



POMPEU FABRA UNIVERSITY
Department of Experimental and Health Sciences

**Capture and identification of
carbohydrate-binding proteins
by SPR and CREDEX-MS**

PhD thesis

Carmen Jiménez Castells

Barcelona, November 2009



POMPEU FABRA UNIVERSITY
Department of Experimental and Health Sciences

Doctoral Programme: Health and Life Sciences

Capture and identification of carbohydrate-binding proteins by SPR and CREDEX-MS

Memoria presentada por CARMEN JIMÉNEZ CASTELLS para optar al título de Doctor por la Universidad Pompeu Fabra. Esta tesis ha sido realizada bajo la codirección del Dr. David Andreu Martínez y el Dr. Ricardo Gutiérrez Gallego, en los grupos de Síntesis de Péptidos y Proteómica de la Universidad Pompeu Fabra y de Investigación en Bioanálisis y Servicios Analíticos, programa de Neuropsicofarmacología del Instituto Municipal de Investigación Médica (IMIM-Hospital del Mar). Programa de Doctorado en Ciencias de la Salud y de la Vida de la Universidad Pompeu Fabra.

Dr. David Andreu
Martínez
Director de tesis

Dr. Ricardo Gutiérrez
Gallego
Director de tesis

Carmen Jiménez
Castells
Doctorando

ACKNOWLEDGEMENTS

En primer lugar, quiero dar las gracias a mis directores de tesis por su confianza durante estos años. A David Andreu por darme la oportunidad de trabajar en su laboratorio y así, conocer cómo es el mundo de la investigación. A Ricardo Gutiérrez por su constante apoyo, inagotable paciencia y espíritu crítico. He aprendido mucho de ambos.

En segundo lugar, quiero agradecer a toda la gente con la que he tenido el placer de compartir laboratorio estos cinco años. A la gente de Proteómica (Eva, Cristina, Guadalupe, Carolina, Javi y Marta) por hacerme el trabajo en el laboratorio más agradable. Especialmente a Beatriz, por ayudarme en los primeros años cuando me peleaba por mejorar los rendimientos de las conjugaciones y aún no tenía claro si yo servía para la ciencia. A Wioleta, por escucharme y entenderme en muchos momentos, y por supuesto, a Sira, por ser una doctoranda tan entusiasta y con un enfoque diferente de los experimentos. También querría dar las gracias a ciertas personas que ya no están en el laboratorio, pero que sin ellas no podría haber llegado hasta aquí. A Miquel Vila Perelló por aportar la idea inicial de este trabajo y a Panni Jacab por guiarme durante los primeros pasos en la síntesis de péptidos.

En tercer lugar, me gustaría agradecer a la gente de Farmacología del IMIM. Especialmente a los antiguos compañeros de despacho (Esther, Quim, Armand y Gerard), gracias a los cuales, el despacho parecía siempre el camarote de los hermanos Marx (siempre había algo que comentar y eran pocos los momentos de silencio). Ahora, sin ustedes, esto ya no es lo mismo y se les echa mucho de menos. Y claro está, que los desayunos no serían los mismos sin Ester C. y Mitona. Finalmente, me gustaría agradecer especialmente a Beth, por ser la “chica de las 6”, aunque últimamente se estaba convirtiendo en las “6 y media” o “casi 7”. Muchas gracias por esto y otras muchas cosas más.

ACKNOWLEDGEMENTS

Ich wollte auch mich bedanken zu Prof. M. Przybylski für die gute Zeiten in dem Analytische Chemie und Biopolymerchemie Labor und in Konstanz. (Thank to Prof. M. Przybylski for his welcome in the laboratory of Analytical Chemistry and Biopolymer Structure Analysis (Konstanz)).

Particularly, to A. Moise, who teach me what I know about CREDEX-MS.

Agradezco al Ministerio de Educación y Ciencia la financiación de este trabajo a través del proyecto de investigación BES-2006-12879 y de la acción integrada hispano-alemana HA2007-0021.

Igualmente quisiera agradecer a mi familia el apoyo y la confianza que siempre han tenido en mí, y que gracias a ellos he podido llegar al final de esta tesis.

Finalmente quiero darle las gracias a David R. por escucharme, aguantarme y aún así, apoyarme siempre. Por fín, estos últimos meses difíciles ya se han acabado.

ABSTRACT

Carbohydrate-binding proteins of non-immunological origin –lectins– have been recognized over the last decades as decisive players in numerous biological processes, ranging from cell-cell communication, fertilization, pathogen-cell adhesion to metastasis. Consequently, there is an increasing interest in finding powerful and nanosized tools to screen for these molecules and to study their carbohydrate interactions in detail. Here, two complementary approaches are described to characterize lectin-carbohydrate interactions with high sensitivity, low sample consumption, and without the need for sample labelling: SPR and CREDEX-MS. In SPR, we have developed an approach where the sugar is immobilized onto a sensor surface through a tailor-made peptide module that allows (1) to capture the lectin, (2) to characterize the interaction through kinetic and thermodynamic parameters, and (3) to identify the interacted protein by mass spectrometry. In CREDEX-MS, based on proteolytic excision of protein-carbohydrate complexes and mass spectrometric analysis, the peptides conforming the carbohydrate binding domain are identified.

RESUMEN

Las lectinas (proteínas de origen no inmune capaces de reconocer azúcares) se han revelado en las últimas décadas como participantes cruciales en multitud de procesos biológicos, tales como la comunicación célula-célula, la fertilización, la adhesión del patógeno a la célula y la metástasis, entre muchos otros. Por lo tanto, existe un gran interés en el desarrollo de técnicas analíticas potentes para el estudio de las interacciones lectina-carbohidrato. En este trabajo, se describen dos aproximaciones complementarias mediante las cuales se pueden caracterizar las interacciones lectinas-azúcar con gran sensibilidad, poca utilización de muestra y sin la necesidad de ningún marcaje. En la técnica basada en resonancia de plasmón superficial (SPR), el azúcar es inmovilizado sobre una superficie a través de un módulo peptídico, lo cual permite (1) capturar la lectina, (2) caracterizar su interacción mediante parámetros cinéticos y termodinámicos y (3) identificar posteriormente la proteína mediante espectrometría de masas. Complementariamente, la técnica CREDEX-MS, basada en la excisión proteolítica del complejo proteína-azúcar y posterior análisis por espectrometría de masas, nos permite identificar los péptidos que forman parte del dominio de unión al azúcar.

ABBREVIATIONS

ACN	acetonitrile
Ahx	aminohexanoic acid
Ala (A)	alanine
Aoa	aminoxyacetic acid
Boc	<i>tert</i> -butyloxycarbonyl
CREDEX	Carbohydrate-REcognition-Domain-EXcision
DHB	dihydroxybenzoic acid
DIC	<i>N,N'</i> -diisopropylcarbodiimide
DIEA	<i>N,N</i> -diisopropylethylamine
DMF	dimethylformamide
DVS	divinylsulfone
ECA	<i>Erythrina cristagalli</i> agglutinin
EDC	1-ethyl-3-(3-diethylaminopropyl)-carbodiimide
Fmoc	9-fluorenylmethyloxycarbonyl
Fuc	fucose
Gal	galactose
Glc	glucose
GalNAc	<i>N</i> -acetyl-D-galactosamine
GlcNAc	<i>N</i> -acetyl-D-glucosamine
Gly (G)	glycine
HBS	HEPES saline buffer
HBTU	2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium

ABBREVIATIONS

HPLC	high performance liquid chromatography
lac	lactose
lacNAc	<i>N</i> -acetyllactosamine
LTA	<i>Lotus tetragonolobus</i> agglutinin
Lys (K)	lysine
MAA	<i>Maackia amurensis</i> agglutinin
MALDI-TOF	matrix assisted laser desorption ionization time of flight
MS	mass spectrometry
Neu5Ac	<i>N</i> -acetyl-neuraminic acid
NHS	<i>N</i> -hydroxysuccinimide
<i>N</i> [Me]-Aoa	<i>N</i> -methyl-aminoxyacetic acid
NMR	nuclear magnetic resonance
Phe (F)	phenylalanine
pI	isoelectric point
RCA	<i>Ricinus communis</i> agglutinin
rt	room temperature
RU	resonance units
SNA	<i>Sambucus nigra</i> agglutinin
SPPS	solid phase peptide synthesis
SPR	surface plasmon resonance
TFA	trifluoroacetic acid
UEA-I	<i>Ulex europeaus</i> agglutinin I
WGA	wheat germ agglutinin

Table of reagents employed in SPPS

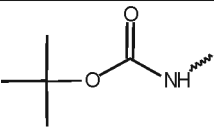
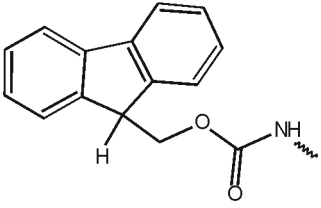
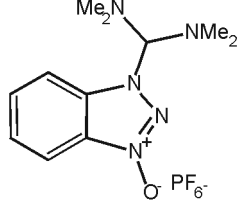
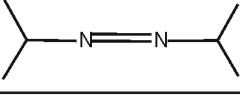
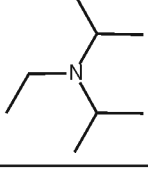
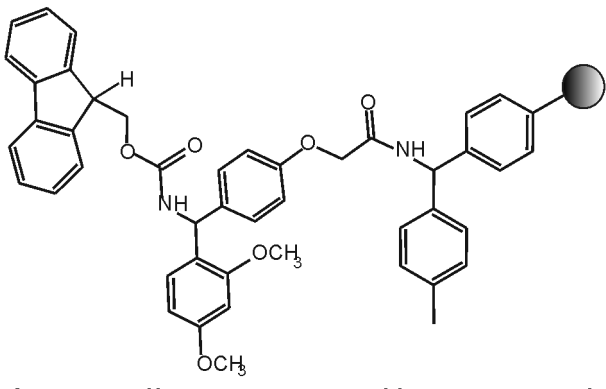
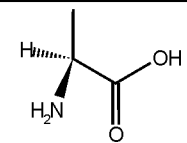
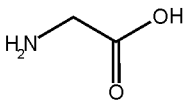
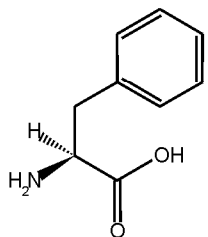
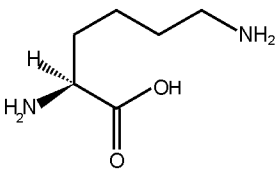
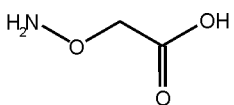
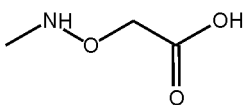
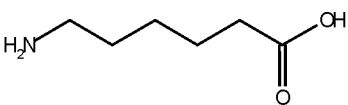
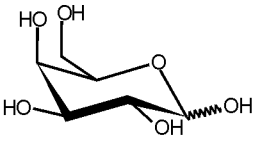
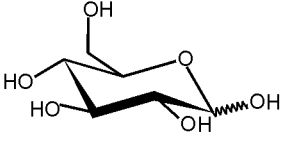
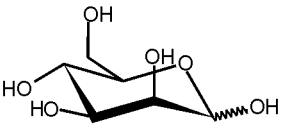
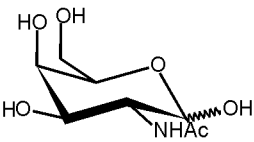
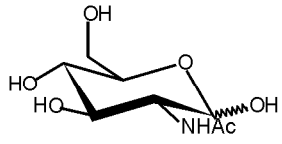
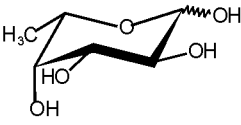
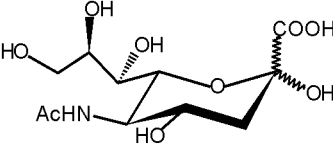
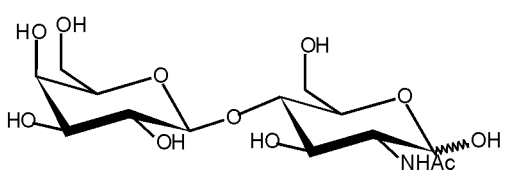
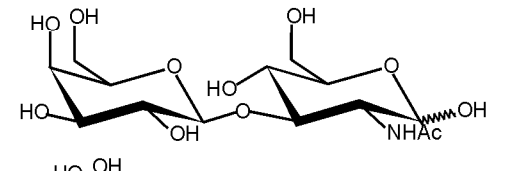
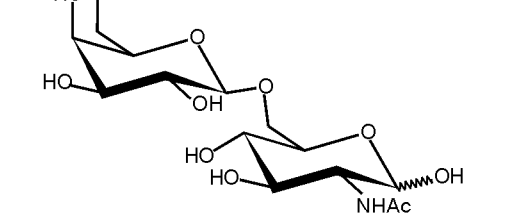
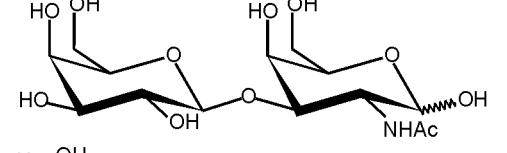
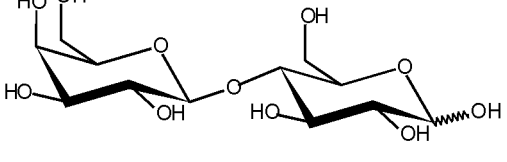
Reagents	Abbreviation	Structure
Protecting groups	Boc	
	Fmoc	
Coupling agents	HBTU	
	DIC	
Additives	DIEA	
Rink amide MBHA resin		 <div style="display: flex; justify-content: space-around; margin-top: 10px;"> ┌──────────┴──────────┐ ┌──────────┴──────────┐ ┌──────────┴──────────┐ </div> <div style="display: flex; justify-content: space-around; margin-top: 5px;"> Fmoc Rink amide p-MBHA </div>

Table of reagents employed in SPPS

Amino acids	Abbreviation	Structure
Alanine	Ala	
Glycine	Gly	
Phenylalanine	Phe	
Lysine	Lys	
Other residues	Abbreviation	Structure
Aminooxyacetic acid	Aoa	
Methyl-aminooxyacetic acid	N[Me]-Aoa	
Aminohexanoic acid	Ahx	

Monosaccharides	Abbreviation	Structure
Galactose	Gal	
Glucose	Glc	
Mannose	Man	
<i>N</i> -acetyl-galactosamine	GalNAc	
<i>N</i> -acetyl-glucosamine	GlcNAc	
Fuc	Fucose	
<i>N</i> -acetyl-neuraminic acid	Neu5Ac	

Disaccharides	
Abbreviation	Structure
Gal- β 1,4-GlcNAc	
Gal- β 1,3-GlcNAc	
Gal- β 1,6-GlcNAc	
Gal- β 1,3-GalNAc	
Gal- β 1,4-Glc	

Disaccharides	
Abbreviation	Structure
Man- α 1,2-Man	
Man- α 1,3-Man	
Man- α 1,6-Man	
Fuc- α 1,3-GlcNAc	
Fuc- α 1,4-GlcNAc	
Fuc- α 1,6-GlcNAc	

Trisaccharides	
Abbreviation	Structure
Neu5Ac- α 2,3-Gal- β 1,4-GlcNAc	
Neu5Ac- α 2,6-Gal- β 1,4-GlcNAc	

CONTENTS

1	INTRODUCTION.....	1
1.1	The sugar code.....	1
1.1.1	Monosaccharides.....	2
1.1.2	Glycoconjugates: glycoproteins, proteoglycans and glycolipids.....	5
1.2	Lectins as decipherers of the sugar code.....	17
1.2.1	History of lectins.....	17
1.2.2	Classification of lectins.....	19
1.2.3	Simple lectins: Plant lectins.....	20
1.2.4	Lectins in other organisms.....	23
1.3	Carbohydrate-lectin interactions.....	24
1.3.1	Structural basis of carbohydrate-lectin interactions.....	24
1.3.2	Glycoside cluster effect.....	25
1.3.3	Analytical methods for studying carbohydrate-lectin interactions: information about specificity, structure, kinetics and thermodynamics.....	27
1.4	Surface Plasmon resonance (SPR).....	30
1.4.1	The SPR principle.....	30
1.4.2	Principal applications: kinetics, affinity, and thermodynamics.....	33
1.4.3	SPR-MS.....	36
1.4.4	Carbohydrate-lectin interactions studied by SPR.....	39
1.4.5	Glycoprobe synthesis via chemoselective ligation.....	42

CONTENTS

1.5	Mass spectrometry	46
1.5.1	Structural information by limited proteolysis coupled with mass spectrometry	47
1.5.2	Carbohydrate-Recognition-Domain-EXcision-MS (CREDEX-MS)	48
2	OBJECTIVES	51
3	RESULTS	53
3.1	Glycprobe synthesis	53
3.1.1	Boc-aminoxyacetic acid coupling	53
3.1.2	The need for Aoa <i>N</i> -methylation	59
3.2	Chemoselective ligation via oxime chemistry	65
3.2.1	Conjugation of mono-, di- and trisaccharides with <i>N</i> [Me]-Aoa-peptide	66
3.2.2	Stability of glycopeptides	73
3.3	Application in SPR	75
3.3.1	Aglycon features: Peptide length	76
3.3.2	<i>N</i> -methylated vs non <i>N</i> -methylated probes	78
3.3.3	Affinity and kinetic studies of carbohydrate-lectin interactions: legume lectins as a proof of principle	82
3.3.4	Determination of thermodynamic parameters by SPR	97
3.3.5	Lectin affinity capture and characterization by SPR-MS	100
3.4	Identification of carbohydrate recognition domain (CRD) by CREDEX-MS	106
4	DISCUSSION	113
5	CONCLUSIONS	125

6	MATERIAL AND METHODS	127
6.1	Materials	127
6.2	Peptide synthesis.....	128
6.2.1	Boc-Aoa-OH coupling.....	128
6.2.2	Boc ₂ -Aoa-OH coupling	129
6.2.3	Boc- <i>N</i> [Me]-Aoa-OH coupling.....	129
6.2.4	Peptide cleavage and work-up.....	129
6.3	Chemoselective ligation of carbohydrates	130
6.3.1	Conjugation between <i>N</i> [Me]-Aoa-peptide and monosaccharides	130
6.3.2	Conjugation between aminoxy-peptides and lactose.....	130
6.3.3	Conjugation between Aoa-peptide and other oligosaccharides	131
6.3.4	Conjugation between <i>N</i> [Me]-Aoa-peptide and oligosaccharides	131
6.4	Acetylation of glycopeptides	133
6.5	UV-quantification of glycoprobes	133
6.6	HPLC	134
6.7	MALDI-TOF mass spectrometry.....	134
6.8	NMR spectroscopy	135
6.9	Surface Plasmon Resonance	135
6.9.1	Immobilization of glycoprobes.....	136
6.9.2	Binding and kinetic experiments.....	138
6.9.3	Thermodynamic experiments	141

CONTENTS

6.9.4	Lectin capture experiments.....	141
6.10	CREDEX-MS	142
7	BIBLIOGRAPHY	145
8	APPENDIX	167

1 INTRODUCTION

1.1 The sugar code

Life is possible thanks to the combined action of four major classes of macromolecules (nucleic acids, lipids, proteins and carbohydrates). In this group of actors, each member plays a specific role in the scene. Even though they are composed mainly of carbon, oxygen, hydrogen and nitrogen, their appearances (structures), and consequently their functions, differ from each other. Whereas linear polymers of nucleic acids are responsible for carrying genetic information between generations, linear polypeptides (proteins) participate in almost every process of the cell, *e.g.* enzymes that mediate in metabolism, proteins with structural and mechanical functions and proteins that mediate cell-cell or cell-organism communication. However, branched polymers of lipids and carbohydrates were, until recent years, unfairly and incorrectly treated as “second-class citizens” [1] because their functions were reduced to generate energy (*e.g.* glucose and glycogen) and be structural components (*e.g.* cellulose and chitin). One of the findings that contributed to changing that perception was the work of Winterburn and Phelps, who calculated the theoretical information storage capacity of carbohydrates as the total number of possible isomers and concluded that “the significance of the glycosyl residues is to impart a discrete recognition role on the protein” [2]. Since then, monosaccharides have been considered as code words in their own right, forming a further code system. In comparison to the other two code systems (nucleotides and proteins) that have only one type of linkage and two points of elongation (5′ and 3′ for DNA/RNA, and N- and C-terminus for proteins), each monosaccharide can generate an α or a β linkage to any one of several positions on another monosaccharide in a chain (four different groups plus the anomeric hydroxyl position (Figure 1.1)).

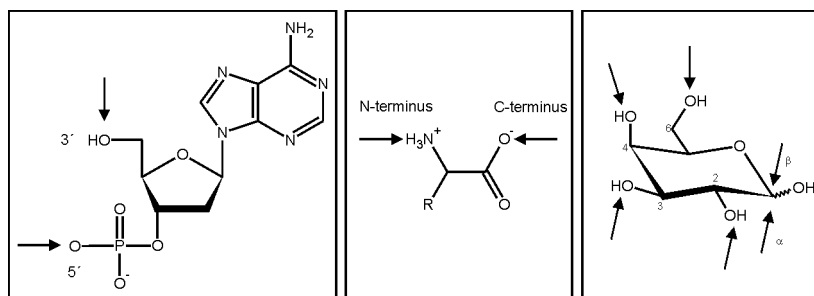


Figure 1.1 Illustration of the three coding systems and their corresponding elongation points (marked by arrows). Nucleotides (left) and amino acids (center) form linear polymers, whereas monosaccharide (right) can be elongated through four hydroxyl groups at C2, C3, and C6 plus the two anomeric hydroxyl positions (α and β).

In addition, further modifications can occur by covalent attachment of sulfate, phosphate, acetyl or methyl groups to the sugar hydroxyl groups. Therefore the theoretical ability to form isomers for oligosaccharides surpasses by far the information-storing capacity of peptides and nucleotides (for an hexasaccharide, there are $1,44 \times 10^{15}$ potential isomers, whereas for a hexapeptide and a hexanucleotide the number of different structures are 6.4×10^7 and 4×10^3 , respectively [3]).

1.1.1 Monosaccharides

Oligosaccharides are formed by conjugation between a number of smallest and non-hydrolyzable units called monosaccharides. These have the common molecular formula $(\text{CH}_2\text{O})_n$ $n=3-9$ and are classified into different groups, based either on the number of carbon atoms (triose, tetraose, pentose, ..., nonose) or on the type of carbonyl group they contain (aldose or ketose). Three main structural features are used to characterize monosaccharides: (1) the D or L configuration which relates to the configurational notation (R or S, respectively) of the chiral centre furthest from the carbonyl group, (2) the anomeric configuration (α - or β -) determined by the relationship (different or identical, respectively),

between the absolute configurations of the anomeric carbon and the chiral center furthest from the carbonyl group (*e.g.*, C5 in hexoses), and (3) the ring conformation (either pyranose or furanose).

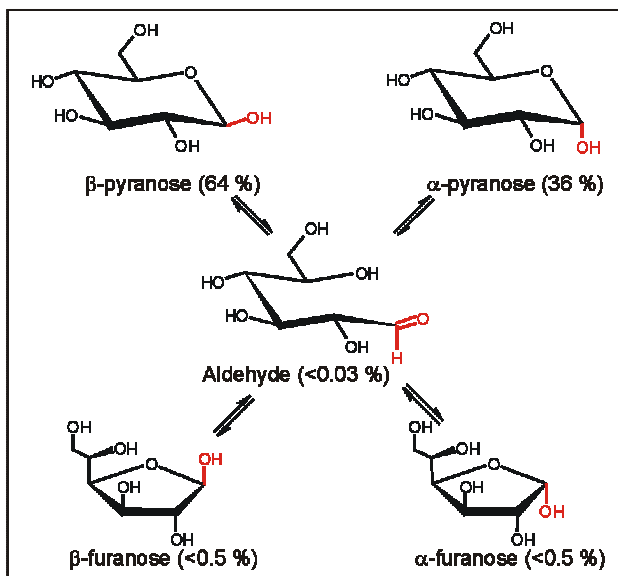


Figure 1.2 Mutarotation scheme of D-glucose.

Thus, a free monosaccharide in solution may exist as an equilibrium between acyclic (carbonyl) and cyclic (hemiacetal: α,β -pyranose and α,β -furanose) structures. The interchange between open and cyclic forms underlies the optical phenomenon of mutarotation, a typical structure attribute of carbohydrates (Figure 1.2). In general, the acyclic forms are only present in trace amounts (<0.03% for glucose), whereas the pyranose ring is the most abundant structure. However, different ratios of anomers are observed for the common hexoses found in nature [4-8]. Moreover, distinct structural orientations are possible in the ring structures. However, pyranoses occur predominantly in chair conformation 4C_1 (opposed to 1C_4 or the envelope conformations), because of energetic reasons.

INTRODUCTION

Table 1.1 Ratio of anomers for the most occurring monosaccharides in mammals.

Monosaccharide	α -pyranose	β -pyranose
D-glucose	36	64
D-galactose	27	73
D-mannose	67	33
<i>N</i> -acetyl-D-glucosamine	61	39
<i>N</i> -acetyl-D-galactosamine	53	47
<i>N</i> -acetylneuraminic acid	<10	>90
L-fucose	30	70

As shown in Figure 1.3, the most common monosaccharides found in higher animals are:

- Neutral sugars: six-carbon monosaccharides such as glucose (Glc), galactose (Gal), mannose (Man) and fucose (Fuc), or five carbon sugars like xylose (Xyl) or ribose (Rib).
- Amino sugars: hexoses with an amino-group at the C2-position such as *N*-acetylglucosamine (GlcNAc) and *N*-acetylgalactosamine (GalNAc).
- Sialic acid: nine-carbon acidic sugars; the most common one is *N*-acetylneuraminic acid (Neu5Ac).
- Uronic acids: hexoses with a carboxylic acid at the C6-position. Glucuronic acid (GlcA) and iudronic acid (IdA) are examples of this group.

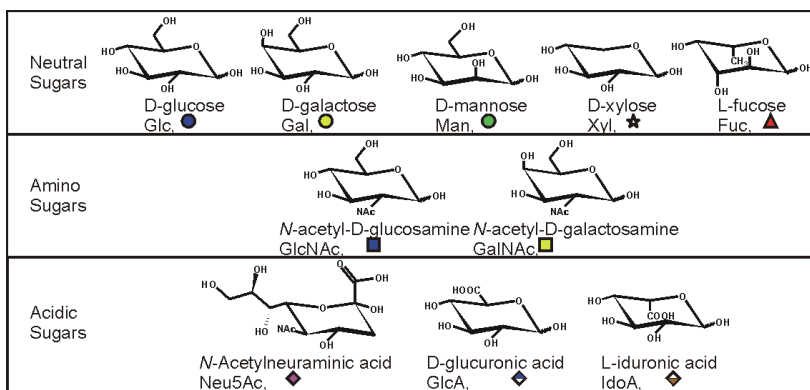


Figure 1.3 Most commonly monosaccharides found in higher animal glycoconjugates are represented using chair conformation. Corresponding abbreviation and symbolic representation accorded by NCBI and Consortium for Functional Glycomics in 2004 are depicted for each monosaccharide.

1.1.2 Glycoconjugates: glycoproteins, proteoglycans and glycolipids

One particular characteristic of carbohydrates, which is not present in other types of biomolecules, is their ability to form branched structures. Every monosaccharide unit has three or four (depending on the number of hydroxyl groups) branching points and therefore complex, branched, structures are formed. Both linear and branched structures can be found either as free-standing entities (*e.g.* polysaccharides), or covalently attached to non-carbohydrate moieties (protein or lipid) (Figure 1.4).

In the mammalian system, over 80% of all (membrane and secretory) proteins are glycosylated. Depending on the linkage between the sugar and the amino acid, the glycan chains are classified in:

- N-glycans: if a β -GlcNAc moiety is linked to an asparagine residue. All N-glycans possess a common trimannosyl core (Man- α 1,6-(Man- α 1,3-)Man- β 1,4-GlcNAc- β 1,4-GlcNAc), to which different monosaccharides are added conforming the three

INTRODUCTION

subtypes of N-glycosylation: high-mannose (with the Man4 and Man4' residues substituted with more Man moieties), hybrid (with the Man4 moiety substituted with a Man and the Man4' substituted with a GlcNAc) or complex (if both mannose residues are substituted with GlcNAc units) (Figure 1.5, Table 1.2). On the GlcNAc residues from hybrid and complex N-glycans, various outer chains, constructed from a series of core portions and modifications (fucosylation, sialylation, galactosylation, glucuronylation, and/or sulfation) of their Gal- β 1,4-GlcNAc units, can be constructed. In addition, further modifications in the trimannosyl core comprise the addition of an α -fucosyl residue linked to the proximal GlcNAc (core-fucose), and the β -GlcNAc linkage to the β -mannosyl residue (bisecting GlcNAc) (Figure 1.5, table 1.2). The vast majority of the membrane and secretory glycoproteins are of the complex-subtype. As an indication of the structural diversity found in N-glycans, nowadays the CarbBank database lists more than 1000 distinct N-glycan structures [9].

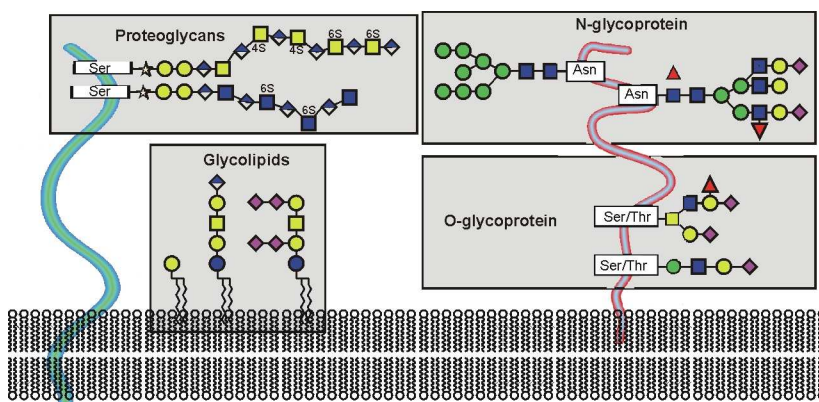


Figure 1.4 Common classes of animal glycoconjugates. The symbol code used is depicted in Figure 1.3.

- O-Glycans: In contrast to the pentasaccharide core in N-glycans, the O-glycosylation cores are structurally simpler (not longer than a trisaccharide), although the monosaccharide unit linked to the Ser/Thr residue is more variable (*e.g.* GalNAc, GlcNAc, Man, Gal, Fuc). However, the most common modification is the attachment of a GalNAc residue to serine or threonine residues. Currently, the O-linked glycans are grouped into seven core subtypes (Figure 1.5, Table 1.3) with the cores 1 and 2 as the most commonly found in mucins and glycoproteins. Further outer modifications, in common with N-glycans, are attached to internal cores, generating a wide structural diversity of glycan epitopes (Table 1.5 and Table 1.6)

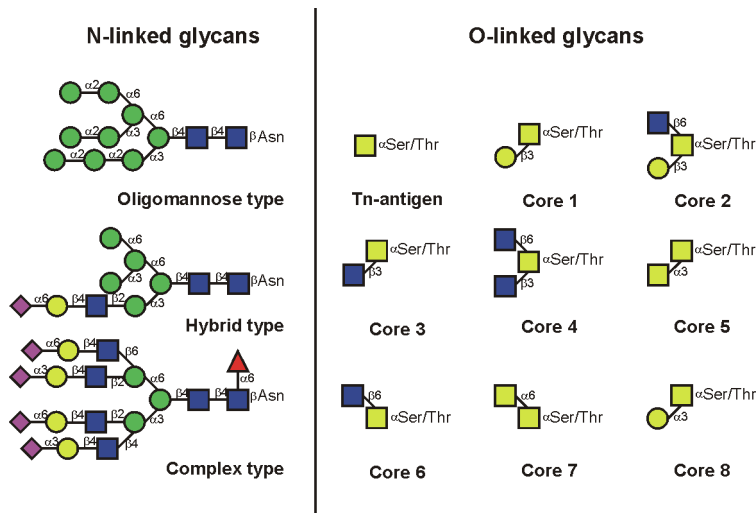


Figure 1.5 Structure of N-glycans and O-glycans.

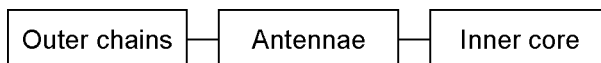
-
- Glycosaminoglycans: Another type of glycosylation found in proteins is the addition of glycosamine glycans to a serine residue. These linear proteoglycans, usually found in the extracellular matrix, are constituted by a core protein with one or more covalently attached linear glycosaminoglycan chains, consisting

in repeating disaccharide units of an amino sugar (either GlcNAc or GalNAc) and an uronic acid (GlcA and IdoA). The most common glycosaminoglycans found in the extracellular matrix are hyaluronan [GlcNAc- β 1,4-GlcA- β 1,3]_n, chondroitin/dermatan [GalNAc- β 1,4-GlcA β 1-3] _n with different levels of sulfation, heparan sulfate [GlcNAc- α 1,4-GlcA- β 1,4]_n and keratan sulfate [Gal- β 1,4-GlcNAc β -1,3]_n (Figure 1.4, Table 1.4). Several biological processes, such as proteolysis control, modulation of angiogenesis during tumor metastasis and the oligomerization of cell growth factors are modulated by sulfated oligosaccharides. Heparin-antithrombin, heparan sulfate-herpes simplex virus are some of the most well-known complexes formed by these interaction partners. The presence of sulfated oligosaccharides in glycosaminoglycans and on N-glycosylated proteins of pathogens (*e.g.* influenza virus, *Trypanosoma cruzi*) make them likely targets for drugs/vaccines/diagnosis developments [10,11].

- Glycophospholipids: Other proteins may employ a glycan bridge to be linked to the cellular membrane via a glycosylphosphatidylinositol (GPI) moiety. As the main biological role of this glycan chain is only the anchoring of protein to the lipid membrane, and it has not been implicated in other processes, this class of glycoproteins will not be further discussed in this chapter.
- Glycosphingolipids: The last class of glycoconjugates found in nature are constituted by glycosphingolipid (often called glycolipid), that is an oligosaccharide usually attached via Glc or Gal to the terminal primary hydroxyl group of the lipid moiety ceramide. This basic unit can be further extended by the linear addition of further monosaccharides (Gal, GlcNAc, and GalNAc) (Figure 1.4, Table 1.4).

1.1.2.1 Relevant sugar epitopes

As glycosylation is a non-template directed process that depends on the availability of donor/acceptor substrate and enzymes, on the migration rate and on environmental factors, the diversity of glycan chains that can be generated on a single glycoconjugate is enormous. However, all glycoconjugates are formed by an inner core which can be modified by the addition of different outer glycan chains.



In N-glycans, the common core region is further elongated by the addition of Man or GlcNAc residues in different linkages (α -1,2, β -1,2, β -1,4, β -1,6) yielding a wide range of structures ranging from mono- to pentaantennary (Table 1.2). By contrast, in O-glycans there are no true antennae built on the inner core, which is generally formed by the attachment of a single GalNAc moiety in α 1- linkage to a serine or threonine residue. This GalNAc α -Ser/Thr determinant, known as Tn antigen, a member of the tumor-associated antigens, constitutes a sensitive and specific biomarker for detection of carcinomas and may be used as therapeutic target for cancer treatment [12]. Different variants of the inner core involve addition of Gal, GlcNAc or GalNAc in β -1,3 and β -1,6 linkage (Table 1.3). The T antigen, or Thomsen-Friedenreich antigen, is represented by the disaccharide Gal- β 1,3-GalNAc. This antigen is used particularly as determinant of malignant carcinomas since, in contrast to normal cells, in most tumor cells it lacks outer chain modifications and is therefore easily detected by specific antibodies [13].

Table 1.2 Inner core structures found in mammalian N-linked glycoconjugates. Currently known names of inner core structures are depicted. The distinctive sugar moiety of each structure is depicted in black, whereas the common precursors are shown in grey.

		Antennae			
Name	Structure	Name	Structure		
N-glycans	Trimannosyl core	High-mannose	Man α 2Man α 6 Man α 2Man α 3 ⁶ Man α 2Man α 3 ⁶ Man β 4GlcNAc β 4GlcNAc-Asn		
		Hybrid-type	Mono- GlcNAc β 2Man α 3 ⁶ Man β 4GlcNAc β 4GlcNAc-Asn		
		Complex-type	Bi-	GlcNAc β 2Man α 6 GlcNAc β 2Man α 3 ⁶ Man β 4GlcNAc β 4GlcNAc-Asn	
			Tri-	2,4	GlcNAc β 2Man α 6 GlcNAc β 4Man α 3 ⁶ Man β 4GlcNAc β 4GlcNAc-Asn
				2,6	GlcNAc β 6Man α 6 GlcNAc β 2Man α 3 ⁶ Man β 4GlcNAc β 4GlcNAc-Asn
		Tetra-	GlcNAc β 6Man α 6 GlcNAc β 2Man α 3 ⁶ Man β 4GlcNAc β 4GlcNAc-Asn		
		Penta-	GlcNAc β 6Man α 6 GlcNAc β 4Man α 3 ⁶ Man β 4GlcNAc β 4GlcNAc-Asn		
		Core fucose	Man α 6Man β 4GlcNAc β 4GlcNAc-Asn Man α 3 ⁶ Fuc α 6		
		Bisecting GlcNAc	GlcNAc β 4Man α 6 Man α 3 ⁶ Man β 4GlcNAc β 4GlcNAc-Asn		

Table 1.3 Inner core structures found in O-linked glycoproteins. The distinctive sugar moiety of each structure is shown in black, whereas the common precursors are coloured in grey.

Inner core	
Name	Structure
Antigen Tn	GalNAc-Ser/Thr
Antigen T, core 1	Gal β 3GalNAc-Ser/Thr
core 2	GlcNAc β 6Gal β 3GalNAc-Ser/Thr
core 3	GlcNAc β 3GalNAc-Ser/Thr
core 4	GlcNAc β 6GlcNAc β 3GalNAc-Ser/Thr
core 5	Gal α 3GalNAc-Ser/Thr
core 6	GlcNAc β 6GalNAc-Ser/Thr
core 7	GalNAc β 6GalNAc-Ser/Thr

By contrast, glycolipids have as inner core structures the disaccharide Gal- β 1,4-Glc or its GlcNAc derivative form GlcNAc β -1,3-Gal- β 1,4-Glc, attached to a ceramide moiety (Table 1.4).

Table 1.4 Inner core structures in glycolipids.

Inner core	
Name	Structure
G _{A3}	Gal β 4Glc-ceramide
Type 4	Gal β 3GalNAc-ceramide
	GlcNAc β 3Gal β 4Glc-ceramide

Inner core	
Name	Structure
Hyaluronan	(GlcNAc β 4GlcA β 3) _n
Chondroitin/Dermatan	(GalNAc β 4GlcA β 3) _n
Heparan	(GlcNAc α 4GlcA β 4) _n
Keratan	(Gal β 4GlcNAc3) _n

Additional glycan structures at “outer” positions of glycan chains are recognized specifically by carbohydrate-binding proteins and therefore known as sugar epitopes, by analogy to the antigen-antibody system (Table 1.5). In hybrid- and complex-type N-glycans, terminal GlcNAc residues can be modified, on the one hand, by addition of a Gal moiety in β -1,3 or β -1,4 linkage yielding type-1 or type-2 chains, respectively. On type-2 chains, glycan elongation with poly-*N*-acetyl-lactosamine structures is described, which can be linear (*e.g.* i-blood antigen) or β 1-6

INTRODUCTION

branched (*e.g.* I-blood antigen). An example of the biological relevance of these poly-*N*-acetyl-lactosamine structures lies in the fact that these epitopes are recognized with high affinity by galectins which are directly implicated in cancer metastasis [14]. On the other hand, the terminal GlcNAc residue can be modified by addition of one GalNAc residue in β -1,4 linkage, generating the family of β -1,4-linked GalNAc epitopes. This specific glycosylation is restricted to only a few cell types, including the pituitary glands that synthesize glycoprotein hormones [15]. Additionally, the GalNAc moiety can be further transformed by modifications such as fucosylation, sulfation and sialylation. A potential function of these GalNAc-epitopes may be the clearance of glycoproteins from plasma, given that the asialoglycoproteinreceptor responsible for such clearance exhibits high affinity for terminal GalNAc [16]. Besides, the core portions act more as substrates for subsequent modifications (see below) than as real epitopes.

Table 1.5 Outer core portions found in mammalian glycoconjugates.

		Name	Structure
Core portions		Type 1	Gal β 3GlcNAc
		Type 2	Gal β 4GlcNAc
		Type 5	Gal β 3Gal
		β 1-6 GlcNAc	GlcNAc β 6Gal
	Blood group II	Linear I	Gal β 3/4GlcNAc β 3Gal β 4GlcNAc β 3Gal β 4GlcNAc
		Branched I	Gal β 4GlcNAc β ₆ Gal β 4GlcNAc β 3Gal β 4GlcNAc
	β 1-4 GalNAc	Terminal	GalNAc β 4GlcNAc
		Fucosylated	GalNAc β 4GlcNAc Fuc α ³
		Sialylated	Sia α 6GalNAc β 4GlcNAc
		Sulfated	SO ₄ GalNAc β 4GlcNAc
		Sd ^a /G _{M3}	GalNAc β 4Gal β 4GlcNAc/Glc Sia α ³

In N- and O-linked glycoproteins and glycolipids, a wide range of structural antigens are built by additional modifications of the core portions. These outer structural epitopes are classified in Table 1.6. The three members of the A, B and H blood group family are elaborated by sequential action of distinct glycosyltransferases on the four precursor structures (type-1, type-2, type-3 and type-4). The H determinant is

characterized by an α Fuc attached to a terminal β -Gal unit in α -1,2 linkage. A and B determinants are formed by subsequent modifications on the H antigen consisting in attachment of a GalNAc or Gal residue (α -1,3)-linked to the terminal β -Gal unit. The A, B and H antigens were first found on the surfaces of red cells and on the epithelia of different organs (*e.g.* digestive, respiratory, urinary and reproductive). However, later on, these epitopes were also found on soluble glycoconjugates as well. Antibodies against the ABO antigens that are not expressed in an individual's red cells constitute the first immune response against bacterial and fungal organisms which expose similar ABO structures. Additionally this innate immune response is responsible for the rejection associated to blood transfusions [17].

As shown in Table 1.6, the Lewis blood group family is comprised by a structural similar set of α -1,3 and α -1,4 fucosylated glycan structures (Le^a , Le^b , Le^x and Le^y). The term Lewis refers to the family name of individuals suffering from blood cell incompatibility that helped to discover this blood group family. Lewis antigens are derived from two precursor structures: the terminal type-1 Gal β 1-3GlcNAc disaccharide in Le^a and Le^b epitopes and the terminal type 2 Gal β 1-4GlcNAc disaccharide in Le^x and Le^y determinants. Le^a epitope is characterized by a Fuc moiety attached to the subterminal GlcNAc in α -1,4 linkage. Additional modifications, such as sialylation and sulfation have been described for this epitope. Antigens from this family play key roles in cell recognition processes, especially in selectin-dependent homing of lymphocytes [18]. On the other hand, Le^b epitope is characterized by the attachment of a second Fuc residue in α -1,2 linkage to terminal Gal. The Le^b antigen may act as a receptor for adhesion of certain microorganisms (*e.g.* *Helicobacter pylori*) to host cell [19]. Le^a and Le^b epitopes are found on soluble glycoproteins secreted by different epithelia (*e.g.* respiratory, digestive,

INTRODUCTION

urinary, reproductive, *etc.*). Red cell membranes acquire glycosphingolipids bearing Lewis blood antigens through passive adsorption. Other members of Le^a and Le^b determinants are found both in glycolipids and soluble glycoproteins. Accumulation of Le^x and Le^y, as well as Le^a antigens has been reported in cancer tissues, suggesting that these Lewis blood group epitopes may play essential roles in tumor metastasis [20].

