affinity molecular tools to detect and isolate antibodies from sera of patients suffering from autoimmune diseases. In particular, most of my Ph.D. work was focused on MS disease that is a neurodegenerative disease presumably involving an antibody-mediated mechanism in the damage of myelin sheath surrounding the axons in the central nervous system. Although the disease-specific autoantigen still remains elusive, the hypothesis that a bacterial infection (NTHi) may play a role in the etiopathogenesis of some sub-forms of MS has been recently investigated. The presence of N-linked glucose epitopes in NTHi C-terminal adhesin fragment HMW1ct (1205-1526) was proven to be essential for the identification of the highest affinity antibodies in MS, showing a very specific recognition. Moreover, the anti-adhesin antibodies that are present in human sera were shown to cross-react with anti-CSF114Glc antibodies, previously characterized as MS specific biomarkers. Finally, the sugar-specific IgG fraction isolated from MS patients' sera was used to confirm myelin specificity by immunohistochemistry. Hence, a correlation between NTHi and MS was established, and hyperglucosylated HMW1ct is the first example of an N-glucosylated antigen that can be considered a relevant candidate for triggering pathogenic antibodies in MS<sup>[93]</sup>.

The debated hypothesis is that IgM antibodies could be reminiscent of an early triggering infection event, but their level tends to decrease with the evolution of the disease, in favour of IgGs that may have a prognostic value instead. Therefore, the challenging task of isolating the specific IgM antibodies is a major goal to understand and treat this complex pathology. In this frame, the relevance of this project is dual.

On one hand, evaluating antibody recognition by specific epitopes and collecting additional information about the role of N-glucosyl asparagine role may provide a great step forward in the comprehension of MS. Therefore, with the aim to identify novel peptide antigens in MS to assess their possible exploitation in antibody binding and trapping, in the following two chapters (3 4) the synthesis of differently glucosylated sequences derived from an adhesin fragment is presented.

On the other hand, the achievement of multivalent, antigen-decorated polymers ideally able to snare circulating antibodies appeared compelling with the future perspective to develop an apheresis-based device for the *selective* removal of autoantibodies from patients' plasma. Therefore, chapter 5 is focused on the development of glucopeptide-grafted polymers for the selective isolation of circulating autoantibodies, especially the IgM type, in MS patients' sera (Figure 9).

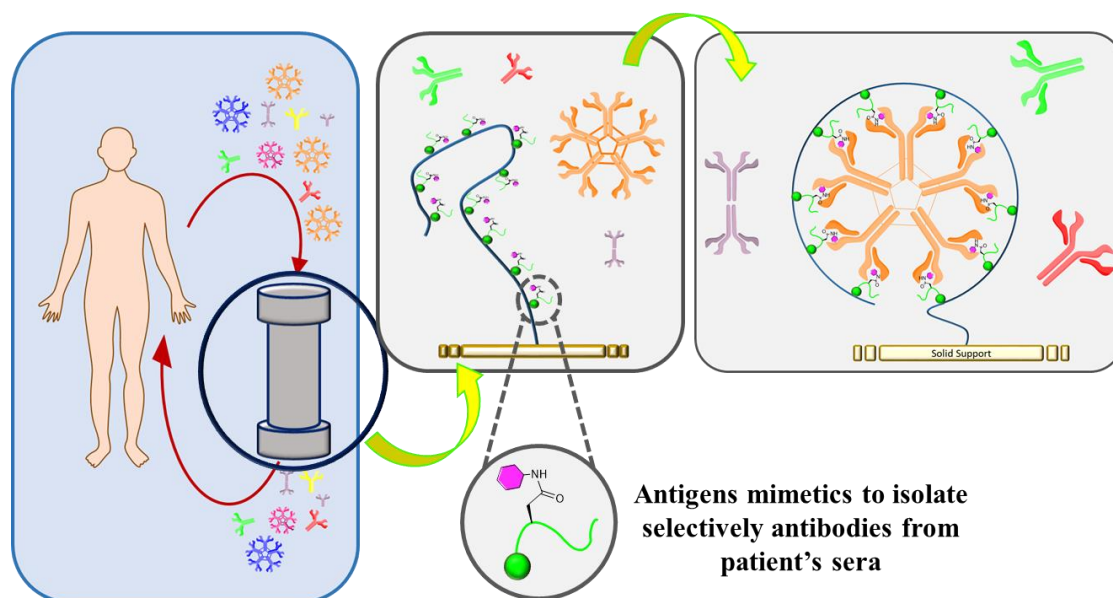


Figure 9 – Schematic representation of the antigen-specific immunoaffinity concept supported in this thesis: the stationary phase of immunoaffinity columns can be equipped with polymer-grafted antigen mimetics, ultimately leading to the selective depletion of specific antibodies, even the elusive IgMs.

As discussed in the introduction, plasma exchange was initially proposed as a valid cotreatment for patients suffering from acute autoimmune attacks. The development of immunoabsorbent columns able to deplete only the antibodies (for example the complete depletion of IgG by grafting protein A or G onto the stationary phase) was a major improvement because it avoids replacement solutions, leaving the other components of the plasma virtually untouched. However, the critical goal nowadays is to achieve more and more selective membranes for the specific isolation of pathogenic antibodies, without compromising the remaining part of the protecting immune components. Finding an alternative to the current treatments with expensive and debilitating combinations of anti-inflammatory and immunosuppressive drugs would have great therapeutic impact. With the intention to emphasize that this concept, supported throughout the thesis, goes beyond MS and can be applied to many anti-glycan antibody-mediated immune disorders, in chapter 6 the synthesis of relevant

glycosyl precursors and the synthetic strategies for the site-specific glycosylation of peptide sequences are discussed.

### **3. HMW1(1347-1354)-DERIVED PEPTIDES**

#### **3.1 INTRODUCTION**

The investigation of the possible role of myelin proteins glycosylation in pathogenic autoantibody generation led to the identification of an antigenic probe named CSF114(Glc), which was used to detect antibodies in a subpopulation of MS patient sera<sup>[83-85]</sup>. This was the first report that an aberrant asparagine-glycosylation (N-Glc) was implicated in triggering formation of autoantibodies in MS.

After that discovery, many attempts have been made to optimize the binding to antibodies and develop new peptide probes which could be used for the diagnosis and therapy of MS.<sup>[94,95]</sup> The challenging task to increase specificity and affinity in MS autoantibody recognition by designing peptides mimicking linear and/or conformational epitopes paved the way for the obtainment of many different analogues of CSF114(Glc)<sup>[96]</sup>, but an unique structure that could outrun the original synthetic peptide was not furtherly employed. Therefore, CSF114(Glc) remains the gold standard peptide probe used in PeptLab, being able to discriminate a group of positive patients, with ca 36% sensitivity and 95% specificity<sup>[38]</sup>.

Nevertheless, a new piece was added to the complex puzzle when the presence of N-Glucosyl epitopes on asparagine residues in consensus sites (N-X-T/S) in *non-typeable* H. influenzae (NTHi) adhesin HMW1ct, corresponding to the C-terminal fragment HMW1(1205-1526), was proven to be essential for the identification of the antibodies in a similar percentage of MS patients. Also, MS antibodies were shown to target the epitope shared by CSF114(Glc) and HMW1(Glc). Glucosylated HMW1ct is the first example of a native N-glycosylated antigen that can be considered a relevant candidate for triggering demyelinating antibodies in MS, supporting the hypothesis of a strong connection between MS and exogenous pathogens<sup>[93]</sup>.

After this ground-breaking finding, we intended to identify novel peptide antigens in MS derived from NTHi HMW1(Glc) and assess their possible exploitation in antibody detection and binding. From a fundamental point of view, finding a minimal epitope recognized with high affinity by antibodies is a crucial task for the investigation of chemical interactions and molecular mechanisms underlying the development of the pathology. From an operational point of view, peptides are more manageable compared to proteins, therefore the long-term search for a potent peptide-based tool might provide unprecedented improvements in diagnostics and therapeutics.

Out of twelve glycosylation sites in HMW1ct, only sites 1,2,3,5,6,7 and 8 are glucosylated in all the variants of HMW1(Glc) when the adhesin HMW1ct and N-glycosyltransferase HMW1C are co-expressed in *E. coli*<sup>[93]</sup>. This is also in accordance with the computational model of the protein, since not all the N-glycosylation sites appear to be exposed on turns and/or on the protein surface (Figure 10). It is hypothesized that sites 1,2,5,6 constitute a relevant cluster of N-glycosyl containing epitopes. In virtue of their spatial proximity and their optimal exposition on the same side of the protein, they could be preferentially targeted by anti-HMW1(Glc) antibodies. Also, as shown in protein primary sequence, sites 5 and 6 are the closest of all and therefore comprise an appealing fragment to start the research for a peptide candidate able to detect MS antibodies, allowing the possibility to synthesize differently glucosylated peptides based on the same epitope sequence.

Based on this assumption, we selected the adhesin fragment HMW1(1347-1354), corresponding to the sequence ANVTLNTT, containing the glycosylation sites 5 and 6 of HMW1ct (Figure 10).

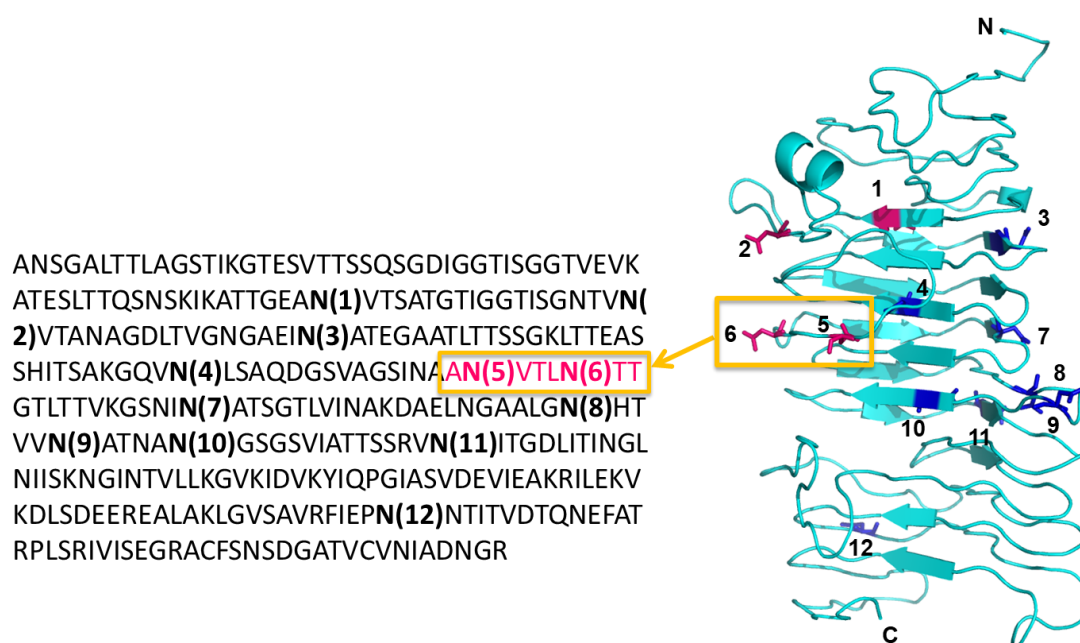


Figure 10 – Structural model of HMW1ct fragment calculated by I-TASSER server<sup>[97]</sup> (right) and the corresponding amino acid sequence (left), with the twelve N-glycosylation sites in evidence. The short sequence containing sites (5) and (6) discussed in this thesis is highlighted in yellow squares.

An N-terminal lysine was envisaged to provide a reactive moiety for conjugation purposes, either to soluble scaffolds or solid supports, as discussed in chapter 5. This

residue is also useful to increase solubility, that is crucial for further characterization such as NMR analysis. Therefore the following peptides were synthesized:

Sequence	Name
Ac-KANVTLNNTT-NH <sub>2</sub>	Peptide 1
Ac-KAN(Glc)VTLNNTT-NH <sub>2</sub>	Peptide 2
Ac-KANVTLN(Glc)TT-NH <sub>2</sub>	Peptide 3
Ac-KAN(Glc)VTLN(Glc)TT-NH <sub>2</sub>	Peptide 4
TPRVERN(Glc)GHSVFLAPYGWMVK	CSF114(Glc)

## 3.2 RESULTS AND DISCUSSION

### 3.2.1 Glycosylated peptide synthesis

Solid Phase Peptide Synthesis (SPPS) methodology has now reached such a level that routinely producing modified peptide sequences has become more and more accessible. Among the orthogonal protection schemes developed in peptide synthesis, the (9H-fluoren-9-ylmethoxy)carbonyl (Fmoc) strategy to protect the *N*<sup>α</sup> position of the amino acid, as temporary protecting group, has definitely gained the widest success.

Fmoc-based SPPS is based on the following steps:

- 1) anchoring of the first amino acid to an insoluble support via the C-terminal carboxyl group
- 2) elongation of the peptide chain through alternate steps of coupling with suitable carboxyl-activating group and Fmoc deprotection in basic conditions (20% piperidine in DMF)
- 3) final cleavage of the peptide from the resin and contemporary deprotection of acid labile side-chain protecting groups

To prepare glycosylated peptides (Figure 11), it is necessary to prepare a conveniently protected building block to be used in solid phase peptide synthesis (SPPS) based on fluorenylmethyloxycarbonyl/*tert*-butyl (Fmoc/*t*Bu) strategy. In the case of glucosylated peptides, the most common employed hydroxyl-protecting group for the glucosyl moiety is the acetyl group.

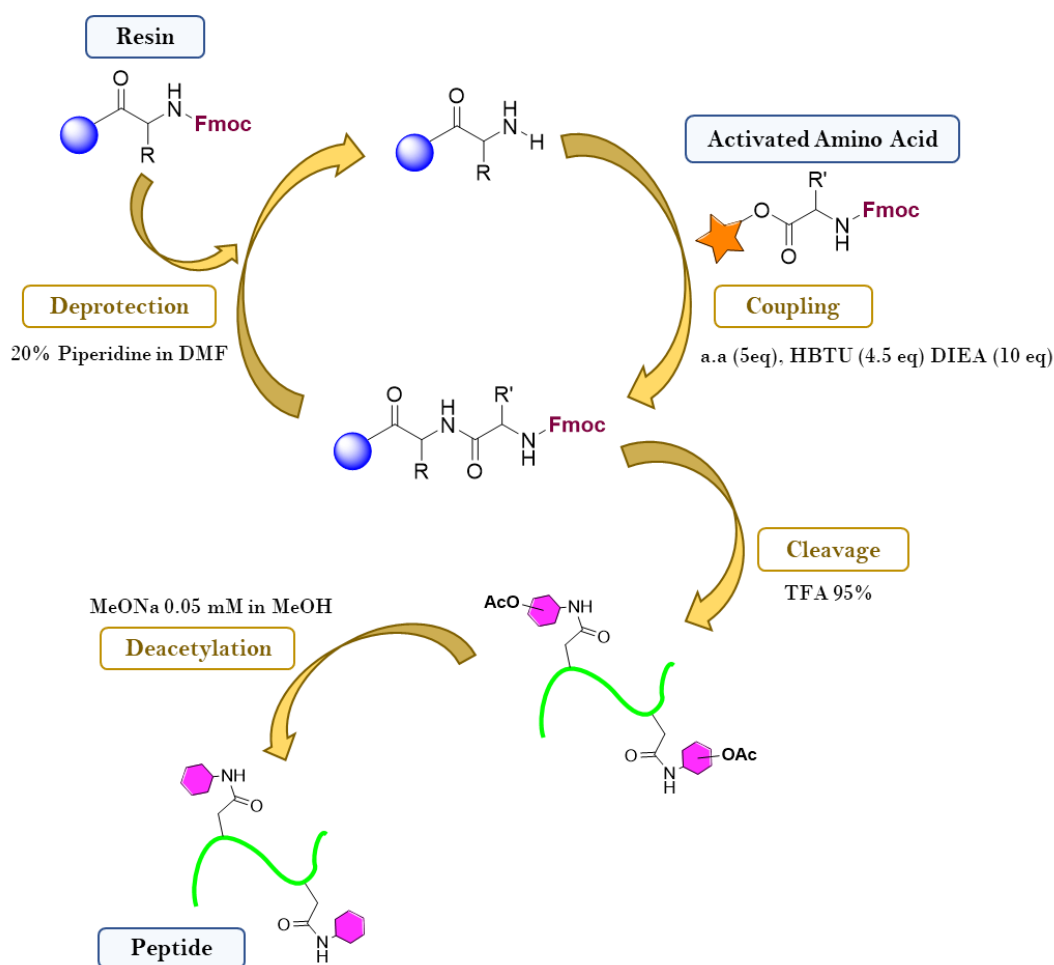


Figure 11 – General procedure for the synthesis of glucosylated peptides on solid-support. After the final acidic cleavage and amino side-chain deprotection, acetylated glucopetide (schematically represented as a green coil exposing two N-linked pink glucose moieties) is deprotected in basic conditions.

To obtain in large quantity the building block (2,3,4,6-Tetra-O-Acetyl-N-[N-Fmoc-L-aspart-4-oyl]- $\beta$ -D-Glucopyranosylamine, i.e. Fmoc-Asn[Glc(OAc<sub>4</sub>)]-OH, different method can be envisaged<sup>[98]</sup>. In any case, they are based on the protection of the hydroxyl functions of glucose bearing amino group on the anomeric position free for the coupling with carboxyl group of amino acids. This can be conveniently achieved starting from glucose to obtain fully protected pentaacetate glucose. Then, the anomeric function was transformed into  $\alpha$ -bromo group *via* S<sub>N</sub>1 with bromhydric acid in acetic acid and acetic anhydride at 0 °C. The corresponding  $\beta$ -azide **I** was obtained *via* S<sub>N</sub>2 by sodium azide. Conversion to the  $\beta$ -amine **II** was accomplished in excellent yield by reduction catalysed by Pd/C in H<sub>2</sub> atmosphere.

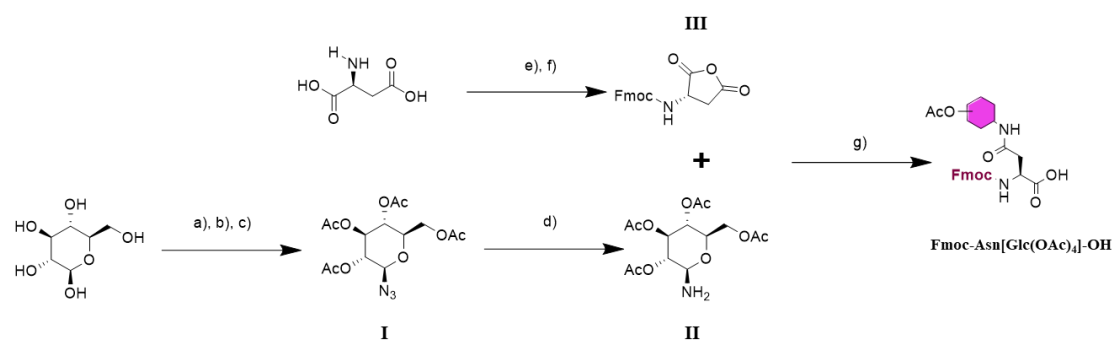


Figure 12 – Reaction scheme for the synthesis of Fmoc-Asn[Glc(OAc<sub>4</sub>)]-OH. a) Ac<sub>2</sub>O in pyridine, b) HBr 33% in AcOH, c) NaN<sub>3</sub> in acetone/H<sub>2</sub>O 6:1 ; d) H<sub>2</sub> on Pd/C in THF ; e) Na<sub>2</sub>CO<sub>3</sub> in H<sub>2</sub>O, Fmoc-OSu in THF, f) Ac<sub>2</sub>O, Δ ; g) DMSO

The fully acetylated glucosyl amine **II** can react with a conveniently protected and activated aspartic acid derivative to obtain the desired building block. In our case, we chose to adopt a previously described method<sup>[99]</sup> that is based on the straightforward production of N-Fmoc protected L-Aspartic anhydride **III** starting from the low-cost unprotected L-Aspartic acid (Asp) *via* a two steps reaction that proceeds easily and with satisfactory yields. This allowed us to obtain large quantities of product **III** which was coupled to product **II** to obtain the desired building block. Contrary to other solvents, the opening of the five-membered anhydride ring is described to be selective in DMSO, and the reaction at room temperature proceeds rapidly to termination as revealed by TLC. However, while the addition of MeOH to reaction mixture is described to provoke the precipitation of the desired product, several attempts to obtain a solid by adding MeOH to reaction mixture failed. Instead, water instantly caused the precipitation of an off-white, slimy aggregate. which was filtered on a frit washing with water for several hours. Any attempt to recrystallize the obtained crude in pure methanol was not effective, hence the dried precipitate was redissolved in the minimum of MeOH and drop-wise addition of H<sub>2</sub>O afforded the desired compound Fmoc-Asn[Glc(OAc<sub>4</sub>)]-OH as a white powder. This compound was used to synthesize all the glucosylated sequences discussed in this thesis. Glucose deprotection occurs via a base-mediated process that leads to the hydrolysis of acetyl groups, and can be performed either on resin or in solution after final cleavage<sup>[96]</sup>. In our case, we opted for the second strategy. At first, the obtained crudes of peptides **2** and **3** were divided in two and then deacetylated by using two already described protocols:

- a) Treatment with MeOH/THF/30% NH<sub>3</sub> aq. (2:2:1)
- b) Treatment with methanolic solution of MeONa (0,05 M)

The reaction mixtures were stirred at room temperature and monitored by HPLC, as described in the experimental section.

In the case of peptide **2**, using method *a* the reaction was considered over after 6h, but the chromatogram showed the presence of two main peaks. Mass analysis by MALDI-TOF revealed that the most intense one is correctly deacetylated peptide **2**, while the other peak mass (+42) suggests the presence of an acetyl group that is probably linked to the lysine (in fact chromatogram of reaction mixture didn't change even after 24h and the addition of a stronger base). It was hypothesized that the spatial proximity between Lys and acetylated glucose moiety on Asn residue makes the unprotonated lysine an optimal competitor of the (weak and diluted) base  $\text{NH}_3$ .

When harsher deprotection conditions were used (method *b* with the remaining half of the crude), only the main peak corresponding to desired product was found, and reaction proceeded faster. In fact, the peak corresponding to acetylated peptide completely disappeared in less than two hours.

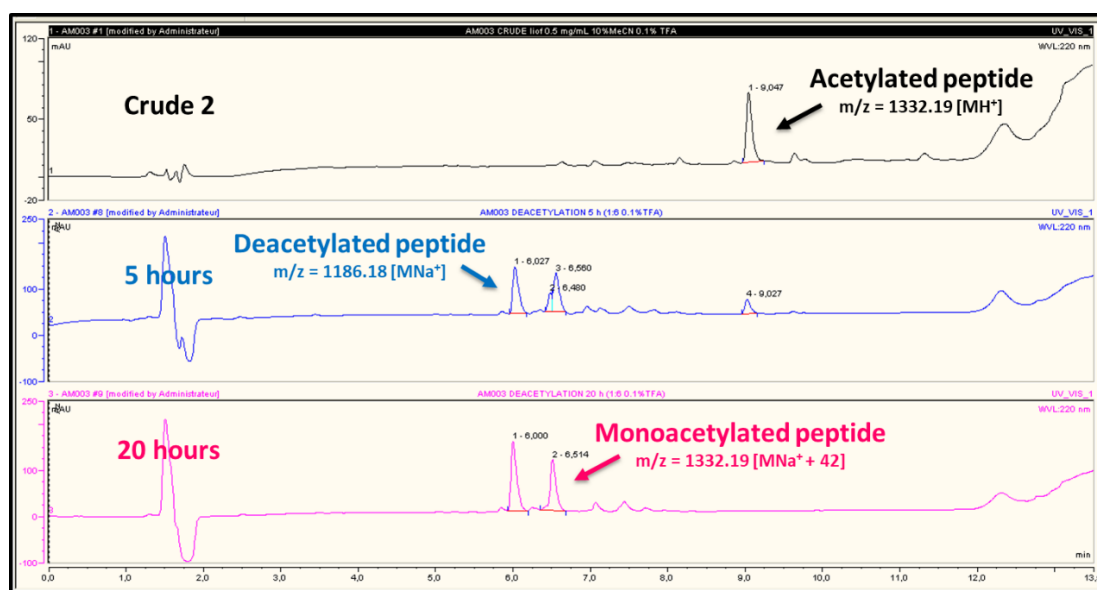


Figure 13 – RP-HPLC chromatograms of deacetylation reaction of peptide **2** by using method *a*.

In the case of peptide **3**, both methods led to desired peptide. The absence of the monoacetylated peptide peak could be due to the larger distance between Lys and Asn(Glc), strengthening the hypothesis that an intramolecular rearrangement may occur.

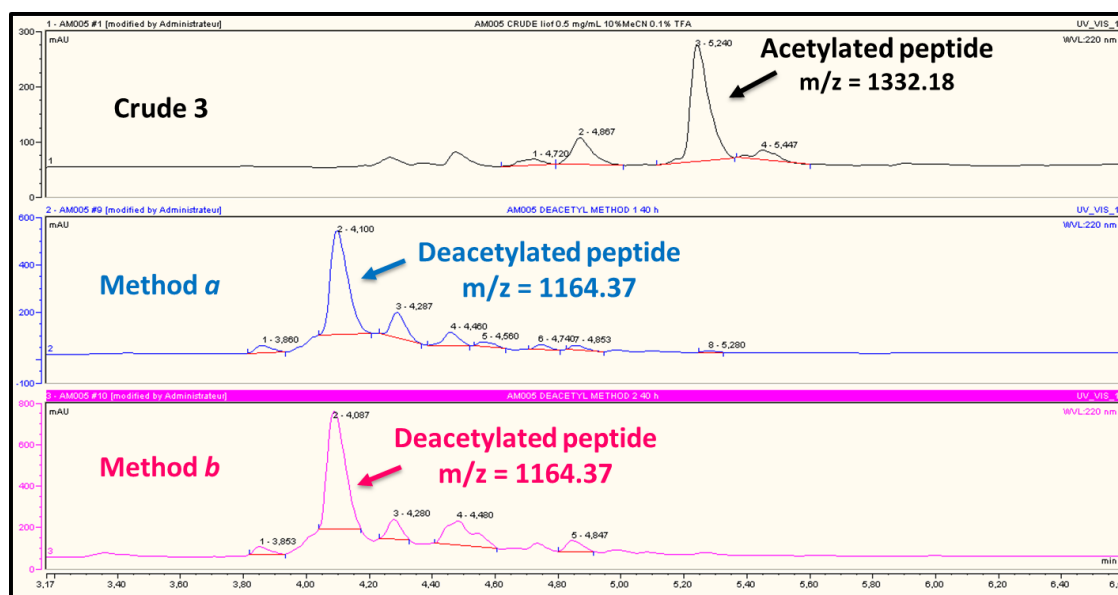


Figure 14 – Comparison of RP-HPLC chromatograms of peptide 3 after cleavage and after deacetylation reactions by using method a and b. Minor peaks were ascribed to sequence deletions or unknown adducts, whereas no acetylated compound was found by MALDI-TOF analysis.

Therefore, for the crude resulting from peptide 4 synthesis only method b was used, and again no presence of acetylated peptide was detected.

Semipreparative RP-HPLC of obtained crudes afforded pure peptides, that were subsequently evaluated by competitive ELISA tests in comparison with CSF114(Glc).

### 3.2.2 Immunological Assays

Immunoassays are based on an antigen/antibody interaction to identify a target compound or a class of compounds. These techniques can exploit the capability of antibodies to detect selectively to a target antigen present in a sample matrix and characterized by a specific chemical structure.

One of the most used immunoassay technique is the Enzyme Linked ImmunoSorbent Assay (ELISA), a technique introduced by Engvall and Perlmann<sup>[100]</sup>, where the detection of antigens or antibodies is based on the use of enzymes, such as alkaline phosphatase (ALP). The antigen from a solution is immobilized on a solid phase, such as a microtiter plate made of rigid polystyrene. Subsequently, the specific antibody-antigen reaction is detected by an enzyme-labeled antibody. The development of color using a chromogenic substrate corresponds to the presence of the antigen. The concentration of the analyte depends on the Lambert-Beer equation and is thus determined by the intensity of colour in the sample. ALP hydrolyzes *p*-nitrophenyl

phosphate to produce *p*-nitrophenol, resulting in a bright yellow colour that can be detected at 405 nm using a microtiter plate reader. The enzyme–substrate reactions are typically completed within 15–60 min, and the reaction is stopped with the addition of an appropriate solution, e.g., sodium hydroxide.

Different types of ELISA have been developed, i.e., direct methods, where the antibody reacting with the antigen is directly linked to the enzyme, or indirect ones, in which an enzyme-linked secondary antibody is employed. Indirect solid-phase (SP) ELISA systems have been demonstrated to be particularly useful to detect the presence of antibodies in sera of patients suffering from antibody-mediated diseases<sup>[37]</sup>. After coating the plate with a known or putative antigen involved in a disease, the key step of this system is the two-binding process of the primary antibody and enzyme-labeled secondary antibody directed to the Fc fragment of human IgGs or IgMs. With the use of sera as source of the primary antibodies, the presence of a disease-associated antibody in the serum can be evaluated; therefore, indirect SP-ELISA can be effectively used to diagnose autoimmune diseases.

Indirect competitive ELISA, or inhibition ELISA, involves the combination of indirect ELISA and competitive ELISA. A specific antigen is immobilized on the solid phase of the microtiter plate. Subsequently, free target antigen at various concentrations is allowed to incubate together with constant concentration of antibody (or serum), resulting in a competition between the immobilized antigen and free antigen against antibodies. After washing steps, the primary antibodies bound to the immobilized antigen are detected by the enzyme-labeled secondary antibody. Therefore, in competitive ELISA, the signal decreases with increasing amount of the free competitive antigen and the concentration required to obtain 50% inhibition of maximum signal (IC<sub>50</sub>, generally expressed as  $-\log$  or pIC<sub>50</sub>) is a crucial parameter for estimating the relevance of a target antigen. Indeed, these assays are particularly useful and to assess the affinity and specificity of target antigens, providing invaluable insights for their optimization and immunological relevance. As discussed in the introduction, the dissociation constant of the circulating antibodies is suggested to be proportional to the partial concentrations of these antibodies in blood serum in equilibrium<sup>[101]</sup>.

IC<sub>50</sub> is inversely related to the  $K_A$ , which describes how much antibody-antigen complex exists at the point when equilibrium is reached. The time taken for this to occur depends on rate of diffusion and is similar for every antibody isotype. However, high-affinity antibodies will bind a greater amount of antigen in a shorter period of time than

low-affinity antibodies.  $K_A$  can therefore vary widely for antibodies from below  $10^5$  mol<sup>-1</sup> to above  $10^{12}$  mol<sup>-1</sup>, and can be influenced by factors including pH, temperature and buffer composition. The affinity of monoclonal antibodies can be measured accurately because they are homogeneous and selective for a single epitope. Serum contains a mixture of heterogeneous polyclonal antibodies of different affinities recognizing several epitopes, hence only an average affinity can be determined<sup>[26]</sup>.

The autoantibody recognition by HMW1ct derived peptides **1-4** as antigens was firstly evaluated by competitive ELISA on four relevant different patients' sera (MS1, MS2, MS3, MS4). The selection of sera was based on previous titration by indirect SP-ELISA with CSF114(Glc). MS1, MS2 and MS3 were found to have a high titer of anti-CSF114(Glc) IgG antibodies, while lacking in IgMs. Serum MS4 was found to possess a high IgM titer instead, and therefore was used for IgM inhibition assays.

As shown in Table 3 and Figure 15-17, glucosylated peptides **2-4** exhibited lower IC<sub>50</sub> values compared to CSF114(Glc), up to nanomolar concentration. On the other hand, nonglucosylated peptide **1** completely failed to compete for the antibody binding, therefore demonstrating once again the crucial importance of the *N*-glycosylic bond between the sugar and the amino acid and the role of the sugar moiety for autoantibody recognition in MS patients' sera.

*Table 3 - Calculated pIC<sub>50</sub> values for inhibitors toward anti-CSF114(Glc) IgG antibodies. Values are reported as 95% confidence intervals for the calculated mean pIC<sub>50</sub> of the adhesin derived peptide antigens 1-4 used as inhibitors of anti-CSF114(Glc) IgG antibodies in 3 representative MS sera.*

Inhibitor	MS 1	MS 2	MS 3
Peptide 1	----	----	----
Peptide 2	9.28 ± 0.55	7.00 ± 0.22	8.02 ± 0.47
Peptide 3	8.04 ± 0.67	7.80 ± 0.56	7.36 ± 0.33
Peptide 4	9.62 ± 0.71	7.51 ± 0.17	8.39 ± 0.36
CSF114(Glc)	7.21 ± 0.29	6.90 ± 0.15	7.93 ± 0.23

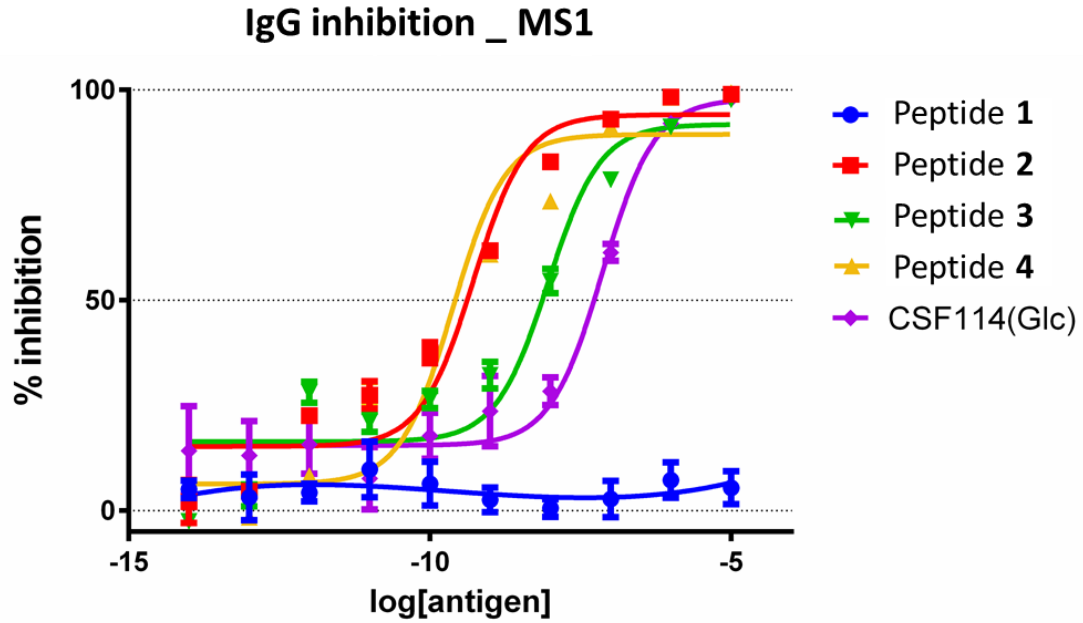


Figure 15 – Inhibition curves of anti-CSF114(Glc) IgG antibodies in serum MS1 with HMWI nonapeptides 1-4 in a competitive indirect ELISA, in comparison with the glucopeptide CSF114(Glc).

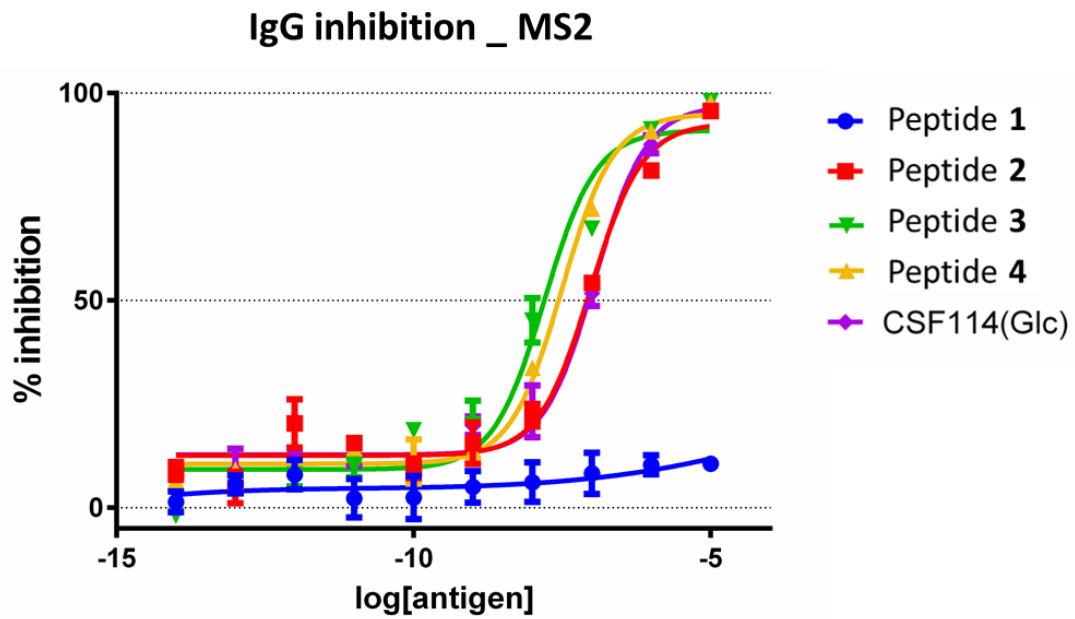


Figure 16 - Inhibition curves of anti-CSF114(Glc) IgG antibodies in serum MS2 with HMWI nonapeptides 1-4 in a competitive indirect ELISA, in comparison with the glucopeptide CSF114(Glc).

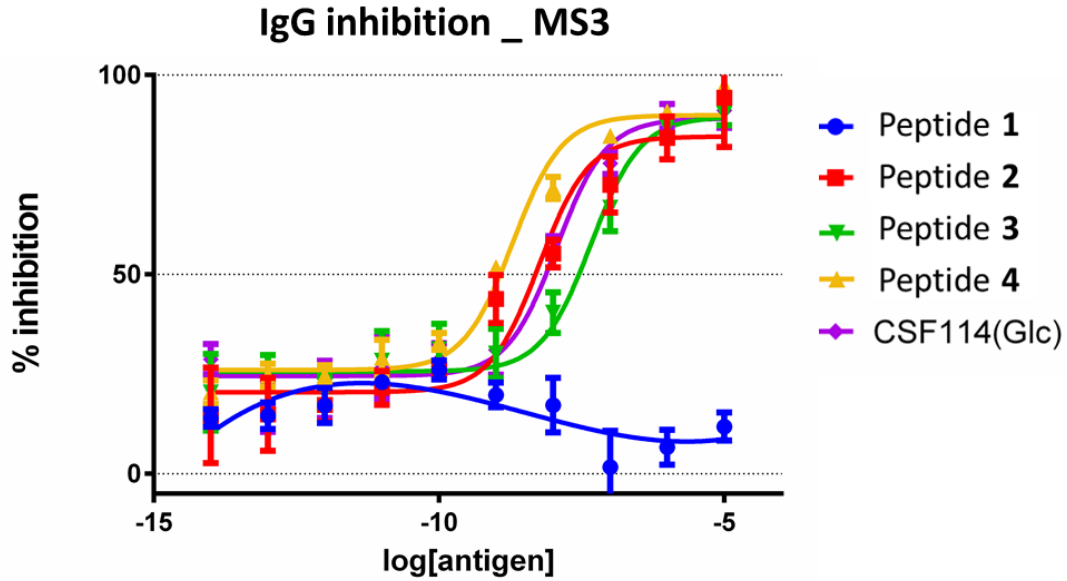


Figure 17 - Inhibition curves of anti-CSF114(Glc) IgG antibodies in serum MS3 with HMW1 nonapeptides 1-4 in a competitive indirect ELISA, in comparison with the glucopeptide CSF114(Glc).

The results of IgG competitions showed that adhesin based peptides are generally more active to bind antibodies in patients' sera. Although very promising, they also show that the recognition is not dependent on the specific position of the sugar moiety. In fact, whereas diglycosylated peptide 4 and N3 monoglycosylated peptide 2 displayed an excellent inhibitory potency in MS1 (Figure 15), peptide 4 and N7 monoglycosylated peptide 5 were the ones exhibiting the best activity in MS2 (Figure 16). In MS3, only peptide 4 resulted in a slightly improved activity compared to CSF114(Glc) (Figure 17). The differences in pIC<sub>50</sub> values among different sera, especially compared to the reference probe CSF114(Glc), may be explained by the presence of glycoprotein-targeting antibodies and their polyclonal nature. In other words, patients differently producing similar but distinct antibodies directed to corresponding N(Glc) epitopes of a protein may present specific predilection toward a glycosylated antigen.

Though able to capture IgM-type antibodies in indirect SP-ELISA, the peptide CSF114(Glc) was never able to inhibit IgMs in competitive ELISA. This is the case also for the numerous antigenic sequences that have been previously developed and investigated in structure-affinity studies aiming to improve antibody recognition. The evidence that only competition tests for IgGs had success may be due to two different reasons. From a practical point of view, anti-N(Glc) IgGs, which possess higher affinity compared to IgMs, are often present in sera of MS patients, therefore hampering the outcome by binding free target antigens and preventing IgM inhibition. Secondly, high-

avidity, pentameric IgM, possessing ten identical antigen binding sites, may be difficult to undermine without a very high affinity ligand. Surprisingly, all the glucosylated peptides 2-4 were able to inhibit anti-CSF114(Glc) IgM binding in serum MS4, although with a low pIC<sub>50</sub> (Table 4 and Figure 18).

Table 4 - Calculated pIC<sub>50</sub> value for inhibitors toward anti-CSF114(Glc) IgM antibodies. Values are reported as 95% confidence interval for the calculated mean pIC<sub>50</sub> of the adhesin derived peptide antigens used as inhibitors of anti-CSF114(Glc) IgM antibodies in a representative serum.

Inhibitor	MS 4
Peptide 2	4.69 ± 0.36
Peptide 3	4.90 ± 0.54
Peptide 4	5.15 ± 0.10
CSF114(Glc)	----

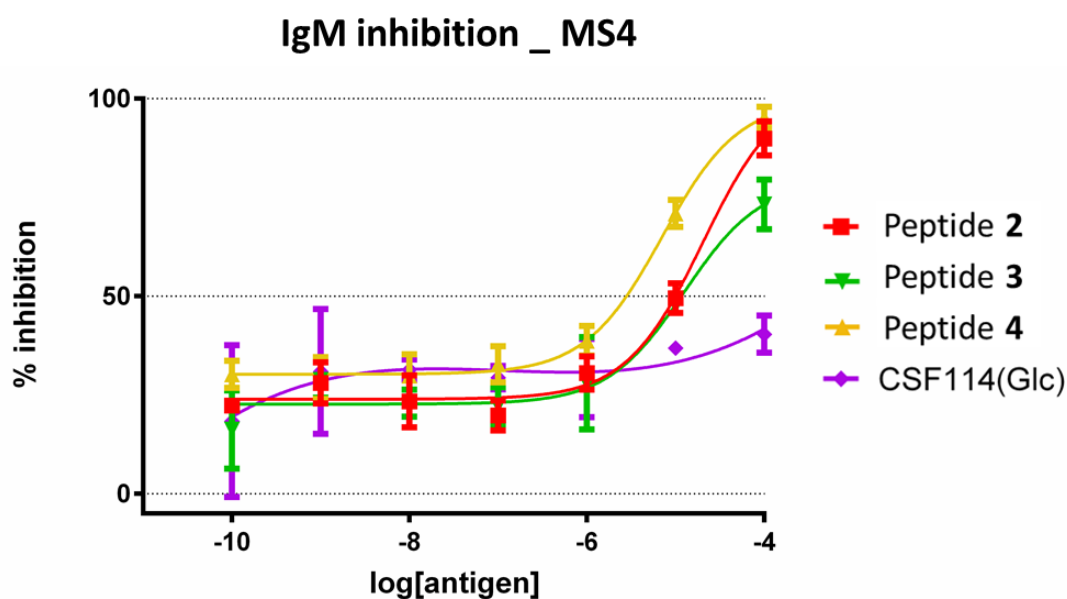


Figure 18 - Inhibition curves of anti-CSF114(Glc) IgM antibodies in serum MS4 with HMW1 nonapeptides 1-4 in a competitive indirect ELISA, in comparison with the glucopeptide CSF114(Glc).

Peptide 4 was found the best candidate for IgM competition, displaying a half maximal inhibiting concentration in the  $\mu\text{M}$  range, and therefore providing a compelling scaffold to develop higher affinity probes.

### 3.2.3 NMR and CD-based conformational investigations

In solution, most of the peptides assume multiple flexible conformations. Theoretical and experimental methods play complementary roles in peptide conformational studies for the determination of the dominant conformers and evaluation of their population<sup>[102]</sup>.

A major goal of conformational investigations is to determine the relationship between conformation and activity of biologically important peptides. Numerous biological results are strongly supported by conformational investigations, clearly indicating that biological peptide activity is determined not only by the presence of specific functions binding to a target protein, but dramatically depends on the conformational properties of the whole peptide structure, too.

The impact of post-translational modifications such as glycosylation on amino acid sequences has been investigated but not fully elucidated yet, and in most cases remains unclear specifically how the carbohydrate moiety interacts with the residues to influence the backbone conformation. Given the tremendous relevance of glycoproteins, which have been implicated in many different cellular processes such as immune response, intracellular targeting, and intermolecular recognition<sup>[36,103,104]</sup>, the ultimate goal of researchers in the field is to understand the significance of glycosylation. This challenging and always trendy topic is closely related to the structural effects that saccharide moieties have on aminoacidic scaffold. In the case of asparagine-linked glycosylation, pioneering studies investigated the wide variety of structural and functional roles that N-glycosylation accomplishes with proteins and peptides. The modification of the asparagine inserted in a consensus sequence was hinted to influence protein rearrangement by altering the local secondary structure proximal to the glycosylation site, facilitating the folding process and stabilizing the mature chain<sup>[105–107]</sup>. A short nonapeptide based on the critical A285 glycosylation site of the hemagglutinin glycoprotein from influenza virus was used as a model system to study the effects of N- glycosylation. Derivatization of this peptide with short carbohydrates revealed that subtle changes in the structure of the carbohydrate have a dramatic impact on peptide conformation<sup>[108]</sup>. The NMR analysis of the chitobiose-linked nonapeptide and its nonglycosylated analogue in water demonstrated that that N-linked glycosylation induces a change in the backbone conformation even for a short sequence, causing the switch from an Asx-turn to a type I  $\beta$ -turn conformation<sup>[109]</sup>.

Aiming to further explain the differences in antibody recognition and unveil the structural consequences of glycosylation, we performed nuclear magnetic resonance (NMR) and circular dichroism (CD) experiments to explore the conformational preferences of the model sequence Ac-KANVTLNNTT-NH<sub>2</sub> containing the two glycosylation sites (peptides **1-4**).

Unfortunately, all these peptides resulted too flexible to provide a reliable tridimensional model. The structural analysis of peptides 1-4 was based on <sup>1</sup>H, <sup>13</sup>C and <sup>15</sup>N chemical shifts, ROE correlations, vicinal coupling constants (see experimental part). The structural data for all these peptides confirmed that they do not possess a defined secondary structure.

The chemical shift analysis of <sup>13</sup>C $\alpha$  and <sup>1</sup>H $\alpha$ , expressed in terms of chemical shift deviations (CSDs) compared to reported random coil shifts, reveals dynamic features for all the sequences. Also, no significant differences ( $\Delta\delta$  <sup>1</sup>H < 0.02 and <sup>13</sup>C < 0.1 ppm) were found by comparing the  $\alpha$  <sup>1</sup>H and <sup>13</sup>C chemical shifts between peptides **2-4** and the unglycosylated sequence **1**, with the only exception of modified asparagine, indicating that they have the same conformational flexibility of the non-glycosylated analogue (Figure 19).

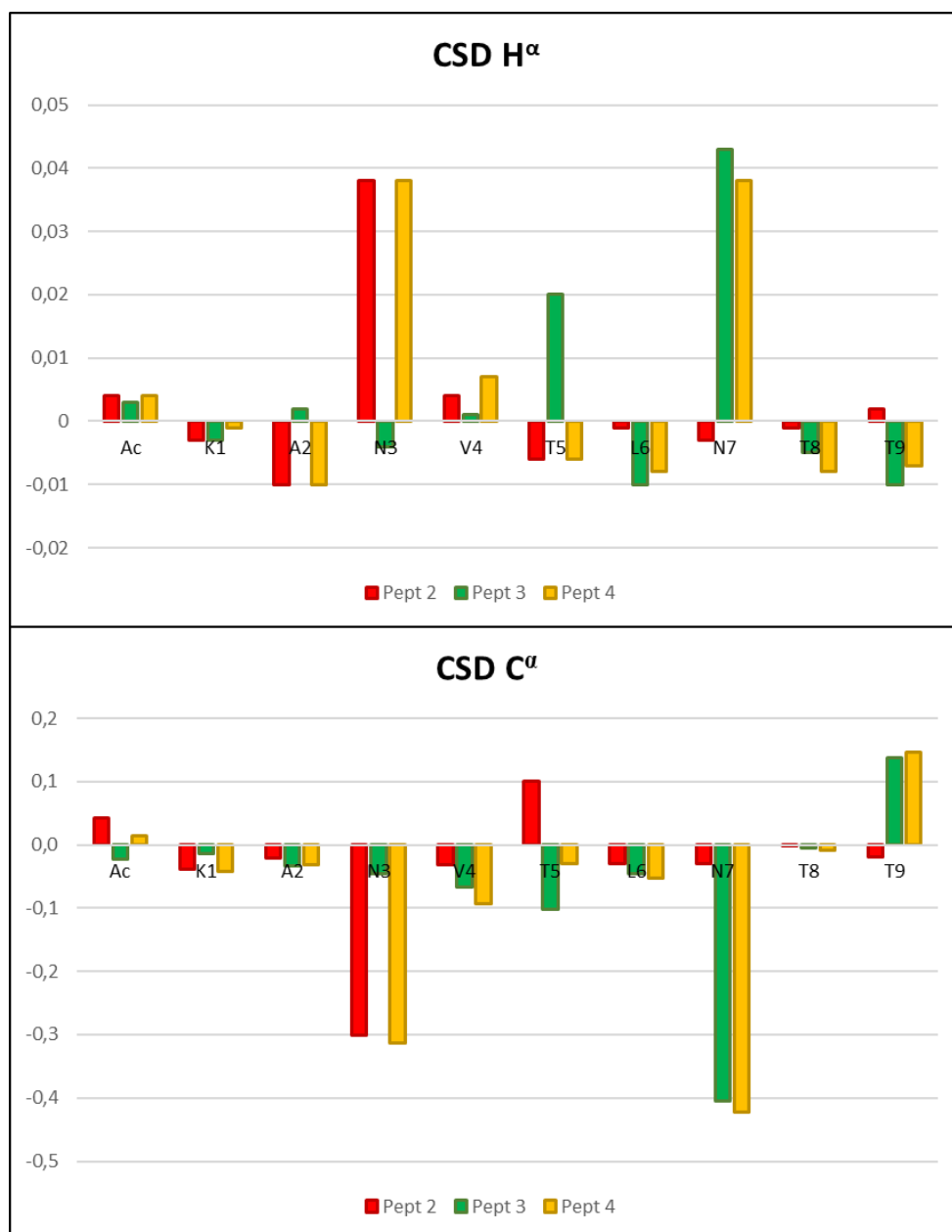


Figure 19 – Chemical Shift Deviation (CSD) for H $\alpha$  (up) and C $\alpha$  (down) of peptides 2,3 and 4 compared to the chemical shift of the corresponding residues in the non-glucosylated peptide 1.

NMR analysis of glycopeptides in water previously revealed that the presence of glucose has a detectable effect only when other constraints such as a disulphide bridge is present, while its impact tends to be more elusive for linear sequences.<sup>[96][4]</sup> Therefore, the lack of a definite conformational effect resulting from the presence of glucose is not surprising for the analysed nonamers. Nevertheless, acquired data may be useful in saturation transfer difference (STD) NMR experiments<sup>[110]</sup> in presence of purified antibodies to gain invaluable information about the epitope-paratope interactions.

CD spectroscopy, which is based on the differential absorption of left- and right-handed circularly polarized light, is widely used to determine the structure of macromolecules, especially the secondary structure of proteins. Within the UV region from 180 up to 240 nm each of the structures  $\alpha$ -helices,  $\beta$ -sheets, and the remaining, unordered part of the polypeptide backbone, usually referred as random coil, contribute in different ways to the peptide spectrum.

Analyzing peptides 1-4, the CD experiments confirmed that the peptides have the same random-coil behavior in water, no matter the position or the number of glucoses. In fact, all the spectra showed a deep negative band at 201 nm that is characteristic of a random coil structure. This confirmed that the effect of glycosylation on the structural arrangements of these short sequences is clearly minimal, meaning that the conformation in water cannot be correlated with their reactivity towards antibodies.

Notably, when a secondary structure inducer such as 2,2,2-trifluoroethanol (TFE) is used as cosolvent, the transition to a more ordered configuration occurred to be more prominent in absence of the glucose moieties. In fact, a more marked appearance of the negative band in the range 205-225 nm and of the positive band in the far UV range ( $\approx 190$  nm) typical of structured peptides was revealed for peptide **1**, compared to peptides **2** and **3**. The impact of TFE addition was even lower in the case of diglycosylated peptide **4**. TFE/water mixtures are considered biomimetic because TFE can displace water, hence removing the solvent hydrogen-bonding molecules and providing a low dielectric environment that promotes intrapeptide hydrogen bonds<sup>[111]</sup>.

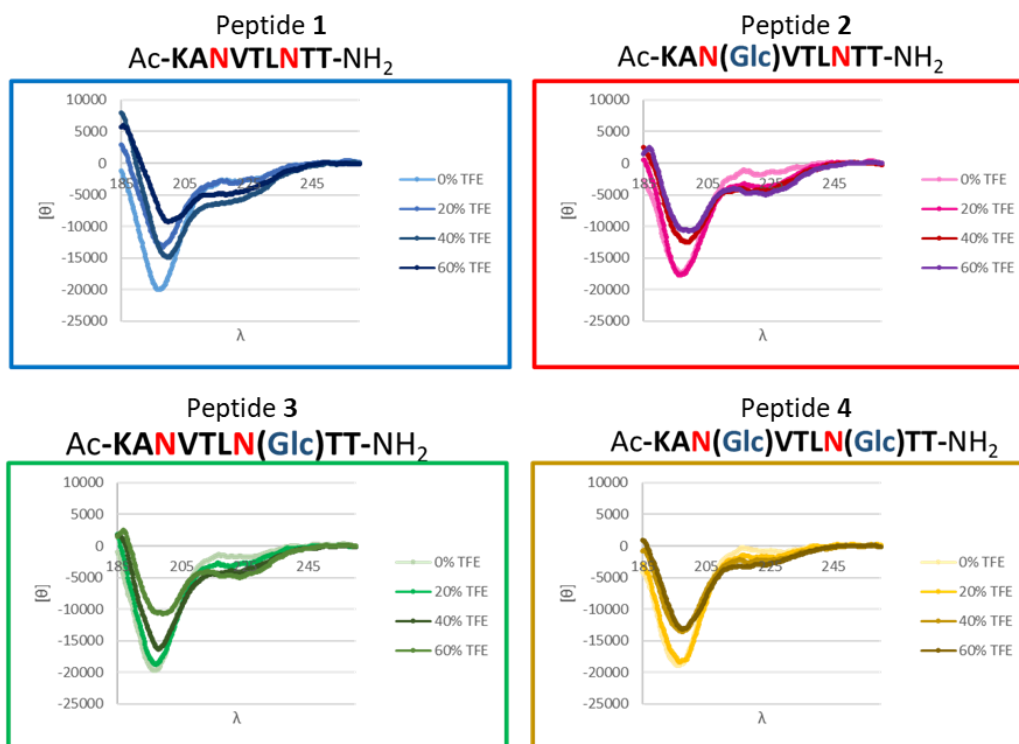


Figure 20 – CD spectra of peptides 1-4 at different concentrations of TFE.

To compare the different peptides, we also used two concentration-independent CD parameters, namely  $R_1$  and  $R_2$ , which are sensitive probes of helix/sheets formation [112,113], and the outcome is a distinct propensity to avoid secondary structures formation for peptides **2**, **3** and **4**.  $R_1$ , the ratio between the intensity of the maximum between 190 and 195 nm and the intensity of the minimum between 195 and 210 nm, is positive for random coils and  $\leq -2$  for fully structured peptides.  $R_2$ , more specific for  $\alpha$ -helices, is the ratio between the ellipticity at 222 nm and the minimum between 195-210 nm, and it is  $\approx 0$  for unstructured sequences, up to 1 for predominantly helical peptides.

In the case of nonglycosylated peptide **1**, the addition of only 20% TFE provoke the change of the  $R_1$  sign, and the effect is even more marked as TFE concentration increases. Instead, calculated  $R_1$  for diglycosylated peptide **4** become  $\approx 0$  only when TFE is the major solvent, with water in minority. Also  $R_2$  trend hints toward similar conclusions, increasing for peptide **1** with higher slope compared to glycosylated sequences.

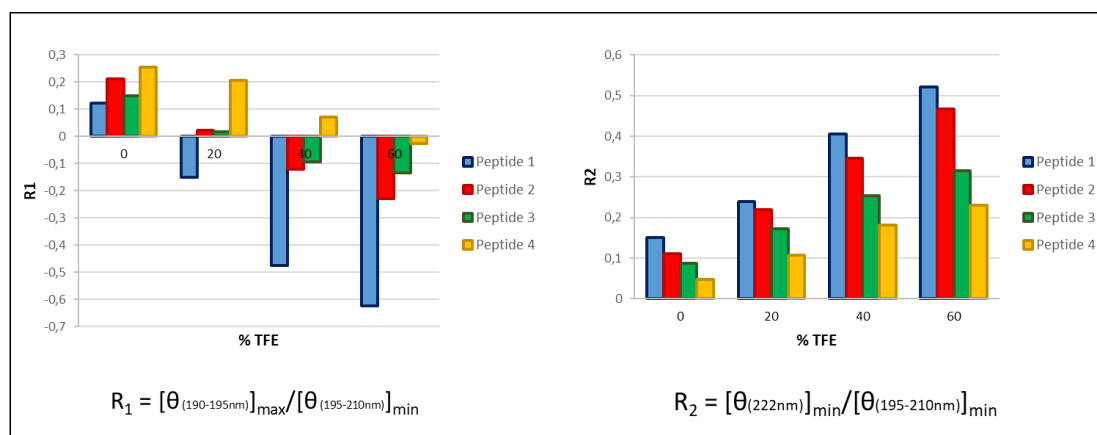


Figure 21 – Bar graphs displaying  $R_1$  (left) and  $R_2$  (right) values versus TFE concentration.  $R_1$  is positive for unstructured peptides while  $R_2 \approx 1$  for  $\alpha$ -helices.

### 3.3 CONCLUSIONS

This chapter is focused on the complex, long-term research for the structure-activity relationship of peptide epitopes involved in antibody recognition. This challenging task can be instrumental to develop mimetic peptide-based molecules improving antibody affinity and specificity to protein antigens. Here, we reported the synthesis and characterization of the HMW1 derived peptides **1-4** containing differently glucosylated consensus sequences. Competitive ELISA tests were performed in order to assess their cross-reactivity with antibodies in patients' sera, using as coating agent the glucopeptide CSF114(Glc), previously designed as a  $\beta$ -turn sequence able to characterize with high specificity and high affinity the antibodies in the sera of a subpopulation of MS patients. The short unstructured nonamers carrying either one or two N-glycosyl moieties were found to be powerful competitive agents for the interaction with IgG antibodies, displaying  $IC_{50}$  between  $10^{-10}$  and  $10^{-8}$  M with relevant patients' sera. Conversely, the non-glucosylated peptide **1** was not able to inhibit the binding. Up to now, these peptides are among the shortest sequences that have been found to have such a low  $IC_{50}$  in competitive ELISA tests for anti-N(Glc) IgG recognition. Moreover, they are the first short peptides that can compete for IgM binding affinity, even though with a high  $IC_{50}$  ( $\approx 10 \mu\text{M}$ ).

However, the recognition of Ab does not appear to be connected to the position of glucose or to the number of glucoses in the peptide sequence. Also, the structural insights reveal high flexibility for all the sequences, without a definite impact of the glucosyl moiety on the secondary structure formation. Hence the reason of antibody

recognition must lay in the presence of Asn(Glc) inserted in a suitable amino acid environment, able to adapt to the structural binding site of the antibodies in the different sera. Since any attempt to cyclize or add some structure inducing constrain was not effective to improve the potency of previously developed sequences<sup>[96]</sup>, this work supports the evidence that antibody binding is not easily correlated to a fixed 3D conformation. More likely it is based on other effects, with Asn(Glc) being the indispensable requisite for the recognition. Therefore, a determined conformation (at least in water) is not necessarily related to the binding properties of the antigens. Although all the sequences show very high mobility and dynamic propensity, the presence of the glucosyl moieties was proven to have an impact in the secondary structure rearrangements in a less hydrophilic environment obtained by the addition of TFE as cosolvent. CD analysis revealed that diglycosylated peptide **4** maintain an extended coil behavior in solutions containing up to 60% TFE, whereas glucose-lacking peptide **1** displayed a clear structure propensity.

With the aim to increase affinity towards anti-Asn(Glc) IgM antibodies, adhesin HMW1(1347-1354) diglycosylated epitope was selected as the best peptide candidate.

#### **4. PEPTIDE HOMODIMERS AND HMW1(Glc) EXPRESSION**

Since the NTHi cell-surface adhesins are extensively glycosylated, the N-glycosyl residues are likely to be presented in a multivalent manner, which would potentially favor the emergence of a robust immunologic response. According to the hypothesis that hyperglucosylated HMW1ct expose different clusters of relevant glucosylated Asn residues, HMW1ct variants were previously prepared by using site-directed mutagenesis<sup>[93]</sup>. All the less-glucosylated forms were able to inhibit IgG binding towards the antibody found in MS sera. Hyperglucosylated HMW1ct and mutants possessing up to four glucose moieties displayed comparable IC<sub>50</sub> values as inhibitors toward anti-CSF114(Glc) IgG antibodies in five MS patients' sera. However, a drop in IgG affinity was observed when the glucosylation sites were diminished from four to two, with IC<sub>50</sub> decreasing from nM range to  $\approx 10$ -1  $\mu$ M. Based on these evidences, it was suggested that the introduction of at least four N-glycosyl moieties on HMW1ct was essential for the identification of the highest affinity antibodies in MS<sup>[93]</sup>.

Intrigued by the possibility that a sequence exposing four N-glycosyl moieties could mimic to a greater degree the activity of the bacterial protein, a homodimer based on di-glucosylated HMW1(1347-1354) fragment was prepared.

Copper (I) catalyzed alkyne-azide cycloaddition (CuAAC), the most known example of "click" reaction<sup>[114]</sup>, was envisaged to introduce a triazolyl linker, isostere of amide bond, connecting two identical sequences bearing two Asn(Glc) each. Fmoc-protected propargylglycine (Pra) and lysine azide (Lys(N<sub>3</sub>)) are commercially available building blocks that can be conveniently used in the SPPS route to produce "clicked" peptides. Therefore, with the same synthetic strategy discussed in the previous chapter, we prepared the two 11-mer sequences Ac-Lys-Ala-Asn(Glc)-Val-Thr-Leu-Asn(Glc)-Thr-Thr-Gly-Pra-NH<sub>2</sub> (peptide **5**) and Ac- Ac-Lys-Ala-Asn(Glc)-Val-Thr-Leu-Asn(Glc)-Thr-Thr-Gly-Lys(N<sub>3</sub>)-NH<sub>2</sub> (peptide **6**). These sequences are very similar to peptide **4** but each of them contains in addition the native glycine residue of HMW1 sequence, as a short spacer between the two relevant glycosylated sites and the grafting moiety.

After SPPS synthesis, the obtained crudes were deacetylated with basic treatment by NaOMe (0.05 M in MeOH). Interestingly, while deacetylation of peptide **5** provided the desired product as the major component, in the case of peptide **6** the conspicuous

formation of a secondary by-product was encountered. In fact, the RP-HPLC analysis of reaction crude after two hours showed the presence of two main peaks with similar intensities (Figure 22).

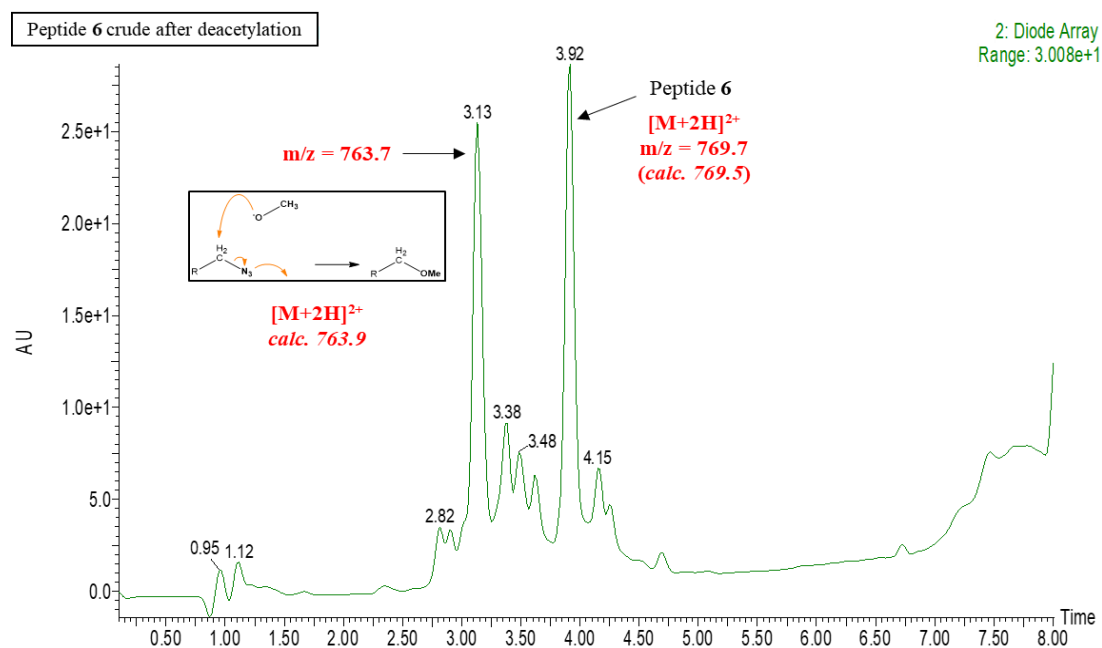


Figure 22 – RP-HPLC chromatogram of deacetylated crude 6, showing the presence of two major peaks. ESI-MS analysis revealed that one ( $t_r = 3.92$  min) corresponds to desired deacetylated peptide 6, whereas the other ( $t_r = 3.13$  min) may derive from the  $S_N2$  by-product, in which a methoxide group has replaced the azide group. Other minor peaks were ascribed to sequence deletions or unidentified by-products. (Analytical RP-HPLC gradient 20-80% B in 5 min at  $0.6 \text{ mL min}^{-1}$ ; solvent system A: 0.1% TFA in  $\text{H}_2\text{O}$ , B: 0.1% TFA in MeCN)

Beside the desired peptide 6, it is present a side-product peak whose mass corresponds to that of the methoxy-derivative. In fact, it is hypothesized that in virtue of its good leaving group properties, azide group is substituted by the abundant nucleophile  $\text{MeO}^-$  during deacetylation, resulting the  $S_N2$  product. To avoid this side-reaction, several parameters may be changed in deacetylation conditions. For example, shorter reaction times may result in a decrease of  $S_N2$  reaction occurrence. Also, diluted methoxide solution could help to prevent such a massive formation of side-product. Finally, changing deacetylation reagents may provide a better yield in desired product. Nevertheless, the obtained amount of pure peptide 6 was sufficient for our purposes, therefore no other deprotection condition was tested. The pure by-product was also collected, and further analysis could confirm its structure.

The purified peptides **5** and **6** were solubilized in water and *tert*-butanol in presence of Cu<sup>I</sup> to catalyse the regioselective alkyne-azide cycloaddition, providing quantitatively (by HPLC reaction analysis) peptide **7** (Figure 23).

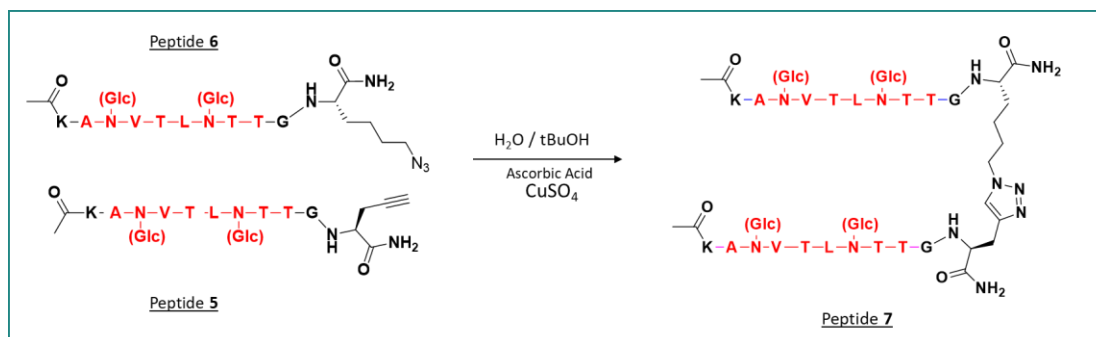


Figure 23 – Scheme of the CuAAC “click” reaction performed to obtain homodimer peptide **7**, starting from *Pra*-containing monomer **5** and *Lis*(N<sub>3</sub>)-containing peptide **6**. Diglycosylated HMW1(1347-1354) sequence is indicated in red.

Copper is often easily complexed by proteins and peptides and its removal is crucial for biological applications. EDTA, a strong chelating agent for copper, was added at the end of the reaction to avoid peptide complexes formation. After a two-step purification process, pure peptide **7** was obtained and used in ELISA experiments.

Additionally, we compared peptide **7** exposing four N-Glc moieties with the two-branched homodimer **8** that is carrying one *Asn*-linked β-D-glucopyranosyl function on each “arm” (Figure 24).

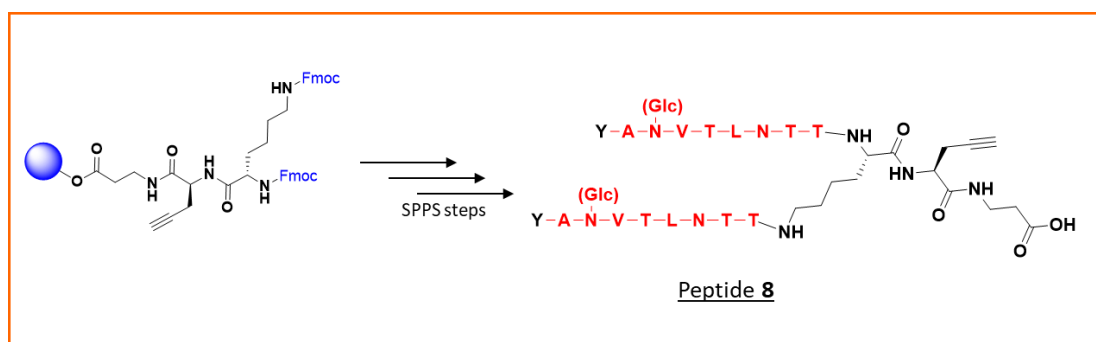


Figure 24 – Sequence of peptide **8**, a two-branched sequence synthesised using non-canonical *Fmoc*-Lys(*Fmoc*)-OH in standard SPPS procedure. Wang resin is represented as a blue sphere, whereas in red is highlighted the HMW1(1347-1354) sequence shared by all the adhesin-derived peptides discussed in this thesis.

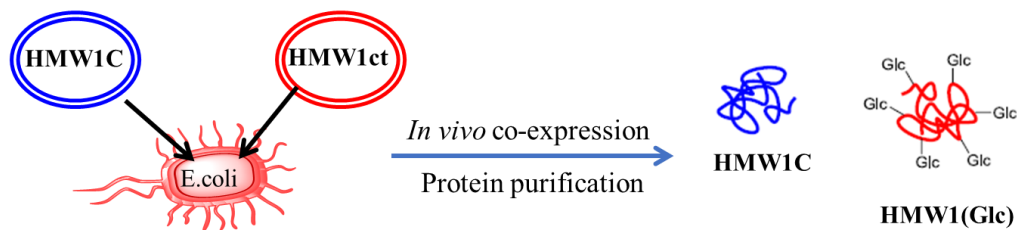
Homodimers **7** and **8** share the repeated primary sequence of HMW1(1347-1354) fragment, that is dimerized with two different strategies. With a similar approach to the

synthesis of multiple antigen peptide (MAP) dendrimers<sup>[115]</sup>, the synthesis of peptide **8** was based on double N-Fmoc protected Lys as the starting point for the ramification and the concurrent growth of the two branches directly on resin. The sequence includes a Pra for the CuAAC conjugation to peptides able to penetrate the blood brain barrier (BBB)<sup>[116]</sup> that will be performed at the Institut for Research in Biomedicine (IRB, Barcelona, Spain) in collaboration with the group of Dr Meritxell Teixidó. For this project, peptide **8** and a Pra-containing derivative of CSF114(Glc), as well as their non-glucosylated analogues, were synthesised. Conjugation to molecular BBB “shuttles”<sup>[117]</sup> and their evaluation in a human BBB cell-based model may lead to small peptide-probes for the detection of antibodies directly in the CNS. This achievement could provide important insights for the comprehension of MS, however the experiments are still ongoing and therefore they will not be discussed in this thesis. Nevertheless, it is important to mention that in competitive indirect ELISAs for anti-N(Glc) antibody binding, peptide **7** and **8** displayed the same inhibition curves with identical IC<sub>50</sub>.

For these experiments, both CSF114(Glc) and hyperglucosylated HMW1ct, i.e. HMW1(Glc), were used as coating antigens.

The peptide probe CSF114(Glc) has similar number of residues, i.e. sequence length (21 amino acids), and display only one glucosyl-asparagine residue inserted a  $\beta$ -turn conformation.

The protein HMW1(Glc) was expressed according to a previously described protocol, resulting in a mixture of variants that was shown to bear an average of eight N-Glc<sup>[97]</sup>. Expression of HMW1(Glc) was performed by growing cultures of *E. coli* cells (strain BL21) previously transfected with two plasmid vectors (Figure 25). One contained the genes encoding for the protein fragment HMW1(1205 – 1536) and the gene for carbenicillin resistance. A second plasmid encoded for the glucosyltransferase enzyme ApHMW1C and contained the gene for a different antibiotic (kanamycin) resistance.



AHHHHHHVWWTANSGALTTLAGSTIKGTESVTTSSQSGDIGGTISGGTVEVKATESLTTQ  
 SNSKIKATTGEAN(1)VTSATGTIGGTISGNTVN(2)VTANAGDLTVGNGAEIN(3)ATEGAA  
 TLTSSGKLTTEASSHITSAKGQVN(4)LSAQDGSVAGSINAAN(5)VTLN(6)TTGTLTTVK  
 GSNIN(7)ATSGTLVINAKDAELNGAALGN(8)HTVVN(9)ATNAN(10)GSGSVIATTSSRVN  
 (11)ITGDLITINGLNIISKNGINTVLLKGVKIDVKYIQPGIASVDEVIEAKRILEKVKDLSD  
 EEREALAKLGVSAVRFIEPN(12)NTITVDTQNEFATRPLSRIVISEGRACFSNSDGATVCV  
 NIADNGR

Figure 25 – Representation of HMW1(Glc) expression and purification starting from recombinant E. coli. Adhesin sequence is indicated in the red square: His<sub>6</sub>-tag, inserted Trp residue for concentration determination, and the twelve Asn residues in glycosylation sites are highlighted.

A nucleotide sequence encoding for a His<sub>6</sub>-tag, i.e., a N-terminal six-histidine fusion tag, is contained in the vector for HMW1ct expression. Polyhistidine tags are the most widely used affinity tags for purifying recombinant proteins because of the great affinity of histidine toward Ni<sup>2+</sup> that allows purification by affinity chromatography, and display advantages such as low immunogenicity and small size<sup>[118]</sup>. A tryptophane residue is also inserted after the His<sub>6</sub>-tag (replacing G<sup>1204</sup> of HMW1ct) to provide a UV-active moiety for protein quantification.

Due to the poor efficiency of transformation process, the first step is the selection of those cells containing both desired plasmid vectors through mild growth steps in presence of antibiotics. Bacterial glycerol stocks were initially plated on Luria-Bertani (LB) agar plates soils containing both carbenicillin and kanamycin, allowing the growth of isolated colonies on solid media. This process ensures that each cell in a population is descended from a single founder cell, and thus to help ensure that each cell in the culture has the same genetic makeup. Only those populations of cells containing both plasmids (each encoding the gene for antibiotic resistance) are able to grow.

Colony picking was performed manually, transferring cells in small-volume (5 mL) liquid LB media and allowing the overnight growth by shaking at 37°C. This pre-culture step is necessary to avoid cell stress that would originate by the passage from the solid soil to a large volume liquid soil, resulting in cell death and precipitation. Cell cultures were then incubated at 37°C in 1 L media and the optical density at 600 nm (OD<sub>600</sub>) was checked to assess bacterial growth. Typically, bacterial growth curve

displays an initial phase (Lag phase) where cells growing in nutrient rich media are increasing in size but not in number, therefore the OD<sub>600</sub> is low and almost constant. This stage can last for several hours (4-5 h for cells expressing both HMW1ct and the enzyme, whereas was shorter when cultures of cells expressing only the adhesin protein were grown). It is followed by an exponential phase (Log phase), where cell division and expansion provide a rapid increase of the optical density. However, bacterial cell growth reaches a plateau, or stationary phase, where the number of dividing cells equal the number of dying cells, because of the depletion of the available nutrients and the accumulation of waste product. Eventually, in the death phase, the number of living cells decreases exponentially and population growth (and OD<sub>600</sub>) experiences a sharp decline.

Massive production of heterologous proteins would hamper bacterial growth, therefore the optimal stage to induce desired protein expression is in the mid-late exponential phase when the OD<sub>600</sub>  $\approx$  0.6 and bacterial cells have not started to die yet. In our case, this occurred after 6-7 h since the initial incubation (4-5 h when only non-glucosylated HMW1ct adhesin was expressed).

The induction of HMW1ct and HMW1C expression is activated through isopropyl  $\beta$ -D-1-thiogalactopyranoside. This lactose analogue is able to stimulate the operon *lac* transcription machinery, which ultimately controls the plasmid gene expression. After induction, cell suspensions are left stirring overnight at 16°C and harvested by centrifugation. Mechanical lysis of the cell membrane was obtained by using an ultrasonic processor that disrupt lipid layers and provoke the release of soluble components, including overexpressed HMW1(Glc) adhesin.

The purification of the obtained solution was performed by FPLC using prepacked columns in two steps. The first one, by immobilized metal affinity chromatography (IMAC), is based on the high affinity of His tag toward divalent nickel ions.

Supports such as beaded agarose can be derivatized with chelating groups to immobilize the desired metal ions, which then function as ligands for binding and purification of biomolecules of interest. One of the most commonly used chelators as ligands for IMAC is nitrilotriacetic acid (NTA), which is then "loaded" with Ni<sup>2+</sup>, providing Ni-NTA resins<sup>[119]</sup>. These solid supports are able to selectively retain polyHis containing proteins, that are subsequently eluted by increasing concentrations of imidazole, which competes with His residues for nickel binding. By flowing the obtained solution after cell lysis through the column, theoretically only HMW1(Glc),

being the only protein with fusion His-tag, should remain bound to the resin. However, when a concentrate solution imidazole was used to remove HMW1(Glc), the eluted fractions analyzed by SDS-PAGE were found to contain a mixture of adhesin protein and HMW1C enzyme (Figure 26). This not surprising<sup>[97]</sup> evidence results from the formation of a stable complex between the two proteins.

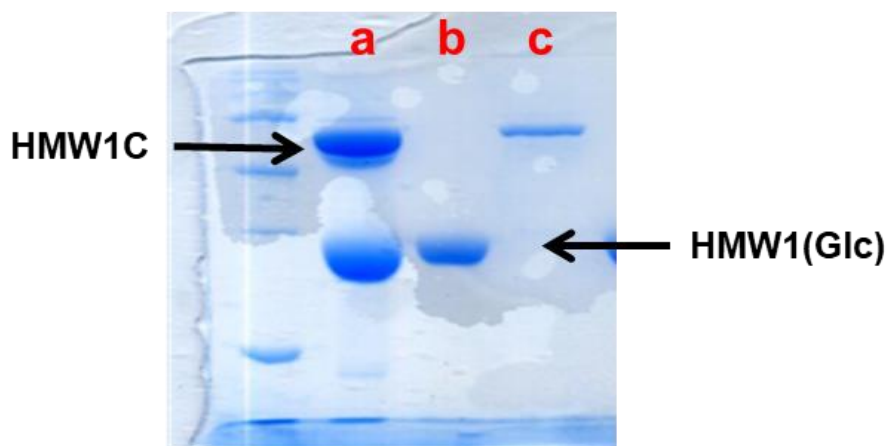


Figure 26 – Picture of stained fractions after SDS-PAGE: a) fraction after Ni affinity column containing both the enzyme HMW1C (71.7 kDa) and hyperglucosylated adhesin (35.8 kDa); b) fraction after anion exchange column containing HMW1(Glc); c) fraction after anion exchange column containing only HMW1C.

Therefore, the separation of HMW1ct(Glc) from HMW1C was obtained in a second purification step through anion exchange chromatography. In this case, cross-linked agarose beads constituting the stationary phase of the prepacked column are functionalized with cationic groups and therefore retains negatively charged molecules by coulombic interaction. The bound molecules are eluted with an anion gradient. Through the increase of ionic strength that breaks the electrostatic interactions between the two proteins, the complex was separated and purified HMW1(Glc) was collected. Obtained fractions were analyzed through SDS-PAGE and protein concentration was determined by UV analysis after buffer exchange and sample concentration.

Hyperglucosylated HMW1ct, CSF114(Glc), peptide **7** and peptide **8** were compared in competitive indirect ELISAs for anti-N(Glc) antibody binding. IgG inhibition was evaluated using representative serum MS1 (Figure 27), whereas for IgM competition serum MS4 was used (Figure 28).

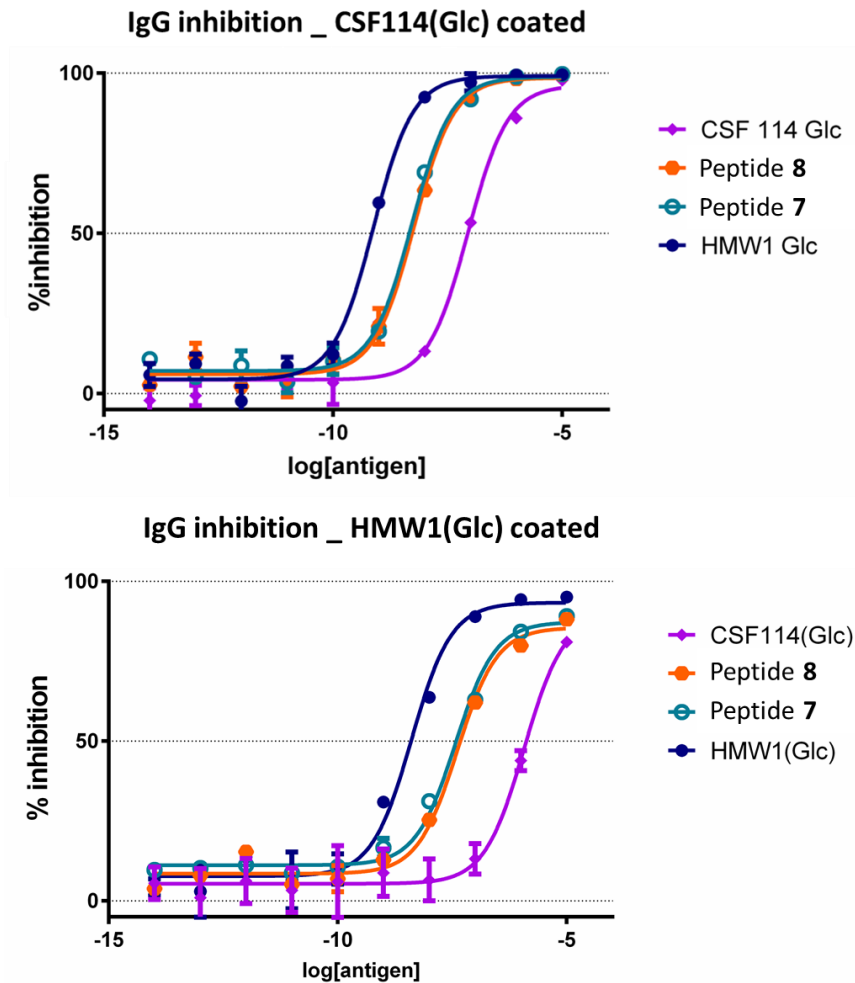


Figure 27 - Inhibition curves of anti-N(Glc) IgG antibodies in serum MSI in a competitive indirect ELISA, using CSF114(Glc) (up) or HMW1(Glc) (down) as coating antigen.

The inhibition curves in Figure 27 show that the HMW1(1347-1355) homodimers have higher affinity for IgGs compared to the previously developed probe CSF114(Glc), despite they do not retain the nanomolar IC<sub>50</sub> of HMW1(Glc) protein. When a multiepitope displaying antigen is presented to polyclonal antibodies that are contained in serum, multivalent interactions may lead to the formation of large, stable (high avidity) structures. This is because the same antigen may be bound by several antibodies, each recognizing a different epitope, and it is probably the case for HMW1(Glc), while peptide 7 and 8 did not display any similar tendency. This clearly highlights that the two epitopes in peptide 7 (or in peptide 8) are too close to interact with two different paratopes, and therefore only one is used to bind an antibody, hampering the availability of the adjacent one.

Moreover, peptides 7 and 8 present the same affinity, meaning that the presence of the two additional N-linked glucoses in peptide 7 is not affecting the interaction with

antibodies. In other words, the higher affinity IgGs are probably recognizing the epitope shared by the two peptides, that is also the same of peptide **2** and **4** (see Figure 15 for IgG competition in the same serum).

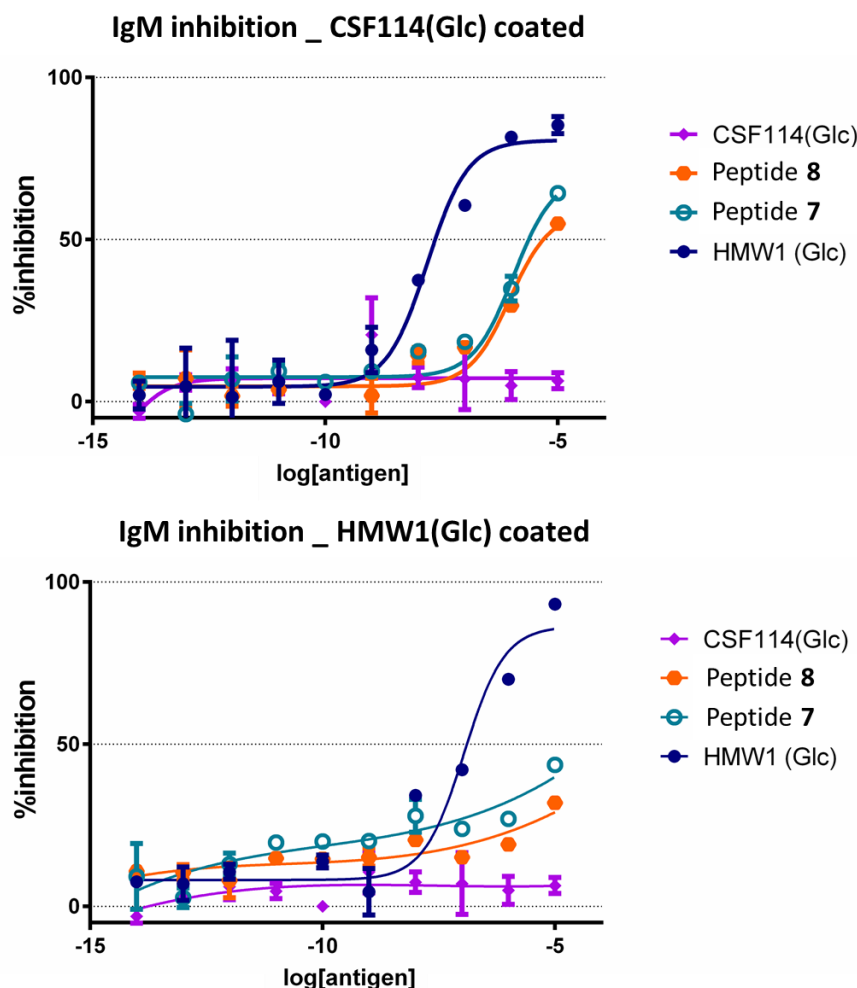


Figure 28 - Inhibition curves of anti-N(Glc) IgM antibodies in serum MS4 with homodimers **7** and **8** in a competitive indirect ELISA, using CSF114(Glc) (up) or HMW1(Glc) (down) as coating antigen.

IgM competition experiments also confirmed that peptides **7** and **8** display the same trend as the corresponding monomers (peptides **4** and **2** respectively, compare Figure 28 and Figure 18). In particular, peptide **7** and peptide **4** have the same IC<sub>50</sub> toward anti-CSF114(Glc) antibodies, which is only slightly better than the one displayed by peptides **2** and **8**.

Calculated mean pIC<sub>50</sub> values for the two novel homodimers are reported in Table 5. When HMW1(Glc) protein was used as coating antigen, these peptides failed to compete for anti-N(Glc) IgM binding in the tested range of concentration (up to 10<sup>-5</sup>

M) and displayed a  $pIC_{50}$  for IgGs that is lower than the one obtained when CSF114(Glc) is coated.

*Table 5 - Calculated  $pIC_{50}$  value for inhibitors toward anti-N(Glc) IgG and IgM antibodies. Values are reported as 95% confidence interval for the calculated mean  $pIC_{50}$  of the four antigens used as inhibitors of anti-CSF114(Glc) and anti-HMW1(Glc) antibodies in representative sera MS1 and MS4.*

Antigen	CSF114(Glc) coated		HMW1(Glc) coated	
	$pIC_{50}$ IgG	$pIC_{50}$ IgM	$pIC_{50}$ IgG	$pIC_{50}$ IgM
<b>HMW1(Glc)</b>	$9.12 \pm 0.17$	$7.97 \pm 0.62$	$8.42 \pm 0.22$	$6.95 \pm 0.68$
<b>Peptide 7</b>	$8.28 \pm 0.13$	$6.08 \pm 0.71$	$7.43 \pm 0.26$	----
<b>Peptide 8</b>	$8.23 \pm 0.16$	$5.55 \pm 0.64$	$7.42 \pm 0.20$	----
<b>CSF114(Glc)</b>	$7.06 \pm 0.37$	----	$5.93 \pm 0.15$	----

As found for glucosylated peptides **2**, **3**, and **4**, they are better peptide candidates for antibody detection than CSF114(Glc), but they do not reach HMW1(Glc) protein performances in antibody binding. Moreover, they do not attain an improvement over the previously described corresponding monomers, probably because the two important epitopes are spatially too close. Therefore, a longer spacing linker or multiple repeated units grafted onto a polymer scaffold may be critical to improve the efficacy in antigen presentation. This pushed us to develop a peptide-polymer conjugate that will be discussed in the next chapter.

## 5. DEXTRAN-BASED TENTACLES

### 5.1 INTRODUCTION

Biosynthetic polymers are materials that combine polymeric scaffolds with bio-oligomers or moieties prepared as mimics of those found in nature. They are now used for a multitude of applications such as novel biomolecule stabilizers, drug-delivery vehicles, therapeutics, biosensors, biomedical adhesives, antifouling materials, and biomimetic scaffolds<sup>[120]</sup>. Biomolecules conjugated to synthetic or natural polymers produced by various grafting strategies enable more and more precise control over advanced architectures, functionalization, and subsequently dynamic function. These bioinspired or fully synthetic polymers act as biopolymer surrogates, executing similar functions and occasionally exceeding the performance of the molecules they mimic<sup>[120]</sup>. In particular, peptide–polymer conjugates can be designed either to benefit from the synergistic behaviour of both the components or to overcome shortcomings inherent to the elements alone. Currently, among the best studied examples of this class of hybrid materials are conjugates of peptides or proteins with poly(ethylene glycol) (PEG). Indeed, PEGylation is now established as a powerful strategy to improve the in vivo properties of therapeutic peptides/proteins, such as solubility, circulation half-life, stability, immunogenicity, and so on<sup>[121]</sup>. Peptide-polymer conjugates make up a new class of soft matter comprising natural and synthetic building blocks. They have the potential to combine the advantages of peptides and synthetic polymers, i.e. the precise chemical structure and diverse functionalities of peptides and the stability and processability of synthetic polymers. The variety of biomolecule-polymer conjugates is immense, as there is flexibility in the length and complexity of the amino acid sequence, the chemical nature, the length and architecture of the polymer, and the overall architecture of the conjugate. The era of “click”-type chemistries<sup>[122,123]</sup>, i.e. CuAAC reaction, oxime, Staudinger ligation, thiol–ene, Glaser reaction, and many more biorthogonal and biomimetic coupling reactions, promoted the access to complex and well-defined peptide–polymer conjugates. Numerous innovative and elegant applications of this class of materials found their way in a huge variety of fields<sup>[124–126]</sup>. However, in spite of the tremendous advances in biopolymer conjugates synthesis, that allowed the generation of conjugates of various sizes, compositions, and architectures described in literature<sup>[120,121,127]</sup>, a few synthetic hurdles still remain to prepare hybrid

conjugate materials. Achieving complex conjugates is challenging, mainly because they are difficult to purify, which limits the quantity and yield of the target conjugate and increases the cost and time of production<sup>[121]</sup>. In particular, standard purification and characterization methods, such as RP-HPLC and MALDI-TOF or ESI MS, respectively, are not trivial for large and sophisticated conjugates<sup>[121]</sup>.

The development of a tentacle-like, antigen-decorated polymer is a promising stratagem for the isolation and characterization of reactive antibodies found in the sera of patients suffering from autoimmune diseases. Specifically, the synthesis of peptide-grafted polymers could finally provide the purification of IgM from patients' sera, that is an unmet need in chemical immunology. These antibodies are considered promising tools for therapy and diagnostic approaches in many pathologies<sup>[128]</sup>. However, the commercial success of IgMs is hampered due to bottlenecks in recombinant production and downstream processing. IgMs are large, complex and highly glycosylated proteins that are only stable in a limited range of conditions. Because of the antibody size and complexity, IgM purification is much more difficult than IgG. The industrial processes of IgG purification from human serum are mainly focused on Protein A or Protein G affinity chromatography, whereas traditional methods for IgM purification are based on the combination of precipitation and chromatographic techniques, including size-exclusion chromatography, ion-exchange chromatography and others, but with scarcity in yields and purity<sup>[129]</sup>. Therefore, prompt efforts are required to investigate these sensitive IgM antibodies.

Although the coating of a solid support with native antigenic proteins or peptides could be envisaged in the effort of detecting and isolating antibodies, a synthetically accessible microarchitecture that is able to snare circulating autoantibodies has enormous advantages and may provide a great step forward both in the treatment and in the comprehension of many autoimmune diseases. Since antibodies are at least bivalent, higher affinity recognition can be achieved through avidity effects in which a construct containing two or more copies of the ligand engages both arms of the monomeric immunoglobulin simultaneously. The immobilization of multivalent antibody ligands at high density on solid surfaces, such as ELISA plates or affinity column stationary phases, is of pivotal interest in medicine and immunology.

Regarding MS disease, an urgent commitment is to fish out and characterize IgM from sera, since these antibodies are reckoned as potent diagnostic tools and their therapeutic value must be elucidated<sup>[83,85]</sup>. With this aim, the synthetical task required for the

construction of antigen-based tentacles consists of three steps: 1) functionalization of the polymer with selectively reactive moieties 2) functionalization of the peptides by adding a sticky end and 3) the final coupling reaction between the two components.

The primary goal is to exploit multivalent interactions between antibodies and glucosylated asparagine in candidate antigens. Conjugating a macromolecular scaffold to short peptide sequences could increase the affinity of antibodies towards these constructs. In fact, this system could roughly reproduce the native framework of membrane-exposed hyperglucosylated proteins, with the crucial, additional advantage that the epitopes are not installed on a fixed, conformationally stabilized tertiary structure but instead the global arrangement is completely free to adapt to antibodies, ideally exposing glucosylated asparagines on side-branches. Moreover, the lack of a protein core could, on the one hand, increase stability and resistance to proteases and on the other hand dramatically reduce nonspecific interactions with the multitude of antibodies and proteins present in patients' sera.