Another family of tumor-associated antigens is represented by sialylated structures in N-glycans, O-glycans and glycolipids (Table 1.6). Diverse sialyltransferases act on different precursor core portions (type 1-4, type 6) and attach a sialic acid moiety to a terminal Gal residue in α -2,3 linkage. α -2,3 sialic acid linkages may participate in a wide range of biological processes, such as leukocyte homing [21], plasma glycoprotein turn-over [22] or cell adhesion of a variety of pathogens (*e.g.* influenza virus, *Helicobacter pylori*) [23,24]. However, the α -2,6 linkage is not as common as the α -2,3 linkage. α -2,6 linkages are found both on terminal and subterminal Gal as well as on internal GalNAc moiety, and can be further modified by α -2,8 sialylation. Neu5Ac α -2,6-linked structures participate in similar biological processes as the α -2,3 sialic acid linkages, such as pathogen cell adhesion and plasma glycoprotein turn-over [25]. In addition, as ligands of the I-type lectin family they participate in the immune response. A more restricted function is carried out by α -2,8 sialylated structures. They modulate cell-cell and cell-matrix interactions important in neural development [26].

Table 1.6 Outer chain modifications found in mammalian glycoconjugates.

		Name	Structure	
Core portions	Type 1		Gal β 3GlcNAc	
	Type 2		Gal β 4GlcNAc	
	Type 5		Gal β 3Gal	
	β 1-6 GlcNAc		GlcNAc β 6Gal	
	Blood group ii	Linear i	Gal β 3/4GlcNAc β 3Gal β 4GlcNAc β 3Gal β 4GlcNAc	
		Branched I	Gal β 4GlcNAc β 6Gal β 4GlcNAc β 3Gal β 4GlcNAc	
	β 1-4 GalNAc	Terminal	GalNAc β 4GlcNAc	
		Fucosylated	GalNAc β 4GlcNAc Fuc α 3	
		Sialylated	Sia α 6GalNAc β 4GlcNAc	
		Sulfated	SO $_4$ GalNAc β 4GlcNAc	
Sd'/G $_{M3}$		GalNAc β 4Gal β 4GlcNAc/Glc Sia α 3		
Modifications	A, B, and H blood group	H	Gal β 3(4)GlcNAc/GalNAc Fuc α 2	
		A	A $_1$	GalNAc α 3Gal β 3(4)GlcNAc/GalNAc
			A	GalNAc α 3Gal β 3(4)GlcNAc/GalNAc Fuc α 2
			A $_2$	Gal β 3GalNAc α 3Gal β 4GlcNAc Fuc α 2
			A $_3$	GalNAc α 3Gal β 3GalNAc α 3Gal β 4GlcNAc Fuc α 2
		B	B $_1$	Gal α 3Gal β 3(4)GlcNAc/GalNAc
	B		Gal α 3Gal β 3(4)GlcNAc/GalNAc Fuc α 2	
	Lewis Blood group	Le a	Le a	Fuc α 4 Gal β 3GlcNAc
			Sialyl-Le a	Fuc α 4 Sia α 3Gal β 3GlcNAc
			3sulfo-Le a	Fuc α 4 SO $_3$ Gal β 3GlcNAc
		Le x	Le b	Fuc α 4 Fuc α 2Gal β 3GlcNAc
			Le c	Gal β 4GlcNAc Fuc α 3
			3sulfo Le c	SO $_3$ Gal β 4GlcNAc Fuc α 3
			Sialyl Le c	Sia α 3Gal β 4GlcNAc Fuc α 3
			6sulfo sialyl Le c	SO $_6$ Sia α 3Gal β 4GlcNAc Fuc α 3
	Le y	Fuc α 2Gal β 4GlcNAc Fuc α 3		
	P Blood group	P'	Gal α 4Gal β 4Glc	
		P	GalNAc β 3Gal α 4Gal β 4Glc	
		P $_1$	Gal α 4Gal β 4GlcNAc β 3Gal β 4Glc	
	Forssman group	F $_1$	GalNAc α 3GalNAc α	
		F $_2$	GalNAc α 3GalNAc β	
	Gallii epitope		Gal α 3Gal β 4GlcNAc	
Sialylated structures group	α -2-3		Sia α 3Gal β 3GlcNAc	
			Sia α 3Gal β 4GlcNAc	
			Sia α 3Gal β 3GalNAc α	
			Sia α 3Gal β 3GalNAc β	
			Sia α 3Gal β 4Glc	
	α -2-6		Sia α 6Gal β 4GlcNAc	
			Sia α 6GalNAc	
			Sia α 6 Sia α 3Gal β 3GalNAc	
α -2-8		Sia α 8Sia α 3Gal β 4GlcNAc		
		Sia α 8Sia α 3Gal β 3GalNAc β		
Sulfated structures group	HNK-1		SO $_3$ GlcA β 3Gal β 4GlcNAc	
			SO $_4$ GalNAc β 4GlcNAc	
	Keratan sulfate		SO $_6$ SO $_6$ SO $_6$ Gal β 4GlcNAc β 3Gal β 4GlcNAc	
			(SO $_6$ GlcNAc α 4IdoA β 4)	
	Chondroitin sulfate		(SO $_4$ /6GalNAc β 4IdoA β 3)	

INTRODUCTION

P blood antigens are found only in membrane-associated glycolipids on red cells and other tissues. All of them start on the lactosyl-ceramide precursor, which is transformed by addition of either Gal in α -1,4 linkage (Pk antigen) or GalNAc in β -1,3-linkage (P antigen). However, for P₁ antigen synthesis, three sequential glycosylation reactions are needed (Table 1.6). It is reported that the blood group P antigens act as receptor for various uropathogenic bacteria (*e.g. Escherichia coli, Pseudomonas aeruginosa*) and virus (*e.g. members of the parvovirus family*) [27].

Another antigen found in glycolipids is the Forssman antigen, formed by addition of GalNAc in a α -1,3 linkage to a terminal GalNAc moiety (Table 1.6). This epitope, synthesized in many different mammals, has been found in tumor but not in normal human cells. Because of its localization mainly in tight junctions and apical/basal membranes, it is assumed that it may participate in cell-cell adhesion and other communication processes [28].

Finally, another antigen not synthesized in humans is the Galili epitope. In non-primate mammals, this epitope is synthesized by a specific α -1,3 galactosyltransferase that acts on type-2 precursor, present in glycoproteins and glycolipids. Due to their lack of Galili antigens, humans display high titers of anti-Galili antibodies. This constitutes a major barrier to the use of xenotransplants in humans, as the Galili antigens of the transplanted cells cause a hyperacute response culminating in cell lysis [29]. On the other hand, anti-Galili antibodies have proven quite useful for diminishing the infection efficacy of some recombinant retroviruses [30].

1.2 Lectins as decipherers of the sugar code

The wealth of biological information encoded by the complex glycan structures found in glycoconjugates requires a code-reader capable of differentiating carbohydrate epitopes acting as recognition determinants in diverse biological communication processes, such as fertilization, microbial infection, immune response and metastasis, among others. This unusual ability to decode the information contained in carbohydrates is present in lectins, a class of proteins discovered during the last century.

1.2.1 History of lectins

The history of lectinology started in 1888, when Stillmark described for the first time the toxicity and agglutination ability of seed extracts of four *Euphorbiacea* plants (*e.g. Ricinus communis*) [31]. Subsequently, a similar toxic hemagglutinin activity was found in extracts from *Abrus precatorius* by Helli [32]. Both extracts were employed in 1891 by Ehrlich to establish the fundamental principles of immunology (specificity, immunological memory and transfer of humoral immunology from a mother to her offspring). Due to the slow progress in the development of fractionation techniques, the first isolated hemagglutinin was not described until 1919, when Sumner and Howell isolated concanavalin A and reported its crystal structure [33]. In 1936, as a consequence of the discovery of the carbohydrate nature of the ligand for *Concanavalin A*, it was suggested that the agglutinating activity might be a consequence of a reaction between the plant protein and carbohydrates on the surface of red cells. Although in 1900 Stillmark had already observed different relative hemagglutinating activities of various seed extracts for red blood cells from different species, it was not until 1940 that the human ABO blood group system was established by Boyd and Renkonen. The differential hemagglutination activity by lectins, such as

INTRODUCTION

Phaseolus limensis, *Vicia cracca* and *Lotus tetragonolobus*, in combination with further inhibition studies by monosaccharides, provided new insights into the chemical nature of the ABO and Lewis blood group antigens. Finally, in 1952, Morgan and Watkins described three types of blood cells, depending on the specific lectin that induces their agglutination and the monosaccharide that inhibits the reaction. Whereas A type red cells were agglutinated by *Phaseolus limensis* and inhibited by α -GalNAc, B type cells and O type cells were agglutinated by *Vicia cracca* and *Lotus tetragonolobus* and inhibited by β -Gal and α -Fuc, respectively. This ability of plant hemagglutinins to distinguish between red blood cells led Boyd and Shapleigh in 1954 to propose the name lectins, from the Latin *legere* that means to pick out or choose [34]. In order to differentiate lectins, which participate exclusively in communication processes, from other carbohydrate-binding proteins such as antibodies, receptors and enzymes, it was necessary to refine the definition. Thus today, to qualify as a lectin, a protein must meet three requirements: (1) it must be a (glyco)protein of non-immune origin, (2) binding carbohydrates, and (3) not modifying the carbohydrates (such as glycosyltransferases or glycosidases), nor transporting them through membranes. Three more discoveries, in the 1960s, brought lectins into the limelight. Nowell found that the lectin of *Phaseolus vulgaris* and *concanavalin A* promoted lymphocyte proliferation [35]. Other lectins such as wheat germ agglutinin were found to agglutinate preferentially malignant cells. These and similar findings demonstrated the usefulness of lectins as investigational tools for different biological events, particularly cancer-related processes.

Two main technical improvements facilitated lectin studies. First, the introduction of affinity chromatography for isolation of lectins by Goldstein and Agrawal in 1965 increased the number of isolated lectins,

from 1 to more than 500 in 2004 [32]. Aside from plants, lectins were found in other sources such as microorganisms (*e.g.* influenza virus hemagglutinin [36]) or animals (*e.g.* galactose-specific hepatic asialoglycoprotein receptor [37]). The isolation and characterization in 1974 of this mammalian lectin highlighted the role of animal lectins in biological recognition. Another discovery that helped to understand lectins as recognition molecules was the isolation by Sly in 1977 of the mannose-6-phosphate receptor involved in intracellular trafficking of lysosomal enzymes [38]. Since then, several animal lectins have been isolated predominantly through methods based on synthetic neoglycoconjugates. Apart from their role in intracellular glycoprotein transport, lectins have been reported to be involved in cell routing, cell-cell and cell-matrix interactions, especially during immune response. Particularly interesting is the finding of a role for lectins in tumor metastasis (*e.g.* galectins) [39-42]. Finally, a particularly powerful insight on the structural determinants of lectin-ligand complexes has been achieved by means of NMR spectroscopy studies in solution, leading to detailed molecular characterization of carbohydrate-binding sites [43].

1.2.2 Classification of lectins

Given that lectins were ubiquitous in nature, the original classification was based on their origin (*e.g.*, plants, animals and microorganisms). Among organisms, lectins vary in size, amino acid composition, metal requirement, domain organization, subunit number and assembly, as well as in their three dimensional structure, and in the constitution of their carbohydrate-recognition sites. It was Mäkela who in 1957 suggested that lectins can be classified in four different classes, based on the configuration at C-3 and C-4 of the monosaccharide that best inhibited cell-agglutination. Thus, lectins that bound Man and Glc belonged to

INTRODUCTION

group III; Gal and GalNAc-binding lectins were classified as group II lectins and Fuc-binding lectins were placed in group I [44].

Recently, it has been found that in general lectins show higher affinity for oligosaccharides (di-, tri- and tetrasaccharides) than for monosaccharides (K_D mM vs μ M) and, as the crystal structure of several lectins has been elucidated, a new classification has been created based on common structural features [45]. Thus, lectins can be conveniently classified as follows:

-simple lectins, consisting of a small number of subunits (not more than four), with molecular weights below 40 kDa. Practically all known plant lectins and some members of the galectin family belong to this group.

-mosaic lectins, composed by several kinds of protein domains, only one of them with carbohydrate-binding ability. Viral hemagglutinins and some animal lectins are examples of this type.

-macromolecular lectins, consisting in polymeric subunits helically arranged and assembled in the form of *fimbriae*, very common in bacteria.

1.2.3 Simple lectins: Plant lectins

Within this class, plant lectins are the largest and most thoroughly studied family, with more than 900 identified members in 2006 [46]. The best-characterized family is the *Leguminosae*. Typically, legume lectins consist of two or four subunits of 25-30 kDa, made up of a single polypeptide chain of about 250 amino acids and with one or two N-linked oligosaccharides. They are synthesized in ribosomes as precursor peptides of approximately 30 kDa, enter the lumen of the RE and are further processed through the Golgi apparatus. Several post-translational modifications affect the initial peptide precursor. Into the lumen of the RE, the N-terminal signal peptide of about 20 to 30 amino acids is split off, and subsequently modified by the attachment of one or two N-linked

oligosaccharide to the consensus sequence Asn-X-Ser/Thr [47]. Finally, legume lectins end up mostly in storage vacuoles (or protein bodies), where they can undergo further proteolysis and transpeptidation reactions. The storage organs of plants (seeds, roots, bulbs, bark and leaves) are the richest sources of legume lectins.

Within the family, remarkable sequence homologies are found, with about 20 % of invariant amino acids, and close to 20% similar ones. Most of the invariant amino acids constitute the unique carbohydrate binding site per subunit. Despite this high sequence homology, the legume family displays enormous diversity in carbohydrate specificity. Specific lectins towards almost every monosaccharide present in mammals (Glc, Gal, Man, Fuc, GlcNAc, GalNAc, Neu5Ac), have been isolated from natural sources and characterized. Examples are shown in Table 1.7.

Table 1.7 Examples of plant lectins and their carbohydrate specificity.

Monosaccharide	Plant lectin source	Abreviation
Glc/Man	<i>Canavalia ensiformis</i>	Con A
Gal	<i>Erythrina cristagalli</i>	ECA
	<i>Ricinus communis</i>	RCA
	<i>Griffonia simplicifolia</i>	GSA
GlcNAc	<i>Triticum vulgare</i>	WGA
GalNAc	<i>Dolichos biflorus</i>	DBA
	<i>Psophocarpus tetragonolobus</i>	PTA
Neu5Ac	<i>Maackia amurensis</i>	MAA
	<i>Sambucus nigra</i>	SNA
Fuc	<i>Lotus tetragonolobus</i>	LTA
	<i>Ulex europeus</i>	UEA

INTRODUCTION

Another particular aspect of the legume family in contrast to other families (*Gramineae* and *Solanaceae*), is its requirement for a tightly bound Ca^{2+} and a transition metal, usually Mn^{2+} . However, it has been described that in plant lectins, the ions are not directly involved in the interaction but responsible of maintaining subunit integrity and of fixing residues for ligand binding [48].

Plant lectins appear to have many different, important functions [49]. They may act as protection against animal predators, such as lectins from *Ricinus communis* and *Phaseolus vulgaris*, which are toxic for mammals in general, or *Galanthus nivalis* (GNA) lectin and WGA, which are toxic for insects and fungi, respectively. Aside from their role in plant defense, lectins also participate in beneficial interactions such as the symbiotic contact with bacteria from the genus *Rhizobium* that allows plants to fix atmospheric nitrogen. In addition to these external functions, legume lectins can also act within the plant as seed storage proteins and modulators of enzyme activity. The large number of saccharides with which legume lectins can interact has led to their extensive use for analytical purposes in the laboratory. Blood typing, cell separation, definition of the glycosylation status of a cell, mapping of neuronal pathways, purification and identification of glycoconjugates are some of the most common applications. Several plant lectins are commonly used in glycobiology research (Table 1.7) [50].

More recently, plant lectins have also been used in medical fields, as mitogenic agents of immune cells [51], diagnostic agents or due to its toxicity as antineoplastic drugs [52]. Lectins such as Con A and WGA are nowadays in clinical trials to combat cancer. However, due to their plant origin and strong cytotoxic properties, immune response and other side-effects are expected, suggesting that considerable efforts will be required if their useful application as anti-cancer drugs is to become a reality.

1.2.4 Lectins in other organisms

Although lectins were first found in plants more than 100 years ago, they are now known to be present throughout nature, from bacteria and viruses to animals. In the microbial world, both bacterial and viral lectins are involved in adhesion to epithelial cells of the respiratory and gastrointestinal tract, as starting point of infection processes. Bacterial adhesins bind preferentially to glycans (Gal- β 1,4-Glc, Gal- α 1,4-Gal) found in glycolipids, although other adhesins have been described that bind specifically to proteoglycans such as heparan sulfate [53]. Viral hemagglutinins, on the other hand, bind mostly to glycans present in glycoproteins. The best studied example is the influenza hemagglutinin which binds to sialic acid-containing glycans, with Neu5Ac- α 2,(3,6)-Gal as the minimal epitope necessary for binding.

More diverse carbohydrate specificities are found in animal lectins. These are classified according to structure in five main families:

- C-type, composed by over 50 members with highly variable carbohydrate specificity and characterized by the strict Ca^{2+} requirement for binding. The mammalian hepatic asialoglycoprotein receptor (Gal/GalNAc-specific), the mannose binding protein (Man-specific) and the selectins (Neu5Ac- α 2,3-Gal- β 1,(3,4)[Fuc- α 1,(3,4)-GlcNAc]-specific) are some examples of this family.
- S-type, with fifteen different galectins identified to date. All of them β -galactose-specific and appearing to be mainly involved in metastatic processes.
- P-type, hitherto with a single member, the mannose 6-phosphate receptor that participates in targeting lysosomal enzymes to their subcellular compartments.

- I-type, consisting of a group of structurally immunoglobulin-related proteins. The Singlec family of sialic acid-specific lectins is the unique well-characterized group of I-type lectins that participate in cell-cell interactions during the immune response.
- Other lectins such as calnexin and calreticulin that participate in the correct folding of nascent glycoproteins, as well as in their intracellular trafficking and delivery to the cell surface.

1.3 Carbohydrate-lectin interactions

1.3.1 Structural basis of carbohydrate-lectin interactions

Structural information on carbohydrate-lectin interactions has been mainly obtained by X-ray crystallography or NMR. Several complexes have been elucidated and their common feature is a glycan-binding site formed by a shallow depression on the surface of the protein (~ 2.5 sugar residues) [45,48]. As depicted in Figure 1.6, hydrogen and metal coordination bonds in combination with hydrophobic and Van der Waal's interactions are involved in the recognition event taking place in the primary binding site. On the monosaccharide side, key hydroxyl groups can act as both acceptors and donors of hydrogen bonds, often mediated by water molecules. In particular, for lectins that recognize neutral hexoses, such as Glc, Man or Gal, the hydroxyl at the C4 position seems to be a decisive player, as differences in the spatial orientation of C4 (axial in Gal, equatorial in Glc and Man) cause significant changes in recognition [54]. Hydrogen bonds are formed between carbohydrate hydroxyl groups and amine groups or hydroxyl of the lectin. A characteristic bidentate hydrogen bond is often formed between two adjacent carbohydrate hydroxyls and the two oxygens of aspartic or glutamic acid residues [55] (Figure 1.6). Although carbohydrates have a highly polar nature, the steric disposition of hydroxyl groups creates hydrophobic patches that can

contact with hydrophobic amino acids, such as phenylalanine, tyrosine, or tryptophan, on the protein surface [45]. Moreover, numerous Van der Waal's interactions contribute significantly to binding. Metal ions, such as Ca^{2+} and Mn^{2+} are situated in close proximity to the binding site ($\sim 4 \text{ \AA}$ apart), but, except for C-type lectins, are not always directly involved in the carbohydrate binding. Their function consists in positioning correctly the amino acid residue interacting with the carbohydrate [56]. Taking into account all the forces involved in carbohydrate-lectin binding, the result is still a relatively weak interaction, with dissociation constants usually in the millimolar range for monosaccharides. Higher affinities towards di- and trisaccharides (in the micromolar range) are possible through secondary binding sites that fit extra monosaccharides. In addition, one or several hydrophobic binding site other than the carbohydrate-binding one and involving protein-protein (*vs.* protein-carbohydrate) interactions have been also reported for lectins [57].

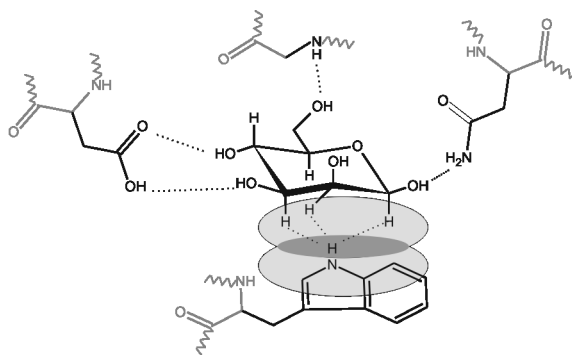


Figure 1.6 Schematic representation of carbohydrate-lectin interactions. Hydrogen bonds between sugar hydroxyl groups and side chains as well as with backbone amide groups are depicted. Hydrophobic interactions (shaded areas) occur between a sugar B-face rich in C-H bonds and aromatic residues of the lectin. Wavy lines depict the rest of the protein backbone.

1.3.2 Glycoside cluster effect

In order to be biologically relevant, carbohydrate-lectin interactions need to improve their binding affinity through multivalency. This enhancement

INTRODUCTION

due to multivalent interaction is called “glycoside cluster effect” [58] and invariably involves an increased binding affinity relative to that of the monomeric ligand.

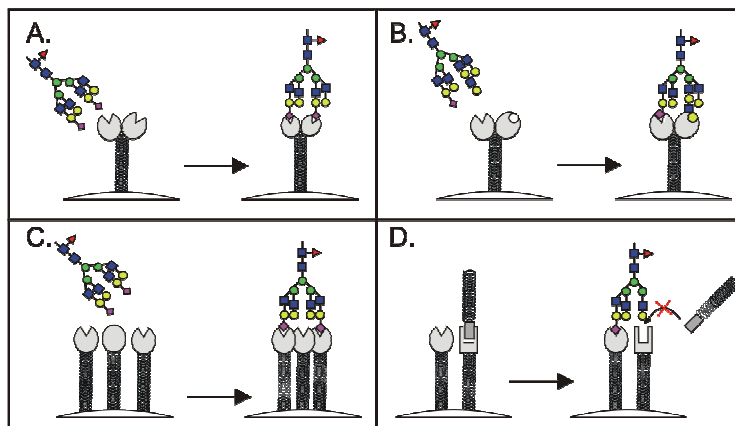


Figure 1.7 Schematic representation of different mechanisms responsible for glycoside cluster effects: (A) intramolecular binding between a lectin with two CRDs and a multivalent homogeneous glycan chain; (B) heterocluster effect involving a lectin with different CRDs and a multivalent heterogeneous glycan chain; (C) intermolecular binding between several single CRD lectins and a multivalent homogeneous glycan; (D) lectin-glycan interaction sterically occluding a protein-protein interaction which otherwise would destabilize the original lectin-glycan binding event.

Lunquist and Toone consider three mechanisms for multivalency: (1) intramolecular or chelate binding, (2) intermolecular aggregative process, and (3) steric stabilization [59]. An intramolecular binding event (“intramolecular” refers to the lectin moiety) involves the clustering of sugar epitopes from a glycan chain in concert with the clustering of carbohydrate recognition units (CRD) from a single lectin (Figure 1.7, panel A). A variant of this binding pattern is the so called “heterocluster effect” [60], whereby mixed-type oligosaccharides –*i.e.*, with different terminal epitopes– are recognized by lectins endowed with two binding sites of different sugar specificity (Figure 1.7, panel B); as a result, binding affinities higher than those of homogeneous conjugates with

identical valency [61] are observed. On the other hand, in lectins with only one CRD per subunit (*e.g.* viral hemagglutinins) the clustering effect is necessarily intermolecular [62] (Figure 1.7, panel C). Finally, the steric stabilization refers to the fact that carbohydrate binding near the protein surface may prevent the approach of other macromolecules, hence the contact with other proteins that would disturb the carbohydrate-protein interaction under primary consideration (Figure 1.7, panel D)

1.3.3 Analytical methods for studying carbohydrate-lectin interactions: information about specificity, structure, kinetics and thermodynamics

There are many different ways to study binding of glycans to lectins, each with its advantages and disadvantages in terms of amount of sample (both protein and glycan) needed, purity required, type and speed of analysis. These approaches are classified into three categories: (1) indirect methods such as inhibition of hemagglutination (HIA) and enzyme-linked lectin assay (ELLA), (2) structural methods such as X-ray crystallography and nuclear magnetic resonance (NMR), and (3) other biophysical methods such as isothermal titration calorimetry (ITC), crystal quartz microbalance (QCM) and surface plasmon resonance (SPR).

Historically, carbohydrate-protein interaction affinities have been evaluated by HIA. The test is based on the ability of a soluble hapten to block the lectin ability of aggregating blood cells. The concentration of hapten causing 50% inhibition of precipitation is taken as the inhibitory concentration, IC_{50} . As absolute IC_{50} values can vary widely among experiments, HIAs are mainly used to compare binding affinities of different soluble ligands, but actual binding affinities are best determined by other methods [63]. Another technique for determine relative binding affinities of a panel of glycoconjugates, ELLA, relies on competitive

INTRODUCTION

binding to a horseradish peroxidase-conjugated lectin between soluble and immobilized carbohydrates. Similar to some ELISA formats, enzyme-conjugated lectin and soluble ligand are added to microtiter plate wells coated with a polymeric saccharide and, after a washing step, the amount of retained lectin is determined spectrophotometrically and from this the corresponding IC_{50} is derived. ELLA avoids HIA-associated problems such as aggregation; however, it has its own disadvantages such as non-specific binding to the microtiter plate, the requirement for lectin-enzyme conjugates, and reduced reproducibility of IC_{50} values due to glycoprotein microheterogeneity. This latter problem can be avoided by covalent immobilization of defined glycoconjugates instead of complex structural glycoproteins [64].

To study carbohydrate-lectin interaction at the molecular level, two main biostructural methods are used: X-ray crystallography and NMR. Originally, carbohydrate-lectin complexes have been elucidated by the former technique, since the often high molecular weight of lectins (> 30 kDa) hampers their analysis in solution by NMR. As of this writing (April 2010), around 200 different crystal carbohydrate-lectin complexes are listed in the RCSB Protein Data Bank (PDB). In order to identify the position and mode of binding of glycans, a resolution of at least 2-2.5 Å is required which is not easy to obtain. Lacking this, carbohydrate-binding sites can be figured out by computer-assisted molecular modeling techniques such as docking [65]. As an alternative to X-ray 3D crystal structure, NMR not only provides structural information but also allows to determine equilibrium association constants [66] by 1D 1H NMR titration experiments where chemical shift variation induced by increasing amounts of ligand are monitored [67]. As X-ray crystallography, NMR has the main disadvantage of requiring substantial amounts of highly pure

sample (both lectin and carbohydrate), often difficult to obtain from natural sources.

Thermodynamic parameters of carbohydrate-lectin interactions are usually determined by means of ITC, QCM and SPR, with the latter two techniques providing also kinetic data. ITC evaluates the change in free energy resulting from binding of a glycan to a lectin. Most ITC instruments work with a continuous power compensation design, where the heat evolved from the binding is determined by measuring the compensating voltage necessary to return the cell to its initial temperature. Control experiments are done in which glycans are added to cells lacking lectin. ITC provides information on binding constant, enthalpy and stoichiometry of binding. In practice, ITC poses problems such as high sample amount (> 10 mg), plus solubility problems when low-affinity interactions are studied, as high lectin concentrations are needed for reliable data. Microcalorimetry, a recent improvement, still requires amounts in the micromolar and millimolar range for lectin and ligand, respectively [68]. These problems notwithstanding, the thermodynamics of several carbohydrate-lectin interactions has been elucidated by means of ITC [69].

To avoid some of the above limitations associated with ITC, two surface-based analytical techniques (SPR, QCM) have been applied, both with significantly lower sample requirements and with the added benefit of allowing to determine kinetics parameters [70,71]. Both techniques are based on the immobilization of one molecule on a surface whilst the interacting partner is passed through as a free molecule. In both techniques binding is detected in real time by means of the difference in mass occurring at the surface upon formation of interacting complexes; detection, however, is based on different physical principles. QCM has a piezoelectric detection mode that measures the frequency shift resulting

from the change in mass on the surface. In addition, changes in viscosity are also measured to help analysis. SPR uses an optical phenomenon to detect refractive index changes, which are proportional to the mass change. In SPR, both association and dissociation rate constants can be determined separately, hitherto the only technique that allows so. In addition, thermodynamic parameters, such as enthalpies and entropies, can be determined by calculating the temperature dependence of the binding constants [72].

1.4 Surface Plasmon resonance (SPR)

In 1983 Liedberg *et al* described the first application of SPR-based sensors to monitor biomolecular interactions [73]. Although there are several SPR-based systems, the technology developed by BIAcore (GE Healthcare, United Kingdom) is the most widely used [74]. The basis of the SPR-biosensor relies on the functionalization of a gold surface with one component of the interaction (called ligand), whereas the other (called analyte) flows over the surface free in solution. To immobilize biomolecules, the gold layer is coated with a matrix of carboxymethylated dextran that provides a three-dimensional hydrophilic matrix, where biomolecules can be immobilized covalently by using different well-characterized chemistries. SPR-based instruments have made it possible to obtain kinetic and thermodynamic data for a large number of protein-protein, protein-peptide and protein-DNA systems. As this technique has been employed in this thesis to study carbohydrate-protein interactions, its principle will be elaborated further.

1.4.1 The SPR principle

When a beam of light passes, at a specific angle of incidence θ_{spr} , from a material with a high refractive index (*e.g.* glass $\mu \sim 1.26$ - 1.38) into a material with a low refractive index (*e.g.* aqueous buffer $\mu \sim 1$), total

internal reflection occurs. However, if the surface of the glass is coated with a thin gold film, the internal reflection is not total, and some light photons are transferred into the metallic film, exciting its electrons and generating surface plasmon waves. This phenomenon is called surface plasmon resonance and, since it entails a loss of energy, it is characterized by a dip in the intensity of the reflected light (Figure 1.8, panel A). As binding of biomolecules results in an increase in the refractive index and θ_{spr} is sensitive to changes in refractive index, binding events are detected as a shift in the θ_{spr} (Figure 1.8, panel B). In SPR-based instruments both the θ_{spr} , and dips are monitored in real time. Therefore, as depicted in (Figure 1.8, panel C), the result of an interaction between two biomolecules can be plotted in a sensorgram as changes in θ_{spr} in resonance units or response units (RU) *vs* time. The binding response in RU (1 RU is equivalent to a shift in θ_{spr} of 10^{-4} degrees) depends on the bound analyte and its molecular mass. Empirical measurements have shown that binding of 1 pg protein to the sensor surface leads to a response of 1 RU. Glycoproteins and lipoproteins have lower specific refractive index increments, whereas nucleic acids have a slightly higher specific refractive index increment than proteins [75]

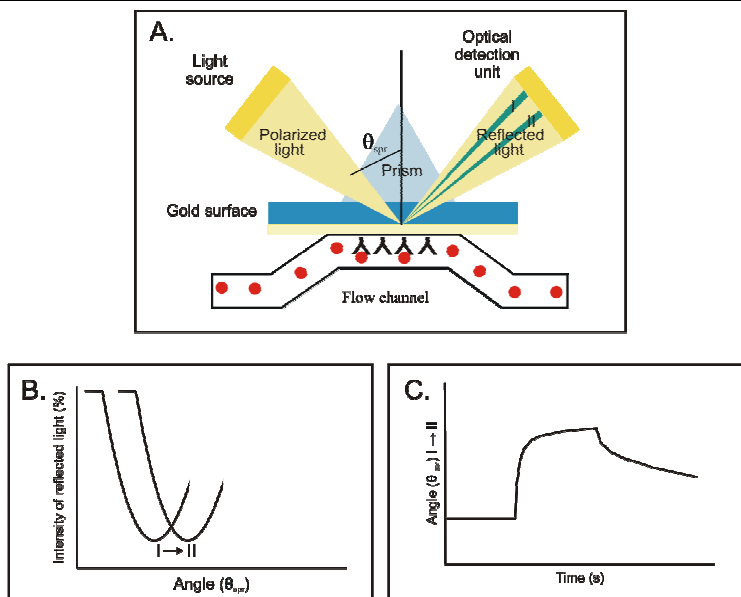


Figure 1.8 (A) Schematic representation of the principle of the SPR technology with a prism as one medium and a solution of analyte (●) as the other. The ligand (▲) is immobilized on the gold surface in contact with the solution. (B) Analyte-binding is measured by monitoring changes in the surface plasmon resonance angle (θ_{spr}), at which the intensity of reflected light reaches a dip. (C) In a sensorgram, the angle at which the dip is observed is plotted vs time.

As shown in Figure 1.9, four different phases can be distinguished in a sensorgram:

- 1) Baseline phase, when only buffer flows across
- 2) Association phase, when the analyte-containing sample is in contact with the sensor and binding occurs
- 3) Dissociation phase, when only buffer flows across and dissociation of the bound analyte occurs
- 4) Regeneration phase, when a solution is injected, which breaks the specific binding between ligand and analyte, and let the surface ready for the next interaction experiment.

Suitable analysis of the data displayed by a sensorgram, during association and dissociation steps, provides quantitative information on active concentration of molecule in a sample, as well as on the affinity and kinetics of the interaction.

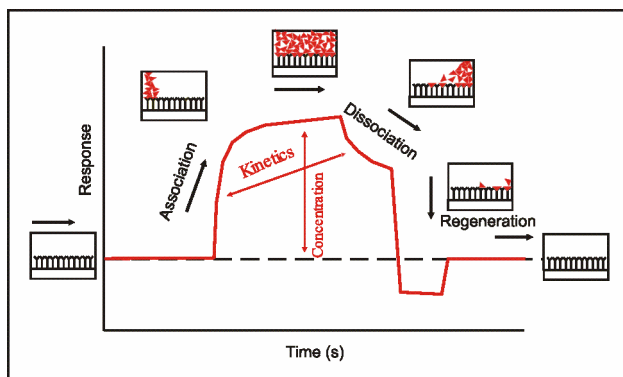
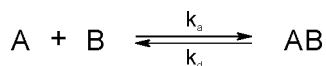


Figure 1.9 A representative sensorgram showing the four steps of an interaction experiment. In the baseline phase, only buffer flows across. In the association phase, the analyte in running buffer is flown over the surface. In the dissociation phase, the running buffer alone is passed over the surface. Finally, a regeneration step is applied to recover the initial surface before running a new cycle. Whereas the shape of the curve gives information about kinetics, the height of the curve quantifies the binding capacity. Reproduced from the website of Biacore AB at www.biacore.com.

1.4.2 Principal applications: kinetics, affinity, and thermodynamics

The best known benefit of direct detection using SPR is the determination of kinetics of biomolecular interactions. In theory, a simple 1:1 interaction between the analyte (A) and the immobilized ligand (B) is described by:



As it is considered that the reaction follows a pseudo-first order kinetics, the formation of the complex as a function of time can be expressed by equation 1.1 where the association rate constant k_a , expressed in $M^{-1}\cdot s^{-1}$ units, describes the rate of complex formation, *i.e.* the number of AB

INTRODUCTION

complexes formed per second in a 1 M solution of A and B. The dissociation rate constant k_d , expressed in s^{-1} units, describes the stability of the complex, *i.e.* the fraction of complexes that decays per second.

$$d[AB]/dt = k_a [A] [B] - k_d [AB] \quad (1.1)$$

Typically, duplicates of a two-fold dilution series of analyte concentration ranging from $8 \cdot K_D$ to $0.25 \cdot K_D$ are injected onto surfaces with different immobilization levels and at different flow rates. Finally, the experimental data are fitted to different pre-programmed kinetic models in order to determine the rate constants of the interaction by using different commercial available software packages (*e.g.* BIAevaluation, CLAMP99, Scrubber).

Additionally, it is very common to describe the strength of an interaction by its affinity, which can be expressed as affinity constant (K_A) and is the ratio of $[AB]$ *versus* $[A]$ and $[B]$ at equilibrium (equation 1.2), or alternatively as the dissociation constant (K_D), which is simply the inverse of the K_A (equation 1.3). Thus, a high K_A describes an interaction with a high affinity to associate, whereas a high K_D means low stability of the complex AB. Equilibrium is reached when the rates of the association and dissociation reactions are equal.

$$K_A = [AB] / [A][B] \quad (1.2) \quad K_D = [A][B] / [AB] \quad (1.3)$$

In SPR, the affinity constant can be measured directly by steady state binding experiment. Typically, duplicate series of analyte concentrations varied over 2-3 orders of magnitude between $10 \cdot K_D$ to $100 \cdot K_D$ are injected and the level of binding in equilibrium (R_{eq}) is measured. It is important to ensure that the analyte injections reach equilibrium. Therefore, an approximate time is calculated by equation 1.4, when the analyte concentration is equal to the K_D .

$$t_{0.9} \approx 1/k_d \quad (1.4)$$

In steady state (or equilibrium), the affinity constant is related to the concentration and binding level according to the equation 1.5.

$$R_{eq} = (K_A * C * R_{max}) / (1 + K_A * C * n) \quad (1.5)$$

where n specifies the number of binding sites per analyte molecule. R_{max} is the theoretical binding capacity. Thus, the affinity constant K_A and R_{max} can be calculated by plotting R_{eq} against C and fitting it to a kinetic model using standard BIAevaluation software. Additionally, under saturation conditions, if it is assumed that all the ligand is oriented correctly and 100% active, the stoichiometry of the interaction (n) can be calculated by equation 1.6.

$$n = (M_B/M_A) * (R_{max}/B) \quad (1.6)$$

where M_A and M_B specify the molecular weight of analyte and ligand respectively. B is the ligand immobilization level and R_{max} the maximal response obtained experimentally at saturation conditions. Alternatively, the affinity constant can be calculated directly from association (k_a) and dissociation (k_d) rate constants (equation 1.7 and 1.8).

$$K_A = k_a/k_d \quad (1.7)$$

$$K_D = k_d/k_a \quad (1.8)$$

However, it must be taken into account that interactions with the same affinity can result from different kinetics. As an example, a K_D of 10^{-9} M can result either from an interaction with k_a 10^6 $M^{-1}s^{-1}$ and k_d $10^{-2}s^{-1}$ or from another with k_a 10^3 $M^{-1}s^{-1}$ and k_d 10^{-5} s^{-1} . Thus K_A and K_D must be used with care when describing complex formation.

Apart from the affinity and kinetic parameters that describe the strength and speed of an interaction, SPR-based technology is able to provide thermodynamic information. The binding free energy described by equation 1.9 is experimentally determined according to the Van't Hoff equation 1.10.

$$\Delta G^\circ = \Delta H^\circ - T\Delta S^\circ \quad (1.9)$$

$$\Delta G^\circ = -R * T * \ln K_A \quad (1.10)$$

ΔH° and ΔS° parameters can be calculated by measuring the dissociation constant over a range of temperatures and fitting them to a more rigorous integrated form of the Van't Hoff equation [186].

1.4.3 SPR-MS

SPR-based technology is mainly used to characterize interaction between two pure biomolecules. However, when analyzing complex biological fluids by SPR, a new step where the identity of the SPR-captured molecule is unambiguously characterized becomes necessary. As SPR is non-destructive, bound proteins on the sensor surface can be further analyzed via other techniques such as mass spectrometry (MS), which is one of the most sensitive and specific techniques for the identification and characterization of peptides and proteins. The fact that the quantities that can be bound to the sensor chip are of the same order of magnitude (*e.g.* 100 fmol) as those typically needed for MS analysis suggests that coupling of these two analytical techniques is possible [76]. MALDI-TOF MS was the first MS format to be coupled with SPR, both because of its high sensitivity and relative tolerance for contaminants such as salts, detergents and buffer components. Thus, Krone *et al.* [77] detected femtomole quantities of myotoxin-a by MALDI-TOF MS after SPR-based analysis using an anti- myotoxin-a antibody immobilized on the sensor surface. After detection of the binding event, flow cells were extensively washed and the sensor chip was taken out from the SPR instrument. MALDI-TOF MS analysis was performed directly on the sensor surface after application of matrix to each flow cell. The ability to detect myotoxin-a either in pure form or as component of a complex mixture (*i.e.*, a snake venom) demonstrated the applicability of this combined approach for combining the real-time detection capabilities of SPR with the accurate molecular mass determination of MS. Other biomarkers such as cystatin C and urinary protein 1 have been quantified and identified

from biological fluids (*e.g.* plasma and urine) by SPR-MS, despite being present at very low concentrations (*e.g.* 3 ug L^{-1}) [78]. However, this approach suffers from two main disadvantages, namely the necessity of an appropriately configured mass spectrometer and the cost of the analysis, as sensor chips were only used for one unique SPR-MS measurement.

In order to gain more flexibility and sensitivity, Sönksen *et al.* developed a sandwich microrecovery method to elute affinity-bound molecules from the sensor surface in few microliters for subsequently analysis by MALDI-TOF MS [79]. After cleaning the whole SPR system to remove any nonspecifically bound proteins, the elution step was carried out between two air bubbles. These two air bubbles have two main functions: (1) detection of the recovery event in the sensorgram, and (2) allowing the switch to a volatile elution buffer (*e.g.*, ammonium bicarbonate). This approach was tested in a BIAcore X instrument for both antibody-antigen and protein-DNA systems, with a detection limit in the femtomole range. It must be taken into account that using this “sandwich elution” there is inevitable protein loss during post-elution steps that decreases sensitivity that can be critical for subsequent MS measurements, especially noticeable when protein concentrations are low. In addition, determination of the intact molecular mass by MALDI-TOF analysis is not a realistic method for protein identification, for reasons that include not only the mass resolution of current MALDI-TOF instruments but also the fact that post-translational modifications cause changes in the molecular mass and thus make identification unreliable [80]. Therefore, in order to identify unequivocally the retained molecule, a further step that includes protein identification by peptide mass fingerprint becomes necessary. Nelson and Natsume reported on a method in which a proteolytic digestion step is incorporated. After specific capture of the protein on the first flow cell, the specific bound molecule can either be digested by a proteolytic

INTRODUCTION

enzyme that is delivered subsequently to the same flow cell [81] or be routed into another flow cell, enzymatically active, where digestion occurs [82]. Both accurate molecular masses of the native protein and the proteolytic fragments provide information for the rigorous protein identification.