To fulfill this goal, the polymeric scaffold must meet several essential features. First, hydrophilicity and high solubility are required for both experimental and biomimetic purposes. Plus, low viscosity and minimum jellification are desirable for free diffusion of antibodies and macromolecules through the 'tentacles-rich' environment. Indeed, the macromolecular arrangement must be stable as longer as possible, i.e. resistant to proteases and other degrading enzymes present in complex biological media. To reduce nonspecific interactions, charges should be avoided. For entropy reasons, the polymer should not be excessively flexible like PEG, but still readily adaptable to paratope shape.

Finally, readily accessible in large molecular weight polymers (i.e. either easily synthesizable or commercially available) are desirable for antibody trapping and isolation.

Among the commercially available polymeric scaffolds that can be found in large size and can be easily functionalized for the conjugation to reactive groups on peptide epitopes, we selected dextran because it is a high molecular weight, inert, water-soluble polymer that has already been used in a wide variety of bio-medical applications. Traditionally, dextran has been used for decades for plasma extenders, cell/organelle separation, protein precipitation, platelet aggregation inhibitor. More recently this class of polysaccharide has found substantial applicability in modern biotechnological investigations such as drug delivery and imaging<sup>[130–132]</sup>. Dextran is a natural polymer

of glucose, composed of approximately 95% linear  $\alpha$ -D-(1-6) linkages. The remaining  $\alpha$ (1-3) linkages account for the branching of dextran. Dextran is found as bacterial extracellular polysaccharides. They are synthesized from sucrose by beneficial lactic acid bacteria, such as *Leuconostoc mesenteroides* and *Lactobacillus brevis*, but also by the dental plaque-forming species *Streptococcus mutans*. Bacteria employ dextran in biofilm formation or as protective coatings, e.g., to evade host phagocytes in the case of pathogenic bacteria. These highly water-soluble polymers are available commercially as different molecular weights (MW), from 1 to 2000 kDa, with a relatively narrow MW distribution. Additionally, dextran contains a large number of hydroxyl groups which can be easily conjugated to drugs and proteins by either direct attachment or through a linker<sup>[133]</sup>.

Interestingly, conjugation of three different ligands to dextran provided an approximately 1000-fold increase in affinity of the dextran conjugate for the specific antibody compared to each monomer, whereas ligand dimers with PEG spacers were shown to be unreliable as high affinity ligands<sup>[134]</sup>.

With the aim to develop a general and reliable method to rapidly transform modest affinity anti-N(Glc) antibody ligands into much higher affinity constructs by taking advantage of avidity effects, the part of work discussed in this chapter is focused on the preparation and characterization of antigen-carrying, dextran-based tentacles.

## **5.2 RESULTS AND DISCUSSION**

### **5.2.1 Functionalization of dextran**

Dextran can be functionalized in several ways to obtain reactive moieties onto which different types of cargos, i.e. peptides, proteins, cyclodextrins and small molecules can be grafted<sup>[135–137]</sup>.

We decided to start from dextran with an average molecular weight of 40 kDa (in the middle of the range of soluble dextran), that could ideally adapt to IgM structure. For the multivalent presentation of the N-Glc epitopes, we opted for CuAAC reaction between glycidyl-propargyl functionalized dextran (Dex40-GP) and azide derivatives of glucosylated HMW1(1347-1355) fragment.

The choice of alkyne-grafted dextran and azido-bearing peptides was more attractive than the reverse since alkyne groups are more stable and it is possible to prepare and

store a large amount of polymer onto which different azido-substituted molecules can be attached.

Dextran of 40 kDa was activated with alkynes through the epoxide opening of glycidyl propargyl ether (GPE) in basic aqueous solution (Figure 29), according to a procedure already described in literature<sup>[136]</sup>. The glycidyl spacer does not impair the solubility and the linker is sufficiently long to presumably provide a suitable distance between the grafted peptides and the glucan backbone. Moreover, the ether bond is hydrolytically stable. Basic water is used as solvent to avoid homopolymerization of the GPE, which is applied in excess due to a competing hydrolysis by hydroxide ions.

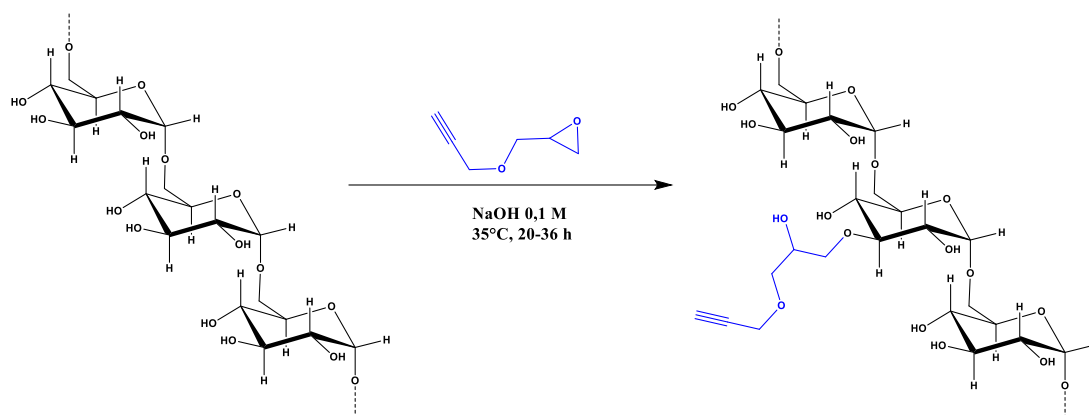


Figure 29 – Reaction scheme for the functionalization of dextran with glycidyl-propargyl groups. Out of simplicity, final product (**Dex40-GP**) is represented as a homogeneously O3-modified construct.

Approximating dextran as a linear polymer composed of  $\alpha$ -1,6 linked glucose units, GPE can react with one of the three free hydroxyl functions (O2, O3, O4), and the literature is not unanimous about which of those is the one carrying the new ether bond<sup>[136,138,139]</sup>. The less hindered position should be O3, however the H1 chemical shift of the newly modified units likely corresponds to O2/O4 functionalized glucose (Figure 30). Nevertheless, for simplicity's sake the modified dextran is always represented as in Figure 29.

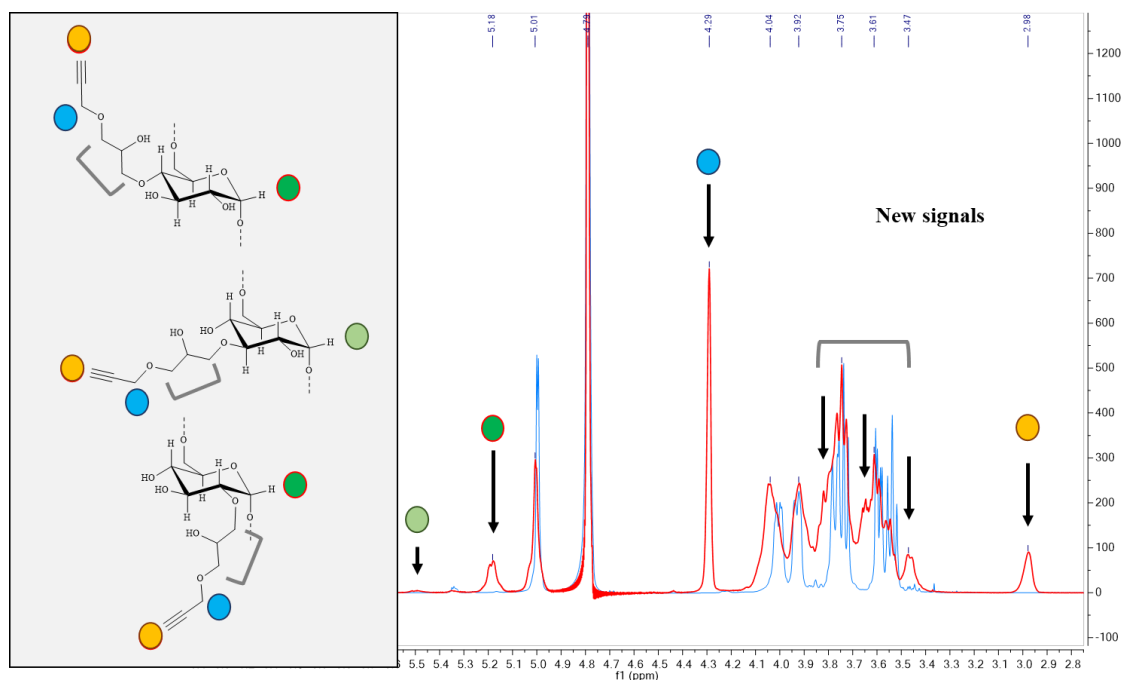


Figure 30 – Superimposition of  $^1\text{H-NMR}$  spectra of commercial dextran 40 kDa (blue spectrum) and functionalized Dex40-GP (red spectrum) evidences new signals appearance that were assigned to corresponding protons of the GP-functionalized glucose units.

The degree of substitution (DS) was calculated by  $^1\text{H-NMR}$  (Figure 31). By the ratio of integrated signals of alkyne proton (2.98 ppm) and anomeric protons (5.60 - 4.95 ppm range), it was possible to retrieve the percentage of propargyl moieties (terminal alkynes) per dextran molecule, that is 29%. The DS in modified units was calculated by the ratio of new H1 signals of modified glucoses and total anomeric protons area, showing a DS of 31%. This tiny discrepancy indicates that not all the modified units include a terminal alkyne in the hanging moiety. Moreover, the area of signal at 4.29 ppm (methylene protons of propargyl group) and the total area of peaks between 4.15-3.35 ppm are both higher than expected. This unreported outcome was ascribed to the minor presence of GPE oligomerization and alkyne-alkyne coupling.

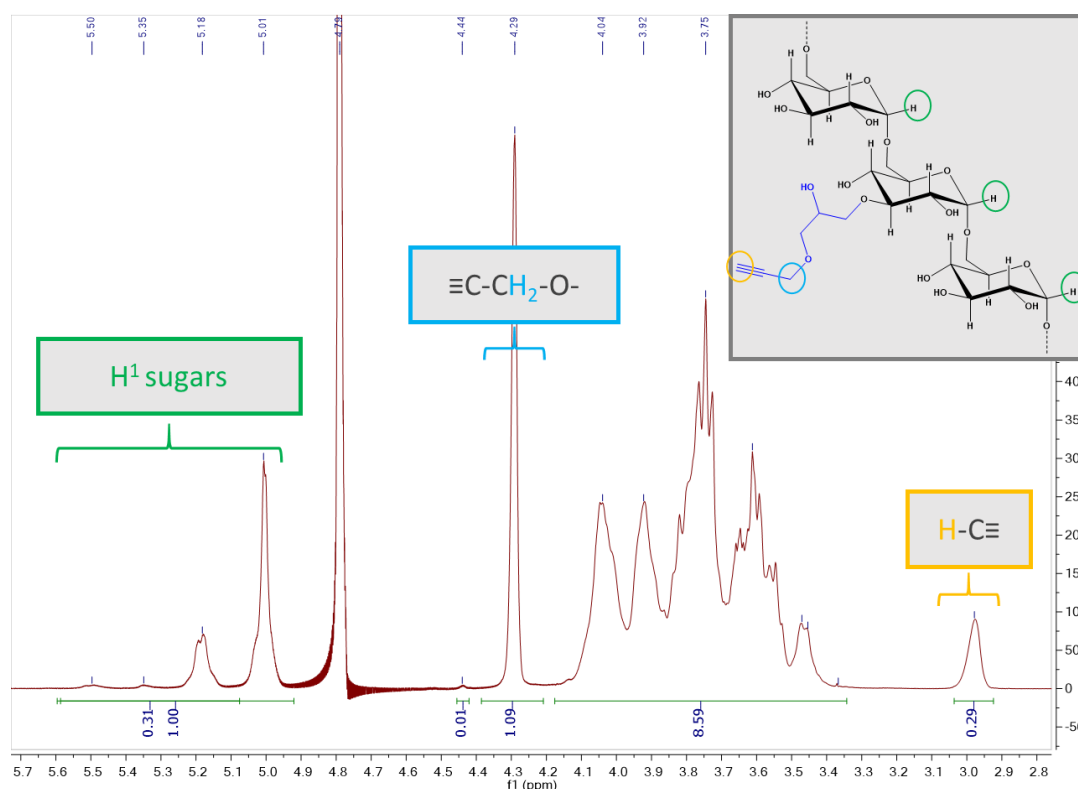


Figure 31 – <sup>1</sup>H-NMR spectrum of Dex40-GP, with the integral of relevant signals that were used to characterize the novel scaffold.

Although it is unlikely that terminal alkyne is deprotonated in reaction conditions ( $pK_a \approx 26$ ), our hypothesis is that the GP alkynes partially reacted together, possibly because of the presence of highly reactive epoxide groups or thanks to favorable steric conditions (Figure 32). Therefore, we concluded that the area of signal at 4.29 ppm accounts for methylene groups linked to both terminal and coupled alkynes, and from its value we could retrieve another fundamental parameter: the average number of GP groups/glucose-GP unit ( $n_{GP} = 1,8$ ). This value is also in agreement with the integration of peaks in the 4.15-3.35 ppm range (see experimental section, chapter 7.4.3 for further information).

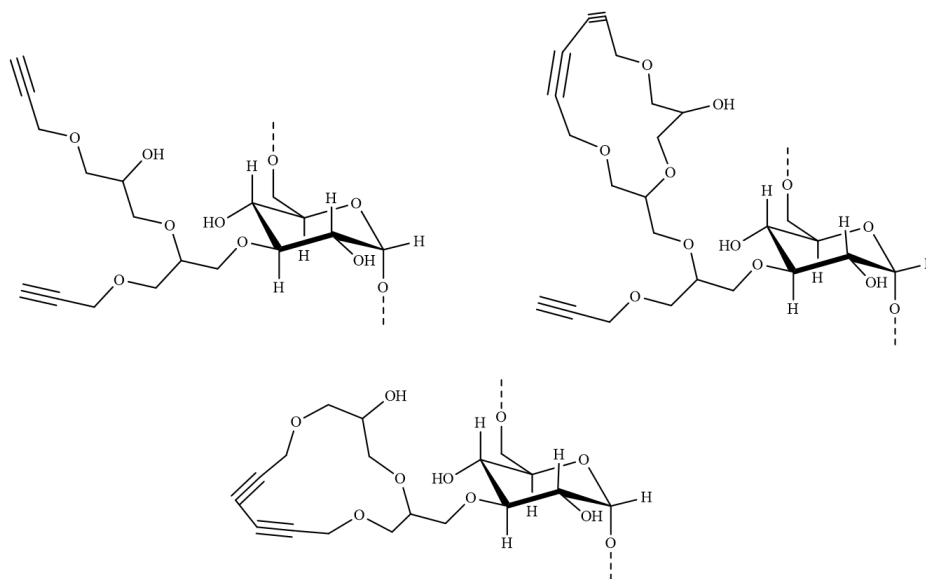


Figure 32 – Example of minor forms of modified units that can result from oligomerization of GP groups ( $n = 2, 3..$ ) and alkyne-alkyne coupling.

Moreover, the  $^{13}\text{C}$ -NMR spectrum evidenced the presence of shifted minor signal that may be due to the presence of different types of alkynes (Figure 33). Although the exact assignment is only hypothetical, the number of  $^{13}\text{C}$  signals is in agreement with forms such as those in Figure 32, and this is a further evidence supporting our conclusions.

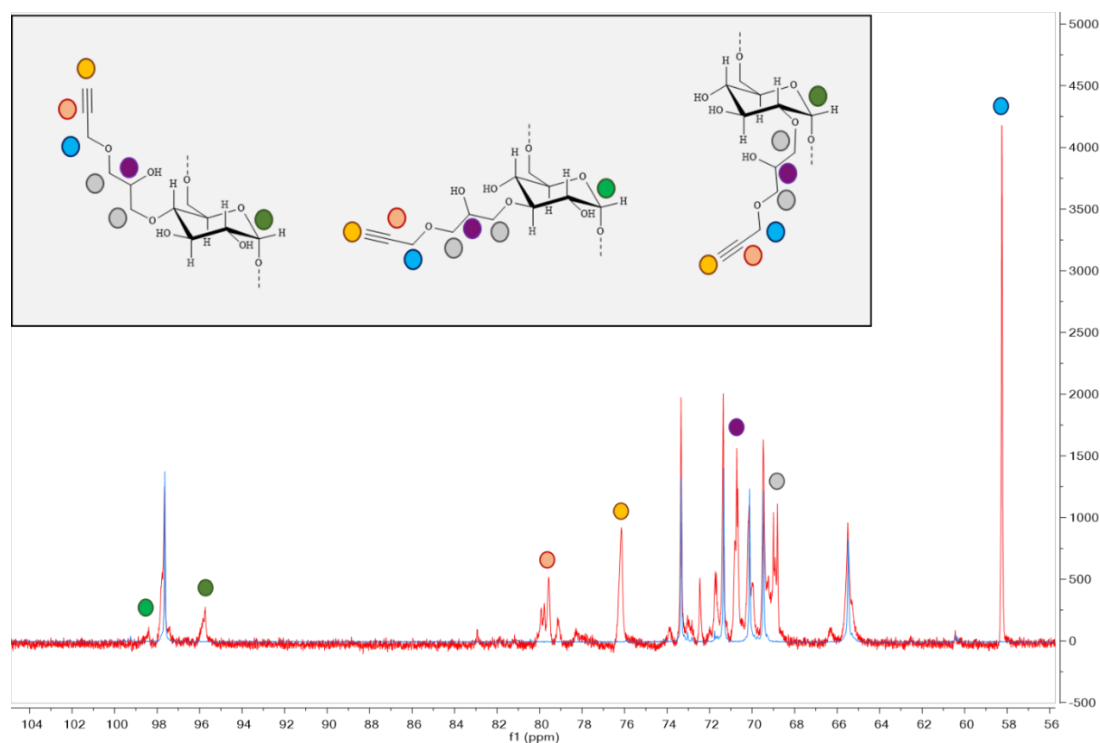


Figure 33 –  $^{13}\text{C}$ -NMR spectrum of Dex40-GP (in red, superimposed to Dex40 spectrum in blue) shows that beside mono-GP modified units other forms do exist. The presence of shifted  $^{13}\text{C}$  signals supports the hypothesis that a minor occurrence of oligo-GP modified units and/or alkyne-alkyne coupling is present.

Considering that each molecule of dextran contains an average of 246.9 glucose units, and that it is unlikely that more than one GP groups reacted with the same unit, 170.4 units remains as simple  $\alpha$ -1,6 glucoses (MW=162 Da) and 76.5 units are  $\alpha$ -1,6 glucoses modified with 1.8 GP groups (MW= 113 Da) each on average. Among these, 71.6 terminal alkyne protons are present. The estimated molecular weight of Dex40-GP, according to these conclusions, is 55.2 kDa.

### 5.2.2 Dextran-peptide conjugates

With the aim to graft peptide sequences to alkyne-functionalized dextran, azide-bearing peptides were prepared to be used as substrates in CuAAC “click” reaction. The selected sequence is the diglycosylated HMW1A (1347-1354) peptide (peptide **4**) already discussed in chapter 3, as it provided the best binding abilities for IgG and IgM antibodies.

At first, we envisaged to prepare N-terminal 4-azido-benzoylated peptides. Instead of acetic anhydride, commercial 4-azido-benzoyc acid ( $pN_3Bz-OH$ ) activated with HATU/DIPEA was used to perform the N-terminus capping prior to final deprotection and cleavage. This strategy, which has never been described before for peptides, allowed the obtainment of peptides **9** and **10**. An advantage of obtaining such compounds are the low-price availability of  $pN_3Bz-OH$  compared to azide-modified amino acid such as the commercially available but more expensive Fmoc-Lys( $N_3$ )-OH. Moreover, the aromatic ring can be exploited as UV-active probe as an additional characterization for conjugates.

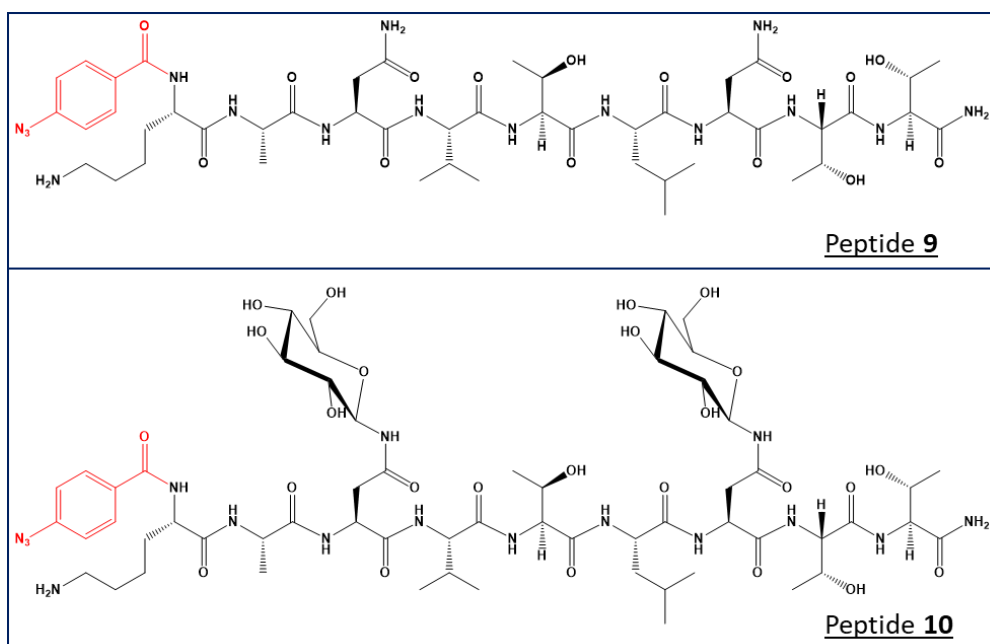


Figure 34 – Structure of peptides **9** ( $pN_3Bz$ -KANVTLNNT-NH<sub>2</sub>) and **10** ( $pN_3Bz$ -KAN(Glc)VTLN(Glc)TT-NH<sub>2</sub>). N-terminal 4-azido-benzoyl group is colored in red.

We also synthesized 4-azido-*N*-(3-hydroxypropyl)-benzamide (**IV**) with a double purpose: on one hand, the compound was used in a CuAAC reaction model with propargyl alcohol, leading to the obtainment of 4-(4-(hydroxymethyl)-1H-1,2,3-triazol-1-yl)-*N*-(3-hydroxypropyl)benzamide (**V**).

This compound, whereas critically far from the polymer structure we aimed to obtain, has the same triazolyl-benzamide moiety as peptide-grafted dextran, thus providing a compelling model for UV analysis.

Secondly, **IV** was clicked to Dex40-GP to optimize the CuAAC reaction conditions with the prepared polymer, both to maintain the amide bond as in peptides **9** and **10** and to avoid the presence of negative charges by using  $pN_3BzOH$ .

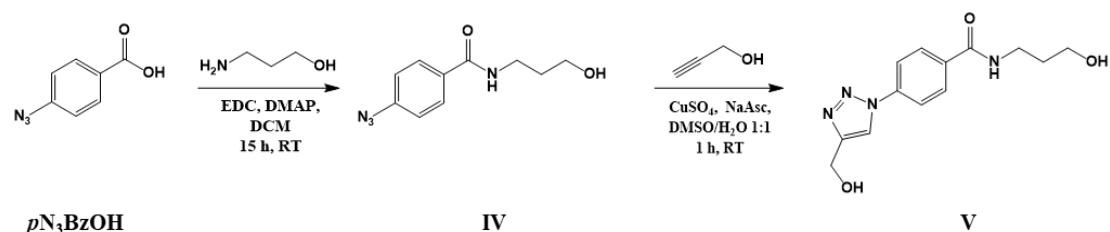


Figure 35 – Synthesis of compounds **IV** and **V** starting from commercial 4-azido-benzoic acid ( $pN_3BzOH$ ). (EDC = 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide; DMAP = 4-Dimethylaminopyridine; NaAsc = Sodium Ascorbate)

Since its discovery in 2002<sup>[140,141]</sup>, the success of CuAAC has been outstanding for myriad of applications, thanks to its selectivity, efficiency and simplicity. Indeed, this type of ‘click’ reaction is highly versatile and can be performed under a variety of reaction conditions including various solvents, a wide pH and temperature range, using different copper sources, with or without additional ligands or reducing agents, and in the presence of other functional moieties. For bioconjugation purposes, this powerful ligation method has been reported to address the numerous challenges of dealing with peptides, proteins, glycans and polynucleotides. However, in spite of the many advancements that have been made in the field, in some cases several issues must be taken into account to optimize CuAAC reaction, especially when dealing with biomolecules, mainly because of their fragile nature and the low amounts at which they are generally manipulated<sup>[142,143]</sup>. Difficulties in CuAAC involve undesired side-reactions such as alkyne cross-couplings by Glaser, Straus, or Eglinton mechanisms, azido stability to reaction conditions, the formation of highly reactive oxygen species (ROS). Steric factors and electronic effects are also reckoned to play a role in the performance of this click chemistry<sup>[144]</sup>.

Several attempts to produce dextran conjugates by using the aryl azide derivatives (compound **IV**, peptides **9** and **10**) lead to disappointing results.

Different conditions were used to optimize the reaction conditions, including different solvent mixtures (DMF/H<sub>2</sub>O 1:1, DMSO/H<sub>2</sub>O 1:1 or 2:1 ratio) and increasing excess of CuSO<sub>4</sub>/sodium ascorbate catalyst. Unfortunately, prompt formation of sticky aggregates and unknown precipitation was always observed. Different work-up procedures, such as centrifugation, dialysis and size-exclusion chromatography, were also employed, but none of the obtained products could be fully characterized by UV, MALDI-TOF or NMR analysis.

The reason of this lack of success could be ascribed to the formation of highly reactive species such as nitrenes. Azido compounds are known to fragment under light or at elevated temperature and to generate nitrenes<sup>[145,146]</sup>. Moreover, aryl azides were observed to undergo a copper-catalyzed reduction through the formation of a nitrene intermediate, leading to formation of a mixture of amines and sulfoxide conjugates, when an excess amount of DMSO was present<sup>[147]</sup>. In our experience, aggregate precipitation was observed even performing reaction in the dark. Future attempts could involve the use of additives and Cu(I) stabilizers, such as nitrogen-type donors,

including bases such as DIPEA and 2,6-lutidine and cosolvents such as acetonitrile, that may help to prevent degradation of CuI by oxidation or disproportionation<sup>[144]</sup>. Conversely, when peptide **6** was used for CuAAC conjugation to Dex40-GP, no precipitation occurred (Figure 36). The reaction proceeded smoothly in previously tested conditions that were found not suitable for aryl-azide compounds. Peptide **6** was the limiting reactant in anticipation of exploiting the unreacted alkyne moieties of the dextran-peptide conjugate for the CuAAC reaction to azido-functionalized surfaces or molecules.

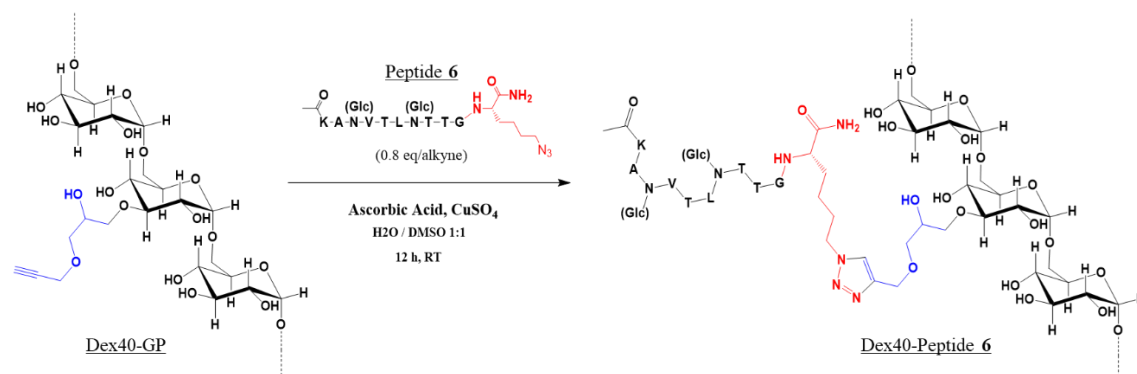


Figure 36 - Click reaction scheme involving the conjugation of diglucosylated peptide **6** to Dex40-GP, leading to the obtainment of Dex40-Pept**6**

After a thorough work-up for copper removal, the obtained Dex40-Pept**6** was characterized by NMR spectroscopy. To facilitate peaks assignment and spectra interpretation, <sup>1</sup>H and <sup>13</sup>C-NMR of peptide **6** were also recorded (see experimental part for further information). <sup>1</sup>H-NMR spectra of Dex40-pept**6** provides crucial information (Figure 37).

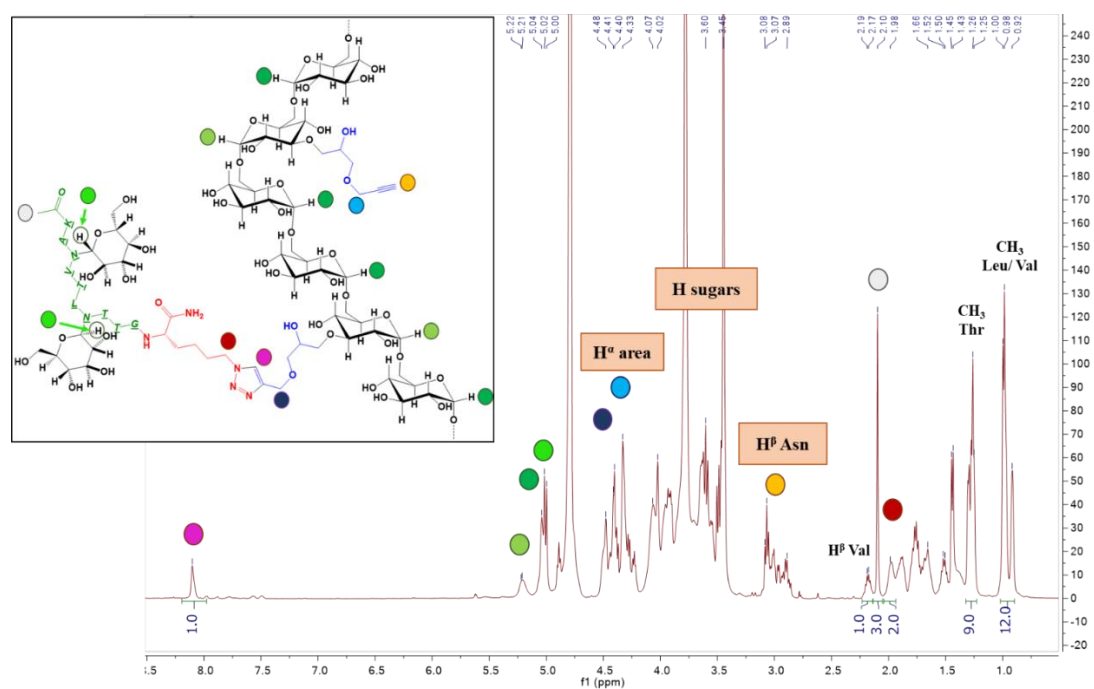


Figure 37 –  $^1\text{H-NMR}$  of Dex40-pept6. Important signals and areas used to characterize the novel conjugate are indicated by colored circles.

By comparing the integrals of the newly formed triazole proton signal (8.10 ppm) and other isolated signals originating from peptide **6** protons, such as  $\text{H}^{\text{b}}$  of valine (2.18 ppm) or methyl protons of acetyl group, threonines, leucine and valine, it is clear that all the peptide component in the sample is covalently linked to dextran backbone (Figure 37). This evidence is important to prove that peptide has not been simply enveloped by the polymer, but the reaction has successfully carried out.

Triazole formation is also evidenced from the shift of  $\text{H}^{\text{e}}$  signal of Lys( $\text{N}_3$ ) from 3.35 ppm to 1.98 ppm. Instead, the signal due to methylene protons of propargyl group (4.29 ppm) decrease in intensity without disappearing, while the new signal of triazole-linked methylene appears at 4.50 ppm. Terminal alkynes signal (2.98 ppm) also decreases without disappearing completely, as expected (see Figure 58 in the experimental section for further information).

Most importantly, using the isolated triazole proton as reference signal, the DS in peptide can be calculated by comparing it with integrals of relevant peaks/ranges, such as  $\text{H}^{\text{1}}$  proton of modified dextran units. The peptide loading is calculated between 19% and 20%, meaning that  $\approx 10\%$  terminal alkynes remain and can be exploited for future applications (see experimental part, chapter 7.4.3 for further information). According to these considerations, each molecule of Dex40-pept6 carries on average 48 peptide branches and it has a final estimated MW of  $\approx 129$  kDa (Figure 38).

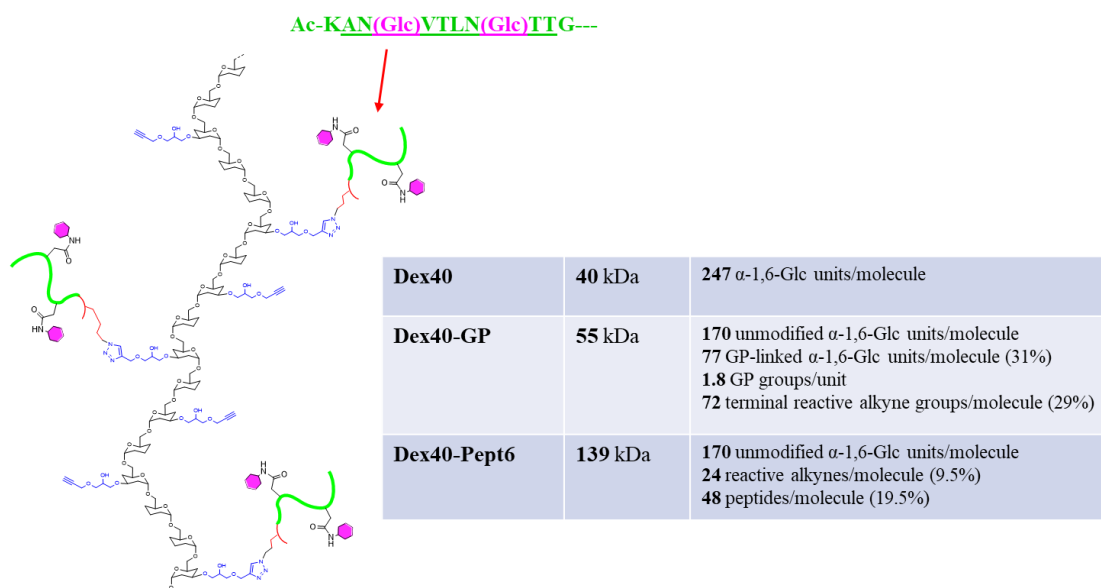


Figure 38 – Schematic representation of the novel conjugate *Dex40-Pept6*, carrying HMW1(1347-1354) diglycosylated epitopes shared by peptides **4**, **5**, **6**, **7** and **10** discussed in this thesis. On the right, main features of dextran and dextran derivatives, based on NMR characterization, are summarized.

### 5.2.3 Applications to antibody detection and capture

The novel peptide-dextran conjugate was used in competitive ELISA tests for a preliminary assessment of antibody binding, using HMW1(Glc) as coating antigen.

As shown in Figure 39, anti-HMW1(Glc) antibody affinity of *Dex40-Pept6* is even greater than the protein itself. Both for IgGs and IgMs, the conjugate displays lower IC<sub>50</sub>, while commercial dextran 40 kDa (*Dex40*) and its alkyne-functionalized version (*Dex40-GP*) do not compete for antibody binding (Figure 39), which can hence be ascribed only to peptide **6**.

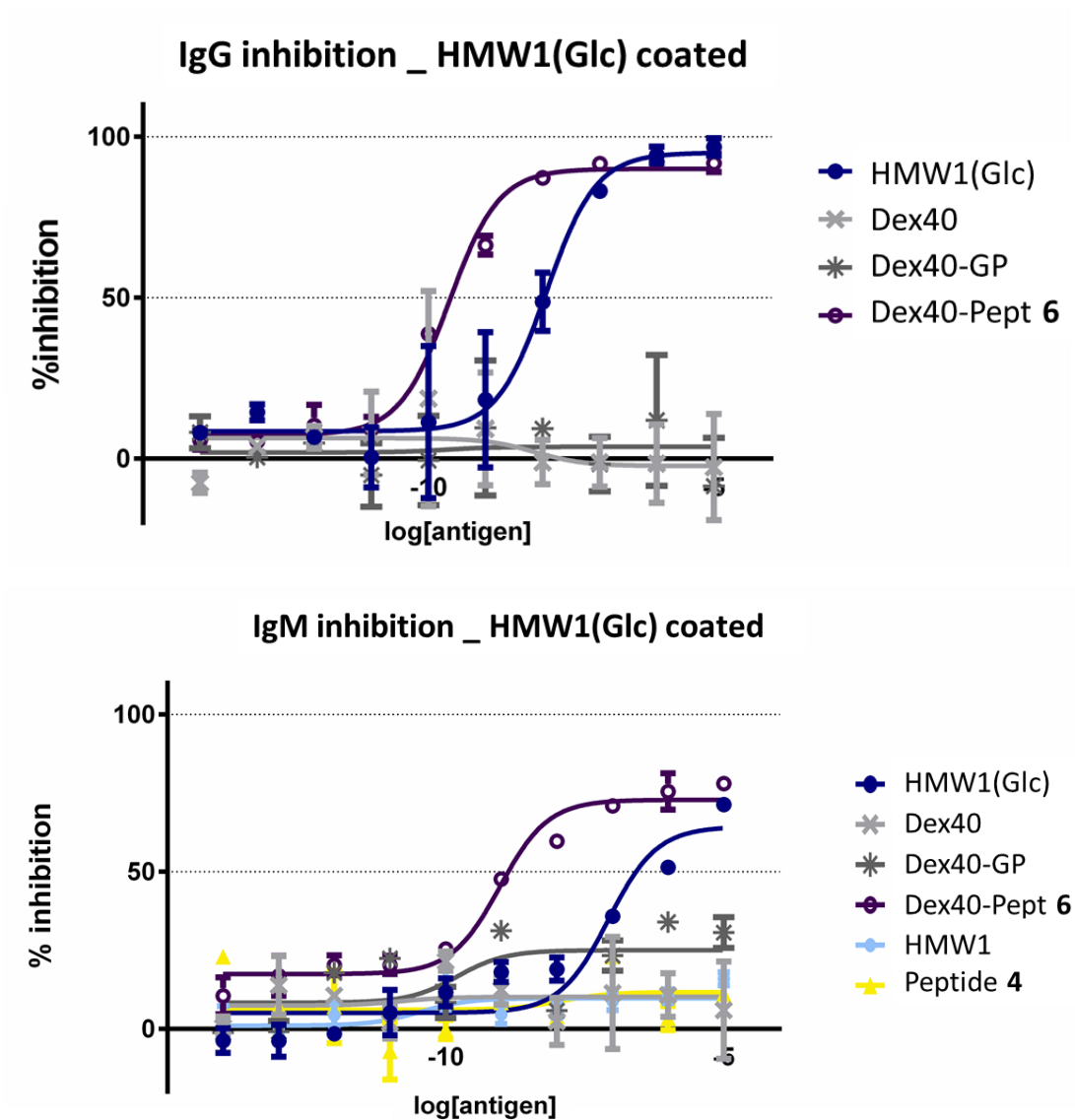


Figure 39 – Inhibition curves of anti-HMW1(Glc) IgG and IgM antibodies in competitive indirect ELISAs.

Although testing more sera will be crucial to assess if this trend is conserved among different patients, this is a first experimental evidence that an octopus-like macromolecule grafted with N-glycosylated peptides can be useful to increase affinity and detect antibodies, acting as a protein surrogate with even better performances. Especially for IgMs, the difference between the two  $IC_{50}$  (Table 6) can be ascribed to an increased avidity because of a multivalency effect. This hypothesis is strongly supported because diglycosylated peptide 4, though displaying the identical epitope of the dextran conjugate, is not able to compete for anti-HMW1(Glc) binding.

*Table 6 - Calculated pIC<sub>50</sub> value for HMW1(Glc) and Dex40-Pept6 toward anti-HMW1(Glc) IgG and IgM antibodies. Values are reported as 95% confidence interval for the calculated mean pIC<sub>50</sub> of in representative sera MS1 (IgG) and MS4 (IgM).*

<b>Antigen</b>	<b>pIC<sub>50</sub> (IgG)</b>	<b>pIC<sub>50</sub> (IgM)</b>
<b>HMW1(Glc)</b>	8.14 ± 0.45	7.07 ± 0.72
<b>Dex40-Pept6</b>	9.70 ± 0.21	9,03 ± 0.39

It is also important to evaluate the possibility to exploit Dex40-pept6 in SP-ELISA for the detection of anti-N(Glc) antibodies in patients' sera. Therefore, preliminary tests were carried out with the aim to find the best conditions to perform indirect SP-ELISA screening of a larger batch of sera. The novel antigenic polymer was successfully coated on polystyrene plates and efficiently detected both IgG and IgM-type antibodies, with similar results regardless of the conditions tested. (see experimental part 7.7). Phosphate buffer saline (PBS) has been selected as the best coating buffer and fetal bovine serum (FBS) diluted 1:10 in NaCl 0,09% solution has been chosen as blocking buffer for future tests involving dextran-peptide conjugates. Measured absorbance values at 405 nm using these same conditions for peptide **4**, peptide **7** and Dex40-pept6 are reported in Figure 40.

Five MS sera were used: MS1 and MS2, already tested for IgG competition, were previously found to have high titer of anti-HMW1(Glc) IgG antibodies, whereas lacking in IgMs. Conversely, MS4 and MS5 were found to have lower IgG titer compared to the IgM one. Finally, MS6 had high titer of both types of antibodies in experiments where HMW1(Glc) was the coated antigen.

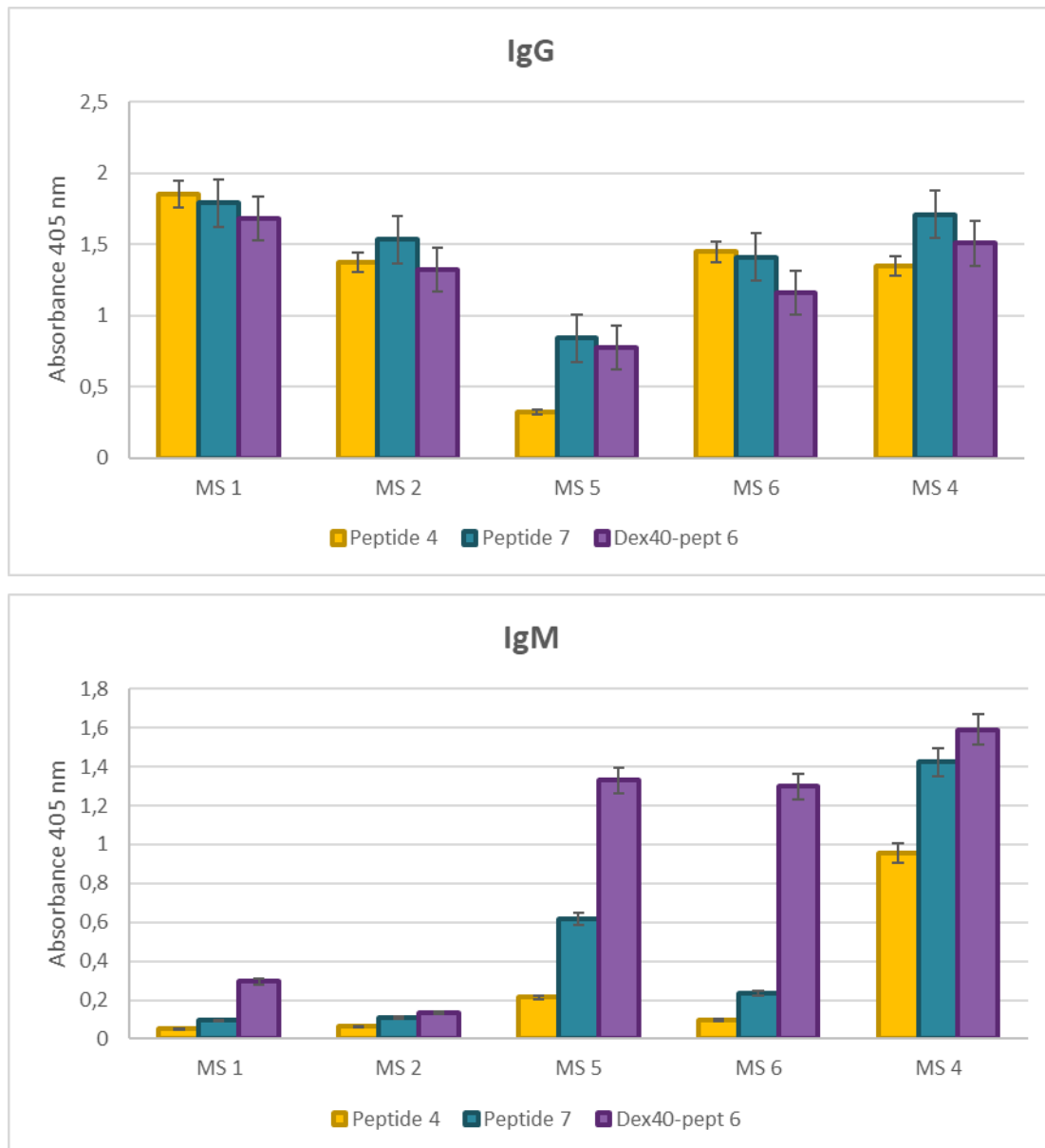


Figure 40 – Comparison of IgG (up) and IgM (down) detection in five MS sera using the three different antigens (diglycosylated peptide 4, homodimer peptide 7, and Dex40-pept6), as determined by SP-ELISA using PBS buffer as coating buffer and FBS buffer as blocking solution. Mean absorbance values at 405 nm are reported, and standard deviation for  $n=3$  independent measures at the same conditions is indicated by error bars.

The results for the three antigens based on diglycosylated HMW1(1347-1354) are in accordance with the previous titration, except for anti-N(Glc) IgG levels in MS4 that were found unexpectedly high. Noteworthy, while IgG detection ability is comparable among the three antigens, IgM capture appears to benefit significantly from the use of a multivalent structure. As already found for short sequences mimicking CSF114(Glc)<sup>[148]</sup>, peptides 4 and 7 were able to detect IgG-type antibodies, but unfortunately displayed a drop in IgM antibodies recognition.

The screening of a larger batch of sera is required to assess the potentiality of dextran-peptide conjugates for IgM detection in SP-ELISA.

Nevertheless, all the experimental evidences strengthen the hypothesis that Dex40-Pept6 is an optimal candidate that can be used to bind antibodies, especially the IgMs. Therefore, anti-N(Glc) IgM-rich serum MS5 was selected to perform preliminary attempts of antibody depletion.

The first step to achieve this goal was the immobilization of Dex40-Pept6 onto CNBr-preactivated sepharose resin beads, allowing to set up the stationary phase for an immunoaffinity-based column. Ideally, the gel substance forming the stationary phase should exhibit mechanical and chemical stability to the coupling and elution conditions, minimal nonspecific interactions with proteins and other sera components and assemble a loose porous network which allows the free flow of large molecules. Polysaccharides such as cross-linked dextran (Sephadex) and derivatives of agarose (Sephарose) have many of these features and therefore are widely used as coupling gels<sup>[149,150]</sup>. CNBr sepharose is a tradename for a crosslinked, beaded-form of agarose, which exposes cyanate ester groups on its surface without an intermediate spacer arm<sup>[151]</sup>. This pre-activated resin is stable in a wide range of pH (2-11) and can be used for coupling peptides or proteins containing primary amino groups by forming an isourea derivative in basic conditions. The coupling reaction is spontaneous and easy to carry out, providing a very convenient way to immobilize ligands through multipoint attachment and resulting in a chemically stable product<sup>[152]</sup>. However, it is important that these multipoint connections do not affect the epitope presentation and thus compromise the effectiveness of the interaction of the immobilized antigen with the antibodies.

CNBr sepharose was successfully used to immobilize both non-glycosylated and hyperglycosylated HMW1ct proteins for sequential immunoaffinity purification of IgG antibodies from sera<sup>[97]</sup>.

By adopting a similar strategy, we incubated 1 mg/mL solution of the novel conjugate /100 mg of beads in basic salted buffer (pH=8.3), allowing the mixture shaking (directly in the column) at room temperature overnight. Each peptide sequence grafted onto dextran chain (48 peptides/chain according to the NMR results) contains a N-terminal lysine residue, therefore ca 0.4  $\mu\text{mol}$  of primary amine group/mg of compound. Because of the difficulty of quantitative measurements at 210-220 nm in the presence of interfering substances, and due to the lack of aromatic residues or UV active moieties, it was not possible to calculate the exact amount of dextran-peptide that was

successfully linked to the beads. . With the aim to elucidate critical features such as the exact percentage of functionalization and, most importantly, the local arrangement of the dextran-peptide conjugate in the gel, future attempts must be done to optimize the control of the coupling reaction step. However, the success of immobilization was assessed by comparing UV spectra of initial solution and eluted reacted mixture. Although the analysis of the solution of Dex40-pept6 before coupling showed the absence of clear peaks (i.e., points of maximum intensity), the absorbance in the same range resulting from the same solution after coupling dropped to zero, and this allowed us to consider the reaction completed.

Unreacted cyanate sites were blocked with a basic saline solution containing an excess of glycine (0.2 M, pH=8).

After thorough washings and equilibration of the functionalized stationary phase, serum MS5 (1 mL diluted 1:10 in saline buffer at pH=7.2) was eluted twice through the column and a small volume of collected fraction was kept for analysis (FT1). The third time, eluted serum was incubated in the column for 1 h at room temperature to promote the binding of kinetically slow antibodies, then final flow through fraction was collected (FT2). Retained antibodies were eluted from the column by using acid glycine solution (pH=2.3). which was immediately neutralized by addition of NaHCO<sub>3</sub> buffer, and their activities were determined by SP-ELISA (Figure 41).

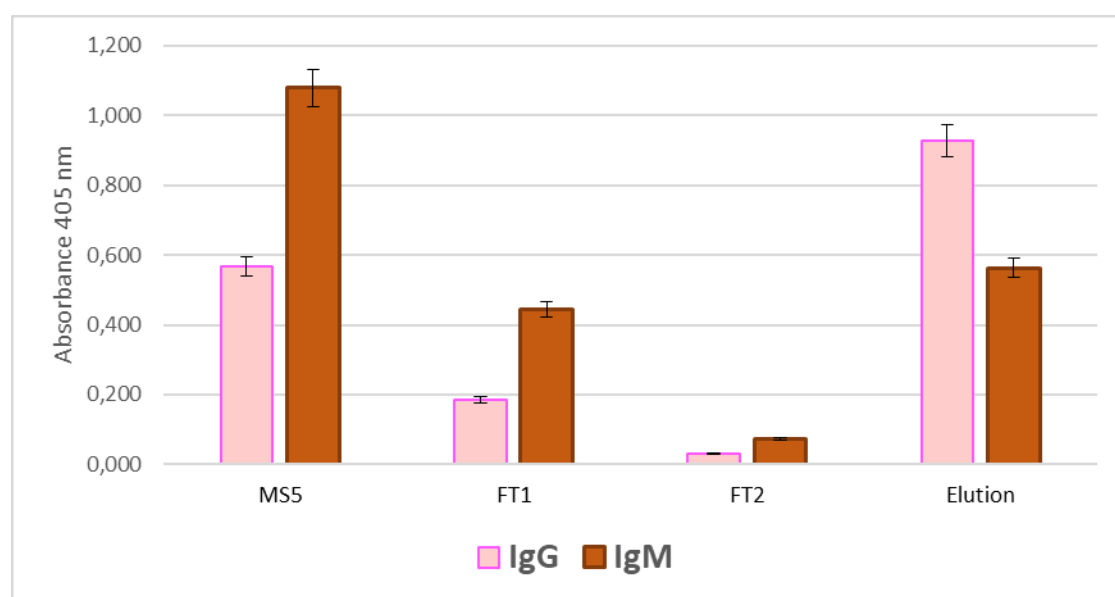


Figure 41 – Immunoaffinity purification of antibodies from serum MS5. Measured SP-ELISA mean absorbances ( $\pm$ standard deviation for  $n = 3$  independent experiments) of IgG and IgM of the fractions obtained by using a sepharose column containing the immobilized Dex40-pept6. Serum MS5, the initial flow through (FT1), the final flow through (FT2) were tested respecting the same dilution ratio, whereas

*eluted fraction containing isolated antibodies (Elution) is approximately 5 times more concentrated. Coated antigen is Dex40-pept6 in PBS buffer; 10% FBS was used as blocking buffer.*

Although highly encouraging, these preliminary results have strengths and weaknesses. Both FT1 and FT2 fractions displayed a remarkable decrease in antibody titer. Particularly FT2 fraction was almost completely depleted from both IgG and IgM antibodies. This assessed the importance of a third, longer-time loading cycle to allow most of the antibodies binding the antigens in the column. Most importantly, it constituted the proof-of-concept that the selective depletion of the anti-N(Glc) antibody component, including IgM type, can be achieved by using Dex40-Pept6.

Unfortunately, antibody content in the eluted fraction was only a small part of the initial component. In fact, initial serum and flow-through fractions were analyzed at the same dilution ratio (1:100 in each well of the plates). Conversely, eluted antibody fraction in each well is  $\approx 5:100$  v:v (see experimental section 7.8 for further information). Additionally, in contrast to the detected immunoglobulin content in MS5, the eluted antibody fraction contained more active IgGs than IgMs. This suggests that the elution conditions are not suitable for IgM elution, that are known to be very sensitive to buffer alterations. Their structure could be affected by the strong acidity of the eluent, hampering their activity and hence their detection. Alternatively, they could have remained strongly attached to the stationary phase of the column, because of the great affinity displayed by Dex40-Pept6 towards anti N-Glc antibodies. In this case, even harsher conditions would be required to isolate and characterize the antibody fraction. Future experiments will confirm the versatility of this novel conjugate as molecular “bait” to fish antibodies from patients’ sera, however these preliminary attempts provided promising hints of its value.

### **5.3 CONCLUSIONS**

In this chapter, we reported the synthesis and characterization of a new dextran-peptide conjugate that holds great potential for future applications. Since multiple interactions are fundamental in many biological processes, the display of bioactive peptides in the polymer side chain can be beneficial for cooperative binding events. Therefore, the assembly of multiple N-glycosylated peptide moieties along a polymer backbone is of crucial interest for the detection and capture of antibodies in sera of MS patients.

Peptide **6**, containing the diglycosylated HMW1(1347-1354) epitope, was loaded onto an alkyne-modified dextran (initial MW = 40 kDa,  $\approx$  247 glucose units/molecule) by CuAAC click chemistry. Quantitative NMR characterization allowed to measure the loading in peptides (19.5%), meaning that each peptide-grafted dextran molecule contains  $\approx$  48 modified units, hence  $\approx$  96 N(Glc) moieties. The concept of these polymeric multiple antigenic peptides was clearly demonstrated. This novel polymeric structure was shown to dramatically increase the binding potency of IgG and IgM in MS sera. The effect of cooperativity resulted in an increase in antigenicity when compared to the free epitope by competitive ELISA tests. Also, the best coating conditions to use Dex40-Pept**6** in SP-ELISA were selected, and the comparison with the analogue peptide **4** and peptide **7** highlighted the enhancement in IgM binding ability.

The natural evolution of this project points toward the exploitation of this synthetic polymer for the isolation and characterization of antibodies. Antibodies from a representative MS serum were successfully depleted on a sepharose resin specifically modified with the adhesin peptide-dextran conjugate, as confirmed by SP-ELISA. This preliminary result is a proof-of-concept of the selective entrapment of circulating autoantibodies (possibly perpetuating nonself recognition) that could hopefully lead to develop a specific apheresis-based device.

However, in order to exploit and characterize the antibody fraction for biological assays, the optimization of the immunoaffinity purification procedure is required since the antibody content in the recovered fraction was disappointingly low.

A critical step is the purification of single-isotype antibodies. As expected, the conjugate Dex40-Pept**6** exhibits weak capability to distinguish between IgGs and IgMs when they are both present. Therefore, with the final aim to purify and characterize IgMs from MS sera, the first step should be IgG removal, by using protein A- or protein G-based media<sup>[153]</sup>. Once the serum is cleaned from IgGs, anti-N(Glc) IgMs could be isolated from the other immunoglobulin and protein components thanks to this novel functionalized resin. On the other hand, recovering intact antibodies is a pivotal issue. Many improvements must be achieved to isolate active autoantibodies and especially the IgMs. Recently, investigational studies of adsorption selectivity of IgM, IgA and IgG by mixed-mode chromatography with specially-designed ligands were performed to improve adsorption selectivity and process efficiency of IgM purification from complex feedstock such as human serum<sup>[129]</sup>. Resin adsorption capacity was

demonstrated to depend critically from pH and a method for IgM, IgA and IgG selective separation was developed by controlling the loading and elution pH. Notably, the presence of elution additives such as 1 M arginine addition was proposed to help improving the selectivity and maintaining the antibody activity of IgM after purification<sup>[129]</sup>.

In our case, human serum containing higher titer of IgM than IgG was flowed through the antigen-functionalized column, and SP-ELISA analysis of the resulting fraction confirmed the depletion of the specific antibody component. However, elution in acid glycine buffer (pH=2.3) afforded a poorly active antibody fraction, meaning either that the antibodies lost their structure during elution (i.e., their activity toward the coated antigens in the plates) or that most of the antibody component remained tightly bound to the column. Moreover, the eluted fraction was richer in IgG than IgM antibodies. In solid-phase, IgM may interact with several peptide-dextran forming a high-avidity antibody-antigen complex that is not trivial to break.

With the aim to set up antigen-specific matrixes for plasmapheresis columns, recycling the stationary phase is a main issue. To regenerate the column by releasing the bound anti-N(Glc) immunoglobulins, harsh pH variations can be used, in the range of the solid phase stability (pH=2-11). Acid treatments to elute the retained components and then basic washings with diluted NaOH allow the reusability of the column for clinical practice. To disrupt the non-covalent antigen-antibody interactions, also non-ionic detergents or denaturing solvents, such as urea and guanidine hydrochloride, can be envisaged since the dextran-peptide conjugate is chemically stable.

Conversely, if the goal is to characterize the specific antibodies by recovering intact and functional immunoglobulins, precautions must be taken, especially when IgM are targeted. The release of immunoglobulins may be achieved in alternative by selectively targeting the dextran scaffold of the antigen-based column, using dextran-hydrolyzing enzymes<sup>[154]</sup>. However, the use of dextranases or other reagents such as sodium periodate would lead to the destruction of the stationary phase and the obtainment of a complex mixture of eluted substances that must be subsequently purified.

Also, quantitative mono-dimensional <sup>1</sup>H-NMR allowed us to characterize the peptide content per macromolecule, but thermal and biophysical properties must be investigated to assess the suitability of the dextran conjugate. When applying these classes of materials to translational medicine, it is important to engineer well-defined materials whose structure and behaviour are well characterized.

A detailed structural characterization may also be important to correlate the three-dimensional conformation of the peptide-polymer chain with antibody binding activity. With fundamental structural and behavioral properties understood, the knowledge gained can then be used to rationally design peptide-polymer conjugates with desired properties. Many parameters can be tailored to achieve a target, including the peptide sequence and length, the chemical nature of the polymer, the length of the polymer, the solvent, and the architecture of the conjugate<sup>[121]</sup>. This first achievement paves the way for future developments involving the production and characterization of new conjugates with different peptide distribution (i.e., the degree of substitution) and polymer size (i.e., dextran molecular weight).

Finally, it is important to note that CuAAC reaction between peptide **6** and Dex40-GP was intentionally performed in order to preserve a minor percentage of unreacted alkynes that could be useful for future applications. The use of 0.8 eq of azide (i.e., peptide)/ alkyne group lead to the obtainment of Dex40-Pept**6** containing 19.5 peptide chains and 9.5 terminal alkynes/ molecule.

Each peptide sequence contains a N-term lysine carrying a primary amine moiety useful for bioconjugation purposes. In our preliminary attempts, we exploited the primary amine groups for the immobilization of Dex40-Pept**6** to CNBr-activated sepharose beads, and this approach allowed us to verify the efficacy our conjugate for antibody removal. Alternatively, the remaining alkynes can be coupled to azide-bearing surfaces or membranes, hopefully providing more favourable antibody-peptide interactions because of the lower steric hindrance around the peptides.

However, the reactivity of alkyne-terminal groups is not restricted to the development of stationary phases for immunoaffinity purification. Several bioconjugation strategies can take advantage of the orthogonality between the alkyne and the amino groups. Indeed, the reaction with other molecules could in principle provide multifunctional architectures. For example, we can envisage the use of lipids for immunostimulation<sup>[155]</sup>, thiol containing molecules for gold nanoparticle-based delivery<sup>[156]</sup> or gold surface functionalization in sensor chips<sup>[157]</sup>, as well as biotin, maleimide and many other functionally reactive derivatives that can be exploited for the detection and characterization of anti-N(Glc) antibodies. In summary, Dex40-Pept**6** represents a chemical playground that allows to conceive complex molecular systems with new and exciting applications in chemical immunology.

## **6. S-ALKYLATION: TOWARD NEW GLYCOPEPTIDE CONJUGATES**

### **6.1 INTRODUCTION**

In the large frame of glycans and glycoproteins involved in autoimmune diseases, the scope of this thesis project was to evaluate the antibody-antigen recognition in patients' sera by exploiting glycopeptide probes, and ultimately to develop a synthetic structure able to trap antibodies with high affinity. As discussed in previous chapters, this general conception was validated in the case of selected MS patients, whose sera contain antibodies targeting N-linked glucosyl moieties on Asn residues.

Nonetheless, anti-glycan antibodies are an abundant subpopulation of serum antibodies with critical functions in many immune processes<sup>[158]</sup> and several glycostructures are known to be specifically recognized in immune disorders<sup>[36]</sup>. Establishing a correlation between antibodies to certain defined mono, di and oligosaccharide conjugates that are common in bacterial, fungal and parasite cells and their relevance in human sera holds a compelling value for diagnosis, prognosis, and therapeutic implications<sup>[36,159]</sup>. As a result, achieving synthetic high-affinity probes to detect and capture specific antibodies would be a great success.

Glycosylation of proteins is the most complex form of co- and posttranslational modification. The determination of structure-function relationships, however, remains problematic because of both the microheterogeneity of glycoproteins that exist in many different glycoforms and their accessibility for investigational studies<sup>[160]</sup>. As it is very difficult to obtain them from natural sources, the chemical synthesis of glycoproteins and glycopeptides with defined glycan structures plays a pivotal role for the detailed determination of the role of protein glycosylation and the design of efficient and reproducible immunoassays. Yet, getting access to large amounts of pure glycoconjugates remains one of the most critical issues in translational research. Although the impressive progresses in chemical and enzymatic methods that allowed the generation of relevant glycoproteins and glycoconjugates, the field is still expanding, responding to the ever-growing demands and challenges of glycoscience<sup>[161]</sup>. Chemical synthesis of glycopeptides is a powerful tool because it allows the large-scale production of homogeneous structures that are able to mimic glycoconjugates' bioactivity. This concept inspired the "chemical reverse approach",

already discussed in the introduction, and its ultimate aim to select and optimize synthetic peptide probes in immune diseases. With the idea that mimicking immunogenic modifications could be even more effective and specific than using native molecules in antibody detection, the screening of focused libraries of unique modified sequences exposing the minimal epitope with the relevant modification may be crucial to detect at the best autoantibodies specific of the autoimmune disease under investigation.

Unfortunately, even the chemical synthesis of glycoconjugates is seldom trivial and expeditious, and the obtainment of large quantities necessary to address the critical medical needs require a lot of efforts. The design of glycopeptides requires a combination of sugar and peptide chemistry, a substantial part being the installation of the glycosidic bond between carbohydrate units and between the saccharide and the peptide chain<sup>[162]</sup>. In fact, the real bottleneck is the obtainment of complex oligosaccharides and their conjugation to the peptide scaffold, since the developments in peptide synthesis has now made readily accessible the large-scale production of amino acid sequences, thanks to the optimizations of solid-phase strategies and to the expansion of powerful methods of fragment condensations of deprotected peptides, i.e. the chemical ligation approaches<sup>[163–165]</sup>.

The synthetic route commonly used for the synthesis of *N*-glycosides exploits the formation of a peptidic bond between a glycosylamine and a protected aspartic acid derivative. As glycosylamine precursors, usually glycosylazides are employed, which can be obtained with the strategy already shown in chapter 3.2.1, by treatment of glycosyl halides with azide salts. However, the obtainment of glycopeptides fully synthesized through SPPS strategy is not trivial for complex conjugates. Peracylated monosaccharides are stable to acid conditions, and Fmoc-Asn[Glc(OAc)<sub>4</sub>]-OH is a suitable building block, as already shown. Also many other *N*-Glycosyl peptides are accessible through SPPS in the case of monosaccharides or simple disaccharides<sup>[41,98]</sup>. Conversely, when dealing with more complex disaccharide or oligosaccharides, two main problems are faced: the chemical incompatibility between sugar and peptide chemistry, such as the poor stability of the inter-glycosidic bond in strongly acid conditions during final cleavage or the presence of moieties sensitive to other SPPS conditions, and the necessity of large quantities of precious glycosylated amino-acid, because of the considerable excess of the activated residue during the coupling. Therefore, different strategies are more often envisaged<sup>[161,166,167]</sup>. These involve the

combination of chemical synthesis, chemoenzymatic synthesis, and bioconjugation strategies to obtain desired glycan targets. As example, the insertion of the first monosaccharide to peptide chain and the enzymatic extension of the glycan moiety is one of the several “mixed” approaches that can be conceived<sup>[162]</sup>. Nevertheless, enzyme accessibility, the small quantities obtained and the impossibility to study aberrant glycosylations that may be involved in immune disfunctions (i.e. the non-native sequences linked to native saccharides or *vice-versa*), are all critical issues to consider. In the chemical solution-phase synthesis of *N*-glycopeptides, it is practical to employ a block glycosylation approach in which the full-length carbohydrate is coupled to peptide side chains<sup>[160]</sup>. Starting from non-glycosylated peptides, the final glycosylation in solution cannot be performed selectively with the acylation of aspartic acid, since most of the times other reactive functions are present. Alternatively, the coupling of glycans to peptides requires the exploitation of selective “click-type” chemistry<sup>[122]</sup>. For example, an often used strategy consists in starting from glycosyl-azides and peptide sequences in which propargylglycine is replacing asparagine to enable CuAAC conjugation<sup>[168]</sup>. Although triazole ring linker is suggested to be isostere of the amide bond<sup>[143]</sup>, the CSF114(Glc) analogue, i.e. [Gly<sup>7</sup>(CH<sub>2</sub>triazolylGlc)]CSF114 displayed negligible inhibitory activity in competitive ELISA and failed to detect autoantibodies in MS sera by SP-ELISA<sup>[41]</sup>. Therefore, synthetic strategies for glycopeptide exploitation in chemical immunology must be carefully assayed, especially when the linkage between peptide backbone and sugar moiety is important for antibody recognition. Nonetheless, with the purpose to use azide-bearing peptides for the development of polymer conjugates through CuAAC conjugation, an additional degree of orthogonality must be envisaged to couple complex glycans to peptide sequences. Thiol substitution to halogen derivatives is another “click” strategy (Figure 42) that offers various practical advantages as it reduces dramatically the synthetic steps and gives an easy access to glycopeptide libraries through a copper-free reaction. Bromoacetamido derivatives are interesting building blocks to link a glycan residue to a peptide sequence by forming a thioether linkage with a free thiol (Cys) residue<sup>[169][170]</sup>. The glucosylated thioether product is not a real mimetic of a glycosyl asparagine, as the distance between the sugar and the peptide backbone is increased, but it retains the essential glycosylated amide function. Moreover, the reaction is compatible with the presence of other functionalities, i.e. it is selective and orthogonal to other conjugation strategies such as CuAAC.

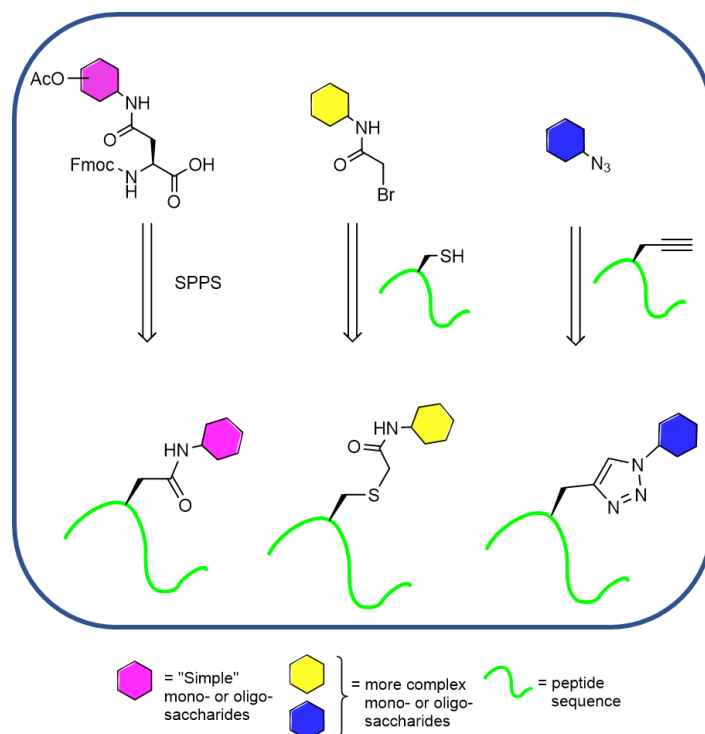


Figure 42 – Three examples of synthetic strategies to obtain N-glycosylated peptide mimetics. For sugars such as Glc, Gal or other simple mono and disaccharides, conveniently protected glycosyl-Asn can be directly inserted into the peptide sequence following the SPPS cyclic scheme. When more sensitive saccharide moieties are employed, e.g. sulphated or branched glycans, other approaches must be considered. These include the replacement of Asn residue with Cys or Pra for site-selective “click” modifications.

With these considerations in mind, in the present final chapter we present our investigational studies to optimize the conditions for this synthetic strategy, together with the synthesis of relevant saccharide building blocks for the development of new glycan mimetics. For their conjugation to peptide or polymer scaffolds, the aforementioned orthogonal schemes will be necessary.

## 6.2 RESULTS AND DISCUSSION

### 6.2.1 Synthesis of relevant glycosyl building blocks

With the aim to optimize reaction conditions for the S-alkylation of cysteine residue in peptide chains, we synthesized the bromo-acetamido derivative of glucose (Glc) and galactose (Gal). The fastest way to synthesize these novel building blocks is a method described in literature based on the use of MsOH and MeCN (v/v 1:4) as solvents for

the obtainment of N-(2,3,4,6-Tetra-O-acetyl- $\beta$ -D-glycosyl) acetamides <sup>[171]</sup>. Unfortunately, in our case starting from peracetylated monosaccharides and following the described protocol using BrCH<sub>2</sub>CN instead of MeCN, the reaction did not produce characterizable products by TLC and NMR analysis. Thus, we envisaged to use a multi-step but more reliable synthesis to obtain the bromo-acetamido derivatives. This protocol involves the well-known carbodiimide activation<sup>[172]</sup> of carboxylic group to form an amide bond with glycosyl amines (Figure 43).

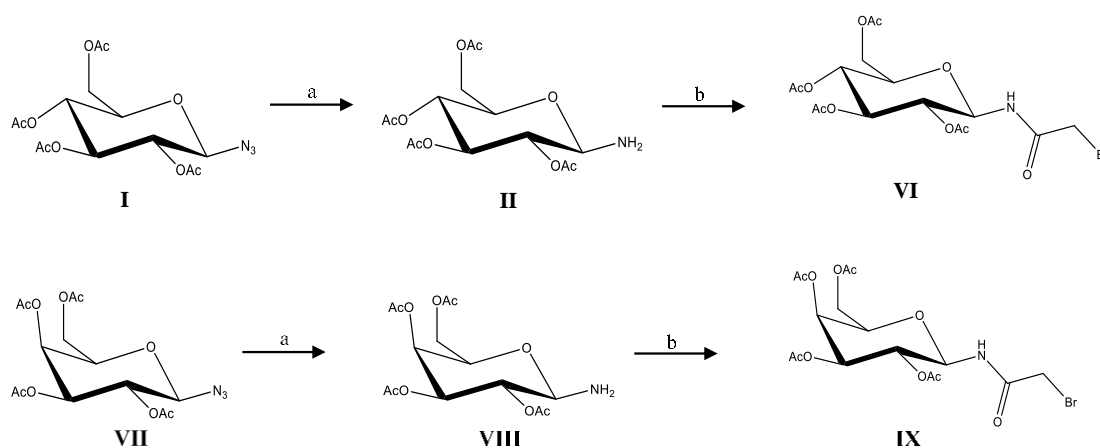


Figure 43 – Synthetic scheme adopted for the synthesis of bromo-acetamide derivatives of Glc (**VI**) and Gal (**IX**). a) H<sub>2</sub>, Pd/C in MeOH; b) 2-Br-acetic acid, DCC, DMAP in DCM

Large-quantity production of glucose derivative **VI** was useful not only to explore the selectivity of the reaction with Cys residue in our model peptide sequence, but also to assess the mimicking ability of thio-glucosylated peptides in reproducing the N-linkage, as discussed in the next paragraph. Galactose derivative **IX** can be useful for comparison in immunological experiments, but its true relevance lies as starting building block for other glycan structures. In fact, some derivatives of galactose are associated to immune disorders. These include several O3-linked derivatives such as HSO<sub>3</sub>-3Gal (sulfatide epitope), in which the sulphated form of galactose plays as fundamental target of prevention of and therapy for nervous disorders, diabetes mellitus, immunological diseases, cancer, and infectious diseases<sup>[173]</sup>. Gal- $\alpha$ 1,3-Gal ( $\alpha$ -Gal epitope) is a nonhuman carbohydrate structure that is known to induce a strong immune response in humans. This disaccharide was identified in several therapeutic monoclonal antibodies and therefore holds potential in establishing correlation with adverse effects of antibody therapies, that may lead to allergies and autoimmunity<sup>[174]</sup>.

Gal is also an essential unit of the the human natural killer-1 (HNK-1) carbohydrate epitope, a unique saccharide possessing sulfated glucuronic acid in a non-reducing terminus (HSO<sub>3</sub>-3GlcAβ1-3Galβ1-4GlcNAc-) that is highly expressed in the nervous system<sup>[175]</sup>. This HNK-1 carbohydrate is the putative autoantigen associated with peripheral demyelinating neuropathy, which relates to IgM paraproteinemia (monoclonal gammopathy), reacting with serum autoantibodies in IgM anti-MAG (myelin-associated glycoprotein) antibody-associated neuropathy patients<sup>[176-178]</sup>.

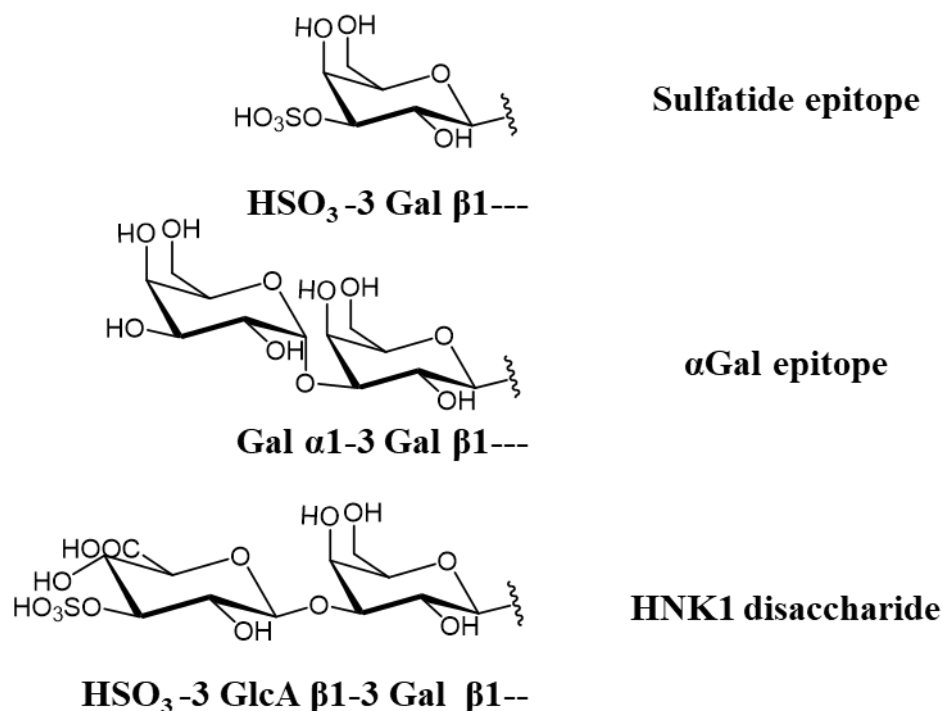


Figure 44 – Examples of epitopes containing Galβ1-unit that have modified O3 positions. Many other complex glycans and glycoconjugates include galactose in their terminal moieties, such as branched N-glycans and gangliosides.

These important epitopes share the Galβ1-unit (Figure 44) that is contained in compound **IX**. However, to obtain the selective O3 modification and reproduce the essential terminal structures, an orthogonal protecting group scheme must be adopted. In fact, the most reactive hydroxyl groups are OH-3 and OH-6, with the axial OH-4 being the least nucleophilic one, but this general consideration is highly dependent on protecting group features and on the donor nature<sup>[179]</sup>. The purification of product mixture obtained in non-selective conditions, whenever possible, can be tedious and lead to a huge loss of precious material. Therefore, we synthesized the 2,6-O-protected compound **X**, that serves as promising building block for future reactions, e.g. the selective sulfation or the formation of glycosidic bonds. Compound **X** was synthesized starting from fully

deprotected  $\beta$ -D-galactopyranosyl-azide that can be obtained uneventfully by deacetylating compound **VII** in presence of MeONa catalytic amount. Notably, initial attempts to selectively protect O6 (in virtue of its less hindered position) with bulky pivaloyl (Piv) or benzoyl (Bz) groups led to a mixture of products that could not be purified, with O3 and O6 being the most abundant ones as evidenced by  $^1\text{H-NMR}$ .

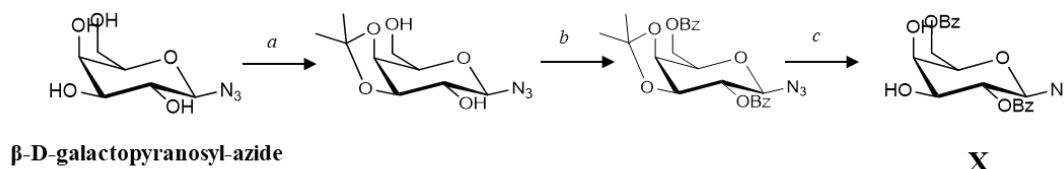


Figure 45 – Synthetic scheme that led to obtainment of 2,6-O-benzyl- $\beta$ -D-galactopyranosylazide **X**. a) CSA, 2,2-dimethoxypropane, 48 h RT;  $\text{Et}_3\text{N}$ , 15 min RT; MeOH/ $\text{H}_2\text{O}$  10:1, 2 h reflux. b) BzCl, pyridine, 16 h RT. c) AcOH 80%, 1h 80°C.

So, we adopted to a more reliable method to selectively protect O-6, which is based on a multi-step scheme developed by *Catelani*<sup>[180]</sup>. By acidic catalysis in 2,2-dimethoxypropane as solvent, the selective protection of of OH-3 and OH-4 of starting  $\beta$ -D-galactopyranosyl-azide was obtained. The hydroxyl groups at the 2 and 6 positions were protected as benzoyl esters, then final acetal hydrolysis afforded compound **X**. This longer but very effective synthetic strategy allowed the obtainment of large quantities of pure compound **X** with remarkably good yields (70-78% over three steps in which only one purification step by column chromatography is involved).

This galactose derivative contains a 1- $\beta$ -linked azide moiety that can be either exploited for CuAAC reaction, or may be reduced through hydrogenation to a primary amine group for the subsequent conjugation to carboxyl functions, e.g. bromoacetic acid for Cys alkylation. It is protected on O2 and O6 positions by benzoate groups (that can be removed by saponification in basic conditions), hence the O3 modification is likely to be selective in controlled conditions. Alternatively, a two-step protocol involving a orthoester-mediated protection to form of the 3,4-orthobenzoate intermediate, which is then hydrolyzed by aqueous AcOH with quantitative yield, can be envisaged to selectively protect axial O4 position<sup>[181]</sup>.

Thanks to these features, compound **X** is a promising building block to be used for the obtainment of sulfatide epitope mimetics, or as galactosyl acceptor in glycosidic bond formation. For this purpose, during my Ph.D. research a major effort was devoted also to the synthesis of a challenging but crucial monosaccharide building block (compound

**XIV**). This fully protected glucuronic acid (GlcA) moiety represents the essential scaffold to synthesise glycomimetics of the HNK-1 epitope, ultimately leading to the obtainment of terminal HSO<sub>3</sub>-3GlcAβ1- unit (Figure 46).

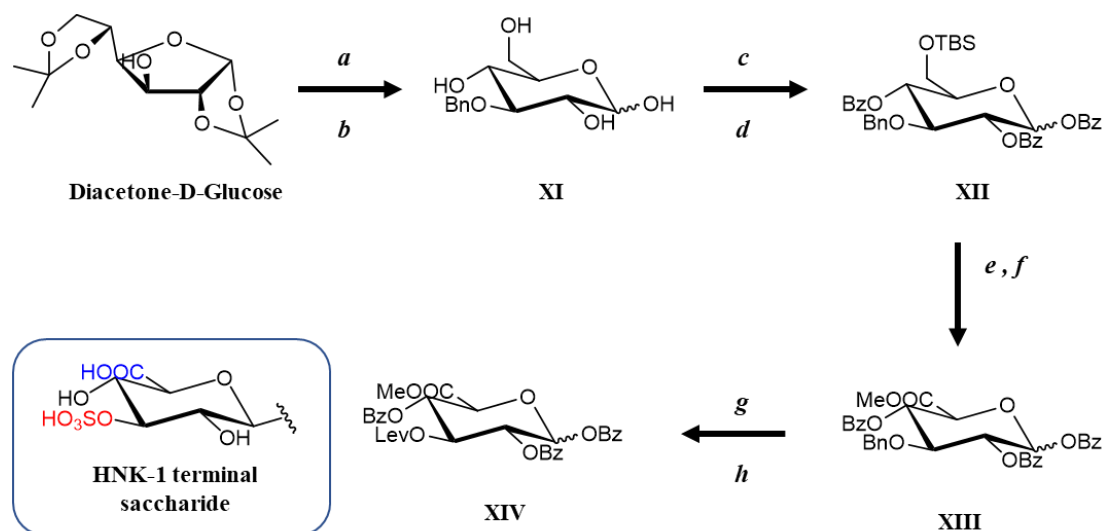


Figure 46 – Eight-step reaction scheme to obtain building block **XIV**, a crucial unit for HNK-1 mimetics synthesis. a) *BnBr*, *NaH*, 2h RT in DMF. b) *TFA* 50%, 1 h RT. c) *TBDMSCl*, 2h RT in pyridine. d) *BzCl*, 16 h RT in pyridine. e) Jones' reagent, 1 h RT in acetone. f) *MeI*, *KHCO<sub>3</sub>*, 16 h RT in DMF. g) *H<sub>2</sub>*, *Pd/C*, 3 h RT in THF. h) *LevOH*, *DCC*, *DMAP*, 16 h RT in DCM.

Compound **XIV** had already been synthesised by our group before<sup>[182]</sup>, whereas several groups in the field have reported the synthesis of similar precursors to obtain HNK1-related oligosaccharides<sup>[177,181,183,184]</sup>. Although challenging, the synthesis of this glucuronyl donor and the subsequent glycoside-bond formation with galactosyl-containing acceptors is pivotal for the large-scale production of specific HNK-1 mimetics and their exploitation in bioassays.

As a result, we performed a multi-step synthesis starting from commercially available diacetone-D-glucose, whose free OH-3 was benzylated in the presence of sodium hydride and benzyl bromide in DMF. The fully protected furanoside intermediate was converted into 3-O-benzyl-D-glucopyranose (compound **XI**, mixture of  $\alpha/\beta$  anomers) in acidic conditions. In the next two steps, the primary alcohol function (O6) was protected by bulky acid-labile *t*-butyldimethylsilyl chloride (*TBDMSCl*) and the remaining free hydroxyls were benzoylated, obtaining complex crudes that were purified by column chromatography to afford desired product **XII** with variable and disappointingly low yields (30%-48%). After oxidation using Jones' reagent (*CrO<sub>3</sub>* and sulfuric acid) in acetone and methylation by methyl iodide in the presence of potassium

hydrogenocarbonate in DMF, compound **XIII** was obtained. The benzyl group was removed by hydrogenolysis and replaced by levulinate protecting group.

This not always straightforward synthetic procedure was introductory for the obtainment of conveniently protected compound **XIV**. This precursor can be converted into the 1 $\beta$ -azide derivative for subsequent click-type reactions with various scaffolds to elaborate innovative architectures based on the terminal sulfated HNK1 monosaccharide. Otherwise, it can be employed as glucuronyl donor for the glycoside bond formation with galactosyl acceptor **X**, allowing the achievement of a disaccharide building block that could interact better with antibodies<sup>[177]</sup>, hence holding more potential for the development of high-affinity probes.

Both scenarios envisage the involvement of an intermediate activated form, i.e. the 1- $\alpha$ -bromo derivative **XV**, that is unstable and not suitable for long-term storage.

In a preliminary attempt, we performed bromination of anomeric carbon on a small quantity of compound **XIV**, and we obtained the final activated  $\alpha$ -donor **XV** with an excellent purity but a less comforting yield (48%). Unfortunately, the subsequent reaction to obtain the desired disaccharide by coupling pure glucuronyl donor **XV** and galactosyl derivative **X** was not successful (Figure 47).

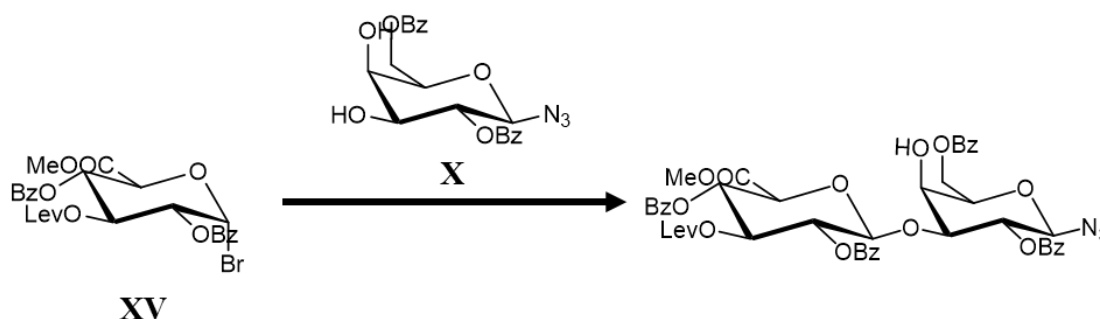


Figure 47 – Reaction scheme of the glycoside bond formation between the saccharides **XV** and **X**. By using AgOTf as promoter and anhydrous DCM as solvent, the reaction did not occur.

Lewis acid activation by silver triflate in anhydrous dichloromethane at 0°C did not lead to the formation of expected disaccharide. Instead, the TLC plate evidenced the disappearance of compound **XV** but no other spots could be clearly identified. Purification was also pointless, and even NMR analysis of obtained fractions failed in providing relevant hints to determine the nature of product mixture.

As a result, future efforts must be directed toward the optimization of the glycosidic bond formation, with the crucial interest of developing HNK1 disaccharide analogues.

Different activating groups could be envisaged, such as trichloroacetoimidates or thioglycosides<sup>[185]</sup>. These schemes in HNK-1 related synthesis were not found suitable for oligosaccharide linkage<sup>[182]</sup>, whereas they have been used in the case of galactosyl acceptors<sup>[186][187]</sup>.

Through a multi-step synthesis, the glycomimetics based on the final monosaccharide HSO<sub>3</sub>-3GlcA- and the final disaccharide HSO<sub>3</sub>-3GlcAβ1-3Gal- moieties can be synthesized, however the procedure is time consuming and skill-demanding.

With the aim to conjugate these valuable sugars to peptides and peptide-based polymers, eventually leading to form the stationary phase of columns for plasmapheresis, large-scale production of precious conjugates shall be achieved, therefore the optimization of current synthetic strategies is crucial.

### 6.2.2 Alkylation of a model cysteinyl peptide

Cysteine carbamidomethylation is a deliberate post-translational modification introduced to cysteine residues by reacting with a haloacetamide. This modification is mainly used in proteomics studies for the identification and characterization of proteins and peptides by mass spectrometry or to prevent cysteine from oxidation by treating proteins in basic aqueous buffer<sup>[188]</sup>.

For our investigational studies, we synthesised Ac-KACVTLNTT-NH<sub>2</sub> (peptide **11**) analogue of peptide **2** previously discussed in chapter 3, where native Asn1348 of HMW1(1347-1354) fragment is replaced by a thiol containing Cys. The synthesis holds a double purpose. On one hand, we aim to assess the thioether derivative ability to inhibit anti-N(Glc) antibodies binding to CSF114(Glc). In fact, no similar compound was ever tested before with human sera, and evaluating the capability of the glucosylated derivative to mimic Asn(Glc) containing peptides could provide insights about the specificity of the antibody recognition. On the other hand, optimizing reaction conditions is important for the future plan of conjugating more complex oligosaccharides to peptides.

Purified compound **VI** was used for the selective glucosylation of peptide **11** (Figure 48).

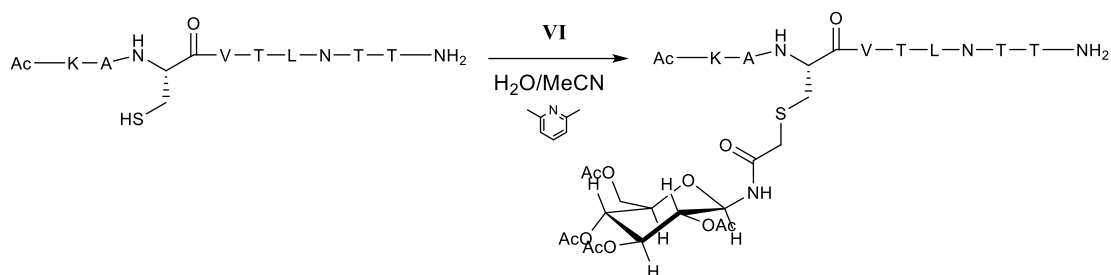


Figure 48 – Reaction scheme for the synthesis of acetylated peptide **12**. Reaction conditions are described in the experimental section.

There are few reports in literature involving the S-alkylation on cysteine residues reacting with glycosyl haloacetamides for the synthesis of N-linked glycoprotein mimetics. A solid-phase coupling method using iodoacetamides for protein semisynthesis was described<sup>[169]</sup>, while in-solution methods for thiol alkylation of biomolecules<sup>[170,189,190]</sup> reported the use of aqueous buffers to perform the reaction.

The model reaction is therefore of great interest because it allows exploring reaction conditions in presence of a competing nucleophile, i.e. the amine moiety of N-terminal Lys side chain. First attempts to couple peptide **11** to compound **VI** envisaged the use of DIEA ( $pK_a=10.75$ ) as base (from 0 up to 4 eq), in virtue of its poor nucleophilicity. Small amounts of peptide were dissolved in the minimum amount of MeCN/H<sub>2</sub>O (3:1 v/v, final concentration ca 3 mg peptide/mL) and an excess the glucose derivative was added (from 2 up to 8 eq).

In neutral environment, reaction proceeded very slowly during one day stirring and did not get to completion (with the pH stabilized  $\approx 6-7$ , as assessed by pH paper), yet the mass of the new peak found in the chromatogram corresponded to the desired product ( $t_r = 5.38$  min, found  $m/z = 1378.17$ , calculated 1378.63) (Figure 49).

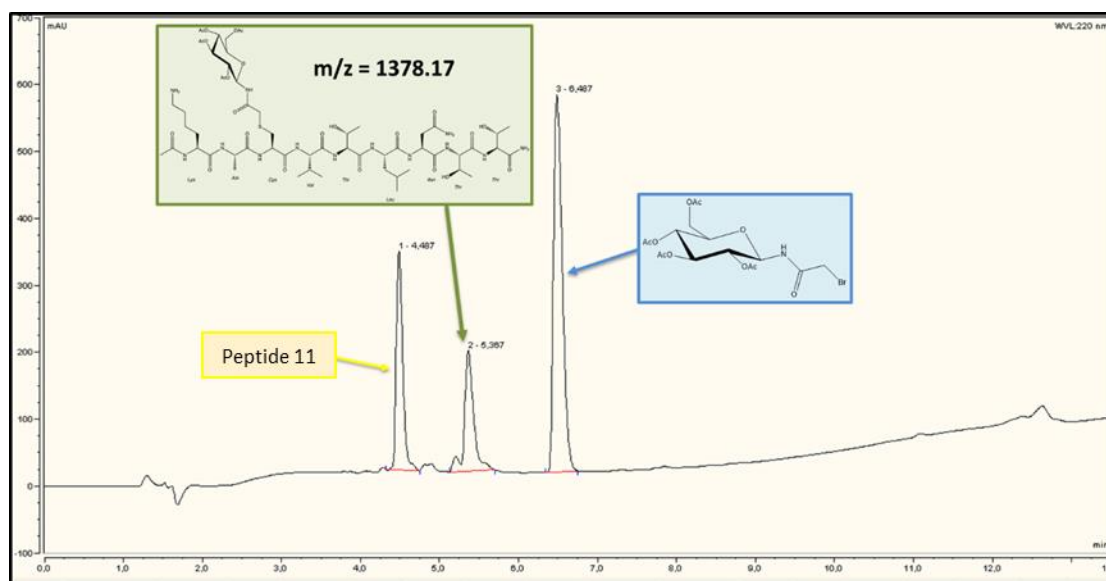


Figure 49 – RP-HPLC chromatogram of the S-alkylation of peptide **11** ( $t_r = 4.50$  min) to form acetylated peptide **12** ( $t_r = 5.38$  min), after 24 h reaction without the addition of base. Analytical RP-HPLC gradient 5-100% B in 10 min at  $1 \text{ mL min}^{-1}$ ; solvent system A: 0.1% TFA in  $\text{H}_2\text{O}$ , B: 0.1% TFA in MeCN

Disappointingly, any small addition of DIEA (pH =9-10 by pH paper) caused the disappearance of the starting peptide but also led to a crowded chromatogram with several peaks due to side-reactions. In particular, a peak of comparable intensity at  $t_r = 6.09$  min was ascribed to a second alkylation, on lysine side chain (found  $m/z = 1764.94$ , calculated 1765.65). The several other peaks presumably resulted from the loss of one or more acetyl groups bound to the glycan moiety, and possibly disulfide bridge formation, however the mass analysis by MALDI-TOF of corresponding fractions was not trivial.

Therefore, we performed the reaction in already reported pH-controlled conditions by two different aqueous buffers, i.e. phosphate buffer (100 mM, pH = 7)<sup>[190]</sup> and  $\text{NH}_4\text{CO}_3$  buffer (pH=8)<sup>[191]</sup>. In both cases, the peptide **11** peak disappearance correlated with the formation of a major peak, corresponding to acetylated peptide **12** by MALDI-TOF analysis. However, the insolubility of both the Br-acetamido derivative **IV** and the acetylated peptide **12** in saline buffers, required the dilution in acetonitrile and longer reaction times. Conversely, by the addition of 3 equivalents of a weak organic base (lutidine,  $pK_a = 6.72$ ) to a 3 mg/mL solution of peptide **11** in 50:50 MeCN/ $\text{H}_2\text{O}$ , HPLC chromatogram showed the quantitative conversion in less than three hours. These conditions were selected as the best ones to perform the S-alkylation reaction between free Cys-containing peptides and acyl-protected glycosylamides. In alternative, the preliminary deprotection of sugar derivatives could be performed before the coupling

reaction, but this could lead to the partial loss of Br-acetamido glycosyl donors that is detrimental in the case of precious saccharides.

This was separated from remaining excess of compound **VI** (that can be in this way recuperated free of salts), then deacetylation and semi-preparative RP-HPLC afforded pure peptide **12**.

By coating CSF114(Glc) in ELISA plates, the competition for IgG binding was evaluated by using serum MS1. Peptides **11** and **12** were compared with peptide **1** and **2**, respectively, as they represent their Cys-containing analogues (Figure 50).