Recently, as a result from the collaboration between Biacore and Bruker Daltonics, an SPR-MS functionality for the BIAcore 3000 instrument has been developed in order to automate and increase the capacity of the recovery process [83]. The micro-recovery function operates, as the manual sandwich elution, by injecting a series of volumes separated by air bubbles. Three special commands (MS_INJECT, MS_WASH and MS_RECOVER) were created in order to minimize the contamination of the recovered material with components of the running buffer, by washing of the fluidic system after sample injection (MS_WASH) and recovery the bound analyte in a small volume between injections of air bubbles (MS_RECOVER) [84]. To elute the specifically bound analyte in the smallest possible volume, 2 μ l of recovery solution is placed over the flow cells and left in contact with the surface for 30 sec. The solution is then returned back to the injection needle and deposited either into a vial or directly on a MALDI target. Throughout the course of such experiment, SPR detection is used to monitor and quantify the capture and recovery of the analyte. Figure 1.10 shows a typical microrecovery sensorgram obtained after the sample injection (A), wash of the fluidic system (B) and recovery of the bound material (C). In most examples, the experimental system was based on well characterized, high-affinity analyte-antibody interactions, or used samples at higher concentrations than those of native biological fluids. For lower affinity interactions or lower analyte concentrations, a larger sensor surface is required. Thus, a prototype of Surface Prep Unit (SPU) was designed that increases the available surface

from 1.2 to 16 mm² [85]. In addition, it has short injections and recovery channels that minimize losses of sample. However, real-time monitoring of the binding event is not available on the SPU, so the optimal conditions for immobilization, binding and elution must be established beforehand in a standard chip. In addition, this current SPR-MS approach is low in throughput (4 analyses on a single sensor surface). To overcome this limitation and perform SPR-MS experiments in a high-throughput manner, an SPR-MS array platform was designed, where binding on multiple protein spots was monitored by SPR Imaging, and characterized afterwards by MALDI-TOF analysis on the same array platform [86].

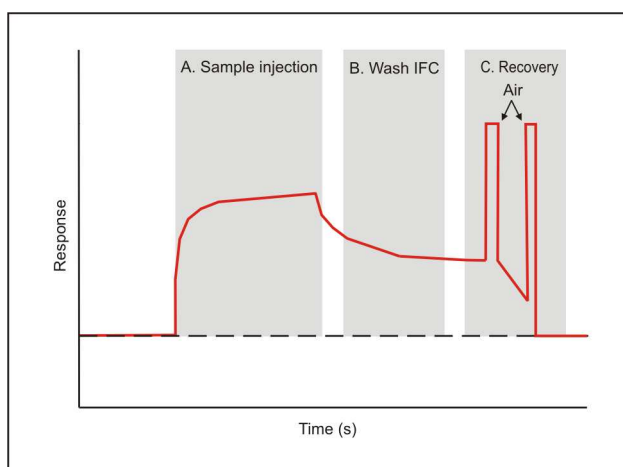


Figure 1.10 Example of a sensorgram designed to recover material for MS analysis.

1.4.4 Carbohydrate-lectin interactions studied by SPR

Although SPR-based technology was originally conceived for the study of protein-protein interactions, its field of application has been extended to other biomolecules (*e.g.* nucleic acids and carbohydrates) in recent times. To study carbohydrate-lectin interaction, either the carbohydrate or the protein must be immobilized. Most SPR-based studies on carbohydrate-lectin interactions reported to date rely on lectin immobilization. Protein

INTRODUCTION

immobilization can be carried out either covalently through amine or thiol groups, or indirectly through capturing systems (*e.g.* biotin-streptavidin) (Figure 1.11, panel A,B,C). Generally, the non-covalent immobilization (*e.g.* biotinylated lectins on streptavidin-activated surfaces) is preferred in front of classical covalent chemistry because it creates a more homogenous surface and preserves ligand native structure, although the coupling of biotin to the protein is a random process [87].

However, with immobilized lectins, only large glycan structures (*e.g.* labeled oligomannose-type N-glycans with MW 1354 [88], tetrasialylated tetraantennary complex type N-glycan with MW 3855 [87], or whole glycoproteins (*e.g.* fetuin with MW 48400)) have been successfully monitored by SPR. It has been reported that the affinity constants calculated were in good agreement with those obtained by classical methods [89]. However detection of mono- (*e.g.* galactose with MW 180) or disaccharide (*e.g.* lactose with MW 342) interactions requires high ligand immobilization levels that are both difficult to obtain and may cause mass transport limitations in kinetics measurements. The alternative approach, direct sugar immobilization, requires chemically demanding and not universally successful procedures. BIAcore has established a method to immobilize oxidized glycoproteins through aldehyde coupling chemistry (Figure 1.11, panel D). Reaction occurs spontaneously between the aldehyde group in the carbohydrate unit and a hydrazide group on the chip surface [90]. As the increase in signal depends, on the one hand, on the amount of immobilized ligand but on the other hand also on its MW and three-dimensional structure, this protocol may be suited only for high-molecular polysaccharides [91] but not for small sugars (mono- di- and trisaccharides).

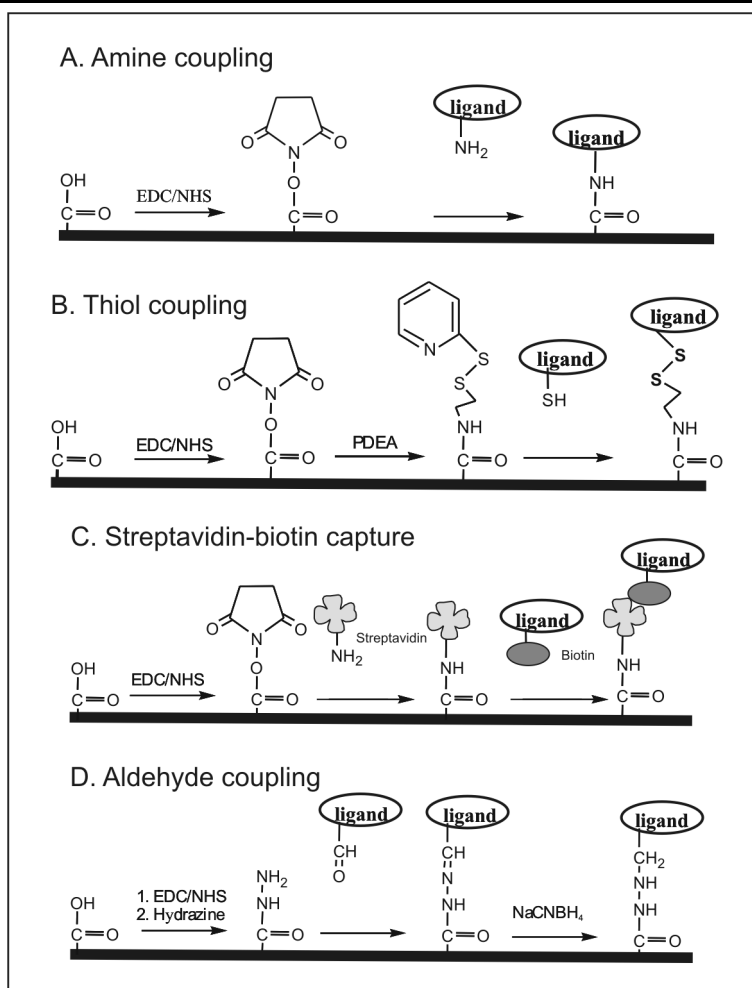


Figure 1.11 (A,B) The amine ($-\text{NH}_2$) and thiol ($-\text{SH}$) coupling chemistries for immobilizing covalently proteins to a carboxymethylated dextran coated surface are well established procedures. (C) Streptavidin-biotin capture is a typical capturing approach to immobilize non-covalently proteins. (D) For immobilization of glycoproteins, the aldehyde coupling chemistry is an alternative.

Very high densities of immobilized oligosaccharides are required to monitor properly the immobilization process, and these are often difficult to achieve, because large amounts of sugar are needed. In order to circumvent this problem, most SPR studies of carbohydrate-lectin interactions have been done on immobilized glycoproteins [92,93] or further processed glycopeptides [94]. Unfortunately, glycoproteins are

INTRODUCTION

found in nature not as pure entities, but as a heterogeneous array of glycosylated forms [95], and therefore purifying large amounts of well-defined glycoproteins from natural sources remains a nearly impossible task. For that reason, chemical synthesis of glycoconjugates is an attractive alternative. Carbohydrates can be conjugated to different kinds of molecules, ranging from whole proteins such as human serum albumin [96], biotin [97-99], lipids [100], and peptides [101,102]. The non-glycosidic part improves the sugar immobilization and disposition onto the surface. Protein and peptide derivatives are immobilized largely through classical amine coupling chemistry. As shown in Figure 1.11 (panel A), the first step is to activate the carboxymethyl groups with *N*-hydroxysuccinimide (NHS), creating a highly reactive succinimide ester which reacts with primary amines on peptides and proteins (usually lysine side chains). After immobilization, the remaining activated carboxymethyl groups are blocked by injecting a high concentration of ethanolamine. However, it must be taken into account that this amine coupling is relatively indiscriminate, and the protein will be coupled randomly through every lysine present on protein generating a heterogeneously surface. In order to avoid this heterogeneity on the surface, synthetic glycoconjugates represent a good alternative. We have recently reviewed in depth the current developments in glycoconjugate synthesis. This review is included in this dissertation as Appendix.

1.4.5 Glycprobe synthesis via chemoselective ligation

A number of coupling methods (see Appendix 1) have been developed to link carbohydrate to platforms in order to facilitate immobilization. Reductive amination was one of the first methods to be implemented in the preparation of glycoconjugates (Figure 1.12, panel A). Under suitable conditions glycans with a free reducing end can be covalently attached to protein amine groups (usually lysine side chains) forming an iminium ion

that can be reduced to an amine. However, this chemistry has been limited to sugars that are readily available in large quantities (mg), as the reaction is slow and inefficient. Typically, large excesses of sugar (300 equivalents), high concentrations (60 mM) and long reaction times (> 24 h) are usually employed in this condensation reaction. In addition, reductive amination confers an open, acyclic structure to the reducing (carbonyl) end of the glycan, rendering it different from the natural counterpart, a fact that may have consequences on its recognition [103]. A new approach for glycoconjugate synthesis takes advantages of recent developments in *chemospecific ligation* [104], a term that includes a series of methodologies (thioether, oxime, hydrazone, *etc.*) originally developed to obtain large-size peptides or even proteins by strictly chemical methods [105],[106]. Other chemistries that can be similarly classified as ligation processes and have been applied to glycoconjugates rely on long-established organic reactions, notably the Huisgen 1,3-dipolar cycloaddition of azides and alkynes (Figure 1.12, panel B), also known by its shorthand name of “click chemistry”. In this variety of ligation, the required derivatization of the glycan component either as an azide or an alkyne, plus the need for (toxic) Cu(I) catalysis may limit applicability [107]. Other electrophile-nucleophile pairs can be selectively coupled in the presence of other reactive components and therefore used in chemoselective ligations. They are classified into two broad categories: (1) reaction of thiol groups with a variety of electrophiles, and (2) reaction of carbonyl groups (ketones or aldehydes) with strong nucleophiles [108]. For the first type, a variety of electrophiles can be used under mild conditions, provided that competing thiol groups are not present (Figure 1.12, panel C). A further, serious limitation of both approaches is that, if the glycan is available in limited amounts (as is often the case), the necessary derivatization steps may be unfeasible. In approach (2), most synthetic glycopeptides are created through condensation of the

INTRODUCTION

unmodified (electrophilic) carbonyl of the anomeric carbon with a strong non-natural nucleophile group (*e.g.* hydrazide, hydroxylamine, thisemicarbazide) previously built into the peptide chain [109]. In these reactions, the so-called alpha effect renders the nitrogen atom of hydroxylamines, hydrazides and thiosemicarbazides less basic but more nucleophilic than that of a primary amine and thus favors the condensation. Under mildly acidic conditions (pH 4-5.5) these highly reactive nucleophiles are unprotonated while other potentially competing amino, guanidine or imidazole groups are protonated and therefore non nucleophilic [104].

Among the chemoselective ligations of type (2) above, conjugation of unmodified carbohydrates with hydroxylamines has been widely used for the synthesis of glycoconjugates. In contrast to reductive amination, in oxime-linked glycopeptides the attached sugar at the reducing end is in equilibrium between closed and open ring forms [110] (Figure 1.12, panel D). To overcome this limitation and drive the equilibrium exclusively towards the closed ring form, Peri *et al.* proposed the use of *N*-methylhydroxylamine instead of regular hydroxylamine, and succeeded in retaining the linked sugar in cyclic pyranose form (Figure 1.12, panel E). These authors argue that *N,O*-substituted hydroxyl amines react with reducing sugars to form intermediate oxy-imminium species which then undergo ring closure with the 5-OH with high diastereoselectivity, with the neoglycopeptide thus formed resembling more closely a native *N*-linked glycopeptide [111-113]. Another approach to maintaining the cyclic form of the attached sugar is through condensation of unmodified carbohydrates with a hydrazide group (Figure 1.12, panel F). NMR analysis of hydrazine-linked glycoconjugates showed that the β -anomeric products were generated predominantly [114]. However, the lower stability at low pH and the presence of *syn* and *anti* isomers in hydrazine-

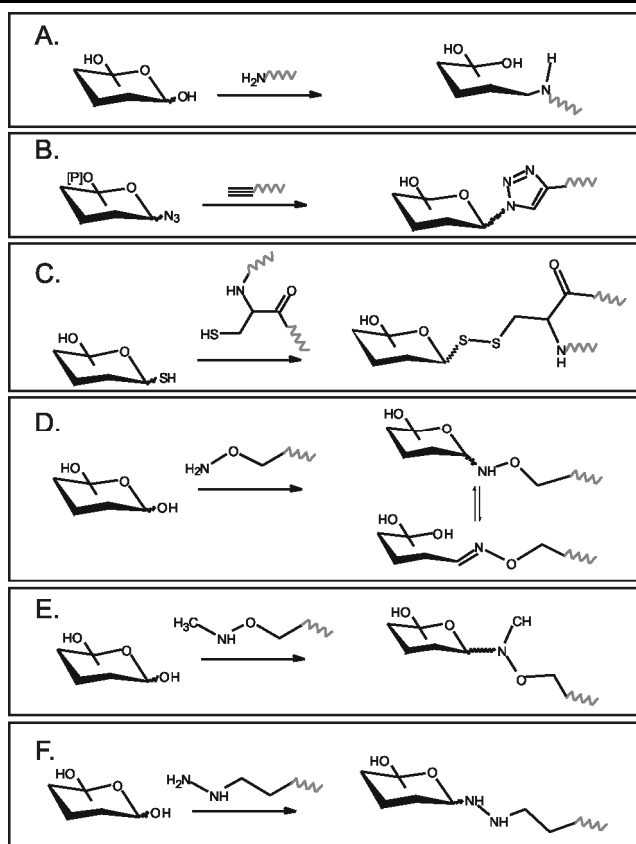


Figure 1.12 Examples of ligation reactions used in the synthesis of glycoconjugates. (A) Reductive amination. (B) Huisgen 1,3 dipolar cycloaddition. (C) Example of thiol-based chemoselective ligation (D,E,F) Carbonyl-based chemoselective ligation with an (D) aminoxy, (E) *N*-methyl-aminoxy and (F) hydrazide group.

linked conjugates still limit the broad utilization of these procedures [115]. Chemoselective ligations with aminoxy or hydrazide groups usually proceed with modest reaction rates (mM^{-1} to $\text{M}^{-1}\cdot\text{s}^{-1}$). In order to accelerate them, high concentrations (mM), elevated temperatures or nucleophilic catalysts (*e.g.*, aniline) have been used. [116-118].

In our group, the oxime ligation was explored to synthesize a neoglycoconjugate that serves as proper glycoprobe ligand in SPR interaction studies [119]. For this approach, a peptide sequence with

specific properties was designed. The aminoxy functionality was incorporated at the N-terminus for the chemoselective ligation with the sugar. As a result of the reaction between the aminoxy group and the reducing sugar group, an oxime bond is formed that maintains the closed ring structure of the reducing end of the sugar component in equilibrium with the open form. The peptide moiety of the neoglycoconjugate is large enough (MW > 700) to provide a substantial mass enhancement effect on the refractive index near the dextran surface compared to that of a simple sugar epitope, thereby facilitating the monitoring of the immobilization.

1.5 Mass spectrometry

Mass spectrometry, in particular electrospray ionization MS (ESI-MS), has been recently used by protein chemists to study noncovalent protein interactions, because the mild ionization process allows detection of intact, weakly bound, complexes without causing molecular fragmentation [120]. Although ESI-MS does not provide high-resolution structural data in the same sense as NMR or X-ray crystallography, it nevertheless allows obtaining stoichiometric information of the complex from less material and in a shorter time. The fact that the MW of the complex can be directly measured allows determining its stoichiometry. Several protein interactions such as protein-cofactor [121], enzyme-substrate [122] and protein-DNA [123], have been studied by ESI-MS. The study by Tuong *et al.* is one of the few where binding of heparin and antithrombin III is described by ESI-MS [124]. This approach has in recent years received renewed impetus by the introduction of improved instrumentation that has been applied to the study of complex supramolecular entities [125].

1.5.1 Structural information by limited proteolysis coupled with mass spectrometry

Recently, a number of MS-based approaches have been proposed as tools to map the interacting domains of sugar-lectin complexes. Though not providing the level of structural detail of NMR and X-ray crystallography, these MS-based approaches, originally developed for antigen-antibody interactions, have the advantages of high sensitivity, rapid analysis time and low sample consumption. The technique, which combines limited proteolysis and MS, is based on the principle that formation of antigen-antibody complex shields some amino acids, particularly those in regions involved in interaction, from protease attack and thus decreases the extent of proteolytic cleavage of the antigen. The epitope excision approach [126] involves (1) antibody immobilization, (2) affinity binding of the antigen followed by limited proteolytic digestion of the immune complex, (3) elution, and mass spectrometric analysis of the affinity-bound peptides (Figure 1.13, panel A). Immobilization of the antibody can be carried out either covalently on *N*-hydroxysuccinimide activated [127] or CNBr activated [128] Sepharose beads or by capturing with a secondary immobilized antibody and subsequently cross-linking [129]. Nonspecific binding can be avoided by using amino modified polystyrene beads instead of standard Sepharose [130]. As antibody binding protects the epitope from the digestion, it can be subsequently eluted in essentially unaltered form and analyzed by MS. Both linear [131] and discontinuous [132,133] epitopes have been successfully characterized by means of this epitope excision approach. Consecutive enzymatic reactions on the bound epitope can be performed to further fine-tune the epitope mapping process [134]. In a variation of this approach, called epitope extraction, the antigen is first subjected to enzymatic digestion, and then the digested fragments are captured by binding to the antibody (Figure 1.13, panel B)

INTRODUCTION

[132,135]. Moreover both approaches have been carried out successfully in solution [136,137]. However, these approaches based on limited proteolysis coupled with mass spectrometry are highly dependent on the availability of protease-specific cleavage sites within the antigen.

1.5.2 Carbohydrate-Recognition-Domain-EXcision-MS (CREDEX-MS)

The epitope excision/extraction techniques described above can be extended to carbohydrate-protein interactions [138]. In contrast to antigen-antibody, in carbohydrate-protein interactions only one interacting partner, the lectin, can be proteolytically digested, so in this case the carbohydrate moiety (epitope) is immobilized on beads and lectin is passed through in solution. Several chemistries have been described to immobilize covalently the sugar to surfaces (see 1.4.4) [139]. Among them, immobilization on divinylsulfone-activated Sepharose is the preferred one for short sugar epitopes, because it appears to maintain the glycan unit next to the divinylsulfone group mainly in its naturally occurring closed-ring form (without excluding partial equilibrium with the open one), while non-anomeric sugar hydroxyl groups are randomly linked to the divinylsulfone-activated surface [140]. The applicability of this CREDEX-MS method was demonstrated by the determination of the CRD of galectin-1 and galectin-3, two β -galactose specific lectins that play key roles during cancer metastasis [141].

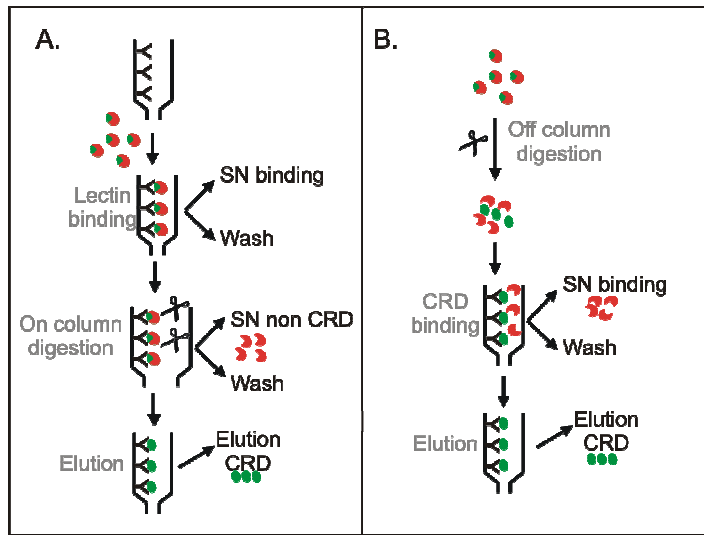


Figure 1.13 Scheme of the combined methodology based on limited proteolysis and MS. (A) Excision experiment. (B) Extraction experiment.

2 OBJECTIVES

There is a growing awareness of the biological importance of carbohydrates on cell surfaces and circulating proteins, on account of their role in cellular communication and cell-pathogen recognition. However, only a number of these interactions have been elucidated and their interacting partners characterized. For this reason, there is considerable interest in devising powerful analytical tools to study these molecular recognition events in detail. SPR is one of the best suited techniques for the study of carbohydrate-protein interactions, due to its sensitivity, low sample consumption and real-time monitoring.

The work described in this thesis builds on previous work in our group, where a glycopeptide probe (Aoa-peptide-amide) was designed for displaying carbohydrate epitopes on SPR sensor surfaces. From this starting point, the present work has been developed in order to extend the application field of this initial approach to other sugars relevant in biological processes. Thus, the following main objectives were set:

1. Optimization of the original peptide scaffold to ensure a broad application and an efficient synthesis.
2. Application of the system to a number of key glycan epitopes, complementing the analysis with mass spectrometric identification of the bound material.

OBJECTIVES

These objectives have been further elaborated as follows:

- Optimization of Aoa-peptide synthesis and conjugation to different mono-, di- and trisaccharides.
- Structural characterization of glycoprobes displaying GlcNAc and Glc by NMR and MS, with the aim of evaluating conformational aspects of the aglycon-bound monosaccharide.
- SPR-based kinetic, affinity and thermodynamic studies of interactions between well-known legume lectins and different sugar epitopes.
- Interfacing SPR with MS for lectin capture and characterization.
- Characterization of lectin carbohydrate recognition domains (CRD) by limited proteolysis and MS (CREDEX-MS)

3 RESULTS

3.1 Glycophage synthesis

In a previous work of our group [119] a short peptide (Aoa-GFAKKG-amide) was developed as an efficient tool for immobilizing carbohydrate epitopes on the sensor surface. The usefulness of this approach was demonstrated in the analysis of a well-known lectin-sugar interaction, namely wheat germ agglutinin and chitopentaose. With this original work as a starting point, the first goal of this section was to critically assess and modify the experimental design in order to improve the presentation of smaller epitopes. Starting from the original sequence - GFAKKG-amide – first the Aoa coupling during SPPS was investigated.

3.1.1 Boc-aminoxyacetic acid coupling¹

As previously described, oxime chemistry is one of the most successful approaches to synthesize glycopeptides. Generally, the oxime bond is formed between the carbonyl group constitutively present at the reducing end of all monosaccharides and an aminoxy group introduced in a peptide sequence. The aminoxy group has a pK_a of 4.6 that makes it suitable for conjugation under mild acidic conditions where other primary amino groups in the peptide are protonated and therefore excluded from the reaction. Despite the fact that this chemoselective ligation has proven to be very efficient for synthesis of neoglycopeptides, it has as main drawback the over-acylation of the NH-O nitrogen during Aoa coupling,

¹ Work reported in Jiménez-Castells, C.; de la Torre, B.G.; Gutiérrez-Gallego R.; Andreu, D. (2007) Optimized synthesis of aminoxy-peptides as glycophage precursors for surface-based sugar-protein interaction studies. *Bioorg. Med. Chem. Lett.*, **17**, 5155-5158.

RESULTS

leading to undesired byproducts and increased heterogeneity in the crude sample [142]. Different protection groups, such as allyloxycarbonyl [143], t-butoxycarbonyl (Boc) [144] or trityl [145], have been tested to overcome this problem. However, none of these protecting groups is able to prevent completely the overacylation reaction under standard Fmoc-tBu SPPS conditions. Even with the bis-protected Boc₂-Aoa-OH derivative, the crude sample does not yield a simple product except at the end of the peptide chain, because of its instability under Fmoc deprotection conditions [146]. Alternative protection schemes using substituents such as phthaloyl [147] or 1-ethoxyethylidene [148] that mask completely the nucleophilicity of the nitrogen atom have been tested under different coupling conditions and shown to completely prevent the *N*-overacylation side reaction. However, at present only Boc-Aoa-OH, Boc₂-Aoa-OH and Fmoc-Aoa-OH derivatives are commercially available [149] and so it was decided to focus our efforts into them, particularly on the former derivative. Different coupling conditions for Boc-Aoa-OH were therefore tested with a view to minimize the overacylation side-reaction. With the original platform (GFAKKG-amide) chosen as substrate, Boc-Aoa-OH was coupled manually using different coupling agents, reaction times and stoichiometries (Figure 3.1, Table 3.1). Prior to the acylation, an aliquot of the peptide-resin was cleaved to ensure the sequence was correctly assembled (>95% by HPLC).

The first coupling chemistry to be evaluated was activation using uronium salts (*e.g.* HBTU), commonly used in automated peptide synthesizers due to its reduced reaction time (45 min) *vs* methods such as carbodiimides (60 min) or preformed active esters (>3 h). HBTU-mediated couplings require base to ensure a fully deprotonated α -amino group on the growing peptide chain and, equally important, to deprotonate the carboxyl group of the amino acid for its reaction with the uronium salt. A tertiary amine,

usually DIEA, is used for these purposes, and added to the reaction mixture in two-fold excess over the protected amino acid. Thus, Boc-Aoa-OH was coupled using 3 equivalents of protected amino acid, 3 equivalents of HBTU and 6 equivalents of DIEA during 45 min. As the ninhydrin test after the first coupling was positive, a second coupling was performed with half number of equivalents. After the coupling, the peptide-resin was cleaved under standard conditions (95% TFA) and the resulting product analyzed by HPLC and MS. Similar to earlier reports [143], this first Aoa coupling method generated a complex mixture of products, as depicted in Figure 3.1, panel A(i). All products were isolated for identification by MALDI-TOF MS. Peak “a”, with a mass of 678.4 Da, was assigned to the Aoa-peptide (m/z 679.8); peak “b”, with mass 605.4 Da, was attributed to the starting hexapeptide GFAKKG-amide (m/z 606.5). Peaks “c” (m/z 752.8) and “d” (m/z 825.7) were respectively assigned to the di- and tri-acylated byproducts. Finally, peak “e” with a mass of 703.4 Da was assigned to an *N*-terminal tetramethylguanidinium byproduct (m/z 704.7) [150]. Other peaks (*e.g.* “f”) in the chromatogram were assumed to be higher-order oligomers.

In view of the difficulties encountered in the HBTU-mediated coupling of Aoa, it was decided to test a somewhat less energetic coupling agent, namely the very conventional *N,N*-diisopropylcarbodiimide (DIC). In a first experiment using 10 equivalents of both Boc-Aoa-OH and DIC for 60 min (Figure 3.1, panel A(ii)), only three different products were obtained after cleavage and identified as Aoa-GFAKKG-amide (m/z 679.8), the over-acylated peptide (m/z 752.6) and the starting GFAKKG-amide (m/z 606.7), with a substantial improvement in the yield of Aoa-peptide over the HBTU method, 93% *vs.* 45% (Table 3.1). The overacylation is assumed to be largely a solution (as opposed to solid phase) process [143], with the overacylated species being subsequently

RESULTS

transferred onto the H₂N-bearing peptide-resin. For that reason, a new coupling condition was tested with a minimal time of incubation (10 min). In this case, 97% of the crude mixture was the target Aoa-peptide, and the remainder 3% corresponded to the over-acylated peptide (Table 3.1, Figure 3.1, panel A(iii)).

Table 3.1 Acylation of GFAKKG-resin with Boc-Aoa-OH using different coupling conditions.

Entry	Boc-Aoa-OH	Coupling agent (equiv)	DIEA (equiv)	Tim (min)	Products			
	(equiv)				a	b	c	d-f
i	3+1,5	HBTU (3+1,5)	6	40	45	4	22	29
ii	10	DIC (10)	-	60	93	2	5	-
iii	8	DIC (8)	-	10	97	0	3	-

The same coupling conditions were tested also in an N-terminally elongated version with an ε-amino hexanoic acid spacer of the GFKKG-amide peptide substrate (Ahx-GFAKKG-amide) (Figure 3.1, panel B, Table 3.2). In both examples, the best results were achieved with 8 eq of protected Boc-Aoa-OH and DIC and a shorter reaction time (10 min)..

Table 3.2 Acylation of Ahx-GFAKKG-resin with Boc-Aoa-OH using different coupling conditions

Entry	Boc-Aoa-OH	Coupling agent (equiv)	DIEA (equiv)	Time (min)	Products			
	(equiv)				a	b	c	d-f
i	3+1,5	HBTU (3+1,5)	6	40	63	17	7	13
ii	10	DIC (10)	-	60	80	16	4	-
iii	8	DIC (8)	-	10	94			

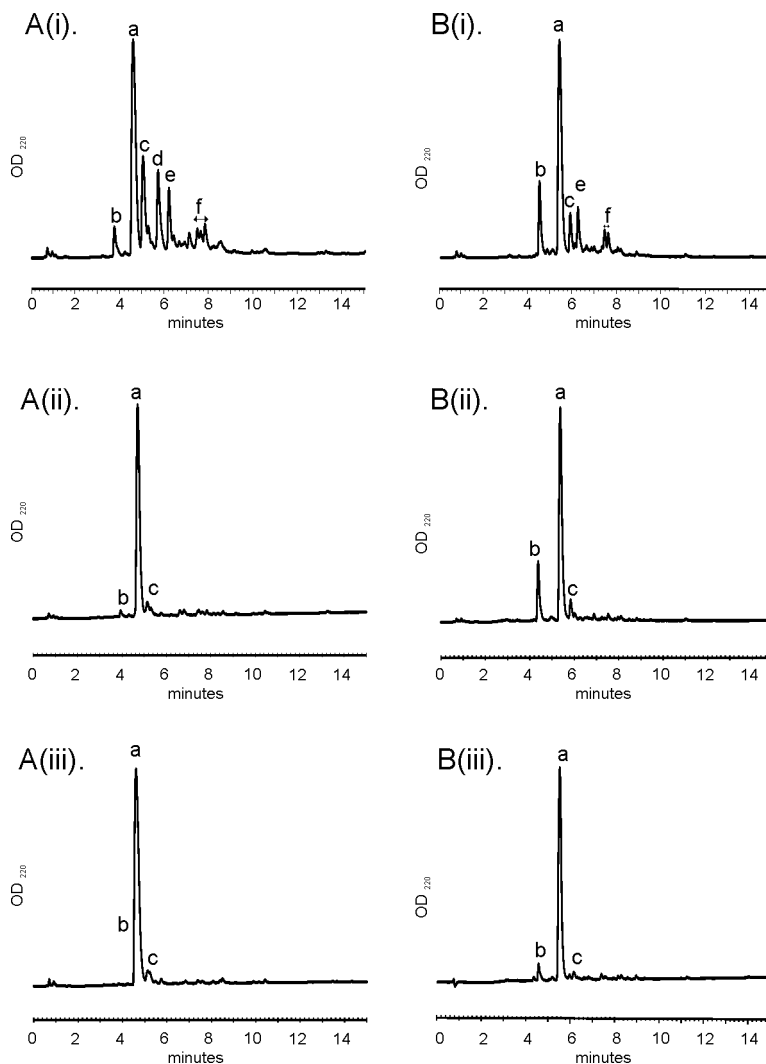


Figure 3.1 HPLC analysis of different conditions for Aoa-GFAKKG-amide [entries A(i-iii)] and Aoa-Ahx-GFAKKG-amide [entries B(i-iii)] synthesis. (i) Boc-Aoa-OH/HBTU/DIEA (3:3:6 equiv), 40min; (ii) Boc-Aoa-OH/DIC (10:10 equiv), 60 min; (iii) Boc-Aoa-OH/DIC (8:8 equiv), 10min. Peak assignments: a: target Aoa-peptides; b: starting peptide; c: diacylated peptide; d: triacylated peptide; e: *N*-terminal guanidine byproduct (uronium capping); f: oligomeric impurities.

A further complication reported for the synthesis of Aoa peptides [142] is that, during purification, several impurities are formed due to the high reactivity of the aminoxy group with carbonyl compounds, even if

RESULTS

present at trace levels such as acetone impurities in the ACN used in HPLC eluents. Moreover, degradation of the aminoxy group to the corresponding hydroxyl compound by acid hydrolysis is evidenced by a mass difference of -15 Da (loss of NH) in the MALDI TOF MS analysis [151]. Therefore, after HPLC purification, quick lyophilization removal of the (volatile) species posing these risks is practically mandatory. A simpler alternative, considering that the optimized (DIC coupling) procedure described above provides the Aoa peptide in an excellent level of purity (97%) is simply to do away with any HPLC purification and use the product directly obtained from the cleavage reaction, without any additional workup. As shown in Figure 3.2, the chemoselective ligation (see Section 3.2) with sugars can be carried out quite successfully with the non-purified peptide thus obtained.

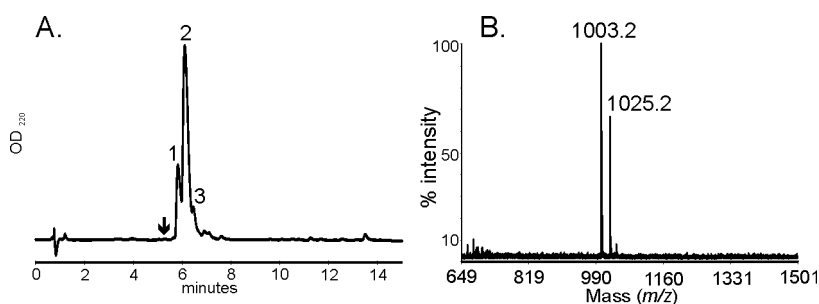


Figure 3.2 A. HPLC analysis after conjugation of Aoa-GFAKKG-amide (prepared as indicated above and used without purification) and lactose during 48 h. Arrow shows the elution position of the starting peptide. B. MALDI-TOF analysis of the purified glycopeptide (lactose-AoaGFAKKG-amide). Equivalent mass spectra (m/z 1003, $[M+H]^+$ and m/z 1025, $[M+Na]^+$) were observed for all three peaks (1-3), suggesting structural isomers (*e.g.* syn/anti oximes, α/β anomers) at the lactose-Aoa linkage.

3.1.2 The need for Aoa *N*-methylation

Several glycopeptides have been synthesized taking advantage of the high reactivity of the aminoxy group towards the aldehyde function in the acyclic (open) form of the sugar. In this chemoselective ligation, the reducing sugar unit attached directly to the peptide is reported to be in equilibrium between the cyclic and the open form [152]. This fact is not adequately taken into account in carbohydrate-lectin interaction studies using oxime-linked glycoprobes [153] where biomolecular interaction parameters are referred to the native (ring) form but should be instead attributed to the mixture. Given our interest in glycoprobes that display small sugar epitopes (*e.g.* mono- to trisaccharide) for lectin recognition, a well-defined (native-like, *i.e.*, ring) conformation of the sugar unit linked to the peptide is of crucial importance. Thus, our first objective in this direction was to structurally characterize the peptide-linked monosaccharide.

To this end, we conducted monodimensional ^1H NMR experiments to quantify the relative monosaccharide populations (*syn* and *anti* open forms, α and β cyclic ring form) in the oxime-linked glycoprobe, using GlcNAc-Aoa-peptide-amide as model. Assignments were done by reference to literature data [154] and chemical shift predictions using the ChemDraw software (version 8, CambridgeSoft, Cambridge, USA). Figure 3.3 shows part of the experimental ^1H NMR spectrum and the corresponding data obtained. Three regions of the spectrum were of particular interest. In the aromatic region (δ 6.40-8.00 ppm), the signals at δ 7.62 ($^3J_{1,2}$ 5.86) and δ 6.86 ($^3J_{1,2}$ 7.32) were attributed to the imino proton of the *anti* and *syn* isomers of the oxime bond in the open form [155] (Figure 3.3, panel A). The group of signals between these two down-field doublets (δ 7.20-7.40 ppm) was assigned to the five protons of the aromatic ring of the Phe residue. In the anomeric region (δ 4.2-5.5 ppm),

RESULTS

the doublets at δ 5.18 ppm ($^3J_{1,2}$ 3.42) and δ 4.36 ppm ($^3J_{1,2}$ 9.77) were assigned to the α and β anomeric protons of the closed ring conformation, respectively (Figure 3.3, panel B). Anomeric configurations are assigned from the two measured coupling constant between H1 and H2 protons. In pyranose ring structures such as GlcNAc, the β -configuration is represented by the axial configuration of both H1 and H2 protons. The 180° dihedral angle formed between these two protons is responsible for the large coupling constant (7-9 Hz) observed. However, in the α -configuration, the protons are equatorial-axial orientated, and therefore the dihedral angle (60°) and the corresponding coupling constant are smaller (2-4 Hz). In the aliphatic region, the two singlets at δ 2.03 and δ 2.00 were assigned to the methyl protons of the *N*-acetyl group in position 2 of the GlcNAc sugar (Figure 3.3, panel C). The co-existence of two signals for the *N*-acetyl group confirmed that both the open and closed ring forms of the sugar exist and provided a complementary means of quantification.

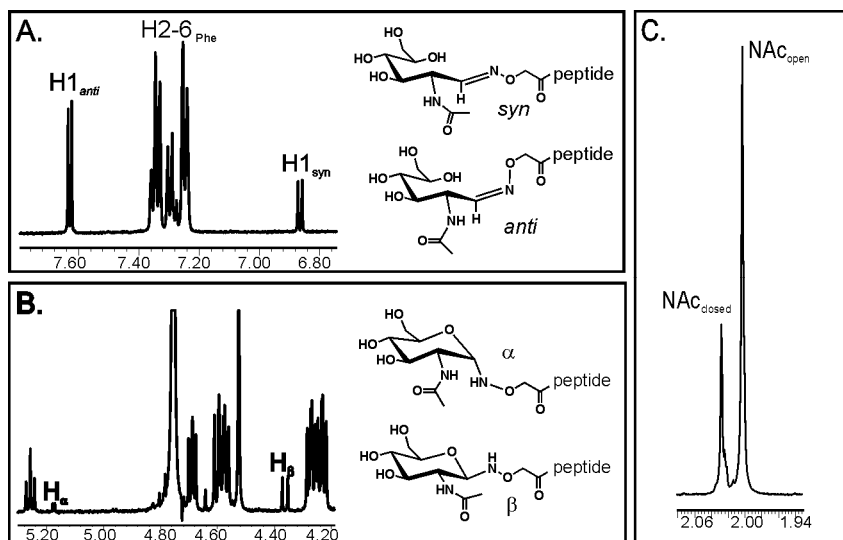


Figure 3.3 1D ^1H NMR spectrum (500 MHz) of GlcNAc-Aoa-peptide recorded in D_2O at 25°C . Panels A, B and C are enlarged areas of the full spectrum corresponded to the aromatic, anomeric and aliphatic regions, respectively.

In order to quantify the relative population of each monosaccharide form (α , β , *syn*, *anti*) in the equilibrium, signals assignable to each species were integrated. The five aromatic protons of the Phe residue were taken as reference for peak integration.

Table 3.3 1D ^1H NMR chemical shifts (500 MHz, at 25 °C) of the most relevant exchangeable protons for GlcNAc-Aoa-peptide in D_2O .

Residue	Exchangeable proton	Chemical shift (δ , ppm)	Coupling constant ($^3J_{1,2}$, Hz)	Signal integration
GlcNAc	H1 <i>anti</i>	7.62	5.86	0.45
	H1 <i>syn</i>	6.86	7.32	0.25
	H1 α	5.18	3.42	0.06
	H1 β	4.36	9.77	0.24
	CH ₃ (open)	2.00	-	2
	CH ₃ (closed)	2.03	-	1
Phe	H2-6	7.24-7.33	-	5.00

As shown in Figure 3.3 and Table 3.3, the peaks assigned to the four possible situations of H1 (α , β anomers in the ring form, panel B; *syn*, *anti* isomers in the open form, panel A) were integrated and the relative populations determined as 6:24:25:45, respectively. Thus, the total fraction of open forms was 70% whereas the remaining 30% corresponded to the closed ring form. This ratio was confirmed by the integration of the acetyl peaks at δ 2.01 (closed) and δ 2.00 (open). In sum, the oxime-based chemoselective ligation using Aoa as nucleophile yields a glycopeptide adduct in which the pyranose form of the sugar is present only at ~ 30%, and of this 30%, an 80% of the the ring is in the β -configuration. This analysis suggests that, for interaction studies of unknown carbohydrate-

RESULTS

binding entities employing relatively short (mono- to trisaccharide) sugar epitopes, the oxime-based chemistry possible does not afford sufficient representation of active, native-like structures, as in the oxime-linked glycoprobe at least four different species coexist (*anti*, *syn*, α , β), of which the two most abundant ones (*anti*, *syn*) are essentially non-native, and hence their presence and potential interaction may affect the interpretation of results.

It has been reported [111,113] that the cyclic form of the sugar is maintained when the primary hydroxylamino group ($\text{H}_2\text{N-O-}$) of Aoa is replaced by a secondary hydroxylamine [$\text{HN}(\text{CH}_3)\text{-O-}$] function. This substitution precludes the possibility of an open oxime form. In order to corroborate this point, the *N*-methylated derivative of the glycoprobe GlcNAc-*N*[Me]-Aoa-GFKKG was prepared and analyzed by NMR.

Figure 3.4 shows a comparison of NMR spectra of both GlcNAc-Aoa-peptide and GlcNAc-*N*[Me]-Aoa-peptide. Important differences are observed in the aromatic and aliphatic region. As shown in panel B, the signals previously assigned to the imino proton of the oxime group in both *syn* and *anti* open forms have disappeared, and only the signals corresponding to the aromatic protons of the Phe residue are observed. The same occurs for the signals of H2 and H3 in the open form (not highlighted in Figure 3.4). In the aliphatic area, the two singlets assigned to the methyl protons of the *N*-acetyl group of GlcNAc-Aoa-GFKKG are converted into one singlet in the spectrum of the GlcNAc-*N*[Me]-Aoa-peptide. This confirms that a secondary amine in the aminoxy functionality ensures that the conjugated monosaccharide remains in a closed ring form, essential for binding to carbohydrate-binding proteins.

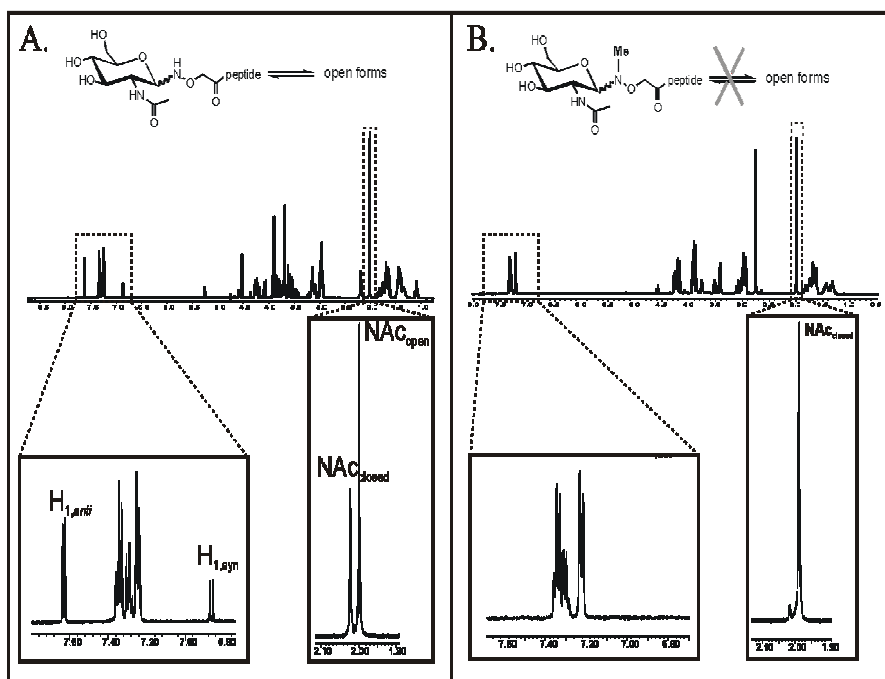


Figure 3.4 Comparison of enlarged areas of the 1D ^1H NMR spectrum of both GlcNAc-Aoa-peptide (A) and GlcNAc-*N*[Me]-Aoa-peptide (B).

In order to determine the proportion of each isomer present in the GlcNAc-*N*[Me]-Aoa-peptide, the doublets corresponding to the α - and β -anomeric protons were assigned by comparison of chemical shift and coupling constant with the non-methylated version. The signals at δ 5.17 ppm ($^3J_{1,2}$ 3.72) and δ 4.26 ppm ($^3J_{1,2}$ 9.92) were assigned to the α - and β -anomeric protons of the closing ring conformation and their relative ratio was determined as 0.1:0.9, respectively. Thus, the most abundant anomer in the GlcNAc-*N*[Me]-Aoa-peptide is the β -form. Likewise, NMR analysis of the Glc-*N*[Me]-Aoa-peptide indicated the β -form as the predominant anomer ($\sim 90\%$). In carbohydrate-lectin interactions, the anomeric configuration of the sugar at the reducing end is of crucial significance, as many lectins are able to discriminate between anomers (*e.g.* galectins binds to β -Gal but not to α -Gal) [156]. Also, Chen *et al.* described that the anomeric configuration of glycoconjugate depends on

RESULTS

the nature of the monosaccharide [157]. Our glycopeptide platform thus becomes an excellent tool to explore β -GlcNAc and β -Glc specific lectin.

Another way to confirm the ring closure of the sugar unit linked to the *N*[Me]-Aoa-peptide, requiring less material and purification efforts, is based on differential derivatization with subsequent analysis by mass spectrometry. Both glycopeptides have in common two Lys side chains that can be acetylated, but depending on whether the sugar unit is in open or closed form, the number of hydroxyl groups available for acetylation is different, namely six or five for the open and closed ring forms, respectively.

Figure 3.5, panel A shows the MALDI-TOF mass spectra of GlcNAc-Aoa-peptide and its peracetylated product. Two product clusters can be observed in the second spectrum and readily assigned to the peracetylated versions of the glycopeptides with both open and closed forms of GlcNAc. A rough estimation of the peak areas of both species corroborated the NMR findings, *i.e.*, 25% in the closed form and 75% in the open form.

In contrast, Figure 3.5, panel B shows a much simpler spectrum for the peracetylated derivative of GlcNAc-*N*[Me]-Aoa-peptide as the peaks correspond to the closed form of GlcNAc only. As already indicated, the advantage of using this technique relative to NMR is that it requires less amount of sample (ng vs μ g). However, no information can be obtained about the anomeric configuration or the ring form of the sugar unit.

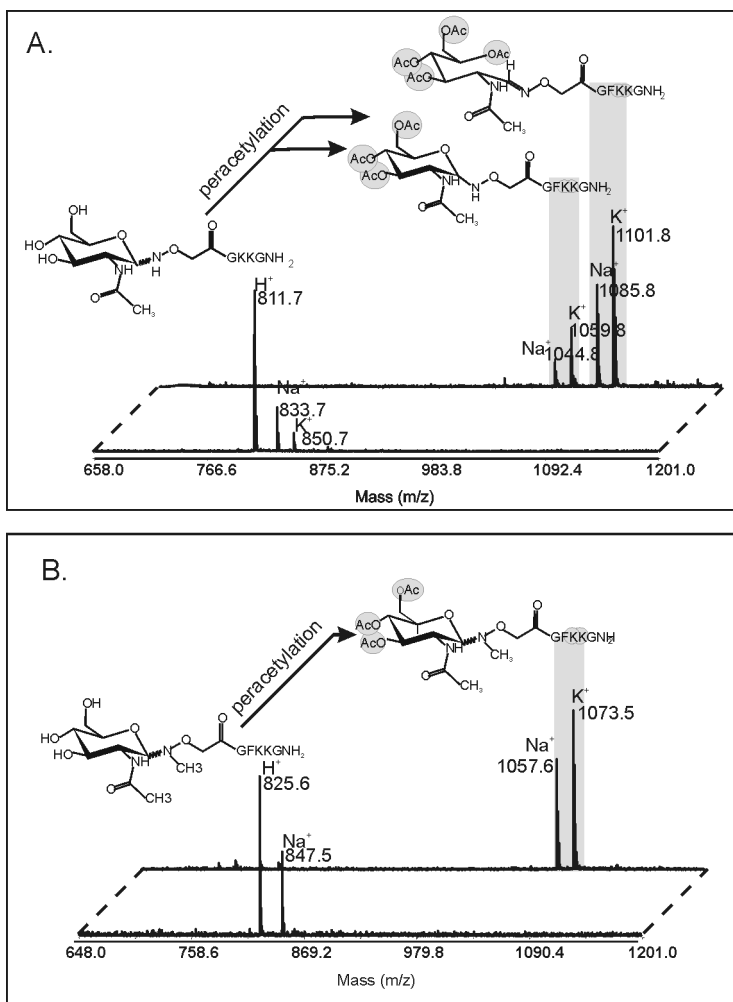


Figure 3.5 MALDI-TOF MS spectra of glycopeptides (front trace) and their peracetylated equivalents (back trace). Panel A. GlcNAc-Aoa-peptide. Panel B. GlcNAc-N[Me]-Aoa-peptide.

3.2 Chemoselective ligation via oxime chemistry

In our previous work [119], the glycoprobe was synthesized by chemoselective ligation between the Aoa-peptide-amide and the pentasaccharide (GlcNAc)₅. Several reaction conditions were tested and the best results were obtained utilizing approximately equimolar conditions and high concentrations of both sugar and peptide (25 mM and 20 mM, respectively, in 0.1M NaOAc pH 4.6 at 37 °C for 72 h). In the

RESULTS

preceding section we have established that conjugation to an *N*[Me]-Aoa-peptide preserves the sugar epitope in cyclic form. As our focus is on well-defined glycoprobes displaying small (mono- to trisaccharide) epitopes, the choice of the *N*[Me]-Aoa-peptide is of utmost importance.

3.2.1 Conjugation of mono-, di- and trisaccharides with *N*[Me]-Aoa-peptide

Crystal structures of carbohydrate-lectin complexes have elucidated that the carbohydrate recognition domain is located in a shallow depression on the protein surface [45,48]. This pocket accommodates oligosaccharides of different size ranging from mono- to trisaccharides, frequently showing increasing affinity with size. Thus, lectins are capable of recognizing even monosaccharides, the smallest and non-hydrolyzable units of glycan chains, although with affinity in the mM range. In this section we discuss the preparation and applicability of *N*[Me]-Aoa-peptide platforms of this kind.

Monosaccharides- Glycoprobes exposing some of the most commonly occurring six-carbon monosaccharides, including Glc, Man, Gal, Fuc, GlcNAc, GalNAc plus the nine-carbon, biologically very relevant sialic acid were prepared by oxime ligation. Conjugations were performed at pH 3.5 and 37 °C. Given the frequent chromatographic similarity between the starting *N*[Me]-Aoa-peptide and its glycoderivatives, a high yield of the ligation reaction was desirable not only *per se* but especially to facilitate HPLC purification. Fortunately, monosaccharides were not a limiting factor of the reaction, and a large excess of carbohydrates could be used to react with the *N*[Me]-Aoa-peptide. Although glycosylation occurred generally to > 60% yields, different rates for various monosaccharides were observed. The highest yield was obtained with Gal (~ 80%), probably because from the three D-hexoses tested (Gal, Glc, Man), it is

the one with the highest aldehyde population in solution [158]. Reaction yields decreased considerably for hexoses with hydroxyl at C2 in axial position such as Man (~ 64%). The deactivating effect of an *N*-acetyl group at the carbon next to the reducing unit (*i.e.* GalNAc (65%)) has been described [159]. For sialic acid, on the other hand, the conjugation with *N*[Me]-Aoa-peptide regrettably did not result in any appreciable product, even with excess of sugar and long incubation time (up to 144 h). In contrast, the conjugation of sialic acid to the non-methylated Aoa-peptide did take place, albeit with lower yields relative to other monosaccharides (15% *vs* ca. 100%), probably due to the higher reactivity of aldehyde (in the tested hexoses) *vs.* the ketone group at C2 of sialic acid.

Disaccharides- A quick glance to the glycan structures in biological glycoconjugates reveals that the monosaccharides covalently attached to proteins and lipids are most often Glc, GlcNAc, GalNAc. With this in mind, glycopeptides displaying disaccharides with either of those three units at the reducing end were prepared by oxime ligation. Unlike monosaccharides, pure disaccharides are costly to the point that using excess of carbohydrate (*e.g.* 200 mM) can become onerous. Therefore, new reaction conditions improving conjugation yields had to be found. In a first attempt, *N*-[Me]-Aoa-peptide and the disaccharide *N*-acetyl-lactosamine (Gal- β 1,4-GlcNAc) were conjugated under the same reaction conditions of the non-methylated Aoa-peptide. Non-methylated Aoa-peptide-amide was also reacted with the disaccharide as a control. After 72 h, the reaction was almost quantitative (87%) for the non-methylated peptide but extremely slow (4%) for the *N*-[Me]-Aoa-peptide. This can be explained by the lower nucleophilicity of a secondary relative to a primary aminoxy group, an effect already observed in the conjugation to monosaccharides, notably sialic acid (see above). Nonetheless, for monosaccharides the more sluggish reactivity of the *N*-[Me]-Aoa-peptide

RESULTS

was counterbalanced by the excess of sugar used and the impact on yields was less evident. For some disaccharides, as mentioned above, excess of sugar was not an option and other parameters known to affect ligation rates (pH, stoichiometry, reaction time) [144].had to be screened in order to improve yields. Another parameter, namely addition of organic cosolvent to the conjugation reaction, was not explored to avoid potential side-reactions (see 3.1.1.) with contaminants present in organic solvent (*e.g.* acetone, formaldehyde). Temperature was neither tested as a parameter, but kept at physiological values (37 °C). Conjugations were done with *N*-[Me]-Aoa-peptide-amide and included as disaccharides *N*-acetyl-lactosamine (Gal- β 1,4-GlcNAc), lacto-*N*-biose (Gal- β 1,3-GlcNAc), galacto-*N*-biose (Gal- β 1,3-GalNAc), and lactose (Gal- β 1,4-Glc). The pH dependence was first studied, as both oxime formation and mutarotation equilibrium are pH-dependent. It is known that oxime formation occurs at acidic pH, with a maximum between pH 3.5 and 5 [160]. So, various pH values in the 3 to 5 range were tested, without changing the stoichiometry and reaction time. After 72 h an aliquot of each condition was analyzed by HPLC and all peaks were isolated for characterization by MALDI-TOF MS. The yields for each conjugation condition were estimated by integration of HPLC peaks and are shown in Table 3.4.

Given the pK_a 4.6 of Aoa, one could expect that the optimum for the reaction was a pH fairly above 4.6. As shown in Table 3.4, however, the optimal pH found experimentally is much lower, around 3.5. This can be interpreted by considering another factor involved in the reaction, namely the rate of mutarotation, which determines the relative populations of open and ring forms in solution, and which is known to be minimal at pH 4.6 [161] but to increase at lower (and higher) pH. In the present case, this latter factor would appear to predominate over the extent of Aoa ionization. In addition, the nature of the monosaccharide at the reducing

end also played an important role on the conjugation yields (Table 3.4). Thus, for N-acetyl-hexosamines, those with a terminal GlcNAc unit, (Gal- β 1,4-GlcNAc and Gal- β 1,3-GlcNAc) had yields in the 39-41% range, while for Gal- β 1,3-GalNAc a considerably better yield (64%) was found. On the other hand, the disaccharide lacking the 2-acetamido group at the terminal glycan unit, *i.e.*, Gal- β 1,4-Glc, was the one obtained in higher yields (~75%) (in this case with an optimum pH of 4.6, not 3.5), evidencing (as already mentioned above in this section) the deactivating effect of the acetamido group.

Table 3.4 Optimization of ligation conditions. All ligations were carried out at 37 °C in 0.1 M NaOAc for 72 h at 20 mM and 25 mM for peptide (*N*[Me]-Aoa-Ahx-GFAKKG) and disaccharides, respectively

Conjugates	pH				
	5	4.6	4	3.5	3
Gal- β 1,4-GlcNAc	-	5	21	41	32
Gal- β 1,3-GlcNAc	-	8	20	39	15
Gal- β 1,3-GalNAc	-	37	51	64	49
Gal- β 1,4-Glc	68	89	75	71	-

% Estimated of by integration of HPLC peaks at 220nm

The effect of stoichiometry at pH 3.5 was next tested and shown to be inconsequential. Thus, reducing the concentration of *N*-[Me]-Aoa-peptide from 20 mM to 12 mM (25 mM disaccharide) did not have any significant effect on the yields.

Table 3.5 shows the results of *N*-[Me]-Aoa-peptide conjugation to disaccharides containing either GlcNAc or Man as the non terminal glycan unit. Although these data were obtained with a different peptide platform (*N*[Me]-Aoa-GFKKG) than those of Table 3.4 (*N*[Me]-Aoa-Ahx-GFAKKG), they are equally illustrative of the subtle substituent effects encountered in this type of reactions. Thus, mannobioses (with an axial

RESULTS

hydroxyl at C2) react significantly more poorly than bioses with terminal GlcNAc, and within this latter group the nature of the non-reducing sugar unit (compare Fuc with Gal, Table 3.5) is also quite determining. Finally, the type of glycosidic linkage between both glycan units also influences the outcome of the conjugations, as shown by the generally better yields observed for β 1,6 over β 1,4 or β 1,3-linked disaccharides in Table 3.5. All these data point to a complex array of electronic and steric effects at play in this type of reactions illustrate clearly the effect of an *N*-acetyl group in the 2-position of reducing unit on the ligation reaction, interfering negatively on the coupling yield between the anomeric centre and the *N*[Me]-Aoa group. With respect to the C2 a more negative effect is observed in conjugations with α -mannosides that differ from the other hexoses (Gal, Glc) in the position axial of the hydroxyl group at C2. As shown in Table 3.5, for mannosides, only a 15% of yield is obtained, whereas for *N*-acetyl-disaccharides, the yield increased to \sim 28-30%). Remarkably is the fact that the presence of Ahx spacer between the aminoxy group and the rest of the peptide improved conjugation yield (Table 3.4, Table 3.5).

Moreover, as depicted in Table 3.5, different reactivity was observed depending of the subsequent non-reducing sugar. For example, under same conditions (pH 3.5, stoichiometry and incubation time, higher yields were obtained with Gal- β 1,6-GlcNAc than with the other two *N*-acetyl-glucosamine-containing disaccharides (33% vs 20-21%) . The same effect 1,6 > 1,3 \sim 1,4, is observed when Gal is substituted with a Fuc. Similar yields were observed for Fuc- α 1,3-GlcNAc and Fuc- α 1,4-GlcNAc (7-8%). By contrast, almost 4-times more was obtained for Fuc- α 1,6-GlcNAc (26%). Nevertheless, this effect of the glycosidic linkage was not observed for mannosides. (15-16 %).

Table 3.5 Effect of the glycosidic linkage on the conjugation yield. All ligation were carried out at 37 °C in 0.1 M NaOAc pH 3.5 at 20 mM and 25 mM for peptide (*N*[Me]-Aoa-GFKKG) and disaccharides, respectively.

	Conjugate
Gal-β1,4-GlcNAc	20
Gal-β1,3-GlcNAc	21
Gal-β1,6-GlcNAc	33
Fuc-α1,3-GlcNAc	7
Fuc-α1,4-GlcNAc	8
Fuc-α1,6-GlcNAc	26
Man-α1,2-Man	11
Man-α1,3-Man	16
Man-α1,6-Man	15

% Estimated of by integration of HPLC peaks at 220nm

Furthermore, long incubation times, *e.g.* ≥ 72 h, are usually needed to form oximes. In order to increase the reaction rate and thus shorten the reaction time, Dirksen *et al.* studied the effect of aniline as catalyst for oxime formation. Aniline was first shown to improve up to 400-fold the oxime rate formation between two unprotected peptides [116]. However, for reducing sugars, it showed lower catalytic efficiency than for strictly open-chain aldehydes, because the kinetics of the reaction with reducing sugars is also influenced by their mutarotation equilibrium (see above). Even so, Thygesen *et al.* obtained up to 2-3 fold increase in conversion levels for the reaction of a non-methylated Aoa-peptide with various monosaccharides in the presence of aniline [162]. Thus, with the same goal of decreasing reaction time, reactions with different disaccharides and the *N*-[Me]-Aoa-peptide were performed in the presence of 100 mM aniline. The effect was evaluated by comparing oxime formation in the

RESULTS

presence and the absence of catalyst. As depicted in Table 3.6, after 24 h incubation the reaction time decreases from 72 h to 24 h for conjugations. Longer incubation times did not show any improvement in yield.

Table 3.6 Effect of aniline on the conjugation yield. All ligations were carried out at 25°C in 0.1M NaOAc pH 4.6 at 20 mM and 25 mM for peptide and disaccharide, respectively. After 24 h an aliquot of each condition was analyzed by HPLC and isolated peaks were characterized by MALDI-TOF MS.

		+ aniline
Gal- β 1,3-GlcNAc	1	7
Gal- β 1,3-GalNAc	6	27
Man- α 1,3-Man	4	13

% Estimated of by integration of HPLC peaks at 220nm

Another way to improve the reaction rate is using a higher excess of one of the reactants. Given its accessibility, the *N*[Me]-peptide and not the carbohydrate is the best choice in this respect. In practice, however, the attachment of mono, di- and trisaccharides to the peptide caused a change in the resulting hydrophobicities that complicated their separation from the unreacted peptide. The obvious solution, *i.e.*, making unreacted Aoa-peptide more hydrophobic, was attempted in two ways. Initially, capping of the aminoxy group by reaction with carbonyls was tried by incubating the conjugation mixture for 1 h with excess of acetone or formaldehyde at room temperature [119]. Neither of the two options was successful in capping the *N*[Me]-Aoa-peptide, confirming the lack of reactivity of *N*[Me]-aminoxy groups with ordinary aldehydes and ketones [163]. Next, *N*-ethylmaleimide, an alkylating agent described as highly reactive with *N*-alkylaminoxy-containing peptides was tested [164]. Added at 100 mM concentration to the conjugation mixture and incubated for 20 h at 37 °C, it caused almost complete conversion (89%) to an *N*-ethylmaleimido derivative well resolved by HPLC from the glycopeptides. Interestingly,

the reaction was also totally chemoselective with regard to (the Lys side chains of) the glycoconjugate.

Trisaccharides- As model trisaccharides, disaccharides with an end-capping sialic acid residue were conjugated to *N*-[Me]-Aoa-peptide. Commercially available compounds of this type include α -2,3- and α -2,6-sialyl lactosamine, as well as a mixture of α -2,(3,6)-sialyl lactoses. A particular concern with these sialylated sugars is that they are fairly acid-labile and thus could suffer from desialylation during conjugation to the peptide module. Thus, it was necessary to evaluate the applicability of the above-mentioned conjugation, which took place at acidic pH, to this kind of sugars. Fortunately, the misgivings proved to be unwarranted, since all three trisaccharides were readily conjugated to both the unmethylated and the methylated versions of the Aoa-peptide in good yields (74% and 28%, respectively).

In summary, reaction conditions for conjugation of *N*[Me]-Aoa-peptide to mono-, di- and acid-labile trisaccharides were established. For monosaccharides, good yields (*i.e.* 64-80%) were obtained using excess of carbohydrate (200 mM). By contrast, for disaccharides, the conjugation yield could be improved by acidifying the reaction mixture to pH 3.5, except for lactose whose optimum pH was 4.6. As an alternative to pH reduction, addition of aniline was also successful in reducing the reaction time from 72 h to 24 h. Finally, trisaccharides were coupled to *N*[Me]-peptide in trouble-free way by the standard procedure. All glycopeptides synthesized and characterized are listed in Section 6.3.

3.2.2 Stability of glycopeptides

When evaluating by HPLC and MS the purity of glycoconjugates of *N*-[Me]-Aoa-peptide and different carbohydrates they were found to be unstable to the pH 2 conditions of the HPLC solvent systems. Thus, after

RESULTS

a few hours in HPLC solutions the glycopeptides underwent acid-catalyzed hydrolysis to the non-conjugated *N*-[Me]-Aoa-peptide (Figure 3.6). Gudmundsdottir studied oxime-based glycoconjugates under acidic conditions and established that acid-catalyzed hydrolysis is accelerated below pH 6.0 [115]. Therefore, in our case it was decided to neutralize HPLC solutions to pH 6 immediately after purification, then lyophilize. This precaution proved effective not only for typical glycoconjugates but also for preventing desialylation of the acid-labile sialyl-containing glycoprobes discussed above.

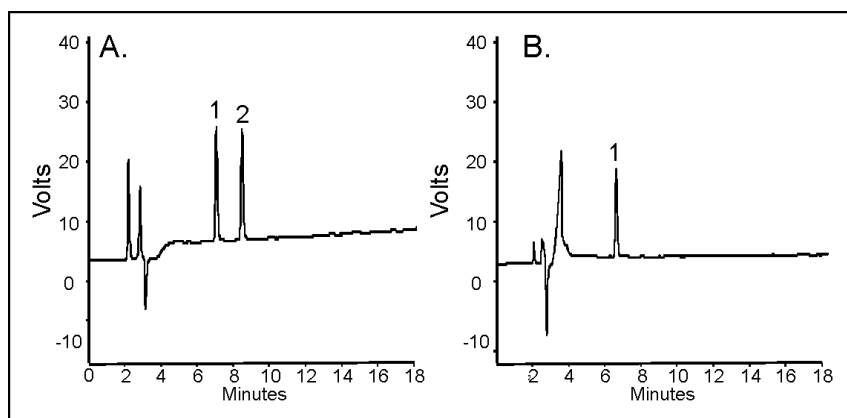


Figure 3.6 HPLC chromatogram corresponding to purified lac-*N*[Me]-Aoa-glycopeptides. After 12 h in HPLC solution, without a neutralization step (panel A), and with neutralization step (panel B). Peak 1 corresponds to glycopeptide and peak 2 to the *N*[Me]-Aoa-peptide.

The long-term stability of lyophilized Gal- β 1,4-Glc-*N*[Me]-Aoa-peptide and Gal- β 1,4-Glc-Aoa-peptide was also tested by HPLC after 2-year storage and shown to be substantially preserved, with less than 5% deglycosylation and no other byproducts identified (Figure 3.7).

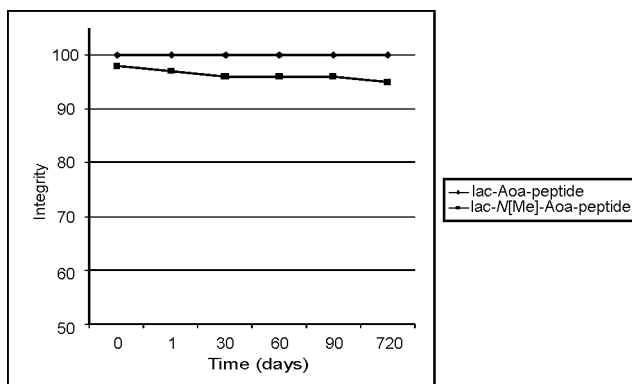


Figure 3.7 Long-term stability study of lyophilized *N*[Me]-Aoa glycopeptide (*i.e.* Gal- β 1,4-Glc-*N*[Me]-Aoa peptide).

3.3 Application in SPR

In previous work [119], the glycopeptide GlcNAc₅-Aoa-GFAKKG was immobilized onto the SPR sensor surface for interaction studies between the pentasaccharide GlcNAc₅ and the lectin WGA. The kinetic data obtained by SPR were comparable to those obtained by other methods, and therefore the proof of principle of this alternative approach to study carbohydrate-lectin interactions was established. The GFAKKG sequence used to immobilize sugar epitopes onto the sensor surface was designed with several criteria in mind: (1) Two mandatory amino acids, the N-terminal Aoa for chemoselective sugar ligation and two Lys residues near the C-terminus to ensure both proper guiding to the negatively charged sensor surface and efficient covalent (peptide bond) binding to the activated carboxyl group on the SPR chip surface were required. (2) In addition, an aromatic Phe residue was included to increase hydrophobicity and facilitate purification by HPLC and quantification by UV-spectrophotometry, while (3) Gly and Ala residues were added to provide flexibility and distance between the dextran surface and the glycan epitope.

3.3.1 Aglycon features: Peptide length

As our interest was focused on presenting sugar epitopes (mono-, di- and trisaccharides) smaller than the chitopentaose used in our previous work, a preliminary study was conducted on the length of the peptide module in the glycoprobe and its effect on sugar recognition by lectins. Thus, three glycoprobes with lactose as sugar epitope and different peptide lengths were tested: (1) the standard prototype (*i.e.*, Aoa-GFAKKG-amide), (2) an elongated (Aoa-GAGAGFAKKG-amide) and (3) a shorter (Aoa-GFKKG-amide) version. On attachment to a Biacore CM5 chip, immobilization levels of 330, 1250 and 240 RU were respectively obtained, which could also be expressed as 0.34, 1.0 and 0.26 pmol of immobilized glycoprobe per mm², respectively¹.

In order to evaluate the binding response of each of the above glycoprobes, two β -galactose specific lectins from *Erythrina cristagalli* (ECA) and *Ricinus communis* (RCA) were flown across the surface.

¹ We questioned if the manufacturer's suggested equivalence of 1 RU= 1 pg of immobilized analyte, devised for globular proteins, was indeed applicable to the present case, where the ligand is a short glycopeptide. To clarify this point, specific refractive indexes for each peptide were calculated from amino acid compositions by the Lorenz and Lorents equation [165,166] and indeed found rather similar to the above value. Thus, 0.981 RU= 1 pg equivalences were calculated for both standard GFAKKG and shorter GFKKG, and 0.977 RU= 1 pg for longer GAGAGFAKKG.

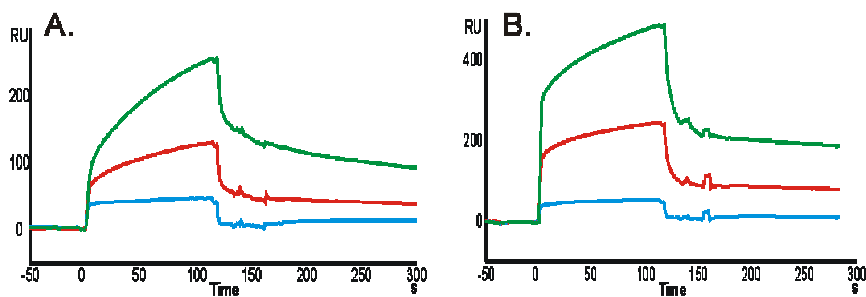


Figure 3.8 Sensorgrams of the interaction of ECA (panel A) and RCA (panel B) with the immobilized glycopeptides (blue trace: Aoa-GAGAGFAKKG-amide; green trace: Aoa-GFKKG-amide; red trace: Aoa-GFAKKG-amide; all conjugated with lactose)

As depicted in Figure 3.8, for both lectins the highest response was observed for the channel containing lactose-Aoa-GFKKG-amide. Despite the similar immobilization levels (*i.e.* $\sim 0.3 \text{ pmol mm}^{-2}$) of both surfaces, an almost three-fold increase in absolute binding response of ECA was observed for the shortest glycoprobe in comparison to the standard version. The same difference was found when the affinity constants from both surfaces were determined (*e.g.*, for ECA, $3.3 \mu\text{M}^{-1}$ and $1.2 \mu\text{M}^{-1}$, respectively). For the elongated version, in contrast, no binding was detected. As this could be due not only to the larger peptide size but also to steric hindrance, which on surfaces with high immobilization level (*e.g.* 1 pmol mm^{-2}) may also play an important role, no valid conclusions could be drawn and the surface subsequently been used as reference for non-specific binding events in further kinetic studies.

In summary, a peptide longer than GFAKKG was not necessary for efficient display of disaccharides on the sensor surface. In contrast, at least for ECA and RCA, the one-residue shorter glycopeptide Aoa-GFKKG-amide provided significantly better recognition than standard Aoa-GFAKKG-amide. These results emphasize the importance of the aglycon feature, in the sense that slight changes in peptide sequence (one Ala residue) can dramatically affect binding responses and consequently affinity constants. As the glycoprobe providing the highest binding

RESULTS

responses, the shorter Aoa-GFKKG-amide peptide was the version employed in further experiments.

Additionally, in order to ensure that lectins were binding specifically to the sugar and that the peptide module did not have any effect in the binding event, competition experiments were done with both lectins (ECA and RCA) at a fixed concentration in the presence of increasing concentration of peptide (Aoa-GFKKG-amide) in solution. As shown in Figure 3.9, no decrease in binding response was observed, indicating that the immobilized peptide module could be a good reference surface.

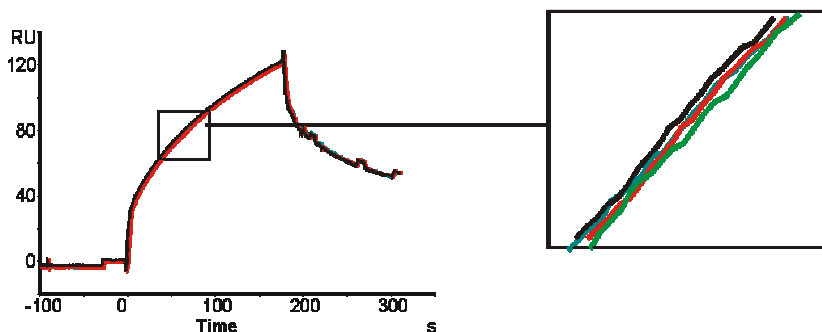


Figure 3.9 Sensorgrams of ECA interaction (250 nM) with immobilized lactose-Aoa-GFKKG-amide in presence of increasing peptide concentrations (0, 250, 500, 1000 nM).

3.3.2 *N*-methylated vs non *N*-methylated probes

As shown in 3.1.2, remarkable differences (100% vs. 30%) in the relative abundance of the native-like ring form of the peptide-proximal glycan are observed between *N*-methylated and non *N*-methylated versions of the glycoprobe. An obvious next step was to evaluate the impact of this situation in the recognition of small sugar epitopes (*i.e.* disaccharides) by lectins.

For this purpose, both peptide modules (with and without *N*-methylation) were conjugated with the same lacNAc disaccharide (*i.e.* Gal- β 1,4-GlcNAc) and immobilized onto a CM5 sensor chip under identical

coupling chemistries providing similar immobilization levels (~ 150 RU) enabling accurate comparison. Then three lectins (ECA, RCA and WGA) with different carbohydrate requirements (mono- or disaccharide) were passed at 250 nM. Unglycosylated *N*-methylated peptide was immobilized as reference surface to correct for matrix effects and non-specific binding. As shown in Figure 3.10 clearly different responses of the two glycoprobes were observed for each lectin.

For ECA, (Figure 3.10, panel A), a lectin described to interact strongly with lacNAc and requiring the entire disaccharide in a ring closed Gal- β 1,4 configuration for binding [167], fairly low and relatively strong responses was respectively observed for the non-methylated (red trace) and methylated (green trace) glycopeptides. Measurements just before dissociation showed the latter probe (fully closed ring structure) to have approximately three-fold higher absolute response values (~ 460 RU vs ~ 180 RU). Equally important, and in contrast to previous reports [168], the methyl group at the exo-anomeric center does not interfere with binding. In an attempt to achieve similar response levels for both probes, a higher immobilization level of approximately 400 RU (*i.e.*, threefold) for the non-methylated probe was tested. However, under these conditions ECA binding at 250 nM was completely abolished. Evidently, at these surface densities, steric hindrance seriously hampers the binding. It seems safe to assume that for other oxime-linked platforms, similar effects will play a role and, therefore, validation of the surface with a standardized lectin will be of utmost importance to accurately interpret binding results.

RESULTS

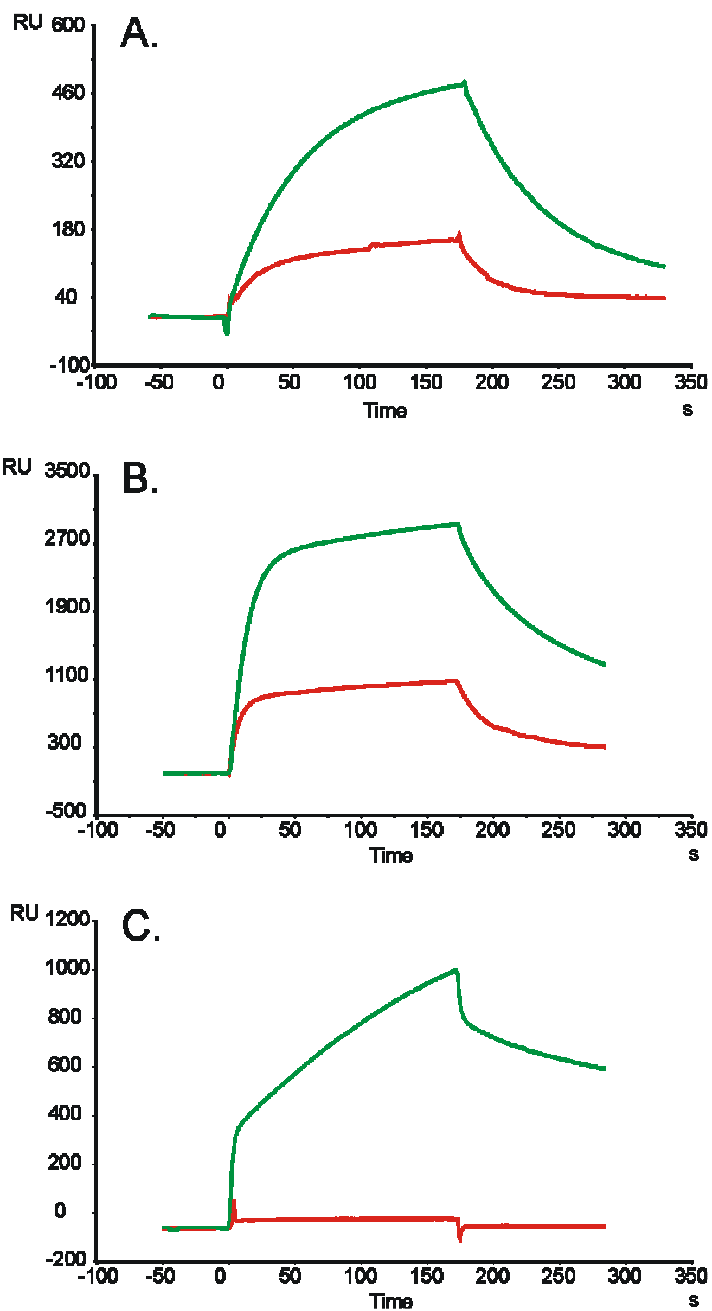


Figure 3.10 SPR sensorgrams of the binding of different lectins (A. ECA; B. RCA; C. WGA) at 250 nM to lacNac-glycoprobes. lacNac-Aoa-peptide: red trace and lacNac-N[Me]-Aoa-peptide: green trace)

Binding of the other lactosamine-binding lectin, RCA, to both glycoprobes is depicted in Figure 3.10, panel B. This lectin has a described specificity for terminal galactose [167], but its binding is enhanced when the monosaccharide is engaged in a type-2 structure (*i.e.* Gal- β 1,4-GlcNAc). In our system, RCA displayed significant binding to the non-methylated glycopeptides (~ 1100 RU) but nonetheless of lower magnitude than the methylated variant, with response levels again around three-fold higher (> 3000 RU absolute binding).

As a final test, WGA was passed across the surface (Figure 3.10, panel C). WGA reportedly binds specifically to GlcNAc, preferentially when terminally exposed [169]. However, on the basis of the sensorgram in Figure 3.10 (panel C), this lectin appears capable of also recognizing non-terminal (*i.e.*, internal GlcNAc) residues in its carbohydrate binding site [170]. The differences between non-methylated (zero binding) and methylated (tight interaction) are quite remarkable in this particular case. The absolute lack of recognition for the non-methylated variant is puzzling, as ~ 30% of the structure is expected to be in closed ring form at a given time. The only plausible explanation is that the dynamic equilibrium between ring and open sugar forms in the non-methylated glycopeptide renders the interaction so unstable that no binding can be observed. This observation is of extreme importance when this or other similar glycoprobes displaying non-native (*i.e.*, open sugar) epitopes are applied for screening purposes, as zero-binding could lead to a false interpretation.

The different binding responses between *N*-methylated and non-methylated Aoa glycopeptides were further corroborated by the association and dissociation constants derived for the three lectins (ECA, RCA and WGA). As depicted on Table 3.7, in all three cases higher association and lower dissociation constants were obtained for the

RESULTS

methylated glycoprobe, strongly suggesting that the existence of four different (*syn* and *anti* open forms, plus α and β anomers) glycan structures on the non-methylated glycoprobe not only influences absolute binding but also the equilibrium. For ECA, the lectin that requires the closed ring Gal- β 1,4 configuration, a remarkable (again) three-fold higher affinity constant (K_A) was determined for the methylated glycoprobe over its non-methylated variant. In contrast, for RCA the differences in affinity were not so evident.

Table 3.7 Affinity (K_A) and dissociation (K_D) constants of lectins (ECA, RCA, WGA) to lacNAc-Aoa-peptide and lacNAc-N[Me]-Aoa-peptide.

Lectin	Immobilized ligand	K_A (M^{-1})	K_D (M)
ECA	lacNAc-Aoa-peptide	1.93×10^6	5.19×10^{-7}
	lacNAc-N[Me]-Aoa-peptide	5.85×10^6	1.71×10^{-7}
RCA	lacNAc-Aoa-peptide	4.75×10^7	2.11×10^{-8}
	lacNAc-N[Me]-Aoa-peptide	5.29×10^7	1.89×10^{-8}
WGA	lacNAc-Aoa-peptide	-	-
	lacNAc-N[Me]-Aoa-peptide	2.31×10^7	4.33×10^{-8}

3.3.3 Affinity and kinetic studies of carbohydrate-lectin interactions: legume lectins as a proof of principle

With the peptide N[Me]-Aoa-GFKKG-amide chosen as the best option for native-like display of small carbohydrate epitopes at the sensor surface, the next phase of our study focused on naturally-occurring (mammalian) carbohydrates and on the lectins to those sugar units.

A first concern in this regard was efficient ligand immobilization (Figure 3.11). For this purpose, a pre-concentration of the ligand (in our case, the glycopeptides probe) close to the activated surface is necessary. A 10-50 $\mu\text{g ml}^{-1}$ ligand concentration in 10 mM NaOAc pH 4.0 is recommended for proteins [171]. At this pH (slightly below the isoelectric point of the

protein), the ligand is sufficiently charged to be concentrated on the surface, but still preserves uncharged amine groups capable of binding covalently to the activated NHS-esters (Figure 3.11).

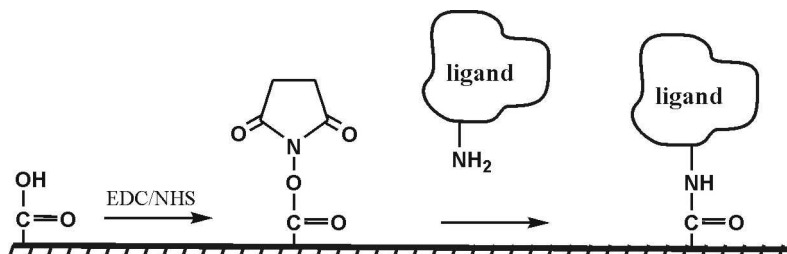


Figure 3.11 Scheme of amine coupling chemistry of ligands to the dextran surface.

In a first attempt, glycopeptide immobilization was done at pH 4.0. Although the immobilization level achieved was acceptable (~ 100 RU, equivalent to ~ 98 pg and thus ~ 0.08 pmols), when lectin was subsequently passed over the surface no significant binding was observed (Figure 3.12, red trace). In the next attempt, and considering that the theoretical isoelectric point of the glycopeptides is ~ 10 , significantly higher than a “standard” protein, a pH 6.0 buffer was used, and binding to the corresponding lectin could be appreciated (Figure 3.12, green trace), corroborating the appropriateness of the generated surface.

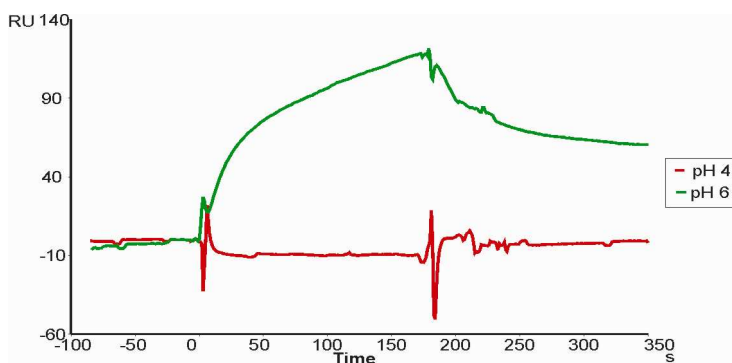


Figure 3.12 SPR sensorgram of the binding of SNA at $1 \mu\text{M}$ to Sia- α 2,6-lac-glycophores immobilized at pH 6 (green trace) or at pH 4 (red trace).

RESULTS

A plausible explanation for the unproductive glycoprobe immobilization (*i.e.*, lack of lectin interaction) at pH 4 may be that, at this pH, with its two Lys side chains positively charged, the glycoprobe is strongly cationic and its electrostatic interaction with the carboxylate dextran matrix is equally strong, driving the peptide “deep down” into the matrix, close to the gold surface, where surface accessibility is suboptimal. In contrast, at pH 6, the electrostatic attraction would be less strong and the display of the sugar epitope would be more confined to the dextran surface, hence favoring lectin recognition. For further experiments, NaOAc pH 6.0 was thus chosen as recommended buffer for immobilization.

An equally important goal of immobilization chemistry is obtaining similar levels at all four flowcells to ensure appropriate comparison between different sugar epitopes. For this purpose an accurate estimation of glycopeptide concentration is of utmost importance. To this end, small glycopeptide amounts (*i.e.* 100-200 μg) were quantified by UV at 258 nm (λ_{max} of a Phe residue, [172]) in a nanodrop system. An experimental extinction coefficient of $\epsilon=0.1438 \text{ mM}^{-1} \text{ cm}^{-1}$ for the non-conjugated N[Me]-Aoa-GFKKG-amide peptide was in this way obtained and used for UV quantitation of all glycoprobes before immobilization.