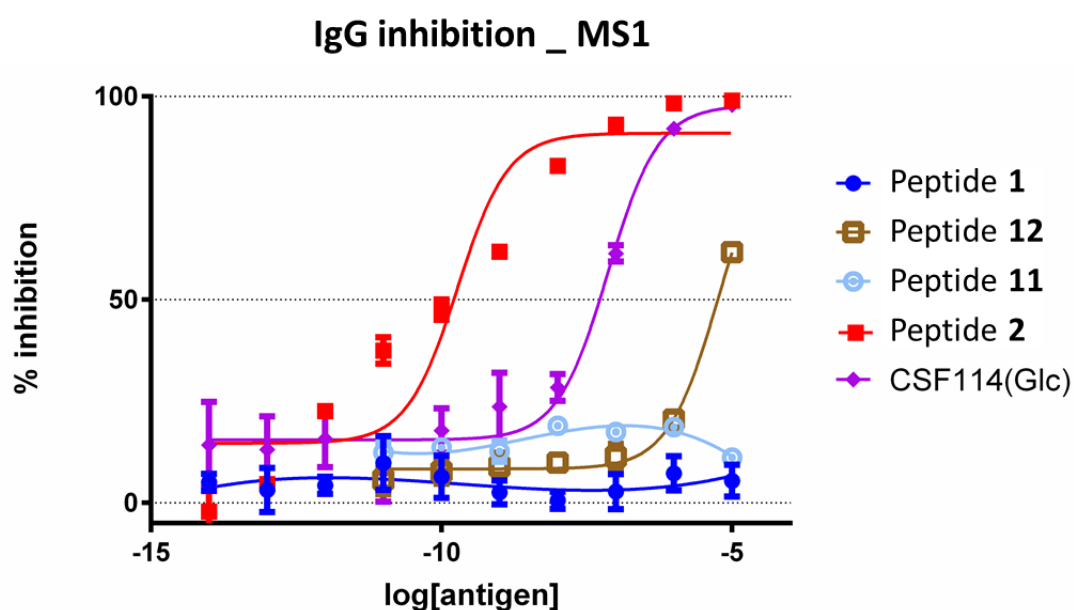


Figure 50 - Inhibition curves of anti-CSF114(Glc) IgG antibodies in serum MS1 with nonapeptides **1-2** and analogues **11-12** in a competitive indirect ELISA, in comparison with the glucopeptide CSF114(Glc).

Table 7 – Calculated  $pIC_{50}$  values for inhibitors toward anti-CSF114(Glc) IgG antibodies. Values are reported as 95% confidence intervals for the calculated mean  $pIC_{50}$  of the peptide antigens **1-2** and **11-12** used as inhibitors of anti-CSF114(Glc) IgG antibodies in MS1 sera.

Inhibitor	MS 1
Peptide 1	----
Peptide 2	$9.28 \pm 0.55$
Peptide 11	----
Peptide 12	$5.32 \pm 0.21$

**CSF114(Glc)**

7.21 ± 0.29

As already reported for other glycopeptides lacking the native N-glycosyl function<sup>[41]</sup>, Peptide **12** failed to display a satisfactory IC<sub>50</sub> in competitive ELISA tests using MS1, decreasing in almost four orders of magnitude compared to its Asn(Glc) analogue peptide **3** (Table 7). Although other sera should be tested, and different antibody isotypes, these experiments confirmed the high specificity of IgG antibodies targeting Asn(Glc). The recognition is likely driven by direct interactions of the antibody binding site with the Asn-linked sugar moiety. Therefore, possessing a glucosyl-amide portion is not enough to make this aberrantly modified residue a good surrogate of glucosyl-asparagine.

However, S-alkylation on cysteine residues with bromoacetamides to obtain N-linked glycoprotein mimetics as probes in autoimmunity was successfully performed in the case of a simple glucose moiety. The found conditions will be useful for the linkage of complex and precious saccharides that cannot be used in SPPS, whenever the specific Asn-glycan distance is less critical for the paratope interaction.

### **6.3 CONCLUSIONS**

This concluding chapter evaluates the exploitation of selective and efficient synthetic strategies that will hopefully lead soon to produce large quantities of neoglycopeptide epitopes. Here, we reported the use of bromo-acetamido derivative of glucose as alkylating reagent for the obtainment of a novel glycosylated peptide.

Glycosylation is a crucial modification that is supposed to be involved in the disruption of immune tolerance for many immune-mediated pathologies, not only in the case of  $\beta$ -glucosyl asparagine residues that have been associated to MS, and, more recently, to Rett Syndrome<sup>[192,193]</sup>. Glycopeptide synthesis can be extremely challenging, often resulting in the obtainment of small quantities of product and limited to robust, simple glycosides. It is important to widen our chemical tools in order to obtain sequences modified with more complex saccharides, such as 3-O-sulfogalactose (sulfatide epitope) to trap anti-sulfatide antibodies that are associated to various neuropathies and diabetes<sup>[173]</sup>, and 3-O sulfoglucuronyl  $\beta$ -1,3- galactose (HNK1 disaccharide epitope)

that is related to monoclonal neuropathies<sup>[175,178]</sup>. The synthesis of very complex targets is still a challenge and it usually is the speed of glycan synthesis that is rate limiting<sup>[160][182,194]</sup>. Then, efficient conjugation strategies must be adopted to obtain glycan conjugates in large amount for clinical applications, e.g., to develop stationary phases for plasmapheresis columns.

We performed the synthesis of fundamental glycosyl precursors, but the final saccharides have not been achieved yet. However, our goal to produce a glycopeptide library fuelled the exploration of robust chemical site-selective modification strategies, with the future aim to obtain and apply homogeneous conjugates with improved properties to antibody binding and capture in human sera.

By using the bromo-acetamido derivative of glucose (**VI**) we performed the thiol alkylation on a Cys containing peptide and selected our best conditions to perform the reaction, that offers useful advantages. First, the alkylation of peptide **11** using the bromoacetamide derivative **VI** in acetonitrile/water solution with only 2-3 eq of lutidine was found simple, fast and effective. Aqueous salted buffers can also be used, but prior deprotection of glycosyl moieties is required for solubility. Also, the attachment of bromoacetamide derivatives is selective to cysteine residues in neutral or slightly basic pH. When pH was increased above 8-9, nucleophilic competition by the deprotonated amine of Lys side chain resulted in the formation of the di-glucosylated adduct, and other side-products. Therefore, this method of conjugating a defined oligosaccharide to cysteine side chains on a peptide backbone provides a compelling fine-tuned strategy for synthetic glycosylation of peptides to reproduce glyconeopitopes. The synthesis of novel glycopeptide and glycoprotein mimetics with this method would allow for structure–activity relationship studies, providing many practical applications as well. Thioethers are stable and have not been shown to be immunogenic or antigenic<sup>[191]</sup>. The formation of a stable *S*-linkage generates compounds with longer half-lives and simplifies their preparation by “click-type” assembly strategies. There is yet no simple biological or chemical means of obtaining glycoproteins with homogeneous carbohydrate structures attached. Although pioneering studies assessed that the resultant thioether linkage does not appear to change the conformation of the sugar moieties<sup>[189]</sup> and that oligosaccharide-linked glycoproteins obtained by using this method maintained their activity in lectin- and antibody-binding assays<sup>[191]</sup>, our current preliminary investigations showed that synthetic cysteine linkage is not equivalent to natural N-linkage, as the extension by an extra CH<sub>2</sub>-S group in glucose-backbone

linkage caused a dramatic decrease in IgG-type antibody affinity in a MS serum. However, in the case of oligosaccharide-linked structures the epitope-paratope interaction is unlikely to be affected by such a little alteration. As a result, the scope of producing neoglycopeptides by this approach is still valid and the preparation of native haloacetamidyl monosaccharide and oligosaccharides is now very essential for mimetic glycopeptide engineering. With this strategy we added a third level of orthogonality to the possible conjugation schemes, meaning that complex glycopeptide libraries possessing an azide and a free amine could be prepared and grafted to previously discussed Dex40-GP or other types of architectures, and this will allow expanding our plethora of glycopeptides probes in autoimmune research.

## 7. EXPERIMENTAL PART

### 7.1 ORGANIC SYNTHESIS

#### 7.1.1 Materials

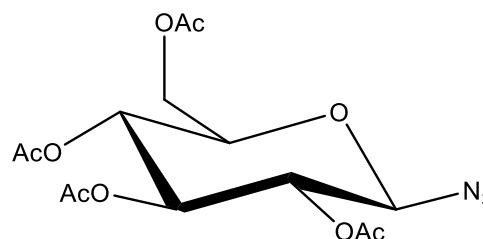
All the chemicals and solvents were obtained from Acros Organics, Carlo Erba, SDS, Merck, Iris Biotech GmbH, Novabiochem, and VWR and used without further purification.

NMR spectra of compounds were recorded on a Bruker spectrometer (300 MHz spectrometer was used for  $^1\text{H}$  spectra and 75 MHz for  $^{13}\text{C}$  spectra), with solvent resonance used as reference.

Thin layer chromatography (TLC) was performed using Macherey-Nagel AlugramSil G/UV254 plate. Spots revelation was performed in three ways: (a) UV (254nm); (b) 10% sulfuric acid in ethanol; (c) ninhydrin in acetone.

#### 7.1.2 (2,3,4,6-Tetra-O-Acetyl)- $\beta$ -D-Glucopyranosyl-Azide (I)

To a stirring solution of  $\alpha/\beta$ -D-Glucose (11 g, 61 mmol, 1 eq) in Pyridine (60 mL) under Ar, Acetic Anhydride (30 mL, 318 mmol, 5.5 eq) was added dropwise. The mixture was left stirring over night at RT. Then solvent was



evaporated and the residue was dissolved in DCM, then washed with HCl 1M,  $\text{NaHCO}_3$  sat., NaCl sat. The organic layer was finally dried over  $\text{MgSO}_4$ , filtered and solvent is evaporated under vacuum. The obtained residue was then crystallized in EtOH, obtaining 19 g (80% yield) of (1,2,3,4,6-Penta-O-Acetyl)- $\alpha/\beta$ -D-Glucopyranose as white crystals.

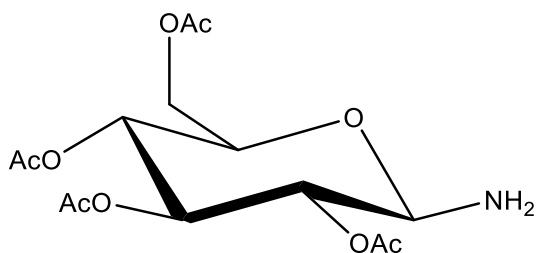
The obtained compound (19 g, 49 mmol) was dissolved and stirred in anhydrous DCM (30 mL) at  $0^\circ\text{C}$  under Ar, then HBr (33% in AcOH, 60 mL) was added dropwise. The solution was left stirring for 2 h while slowly heating up to RT. Then solvent was evaporated under vacuum and the residue was neutralized with  $\text{NaHCO}_3$  sat., then DCM was added and extracted twice with  $\text{NaHCO}_3$  sat., once with NaCl sat., dried over  $\text{MgSO}_4$  and filtered. Organic layer was evaporated under vacuum to give 19.4 g (2,3,4,6-Tetra-O-Acetyl)- $\alpha$ -D-Glucopyranosyl bromide (96% yield) as a foaming sticky residue which was used without further purification.

Previous crude was dissolved in acetone (90 mL) and then water (15 mL) was added. NaN<sub>3</sub> (5.6 g, 1.6 eq) was finally added and the suspension was left stirring overnight. In the morning, acetone was evaporated under vacuum and residue was extracted in DCM, washed with water, brine and dried over MgSO<sub>4</sub>. After evaporating solvent, desired product **I** was crystallized from EtOH (14.9 g as white crystals, 84% yield).

TLC *R<sub>f</sub>* 0.45 [CHCl<sub>3</sub>—MeOH 97/3; revealed with (b)]. IR: 2119 (N<sub>3</sub>) cm<sup>-1</sup>. <sup>1</sup>H NMR (300 MHz, Chloroform-*d*) δ 5.21 (t, *J* = 9.4 Hz, 1H, H<sup>3</sup>), 5.09 (t, *J* = 9.7 Hz, 1H, H<sup>4</sup>), 4.95 (t, *J* = 9.1 Hz, 1H, H<sup>2</sup>), 4.64 (d, *J* = 8.8 Hz, 1H, H<sup>1</sup>), 4.27 (dd, *J* = 12.4, 4.7 Hz, 1H, H<sup>6</sup>), 4.16 (dd, *J* = 12.5, 2.4 Hz, 1H, H<sup>6</sup>), 3.79 (ddd, *J* = 10.0, 4.8, 2.4 Hz, 1H, H<sup>5</sup>), 2.09 (s, 3H, OAc), 2.07 (s, 3H, OAc), 2.02 (s, 3H, OAc), 2.00 (s, 3H, OAc).

Data are in agreement with those reported in literature<sup>[195]</sup>.

### 7.1.3 (2,3,4,6-Tetra-O-Acetyl)-β-D-Glucopyranosyl-Amine (II)



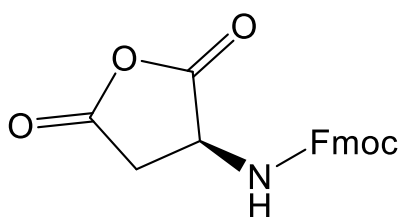
Crystals of **I** (3.7 g, 10 mmol) were dissolved in 20 mL THF, NEt<sub>3</sub> (1.5 mL, 1.1 eq) was added and finally Pd/C. Suspension was left stirring under H<sub>2</sub> atmosphere (6 bar) for 1-2 h at RT, then it

was filtered through *celite* and evaporated. The pale-yellow residue (3.5 g) was recrystallized in MeOH, yielding white crystals of product **II** (3.1 g, 89% yield).

TLC *R<sub>f</sub>* = 0.37 [AcOEt/CyHex 4:1; revealed with (b), (c)]. <sup>1</sup>H NMR (300 MHz, Chloroform-*d*) δ 5.23 (t, *J* = 9.5 Hz, 1H, H<sup>3</sup>), 5.03 (dd, *J* = 10.0, 9.4 Hz, 1H, H<sup>4</sup>), 4.82 (dd, *J* = 9.7, 9.0 Hz, 1H, H<sup>2</sup>), 4.22 (dd, *J* = 12.3, 4.8 Hz, 2H, H<sup>6</sup>), 4.17 (superimposed d, 1H, H<sup>1</sup>), 4.09 (dd, *J* = 12.3, 2.4 Hz, 1H, H<sup>6</sup>), 3.68 (ddd, *J* = 10.0, 4.8, 2.3 Hz, 1H, H<sup>5</sup>), 2.08 (s, 3H, OAc), 2.06 (s, 3H, OAc), 2.01 (s, 3H, OAc), 2.00 (s, 3H, OAc).

Data are in agreement with those reported in literature<sup>[98]</sup>.

### 7.1.4 N-Fmoc-L-Aspartic Anhydride (III)



To a solution of Fmoc-OSu (5 g, 15 mmol) in THF (25 mL) was added a solution of Asp (2.2 g, 15 mmol) in 30 mL water containing Na<sub>2</sub>CO<sub>3</sub> (3.5 g, 33 mmol) and the mixture was left stirring vigorously for 20-24 h at RT, then washed with ether (2x20 mL).

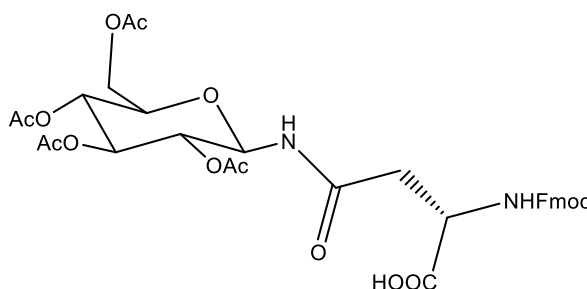
The aqueous phase was acidified to pH 2 using HCl, and then extracted with EtOAc

(2x40 mL). The organic phase was dried with MgSO<sub>4</sub> and evaporated. The white solid (4.7 g, 89% yield) was suspended in Ac<sub>2</sub>O (15 mL). By using a heating gun with rapid heating and shaking for around 1 min, the residue was dissolved and solution was quickly cooled to RT to give a white foaming precipitate that was filtered off, washed with dry and cold ether and finally dried under vacuum for 24 h (4.5 g, 70% yield over two steps).

TLC *R<sub>f</sub>* 0.39 [AcOEt/CyHex 4:1; revealed with (a)]. <sup>1</sup>H NMR (300 MHz, DMSO-*d*<sub>6</sub>) δ 8.21 (d, *J* = 7.7 Hz, 1H), 7.89 (d, *J* = 7.5 Hz, 2H), 7.67 (d, *J* = 7.4 Hz, 2H), 7.48 – 7.36 (m, 2H), 7.33 (tt, *J* = 7.4, 1.5 Hz, 2H), 4.65 (ddd, *J* = 10.0, 7.7, 6.2 Hz, 1H), 4.40 (dd, *J* = 6.6, 3.0 Hz, 2H), 4.25 (t, *J* = 6.5 Hz, 1H), 3.24 (dd, *J* = 18.5, 10.0 Hz, 1H), 2.85 (dd, *J* = 18.5, 6.2 Hz, 1H).

Data are in agreement with those reported in literature<sup>[99]</sup>.

### 7.1.5 (2,3,4,6-Tetra-O-Acetyl-N-[N-Fmoc-L-aspart-4-oyl]-β-D-Glucopyranosylamine (Fmoc-Asn[Glc(OAc)<sub>4</sub>]-OH)



Compound **III** (1.75 g, 5 mmol) was added to a stirred solution of compound **II** (1.8 g, 5 mmol) in 3 mL of DMSO under argon atmosphere. The reaction mixture was kept at RT for 3 h and then water was added. The

white, sticky precipitated solid was washed thoroughly with water and dried over-night. Recrystallization in MeOH/H<sub>2</sub>O 2:1 afforded 2.4 g (69% yield) of desired product.

TLC *R<sub>f</sub>* 0.19 [AcOEt—CyHex 2:1; revealed with (a) and (b)].

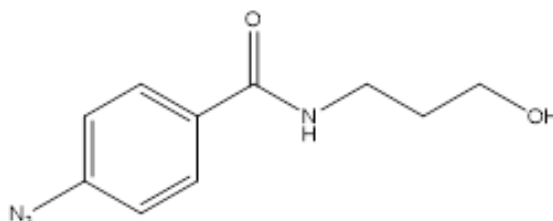
<sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 7.75 (d, *J* = 7.4 Hz, 2H, *fluorenyl* H<sup>4</sup> and H<sup>5</sup>), 7.58 (d, *J* = 7.4 Hz, 2H *fluorenyl* H<sup>1</sup> and H<sup>8</sup>), 7.39 (t, *J* = 7.3 Hz, 2H, *fluorenyl* H<sup>3</sup> and H<sup>6</sup>), 7.30 (t, *J* = 7.4 Hz, 2H *fluorenyl* H<sup>2</sup> and H<sup>7</sup>), 6.89 (d, *J* = 9.2 Hz, 1H, *glucose* -NH), 6.18 (d, *J* = 8.2 Hz, 1H *Fmoc* -NH), 5.33 (t, *J* = 9.5 Hz, 1 H, *glucose* H<sup>1</sup>), 5.29 (t, *J* = 9.4 Hz, 1H, *glucose* H<sup>3</sup>), 5.08 (t, *J* = 9.7 Hz, 1H, *glucose* H<sup>4</sup>), 4.95 (t, *J* = 9.6 Hz, 1H, *glucose* H<sup>2</sup>), 4.59 (dd, *J* = 8.0, 4.8 Hz, 1H, H<sup>a</sup>), 4.37 (t, *J* = 7.6 Hz, 2H, *Fmoc* CH<sub>2</sub>-O), 4.29 (dd, *J* = 12.6, 4.3 Hz, *glucose* H<sup>6</sup>), 4.21 (t, *J* = 7.1 Hz, 1H, *fluorenyl* H<sup>9</sup>), 4.06 (dd, *J* = 12.6, 2.2 Hz, *glucose* H<sup>6'</sup>), 3.83 (ddd, *J* = 10.2, 4.3, 2.1 Hz, *glucose* H<sup>5</sup>), 2.98 – 2.69 (m, 2H, H<sup>β</sup>), 2.04 (s, 3H, OAc), 2.03 (s, 6H, 2x OAc), 2.01 (s, 3H, OAc).

$^{13}\text{C}$  NMR (75 MHz,  $\text{CDCl}_3$ )  $\delta$  173.10, 171.43, 170.71, 170.02, 169.60, 156.36, 143.73, 141.27, 127.78, 127.13, 125.15, 120.03, 78.04, 77.46, 77.04, 76.62, 73.76, 72.66, 70.61, 68.01, 67.48, 61.62, 50.32, 47.04, 40.35, 37.65, 20.72, 20.61.

Data are in agreement with those reported in literature<sup>[98,99]</sup>.

### 7.1.6 4-azido-N-(3-hydroxypropyl)-benzamide (IV)

To synthesize compound **IV** we used a protocol described for a similar compound<sup>[196]</sup>.



1-ethyl-3-(3-dimethylaminopropyl)carbodiimide

(1.01 g, 5.3 mmol) and 4-dimethylaminopyridine (645 mg, 5.3 mmol) were added to a stirring solution of 4-azidobenzoic acid (812 mg, 5.0 mmol) in dichloromethane (15 mL) at room temperature. After 30 min, 3-aminopropan-1-ol (465  $\mu\text{L}$ , 6.0 mmol) was added and the mixture was left stirring overnight at room temperature.

Then solvent was removed under vacuum and obtained crude was dissolved in ethyl acetate (30 mL), washed with HCl 1M (15 mL),  $\text{NaHCO}_3$  sat. (15 mL) and NaCl sat., and finally dried over  $\text{MgSO}_4$ . Organic layer was evaporated under vacuum and obtained crude (a yellow pasty solid) was purified by silica gel chromatography (ethyl acetate/cyclohexane 4:1).

The pale orange solid **IV** obtained after flash chromatography (456 mg, 41% yield) was characterized by NMR and FT-IR.

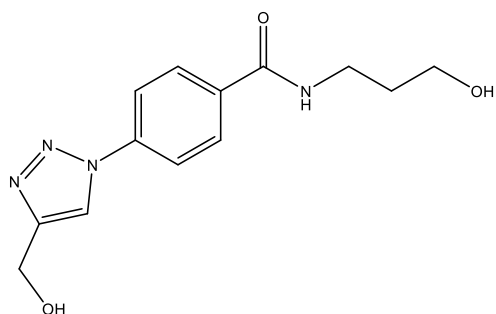
TLC  $R_f$  = 0,20 [EtOAc/CyHex 1:4 ; revealed with (a)]

FT-IR ( $\text{cm}^{-1}$ ) : 3400-3200 (with a sharp peak at 3287, secondary amide N-H / primary alcohol O-H stretching) , 2957-2877 (aromatic and aliphatic C-H stretching), 2130 (N=N stretching of azido moiety), 1627 (amide C=O stretching)

$^1\text{H}$ -NMR (300 MHz,  $\text{CDCl}_3$ )  $\delta$  7.77 (d,  $J$  = 8.6 Hz, 2H,  $\text{H}^2$ ), 7.05 (d,  $J$  = 8.6 Hz, 2H,  $\text{H}^3$ ), 6.75 (s, *broad*, 1H,  $\text{H}^N$ ), 3.73 (m, 2H, C- $\text{CH}_2$ -OH), 3.63 (q,  $J$  = 6.0 Hz, 2H, N- $\text{CH}_2$ -C), 1.80 (m, 2H, C- $\text{CH}_2$ -C)

$^{13}\text{C}$ -NMR (75 MHz,  $\text{CDCl}_3$ )  $\delta$  167.5 (C=O), 143.6 (C- $\text{N}_3$ ), 130.8 (C-CONH), 128.9 (C- $\text{H}^2$ ), 119.2 (C- $\text{H}^3$ ), 60.2 ( $\text{CH}_2$ -OH), 37.7 ( $\text{CH}_2$ -NH), 32.0 ( $\text{CH}_2$ )

### 7.1.7 Synthesis of 4-(4-(hydroxymethyl)-1H-1,2,3-triazol-1-yl)-N-(3-hydroxypropyl)benzamide (V)



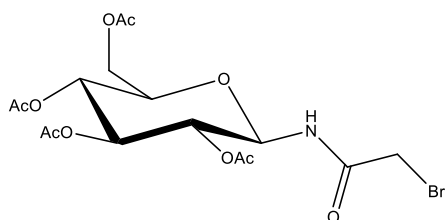
To a stirring solution of compound **IV** (21 mg, 95  $\mu\text{mol}$ ) in DMSO (300  $\mu\text{L}$ ), propargyl alcohol (6  $\mu\text{L}$ , 110  $\mu\text{mol}$ , 1.2 eq) was added.  $\text{H}_2\text{O}$  (150  $\mu\text{L}$ ) was then added, followed by the addition of the catalyst mixture (50  $\mu\text{L}$   $\text{CuSO}_4$  1 M, 0.5 eq, and 100  $\mu\text{L}$  sodium ascorbate 1 M, 1 eq). The mixture was stirred at room temperature and monitored by TLC. After 1h, 10 mL 0.1 EDTA solution were added. Desired product was extracted in ethyl acetate (2 x 10 mL). The organic layer was dried over  $\text{MgSO}_4$  and the solvent was removed on a rotary evaporator, affording pure compound **V** as an orange powder (20 mg, 75% recovered product).

TLC  $R_f$ : 0,18 [1:10 MeOH/EtOAc; revealed with (a)]

$^1\text{H-NMR}$  (300 MHz, Methanol- $d_4$ )  $\delta$  8.56 (s, 1H,  $\text{H}^{\text{Triaz}}$ ), 8.12 - 7.92 (m, 4H,  $\text{H}^{\text{Phen}}$ ), 4.79 (s, 2H,  $\text{triaz-CH}_2\text{-OH}$ ), 3.68 (t,  $J = 6.3$  Hz, 2H,  $\text{C-CH}_2\text{-OH}$ ), 3.52 (t,  $J = 7.0$  Hz, 2H,  $\text{HN-CH}_2\text{-C}$ ), 1.87 (p,  $J = 6.6$  Hz, 2H,  $\text{C-CH}_2\text{-C}$ )

$^{13}\text{C-NMR}$  (75 MHz, Methanol- $d_4$ )  $\delta$  168.9 ( $\text{C=O}$ ), 140.5 ( $\text{C-N}^{\text{triaz}}$ ), 136.0 ( $\text{C-CONH}$ ), 130.1 ( $\text{CH}^{\text{Ar}}$ ), 122.2 ( $\text{C}^{\text{Ar}}$ ), 121.2 ( $\text{CH}^{\text{Ar}}$ ), 60.6 ( $\text{CH}_2\text{-OH}$ ), 56.5 ( $\text{CH}_2\text{-OH}$ ) 38.3 ( $\text{CH}_2\text{-NH}$ ), 33.2 ( $\text{C-CH}_2\text{-C}$ )

### 7.1.8 (2,3,4,6-Tetra-O-Acetyl)- $\beta$ -D-Glucopyranosyl- (2'-Br)-N-Acetyl-Amine (VI)



Compound **II** (1,3 g, 3 mmol) is dissolved in anhydrous DCM (20 mL), and a previously prepared solution of 2-bromo-acetic acid (1.66 g, 12 mmol, 4 eq) and DCC (2.47 g, 12 mmol, 4 eq) is added with DMAP (73 mg, 0.2 eq). The mixture is left stirring at RT. After 1 h, TLC showed the completion of the reaction. Obtained suspension is filtered through Celite, and collected solution is then washed with HCl 1M,  $\text{NaHCO}_3$  sat, NaCl sat and dried over  $\text{MgSO}_4$ . After evaporating solvent, residue is purified by column chromatography using EtOAc/CyHex (3:7) as eluents. Pure

fractions rapidly crystallized in eluent mixture, affording product **VI** (1.12 g, 80% yield) as white crystals.

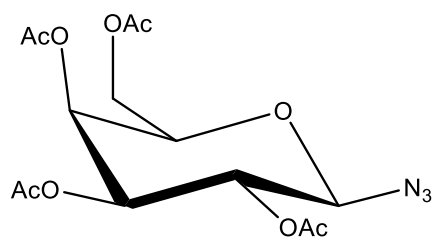
TLC  $R_f = 0.5$  [3:2 EtOAc/CyHex, revealed with (b)]

$[M+Na]^+$   $m/z = 490.31$  (calc. 490.03)

$^1H$  NMR (300 MHz, Chloroform-*d*)  $\delta$  7.15 (d,  $J = 8.9$  Hz, 1H,  $H^N$ ), 5.31 (t,  $J = 9.5$  Hz, 1H,  $H^3$ ), 5.18 (t,  $J = 9.2$  Hz, 1H,  $H^1$ ), 5.07 (t,  $J = 9.7$  Hz, 1H,  $H^4$ ), 4.99 (t,  $J = 9.4$  Hz, 1H,  $H^2$ ), 4.29 (dd,  $J = 12.5, 4.3$  Hz, 1H,  $H^6$ ), 4.08 (dd,  $J = 12.5, 2.4$  Hz, 1H,  $H^6$ ), 3.85-3.78 (m, 3H,  $H^5 + CH_2Br$ ), 2.07 (s, 3H, OAc), 2.05 (s, 3H, OAc), 2.02 (s, 3H, OAc), 2.01 (s, 3H, OAc).

$^{13}C$  NMR (75 MHz,  $CDCl_3$ )  $\delta$  170.82 ( $\underline{C}O$  Ac), 170.60 ( $\underline{C}O$  Ac), 169.87 ( $\underline{C}O$  Ac), 169.51 ( $\underline{C}O$  Ac), 166.49 (NH $\underline{C}OCH_2Br$ ), 78.72 ( $C^1$ ), 73.76 ( $C^5$ ), 72.53 ( $C^3$ ), 70.15 ( $C^2$ ), 68.05 ( $C^4$ ), 61.56 ( $C^6$ ), 28.11 ( $\underline{C}H_2Br$ ), 20.73 ( $CH_3$  Ac), 20.63 ( $CH_3$  Ac), 20.58 ( $CH_3$  Ac).

### 7.1.9 (2,3,4,6-Tetra-O-Acetyl)- $\beta$ -D-Galactopyranosyl-Azide (**VII**)



$\beta$ -D-Galactose pentaacetate (15 g, 38 mmol) was dissolved and stirred in anhydrous DCM (50 mL) at 0°C under Ar, then HBr (33% in AcOH, 60 mL) was added dropwise. The solution was left stirring for 2 h while slowly heating up to RT. Then

solvent was evaporated under vacuum and the residue was neutralized with  $NaHCO_3$  sat., then DCM was added and extracted twice with  $NaHCO_3$  sat., once with  $NaCl$  sat., dried over  $MgSO_4$  and filtered. Organic layer was evaporated under vacuum to give 12.5 g (2,3,4,6-Tetra-O-Acetyl)- $\alpha$ -D-Glucopyranosyl bromide (79% yield) which was used without further purification.

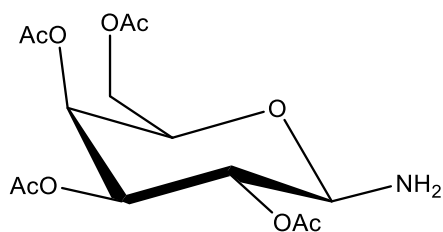
Previous crude was dissolved in acetone (55 mL) and then water (10 mL) was added.  $NaN_3$  (3.3 g, 1.5 eq) was added and the suspension was left stirring overnight. In the morning, acetone was evaporated under vacuum and residue was extracted in DCM, washed with water, brine and dried over  $MgSO_4$ . After evaporating solvent, desired product **I** was crystallized from EtOH (7.5 g as white crystals, 53% yield over two steps). TLC  $R_f$  0.5 [1:1 EtOAc/CyHex; revealed with (b)].

$^1H$  NMR (300 MHz, Chloroform-*d*)  $\delta$  5.43 (dd,  $J = 3.3, 1.1$  Hz, 1H,  $H^4$ ), 5.18 (dd,  $J = 10.3, 8.7$  Hz, 1H,  $H^2$ ), 5.05 (dd,  $J = 10.4, 3.3$  Hz, 1H,  $H^3$ ), 4.61 (d,  $J = 8.7$  Hz, 1H,  $H^1$ ),

4.18 (dd,  $J = 6.5, 2.1$  Hz, 2H,  $H^6 + H^6'$ ), 4.03 (*pseudo-t*,  $J = 6.5$  Hz, 1H,  $H^5$ ), 2.18 (s, 3H, OAc), 2.10 (s, 3H, OAc), 2.07 (s, 3H, OAc), 2.00 (s, 3H, OAc).

Data are in agreement with those reported in literature<sup>[195]</sup>.

#### 7.1.10 (2,3,4,6-Tetra-O-Acetyl)- $\beta$ -D-Galactopyranosyl-Amine (VIII)



Compound **VII** (4.0 g, 10.7 mmol) were dissolved in 20 mL THF,  $NEt_3$  (1.5 mL, 1.1 eq) was added and finally Pd/C. Suspension was left stirring under  $H_2$  atmosphere (6 bar) for 2 h at RT, then it was filtered through *celite* and evaporated.

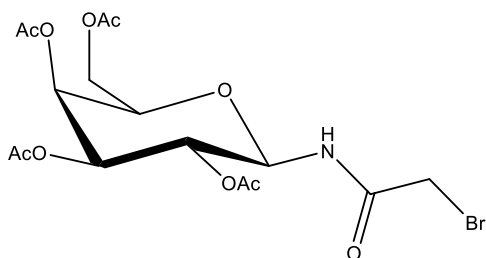
The off-white residue (3.7 g) slowly recrystallized in MeOH, yielding white crystals of product **VIII** (2.1 g, 57% yield).

TLC  $R_f = 0.4$  [AcOEt—cyclohexane 1/1; revealed with (b), (c)].

$^1H$  NMR (300 MHz, Chloroform-*d*)  $\delta$  5.46 – 5.34 (m, 1H,  $H^1$ ), 5.10 – 4.97 (m, 2H,  $H^2 + H^3$ ), 4.21 – 4.07 (m, 4H,  $H^4 + H^5 + H^6$ ), 2.16 (s, 3H, OAc), 2.09 (s, 3H, OAc), 2.06 (s, 3H, OAc), 1.99 (s, 3H, OAc)

Data are in agreement with those reported in literature<sup>[197]</sup>.

#### 7.1.11 (2,3,4,6-Tetra-O-Acetyl)- $\beta$ -D-Galactopyranosyl- (2'-Br)-N-Acetyl-Amine (IX)



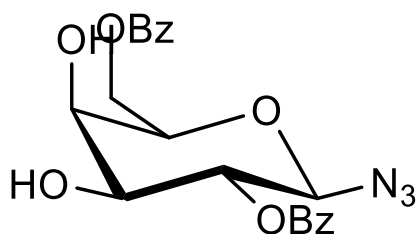
Compound **VIII** (671 mg, 1.9 mmol) is dissolved in anhydrous DCM (7 mL), and a previously prepared solution of 2-bromoacetic acid (950 mg, 6.8 mmol) and DCC (1.41 g, 6.8 mmol) is added with DMAP (50 mg, 0.4 mmol). The mixture is left stirring at RT in the dark. After 1 h, TLC showed the completion of the reaction. Obtained suspension is filtered through Celite, and collected solution is then washed with HCl 1M,  $NaHCO_3$  sat, NaCl sat and dried over  $MgSO_4$ . After evaporating solvent, residue is purified by column chromatography with a linear gradient of EtOAc/CyHex as eluents. Pure fractions are collected, and solvent is evaporated, affording product **IX** (410 mg, 46% yield) as a white powder.

TLC  $R_f = 0.6$  [3:2 EtOAc/CyHex, revealed with (b)]

$^1\text{H}$  NMR (300 MHz, Chloroform-*d*)  $\delta$  7.15 (d,  $J$ = 7.9 Hz, 1H,  $\text{H}^{\text{N}}$ ), 5.44 (pseudo-d,  $J$ = 2.7 Hz, 1H,  $\text{H}^3$ ), 5.22 – 5.11 (m, 3H,  $\text{H}^1 + \text{H}^2 + \text{H}^3$ ), 4.17 - 4.01 (m, 3H,  $\text{H}^5 + \text{H}^6$ ), 3.90-3.77 (m, 2H,  $\text{CH}_2\text{Br}$ ), 2.15 (s, 3H, OAc), 2.07 (s, 3H, OAc), 2.03 (s, 3H, OAc), 2.00 (s, 3H, OAc).

$^{13}\text{C}$  NMR (75 MHz,  $\text{CDCl}_3$ )  $\delta$  171.2 ( $\text{CO Ac}$ ), 170.5 ( $\text{CO Ac}$ ), 170.1 ( $\text{CO Ac}$ ), 169.9 ( $\text{CO Ac}$ ), 166.5 ( $\text{NHCOCH}_2\text{Br}$ ), 79.2 ( $\text{C}^1$ ), 72.7 ( $\text{C}^5$ ), 70.8 ( $\text{C}^3$ ), 68.0 ( $\text{C}^2$ ), 67.2 ( $\text{C}^4$ ), 61.2 ( $\text{C}^6$ ), 28.3 ( $\text{CH}_2\text{Br}$ ), 20.86 ( $\text{CH}_3 \text{Ac}$ ), 20.81 ( $\text{CH}_3 \text{Ac}$ ), 20.75 ( $\text{CH}_3 \text{Ac}$ ), 20.68 ( $\text{CH}_3 \text{Ac}$ ).

### 7.1.12 2,6-di-O-benzoyl- $\beta$ -D-Galactopyranosyl-Azide (X)



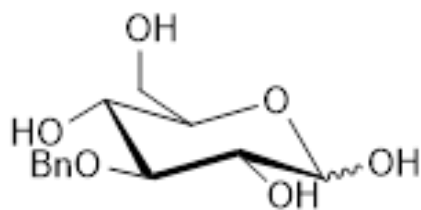
Camphorsulfonic acid (350 mg, 1.5 mmol) was added to a solution of  $\beta$ -D-galactopyranosyl-Azide (7.0 g, 34.1 mmol) in 2,2-dimethoxypropane (350 mL). The mixture was stirred for 48 h at RT under argon atmosphere.  $\text{Et}_3\text{N}$  (4.5 mL, 32.5 mmol) was then added and the mixture was stirred for 15 min. The mixture was concentrated to dryness, adding toluene (10 mL x3) to remove completely  $\text{Et}_3\text{N}$ . The residue was dissolved in MeOH/ $\text{H}_2\text{O}$  (10:1 v/v, 220 mL) and refluxed for 2 hours, until TLC (CycloHex/ $\text{EtOAc}$  3:2) showed complete conversion of the intermediate 6-O-(2-methoxy-2-propyl)-3,4-O-isopropylidene- $\beta$ -D-galactopyranosyl-azide ( $R_f = 0.47$ ) to the desired product 3,4-O-Isopropylidene- $\beta$ -D-Galactopyranosyl-Azide ( $R_f = 0.11$ ). Then the solution was concentrated, adding toluene (10 mL x3) to remove residual water. After complete dryness, a crude of 3,4-O-Isopropylidene- $\beta$ -D-Galactopyranosyl-Azide (8.6 g, quantitative yield) is obtained as an off-white powder. This was dissolved in pyridine (100 mL), and benzoyl chloride (10 mL, 86 mmol, 2.5 eq) was added dropwise. After 16 h stirring, solvent was evaporated under vacuum, and the resulting 2,6-di-O-benzoyl-3,4-O-isopropylidene-  $\beta$ -D-Galactopyranosyl-Azide crude was dissolved in DCM (50 mL) and washed with diluted HCl (30 mL x2),  $\text{NaHCO}_3$  sat (30 mL x2), NaCl sat (30 mL x2), dried over  $\text{MgSO}_4$  and filtered. DCM was evaporated and the crude was then stirred in acetic acid (80 mL) at  $80^\circ\text{C}$  for 5 minutes. Water (20 mL) was added to the solution that was stirred for an additional hour at  $80^\circ\text{C}$ . Solvent mixture was concentrated, and desired product was extracted in DCM/water (50 mL/40 mL). Organic layer was washed with  $\text{NaHCO}_3$  sat (30 mL x2),

NaCl sat (30 mL x2), dried over MgSO<sub>4</sub>, filtered, and dried under vacuum. Resulting oily crude (13.3 g) was purified by column chromatography (30% EtOAc in CyHex) affording pure compound **X** (11.13 g, 78% yield over three steps), that easily crystallize by concentrating in eluted fractions.

TLC R<sub>f</sub> = 0,15 (EtOAc/CyHex 3:7) [revealed with (a) and (b)]

<sup>1</sup>H NMR (300 MHz, Chloroform-*d*) δ 8.13 – 7.96 (m, 4H, H<sup>Bz</sup> arom), 7.69 – 7.53 (m, 2H, H<sup>Bz</sup> arom), 7.51 – 7.38 (m, 4H, H<sup>Bz</sup> arom), 5.23 (dd, *J* = 9.6, 8.8 Hz, 1H, H<sup>2</sup>), 4.80 – 4.70 (superimposed dd, 1H, H<sup>6</sup>), 4.74 (d, *J* = 8.8 Hz, 1H, H<sup>1</sup>), 4.56 (dd, *J* = 11.6, 6.5 Hz, 1H, H<sup>6'</sup>), 4.08 (pseudo-d, *J* = 3.5 Hz, 1H, H<sup>4</sup>), 3.98 (td, *J* = 6.5, 1.1 Hz, 1H, H<sup>5</sup>), 3.87 (dd, *J* = 9.6, 3.5 Hz, 1H, H<sup>3</sup>), 3.52 (broad s, 1H, OH), 3.36 (broad s, 1H, OH).

### 7.1.13 3-O-Benzyl- $\alpha/\beta$ -D-Glucose (XI)



25 g of 1,2,5,6-Di-*O*-isopropylidene- $\alpha$ -D-glucopyranose (diacetone-D-glucose, 96 mmol, 1 eq) were dissolved in 250 mL of anhydrous DMF under Ar atmosphere. 22.8 mL of benzyl bromide (BnBr, 192 mmol, 2 eq) were added to the stirring solution that was cooled down to 0°C in an ice bath. Then 7.67 g of sodium hydride (NaH, 60% in mineral oil, 192 mmol, 2 eq) were added, and the mixture was stirred for 2 h at room temperature. When TLC showed reaction completion (R<sub>f</sub> = 0,40, ethyl acetate/cyclohexane 1:4), 10 mL of anhydrous methanol were added dropwise at 0°C, and solution was stirred for 1 h at room temperature. The solvent was evaporated, and the residue was diluted in diethyl ether (60 mL). The organic layer was washed with water (100 mL x 3), dried over MgSO<sub>4</sub>, filtered and evaporated, affording 3-O-Bn derivative as a honey-like oil (32.3 g, 95% crude yield).

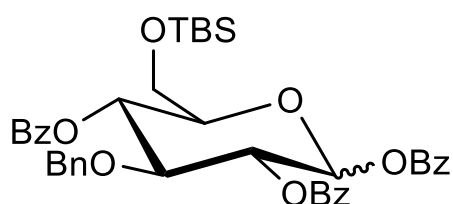
100 mL of H<sub>2</sub>O/TFA solution (1:1 v/v) were added to obtained crude and the mixture was stirred at room temperature overnight. The solvent was evaporated to dryness and the residue was then dissolved in water (100 mL). The obtained solution was washed with a mixture of ethyl acetate and ether (5:1 v/v, 30 mL x 2). Solvent was evaporated to dryness to give **XI** as a white foam (17.5 g, 68% yield over two steps).

TLC R<sub>f</sub> = 0,20 (DCM/MeOH 9:1) [revealed with (a) and (b)]

$^1\text{H}$  NMR (300 MHz, Deuterium Oxide)  $\delta$  7.44 – 7.27 (m,  $\text{H}^{\text{Ar}}$  Ph group), 5.13 (d,  $J = 3.7$  Hz,  $\text{H}^1$   $\alpha$  anomer), 4.79 (s, H PhCH<sub>2</sub>O), 4.56 (d,  $J = 7.9$  Hz,  $\text{H}^1$   $\beta$  anomer), 3.86 – 3.19 (m, H glucose).

NMR data are not reported in literature; however, they are in agreement with the desired structure and evidence the disappearance of isopropylidene signals of starting compound.

#### 7.1.14 1,2,4-Tri-O-Benzoyl-3-O-Benzyl-6-O-*tert*-butyldimethylsilyl- $\alpha/\beta$ -D-Glucopyranose (XII)



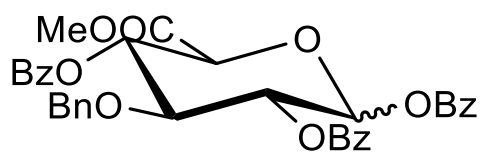
Tert-butyldimethylsilyl chloride (TBDMSCl, 600 mg, 4 mmol, 1.3 eq) was added to a stirred solution of **XI** (820 mg, 3 mmol, 1 eq) in pyridine (10 mL) at 0°C under Ar. The mixture was stirred at room temperature and monitored by TLC. After 2 h, benzoyl chloride (BzCl, 2.7 mL, 21 mmol, 2.3 eq/OH) was added dropwise at room temperature. After one night at RT, solvent was removed in vacuum. The residue was diluted with DCM (10 mL), washed with HCl 1M, sat. NaHCO<sub>3</sub>, brine, then dried with MgSO<sub>4</sub> and evaporated in vacuum. Resulting dark yellow oil was purified by flash chromatography (EtOAc/CyHex) to afford **XII** as a white powder (1.012 g, 48% yield).

TLC  $R_f = 0,34$  (EtOAc/CyHex 1:4) [revealed with (a) and (b)]

$^1\text{H}$  NMR (300 MHz, Chloroform-*d*)  $\delta$  8.43 – 6.99 (m, H *arom.*), 6.80 (d,  $J = 3.7$  Hz,  $\text{H}^1$   $\alpha$  anomer), 6.20 (d,  $J = 7.8$  Hz,  $\text{H}^1$   $\beta$  anomer), 5.74 (dd,  $J = 8.8, 7.8$  Hz), 5.68 – 5.52 (m), 4.86 – 4.57 (m), 4.50 (t,  $J = 9.6$  Hz), 4.22 (t,  $J = 8.8$  Hz), 4.01 – 3.85 (m), 0.88 (s, Si-C(CH<sub>3</sub>)<sub>3</sub>), 0.86 (s, Si-C(CH<sub>3</sub>)<sub>3</sub>), 0.03 (s, Si(CH<sub>3</sub>)<sub>2</sub>), 0.00 (s, Si(CH<sub>3</sub>)<sub>2</sub>).

Data are in accordance with those reported in literature<sup>[182]</sup>.

#### 7.1.15 Methyl 1,2,4-Tri-O-Benzoyl-3-O-Benzyl- $\alpha/\beta$ -D-Glucopyranosuronate (XIII)



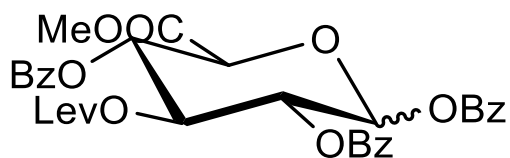
10 mL solution of Jones' reagent (1.56 g CrO<sub>3</sub> in 10 mL H<sub>2</sub>SO<sub>4</sub> 3,5 M, 4 eq) were added dropwise to a stirred solution of **XII** (1.9 g, 2.7 mmol, 1 eq) in acetone (30 mL) at 0°C. The mixture was warmed at RT and stirred for 5 h. H<sub>2</sub>O was added and extracted three times

with DCM. Combined organic fractions were dried (MgSO<sub>4</sub>), filtered and the solvent was removed in vacuum. The residue was dissolved in DMF (15 mL), CH<sub>3</sub>I (1.2 mL, 20 mmol) and KHCO<sub>3</sub> (2.4 g, 24 mmol) were added and the mixture was stirred overnight at RT. The suspension was filtered through Celite, the filtrate was diluted with H<sub>2</sub>O (20 mL) and extracted three times with AcOEt/Et<sub>2</sub>O 2:1 (15 mL), dried over MgSO<sub>4</sub> and evaporated under vacuum. The residue was purified by flash chromatography (EtOAc/CyHex 1:4) to afford **XIII** (1.26 g, 76% yield over two steps) as a white powder.

TLC R<sub>f</sub> = 0,20 (EtOAc/CyHex 1:4) [revealed with (a) and (b)]

<sup>1</sup>H NMR (300 MHz, Chloroform-*d*) δ 8.09 – 7.88 (m, H<sup>Ar</sup>), 7.69 – 7.31 (m, H<sup>Ar</sup>), 7.21 – 7.08 (m, H<sup>Ar</sup>), 6.83 (d, *J* = 3.5 Hz, H<sup>1</sup> α anomer), 6.25 (d, *J* = 6.1 Hz, H<sup>1</sup> β anomer), 5.82 (t, *J* = 7.4 Hz), 5.68 – 5.55 (m), 4.82 – 4.63 (m), 4.52 (d, *J* = 7.4 Hz), 4.48 (d, *J* = 9.0 Hz), 4.24 (t, *J* = 7.3 Hz), 3.67 (s, COOCH<sub>3</sub> α anomer), 3.51 (s, COOCH<sub>3</sub> β anomer). Data are in accordance with those reported in literature<sup>[182]</sup>.

#### 7.1.16 Methyl 1,2,4-Tri-O-Benzoyl-3-O-Levulinyl-α/β-D-Glucopyranosuronate (**XIV**)



Compound **XIII** (4.2 g, 6.9 mmol) was solubilized in 10 mL anhydrous DCM and hydrogenated in the presence of 10% Pd/C (400 mg) for 2 h under H<sub>2</sub> pressure (6 bar). The suspension was filtered through Celite and dried under vacuum. The oily residue was dissolved in anhydrous DCM (40 mL), then levulinic acid (3.15 mL, 24 mmol), DCC (4.96 g, 24 mmol) and DMAP (293.2 mg, 2.4 mmol) were added in this order and the mixture was stirred at RT overnight. Suspension was filtered through celite with DCM, washed with HCl 1 M (30 mL), NaHCO<sub>3</sub> sat. (30 mL x2), and brine (30 mL x2). The organic layer was dried over MgSO<sub>4</sub>, filtered and evaporated. The residue was purified by flash chromatography (EtOAc/CyHex 1:3) to afford **XIV** (3.65 g, 84% yield over two steps) as a white powder.

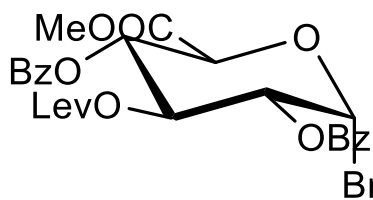
TLC R<sub>f</sub> = 0,20 (EtOAc/CyHex 1:3) [revealed with (a) and (b)]

<sup>1</sup>H NMR (300 MHz, Chloroform-*d*) δ 8.17 – 7.90 (m, H<sup>Bz</sup> arom.), 7.72 – 7.36 (m, H<sup>Bz</sup> arom.), 6.88 (d, *J* = 3.7 Hz, H<sup>1</sup> α anomer), 6.24 (d, *J* = 7.3 Hz, H<sup>1</sup> β anomer), 6.09 (t, *J* = 9.9 Hz, H<sup>3</sup> α anomer), 5.82 (t, *J* = 8.9 Hz, H<sup>3</sup> β anomer), 5.76 – 5.51 (m), 4.70 (d, *J* =

10.1 Hz), 4.53 (d,  $J = 9.0$  Hz), 3.70 (s, COOCH<sub>3</sub>  $\alpha$  anomer), 3.65 (s, COOCH<sub>3</sub>  $\beta$  anomer), 2.60 – 2.38 (m, CH<sub>2</sub> levulinate), 1.96 (s, CH<sub>3</sub> levulinate).

Data are in accordance with those reported in literature<sup>[182]</sup>.

### 7.1.17 Methyl 1-bromo-2,4-di-O-Benzoyl-3-O-Levulinyl- $\alpha$ -D-Glucopyranosuronate (XV)



HBr 33% in AcOH (10 mL) was added to a stirred solution of **XIV** (1.2 g, 2 mmol) in anhydrous DCM (10 mL) at 0°C under argon atmosphere. The mixture was immediately warmed up to RT and stirred for 2 h. NaHCO<sub>3</sub> sat. (20 mL) was slowly added to neutralize the solution, then organic layer was collected and washed with H<sub>2</sub>O, dried over MgSO<sub>4</sub>, filtered and evaporated in vacuum. The residue was purified by column chromatography (EtOAc/CyHex 1:4) to afford **XV** (550 mg, 48% yield) as a colorless foam. Pure starting compound **XIV** ( $\alpha$  anomer, 210 mg, 16%) was also recovered.

TLC R<sub>f</sub> = 0,45 (EtOAc/CyHex 1:3) [revealed with (a) and (b)]

<sup>1</sup>H NMR (300 MHz, Chloroform-*d*)  $\delta$  8.01 – 7.91 (m, 4H, H<sup>Ar</sup> Bz group), 7.61 – 7.48 (m, 2H, H<sup>Ar</sup> Bz group), 7.45 – 7.35 (m, 4H, H<sup>Ar</sup> Bz group), 6.76 (d,  $J = 4.0$  Hz, 1H, H<sup>1</sup>), 5.94 (t,  $J = 9.8$  Hz, 1H, H<sup>3</sup>), 5.48 (dd,  $J = 10.2, 9.8$  Hz, 1H, H<sup>4</sup>), 5.12 (dd,  $J = 10.0, 4.0$  Hz, 1H, H<sup>2</sup>), 4.70 (d,  $J = 10.2$ , 1H, H<sup>5</sup>), 3.60 (s, 3H, COOCH<sub>3</sub>), 2.49 – 2.43 (m, 2H, CH<sub>2</sub> levulinate), 2.34 (m, 2H, CH<sub>2</sub> levulinate), 1.85 (s, 3H, CH<sub>3</sub> levulinate).

Data are in accordance with those reported in literature<sup>[182]</sup>.

## 7.2 PEPTIDE SYNTHESIS

### 7.2.1 General procedure for peptides synthesis

All peptides were synthesized on a manual batch synthesizer (PLS 4x4, Advanced ChemTech) by Fmoc-based solid phase strategy in fritted syringes.

Generally, peptide chains were assembled by sequential coupling of activated Fmoc-amino acid (5 equiv) in DMF (1 mL x 100 mg resin) in the presence of HBTU (4.5-5 equiv) and DIEA (7-10 equiv) for 20 min at room temperature. Resins were then washed with DMF (x2) and CH<sub>2</sub>Cl<sub>2</sub> (x2) and the completeness of each coupling was

ascertained by the Kaiser's test. Fmoc deprotection was carried out twice (5 min, 10 min) by treatment with piperidine (20% v/v in DMF) followed by washing with DMF, DCM, and DMF. Fmoc-Asn[Glc(OAc<sub>4</sub>)]-OH, Fmoc-Lys(N<sub>3</sub>)-OH and Fmoc-Pra-OH couplings (2.5 eq) were performed in DMF (1 mL x 100 mg resin) in the presence of HATU (2.5 equiv) and DIEA (5 equiv) for 60 min at room temperature. N-terminus acetylation reactions were performed by double treatment for 10 min in DMF/Ac<sub>2</sub>O (4:2), 1 mL/100 mg resin. After complete elongation of the peptide chains, the resins were washed with DCM, then methanol, and finally dried.

The final cleavage from the resin and side-chain deprotection were performed by treatment for 3 h with TFA/TIS/H<sub>2</sub>O (95:2.5:2.5), 1 mL/ 100 mg resin. Precipitation in diethyl ether at 4°C and lyophilization afforded crude peptides.

The removal of the acetyl protecting groups of the β-D-glucopyranosyl moiety linked to the asparagine residue side chain was carried out by two different protocols<sup>[41,148]</sup>:

a) treatment with MeOH/THF/30% NH<sub>3</sub> aq. (2:2:1 ratio, 5 mg lyophilized crude/mL solution, pH 9). The reaction was quenched by adding conc. HCl to neutrality and the solvents were evaporated under vacuum. Water was added and the residue was lyophilized.

b) treatment with 0.05 M MeONa in dry MeOH (pH 12, 5-10 mg lyophilized crude/mL solution, 2-3 h). The reaction was quenched by adding conc. HCl to neutrality, the solvent was evaporated under vacuum and the residue lyophilized.

All peptides were purified through semipreparative RP-HPLC (>95%) to be used for autoantibody detection and characterized through RP-HPLC and mass spectrometry using methods and solvent system reported in tables.

Peptides **1-4** and **9-11** were purified by RP-HPLC on a ACE 5 C18-300" (250 × 10 mm) column (Waters, Saint Quentin en Yvelines, France) at 28°C using a Waters instrument (Separation Module 2695, detector diode array 2996) working at 10 mL/min, with the indicated linear gradients. The solvent systems used were: A (0.1% TFA in H<sub>2</sub>O) and B (0.1 % TFA in MeCN). Characterization was performed by RP-HPLC (UltiMate, Thermo Scientific) equipped with a ACE 5 C18-300" (250 × 4.6 mm) column working at 1 mL/min, with UV detection at 220 and 280 nm. The solvent systems used were: A (0.1% TFA in H<sub>2</sub>O) and B (0.1 % TFA in MeCN). Mass spectral analysis was performed by MALDI-TOF (Voyager-DETM PRO Workstation, Applied Biosystems) in positive ion reflector mode using the matrix α-Cyano-4-hydroxycinnamic acid (CHCA).

For peptides **5-8**, **12** and **CSF114(Glc)**, semi-preparative purifications via RP-HPLC were performed by a Phenomenex Jupiter C-18 (250 × 4.6 mm) column at 28 °C using a Waters instrument (Separation Module 2695, detector diode array 2996) working at 4 mL/min. The solvent systems used were: A (0.1% TFA in H<sub>2</sub>O) and B (0.1 % TFA in MeCN). Characterization was performed by RP-HPLC ESI-MS. HPLC system is an Alliance Chromatography (Waters) with a Phenomenex Kinetex C-18 column 2.6µm (100 × 3.0 mm) working at 0.6 mL/min, with UV detection at 215nm, coupled to a single quadrupole ESI-MS (Micromass ZQ). The solvent systems used were: A (0.1% TFA in H<sub>2</sub>O) and B (0.1 % TFA in MeCN).

The products were lyophilized with an Edwards apparatus, model Modulyo.

### 7.2.2 Synthesis of Peptides 1-4

Peptides were synthesized manually in a fritted syringe as described in the general procedure, on a Rink amide (p-methylbenzhydrylamine)-resin (MBHA-resin, 100–200 mesh, 0.62 mmol/g), in 0.05 mmol scale.

The conditions used for the purification and analysis of peptides 1-4, corresponding found MS values (calculated in square brackets) and yields are reported in table Table 8

Table 8 – Analytical data for peptides 1-4

Peptide	Purification Method	t <sub>r</sub> (min) <sup>a</sup>	[M+H] <sup>+</sup> m/z	Mass (mg)	Yield
<b>1</b>	5-40 % MeCN, 10 mL/min in 10 min	6.32 <sup>b</sup>	1002.28 [1002.55]	36	66%
<b>2</b>	10-30 % MeCN, 10 mL/min in 10 min	5.80 <sup>b</sup>	1186.18 [1164.60]	14	22%
<b>3</b>	10-30 % MeCN, 10 mL/min in 10 min	5.95 <sup>b</sup>	1326.35 [1326.66]	17	24%
<b>4</b>	10-30 % MeCN, 10 mL/min in 10 min	5.90 <sup>b</sup>	1164.37 [1164.60]	23	36%

<sup>a</sup>Analytical RP-HPLC gradients at 1 mL min<sup>-1</sup> in 10 min; solvent system A: 0.1% TFA in H<sub>2</sub>O, B: 0.1% TFA in MeCN. <sup>b</sup> 10-30% B in 10 min.

### 7.2.3 Synthesis of Peptides 5-6

Peptides **5** and **6** were synthesized manually in a fritted syringe as described in the general procedure, on a Rink amide (p-methylbenzhydrylamine)-resin (MBHA-resin, 100–200 mesh, 0.62 mmol/g), in 0.1 mmol scale.

The conditions used for the purification and analysis of peptides **5** and **6**, corresponding found MS values (calculated in square brackets) and yields are reported in Table 9.

Table 9 – Analytical data for peptides **5** and **6**

Peptide	Purification Method	t <sub>r</sub> (min) <sup>a</sup>	m/z	Mass (mg)	Yield
<b>5</b>	20-40 % MeCN, 4 mL/min in 30 min	3.28 <sup>b</sup>	1478.9 [1478.7]	40	25%
<b>6</b>	20-35 % MeCN, 4 mL/min in 30 min	3.90 <sup>b</sup>	1538.2 [1537.8]	22	13%

<sup>a</sup>Analytical RP-HPLC gradients at 0.6 mL min<sup>-1</sup> in 5 min; solvent system A: 0.1% TFA in H<sub>2</sub>O, B: 0.1% TFA in MeCN. <sup>b</sup> 10-60% B in 5 min.

### 7.2.4 Synthesis of Peptide 7

Peptide **5** (5.1 mg, 3.4 μmol) and peptide **6** (5.2 mg, 3.4 μmol) were dissolved in 2 mL of water and then 1 mL of *t*BuOH was added. 100 μL of 1 M ascorbic acid solution and 50 μL of 1M copper sulfate solution were added, and the mixture was stirred at room temperature. The reaction was monitored by RP-HPLC and stopped after 3 h. 0.1 M EDTA solution was added to chelate copper and the resulting solution was freeze-dried. Lyophilized crude was dissolved in 3 mL of water and absorbed through a 10 mL C<sub>18</sub> column, previously equilibrated with water. The column was washed with increasing concentrations of MeCN in water (0-50% v/v), and eluted fractions were analyzed by RP-HPLC. Fractions containing desired product (25% and 30%) were collected and lyophilized. Obtained peptide **7** (9 mg) was furtherly purified by semi-preparative RP-HPLC.

The conditions used for the purification and analysis of peptide **7**, corresponding found MS value (calculated in square brackets) and yield are reported in Table 10.

Table 10 – Analytical data for peptide 7

Peptide	Purification Method	t <sub>r</sub> (min) <sup>a</sup>	[M+3H] <sup>3+</sup> m/z	Mass (mg)	Yield
7	20-40 % MeCN, 4 mL/min in 30 min	3.50 <sup>b</sup>	1006.5 [1006.2]	7	70%

<sup>a</sup>Analytical RP-HPLC gradients at 0.6 mL min<sup>-1</sup> in 5 min; solvent system

A: 0.1% TFA in H<sub>2</sub>O, B: 0.1% TFA in MeCN. <sup>b</sup> 10-60% B in 5 min.

### 7.2.5 Synthesis of Peptide 8

Peptide **8** was synthesized manually in a fritted syringe as described in the general procedure, on a Fmoc-βAla-Wang resin (100–200 mesh, 0.7 mmol/g), in 0.1 mmol scale.

The conditions used for the purification and analysis of peptide **8**, corresponding found MS value (calculated in square brackets) and yield are reported in Table 11.

Table 11 – Analytical data for peptide 8

Peptide	Purification Method	t <sub>r</sub> (min) <sup>a</sup>	[M+2H] <sup>2+</sup> m/z	Mass (mg)	Yield
8	20-35 % MeCN, 4 mL/min in 30 min	2.82 <sup>b</sup>	1296.9 [1296.6]	25	9%

<sup>a</sup>Analytical RP-HPLC gradients at 0.6 mL min<sup>-1</sup> in 5 min; solvent system

A: 0.1% TFA in H<sub>2</sub>O, B: 0.1% TFA in MeCN. <sup>b</sup> 20-80% B in 5 min.

### 7.2.6 Synthesis of Peptide 9-10

Peptides **9-10** were synthesized manually in a fritted syringe as described in the general procedure, on a Rink amide (p-methylbenzhydrylamine)-resin (MBHA-resin, 100–200 mesh, 0.62 mmol/g), in 0.05 mmol scale. Instead of N-terminal acetylation, after the final Fmoc group removal 4-azido-benzoic acid (6 eq), HBTU (5 eq) and DIEA (10 eq) in 2 mL DMF were added for 1 h. The resins were dried and final cleavage was performed according to the general procedure.

Peptide **10** was then deacetylated as described in the general procedure. Peptide crudes were purified and analytical data are reported in Table 12.

Table 12 – Analytical data for peptide 9 and 10

Peptide	Purification Method	t <sub>r</sub> (min) <sup>a</sup>	[M+H] <sup>+</sup> m/z	Mass (mg)	Yield
<b>9</b>	20-50 % MeCN, 10 mL/min in 10 min	6.63 <sup>b</sup>	1105.10 [1105.57]	24	42%
<b>10</b>	20-50% MeCN, 10 mL/min in 10 min	4.81 <sup>c</sup>	1429.25 [1429.67]	14	18%

<sup>a</sup>Analytical RP-HPLC gradients at 1 mL min<sup>-1</sup> in 10 min; solvent system A: 0.1% TFA in H<sub>2</sub>O, B: 0.1% TFA in MeCN. <sup>b</sup> 15-45% B in 10 min. <sup>c</sup> 5-90% B in 10 min.

### 7.2.7 Synthesis of Peptide 11-12

Peptides **11** was synthesized manually in a fritted syringe as described in the general procedure, on a Rink amide (p-methylbenzhydrylamine)-resin (MBHA-resin, 100–200 mesh, 0.62 mmol/g), in 0.05 mmol scale. For the final cleavage, a mixture of TFA/TIS/H<sub>2</sub>O/EDT (94:2:2:2) was used to prevent Cys oxidation.

Peptide crude was purified, and analytical data are reported in

Peptide **11** (6.2 mg, 6.2 μmol) was solubilized in 800 μL H<sub>2</sub>O/MeCN (1:1 v/v, 3 mg/mL) and 200 μL of a solution 60 mM of compound **VI** in MeCN (12 μmol) was added.

2 μL of 2,6-lutidine (17 μmol, pH≈7-8 by pH paper) were added and the solution monitored by RP-HPLC. After 3 h, water (≈10 mL) was added to freeze-dry.

Lyophilized crude was dissolved in minimum of H<sub>2</sub>O/MeCN (1:1) and absorbed through a 10 mL C<sub>18</sub> column, previously equilibrated with water. The column was washed with increasing concentrations of MeCN in water (0-100% v/v). Obtained peptide fraction (5 mg) was deacetylated as described in the general procedure and purified by semi-preparative RP-HPLC, affording pure peptide **12** (2 mg, 27% yield).

Table 13 – Analytical data for peptides 11-12

Peptide	Purification Method	t <sub>r</sub> (min) <sup>a</sup>	[M+H] <sup>+</sup> m/z	Mass (mg)	Yield
<b>11</b>	5-40 % MeCN, 10 mL/min in 10 min	5.96 <sup>b</sup>	991.43 [991.52]	23	42%

<b>12</b>	20-50% MeCN, 4 mL/min in 30 min	3.95 <sup>c</sup>	1210.30 [1210.59]	2	27%
-----------	------------------------------------	-------------------	----------------------	---	-----

<sup>a</sup>Analytical RP-HPLC gradients; solvent system A: 0.1% TFA in H<sub>2</sub>O, B: 0.1% TFA in MeCN. <sup>b</sup> 15-45% B at 1 mL min<sup>-1</sup> in 10 min. <sup>c</sup> 10-90% B in at 0.6 mL min<sup>-1</sup> in 5 min.

### Synthesis of CSF114(Glc)

CSF114(Glc) was synthesized manually in a fritted syringe as described in the general procedure, using a Fmoc-Lys(Boc)-Wang resin (100–200 mesh, 0.24 mmol/g), in 0.05 mmol scale.

The conditions used for the purification and analysis of deacetylated CSF114(Glc), corresponding found MS value (calculated in square brackets) and yields are reported in Table 14.

Table 14 – Analytical data for CSF114(Glc)

Peptide	Purification Method	t <sub>r</sub> (min) <sup>a</sup>	[M+3H] <sup>3+</sup> m/z	Mass (mg)	Yield
CSF114(Glc)	20-35 % MeCN, 4 mL/min in 30 min	2.80 <sup>b</sup>	869.9 [870.0]	27	21%

<sup>a</sup>Analytical RP-HPLC gradients at 0.6 mL min<sup>-1</sup> in 5 min; solvent system A: 0.1% TFA in H<sub>2</sub>O, B: 0.1% TFA in MeCN. <sup>b</sup> 20-80% B in 5 min.

## 7.3 PEPTIDE CONFORMATIONAL ANALYSIS

NMR and CD experiments of peptides **1-4** were conducted under the supervision of Prof. Olivier Lequin (Laboratoire de Biomolécules) at the Sorbonne Université scientific campus (Jussieu, Paris).

### 7.3.1 NMR of peptides 1-4

NMR spectra were recorded on a Bruker Avance III spectrometer (Wissembourg, France) equipped with a TCI cryoprobe at a <sup>1</sup>H frequency of 500 MHz. Lyophilized peptides were dissolved in 550 μL of 90% H<sub>2</sub>O/10% D<sub>2</sub>O at ≈ 5 mM concentration. The pH was set to 5.0 using a buffered solution of sodium succinate (25 mM final

concentration). Sodium 2,2-dimethyl-2-silapentane-5-sulfonate (DSS, 0.1 mM) was used as an internal reference for chemical shift calibration.  $^1\text{H}$ ,  $^{13}\text{C}$ , and  $^{15}\text{N}$  resonances of all atoms were assigned by recording 2D  $^1\text{H}$ - $^1\text{H}$  TOCSY (72 ms mixing time), 2D  $^1\text{H}$ - $^1\text{H}$  ROESY (400 ms), 2D  $^1\text{H}$ - $^1\text{H}$  NOESY (300 ms), heteronuclear 2D  $^1\text{H}$ - $^{13}\text{C}$  HSQC, 2D  $^1\text{H}$ - $^{13}\text{C}$  HSQC-TOCSY, 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC and 2D  $^1\text{H}$ - $^{13}\text{C}$  HMBC spectra (Table 15).

NMR experiments were processed with TOPSPIN software (Bruker). Spectra were analyzed using Sparky software<sup>[198]</sup>. Exact peptide concentrations were calculated on 1D  $^1\text{H}$  WATERGATE (D1=15s) by integration of 9 HN x 110% in 8.65-8.08 ppm region or 4 Me L/V in 0.99-0.80 ppm region. Identified non-sequential ROESY correlation are reported in Table 16.  $^3J_{\text{HN-H}\alpha}$  and  $^3J_{\text{H}\alpha\text{-H}\beta}$  coupling constants were measured on 1D  $^1\text{H}$  WATERGATE spectra (Table 17). The chemical shift deviations of  $\text{H}^{\alpha}$  protons and  $\text{C}^{\alpha}$  carbons were calculated as the differences between observed chemical shifts and random coil values reported in water.

Table 15 - Complete  $^1\text{H}$ ,  $^{13}\text{C}$ ,  $^{15}\text{N}$  NMR assignments of peptides 1-4.

Residue	Atom	Nucleus	PEPTIDE 1	PEPTIDE 2	PEPTIDE 3	PEPTIDE 4
Ac	C	$^{13}\text{C}$	177.2	177.2	177.2	177.2
Ac	CA	$^{13}\text{C}$	24.4	24.5	24.4	24.5
Ac	HA	$^1\text{H}$	2.03	2.03	2.03	2.03
K1	C	$^{13}\text{C}$	176.8	176.8	176.8	176.8
K1	CA	$^{13}\text{C}$	56.4	56.4	56.4	56.4
K1	CB	$^{13}\text{C}$	33.3	33.3	33.3	33.3
K1	CD	$^{13}\text{C}$	29.2	29.2	29.2	29.2
K1	CE	$^{13}\text{C}$	42.2	42.2	42.1	42.1
K1	CG	$^{13}\text{C}$	24.8	24.8	24.8	24.8
K1	HA	$^1\text{H}$	4.26	4.25	4.25	4.26
K1	HB2	$^1\text{H}$	1.81	1.81	1.80	1.81
K1	HB3	$^1\text{H}$	1.71	1.72	1.72	1.72
K1	HD	$^1\text{H}$	1.69	1.69	1.69	1.69
K1	HE	$^1\text{H}$	3.00	3.00	3.00	3.00
K1	HG	$^1\text{H}$	1.45	1.45	1.45	1.45
K1	HN	$^1\text{H}$	8.32	8.32	8.32	8.32

<b>K1</b>	<b>HZ</b>	<b><sup>1</sup>H</b>	7.57	7.57	7.57	7.57
<b>K1</b>	<b>N</b>	<b><sup>15</sup>N</b>	127.2	127.2	127.2	127.2
<b>K1</b>	<b>NZ</b>	<b><sup>15</sup>N</b>	62.8	62.8	62.8	62.8
<b>A2</b>	<b>C</b>	<b><sup>13</sup>C</b>	177.5	177.5	177.5	177.6
<b>A2</b>	<b>CA</b>	<b><sup>13</sup>C</b>	52.5	52.5	52.4	52.5
<b>A2</b>	<b>CB</b>	<b><sup>13</sup>C</b>	19.3	19.4	19.3	19.3
<b>A2</b>	<b>HA</b>	<b><sup>1</sup>H</b>	4.31	4.30	4.31	4.30
<b>A2</b>	<b>HB</b>	<b><sup>1</sup>H</b>	1.38	1.38	1.38	1.38
<b>A2</b>	<b>HN</b>	<b><sup>1</sup>H</b>	8.45	8.44	8.45	8.44
<b>A2</b>	<b>N</b>	<b><sup>15</sup>N</b>	125.3	125.3	125.3	125.3
<b>Glc3</b>	<b>C1</b>	<b><sup>13</sup>C</b>	---	82.1	---	82.0
<b>Glc3</b>	<b>C2</b>	<b><sup>13</sup>C</b>	---	74.6	---	74.6
<b>Glc3</b>	<b>C3</b>	<b><sup>13</sup>C</b>	---	79.4	---	79.3
<b>Glc3</b>	<b>C4</b>	<b><sup>13</sup>C</b>	---	71.9	---	72.0
<b>Glc3</b>	<b>C5</b>	<b><sup>13</sup>C</b>	---	80.3	---	80.3
<b>Glc3</b>	<b>C6</b>	<b><sup>13</sup>C</b>	---	63.3	---	63.3
<b>Glc3</b>	<b>H1</b>	<b><sup>1</sup>H</b>	---	4.94	---	4.95
<b>Glc3</b>	<b>H2</b>	<b><sup>1</sup>H</b>	---	3.40	---	3.39
<b>Glc3</b>	<b>H3</b>	<b><sup>1</sup>H</b>	---	3.54	---	3.54
<b>Glc3</b>	<b>H4</b>	<b><sup>1</sup>H</b>	---	3.44	---	3.43
<b>Glc3</b>	<b>H5</b>	<b><sup>1</sup>H</b>	---	3.49	---	3.49
<b>Glc3</b>	<b>H6A</b>	<b><sup>1</sup>H</b>	---	3.73	---	3.74
<b>Glc3</b>	<b>H6B</b>	<b><sup>1</sup>H</b>	---	3.86	---	3.87
<b>Glc3</b>	<b>HN</b>	<b><sup>1</sup>H</b>	---	8.99	---	8.98
<b>Glc3</b>	<b>N</b>	<b><sup>15</sup>N</b>	---	132.2	---	132.4
<b>N3</b>	<b>C</b>	<b><sup>13</sup>C</b>	175.3	175.2	175.3	175.2
<b>N3</b>	<b>CA</b>	<b><sup>13</sup>C</b>	53.3	53.0	53.3	53.0
<b>N3</b>	<b>CB</b>	<b><sup>13</sup>C</b>	38.7	39.1	38.8	39.1
<b>N3</b>	<b>CG</b>	<b><sup>13</sup>C</b>	177.2	175.5	177.2	175.5
<b>N3</b>	<b>HA</b>	<b><sup>1</sup>H</b>	4.70	4.74	4.69	4.74
<b>N3</b>	<b>HB2</b>	<b><sup>1</sup>H</b>	2.83	2.91	2.75	2.91
<b>N3</b>	<b>HB3</b>	<b><sup>1</sup>H</b>	2.75	2.81	2.83	2.81
<b>N3</b>	<b>HD21</b>	<b><sup>1</sup>H</b>	7.67	---	7.67	---

<b>N3</b>	<b>HD22</b>	<b><sup>1</sup>H</b>	6.97	---	6.96	---
<b>N3</b>	<b>HN</b>	<b><sup>1</sup>H</b>	8.52	8.55	8.52	8.55
<b>N3</b>	<b>N</b>	<b><sup>15</sup>N</b>	118.7	118.7	118.7	118.7
<b>N3</b>	<b>ND2</b>	<b><sup>15</sup>N</b>	113.1	---	113.2	---
<b>V4</b>	<b>C</b>	<b><sup>13</sup>C</b>	176.3	176.3	176.3	176.3
<b>V4</b>	<b>CA</b>	<b><sup>13</sup>C</b>	62.4	62.4	62.3	62.3
<b>V4</b>	<b>CB</b>	<b><sup>13</sup>C</b>	32.9	32.9	32.9	32.9
<b>V4</b>	<b>CG1</b>	<b><sup>13</sup>C</b>	21.2	21.1	21.1	21.2
<b>V4</b>	<b>CG2</b>	<b><sup>13</sup>C</b>	20.4	20.4	20.4	20.4
<b>V4</b>	<b>HA</b>	<b><sup>1</sup>H</b>	4.20	4.20	4.20	4.21
<b>V4</b>	<b>HB</b>	<b><sup>1</sup>H</b>	2.12	2.12	2.11	2.11
<b>V4</b>	<b>HG1</b>	<b><sup>1</sup>H</b>	0.93	0.94	0.93	0.93
<b>V4</b>	<b>HG2</b>	<b><sup>1</sup>H</b>	0.93	0.93	0.93	0.93
<b>V4</b>	<b>HN</b>	<b><sup>1</sup>H</b>	8.16	8.15	8.16	8.14
<b>V4</b>	<b>N</b>	<b><sup>15</sup>N</b>	120.7	120.4	120.7	120.4
<b>T5</b>	<b>C</b>	<b><sup>13</sup>C</b>	174.4	174.3	174.3	174.3
<b>T5</b>	<b>CA</b>	<b><sup>13</sup>C</b>	62.1	62.2	62.0	62.0
<b>T5</b>	<b>CB</b>	<b><sup>13</sup>C</b>	69.8	69.8	69.8	69.8
<b>T5</b>	<b>CG2</b>	<b><sup>13</sup>C</b>	21.7	21.7	21.7	21.7
<b>T5</b>	<b>HA</b>	<b><sup>1</sup>H</b>	4.33	4.32	4.35	4.32
<b>T5</b>	<b>HB</b>	<b><sup>1</sup>H</b>	4.17	4.15	4.17	4.16
<b>T5</b>	<b>HG2</b>	<b><sup>1</sup>H</b>	1.20	1.21	1.21	1.21
<b>T5</b>	<b>HN</b>	<b><sup>1</sup>H</b>	8.37	8.39	8.37	8.40
<b>T5</b>	<b>N</b>	<b><sup>15</sup>N</b>	119.1	119.4	119.2	119.4
<b>L6</b>	<b>C</b>	<b><sup>13</sup>C</b>	176.9	176.9	176.9	176.9
<b>L6</b>	<b>CA</b>	<b><sup>13</sup>C</b>	55.2	55.2	55.1	55.1
<b>L6</b>	<b>CB</b>	<b><sup>13</sup>C</b>	42.5	42.5	42.6	42.6
<b>L6</b>	<b>CD1</b>	<b><sup>13</sup>C</b>	24.8	23.5	24.8	24.8
<b>L6</b>	<b>CD2</b>	<b><sup>13</sup>C</b>	23.5	21.7	23.5	23.5
<b>L6</b>	<b>CG</b>	<b><sup>13</sup>C</b>	27.0	27.0	27.0	27.0
<b>L6</b>	<b>HA</b>	<b><sup>1</sup>H</b>	4.38	4.38	4.37	4.37
<b>L6</b>	<b>HB2</b>	<b><sup>1</sup>H</b>	1.64	1.65	1.64	1.64
<b>L6</b>	<b>HB3</b>	<b><sup>1</sup>H</b>	1.60	1.60	1.57	1.59

<b>L6</b>	<b>HD1</b>	<b><sup>1</sup>H</b>	0.92	0.93	0.92	0.92
<b>L6</b>	<b>HD2</b>	<b><sup>1</sup>H</b>	0.86	0.87	0.86	0.86
<b>L6</b>	<b>HG</b>	<b><sup>1</sup>H</b>	1.61	1.61	1.61	1.61
<b>L6</b>	<b>HN</b>	<b><sup>1</sup>H</b>	8.37	8.39	8.38	8.40
<b>L6</b>	<b>N</b>	<b><sup>15</sup>N</b>	125.5	125.6	125.6	125.7
<b>Glc7</b>	<b>C1</b>	<b><sup>13</sup>C</b>	---	---	82.0	82.0
<b>Glc7</b>	<b>C2</b>	<b><sup>13</sup>C</b>	---	---	74.6	74.6
<b>Glc7</b>	<b>C3</b>	<b><sup>13</sup>C</b>	---	---	79.3	79.2
<b>Glc7</b>	<b>C4</b>	<b><sup>13</sup>C</b>	---	---	72.0	72.0
<b>Glc7</b>	<b>C5</b>	<b><sup>13</sup>C</b>	---	---	80.3	80.3
<b>Glc7</b>	<b>C6</b>	<b><sup>13</sup>C</b>	---	---	63.3	63.3
<b>Glc7</b>	<b>H1</b>	<b><sup>1</sup>H</b>	---	---	4.94	4.94
<b>Glc7</b>	<b>H2</b>	<b><sup>1</sup>H</b>	---	---	3.38	3.38
<b>Glc7</b>	<b>H3</b>	<b><sup>1</sup>H</b>	---	---	3.53	3.53
<b>Glc7</b>	<b>H4</b>	<b><sup>1</sup>H</b>	---	---	3.41	3.41
<b>Glc7</b>	<b>H5</b>	<b><sup>1</sup>H</b>	---	---	3.49	3.50
<b>Glc7</b>	<b>H6a</b>	<b><sup>1</sup>H</b>	---	---	3.71	3.73
<b>Glc7</b>	<b>H6b</b>	<b><sup>1</sup>H</b>	---	---	3.87	3.86
<b>Glc7</b>	<b>HN</b>	<b><sup>1</sup>H</b>	---	---	8.98	8.97
<b>Glc7</b>	<b>N</b>	<b><sup>15</sup>N</b>	---	---	132.0	132.0
<b>N7</b>	<b>C</b>	<b><sup>13</sup>C</b>	175.6	175.6	175.7	175.7
<b>N7</b>	<b>CA</b>	<b><sup>13</sup>C</b>	53.3	53.3	52.9	52.9
<b>N7</b>	<b>CB</b>	<b><sup>13</sup>C</b>	38.8	38.8	39.1	39.1
<b>N7</b>	<b>CG</b>	<b><sup>13</sup>C</b>	177.2	177.2	175.5	175.5
<b>N7</b>	<b>HA</b>	<b><sup>1</sup>H</b>	4.80	4.79	4.84	4.83
<b>N7</b>	<b>HB2</b>	<b><sup>1</sup>H</b>	2.88	2.88	2.84	2.97
<b>N7</b>	<b>HB3</b>	<b><sup>1</sup>H</b>	2.78	2.78	2.97	2.84
<b>N7</b>	<b>HD21</b>	<b><sup>1</sup>H</b>	7.67	7.67	---	---
<b>N7</b>	<b>HD22</b>	<b><sup>1</sup>H</b>	6.95	6.96	---	---
<b>N7</b>	<b>HN</b>	<b><sup>1</sup>H</b>	8.58	8.57	8.61	8.61
<b>N7</b>	<b>N</b>	<b><sup>15</sup>N</b>	120.2	120.2	120.5	120.4
<b>N7</b>	<b>ND2</b>	<b><sup>15</sup>N</b>	113.0	113.0	---	---
<b>T8</b>	<b>C</b>	<b><sup>13</sup>C</b>	174.9	174.9	175.0	175.0

T8	CA	<sup>13</sup> C	61.9	61.9	61.9	61.9
T8	CB	<sup>13</sup> C	69.7	69.6	69.6	69.6
T8	CG2	<sup>13</sup> C	21.5	21.5	21.5	21.6
T8	HA	<sup>1</sup> H	4.43	4.43	4.42	4.42
T8	HB	<sup>1</sup> H	4.33	4.33	4.35	4.35
T8	HG2	<sup>1</sup> H	1.20	1.20	1.21	1.21
T8	HN	<sup>1</sup> H	8.27	8.27	8.31	8.31
T8	N	<sup>15</sup> N	114.7	114.7	114.7	114.6
T9	C	<sup>13</sup> C	177.3	177.3	177.3	177.4
T9	CA	<sup>13</sup> C	61.9	61.9	62.0	62.0
T9	CB	<sup>13</sup> C	69.8	69.8	69.7	69.7
T9	CG2	<sup>13</sup> C	21.6	21.6	21.6	21.6
T9	HA	<sup>1</sup> H	4.34	4.34	4.33	4.33
T9	HB	<sup>1</sup> H	4.28	4.28	4.28	4.28
T9	HG2	<sup>1</sup> H	1.22	1.22	1.22	1.22
T9	HN	<sup>1</sup> H	8.18	8.18	8.22	8.22
T9	HT1	<sup>1</sup> H	7.63	7.63	7.60	7.60
T9	HT2	<sup>1</sup> H	7.23	7.23	7.25	7.25
T9	N	<sup>15</sup> N	116.2	116.2	116.3	116.3
T9	NT	<sup>15</sup> N	109.8	109.8	109.8	109.8

Table 16 - ROESY *i/i+2* correlations. Beside sequential  $H^{\alpha}$ - $H^N$  and  $H^N$ - $H^N$  ROESY correlations, ROESY *i/i+2* weak correlations, indicative of turn propensity, were detected.

Peptide 1	Peptide 2	Peptide 3	Peptide 4
A2 <sup>HB</sup> - V4 <sup>HN</sup>	A2 <sup>HB</sup> - V4 <sup>HN</sup>	A2 <sup>HB</sup> - V4 <sup>HN</sup>	A2 <sup>HB</sup> - V4 <sup>HN</sup>
N3 <sup>HA</sup> - T5 <sup>HN</sup>	---	N3 <sup>HA</sup> - T5 <sup>HN</sup>	N3 <sup>HA</sup> - T5 <sup>HN</sup>
T5 <sup>HB</sup> - N7 <sup>HN</sup>	T5 <sup>HB</sup> - N7 <sup>HN</sup>	T5 <sup>HB</sup> - N7 <sup>HN</sup>	T5 <sup>HB</sup> - N7 <sup>HN</sup>
T5 <sup>HG2</sup> - N7 <sup>HN</sup>	T5 <sup>HG2</sup> - N7 <sup>HN</sup>	T5 <sup>HG2</sup> - N7 <sup>HN</sup>	T5 <sup>HG2</sup> - N7 <sup>HN</sup>
L6 <sup>HB3</sup> - T8 <sup>HN</sup>	L6 <sup>HB3</sup> - T8 <sup>HN</sup>	---	---

Table 17 - <sup>3</sup>J(HN-H $\alpha$ ) measured on 1D <sup>1</sup>H WATERGATE. Superimposed signals provided less certainty in the measurement, and corresponding values are reported in red.

	Peptide 1	Peptide 2	Peptide 3	Peptide 4
<b>K1</b>	6,6	6,7	7	6,4

<b>A2</b>	5,7	5,6	5,7	5,6
<b>Glc3</b>	---	8,7	---	8,7
<b>N3</b>	7,1	7	7,1	7,1
<b>V4</b>	8	7,7	7,8	7,7
<b>T5</b>	7,9	7,9	7,9	7,9
<b>L6</b>	6,7	6,7	6,8	6,7
<b>Glc7</b>	---	---	8,8	8,7
<b>N7</b>	7,3	7,3	7,3	7,3
<b>T8</b>	7,9	7,9	8,0	7,9
<b>T9</b>	8,0	8,0	8,0	7,2

### 7.3.2 CD of peptides 1-4

All CD spectra were recorded on a Jasco J-810 spectropolarimeter using cells of 1 mm path length. Peptide buffered solutions (exactly titrated by <sup>1</sup>H-NMR) were diluted in TFE 0%, 20%, 40% and 60% in water, to obtain a final peptide concentration of 85-100 μM.

Spectra were the average of ten scans from 185 to 260 nm, recorded with a band width of 0.5 nm at scan rate of 5 nm/min. Collected data were normalized considering the concentration and the number of amide bonds for each peptide (Figure 51).

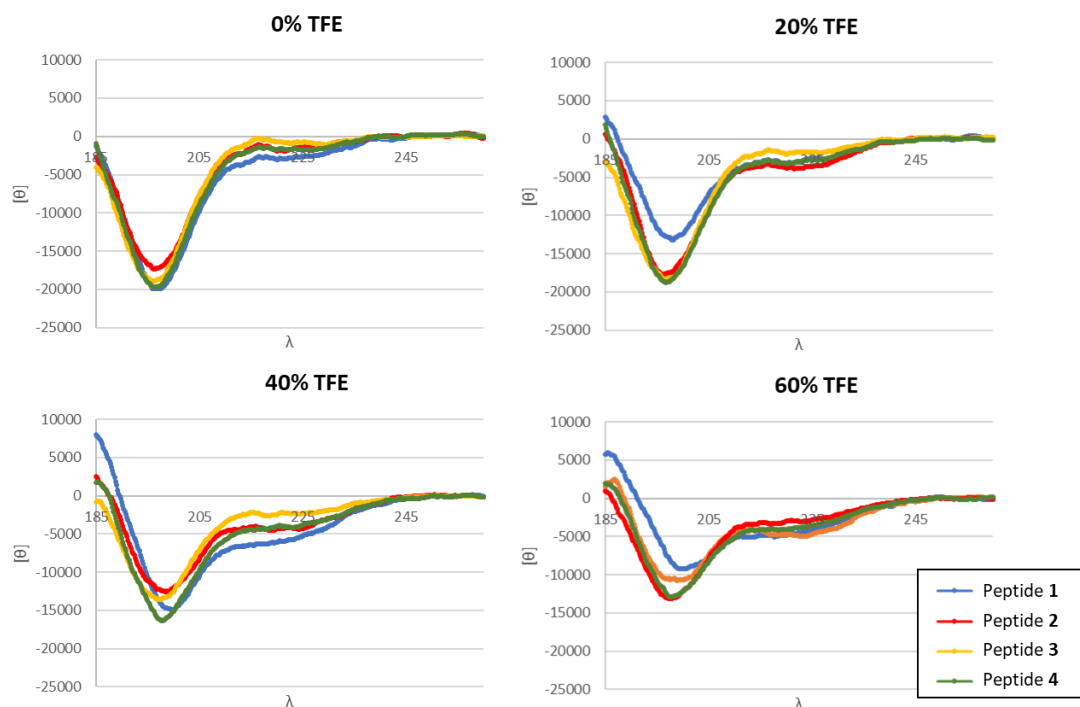


Figure 51 – Normalized CD spectra of peptides 1-4 recorded between 185-260 nm at increasing concentrations of TFE in water (0%-60%)

## 7.4 DEXTRAN CONJUGATES

### 7.4.1 Synthesis of Dex40-GP

To a stirring solution of dextran 40 kDa (1 g, 25  $\mu\text{mol}$  dextran, 6.2 mmol glucose units) in 0.1 M aq. NaOH (7 mL) was added glycidyl propargyl ether (1 mL, 9.3 mmol). The mixture was stirred at 35°C for 20 h then added dropwise in isopropyl alcohol (100 mL). The white precipitated compound was thoroughly filtered through a sintered glass filter and washed with 1-propanol. The crude solid was dissolved in water ( $\approx$  9 mL) and dialyzed in a cassette (Slide-A-Lyzer 10K MWCO, Thermo Fisher Scientific) against Milli-Q water until the conductivity of the solution was  $\approx$  0 (48 h). The alkyne dextran (1.103 g, Final estimated MW  $\approx$  55 kDa, 80% recovery yield) was obtained as a white dense powder after lyophilization. The degree of substitution (DS) was determined by  $^1\text{H-NMR}$ .

### 7.4.2 Synthesis of Dex40-Pept6

To a stirring solution of Dex40-GP (10 mg, 13.0  $\mu\text{mol}$  propargyl groups, 1 eq) and peptide **6** (15 mg, 9.8  $\mu\text{mol}$ , 0.75 eq) in DMSO (800  $\mu\text{L}$ ) and water (500  $\mu\text{L}$ ), a mixture of 100  $\mu\text{L}$  of  $\text{CuSO}_4$  1M (5 eq) and 200  $\mu\text{L}$  Ascorbic acid 1M (10 eq) was added.

The pale-orange solution was left under stirring overnight. Then 20 mL of a solution EDTA 10 mM were added and the solution was lyophilized. The dried crude was dissolved in 15 mL H<sub>2</sub>O Milli-Q and centrifuged with Amicon Ultra Centrifugal Filters (MWCO = 30 kDa), then buffer-exchange steps with H<sub>2</sub>O Milli-Q were performed to remove all EDTA (final washing volume  $\approx$ 100 mL, Dex40-Pept6  $\approx$ 1 mL in H<sub>2</sub>O). After lyophilization the desired product is obtained as a white powder (23 mg, Degree of Substitution = 19.5 % (e.g. 1 peptide chain / 5 glucose units), Final estimated MW  $\approx$  130 kDa, 89% yield). The degree of substitution (DS) was determined by proton NMR.

### 7.4.3 NMR analysis of dextran-based conjugates

NMR analysis for substitution degree calculations were conducted in collaboration with Dr Monica Bertoldo and Dr Paolo Dambruoso at the Institute of Organic Synthesis and Photoreactivity (ISOF) in Bologna, Italy.

NMR spectra of Dex40, Dex40-GP, Peptide 6 and Dex40-Pept6 were acquired in D<sub>2</sub>O (15-25 mg/mL) with a Bruker 500 MHz spectrometer. Preliminary inversion recovery experiments were conducted for each sample in order to find the maximum longitudinal relaxation time (T<sub>1</sub>) and set the appropriate delay time (d<sub>1</sub>) for quantitative <sup>1</sup>H-NMR analyses. <sup>1</sup>H spectra were calibrated using the residual water signal in accordance with literature<sup>[199]</sup>. <sup>13</sup>C chemical shifts were calibrated indirectly based on the <sup>1</sup>H calibration.

#### 1.1.1.1 NMR of Dextran 40 kDa (Dex40)

NMR spectra of Dex40 present the main peaks originating from protons and carbons of the  $\alpha$ -1,6 linked glucoses (most abundant form). Other minor peaks are assigned in Figure 52.

<sup>1</sup>H NMR (500 MHz, Deuterium Oxide)  $\delta$  5.00 (d,  $J$  = 3.6 Hz, 1H, H<sup>1</sup>), 4.01 (dd,  $J$  = 11.5, 4.6 Hz, 1H, H<sup>6</sup>), 3.93 (ddd,  $J$  = 10.2, 4.5, 2.2 Hz, 1H, H<sup>5</sup>), 3.77 (*pseudo*-d,  $J$  = 11.5 Hz, 1H, H<sup>6</sup>), 3.74 (*pseudo*-t,  $J$  = 10.0 Hz, 1H, H<sup>4</sup>), 3.59 (dd,  $J$  = 9.9, 3.7 Hz, 1H, H<sup>2</sup>), 3.54 (*pseudo*-t,  $J$  = 9.5 Hz, 1H, H<sup>3</sup>).

<sup>13</sup>C NMR (126 MHz, d<sub>2</sub>o)  $\delta$  100.5 (C1), 76.2 (C3), 74.2 (C2), 73.0 (C4), 72.3 (C5), 68.4 (C6) ppm.