Monosaccharides- Once a cyclic structure for the peptide-linked glycan was ensured by conjugation to the *N*[Me]-Aoa-peptide, the next step was to test whether the glycoprobe, exposing a minimal sugar epitope, was sufficient to distinguish between lectins with different primary specificities. As a proof of principle, the Gal-*N*[Me]-Aoa-peptide was tested with RCA and ECA. As depicted in Figure 3.13, binding could be observed, although with the intrinsically low affinity that characterizes

monosaccharide-driven interactions. The modest but nonetheless analytically useful response of this Gal-displaying surface is worth noting, as well as the possibility (not explored in this work) of extending the principle to sensor chips with the most common monosaccharides as a first step in screening for new lectins.

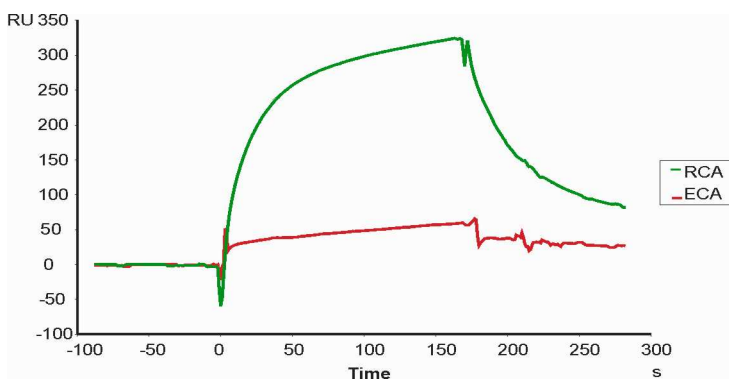


Figure 3.13 SPR reponse for the binding of β -galactose specific lectins (green trace: RCA and red trace: ECA) to Gal-*N*[Me]-Aoa-glycprobe.

Disaccharides- A further step in characterizing the carbohydrate specificity of lectins by our SPR platform was a sensor chip where three disaccharides only differing in the glycosidic linkage between the first and second monosaccharide units were compared. Thus, glycopeptides with *N*-acetyl-lactosamine (Gal- β 1,4-GlcNAc, type II glycans), lacto-*N*-biose (Gal- β 1,3-GlcNAc, type I glycans) and β 1,6 galactosyl-*N*-acetyl glucosamine (Gal- β 1,6-GlcNAc) conjugated to the *N*[Me]-Aoa-peptide were immobilized. In addition, a glycprobe with lactose (Gal- β 1,4-Glc) was also tested, to evaluate the importance of the *N*-acetyl group at position C2. For reference surface, and given that all lectins have a certain degree of promiscuity towards other glycosidic entities, a non-glycan ligand such as unconjugated *N*[Me]-Aoa-peptide was preferred as universal probe. Immobilization levels of ~ 500 RU and ~ 120 RU were obtained for reference and glycprobes, respectively. Several lectins

RESULTS

extensively studied in terms of carbohydrate binding such as WGA, ECA and RCA were flown at different concentrations over the sensor chip surfaces.

ECA and RCA are described to interact more strongly with Gal- β 1,4-GlcNAc than with other *N*-acetyl-disaccharides [167]. However, the nature of the glycosidic linkage and of the proximal monosaccharide turned out to be of decisive importance. As depicted in Figure 3.14, panel A, for ECA a strong preference to Gal- β 1,4-GlcNAc was clearly seen, along with 9-fold lower binding to Gal- β 1,4-Glc, revealing a crucial role for the *N*-acetyl group at position C2. Moreover, by comparing the structures of the three *N*-acetyl-disaccharides a more accurate analysis of the role of non-terminal sugar hydroxyl groups involved in the interaction was possible. Thus, binding was dramatically reduced for β 1,3 (21%) and β 1,6 (5%) when compared to β 1,4 linkage (100%). For the two former disaccharides, whose C3 and C6 hydroxyls participate in the glycosidic bond and are thus unavailable for ECA interaction, this dramatic decrease in binding response would indirectly suggest a relevant function for both hydroxyls in the binding to ECA. In summary, from the SPR sensorgrams it can be confirmed that the *N*-acetyl group, and to a lesser degree the hydroxyls at C3 and C6, participate in sugar-ECA recognition, in good agreement with structural data showing the O3, N2, O6, O_{NAC} atoms to be involved in the interaction [173].

The recognition pattern for RCA was similar to that of ECA (Figure 3.14, panel B), although the participation of the C3 and C6 hydroxyls of the proximal sugar seemed not to be as determinant as for ECA, suggesting that in this case the interacting entities are located principally in the Gal residue. Also, the *N*-acetyl-group at C2 appears to be less relevant than

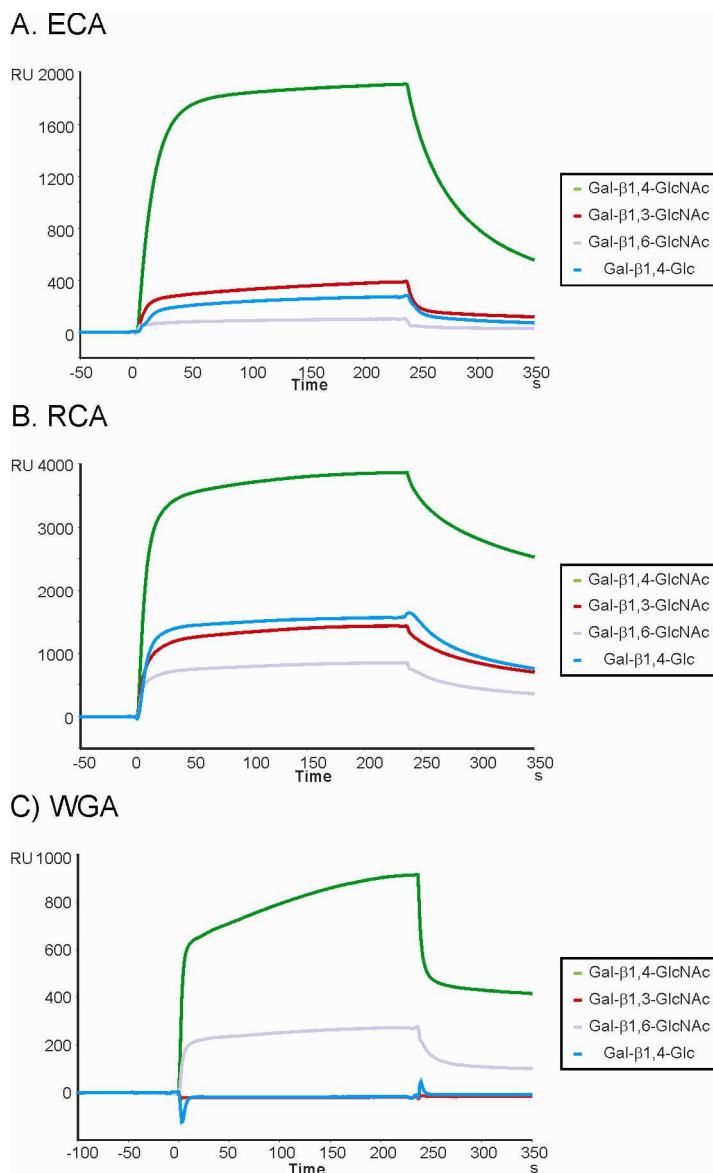


Figure 3.14 SPR sensorgrams of three legume lectins on different β -galactosides exposing glycoprobes. Glycoprobes displaying Gal- β 1,4-GlcNAc, Gal- β 1,3-GlcNAc, Gal- β 1,6-GlcNAc and Gal- β 1,4-Glc were coupled to the sensor surface with similar immobilization levels. The sensorgrams shown corresponded to the differential curves after subtracting the reference channel with the peptide without sugar epitope immobilized.

RESULTS

for ECA, since binding decreased when Gal- β 1,4-Glc instead of Gal- β 1,4-GlcNAc was employed, but to a lesser extent than observed for ECA. Therefore, comparing the binding response observed for each of the three glycoprobes, the O and N atoms of the peptide-linked sugar can be ranked as $O_6 > N_2 \approx O_3$ in terms of binding importance.

Like ECA and RCA, WGA interacts preferentially with Gal- β 1,4-GlcNAc, but also binds to Gal- β 1,6-GlcNAc (Figure 3.14, panel C), underlying its ability to recognize internal GlcNAc. The fact that neither Gal- β 1,3-GlcNAc nor Gal- β 1,4-Glc are recognized by WGA corroborates the requirements for an *N*-acetyl function at C2 and an hydroxyl at C3, this latter feature hitherto unreported for WGA. In this case, the order of importance between atoms of the proximal glycan is $N_2 \approx O_3 > O_6$ (Figure 3.14, panel C).

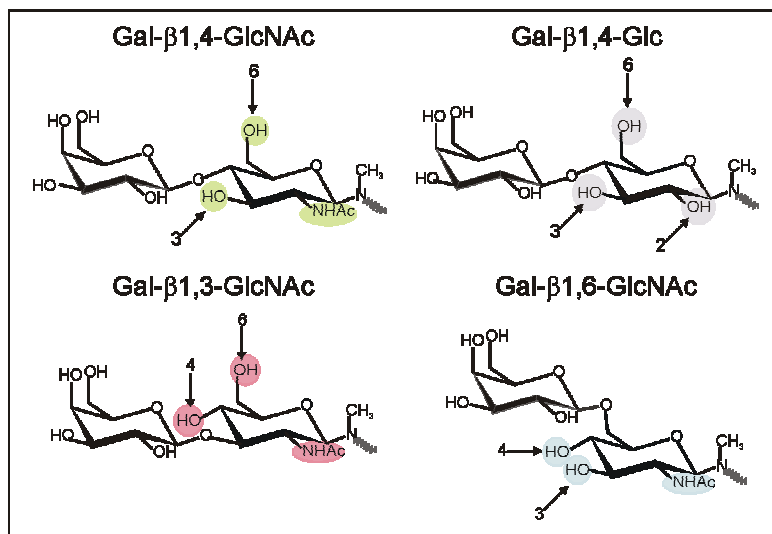


Figure 3.15 Schematic representation of sugar epitopes immobilized on the sensor surface. Hydroxyl and *N*-acetyl groups that may participate in the interaction are indicated by circles.

While simple visual inspection of the sensorgrams allows a qualitative assessment of the relevance for interaction of each hydroxyl and *N*-acetyl group of the peptide-linked glycan, a more complete, quantitative analysis is also possible by monitoring both the association and dissociation events and their kinetic rate constants (k_a , k_d) plus the derived affinity constant (K_A) (Table 3.8).

Table 3.8 Kinetic rate constants and the derived affinity constant (K_A) of lectins (ECA, RCA, WGA) to different glycoprobes exposing terminal β -galactosyl-disaccharides.

Lectin	Disaccharide	k_a	k_d	K_A
ECA	Gal- β 1,4-GlcNAc	4.7×10^4	5.3×10^{-3}	8.9×10^6
	Gal- β 1,4-Glc	3.5×10^3	5.8×10^{-3}	6.0×10^5
	Gal- β 1,3-GlcNAc	2.3×10^3	3.2×10^{-3}	7.2×10^5
	Gal- β 1,6-GlcNAc	-	-	-
RCA1	Gal- β 1,4-GlcNAc	9.0×10^4	1.3×10^{-3}	6.9×10^7
	Gal- β 1,4-Glc	1.7×10^5	3.8×10^{-3}	4.5×10^7
	Gal- β 1,3-GlcNAc	7.5×10^4	3.6×10^{-3}	2.1×10^7
	Gal- β 1,6-GlcNAc	2.4×10^4	4.54×10^{-3}	5.3×10^6
WGA	Gal- β 1,4-GlcNAc	1.4×10^4	$6,6 \times 10^{-4}$	2.1×10^7
	Gal- β 1,4-Glc	-	-	-
	Gal- β 1,3-GlcNAc	-	-	-
	Gal- β 1,6-GlcNAc	1.8×10^4	2.4×10^{-3}	7.5×10^6

Another group of disaccharides, *i.e.* mannobioses, and their interaction with Con A has been studied as well- Glycoprobes correctly displaying mannobioses that only differ in their glycosidic linkage were immobilized on a sensor chip, with 903 RU, 341 RU and 344 RU levels obtained for Man- α 1,2-Man, Man- α 1,3-Man and Man- α 1,6-Man, respectively. Since an exactly equal amount of glycoprobe (*e.g.* 60 μ l, 1 g L⁻¹ in NaOAc pH6)

RESULTS

was injected onto the three surfaces (see Section 3.3.3), the high value for Man- α 1,2-Man-N[Me]-Aoa-GFKKG could only be due to a sensor surface of particularly high substitution level, resulting in anomalously high activation levels (470 RU vs 200 RU). Even so, the Man- α 1,2-Man surface turned out to be appropriate for interaction with the mannose-specific lectin Con A, to be detected. Given the reported ability of Con A to bind carboxymethylated dextran surfaces [174], subtraction of the signal from the reference cell (non-conjugated N[Me]-Aoa-peptide, see Section 3.3.2) was of crucial importance to assess the glycan-specific binding of Con A.

Con A, a relatively complex lectin, at pH > 7 is organized as a tetramer with two CRDs per dimer situated on opposite faces of the protein (Figure 3.16).

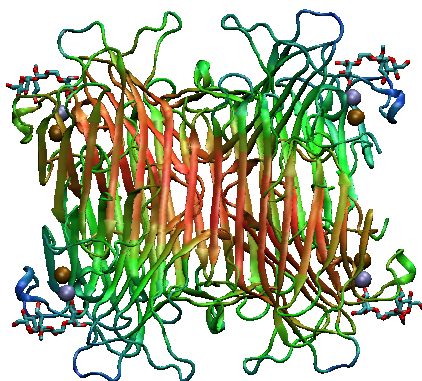


Figure 3.16 Biological assembly of Con A (PDB 3D4K) in complex with mannose.generated with VMD software. [175].

In order to evaluate differences in Con A affinity between mannobioses, sensorgrams of different Con A concentrations were recorded and kinetic constants determined by simultaneously fitting the experimental curves to a bivalent kinetic model (Figure 3.17). This results in the first and second binding events being separately described by two sets of rate constants

(Table 3.9). The unconventional units (*i.e.* RU⁻¹) used to describe the second association event make comparison with constants derived from other methods (*e.g.* ITC) difficult, and only the affinity constants for the first event, K_{A1} ($K_{A1} = k_{a1}/k_{d1}$) can be compared. For this event, higher affinity Man- α 1,3-Man ($K_{A1} = 5.9 \times 10^6 \text{ M}^{-1}$) over Man- α 1,6-Man ($K_{A1} = 2.2 \times 10^6 \text{ M}^{-1}$) was found in agreement with previous ITC data [176]. The third mannobiose, Man- α 1,2-Man, had a lower K_{A1} than the other two, possibly due to mass transport problems common on surfaces with high immobilization levels.

Comparison of SPR- and ITC derived affinity values for the first binding event is noteworthy. For the disaccharides - Man- α 1,3-Man and Man- α 1,6-Man, Con A K_{A1} values obtained with the first method ($\sim 10^6$) are two orders of magnitude higher than with the second ($\sim 10^4$) [176]. One possible explanation for these high SPR values, keeping in mind the two-CRDs-per-dimer model of Con A, is that although the bivalent binding model used in SPR only allows to derive standard-unit values for K_{A1} (see above), the apparent k_{d1} , used for K_{A1} determination is also influenced by the second event. If this contribution is not factored in, a lower k_{d1} , and consequently a higher K_{A1} results.

Table 3.9 Kinetic constants of ConA to the three different mannobioses-containing glycoprobes determined by SPR

Ligand	$k_{a1} (\text{M}^{-1}\text{s}^{-1})$	$k_{d1} (\text{s}^{-1})$	$k_{a2} (\text{RU}^{-1} \text{s}^{-1})$	$k_{d2} (\text{s}^{-1})$
Man- α 1,2-Man	2.07×10^4	4.58×10^{-2}	1.54×10^{-3}	2.83×10^{-2}
Man- α 1,3-Man	5.01×10^4	8.50×10^{-3}	1.46×10^{-5}	3.97×10^{-3}
Man- α 1,6-Man	2.44×10^4	1.07×10^{-2}	9.81×10^{-6}	6.08×10^{-3}

RESULTS

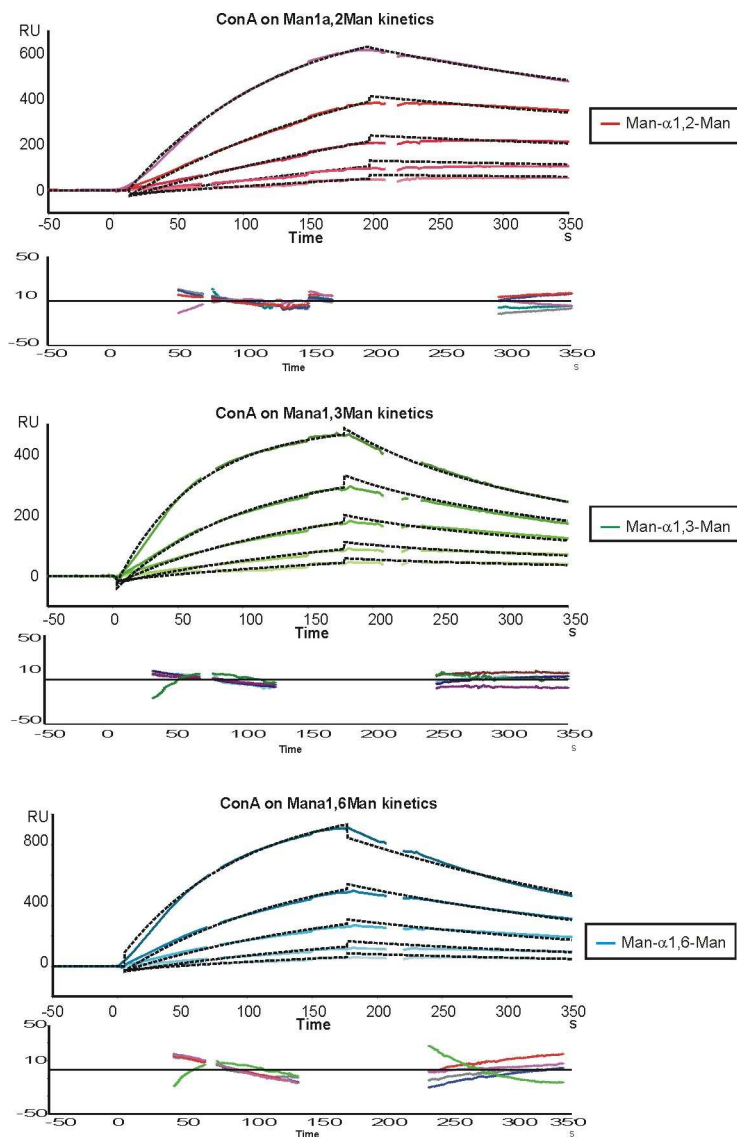


Figure 3.17 Binding of ConA to immobilized mannobioses at five different concentrations (5, 10, 20, 40 and 80 nM). Fitting curves using a bivalent model are shown in discontinuous black lines. Below each set of sensorgrams, the corresponding plot of residuals, indicating the difference between experimental and fitted data is shown.

A last group of disaccharides to be analyzed with our SPR platform included those with terminal Fuc units (Fuc- α 1,3-GlcNAc, Fuc- α 1,4-GlcNAc and Fuc- α 1,6-GlcNAc). The latter structure is a potential fucosylation core of N-glycans, and the former two are partial epitopes of Lewis X and Lewis A structures, respectively. Immobilization levels of 73 (~ 0.08 pmol), 236 (~ 0.25 pmol) and 116 RU (~ 0.12 pmol) for the α 1,2, α 1,3 and α 1,6-linked glycopeptides were respectively obtained and the resulting surfaces were tested against two Fuc-specific lectins from *Lotus tetragonolobus* (LTA) and *Ulex europeaus* (UEA) (Figure 3.18, panel A and C). In order to test the specificity of the interaction, a competition experiment with soluble α Me-Fuc and α Me-Man was carried out. For both lectins, the sensorgrams demonstrated that the binding response decreased with increasing amounts of α Me-Fuc, but not of α Me-Man (Figure 3.18, panel B and D). Carbohydrate affinity for LTA has been described to be strongly specific for the blood group H determinant (Fuc- α 1,2-Gal) [177] and Le^x trisaccharide (*i.e.* Gal- β 1,4[Fuc- α 1,3]GlcNAc) [178]. In larger fucosylated oligosaccharides, the position and the linkage become critical [179]. On our sensor chip no clear binding was observed to any of the fucosylated glycoprobes, confirming the weak binding reported previously for structures carrying Fuc- α 1,(3,4)-linked to GlcNAc [88]. Therefore, we tentatively conclude that LTA needs the entire Le^x trisaccharide for binding to be clearly observed.

RESULTS

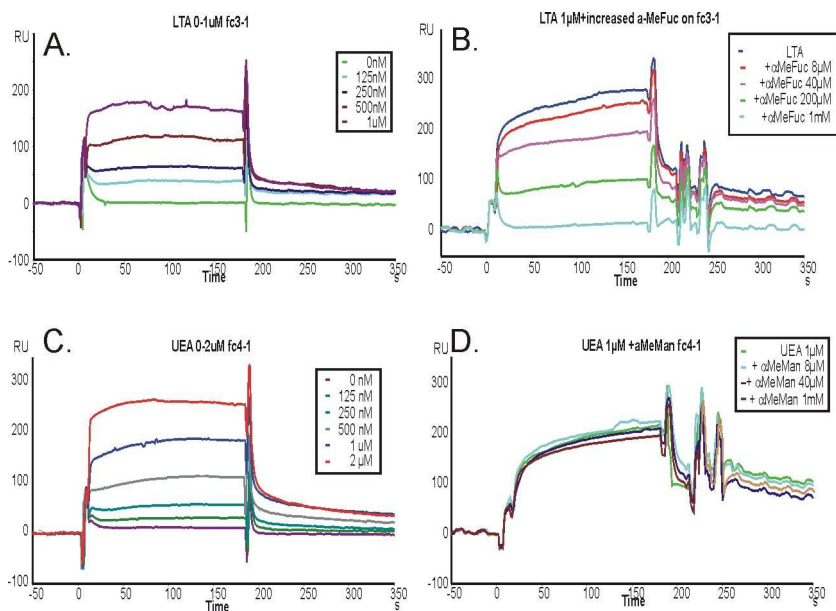


Figure 3.18 Binding sensorgrams of two fuc-binding lectins (A.: LTA; C.: UEA) at five different concentrations from 125 nM to 1 μ M, to Fuc- α 1,4-GlcNAc-glycoprobe (A, B) and Fuc- α 1,6-GlcNAc-glycoprobe (C, D). Competition experiments with increasing concentrations of α Me-Fuc (B) and α Me-Man (D).

Trisaccharides- A protein that reportedly requires three intact sugar units for binding is the sialic acid-specific lectin from *Maackia amurensis* (MAA) [180]. Sialic acids are generally found at the non-reducing end of a glycan and therefore are of the utmost importance for relevant cell communication processes [181]. A sensor chip was prepared with the trisaccharides Sia- α 2,3- and Sia- α 2,6-lacNAc, representative of the two existing linkage types in adult mammalian glycoproteins. As a reference surface, in this setting a lactose-displaying probe (instead of the hitherto used unglycosylated peptide) was employed. The three glycoprobes were attached to the sensor surface with similar immobilization levels: 80 RU (~ 82 pg ~ 0.06 pmol) for the α 2,3-linked glycoprobe, and 56 RU (~ 57 pg ~ 0.04 pmols) for the two α 2,6-linked glycopeptide. After immobilization, two sialyl-specific lectins, from *Maackia amurensis* (MAA) and

Sambucus nigra (SNA), were flown over the surfaces. The different binding responses observed for the two lectins were in perfect agreement with their reported carbohydrate specificity. Thus, whereas SNA showed a marked preference for the Sia- α 2,6-lacNAc isomer, MAA recognized only the Sia- α 2,3-lacNAc-containing glycoprobe (Figure 3.19).

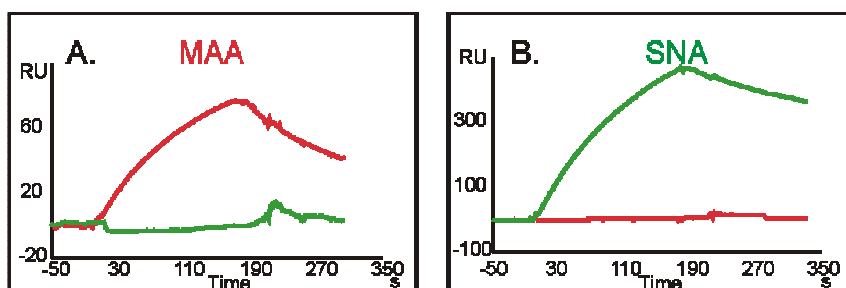


Figure 3.19 Binding sensorgrams showing the different carbohydrate specificity of both lectins (A: MAA; B: SNA) to their specific sugar epitopes (Sia- α 2,3-lacNAc and Sia- α 2,6-lacNAc in red and green trace, respectively)

For kinetic analysis, several concentrations in the 250 nM to 1.9 μ M range for MAA and in the 74 to 563 nM range for SNA were injected. A buffer injection without lectin was used as an internal standard and subtracted from the binding results for each flow before analysis. No mass-transport limiting effects were observed, as similar affinity constants were derived from experiments at either 10 μ l min^{-1} or 40 μ l min^{-1} flowrates. A 1:1 Langmuir binding model was chosen for sensorgram fitting (Figure 3.20, Table 3.10) because, although both lectins are divalent, in this kind of assay only one binding site is actually involved [89]. As expected for interactions involving trisaccharides, the equilibrium constants K_A obtained were in the range of 10^6 - 10^7 M^{-1} . For MAA, the apparent affinity constants determined by our approach was 9.12×10^5 M^{-1} . This constant is of the same magnitude as that reported previously [87], but one order lower than other constants determined using complex neoglycoproteins [182], in both cases by SPR. The higher affinity for neoglycoproteins can

RESULTS

easily be explained by their multivalent nature. For SNA, value of $6.27 \times 10^6 \text{ M}^{-1}$ was calculated for K_A , consistent with previously reported K_A $6.7 \times 10^6 \text{ M}^{-1}$ [89].

Table 3.10 Kinetic constants of MAA and SNA to their specific sialic-containing glycoprobes determined by SPR.

	glycoprobe	$k_a (\text{M}^{-1} \text{s}^{-1})$	$k_d (\text{s}^{-1})$	$K_A (\text{M}^{-1})$
MAA	Sia- α 2,3-lacNAc	5.55×10^3	6.08×10^{-3}	9.12×10^5
SNA	Sia- α 2,6-lacNAc	1.02×10^4	1.63×10^{-3}	6.27×10^6

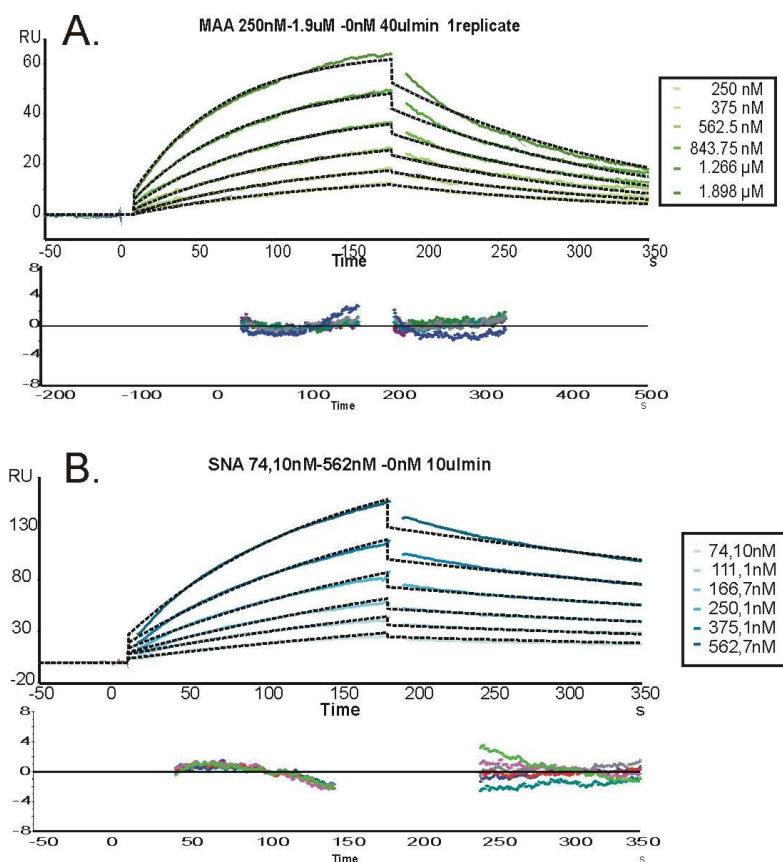


Figure 3.20 SPR sensorgrams of two Sia-specific lectins at six different concentrations. Binding of MAA to Sia- α 2,3-lacNAc glycoprobe (panel A) and SNA to Sia- α 2,6-lacNAc glycoprobe (panel B). Fitted curves are indicated as discontinuous black trace. The corresponding plot of residuals is shown below each set of sensorgrams.

3.3.4 Determination of thermodynamic parameters by SPR

SPR allows evaluating in real time the effect of temperature and thus deriving thermodynamical parameters for the interaction (see section 1.4.2). In order to explore the performance of our analytical platform in this regard, rate constants for the ECA-Gal- β 1,4-GlcNAc interaction were determined at different temperatures from 10 to 25 °C with 2.5 °C intervals. An overlay plot shows how temperature affects the binding profiles (Figure 3.21): both association and dissociation rate increase with temperature.

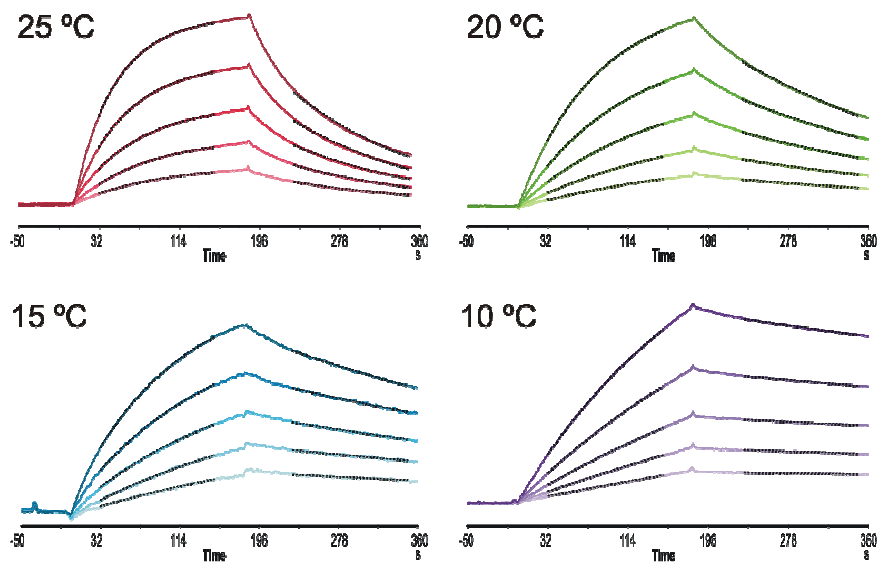


Figure 3.21 Effect of temperature on ECA binding to Gal β 1-4GlcNAc glycoprobe. Sensorgrams obtained for five concentrations of ECA (from 100 to 700 nM) at five different temperatures are shown. At each temperature, colored lines represent experimental data and black lines the local fitting to a 1:1 Langmuir model.

At each temperature, the association and dissociation rate constants were determined by locally fitting the five ECA concentrations during the same association and dissociation periods. As shown in Figure 3.21 and Table

RESULTS

3.11, ECA underwent approximately 5- and 10-fold increases in association and dissociation rates with temperature, respectively, which in turn caused a 2-fold decrease in the K_A .

Table 3.11 Kinetic data derived by locally fitting of the SPR sensorgrams at seven temperatures. The values of rate constants were calculated by the mean of the five concentrations measured by duplicate.

T / °C	$k_a / M^{-1} s^{-1} \times 10^{-4}$	$k_d / s^{-1} \times 10^3$	$K_A / M^{-1} \times 10^{-6}$	Chi ²
10	0.59	0.68	8.68	0.98
12.5	0.78	1.12	6.96	0.21
15	0.99	1.75	5.66	0.08
17.5	1.18	2.41	4.90	0.08
20	1.46	3.50	4.17	0.37
22.5	1.68	5.34	3.15	0.52
25	2.82	7.13	3.96	0.76

The equilibrium association constants were used to determine Van't Hoff enthalpies by plotting $\ln(K_A)$ vs $1/T$. As shown in Figure 3.22 panel A, the plot was fitted to a straight line and the enthalpy was calculated using the non-integrated form of the van't Hoff equation (eq. 3.1):

$$\ln K_A = -\Delta H^\circ/RT + \Delta S^\circ/R \quad (3.1)$$

where $R = 8.314 \text{ J K}^{-1} \text{ mol}^{-1}$. The ΔH value was derived from the slope of the line.

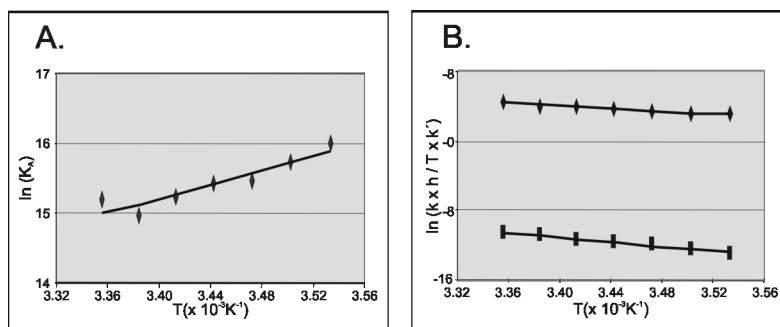


Figure 3.22 Van't Hoff (A) and Eyring (B) analysis of ECA binding to Gal-β1,4-GlcNAc-glycoprobe: ♦, association; ■, dissociation.

In addition, the possibility of monitoring independently the temperature effect on the association and dissociation phases, allows analyzing thermodynamic parameters using the Eyring equation (eq. 3.2),

$$\ln(k/T) = -\Delta H/RT + -\Delta S/R + \ln(k'/h) \quad (3.2)$$

where k is the appropriate rate constant (association or dissociation), and k' and h are the Boltzmann ($k' = 1.38 \times 10^{-23} \text{ J K}^{-1}$) and the Planck constants ($h = 6.63 \times 10^{-34} \text{ J s}$), respectively.

Figure 3.22, panel B, shows the overlay of Eyring plots represented by $\ln(k/T)$ vs $1/T$. Both plots could be fitted to a linear model for association and dissociation, respectively. As depicted on Table 3.12, ΔH values derived from Van't Hoff ($-41.9 \text{ kJ mol}^{-1}$) and from Eyring analyses ($-42.3 \text{ kJ mol}^{-1}$) were in agreement to the value of $-45.6 \text{ kJ mol}^{-1}$, reported by Gupta *et al.* by means of ITC [183].

In both technologies (SPR and ITC), the entropic value is calculated from the Gibbs free energy equation (eq. 3.3) as recommended [186].

$$\Delta G^\circ = \Delta H^\circ - T\Delta S^\circ = -R T \ln K_A \quad (3.3)$$

RESULTS

Table 3.12 Binding constants and thermodynamic parameters for ECA-Gal- β 1,4-GlcNAc interactions determined using SPR, ITC and NMR

	SPR		ITC ^{[[184]]}	NMR ^{[[185]]}
	Van 't Hoff	Eyring		
K_A / M^{-1}	3.9×10^6	4.0×10^6	1.1×10^4	6×10^3
$\Delta H / kJ mol^{-1}$	-41.9	-42.3	-45.6	-54.5
$\Delta S / J mol^{-1} K^{-1}$	-14.2	-15.7	-75.3	-102

The equilibrium association constants derived from SPR experiments (*e.g.* $K_A = 3.9 \times 10^6 M^{-1}$ at 25 °C) are two orders of magnitude higher than the binding constant reported by ITC for free lacNAc ($K_A = 1.1 \times 10^4 M^{-1}$ at 27 °C). These differences between surface- and solution-based methods have been already observed for other carbohydrate-lectin interactions [100,187], probably because the lectin multivalency in addition to the relatively high sugar immobilization level may increase the apparent affinity through secondary interactions. These differences in equilibrium association constants plus little differences in ΔH derived from both methods are responsible from the high differences in the corresponding entropic values (-15 vs -75 J mol⁻¹ K⁻¹). Even though for small molecule inhibitor-enzyme interactions the affinity and thermodynamic constants derived from both methods (SPR and ITC) were similar [70], with our SPR-based approach, where the sugar is immobilized, only equivalent enthalpy values were obtained.

3.3.5 Lectin affinity capture and characterization by SPR-MS

Although the primary application of the SPR-based technology is to characterize biomolecular interactions through the direct monitoring of binding and the determination of kinetic parameters, recently SPR-based instruments have been also used to capture interacting proteins from

complex mixtures for subsequent identification by mass spectrometry. Basically, in a standard ligand-capture experiment, performed in a BIAcore 3000 instrument, after injecting the analyte onto the glycoprobe-immobilized surface, the flow cell and the fluidic system are automatically washed with a MS-compatible buffer (*e.g.* 25 mM NH_4HCO_3) in order to clean the system and avoid carry-over of unreacted protein material to the recovered elution. Then, the captured lectin is specifically eluted through a recovery function, named MS recover. This command results in the collection of bound lectin in very small volume (2 μl), separated by air bubbles to avoid sample diffusion or cross-contamination and returning the captured protein automatically into a specific vial or even a MALDI target. Meanwhile during the capturing experiment, the SPR signal is used to monitor and quantify both the capture and recovery of the bound interacting partner.

In order to test the capacity of our SPR-based platform to capture carbohydrate-binding proteins for subsequent identification by mass spectrometry, three legume lectins with different molecular weights and different affinities were passed over their specific glycopeptide platform. In order to capture the highest amount of carbohydrate-binding protein on a single surface, the surface that displayed the highest affinity (*i.e.* Gal- β 1,4-GlcNAc) was selected to perform the capture experiment. The glycoprobe immobilization was carried out as described in Section 6.9.1 and an immobilization level of 165 RU was obtained, equivalent to 165 pg or 0.17 pmol of the glycoprobe immobilized. As with other sensor chips, the non-glycosylated peptide platform was immobilized as reference surface, albeit that this material is not collected.

The smallest lectin WGA, a dimer with a molecular weight of 34 kDa, was injected at 500 nM during 3 min. After washing the flow cells and the fluidic system with running buffer and 25 mM NH_4HCO_3 , a 400 RU

RESULTS

binding response, corresponding to 12 fmol bound material, was observed. For the next step (lectin elution and MS analysis, with concomitant surface regeneration), an appropriate recovery solution was required that (1) was compatible with MS-analysis, (2) was able to elute the bound protein in a small volume. As the solutions used thus far to regenerate the surface (*e.g.* 10 μ l of 10 mM lac, or 10 μ l of 0.5 M GlcNAc) did not fulfil either requirement, a solution with a lower concentration of a sugar with higher affinity (*i.e.* GlcNAc₂) was chosen. Upon 30s-incubation with 2 μ l of 10 mM GlcNAc₂, the bound lectin was completely eluted from the surface. The 2 μ l-sample was next concentrated by vacuum centrifugation and dissolved in 0.5 μ l of matrix solution for MALDI-TOF analysis. The recovered material (*i.e.* 12 fmol) was sufficient to visualise the molecular weight of the lectin by MALDI-TOF MS. Figure 3.23, panel A shows the capture sensorgram and the corresponding MALDI-TOF MS spectrum of the recovered material for WGA. A more exhaustive characterization was prevented by the distinctive complexity of WGA structure (170 residues and 16 disulfide bonds), which made proteolytic digestion of the recovered WGA impossible without further work-up.

A lectin of higher molecular weight, RCA, a 120 kDa tetramer, was passed at 80 nM over the same immobilized glycoprobe surface for 3 min, and the bound lectin (3800 RU = 3800 pg = 32 fmol) was recovered with 2 μ l of 10 mM lacNAc. As with WGA, the retained lectin could be subsequently characterized by MALDI-TOF MS, although due to its high molecular weight, signal intensity was much lower than for WGA (Figure 3.23, panel B).

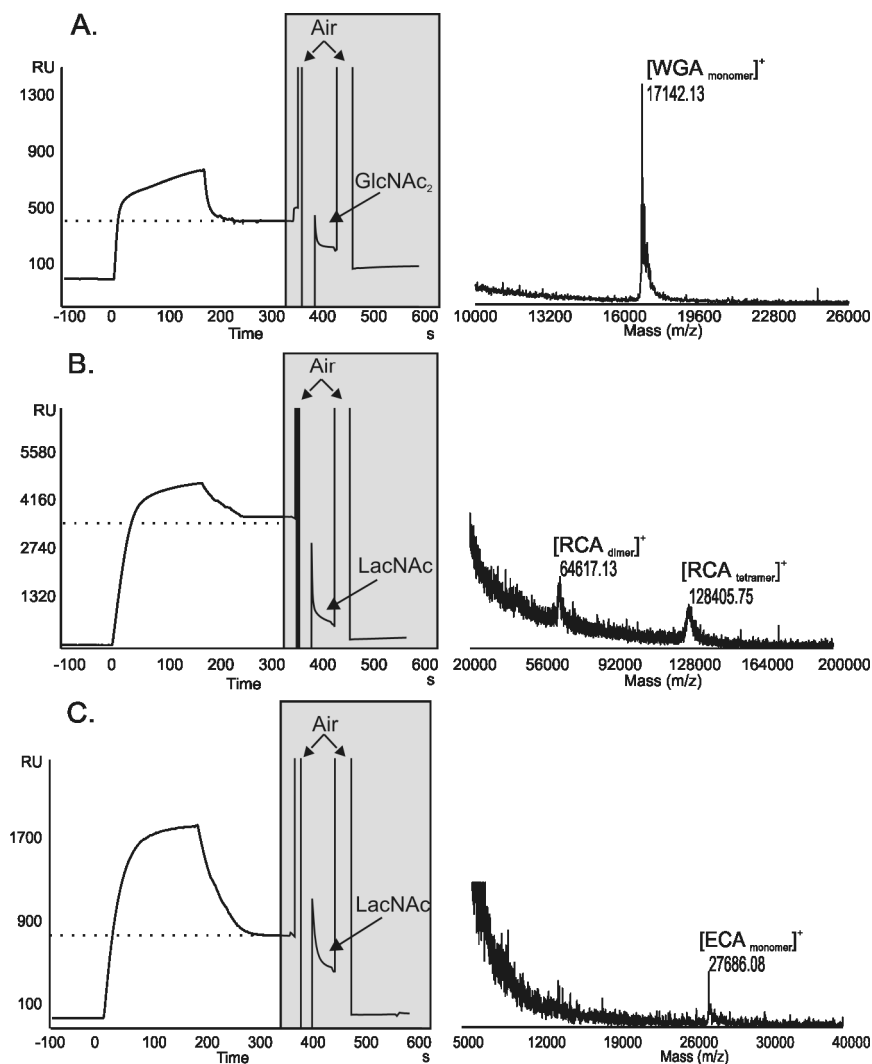


Figure 3.23 SPR sensorgrams of three recovery experiments over lacNAc-glycoprobe and MALDI-TOF MS spectra of the protein recovered. A.: a solution of WGA 500 nM were injected and the captured material was recovered with 2 μ l of GlcNAc₂; B.: RCA at 80 nM was passed on the surface and the recovery was done with 2 μ l lacNAc. C. ECA at 1 μ M was passed and the capture material eluted with 2 μ l lacNAc.

The capture of a third lectin, ECA, a 55 kDa dimer, was the least successful. As with WGA and RCA, ECA was injected over the glycosylated surface for 3 min. However, as its dissociation rate is larger than the other two lectins, ECA was injected at higher concentration (1

RESULTS

μM vs 80 nM). After washing, the surface-captured material was determined as 790 RU (790 pg, 14 fmol). Although efficiently eluted with 2 μl of 10 mM lacNAc, it was difficult to recover sufficient material for unambiguous MALDI-TOF MS analysis. One partial solution to the problem, considering the high dissociation rate of this lectin, was to perform the experiment at a lower temperature. Indeed, at 10 °C, both association and dissociation rates were diminished (data not shown). However, the slower dissociation required higher sugar concentration for the elution step, and this in turn complicated observation of the released lectin by MALDI-TOF MS.

In summary, a successful combination of SPR and MALDI-TOF analysis to characterize sugar-lectin interactions appears largely dependent on the dissociation rate constant of the complex and the ionization capacity of the lectin. Thus, for WGA both a low dissociation rate ($6.6 \times 10^{-4} \text{ s}^{-1}$) and a small molecular weight (34 kDa) facilitate capture and MS identification. In contrast, the high dissociation rate constant of ECA ($5.3 \times 10^{-3} \text{ s}^{-1}$) made recovery and subsequent analysis more complicated, especially at room temperature.

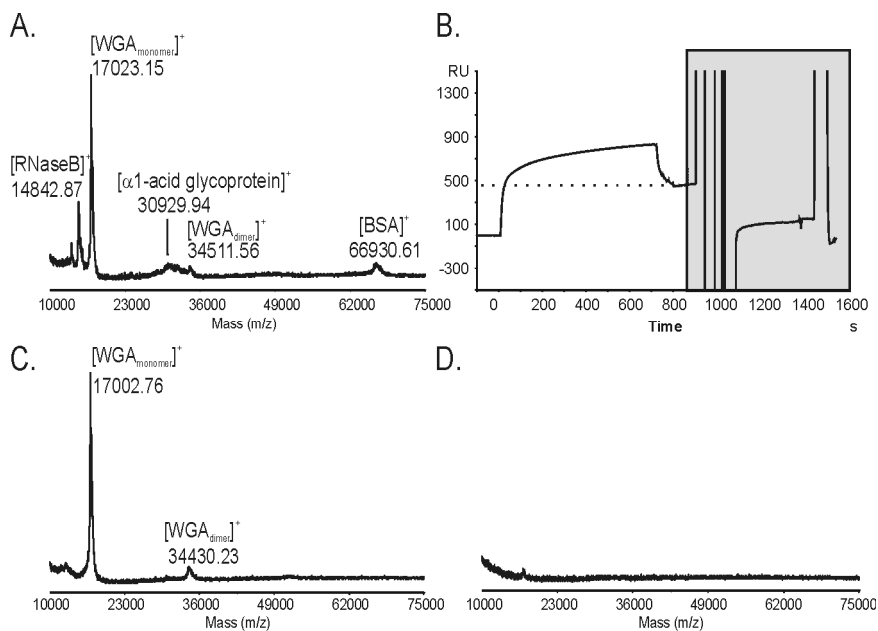


Figure 3.24 A. MALDI-TOF MS spectrum of the protein mixture containing BSA, RNase B, α 1-acid glycoprotein and WGA in equimolar concentrations. B. SPR sensorgram showing the corresponding recovery experiment over the lacNAc-immobilized glycoprobe. C and D. MALDI-TOF MS spectrum of the material eluted from the immobilized glycoprobe surface (panel C) and from the immobilized *N*[Me]-Aoa-peptide reference surface (panel D).

Finally, in order to evaluate the efficiency of our system for capturing carbohydrate-binding proteins in the presence of other proteins, a mixture containing BSA and glycoproteins such as RNase B and α 1-acid glycoprotein in equimolar concentrations was prepared. To this mixture, the lectin WGA was added and the solution passed over the reference (*N*[Me]-Aoa-peptide) and the glycosylated (lacNAc-*N*[Me]-Aoa-peptide) surfaces. After four consecutive injection and recovering cycles, the sample was concentrated and dissolved in 1 μ l of MALDI-TOF matrix solution. As seen in Figure 3.24, WGA is efficiently captured from the protein mixture and distinctly recovered for MS identification in the presence of other glycoproteins in the mixture.

3.4 Identification of carbohydrate recognition domain (CRD) by CREDEX-MS

As mentioned in 1.3.3 although X-ray crystallography and NMR are primarily used for structural studies of carbohydrate-lectin complexes, the sample amounts required in these techniques are not always easy to obtain. For scarce sample situations, the CREDEX-MS approach (Section 1.5.2) based on limited proteolysis and mass spectrometry, can become a viable alternative for elucidating carbohydrate-lectin interactions [138]. For instance, the CRDs of both galectin-1 and galectin-3 have recently been successfully delimited by CREDEX-MS [141]. In this instance, lactose was immobilized on Sepharose and both excision and extraction experiments were carried out. The affinity-bound peptides, eluted and identified by MS, were in perfect agreement with the crystal structure for galectin-3 in complex with lacNAc [188].

In order to test the complementarity of CREDEX-MS with the SPR-based glycoprobe approach for other carbohydrate-binding proteins, ECA was chosen. Several primary sequences, apparently reflecting point mutations, have been described for this 55 kDa protein, active as a dimer [189]. As many other legume lectins, ECA is N-glycosylated, at both Asn17 and Asn 113 [190]. From its crystal structure in complex with lactose, residues Leu86, Ala88, Asp89, Tyr106, Gly107, Phe131, Asn133, Ala218 and Gln219 can be identified as directly interacting with the sugar moiety (Figure 3.25) [191].

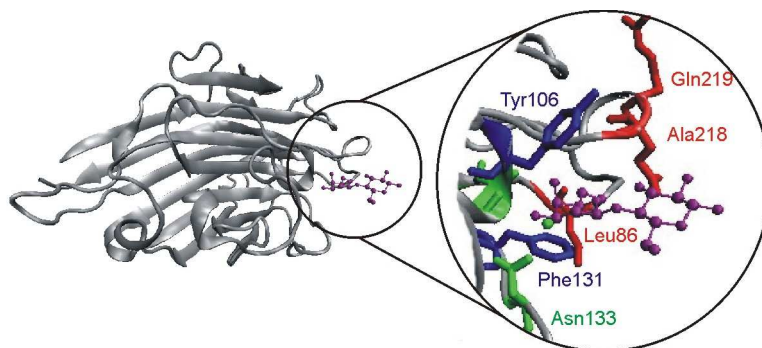


Figure 3.25 Ribbon representation of the crystal structure of ECA in complex with lactose (PDB 1GZ9). On the right, the carbohydrate recognition domain is enlarged. Key residues in the interaction are indicated in different colours: weaker hydrogen bonds (red), strong hydrogen bond network (green) and hydrophobic interactions (blue). Lactose is shown in purple (balls and sticks representation). This figure was generated with VMD OpenGL Display (version 1.8.3).

In this type of interactions, the CRD is generally made up of amino acid residues close in space though not necessarily in sequence. Thus, optimal digestion conditions to achieve highest sequence coverage are mandatory. Such conditions were established in solution digestions of ECA with common proteolytic enzymes (trypsin, chymotrypsin and glu-C). As shown in Figure 3.26, trypsin (1:20 w:w enzyme-ECA, 37 °C overnight) afforded a good sequence coverage (86%).

Before the excision experiment, the sugar epitope Gal- β 1,4-GlcNAc was immobilized onto divinylsulfonyl-functionalized Sepharose (DVS-S) [140]. This immobilization chemistry was chosen because of its expediency, although it is known to cause random carbohydrate immobilization [139] as it targets any possible hydroxyl function. In the excision experiment, ECA was incubated with the lacNAc-DVS-S during 24 h, the column was washed until no protein was observed by MALDI-TOF MS, then the lectin-sugar complex was digested with trypsin at 37 °C overnight (Figure 3.27, panel A).

RESULTS

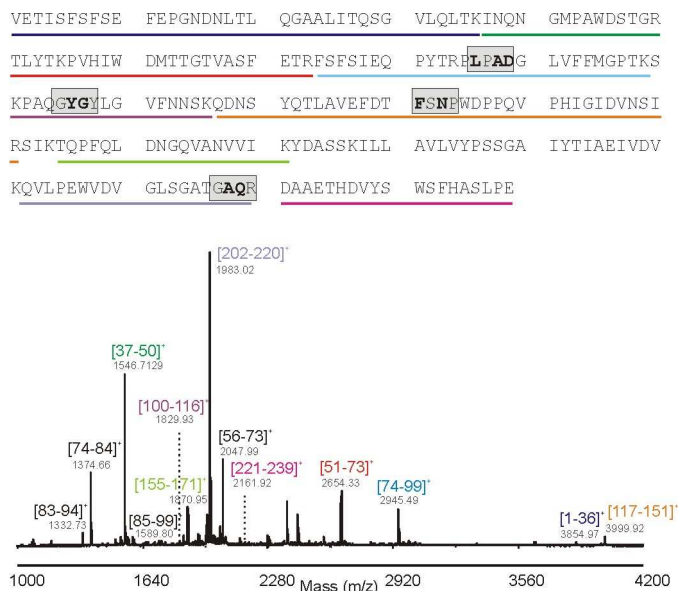


Figure 3.26 Primary sequence of ECA (Svensson 2002) and MALDI-TOF MS spectrum corresponding to trypsin digestion in solution. No peaks attributable to the two N-glycosylated Asn species (see previous page) could be detected.

Again the column was washed until no peptides were observed by MALDI-TOF (Figure 3.27, panel B). Finally, the bound peptides were eluted with ACN: H₂O 2:1 TFA 0.1% and identified by MALDI-TOF MS (Figure 3.27, panel C; Table 3.13 Peptides identified in the elution fraction of excision experiment carried out with trypsin. Amino acid residues that participate directly in carbohydrate-lectin interactions are highlighted.) Peptides identified in the elution fraction of excision experiment carried out with trypsin. Amino acid residues that participate directly in carbohydrate-lectin interactions are highlighted

In Figure 3.28, the peptides identified in the elution fraction are highlighted within the crystal structure of the complex [192]. Five out of eight peptides [74-99], [83-94], [100-116], [117-151] and [202-220] (highlighted in green) correlate directly with residues from the X-ray data

described as interacting. In addition, three other peptides [37-50], [51-73], and [74-84] (highlighted in red) were observed in the elution fraction.

Table 3.13 Peptides identified in the elution fraction of excision experiment carried out with trypsin. Amino acid residues that participate directly in carbohydrate-lectin interactions are highlighted.

Peptide	Sequence	m/z _{calc}	m/z _{found}
[37-50]	INQNGMPAWDSTGR	1546.71	1546.70
[51-73]	TLYTKPVHIWDMTTGTVASFETR	2654.33	2654.33
[74-84]	SFSIEQPYTR	1374.67	1374.66
[83-94]	TRPLPADGGLVFF	1332.73	1332.73
[74-99]	FSFSIEQPYTRPLPADGLVFFMGPTK	2945.49	2945.49
[100-116]	SKPAQGYGYLGVFNNSK	1829.92	1829.90
[117-151]	QDNSYQTLAVEFDTFSNPWDPQVPH IGIDVNSIR	3999.92	3999.91
[202-220]	QVLPEWVDVGLSGATGAQR	1983.03	1983.00

As can be seen in Figure 3.28, panel B, these last are not directly related to the CRD, but could correspond to peptide-peptide interaction with the peptide [202-220], that do interact with the glycan.

RESULTS

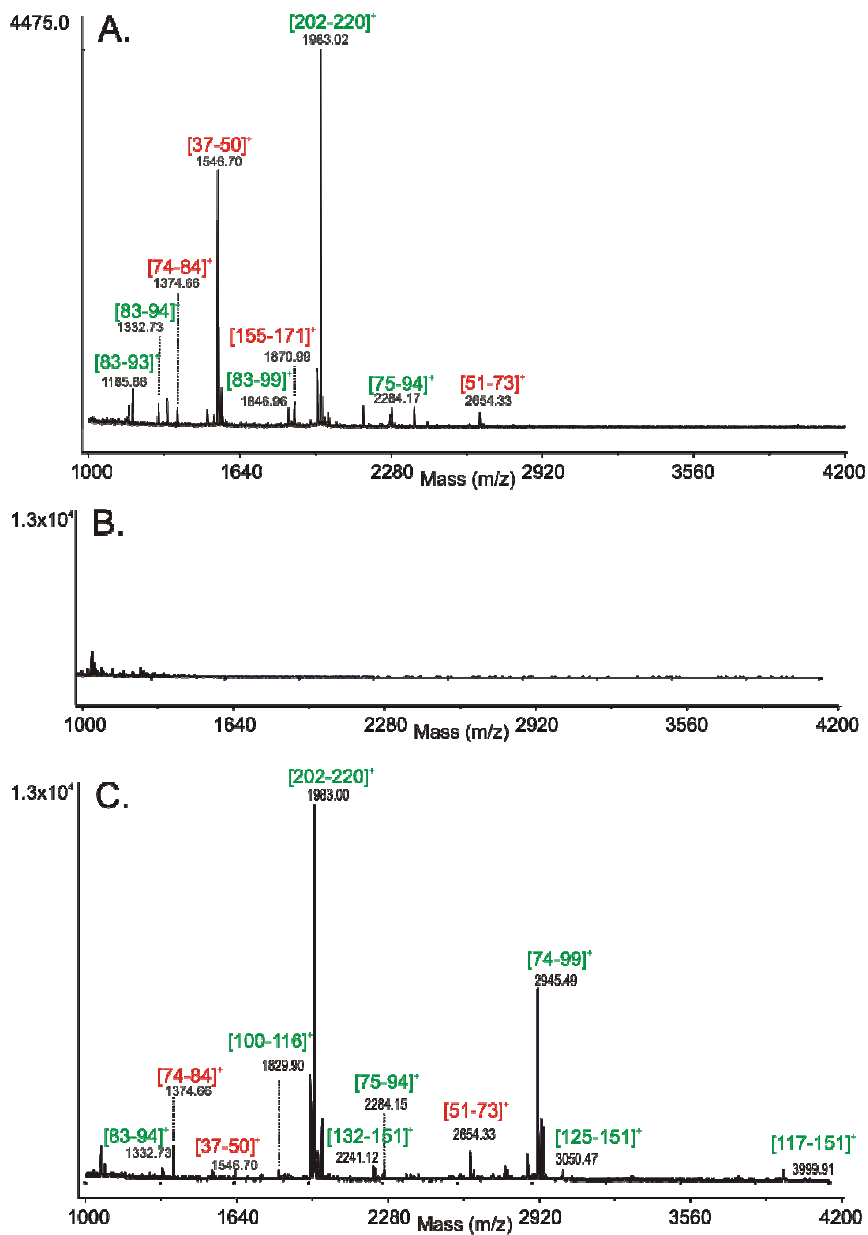


Figure 3.27 MALDI-TOF MS spectra of the different fractions of the excision experiment. A. On-column digestion. B. Supernatant after washing. C. Elution fraction. Peptides identified in the elution fraction are coloured in green (sugar-peptide interaction) and red (peptide-peptide interaction).

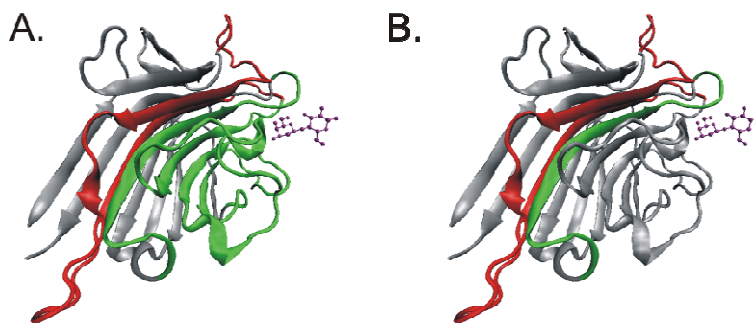


Figure 3.28 Ribbon representation of the ECA-lacNAc complex. A. Peptides identified in the elution fraction are indicated in green (carbohydrate-peptide interaction) and in red (peptide-peptide interaction). B. Peptides [37-50], [51-73] and [74-84] (in red) are situated closely to the peptide [202-220] (in green). This figure was generated by VMD OpenGL Display (version 1.8.3).

In order to discard the non-specific peptides and obtain shorter true-binding peptides, an excision experiment was performed in which, after trypsin digestion, another digestion with chymotrypsin was carried out with washing steps after each digestion. Figure 3.29 show the mass spectra corresponding to the different fractions. With the two consecutive digestions, the longest peptide [202-220] was digested into shorter peptides, thereby hampering potential (and predictably stronger) peptide-peptide interactions. As a result, non-specific peptides (*i.e.* [33-50], [51-73], [74-84]) were absent from the elution fraction, whereas the specific ones ([74-99], [117-151] and, to lesser extent, [202-220]) were preserved (Figure 3.29 panel E). From this example it can be concluded that, at least for excision experiments yielding long peptides, tandem proteolysis can provide better results than single enzyme digestions.

RESULTS

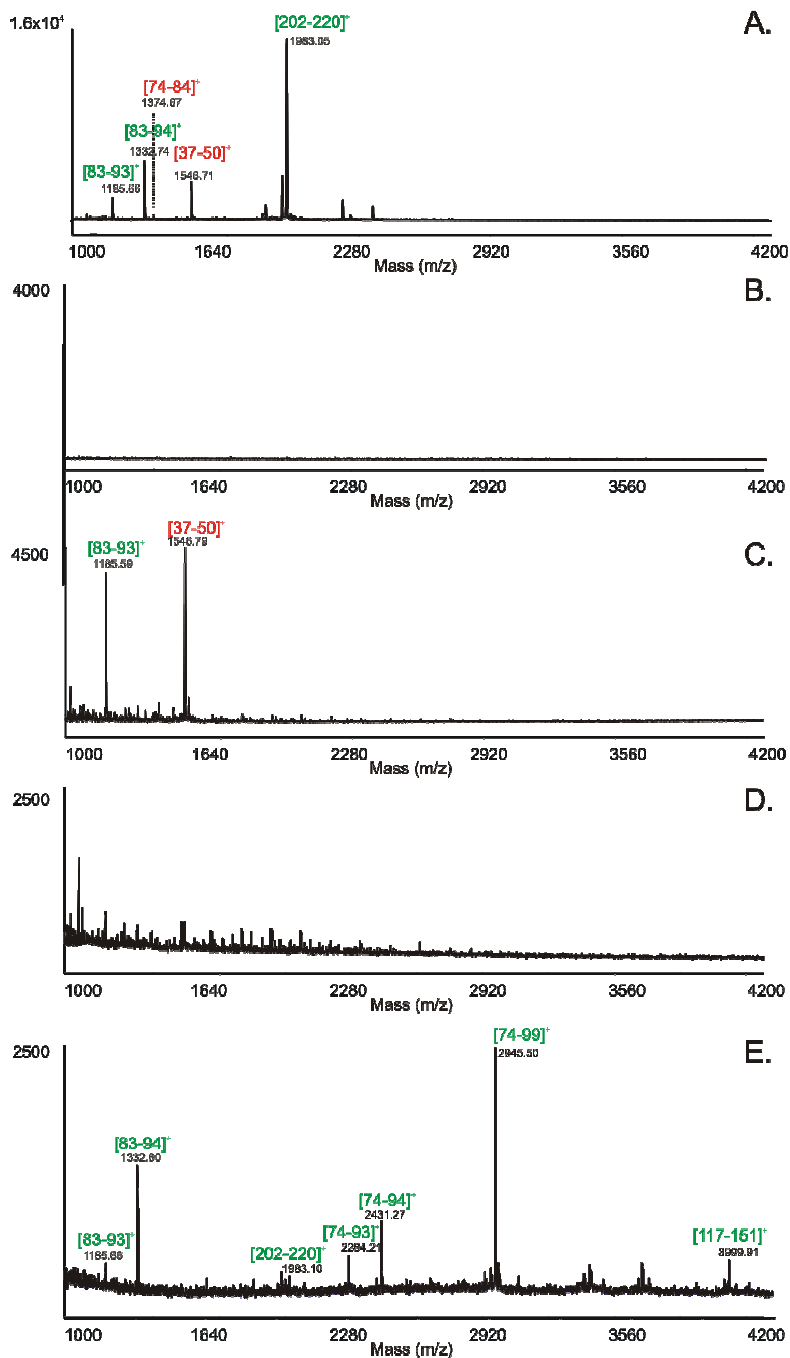


Figure 3.29 MALDI-TOF MS spectra corresponding to different fractions of excision experiment with two enzyme digestions on chain. A. First on column digestion with trypsin. B. Supernatant after washing. C. Second on column digestion with chymotrypsin. D. Supernatant after washing. E. Elution fraction.

4 DISCUSSION

This thesis focuses mainly on a novel approach by our group to study carbohydrate-lectin interactions by SPR. After the proof-of-principle [119], our main goal has been to extend the methodology to relevant sugars, thus establishing SPR as a viable alternative for monitoring carbohydrate-lectin interactions.

Among the tools used to study the mechanism of these interactions, HIA and ELLA (see Section 1.3.3), very useful in the early stages of lectinology to determine carbohydrate affinities [193], have been gradually supplanted by techniques such as microarrays, which allow to determine carbohydrate specificity in a more straightforward manner [194]. Although microarrays possess such desirable features as nanoscale and high throughput screening, the need to label at least one of the interacting partners substantially reduces their appeal. For in-depth probing the structural details of carbohydrate-lectin interactions, NMR and X-ray crystallography are the undisputed techniques of choice [195,196]. While highly informative, they are also quite demanding in terms of the amounts (in the mg range) and purity required. Lower amounts but still highly pure materials are required to characterize thermodynamically the interaction by ITC [69]. The information obtained by these techniques usually refers to free, relatively small, glycans (mono-, di- at most trisaccharides) interacting with a full-length protein or its truncated version. Some of these situations are rather unlike native scenarios, where a lectin interacts with glycans that are part of glycoproteins or glycolipids, usually present at high densities.

In recent years, SPR has emerged as a very suitable technique for studying biomolecular interactions. In comparison to other techniques, in SPR one of the interacting entities is immobilized onto a surface and the other is flown across. Interactions are monitored directly, with no need for

DISCUSSION

labelling, and in real time, altogether rendering the technique straightforward, fast, and very sensitive. Furthermore, it allows determining kinetic rate constants, not easily obtained by other techniques, with the exception of the similar QCM biosensor [197]. In addition, the use of a flexible hydrophilic polymer such as carboxymethylated dextran as surface provides a three-dimensional scaffold that may mimic the extracellular matrix [198]. On the down side, the matrix also enhances phenomena such as mass transport-limited binding, causing low diffusion rates from bulk solution to surface-immobilized ligand, a problem particularly important for large analytes.

Although initially designed for protein-protein interactions, a large number of other biomolecular interactions (*i.e.* peptide-protein [199], DNA-protein [200], lipid-protein [201] and sugar-protein [202] have been also reported. Carbohydrate-protein interaction studies by SPR have been addressed in several ways. The initial choice was to immobilize the lectin moiety, for two obvious reasons: (1) immobilization chemistry already well-established for proteins (2) immobilizing the larger molecule avoids potential mass transport problems [89]. While this approach has allowed reliable monitoring of numerous binding events, the analytes used as soluble ligands are complex glycans [87] or glycoproteins [203], and the affinities thus determined refer therefore to the entire glycoconjugate. Therefore, the primary carbohydrate specificity of the lectin cannot strictly be extrapolated from these experiments. In addition, in the immobilized lectin mode, small (< 500 Da) carbohydrates are difficult to quantify because of their low MW and refractive index. The alternate mode, sugar immobilization, has been addressed in several ways. Direct carbohydrate attachment via aldehyde chemistry (possible in all reducing sugars) has proven successful only with large polysaccharides [204]. Sugar epitopes have also been immobilized as glycopeptides derived from

proteolysis of native glycoproteins [94]. However, the macro- and microheterogeneity found in glycoproteins often makes unequivocal interpretation of the results a complex task. A third approach, relying on well-characterized synthetic glyconjugates, is in several ways the best manner to prepare sensor surfaces of precisely defined glycosylation. Thus, sugar epitopes have been immobilized onto dextran matrices using non-natural derivatives [205], as well as fluorescent [206,207] and biotin groups [97,98,208]. In lieu of three-dimensional dextrans, other authors have chosen glycoconjugate self-assembled-monolayers that resemble biological lipid membranes [100,209,210]. In all these approaches, the immobilization process is well addressed and a homogenous surface is prepared. However, the use of non-native structures in close proximity to the glycan may arguably affect the interaction. As reported elsewhere, although lectins and other carbohydrate-binding proteins are highly specific to their sugar epitopes, the aglycon, particularly its conformation, may play an important role as well [211]. In our approach, the carbohydrate moiety is attached to the dextran matrix through a peptide module. In this way, the sugar epitope is immobilized in a well-defined manner and since the aglycon is a peptide, it can in principle be readily adapted if a sequential or conformational feature needs to be reproduced.

In our original prototype [119], the peptide sequence was designed to fulfil specific requirements. Thus, the two Lys residues near the C-terminus and the Phe residue were introduced to obtain a well-oriented immobilization and facilitate UV-quantification, respectively. In view of the limited sources of well-defined, pure carbohydrates, a simple, straightforward synthetic route to the neoglycopeptides was desirable. Chemoselective approaches have proven extremely useful in this regard [104]. From the different strategies available, only a few do not require previous derivatization of the sugar [113]. Thus, we chose conjugation of

unmodified carbohydrates with hydroxylamine-functionalized peptides, which target specifically the single aldehyde group in the glycan chain.

In order to validate this SPR-based approach as an alternative and robust technique to study carbohydrate-lectin interactions, the different steps required to create an appropriate glycosylated surface were carefully studied. First, the synthesis of the peptide platform was optimized to achieve a high purity product, minimizing reaction time and without the requirement of HPLC purification [212]. Next, we focused on the sugar epitope. In oxime-linked glycopeptides, it is assumed that the non-terminal sugar unit (linked to peptide) is found in equilibrium between open and closed ring forms [153]. However, in many interaction studies, particularly those based on microarrays [153] or photoreactive glycoprobes [213], and involving both short and long oligosaccharides, this equilibrium is usually ignored. As a result, many literature reports, in fact, describes the result of a mixture of interactions. This is particularly serious when small sugar epitopes are the target of lectin recognition, since lectins bind only to native-like (ring) sugar forms, not to open acyclic conformers. In this context, we used NMR structural analysis to determine the relative abundance of each oxime-linked form and showed that, when ligation involves a primary hydroxylamine-functionalized peptide, only 30% of the resulting glycopeptide has the sugar unit closest to the peptide in cyclic form, and of this 30% of ring form 20% and 80% are in α - and β -anomeric configuration, respectively. This poor exposure of ring forms prompted a modification of the original platform with the goal of native-like display small sugar epitopes.

To improve the representation of the ring-closed form in the sample and, at the same time, minimize changes in the peptide platform, the primary ($\text{H}_2\text{N-O-}$) was replaced by a secondary hydroxylamino [$\text{HN}(\text{CH}_3)\text{-O-}$] group, which prevents the formation of the open oxime form [112]. This

hypothesis was confirmed by differential derivatization and subsequent analysis by MS of the reaction products (See Section 3.1.2). This protocol allows determining the ratio of linear-*vs*-ring peptide-linked sugar, and is fast, easy, and generally applicable to all conjugation products. It was additionally corroborated by NMR analysis, which showed that the β -anomer is predominant (> 90%) in the conjugates exposing GlcNAc and Glc. Therefore, at least for these two monosaccharides, the modification of secondary [HN(CH₃)-O-] instead of primary (H₂N-O-) hydroxylamino function in the peptide module is the preferred way to expose the sugar epitopes in the native-like β -configuration. The importance of ensuring ring closure in interaction studies was demonstrated by comparing the relative binding response of different lectins to both glycoprobes (*N*[Me]-Aoa *vs* Aoa) by means of SPR (See Section 3.3.2) [214]. As a general rule, similar NMR analysis of glycopeptides exposing other monosaccharides would be recommended to substantiate accurate read-out of lectin-carbohydrate interaction studies.

Once the *N*[Me]-Aoa-peptide was established as an efficient platform for exposing small sugar epitopes with the peptide-proximal sugar unit in native-like ring form, conjugations to a broader range of mono- and oligosaccharides were explored in order to optimize the step prior to immobilization onto the sensor chip. First of all, the *N*[Me]-peptide was conjugated to the most commonly-occurring monosaccharides found in mammalian organisms (Glc, Man, Gal, Fuc, GlcNAc, GalNAc, and Neu5Ac). Conjugation reactions were carried out using excess of sugar, taking advantage of the fact that, in these cases, the monosaccharide was not a limiting factor. Although good yields (> 60%) were obtained for all monosaccharides, clear differences in reaction rates were observed between monosaccharides. Similar differences have been reported by other authors for Aoa-containing peptides [162,215] and *N*[Me]-Aoa-peptides

DISCUSSION

[112,216]. As all reactions were carried out under similar conditions of temperature, and peptide-to-sugar stoichiometry, with the monosaccharide as the sole variation, significant differences between conjugation yields (from 60% to 80%) could be attributed exclusively to the nature of the carbohydrate. Under these conditions, three factors seem to affect the conjugation rate with the *N*[Me]-Aoa- peptide: (1) the abundance of the reactive acyclic form in solution, (2) the position of the hydroxyl group at C2, and (3) the presence of a 2-acetamido substituent at C2. Remarkably, the presence of a ketone instead of an aldehyde group, a less frequent feature in mammalian carbohydrates but relevantly present in sialic acids, appears to be decisive. For this monosaccharide, no product was observed even with excess of sugar and long incubation times (See Section 3.2.1). This lack of reactivity has been also observed in the conjugation to other ketoses such as fructose [217].

Subsequently, glycoprobes displaying a wide range of disaccharides, including Gal- β 1,4-Glc and *N*-acetyl-containing disaccharides such as Gal- β 1,(3,4)-GlcNAc, and Gal- β 1,3-GalNAc were synthesized (See Section 3.2.1) In this part of our work, efforts were directed to find appropriate reaction conditions to improve yields, given that in this case the amount of carbohydrate was a limiting factor. As expected, lower yields were observed for conjugations with mannobioses and *N*-acetyl-containing disaccharides in comparison to non-acetylated sugars, corroborating the results obtained previously with monosaccharides and those already reported for other glycoconjugates [112]. The negative effect observed in both cases could be partially offset by acidifying the reaction mixture, thus accelerating the mutarotational equilibrium and consequently increasing the relative amount of reactive open aldehyde form. For these sugars, optimum pH values of 3.5 and 4.6 for *N*-acetyl-disaccharides and non-acetylated sugars, respectively, could be

established. Additionally, the preparation of a wide range of glycopeptides, different only in their glycosidic linkages and/or the terminal sugar unit, under the same reaction conditions, allowed evaluating the effect of these differences in the conjugation rate. Thus, within each set of disaccharides (*i.e.* Gal- β 1,(3,4,6)-GlcNAc, Fuc- α 1,(3,4,6)-GlcNAc and Man- α 1,(2,3,6)-Man), higher reactivities were observed for disaccharides with the glycosidic linkage in 1,6 compared to 1,3 and 1,4 for those with non-terminal GlcNAc. However, for mannobioses (*i.e.* Man- α 1,(2,3,6)-Man) no differences between linkages were observed, probably due to stronger negative effect of the axial disposition of the hydroxyl at C2. Complementary to carrying out the conjugation at an optimum pH, the addition of an organic catalyst such as aniline was evaluated with the aim of reducing incubation time. Other authors have shown a significant improvement in yield (from 62% to 100% in 3 h) for conjugations involving non-methylated Aoa-containing peptides and monosaccharides [162]. However, in our case and employing the less reactive secondary amine *N*[Me]-Aoa-peptides, the effect was not that significant. Even so, a slight improvement (from 6 to 27% for Gal- β 1,3GalNAc) was observed after 24 h incubation, leaving open speculation on the addition of other catalysts for further improvements in conjugation yields.

Another major problem when synthesizing glycopeptides is related to HPLC purification, as the little change in hydrophobic properties between glycopeptides and non-reacted peptides render this a difficult separation, particularly in conjugations with low yields (*i.e.* < 20%). Several attempts were undertaken to increase the hydrophobic character of non-reacted peptide and consequently increase its retention time to facilitate HPLC purification. However, addition of ordinary aldehydes and ketone scavenger agents such as formaldehyde and acetone, described before as

DISCUSSION

beneficial [163], in our hands did not provide desired results. In a subsequent attempt, however, a well-known alkylating agent, *N*-ethylmaleimide [218], was successful in enabling better separation between glycopeptides and nonreacted peptides. An additional aspect related to purification, namely the presence of (trifluoroacetic) acid in typical HPLC eluents causing rapid degradation of the *N*[Me]-Aoa-peptides during lyophilization steps, was also addressed. We showed that immediate neutralization of the solution containing the glycoprobe to pH 5 is mandatory to preserve the product's integrity. Finally, the suitability of the established conjugation conditions was tested for the most acid-labile sugars, namely those containing terminal sialic acid units. These trisaccharides were successfully conjugated to the *N*[Me]-Aoa-peptide at mildly acidic pH, purified and characterized by MALDI-TOF MS, similar to other glycoconjugates. Additionally, their long-term stability was tested and shown to be more than acceptable (> 95% of purity after two years) provided they are stored in lyophilized form.

Although several glycoconjugates have been synthesized using *N*[Me]-Aoa peptide versions as mimics of natural glycoproteins [111,219], only few of them have been used as probes for interaction studies, and not always with conclusive results. Thus, whereas interactions between mannose- and *N*-acetyllactosamine- *N*[Me]derivatives and lectins such as ConA and RCA have been successfully detected using microarrays [157,220], others laboratories, also employing microarrays, failed to detect interactions with the methylated glycoprobe using carbohydrate-specific antibodies, suggesting that the methyl group in close proximity to a recognition motif could interfere with the interaction [221]. Having synthesized and characterized a broad collection of different glycopeptides, we were able to immobilize several mono-, di- and

trisaccharide-displaying probes onto the sensor surface and to monitor by SPR their interaction with relevant lectins.

One crucial aspect of the immobilization process is worth mentioning, even if its underlying mechanistic reasons have not been fully understood. When two surfaces immobilized with the same probe but at different pHs were compared, only the surface prepared at a pH closer to the pI of the ligand turned out to be functional, while the other surface did not display any interaction whatsoever. From this finding it follows that immobilization at low pH (4 instead of 6) renders non-functional surface deposits. Presumably, the small size of our glycoprobe in comparison to whole proteins, in combination with the high electrostatic attraction caused by the low pH, facilitates an excessive penetration of the glycoprobe into the dextran matrix that makes the conjugate less accessible for lectin binding.

The appropriateness of our approach for exposing accessibly and in native-like form a minimal monosaccharide epitope was successfully tested with a glycoprobe displaying Gal. Although strictly speaking this experiment must be considered as only a proof-of-principle for monosaccharides, it can be plausibly argued that our approach can be used as a first step to screen new lectins and characterize their primary carbohydrate specificity (See Section 3.3.3).

As mentioned before (Section 1.4.4), SPR-based interaction studies have been used to determine kinetic and affinity constants of different lectins for their specific carbohydrate ligands. Our approach was similarly evaluated in this sense and kinetic studies were performed with a variety of sugar epitopes and their specific lectins. Theoretically, in order to avoid mass transport effects, a maximal 100 RU of total analyte response is desirable [171]. However, in our approach, where a rather small molecule is immobilized onto the sensor surface and a much larger molecule is

DISCUSSION

passed as analyte, such low immobilization levels are virtually impossible to monitor with the currently available SPR instrumentation. Consequently, the larger the glycoprobe, the easier it is to monitor the immobilization process. Accordingly, for glycoprobes exposing larger sugar epitopes (*i.e.* trisaccharides), lower immobilization levels could be used and more accurate kinetic constants could be derived. In contrast, for disaccharide-exposing glycoprobes, even though the response was in accordance to the carbohydrate specificity described for each lectin, mass transport effects could significantly affect the kinetic evaluation.

During the course of this thesis, other applications of the SPR technology not yet exploited for carbohydrate-lectin interactions were evaluated. First, a detailed analysis and ranking of the carbohydrate hydroxyls that may participate in the interaction was performed, by comparing the binding response of specific lectins to disaccharides differing only in glycosidic linkage. In this way, in the interaction of lacNAc with ECA, the O3, N2, O6 and O_{NAC} atoms were shown to play relevant roles, in good agreement with previously reported X-ray data [173].

Secondly, the effect of the temperature on the carbohydrate-lectin interaction was evaluated and used to extract thermodynamic parameters of the interaction. Similar experiments have been recently performed with other type of interactions, generally involving small molecules [70], but only little information is available for carbohydrate-lectin interactions, apart from those derived by ITC [69]. As a representative example, the ECA-lacNAc interaction was monitored at different temperatures and its thermodynamic parameters determined from the van't Hoff equation. The enthalpy value obtained by SPR was similar to that reported previously by ITC [222] and within the range of enthalpies described for lectins and disaccharides [223]. In contrast, significant differences in entropic data derived from SPR and ITC were observed, which may reflect differences

in the way the interaction studies are performed. Thus, whereas in ITC interaction occurs in a solution where all molecules are totally free, in SPR the surface immobilization of one reactant may restrict its rotational freedom and diffusional properties, consequently altering the thermodynamics. The fact that SPR-derived thermodynamic parameters were only fully comparable to ITC-derived values for interactions involving small analytes leads us to conclude that the discrepancies encountered in our studies may be due to mass transport limitations.

Thirdly, we have applied our approach in order to capture lectins for subsequent MS characterization. While this application of the SPR technology has been tested for other interactions such as antigen-antibody [79], in this thesis we have tested for weaker interactions such as carbohydrate-lectin. Unfortunately, the success in lectin capture and MS characterization seems to depend highly from the molecular weight of the lectin and the dissociation rate of the complex. Even so, we have been able to show that careful temperature control favours weak interactions with high dissociation rates which in turn facilitate ligand capture (See Section 3.3.5).

A thorough study of the carbohydrate-lectin interaction mechanism must ideally include determination of the CRD of the lectin. Although for solid-phase based studies limited proteolysis has provided valuable insights in different fields [224], for this purpose our SPR-based approach is limited. Particularly, the presence of crucial lysine residues in the peptide platform makes this determination impractical when employing proteases that target these residues. Thus, in order to cover this limitation, the CRD determination was performed by CREDEX-MS. Similar to the SPR approach, the sugar is immobilized to a hydrophylic sugar matrix and after incubation with lectin, the complex is digested with a proteolytic enzyme. This technique has been successfully used to determine CRD of galectins

DISCUSSION

[141]. In this thesis, the technique was tested for the complex ECA-lacNAc. As the X-ray crystal structure of this complex is already known [225], the results obtained by CREDEX-MS could be compared directly with those obtained by X-ray crystallography (See Section 3.4).. In contrast to the good results obtained with only one step of trypsin digestion for the complex galectin-lactose, the experiments with ECA were not so straightforward and a second digestion was necessary to shorten peptides and avoid peptide-peptide interactions that could give to incorrect interpretations. From these studies it became clear that the interactions based on carbohydrates are weak but durable as two consecutive digestion could be performed, including intermediate washing, *etc.* without complete loss of all interacted peptides.

Taken together, SPR and CREDEX-MS could be used to study the mechanism of carbohydrate-lectin in detail. The application of our SPR-based approach has been tested here with a wide range of sugar epitopes and their specific lectins. In this thesis, special attention was put in the preparation of well-defined glycopeptide probes that expose the sugar epitope in their native-like structure. However, there are still many other sugar epitopes to be explored (*e.g.* Galili epitope, A/B/H blood group, *etc.*) in order to cover completely the sugar epitope repertoire found in mammalian organisms. Thus, this work represents a step further to establish our SPR-approach as a serious alternative approach to study carbohydrate-lectin interactions. Additionally, CREDEX-MS allows covering the potential limitations of our SPR-approach, making this combination a perfect duet to perform a detailed study of carbohydrate-lectin interactions.

5 CONCLUSIONS

The research conducted in the course of this doctoral thesis has resulted in the scientific publications included in Appendix. The main conclusions from these papers, as well as from still unpublished results, are:

1. Carbodiimide-based Boc-Aoa coupling and short reaction times enhance yields in the synthesis of aminoxy-containing peptides. Using this optimized procedure it is not even necessary to purify the Aoa-peptide prior to chemoselective sugar ligation.
2. Sugar ligation to Aoa-peptides leads to structures whose equilibria between ring and open sugar forms is different from native monosaccharides. Therefore, methylation of the Aoa-peptide is required to ensure a ring structure in the peptide-linked sugar. Glycoprobes exposing only cyclic forms of sugar epitopes (*i.e.* GlcNAc, Glc) are predominantly found in β -configuration.
3. The requirement for a closed ring structure of the peptide-attached monosaccharide in the lectin binding event dictates optimal ligation conditions. For large (*e.g.* pentasaccharide) glycans, the standard conditions (pH 4.6) defined for Aoa are suitable. However, where the ring form of the peptide-linked sugar must be preserved, the Aoa-peptide should be replaced by the *N*[Me]-Aoa variant. This modification is slightly detrimental of the aminoxy reactivity.
4. The conjugation of *N*[Me]-Aoa-peptides to sugars containing an *N*-acetyl group at C2 requires a lower pH (3.5) in order to enhance the mutarotational exchange rate. This pH is compatible with even the most acid-labile sugars (*i.e.* sialic acid). Below pH 3, hydrolysis of the oxime linkage and other possible side

CONCLUSIONS

reactions call for low pH avoidance during post-synthesis manipulation (*e.g.* HPLC work-up, SPR regeneration, *etc.*).

5. The protocols established in this work are of general application for preparing a broad range of glycoprobes that can efficiently and reproducibly be immobilized for carbohydrate binding studies. In particular, the methodology allows differences in glycosidic linkage to be fully resolved employing lectin amounts in the nM to μ M range.
6. For optimal display of glycoprobes on SPR sensor surfaces, immobilization should be performed at pH 6. Immobilizations at lower pH cause embedding of the glycoprobe within the dextran matrix and poor lectin recognition.
7. The application of our SPR-based approach for thermodynamic analysis resulted in enthalpy values comparable to those obtained by ITC. However, significant differences were observed in entropic values, probably due to mass transport limitations.
8. The SPR-based approach is fully compatible with affinity capture of lectins and subsequent characterization by MALDI-TOF MS. The specific capture of lectins in the presence of other glycoproteins makes our methodology applicable for pull-down assays.
9. CREDEX-MS, a novel approach to evaluate carbohydrate binding proteins, is fully complementary to the SPR-based glycoprobe method. The main advantage of CREDEX-MS is its ability to define the CRD by limited proteolysis.