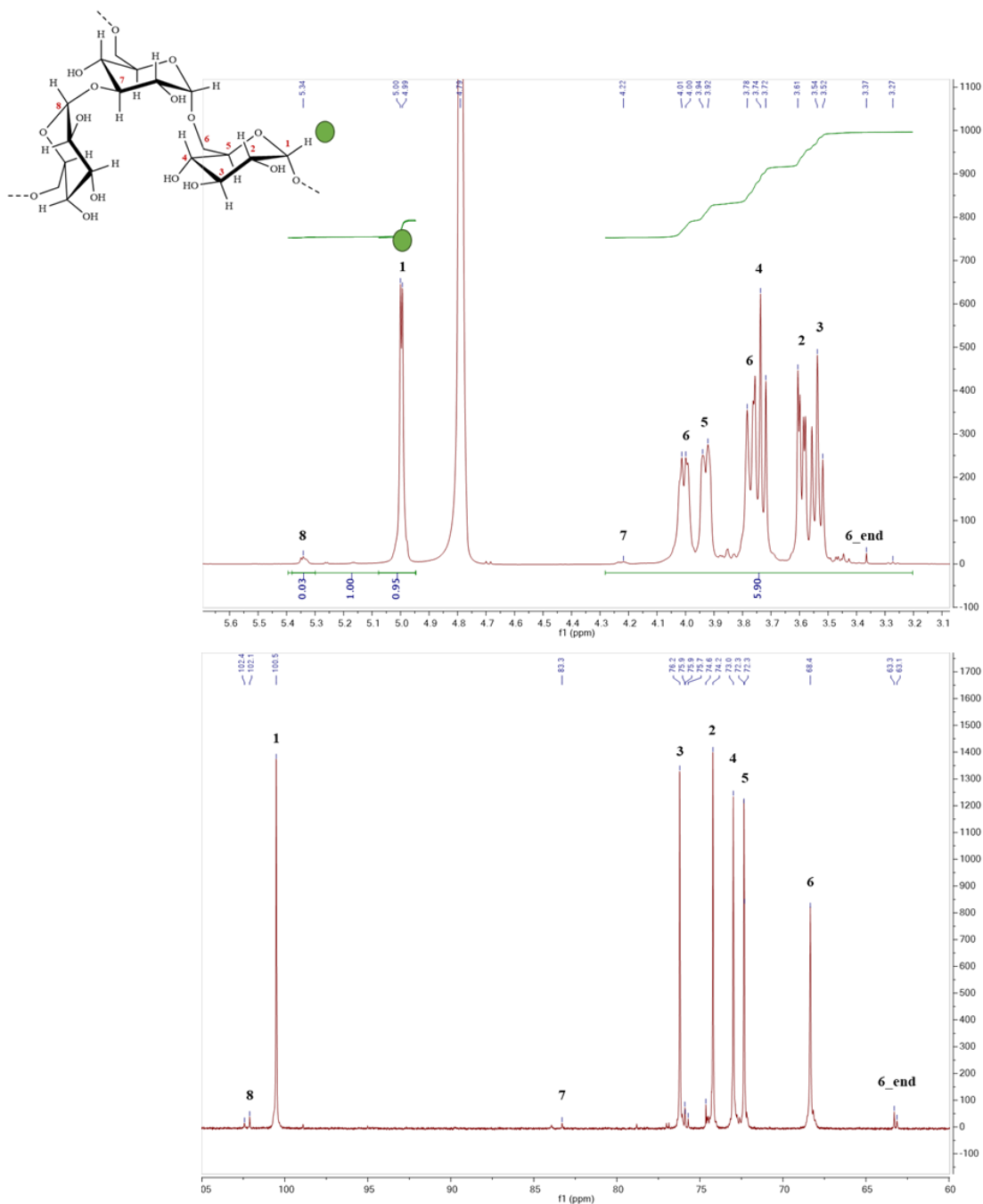


Figure 52 – Recorded NMR spectra ( $^1\text{H}$  up and  $^{13}\text{C}$  down) of Dex40 with relevant found signals.

### 1.1.1.2 NMR of Dex40-GP

NMR spectra of Dex40-GP show the success of alkyne functionalization of dextran molecules. Proton spectrum was used to characterize the novel construct which was found to carry a minor component of oligomerized and alkyne-alkyne coupled GP groups.

Considering the area of  $\text{H}^1$  signals as 100 (total glucose units):

- Percentage of ether-modified glucose units/dextran molecule:  $DS_{GP} = 30.9/100 \approx 31\%$
- Percentage of terminal alkynes/dextran molecule:  $DS_{alk} = 28.7/100 \approx 29\%$
- Average number of GP groups/glucose-GP unit:  $n_{GP} = (109.6/2)/30.9 \approx 1.8$

These calculations can be used to estimate the expected area of 4.15 – 3.38 ppm range:  $6 \cdot (69.1 + 31.9) + 5 \cdot (30.9 \cdot 1.77) = 873.5$ . This value is in accordance with the observed one.

$^1H$  NMR (500 MHz, Deuterium Oxide)  $\delta$  5.52 – 4.95 ( $H^1$  signals), 4.29 ( $\equiv C-CH_2-O-$  signals), 4.15 – 3.38 (remaining signals), 2.98 ( $H-C\equiv C-$  signals).

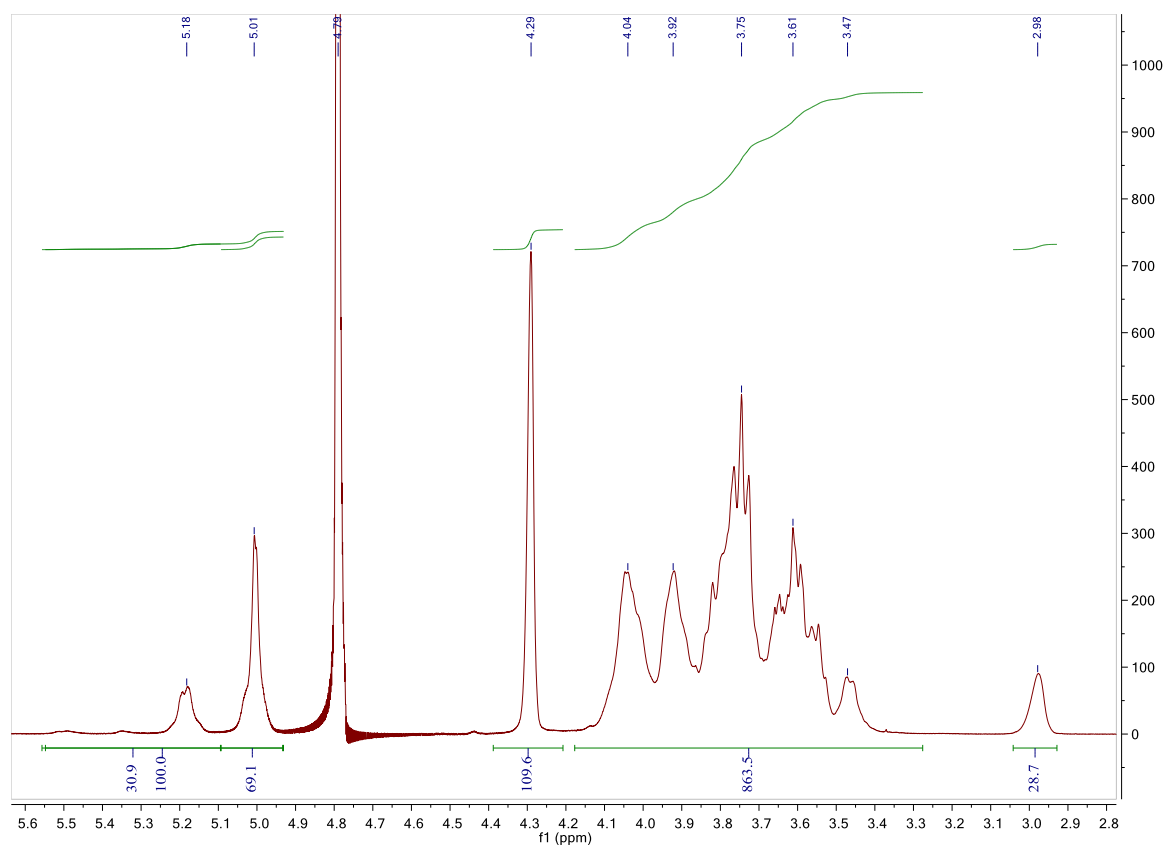


Figure 53 –  $^1H$ -NMR spectrum of Dex40-GP with relevant areas used for the characterization, as described in the results section

$^{13}C$  NMR (126 MHz,  $d_2O$ )  $\delta$  100.51, 98.58, 82.42, 79.02, 76.19, 75.30, 74.59, 74.53, 74.21, 73.57, 73.52, 72.99, 72.32, 72.08, 71.85, 71.76, 71.66, 68.35, 61.09 ppm.

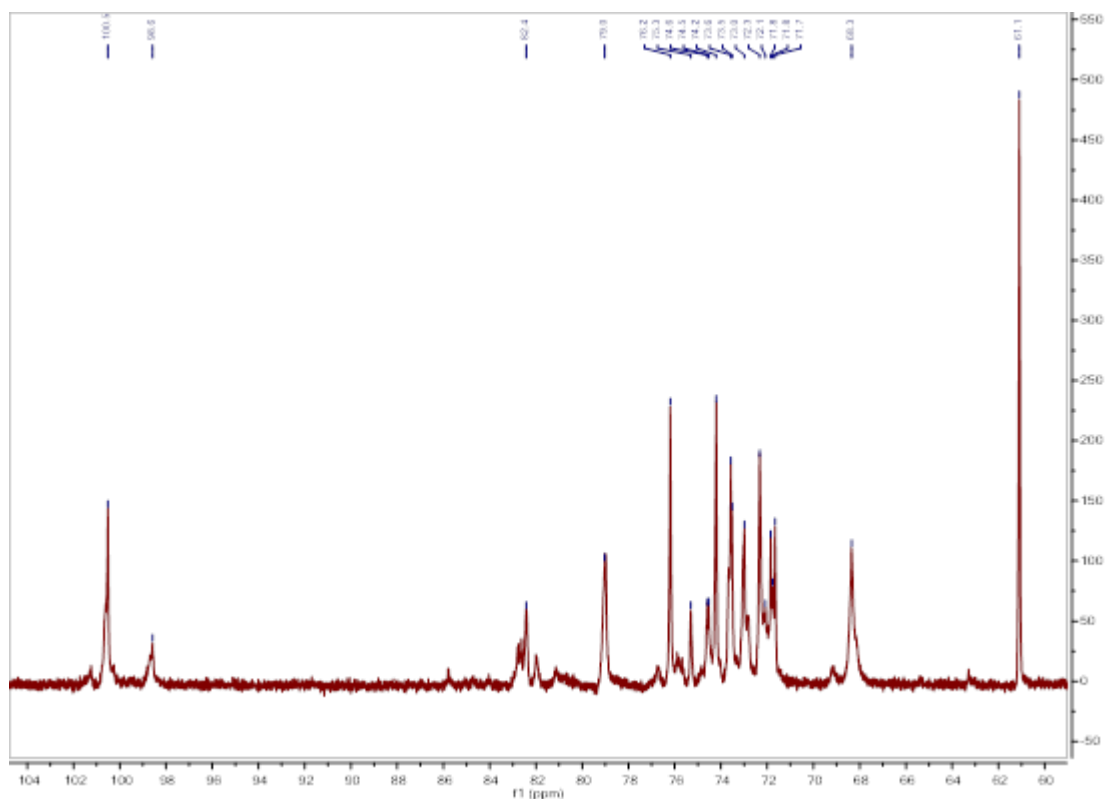


Figure 54 –  $^{13}\text{C}$ -NMR spectrum of Dex40-GP. As discussed in the results section, hypothetical assignment of several minor peaks is in agreement with the presence of GP oligomers and alkyne-alkyne couplings.

### 1.1.1.3 NMR of peptide **6**

$^1\text{H}$  and  $^{13}\text{C}$  spectra of peptide **6**, whose structure is shown in Figure 55, were acquired to allow the spectra interpretation of Dex40-Pept**6**.

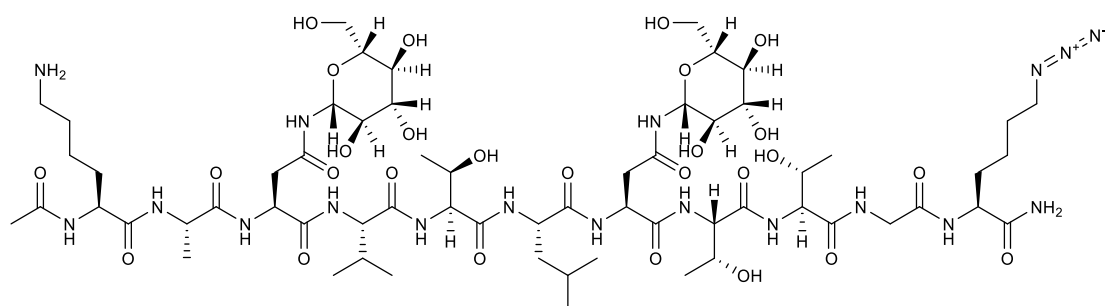


Figure 55 – Structure of peptide **6**: Ac – Lys – Ala – Asn (Glc) – Val – Thr – Leu – Asn (Glc) – Thr – Thr – Gly – Lys( $\text{N}_3$ ) –  $\text{NH}_2$

The peptide has 80 non-exchangeable protons, but the sum of the integrated signals in  $^1\text{H}$  spectrum is 78 because the  $\text{H}^a$  signal from the two Asn(Glc) falls in the solvent range ( $\approx 4.70$  ppm) (Figure 56).

$^1\text{H}$  NMR (500 MHz,  $\text{D}_2\text{O}$ )  $\delta$  4.97 (d,  $J = 9.2$  Hz, 2H,  $\text{H}^1$  Glc), 4.49 – 4.14 (superimposed signals, 11H,  $\text{H}^a$  area +  $\text{H}^b$  Thr), 4.00 (s, 2H,  $\text{H}^a$  Gly), 3.87 (superimposed dd, 2H,  $\text{H}^6$  Glc), 3.74 (dd,  $J = 12.4, 5.1$  Hz, 2H,  $\text{H}^{6'}$  Glc), 3.56 (t,  $J = 9.1$  Hz, 2H,  $\text{H}^3$  Glc), 3.50 (m, 2H,  $\text{H}^5$  Glc), 3.44 (t,  $J = 9.4$  Hz, 2H,  $\text{H}^4$  Glc), 3.40 (t,  $J = 9.2$  Hz, 2H,  $\text{H}^2$  Glc), 3.35 (t,  $J = 6.8$  Hz, 2H,  $\text{H}^e$  Lys( $\text{N}_3$ )), 3.05 – 2.80 (superimposed signals, 6H, t  $\text{H}^e$  Lys + dd  $\text{H}^b$  Asn), 2.14 (h,  $J = 6.8$  Hz, 1H,  $\text{H}^b$  Val), 2.05 (s, 3H,  $-\text{CH}_3$  Ac), 1.94 – 1.57 (m, 11H,  $\text{H}^{\beta 1,2}$   $\text{H}^\delta$  Lys +  $\text{H}^{\beta 1,2}$   $\text{H}^\delta$  Lys( $\text{N}_3$ ) +  $\text{H}^{\beta 1,2}$   $\text{H}^\gamma$  Leu), 1.54 – 1.41 (m, 4H,  $\text{H}^\gamma$  Lys +  $\text{H}^\gamma$  Lys( $\text{N}_3$ )), 1.40 (d,  $J = 7.2$  Hz, 3H,  $\text{H}^b$  Ala), 1.27-1.22 (superimposed d,  $J \approx 6.4$  Hz, 9H,  $\text{H}^2$  Thr), 0.97-0.93 (superimposed d, 9H,  $\text{H}^{\delta 2}$  Leu +  $\text{H}^\gamma$  Val), 0.89 (d,  $J = 5.6$  Hz, 3H,  $\text{H}^{\delta 1}$  Leu).

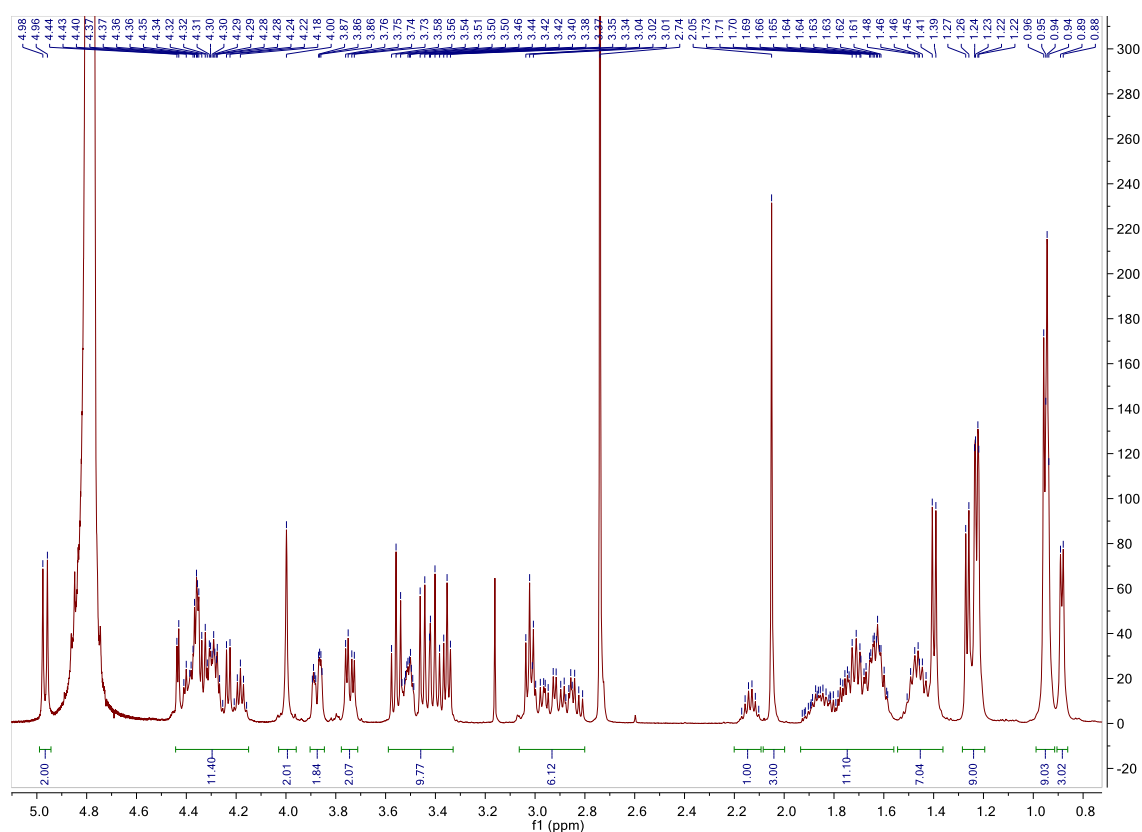


Figure 56 –  $^1\text{H}$ -NMR spectrum of Peptide 6

Peptide 6 possesses 62 Carbons, each one originating a distinct signal in  $^{13}\text{C}$ -NMR spectrum, except  $\text{C}^1$  and  $\text{C}^2$  of the two glucoses (superimposed signals) (Figure 57).

$^{13}\text{C}$  NMR (126 MHz,  $\text{d}_2\text{o}$ )  $\delta$  179.5 – 174.0 (14 C,  $\text{C}^{\text{amides}}$ ), 80.0 (2 C,  $\text{C}^1$  glucoses), 78.3 – 69.9 (8 C,  $\text{C}^{2,3,4,5}$  glucoses), 67.6 (3 C,  $\text{C}^\beta$  Thr), 61.3 – 59.9 (6 C,  $\text{C}^6$  glucoses +  $\text{C}^\alpha$

Thr + C<sup>α</sup> Val), 56.3 – 52.4 (7 C, C<sup>α</sup> area + C<sup>ε</sup> Lys(N<sub>3</sub>)), 45.3 – 39.1 (5 C, C<sup>α</sup> Gly + C<sup>ε</sup> Lys + C<sup>β</sup> Lys + C<sup>β</sup> Asn), 33.3 – 19.4 (17 C, remaining signals) ppm.

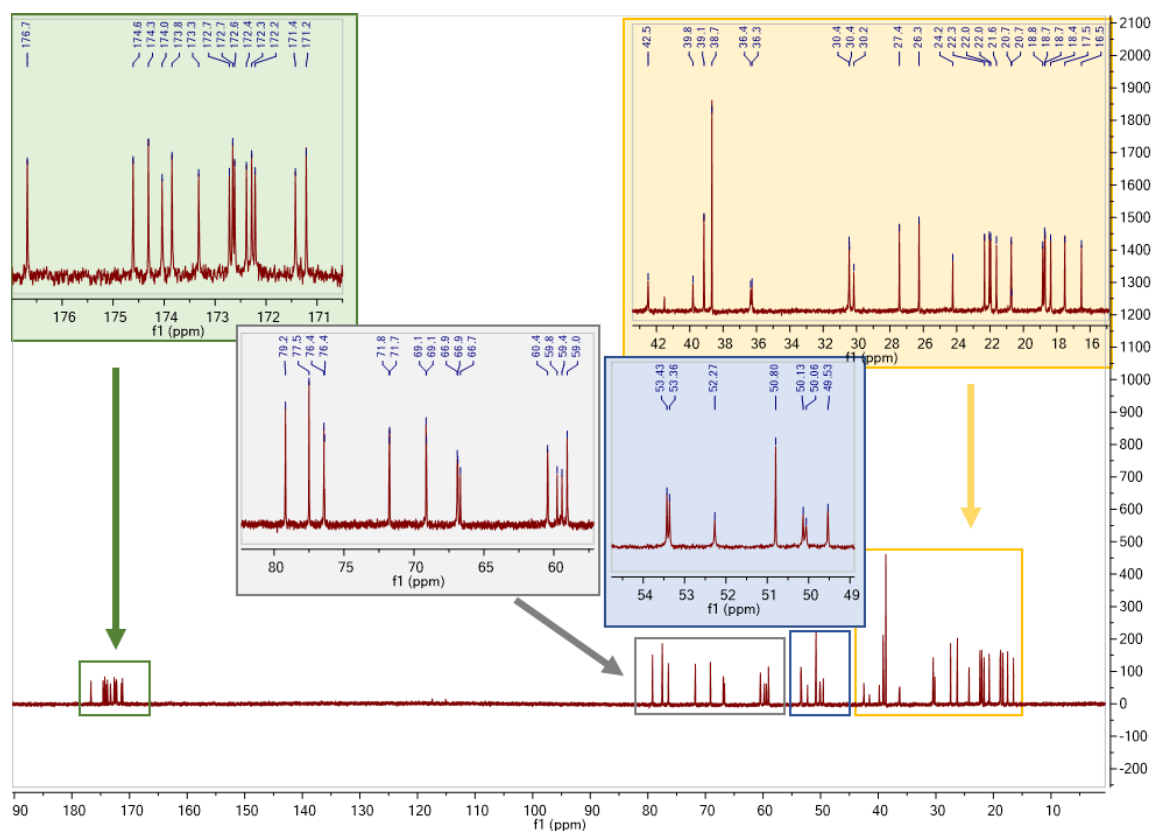


Figure 57 – <sup>13</sup>C-NMR spectrum of Peptide 6.

#### 1.1.1.4 NMR of Dex40-Pept6

Peaks assignment of <sup>1</sup>H-NMR spectrum for Dex40-Pept6 was based on previous spectra (Figure 58).

<sup>1</sup>H NMR (500 MHz, Deuterium Oxide) δ 8.10 (s, 1H, H triazole), 5.31 – 4.95 (m, H<sup>1</sup> area), 4.54 - 4.19 (m), 4.10 – 3.40 (m), 3.12 – 2.84 (m), 2.20 – 2.16 (m, 1H, H<sup>β</sup> Val), 2.09 (s, 3H, -CH<sub>3</sub> Ac), 2.04 – 1.35 (m, 18H, H<sup>β1,2</sup> H<sup>δ</sup> H<sup>γ</sup> Lys + H<sup>β1,2</sup> H<sup>δ</sup> H<sup>γ</sup> Lys(N<sub>3</sub>) + H<sup>β1,2</sup> H<sup>γ</sup> Leu + H<sup>β</sup> Ala), 1.33-1.23 (m, 9H, H<sup>γ2</sup> Thr), 1.01-0.88 (m, 12H, H<sup>δ1,2</sup> Leu + H<sup>γ</sup> Val).

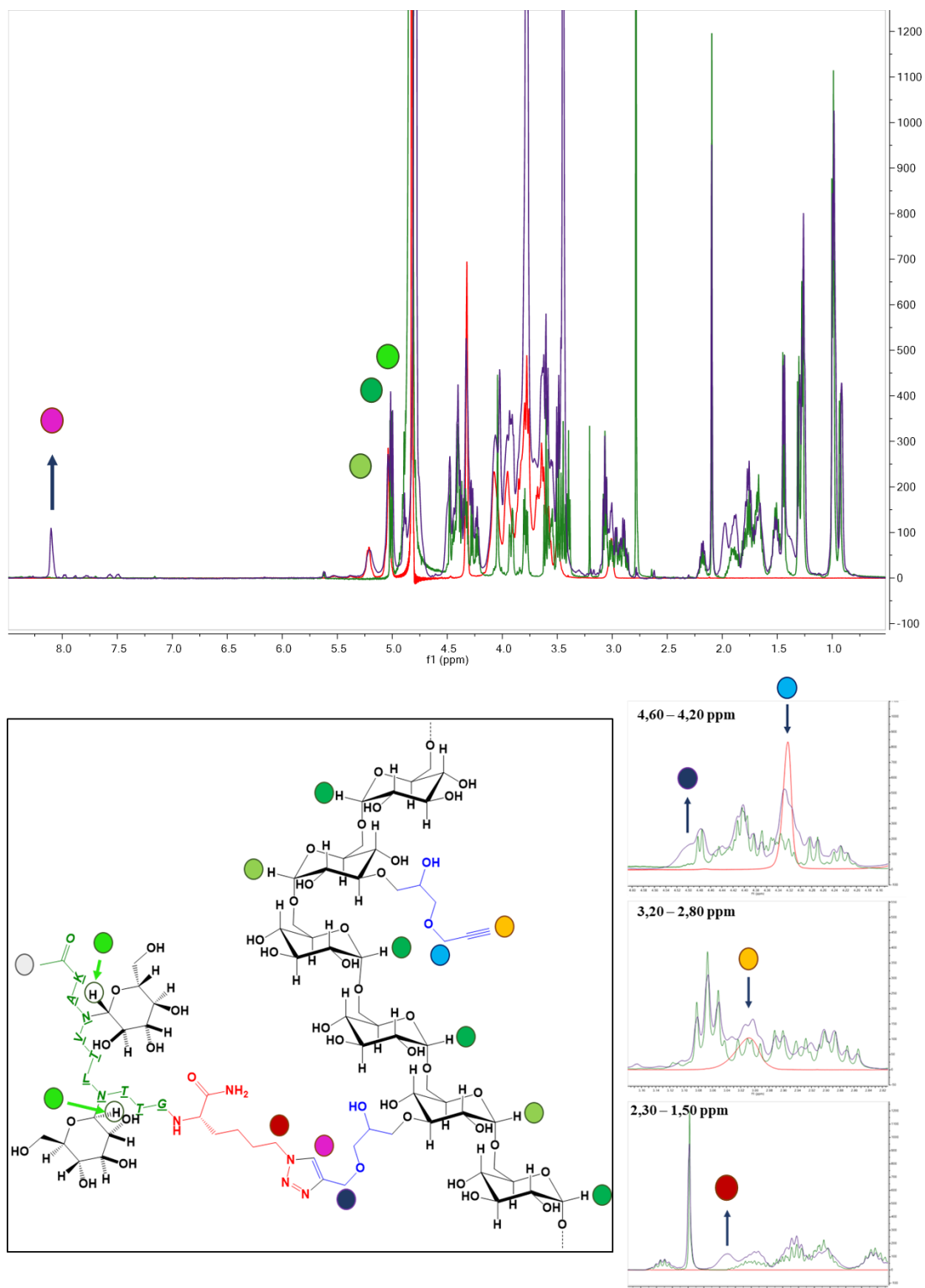


Figure 58 – Superimposition of  $^1\text{H-NMR}$  spectra of Dex40-GP (red), Peptide 6 (green) and Dex40-Peptide 6 (violet). The zoom of relevant ppm ranges shows the diagnostic changes in the spectra after the CuAAC reaction (down).

Relevant integral ratios were used to calculate the DS in peptide and in unreacted alkynes, as discussed in the results section, chapter 5.2.2.

Specifically, ratio between the area of triazole proton (8.10 ppm) and the area of other ranges (Figure 59) were used to calculate X ( $DS_{\text{pept6}}$ ):

$$\text{A) } 0.62/1.00 = X/0.31 \quad \rightarrow X = 19.22 \%$$

$$\text{B) } 0.14/1.00 = X/(1+2X) \quad \rightarrow X = 19.44 \%$$

$$\text{C) } 0.06/1.00 = X/(2*0.31*1.77 + 11X) \quad \rightarrow X = 19.42 \%$$

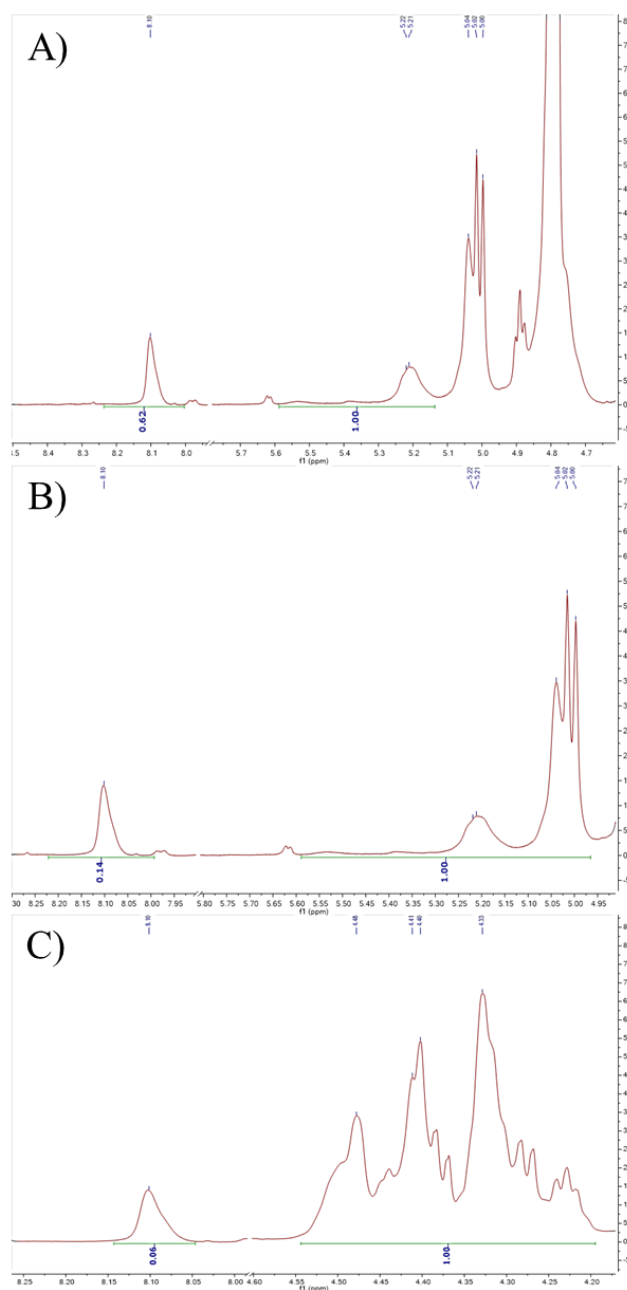


Figure 59 – Zoom of relevant ppm ranges in <sup>1</sup>H-NMR spectrum of Dex40-Pept6 used for peptide DS calculation, with corresponding integration values. Triazole proton signal (8.10 ppm) was used as reference in comparison to: A) H1 of modified glucose units (5.60-5.15 ppm); B) H1 of total glucoses, i.e., glucose units of dextran and N-(βGlc) moieties of peptide 6 (5.60-4.95 ppm); C) the range containing 8 H<sub>α</sub> + 3H<sub>β</sub> of peptide residues and methylene protons of GP linkers (4.55-4.20 ppm).

By reasonably varying spectrum parameters such as baseline, phase and integration limits, obtained values for  $DS_{\text{peptide6}}$  are always between 19% and 20%, therefore 19.5% was used as practical value for further calculations. Consequently, 9.5% is accepted as a good estimation of remaining terminal alkynes.

$^{13}\text{C}$  NMR (126 MHz,  $\text{d}_2\text{o}$ )  $\delta$  179.2 – 173.9 (14 C,  $\text{C}^{\text{amides}}$ ), 146.7 (1C,  $\text{C}^{\text{triaz}}$ ), 127.7 (1C,  $\text{C}^{\text{triaz}}$ ), 100.5 – 61.1 (glucose signals, glycidyl signals, propargyl signals,  $\text{C}^\alpha$   $\text{C}^\beta$  Thr,  $\text{C}^\alpha$  Val), 56.3 -19.4 (peptide signals).

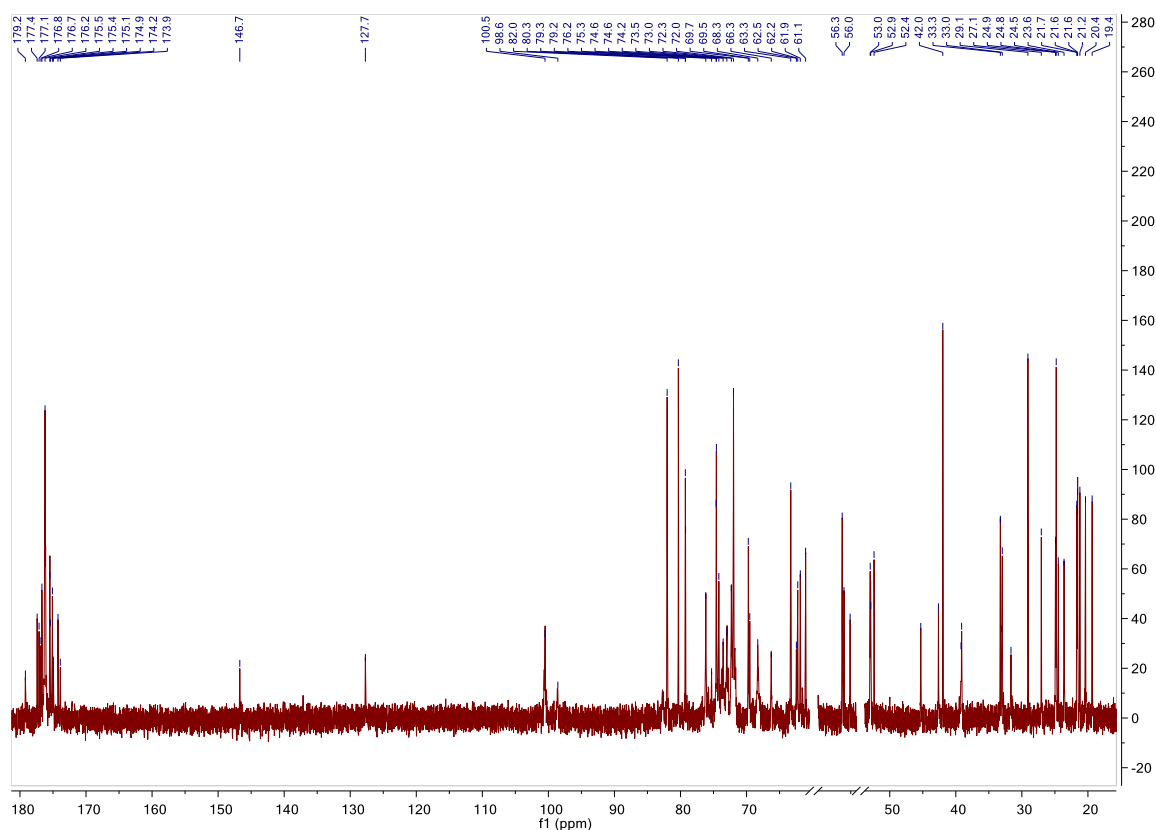


Figure 60 –  $^{13}\text{C}$ -NMR spectrum of Dex40-Pept6. The two isolated peaks at 146.7 and 127.7 originate from the triazole ring and are diagnostic for the success of the reaction.

## 7.5 PROTEIN EXPRESSION

### 7.5.1 General procedure for protein expression

Protein fragment HMW1ct and the enzyme HMW1C were expressed similarly to a described protocol<sup>[97]</sup>, using *E. coli* BL21 cells previously engineered with plasmid pET-45b (+) (Novagen), encoding for the fragment HMW1ct and equipped with the

gene for carbenicillin resistance, and plasmid pET-24a (+) (Novagen), encoding for the glucosyltransferase enzyme ApHMW1C and equipped with the gene for kanamycin resistance. Cells cultures were prepared using Luria-Bertani (LB) culture soils; the LB medium (SOC) liquid soil was composed by 10 g of tryptone, 5 g of yeast extract and 10 g of NaCl dissolved in 1 L of H<sub>2</sub>O Milli-Q, while the LB Agar Plates soil was composed by 2 g of tryptone, 1 g of yeast extract, 2 g of NaCl and 3 g of Agar dissolved in 200 ml of H<sub>2</sub>O Milli-Q. Stock solutions of antibiotics were prepared in H<sub>2</sub>O Milli-Q and stored at -20°C. Working concentration in cell media is 50 µg/mL for kanamycin (only for hyperglucosylated HMW1ct(Glc) and 100 µg/mL for carbenicillin. Lysis buffer (pH 7.5) was composed by 5.96 g of HEPES (50 mM), 2.92 g of NaCl (100 mM) and 50 ml of glycerol (10%) dissolved in 0.5 L of H<sub>2</sub>O Milli-Q. Cell were coated on Petri dishes (Nunc, ThermoFisher Scientific) with LB Agar Plates soil, containing the antibiotic(s), and incubated overnight at 37°C to allow the growth of the bacterial colonies. The pre-culture phase was performed by picking up one single circular and isolated colony and transferring it in 5 ml of LB medium (SOC) liquid soil containing the antibiotic(s). The solution was incubated overnight at 37°C under shaking. The pre-culture solution was then transferred in 1 L of the same LB medium (SOC) liquid soil containing the antibiotic(s). The solution was incubated under shaking at 37 °C. Cells growth was monitored measuring the optical density at 600 nm (OD<sub>600</sub>) with an UV instrument (Amersham Biosciences, Little Chalfont, UK). The same LB medium (SOC) liquid soil was used as blank. When the OD value reached 0.6, the induction of the expression was performed adding 1 mL of isopropyl-β-D-1-thiogalactopyranoside (IPTG) (1 mg/mL solution). Cell suspension was incubated overnight at 16°C under shaking. Cells were recovered through centrifugation at 4000 rpm for 30 min at 4°C. The supernatant was removed, and the pellet was suspended in 20 mL of lysis buffer, recentrifuged again and stored at -20°C after supernatant removal.

### **7.5.2 General procedure for protein purification**

The pellet was suspended in 30 ml of lysis buffer adding 10 µL/g of cells of protease inhibitor (cocktail Set III EDTA-free, Merk). Mechanical lysis of the cell membrane was obtained by using an ultrasonic processor. The lysis solution was then centrifuged for 110 min at 35000 rpm and the supernatant containing the product(s) was recovered. The purification was performed using an Äkta FPLC system. During the first purification step a Hi Trap-His column (HisTrap HP 5 mL, GE Healthcare) was used

with the binding buffer A1 for Hi Trap-His (20 mM Tris buffer, 0.5 M NaCl, 30 mM imidazole, pH 7.4) and the elution buffer B1 for Hi Trap-His (20 mM Tris buffer, 0.5 M NaCl, 0.5 M imidazole, pH 7.4).

The conditioning of the column was performed using buffer A1 for 10 minutes. The supernatant containing the products was then injected and eluted with a gradient from 0% to 100% of buffer B1. The UV detector was set to 280 nm and 215 nm. All the fraction obtained were analyzed through Sodium Dodecyl Sulphate - PolyAcrylamide Gel Electrophoresis technique (SDS-PAGE).

The separation of HMW1ct(Glc) from ApHMW1C was obtained in the second purification step through the ion exchange technique. A Hi Trap Q-FF column was used. A buffer exchange in order to substitute buffer B1 with binding buffer A2 (20 mM Tris buffer, 20 mM NaCl, pH 8) for Hi Trap Q-FF was performed using Amicon Ultra Centrifugal Filters (MWCO = 10 kDa). The Hi Trap Q-FF column was then conditioned with buffer A2 for 10 minutes. The sample was injected and eluted using a gradient from 0% to 100% of elution buffer B2 (20 mM Tris buffer, 1 M NaCl, pH 8) for Hi Trap Q-FF. The UV detector was set to 280 nm and 215 nm. All the fraction obtained were analyzed through Sodium Dodecyl Sulphate - PolyAcrylamide Gel Electrophoresis technique (SDS-PAGE).

Both HMW1ct and HMW1ct(Glc) were stocked in PBS buffer (8 g of NaCl, 0.2 g of KCl, 1.44 g of Na<sub>2</sub>HPO<sub>4</sub> e 0.24 g of KH<sub>2</sub>PO<sub>4</sub> dissolved in 1 L of H<sub>2</sub>O Milli-Q) at -20°C. Their concentration was calculated using the Lambert-Beer law after an absorption measure performed using an UV spectrometer set to a range from 320 and 240 nm.

### **7.5.3 General procedure for SDS-PAGE**

The SDS-PAGE gel was prepared by depositing between two glasses the running gel 16% solution, composed by 1.6 ml H<sub>2</sub>O Milli-Q, 4.27 ml 30% acrylamide, 2 ml 1.5M tris buffer pH 8.8, 80 µl 10% SDS, 80 µl 10% ammonium persulfate (APS), 10 µl tetramethylethylenediamine (TEMED). After the polymerization, the stacking gel 4% solution (1.8 ml H<sub>2</sub>O Milli-Q, 0.4 ml 30% acrylamide, 0.750 ml 0.5M tris buffer pH 6.8, 30 µl 10% SDS, 30 µl 10% APS, 6 µl TEMED) was deposited above the previous one inserting the comb for the formation of the wells. After the polymerization, the gel was positioned inside the SDS-PAGE apparatus and the tank buffer 1x (100 mL of Tris buffer/Glycine/SDS (10x) in 1 L) was added. 10 µL of each sample were combined

with 5  $\mu$ L of loading buffer 5x (200mM of Tris-Cl (pH 6,8), 400mM of DTT, 8% of SDS, 0,4% of bromophenol blue and 40% of glycerol), treated to 100°C for a few minutes and centrifuged. Each sample was then loaded in the dedicate well. The commercial marker PageRuler Plus Prestained Protein Ladder, 10 to 250 kDa, was used as reference. The electrophoresis was performed for 90 min at 140 mV and subsequently the gel subjected to stain using a Comassie solution (25 ml H<sub>2</sub>O Milli-Q, 20 ml MeOH, 5 ml AcOH, 0.05 g Comassie blue dye) for 30 min. In order to remove the excess dye, the stained gel was treated overnight with a destaining solution (700 ml H<sub>2</sub>O Milli-Q, 200 ml MeOH, 100 ml AcOH) under gentle shaking.

## 7.6 INDIRECT COMPETITIVE ELISA

Coating antigen was dissolved in H<sub>2</sub>O (1 mg/mL) and then diluted 1:100 in carbonate buffer 0.05 M (pH 9.6). 100  $\mu$ L of coating antigen solution were added to each well (1  $\mu$ g antigen/well) of a 96-Well activated Polystyrene ELISA plate (Limbro Titertek, ICN Biomedicals, Inc., Aurora, Ohio, USA), and left to incubate overnight at 4°C. The plate was then emptied, washed three times with 0,9% w/w saline solution containing 0.05% v/v Tween 20, then emptied again. 100  $\mu$ L of Fetal Bovine Serum solution (FBS 10% in saline Tween) /well were added and the plate was left to incubate 2 h at room temperature. Then plate was emptied and a mixture of sera and competing antigen was added in each well. Serum concentration was constant (dilution 1:300) while competing antigen concentration was between 10<sup>-14</sup> and 10<sup>-5</sup>.

The plate was incubated 1h at room temperature, then emptied, washed three times, and emptied again.

100  $\mu$ L of secondary antibody solution (alkaline phosphatase conjugated anti human IgM or IgG Fab2-specific affinity purified antibodies) /well were added and incubated 3 h at room temperature. Each plate was emptied, washed three times then emptied.

100  $\mu$ L/well of substrate solution (*p*-nitrophenylphosphate 0.1% w/v in carbonate buffer and MgCl<sub>2</sub> 10 mM) were added and absorbance was measured at regular intervals with a plate reader (Tecan-Sunrise spectrophotometer working at 405 nm). After 30-60 min, the reaction was blocked by adding 50  $\mu$ L NaOH 1M/well and the final absorbance value was measured.

Each competing antigen was tested in triplicates (three rows/plate), and each experiment was performed at least twice in different days. Within-assays and between-

assays coefficient of variations were below 10%. Peptides concentration-absorbance relationship was represented graphically as signal inhibition percentage, and half-maximal response concentration values (IC<sub>50</sub>) were calculated with GraphPad Prism.

## 7.7 INDIRECT SP-ELISA

To find the best conditions in order to perform indirect SP-ELISA screening of a larger batch of sera, preliminary tests were carried out for peptides **4** and **7**, and Dex40-Pept**6**.

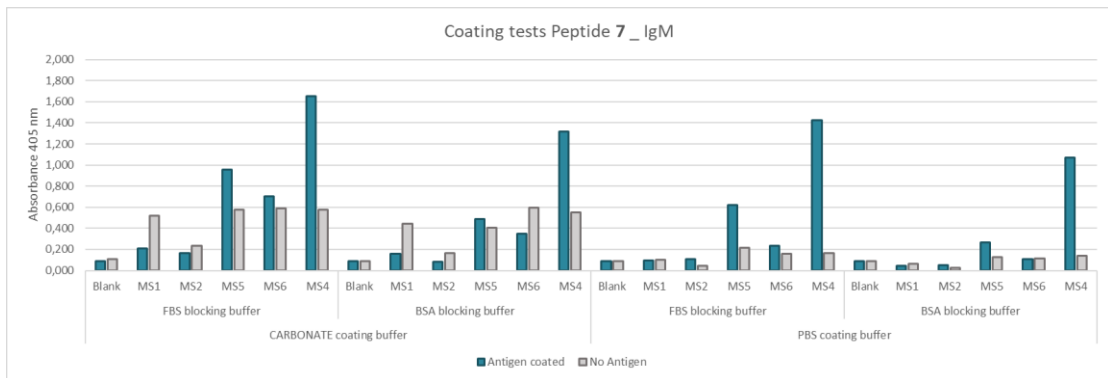
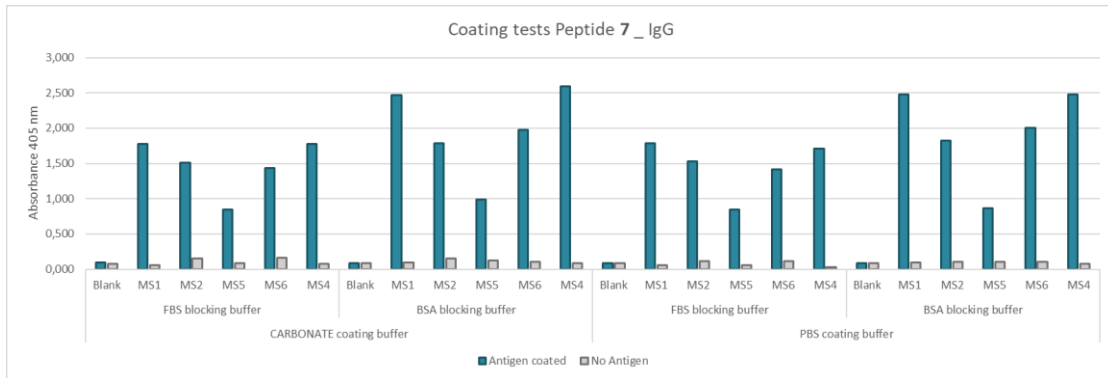
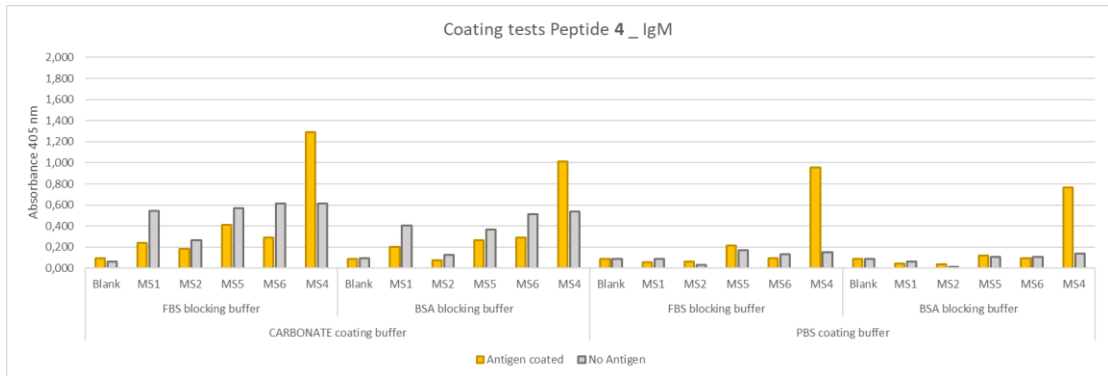
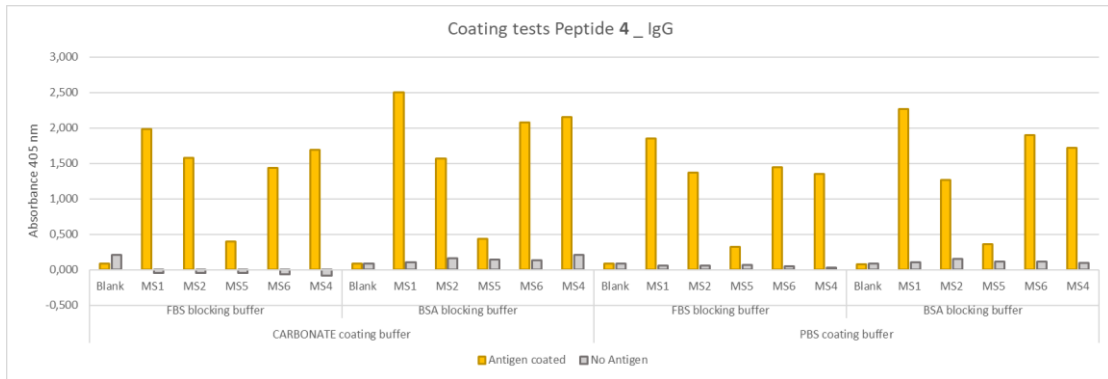
The following parameters were evaluated:

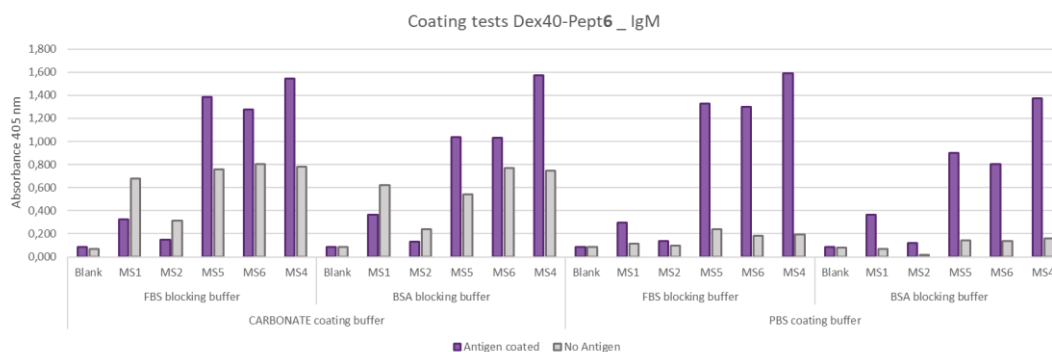
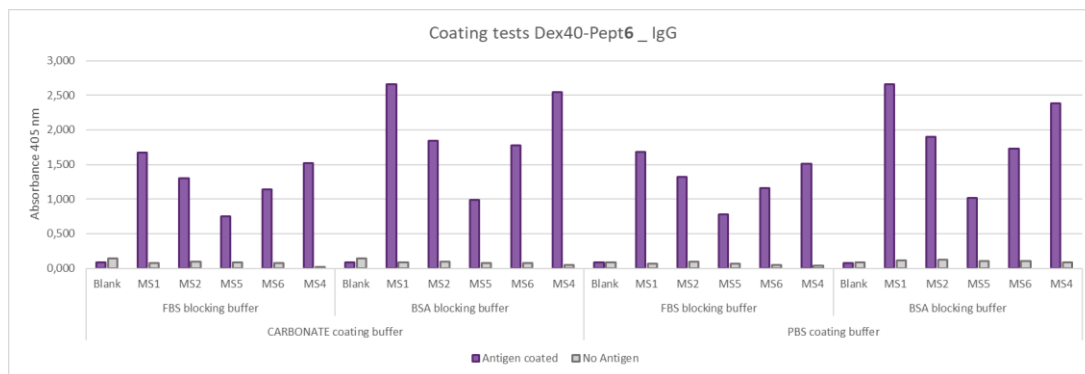
- Presence or absence of the coated antigen
- Coating buffer (Na<sub>2</sub>CO<sub>3</sub> 0.05 M buffer pH 9.6 or PBS buffer pH 7.2)
- Blocking buffer (10% fetal bovine serum (FBS) or 5 % bovine serum albumin (BSA))
- IgG and IgM detection
- 5 different representative sera + no serum (blank)

ELISA plates, sera dilutions and incubation times were not tested, and used according to previously optimized procedures employing glucosylated antigens for the detection of antibodies in MS sera<sup>[83,97]</sup>.