6 MATERIAL AND METHODS

6.1 Materials

Fmoc (N^{α} -(9-fluorenylmethyloxycarbonyl)) protected amino acids were purchased from Senn Chemichals (Dielsdorf, Switzerland). Boc (*tert*-butyloxycarbonyl)-protected aminooxyacetic acid (Boc-Aoa), Bis-Boc-protected aminooxyacetic acid (Boc₂-Aoa) and Rink amide MBHA resin were from Novabiochem (Läufelfingen, Switzerland). Boc-methylaminooxyacetic acid DCHA was from NeoMPS (Strasbourg, France). 2-(1*H*-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU) was obtained from Iris Biotech (Marktredwitz, Germany) *N,N'*-diisopropylcarbodiimide (DIC) and acetic anhydride were from Fluka (Madrid, Spain). *N,N'*-diisopropylethylamine (DIEA) was from Merck Biosciences (Darmstadt, Germany), and triisopropylsilane was from Sigma-Aldrich (Madrid, Spain). HPLC-grade acetonitrile (ACN), *N,N*-dimethylformamide (DMF), trifluoroacetic acid (TFA), diethyl ether, and pyridine were from SDS (Peypin, France).

Monosaccharides (Gal, Glc, Man, Fuc, GlcNAc, GalNAc and α Me-Man) and lactose (Gal- β 1,4-Glc) were purchased from Sigma-Aldrich (Madrid, Spain). Other carbohydrates (Gal- β 1,(3,4,6)-GlcNAc, Gal- β 1,3-GalNAc, GlcNAc- β 1,4-GlcNAc, Man- α 1,(2,3,6)-Man, Sia- α 2,3-Gal- β 1,4-GlcNAc and Sia- α 2,6-Gal- β 1,4-GlcNAc) employed in this work for glycopeptide synthesis were from Dextra (Reading, United Kingdom). Disaccharides with terminal Fuc (Fuc- α 1,(3,4,6)-GlcNAc) were obtained from Toronto Research Chemicals (Toronto, Canada) and α Me-Fuc was from Iris Biotech GmbH (Marktredwitz, Germany).

Lectins from *Triticum vulgare* (WGA), *Erythrina cristagalli* (ECA), *Maackia amurensis* (MAA), *Sambucus nigra* (SNA), *Canavalia ensiformis* (Con A), *Lotus tetragonolobus* (LTA) and *Ulex europaeus*

(UEA) were purchased from Sigma-Aldrich (Madrid, Spain). Lectin from *Ricinus communis* was from Vector laboratories (Burlingame, USA).

Aniline and *N*-ethylmaleimide were from Sigma-Aldrich (Madrid, Spain). 2,5-dihydrobenzoic acid (DHB) and α -cyano-4-hydroxycinnamic acid (CHCA) were obtained from Sigma-Aldrich (Madrid, Spain). Sinapinic acid was from Fluka (Madrid, Spain).

CM5 sensor chips, 1-ethyl-3-(3-diethylaminopropyl)-carbodiimide (EDC), *N*-hydroxysuccinimide (NHS), ethanolamine hydrochloride pH 8.5, and HBS-P (0.01 M HEPES pH 7.4; 0.15 M NaCl; 0.005% v/v surfactant P20) buffer were from Biacore (GE Healthcare, Uppsala, Sweden)

Sepharose-4B and divinylsulfone was purchased from Sigma-Aldrich (Madrid, Spain). Microcolumn and 35- μ m pore size filters were from MoBiTec (Goettingen, Germany).

Sequencing-grade modified porcine trypsin was from Promega (Madison USA). Sequencing grade chymotrypsin, and Gluc-C were obtained from Roche Diagnostics GmbH (Penzberg, Germany).

Poros 20 R2 was obtained from Applied BioSystems (Foster City, USA)

6.2 Peptide synthesis

Peptide substrates GFAKKG-amide, Ahx-GFAKKG-amide, GAGAGFAKKG-amide and GFKKG-amide were synthesized by Fmoc-based solid phase synthesis on a Rink MBHA resin (0.70 mmol g⁻¹) using Fmoc chemistry at a 0.1 mmol scale in an Applied Biosystems 430A automated synthesizer.

6.2.1 Boc-Aoa-OH coupling

Boc-aminoxyacetic acid (Boc-Aoa-OH) was coupled manually under different conditions. (i): Boc-Aoa-OH/HBTU/DIEA (3:3:6 equiv), 40

min; (ii): Boc-Aoa-OH/DIC (10:10 equiv), 60 min; (iii) Boc-Aoa-OH/DIC (8:8 equiv), 10 min; (all incubations were performed at rt).

6.2.2 Boc₂-Aoa-OH coupling

Boc₂-Aoa-OH was coupled manually using 3 equivalents of the protected amino acid and 3 equiv of DIC in DCM for 60 min at rt. A double coupling was carried out with half number of equivalents (1.5 equiv) to ensure a high purity (> 95%) in the peptide crude.

6.2.3 Boc-*N*[Me]-Aoa-OH coupling

Prior to coupling, Boc-*N*[Me]-Aoa-OH/DCHA was converted to the free acid by acid extraction. A total of 700 mg of Boc-*N*[Me]-Aoa-OH/DCHA were dissolved in 45 ml of 0.2 M HCl, mixed with equal volume of ethyl acetate and strongly shaken. The organic phase containing Boc-*N*[Me]-Aoa-OH was recovered from three consecutive extractions. The ethyl acetate was evaporated and the residue weighed (~ 330 mg). Coupling with Boc-*N*[Me]-Aoa-OH was performed with HBTU (3 equiv each), and DIEA (6 equiv) in DMF for 45 min.

6.2.4 Peptide cleavage and work-up

All peptides were fully deprotected and cleaved from the resin with TFA-water-triisopropylsilane (95:2.5:2.5, v/v, 90 min, rt). Peptides were isolated by precipitation with cold diethyl ether and separated by centrifugation, solubilized in water and lyophilized.

The products were analyzed by analytical HPLC (Phenomenex Luna C₈ column; 0-30% B for Aoa-GFAKKG-amide and Aoa-GFKKG-amide and 0-40% B for Aoa-Ahx-GFAKKG-amide and Aoa-GAGAGFAKKG-amide. All peptides were characterized by MALDI-TOF MS).

Table 6.1 Sequences and molecular weight of all Aoa-containing peptides synthesized in this work.

Peptides	Peptide sequence	$[M+H]^+_{\text{calc}}$	$[M+H]^+_{\text{exp}}$
1	Aoa-GFAKKG-amide	679.39	679.77
2	Aoa-Ahx-GFAKKG-amide	792.47	792.90
2'	<i>N</i> [Me]-Aoa-Ahx-GFAKKG-amide	806.53	806.64
3	Aoa-GAGAGFAKKG-amide	935.51	935.54
4	Aoa-GFKKG-amide	608.35	608.36
4'	<i>N</i> [Me]-Aoa-GFKKG-amide	622.37	622.10

6.3 Chemoselective ligation of carbohydrates

6.3.1 Conjugation between *N*[Me]-Aoa-peptide and monosaccharides

Conjugation between **4'** and monosaccharides (Gal, Glc, Man, GlcNAc, GalNAc, Fuc, and Sia) were performed in 0.1 M NaOAc pH 3.6 at 20 mM and 200 mM respectively for 144 h at 37 °C. Glycopeptides were purified by semi-preparative HPLC (Sphereclone C₁₈ column; 10-20% B (See HPLC conditions) and pure glycopeptides were characterized by MALDI-TOF MS (Table 6.2).

6.3.2 Conjugation between aminoxy-peptides and lactose

Ligation between Aoa-peptides (**1**, **3** and **4**) and lactose (Gal-β1,4-Glc) were carried out at 20 mM and 200 mM respectively in 0.1 M NaOAc pH 4.6 at 37 °C for 48 h. Before purification, the excess of nonreacted Aoa-peptide was capped by addition of acetone (1 g L⁻¹) to the reaction mixture. Glycopeptides were purified by semi-preparative HPLC (Phenomenex Luna C₈, 0-15% B for short peptides, and 0-30% B for

longer peptides (See HPLC conditions), lyophilized and characterized by MALDI-TOF MS (Table 6.2).

6.3.3 Conjugation between Aoa-peptide and other oligosaccharides

Gal- β 1,4-GlcNAc was reacted with **4** at 25 mM and 20 mM respectively in 0.1 M NaOAc pH 4.6 at 37 °C. After 72 h, the glycopeptide was purified by semi-preparative HPLC (Sphereclone C₁₈ column; 10-20% B (See HPLC conditions) and pure glycopeptides were characterized by MALDI-TOF MS (Table 6.2).

6.3.4 Conjugation between N[Me]-Aoa-peptide and oligosaccharides

- Disaccharides (Gal- β 1,4-Glc, Gal- β 1,(3,4)-GlcNAc and Gal- β 1,3-GalNAc, were conjugated to **2'** at 25 mM and 20 mM, respectively, in 0.1 M NaOAc buffer for 72 h at 37 °C. Different pH values (5.0; 4.6; 4.0; 3.5; 3.0) were explored. Two different ratios (25:20 mM and 25:12 mM) for disaccharides and peptides, respectively, were tested.
- Disaccharides (Gal- β 1,(3,4,6)-GlcNAc, Fuc- α 1,(3,4,6)-GlcNAc and Man- α 1,(2,3,6)-Man) were conjugated to **4'** at 25 mM and 20 mM, respectively in 0.1 M acetate buffer pH 3.5 at 37 °C for 72 h.
- Conjugation reactions with peptide **4'** (20 mM) and some disaccharides (Gal- β 1,3-GlcNAc, Gal- β 1,3-GalNAc and Man- α 1,3-Man, 25 mM) were performed either in the absence or presence of aniline (100 mM) in 0.1 M NaOAc pH 3.5 at 37 °C for 24 h.
- In order to alkylate the excess of nonreacted **4'**, *N*-ethylmaleimide 100 mM was added to the reaction mixture and reacted for 20 h.

MATERIAL AND METHODS

Table 6.2 Glycoprobes synthesized

Glycopeptide	$[M+H]^+$ _{calc}	$[M+H]^+$ _{exp}
Gal- 4'	784.42	784.39
Glc- 4'	784.42	784.36
Man- 4'	784.42	784.27
GlcNAc- 4'	825.45	825.36
GalNAc- 4'	825.45	825.57
Fuc- 4'	768.43	768.19
Sia- 4'	913.47	-
Gal- β 1,4-GlcNAc- 2'	1171.77	1171.71
Gal- β 1,3-GlcNAc- 2'	1171.77	1171.55
Gal- β 1,3-GalNAc- 2'	1171.77	1171.16
Gal- β 1,4-Glc- 2'	1130.75	1130.72
Gal- β 1,4-Glc- 1	1003.50	1003.73
Gal- β 1,4-Glc- 3	1259.62	1259.86
Gal- β 1,4-Glc- 4	932.46	932.83
Gal- β 1,4-Glc- 4'	946.48	946.78
Gal- β 1,4-GlcNAc- 4	973.48	973.81
Gal- β 1,4-GlcNAc- 4'	987.50	987.86
Gal- β 1,3-GlcNAc- 4'	987.50	987.94
Gal- β 1,6-GlcNAc- 4'	987.50	987.85
Man- α 1,2-Man- 4'	946.48	946.41
Man- α 1,3-Man- 4'	946.48	946.39
Man- α 1,6-Man- 4'	946.48	946.62
Fuc- α 1,3-GlcNAc- 4'	971.51	971.35
Fuc- α 1,4-GlcNAc- 4'	971.51	971.36
Fuc- α 1,6-GlcNAc- 4'	971.51	971.58
Neu5Ac- α 2,(3,6)-Gal- β 1,4-GlcNAc- 4	1223.55	1223.79
Neu5Ac- α 2,3-Gal- β 1,4-GlcNAc- 4'	1279.60	1279.16
Neu5Ac- α 2,6-Gal- β 1,4-GlcNAc- 4'	1279.60	1279.14

All these glycoconjugates were purified by semipreparative HPLC (SphereClone C₁₈, 10-20% B). Immediately after purification, glycopeptide solutions were neutralized with 10 mM NH₄HCO₃ to pH ~ 5 and lyophilized. Purified glycopeptides were characterized by HPLC and MALDI-TOF MS (Table 6.2)

6.4 Acetylation of glycopeptides

Glycopeptides GlcNAc-Aoa- **4** and GlcNAc- **4'** were acetylated with acetic anhydride and pyridine (1:1 v/v; 100 equiv) overnight at rt. The excess of reagents was evaporated in a speed-vac centrifuge and the resulting acetylated glycopeptides were characterized by MALDI-TOF mass spectrometry.

- (Ac)_n [GlcNAc- **4**]

m/z a) closed form: 1021.8 [M+H]⁺, 1043.8 [M+Na]⁺, 1059.8 [M+K]⁺

b) open form: 1063.8 [M+H]⁺, 1085.8 [M+Na]⁺, 1101.8 [M+K]⁺

- (Ac)_n [GlcNAc- **4'**]

m/z 1035.6 [M+H]⁺, 1057.6 [M+Na]⁺, 1073.6 [M+K]⁺

6.5 UV-quantification of glycoprobes

Pure glycoprobes were quantified by UV-spectroscopy by measuring the absorbance at 258 nm using a Nanodrop device (Nanodrop Technologies, Inc., Wilmington, USA). An operational extinction coefficient for the peptide- **4'** was experimentally obtained (ϵ 0.1438 mM⁻¹ cm⁻¹; *l* = 1 cm).

6.6 HPLC

HPLC analysis and purifications were performed using the following systems:

- Compact Shimadzu LC-2010A for analytical purposes with a Phenomenex Luna C₈ (3 μm, 0.46 x 5 cm) for peptides and a Sphereclone C₁₈ (5 μm, 0.1 x 25 cm) for glycopeptides. Linear gradients of buffer B (0.036% TFA in ACN) into buffer A (0.045% TFA in H₂O) were used over 15 min for Phenomenex Luna column and over 20 min for Sphereclone column at a flow rate of 1 ml min⁻¹.
- Shimadzu SCL-10A for semi-preparative purifications. Phenomenex Luna C₈ (10 μm, 0.1 x 25 cm) for peptides and SphereClone C₁₈ (10 μm, 0.1 x 25 cm) for glycopeptides. Linear gradients of buffer B (0.1% TFA in ACN) into buffer A (0.1% TFA in ACN) over 30 min for Phenomenex Luna column and over 20 min for Sphereclone column at a flow rate of 5 ml min⁻¹.

Under each condition the yield was estimated by integration of analytical HPLC peaks at 220 nm.

6.7 MALDI-TOF mass spectrometry

Peptides and glycopeptides were dissolved in water and mixed with the corresponding matrix solution (1:1 v/v) and 1 μl of the mixture was applied to the MALDI target and allowed to dry at room temperature. For both synthetic peptide and conjugate analysis, a solution of DHB (10 mg ml⁻¹) in ACN:water:TFA (70:30:0.1 v/v/v) was chosen. For peptide mixtures generated after digestion, a saturated solution of CHCA in ACN:water:TFA (70:30:0.1 v/v/v) was used. For protein analysis, a solution of sinapinic acid (10 mg ml⁻¹) in ACN:water:TFA (70:30:0.1

v/v/v) was used. Experiments were carried out on a Voyager-DETM STR Biospectrometry workstation (Applied Biosystems, Foster City, USA), equipped with a N₂ laser (337 nm). Peptides and glycopeptides were measured in the reflectron mode and positive polarity, except for sialic acid containing probes that were measured both in the positive and negative mode. Proteins were measured in the linear mode and positive polarity. External calibration of the spectrometer was performed using SequazymeTM Peptide Mass Standard Kit (PerSeptive Biosystems) of the desired range. Recorded data were processed with Data ExplorerTM Software (Applied Biosystems, Foster City, USA).

6.8 NMR spectroscopy

NMR experiments were performed on a Varian Inova VXR-500 spectrometer (Parc Cientific de Barcelona, Barcelona, Spain).

For NMR experiments, GlcNAc- **2'**, GlcNAc- **4**, GlcNAc- **4'** and Glc-peptide **4'** were repeatedly exchange in ²H₂O with intermediate lyophilizations, and finally dissolved in 500 μl ²H₂O. ¹H 1D and 2D NMR spectra (COSY: cosygpqf) were recorded at 500 MHz at 25 °C. Chemical shifts (δ) were expressed in parts per million relative to ²H₂O (²H₂O δ 4.754). Spectra were processed using MestRe-C software (version 2.0.0.1, MestreLab Research, Santiago de Compostela, Spain).

6.9 Surface Plasmon Resonance

All SPR experiments were performed on a BIAcore 3000 (GE Healthcare, Uppsala, Sweden) by using CM5 sensor chips and HBS-P (0.01 M HEPES pH 7.4; 0.15 M NaCl; 0.005% v/v surfactant P20) as running buffer, supplemented with CaCl₂ (5 mM) and MnCl₂ (1 mM).

6.9.1 Immobilization of glycoprobes

All immobilizations were done at $5 \mu\text{l min}^{-1}$. The carboxyl functionalities of the CM5 sensor chip were activated by injecting a mixture of NHS and EDC (1:1 v/v; $50 \mu\text{l}$). Then, 1 mg ml^{-1} glycopeptide solution in 10 mM NaOAc was injected for approximately 12 min over the activated surface. Two different pH values were tested (pH 4 vs pH 6) for the immobilization process. Afterwards, unreacted groups on the chip surface were blocked by injection of ethanolamine-HCl (1 M ; pH 8.5; $60 \mu\text{l}$). The difference between the resonance units after the surface activation and the final response corresponded to the amount of immobilized glycopeptide. ($0.981 \text{ RU} \sim 1 \text{ pg glycopeptide mm}^{-2}$, see section 3.3.1). The same protocol (performed at pH 6) was followed for the immobilization of all glycoprobes.

On each chip, a reference surface (preferentially, first flow cell) was created with a non-specific probe immobilized (glycoconjugate or *N*-[Me]-peptide) to subtract the non-specific lectin binding to the dextran surface. Glycoprobes immobilized on each sensor chip (plus their immobilization level) are indicated on Table 6.3.

Table 6.3 Sensor chips prepared during this work.

Sensor chip	Ligand immobilized	Amount immobilized	
		RU	pmol mm ⁻²
1	lac- 1	330	0.34
	lac- 3	1250	0.26
	lac- 4	240	1
2	4'	81	0.13
	Gal- β 1,4-GlcNAc- 4	132	0.14
	Gal- β 1,4-GlcNAc- 4'	165	0.17
3	lac- 4	520	0.57
	Sia- α 2,(3,6)-lac- 4 ^a	106	0.09
	Sia- α 2,(3,6)-lac- 4	89	0.07
4	4'	450	0.74
	Gal- 4'	875	1.1
5	4'	500	0.82
	Gal- β 1,4-GlcNAc- 4'	120	0.12
	Gal- β 1,3-GlcNAc- 4'	124	0.13
	Gal- β 1,6-GlcNAc- 4'	146	0.15
6	4'	819	1.3
	Gal- β 1,4-Glc- 4'	126	0.14
7	4'	277	0.45
	Man- α 1,2-Man- 4'	903	0.97
	Man- α 1,3-Man- 4'	341	0.37
	Man- α 1,6-Man- 4'	344	0.37
8	4'	277	0.45
	Fuc- α 1,3-GlcNAc- 4'	73	0.08
	Fuc- α 1,4-GlcNAc- 4'	236	0.25
	Fuc- α 1,6-GlcNAc- 4'	116	0.12
9	lac- 4'	27	0.03
	Sia- α 2,3-lacNAc- 4'	80	0.06
	Sia- α 2,6-lacNAc- 4'	56	0.04

^a Glycoprobe immobilized at pH 4.0.

6.9.2 Binding and kinetic experiments

All lectin binding experiments were performed at 25 °C. The signal of the reference flow cell was subtracted from the specific interaction before analysis. When necessary a double referencing was applied.

Regeneration of the sensor surface was accomplished by injecting the complementary carbohydrate (10 mM lac, 100 mM α Me-Man, 50 mM α Me-Fuc, 0.5 M GlcNAc) or, alternatively, glycoprotein solutions (*e.g.* fetuin at 2 mg ml⁻¹).

All kinetic measurements were carried out by running a manually programmed sequence.

All sensorgrams were analyzed using the BIAevaluation 4.1.1 software package (GE Healthcare, Uppsala, Sweden).

- Experiments carried out on sensor chip **1**:

For binding experiments, each lectin solution (ECA (1 μ M) and RCA (500 nM)) was injected over the three immobilized glycoprobes during 2 min at 10 μ l min⁻¹. After lectin injection, sample solution was replaced by running buffer and the dissociation phase was monitored for 3 min. After each binding cycle, the surface was regenerated by injection of 10 μ l of 10 mM lac in running buffer.

For kinetic studies, a 30 μ l-pulse of ECA was injected at several concentrations (62.5, 250, 500, and 1000 nM) over the three immobilized glycopeptides surfaces at 10 μ l min⁻¹. In this case, due to the zero-response observed for immobilized lac- **3**, this channel was used as reference for non-specific binding events.

For competition experiments with **4**, ECA at 250 nM (30 μ l) and RCA at 250 nM (30 μ l) were injected in presence of increasing concentrations of **4** (0, 250, 500, and 1000 nM) over the immobilized lac- **4** at 10 μ l min⁻¹.

- Experiments performed on sensor chip **2**:

Each lectin solution (ECA, RCA and WGA) was injected at 250 nM during 3 min over both immobilized glycoprobes (Gal- β 1,4-GlcNAc- **4** and Gal- β 1,4-GlcNAc-peptide **4'**) and reference flow cell (peptide **4'**). Surface channels were regenerated by 10 μ l-injection of 10 mM lac (for ECA and RCA) or 0.5 M GlcNAc (for WGA).

For kinetic experiments, ECA and RCA (at 5, 10, 20, and 60 nM) and WGA (at 3.75, 7.5, 15, and 30 nM) were prepared by dilution of the original solution with running buffer. Each lectin solution was injected over the three channels at 20 and 40 μ l min⁻¹.

- Experiments carried out on sensor chip **3**:

For the immobilization of Sia- α 2,(3,6)-lac- **4** glycoprobe, two pH (pH 4 vs pH 6) were tested. Binding of SNA was tested by an injection of SNA at 1 μ M during 3 min over both immobilized glycoprobes and reference lac-peptide **4** channels, at 10 μ l min⁻¹. Surface regeneration was accomplished by injection of 10 μ l of fetuin at 2 mg ml⁻¹.

- Experiments performed on sensor chip **4**:

Each lectin solution (ECA at 500 nM, RCA at 62.5 nM) was passed over immobilized Gal- **4'** and reference **4'** flow cells at 10 μ l min⁻¹. Regeneration was performed with a 10 μ l-pulse of 10 mM lac.

- Experiments performed on sensor chip **5** and **6**:

All SPR measurements performed on sensor chip **5** and **6** were carried out with a running buffer with lower NaCl concentration than HBS-P (25 mM NaCl).

For binding experiments, a 40 μ l-pulse of each lectin solution (ECA at 2.5 μ M, RCA at 670 nM and WGA at 500 nM) was injected on the immobilized Gal- β 1,(3,4,6)-GlcNAc- **4'** glycoprobes and the reference **4'**

MATERIAL AND METHODS

flow cells at $10 \mu\text{l min}^{-1}$. The same measurements were performed on the immobilized Gal- β 1,4-Glc- 4' glycoprobe and corresponding reference 4' flow cells of sensor chip 6.

For kinetic studies, several concentrations of each lectin were prepared following a set of 2/3-fold dilutions of the most concentrated sample in running buffer. Thus, concentrations varied from 66 nM to 2.5 μM for ECA, from 17 nM to 667 nM for RCA, and from 29 nM to 1.11 μM for WGA. Analyses were performed at $50 \mu\text{l min}^{-1}$. After protein injection, sample solution was replaced by running buffer and the dissociation phase was monitored for 6 min. Sensor surface was regenerated with a 50 μl -injection of 10 mM lac (for ECA and RCA) or 0.5 M GlcNAc (for WGA). Two replicates were performed for each injection.

- Experiments carried out on sensor chip 7:

For kinetic experiments, different concentrations were prepared for Con A (5, 10, 20, 40, and 80 nM) and passed over the sensor surface for 3 min at $10 \mu\text{l min}^{-1}$. After lectin injection, running buffer flows across the surface and dissociation phase was monitored for 3 min. Regeneration of the sensor surface was accomplished with a 10- μl injection of 100 mM $\alpha\text{Me-Man}$.

- Experiments performed on sensor chip 8:

All SPR measurements performed on sensor chip 8 were carried out with a running buffer with lower NaCl concentration than HBS-P (25 mM NaCl).

For binding experiments, five concentrations (125, 250, 500, and 1000 nM; 30 μl) of lectins LTA and UEA-I were injected over the functionalized surface for 3 min at $10 \mu\text{l min}^{-1}$. The sensor surface was regenerated with a 10- μl pulse of 50 mM $\alpha\text{Me-Fuc}$.

Competition experiments were performed with 1 μM lectin solutions and increasing concentration of either $\alpha\text{Me-Fuc}$ or $\alpha\text{Me-Man}$ (8, 40, 200, and 1000 μM). The mixture (lectin and monosaccharide) was injected for 3 min at 10 $\mu\text{l min}^{-1}$. After each binding, the sensor surface was regenerated with 10 μl of 50 mM $\alpha\text{Me-Fuc}$.

- Experiments carried out on sensor chip **9**:

For kinetic experiments, several concentrations of MAA (from 0.250 to 4.3 μM) and SNA (from 0.05 to 1.3 μM) were prepared in running buffer following a set of 2/3-fold dilutions of the most concentrated sample. Two replicates of each lectin solution were passed over the glycosylated surface for 3 min at 40 $\mu\text{l min}^{-1}$. After the end of sample injection, the dissociation phase was monitored for 4 min. After each binding cycle, the glycosylated surface was regenerated with an injection of fetuin 5 g L^{-1} for 1 min.

6.9.3 Thermodynamic experiments

Thermodynamic experiments were performed in the 10-25 $^{\circ}\text{C}$ temperature interval, with 2.5 $^{\circ}\text{C}$ increments. ECA solutions at concentrations from 100 to 700 nM were prepared from a 17.6 μM protein stock in HEPES buffer. Two replicates of each solution were injected (150 μl) over the functionalized sensor surface at 50 $\mu\text{l min}^{-1}$. After lectin injection, the sample was replaced by running buffer and the dissociation phase was monitored for 6 min. Regeneration of the sensor surface was accomplished with 2 consecutive 50- μl injections of 25 mM lac.

6.9.4 Lectin capture experiments

Recovery experiments were performed on sensor chip **2** at 10 $\mu\text{l min}^{-1}$.

A solution of each lectin (500 nM WGA, 80 nM RCA, and 1 μM ECA) was injected for 3 min at 10 $\mu\text{l min}^{-1}$. By means of the function MS

recover, the captured material was eluted in 2 μl of regeneration solution (10 mM GlcNAc- β 1,4-GlcNAc for WGA, and 10 mM Gal- β 1,4-GlcNAc for ECA and RCA), and concentrated by vacuum centrifugation.

For recovery experiments in the presence of other proteins, a mixture was prepared containing equimolar concentrations of BSA, RNase B, α 1-acid glycoprotein and WGA and passed over both the specific Gal- β 1,4-GlcNAc- **4'** immobilized glycoprobe and the reference **4'** flow cells of the sensor surface at 10 $\mu\text{l min}^{-1}$. The captured sample was recovered from four consecutive injection cycles and concentrated by vacuum centrifugation.

For MALDI-TOF analysis, the sample was dissolved in 1 μl matrix solution and deposited on the MALDI plate.

6.10 CREDEX-MS

Prior the CREDEX-MS experiment, ECA (5 μg) was digested in solution with three different proteolytic enzymes. The following incubation conditions were used:

- Trypsin: 0.25 μg trypsin in 25 mM NH_4HCO_3 , pH 8.5 (1:20 enzyme:substrate ratio) at 37 °C overnight.
- Chymotrypsin: 0.25 μg chymotrypsin in 100 mM Tris-HCl, pH 7.8, 10 mM CaCl_2 (1:20 enzyme:substrate ratio) at 35 °C for 24 h.
- Glu-C: 0.25 μg Glu-C in PBS pH 7.8 (1:20 enzyme:substrate ratio) at 35 °C for 24 h.

Peptide mixtures were desalted by reverse phase micro-column (Poros 20 R2). For each sample, a 2 μl column was packed in a GELoader tip and equilibrated with 10 μl H_2O :TFA 0.1%. The sample was dissolved in 20 μl of H_2O :TFA 0.1% and loaded onto the column. After slowly passing the sample through the column, a washing step was performed with 10 μl

of H₂O:TFA 0.1%. Finally, the bound material was eluted with 10 µl of ACN:H₂O: TFA (80:20:0.1 v/v/v). For MALDI-TOF MS analysis, 0.5 µl of sample was mixed with 0.5 µl of matrix solution and applied to the MALDI plate.

For disaccharide immobilization, 5 mg of Gal-β1,4-GlcNAc were dissolved in 50 µl of 0.5 M K₂CO₃, pH 11 and the solution was incubated with 50 µg of dry divinylsulfonyl-activated Sepharose. The coupling reaction was carried out overnight at rt under continuous shaking (800 rpm). The microcolumn was washed sequentially with 50 mM NH₄OAc, pH 4 and 50 mM NH₄HCO₃, pH 8. Finally, the column was equilibrated with binding buffer (10 mM HEPES, 25 mM NaCl, 5 mM CaCl₂ and 1 mM MnCl₂) and stored at 4 °C.

Before performing the excision experiment, the Gal-β1,4-GlcNAc-Sepharose column was tested for binding. 20 µg of ECA were added onto the Gal-β1,4-GlcNAc-Sepharose and incubated in binding buffer for 24 h at 37 °C. Unbound material was removed by extensive washing with running buffer. Bound lectin was eluted with 60% ACN, 0.1% TFA, and the protein content of each fraction (flow through; wash and elution) was analyzed by 1D-SDS-PAGE electrophoresis.

For the excision experiment, 20 µg of ECA were added to the Gal-β1,4-GlcNAc-Sepharose and incubated for 24 h at 37 °C. Unbound lectin was removed by washing with running buffer until no signal could be detected by MS. Then, the sugar-lectin complex was first digested overnight with trypsin (1 µg) in 25 mM NH₄HCO₃, pH 8.5 at 37 °C. After 15 h digestion, the flowthrough containing digestion products was removed and the column was washed with binding buffer. For the second digestion with chymotrypsin, 1 µg chymotrypsin was incubated with the complex in 100 mM Tris-HCl, 10 mM CaCl₂, pH 7.8 at 35 °C for 24 h. After washing with binding buffer, the specific-bound peptides were eluted with 60 µl of 60%

MATERIAL AND METHODS

ACN, 0.1% TFA, concentrated and lyophilized for MALDI-TOF MS analysis.

Prior to MALDI-TOF analysis, each sample was desalted by means of micro-column purification (Poros R2).

7 BIBLIOGRAPHY

1. Lieth C-W von der, L.E.K.T.: Carbohydrates: secondclass citizens in biomedicine and in bioinformatics? Lecture notes in computer science **1278**, 147-155 (1997)
2. Winterburn, P.J., Phelps, C.F.: The significance of glycosylated proteins. *Nature* **236**, 147-151 (1972)
3. Gabius, H.J.: Biological information transfer beyond the genetic code: the sugar code. *Naturwissenschaften* **87**, 108-121 (2000)
4. Fujita, K., Oura, F., Nagamine, N., Katayama, T., Hiratake, J., Sakata, K., Kumagai, H., Yamamoto, K.: Identification and molecular cloning of a novel glycoside hydrolase family of core 1 type O-glycan-specific endo-alpha-N-acetylgalactosaminidase from *Bifidobacterium longum*. *J. Biol. Chem.* **280**, 37415-37422 (2005)
5. Kamerling J.P., Boons G-J, Lee Y.C., Suzuki A., & Taniguchi N.Voragen A.G.J.: *Comprehensive glycoscience. From Chemistry to Systems Biology*. Elsevier, 2007)
6. Moustafa, I., Connaris, H., Taylor, M., Zaitsev, V., Wilson, J.C., Kiefel, M.J., von Itzstein, M., Taylor, G.: Sialic acid recognition by *Vibrio cholerae* neuraminidase. *J. Biol. Chem.* **279**, 40819-40826 (2004)
7. Pigman, W., Isbell, H.S.: Mutarotation of sugars in solution. 1. History, basic kinetics, and composition of sugar solutions. *Adv. Carbohydr. Chem. Biochem.* **23**, 11-57 (1968)
8. Ryu, K.S., Kim, C., Park, C., Choi, B.S.: NMR analysis of enzyme-catalyzed and free-equilibrium mutarotation kinetics of monosaccharides. *J. Am. Chem. Soc.* **126**, 9180-9181 (2004)
9. Vliegthart, J.F., Casset, F.: Novel forms of protein glycosylation. *Curr. Opin. Struct. Biol.* **8**, 565-571 (1998)
10. Kovensky, J.: Sulfated oligosaccharides: new targets for drug development? *Curr. Med. Chem.* **16**, 2338-2344 (2009)
11. Raman, R., Sasisekharan, V., Sasisekharan, R.: Structural insights into biological roles of protein-glycosaminoglycan interactions. *Chem. Biol.* **12**, 267-277 (2005)

BIBLIOGRAPHY

12. Kanska, G., Guerry, M., Caldefie-Chezet, F., De Latour, M., Guillot, J.: Study of the expression of Tn antigen in different types of human breast cancer cells using VVA-B4 lectin. *Oncol. Rep.* **15**, 305-310 (2006)
13. Yu, L.G.: The oncofetal Thomsen-Friedenreich carbohydrate antigen in cancer progression. *Glycoconj. J.* **24**, 411-420 (2007)
14. Do, K.Y., Smith, D.F., Cummings, R.D.: LAMP-1 in CHO cells is a primary carrier of poly-N-acetyllactosamine chains and is bound preferentially by a mammalian S-type lectin. *Biochem. Biophys. Res. Commun.* **173**, 1123-1128 (1990)
15. Manzella, S.M., Hooper, L.V., Baenziger, J.U.: Oligosaccharides containing beta 1,4-linked N-acetylgalactosamine, a paradigm for protein-specific glycosylation. *J. Biol. Chem.* **271**, 12117-12120 (1996)
16. Baenziger, J.U., Kumar, S., Brodbeck, R.M., Smith, P.L., Beranek, M.C.: Circulatory half-life but not interaction with the lutropin/chorionic gonadotropin receptor is modulated by sulfation of bovine lutropin oligosaccharides. *Proc. Natl. Acad. Sci. U. S. A* **89**, 334-338 (1992)
17. Garratty, G.: Problems associated with passively transfused blood group alloantibodies. *Am. J. Clin. Pathol.* **109**, 769-777 (1998)
18. Kansas, G.S.: Selectins and their ligands: current concepts and controversies. *Blood* **88**, 3259-3287 (1996)
19. Heneghan, M.A., Moran, A.P., Feeley, K.M., Egan, E.L., Goulding, J., Connolly, C.E., McCarthy, C.F.: Effect of host Lewis and ABO blood group antigen expression on *Helicobacter pylori* colonisation density and the consequent inflammatory response. *FEMS Immunol. Med. Microbiol.* **20**, 257-266 (1998)
20. Thurin, M., Kieber-Emmons, T.: SA-Lea and tumor metastasis: the old prediction and recent findings. *Hybrid. Hybridomics.* **21**, 111-116 (2002)
21. Misugi, E., Kawamura, N., Imanishi, N., Tojo, S.J., Morooka, S.: Sialyl Lewis X moiety on rat polymorphonuclear leukocytes responsible for binding to rat E-selectin. *Biochem. Biophys. Res. Commun.* **215**, 547-554 (1995)
22. Sorensen, A.L., Rumjantseva, V., Nayeb-Hashemi, S., Clausen, H., Hartwig, J.H., Wandall, H.H., Hoffmeister, K.M.: Role of sialic acid for platelet life span: exposure of beta-galactose results in the rapid clearance

of platelets from the circulation by asialoglycoprotein receptor-expressing liver macrophages and hepatocytes. *Blood* **114**, 1645-1654 (2009)

23. Simon, P.M., Goode, P.L., Mobasser, A., Zopf, D.: Inhibition of *Helicobacter pylori* binding to gastrointestinal epithelial cells by sialic acid-containing oligosaccharides. *Infect. Immun.* **65**, 750-757 (1997)

24. Takemoto, D.K., Skehel, J.J., Wiley, D.C.: A surface plasmon resonance assay for the binding of influenza virus hemagglutinin to its sialic acid receptor. *Virology* **217**, 452-458 (1996)

25. Park, E.I., Manzella, S.M., Baenziger, J.U.: Rapid clearance of sialylated glycoproteins by the asialoglycoprotein receptor. *J. Biol. Chem.* **278**, 4597-4602 (2003)

26. Rutishauser, U.: Polysialic acid and the regulation of cell interactions. *Curr. Opin. Cell Biol.* **8**, 679-684 (1996)

27. Stroud, M.R., Stapleton, A.E., Levery, S.B.: The P histo-blood group-related glycosphingolipid sialosyl galactosyl globoside as a preferred binding receptor for uropathogenic *Escherichia coli*: isolation and structural characterization from human kidney. *Biochemistry* **37**, 17420-17428 (1998)

28. Ono, K., Hattori, H., Uemura, K., Nakayama, J., Ota, H., Katsuyama, T.: Expression of Forssman antigen in human large intestine. *J. Histochem. Cytochem.* **42**, 659-665 (1994)

29. Sandrin, M.S., McKenzie, I.F.: Gal alpha (1,3)Gal, the major xenoantigen(s) recognised in pigs by human natural antibodies. *Immunol. Rev.* **141**, 169-190 (1994)

30. Abdel-Motal, U.M., Guay, H.M., Wigglesworth, K., Welsh, R.M., Galili, U.: Immunogenicity of influenza virus vaccine is increased by anti-gal-mediated targeting to antigen-presenting cells. *J. Virol.* **81**, 9131-9141 (2007)

31. Stillmark, H.: Ricin, ein giftiges Ferment. (1988)

32. Sharon, N., Lis, H.: History of lectins: from hemagglutinins to biological recognition molecules. *Glycobiology* **14**, 53R-62R (2004)

33. Sumner, J.B., Howell, S.F., Zeissig, A.: Concanavalin A and hemagglutination. *Science* **82**, 65-66 (1935)

BIBLIOGRAPHY

34. Boyd, W.C., Shapleigh, E.: Specific Precipitating Activity of Plant Agglutinins (Lectins). *Science* **119**, 419 (1954)
35. Nowell, P.C., Daniele, R.P., Winger, L.A.: Kinetics of human lymphocyte proliferation: proportion of cells responsive to phytohemagglutinin and correlation with E rosette formation. *J. Reticuloendothel. Soc.* **17**, 47-56 (1975)
36. Becht, H., Rott, R.: Purification of influenza virus hemagglutinin by affinity chromatography. *Med. Microbiol. Immunol.* **158**, 67-70 (1972)
37. Stockert, R.J., Morell, A.G., Scheinberg, I.H.: Mammalian hepatic lectin. *Science* **186**, 365-366 (1974)
38. Sly, W.S., Fischer, H.D.: The phosphomannosyl recognition system for intracellular and intercellular transport of lysosomal enzymes. *J. Cell Biochem.* **18**, 67-85 (1982)
39. Califice, S., Castronovo, V., van den, B.F.: Galectin-3 and cancer (Review). *Int. J. Oncol.* **25**, 983-992 (2004)
40. Danguy, A., Camby, I., Kiss, R.: Galectins and cancer. *Biochim. Biophys. Acta* **1572**, 285-293 (2002)
41. Gorelik, E., Galili, U., Raz, A.: On the role of cell surface carbohydrates and their binding proteins (lectins) in tumor metastasis. *Cancer Metastasis Rev.* **20**, 245-277 (2001)
42. Kilpatrick, D.C.: Animal lectins: a historical introduction and overview. *Biochim. Biophys. Acta* **1572**, 187-197 (2002)
43. Poveda, A., Asensio, J.L., Espinosa, J.F., Martin-Pastor, M., Canada, J., Jimenez-Barbero, J.: Applications of nuclear magnetic resonance spectroscopy and molecular modeling to the study of protein-carbohydrate interactions. *J. Mol. Graph. Model.* **15**, 9-17, 53 (1997)
44. Mäkelä, O.: Studies in hemagglutinins of leguminosae seeds. *Ann. Med. Exp. Biol. Fenn.* **35**, 1-133 (1957)
45. Lis, H., Sharon, N.: Lectins: Carbohydrate-Specific Proteins That Mediate Cellular Recognition. *Chem. Rev.* **98**, 637-674 (1998)
46. Chandra, N.R., Kumar, N., Jeyakani, J., Singh, D.D., Gowda, S.B., Prathima, M.N.: Lectindb: a plant lectin database. *Glycobiology* **16**, 938-946 (2006)

47. Ashford, D.A., Dwek, R.A., Rademacher, T.W., Lis, H., Sharon, N.: The glycosylation of glycoprotein lectins. Intra- and inter-genus variation in N-linked oligosaccharide expression. *Carbohydr. Res.* **213**, 215-227 (1991)
48. Weis, W.I., Drickamer, K.: Structural basis of lectin-carbohydrate recognition. *Annu. Rev. Biochem.* **65**, 441-473 (1996)
49. Rudiger, H., Gabius, H.J.: Plant lectins: occurrence, biochemistry, functions and applications. *Glycoconj. J.* **18**, 589-613 (2001)
50. Wu, A.M., Lisowska, E., Duk, M., Yang, Z.: Lectins as tools in glycoconjugate research. *Glycoconj. J.* **26**, 899-913 (2009)
51. Ashraf, M.T., Khan, R.H.: Mitogenic lectins. *Med. Sci. Monit.* **9**, RA265-RA269 (2003)
52. Liu, B., Bian, H.J., Bao, J.K.: Plant lectins: potential antineoplastic drugs from bench to clinic. *Cancer Lett.* **287**, 1-12 (2010)
53. Chen, Y., Maguire, T., Hileman, R.E., Fromm, J.R., Esko, J.D., Linhardt, R.J., Marks, R.M.: Dengue virus infectivity depends on envelope protein binding to target cell heparan sulfate. *Nat. Med.* **3**, 866-871 (1997)
54. Elgavish, S., Shaanan, B.: Lectin-carbohydrate interactions: different folds, common recognition principles. *Trends Biochem. Sci.* **22**, 462-467 (1997)
55. Quijcho, F.A.: Probing the atomic interactions between proteins and carbohydrates. *Biochem. Soc. Trans.* **21**, 442-448 (1993)
56. Garcia-Pino, A., Buts, L., Wyns, L., Loris, R.: Interplay between metal binding and cis/trans isomerization in legume lectins: structural and thermodynamic study of *P. angolensis* lectin. *J. Mol. Biol.* **361**, 153-167 (2006)
57. Komath, S.S., Kavitha, M., Swamy, M.J.: Beyond carbohydrate binding: new directions in plant lectin research. *Org. Biomol. Chem.* **4**, 973-988 (2006)
58. Lee Y.C., Lee R.T.: Carbohydrate-Protein Interactions: Basis of Glycobiology. *Acc. Chem. Res.* **28**, 321-327 (1995)