Briefly, 96-Well activated Polystyrene ELISA plates (Limbro Titertek, ICN Biomedicals, Inc., Aurora, Ohio, USA) were coated with 1 µg/100 µL/well of antigen in coating buffer (or pure coating buffer without antigen) and incubated at 4 °C overnight. After 3 washes with 0,9% w/w saline solution containing 0.05% v/v Tween 20, non-specific binding sites were blocked with blocking buffer at r.t. for 60 minutes. Sera diluted 1:100 in blocking buffer (100 µL/well) were applied at 4 °C for 16 h. After 3 washes, 100 µL/well of secondary antibody solution (alkaline phosphatase conjugated anti human IgM or IgG Fab2-specific affinity purified antibodies diluted in blocking buffer) were added. After 3 h incubation at room temperature, plates were washed 3 times and then 100 µL/well of substrate solution (*p*-nitrophenylphosphate 0.1% w/v in carbonate buffer and MgCl<sub>2</sub> 10 mM) were added. After 15 minutes (IgG plates) or 40 minutes (IgM plates), the reaction was blocked with 50 µL of 1 M NaOH and the absorbance read in a plate reader (SUNRISE, TECAN, Austria) at 405 nm.

Median absorbance values at 405 nm are reported and differentiated according to the different parameters.





## 7.8 IMMUNOAFFINITY COLUMN

Preliminary attempts of antibody purification were performed similarly to a previously described method<sup>[97]</sup> using CNBr-Sepharose resin (Sigma). The resin (100 mg) was washed twice with 1 mL HCl 1 mM and centrifuged at 4000 rpm for 3 minutes. The washing step was repeated three times with H<sub>2</sub>O Milli-Q and once with coupling buffer (NaHCO<sub>3</sub> 0.1 M, NaCl 0.5 M, pH = 8.3). Resin was then transferred in a fritted column and washed one last time with coupling buffer (gravity flow). Dex40-Pept6 (1 mg) was dissolved in 1 mL of coupling buffer and the solution was applied to the resin overnight at room temperature and vigorous shaking. Then the resin was washed with coupling buffer twice, and 1 mL glycine solution (0.2 M, pH = 8.0) was applied to the resin for 2 h at room temperature. The resin was then washed twice with 1 mL coupling buffer, twice with 1 mL acetate buffer (sodium acetate 0.1 M, NaCl 0.5 M, pH = 4.3) and equilibrated with Dulbecco's phosphate buffer saline (D-PBS) at pH = 7.2. Serum MS5 (1 mL) was diluted 1:10 in D-PBS, passed through a 0.22 μm filter and applied to the sepharose column. 10 mL eluted fraction by gravity flow was recirculated one more time through the column and 50 μL from the eluted fraction were taken for ELISA test (FT1 fraction in the results section). The eluted fraction was then applied a third time and left to incubate for 1 h. Final flow through fraction was collected (FT2 fraction in

the results section) and then column was washed thoroughly with D-PBS (10 mL) and coupling buffer (10 mL). Adsorbed antibodies were eluted using 10 mL of glycine 0.2 M at pH=2.6. Eluted fraction was immediately neutralized drop-by-drop by adding 10% of NaHCO<sub>3</sub> 0.5 M, constantly monitoring the pH with pH paper. The final volume of neutralized eluted fraction is  $\approx$  15 mL, which is then concentrated by centrifuging with Amicon ultracentrifugal filter units (Merck, MWCO = 50 kDa) and recovered in D-PBS pH = 7.2 ( $\approx$ 1 mL final volume, A<sub>280</sub> = 0.909). 50  $\mu$ L of eluted fraction were diluted up to 1 mL in FBS and their activity was checked by SP-ELISA (Elution fraction in the results section).

After use, the column was thoroughly washed with additional glycine 0.2 M at pH=2.6 (5 mL), D-PBS buffer (10 mL), coupling buffer (10 mL) and finally with EtOH 20% solution in water, then stored at 4°C to be reused.

## **8. ABBREVIATIONS**

### Amino acids:

A	(Ala)	Alanine
C	(Cys)	Cysteine
D	(Asp)	Aspartic acid
E	(Glu)	Glutamic acid
F	(Phe)	Phenylalanine
G	(Gly)	Glycine
H	(His)	Histidine
I	(Ile)	Isoleucine
K	(Lys)	Lysine
L	(Leu)	Leucine
M	(Met)	Methionine
N	(Asn)	Asparagine
P	(Pro)	Proline
Q	(Gln)	Glutamine
R	(Arg)	Arginine
S	(Ser)	Serine
T	(Thr)	Threonine
V	(Val)	Valine
Pra		Propargylglycine

Ab: antibody; Ac: Acetyl; AcOH: Acetic Acid; Ag: Antigen; AP: alkaline phosphatase; APC: Antigen presenting cell; Boc: tert-butoxy carbonyl; BSA: bovine serum albumin; CD: Circular Dichroism; CNS: Central Nervous System; DCM: dichloromethane; DIC: N,N'-diisopropylcarbodiimide; DIEA/DIPEA: N,N-Diisopropylethylamine; DMF Dimethylformamide; DMSO Dimethyl sulfoxide; DS: degree of substitution; EDC: 3-(Ethyliminomethyleneamino)-N,N-dimethylpropan-1-amine; EDT: ethanedithiol; EDTA: ethylenediaminetetraacetic acid; ELISA: enzyme-linked immunosorbent assay; ESI-MS: electrospray ionization-mass spectrometry; Fab: fragment antigen-binding; Fc : fragment crystallizable; FBS: fetal bovine serum; FDA: Food and Drugs Administration agency; Fmoc: 9-H-fluoren-9-ylmethoxycarbonyl; Gal: galactose; Glc: glucose;

GlcA: glucuronic acid; HATU: 1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxide hexafluorophosphate; HBTU: 3-[Bis(dimethylamino)methylumyl]-3H-benzotriazol-1-oxide hexafluorophosphate; HLA: human leukocyte antigen; HMW1: high molecular weight adhesin 1; HOBt: 1-hydroxybenzotriazole; HPLC: high performance liquid chromatography; Ig: immunoglobulin (IgG, IgM, etc.); MALDI-TOF: matrix-assisted laser desorption ionization- time of flight; MAP: multiple antigenic peptide; MeCN (or CH<sub>3</sub>CN or ACN): acetonitrile; MeOH: methanol; MHC: major histocompatibility complex; MOG: myelin oligodendrocyte glycoprotein; MS: Multiple Sclerosis; MW: microwaves; MWCO: molecular weight cut-off; NMM: N-methyl morpholine; NMR: nuclear magnetic resonance; OGT: oligoglycosyl transferase; Pbf: 2,2,4,6,7-Pentamethyldihydrobenzofuran-5-sulfonyl; PBS: phosphate buffered saline; PDB: protein data bank; PEG: polyethylene glycol; PNPP: para-nitrophenylphosphate; PS: polystyrene; RP-HPLC: Reverse-Phase High-performance Liquid Chromatography; SDS-PAGE: Sodium Dodecyl Sulphate - PolyAcrylamide Gel Electrophoresis; SP-ELISA: solid-phase enzyme-linked immunosorbent assay; SPPS: solid phase peptide synthesis; *t*Bu: tert-butyl; TBDMS (or TBS): tert-Butyl dimethyl silyl; TFA: trifluoroacetic acid; TIS: triisopropylsilane; THF: tetrahydrofurane; TLC: Thin layer chromatography; Trt: trityl.

## 9. REFERENCES

- [1] L. B. Nicholson, *Essays Biochem* **2016**, *60*, 275–301.
- [2] D. D. Chaplin, *J Allergy Clin Immunol* **2010**, *125*, S3-23.
- [3] L. Wang, F.-S. Wang, M. E. Gershwin, *J. Intern. Med.* **2015**, *278*, 369–395.
- [4] M. Mahler, M. J. Fritzler, *Annals of the New York Academy of Sciences* **2010**, *1183*, 267–287.
- [5] E. Maverakis, K. Kim, M. Shimoda, M. E. Gershwin, F. Patel, R. Wilken, S. Raychaudhuri, L. R. Ruhaak, C. B. Lebrilla, *J Autoimmun* **2015**, *0*, 1–13.
- [6] Y. van Kooyk, G. A. Rabinovich, *Nat. Immunol.* **2008**, *9*, 593–601.
- [7] G. A. Rabinovich, Y. van Kooyk, B. A. Cobb, *Ann. N. Y. Acad. Sci.* **2012**, *1253*, 1–15.
- [8] J. D. Marth, P. K. Grewal, *Nat. Rev. Immunol.* **2008**, *8*, 874–887.
- [9] C. Münz, J. D. Lünemann, M. T. Getts, S. D. Miller, *Nat. Rev. Immunol.* **2009**, *9*, 246–258.
- [10] A. M. Ercolini, S. D. Miller, *Clin. Exp. Immunol.* **2009**, *155*, 1–15.
- [11] G. Sherbet, *British Journal of Medical Practitioners* **2009**, *2*, 6–13.
- [12] J. Poole, C. J. Day, M. von Itzstein, J. C. Paton, M. P. Jennings, *Nature Reviews Microbiology* **2018**, *16*, 440.
- [13] N. Yuki, K. Susuki, M. Koga, Y. Nishimoto, M. Odaka, K. Hirata, K. Taguchi, T. Miyatake, K. Furukawa, T. Kobata, et al., *PNAS* **2004**, *101*, 11404–11409.
- [14] K. Kaida, S. Kusunoki, *Journal of Neuroimmunology* **2010**, *223*, 5–12.
- [15] M. C. Dalakas, in *Clinical Immunology (Fifth Edition)* (Eds.: R.R. Rich, T.A. Fleisher, W.T. Shearer, H.W. Schroeder, A.J. Frew, C.M. Weyand), Content Repository Only!, London, **2019**, pp. 903-915.e1.
- [16] C. Wim Ang, B. C. Jacobs, J. D. Laman, *Trends in Immunology* **2004**, *25*, 61–66.
- [17] D. R. Getts, E. M. L. Chastain, R. L. Terry, S. D. Miller, *Immunol. Rev.* **2013**, *255*, 197–209.
- [18] K. Kessenbrock, M. J. Fritzler, M. Groves, P. Eissfeller, C. A. von Mühlen, P. Höpfl, M. Mahler, *J. Mol. Med.* **2007**, *85*, 953–959.
- [19] M. M. Ahmed, S. M. Berney, R. E. Wolf, M. Hearth-Holmes, S. Hayat, E. Mubashir, H. Vanderheyde, W.-L. Chang, J. W. King, *Am. J. Med. Sci.* **2006**, *331*, 252–256.
- [20] I. Fattal, N. Shental, D. Mevorach, J.-M. Anaya, A. Livneh, P. Langevitz, G. Zandman-Goddard, R. Pauzner, M. Lerner, M. Blank, et al., *Immunology* **2010**, *130*, 337–343.
- [21] M. B. A. Oldstone, *Faseb J* **1998**, 1255–1265.
- [22] M. B. A. Oldstone, *Monoclonal Antibodies in Immunodiagnosis and Immunotherapy* **2014**, *33*, 158–165.
- [23] J. Varley, J. Taylor, S. R. Irani, *Neuropharmacology* **2018**, *132*, 71–82.
- [24] I. Sela-Culang, V. Kunik, Y. Ofran, *Front Immunol* **2013**, *4*, DOI 10.3389/fimmu.2013.00302.
- [25] A. Janda, A. Bowen, N. S. Greenspan, A. Casadevall, *Front Microbiol* **2016**, *7*, 22.
- [26] Bio-Rad, “Affinity and Avidity of Antibodies,” can be found under <https://www.bio-rad-antibodies.com/antigen-antibody-interactions.html>, **n.d.**
- [27] N. S. Lipman, L. R. Jackson, L. J. Trudel, F. Weis-Garcia, *ILAR J* **2005**, *46*, 258–268.
- [28] H. W. Schroeder, L. Cavacini, *J Allergy Clin Immunol* **2010**, *125*, S41–S52.

- [29] S. J. Beasley, J. M. Greer, *Clinical and Experimental Neuroimmunology* **2015**, *6*, 370–386.
- [30] C. Grönwall, G. J. Silverman, *J Clin Immunol* **2014**, *34*, S12–S21.
- [31] M. Boes, *Mol. Immunol.* **2000**, *37*, 1141–1149.
- [32] G. J. Silverman, J. Vas, C. Grönwall, *Nat Rev Rheumatol* **2013**, *9*, 291–300.
- [33] T. T. T. Nguyen, N. Baumgarth, *Crit Rev Immunol* **2016**, *36*, 163–177.
- [34] “Synthetic Polypeptides as Antigens, Volume 19 - 1st Edition,” can be found under <https://www.elsevier.com/books/synthetic-polypeptides-as-antigens/briand/978-0-444-80975-9>, **n.d.**
- [35] H. M. Geysen, S. J. Rodda, T. J. Mason, *Mol. Immunol.* **1986**, *23*, 709–715.
- [36] A. Mazzoleni, J.-M. Mallet, P. Rovero, A. M. Papini, *Archives of Biochemistry and Biophysics* **2019**, *663*, 44–53.
- [37] A. M. Papini, *Journal of Peptide Science* **2009**, *15*, 621–628.
- [38] F. Real-Fernández, I. Passalacqua, E. Peroni, M. Chelli, F. Lolli, A. M. Papini, P. Rovero, *Sensors (Basel)* **2012**, *12*, 5596–5607.
- [39] N. Schall, N. Page, C. Macri, O. Chaloin, J.-P. Briand, S. Muller, *J. Autoimmun.* **2012**, *39*, 143–153.
- [40] J. Vagner, H. Qu, V. J. Hruby, *Current Opinion in Chemical Biology* **2008**, *12*, 292–296.
- [41] F. Nuti, E. Peroni, F. Real-Fernández, M. A. Bonache, A. Le Chevalier-Isaad, M. Chelli, N. Lubin-Germain, J. Uziel, P. Rovero, F. Lolli, et al., *Biopolymers* **2010**, *94*, 791–799.
- [42] N. Auberger, M. Di Pisa, M. Larregola, G. Chassaing, E. Peroni, S. Lavielle, A.-M. Papini, O. Lequin, J.-M. Mallet, *Bioorg. Med. Chem.* **2014**, *22*, 6924–6932.
- [43] M. D. Rosenblum, I. K. Gratz, J. S. Paw, A. K. Abbas, *Sci Transl Med* **2012**, *4*, 125sr1.
- [44] C. Yu, J. Xi, M. Li, M. An, H. Liu, *Bioconjug. Chem.* **2018**, *29*, 719–732.
- [45] D. L. Hirsch, P. Ponda, *Immunotargets Ther* **2014**, *4*, 1–11.
- [46] J. Pozsgay, Z. Szekanecz, G. Sármay, *Nature Reviews Rheumatology* **2017**, *13*, 525–537.
- [47] P. Serra, P. Santamaria, *Nat Biotechnol* **2019**, *37*, 238–251.
- [48] A. P. Sanchez, R. Cunard, D. M. Ward, *J Clin Apher* **2013**, *28*, 20–29.
- [49] G. J. Pons-Estel, G. E. Salerni, R. M. Serrano, J. A. Gomez-Puerta, M. A. Plasin, E. Aldasoro, M. Lozano, J. Cid, R. Cervera, G. Espinosa, *Autoimmunity Reviews* **2011**, *10*, 679–684.
- [50] D. Aguirre-Valencia, J. Naranjo-Escobar, I. Posso-Osorio, M. C. Macía-Mejía, I. Nieto-Aristizábal, T. Barrera, M. A. Obando, G. J. Tobón, “Therapeutic Plasma Exchange as Management of Complicated Systemic Lupus Erythematosus and Other Autoimmune Diseases,” DOI 10.1155/2019/5350960 can be found under <https://www.hindawi.com/journals/ad/2019/5350960/>, **2019**.
- [51] A. P. Batocchi, A. Evoli, C. D. Schino, P. Tonali, *Therapeutic Apheresis* **2000**, *4*, 275–279.
- [52] T. Burnouf, H. Goubran, M. Radosevich, *Journal of Chromatography B: Biomedical Sciences and Applications* **1998**, *715*, 65–80.
- [53] P. M. Honoré, R. Jacobs, E. D. Waele, V. V. Gorp, H. D. Spapen, *BPU* **2014**, *38*, 158–159.
- [54] S. Oji, K. Nomura, *Transfusion and Apheresis Science* **2017**, *56*, 671–676.
- [55] C. Fassbender, R. Klingel, W. Köhler, *Atheroscler Suppl* **2017**, *30*, 257–263.
- [56] K. K. R. Tetala, A. P. Heikema, A. V. Pukin, C. A. G. M. Weijers, A. P. Tio-Gillen, M. Gilbert, H. P. Endtz, A. van Belkum, H. Zuilhof, G. M. Visser, et al., *J. Med. Chem.* **2011**, *54*, 3500–3505.

- [57] H. Alkan, N. Bereli, Z. Baysal, A. Denizli, *Biochemical Engineering Journal* **2010**, *51*, 153–159.
- [58] “Plasmapheresis | Asahi Kasei Medical Co., Ltd.,” can be found under [http://www.asahi-kasei.co.jp/medical/en/personal/cure/cure\\_01.html](http://www.asahi-kasei.co.jp/medical/en/personal/cure/cure_01.html), **n.d.**
- [59] H. J. Willison, K. Townson, J. Veitch, J. Boffey, N. Isaacs, S. M. Andersen, P. Zhang, C.-C. Ling, D. R. Bundle, *Brain* **2004**, *127*, 680–691.
- [60] J. Weinmann-Menke, S. Holtz, D. Sollinger, M. Dörken, S. Boedecker, B. Schamberger, F. Pfister, K. Amann, J. Lutz, *American Journal of Kidney Diseases* **2019**, DOI 10.1053/j.ajkd.2019.05.021.
- [61] K. Lazaridis, I. Dalianoudis, V. Baltatzidi, S. J. Tzartos, *Journal of Neuroimmunology* **2017**, *312*, 24–30.
- [62] B. M. Segal, in *Clinical Immunology (Fifth Edition)* (Eds.: R.R. Rich, T.A. Fleisher, W.T. Shearer, H.W. Schroeder, A.J. Frew, C.M. Weyand), Content Repository Only!, London, **2019**, pp. 891-902.e1.
- [63] M. S. Weber, B. Hemmer, S. Cepok, *Biochim. Biophys. Acta* **2011**, *1812*, 239–245.
- [64] A. J. Thompson, B. L. Banwell, F. Barkhof, W. M. Carroll, T. Coetzee, G. Comi, J. Correale, F. Fazekas, M. Filippi, M. S. Freedman, et al., *The Lancet Neurology* **2018**, *17*, 162–173.
- [65] A. D’Ambrosio, S. Pontecorvo, T. Colasanti, S. Zamboni, A. Francia, P. Margutti, *Autoimmunity Reviews* **2015**, *14*, 1097–1110.
- [66] A. Minagar, “Current and Future Therapies for Multiple Sclerosis,” DOI 10.1155/2013/249101 can be found under <https://www.hindawi.com/journals/scientifica/2013/249101/abs/>, **2013**.
- [67] F. D. Angelis, N. A. John, W. J. Brownlee, *BMJ* **2018**, *363*, k4674.
- [68] D. D. Mitsikostas, D. S. Goodin, *Multiple Sclerosis and Related Disorders* **2017**, *18*, 109–116.
- [69] “Mechanism and adverse effects of multiple sclerosis drugs: a review article. Part 1,” can be found under <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6737429/>, **n.d.**
- [70] “Mechanism and adverse effects of multiple sclerosis drugs: a review article. Part 2,” can be found under <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6737425/>, **n.d.**
- [71] A.-K. Pröbstel, N. S. R. Sanderson, T. Derfuss, *Int J Mol Sci* **2015**, *16*, 16576–16592.
- [72] B. G. Weinshenker, P. C. O’Brien, T. M. Petterson, J. H. Noseworthy, C. F. Lucchinetti, D. W. Dodick, A. A. Pineda, L. N. Stevens, M. Rodriguez, *Annals of Neurology* **1999**, *46*, 878–886.
- [73] R. Deschamps, A. Gueguen, N. Parquet, S. Saheb, F. Driss, M. Mesnil, C. Vignal, J. Aboab, R. Depaz, O. Gout, *J. Neurol.* **2016**, *263*, 883–887.
- [74] M. Keegan, F. König, R. McClelland, W. Brück, Y. Morales, A. Bitsch, H. Panitch, H. Lassmann, B. Weinshenker, M. Rodriguez, et al., *The Lancet* **2005**, *366*, 579–582.
- [75] S. Kinzel, M. S. Weber, *Brain Sci* **2017**, *7*, DOI 10.3390/brainsci7070070.
- [76] “Protein post-translational modifications in bacteria | Nature Reviews Microbiology,” can be found under <https://www.nature.com/articles/s41579-019-0243-0>, **n.d.**
- [77] C. Schäffer, P. Messner, *FEMS Microbiol. Rev.* **2017**, *41*, 49–91.
- [78] H. A. Doyle, M. J. Mamula, *Trends in Immunology* **2001**, *22*, 443–449.
- [79] S. M. Anderton, *Current Opinion in Immunology* **2004**, *16*, 753–758.
- [80] J. Cui, F. Shao, *Trends in Biochemical Sciences* **2011**, *36*, 532–540.

- [81] Q. Lu, S. Li, F. Shao, *Trends Microbiol.* **2015**, *23*, 630–641.
- [82] S. Mazzucco, S. Matà, M. Vergelli, R. Fioresi, E. Nardi, B. Mazzanti, M. Chelli, F. Lolli, M. Ginanneschi, F. Pinto, et al., *Bioorg. Med. Chem. Lett.* **1999**, *9*, 167–172.
- [83] F. Lolli, B. Mazzanti, M. Pazzagli, E. Peroni, M. C. Alcaro, G. Sabatino, R. Lanzillo, V. Brescia Morra, L. Santoro, C. Gasperini, et al., *J. Neuroimmunol.* **2005**, *167*, 131–137.
- [84] A. Carotenuto, A. M. D’Ursi, E. Nardi, A. M. Papini, P. Rovero, *J. Med. Chem.* **2001**, *44*, 2378–2381.
- [85] F. Lolli, B. Mulinacci, A. Carotenuto, B. Bonetti, G. Sabatino, B. Mazzanti, A. M. D’Ursi, E. Novellino, M. Pazzagli, L. Lovato, et al., *PNAS* **2005**, *102*, 10273–10278.
- [86] A. Carotenuto, A. M. D’Ursi, B. Mulinacci, I. Paolini, F. Lolli, A. M. Papini, E. Novellino, P. Rovero, *J. Med. Chem.* **2006**, *49*, 5072–5079.
- [87] A. Carotenuto, M. C. Alcaro, M. R. Saviello, E. Peroni, F. Nuti, A. M. Papini, E. Novellino, P. Rovero, *J. Med. Chem.* **2008**, *51*, 5304–5309.
- [88] C. Guardiani, G. F. Signorini, R. Livi, A. M. Papini, P. Procacci, *J. Phys. Chem. B* **2012**, *116*, 5458–5467.
- [89] R. Schreiner, E. Schnabel, F. Wieland, *The Journal of Cell Biology* **1994**, *124*, 1071–1081.
- [90] D. Calo, L. Kaminski, J. Eichler, *Glycobiology* **2010**, *20*, 1065–1076.
- [91] J. Gross, S. Grass, A. E. Davis, P. Gilmore-Erdmann, R. R. Townsend, J. W. St. Geme, *J Biol Chem* **2008**, *283*, 26010–26015.
- [92] S. Grass, C. F. Lichti, R. R. Townsend, J. Gross, J. W. St Geme, *PLoS Pathog.* **2010**, *6*, e1000919.
- [93] M. T. C. Walvoort, C. Testa, R. Eilam, R. Aharoni, F. Nuti, G. Rossi, F. Real-Fernandez, R. Lanzillo, V. Brescia Morra, F. Lolli, et al., *Sci Rep* **2016**, *6*, DOI 10.1038/srep39430.
- [94] S. Pandey, I. Dioni, D. Lambardi, F. Real-Fernandez, E. Peroni, G. Pacini, F. Lolli, R. Seraglia, A. M. Papini, P. Rovero, *Mol Cell Proteomics* **2013**, *12*, 277–282.
- [95] “Designed Glucopeptides Mimetics of Myelin Protein Epitopes As Synthetic Probes for the Detection of Autoantibodies, Biomarkers of Multiple Sclerosis | Journal of Medicinal Chemistry,” can be found under <https://pubs.acs.org/doi/abs/10.1021/jm301031r>, **n.d.**
- [96] F. R. Fernández, M. D. Pisa, G. Rossi, N. Auberger, O. Lequin, M. Larregola, A. Benchohra, C. Mansuy, G. Chassaing, F. Lolli, et al., *Peptide Science* **2015**, *104*, 560–576.
- [97] M. T. C. Walvoort, C. Testa, R. Eilam, R. Aharoni, F. Nuti, G. Rossi, F. Real-Fernandez, R. Lanzillo, V. Brescia Morra, F. Lolli, et al., *Sci Rep* **2016**, *6*, DOI 10.1038/srep39430.
- [98] I. Paolini, F. Nuti, M. de la Cruz Pozo-Carrero, F. Barbetti, B. Kolesinska, Z. J. Kaminski, M. Chelli, A. M. Papini, *Tetrahedron Letters* **2007**, *48*, 2901–2904.
- [99] F. M. Ibatullin, S. I. Selivanov, *Tetrahedron Letters* **2009**, *50*, 6351–6354.
- [100] E. Engvall, P. Perlmann, *J. Immunol.* **1972**, *109*, 129–135.
- [101] A. Palermo, A. Nesterov-Mueller, *Int J Mol Sci* **2019**, *20*, DOI 10.3390/ijms20030604.
- [102] A. Bierzyński, *Acta Biochim Pol* **2001**, *48*, 1091–1099.
- [103] D. H. Dube, C. R. Bertozzi, *Nat Rev Drug Discov* **2005**, *4*, 477–488.
- [104] E. I. Buzás, B. György, M. Pásztói, I. Jelinek, A. Falus, H.-J. Gabius, *Autoimmunity* **2006**, *39*, 691–704.

- [105] S. E. O'Connor, B. Imperiali, *Chem. Biol.* **1996**, *3*, 803–812.
- [106] B. Imperiali, K. W. Rickert, *Proc Natl Acad Sci U S A* **1995**, *92*, 97–101.
- [107] B. Imperiali, S. E. O'Connor, *Current Opinion in Chemical Biology* **1999**, *3*, 643–649.
- [108] S. E. O'Connor, B. Imperiali, *Chemistry & Biology* **1998**, *5*, 427–437.
- [109] S. E. O'Connor, B. Imperiali, *J. Am. Chem. Soc.* **1997**, *119*, 2295–2296.
- [110] A. Viegas, J. Manso, F. L. Nobrega, E. J. Cabrita, *J. Chem. Educ.* **2011**, *88*, 990–994.
- [111] “Mechanism by which 2,2,2-trifluoroethanol/water mixtures stabilize secondary-structure formation in peptides: A molecular dynamics study | PNAS,” can be found under <https://www.pnas.org/content/99/19/12179>, **n.d.**
- [112] “Side chain–backbone hydrogen bonding contributes to helix stability in peptides derived from an  $\alpha$ -helical region of carboxypeptidase A - Bruch - 1991 - Proteins: Structure, Function, and Bioinformatics - Wiley Online Library,” can be found under <https://onlinelibrary.wiley.com/doi/abs/10.1002/prot.340100206>, **n.d.**
- [113] “Helical Structure of Dermaseptin B2 in a Membrane-Mimetic Environment | Biochemistry,” can be found under <https://pubs.acs.org/doi/full/10.1021/bi034401d>, **n.d.**
- [114] J. Thundimadathil, *Chimica Oggi* **2013**, *31*, 4.
- [115] J. P. Tam, *Journal of Immunological Methods* **1996**, *196*, 17–32.
- [116] B. Oller-Salvia, M. Sánchez-Navarro, S. Ciudad, M. Guiu, P. Arranz-Gibert, C. Garcia, R. R. Gomis, R. Cecchelli, J. García, E. Giralt, et al., *Angewandte Chemie International Edition* **2016**, *55*, 572–575.
- [117] B. Oller-Salvia, M. Sánchez-Navarro, E. Giralt, M. Teixidó, *Chem Soc Rev* **2016**, *45*, 4690–4707.
- [118] C. L. Young, Z. T. Britton, A. S. Robinson, *Biotechnology Journal* **2012**, *7*, 620–634.
- [119] A. Spriestersbach, J. Kubicek, F. Schäfer, H. Block, B. Maertens, *Meth. Enzymol.* **2015**, *559*, 1–15.
- [120] A. S. Carlini, L. Adamiak, N. C. Gianneschi, *Macromolecules* **2016**, *49*, 4379–4394.
- [121] J. Y. Shu, B. Panganiban, T. Xu, *Annual Review of Physical Chemistry* **2013**, *64*, 631–657.
- [122] H. Nandivada, X. Jiang, J. Lahann, *Advanced Materials* **2007**, *19*, 2197–2208.
- [123] K. Nwe, M. W. Brechbiel, *Cancer Biotherapy and Radiopharmaceuticals* **2009**, *24*, 289–302.
- [124] Y. Zhang, R. G. Carbonell, O. J. Rojas, *Biomacromolecules* **2013**, *14*, 4161–4168.
- [125] D. Duret, A. Grassin, M. Henry, T. Jacquet, F. Thoreau, S. Denis-Quanquin, J.-L. Coll, D. Boturyn, A. Favier, M.-T. Charreyre, *Bioconjugate Chem.* **2017**, *28*, 2241–2245.
- [126] C. Ngambenjawong, S. H. Pun, *ACS Biomater. Sci. Eng.* **2017**, *3*, 2050–2053.
- [127] M. A. Gauthier, H.-A. Klok, *Chem. Commun.* **2008**, 2591–2611.
- [128] J. Hennicke, A. M. Lastin, D. Reinhart, C. Grünwald-Gruber, F. Altmann, R. Kunert, *Analytical Biochemistry* **2017**, *539*, 162–166.
- [129] Y.-D. Luo, Q.-L. Zhang, S.-J. Yao, D.-Q. Lin, *Journal of Chromatography A* **2018**, *1533*, 77–86.
- [130] R. Mehvar, *Journal of Controlled Release* **2000**, *69*, 1–25.
- [131] R. Wilson, D. G. Spiller, A. Beckett, I. A. Prior, V. Sée, *Chem. Mater.* **2010**, *22*, 6361–6369.

- [132] J.-P. Meyer, K. M. Tully, J. Jackson, T. R. Dilling, T. Reiner, J. S. Lewis, *Bioconjugate Chem.* **2018**, *29*, 538–545.
- [133] E. C. Gil, A. I. Colarte, A. El Ghzaoui, D. Durand, J. L. Delarbre, B. Bataille, *European Journal of Pharmaceutics and Biopharmaceutics* **2008**, *68*, 319–329.
- [134] J. Morimoto, M. Sarkar, S. Kenrick, T. Kodadek, *Bioconjug Chem* **2014**, *25*, 1479–1491.
- [135] S. Oliver, A. Jofri, D. S. Thomas, O. Vittorio, M. Kavallaris, C. Boyer, *Carbohydr Polym* **2017**, *169*, 480–494.
- [136] T. T. Nielsen, V. Wintgens, C. Amiel, R. Wimmer, K. L. Larsen, *Biomacromolecules* **2010**, *11*, 1710–1715.
- [137] Y. Chau, F. E. Tan, R. Langer, *Bioconjugate Chem.* **2004**, *15*, 931–941.
- [138] L. W. Ståde, T. T. Nielsen, L. Duroux, M. Hinge, K. Shimizu, L. Gurevich, P. K. Kristensen, C. Wingren, K. L. Larsen, *ACS Appl Mater Interfaces* **2015**, *7*, 4160–4168.
- [139] I. Antoniuk, V. Wintgens, G. Volet, T. T. Nielsen, C. Amiel, *Carbohydrate Polymers* **2015**, *133*, 473–481.
- [140] “Peptidotriazoles on Solid Phase: [1,2,3]-Triazoles by Regiospecific Copper(I)-Catalyzed 1,3-Dipolar Cycloadditions of Terminal Alkynes to Azides | The Journal of Organic Chemistry,” can be found under <https://pubs.acs.org/doi/10.1021/jo011148j>, **n.d.**
- [141] V. V. Rostovtsev, L. G. Green, V. V. Fokin, K. B. Sharpless, *Angewandte Chemie International Edition* **2002**, *41*, 2596–2599.
- [142] V. Hong, S. I. Presolski, C. Ma, M. G. Finn, *Angewandte Chemie International Edition* **2009**, *48*, 9879–9883.
- [143] A. A. H. Ahmad Fuaad, F. Azmi, M. Skwarczynski, I. Toth, *Molecules* **2013**, *18*, 13148–13174.
- [144] V. D. Bock, H. Hiemstra, J. H. van Maarseveen, *European Journal of Organic Chemistry* **2006**, *2006*, 51–68.
- [145] L. Horner, A. Christmann, *Angewandte Chemie International Edition in English* **1963**, *2*, 599–608.
- [146] *Reactive Intermediate Chemistry*, John Wiley & Sons, Ltd, **2003**.
- [147] H. Peng, K. H. Dornevil, A. B. Draganov, W. Chen, C. Dai, W. H. Nelson, A. Liu, B. Wang, *Tetrahedron* **2013**, *69*, 5079–5085.
- [148] F. Real Fernández, M. Di Pisa, G. Rossi, N. Auberger, O. Lequin, M. Larregola, A. Benchohra, C. Mansuy, G. Chassaing, F. Lolli, et al., *Biopolymers* **2015**, *104*, 560–576.
- [149] P. Cuatrecasas, *J. Biol. Chem.* **1970**, *245*, 3059–3065.
- [150] E. A. dos Santos, A. S. de Oliveira, L. M. A. Rabêlo, A. F. Uchôa, A. H. A. Morais, *Affinity Chromatography* **2012**, DOI 10.5772/34982.
- [151] J. Kohn, M. Wilchek, *Enzyme and Microbial Technology* **1982**, *4*, 161–163.
- [152] G. S. David, T. H. Chino, R. A. Reisfeld, *FEBS Letters* **1974**, *43*, 264–266.
- [153] A. C. A. Roque, C. S. O. Silva, M. Â. Taipa, *Journal of Chromatography A* **2007**, *1160*, 44–55.
- [154] E. Khalikova, P. Susi, T. Korpela, *Microbiol Mol Biol Rev* **2005**, *69*, 306–325.
- [155] M. Zaman, I. Toth, *Front Immunol* **2013**, *4*, DOI 10.3389/fimmu.2013.00318.
- [156] P. Ghosh, G. Han, M. De, C. K. Kim, V. M. Rotello, *Advanced Drug Delivery Reviews* **2008**, *60*, 1307–1315.
- [157] J. W. Lee, S. J. Sim, S. M. Cho, J. Lee, *Biosensors and Bioelectronics* **2005**, *20*, 1422–1427.
- [158] S. M. Muthana, L. Xia, C. T. Campbell, Y. Zhang, J. C. Gildersleeve, *PLoS One* **2015**, *10*, DOI 10.1371/journal.pone.0119298.

- [159] N. Dotan, R. Altstock, M. Schwarz, A. Dukler, *Lupus* **2006**, *15*, 442–450.
- [160] C. Haase, O. Seitz, in *Glycopeptides and Glycoproteins: Synthesis, Structure, and Application* (Ed.: V. Wittmann), Springer, Berlin, Heidelberg, **2007**, pp. 1–36.
- [161] L. Krasnova, C.-H. Wong, *Annu. Rev. Biochem.* **2016**, *85*, 599–630.
- [162] M. Brito-Arias, in *Synthesis and Characterization of Glycosides* (Ed.: M. Brito-Arias), Springer International Publishing, Cham, **2016**, pp. 311–353.
- [163] P. E. Dawson, T. W. Muir, I. Clark-Lewis, S. B. Kent, *Science* **1994**, *266*, 776–779.
- [164] D. M. M. Jaradat, *Amino Acids* **2018**, *50*, 39–68.
- [165] B. L. Bray, *Nat Rev Drug Discov* **2003**, *2*, 587–593.
- [166] Y. Kajihara, N. Yamamoto, R. Okamoto, K. Hirano, T. Murase, *Chem Rec* **2010**, *10*, 80–100.
- [167] R. Okamoto, M. Izumi, Y. Kajihara, *Current Opinion in Chemical Biology* **2014**, *22*, 92–99.
- [168] B. H. M. Kuijpers, S. Groothuys, A. C. Soede, P. Laverman, O. C. Boerman, F. L. van Delft, F. P. J. T. Rutjes, *Bioconjug. Chem.* **2007**, *18*, 1847–1854.
- [169] D. Macmillan, A. M. Daines, M. Bayrhuber, S. L. Flitsch, *Org. Lett.* **2002**, *4*, 1467–1470.
- [170] L. A. Marcaurelle, C. R. Bertozzi, *J. Am. Chem. Soc.* **2001**, *123*, 1587–1595.
- [171] X. Song, R. I. Hollingsworth, *Synlett* **2006**, *2006*, 3451–3454.
- [172] D. H. Rich, J. Singh, in *Major Methods of Peptide Bond Formation* (Eds.: E. Gross, J. Meienhofer), Academic Press, **1979**, pp. 241–261.
- [173] T. Takahashi, T. Suzuki, *J Lipid Res* **2012**, *53*, 1437–1450.
- [174] A. Kuriakose, N. Chirmule, P. Nair, *J Immunol Res* **2016**, *2016*, DOI 10.1155/2016/1298473.
- [175] J. Morise, H. Takematsu, S. Oka, *Biochimica et Biophysica Acta (BBA) - General Subjects* **2017**, *1861*, 2455–2461.
- [176] V. E. Shashoua, P. F. Daniel, M. E. Moore, F. B. Jungalwala, *Biochemical and Biophysical Research Communications* **1986**, *138*, 902–909.
- [177] Y. E. Tsvetkov, M. Burg-Roderfeld, G. Loers, A. Ardá, E. V. Sukhova, E. A. Khatuntseva, A. A. Grachev, A. O. Chizhov, H.-C. Siebert, M. Schachner, et al., *J. Am. Chem. Soc.* **2012**, *134*, 426–435.
- [178] M. Ieronymaki, F. Nuti, D. Brancaccio, G. Rossi, F. Real-Fernández, Y. Cao, O. Monasson, M. Larregola, E. Peroni, J. Uziel, et al., *ChemMedChem* **2017**, *12*, 751–759.
- [179] S. van der Vorm, J. M. A. van Hengst, M. Bakker, H. S. Overkleeft, G. A. van der Marel, J. D. C. Codée, *Angewandte Chemie International Edition* **2018**, *57*, 8240–8244.
- [180] G. Catelani, F. Colonna, A. Marra, *Carbohydrate Research* **1988**, *182*, 297–300.
- [181] F. Belot, A. Otter, M. Fukuda, O. Hindsgaul, *Synlett* **2003**, *2003*, 1315–1318.
- [182] R. Chevalier, B. Colsch, C. Afonso, N. Baumann, J.-C. Tabet, J.-M. Mallet, *Tetrahedron* **2006**, *62*, 563–577.
- [183] A. V. Kornilov, E. V. Sukhova, N. E. Nifantiev, *Carbohydrate Research* **2001**, *336*, 309–313.
- [184] R. Herrendorff, P. Hänggi, H. Pfister, F. Yang, D. Demeestere, F. Hunziker, S. Frey, N. Schaeren-Wiemers, A. J. Steck, B. Ernst, *PNAS* **2017**, *114*, E3689–E3698.
- [185] X. Zhu, R. R. Schmidt, *Angew. Chem. Int. Ed. Engl.* **2009**, *48*, 1900–1934.

- [186] Y. E. Tsvetkov, M. Burg-Roderfeld, G. Loers, A. Ardá, E. V. Sukhova, E. A. Khatuntseva, A. A. Grachev, A. O. Chizhov, H.-C. Siebert, M. Schachner, et al., *J. Am. Chem. Soc.* **2012**, *134*, 426–435.
- [187] R. Herrendorff, P. Hänggi, H. Pfister, F. Yang, D. Demeestere, F. Hunziker, S. Frey, N. Schaeren-Wiemers, A. J. Steck, B. Ernst, *PNAS* **2017**, *114*, E3689–E3698.
- [188] I. Rombouts, B. Lagrain, M. Brunnbauer, J. A. Delcour, P. Koehler, *Sci Rep* **2013**, *3*, 1–11.
- [189] N. J. Davis, S. L. Flitsch, *Tetrahedron Letters* **1991**, *32*, 6793–6796.
- [190] N. Yamamoto, T. Sakakibara, Y. Kajihara, *Tetrahedron Letters* **2004**, *45*, 3287–3290.
- [191] S. Y. Wong, G. R. Guile, R. A. Dwek, G. Arsequell, *Biochem J* **1994**, *300*, 843–850.
- [192] A. M. Papini, F. Nuti, F. Real-Fernandez, G. Rossi, C. Tiberi, G. Sabatino, S. Pandey, S. Leoncini, C. Signorini, A. Pecorelli, et al., *Journal of Immunology Research, Journal of Immunology Research* **2014**, *2014*, 2014, e260973.
- [193] C. De Felice, S. Leoncini, C. Signorini, A. Cortelazzo, P. Rovero, T. Durand, L. Ciccoli, A. M. Papini, J. Hayek, *Autoimmun Rev* **2016**, *15*, 411–416.
- [194] K. Hirano, M. Izumi, D. Macmillan, K. Tezuka, T. Tsuji, Y. Kajihara, *Journal of Carbohydrate Chemistry* **2011**, *30*, 306–319.
- [195] O. G. Adesoye, I. N. Mills, D. P. Temelkoff, J. A. Jackson, P. Norris, *J. Chem. Educ.* **2012**, *89*, 943–945.
- [196] Y. Chen, A. S. Kamlet, J. B. Steinman, D. R. Liu, *Nature Chem* **2011**, *3*, 146–153.
- [197] M. A. Maier, C. G. Yannopoulos, N. Mohamed, A. Roland, H. Fritz, V. Mohan, G. Just, M. Manoharan, *Bioconjugate Chem.* **2003**, *14*, 18–29.
- [198] “NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy,” can be found under <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4393527/>, **n.d.**
- [199] H. E. Gottlieb, V. Kotlyar, A. Nudelman, *J. Org. Chem.* **1997**, *62*, 7512–7515.

## RÉSUMÉ

---

Le but principal de ce travail de thèse a été de développer des sondes peptidiques pour détecter et isoler des auto-anticorps spécifiques à haute affinité à partir du sérum de patients souffrant de sclérose en plaques (SEP). Nous avons sélectionné un peptide di-glucosylé de l'adhésine HMW1 comme séquence minimale capable de détecter les anti-N(Glc) IgM. Ainsi, un dextrane de 40 kDa a été modifié avec des groupes propargyle et a été utilisé comme échafaudage pour conjuguer par CuAAC le peptide N-glucosylé. On a prouvé que ce nouveau polymère augmente considérablement la puissance de la liaison des IgG et IgM caractéristique de la forme de SEP que nous venons de caractériser. Les anticorps d'un sérum représentatif ont été purifiés avec succès, comme confirmé par test ELISA. Ce résultat semble prometteur en tant que preuve de concept de la possible déplétion sélective des auto-anticorps circulants dans la SEP, qui pourrait conduire à développer un dispositif thérapeutique spécifique basé sur aphérèse.

## MOTS CLÉS

---

peptides ; maladies auto-immunes ; glycosylation ; anticorps ; glycoconjugués ; sclérose en plaque

## ABSTRACT

---

The main purpose of this thesis was to develop peptide probes to detect and isolate specific and high affinity antibodies from sera of patients suffering from multiple sclerosis (MS). We selected a di-glucosylated adhesin HMW1 peptide as the shortest sequence up to now able to compete with the highest affinity with anti-N(Glc) IgM binding. 40 kDa dextran was modified with propargyl groups and used as a scaffold to conjugate by CuAAC the di-glucosylated peptide. This novel polymeric structure was proven to dramatically increase binding potency of IgGs and IgMs in MS sera. Abs from a representative MS serum, were successfully purified on a sepharose resin specifically modified with the adhesin peptide-dextran conjugate, as confirmed by ELISA. This result appears promising as a proof-of-concept of the selective removal of circulating autoantibodies (possibly perpetuating nonself recognition) that could likely lead to develop a specific apheresis-based device.

## KEYWORDS

---

peptides ; auto-immune diseases ; glycosylation ; antibodies ; glycoconjugates ; multiple sclerosis