BIBLIOGRAPHY

59. Lundquist, J.J., Toone, E.J.: The cluster glycoside effect. *Chem. Rev.* **102**, 555-578 (2002)
60. Horan, N., Yan, L., Isobe, H., Whitesides, G.M., Kahne, D.: Nonstatistical binding of a protein to clustered carbohydrates. *Proc. Natl. Acad. Sci. U. S. A* **96**, 11782-11786 (1999)
61. Gomez-Garcia, M., Benito, J.M., Rodriguez-Lucena, D., Yu, J.X., Chmurski, K., Ortiz, M.C., Gutierrez, G.R., Maestre, A., Defaye, J., Garcia Fernandez, J.M.: Probing secondary carbohydrate-protein interactions with highly dense cyclodextrin-centered heteroglycoclusters: the heterocluster effect. *J. Am. Chem. Soc.* **127**, 7970-7971 (2005)
62. Lee, R.T., Lee, Y.C.: Affinity enhancement by multivalent lectin-carbohydrate interaction. *Glycoconj. J.* **17**, 543-551 (2000)
63. Silva, J.A., Damico, D.C., Baldasso, P.A., Mattioli, M.A., Winck, F.V., Fraceto, L.F., Novello, J.C., Marangoni, S.: Isolation and biochemical characterization of a galactoside binding lectin from *Bauhinia variegata* candida (BvcL) seeds. *Protein J.* **26**, 193-201 (2007)
64. Maierhofer, C., Rohmer, K., Wittmann, V.: Probing multivalent carbohydrate-lectin interactions by an enzyme-linked lectin assay employing covalently immobilized carbohydrates. *Bioorg. Med. Chem.* **15**, 7661-7676 (2007)
65. Bush, C.A., Martin-Pastor, M., Imberty, A.: Structure and conformation of complex carbohydrates of glycoproteins, glycolipids, and bacterial polysaccharides. *Annu. Rev. Biophys. Biomol. Struct.* **28**, 269-293 (1999)
66. Asensio, J.L., Canada, F.J., Bruix, M., Gonzalez, C., Khair, N., Rodriguez-Romero, A., Jimenez-Barbero, J.: NMR investigations of protein-carbohydrate interactions: refined three-dimensional structure of the complex between hevein and methyl beta-chitobioside. *Glycobiology* **8**, 569-577 (1998)
67. Aboitiz, N., Vila-Perello, M., Groves, P., Asensio, J.L., Andreu, D., Canada, F.J., Jimenez-Barbero, J.: NMR and modeling studies of protein-carbohydrate interactions: synthesis, three-dimensional structure, and recognition properties of a minimum hevein domain with binding affinity for chitoooligosaccharides. *Chembiochem.* **5**, 1245-1255 (2004)
68. Gourdine, J.P., Cioci, G., Miguet, L., Unverzagt, C., Silva, D.V., Varrot, A., Gautier, C., Smith-Ravin, E.J., Imberty, A.: High affinity

interaction between a bivalve C-type lectin and a biantennary complex-type N-glycan revealed by crystallography and microcalorimetry. *J. Biol. Chem.* **283**, 30112-30120 (2008)

69. Dam, T.K., Brewer, C.F.: Thermodynamic studies of lectin-carbohydrate interactions by isothermal titration calorimetry. *Chem. Rev.* **102**, 387-429 (2002)

70. Navratilova, I., Papalia, G.A., Rich, R.L., Bedinger, D., Brophy, S., Condon, B., Deng, T., Emerick, A.W., Guan, H.W., Hayden, T., Heutmekers, T., Hoorelbeke, B., McCroskey, M.C., Murphy, M.M., Nakagawa, T., Parmeggiani, F., Qin, X., Rebe, S., Tomasevic, N., Tsang, T., Waddell, M.B., Zhang, F.F., Leavitt, S., Myszka, D.G.: Thermodynamic benchmark study using Biacore technology. *Anal. Biochem.* **364**, 67-77 (2007)

71. Pei, Z., Anderson, H., Aastrup, T., Ramström, O.: Study of real-time lectin-carbohydrate interactions on the surface of quartz crystal microbalance. *Biosens. Bioelectron.* **21**, 60-66 (2005)

72. Day, Y.S., Baird, C.L., Rich, R.L., Myszka, D.G.: Direct comparison of binding equilibrium, thermodynamic, and rate constants determined by surface- and solution-based biophysical methods. *Protein Sci.* **11**, 1017-1025 (2002)

73. Liedberg, B., Nylander, C., Lönström, I.: Surface plasmon resonance for gas detection and biosensing. *Sensors and Actuators* **4**, 299-304 (1983)

74. www.biacore.com2010)

75. Armstrong, S.H., Budka, M.J.: Preparation and properties of serum and plasma proteins; the refractive properties of the proteins of human plasma and certain purified fractions. *J. Am. Chem. Soc.* **69**, 1747-1753 (1947)

76. Nelson, R.W., Krone, J.R.: Advances in surface plasmon resonance biomolecular interaction analysis mass spectrometry (BIA/MS). *J. Mol. Recognit.* **12**, 77-93 (1999)

77. Krone, J.R., Nelson, R.W., Dogruel, D., Williams, P., Granzow, R.: BIA/MS: interfacing biomolecular interaction analysis with mass spectrometry. *Anal. Biochem.* **244**, 124-132 (1997)

78. Nedelkov, D., Nelson, R.W.: Analysis of native proteins from biological fluids by biomolecular interaction analysis mass spectrometry

BIBLIOGRAPHY

(BIA/MS): exploring the limit of detection, identification of non-specific binding and detection of multi-protein complexes. *Biosens. Bioelectron.* **16**, 1071-1078 (2001)

79. Sonksen, C.P., Nordhoff, E., Jansson, O., Malmqvist, M., Roepstorff, P.: Combining MALDI mass spectrometry and biomolecular interaction analysis using a biomolecular interaction analysis instrument. *Anal. Chem.* **70**, 2731-2736 (1998)

80. Natsume, T., Nakayama, H., Isobe, T.: BIA-MS-MS: biomolecular interaction analysis for functional proteomics. *Trends Biotechnol.* **19**, S28-S33 (2001)

81. Natsume, T., Nakayama, H., Jansson, O., Isobe, T., Takio, K., Mikoshiba, K.: Combination of biomolecular interaction analysis and mass spectrometric amino acid sequencing. *Anal. Chem.* **72**, 4193-4198 (2000)

82. Nelson, R.W., Nedelkov, D., Tubbs, K.A.: Biosensor chip mass spectrometry: a chip-based proteomics approach. *Electrophoresis* **21**, 1155-1163 (2000)

83. www.biocore.com: Technical Note 18. 2010)

84. Zhukov, A., Schurenberg, M., Jansson, O., Areskoug, D., Buijs, J.: Integration of surface plasmon resonance with mass spectrometry: automated ligand fishing and sample preparation for MALDI MS using a Biacore 3000 biosensor. *J. Biomol. Tech.* **15**, 112-119 (2004)

85. Catimel, B., Rothacker, J., Catimel, J., Faux, M., Ross, J., Connolly, L., Clippingdale, A., Burgess, A.W., Nice, E.: Biosensor-based micro-affinity purification for the proteomic analysis of protein complexes. *J. Proteome. Res.* **4**, 1646-1656 (2005)

86. Nedelkov, D.: Development of surface plasmon resonance mass spectrometry array platform. *Anal. Chem.* **79**, 5987-5990 (2007)

87. Haseley, S.R., Talaga, P., Kamerling, J.P., Vliegthart, J.F.: Characterization of the carbohydrate binding specificity and kinetic parameters of lectins by using surface plasmon resonance. *Anal. Biochem.* **274**, 203-210 (1999)

88. Gutiérrez Gallego R., Haseley, S.R., van Miegem, V.F., Vliegthart, J.F., Kamerling, J.P.: Identification of carbohydrates binding to lectins by

- using surface plasmon resonance in combination with HPLC profiling. *Glycobiology* **14**, 373-386 (2004)
89. Shinohara, Y., Hasegawa, Y., Kaku, H., Shibuya, N.: Elucidation of the mechanism enhancing the avidity of lectin with oligosaccharides on the solid phase surface. *Glycobiology* **7**, 1201-1208 (1997)
90. O'Shannessy, D.J., Wilchek, M.: Immobilization of glycoconjugates by their oligosaccharides: use of hydrazido-derivatized matrices. *Anal. Biochem.* **191**, 1-8 (1990)
91. Satoh, A., Matsumoto, I.: Analysis of interaction between lectin and carbohydrate by surface plasmon resonance. *Anal. Biochem.* **275**, 268-270 (1999)
92. Barre, A., Van Damme, E.J., Peumans, W.J., Rouge, P.: Structure-function relationship of monocot mannose-binding lectins. *Plant Physiol* **112**, 1531-1540 (1996)
93. van der Merwe, P.A., Crocker, P.R., Vinson, M., Barclay, A.N., Schauer, R., Kelm, S.: Localization of the putative sialic acid-binding site on the immunoglobulin superfamily cell-surface molecule CD22. *J. Biol. Chem.* **271**, 9273-9280 (1996)
94. Shinohara, Y., Kim, F., Shimizu, M., Goto, M., Tosu, M., Hasegawa, Y.: Kinetic measurement of the interaction between an oligosaccharide and lectins by a biosensor based on surface plasmon resonance. *Eur. J. Biochem.* **223**, 189-194 (1994)
95. Brooks, S.A.: Strategies for analysis of the glycosylation of proteins: current status and future perspectives. *Mol. Biotechnol.* **43**, 76-88 (2009)
96. Otamiri, M., Nilsson, K.G.: Analysis of human serum antibody-carbohydrate interaction using biosensor based on surface plasmon resonance. *Int. J. Biol. Macromol.* **26**, 263-268 (1999)
97. Linman, M.J., Taylor, J.D., Yu, H., Chen, X., Cheng, Q.: Surface plasmon resonance study of protein-carbohydrate interactions using biotinylated sialosides. *Anal. Chem.* **80**, 4007-4013 (2008)
98. Munoz, F.J., Rumero, A., Sinisterra, J.V., Santos, J.I., Andre, S., Gabius, H.J., Jimenez-Barbero, J., Hernaiz, M.J.: Versatile strategy for the synthesis of biotin-labelled glycans, their immobilization to establish a bioactive surface and interaction studies with a lectin on a biochip. *Glycoconj. J.* **25**, 633-646 (2008)

BIBLIOGRAPHY

99. Uzawa, H., Ohga, K., Shinozaki, Y., Ohsawa, I., Nagatsuka, T., Seto, Y., Nishida, Y.: A novel sugar-probe biosensor for the deadly plant proteinous toxin, ricin. *Biosens. Bioelectron.* **24**, 929-933 (2008)
100. Critchley, P., Clarkson, G.J.: Carbohydrate-protein interactions at interfaces: comparison of the binding of *Ricinus communis* lectin to two series of synthetic glycolipids using surface plasmon resonance studies. *Org. Biomol. Chem.* **1**, 4148-4159 (2003)
101. Zeng, X., Murata, T., Kawagishi, H., Usui, T., Kobayashi, K.: Synthesis of artificial N-glycopolypeptides carrying N-acetylglucosamine and related compounds and their specific interactions with lectins. *Biosci. Biotechnol. Biochem.* **62**, 1171-1178 (1998)
102. Zeng, X., Murata, T., Kawagishi, H., Usui, T., Kobayashi, K.: Analysis of specific interactions of synthetic glycopolypeptides carrying N-acetylglucosamine and related compounds with lectins. *Carbohydr. Res.* **312**, 209-217 (1998)
103. Gildersleeve, J.C., Oyelaran, O., Simpson, J.T., Allred, B.: Improved procedure for direct coupling of carbohydrates to proteins via reductive amination. *Bioconjug. Chem.* **19**, 1485-1490 (2008)
104. Langenhan, J.M., Thorson, J.S.: Recent carbohydrate-based chemoselective ligation applications. *Current Organic Synthesis* **2**, 59-81 (2005)
105. Canne, L.E., Ferredamare, A.R., Burley, S.K., Kent, S.B.H.: Total Chemical Synthesis of A Unique Transcription Factor-Related Protein - Cmyc-Max. *Journal of the American Chemical Society* **117**, 2998-3007 (1995)
106. Kochendoerfer, G.G., Tack, J.M., Cressman, S.: Total chemical synthesis of a 27 kDa TASP protein derived from the MscL ion channel of *M. tuberculosis* by ketoxime-forming ligation. *Bioconjug. Chem.* **13**, 474-480 (2002)
107. Lee, D.J., Mandal, K., Harris, P.W., Brimble, M.A., Kent, S.B.: A one-pot approach to neoglycopeptides using orthogonal native chemical ligation and click chemistry. *Org. Lett.* **11**, 5270-5273 (2009)
108. Peri, F., Cipolla, L., La Ferla, B., Nicotra, F.: Glycoconjugate and oligosaccharide mimetics by chemoselective ligation. **6**, 635-644 (2003)

109. Hang, H.C., Bertozzi, C.R.: Chemoselective approaches to glycoprotein assembly. *Acc. Chem. Res.* **34**, 727-736 (2001)
110. Lemieux, G.A., Bertozzi, C.R.: Chemoselective ligation reactions with proteins, oligosaccharides and cells. *Trends Biotechnol.* **16**, 506-513 (1998)
111. Carrasco, M.R., Nguyen, M.J., Burnell, D.R., MacLaren, M.D., Hengel, S.M.: Synthesis of neoglycopeptides by chemoselective reaction of carbohydrates with peptides containing a novel N¹-methyl-aminooxy amino acid. *Tetrahedron Letters* **43**, 5727-5729 (2002)
112. Peri, F., Dumy, P., Mutter, M.: Chemo- and stereoselective glycosylation of hydroxylamino derivatives: A versatile approach to glycoconjugates. *Tetrahedron* **54**, 12269-12278 (1998)
113. Peri, F., Nicotra, F.: Chemoselective ligation in glycochemistry. *Chem. Commun. (Camb.)* 623-627 (2004)
114. Lee, M.R., Shin, I.: Facile preparation of carbohydrate microarrays by site-specific, covalent immobilization of unmodified carbohydrates on hydrazide-coated glass slides. *Org. Lett.* **7**, 4269-4272 (2005)
115. Gudmundsdottir, A.V., Paul, C.E., Nitz, M.: Stability studies of hydrazide and hydroxylamine-based glycoconjugates in aqueous solution. *Carbohydr. Res.* **344**, 278-284 (2009)
116. Dirksen, A., Hackeng, T.M., Dawson, P.E.: Nucleophilic catalysis of oxime ligation. *Angew. Chem. Int. Ed Engl.* **45**, 7581-7584 (2006)
117. Dirksen, A., Dirksen, S., Hackeng, T.M., Dawson, P.E.: Nucleophilic catalysis of hydrazone formation and transimination: implications for dynamic covalent chemistry. *J. Am. Chem. Soc.* **128**, 15602-15603 (2006)
118. Dirksen, A., Dawson, P.E.: Rapid oxime and hydrazone ligations with aromatic aldehydes for biomolecular labeling. *Bioconjug. Chem.* **19**, 2543-2548 (2008)
119. Vila-Perello, M., Gallego, R.G., Andreu, D.: A simple approach to well-defined sugar-coated surfaces for interaction studies. *Chembiochem* **6**, 1831-1838 (2005)
120. Loo, J.A.: Studying noncovalent protein complexes by electrospray ionization mass spectrometry. *Mass Spectrom. Rev.* **16**, 1-23 (1997)

BIBLIOGRAPHY

121. Pramanik, B.N., Bartner, P.L., Mirza, U.A., Liu, Y.H., Ganguly, A.K.: Electrospray ionization mass spectrometry for the study of non-covalent complexes: an emerging technology. *J. Mass Spectrom.* **33**, 911-920 (1998)

122. Drummond, J.T., Loo, R.R., Matthews, R.G.: Electrospray mass spectrometric analysis of the domains of a large enzyme: observation of the occupied cobalamin-binding domain and redefinition of the carboxyl terminus of methionine synthase. *Biochemistry* **32**, 9282-9289 (1993)

123. Kapur, A., Beck, J.L., Brown, S.E., Dixon, N.E., Sheil, M.M.: Use of electrospray ionization mass spectrometry to study binding interactions between a replication terminator protein and DNA. *Protein Sci.* **11**, 147-157 (2002)

124. Tuong, A., Uzabiaga, F., Petitou, M., Lormeau, J.C., Picard, C.: Direct observation of the non-covalent complex between human antithrombin III and its heparin binding sequence by capillary electrophoresis and electrospray mass spectrometry. *Carbohydr. Lett.* **1**, 55-60 (1994)

125. Zhou, M., Robinson, C.V.: When proteomics meets structural biology. *Trends Biochem. Sci.* **35**, 522-529 (2010)

126. Suckau, D., Kohl, J., Karwath, G., Schneider, K., Casaretto, M., Bitter-Suermann, D., Przybylski, M.: Molecular epitope identification by limited proteolysis of an immobilized antigen-antibody complex and mass spectrometric peptide mapping. *Proc. Natl. Acad. Sci. U. S. A* **87**, 9848-9852 (1990)

127. Macht, M., Marquardt, A., Deininger, S.O., Damoc, E., Kohlmann, M., Przybylski, M.: "Affinity-proteomics": direct protein identification from biological material using mass spectrometric epitope mapping. *Anal. Bioanal. Chem.* **378**, 1102-1111 (2004)

128. Iacob, R.E., Keck, Z., Olson, O., Fong, S.K., Tomer, K.B.: Structural elucidation of critical residues involved in binding of human monoclonal antibodies to hepatitis C virus E2 envelope glycoprotein. *Biochim. Biophys. Acta* **1784**, 530-542 (2008)

129. Peter, J.F., Tomer, K.B.: A general strategy for epitope mapping by direct MALDI-TOF mass spectrometry using secondary antibodies and cross-linking. *Anal. Chem.* **73**, 4012-4019 (2001)

130. Pimenova, T., Meier, L., Roschitzki, B., Paraschiv, G., Przybylski, M., Zenobi, R.: Polystyrene beads as an alternative support material for epitope identification of a prion-antibody interaction using proteolytic excision-mass spectrometry. *Anal. Bioanal. Chem.* **395**, 1395-1401 (2009)
131. Zhao, Y., Muir, T.W., Kent, S.B., Tischer, E., Scardina, J.M., Chait, B.T.: Mapping protein-protein interactions by affinity-directed mass spectrometry. *Proc. Natl. Acad. Sci. U. S. A* **93**, 4020-4024 (1996)
132. Hochleitner, E.O., Gorny, M.K., Zolla-Pazner, S., Tomer, K.B.: Mass spectrometric characterization of a discontinuous epitope of the HIV envelope protein HIV-gp120 recognized by the human monoclonal antibody 1331A. *J. Immunol.* **164**, 4156-4161 (2000)
133. Michel, S., Forest, E., Petillot, Y., Deleage, G., Heuze-Vourc'h, N., Courty, Y., Lascoux, D., Jolivet, M., Jolivet-Reynaud, C.: Involvement of the C-terminal end of the prostrate-specific antigen in a conformational epitope: characterization by proteolytic degradation of monoclonal antibody-bound antigen and mass spectrometry. *J. Mol. Recognit.* **14**, 406-413 (2001)
134. Parker, C.E., Tomer, K.B.: MALDI/MS-based epitope mapping of antigens bound to immobilized antibodies. *Mol. Biotechnol.* **20**, 49-62 (2002)
135. Zhao, Y., Chait, B.T.: Protein epitope mapping by mass spectrometry. *Anal. Chem.* **66**, 3723-3726 (1994)
136. Lu, X., DeFelippis, M.R., Huang, L.: Linear epitope mapping by native mass spectrometry. *Anal. Biochem.* **395**, 100-107 (2009)
137. Kiselar, J.G., Downard, K.M.: Direct identification of protein epitopes by mass spectrometry without immobilization of antibody and isolation of antibody-peptide complexes. *Anal. Chem.* **71**, 1792-1801 (1999)
138. Przybylski, M., Moise, A., Gabius, H.: Identification of ligand recognition domains. **EP2009/003495**, 2009)
139. Satoh, A., Fukui, E., Yoshino, S., Shinoda, M., Kojima, K., Matsumoto, I.: Comparison of methods of immobilization to enzyme-linked immunosorbent assay plates for the detection of sugar chains. *Anal. Biochem.* **275**, 231-235 (1999)

140. Hatakeyama, T., Murakami, K., Miyamoto, Y., Yamasaki, N.: An assay for lectin activity using microtiter plate with chemically immobilized carbohydrates. *Anal. Biochem.* **237**, 188-192 (1996)
141. Przybylski, M., Moise, A., Siebert, H., Gabius, H.: CREDEX-MS: Molecular elucidation of carbohydrate recognition peptides in lectins and related proteins by proteolytic excision-mass spectrometry. *J. Pept. Sci. Supplement to Vol. 14*, 40 (2008)
142. Brask, J., Jensen, K.J.: Carbopeptides: Chemoselective Ligation of Peptide Aldehydes to an Aminoxy-functionalized D-galactose Template. *J. Peptide Sci.* **6**, 290-299 (2000)
143. Decostaire, I.P., Lelievre, D., Zhang, H.H., Delmas, A.F.: Controlling the outcome of overacylation of N-protected aminoxyacetic acid during the synthesis of an aminoxy-peptide for chemical ligation. *Tetrahedron Letters* **47**, 7057-7060 (2006)
144. Shao, J., Tam, J.P.: Unprotected Peptides as Building Blocks for the Synthesis of Peptide Dendrimers with Oxime, Hydrazone, and Thiazolidine Linkages. *J. Am. Chem. Soc.* **117**, 3893-3899 (1995)
145. Wahl, F., Mutter, M.: Analogues of Oxytocin with an Oxime Bridge using Chemoselectively Addressable Building Blocks. *Tetrahedron Letters* **37**, 6861-6864 (1996)
146. Foillard, S., Rasmussen, M.O., Razkin, J., Boturyn, D., Dumy, P.: 1-Ethoxyethylidene, a new group for the stepwise SPPS of aminoxyacetic acid containing peptides. *J. Org. Chem.* **73**, 983-991 (2008)
147. Leu, M.-R., Lee, J., Baek, B.-H., Shin, I.: The First Solid-Phase Synthesis of Oligomeric α -Aminoxy Peptides. *Synlett* **3**, 325-328 (2003)
148. Duléry, V., Renaudet, O., Dumy, P.: Ethoxyethylidene protecting group prevents N-overacylation in aminoxy peptide synthesis. *Tetrahedron* **63**, 11952-11958 (2007)
149. From Novabiochem (Läufelfingen, Switzerland). 2009)
150. Gausepohl, H., Kraft, M., Frank, R.W.: Asparagine Coupling in Fmoc Solid-Phase Peptide-Synthesis. *International Journal of Peptide and Protein Research* **34**, 287-294 (1989)
151. Bure, C., Lelievre, D., Delmas, A.: Identification of by-products from an orthogonal peptide ligation by oxime bonds using mass

- spectrometry and tandem mass spectrometry. *Rapid Commun. Mass Spectrom.* **14**, 2158-2164 (2000)
152. Ramsay, S.L., Freeman, C., Grace, P.B., Redmond, J.W., MacLeod, J.K.: Mild tagging procedures for the structural analysis of glycans. *Carbohydr. Res.* **333**, 59-71 (2001)
153. Zhou, X., Zhou, J.: Oligosaccharide microarrays fabricated on aminoxyacetyl functionalized glass surface for characterization of carbohydrate-protein interaction. *Biosens. Bioelectron.* **21**, 1451-1458 (2006)
154. Vliegthart J.F.G.: High-resolution, ¹H-nuclear magnetic resonance spectroscopy as a tool in the structural analysis of carbohydrates related to glycoproteins. *Adv. Carbohydr. Chem. Biochem.* **41**, 209-374 (1983)
155. Andreana, P.R., Xie, W., Cheng, H.N., Qiao, L., Murphy, D.J., Gu, Q.M., Wang, P.G.: In situ preparation of beta-D-1-O-hydroxylamino carbohydrate polymers mediated by galactose oxidase. *Org. Lett.* **4**, 1863-1866 (2002)
156. Sharon, N. & Lis, H.: Specificity and affinity. In: Sharon, N. & Lis, H. (ed.) *Lectins*, pp. 63-103. (2007)
157. Chen, G.S., Pohl, N.L.: Synthesis of fluororous tags for incorporation of reducing sugars into a quantitative microarray platform. *Org. Lett.* **10**, 785-788 (2008)
158. Zhu, Y., Zajicek, J., Serianni, A.S.: Acyclic forms of [1-(¹³C)]aldohexoses in aqueous solution: quantitation by (¹³C) NMR and deuterium isotope effects on tautomeric equilibria. *J. Org. Chem.* **66**, 6244-6251 (2001)
159. Ramsay, S.L., Freeman, C., Grace, P.B., Redmond, J.W., MacLeod, J.K.: Mild tagging procedures for the structural analysis of glycans. *Carbohydr. Res.* **333**, 59-71 (2001)
160. Jensen, K.J.: Studies on the Mechanism of Oxime and Semicarbazone Formation. *J. Am. Chem. Soc.* **81**, 475-481 (1959)
161. Haas, J.W., Kadunce, R.E.: Rates of Reaction of Nitrogen Bases with Sugars. I. Studies of Aldose Oxime, Semicarbazone and Hydrazone formation. *J. Am. Chem. Soc.* **84**, 4910-4913 (1962)

BIBLIOGRAPHY

162. Thygesen, M.B., Munch, H., Sauer, J., Cló, E., Jorgensen, M.R., Hindsgaul, O., Jensen, K.J.: Nucleophilic Catalysis of Carbohydrate Oxime Formation by Aniline. *J. Org. Chem.* **75**, 1752-1755 (2010)
163. Carrasco, M.R.: N-Alkylaminoxy Amino Acids as Versatile Derivatives for Post-Translational Modification of Synthetic Peptides. 19th American Peptide Symposium Proceedings 300-302 (2005)
164. Carrasco, M.R., Silva, O., Rawls, K.A., Sweeney, M.S., Lombardo, A.A.: Chemoselective alkylation of N-alkylaminoxy-containing peptides. *Organic Letters* **8**, 3529-3532 (2006)
165. McMeekin, T.L., Wilenski, M., Groves, M.L.: Refractive indices of proteins in relation to amino acid composition and specific volume. *Biochem. Biophys. Res. Commun.* **7**, 151-156 (1962)
166. Harpaz, Y., Gerstein, M., Chothia, C.: Volume changes on protein folding. *Structure.* **2**, 641-649 (1994)
167. Itakura, Y., Nakamura-Tsuruta, S., Kominami, J., Sharon, N., Kasai, K., Hirabayashi, J.: Systematic comparison of oligosaccharide specificity of *Ricinus communis* agglutinin I and *Erythrina* lectins: a search by frontal affinity chromatography. *J. Biochem.* **142**, 459-469 (2007)
168. Liu, Y., Feizi, T., Campanero-Rhodes, M.A., Childs, R.A., Zhang, Y., Mulloy, B., Evans, P.G., Osborn, H.M., Otto, D., Crocker, P.R., Chai, W.: Neoglycolipid probes prepared via oxime ligation for microarray analysis of oligosaccharide-protein interactions. *Chem. Biol.* **14**, 847-859 (2007)
169. Allen, A.K., Neuberger, A., Sharon, N.: The purification, composition and specificity of wheat-germ agglutinin. *Biochem. J.* **131**, 155-162 (1973)
170. Gallagher, J.T., Morris, A., Dexter, T.M.: Identification of two binding sites for wheat-germ agglutinin on poly-lactosamine-type oligosaccharides. *Biochem. J.* **231**, 115-122 (1985)
171. ed. by Biacore: *Biacore Sensor Surface Handbook*. 2003)
172. Pace, C.N., Vajdos, F., Fee, L., Grimsley, G., Gray, T.: How to measure and predict the molar absorption coefficient of a protein. *Protein Sci.* **4**, 2411-2423 (1995)

173. Elgavish, S., Shaanan, B.: Structures of the Erythrina corallodendron lectin and of its complexes with mono- and disaccharides. *J. Mol. Biol.* **277**, 917-932 (1998)
174. Kalinin, N.L., Ward, L.D., Winzor, D.J.: Effects of solute multivalence on the evaluation of binding constants by biosensor technology: studies with concanavalin A and interleukin-6 as partitioning proteins. *Anal. Biochem.* **228**, 238-244 (1995)
175. Kadirvelraj, R., Foley, B.L., Dyekjaer, J.D., Woods, R.J.: Involvement of water in carbohydrate-protein binding: concanavalin A revisited. *J. Am. Chem. Soc.* **130**, 16933-16942 (2008)
176. Mandal, D.K., Kishore, N., Brewer, C.F.: Thermodynamics of lectin-carbohydrate interactions. Titration microcalorimetry measurements of the binding of N-linked carbohydrates and ovalbumin to concanavalin A. *Biochemistry* **33**, 1149-1156 (1994)
177. Pereira, M.E., Kabat, E.A.: Blood group specificity of the lectin from *Lotus tetragonolobus*. *Ann. N. Y. Acad. Sci.* **234**, 301-305 (1974)
178. Pereira, M.E., Kabat, E.A.: Specificity of purified hemagglutinin (lectin) from *Lotus tetragonolobus*. *Biochemistry* **13**, 3184-3192 (1974)
179. Yan, L., Wilkins, P.P., Alvarez-Manilla, G., Do, S.I., Smith, D.F., Cummings, R.D.: Immobilized *Lotus tetragonolobus* agglutinin binds oligosaccharides containing the Le(x) determinant. *Glycoconj. J.* **14**, 45-55 (1997)
180. Knibbs, R.N., Goldstein, I.J., Ratcliffe, R.M., Shibuya, N.: Characterization of the carbohydrate binding specificity of the leukoagglutinating lectin from *Maackia amurensis*. Comparison with other sialic acid-specific lectins. *J. Biol. Chem.* **266**, 83-88 (1991)
181. Varki, A.: Glycan-based interactions involving vertebrate sialic-acid-recognizing proteins. *Nature* **446**, 1023-1029 (2007)
182. Duverger, E., Coppin, A., Strecker, G., Monsigny, M.: Interaction between lectins and neoglycoproteins containing new sialylated glycosynthons. *Glycoconj. J.* **16**, 793-800 (1999)
183. Gupta, D., Cho, M., Cummings, R.D., Brewer, C.F.: Thermodynamics of carbohydrate binding to galectin-1 from Chinese hamster ovary cells and two mutants. A comparison with four galactose-specific plant lectins. *Biochemistry* **35**, 15236-15243 (1996)

BIBLIOGRAPHY

184. Gupta, D., Cho, M., Cummings, R.D., Brewer, C.F.: Thermodynamics of carbohydrate binding to galectin-1 from Chinese hamster ovary cells and two mutants. A comparison with four galactose-specific plant lectins. *Biochemistry* **35**, 15236-15243 (1996)
185. De Boeck, H., Loontjens, F.G., Lis, H., Sharon, N.: Binding of simple carbohydrates and some N-acetyllactosamine-containing oligosaccharides to *Erythrina cristagalli* agglutinin as followed with a fluorescent indicator ligand. *Arch. Biochem. Biophys.* **234**, 297-304 (1984)
186. Roos, H., Karlsson, R., Nilshans, H., Persson, A.: Thermodynamic analysis of protein interactions with biosensor technology. *J. Mol. Recognit.* **11**, 204-210 (1998)
187. Sharma, S., Bharadwaj, S., Surolia, A., Podder, S.K.: Evaluation of the stoichiometry and energetics of carbohydrate binding to *Ricinus communis* agglutinin: a calorimetric study. *Biochem. J.* **333 (Pt 3)**, 539-542 (1998)
188. Lopez-Lucendo, M.F., Solis, D., Andre, S., Hirabayashi, J., Kasai, K., Kaltner, H., Gabius, H.J., Romero, A.: Growth-regulatory human galectin-1: crystallographic characterisation of the structural changes induced by single-site mutations and their impact on the thermodynamics of ligand binding. *J. Mol. Biol.* **343**, 957-970 (2004)
189. Svensson, C., Teneberg, S., Nilsson, C.L., Kjellberg, A., Schwarz, F.P., Sharon, N., Krenzel, U.: High-resolution crystal structures of *Erythrina cristagalli* lectin in complex with lactose and 2'-alpha-L-fucosyllactose and correlation with thermodynamic binding data. *J. Mol. Biol.* **321**, 69-83 (2002)
190. Bonneil, E., Young, N.M., Lis, H., Sharon, N., Thibault, P.: Probing genetic variation and glycoform distribution in lectins of the *Erythrina* genus by mass spectrometry. *Arch. Biochem. Biophys.* **426**, 241-249 (2004)
191. Svensson, C., Teneberg, S., Nilsson, C.L., Kjellberg, A., Schwarz, F.P., Sharon, N., Krenzel, U.: High-resolution crystal structures of *Erythrina cristagalli* lectin in complex with lactose and 2'-alpha-L-fucosyllactose and correlation with thermodynamic binding data. *J. Mol. Biol.* **321**, 69-83 (2002)
192. Svensson, C., Teneberg, S., Nilsson, C.L., Kjellberg, A., Schwarz, F.P., Sharon, N., Krenzel, U.: High-resolution crystal structures of

- Erythrina cristagalli lectin in complex with lactose and 2'-alpha-L-fucosyllactose and correlation with thermodynamic binding data. *J. Mol. Biol.* **321**, 69-83 (2002)
193. Mo, H., Meah, Y., Moore, J.G., Goldstein, I.J.: Purification and characterization of Dolichos lablab lectin. *Glycobiology* **9**, 173-179 (1999)
194. Park, S., Lee, M.R., Pyo, S.J., Shin, I.: Carbohydrate chips for studying high-throughput carbohydrate-protein interactions. *J. Am. Chem. Soc.* **126**, 4812-4819 (2004)
195. Kogelberg, H., Solis, D., Jimenez-Barbero, J.: New structural insights into carbohydrate-protein interactions from NMR spectroscopy. *Curr. Opin. Struct. Biol.* **13**, 646-653 (2003)
196. Palmer, R.A., Niwa, H.: X-ray crystallographic studies of protein-ligand interactions. *Biochem. Soc. Trans.* **31**, 973-979 (2003)
197. Mori, T., Toyoda, M., Ohtsuka, T., Okahata, Y.: Kinetic analyses for bindings of concanavalin A to dispersed and condensed mannose surfaces on a quartz crystal microbalance. *Anal. Biochem.* **395**, 211-216 (2009)
198. Löfas, S., Johnsson, B.: A novel hydrogel matrix on gold surfaces in surface plasmon resonance sensors for fast and efficient covalent immobilization of ligands. *J. Chem. Soc. Chem. Commun.* 1526-1528 (1990)
199. Spiga, O., Bernini, A., Scarselli, M., Ciutti, A., Bracci, L., Lozzi, L., Lelli, B., Di Maro, D., Calamandrei, D., Niccolai, N.: Peptide-protein interactions studied by surface plasmon and nuclear magnetic resonances. *FEBS Lett.* **511**, 33-35 (2002)
200. Stockley, P.G., Persson, B.: Surface plasmon resonance assays of DNA-protein interactions. *Methods Mol. Biol.* **543**, 653-669 (2009)
201. Besenicar, M., Macek, P., Lakey, J.H., Anderluh, G.: Surface plasmon resonance in protein-membrane interactions. *Chemistry and Physics of Lipids* **141**, 169-178 (2006)
202. Duverger, E., Frison, N., Roche, A.C., Monsigny, M.: Carbohydrate-lectin interactions assessed by surface plasmon resonance. *Biochimie* **85**, 167-179 (2003)

BIBLIOGRAPHY

203. Hutchinson, A.M.: Characterization of glycoprotein oligosaccharides using surface plasmon resonance. *Anal. Biochem.* **220**, 303-307 (1994)
204. Vornholt, W., Hartmann, M., Keusgen, M.: SPR studies of carbohydrate-lectin interactions as useful tool for screening on lectin sources. *Biosens. Bioelectron.* **22**, 2983-2988 (2007)
205. Gondran, C., Dubois, M.P., Fort, S., Cosnier, S., Szunerits, S.: Detection of carbohydrate-binding proteins by oligosaccharide-modified polypyrrole interfaces using electrochemical surface plasmon resonance. *Analyst* **133**, 206-212 (2008)
206. de Boer, A.R., Hokke, C.H., Deelder, A.M., Wuhrer, M.: Serum antibody screening by surface plasmon resonance using a natural glycan microarray. *Glycoconj. J.* **25**, 75-84 (2008)
207. Munoz, F.J., Perez, J., Rumbero, A., Santos, J.I., Canada, F.J., Andre, S., Gabius, H.J., Jimenez-Barbero, J., Sinisterra, J.V., Hernaiz, M.J.: Glycan tagging to produce bioactive ligands for a surface plasmon resonance study via immobilization on different surfaces. *Bioconjug. Chem.* **20**, 673-682 (2009)
208. Zeng, X., Sun, Y., Ye, H., Liu, J., Xiang, X., Zhou, B., Uzawa, H.: Effective chemoenzymatic synthesis of p-aminophenyl glycosides of sialyl N-acetyllactosaminide and analysis of their interactions with lectins. *Carbohydr. Res.* **342**, 1244-1248 (2007)
209. Lienemann, M., Paananen, A., Boer, H., de la Fuente, J.M., Garcia, I., Penades, S., Koivula, A.: Characterization of the wheat germ agglutinin binding to self-assembled monolayers of neoglycoconjugates by AFM and SPR. *Glycobiology* **19**, 633-643 (2009)
210. Maljaars, C.E., de Souza, A.C., Halkes, K.M., Upton, P.J., Reeman, S.M., Andre, S., Gabius, H.J., McDonnell, M.B., Kamerling, J.P.: The application of neoglycopeptides in the development of sensitive surface plasmon resonance-based biosensors. *Biosens. Bioelectron.* **24**, 60-65 (2008)
211. Carotenuto, A., Alcaro, M.C., Saviello, M.R., Peroni, E., Nuti, F., Papini, A.M., Novellino, E., Rovero, P.: Designed glycopeptides with different beta-turn types as synthetic probes for the detection of autoantibodies as biomarkers of multiple sclerosis. *J. Med. Chem.* **51**, 5304-5309 (2008)

212. Jimenez-Castells, C., de la Torre, B.G., Gutierrez, G.R., Andreu, D.: Optimized synthesis of aminoxy-peptides as glycoprobe precursors for surface-based sugar-protein interaction studies. *Bioorg. Med. Chem. Lett.* **17**, 5155-5158 (2007)
213. Hatanaka, Y., Kempin, U., Jong-Jip, P.: One-step synthesis of biotinyl photoprobes from unprotected carbohydrates. *J. Org. Chem.* **65**, 5639-5643 (2000)
214. Jimenez-Castells, C., de la Torre, B.G., Andreu, D., Gutierrez-Gallego, R.: Neo-glycopeptides: the importance of sugar core conformation in oxime-linked glycoprobes for interaction studies. *Glycoconj. J.* **25**, 879-887 (2008)
215. Lohse, A., Martins, R., Jorgensen, M.R., Hindsgaul, O.: Solid-phase oligosaccharide tagging (SPOT): Validation on glycolipid-derived structures. *Angew. Chem. Int. Ed Engl.* **45**, 4167-4172 (2006)
216. Seo, J., Michaelian, N., Owens, S.C., Dashner, S.T., Wong, A.J., Barron, A.E., Carrasco, M.R.: Chemoselective and microwave-assisted synthesis of glycopeptoids. *Org. Lett.* **11**, 5210-5213 (2009)
217. Seo, J., Michaelian, N., Owens, S.C., Dashner, S.T., Wong, A.J., Barron, A.E., Carrasco, M.R.: Chemoselective and microwave-assisted synthesis of glycopeptoids. *Org. Lett.* **11**, 5210-5213 (2009)
218. Carrasco, M.R., Silva, O., Rawls, K.A., Sweeney, M.S., Lombardo, A.A.: Chemoselective alkylation of N-alkylaminoxy-containing peptides. *Org. Lett.* **8**, 3529-3532 (2006)
219. Matsubara, N., Oiwa, K., Hohsaka, T., Sadamoto, R., Niikura, K., Fukuhara, N., Takimoto, A., Kondo, H., Nishimura, S.: Molecular design of glycoprotein mimetics: glycoblotting by engineered proteins with an oxylamino-functionalized amino acid residue. *Chemistry* **11**, 6974-6981 (2005)
220. Bohorov, O., Andersson-Sand, H., Hoffmann, J., Blixt, O.: Arraying glycomics: a novel bi-functional spacer for one-step microscale derivatization of free reducing glycans. *Glycobiology* **16**, 21C-27C (2006)
221. Liu, Y., Feizi, T., Campanero-Rhodes, M.A., Childs, R.A., Zhang, Y., Mulloy, B., Evans, P.G., Osborn, H.M., Otto, D., Crocker, P.R., Chai, W.: Neoglycolipid probes prepared via oxime ligation for microarray analysis of oligosaccharide-protein interactions. *Chem. Biol.* **14**, 847-859 (2007)

BIBLIOGRAPHY

222. Gupta, D., Cho, M., Cummings, R.D., Brewer, C.F.: Thermodynamics of carbohydrate binding to galectin-1 from Chinese hamster ovary cells and two mutants. A comparison with four galactose-specific plant lectins. *Biochemistry* **35**, 15236-15243 (1996)
223. Ambrosi, M., Cameron, N.R., Davis, B.G.: Lectins: tools for the molecular understanding of the glycode. *Org. Biomol Chem.* **3**, 1593-1608 (2005)
224. Hager-Braun, C., Tomer, K.B.: Determination of protein-derived epitopes by mass spectrometry. *Expert. Rev. Proteomics.* **2**, 745-756 (2005)
225. Svensson, C., Teneberg, S., Nilsson, C.L., Kjellberg, A., Schwarz, F.P., Sharon, N., Krenzel, U.: High-resolution crystal structures of *Erythrina cristagalli* lectin in complex with lactose and 2'-alpha-L-fucosyllactose and correlation with thermodynamic binding data. *J. Mol. Biol.* **321**, 69-83 (2002)

8 APPENDIX

Jiménez-Castells, C.; Defaus, S.; Andreu, D. & Gutiérrez-Gallego, R.

"Recent progress in the field of neoglycoconjugate chemistry"

BioMol Concepts. **1**, 85-96 (2010)

Jiménez-Castells C, Defaus S, Andreu D, Gutiérrez-Gallego R. [Recent progress in the field of neoglycoconjugate chemistry](#). BioMol Concepts. 2010; 1(1): 85-96.

Jiménez-Castells, C.; de la Torre, B.G.; Gutiérrez-Gallego, R.; Andreu, D.

"Optimized synthesis of aminoxy-peptides as glycoprobe precursors for surface-based sugar-protein interaction studies."

Bioorg. Med. Chem. Lett. **17**, 5155-5158 (2007)

Jiménez-Castells C, de la Torre BG, Gutiérrez Gallego R, Andreu D. [Optimized synthesis of aminoxy-peptides as glycoprobe precursors for surface-based sugar-protein interaction studies](#). Bioorg Med Chem Lett. 2007; 17(18): 5155-8.

Jiménez-Castells, C.; de la Torre, B.G.; Andreu, D.; Gutiérrez-Gallego, R.

"Neo-glycopeptides: the importance of sugar core conformation in oxime-linked glycoprobes for interaction studies"

Glycoconjugate J. **25**, 879-887 (2008)

Jiménez-Castells C, de la Torre BG, Andreu D, Gutiérrez-Gallego R. [Neo-glycopeptides: the importance of sugar core conformation in oxime-linked glycoprobes for interaction studies.](#) Glycoconj J. 2008; 25(9): 879-87.